EUMETSAT Satellite Application Facility on Climate Monitoring

Validation Report

CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 3 (CLARA-A3)

Top-of-Atmosphere Radiation

	TCDR	ICDR
TOA Reflected Solar Flux (RSF)	**CM-11312**	**CM-6331**
TOA Outgoing Longwave Radiation (OLR)	**CM-11342**	**CM-6321**

Reference Number:	SAF/CM/RMIB/VAL/GAC/TOA
Issue/Revision Index:	1.1
Date:	06.02.2023

Document Signature Table

	Name	**Function**	**Signature**	**Date**
Author	Tom Akkermans Nicolas Clerbaux	CM SAF scientist CM SAF scientist		06/02/2023
Editor	Marc Schröder	CM SAF Science Coordinator		06/02/2023
Approval	CM SAF Steering Group			
Release	Rainer Hollmann	CM SAF Project Manager		

Distribution List

Internal Distribution	
Name	**No. Copies**
DWD / Archive	1
CM SAF Team	1

External Distribution		
Company	**Name**	**No. Copies**
Public		1

Document Change Record

Issue/ Revision	**Date**	**DCN No.**	**Changed Pages/Paragraphs**
1.0	03.10.2022		Initial version for joint DRR3.2/ORR
1.1	06.02.2023		RIDs from DRR 3.2 implemented Adapted ICDR processing

Applicable Documents

The following documents, of the exact issue shown, form part of this document to the extent specified herein. Applicable documents are those referenced in the Contract or approved by the Approval Authority. They are referenced in this document in the form [AD X].

Reference	Title	Code, Version, Date
AD 1	CM SAF CDOP3 Project Plan	SAF/CM/DWD/CDOP3_PP_1_7, Version 1.7, 08/12/2021
AD 2	CM SAF CDOP3 Product Requirement Document	SAF/CM/DWD/PRD, Version 4.1

Reference Documents

The reference documents contain useful information related to the subject of the project. These reference documents complement the applicable ones, and can be looked up to enhance the information included in this document if it is desired. They are referenced in this document in the form [RD X].

Reference	Title	Code, Version, Date
RD 1	CM SAF Requirements Review: AVHRR GAC Edition 3 data records (CLARA Ed.3)	SAF/CM/CDOP3/SMHI/RR32, Version 1.2, 08/05/2020
RD 2	ATBD CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 3 (CLARA-A3): Cloud Products processing chain (level-1 – level-2/2b – level-3).	SAF/CM/DWD/ATBD/GAC v3.3
RD 3	Algorithm Theoretical Basis Document for Cloud Micro Physics of the NWC/PPS	NWC/CDOP3/PPS/SMHI/SCI/ATBD/CMIC, Issue 3.0, 26/04/2021
RD 4	Algorithm Theoretical Basis Document for the Cloud Probability of the NWC/PPS	NWC/CDOP3/PPS/SMHI/SCI/ATBD/CloudProbability, Issue 2.0, 26/04/2021
RD 5	ATBD CM SAF Cloud, Albedo, Radiation data record, AVHRR-based, Edition 3 (CLARA-A3): Surface Radiation	SAF/CM/DWD/ATBD/CLARA/RAD, Issue 3.1
RD 6	The Data Set Generation Capability Description Document, AVHRR GAC Edition 3 (CLARA-A3)	SAF/CM/DWD/DGCDD/GAC/3, version 3.3
RD 7	Algorithm Theoretical Basis Document CLARA Edition 3: TOA Radiation	SAF/CM/RMIB/ATBD/GAC/TOA
RD 8	Product User Manual CLARA Edition 3: TOA Radiation	SAF/CM/RMIB/PUM/GAC/TOA

Reference	Title	Code, Version, Date
RD 9	Algorithm Theoretical Basis Document CLARA Edition 3: Cloud Products (level-1 to level-3)	SAF/CM/DWD/ATBD/CLARA/CLD

Table of Contents

List of Tables

List of Figures

Executive Summary

This CM SAF report provides information on the validation of the Top-Of-Atmosphere (TOA) radiation products from the CM SAF CLARA Edition 3.0 (CLARA-A3) data records derived from AVHRR sensors onboard the series of NOAA and MetOp satellites. More specifically, it presents the validation of TOA Reflected Solar Flux [CM-11312, RSF, Section 5] and TOA Outgoing Longwave Radiation [CM-11342, OLR, Section 6], available from 1979 to 2020. The algorithm theoretical basis document (ATBD) describes the individual parameter algorithms [RD 7]. Both products (RSF, OLR) are validated against available satellite-based reference data records. Regional uncertainty combines accuracy and precision, and is defined by the mean absolute bias derived from validation with the reference data. Stability is defined by the variability in mean bias with the reference data. Both validation metrics are evaluated against the requirements as given in the product requirements document (PRD) [AD 2].

All data records fulfil the most basic requirements (*threshold requirements*) as specified in the Product Requirements Document (PRD) [AD 2]. Table 1-1 and Table 1-2 summarize the validation results of the Daily Mean (DM) and Monthly Mean (MM) RSF and OLR products in view of these *threshold requirements*, including the relative duration this requirement was met (cfr last column).

Table 1-1: Validation results for regional uncertainty (accuracy+precision) during 1979-2020

CLARA-A3 Parameter		Average MAB (W/m²)	Reference data record	Threshold requirement W/m² (% met)		
RSF	MM	**3.2**	CERES	**8**	✓	99.3%
	DM	**9.0**	CERES	**16**	✓	94.5%
OLR	MM	**1.8**	HIRS	**8**	✓	99.8%
	DM	**4.8**	HIRS	**16**	✓	99.6%

Table 1-2: Validation results for stability during 1979-2020

CLARA-A3 Parameter		Reference data record	Threshold requirement W/m² (% met)		
RSF	MM	ERA5	**4**	94.0%	✓
OLR	MM	HIRS	**4**	99.6%	✓
OLR	MM	HIRS-MM	**4**	98.0%	✓

The more demanding requirements (called *target* and *optimal requirements*) are met either for some products, or during some periods, or with regard to some reference records. The ICDR proves to be almost identical to the TCDR during their overlapping time span, being the second half of 2020. All details are highlighted in this report. The requirements are defined in the product requirements document (PRD) [AD 2].

1 The EUMETSAT SAF on Climate Monitoring

The importance of climate monitoring with satellites was recognized in 2000 by EUMETSAT Member States when they amended the EUMETSAT Convention to affirm that the EUMETSAT mandate is also to "contribute to the operational monitoring of the climate and the detection of global climatic changes". Following this, EUMETSAT established within its Satellite Application Facility (SAF) network a dedicated centre, the SAF on Climate Monitoring (CM SAF, http://www.cmsaf.eu).

The consortium of CM SAF currently comprises the Deutscher Wetterdienst (DWD) as host institute, and the partners from the Royal Meteorological Institute of Belgium (RMIB), the Finnish Meteorological Institute (FMI), the Royal Meteorological Institute of the Netherlands (KNMI), the Swedish Meteorological and Hydrological Institute (SMHI), the Meteorological Service of Switzerland (MeteoSwiss), and the Meteorological Service of the United Kingdom (UK MetOffice). Since the beginning in 1999, the EUMETSAT Satellite Application Facility on Climate Monitoring (CM SAF) has developed and will continue to develop capabilities for a sustained generation and provision of Climate Data Records (CDR's) derived from operational meteorological satellites.

In particular the generation of long-term data sets is pursued. The ultimate aim is to make the resulting data sets suitable for the analysis of climate variability and potentially the detection of climate trends. CM SAF works in close collaboration with the EUMETSAT Central Facility and liaises with other satellite operators to advance the availability, quality and usability of Fundamental Climate Data Records (FCDRs) as defined by the Global Climate Observing System (GCOS). As a major task the CM-SAF utilizes FCDRs to produce records of Essential Climate Variables (ECVs) as defined by GCOS. Thematically, the focus of CM SAF is on ECVs associated with the global energy and water cycle.

Another essential task of CM SAF is to produce data sets that can serve applications related to the new Global Framework of Climate Services initiated by the WMO World Climate Conference-3 in 2009. CM SAF is supporting climate services at national meteorological and hydrological services (NMHSs) with long-term data records but also with data sets produced close to real time that can be used to prepare monthly/annual updates of the state of the climate. Both types of products together allow for a consistent description of mean values, anomalies, variability and potential trends for the chosen ECVs. CM SAF ECV data sets also serve the improvement of climate models both at global and regional scale.

As an essential partner in the related international frameworks, in particular WMO SCOPE-CM (Sustained COordinated Processing of Environmental satellite data for Climate Monitoring), the CM SAF - together with the EUMETSAT Central Facility, assumes the role as main implementer of EUMETSAT's commitments in support to global climate monitoring. This is achieved through:

- Application of highest standards and guidelines as lined out by GCOS for the satellite data processing,
- Processing of satellite data within a true international collaboration benefiting from developments at international level and pollinating the partnership with own ideas and standards,
- Intensive validation and improvement of the CM SAF climate data records,
- Taking a major role in data set assessments performed by research organisations such as WCRP. This role provides the CM SAF with deep contacts to research organizations that form a substantial user group for the CM SAF CDRs,

- Maintaining and providing an operational and sustained infrastructure that can serve the community within the transition of mature CDR products from the research community into operational environments.

A catalogue of all available CM SAF products is accessible via the CM SAF webpage, www.cmsaf.eu/. Here, detailed information about product ordering, add-on tools, sample programs and documentation is provided.

2 Introduction

2.1 TOA radiative fluxes in CM SAF

At the Top-Of-Atmosphere (TOA), the following radiative fluxes are defined: the Incoming Solar Radiation (ISR), the Reflected Solar Flux (RSF) and the Outgoing Longwave Radiation (OLR), as illustrated in Figure 2-1.
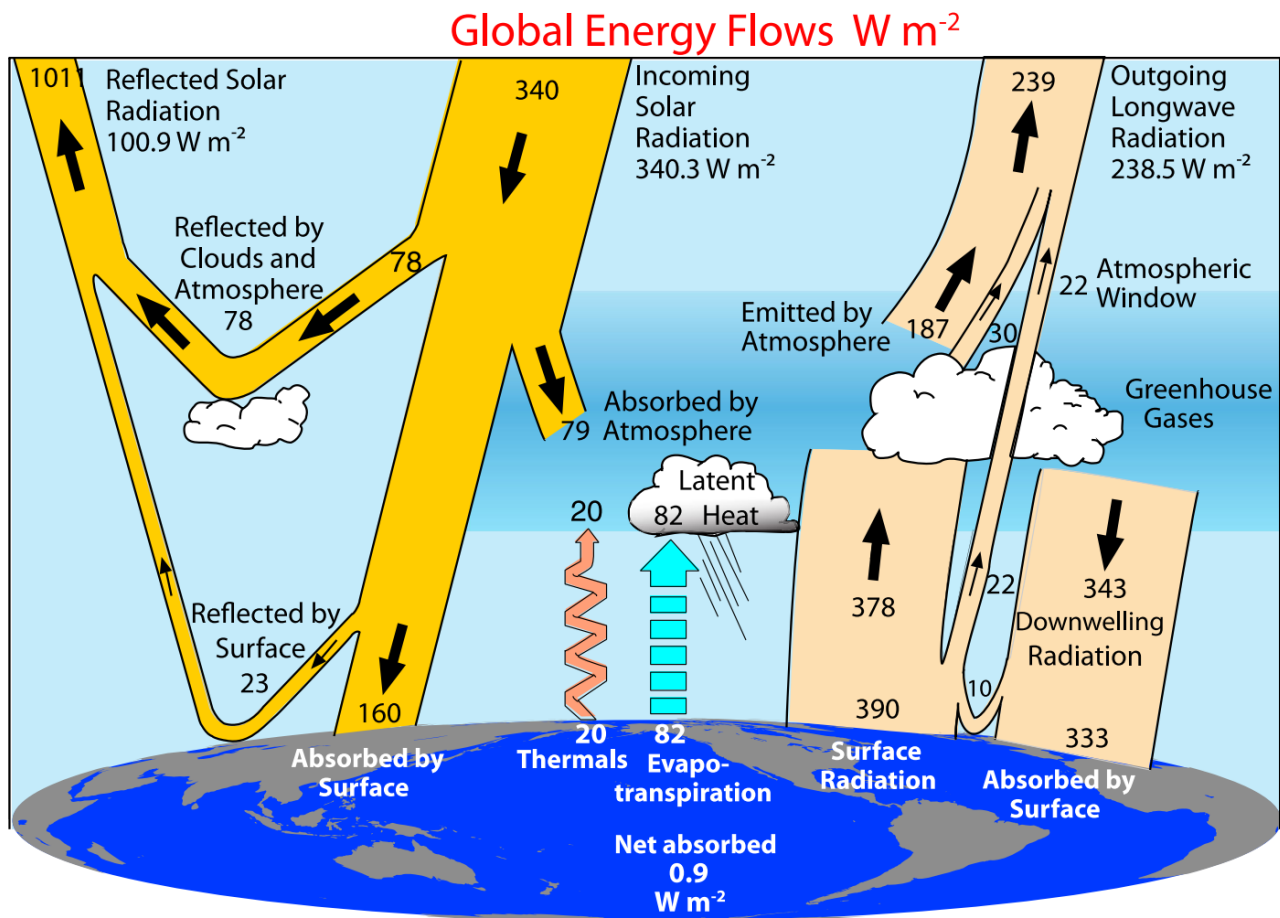
Figure 2-1: An overall schematic of the global annual mean energy flows through the climate system has at the top of atmosphere (TOA) the Incoming and Reflected Solar Radiation, and the Outgoing Longwave Radiation (Trenberth 2020).

These three components of the Earth Radiation Budget (ERB) are the driver of the climate on our planet. In the frame of climate monitoring, the continuous monitoring of these fluxes is of prime importance to understand climate variability and change. The nature of these quantities, which are defined at TOA, makes the use of satellite observations especially useful.

A full global coverage of broadband observations is provided by the Clouds and the Earth's Radiant Energy System (CERES) instruments and derived products (Loeb et al., 2018). Although these are acknowledged to be the golden standard w.r.t. radiative flux data records, two limitations can be identified: (1) the products are relatively recent, e.g. starting in year 2000 for the EBAF product, and (2) the products have a relatively coarse spatial resolution of 1°x1° (lat-lon equal angle grid). The CLARA-A3 TOA RSF and OLR products developed within CM SAF aim to bridge these gaps,

respectively by (1) a prolongation back in time to the late 1970s and (2) by increasing the spatial resolution to 0.25°x0.25°. A third advantage of the new CDRs lies in their synergy and compatibility with the other CDRs from the CM SAF CLARA product family (cloud mask and other cloud parameters, surface radiation, surface albedo, etc.) sharing common algorithms and processing chains.

The two new CLARA-A3 TOA data records are generated with the following CM SAF identifiers:

Content	CM SAF identifier	
	TCDR	*ICDR*
TOA Reflected Solar Flux (RSF)	CM-11312	CM-6331
TOA Outgoing Longwave Radiation (OLR)	CM-11342	CM-6321

After a brief summary of the user requirements (Section 2.2) and of the processing system (Section 2.3), the validation reference data records are presented in Section 3, and the validation methodology is presented in section 4. Then, Sections 5 and 6 provide the detailed validations of respectively the RSF and OLR products. In view of providing the relevant information to the users (through the Product User Manual [RD 8]), a summary table and some conclusions are given in Section 7. Then Sections 8 and 9 provide respectively references and acronyms. Finally, Section 10 contains annexes to the document: these provide complementary information (for advanced users) and are not necessary to read and understand the actual report.

2.2 Summary of user requirements

Basic requirements are defined in the Product Requirements Document [AD 2] and further specified and clarified in the Requirements Review document [RD 1]. They are repeated here as follows:

Table 2-1: Accuracy requirements for TOA radiation products

Identifier	Parameter	Temporal resolution	Accuracy (mean absolute bias) (W/m²)		
			Threshold	*Target*	*Optimal*
CM-11312	RSF	Monthly mean	8	4	2
		Daily mean	16	8	4
CM-11342	OLR	Monthly mean	8	4	2
		Daily mean	16	8	4

Table 2-2: Stability requirements for TOA radiation products

Identifier	Parameter	Temporal resolution	Decadal stability (W/m²/decade)		
			Threshold	*Target*	*Optimal*
CM-11312	RSF	Monthly mean	4	0.6	0.3
CM-11342	OLR	Monthly mean	4	0.6	0.3

The accuracy requirement is defined by calculating the spatial uncertainty, here expressed as the global Mean Absolute Bias (MAB) between the CLARA-A3 product and some reference data records, all re-gridded on the 1°x1° CERES Nested Processing Grid (i.e., global). From this, a time series is

created to track the temporal variability of the performance, which is quantified by the MAB remaining below the predefined requirement thresholds. The stability requirement is defined in the same way, except that here the global Mean Bias (MB) is used as parameter, and that the performance is quantified by the variability in MB remaining within the predefined requirement range (also called "stability envelope").

2.3 Summary of retrievals

A detailed description of the retrievals was provided in [RD 7], and also published by Clerbaux et al. (2020) for OLR, and by Akkermans and Clerbaux (2021) for RSF. A schematic of the CLARA-A3 TOA fluxes processing chain is given in Figure 2-2.

The panels on the right (bluish colors) describe the RSF retrieval. First, the narrowband reflectances are converted to broadband reflectance using empirical regressions with the Clouds and the Earth's Radiant Energy System (CERES) observations, using a large database of collocated, coangular, and simultaneous AVHRR-CERES observations (i.e. spectral conversion). Second, the anisotropy is corrected by applying Angular Distribution Models (ADMs), which convert directional reflectance into a hemispherical albedo (i.e. directional-hemispherical conversion). Third, the instantaneous albedos are temporally interpolated by a flexible diurnal cycle model, capable of ingesting any number of observations at any time of day, making it suitable for any orbital configuration of NOAA and MetOp satellites. Finally, the twilight conditions (when SZA>84°) prevailing near sunrise and sunset are simulated with an empirical model (not shown in figure). The entire day is then integrated into a single daily mean RSF, and subsequently in monthly mean RSF.
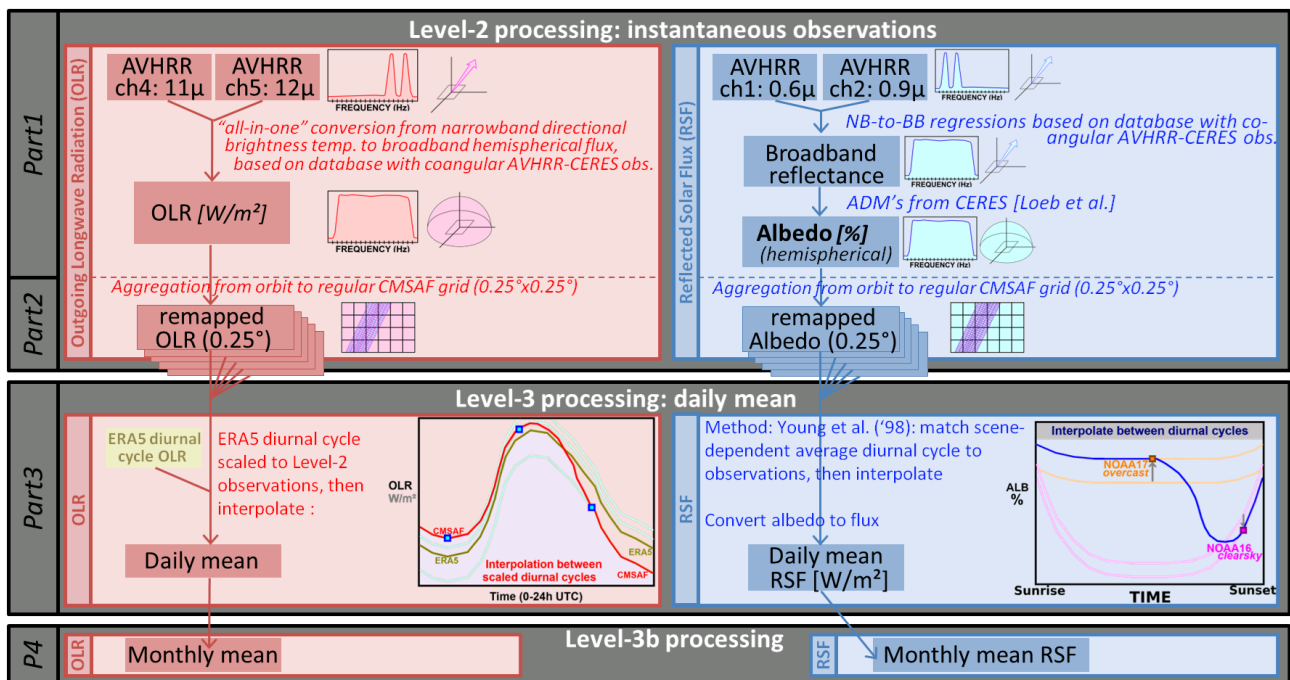
Figure 2-2: Overview of the OLR (left) and RSF (right) processing chain

The panels on the left (reddish colors) describe the OLR retrieval. A first step is the estimation of the instantaneous OLR from the AVHRR observations. This is done by regressions on the same large database of collocated AVHRR-CERES observations (as used for the RSF), but in contrast to RSF,

the OLR regressions are an "all-in-one" conversion which combines the spectral and directional components. A second step concerns the estimation of daily and monthly OLR from the instantaneous AVHRR overpasses. Over clear sky land, the OLR from ERA5 reanalysis is used to estimate the diurnal variation; otherwise, linear regression is applied.

2.4 Generated Output

Because the different CLARA-A3 products use a different terminology for their intermediate processing chains, the TOA processing levels are presented in Table 2-3 to avoid confusion.

Table 2-3: Processing levels of CLARA-A3 TOA flux products RSF and OLR

Processing level	Description	Resulting spatiotemporal and orbital resolution	Resulting product's release info
Level-1c	The PPS output containing geolocated and calibrated reflectance and radiances (5 channels) i.e. the AVHRR FDR	Instantaneous, satellite-specific, on *GAC orbit grid*** (1 file per satellite and per orbit**)	Intermediate product, not publicly released
Level-2	Part 1*: retrieval algorithm of geophysical variables shortwave **TOA albedo** and longwave **TOA OLR**	Instantaneous, satellite-specific, on *GAC orbit grid*** (1 file per satellite and per orbit***)	
Level-2b****	Part 2*: Spatial aggregation from *GAC orbit grid* to a global equal-area 0.25°x0.25° grid	Instantaneous, satellite-specific, on a global 0.25°x0.25° grid (1 file per satellite and per orbit***). *Pixels outside orbital swath are marked invalid.*	
Level-3 (daily)	Part 3*: Processing of (SW) instantaneous albedo to daily mean RSF, or (LW) instantaneous OLR to daily mean OLR	Daily mean composite including orbits from all satellite overpasses during that day, on a global 0.25°x0.25° grid (1 file per day). *Full global coverage.*	Released as part of CLARA-A3 products CM-11312 (RSF) and CM-11342 (OLR)
Level-3b (monthly)	Part 4*: Temporal aggregation to monthly mean RSF and OLR	Monthly mean composed from all daily means, on a global 0.25°x0.25° grid (1 file per month). *Full global coverage.*	
(*) The different components of the software are referred to as 'program parts 1 to 4' [RD 7] *(**) GAC orbit grid: the native grid from the PPS products that are used as input for the TOA retrieval; same as FDR grid (is it?); spatial resolution is roughly 4km* *(***) orbit: roughly 1 orbital period of data, with possibly slight overlap between start and end* *(****) Could be misleading: according to many classifications this would be considered L3*			

3   Reference Data Records used in the validation

The following data records are used for the validation (all have global spatial coverage).

Table 3-1: Data records used for comparison and validation of CLARA-A3

Name	Variable	Period	Temporal resolution	Spatial resolution	Suitable for. (used for.)
CERES EBAF Ed.4.1	OLR, RSF	2000/03-2020/12	Monthly	1°x1° (nested grid)	Spatial pattern validation (regional uncertainty); + Long term global mean validation (stability)
CERES SYN1deg Ed.4.1	OLR, RSF	2000/03-2020/12	Monthly, daily, monthly diurnal cycle	1°x1° (nested grid)	
HIRS OLR Daily v01r02 ("*HIRS*")	OLR	1979/12-2020/12	Daily, *Monthly (derived by aggregation)*	1°x1°	
HIRS OLR Monthly v02r07 ("*HIRS-MM*")	OLR	1979/12-2020/12	Monthly	2.5°x2.5°	Long term global mean *validation* (stability)
ERA5 (reanalysis)	OLR, RSF	1979/12-2020/12	Monthly, daily	0.25°x0.25°	
ISCCP FH MPF	OLR, RSF	1983/07-2017/06	Monthly	1°x1°	Long term global mean
Cloud_CCI L3C AVHRR-PM v3.0	OLR, RSF	1982/01-2020/12	Monthly	0.5°x0.5°	*comparison only (no validation)*

3.1   CERES SYN1deg Ed.4.1 (daily and monthly)

The CERES SYN1deg products Ed4.1 provide estimates of the daily and monthly mean RSF and OLR fluxes from March 2000 onward at a 1°x1° latitude-longitude resolution (in fact, on a *1°x1° nested grid*, see Section 4.3). The products consist of CERES-observed, geostationary enhanced and temporally interpolated TOA radiative fluxes. Given the sun-synchronous orbits of the CERES instruments onboard the Aqua and Terra satellites, the observations are performed only twice a day. Therefore, hourly TOA fluxes and cloud properties from five contiguous geostationary imagers, covering 60°S–60°N at any given time, are used for an improved modelling of the diurnal variability between the CERES observations (Doelling et al., 2013).

While the SYN1deg approach provides improved diurnal coverage by merging CERES and 1-hourly geostationary (GEO) data, artifacts in the GEO imager visible bands over certain regions and time periods can introduce larger regional uncertainties. Spurious jumps in the SW TOA flux record can occur when GEO satellites are replaced, because of changes in satellite position, calibration, visible sensor spectral response, cloud retrieval quality, and imaging schedules. Such artifacts in the GEO data can be problematic in studies of TOA radiation interannual variability and/or trends (Loeb et al.,2018)*.*

In practice, CERES SYN1deg is still the best reference data record to validate daily TOA fluxes. It is used for daily and monthly global mean validation (stability), as well as for processing error validation

(spatial pattern validation, aka regional uncertainty) given its spatiotemporal high resolution (combination of GEO data). The largest disadvantage is the record's time span (from 2000 onwards), a period which is therefore also referred to as the 'CERES era', in contrast to the preceding period 1979-1999 ('pre-CERES era'). The data is downloaded from the 'CERES Ordering Tool' web portal (https://ceres-tool.larc.nasa.gov/ord-tool/).

A few CERES products suffer from data gaps in certain periods (Table 3-2). As a consequence, three months are not used for validation purposes (August 2000, June 2001, March 2002). The impact of gaps after July 2002 is less problematic as the CERES products are composed of both Terra and Aqua CERES.

Table 3-2: Dates of the major data gaps in the Terra and Aqua CERES records.

Terra	Aqua
April 26-27, 2000	July 1-3, July 30 – August 6, 2002
August 6-18, 2000 (*)	October 1-14, 2004 (instrument anomaly)
June 15 – July 2, 2001 (*) due to MODIS anomaly	March 30-31, 2005
March 20-28, 2002 (*)	August 16 – September 3, 2020
December 17-24, 2003	
February 19-27, 2016	
(*) these gaps have a large impact on the monthly regional uncertainty; as a consequence these months were not used for the validation analysis (August 2000, June 2001, March 2002)	

3.2 CERES EBAF Ed.4.1 (monthly)

The CERES Energy Balanced and Filled (EBAF) Ed4.1 data record (Loeb et al., 2018) provides state-of-the-art estimates of monthly mean RSF and OLR fluxes from March 2000 onward at a 1°x1° latitude-longitude resolution.

To maintain the excellent CERES instrument calibration stability but also preserve the diurnal information found in SYN1deg, EBAF Ed4 introduced a new approach involving diurnal correction ratios (DCRs) to convert daily regional mean SSF1deg SW fluxes to diurnally complete values analogous to SYN1deg, but without geostationary artifacts (Loeb et al., 2018).

Furthermore, even with the most recent instrument calibration, the SYN1deg Ed4 net imbalance is still ~4.3 W m² much larger than the expected observed ocean heating rate ~0.71 W m² (Johnson et al. 2016). Therefore, the CERES EBAF dataset uses an objective constrainment algorithm (Loeb et al.,2009) to adjust SW and LW TOA fluxes within their ranges of uncertainty to remove the inconsistency between average global net TOA flux and heat storage in the Earth-atmosphere system.

CERES EBAF is used for monthly global mean validation (stability) as well as for processing error validation (spatial pattern validation, aka regional uncertainty). The record's time span is identical to the SYN1deg product, as is the record's download location (https://ceres-tool.larc.nasa.gov/ord-tool/).

3.3 HIRS OLR Daily v01r02

The NOAA National Centers for Environmental Information (NCEI) provides a high quality CDR of daily OLR (Lee, 2014; Lee et al., 2014). Level-1b all-sky data from the High-resolution Infrared Radiation Sounder (HIRS) instruments are the main input into the daily OLR record. The data record is produced by applying a combination of statistical techniques, including OLR regression, instrument ambient temperature prediction coefficients and inter-satellite bias corrections. The **HIRS OLR Daily** record is featured by: (i) a global coverage, (ii) a 1°x1° equal-angle grid resolution, (iii) a temporal coverage from the 1st January 1979 onwards. The OLR estimated from imagers' radiance observations on-board operational geostationary satellites (via Gridsat CDR and GSIP OLR product) is incorporated to allow an accurate temporal integration of the daily mean OLR. Since polar areas (about 60° polewards) are not covered by geostationary observations, only HIRS observations are used to derive the daily OLR in these regions. The HIRS OLR estimation technique has been vigorously validated against the Earth Radiation Budget Experiment (ERBE) and CERES data (see Ellingson et al., 1994; Lee et al., 2007).

The HIRS OLR Daily data record is used for daily and monthly global mean validation (stability), as well as for processing error validation. In contrast to the state-of-the-art CERES products, it's available for the entire time span of the CLARA-A3 record (1979-2020), making it the main reference data record for OLR. In practice, it is used to verify whether the CERES performance is maintained backward in time, during the pre-CERES era. The monthly mean OLR is calculated by temporally aggregating the daily mean OLR. The data is downloaded from the 'UMD OLR CDR Portal' (https://olr.umd.edu/). In this report, this reference data record is referred to as "*HIRS*" in figures and tables.

3.4 HIRS OLR Monthly v02r07

The **HIRS OLR Monthly** data record share the same basic characteristics as the *Daily* record, described in Section 3.3. The data record uses the Level-1b HIRS data as main input and is produced by applying the same combination of statistical techniques. However, the *Monthly* HIRS OLR is generated on a 2.5°x2.5° equal-angle grid. In addition, the monthly OLR CDR is estimated from the HIRS all-sky radiance observations directly and does not use geostationary observations. This data is used to address the stability of the monthly mean CLARA-A3 OLR products. However, it is not used for spatially-explicit validation (processing error aka regional uncertainty) because of the spatial incompatibility due to the low resolution (more info in Section 4.3). The data is downloaded from the 'UMD OLR CDR Portal' (https://olr.umd.edu/). In figures and tables, this reference data record is referred to as "*HIRS-MM*".

3.5 ERA5

At the time of writing, ERA5 is the latest atmospheric reanalysis from ECMWF (Hersbach et al., 2020). The data record provides a physically consistent blend of forecasts and observations, resulting in a spatially and temporally seamless coverage. The model is the Integrated Forecasting System (IFS) cycle 41r2 with a 4-D variational analysis (4DVAR) assimilation system. The output has a temporal resolution of 1 hour, and a reduced gaussian spatial grid, which is then bilinearly

interpolated on a regular lat/lon grid of 0.25°x0.25°. The radiation scheme from ERA5 is described in Hogan and Bozzo (2018). The record's total time span is 1959-2020.

Given the physical consistency throughout the record, ERA5 is selected for long-term global mean bias validation: it is useful to assess the stability of CLARA-A3's data record, especially when there is no other reference data record available (e.g. RSF). On the other hand, ERA5 is still a reanalysis product with a significant modeling component: it drastically underperforms in short-term spatially-explicit comparisons, making it not useful for processing error validation (regional uncertainty) in this report.

The data have been collected from the Copernicus Climate Data Store, available online at https://cds.climate.copernicus.eu .

3.6 Data records purely used for intercomparison

3.6.1 ISCCP-FH

ISCCP-FH (Young et al.,2018; Zhang et al.,2019) is in essence a cloud product with TOA fluxes calculated from the retrieved cloud properties using a radiative transfer model (*RadH-PRD*). For the cloud retrievals, ISCCP-FH uses a composite of polar and geostationary satellites.

The data is provided on a regular grid of 1°x1°, and has been downloaded from https://isccp.giss.nasa.gov/pub/flux-fh/tar-nc4_MPF/.

The product is only used to **compare** the global mean TOA flux (stability) with CLARA-A3 and other data records, and to make a brief assessment of their differences. They are not used for actual validation given their lower performance w.r.t. the state-of-the-art reference records CERES and HIRS records.

3.6.2 Cloud_CCI

Similar to ISCCP-FH, Cloud_CCI (Stengel et al.,2020) is also a cloud product with TOA fluxes calculated from the retrieved cloud properties using a radiative transfer model (*BUGSrad*). For the cloud retrievals, Cloud_CCI (L3C AVHRR-PM v3.0) is based purely on AVHRR afternoon satellites.

The Cloud_CCI data are provided on a regular grid of 0.5°x0.5° and have been downloaded from https://public.satproj.klima.dwd.de/data/ESA_Cloud_CCI /CLD_PRODUCTS/v3.0/L3C/

Analogous to ISCCP-FH, the Cloud_CCI product is only used to **compare** the global mean TOA flux (stability) with CLARA-A3 and other data records, and to make a brief assessment of their differences. They are not used for actual validation given their lower performance w.r.t. CERES and HIRS.

4   Methodology

4.1   Terminology

The three main uncertainty metrics discussed here are the mean bias, the stability, and the processing error (regional uncertainty) of the CLARA-A3 fluxes with respect to the reference data records.

4.1.1   Mean bias

The CLARA-A3 RSF and OLR products rely on empirical relations with CERES products, and hence their **absolute radiometric level** can be considered 'tuned' (not independent). Consequently, no requirement is defined in the Requirements Review Document [RD 1], and no attempt is done to quantify this metric in this report.

Rather than denoting the absolute radiometric error, this report uses the term *'Mean Bias'* to describe the instantaneous **overall bias with respect to a reference data record**. It is calculated by subtracting the gridded CLARA-A3 flux from a gridded reference data record which produces a gridded bias (a 'bias map'), from which the global spatial average is taken.

Depending on the reference data record, this *Mean Bias* may have several causes, such as a differences in calibration, satellite instruments, time of observation, temporal frequency, etc., which all have in common that they are not random but relatively constant in time and space (although they may slowly evolve in time, e.g. drifting of satellite orbit).

Because of its 'tuned' character (not independent), and given the significant regional bias variations (leading to large compensation effects), the *Mean Bias* itself is not considered a meaningful 'accuracy' metric for the CLARA-A3 TOA flux products. However, it is still interesting to compare the CLARA-A3 mean bias with other data records, i.e. how are CLARA-A3 and these other data records scaled compared to the absolute level of the CERES products.

4.1.2   Stability

The **stability** of the CLARA-A3 data record is evaluated as the maximum variation (max-min) of the global *Mean Bias* over a period of 10 years (decade).

A stable data record consists of a temporally systematic *Mean Bias*. Note that this stability is only relative to the inherent stability of the reference data record (which is, however, very good for some data records). Using different reference records allows attributing observed stability problems to one of these sources.

Variations or discontinuities, caused by several mechanisms mentioned in Section 4.1.1, should remain within acceptable limits to render the data record useful for climate monitoring purposes.

4.1.3 Processing error (regional uncertainty)

The second source of uncertainty comes from the processing of AVHRR observations into TOA fluxes. This includes the conversion of the narrowband (channel) observations (reflectances and brightness temperatures) into broadband quantities, the subsequent integration from these directional to hemispherical quantities using the ADMs, and finally the daily and monthly temporal interpolation of these quantities (see [RD 7] for details). To quantify this error, the CLARA-A3 products are compared with similar products derived from the CERES instruments at a 1°x1° spatial scale. CERES is considered as the best reference data to address this accuracy. For OLR, also HIRS is used to assess the processing error during the pre-CERES era (1979-1999).

