EUMETSAT Satellite Application Facility on Climate Monitoring

Visiting Scientist Report

Evaluation and Assessment of the CM-SAF Surface Solar Radiation Climate Data Records

CDOP AS Study No 18

Date: 29.08.2011

by

Sven Brinckmann,
Bodo Ahrens

Institut für Atmosphäre und Umwelt
Goethe-Universität Frankfurt am Main
email: s.brinckmann@iau.uni-frankfurt.de

Contents

Abstract

A satellite-based climate dataset (named Heliosat) of monthly means of the surface solar irradiance (SIS) from 1983 to 2005, covering the area between -70° and +70° in latitude and longitude, was investigated with regard to possible inhomogeneities. Several BSRN (Baseline Surface Radiation Network) station measurements and ERA-Interim model reanalysis fields have been used as reference data. Using the Standard Normal Homogeneity Test (SNHT) and two other test procedures, we have focused on the detection of break-type inhomogeneities which may occur due to satellite changes. In comparison with the few available station measurements the Heliosat time series do not show any significant inhomogeneities, even though slight discrepancies to the surface measurements appear. The analyses of the complete data field reveal breaks in the time series related to satellite changes for certain regions especially in Africa. For other areas, e.g. Europe, no such inhomogeneities have been found.

1 Introduction

A novel climate data record from the Satellite Application Facility on Climate Monitoring (CM-SAF) of the surface solar incoming radiation has been deduced from satellite measurements. The so-called "Heliosat" data set is based on geostationary satellite observations from the MVIRI instruments aboard the Meteosat satellite series from 1983 to 2005. The current reflectance in the visible range (compared to corresponding maximum and minimum values over one month) and the theoretically determined maximum SIS at each grid point have been used to calculate the current SIS.

High-quality measurements of the SIS at surface stations are limited to a small number of stations and the corresponding time series do not go back further than 1992. Therefore, the highly resolved (0.03°) Heliosat dataset with a large coverage can provide very useful data which allow investigations of regional and global changes in the incoming solar radiation. Such variations can be caused by indirect effects of variabilities in the solar irradiance (e.g. the 11-year solar cycle) as well as by changes in the aerosol abundances as a result of anthropogenic activities (IPCC, 2007). The analysis of global datasets of the SIS can thus help to improve the understanding and quantification of these processes, which is of importance for the overall assessment of the global anthropogenic climate change.

Due to the fact, that the satellite-derived SIS data base on indirect measurements involving potentially large errors, the data must be interpreted very carefully. Several changes of the instruments within the observation period between 1983 and 2005 and sensitivity drifts of the instruments can lead to discontinuities as well as artificial trends, which may superimpose real climatological variations. To evaluate the homogeneity of time series, numerous techniques have been developed (see e.g. Peterson et al., 1998). An important basis of homogeneity testing is the comparison of the investigated time series with reference data of high quality. Depending on the type of climate variable difference or ratio series between candidate and reference are considered, so that climate variations

can ideally be seperated from artificial changes. The different methods are usually designed to evaluate the probability for the presence of a specific type of inhomogeneity. In some cases time series do contain multiple inhomogeneities, whose detection requires advanced strategies of homogeneity testing. Additional complexity can be introduced by the presence of both break-like and trend-like inhomogeneities.

2 Heliosat dataset

The SIS dataset is based on measurements of the Meteosat Visible and InfraRed Imager (MVIRI), a radiometer covering three spectral ranges. Using the visible spectrum (0.45-1μm) of these geostationary satellite instruments, the SIS is derived in a two-step approach. First, the Heliosat algorithm uses the observed reflectance to determine the effective cloud albedo n, also denoted as cloud index in previous literature (Cano et al., 1986; Beyer et al., 1996; Hammer et al., 2003). A clear-sky radiatiative transfer model is used to calculate the clear-sky surface solar irradiance (often termed global radiation) (SIS_{cs}). SIS is then determined by combining the clear sky irradiances with the retrieved effective cloud albedo. In the following, we present an overview of the retrieval method used to generate the data set. For a more detailed description the interested reader is referred to Posselt et al. (2011).

Traditionally, the Heliosat algorithm uses the digital count D of the visible satellite channel. In this way, it neither depends on any calibration information nor on information from other channels. The retrieval of the effective cloud albedo is based on the assumption that the current cloudiness is related to the difference between 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 n is derived from the current normalized counts ρ (i.e., accounting for the dark offset, the solar zenith angle and the sun-earth distance), the clear sky normalized counts ρ_{cs} and the normalized counts of a compact (not convective) cloud deck ρ_{max} as an estimation for the maximum possible normalized count.

$$n = \frac{\rho - \rho_{cs}}{\rho_{max} - \rho_{cs}} \tag{1}$$

The clear-sky normalized count is determined for every pixel separately as the minimum value of ρ during a certain time period (e.g., one 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). In the generation of this data set, the maximum normalized count is used in the new self-calibration method as presented in Posselt et al. (2011).

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 and 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, 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 gnu-public license at http://sourceforge.net/projects/gnu-magic/.

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

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

Thereby, SIS is not fully proportional to $1/n$. Instead, for small and large values of n the factor $k(n)$ is adjusted to the exact range of n (see details in Hammer et al., 2003).

SIS is determined for each MVIRI satellite observation, i.e., every 30 minutes for the full satellite view. The data were regridded to a regular lon-lat grid with a 0.03° spacing using a triangulation technique. The calculation of the monthly means is based on the daily means, using the requirement that a minimum of 20 daily mean values are available. The target accuracy of the monthly mean surface solar radiation products is 10 Wm^{-2}.

In the time period between 1983 and 2005 six generations of satellites have been used (Meteosat 2-7). Table 1 shows the main dates at which the use of the instruments has been switched to another satellite generation. In case that instrument changes have an effect on the quality of the SIS dataset, breaks in the corresponding time series would be expected for these dates. Other types of inhomogeneities such as trends could be caused by sensitivity drifts of the instruments. However, as the SIS values are based on relative reflectance measurements – with ρ_{cs} and ρ_{max} determined repeatedly from month to month – such gradual inhomogeneities related to sensitivity problems are expected to be small.

In addition to the dates listed in table 1 further temporary changes (of a few days) between two consecutive instruments have occurred (see details in Rigollier et al., 2002). These changes might also have a slight impact on the quality of the data for the specific months, but assuming that this influence is small, the affected data are considered in our analyses in order to maintain continuous time series. Due to technical problems with the satellite measurements the data for one month (12/1988) are missing in the Heliosat data set.

3 Reference data

3.1 BSRN station measurements

BSRN is a project of the World Climate Research Program (WCRP) and the Global Energy and Water Experiment (GEWEX). The aim of the BSRN is to provide consistent quality-controlled surface radiation data from a global station network using a defined set of instrumentation and measurement protocols. Currently, about 40 surface stations are operating in different climatic zones. The accuracy of the shortwave measurement is estimated to be 5 Wm^{-2} (Ohmura et al., 1998). Here, we use the surface data obtained

Table 1: History of the Meteosat satellites used for the Heliosat SIS dataset.

Satellite	Beginning of period	End of period
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 1989	19 Apr 1990
Meteosat-4	19 Apr 1990	04 Feb 1994
Meteosat-5	04 Feb 1994	13 Feb 1997
Meteosat-6	13 Feb 1997	03 Jun 1998
Meteosat-7	03 Jun 1998	31 Dec 2005

at 5 stations covering time periods back to at least 1996 as a reference to evaluate the quality of the satellite-derived data sets.

3.2 ERA-Interim data fields

The BSRN station measurements, available for very few sites and a limited time period (starting at the earliest in 1992), alone are not sufficient for a comprehensive evaluation of the Heliosat data. For a comparative homogeneity analysis of the whole Heliosat dataset we made use of the ERA-Interim reanalysis fields provided by the ECMWF (European Centre for Medium-Range Weather Forecasts). This data field with a spatial resolution of 0.7° is available from 1979 until today and is continuously updated (Dee et al., 2011). While satellite data provide the biggest single source of information for the assimilation scheme, also surface observations and soundings are assimilated using a 4D-VAR assimilation system. A bias correction scheme to correct for long-term drifts and calibration errors in satellite data is included in the reanalysis (Dee and Uppala, 2009). While the low-frequency variations of surface humdity, temperature, and precipitation have been shown to be consistent with surface observations (Simmons et al., 2010), an analysis of the temporal homogeneity of the solar surface radiation data from ERA-Interim has not been conducted so far.

4 Test procedures

In this work we mainly used the SNHT described by Alexandersson (1986) to detect possible inhomogeneities in the investigated time series. For each point k between 1 and n, where n is the length of the time series, the test value $T(k)$ is calculated. This value compares the normalized mean of the first k years ($\bar{z_1}$) with that of the last $n-k$ years ($\bar{z_2}$), whereas $\bar{Y}$ and s denote the mean and the standard deviation respectively over the whole period.

$$T(k) = k\bar{z}_1^2 + (n-k)\bar{z}_2^2 \quad k = 1, ..., n \tag{3}$$

with

$$\bar{z}_1 = \frac{1}{k}\sum_{i=1}^{k}(Y_i - \bar{Y})/s \quad ; \quad \bar{z}_1 = \frac{1}{n-k}\sum_{i=k+1}^{n}(Y_i - \bar{Y})/s \tag{4}$$

$T(k)$ grows with the difference between the two means. If the maximum of $T(k)$ (denoted T_0) exceeds a certain level, a shift of the mean at k is detected. These critical values T_c depend on the length of the investigated time series and the chosen significance level. The T_c were derived from simulations with 10.000 normally distributed random time series, while we used a significance of 95% in all analyses presented in this work.

The test procedure should be applied to a given candidate time series only in comparison with a related reference time series of high quality. Following this recommendation, we calculated difference series (candidate minus reference) for all investigated SIS data. For both, candidate and reference, monthly anomalies have been used instead of the absolute values in order to eliminate possible seasonal difference patterns. In case that the reference time series itself includes inhomogeneities, such a relative analysis will detect inhomogeneities as a result of the inconsistency in the reference series. To assess such problems, an additional absolute test of both candidate and reference time series has been carried out for all presented analyses.