In practice, the CM SAF products are first regridded on the same nested 1°x1° lat-lon grid as used for the CERES products. Then, the root mean squar (RMS, bias corrected) of the difference with the CERES products is evaluated. It is interesting to look at time series of the processing error, to check the consistency over the data record extent, in particular to check that the errors obtained with different satellites (different AVHRR instruments) are consistent with each other.

Even after correction for the global *Mean Bias* (Section 4.1.1), the processing error still contains a considerable regional **systematic component**: indeed, each grid box has a surface type that is invariant in time (e.g. ocean, desert, ..), and in some regions also the cloud cover has a preferential state (e.g. clear sky is dominant in the Sahara desert). Therefore, scene type dependent errors can be considered regionally systematic errors. This explains the **"accuracy" part** of the processing error. On the other hand, there is also a **random component** of the processing error. For instance, errors dependent on viewing and illumination geometry (angular dependent errors). For instantaneous fluxes, or for fluxes integrated on short time scales (e.g. daily mean), these errors can be significant. On longer time scales, for a given location (grid box), these errors cancel each other out, since the angles of all observations are not constant but change randomly over time. Indeed, we see that part of the processing error decreases when calculated for longer time scale. This explains the **"precision" part** of the processing error. Accuracy and precision are therefore assessed together in the combined processing error, and globally integrated with the bias-corrected metrics MAB and RMSB, which are calculated spatially, i.e. over all the grid boxes, and for each time step (daily mean flux, monthly mean flux, ..).

The processing error metrics MAB and RMSB are furthermore an expression of the **regional uncertainty** in the spatially-explicit grid of CLARA-A3 fluxes: it describes to which extent the bias deviates from its mean **in the spatial dimension**, i.e. how spatially homogeneous or heterogeneous the bias is (for a given temporal unit, i.e. for a given map depicting daily or monthly mean flux). The CLARA-A3 flux is provided with an uncertainty range of +/- MAB with 57.5% accuracy, or with an uncertainty range of +/- RMSB with 68% accuracy, assuming a gaussian distribution (see also Sections 4.2.3 and 4.2.4).

4.2 Statistical measures

The retrieved daily mean CLARA-A3 flux (F_{CLARA}) is validated against the daily mean flux from a gridded reference data record, denoted by F_{REF}. The following statistical measures are used in the validation report:

4.2.1 Bias defined per grid box ($B_{i,j}$)

Prior to the validation, the spatial resolutions of both F_{CLARA} and F_{REF} are first downgraded to match the CERES nested processing grid (Section 4.3). Maps of their bias are then created (daily "bias maps"), from which a single grid box with indices (*i,j*) is calculated as:

$$B_{i,j} = F_{CLARA,i,j} - F_{REF,i,j}$$

(1)

The grid box specific bias is used to calculate the other statistical measures.

4.2.2 Mean Bias (MB), defined globally

The global Mean Bias (*MB*) is calculated over all grid boxes' biases as follows:

$$MB = \frac{1}{m \cdot n} \cdot \sum_{i=1}^{m}\sum_{j=1}^{n} w_j(B_{i,j}) = \frac{1}{m \cdot n} \cdot \sum_{i=1}^{m}\sum_{j=1}^{n} w_j(F_{CLARA,i,j} - F_{REF,i,j})$$

(2)

Where $B_{i,j}$ is the grid box specific bias (Section 4.2.1), *m* and *n* are the number of grid boxes in longitude (360) and latitude (180) dimension, respectively, and w_j is a meridionally varying weighting factor to correct the equal angle grid boxes to equal area grid boxes (to prevent an over-representation of the polar areas). The weighting factor is normalized so that its global average amounts 1 (i.e. its global sum amounts *m·n*). The *MB* statistic is used in this report to validate the global bias, and hence the stability.

4.2.3 Mean Absolute Bias (MAB), bias-corrected, defined globally

The global Mean Absolute Bias (*MAB*) is calculated by first subtracting the global Mean Bias from every grid box' bias ($B_{i,j}$-*MB*), which removes the general bias ('bias correction'). Subsequently, the absolute value is taken from the result, after which a global average is calculated in the same way as for the global mean bias (Section 4.2.2).

$$MAB = \frac{1}{m \cdot n} \cdot \sum_{i=1}^{m}\sum_{j=1}^{n} w_j \cdot |B_{i,j} - MB| = \frac{1}{m \cdot n} \cdot \sum_{i=1}^{m}\sum_{j=1}^{n} w_j \cdot |F_{CLARA,i,j} - F_{REF,i,j} - MB|$$

(3)

The *MAB* statistic is used in this report to validate the processing error (*regional uncertainty*). Assuming normality, the range between +/-1 *MAB* contains roughly 57% of the data, and the range between +/- 2 *MAB* contains roughly 89% of the data (Figure 4-1).
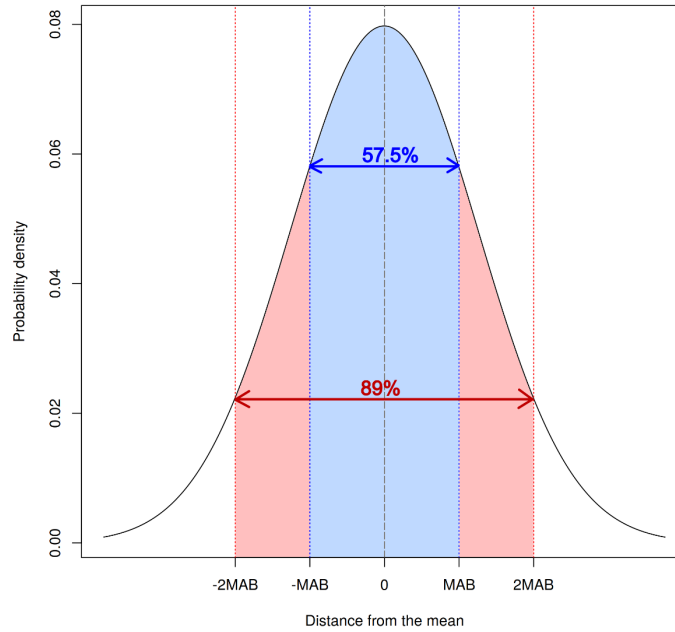
Figure 4-1: Normal distributions with +/-1*MAB and +/-2*MAB

4.2.4 RMS Bias (RMSB), bias-corrected, defined globally

The global Root Mean Square Bias (*RMSB*) is calculated by first subtracting the global Mean Bias from every grid box' bias ($B_{i,j}$-*MB*), which removes the general bias. Subsequently, the *RMS* is taken over the resulting bias-corrected grid boxes, taking into account the equal-area weighting factor (cf Section 4.2.2).

$$RMSB = \sqrt{\frac{1}{m \cdot n} \cdot \sum_{i=1}^{m}\sum_{j=1}^{n} w_j \cdot \left(B_{i,j} - MB\right)^2} = \sqrt{\frac{1}{m \cdot n} \cdot \sum_{i=1}^{m}\sum_{j=1}^{n} w_j \cdot \left(F_{CLARA,i,j} - F_{REF,i,j} - MB\right)^2}$$

(4)

The *RMSB* statistic is used in this report to validate the processing error (*or regional uncertainty*). Assuming normality, the range between +/-1*RMSB* contains roughly 68% of the data, and the range between +/- 2*RMSB* contains roughly 95% of the data. The relation between *MAB* and *RMSB* is in a normal distribution the following:

$$MAB = RMSB \cdot \sqrt{\frac{2}{\pi}}$$

(5)

Or, roughly a factor 0.8 between both metrics.

4.3 Maps and grids

Unlike validation of global means, a spatially-explicit validation (such as MAB or RMSB) requires each data record to be aggregated on a ***common base grid***, typically the coarsest one. It would not be 'fair' to perform a validation on a coarse grid while others on a finer grid, since the aggregation process smooths out (or even masks) fine grained spatial pattern errors. That is also the reason why in this report the very coarse (2.5°x2.5°) HIRS OLR Monthly v02r07 is not used for MAB or RMSB validation. All others data records were already available in (or could be aggregated to) the so-called CERES Nested 1.0° grid[1], which was selected as *common base grid*.[2]

Table 4-1: CERES Nested 1.0° Processing Grid

Latitude Segment	# of zones in segment	Longitude extent (°)	# of regions/zone	# of regions in segment
Equator to 45°	90	1°	360	32400
45° to 70°	50	2°	180	9000
70° to 80°	20	4°	90	1800
80° to 89°	18	8°	45	810
89° to 90°	2	360°	1	2
Total	*180*	*–*	*–*	*44012*
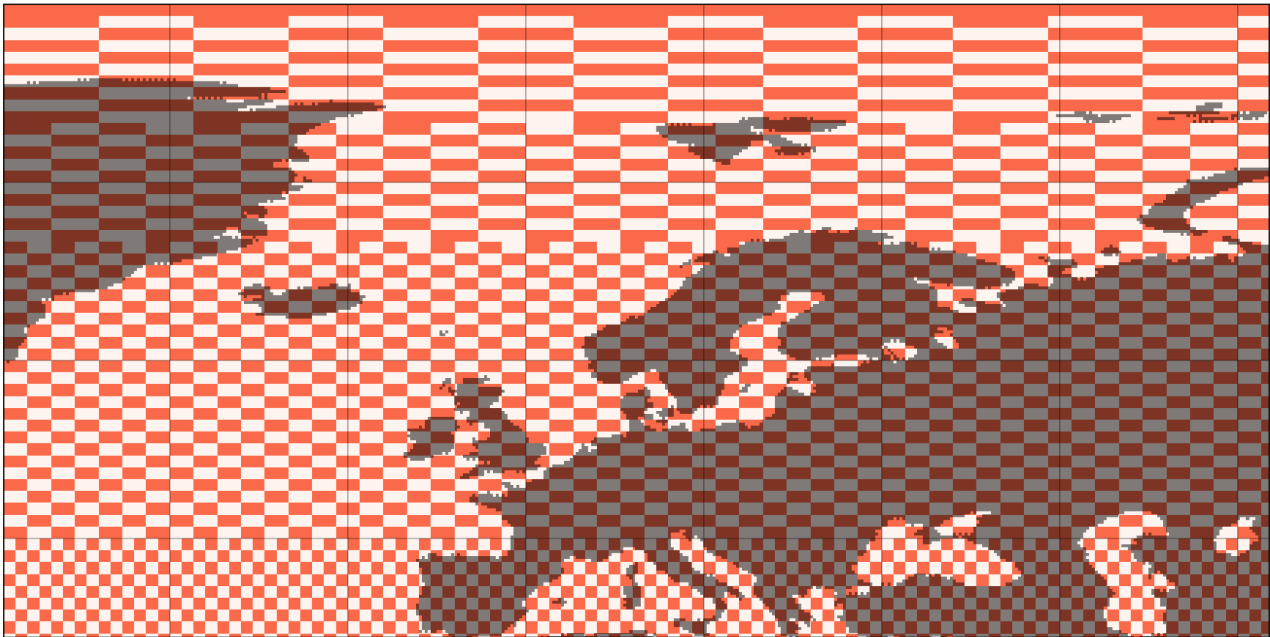
Figure 4-2: CERES Nested 1.0° Processing Grid, zoom-in on Europe, showing transition between 1°x1° (south of 45°N) to 1°x8° (north of 80°N)

Since this is an equal-angle grid, global statistical metrics do not represent the true spatial distribution (in reality pixel area decreases poleward). Therefore, a meridionally varying weighting factor is applied to account for the spatial distortion, thereby in practice converting the grid to an equal-area grid (Section 4.2.2, Equation 6).

---

[1] https://ceres.larc.nasa.gov/data/general-product-info/#ceres-nested-10-processing-grid

[2] The results of a hypothetical separate spatially-explicit validation on the 2.5°x2.5° grid for HIRS v02r07 would not be comparable. Furthermore, selecting the 2.5°x2.5° grid as common base grid was also no option, as too much spatial information would be lost (factor of 6.25 information loss w.r.t. CERES grid, factor 100 w.r.t. CLARA-A3 grid).

4.4   Missing data in CLARA-A3, and gap-filling

For a number of reasons, there are gaps in the data coverage of CLARA-A3 TOA flux products. This section identifies these missing data.

4.4.1   Temporal data gaps

The occurrence of so-called "temporal data gaps" for the RSF product is shown in Figure 4-3. A day is considered 'missing' if there is no (i.e. zero) Level-2b input data to create the daily mean, which means **no** valid input files (i.e. orbital overpasses) during the preceding, actual, or following day[3]. In that case, there is simply no input data to generate at least 1 valid daily mean grid box. The most common reasons are instrument failure and/or temporal gaps between subsequent satellites (between decommissioning and launch).

---

[3] At least 1 inputfile in the preceding, actual or following day is sufficient to generate a valid daily mean.

Temporal gaps in CLARA-A3: number of missing days per month
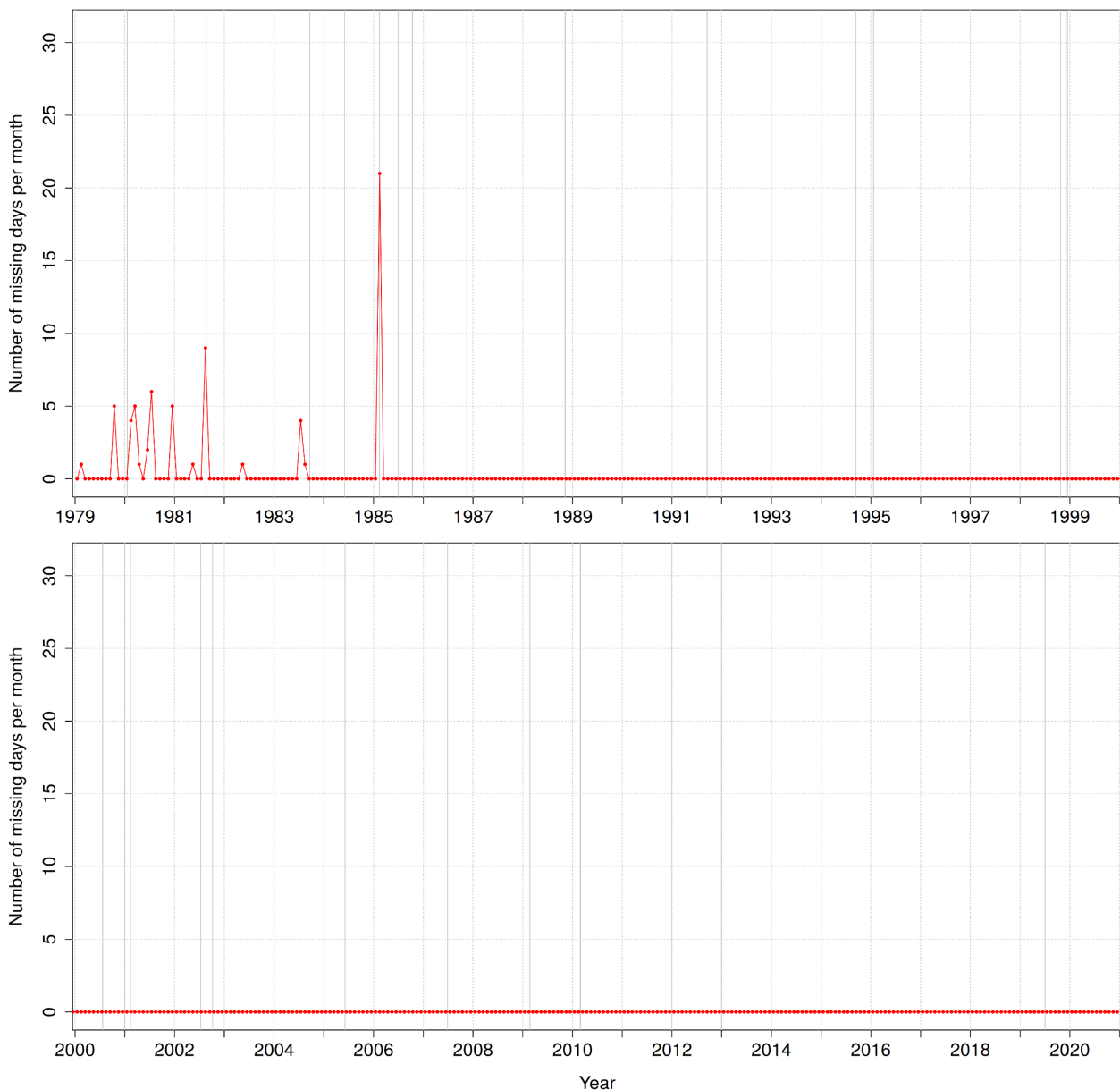
Figure 4-3: Temporal gaps in CLARA-A3

Most important are February 1985 (21 missing days) and August 1981 (9 missing days). There is no attempt to perform any gap filling.

Some of the monthly validation metrics used in this report were found very sensitive to the number of missing days. Therefore, a threshold of maximum 2 missing days/month was applied to filter out the outliers caused by these missing data. For the daily validation metrics, it had no impact, since these daily mean fluxes simply don't exist and hence are omitted in the plots.

4.4.2 Spatial data gaps and filling

The occurrence of spatial data gaps for the RSF product is shown in Figure 4-4. Missing data are considered "spatial data gaps" if it concerns only parts of the globe, more specifically if not 100% the

pixels of a daily mean RSF map are missing. It is acknowledged that the distinction between temporal and spatial data gaps is not really meaningful in some extreme cases (e.g. if almost 100% of pixels are missing).
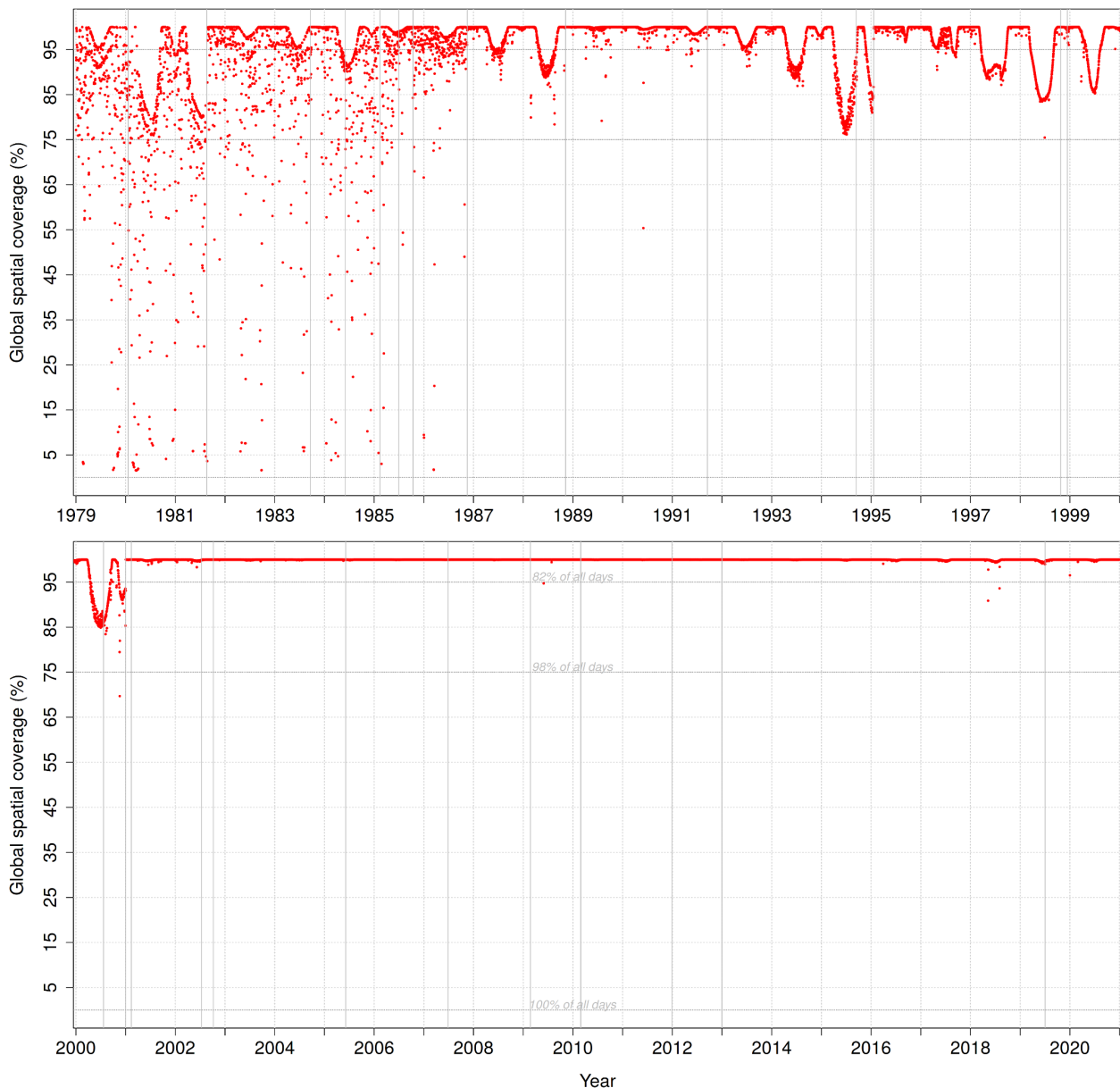
Figure 4-4: Spatial gaps in CLARA-A3 RSF

There are a number of types of spatial data gaps. The most frequent (in terms of time) originates in the fact that for some regions the day length is simply too short to have valid observations between sunrise and sunset. This is quite common during southern wintertime (JJA) in the region where the South Pole "twilight-only regime" borders to the "twilight-and-daylight regime". Even after implementation of a mitigating exception (extension of the twilight model from 84°<SZA<100° to 80°<SZA<100° on condition that the SZA never decreases below 80°), systematic small pockets of missing values remain common, an example of which is shown in Figure 4-5. Their impact on the global mean is, however, negligible.

CLARA-A3 TOA SW radiation (20170701)   Mean = 92.8 W/m^2
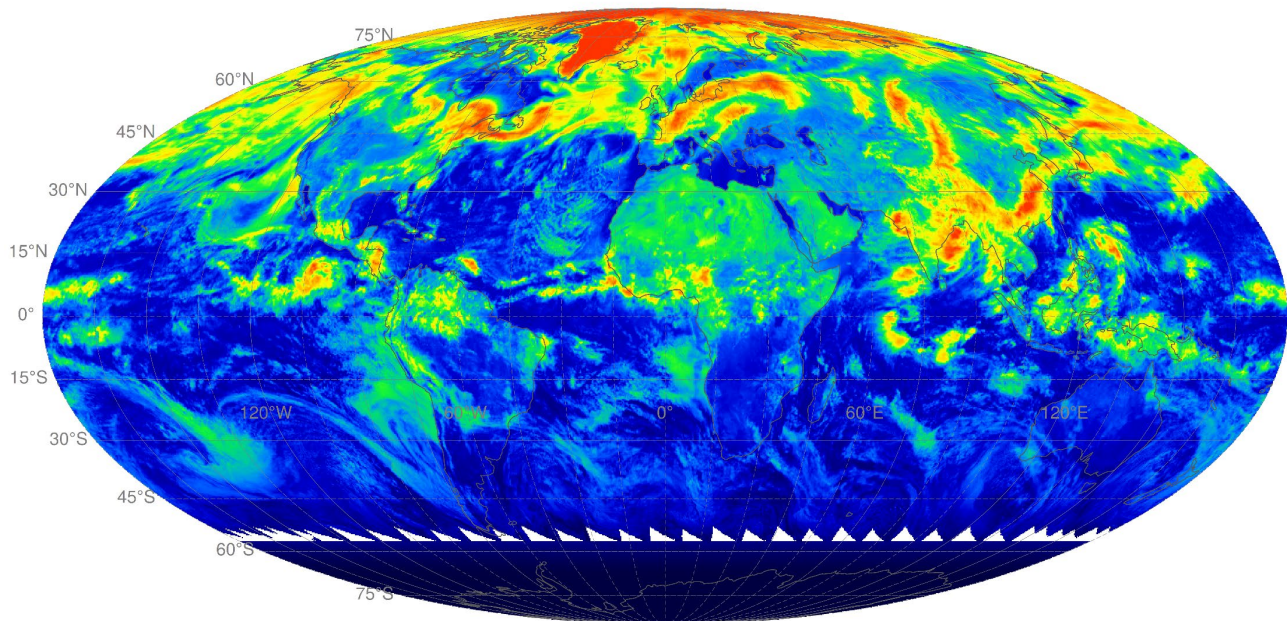
Figure 4-5: Daily CLARA-A3 RSF for 01-07-2017. White area indicates missing spatial data.

A second type consists of missing (parts of) individual orbits, which were filtered out due to insufficient quality, instrument failure, etcetera. An example is given in Figure 4-6.

CLARA-A3 TOA SW radiation (19830103)   Mean = 102.72 W/m^2
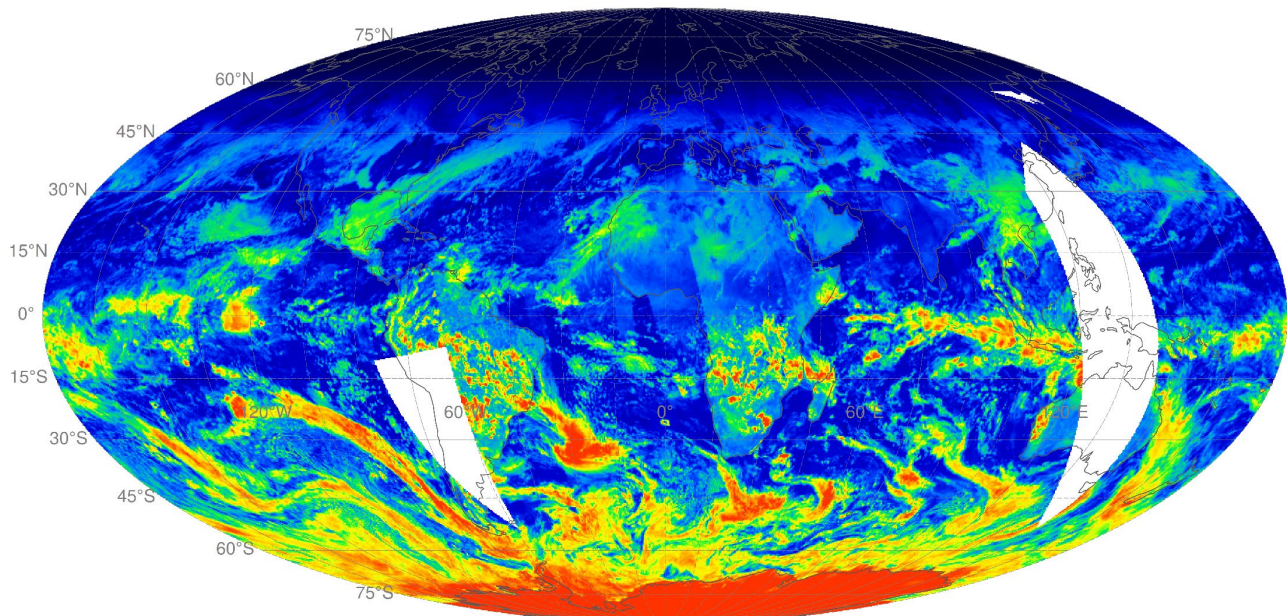
Figure 4-6: Daily CLARA-A3 RSF for 03-01-1983. White area indicates missing spatial data.

The most important type of missing RSF data in terms of spatial extent (of missing area) concerns areas that were flagged as insufficient quality, e.g. due to solar contamination, in the first part of the CLARA-A3 processing chain (i.e. reading the PyGAC output before using the data in PPS, more info

	Validation Report **CLARA Edition 3** **TOA Radiation**	Doc.No: SAF/CM/RMIB/VAL/GAC/TOA Issue: 1.1 Date: 06.02.2023

EUMETSAT
CM SAF
CLIMATE MONITORING

in the ATBD [RD 7] Sections 2.1 and 2.2 therein). An example of 13/07/1994 is given below. Almost 25% of the pixels are missing on that day.
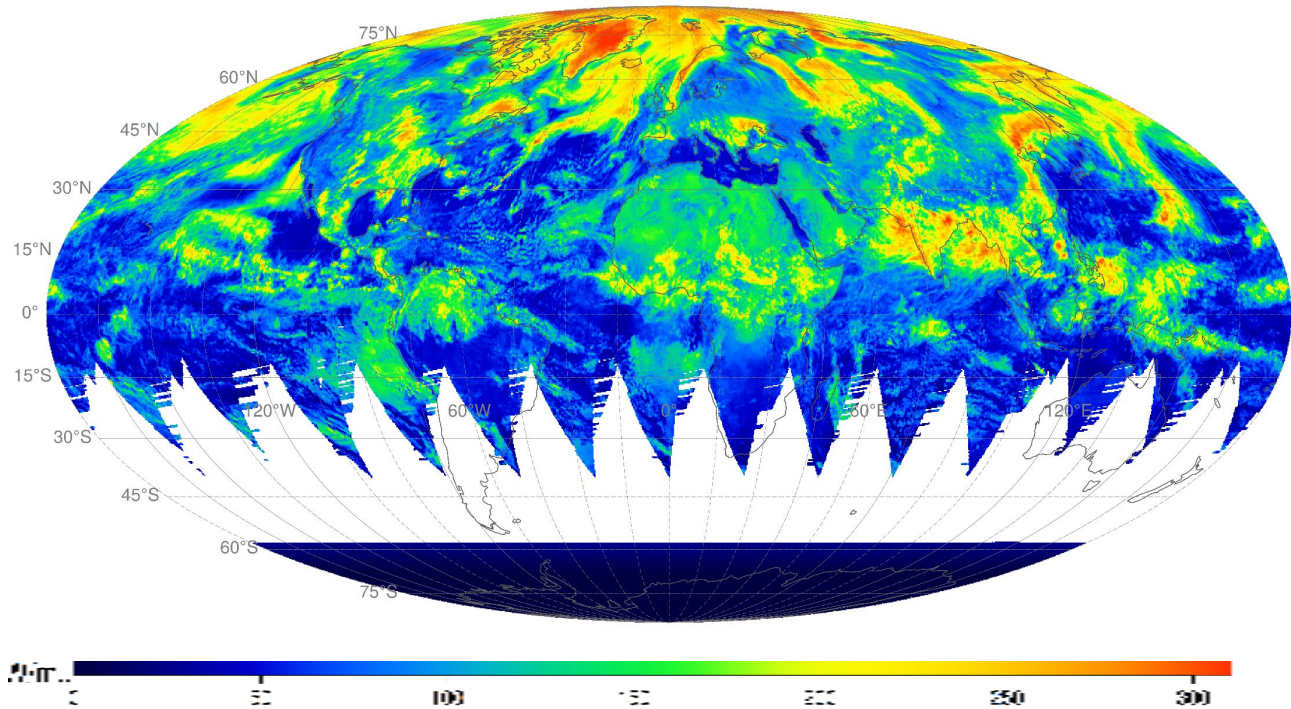
Figure 4-7: Daily CLARA-A3 RSF for 13-07-1994. White area indicates missing spatial data.

The official CM SAF CLARA-A3 TOA flux products are available without gap filling (as in Figure 4-7). It is clear that the user should be aware of these gaps when deriving global average fluxes from these maps. This is shown in Figure 4-8, depicting the global mean daily RSF for CLARA-A3 (orange) and ERA5 (black). During years with increased spatial data gaps (Figure 4-4), the orange CLARA-A3 RSF has peculiar additional peaks compared to the black ERA5 annual cycle (Figure 4-8). Without any reference, context, or warning, this plot could be easily misinterpreted.

Global mean flux of daily CLARA-A3 RSF and other data records
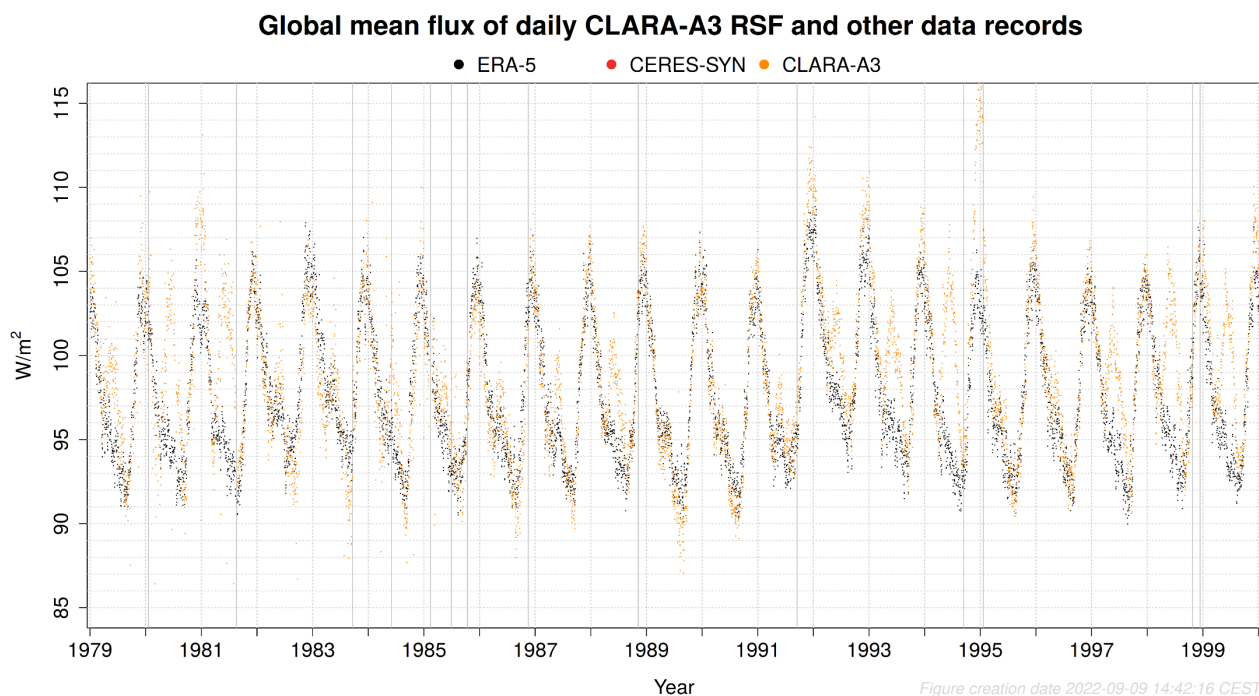
Figure 4-8: Global mean flux of daily RSF during 1979-1999 without spatial gap-filling

In this validation report, however, the spatial gaps in **daily** mean RSF and OLR are filled using daily mean ERA5 as '*surrogate data record*'. This is done <u>after</u> the Level-3 CLARA-A3 fluxes (daily mean flux at 0.25°x0.25°) were aggregated to the CERES nested grid (cfr. Section 4.3). It is only performed if there is at least one valid pixel (i.e. not for entirely missing days, see Section 4.4.1). The result after spatial gap-filling is shown in Figure 4-9. Now, both graphs correspond much better, demonstrating the need of proper spatial gap-filling before validating the product. More information about the spatial gap-filling is provided in Annex 10.1.

Global mean flux of daily CLARA-A3 RSF and other data records (*)
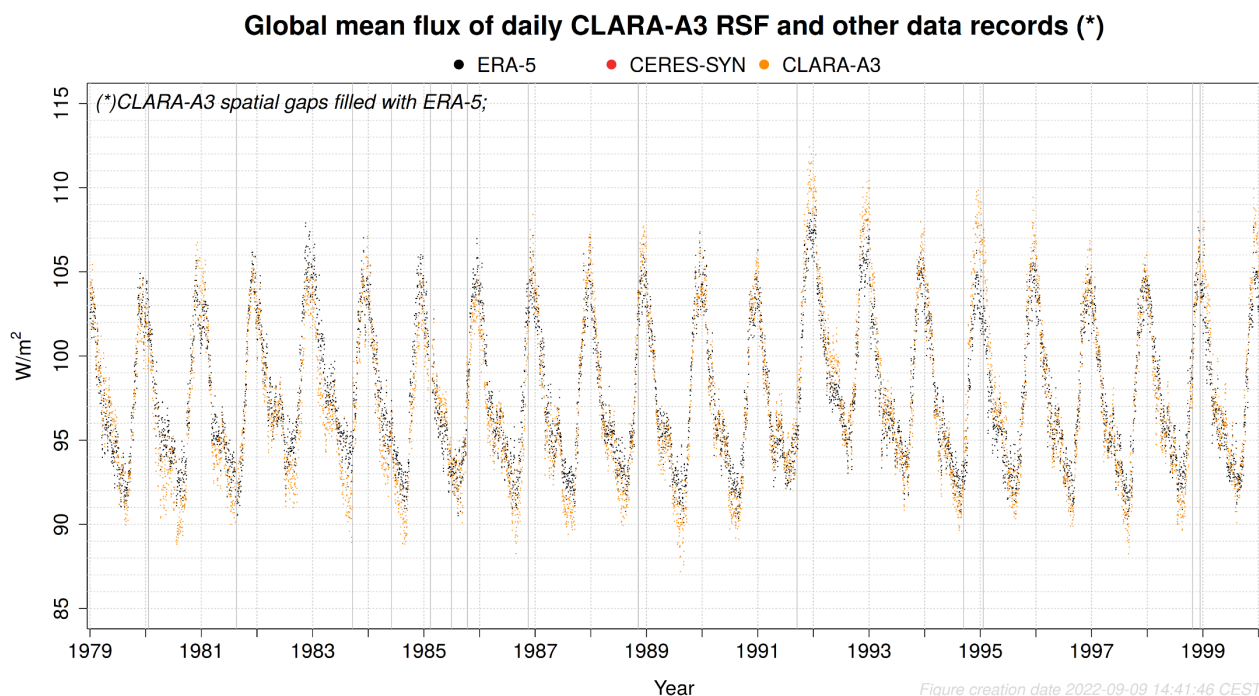
Figure 4-9: Global mean flux of daily RSF during 1979-1999 with spatial gap-filling

For the official **monthly** mean RSF and OLR, the mean flux on a given grid box is calculated without considering the number of contributing days with a valid flux value on that pixel. The number of days for that grid box is provided as ancillary data, so that the user can choose a threshold for using or rejecting the grid box's monthly mean value.

In this validation report, however, the monthly mean RSF and OLR products are derived by averaging all the contributing gap-filled daily mean fluxes (which have, per definition, complete spatial coverage). Hence, these gap-filled monthly mean fluxes also have a complete spatial and temporal coverage. Only exception here is if there are missing days (temporal gaps), as explained in Section 4.4.1). In most cases the gap-filling does not greatly affect the validation results. Nevertheless, to provide full transparency the validation results using non-gap-filled CLARA-A3 fluxes are shown in a separate section in Annex 10.7 (for RSF) and Annex 10.8 (for OLR).

Spatial gap-filling is done only as part of the validation (this report), and not provided in the official product. The main reason for this is that the need for gap-filling strongly depends on the specific application of the data (e.g. local vs global scale). Furthermore, the user is left a choice of *surrogate data record* if gap-filling is deemed necessary, without forcing the use of e.g. ERA5 or any other data record.

4.5 CLARA-A3 orbital configuration and temporal data visualization

The orbital constellation of AVHRR-carrying satellites is not constant but varies in time regarding the number of satellites, and regarding their respective Equator Crossing Time (ECT)[4]. This is referred to as the "**orbital configuration**", which determines the temporal coverage throughout the day (density and spread of observations) for a given location. A single satellite observes a given location at the equator every 12 hours, i.e. two times per day (*ascending and descending node*), from which one during daylight conditions ('daytime') as illustrated in Figure 4-10 (useful for both RSF and OLR), and the other during nighttime, i.e. between 18h and 06h local time (only useful for OLR).

There are certain typical time slots in which the satellites are launched, and subsequently slowly evolve due to orbital drift. Historically these are the morning orbit (around 7h30 ECT at launch) and the afternoon orbit (around 14h00-14h30 ECT at launch). Over time, they each tend to drift towards the terminator, i.e. the morning orbit towards an earlier ECT whereas the afternoon orbit towards a later ECT. It is worth mentioning here that this historical configuration was not symmetrical around noon (12h ECT), i.e. the morning orbit is always closer to the terminator compared to the afternoon orbit.

For some periods in the record, there is only one orbit available, either morning or afternoon. This limited temporal coverage is referred to as "***suboptimal orbital configuration***", as only a part of the day is covered. Note that it is not a binary issue: even in an orbital configuration with 2 satellites, the temporal coverage can be downgraded when one of the orbits has strongly drifted towards the terminator, thereby gradually resembling more and more a suboptimal orbital configuration.

---

[4] Expressed as local time aka solar time (hours)
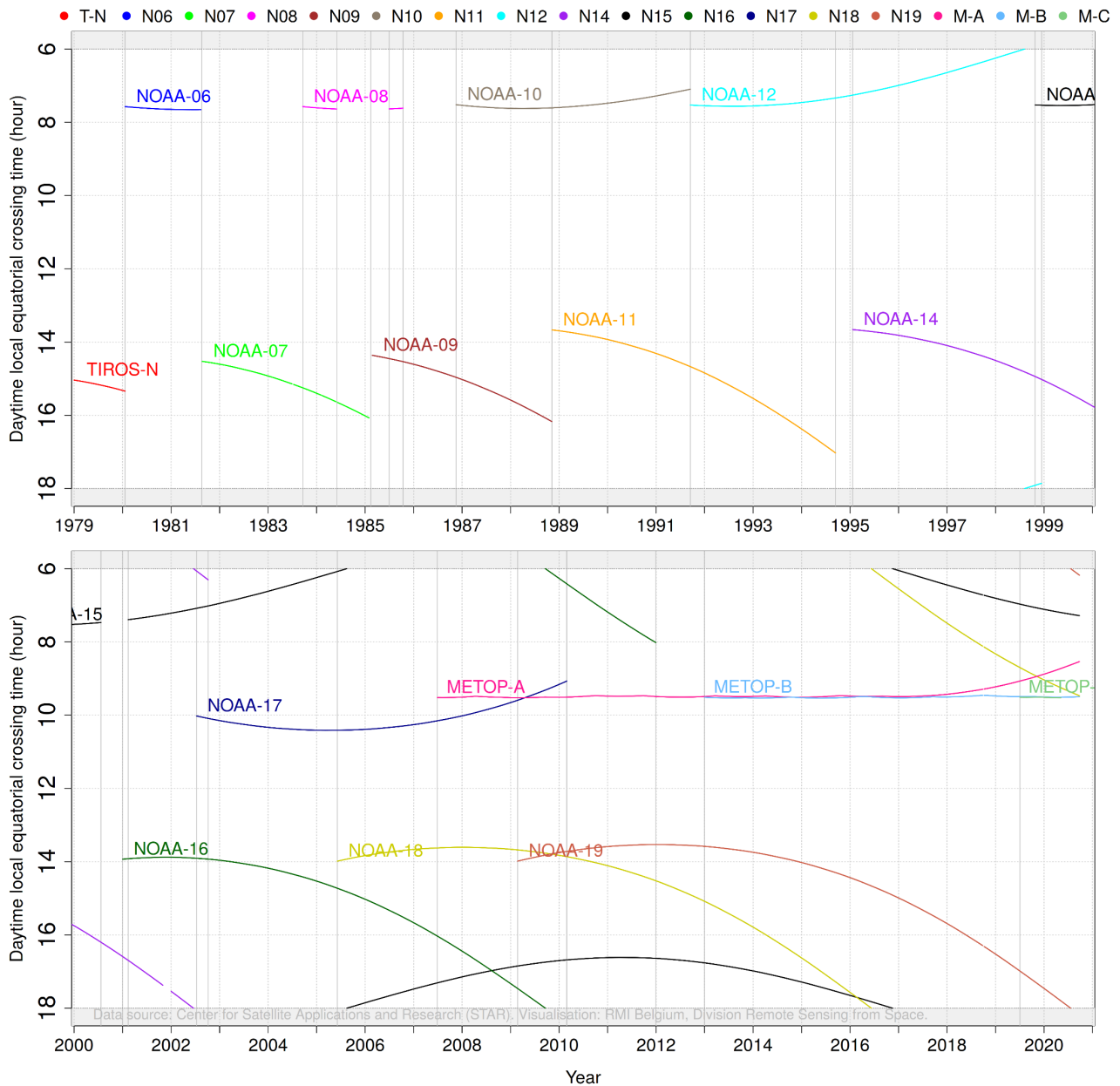
Figure 4-10: Daytime local equator crossing time of satellites used for CLARA-A3

The vertical solid gray lines in Figure 4-10 indicate transitions (discontinuities) in the orbital configuration, which often correspond to changes in (local) time of observation (i.e., ECT). These lines are included in all the temporal plots of the current report, and an overview of all these transitions is provided in Table 4-2.

Table 4-2: Transitions in CLARA-A3 orbital configuration

Date (start)	Date (end)	Satellite	Orbital configuration
1979-01-01	1980-01-20	TIROS-N	Afternoon (=suboptimal)
1980-01-20	1981-08-19	NOAA-6	Morning (=suboptimal)
1981-08-19	1983-09-19	NOAA-7	Afternoon (=suboptimal)
1983-09-19	1984-06-01	NOAA-8 & NOAA-7	Morning & Afternoon
1984-06-01	1985-02-13	NOAA-7	Afternoon (=suboptimal)
1985-02-13	1985-07-01	NOAA-9	Afternoon (=suboptimal)
1985-07-01	1985-10-14	NOAA-8 & NOAA-9	Morning & Afternoon
1985-10-14	1986-11-17	NOAA-9	Afternoon (=suboptimal)
1986-11-17	1988-11-08	NOAA-10 & NOAA-9	Morning & Afternoon
1988-11-08	1991-09-16	NOAA-10 & NOAA-11	Morning & Afternoon
1991-09-16	1994-09-13	NOAA-12 & NOAA-11	Morning & Afternoon
1994-09-13	1995-01-20	NOAA-12	Morning (=suboptimal)
1995-01-20	1998-10-26	NOAA-12 & NOAA-14	Morning & Afternoon
1998-10-26	1998-12-14	NOAA-15 & NOAA-14 & NOAA-12	Morning & Afternoon
1998-12-14	2000-07-22	NOAA-15 & NOAA-14	Morning & Afternoon
2000-07-22	2001-01-01	NOAA-14	Late afternoon (=suboptimal)
2001-01-01	2001-02-12	NOAA-16 & NOAA-14	Afternoon (=suboptimal)
2001-02-12	2002-07-11	NOAA-15 & NOAA-16 & NOAA-14	Morning & Afternoon
2002-07-11	2020-12-31	(multiple satellites)	Morning, **Mid-morning**, afternoon

4.6 Evaluation of ICDR products

Interim Climate Data Records (ICDRs) denote regular extensions of TCDRs with algorithms and processing systems identical to the already generated reference TCDRs, but with shorter time latency. For this validation report, ICDR data was generated and made available for the period July 2020 until December 2020 (i.e. second half of that year). The ICDR performance is assessed based on how closely it resembles the TCDR despite their differences in input data (ERA5 vs ERA5T for NWP variables, and sea ice concentration with a different OSI SAF product, see also [RD 7]) and differences in FDR (exclusion of Metop-C). The ICDR validation results are provided in Sections 5.4 (RSF) and 6.4 (OLR).

Notice that it was decided to exclude results from Metop-C among the ICDR products (*cfr. section 2.1.1 in [RD 7]; chapter 2 in RD 9).* Even if the level-3 product of the TCDR included 1 ½ years of Metop-C data, it was judged as risky to continue processing Metop-C data further for the ICDR due to uncertainties in the time-dependent calibration corrections of the calibration for the visible channels. These corrections need to be estimated from a longer period of historic data than just 1 ½ years. Thus, Metop-C data will be introduced in the ICDR at a later stage when more reliable calibration corrections are available.

5   Results for Reflected Solar Flux (RSF)

5.1   Mean bias and Stability

The global mean flux of monthly RSF from different data records is shown in Figure 5-1, among which CLARA-A3 RSF in orange.
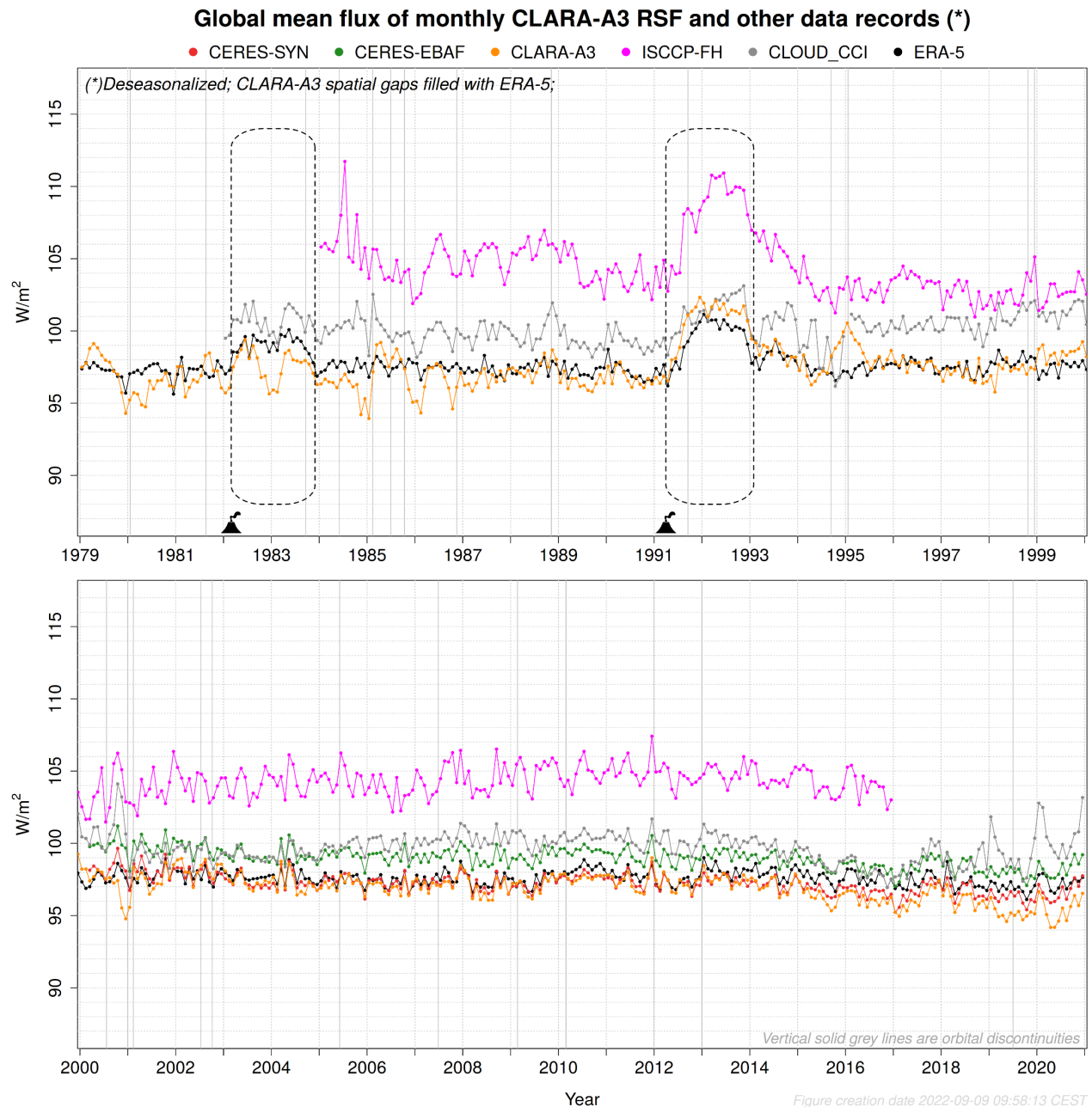
Figure 5-1: Global mean flux of monthly CLARA-A3 RSF and other data records

The ERA5 time series proves to be stable and can be used to assess the stability of other data records in the pre-CERES era (1979-1999). The two major volcanic eruptions El Chichón and Pinatubo are indicated on the time series, and their radiative impact is estimated in Table 5-1. The volcanic eruptions led to a dramatic increase in stratospheric sulfate aerosol loading, causing a large rise in the reflection of solar radiation due to the optical properties of sulfuric acid droplets (Canty et

al., 2013). Unlike the Pinatubo event, CLARA-A3 RSF does not properly capture the radiative impact of the El Chichón event (cf the drop of ~2W/m² w.r.t. ERA5 around January 1983, Figure 5-1); this probably originates from a seasonal bias during the relatively short time of suboptimal orbital configuration, i.e. not removable by deseasonalization of the time series. More information about this issue is given in Annex 10.3.1.

Table 5-1: Volcanic eruptions with global impact on monthly CLARA-A3 RSF

Eruption	**Time span of radiative impact**	**Duration**	**Estimated impact**
El Chichón (Mexico)	March 1982 (eruption) – December 1983	+/- 21 months	**+3W/m² (CLARA-A3),** +2W/m² (ERA5)
Pinatubo (Philippines)	April 1991 (eruption) – January 1993	+/- 21 months	**+5W/m² (CLARA-A3**), +4W/m² (ERA5), or +6W/m² (ISCCP-FH)

The CLARA-A3 RSF trend during 2000-2020 amounts -0.99 W/m²/decade, which is relatively close to the CERES EBAF trend (-0.73 W/m²/decade).

The global mean bias is calculated by subtracting the reference data records from CLARA-A3 RSF, resulting in the time series shown in Figure 5-2. The overall stability of CLARA-A3 RSF is assessed w.r.t. **ERA5**, and this is done by creating a so-called '*stability envelope*' (using the threshold requirement of 4W/m² as a reference), set symmetrically around the (slightly negative) mean bias, which is normally distributed (Figure 5-3). **The overall stability complies with the threshold requirement, as more than 94% of the data is contained within the 4W/m² stability envelope.**

During the CERES era (2000-2020) the CLARA-A3 RSF performance is very good, with a mean bias w.r.t. CERES SYN close to zero for the larger part of the two decades (Figure 5-2).

The largest bias fluctuations, where the monthly RSF bias approaches or exceeds the edges of the *stability envelope*, are situated in the first decade of the data record but also as some isolated peaks in later years (1994-'95, 1999, 2000). These biases are predominantly caused by a "*suboptimal orbital configuration*" (Section 4.5) which affects the ability to obtain the correct diurnal cycle; since the impact is region-specific, it introduces a hemispherical imbalance and hence a seasonal cycle in the global mean bias during the years of suboptimal orbital configuration (which is not removed after deseasonalization, given the limited time span w.r.t. the entire data record's duration). Furthermore, the biases are to some extent also caused by the underlying Level-2b data record (containing instantaneous TOA albedo) which in turn depends on the stability of the underlying Level-1c AVHRR FDR and on the consistency and quality of the different input data such as cloud properties. A closer look to the RSF bias, together with an identification of the different bias fluctuations and an analysis of their probable causes, are provided in Appendix 10.3.

A positive bias can be noticed during the Pinatubo radiative impact period, which increased the bias w.r.t. ERA5 by more than +1W/m² compared to the period before and after. However, this should not be considered to be a bias in CLARA-A3 but rather in ERA5. More information about this issue is provided in Annex 10.3.2.

A slight downward trend of about 1W/m² can be noticed between 2015-2020, which is caused by a trend in one of the satellites' Level-2b data record (Metop-B). The full investigation of this issue is provided in Annex 10.3.3. When this issue is solved in future versions of the data record, the RSF trend would be less negative, so closer to CERES EBAF (-0.73 W/m²/dec).

Global mean bias of monthly CLARA-A3 RSF w.r.t. other data records (*)
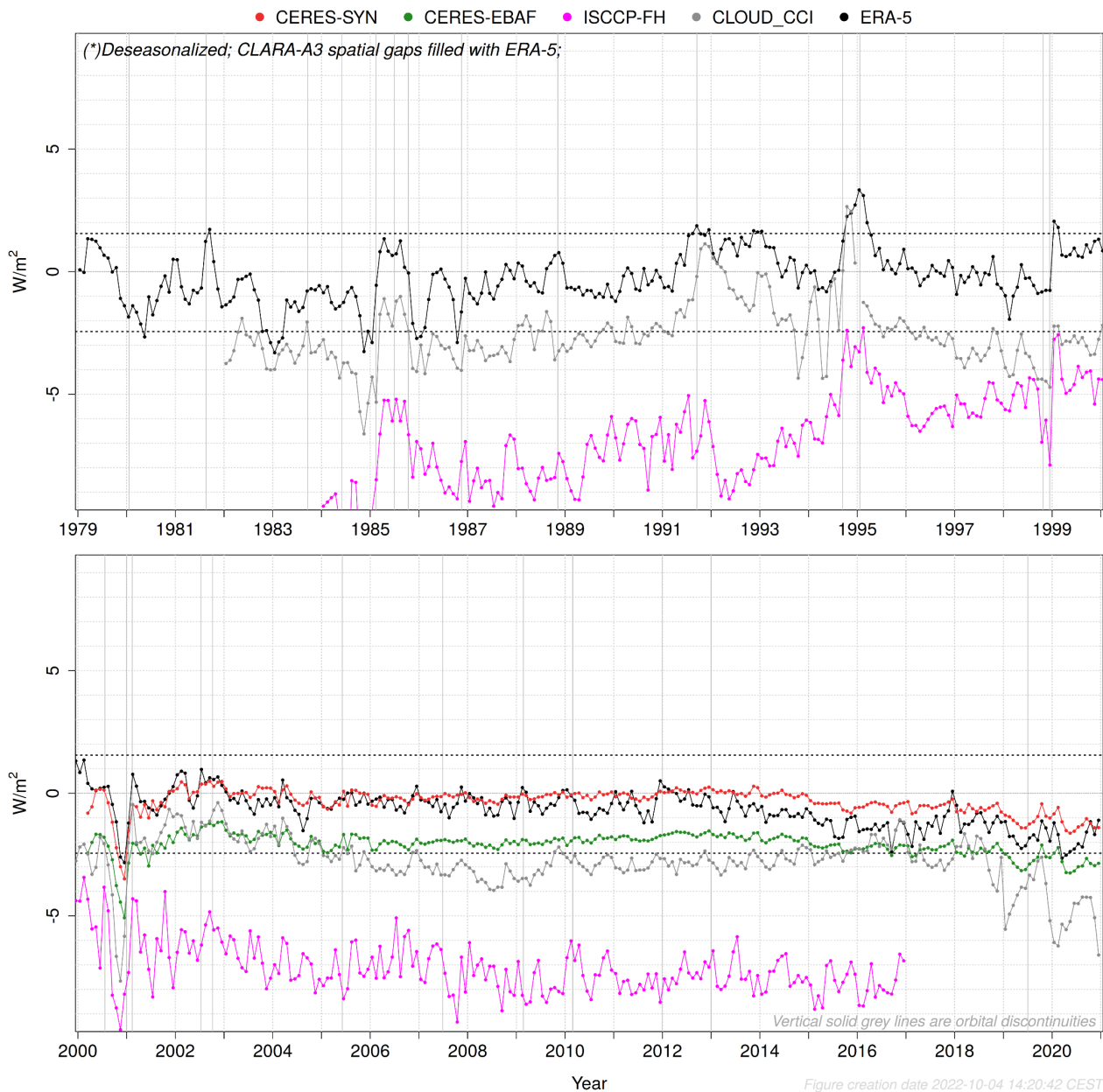
Figure 5-2: Global mean bias of monthly CLARA-A3 RSF w.r.t. other data records. Dotted lines indicate a stability envelope of 4W/m² around the bias w.r.t. ERA5.

In absolute terms, it is not surprising that CLARA-A3 is close to **CERES-SYN1deg** (red curve in Figure 5-2) given the empirical relations between AVHRR and CERES that were first established and then used to derive CERES-like broadband quantities: this could be considered a kind of 'tuning' or 're-calibrating' of the absolute radiometric level. More important is that this time series is relatively flat which indicates a good stability w.r.t. the CERES products. **CERES-EBAF** is consistently ~1.5W/m² lower (green curve in Figure 5-2), which can be explained by the EBAF adjustments made to comply with current estimates of the global energy (im-)balance. Similar to CLARA-A3 RSF, the **Cloud_CCI** data record is based on the AVHRR instrument, but the Cloud_CCI product shown in Figure 5-2 (in gray) is only based on afternoon satellites. Its overall stability is reasonable, mostly hovering around -2 to -3W/m² w.r.t. CERES-SYN and CLARA-A3. Finally, the **ISCCP-FH** data record is considered the least performing, given its seemingly random and large short-term fluctuations (in

the order of 2-3W/m²) as well as long-term instability (oscillating between -10 and -5W/m² w.r.t. CERES-SYN and CLARA-A3).

Figure 5-3: *Histogram of monthly RSF bias from CLARA-A3 w.r.t. ERA5*

The daily mean analysis is not shown here, because the bias' magnitude and fluctuations are very similar and are not affected by the temporal aggregation. For the sake of completeness and transparency, they are provided anyway in Annex 10.2.1.

5.2 Processing error (regional uncertainty)

This validation report measures the processing error (regional uncertainty) with the *Mean Absolute Bias (MAB)* as described in Section 4.2. Nevertheless, all the analyses are also performed with the *Root Mean Square Bias (RMSB)* and available in Annex 10.9.1.

5.2.1 Monthly

The months August 2000, June 2001 and March 2002 are not validated since the CERES products contain data gaps in those months (Table 3-2, Section 3.1). This results in a total number of 247 month. Figure 5-4 shows that for 100% of these months, the MAB w.r.t. CERES SYN1deg-Month is below the threshold requirement of 8W/m², 96.4% is below the target requirement of 4W/m², and 49.8% below the optimal requirement of 2W/m². On average, the MAB amounts 2.3W/m². Much more than for the bias (Section 5.1), the **processing error (regional uncertainty) during the CERES era** is clearly **related to the orbital configuration** (Figure 4-10). Best performance, with MAB around the optimal requirement (2W/m²), is obtained with a maximum number and best spread of observations (i.e. satellite observations) throughout the day, i.e. best temporal coverage. The gradual decrease in performance (i.e. increase of MAB) after 2016 is due to NOAA-19's drift towards an evening orbit (without a replacement satellite for the afternoon orbit). The first years, until halfway 2002, are characterized by a markedly higher MAB, around the target requirement (4W/m²), and again the main reason is the orbital configuration: indeed, the mid-morning orbit is only available since mid-2002 (NOAA-17). The sharp peak during the second half of 2000 represents the worst orbital configuration, being a single late afternoon orbit (NOAA-14).

Global MAB between monthly CLARA-A3 RSF and other data records (*)
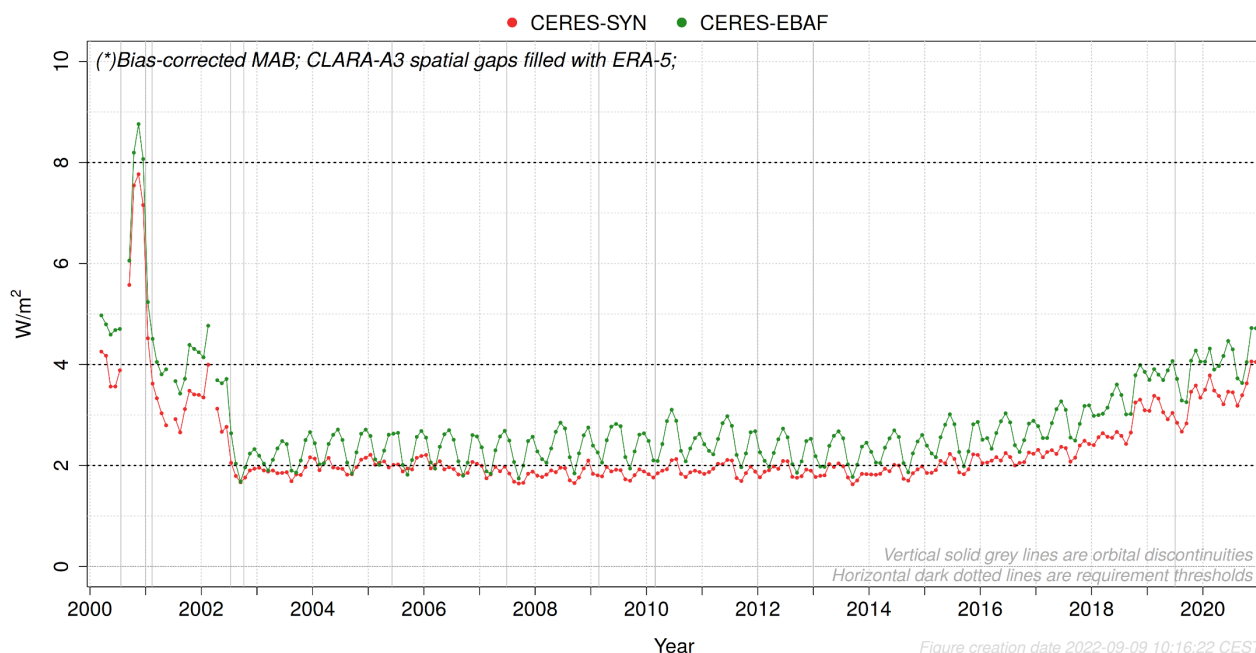
Figure 5-4: Global MAB between monthly CLARA-A3 RSF and other data records

The MAB w.r.t. CERES-SYN1deg (red curve in Figure 5-4) has no significant seasonal cycle during the period with best orbital configuration (2002-2016), whereas first and last few years are characterized by a **seasonal pattern.** Additionally, and unlike the global mean bias, this is also

associated with a decrease in performance (an **increase of MAB**) during those years, which occurs gradually in the last few years (2016-2020).

The following distinct periods can be delineated in the monthly MAB:

1. First half of 2000 with morning + late afternoon satellite, and an MAB of 4W/m²
2. Second half of 2000 with only late afternoon satellite, and an MAB of 6-8W/m²
3. Between 2001-mid2002 with morning+afternoon satellite, and an MAB of 2.5-3.5W/m²
4. Between mid2002-2016 with midmorning+afternoon satellite, and an MAB of 2W/m²
5. After 2016 with midmorning and drifting afternoon satellite, and an MAB increasing from 2 to 4W/m²

As the MAB time series of monthly RSF (this section) shares the same characteristics and features as observed for the *daily* RSF (Figure 5-5), an explanation of these characteristics and features is provided in a common Section 5.2.3.

A consistent seasonal cycle of the MAB w.r.t. CERES EBAF is noticeable, contrary to the absence of such consistent pattern in the MAB w.r.t. CERES SYN1deg. This difference is probably caused by a difference in the processing of CERES products.

Note that until here the MAB validation only concerns the so-called *CERES era (2000-2020)*, roughly corresponding to the second half of CLARA-A3's data record time span. The first half of the record does not have a suitable reference data record to estimate the regional uncertainty. However, since it is clear from the second half of the record that the orbital configuration explains most of the variability, it is possible to estimate the MAB during the pre-CERES era by simulating the associated pre-CERES orbital configurations using carefully selected CERES-era satellites. Annex 10.4 provides the full details of this theoretical exercise, from which the results can be summarized as follows:

Table 5-2: Summary of Monthly RSF uncertainty for the entire CLARA-A3 data record

Time span	Orbital configuration	Frequency of meeting the requirement (%)			Average MAB (W/m²)	Temporal share (%)
		Threshold 8 W/m²	Target 4 W/m²	Optimal 2 W/m²		
Pre-CERES era (1979/01-2000/02)	*(I) Afternoon satellite only*	*100%*	*0.0%*	*0.0%*	*4.8 W/m²*	*62 months (12.4%)*
	(II) Morning satellite only	*84.8%*	*0.0%*	*0.0%*	*6.6 W/m²*	*23 months (4.6%)*
	(III) Afternoon and morning satellite only	*100%*	*73.6%*	*0.0%*	*3.6 W/m²*	*169 months (33.7%)*
CERES era (2000/03-2020/12)	All available satellites	100%	96.4%	49.8%	2.3 W/m²	247 months (49.3%)*
Total record	**All available satellites**	**99.3%**	**72.4%**	**24.6%**	**3.2 W/m²**	**501 months (100%)***
(*) Three months are not taken into account: August 2000, June 2001 and March 2002, since CERES products contain data gaps in those months;						

The result is that for the entire data record time span, the average CLARA-A3 monthly RSF MAB w.r.t. CERES-SYN1deg is estimated at 3.2 W/m², with 99.3% (>2σ) of the days below the threshold requirement and 72.4% (>1σ) below the target requirement.

5.2.2 Daily

Figure 5-5 shows that for 98.3% of these days, the MAB w.r.t. CERES SYN1deg-Day is below the threshold requirement of 16W/m², 80.4% is below the target requirement of 8W/m², and 0% below the optimal requirement of 4W/m². On average, the MAB amounts 6.2W/m².

Global MAB between daily CLARA-A3 RSF and other data records (*)
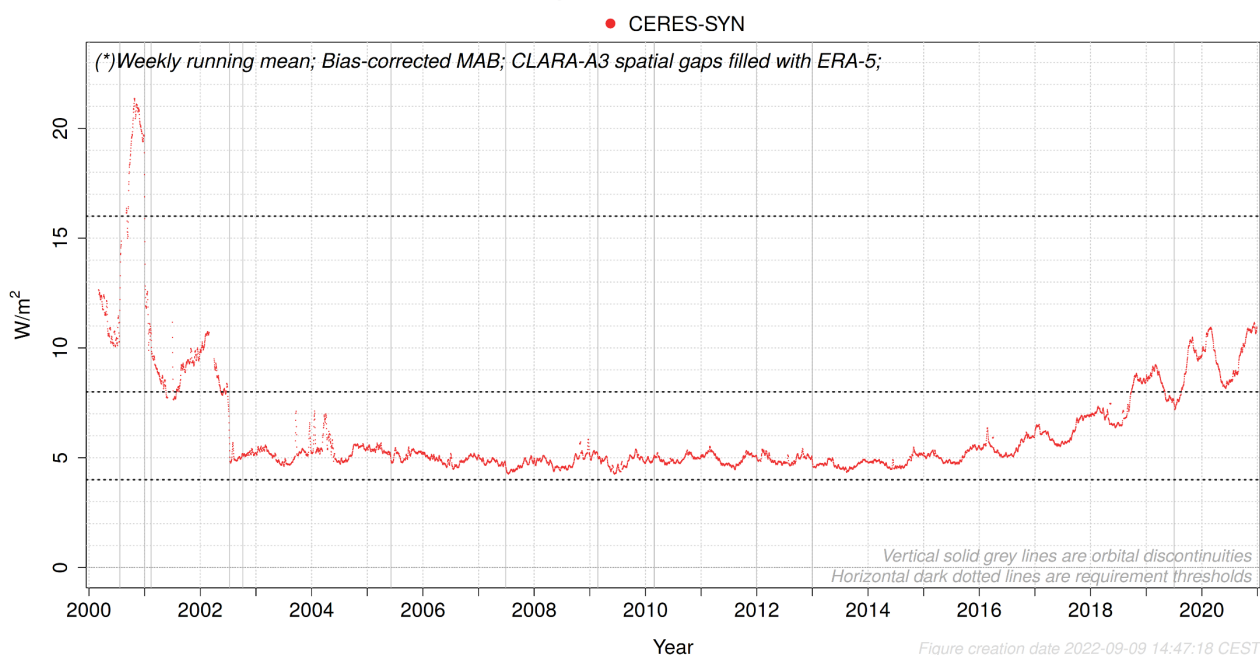
Figure 5-5: Global MAB between daily CLARA-A3 RSF and CERES-SYN1deg-Day

Similar to the monthly MAB, daily MAB exhibits large fluctuations with sharp delineations that are relatable to orbital configuration changes:

1. First half of 2000 with morning + late afternoon satellite, and an MAB of 10-13W/m²
2. Second half of 2000 with only late afternoon satellite, and an MAB of 19-21W/m²
3. Between 2001-mid2002 with morning+afternoon satellite, and an MAB of 8-10W/m²
4. Between mid2002-2016 with midmorning+afternoon satellite, and an MAB of 5W/m²
5. After 2016 with midmorning and drifting afternoon satellite, and an MAB increasing from 5 to 10W/m²

The time series of daily MAB (Figure 5-5) shares the same characteristics and features as observed for the monthly MAB (Figure 5-4) and described in previous Section 5.2.1, which are therefore discussed in the next common Section 5.2.3.