We compared the outcomes of the SNHT with two other test procedures designed for the detection of breaks in time series, which are the Buishand range test (Buishand, 1982) and the Pettitt test (Pettitt, 1979). A description of these tests is given in the appendix. Our analyses with all three tests show a high consistency for these test procedures (compare section 6.2). Discrepancies occur mainly due to the fact, that the SNHT is very sensitive to breaks near the beginning and the end of time series, while the other two methods are more sensitive to shifts in the middle of a series. For many of the analyses presented in the following sections we only relied on the SNHT, while excluding positive results for the first and the last ten months of a time series to take into account the mentioned sensitivity problem of this method.

Figure 1 (left panels) illustrates the SNHT test statistics for an artificial time series with constant mean but one shift of the mean at 06/1994. The test value T peaks right at this date with a value of 137 which is far above the critical level of around 10 determined for the given length of time series and a significance level of 95%. It is clear that the break detection will be more difficult if the magnitude of the shift is small compared to the variability of the time series. In such cases the detected date of the break can deviate slightly from the real time. In case of stronger trends in the investigated time series (which can be a result of both natural and artificial forcings) the SNHT also tends to detect an inhomogeneity, but with a test statistic showing plateau-like curves rather than sharp maxima. Other test methods have been developed to detect and assess such gradual changes (Alexandersson and Moberg, 1997) or other types of inhomogeneities (Bayesian approach, see e.g. Lee et al. (2005)). Nevertheless, as we expect a shift of the mean as a possible result of satellite changes (compare section 2), we do rely on the usual SNHT in this study.

Within the investigated time period from 1983 to 2005 in total 7 major instrument changes (table 1) have occurred. So, it is possible that multiple breaks appear within the

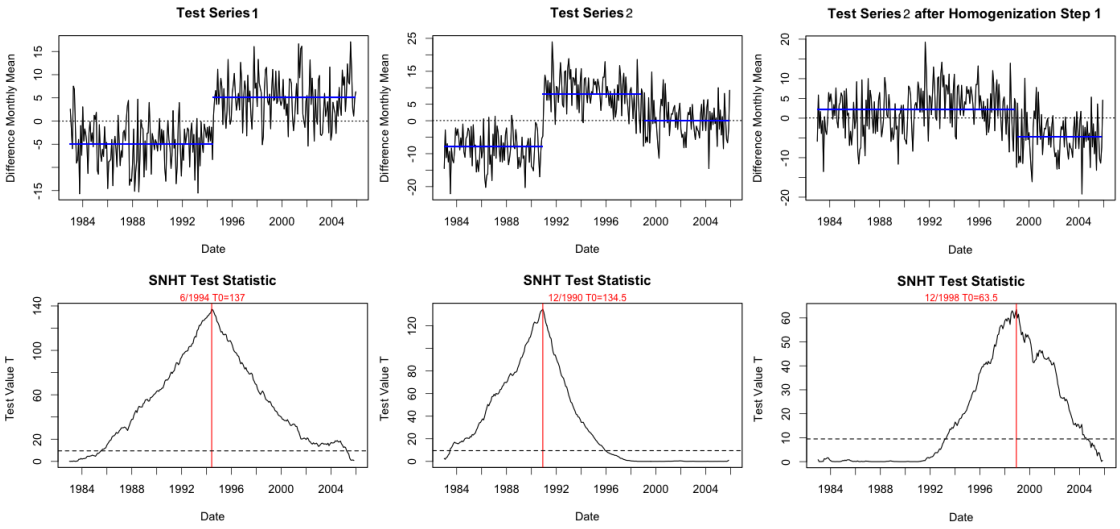

Figure 1: SNHT test statistics for test series 1 with one break (left) and test series 2 with two breaks (middle). After the first homogenization step the second break is detected (right).

investigated time series. To consider multiple breaks, the aforementioned test procedure has to be extended. Easterling and Peterson (1995) have described a method to homogenize climatological time series in consideration of multiple shifts. Using the SNHT, the total time period is divided into subintervals before and after a detected break date. The test is then applied to the subintervals to check for possible further shifts. This procedure with subdivision and testing is repeated as long as new inhomogeneities are found. Finally all breaks are corrected backwards (starting with the last one) by adjusting the data for the two adjacent subintervals with respect to the mean of the younger time period.

We have decided to apply another method, in which the considered time interval remains constant. After a positive SNHT the two subseries of the intervals before and after the first break are both adjusted to means of zero. The corrected series (with respect to the first detected shift) is tested again for a possible second break. This process is repeated 12 times and all significant breaks (dates and magnitudes) are noted. In figure 1 (middle and right panel) the first two steps of the described test method are illustrated for an artificial time series with two breaks (after 96 and 192 months). The dates of the two breaks are well detected at 12/1990 and 12/1998. One problem of this approach is the imprecise adjustment of the two subseries in the first homogenization step, as, in this case, the mean of the second subseries is determined over an interval including the secondary break. This has the consequence that the magnitude of the break may not be well captured in the first step. But as we allow a repeated detection of identical break dates in the following test steps, whereas the new break magnitudes are added to the value from the step it was diagnosed first (even if significance is not reached again), this quantification error is expected to be small.

5 Analyses of Reference Data

5.1 BSRN station measurements

In the first step the reference datasets have been investigated using the methods described above. Concerning the BSRN station measurements, five time series within the Heliosat observation area have been closely investigated, namely Carpentras (France), Payerne (Switzerland), Lindenberg (Germany), Bermuda (Atl., Great Britain) and Florianopolis (Brazil). These time series provide a sufficient length (starting between 1992 and 1996) and quality to be suitable for a comparison with the Heliosat data, even though for the Heliosat time series before 1992 no information can be drawn from this comparison.

In figure 2 the SNHT analyses for Payerne and Bermuda are illustrated. For Payerne we find a test value T slightly exceeding the critical level around 02/2003, while for Bermuda T is clearly below this level over the whole time period. The relatively strong linear trend, as apparent for Payerne BSRN (see blue trend line in the graph), favors the detection of an inhomogeneity. This is an important issue regarding the homogeneity analysis of absolute time series rather than difference time series candidate minus reference. In the difference series such a trend would be removed, if the two time series both exhibit this trend.

Further results of the SNHT analyses and the two other test methods (Buishand, Pettitt) for Payerne, Bermuda as well as the other three BSRN stations are listed in table 2. Given are the maximum test values T_0 (max(T), SNHT), R (Buishand) and X_{k0} (Pettitt), the corresponding critical values for a significance level of 95% and the linear trends of the test series of monthly anomalies. In case of significance (marked bold) the date of the inhomogeneity is added (for comparison in some cases insignificant results are added in brackets). Except for Payerne no other of the four investigated time series was found to be inhomogeneous. These results are consistent for the three considered test methods. Only for Payerne deviating test results (with a negative Buishand test result and positive outcomes for SNHT and the Pettitt test) have been found.

For Payerne a second time series from the so-called ASRB (Alpine Surface Radiation Budget) network was available for additional analyses. In the last two columns of table 2 the outcomes of these calculations, an absolute analysis of the ASRB series as well as a relative analysis ASRB versus BSRN are shown. As noted in section 4, the differences of the monthly anomalies have been calculated to form the test series for the relative analysis. For the absolute tests we find the dates of the maxima coinciding for all three test methods, which is a good indication of the agreement between the two series (maxima caused by the same variation of the measured SIS data) and shows that the inhomogeneity detected for BSRN absolute is apparently a result of real climatological fluctuations. Even though the agreement between BSRN and ASRB is very good over most of the time period, with a mean absolute deviation (between the monthly anomalies) of only 0.93 W/m^2, somewhat higher discrepancies at the very beginning cause a slight indication of a break at 5/1996. Nevertheless, considering the high sensitivity of the SNHT towards beginning and end of a time period, a significant inhomogeneity can be excluded.

Table 2: Results of absolute homogeneity testing of BSRN station measurements and comparison with ASRB measurements at Payerne (last two columns).

Station Start Date	CAR 9/1996	FLO 7/1994	BER 1/1992	LIN 10/1994	PAY 10/1992	PAY ASRB 1/1995	ASRB vs. BSRN 1/1995
SNHT:							
T_0/T_c	3.44/8.53	2.65/8.71	3.39/8.92	5.05/8.67	**11.61**/8.82	8.16/8.67	**8.90**/8.67
Date of Break					**2/2003**	(2/2003)	5/1996
Buishand:							
R/R_c	1.04/1.62	1.32/1.63	1.06/1.63	1.37/1.63	1.39/1.63	1.29/1.63	1.03/1.63
Date of Break					(2/2003)	(2/2003)	(5/1996)
Pettitt:							
X_{k0}/X_c	617/887	700/1220	1067/1649	914/1132	**1582**/1520	1080/1147	**1201**/1147
Date of Break					**1/2003**	(2/2003)	**5/1996**
Linear Trend $[Wm^{-2}yr^{-1}]$	0.59	0.31	0.35	0.23	0.98	0.71	0.06
Mean Deviation $[Wm^{-2}]$							0.93

5.2 ERA-Interim data fields

Regarding the ERA-Interim SIS data fields, a similar absolute inhomogeneity analysis under consideration of all three test methods was applied for each grid point. Figure 3 (left panel) shows the results for the SNHT test value T_0 (first analysis step), whereas only significant values on the 95% level are displayed. For most of the considered regions the time series are indicated to be homogeneous. Relatively high values of T_0 of up to around 80 are reached in certain tropical areas. In the same figure on the righthand side the numbers of independent inhomogeneities (identical dates are counted once) for each single grid point are illustrated. It is interesting to see that the highest numbers - values of 5 and 6 have been detected for a region in the tropical Atlantic and an area in the North of Brazil respectively - do not necessarily coincide with the regions of highest T_0 values.

To assess the types of inhomogeneities found for the most problematic regions according to both figures, we have looked at the time series of specific grid points. For the two regions with very high T_0 values over Africa very broad maxima of T can be found, which indicates the presence of significant long-term trends in the time series (see figure 13 (upper panels) in the appendix). In contrast, for the two mentioned regions with the highest number of inhomogeneities the test statistics show relatively sharp peaks as a result of consecutive break-like inhomogeneities in the time series (same figure, lower panels). Nevertheless, it is apparent from the corresponding data series that even these changes occur more or less continuously (within short time periods) rather than by a sudden shift of the mean.