Similar to the monthly RSF record, the first half of the daily RSF record does not have a suitable reference data record to estimate the regional uncertainty, so again the MAB during the pre-CERES era is estimated by simulating the associated pre-CERES orbital configurations using CERES-era satellites. Annex 10.4 provides the full details, from which the results can be summarized as follows:

Table 5-3: Summary of Daily RSF uncertainty for the entire CLARA-A3 data record

Time span	Orbital configuration	Frequency of meeting the requirement (%)			Average MAB (W/m²)	Temporal share (%)
		Threshold 16 W/m²	Target 8 W/m²	Optimal 4 W/m²		
Pre-CERES era (1979/01-2000/02)	*(I) Afternoon satellite only*	*88.7%*	*0.0%*	*0.0%*	*13.5 W/m²*	*1891 days (12.4%)*
	(II) Morning satellite only	*35.3%*	*0.0%*	*0.0%*	*17.4 W/m²*	*701 days (4.6%)*
	(III) Afternoon and morning satellite only	*99.2%*	*15.7%*	*0.0%*	*10.4 W/m²*	*5155 days (33.7%)*
CERES era (2000/03-2020/12)	All available satellites	98.4%	81.2%	0.0%	6.2 W/m²	7534 days (49.3%)*
Total record	**All available satellites**	**94.5%**	**44.9%**	**0.0%**	**9.0 W/m²**	**15281 days (100%)***
(*) Three months are not taken into account: August 2000, June 2001 and March 2002, since CERES products contain data gaps in those months;						

The result is that for the entire data record time span, the average CLARA-A3 RSF MAB w.r.t. CERES-SYN1deg is estimated at 9.0 W/m², with 94.5% of the days below the threshold requirement.

5.2.3 Explanation of the time series' characteristics and features

For monthly and daily RSF, both the increased MAB as well as the introduction of its seasonality can be explained by a **degrading temporal coverage** over regions characterized by large-scale regional climate phenomena with an asymmetric diurnal cycle (e.g. marine stratus thinning or land convection). This introduces strong regional biases, as documented in Section 5.3.1, which can be positive or negative, depending on the region and the kind of phenomena. Furthermore, a degrading temporal coverage also introduces strong biases with fast moving small-scale or heterogeneous weather systems (e.g. fronts), typically consisting of swirls with positive alongside negative bias, caused by an extrapolation of e.g. the mid-morning observation to the afternoon (when the afternoon satellite has disappeared or drifted toward the evening), or simply put: the weather moves too fast to be accurately observed (Akkermans et al., 2021).

Globally averaged together, all these biases vary seasonally because of a hemispherical imbalance of the associated regional climate features' occurrence and strength, explaining the **seasonal pattern of MAB**. With any degradation of the temporal coverage (orbital configuration) these regional biases grow accordingly, which directly leads to an **increase of the global MAB** regardless the sign of these regional biases.

The sign of these regional biases is also the reason why the *global mean bias* during the CERES era remains relatively stable and without seasonal pattern (cfr. red curves in Figure 5-2 or Figure 10-10), namely because negative and positive regional biases tend to balance (compensate) each other; the global mean bias is only affected with really bad temporal coverage prevailing mainly during the pre-CERES era (and called "*suboptimal orbital configuration*"), as demonstrated in Annex 10.3.1 (cfr rectangles in Figure 10-10); the MAB on the other hand is much more sensitive and only

needs a gradually drifting satellite afternoon orbit (NOAA-19) to introduce this kind of seasonality and performance drop in its time series (Figure 5-4, Figure 5-5).

In addition, it is worth mentioning that observations with low illumination conditions (high solar zenith angle), prevalent close to the terminator, lead to a larger processing error, for instance due to the increased uncertainty of scene type defining parameters (cloud mask, cloud optical thickness, cloud phase,..) which propagates as uncertainty in the narrowband-to-broadband and ADM processes. This effect is also tied to the orbital configuration, as orbital drift typically increases the average solar zenith angle for a given location.

5.2.4 Explanation of the difference between monthly and daily MAB

Besides the overall characteristics and features, the daily MAB is generally higher compared to the monthly MAB. The reason is bias compensation, on different levels and scales.

Firstly, there is a temporal sampling compensation: biases caused by fast moving small-scale or heterogeneous weather systems (e.g. broken cloud fields) vary in sign from day to day, depending on the weather system's morphology and movements (direction, timing, speed..). The aggregation to a monthly mean bias smooths out this daily variability.

Secondly, there are numerous error sources related to the retrieval of instantaneous TOA albedo, which are propagated to the daily mean RSF (and the less satellites, the stronger this propagation). However, averaged over 30 days many of these errors tend to cancel each other out. Examples are the errors related to the ADM (viewing and illumination geometry change every day, this in contrast to geostationary observations) and errors related to scene type identification such as cloud cover and cloud properties (relevant for ADM but also for narrowband-to-broadband conversion, etc). According to the terminology outlined in Section 4.1.3, these kind of compensating errors could for a large part be considered as the random component of the processing error ('precision part'), characterized by the daily MAB, whereas the errors that are still detected in the monthly MAB could be considered the processing error's systematic component ('accuracy part').

5.3 Regional comparison (geographical distribution)

5.3.1 Annual Means

Annual means of CLARA-A3 RSF bias w.r.t. CERES-SYN1deg are shown in Figure 5-6 for the years 2000-2020. Between 2003-2016 the biases are generally relatively low (within +/- 2W/m²) in most regions, with some regions showing systematically (slightly) larger biases, in both negative sense (*bluish colors; e.g. ocean west of African continent, Antarctica, eastern Canada,..*) and positive sense (*reddish colors; e.g. non-desert African and South-East Asian land masses*).

Before 2003, roughly the same bias patterns occur, but much stronger. In addition, also the ocean region west of South America exhibits underestimation. In 2001 and 2002 there is also an artefact visible south of Australia (artificial cluster of pixels with overestimation), probably related to geostationary satellite fusion in the CERES-SYN product.

After 2016, some other bias patterns become apparent. Most remarkable here is the increasing overestimation (red colors) in the ocean region west of South American and African land masses (in contrast to the underestimation in these very regions before 2003).

The next paragraphs discuss the mechanisms behind some of the largest observed biases are discussed.
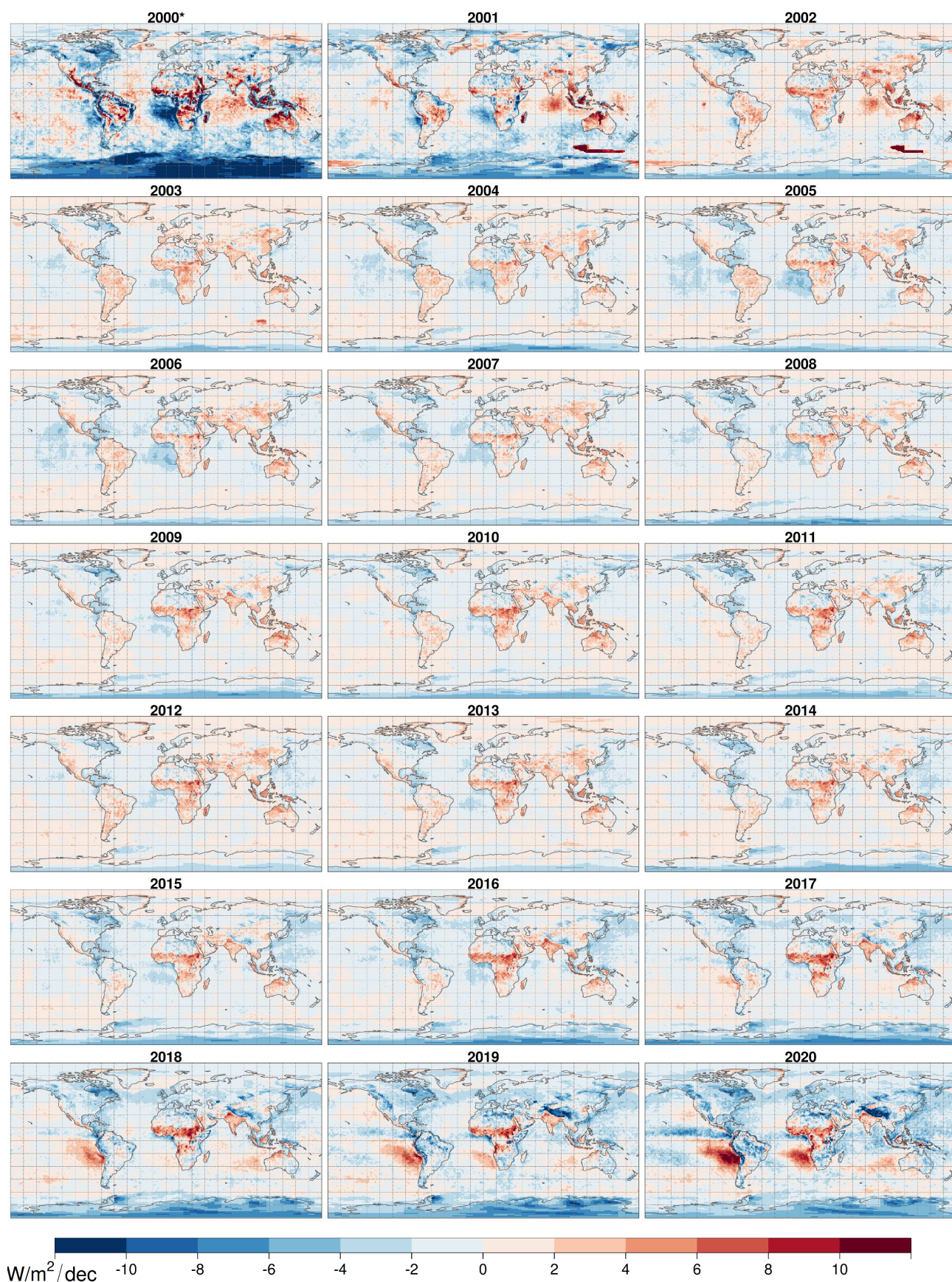
Figure 5-6: Bias of annual CLARA-A3 RSF w.r.t. CERES-SYN from 2000-2020. (*) The year 2000 is only March to December.

The orbital configuration affects the frequency and timing of the observations throughout the day, leading to different kinds of biases (e.g. bias due to absence of afternoon satellite is not necessarily the same as bias due to absence of morning satellite). To obtain the daily mean RSF, a SZA-dependent albedo model (a.k.a. shortwave diurnal cycle model) extrapolates the observations symmetrically w.r.t. noon (cfr [RD 7]). Hence, the sensitivity of the daily mean RSF bias is related to the asymmetry of the actual observed diurnal cycle (*w.r.t. noon*), which, in turn, depends on specific regional climatological phenomena.

For this, an important example consists of the typical **marine stratocumulus** regions located over the eastern parts of tropical ocean basins (Eastman and Warren, 2014), where clouds persist in the morning and dissolve (i.e. become thinner, 'burn off') during the afternoon. As a result, the absence of an afternoon observation causes the thick morning cloud to be wrongly extrapolated into the afternoon, resulting in a positive daily mean RSF bias; this effect can be clearly seen in the ocean regions west of Africa and South America during 2019-2020 (Figure 5-6) and is due to orbital drift of NOAA-19 starting in 2016 (Figure 4-10). The opposite is true in case of a missing morning observation, as is the case before 2002 (Figure 4-10), whereby the thin afternoon cloud is wrongly extrapolated into the morning, leading to a negative daily mean bias; this effect can be seen in the same regions during 2000-2001 (Figure 5-6). Even if they vary seasonally, the aforementioned biases are strong enough to be visible in the annual mean biases. More details about these marine stratocumulus-related biases are provided in Annex 10.5.1.

Besides this example of thinning marine stratocumulus, there are also other regional climate types with their own characteristic cloud diurnal cycle, for instance land regions with high convective activity, which has typically an opposite phase (cloud peak in afternoon).

Finally, we notice for suboptimal orbital configurations with only a late afternoon orbit, that regions with predominantly snow and ice scene types (Antarctic main land and surrounding sea ice) are characterized by significant RSF underestimations during summertime (months DJF for Antarctica). These seasonal errors are large enough to propagate to the annual means, as seen for the year 2000 (only NOAA-14 orbit). This phenomenon occurs mostly when SZA has grown bigger, i.e. after some orbital drift already took place. In these conditions, the retrieved CLARA-A3 cloud parameters deviate a lot from the CERES cloud parameters. More details about this bias are provided in Annex 10.5.2.

5.3.2 Comparison of regional trends

This section seeks to validate the regional trends in RSF. Given some large regional biases in the years 2000-2002 and 2017-2020 (Section 5.3.1), mostly due to orbital configuration of CLARA-A3, the selected period of regional trend comparison is 2003-2016. It should not be considered a comprehensive climatological trend analysis (i.e. science about the physical climate mechanisms).

Figure 5-7 shows the linear trend of RSF between 2003-2016, expressed in W/m²/decade, comparing CLARA-A3 and CERES EBAF. CLARA-A3 performs well as the observed patterns are very similar to CERES. The patterns are roughly similar to the ones discussed in Loeb and Doelling (2020) (spanning a period between 2002-2019), and contain broad regions of significant trends (above the 95% confidence level) including Antarctica, parts of the Southern Ocean, the North Atlantic, and off the west coast of North America. The positive trend over Eastern China is possibly related with increasing air pollutant emissions (so-called *surface dimming* because less sunlight reaches the

surface), however this would contradict with studies observing the opposite trend since measures were taken by the government (which should cause a decreasing RSF trend, and *surface brightening*). Loeb and Doelling (2020) concluded that further modeling studies are needed to determine how much of these trends are associated with climate forcing as opposed to internal variability in the climate system.
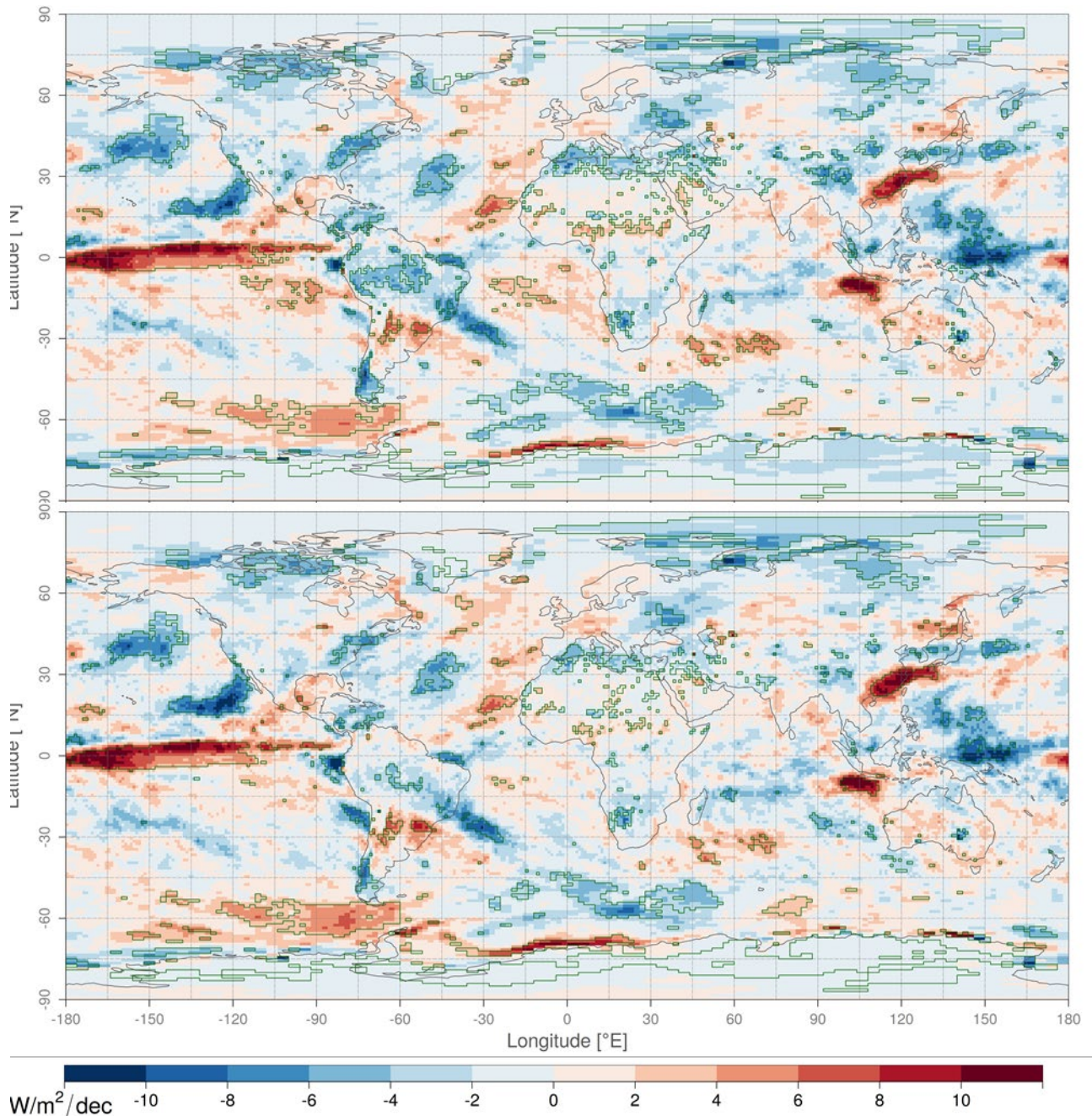
Figure 5-7: Linear trend of monthly RSF between 2003-2016 for CLARA-A3 (top) and CERES EBAF (bottom). Areas delineated in green correspond to trends that exceed the 95% confidence interval.

5.4 Evaluation of ICDR RSF products

The global mean daily RSF (absolute flux) during the year 2020 is shown in Figure 5-8: The red curve (ICDR) has a similar day-to-day variation compared to the black curve (TCDR), although there are differences: the RSF in the ICDR is generally slightly lower, with an average of -0.23 W/m² (Figure 5-9), which is about 0.23%. This difference is mostly caused by the exclusion of Metop-C from the ICDR (the difference due to changing NWP and sea ice input data is an order of magnitude smaller).
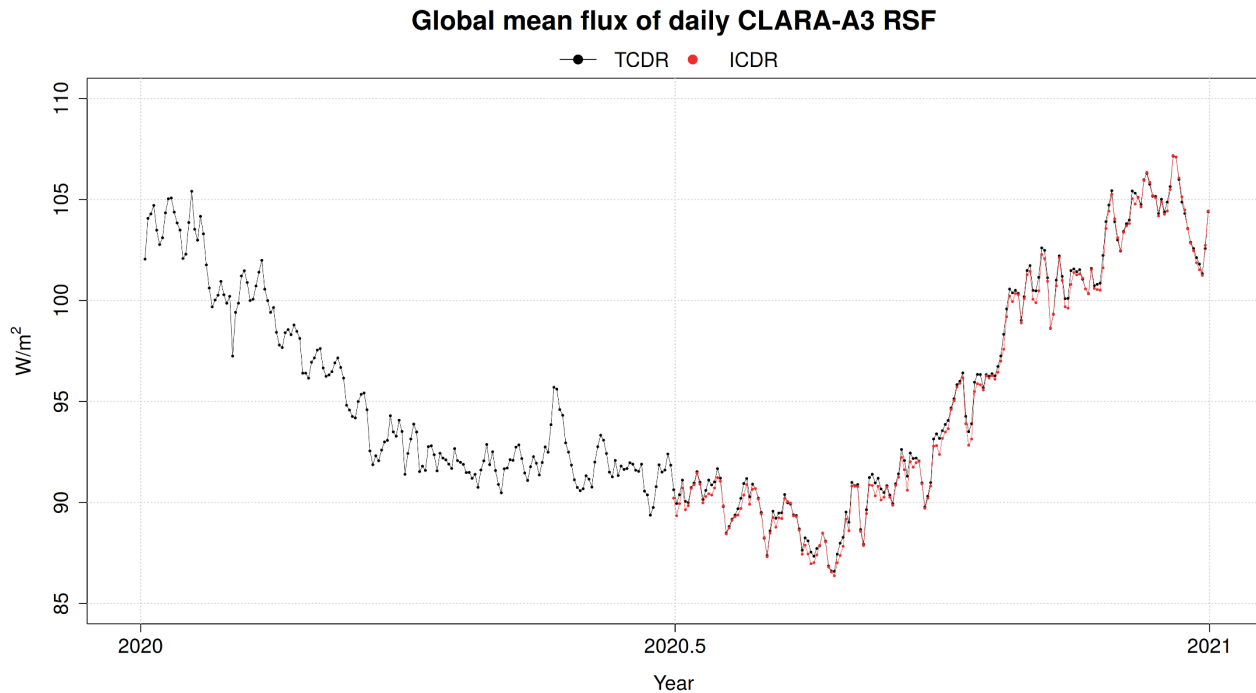
Figure 5-8: Global mean flux of daily CLARA-A3 RSF generated with TCDR and ICDR

The global MAB of daily CLARA-A3 RSF w.r.t. CERES is shown in Figure 5-10, showing that the TCDR and ICDR versions of the CLARA-A3 RSF are similar in terms of processing error (regional uncertainty). The exclusion of Metop-C in the ICDR (2nd half of 2020) leads to a bit higher MAB (w.r.t. CERES SYN) because of the lower temporal coverage (especially in this period without afternoon orbit). However, the calibration error for Metop-C is growing much larger in the subsequent years, which by then cannot be compensated anymore by its added value for temporal coverage. Their difference in MAB ranges between 0 and 1 W/m² (~9%). The MAB cycle of 10 to 12 days in Figure 5-10 is caused by the shifting swath alignment of Metop-C with Metop-B and NOAA-18.

CLARA-A3 Global mean RSF: difference between ICDR and TCDR
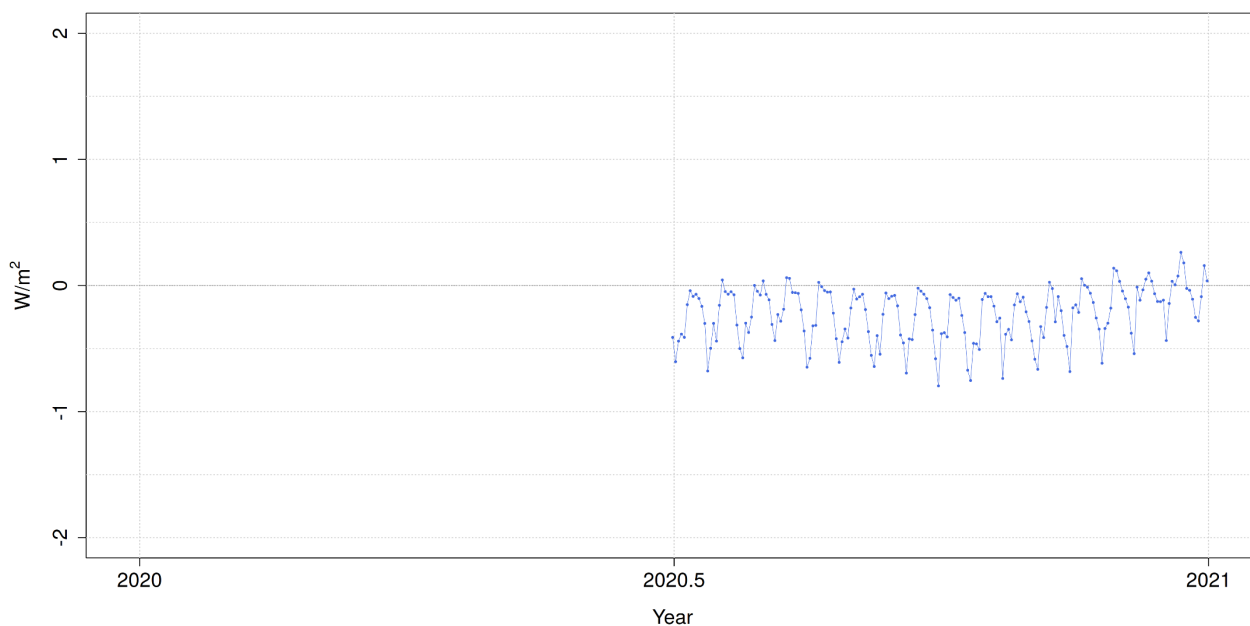
Figure 5-9: Global mean flux of daily CLARA-A3 RSF : difference between TCDR and ICDR. For comparison, the scale range is identical to the OLR equivalent in Figure 6-12.

Global MAB between daily CLARA-A3 RSF and CERES SYN
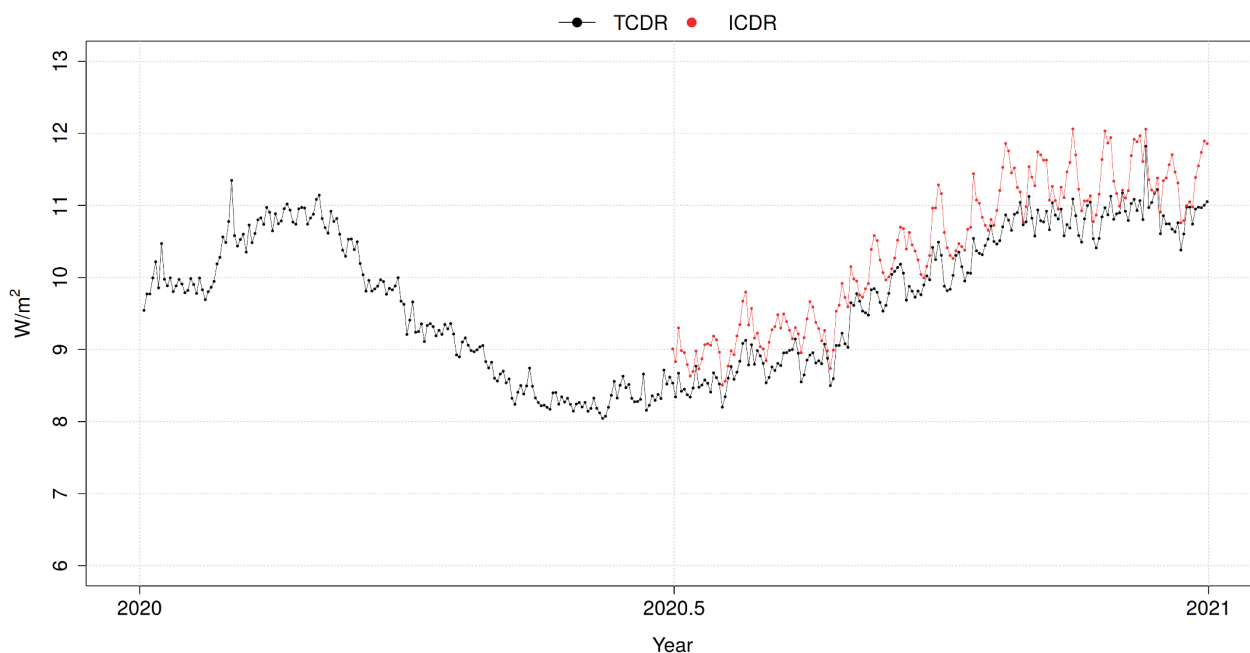
Figure 5-10: Global daily MAB between CERES-SYN1deg-Day and CLARA-A3 RSF generated with TCDR and ICDR. For comparison, the scale range is identical to the OLR equivalent in Figure 6-13

TOA outgoing shortwave flux filled
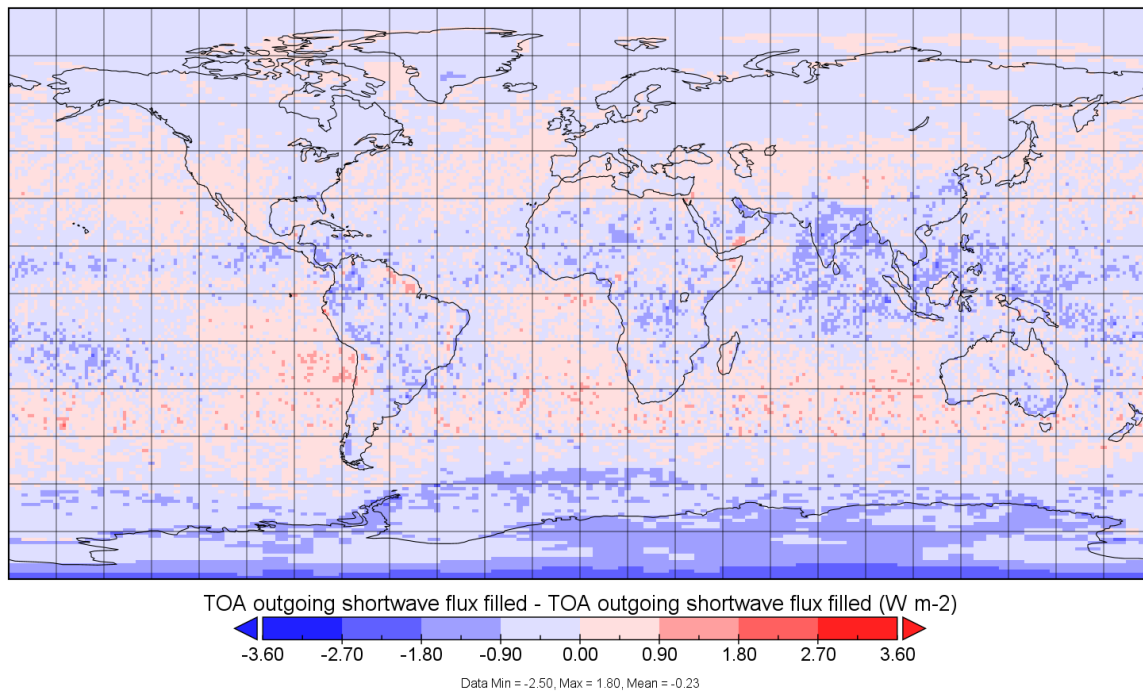
Figure 5-11: CLARA-A3 RSF 6-month mean RSF difference between TCDR and ICDR. For comparison, the scale range is identical to the OLR equivalent in Figure 7-14.

A spatial comparison between TCDR and ICDR is provided in Figure 5-11. The global mean difference is small (-0.23 W/m²) and its spatial distribution is relatively homogeneous, although there are noticeable patterns: the RSF becomes more negative in the tropics, more positive in the subtropics, and again more negative over Antarctica. Note that this is the mean for only 6 months, in which the southern hemisphere receives (and hence reflects) most sunlight. The results would probably be mirrored in the north for the other six months. As mentioned before, these patterns are predominantly caused by the exclusion of Metop-C in the ICDR, and only very slightly impacted by the changed ICDR input data (NWP and sea ice).

Overall, we can conclude that the switch from TCDR to ICDR has a limited impact, given their absolute difference which is relatively small compared to the TCDR's bias variability w.r.t. other reference data records (Figure 5-2) and given the difference in their regional uncertainty (MAB) which is relatively small compared to the TCDR's absolute MAB (Figure 5-5 and Figure 5-10). Furthermore, it does not cause a violation of the threshold requirement for stability (since the ICDR's bias does not fall outside the stability envelope), nor for the processing error (since the ICDR's MAB does not exceed the level of 16 W/m²).

6   Results for Outgoing Longwave Radiation (OLR)

6.1   Mean bias and Stability

The global mean flux of monthly OLR from different data records is shown in Figure 6-1, among which CLARA-A3 OLR in orange.
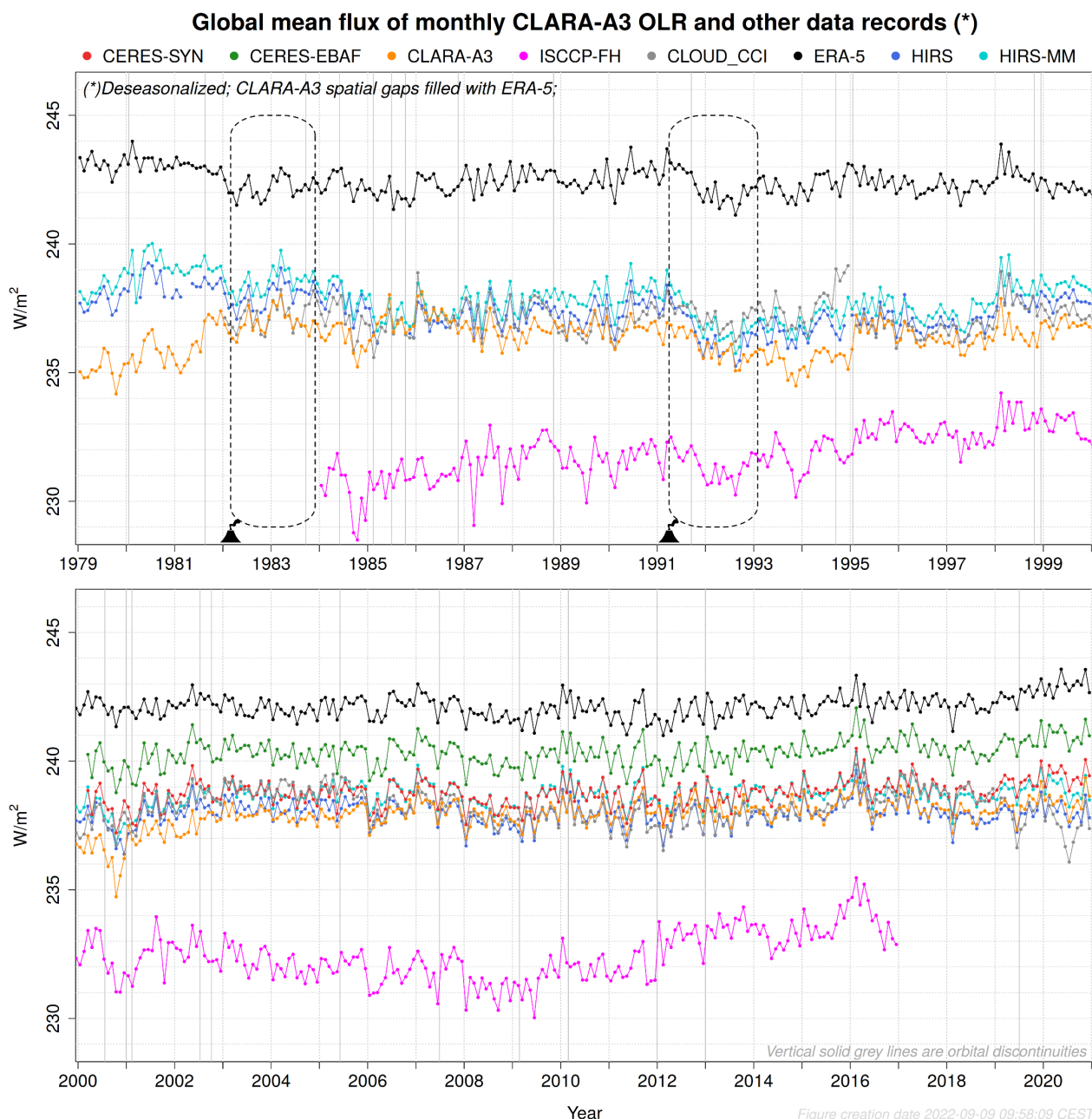
Figure 6-1: Global mean flux of monthly CLARA-A3 OLR and other data records

The HIRS and ERA5 data records are stable with respect to each other, increasing the confidence in their ability to serve as stability benchmark for the other data records. Volcanically induced aerosols also trap thermal radiation, but its total radiative impact is less compared to the shortwave radiative impact (Canty et al., 2013), shown in Section 5.1, so that the net effect is a surface cooling ('dimming'). The two major volcanic eruptions El Chichón and Pinatubo are indicated on the time

series in Figure 6-1. The El Chichón eruption has no clear impact in the CLARA-A3 data record (but it might have caused a small drop of -0.5 W/m² in other data records), whereas the Pinatubo event probably caused a drop in OLR of approximately -1 W/m², which is about half the assumed impact as seen in the HIRS OLR data records (-2 W/m²). Overall, and for most data records, these radiative impacts are almost similar to many other drops and jumps in the time series, making it difficult to assess and quantify them.

The CLARA-A3 OLR trend during 2000-2020 amounts +0.29 W/m²/decade(dec), compared to the CERES EBAF trend of +0.43 W/m²/dec. Combined with the RSF trend (Section 5.1) and TSI trend during that period (-0.07W/m²/dec), the Net flux at TOA is estimated by CLARA-A3 at +0.48 W/m²/dec which is relatively close to the CERES EBAF estimate of +0.36 W/m²/dec. When the Metop-B issue (Section 5.1) is solved in future versions of the data record, the CLARA-A3 RSF trend would be less negative and hence the Net flux at TOA would be closer to CERES EBAF.