Figure 4 shows the frequency distribution of the detected inhomogeneities with respect to the date of their appearances (according to the SNHT). The distribution is found to be relatively uniform, without exhibiting single outstanding dates that would indicate breaks with a presence on a global scale. As mentioned in section 4 the results for the first and the last ten dates are not considered, due to the sensitivity problems of the SNHT.

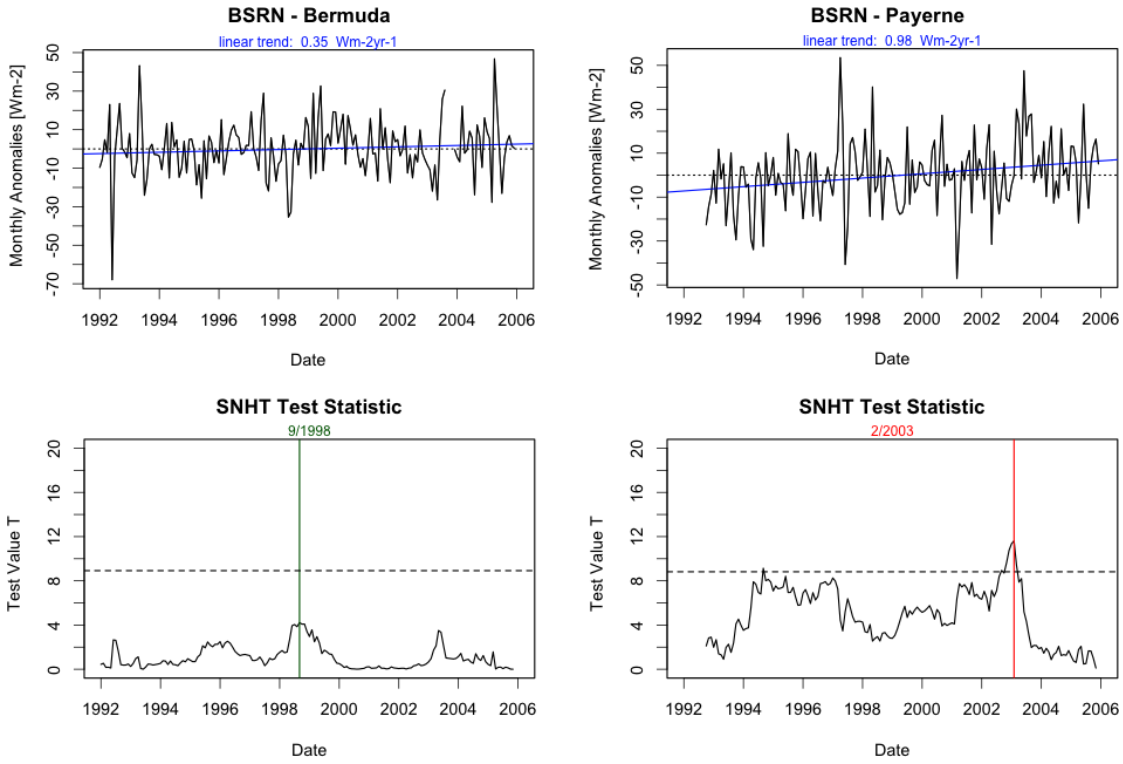

Figure 2: SNHT test statistics for the absolute analyses of Bermuda BSRN (left panels) and Payerne BSRN (right panels).

The red lines mark the dates of the seven satellite changes relevant for the Heliosat data. Instrument changes that have occurred around the middle of a month, are marked by two lines positioned at the two months adjacent to this date. Even though we do not expect inhomogeneities in the ERA-Interim datasets related to these changes, there is a small connection between these satellite measurements and the model reanalyses, as the satellite data contribute to a certain degree to the data assimilation of the model. Our results show very well that such possible inhomogeneities related to satellite changes have no visible impact on the ERA-Interim SIS data.

Additionally, we have carried out a relative analysis of the ERA-Interim time series versus BSRN/ASRB at the grid points nearest to corresponding BSRN/ASRB stations. The outcomes of theses calculations are shown in table 5 in the appendix. We find relatively good agreement between the different datasets, considering that the spatial distances between the station sites and corresponding ERA-Interim grid points are relatively large. Only for Florianopolis an inhomogeneity, as a result of significantly deviating trends, was detected by all three test methods.

Summarizing the outcomes of the above analyses, we do not find a clear indication of inhomogeneities in the investigated time series of the surface measurement network

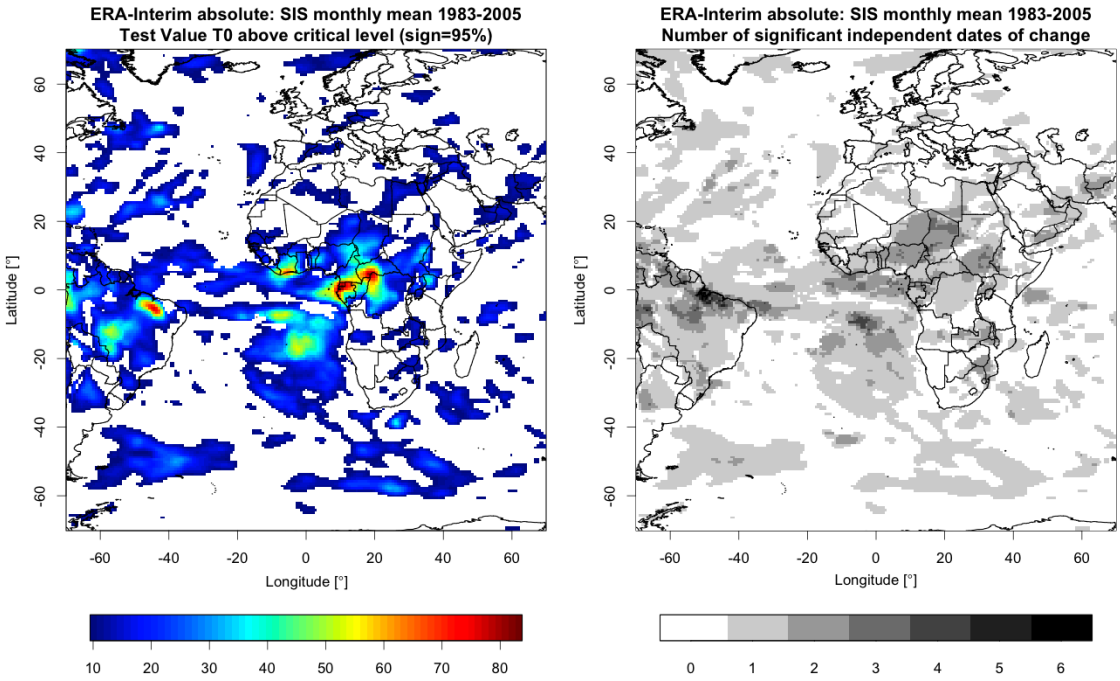

Figure 3: Test values T_0 above significance level of 95% (left panel) and number of detected inhomogeneities (right panel) for an absolute analysis of the ERA-Interim data fields using the SNHT.

BSRN and one additional ASRB time series. The comparison between BSRN and ASRB for Payerne reveals in total very good agreement between the two datasets, which indicates the high quality of these station measurements and their suitability as reference time series for the evaluation of the Heliosat data at single grid points. Concerning the ERA-Interim data fields, for most of the considered regions homogeneity is indicated according to the absolute test analyses. Certain areas, especially in the tropics, exhibit stronger trends (partly on very short timescales), which seem to be of an artificial origin. For these regions the ERA-Interim model data can not be used as a reference for the homogeneity investigations of the Heliosat data fields.

6 Analyses of Heliosat Data

6.1 Comparison with station measurements

To assess the homogeneity of the Heliosat time series at the sites of the station measurements both, an absolute analysis of the candidate series and a relative calculation versus the corresponding BSRN reference time series were conducted. For Payerne the additional ASRB time series was also considered as reference. The outcomes of these analyses are summarized in tables 3 and 4. Regarding the absolute calculations (table 3) none of the five time series – for a better comparison the same time periods as avail-

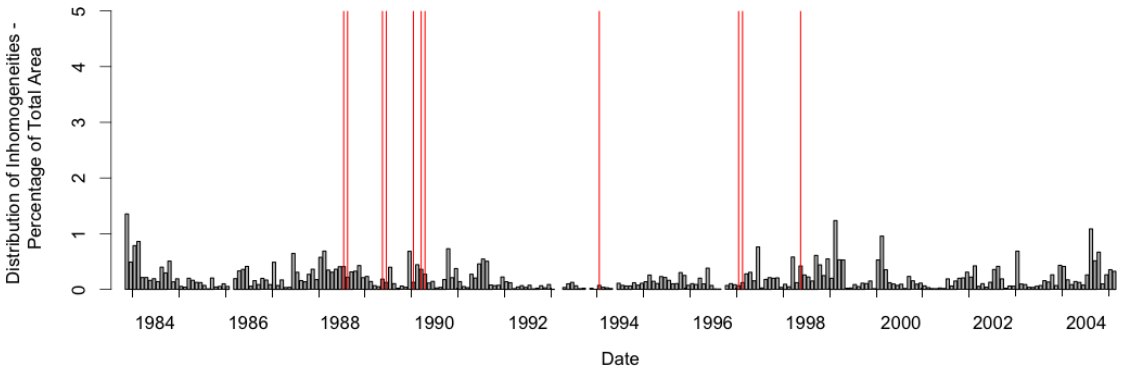

Figure 4: Distribution of the frequency of breaks (in percent of total area) as diagnosed for the ERA-Interim SIS datafield (absolute analysis). The dates of Meteosat satellite changes are marked red.

Table 3: Results of absolute homogeneity testing of Heliosat time series at the grid points nearest to corresponding BSRN stations.

Station Start Date	CAR 9/1996	FLO 7/1994	BER 1/1992	LIN 10/1994	PAY 10/1992
SNHT:					
T_0/T_c	4.78/8.53	2.56/8.71	5.44/8.92	6.04/8.67	8.12/8.82
Date of Break					(1/2003)
Buishand:					
R/R_c	1.12/1.62	1.25/1.63	1.54/1.63	1.45/1.63	1.18/1.63
Date of Break					(1/2003)
Pettitt:					
X_{k0}/X_c	742/887	607/1220	1485/1649	958/1132	1268/1520
Date of Break					(1/2003)
Linear Trend					
$[Wm^{-2}yr^{-1}]$	0.81	0.29	0.31	0.38	0.66

able for the reference series have been considered – shows noticeable inhomogeneities according to the three tests.

In contrast, the relative analyses (table 4) indicate a slight tendency to inhomogeneities for Carpentras, Bermuda and Lindenberg, with one or two positive test results out of three tests for these three time series. For Payerne significant discrepancies between Heliosat and both reference series BSRN and ASRB appear, resulting in test values above the critical levels for all three test methods. In figure 5 these relative analyses are illustrated, exemplary for Bermuda and Payerne (BSRN) (see figure 14 in the appendix for the analyses at the other stations and Payerne ASRB). At Bermuda the slight inhomogeneity is detected for the beginning of the considered time interval, caused by a significant trend in the difference time series. Such a relative trend between Heliosat and BSRN can also be found for Payerne, which results in a broad maximum in the test statistics between around 1996 and 2000. The three satellite changes that occurred

Table 4: Results of relative homogeneity testing of Heliosat time series (versus BSRN) at the grid points nearest to corresponding BSRN stations.

Station	CAR	FLO	BER	LIN	PAY BSRN	PAY ASRB
Start Date	9/1996	7/1994	1/1992	10/1994	10/1992	1/1995
SNHT:						
T_0/T_c	**12.47**/8.53	2.25/8.71	**9.61**/8.92	6.08/8.67	**12.86**/8.82	**15.64**/8.67
Date of Break	**8/2004**		**9/1993**	(4/2002)	**11/1997**	**2/1997**
Buishand:						
R/R_c	1.45/1.62	1.11/1.63	**1.92**/1.63	1.39/1.63	**1.94**/1.63	**1.92**/1.63
Date of Break	(1/2003)		**9/1993**	(4/2002)	**9/1998**	**9/1998**
Pettitt:						
X_{k0}/X_c	809/887	649/1220	1327/1649	**1170**/1132	**1882**/1520	**1461**/1147
Date of Break	(1/2003)		(9/1993)	**4/2002**	**9/1998**	**9/1998**
Linear Trend						
$[Wm^{-2}yr^{-1}]$	0.22	-0.03	0.04	0.15	-0.32	-0.48
Mean Deviation						
$[Wm^{-2}]$	2.9	6.1	5.3	3.0	4.8	4.9

within the analyzed time period are marked with grey vertical lines. It is obvious that none of the slight inhomogeneities detected for the single time series has the characteristics of a break that could be associated with these instrument changes. Since the station measurements are assumed to be of higher quality (concerning both accuracy and precision) than the Heliosat time series, which is also indicated by the very good agreement between the BSRN and the ASRB measurements at Payerne (with a mean absolute deviation of 0.9 W/m^2, table 2), much of the observed differences between Heliosat and BSRN is expected to be associated with problems in the Heliosat data. Nevertheless, mean differences of 2.9 to 6.1 W/m^2, as determined for the monthly anomalies between Heliosat and BSRN for the five stations, are acceptable considering the relatively high uncertainties involved by indirect satellite measurements as used here.

Overall, the comparison of the Heliosat data with the five BSRN time series reveals relatively good agreement. Even though discrepancies near the 95% significance levels appear in four of the five time series, break-type inhomogeneities related to satellite changes are not found.

6.2 Comparison with ERA-Interim data fields

In a first step we analyzed the Heliosat data absolutely performing the same test procedures as used in the previous section. To reduce the computing time and for a better comparison with the ERA-Interim data fields, the Heliosat data were interpolated (using first order conservative remapping) to the 0.7° grid used in ERA-Interim. In figure 6 (left panel) the results of these absolute analyses for the test value T_0 as diagnosed for each grid point are shown. We find certain areas, especially in Africa, with extremly high T_0 values (above 200). For most of these areas a multitude of inhomogeneities (up to 9) was analyzed.

Using the ERA-Interim data fields, we determined difference series between the monthly

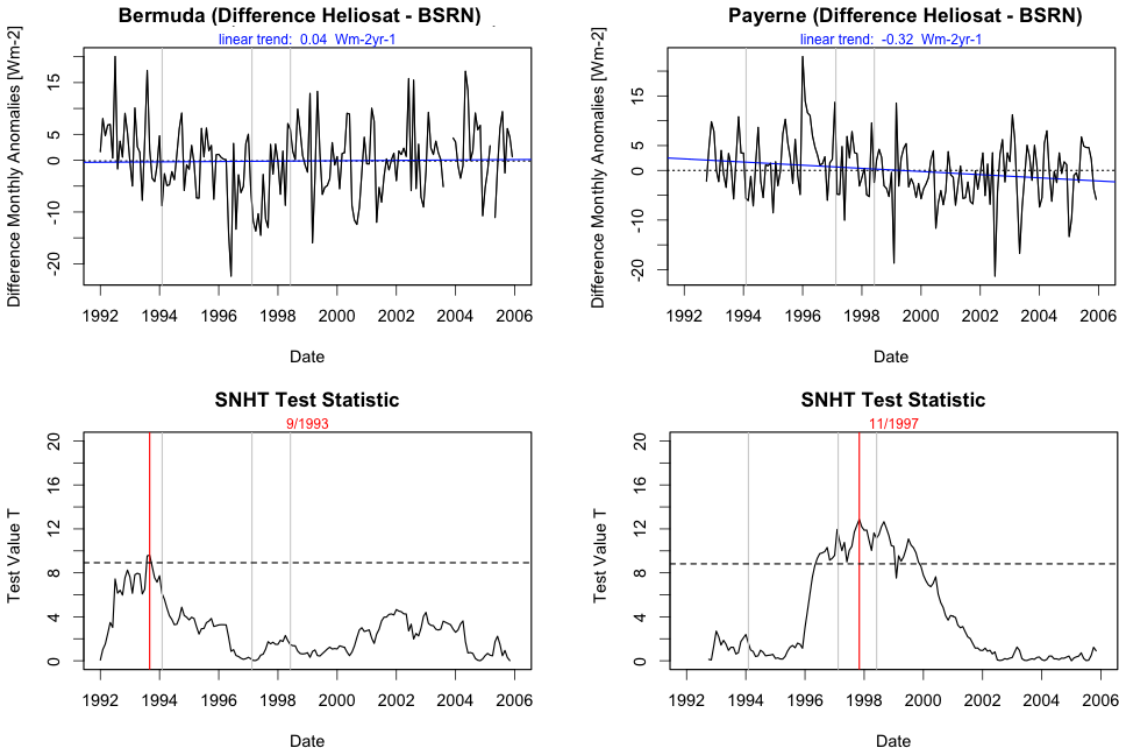

Figure 5: SNHT test statistics for the relative analyses Heliosat versus BSRN for Bermuda (left panels) and Payerne (right panels).

anomalies of the Heliosat and the ERA-Interim SIS. The outcomes for the analyses of these difference series are shown in figure 6 on the right. Even though for most of the considered areas the ERA-Interim data fields are indicated to be homogeneous (as shown in section 5), for some tropical regions a clear indication of inhomogeneities in these reference data is given. To consider this problem we excluded all positive results from the relative analysis (Heliosat versus ERA-Interim) in case no inhomogeneities were detected in the absolute Heliosat analysis. In this way we can avoid the consideration of peaks which are possibly a result of inconsistent model data. Due to the fact that for all regions with strong inhomogeneities in Heliosat absolute the ERA-Interim data are indicated to have a relatively high quality, the outcomes of the relative calculations should give a good picture of the most significant inhomogeneities related to inconsistent satellite data.

Using these filter restraints, we obtain a modified distribution of inhomogeneities as illustrated in figure 7. In the upper two panels the test values T_0 above significance level (left) and the number of detected inhomogeneities (right) are shown. The changes with respect to the absolute analyses and the unfiltered relative calculations are small, which underlines the dominant impact of the Heliosat data. Again, areas with numerous inhomogeneities of very high significance are found especially in some regions of North

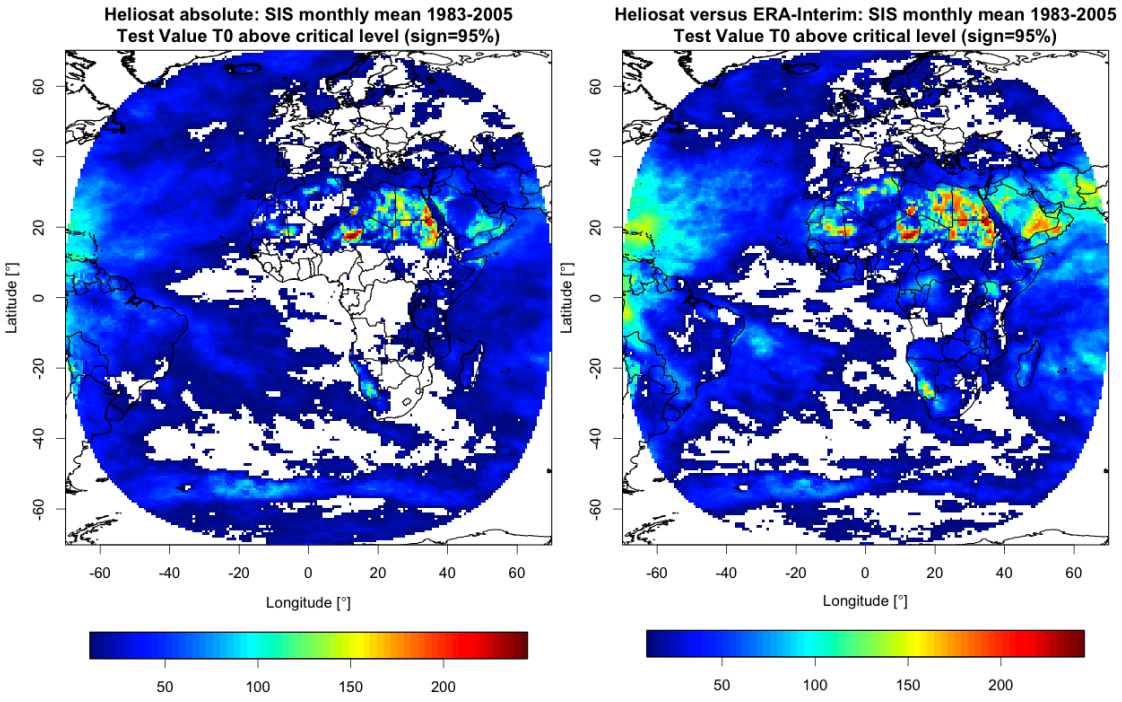

Figure 6: Test values T_0 above significance level of 95% for an absolute analysis of the Heliosat data fields (left) and a relative analysis versus ERA-Interim (right, unfiltered).