The global mean bias is calculated by subtracting the reference data records from CLARA-A3 OLR, resulting in the time series shown in Figure 6-2. The overall stability of CLARA-A3 OLR is assessed w.r.t. **HIRS**, and this is done by creating a so-called '*stability envelope*' (using the threshold requirement of 4W/m² as a reference), which is set to an arbitrarily chosen optimal range of [-3.2 ; +0.8] W/m² because the bias is not normally distributed (Figure 6-3). **The overall stability complies with the threshold requirement, as almost all (99.6%) the data is contained within the 4W/m² stability envelope.** The daily mean analysis is not shown here, because the bias' magnitude and fluctuations are very similar and are not affected by the temporal aggregation. For the sake of completeness, they are provided in Annex 10.2.2.

Global mean bias of monthly CLARA-A3 OLR w.r.t. other data records (*)
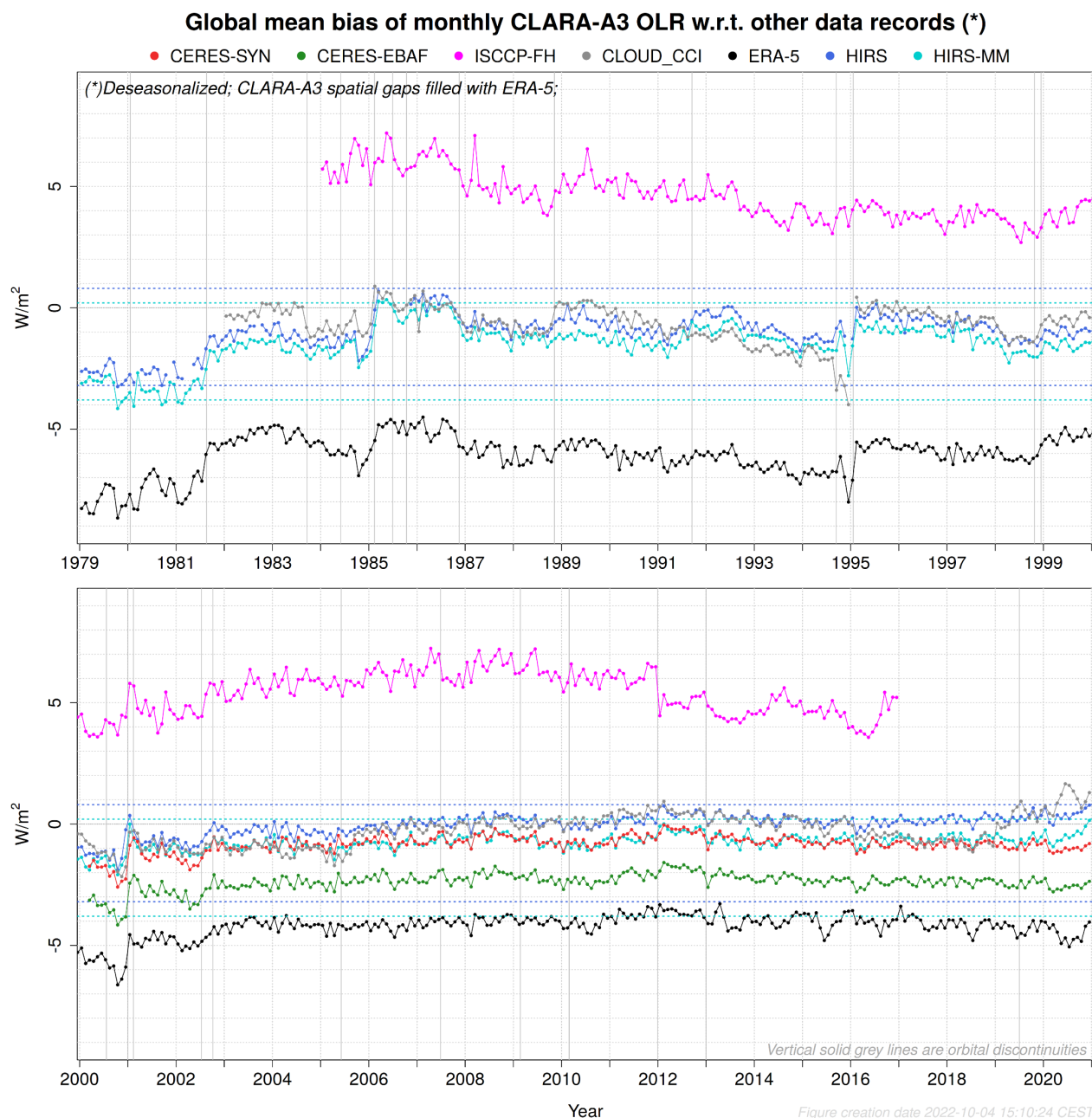
Figure 6-2: Global mean bias of monthly CLARA-A3 OLR w.r.t. other data records. Dotted lines indicate a stability envelope of 4W/m² around the bias w.r.t. HIRS and HIRS-MM

This histogram shows a bimodal pattern, with the first small peak representing the first few years of the records with a distinctly lower mean bias than the rest of the record. This period with distinctly lower fluxes corresponds with the TIROS-N and NOAA-6 satellites (January 1979 – August 1981) and has an average bias of -2.5W/m², which is markedly lower compared to the mean bias between 1982-2002 (around -1W/m²) and between 2002-2020 (around 0W/m²). Some tests were done to investigate this issue, for instance verifying whether the AVHRR instrument (AVHRR/1 vs AVHRR/2) could play a role (the first has one less thermal channel), or the ECT (morning vs afternoon satellite).

Distribution of monthly OLR bias from CLARA-A3 w.r.t. HIRS
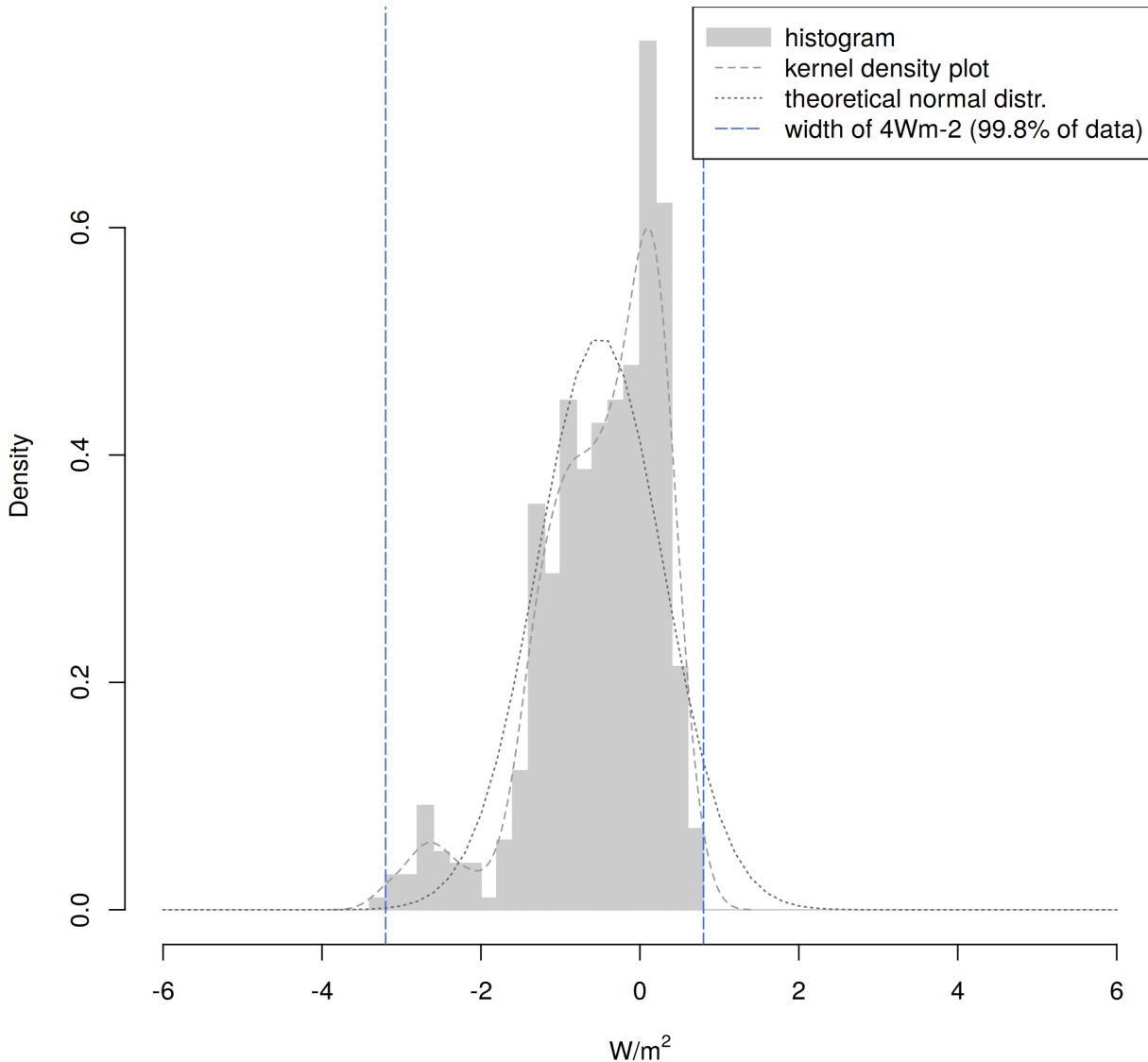
Figure 6-3: Histogram of monthly OLR bias from CLARA-A3 w.r.t. HIRS (1979-2020)

The results are provided in Annex 10.6 and indicate that switching from all-satellite to morning-only orbital configuration only has a small impact (about -0.5W/m² on average), not enough to explain the much more negative bias during the morning-only configuration during 1980-1981 (from NOAA-6). Hence, the latter can be assumed to be caused by the underlying L2b data record which in turn is likely impacted by an issue with the calibration of the FDR, or an issue with the spectral response correction factors.

6.2 Processing error (regional uncertainty)

This validation report measures the processing error (regional uncertainty) with the *Mean Absolute Bias (MAB)* as described in Section 4.2. Nevertheless, all the analyses are also performed with the *Root Mean Square Bias (RMSB)* and available in Annex 10.9.2.

6.2.1 Monthly

Figure 6-4 shows that for 99.8% of the record's time span, the MAB w.r.t. HIRS is below the threshold requirement of 8W/m², 99.2% is below the target requirement of 4W/m², and 77.2% below the optimal requirement of 2W/m². On average, the MAB w.r.t. HIRS amounts 1.8W/m². Many of the missing months in Figure 6-4 are due to temporal gaps in the HIRS data record, for which only monthly means were created when all daily means were found valid. In contrast to the daily MAB (see next Section 6.2.2) the monthly MAB does not exhibit large fluctuations with delineations relatable to e.g. orbital configuration changes. There are however subtle distinctions between roughly 4 periods: (1) between 1979-1981 with suboptimal orbital configuration from the first two satellites TIROS-N and NOAA6 and an MAB of 2.5-3.5W/m²; (2) between 1982-1986 with a suboptimal orbital configuration from satellites NOAA-7/-8/-9 and an MAB of 2.0-2.5W/m²; (3) between 1987-2000 with mostly morning+afternoon satellites and an MAB of 1.5-2.0W/m²; and (4) between 2001-2020 with midmorning+afternoon satellites and an MAB of 1.25-1.75W/m².

The reasons for the dependency of the monthly MAB on the orbital configuration are identical to the ones for RSF, as described in Section 5.2.3. However, the MAB of monthly OLR is much less sensitive to orbital configuration compared to RSF (compare Figure 5-4 with Figure 6-4): the absence of the midmorning orbit NOAA-17 (before mid-2002) and the orbital drift of the afternoon orbit NOAA-19 (after 2016) both have almost no impact on monthly MAB, with a quasi constant MAB of around 1.5W/m² between 2001-2020. On the other hand, large degradations in orbital configurations do have an impact, for instance the late-afternoon-only configuration in the second half of 2000, causing the MAB to double (4W/m²); however, these impacts are still small compared to RSF, where the same degradation leads to a quadrupling of MAB (Figure 5-4). There are multiple reasons for this, for instance the intra-day relative range which is much lower for OLR (than for RSF), thereby lowering the impact of wrong temporal extrapolation due to suboptimal temporal coverage. Another reason is the number of observations per day, which for OLR is double (compared to RSF) because it can also rely on nighttime observations, which again lowers the impact of suboptimal temporal average on the daily mean integration (compared to RSF).

Global MAB between monthly CLARA-A3 OLR and other data records (*)

● CERES-SYN ● CERES-EBAF ● HIRS
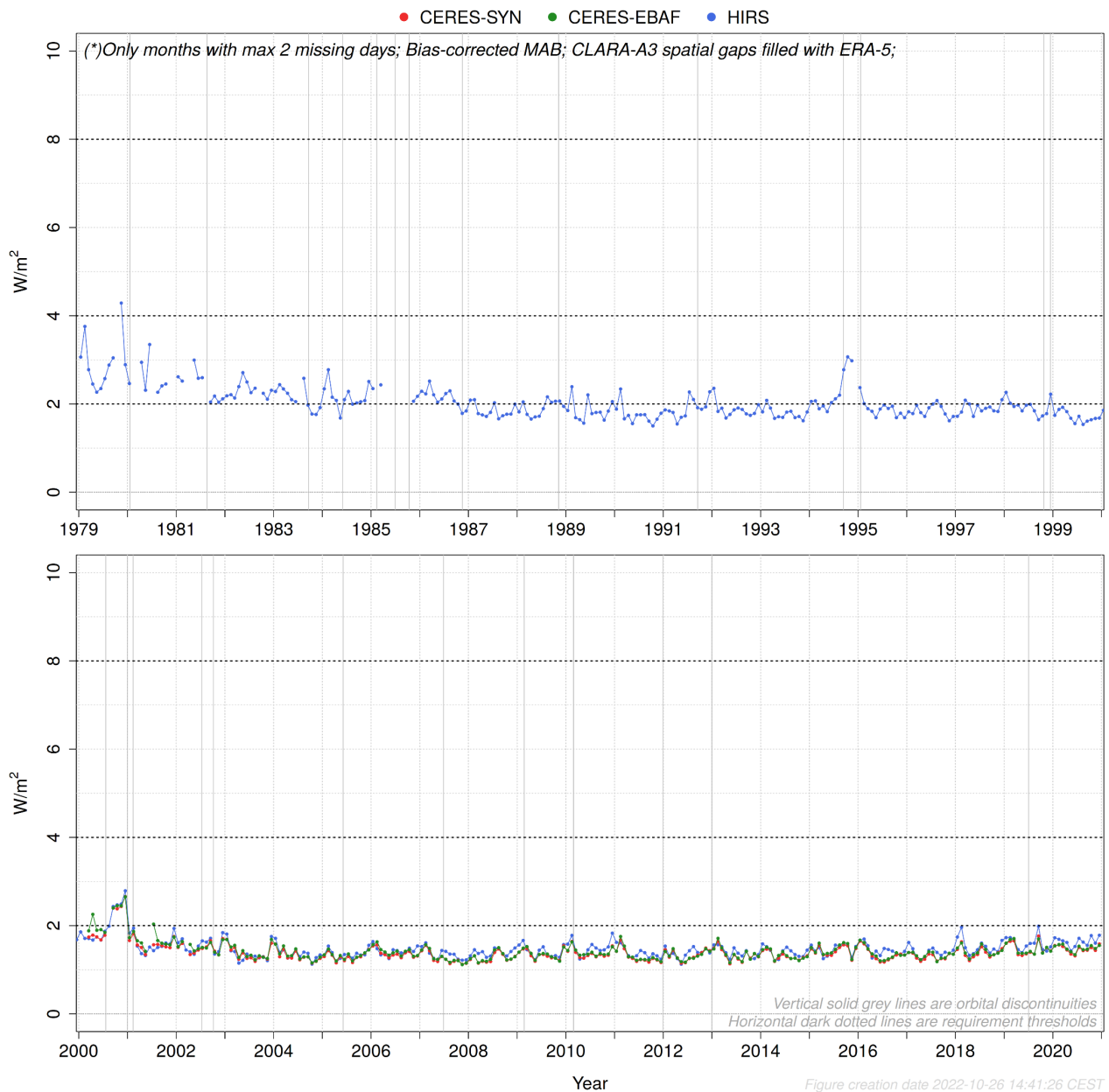
Figure 6-4: Global MAB between monthly CLARA-A3 OLR and other data records

6.2.2 Daily

Figure 6-5 shows that for 99.6% of the record's time span, the MAB w.r.t. HIRS is below the threshold requirement of 16W/m², 97.1% is below the target requirement of 8W/m², and 37.1% below the optimal requirement of 2W/m². On average, the MAB w.r.t. HIRS amounts 4.8W/m². The daily MAB exhibits larger fluctuations with clear delineations that are relatable to orbital configuration changes:

1. between 1979-mid1983 and mid1984-1986 with suboptimal orbital configurations, i.e. morning-only or afternoon-only satellite, and an MAB of 6-8W/m²;
2. between 1987-2002 with mostly morning+afternoon satellites and an MAB of 4-6W/m², slightly varying according to orbital drift;

3. peaks during 1995 and 2000 with respectively an only-early-morning and an only-late-afternoon satellite, and an MAB of around 8W/m²;

4. between 2002-2016 with midmorning+afternoon satellites and an MAB of 3.7W/m²

5. after 2016 with midmorning + drifting afternoon satellite and an MAB increasing to 4.5W/m².
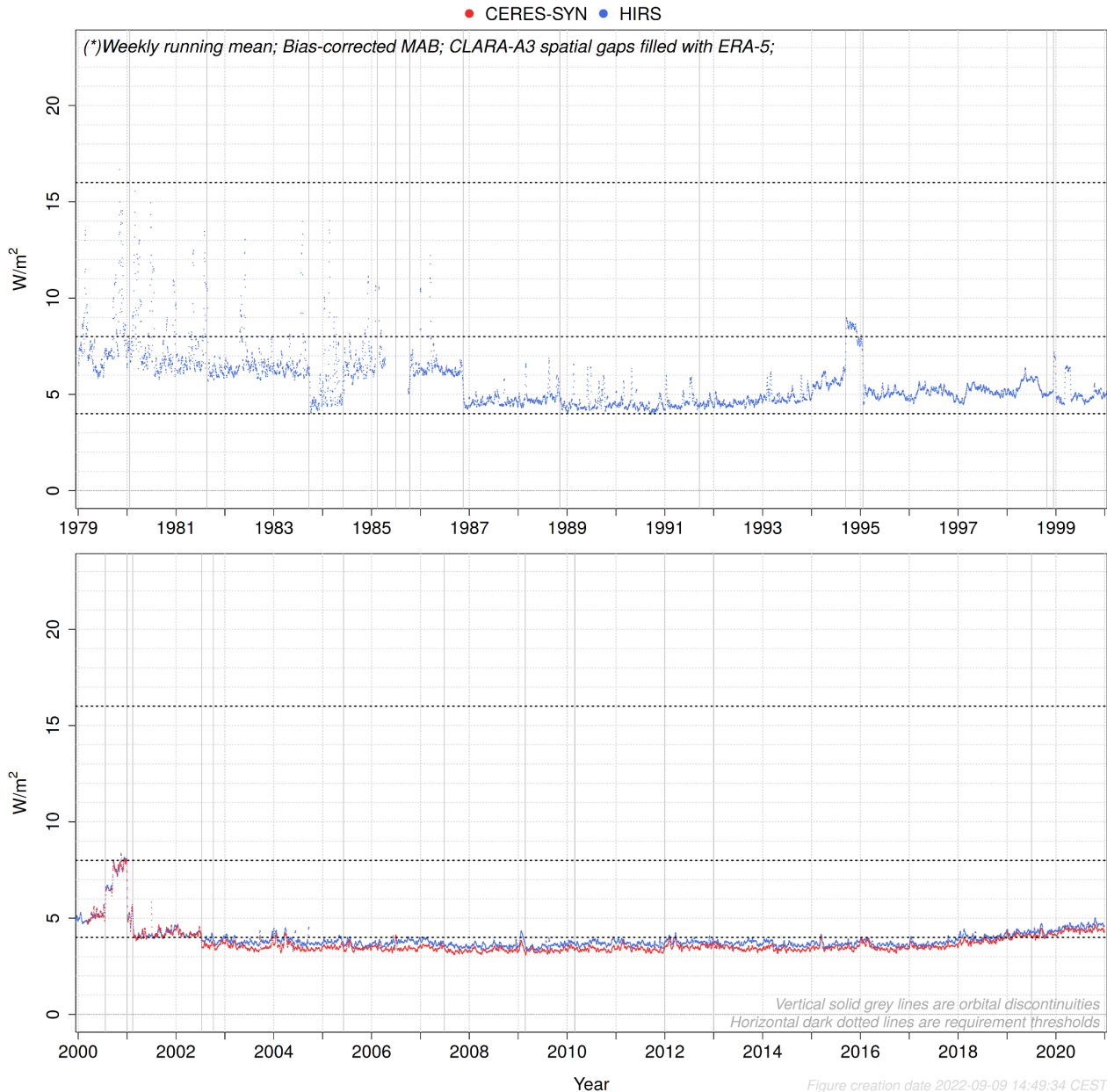
Figure 6-5: Global MAB between daily CLARA-A3 OLR and other data records

The daily MAB is much less sensitive to orbital configuration compared to RSF (compare Figure 5-5 with Figure 6-5): the absence of the midmorning orbit NOAA-17 (before mid-2002) and the orbital drift of the afternoon orbit NOAA-19 (after 2016) both have only a small impact on OLR (+0.5 to +1.0W/m² MAB), which has a quasi constant MAB of around 3.7W/m² between 2002-2016. Large degradations in orbital configurations have a bigger impact, for instance the late-afternoon-only configuration in the second half of 2000, causing the MAB to double (to 8W/m²); however, these impacts are still small compared to RSF, where the same degradation leads to a quadrupling MAB

(Figure 5-5). The reasons for this difference between RSF and OLR were already discussed for the monthly MAB in Section 6.2.1 (intra-day relative range and number of observations per day).

6.3 Regional comparison (geographical distribution)

6.3.1 Annual Means

Figure 6-6 and Figure 6-7 show respectively the annual biases of CLARA-A3 OLR w.r.t. HIRS (1979-1999) and w.r.t. CERES-SYN1deg (2000-2020).

Between 2003-2020 (Figure 6-7) the biases w.r.t. CERES are in generally low in most regions (within +/- 2W/m²), with some regions showing systematically (slightly) larger biases, in both negative sense (*bluish colors; e.g. ocean west of South America*) and positive sense (*reddish colors; e.g. desert regions in North Africa and Middle East*).

In most regions the OLR bias during the years 2000-2002 is larger (similar to what is observed for RSF), which is related to the degraded orbital configuration before 2003, as explained for RSF in Section 5.3.1. However, the OLR bias degradation in the year 2000 is in most regions limited to +/- 4W/m², which is much less compared to RSF's bias in the year 2000 (compare Figure 5-6 and Figure 6-7).

The OLR bias is (unlike the RSF bias) not significantly impacted by the changing orbital configuration after 2016, characterized by NOAA-19's orbital drift from afternoon towards evening orbit. Some small effects can be noticed, such as the changing bias over the desert regions, from overestimation (red colors) to neutral or slightly underestimation (bluish colors).

In the biases w.r.t. HIRS OLR (Figure 6-6), some HIRS-related artefacts are visible, such as the ring-like features around -75°E and 75°E longitudes (edges of geostationary satellite FOV). The significantly larger bias in the year 1985 is related to missing data in the HIRS data record (april-october). Apart from these features, however, the bias of CLARA-A3 w.r.t. HIRS OLR looks relatively good in the majority of the regions and years during the pre-CERES era (1979-1999). The first three years suffer from a negative bias across most regions of the globe, consistent with the findings in Setion 6.1.
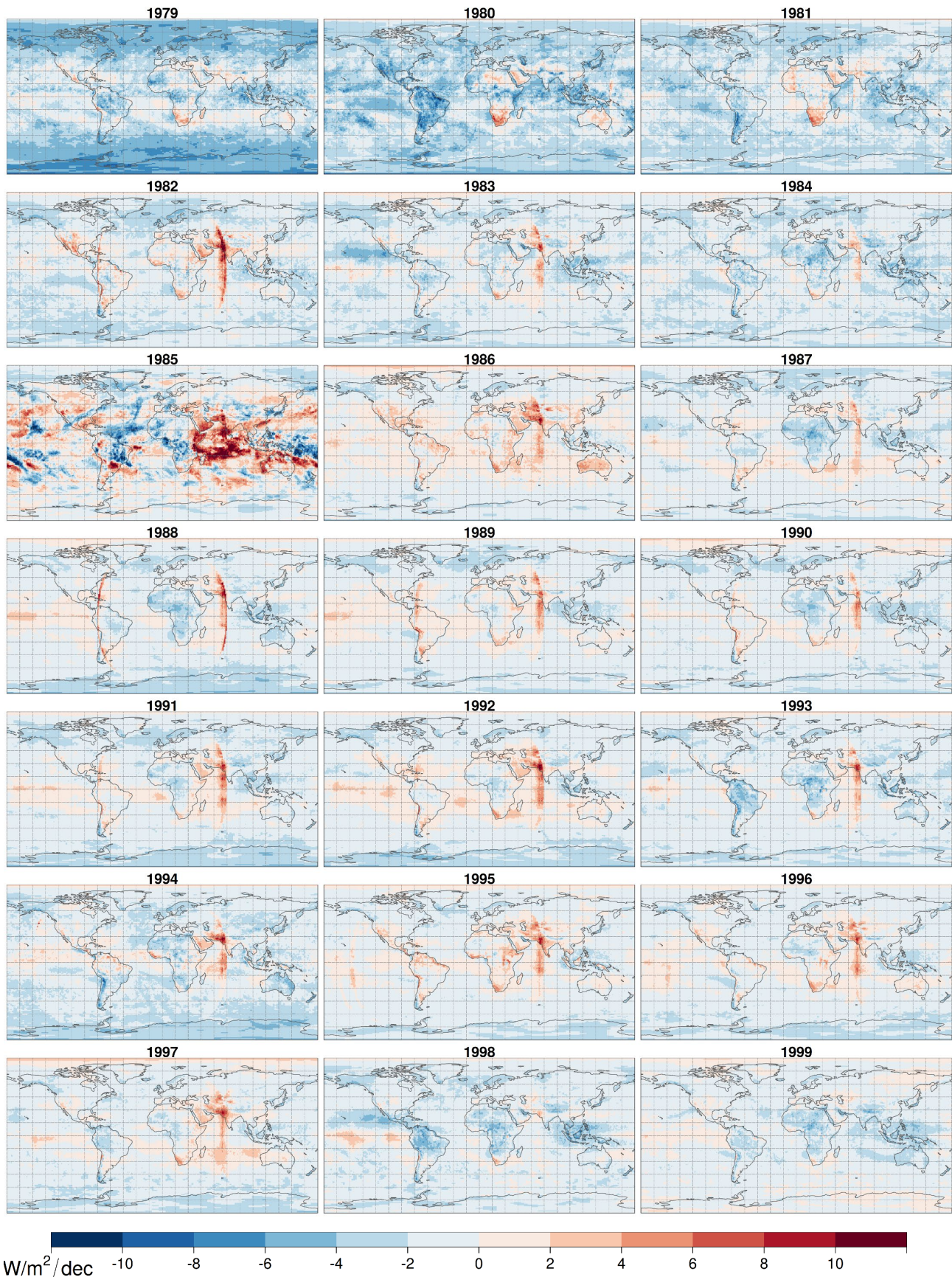
Figure 6-6: Bias of annual CLARA-A3 OLR w.r.t. HIRS-OLR from 1979 to 1999.
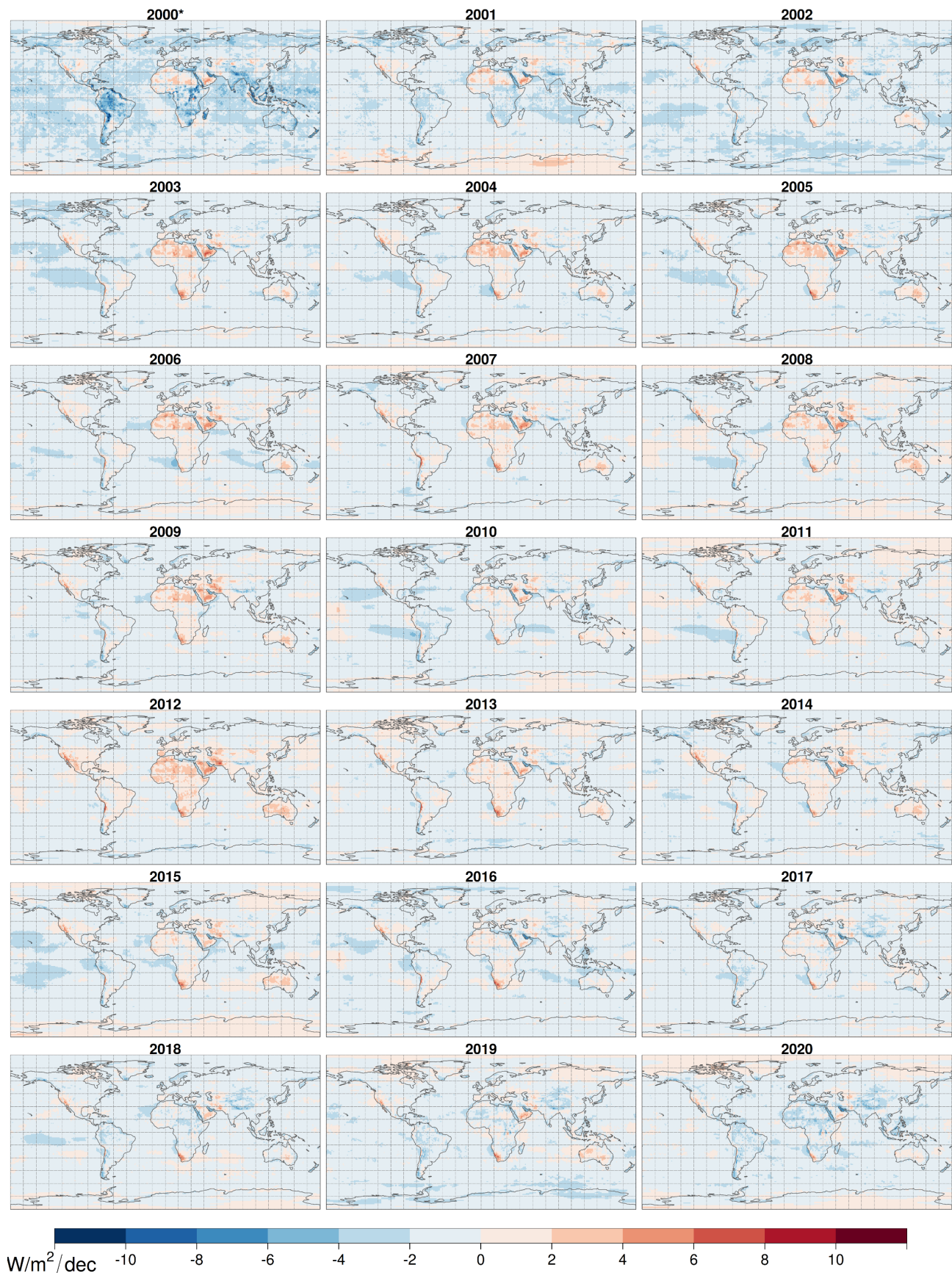
Figure 6-7: Bias of annual CLARA-A3 OLR w.r.t. CERES-SYN from 2000 to 2020. (*) The year 2000 only contains March to December.

6.3.2 Comparison of regional trends

This section seeks to validate the regional trends in OLR. It should not be considered a comprehensive climatological trend analysis (i.e. science about the physical climate mechanisms).

For the CERES era (period 2000-2020), the OLR trends from CLARA-A3 and CERES-EBAF are shown in Figure 6-8. The regional trends are well captured by CLARA-A3, given the good similarities between both maps. Analogous to RSF, it is not yet clear how much of these trends are associated with climate forcing as opposed to internal variability in the climate system.
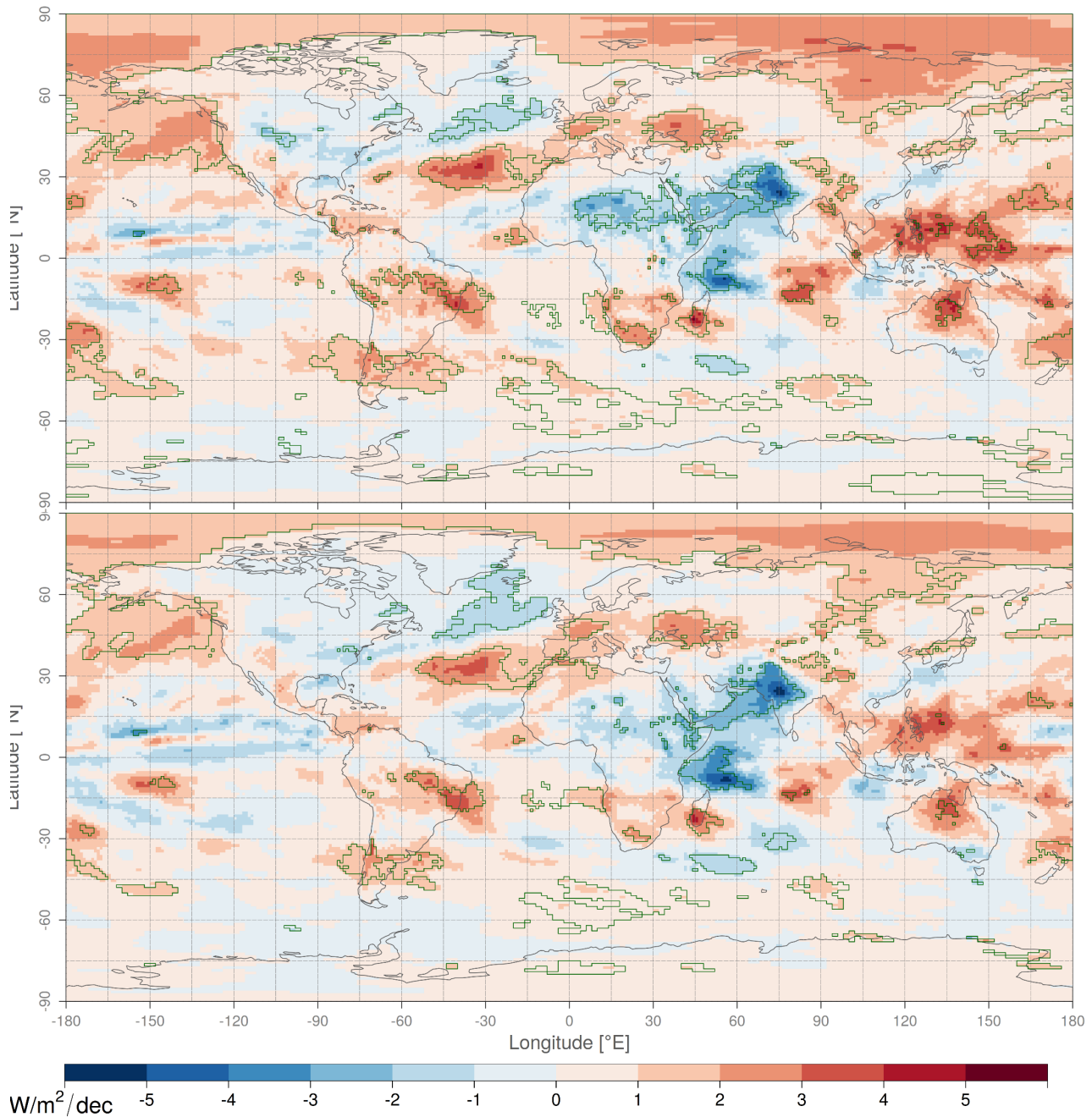
Figure 6-8: Linear trend of monthly OLR between 2000-2020 for CLARA-A3 (top) and CERES EBAF (bottom). Areas delineated in green correspond to trends that exceed the 95% confidence interval.