Africa, while the majority of the observation area is indicated to be homogeneous (e.g. continental Europe) or only slightly problematic (with one detected inhomogeneity, as found for many oceanic regions). In some areas a strong contrast between grid points above land and sea is visible, which refers to problems with the data retrieval in Heliosat for certain surface conditions.

The lower left panel of figure 7 illustrates the distributions of the dates of inhomogeneities as detected in the first SNHT testing step. Many inhomogeneities are diagnosed for dates around 1990, which could indicate significant problems with the satellite measurements at this time. This issue will be picked up again in the discussion to figure 9. In the lower right panel of figure 7, displaying the number of positive test results from the three different test procedures (SNHT, Buishand, Pettitt), the high consistency of these methods is emphasized, as in most of the cases the test results of the three procedures agree.

For two grid points with exceptionally high T_0 values the analysis is illustrated in figure 8. For the grid point in Sudan (left two panels) a strong break occurs at 12/1998 (marked red). This and at least one of the eight inhomogeneities (02/1997) detected in the following analysis steps (marked orange) coincides with instrument changes (marked grey). An additional break at 10/1994 is likely to be a result of a changed retrieval algorithm for the Heliosat dataset (Rebekka Posselt, personal communication). The

second grid point in Niger (right) shows a clear break for the instrument change in June 1989. The other four inhomogeneities detected for this location are less pronounced and apparently not connected with further satellite changes. Nevertheless, for both series additional trend-like inhomogeneities within single satellite periods can be observed. This finding indicates the presence of gradual changes related to sensitivity variations of the instruments, even though the retrieval method is supposed to correct such changes (see section 2).

Figure 9 displays the frequency distribution of the dates of the inhomogeneities analyzed in these filtered relative calculations. The findings clearly indicate the presence of breaks associated with instrument changes (marked red) on a global scale. The largest impact is indicated for the change in 04/1990, which has been detected as a date of break in more than 6 % of the total area. Further breaks caused by instrument changes occur for 05/1998, 02/1997 and 06/1989. The aforementioned change of the retrieval algorithm in October 1994 is also found to have a clear impact on the data quality in certain regions (marked with arrow). For the time period between 1988 and 1990 the number of detected inhomogeneities is very high, which could indicate the presence of significant trends in addition to single breaks. This observation refers to considerable instrumental problems appearing for the Meteosat-3 generation (Richard MÃ$\frac{1}{4}$ller, personal communication) used mainly between August 1988 and June 1989 as well as from February to April 1990. From the seven dates of satellite changes only 01/1990 and 01/1994 are indicated to be inconspicious for the whole observation area.

Figure 10 and 11 show the regions with inhomogeneities occuring at dates coinciding with at least one of the seven instrument changes or the retrieval modification at 10/1994. Largest positive breaks of partly more than 40 W/m^2 occured at 05+06/1989 (panel 2, Niger region only) and 05/1998 (panel 8). Especially for 07+08/1988 (panel 1), 10/1994 (panel 5) and 01+02/1997 (panel 7) significant negative shifts of around 10-20 40 W/m^2 are found. While the changes in 05+06/1989 (panel 2) and 03+04/1990 (panel 4) affect various regions of the observation area, the other four break dates are spatially limited to the aforementioned highly problematic regions in Africa. As illustrated in figure 12, up to five of the breaks detected at grid points in these regions can be attributed to alterations in the satellite measurements. These areas typically exhibit mountainous land surfaces without significant vegetation. This could be an indication that the measurement of the reflectance, as used for the determination of the SIS (see section 2), is more problematic for such surfaces.

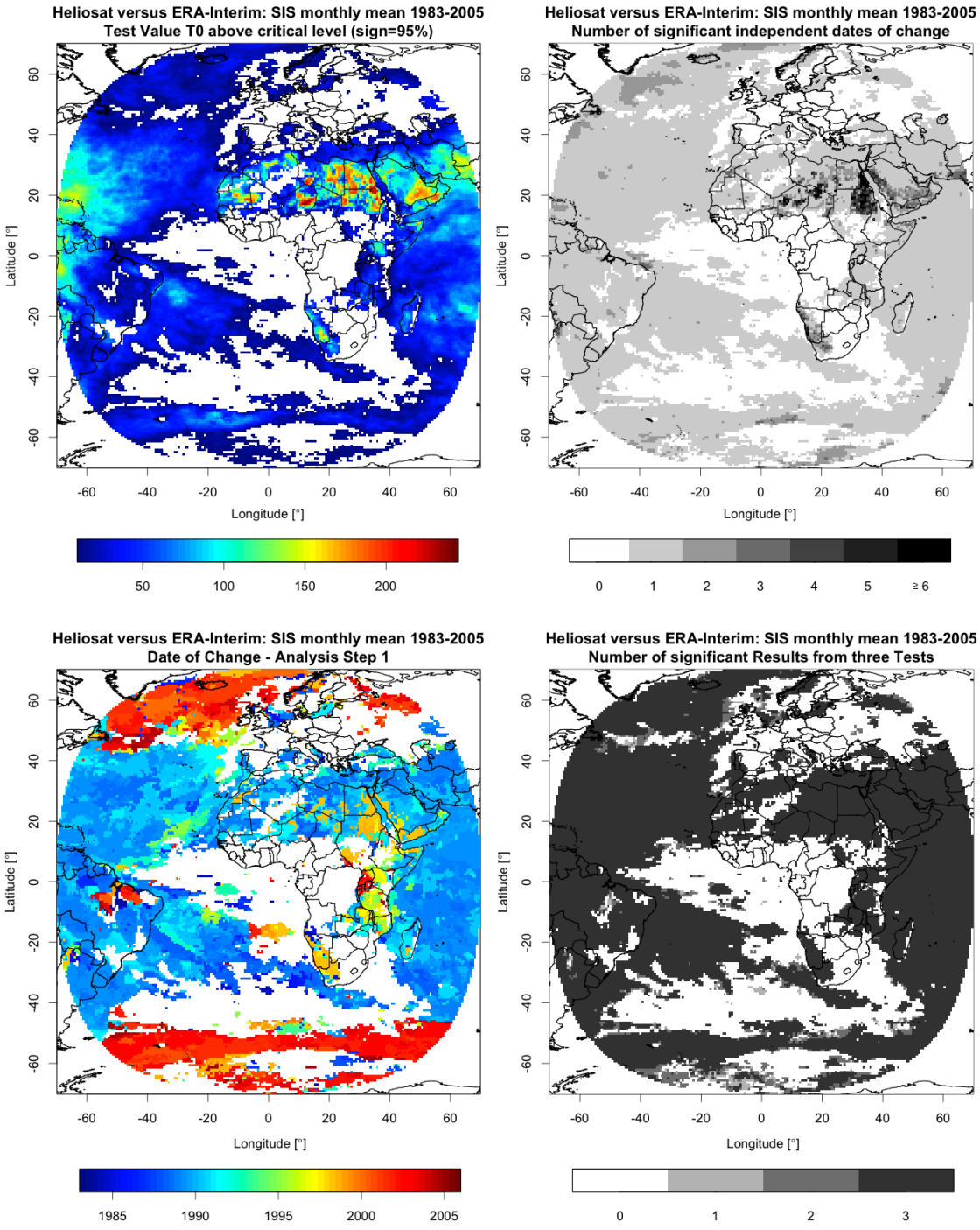

Figure 7: Test values T_0 above significance level of 95% (upper left panel), number of detected inhomogeneities (upper right), the date of change of the first detected break (lower left) and the number of positive results from three inhomogeneity test methods (lower right) for the relative analysis of the Heliosat data fields versus ERA-Interim using filter constraints as mentioned in the text.

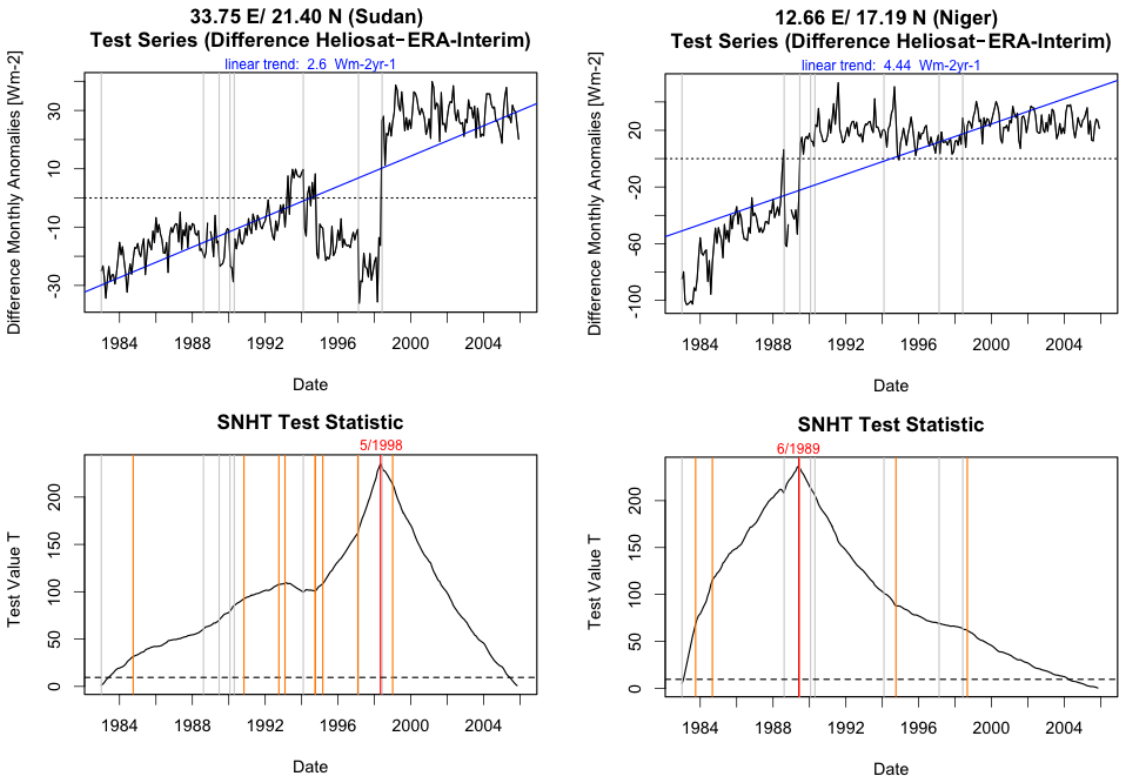

Figure 8: SNHT test statistics for the relative analyses of Heliosat versus ERA-Interim at two grid points with highest T_0 levels over Africa.

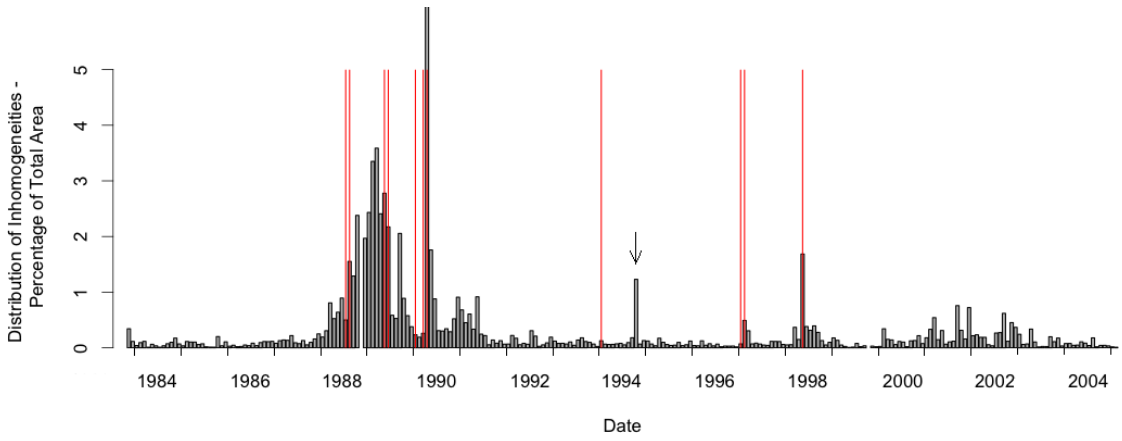

Figure 9: Distribution of the frequency of breaks (in percent of total area) as diagnosed for the Heliosat SIS datafield in comparison with ERA-Interim (relative analysis, filtered). The dates of satellite changes are marked red. At 10/1994 (marked with arrow) the Heliosat retrieval algorithm has been changed.

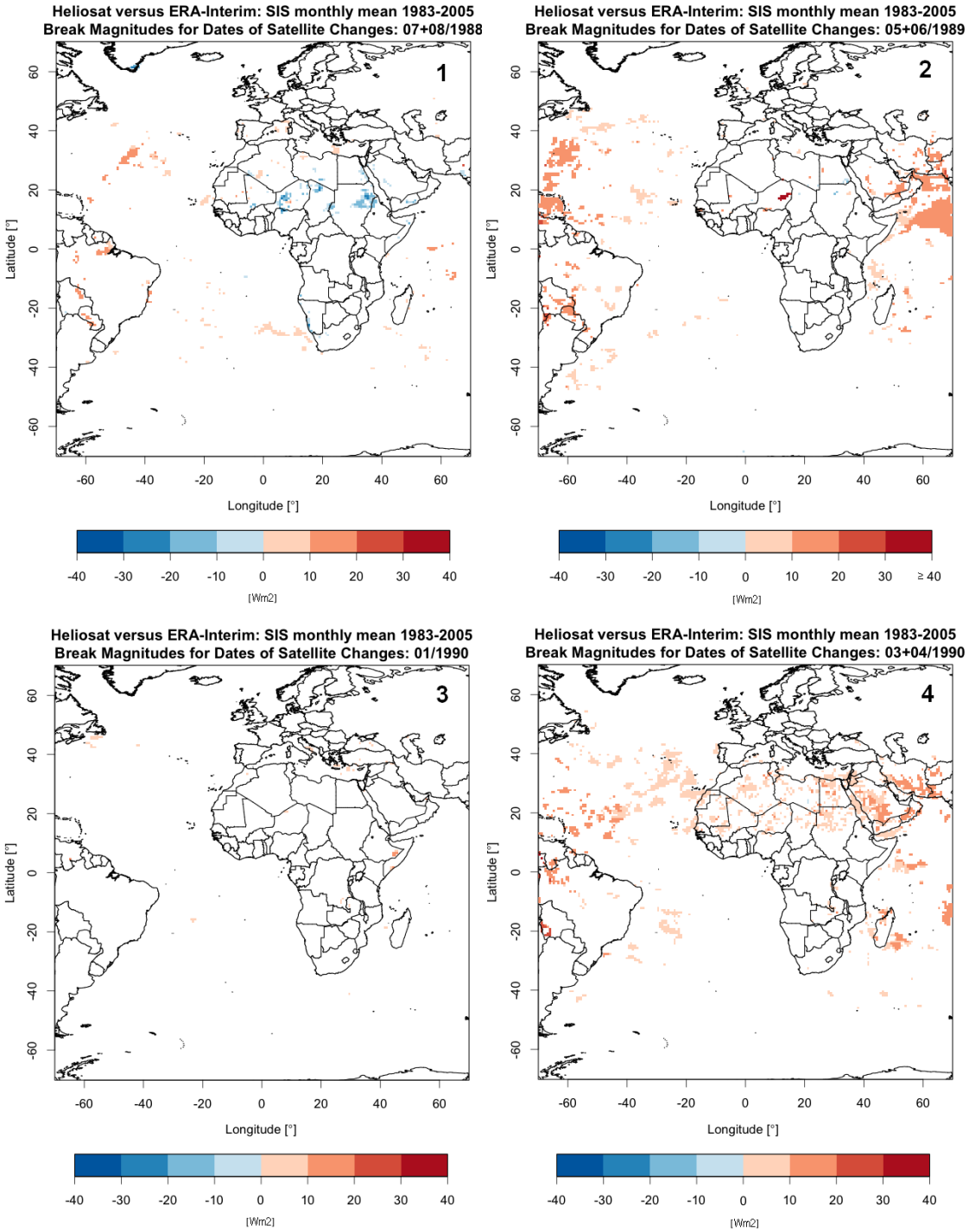

Figure 10: Relative analysis of Heliosat versus ERA-Interim: Break magnitudes for areas with inhomogeneities coinciding with the dates of instrument or retreival changes. (1): 07+08/1988, (2): 05+06/1989, (3): 01/1990, (4): 03+04/1990.

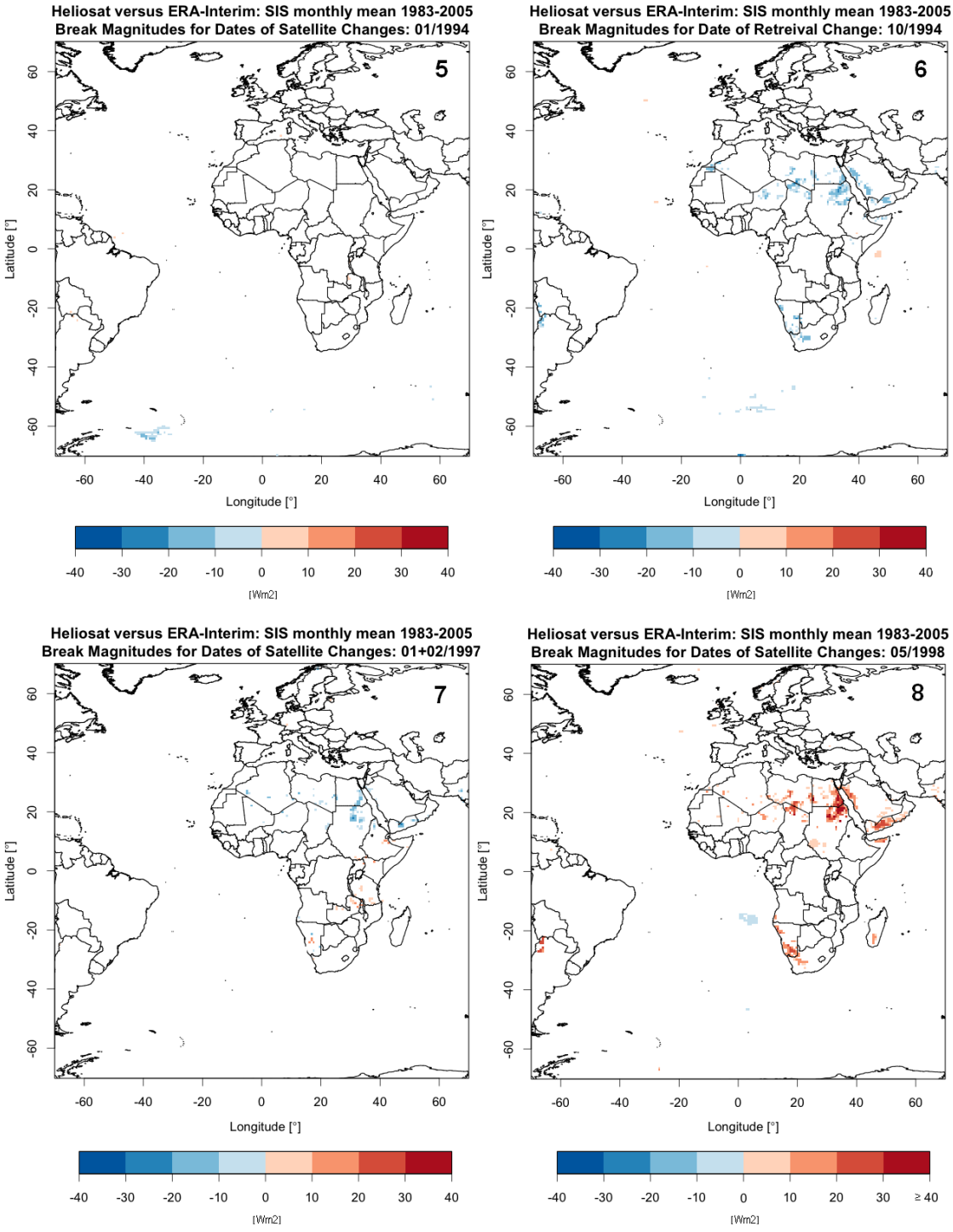

Figure 11: Relative analysis of Heliosat versus ERA-Interim: Break magnitudes for areas with inhomogeneities coinciding with the dates of instrument or retreival changes. (5): 01/1994, (6): 10/1994, (7): 01+02/1997, (8): 05/1998.