Trends for the entire data record (1979-2020) are given in Figure 6-9. Apart from the positive trend in the Arctic region, most regions are characterized by more modest trends compared to the period 2000-2020, as expected. As the time series is long enough (42 years), these regional trends definitely tell us something about real climatological trends. And also here CLARA-A3 OLR performs well given the good correspondence with the trends observed in the HIRS data record. Some of the HIRS-related artefacts (seen in the annual means) also affect the regional trends, such as the ring-like FOV-edge around 75°E.
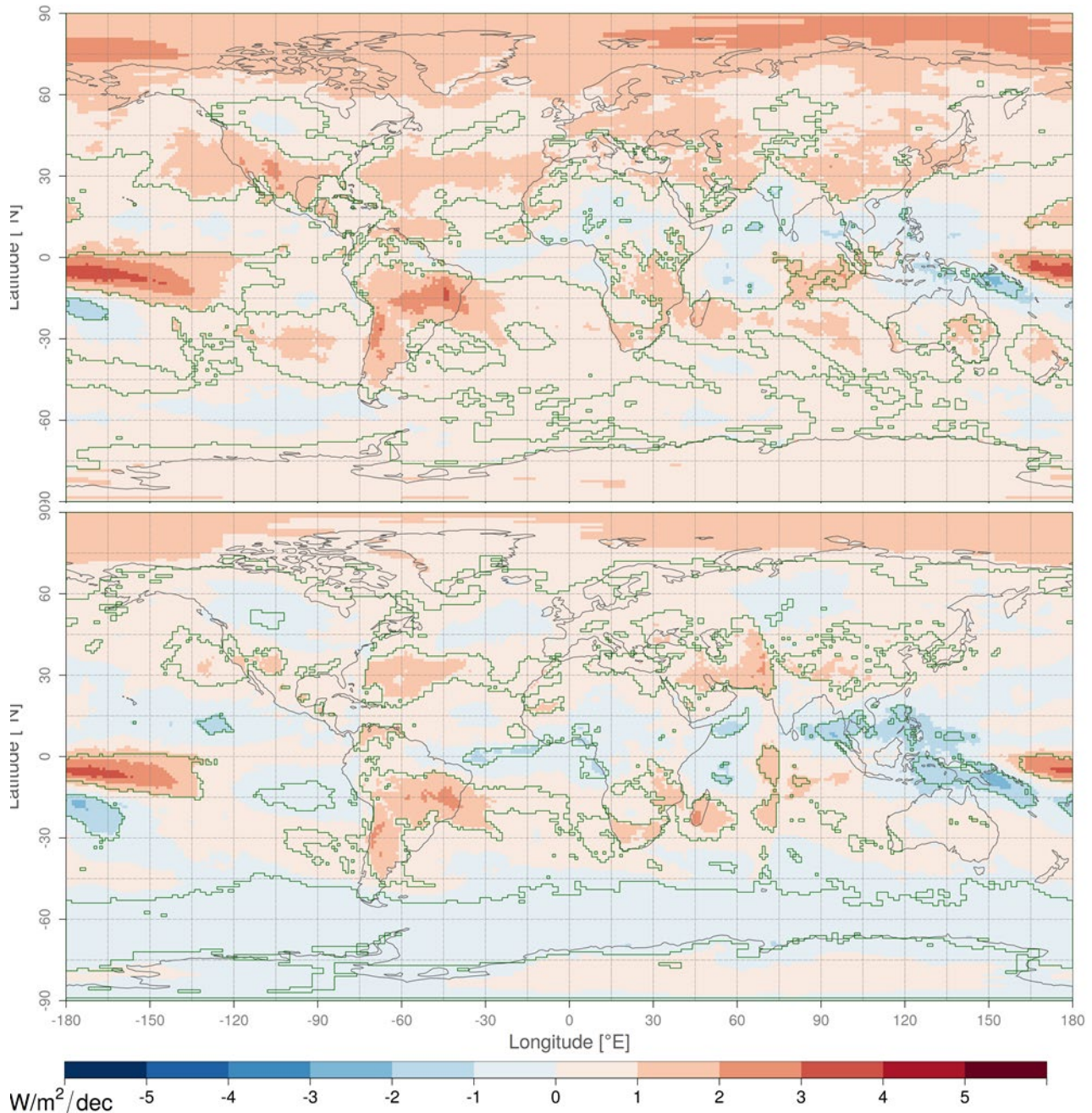
Figure 6-9: Linear trend of monthly OLR between 1979-2020 for CLARA-A3 (top) and HIRS-OLR 1°x1° (bottom). Areas delineated in green correspond to trends that exceed the 95% confidence interval

As summarized by Dewitte and Clerbaux (2017), who compared the mean OLR from 1985-2000 with 2001-2017, an increase of OLR can be seen in the subtropical high pressure areas, part of the North-Hemisphere mid-latitude regions, and the Arctic, where the temperature rise is the highest. Dewitte and Clerbaux (2017) attribute the observed trends to mix of a 'clear-sky effect' and a 'tropical cloud' effect: *concerning the clear sky effect, there are regional changes of OLR which are correlated with surface temperature changes. In the Arctic – where the strongest temperature increase occurs – a strong increase in the OLR is seen. In general, in the Northern Hemisphere – where the surface temperature increase is stronger than in the Southern Hemisphere – there is also an OLR increase. Concerning the tropical cloud effect, there are regional patterns in the changes of the OLR, which are suggesting a relative strengthening of La Niña conditions compared to El Niño conditions. These changes imply societally important regional changes in precipitation* (Dewitte and Clerbaux, 2017).

A final comparison of regional trends is made with the low-resolution version of the HIRS data record (2.5°x2.5°) called HIRS-MM (Figure 6-10). The overall patterns confirm the trends from the high-resolution data records and provide additional trust in the obtained results. Since the HIRS-MM data record does not make use of geostationary satellites, the trend is not affected by the artefacts seen in the daily HIRS data record.
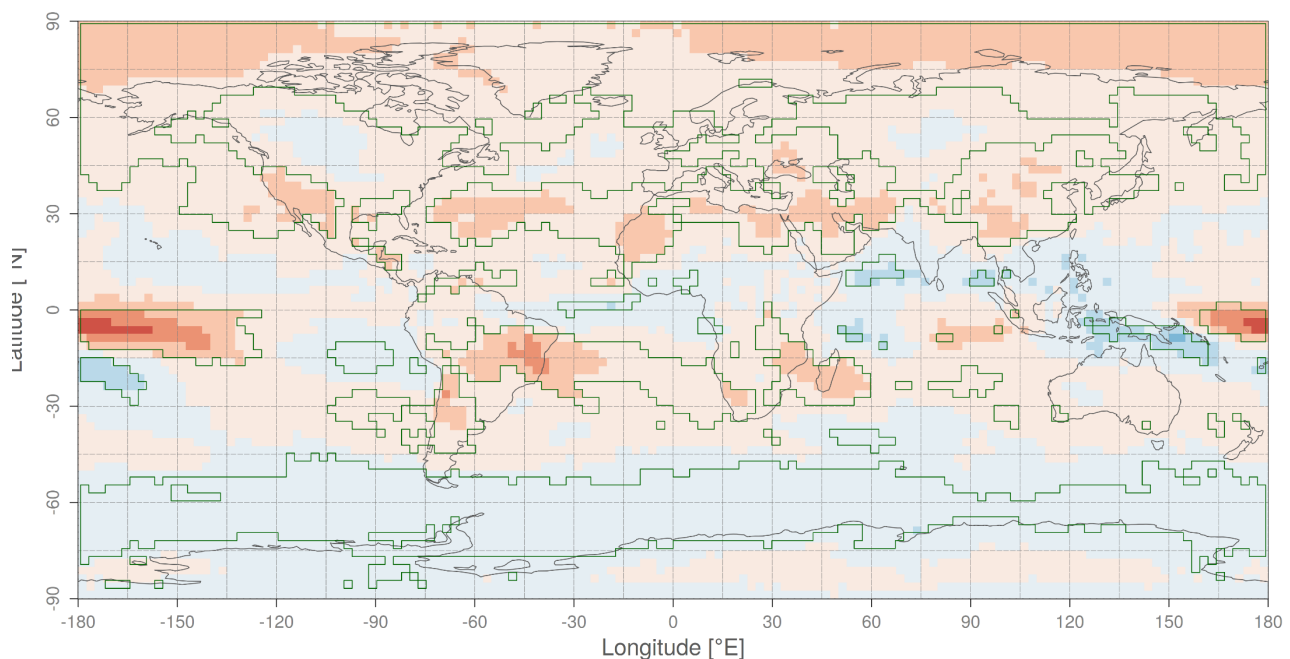
Figure 6-10: Linear trend of monthly OLR between 1979-2020 for HIRS-OLR 2.5°x2.5°. Areas delineated in green correspond to trends that exceed the 95% confidence interval

6.4 Evaluation of ICDR OLR products

The global mean daily OLR (absolute flux) during the year 2020 is shown in Figure 6-11: The red curve (ICDR) has a very similar day-to-day variation compared to the black curve (TCDR), although there are small differences: the RSF in the ICDR is roughly the same (as the TCDR) in July, only to slowly increase (w.r.t. the TCDR) towards December, with a 6-month average difference of +0.05 W/m² (Figure 7-12), which is only about 0.02%. This difference is caused by the exclusion of Metop-C from the ICDR (the difference due to changing NWP and sea ice input data for OLR is non-existent).

Global mean flux of daily CLARA-A3 OLR
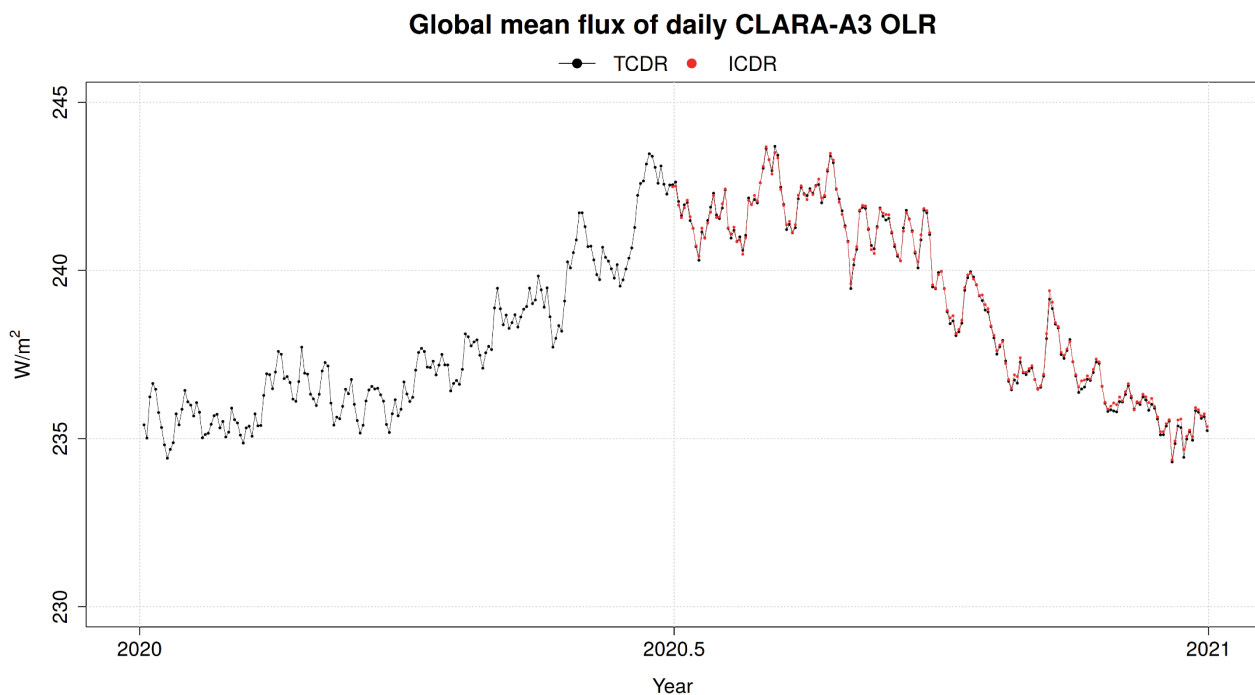
Figure 6-11: Global mean flux of daily CLARA-A3 OLR generated with TCDR and ICDR

The global MAB of daily CLARA-A3 OLR w.r.t. CERES is shown in Figure 6-13, showing that the TCDR and ICDR versions of the CLARA-A3 RSF are similar in terms of processing error (regional uncertainty). The exclusion of Metop-C in the ICDR (2nd half of 2020) leads to only a slightly higher MAB w.r.t. CERES SYN (up to +0.16 W/m², or ~3.5%) because of the lower temporal coverage (especially in this period without afternoon orbit). However, the calibration error for Metop-C is growing much larger in the subsequent years, which by then cannot be compensated anymore by its added value for temporal coverage.

CLARA-A3 Global mean OLR: difference between ICDR and TCDR
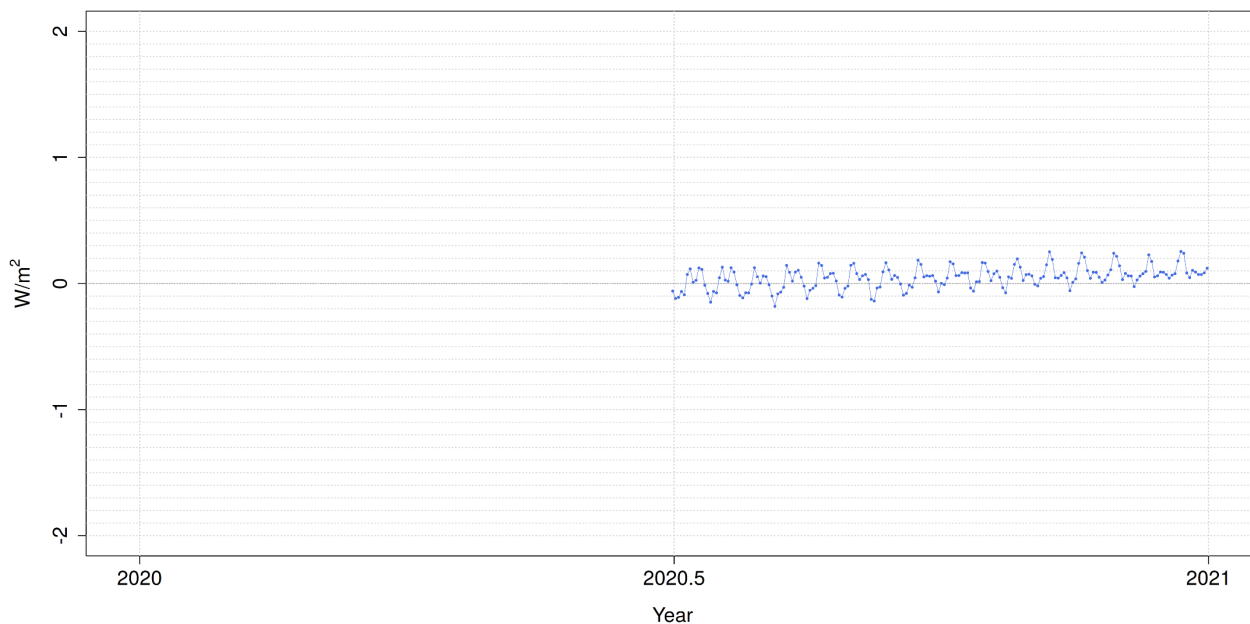
Figure 6-12: Global mean flux of daily CLARA-A3 OLR : difference between TCDR and ICDR. For comparison, the scale range is identical to the RSF equivalent in Figure 5-9.

Global MAB between daily CLARA-A3 OLR and CERES SYN
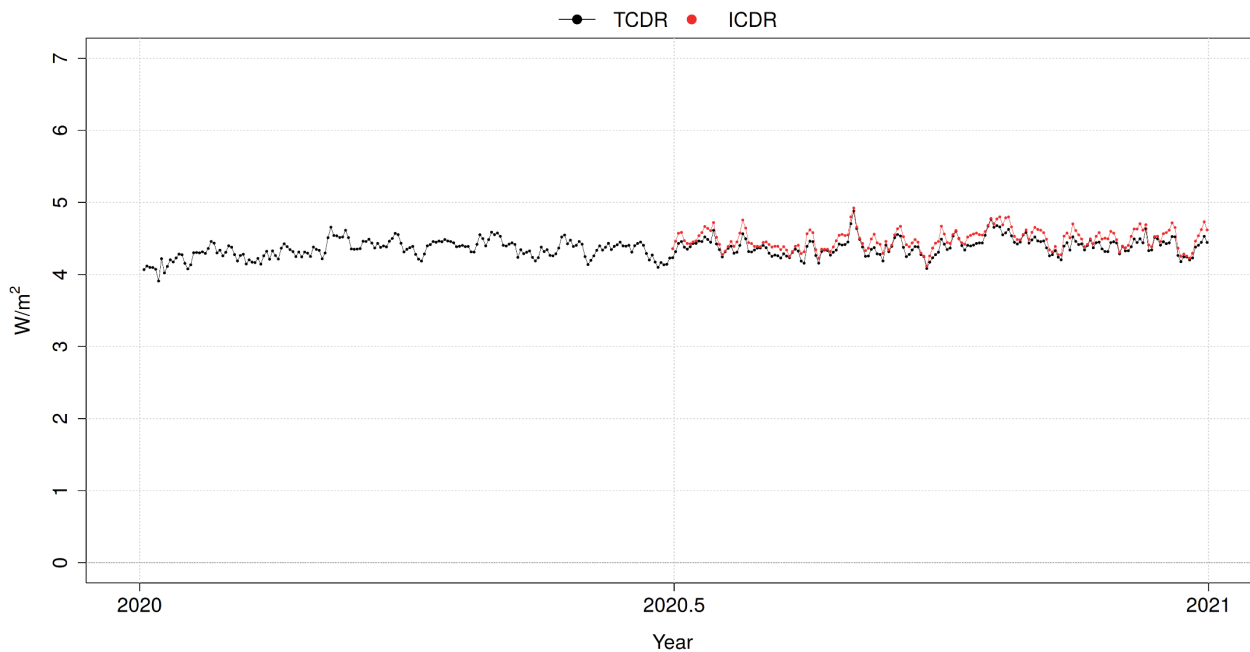
Figure 6-13: Global daily MAB between CERES-SYN1deg-Day and CLARA-A3 OLR generated with TCDR and ICDR. For comparison, the scale range is identical to the RSF equivalent in Figure 5-10.

TOA outgoing longwave flux filled (alias: OLR)
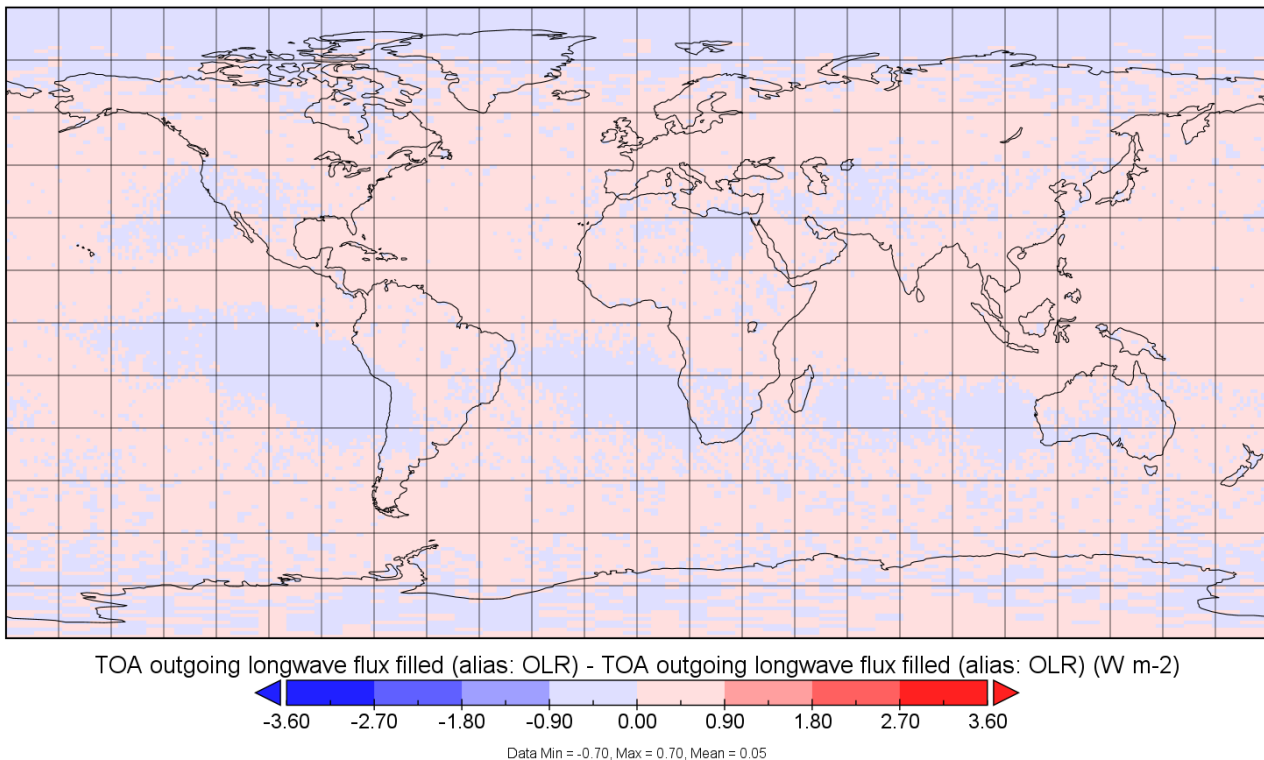
Figure 6-14: CLARA-A3 OLR 6-month mean OLR difference between TCDR and ICDR. For comparison, the scale range is identical to the RSF equivalent in Figure 5-11.

The global mean difference is negligible (-0.05 W/m²) and its spatial distribution has very subtle patterns, however, these are not relevant given their small quantities.

Overall, we can conclude that the switch from TCDR to ICDR has a very limited impact, given their absolute difference which is very small compared to the TCDR's bias variability w.r.t. other reference data records (Figure 6-2) and given the difference in their regional uncertainty (MAB) which is very small compared to the TCDR's absolute MAB (Figure 6-5 and Figure 6-13). Furthermore, it does not cause a violation of the threshold requirement for stability (since the ICDR's bias does not fall outside the stability envelope), nor for the processing error (since the ICDR's MAB does not exceed the level of 16 W/m²).

7 Conclusions

Table 7-1 and Table 7-2 summarize the validation results in view of the requirements.

Table 7-1: Requirements and validation results for processing error (regional uncertainty)

CLARA-A3 Parameter		Averag e MAB (W/m²)	Ref. data record	Threshold requirement W/m² (% met)			Target requirement W/m² (% met)			Optimal requirement W/m² (% met)		
for the entire time series (1979-2020):												
RSF	MM	**3.2***	CERES	8	✓	99.3%*	4	✓	72.4%*	2	✗	24.6%
	DM	**9.0***	CERES	16	✓	94.5%*	8	✗	44.9%*	4	✗	0.0%*
OLR	MM	**1.8**	HIRS	8	✓	99.8%	4	✓	99.2%	2	✓	77.2%
	DM	**4.8**	HIRS	16	✓	99.6%	8	✓	97.1%	4	✗	37.1%
during the CERES era (2000/03-2020/12):												
RSF	MM	**2.3**	CERES	8	✓	100%	4	✓	96.4%	2	✗	49.8%
	DM	**6.2**	CERES	16	✓	98.3%	8	✓	80.4%	4	✗	0.0%
OLR	MM	**1.4**	CERES	8	✓	100%	4	✓	100%	2	✓	98.4%
	DM	**3.7**	CERES	16	✓	100%	8	✓	99.7%	4	✓	81.0%
(*) estimated by mimicking equivalent orbital configurations during CERES era, cfr. Section 11.4												

Table 7-2: Requirements and validation results for stability

CLARA-A3 Parameter		Reference data record	Threshold requirement W/m² (% met)			Target requirement W/m² (% met)			Optimal requirement W/m² (% met)		
for the entire time series (1979-2020):											
RSF	MM	ERA5	4	94.0%	✓	0.6	30.2%	✗	0.3	14.1%	✗
OLR	MM	HIRS	4	99.6%	✓	0.6	37.3%	✗	0.2	15.3%	✗
OLR	MM	HIRS-MM	4	98.0%	✓	0.6	44.4%	✗	0.2	16.5%	✗
during the CERES era (2000/03-2020/12):											
RSF	MM	CERES	4	99.2%	✓	0.6	52.8%	✗	0.3	24.4%	✗
OLR	MM	CERES	4	100%	✓	0.6	69.6%	✗	0.2	28.4%	✗

For **processing error (regional uncertainty)**, Table 7-1 refers to the (spatially) gap-filled, bias-corrected Mean Absolute Bias (MAB) between CLARA-A3 and the reference data record. The threshold, target, and optimal requirements are met when the average MAB (W/m²) does not exceed the respective values set by the Requirements Review document [RD 1], i.e. 8,4,2 W/m² for Monthly Mean (MM), and 16,8,4 W/m² for Daily Mean (DM): when a requirement is met it gets a green tick (✓), otherwise it gets a red cross (✗). Additionally, the temporal coverage of this requirement is also indicated (% of months or days that the requirement is met). Similarly, for **stability**, Table 7-2 refers to the deseasonalized, (spatially) gap-filled, bias-corrected Mean Bias (MB) of CLARA-A3 with respect to the reference data record, which should remain within the stability envelopes of 4, 0.6, and 0.3 W/m² as set by the Requirements Review document [RD 1] for respectively the threshold, target and optimal requirement.

Both tables (i.e. both validation metrics) are split in a part for the entire time series (1979-1999) and for the CERES era (2000-2020). Depending on the variable (OLR or RSF), table (validation metric) and time period, different reference data records are used (CERES-SYN1deg, HIRS-OLR, ERA5):

During the **CERES era (2000-2020)**, both RSF and OLR are validated with CERES.

For the **entire time series (1979-2020)**, OLR is validated with HIRS (processing error and stability) and HIRS-MM (stability). For RSF the stability is validated with ERA5, whereas its (total) processing error is estimated by mimicking the pre-CERES-era processing error using CERES-era equivalent orbital configurations and combining this with the CERES-era processing error.

All data records fulfil the threshold requirements, on both time scales (daily and monthly means) and for both products (OLR and RSF). For processing error (regional uncertainty), the target requirement for OLR is always fulfilled, but for RSF only for the monthly means (for the daily means it is only fulfilled when limiting the time range to the CERES era). For stability, none of the products achieve the target requirement of 0.6W/m²/decade, which was adopted from the very strict GCOS requirements but did not turn out achievable with this data record.

In addition to the evaluation of the TCDR, a study was made of the quality of the ICDR which continuously extends the CLARA-A3 data recordMeteop-C was removed from the ICDR because of its anticipated decay with time of the calibration accuracy for the visible radiances for this satellite. A vicarious calibration technique is used for the visible channels (i.e., no on-board calibration is available) and this requires data from several years to estimate sensor degradation with time and potential biases compared to previous AVHRR instruments. With only 1 ½ years of data available from Metop-C, the calibration information was not considered reliable for this satellite concerning the time-dependent corrections. The ICDR products were compared with TCDR products for an overlap period defined by half a year (July-December 2020). Relatively small differences in the results were found, which still clearly fulfilled the defined threshold requirements. The ICDR production will start with the same calibration version as used in the TCDR (v2017) and data from Metop-C will be added later to the ICDR as soon as a new calibration update is received**.** This update will also lead to better results for the other active satellites due to enhanced quality of time-dependent calibration corrections.

8 References

Akkermans, T., & Clerbaux, N. (2021). Retrieval of Daily Mean Top-of-Atmosphere Reflected Solar Flux Using the Advanced Very High Resolution Radiometer (AVHRR) Instruments. Remote Sensing, 13(18), 3695.

Brient, F., Roehrig, R., & Voldoire, A. (2019). Evaluating marine stratocumulus clouds in the CNRM-CM6-1 model using short-term hindcasts. Journal of Advances in Modeling Earth Systems, 11(1), 127-148.

Canty, T., Mascioli, N. R., Smarte, M. D., & Salawitch, R. J. (2013). An empirical model of global climate–Part 1: A critical evaluation of volcanic cooling. Atmospheric Chemistry and Physics, 13(8), 3997-4031.

Clerbaux, N., Akkermans, T., Baudrez, E., Velazquez Blazquez, A., Moutier, W., Moreels, J., & Aebi, C. (2020). The climate monitoring SAF outgoing longwave radiation from AVHRR. Remote Sensing, 12(6), 929.

Dewitte, S., & Clerbaux, N. (2018). Decadal changes of earth's outgoing longwave radiation. Remote Sensing, 10(10), 1539.

Doelling, D. R., Loeb, N. G., Keyes, D. F., Nordeen, M. L., Morstad, D., Nguyen, C., ... & Sun, M. (2013). Geostationary enhanced temporal interpolation for CERES flux products. Journal of Atmospheric and Oceanic Technology, 30(6), 1072-1090.

Eastman, R., & Warren, S. G. (2014). Diurnal cycles of cumulus, cumulonimbus, stratus, stratocumulus, and fog from surface observations over land and ocean. Journal of climate, 27(6), 2386-2404.

Guilbert, S., Parol, F., Cornet, C., Ferlay, N., & Thieuleux F. (2022). Comparison between CERES/AQUA and POLDER/PARASOL shortwave fluxes: analysis of POLDER/PARASOL diurnal extrapolation. Presentation at CERES Science Team Meeting, 28/04/2022. Accessible online at https://ceres.larc.nasa.gov/ceres-science-team-meetings/2022-04/

Gristey, J. J., Chiu, J. C., Gurney, R. J., Morcrette, C. J., Hill, P. G., Russell, J. E., & Brindley, H. E. (2018). Insights into the diurnal cycle of global Earth outgoing radiation using a numerical weather prediction model. Atmospheric Chemistry and Physics, 18(7), 5129-5145.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... & Thépaut, J. N. (2020). The ERA5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146(730), 1999-2049.

Hogan, R. J., & Bozzo, A. (2018). A flexible and efficient radiation scheme for the ECMWF model. Journal of Advances in Modeling Earth Systems, 10(8), 1990-2008.

Johnson, G. C., Lyman, J. M., & Loeb, N. G. (2016). Improving estimates of Earth's energy imbalance. Nature Climate Change, 6(7), 639-640.

Lee, H.-T. (2014). HIRS Daily OLR Climate Data Record Development and Evaluation, CERES Science Team Meeting, April 22-24, 2014 Hampton, VA.

Lee, H.-T., Schreck, C. J., & Knapp, K. R. (2014). Generation of the Daily OLR Climate Data Record. 2014 EUMETSAT Meteorological Satellite Conference, 22–26 September 2014, Geneva, Switzerland.

Loeb, N. G., Wielicki, B. A., Doelling, D. R., Smith, G. L., Keyes, D. F., Kato, S., ... & Wong, T. (2009). Toward optimal closure of the Earth's top-of-atmosphere radiation budget. Journal of Climate, 22(3), 748-766.

Loeb, N. G., Doelling, D. R., Wang, H., Su, W., Nguyen, C., Corbett, J. G., ... & Kato, S. (2018). Clouds and the earth's radiant energy system (CERES) energy balanced and filled (EBAF) top-of-atmosphere (TOA) edition-4.0 data product. Journal of Climate, 31(2), 895-918.

Loeb, N. G., & Doelling, D. R. (2020). CERES energy balanced and filled (EBAF) from afternoon-only satellite orbits. Remote Sensing, 12(8), 1280.

Stengel, M., Stapelberg, S., Sus, O., Finkensieper, S., Würzler, B., Philipp, D., ... & McGarragh, G. (2020). Cloud_cci Advanced Very High Resolution Radiometer post meridiem (AVHRR-PM) dataset version 3: 35-year climatology of global cloud and radiation properties. Earth System Science Data, 12(1), 41-60.

Trenberth, K. E. (2020). Understanding climate change through Earth's energy flows. Journal of the Royal Society of New Zealand, 50(2), 331-347.

Young, D. F., Minnis, P., Doelling, D. R., Gibson, G. G., & Wong, T. (1998). Temporal interpolation methods for the Clouds and the Earth's Radiant Energy System (CERES) experiment. Journal of Applied Meteorology, 37(6), 572-590.

Young, A. H., Knapp, K. R., Inamdar, A., Hankins, W., & Rossow, W. B. (2018). The international satellite cloud climatology project H-Series climate data record product. Earth System Science Data, 10(1), 583-593.

Zhang, Y., Rossow, W. B., Lacis, A. A., & Oinas, V. (2019). Calculation, evaluation and application of long-term, global radiative flux datasets at ISCCP: past and present. Study of cloud and water processes in weather and climate through satellite observations Submitted for. World Scientific Publishing Company as the second volume in a multi-part series on Earth sciences.

Zhang, Y., Jin, Z., & Sikand, M. (2021). The top-of-atmosphere, surface and atmospheric cloud radiative kernels based on ISCCP-H datasets: method and evaluation. Journal of Geophysical Research: Atmospheres, 126(24), e2021JD035053.

9 Glossary

ADM	Angular Distribution Model *(also: Angular Dependency Model)*
ATBD	Algorithm Theoretical Baseline Document
AVHRR	Advanced Very High Resolution Radiometer
CDR	Climate Data Record *(see also: FCDR)*
CERES	Clouds and the Earth's Radiant Energy System
CLARA-AX	CM SAF cloud, Albedo and Radiation products, AVHRR based, Edition X
CLASS	The Comprehensive Large Array-data Stewardship System (NOAA)
Cloud_CCI	Project covering the cloud component in the European Space Agency's (ESA) Climate Change Initiative (CCI) programme
CDR	Climate Data Record *(see also: FCDR)*
CMAext(ended); CMAprob	Cloud Mask Extended (legacy); Cloud Mask Probabilistic (new)
CM SAF	Climate Monitoring Satellite Application Facility
COT	Cloud Optical Thickness *(also: Cloud Optical Depth)*
DLB	Day Light Block
DM	Daily Mean
EBAF	"*Energy Balanced and Filled*" product from CERES
ERA5	Fifth generation ECMWF reanalysis
F(C)DR	Fundamental (Climate) Data Record
FIDUCEO	Fidelity and uncertainty in climate data records from Earth Observations (EU FP7 project)
GAC	Global Area Coverage
GAC orbit grid	Irregular grid with each pixel representing an AVHRR/GAC observation on its original location
GEO	Geostationary (satellite)
HIRS	High Resolution Infrared Radiation Sounder
ICDR	Intermediate Climate Data Record
IGBP	International Geosphere Biosphere Program
ISCCP-FH	International Satellite Cloud Climatology Project, **H**-series, **F**lux product
IWV	Integrated Water Vapor
LW	Longwave
MB	Mean Bias
MM	Monthly Mean
MAB	Mean Absolute Bias
MODIS	Moderate Resolution Imaging Spectroradiometer (on Aqua and Terra satellites)
NOAA	National Oceanic and Atmospheric Administration
NOAA-X	NOAA satellite numbered X
NTB	Narrowband-to-broadband
NWC SAF	Nowcasting Satellite Application Facility
OLR	Outgoing Longwave Radiation (W/m²)
OSI SAF	Ocean and Sea Ice Satellite Application Facility
PPS	Polar Platform System
RAA (φ)	Relative Azimuth Angle (°)
RMSB	Root Mean Square Bias
RSF	Reflected Solar Flux (W/m²)
SBAF	Spectral Band Adjustment Factors
SNO	Simultaneous Nadir Observations
SSF, SSF1deg	"*Single Scanner Footprint*" product from CERES
SST	Sea Surface Temperatures
SW	Shortwave
SYN, SYN1deg	"Synoptic TOA fluxes and clouds" product from CERES
SZA (θ_0)	Solar Zenith Angle (°)
TCDR	Thematic Climate Data Record
TIROS-N	Television InfraRed Observation Satellite -N
TOA	Top of Atmosphere

TRMM	Tropical Rainfall Measuring Mission
TWL	Twilight
VIRS	Visible Infrared Scanner
VIIRS	Visible Infrared Imaging Radiometer Suite
VZA (θ)	Viewing Zenith Angle (°)

10 Annex

In this chapter, additional information is given, supplemental to the main text. It is not essential to understand the analyses and conclusions provided in the main text. However, it is mostly used to show intermediate analysis steps, as further clarification, and/or for the sake of completeness.