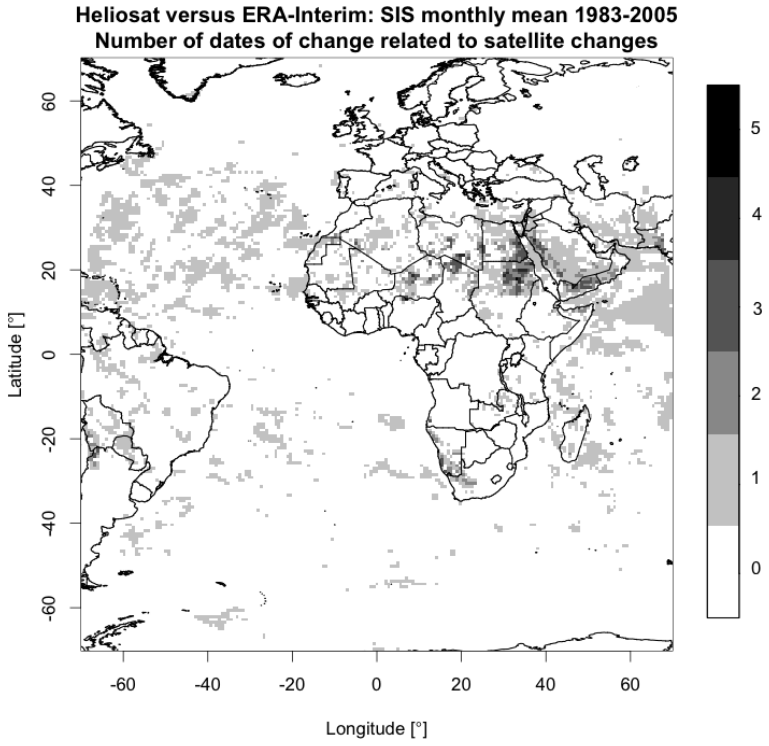

Figure 12: Relative analysis of Heliosat versus ERA-Interim: Number of breaks coinciding with the dates of instrument or retreival changes.

7 Conclusions

The investigation of the Heliosat SIS datafield shows for the majority of the regions no shifts related to satellite changes. Nevertheless, significant discrepancies between the satellite data and the BSRN surface measurement data as well as the ERA-Interim data are apparent for many regions. Most of these inhomogeneities are indicated to be a result of relative trends between candidate and reference time series. As many of these discontinuities occur in the time period between 1988 and 1990, when the problematic Meteosat-3 instrument generation was used in two intervals, some of these inconsistencies are possibly related to changes of the instrument's sensitivity.

Instrument and retreival changes have led to break-like inhomogeneities in certain regions. Especially for narrowly confined areas in Africa a multitude of such shifts has been detected. Its typically mountainous surfaces without vegetation refer to possible problems of the satellite measurements in combination with the given data retreival for such specific surface conditions. This dependancy of the data quality on surface characteristics is also emphasized by the observation of sharp contrasts in the homogeneity distribution between land and sea surfaces in certain areas.

While most of the detected breaks can be well attributed to changes in the instrumentation, the relative trends observed between the Heliosat and the reference data are not necessarily a result of the inaccuracy of satellite measurements. Such gradual changes, especially with respect to surface measurements might also be introduced by the direct aerosol effect on SIS, as this direct impact by changing aerosol abundances would be included in the surface time series but not in the satellite series based on cloud albedo measurements and with the assumption of a constant aerosol climatology. More specific homogeneity investigations are needed to evaluate if these trend-like inhomogeneities have instrumental or real climatological origins.

Appendix

Buishand Range Test

The Buishand range test (Buishand, 1982) considers the partial sums S_k^*, which describe the cumulative deviations from the mean of the time series.

$$S_k^* = \sum_{i=1}^{k} (Y_i - \bar{Y}) \quad k = 1, ..., n \tag{5}$$

For a homogeneous series without a trend S_k^* is expected to fluctuate around zero. In case of a break at date k_0, S_k^* will show an extremum around this date. Significance is given, if the 'rescaled adjusted range' R, defining the difference between minimum and maximum of S_k^* (normalized to the standard deviation s of the data), exceeds a certain level.

$$R = (\max S_k^* - \min S_k^*)/s \tag{6}$$

The critical values R_c (as shown in tables 2 to 4) are derived from simulations with 10.000 normally distributed random time series.

Pettitt Test

In this test (Pettitt, 1979) ranks r_1,...,r_n are determined for the values of the time series Y_1,...,Y_n. The test statistics is given as:

$$X_k = 2 \sum_{i=1}^{k} r_i - k(n+1) \quad k = 1, ..., n \tag{7}$$

X_k will vary around zero for a normally distributed time series. If the absolute value of X_k peaks above a critical level X_c for date k_0, a break around this date is detected.

$$X_{k0} = \max |X_k| \tag{8}$$

Again, we conducted 10.000 simulations to derive critical values for the different lengths of time series.

Additional Tables and Figures

Table 5: Results of relative homogeneity testing of ERA-Interim time series (versus BSRN) at the grid points nearest to corresponding BSRN stations. In the last column are listed the corresponding results versus ASRB Payerne.

Station Start Date	CAR 9/1996	FLO 7/1994	BER 1/1992	LIN 10/1994	PAY 10/1992	PAY ASRB 1/1995
SNHT:						
T_0/T_c	7.53/8.53	**12.55**/8.71	1.04/8.92	3.55/8.67	5.00/8.82	2.95/8.67
Date of Break		5/1999				
Buishand:						
R/R_c	1.21/1.62	**2.21**/1.63	0.83/1.63	1.11/1.63	1.34/1.63	1.28/1.63
Date of Break		5/1999				
Pettitt:						
X_{k0}/X_c	641/887	**1625**/1220	551/1649	442/1132	1310/1520	677/1147
Date of Break		5/1999				
Linear Trend						
$[Wm^{-2}yr^{-1}]$	-0.30	0.58	0.00	-0.03	-0.26	0.00
Mean Deviation						
$[Wm^{-2}]$	4.4	8.6	7.1	5.0	7.4	7.3

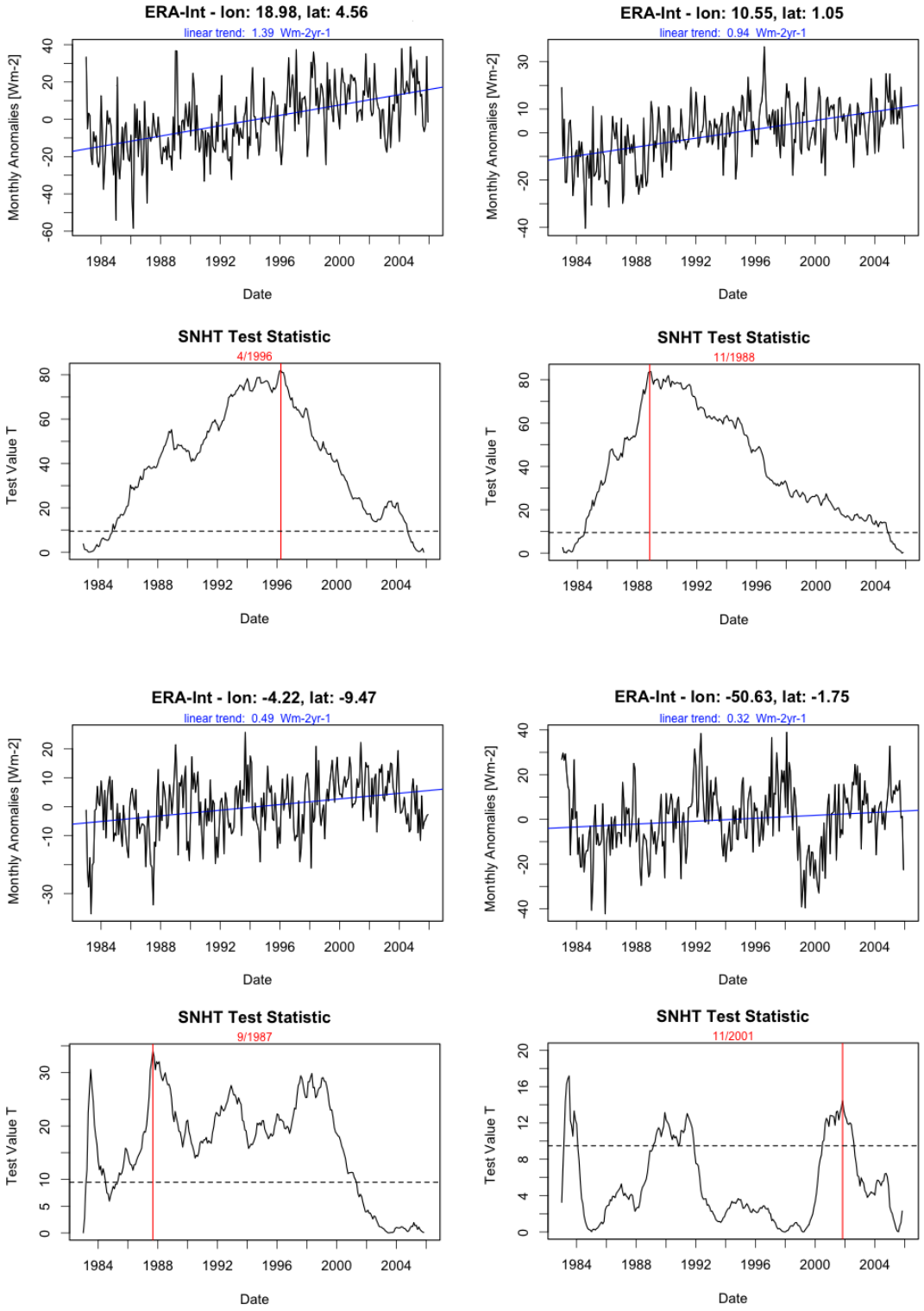

Figure 13: SNHT test statistics for the absolute analyses of ERA-Interim SIS at two grid points with highest T_0 levels over Africa (upper panels) and at two grid points with largest numbers of breaks, above the Atlantic (lower left panels) and in Brazil (lower right panels).