10.1 Missing data treatment: details on spatial gap-filling using ERA5

First the CLARA-A3 products are aggregated to match the CERES Nested 1° Processing grid (cfr Section 4.3), see Figure 10-1.
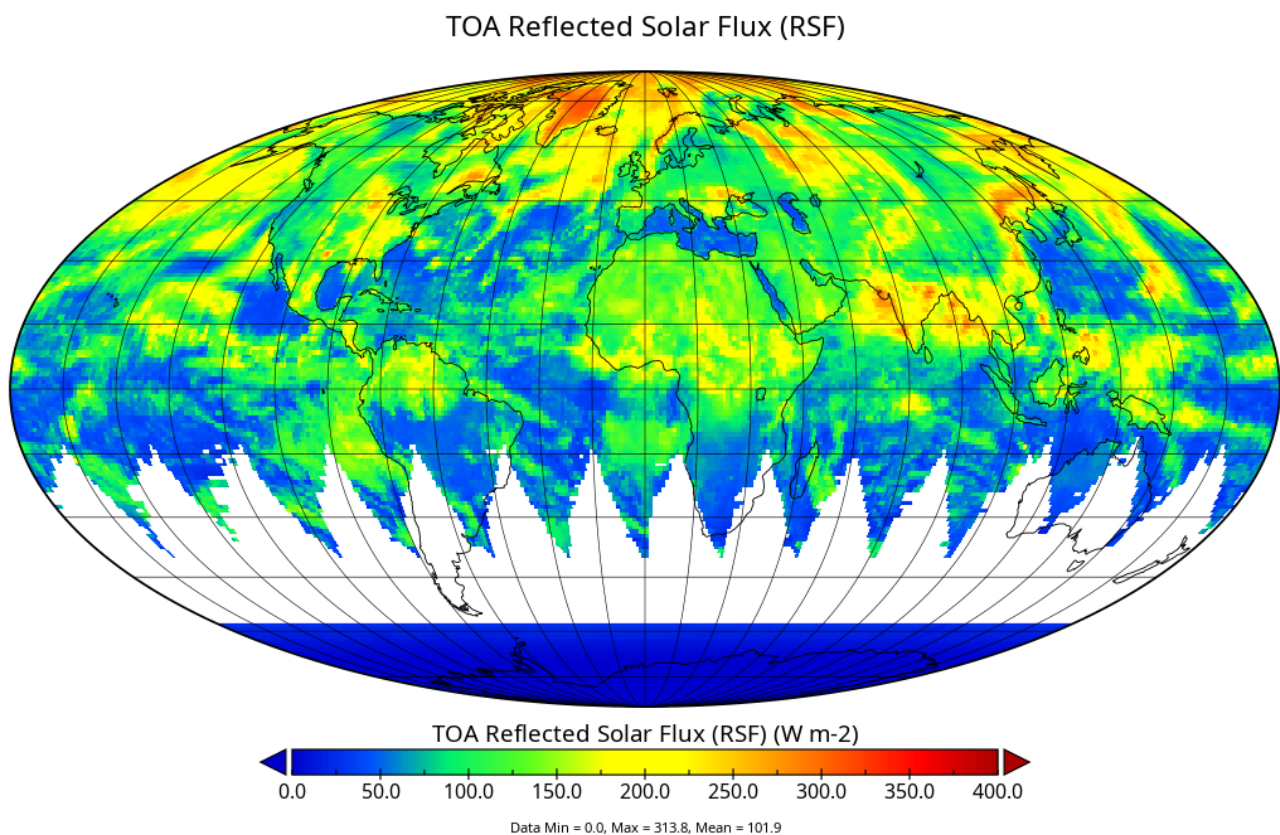
Figure 10-1: Daily CLARA-A3 RSF for 13-07-1994. Aggregated to CERES 1° nested grid.

The white region indicates missing data: each of these pixels is replaced by the ERA5 RSF from the same day. The result is shown in Figure 10-2. Note that the global mean RSF in the non-gap-filled version is 101.9 W/m2 whereas the gap-filled version is only 91.7 W/m², which is a difference of about 10 W/m² (!). Users need to be aware of these gaps and come up with their own solutions for a proper scientific analysis. We recommend filling this area with reanalysis.

TOA outgoing shortwave flux filled

TOA outgoing shortwave flux filled (W m-2)

0.0   50.0   100.0   150.0   200.0   250.0   300.0   350.0   400.0

Data Min = 0.0, Max = 313.8, Mean = 91.7
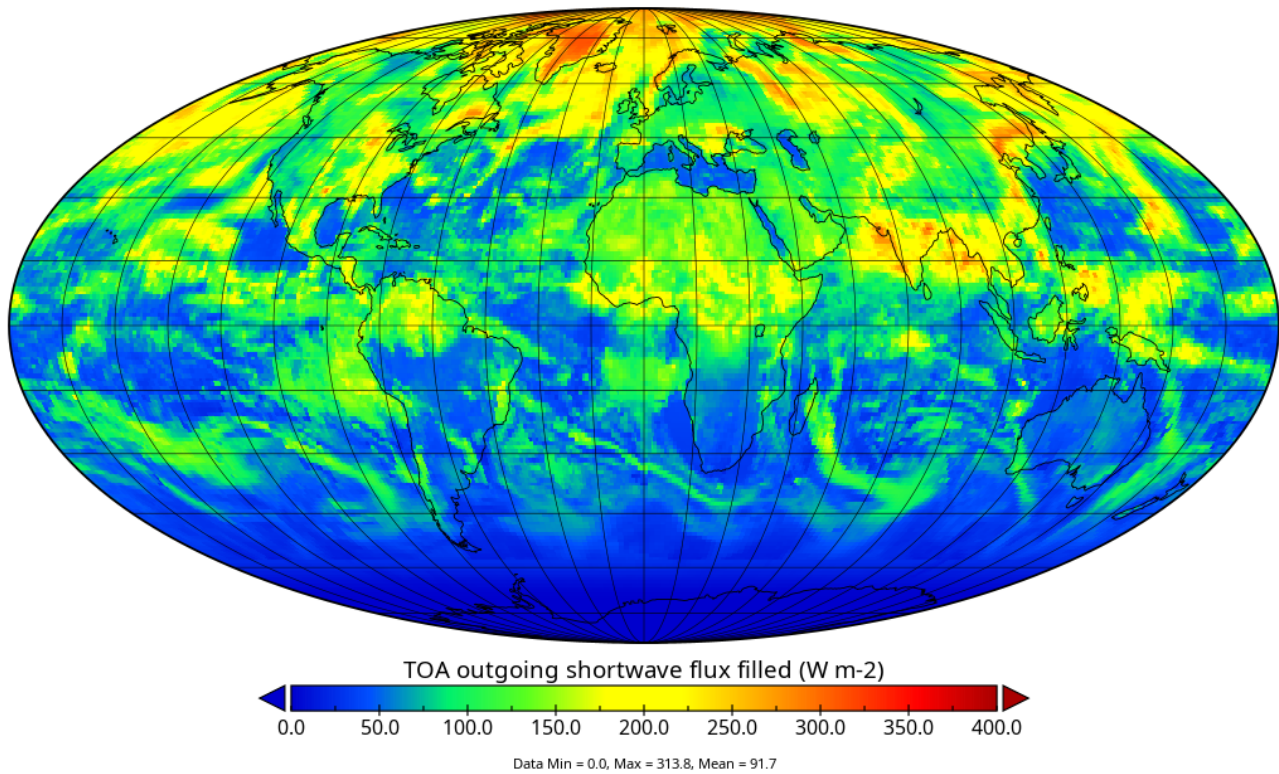
Figure 10-2: Daily CLARA-A3 RSF for 13-07-1994. Aggregated to CERES 1° nested grid. Gaps filled with ERA5 RSF.

10.2 Results for Daily Mean bias and Stability

This Annex contains the results of mean bias and stability on **daily** time scale. They are not included in the main document since these results are roughly similar to the ones on monthly time scale, which means that the monthly mean fluxes are representative for the daily mean fluxes, i.e. that there are no significant intra-month compensating errors. This is demonstrated by the results in this annex.

10.2.1 RSF
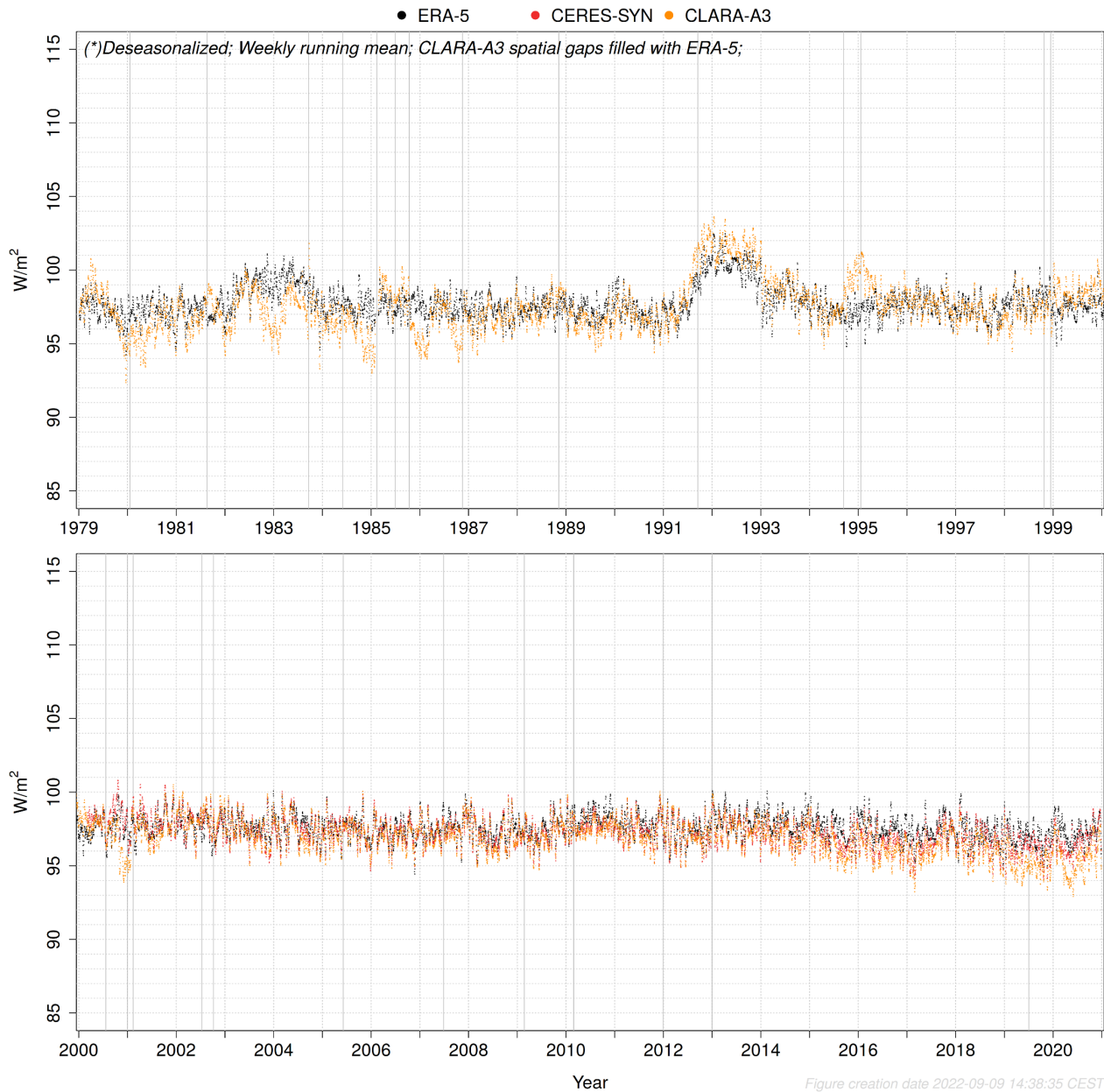
Figure 10-3: Global mean flux of daily CLARA-A3 RSF and other data records

Global mean bias of daily CLARA-A3 RSF w.r.t. other data records (*)

● ERA-5   ● CERES-SYN
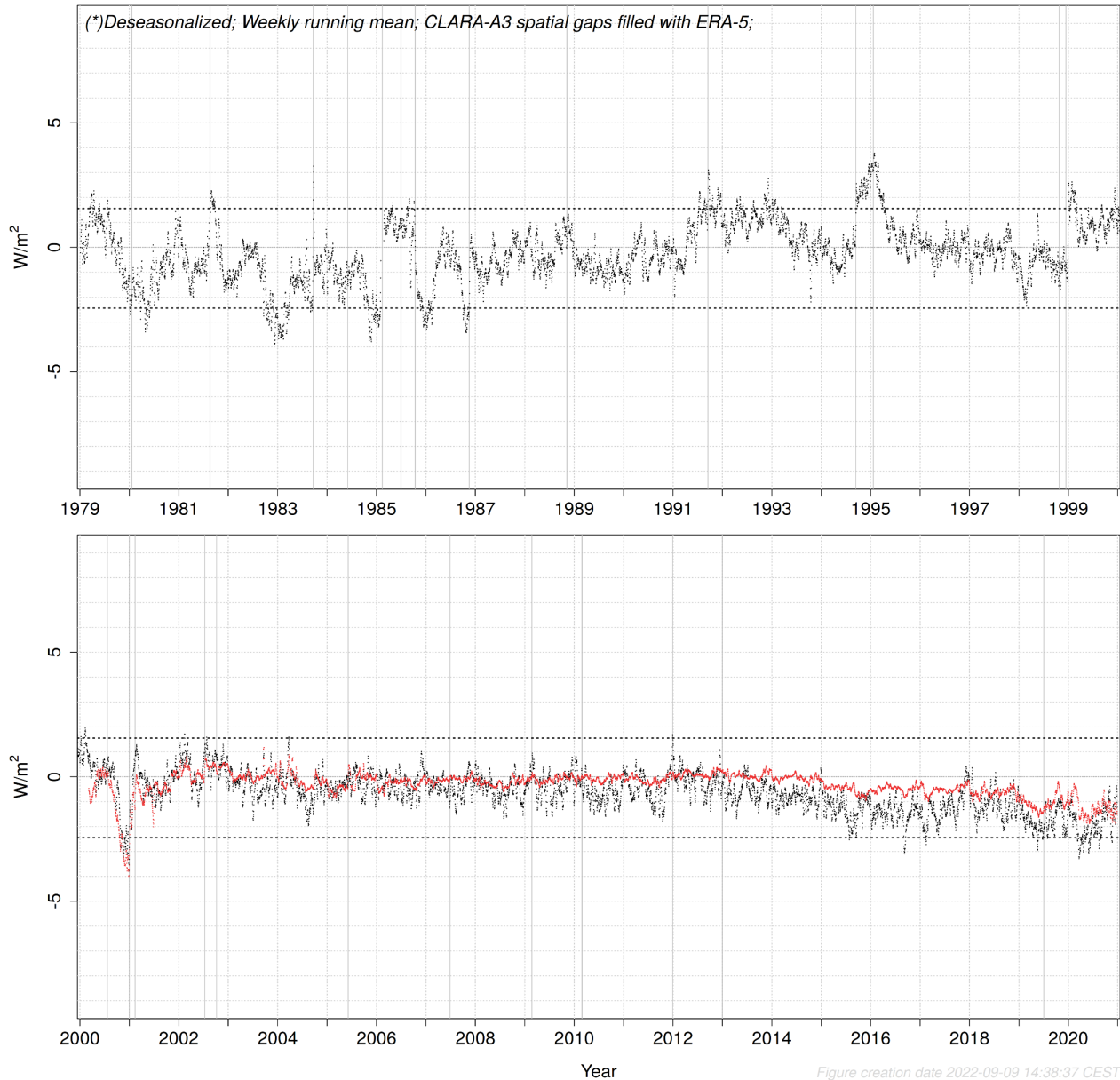
Figure 10-4: Global mean bias of daily CLARA-A3 RSF w.r.t. other data records

Figure 10-5: Histogram of daily RSF bias from CLARA-A3 w.r.t. ERA5

10.2.2 OLR
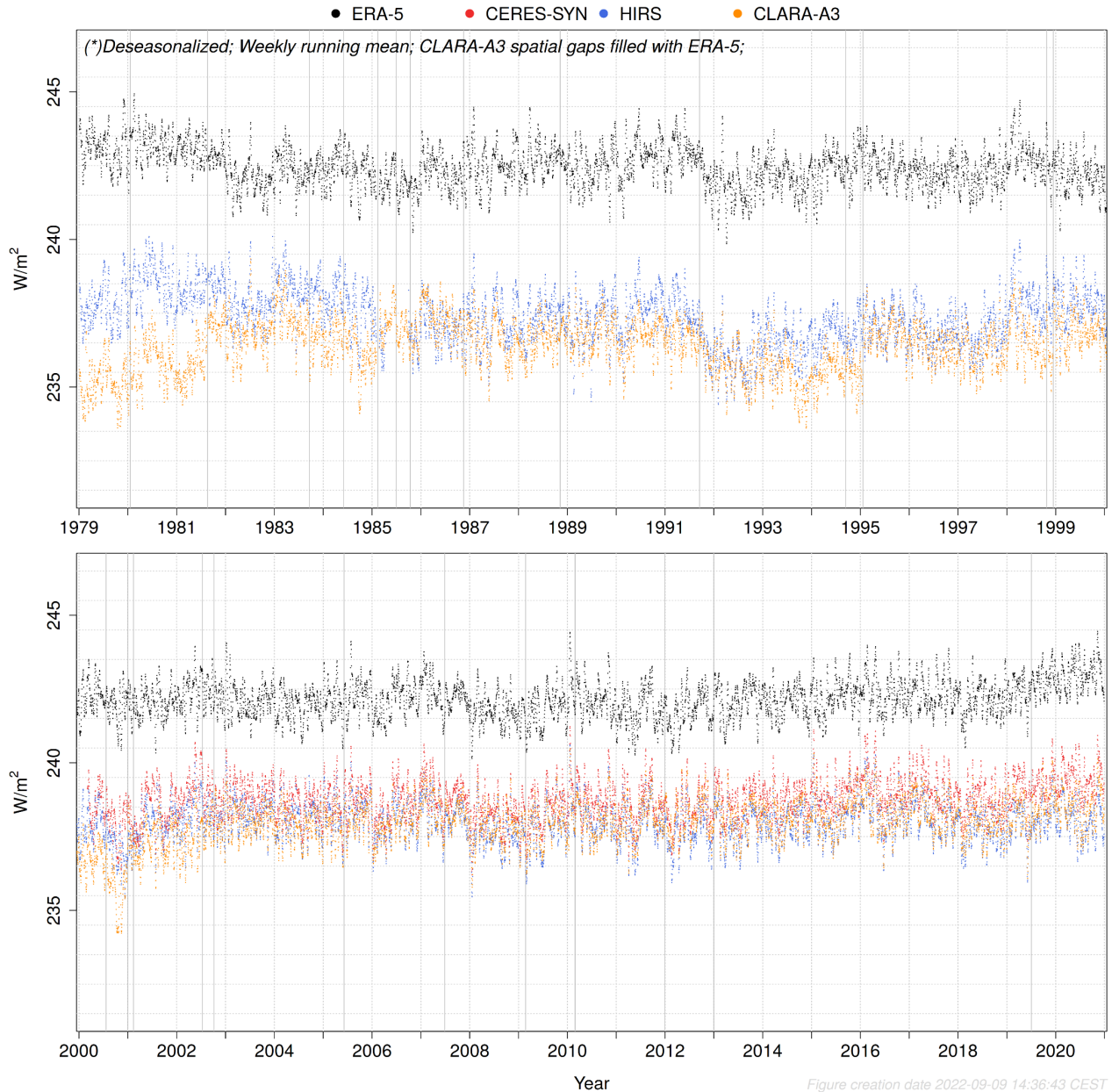
Figure 10-6: Global mean flux of daily CLARA-A3 OLR and other data records
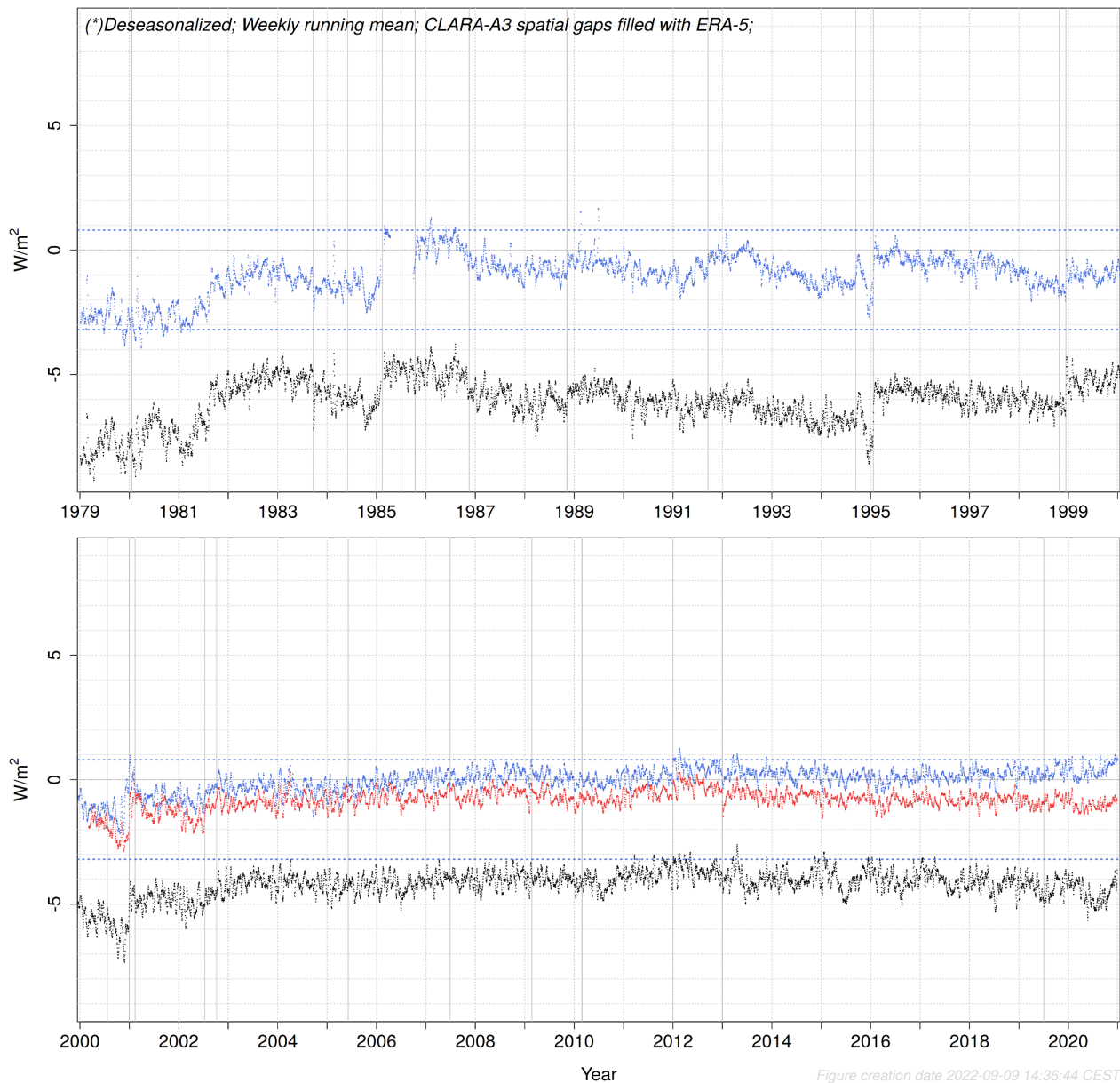
Global mean bias of daily CLARA-A3 OLR w.r.t. other data records (*)

Figure 10-7: Global mean bias of daily CLARA-A3 OLR w.r.t. other data records

Distribution of daily OLR bias from CLARA-A3 w.r.t. HIRS
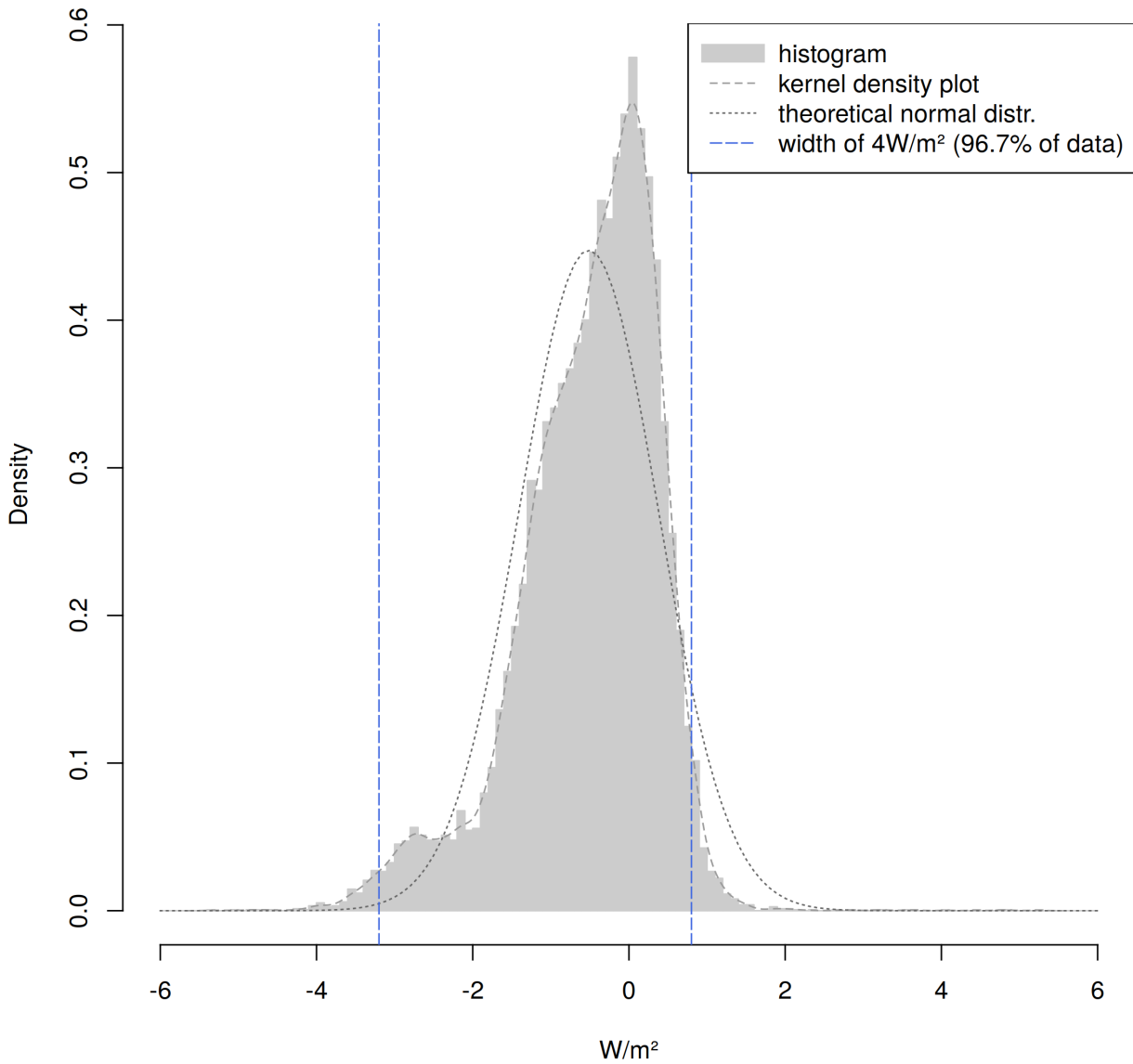
Figure 10-8: Histogram of daily OLR bias from CLARA-A3 w.r.t. ERA5

10.3 RSF mean bias investigation and assessment

In this section, the RSF mean bias is examined more closely. Because the major bias fluctuations (and their causes) are not necessarily aligned with start and end of calendar months, it is better to investigate this with time series of *daily* mean RSF bias. First, the original (non-deseasonalized) bias is shown in Figure 10-9. Seasonal cycles in bias are a combination of seasonal cycles in both data records (so both in CLARA-A3 and in the reference data record). First the focus is on the **CERES era (2000-2020)**: a clear bimodal seasonal cycle is visible in the bias w.r.t. ERA5, but not in the bias w.r.t. CERES-SYN: this indicates that mainly (seasonal deviation in) the ERA5 record is the driver behind their combined bias' seasonal cycle. The bimodal ERA5 seasonal bias pattern consists of a high peak around December, a moderate trough and peak around respectively April and June, and a low trough around August.
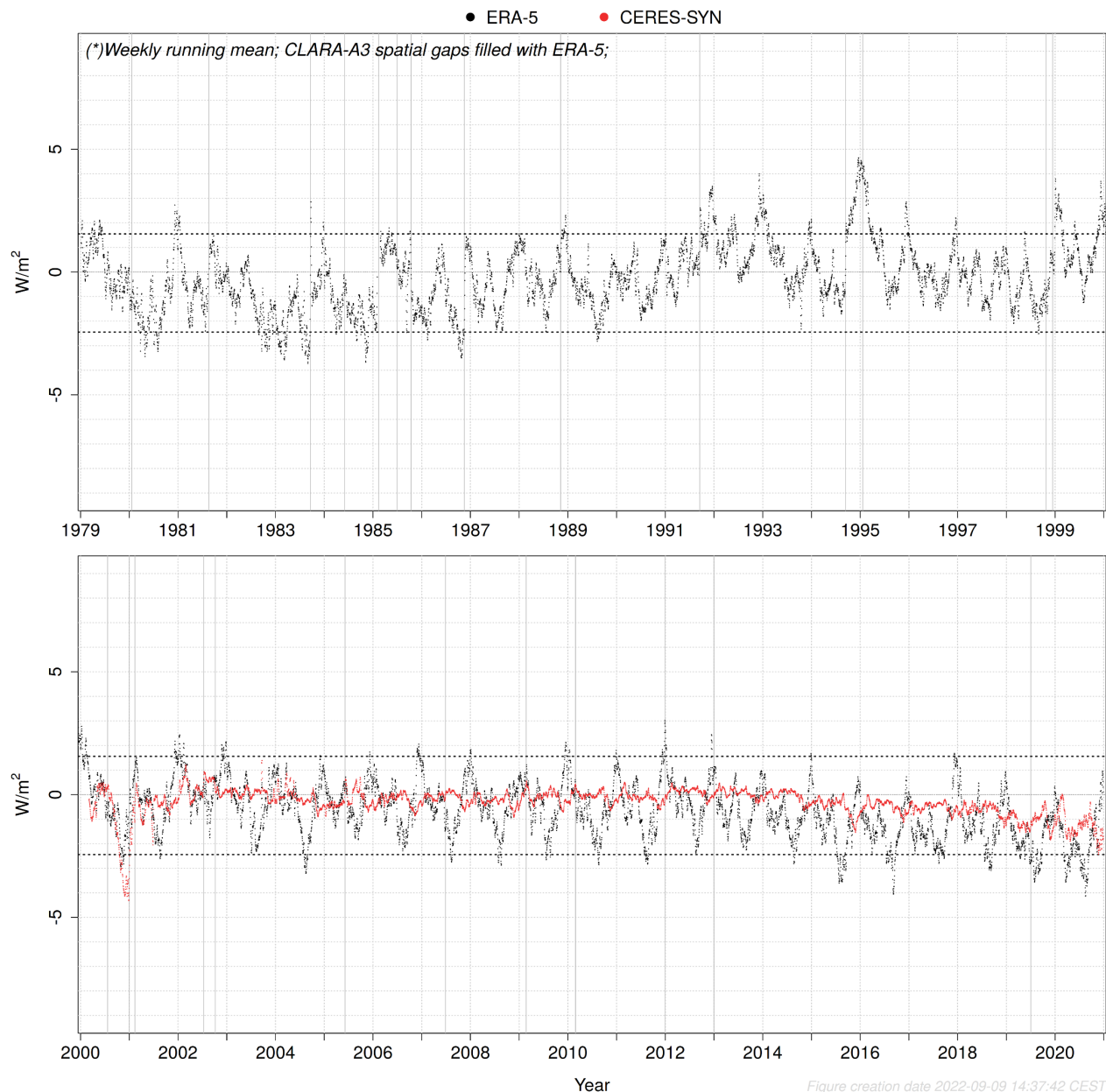
Figure 10-9: Global mean bias of non-deseasonalized daily CLARA-A3 RSF w.r.t. other data records

One exception is the second half of the year 2000, characterized by a suboptimal orbital configuration for CLARA-A3 (i.e. only one late afternoon orbit). This means that besides ERA5, also suboptimal orbital configurations in CLARA-A3 cause fluctuations in the mean bias; but in contrast to ERA5, these are bound to limited periods (i.e. not systematic throughout the entire CLARA-A3 data record). Given the nature of the ERA5 data record (i.e. reanalysis), its seasonal patterns are systematic and similar throughout the entire record (1979-2000); hence, deseasonalizing the bias (of CLARA-A3 w.r.t. ERA5) erases all these systematic seasonal patterns originating from ERA5, leaving only the non-systematic and temporally limited fluctuations visible (coloured rectangles in Figure 10-10).

Figure 10-10: Global mean bias of deseasonalized daily CLARA-A3 RSF w.r.t. other data records. Large rectangles indicate suboptimal orbital configurations: afternoon (yellow) or morning (pink). Green rectangles indicate deviations probably caused by L-2b stability.

The assumption that the bias fluctuations in these coloured rectangles are mainly due to CLARA-A3 (and not to ERA5), is supported by examining the deseasonalized time series of absolute RSF

(Figure 5-1): indeed, the otherwise stable ERA5 RSF only shows significant deviations during the aftermath of the two volcanic eruptions.

The following sections discuss the most important contributors to the RSF mean biases (besides the systematic seasonal deviations from ERA5), i.e. the biases seen in Figure 10-10: (10.3.1) the CLARA-A3 orbital configuration; (10.3.2) the radiative impact of volcanic eruptions, (10.3.3) the underlying TOA albedo Level-2b record stability, and (10.3.4) the applied albedo models.

10.3.1 CLARA-A3 orbital configuration

As seen for the second half of year 2000, a likely reason for non-systematic bias fluctuations are CLARA-A3's suboptimal orbital configurations, i.e. having only a morning or an afternoon orbit (Table 10-1, in which the colours refer to the rectangles in Figure 10-10).

Table 10-1: List of CLARA-A3 suboptimal orbital configurations

Date (start)	Date (end)	Satellite	Orbital configuration*
1979-01-01	1980-01-20	TIROS-N	Afternoon (=suboptimal)
1980-01-20	1981-08-19	NOAA-6	Morning (=suboptimal)
1981-08-19	1983-09-19	NOAA-7	Afternoon (=suboptimal)
1984-06-01	1985-02-13	NOAA-7	Afternoon (=suboptimal)
1985-02-13	1985-07-01	NOAA-9	Afternoon (=suboptimal)
1985-10-14	1986-11-17	NOAA-9	Afternoon (=suboptimal)
1994-09-13	1995-01-20	NOAA-12	Morning (=suboptimal)
2000-07-22	2001-01-01	NOAA-14	Late afternoon (=suboptimal)
2001-01-01	2001-02-12	NOAA-16 & NOAA-14	Afternoon (=suboptimal)
(*) The colors refer to morning-only (yellow) or afternoon-only (reddish) configuration, as in Figure 11-10.			