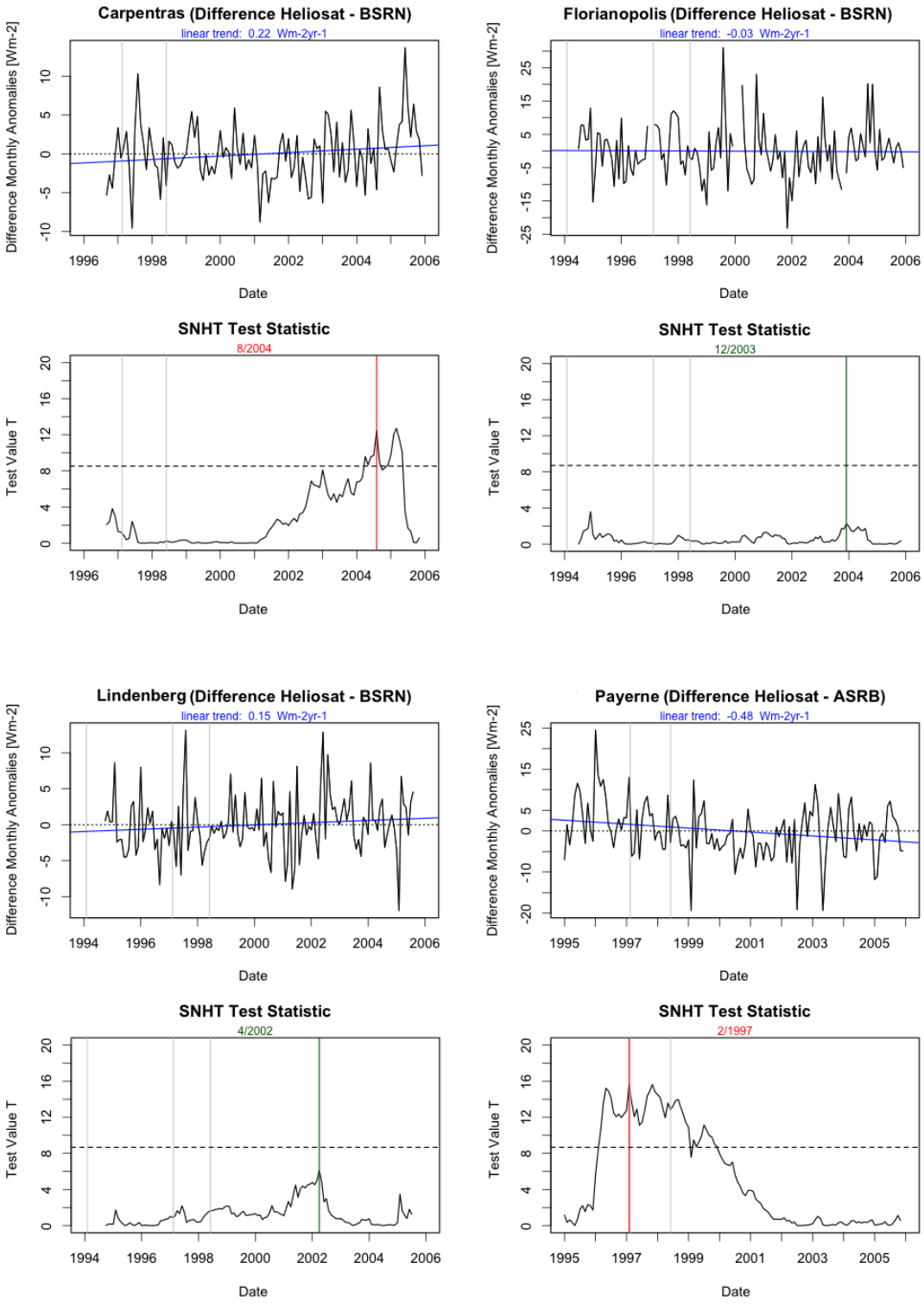

Figure 14: SNHT test statistics for the relative analyses Heliosat versus BSRN for Carpentras, Florianopolis (upper panels), Lindenberg (lower right panels) and Payerne (versus ASRB, lower right panels).

References

Alexandersson, H.: A homogeneity test applied to precipitation data, International Journal of Climatology, 6, 661–675, doi:10.1002/joc.3370060607, 1986.

Alexandersson, H. and Moberg, A.: Homogenization of Swedish Temperature Data. Part i: Homogeneity Test for Linear Trends, International Journal of Climatology, 17, 25–34, doi:10.1002/(SICI)1097-0088(199701)17:1¡25::AID-JOC103¿3.3.CO;2-A, 1997.

Beyer, H. G., Costanzo, C., and Heinemann, D.: Modifications of the Heliosat procedure for irradiance estimates from satellite images, Solar Energy, 56, 207–212, 1996.

Buishand, T.: Some methods for testing the homogeneity of rainfall records, Journal of Hydrology, 58, 11–27, doi:10.1016/0022-1694(82)90066-X, 1982.

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

Dee, D. P. and Uppala, S.: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis, Quart. J. Roy. Meteorol. Soc., 135, 1830–1841, 2009.

Dee, D. P., Uppala, S. M., Simmons, A. J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M. A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A. C. M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A. J., Haimberger, L., Healy, S. B., Hersbach, H., Hólm, E. V., Isaksen, L., Kållberg, P., Köhler, M., Matricardi, M., McNally, A. P., Monge-Sanz, B. M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., and Vitart, F.: The ERA-Interim reanalysis: configuration and performance of the data assimilation system, Quarterly Journal of the Royal Meteorological Society, 137, 553–597, doi: 10.1002/qj.828, 2011.

Hammer, A., Heinemanna, D., Hoyer, C., Kuhlemann, R., Lorenz, E., Müller, R., and Beyer, H. G.: Solar energy assessment using remote sensing technologies, Remote Sens. Environ., 86, 423–432, 2003.

IPCC: Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Cambridge University Press, Cambridge and New York, 2007.

Lee, T. C. K., Zwiers, F. W., Hegerl, G. C., Zhang, X., and Tsao, M.: A Bayesian Climate Change Detection and Attribution Assessment., Journal of Climate, 18, 2429–2440, doi:10.1175/JCLI3402.1, 2005.

Mayer, B. and Kylling, A.: Technical note: The libRadtran software package for radiative transfer calculations - description and examples of use, Atmospheric Chemistry & Physics, 5, 1855–1877, 2005.

Mueller, R. W., Matsoukas, C., Gratzki, A., Behr, H. D., and Hollmann, R.: The CM-SAF operational scheme for the satellite based retrieval of solar surface irradiance - A LUT based eigenvector hybrid approach, Remote Sens. Environ., 113, 1012–1024, 2009.

Ohmura, A., Gilgen, H., Hegner, H., Müller, G., Wild, M., Dutton, E. G., Forgan, B., Fröhlich, C., Philipona, R., Heimo, A., König-Langlo, G., McArthur, B., Pinker, R., Whitlock, C. H., and Dehne, K.: Baseline Surface Radiation Network (BSRN/WCRP): New Precision Radiometry for Climate Research., Bulletin of the American Meteorological Society, 79, 2115–2136, doi:10.1175/1520-0477(1998)079¡2115:BSRNBW¿2.0.CO;2, 1998.

Peterson, T. C., Easterling, D. R., Karl, T. R., Groisman, P., Nicholls, N., Plummer, N., Torok, S., Auer, I., Boehm, R., Gullett, D., Vincent, L., Heino, R., Tuomenvirta, H., Mestre, O., Szentimrey, T., Salinger, J., Førland, E. J., Hanssen-Bauer, I., Alexandersson, H., Jones, P., and Parker, D.: Homogeneity adjustments of in situ atmospheric climate data: a review, International Journal of Climatology, 18, 1493–1517, doi:10.1002/(SICI)1097-0088(19981115)18:13¡1493::AID-JOC329¿3.0.CO;2-T, 1998.

Pettitt, A. N.: A non-parametric approach to the change-point problem, Applied Statistics, 28, 126–135, doi:10.2307/2346729, 1979.

Posselt, R., Mueller, R., Stöckli, R., and Trentmann, J.: Spatial and Temporal Homogeneity of Solar Surface Irradiance across Satellite Generations, Remote Sensing, 3, 1029–1046, doi:10.3390/rs3051029, 2011.

Rigollier, C., Lefèvre, M., Blanc, P., and Wald, L.: The Operational Calibration of Images Taken in the Visible Channel of the Meteosat Series of Satellites, J. Atmos. Ocean. Technol., 19, 1285–1293, 2002.

Simmons, A. J., Willett, K. M., Jones, P., Thorne, P. W., and Dee, D. P.: Low-frequency variations in surface atmospheric humidity, temperature, and precipitation: Inferences from reanalyses and monthly gridded observational data sets, JGR, 115, D01110, doi: 10.1029/2009JD012442, 2010.

Frankfurt, den 29.08.2011

Sven Brinckmann Bodo Ahrens