First, to verify and quantify the impact of these suboptimal orbital configurations, some tests are performed: Alternative CLARA-A3 data is generated using data from the CERES era (2000-2020) but, in contrast to the real CLARA-A3 data record, using only a single satellite at a given time in order to mimic the suboptimal orbital configuration during the pre-CERES era (1979-1999). More specifically, an afternoon-only configuration is mimicked by using either only NOAA-18 (between 2005-2014) or only NOAA-19 (between 2015-2019), cfr. Figure 10-11.
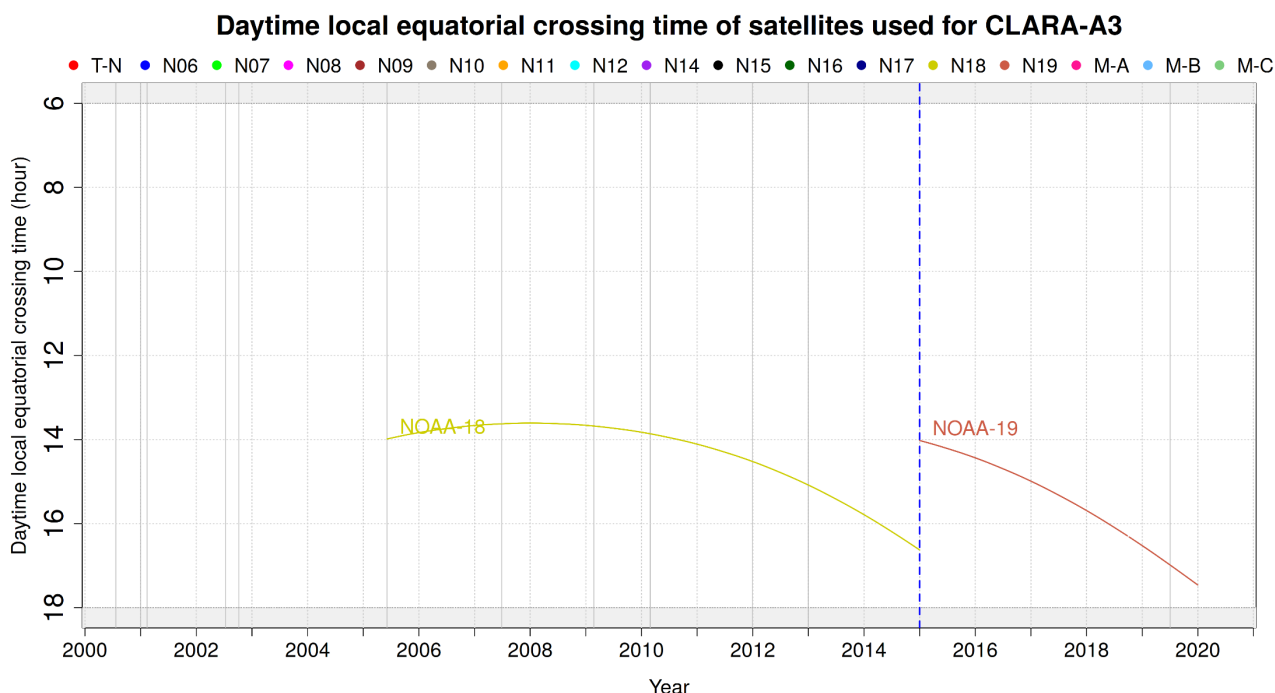
Daytime local equatorial crossing time of satellites used for CLARA-A3

● T-N ● N06 ● N07 ● N08 ● N09 ● N10 ● N11 ● N12 ● N14 ● N15 ● N16 ● N17 ● N18 ● N19 ● M-A ● M-B ● M-C

Figure 10-11: Daytime local equator crossing time of afternoon satellites NOAA-18 (2005-2014) and NOAA-19 (2015-2019)

The resulting daily mean bias is shown in Figure 10-12.
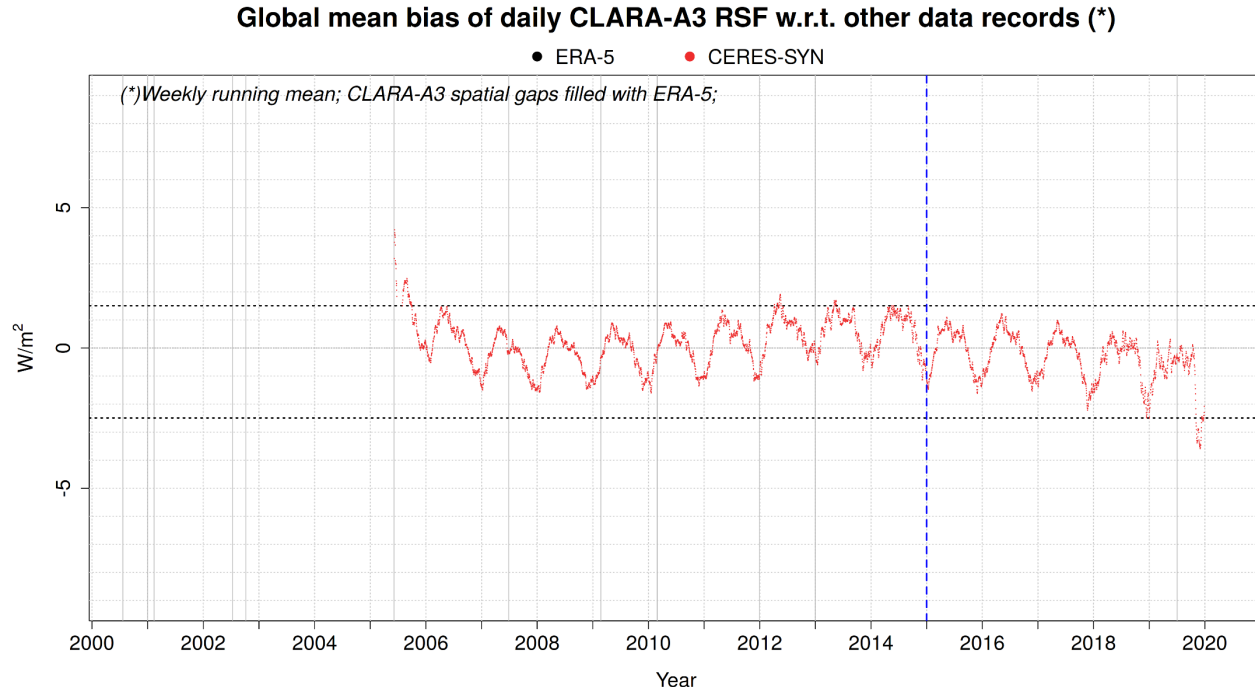
Global mean bias of daily CLARA-A3 RSF w.r.t. other data records (*)

● ERA-5 ● CERES-SYN

(*)Weekly running mean; CLARA-A3 spatial gaps filled with ERA-5;

Figure 10-12: Global mean bias of non-deseasonalized daily CLARA-A3 RSF w.r.t. CERES-SYN, using only afternoon satellites NOAA-18 (2005-2014) or NOAA-19 (2015-2019)

Compared to the non-deseasonalized mean bias of the **all-satellite** orbital configuration (red curve in Figure 10-9), the non-deseasonalized **single-satellite** configuration clearly causes a seasonal cycle (red curve in Figure 10-12). This seasonality has an amplitude of ~2W/m² and a unimodal pattern, with a low/trough around November-February, and a high/peak around April-July. It is precisely this pattern that is also noticeable in the first years of the deseasonalized ERA5 bias,

indicative for the non-systematic CLARA-A3 biases, more specifically in the periods with an afternoon-only orbit (yellow entries in Table 10-1 and yellow rectangles in Figure 10-10). This confirms the **existence of periods with seasonal bias in CLARA-A3 related to their *suboptimal orbital configuration***; Their **limited, i.e. non-systematic, time span (duration)** explains why they remain visible in the deseasonalized time series (i.e. because only seasonality which is systematic across the entire record gets erased by deseasonalization). In fact, we can distinguish two distinct seasonal patterns in the periods with suboptimal orbital configuration:

- **Afternoon-only**: the above described pattern with a low/trough around November-February, high/peak around April-July (yellow rectangles in Figure 10-10; also mimicked by CERES era experiment in Figure 10-12); it is probably this effect that causes the CLARA-A3 RSF radiative effect of the El Chichón event to be not well represented (cf the drop of ~2W/m² w.r.t. ERA5 around January 1983, <u>Figure 5-1</u>).
- **Morning-only**: unimodal: pattern with a low/trough around May and high/peak during December-January (reddish rectangles in Figure 10-10).

**Where does this seasonality come from?**

Incomplete temporal coverage (i.e. having a *suboptimal orbital configuration*) of regional climate phenomena with an asymmetric diurnal cycle (e.g. marine stratus thinning or land convection) introduces strong regional biases, as documented in Section 5.3.1, which can be positive as well as negative, depending on the region and the kind of phenomena. Globally averaged, these biases vary seasonally because of the hemispherical imbalance of the associated regional climate features' occurrence and strength. **As a result, a *suboptimal orbital configuration* is characterized by its own distinct annual cycle of mean bias**.

10.3.2 Radiative impact of Volcanic eruptions

There is a clear positive bias related to the period of the Pinatubo eruption, between April 1991 and January 1993 (Figure 10-10) which increased the bias w.r.t. ERA5 by more than +1W/m² compared to the period before and after. Here it probably concerns a bias in the ERA5 reanalyses, in which an underestimation of the prescribed aerosol optical depth would explain an underestimated RSF; this would be the more likely cause, rather than an apparent overestimation of CLARA-A3 RSF, which is, after all, based on observations.

10.3.3 Underlying TOA albedo Level-2b data record stability

The RSF daily mean bias also depends on the stability of the underlying Level-2b (Table 2-3) data record which contains the instantaneous albedo and is spatially aggregated to the 0.25°x0.25° CLARA-A3 grid. This in turn depends on many factors, such as the Level-1c (Table 2-3) AVHRR FDR stability, but also the quality and stability of several auxiliary input data such as cloud mask and cloud properties. It is beyond the scope of this report to validate each of these input data and assess their individual impact on the CLARA-A3 TOA flux product's validation. To demonstrate the potential impact of Level-2b stability on Daily Mean RSF when only a single satellite is used (i.e. a *suboptimal orbital configuration*), a test is performed using only the midmorning satellites NOAA-17 and MetOp-B (Figure 10-13).
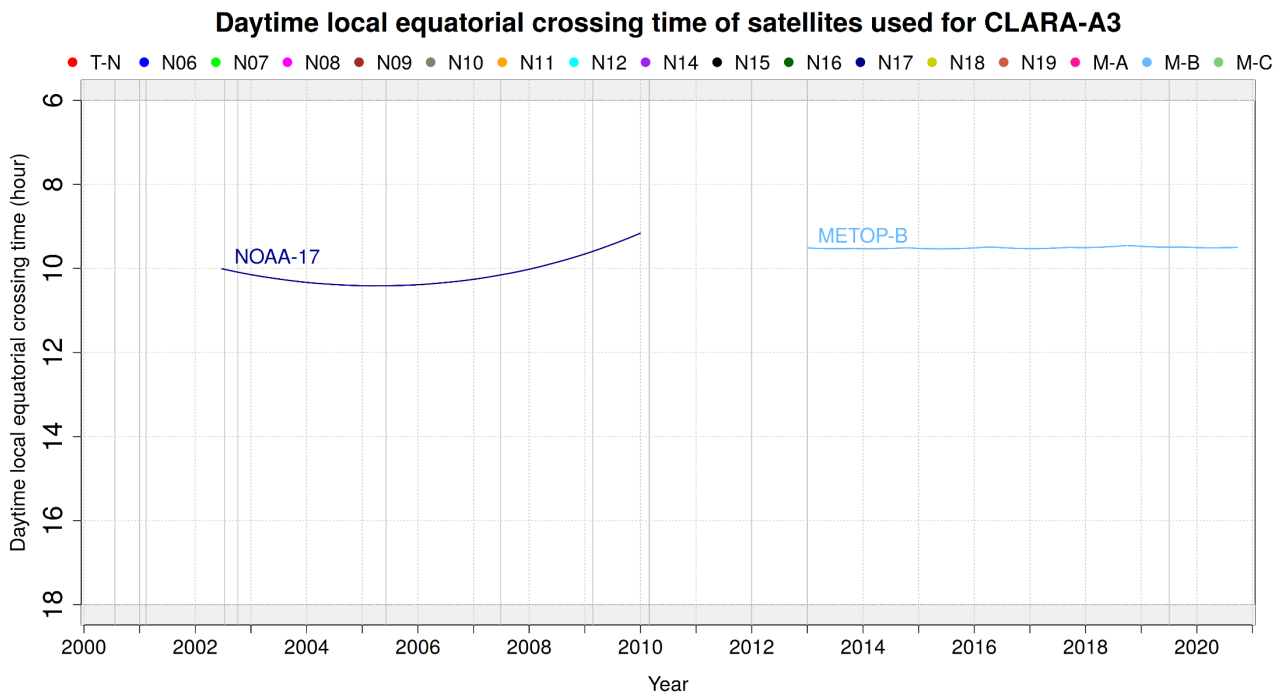
Daytime local equatorial crossing time of satellites used for CLARA-A3

Figure 10-13: Daytime local equator crossing time of afternoon satellites NOAA-17 (2002-2009) and MetOp-B (2013-2020)

The MetOp constellation is characterized by stable orbits, i.e. there is no orbital drift affecting the local equator crossing time, so the average solar zenith angle remains the same during the period 2013-2020. This makes it the ideal case to isolate the Level-2b stability as a bias contributor, because it excludes any SZA dependent bias such as the applied albedo model (which is the last contributor, cfr Section 10.3.4). The resulting daily mean bias is shown in Figure 10-14.
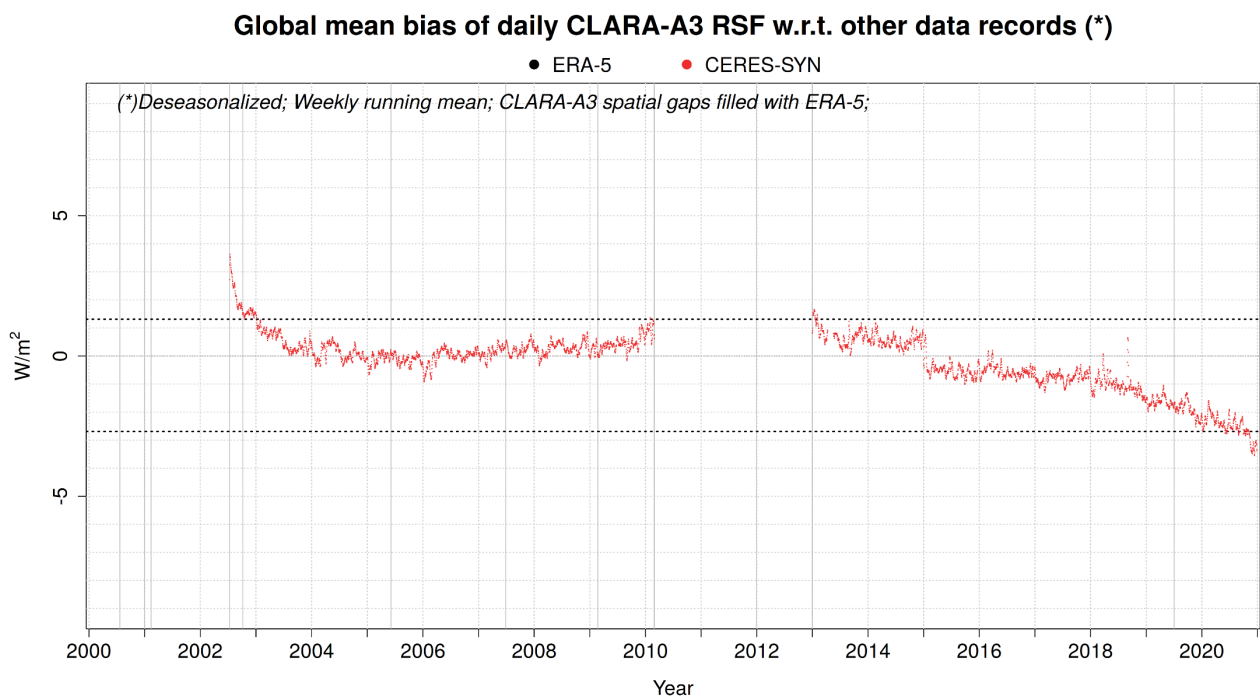
Global mean bias of daily CLARA-A3 RSF w.r.t. other data records (*)

Figure 10-14: Global mean bias of non-deseasonalized daily CLARA-A3 RSF w.r.t. CERES-SYN, using only midmorning satellites NOAA-17 (2002-2009) or MetOp-B (2013-2020)

The resulting bias degrades from 0.5 W/m² to -2.5 W/m² at the end of the data record. This is in line with the observed Level-1c and Level-2b trends of respectively narrowband visible reflectance and broadband albedo (figures not shown). Another test is done to verify whether this **decreasing trend in MetOp-B** is the main reason of the gradually decreasing daily and monthly RSF bias between 2015-2020 in the full RSF product in which all satellites were used (Figure 10-10, Figure 5-2). This is done by using all satellites *except* MetOp-B. The result of this test is shown in Figure 10-15, in which the mean bias w.r.t. CERES-SYN hovers around zero until the end of the data record. This demonstrates that MetOp-B is the main driver behind the decreasing trend in the CLARA-A3 RSF product (compare Figure 10-15 with the all-satellite configuration in Figure 10-16).

Figure 10-15: Global mean bias of monthly CLARA-A3 RSF w.r.t. other data records, using all satellites except MetOp-B

Global mean bias of monthly CLARA-A3 RSF w.r.t. other data records (*)
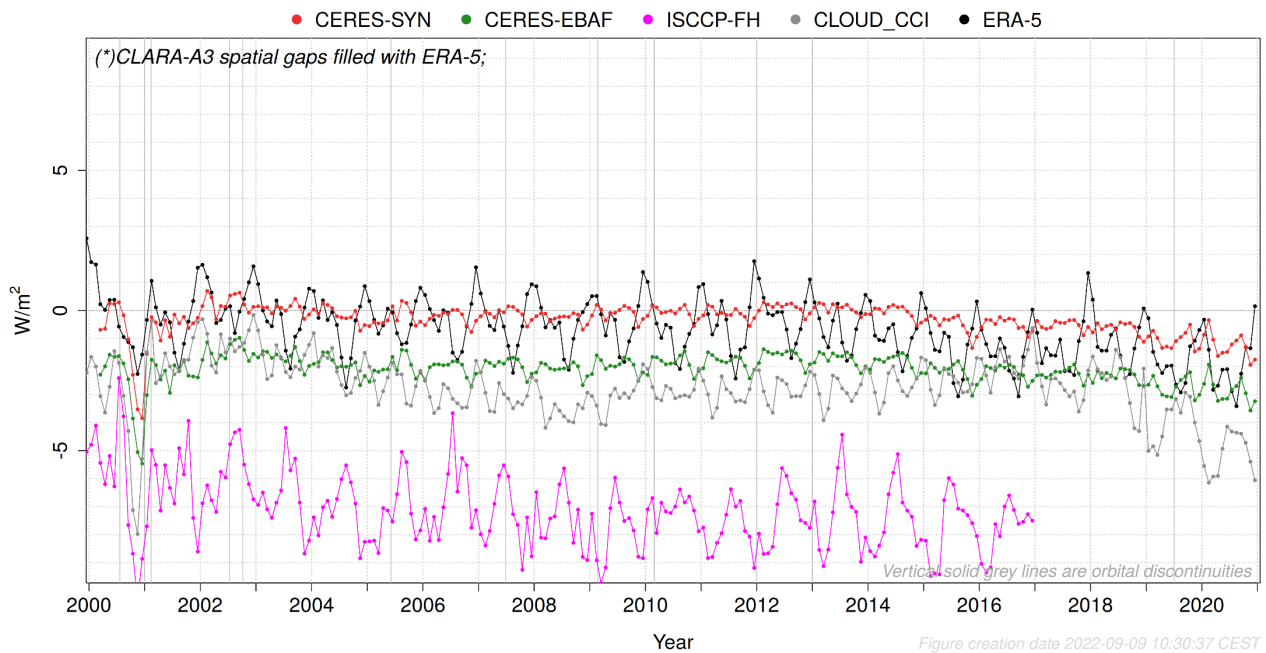
Figure 10-16: Global mean bias of monthly CLARA-A3 RSF w.r.t. other data records, using all satellites (default orbital configuration)

It should be noted that MetOp-B was not well characterized because of its limited historic record when the FDR was generated. At that time, it was difficult to predict such a degradation and anticipate its future calibration parameters.

Another feature that is worth highlighting in this section is the **post-launch peak** in positive RSF bias. In this case it is very significant for the NOAA-17 case, with the first month characterized by a bias of up to +3 W/m² (Figure 10-14). However, this rapidly normalizes to +1W/m² after 2 months. It is probably related to the stability of the underlying TOA albedo Level-2b data record, given the very rapid decrease and the apparent absence of logic relation with solar zenith angle (i.e. with orbital drift or other orbital properties). This, in turn, can be due to the Level-1c FDR stability or other input data such as cloud products.

When comparing these biases with the all-satellite daily mean RSF bias (Figure 10-10), an important conclusion is that these stability issues are more important when only a single satellite used (suboptimal orbital configuration) because this causes a direct propagation of the error from Level-2 (instantaneous) to Level-3 (daily); With multiple satellites combined, the bias' propagation and hence the bias in daily mean RSF is much less outspoken because it only affects part of the diurnal cycle.

Now it becomes possible to attribute some bias features in the early years of the record. For instance, the start of the NOAA-7 period (August 1981) is characterized by a higher RSF compared to the rest of the NOAA-7 period (1981/08/19 – 1983/09/19), cfr green rectangle in Figure 10-10. During that period there is probably a superposition of two effects: (1) it's near the peak of the typical seasonal cycle associated with afternoon-only suboptimal orbital configuration, as discussed in Section 10.3.1, and (2) probably a post-launch positive bias in the underlying Level-2b TOA albedo data record, in turn caused by e.g. Level-1c FDR calibration issue or issues with other input data such as cloud products.

Finally, it should be acknowledged that it is difficult to separate the issue of Level-2b stability (this Section 10.3.3) and the issue of albedo model deficiencies (next Section 10.3.4). The main reason

is that a proper validation of the Level-2b stability requires the reflectance (if narrowband) or albedo (if converted to broadband) to be corrected for the satellite's drift in local solar time (i.e., SZA): however, such a correction is to be performed using… albedo models (!). So disentangling the two issues is difficult. However, for certain cases it is possible, e.g. for non-drifting satellites such as the MetOp constellation (discussed above in this Section 10.3.3).

10.3.4 Albedo models

The daily mean RSF is estimated by applying *albedo models*, also called '*directional models*' or '*diurnal cycle models*'. These are typically part of Angular Dependency Models (ADM's) and describe the relation between solar zenith angle (SZA) and TOA albedo. Since this relation is highly dependent on scene type (i.e. combination of land surface and cloud properties), a large number of scene dependent albedo models is necessary. The temporal interpolation method applied in CLARA-A3 is called the '*constant meteorology method'* has been documented extensively by Young et al.(1998) and used subsequently in the CERES processing, where it is also called the '*CERES-only (CO) method'* (Doelling et al.,2013). Basically, the diurnal cycle associated with the observation's scene type is scaled so that it matches the observed albedo (at the time of observation). This means that not the absolute magnitude of an albedo model is important, but its shape.
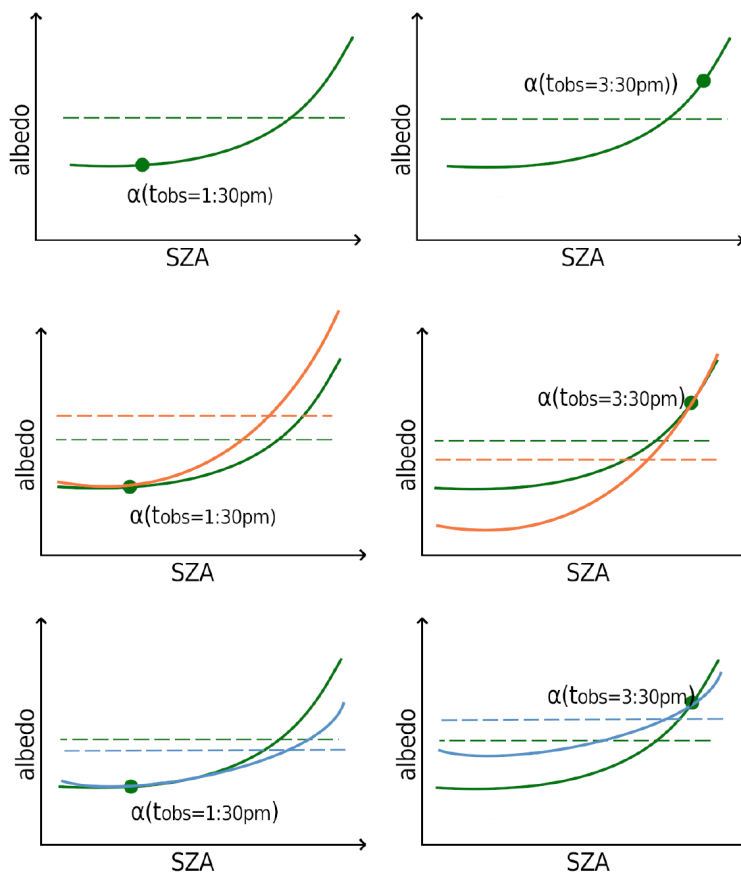
Figure 10-17: (top) observed dependency of albedo on SZA; (centre) too steep albedo model, causing overestimation with low SZA and underestimation with high SZA, i.e. decreasing trend with orbital drift; (bottom) opposite, i.e. too flat albedo model leading to increasing trend with orbital drift. Figure from Guilbert et al., 2022.

For orbital configurations with only a single observation per day, imprecise albedo models could have relatively large consequences on the resulting daily mean RSF. This is shown in Figure 10-17

(Guilbert et al.,2022). Note that in reality there are many scene types, each with its own distinct CERES TRMM albedo model, but for this explanation an aggregated conceptual example is shown.

The potential impact of imprecise albedo models is demonstrated using a CERES-era imitation of a pre-CERES era suboptimal orbital configuration (afternoon-only NOAA-18 orbit).

Daytime local equatorial crossing time of satellites used for CLARA-A3
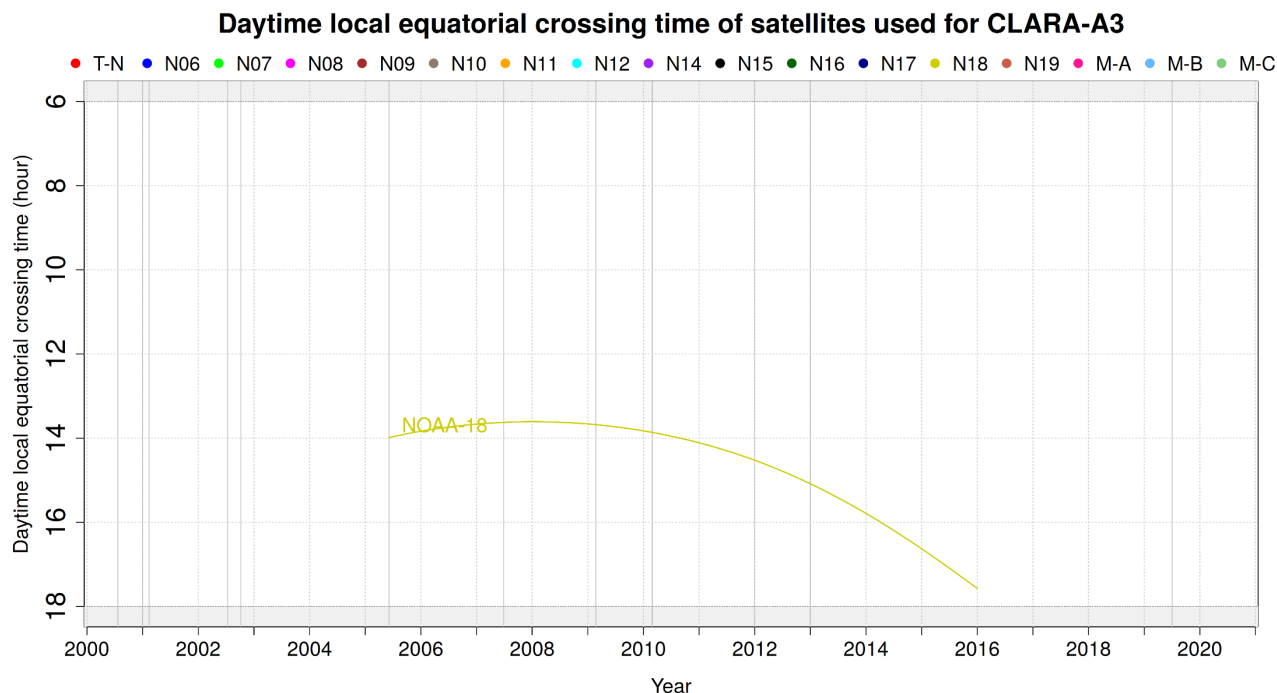
Figure 10-18: Daytime local equator crossing time of afternoon satellite NOAA-18.

The resulting global mean bias is shown in Figure 10-19.

Global mean bias of daily CLARA-A3 RSF w.r.t. other data records (*)
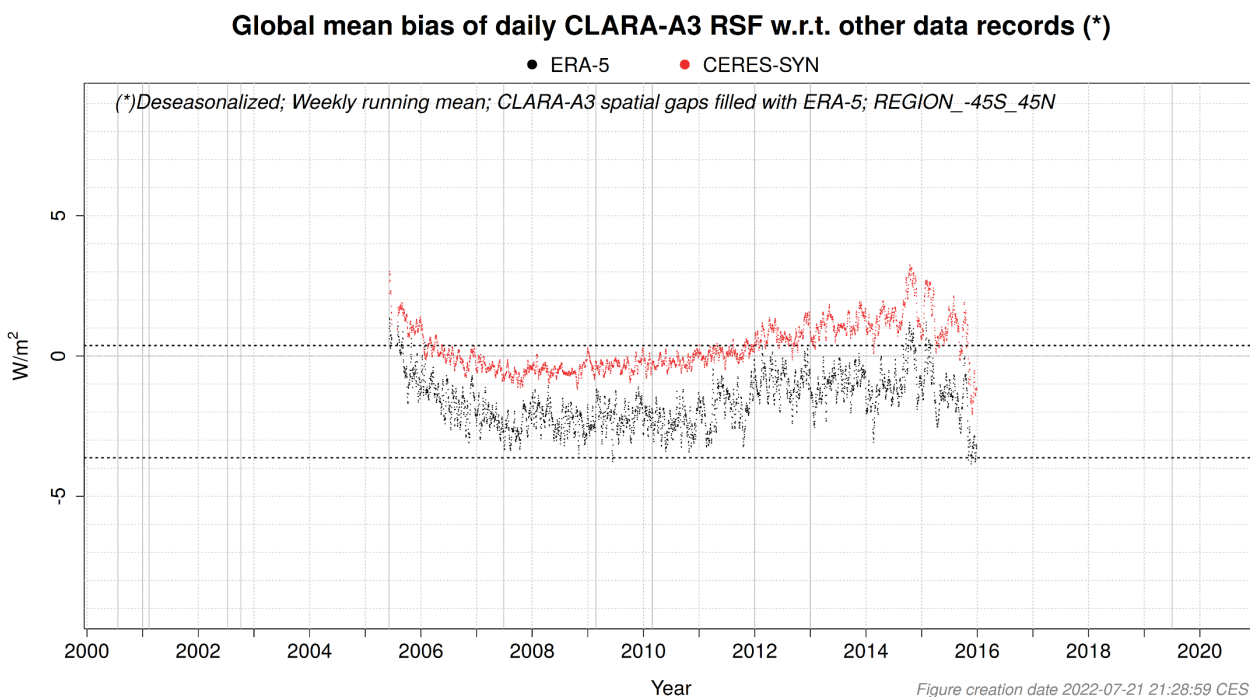
Figure 10-19: Global mean bias of deseasonalized daily CLARA-A3 RSF w.r.t. CERES-SYN, using only the afternoon satellite NOAA-18

The first half year (second half of 2005) presumably suffers from the post-launch positive bias mentioned in Section 10.3.3.

The bias in the subsequent years is probably related to the varying local equator crossing time, resulting in a shifting SZA. This is explained as follows. Figure 10-20 is an adaptation of Guilbert et al. (2022), adjusted for the specific case of CLARA-A3 RSF from NOAA-18: it depicts the subtle mechanisms of how SZA can influence the daily mean RSF bias (note that this is a conceptual figure, not to scale); it is clear that the CERES TRMM albedo model is more curved compared to the relations found in the NOAA-18 observations. Instead of a unidirectional impact (orbital drift -> higher SZA -> higher bias, such as depicted in Figure 10-17), the TRMM-NOAA18 albedo model shape deviation is a bit more complex, as shown in Figure 10-20 which relates local equator crossing time (ECT, Figure 10-18) with model bias (Figure 10-19): Around year 2006 there is bias of 0W/m² associated with a local Equator Crossing Time (ECT) of around 2PM. After two years (in 2008) the ECT shifts a bit toward noon to arrive at 1:30PM with an associated bias of -1W/m². Again two years later (in 2010) the ECT shifts back toward its initial position at 2PM, again associated with a neutral bias. Then, the orbital drift towards the afternoon starts, reaching a peak positive bias four years later (2014) at an ECT of around 3:30PM. The orbital drift continues towards the late afternoon, reaching an ECT of around 5PM in 2015, associated with a net neutral bias. Half a year later, the ECT approaches the terminator (5:30PM) resulting in a negative bias.
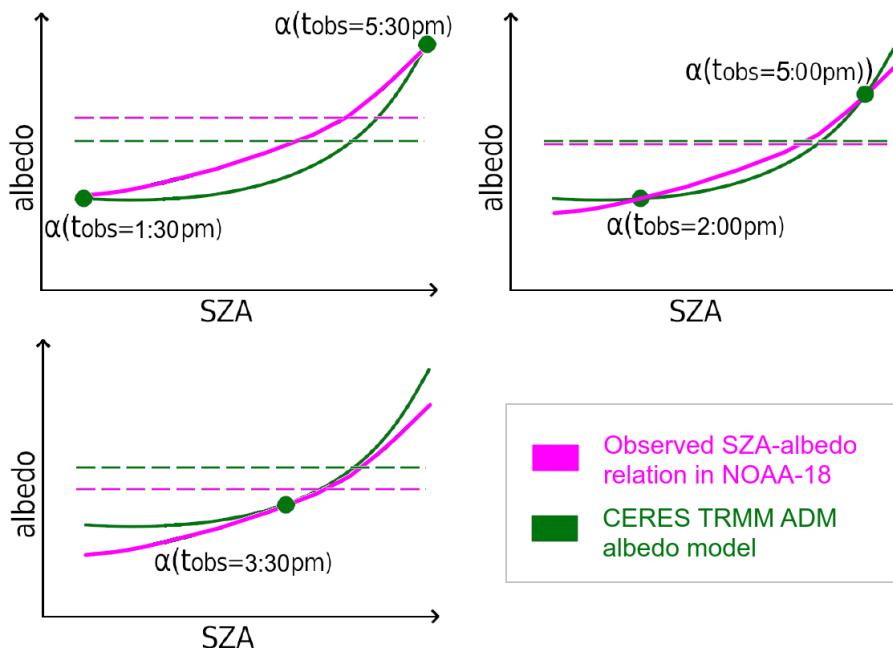
Figure 10-20: Different shape of albedo model (solid lines) leading to different RSF daily mean biases (broken lines) when using CERES TRMM albedo model: (top left) noon or evening observations lead to negative bias; (top right) early or late afternoon observations lead to net zero bias; (bottom) afternoon observations lead to negative bias. Figure adapted from Guilbert et al., 2022.

This effect is most important when only one observation per day is available, and is of minor importance with a larger temporal coverage throughout the day, because it decreases the importance of the albedo model interpolation in favour of direct interpolation between observations. In fact, there are no clear examples of RSF bias (in Figure 10-10) predominantly caused by albedo model deficiencies, even in the periods with suboptimal orbital configuration (most susceptible for these errors). However, it is difficult to separate the Level-2b stability (due to either Level-1c FDR stability,

or other factors such as potential trends in input cloud properties) on the one hand, and albedo model deficiencies on the other hand. The main reason is that a proper validation of the Level-2b stability requires the reflectance (if narrowband) or albedo (if converted to broadband) to be corrected for the satellite's drift in local solar time (i.e., SZA): however, such a correction is to be performed using the albedo models. So disentangling the two is difficult. However, for certain cases it is possible, e.g. for non-drifting satellites such as the MetOp constellation (see Section 11.3.3).

10.4 RSF processing error (regional uncertainty) during 1979-1999

Three orbital configurations exist in the pre-CERES (1979-1999) period of the data record (morning-only, afternoon-only, and morning+afternoon). For each of them, a similar configuration is simulated using more recent satellites (from the CERES era), by creating daily and monthly mean RSF fluxes with only a selection of the available satellites. Since we know the temporal coverage of the diurnal cycle is the largest source of bias, this method will provide an estimate of the uncertainty during the pre-CERES era.

Before this is done, a test is performed to demonstrate the effect of suboptimal orbital configuration on the processing error (MAB), using data from the CERES era. It concerns the same experimental orbital configuration as shown in Figure 10-11 and Figure 10-12, i.e. with afternoon-only satellites NOAA-18 and NOAA-19 (Annex 10.3.1), but here the MAB is discussed, instead of the mean bias. The results are shown in Figure 10-21.

Global MAB between daily CLARA-A3 RSF and other data records (*)
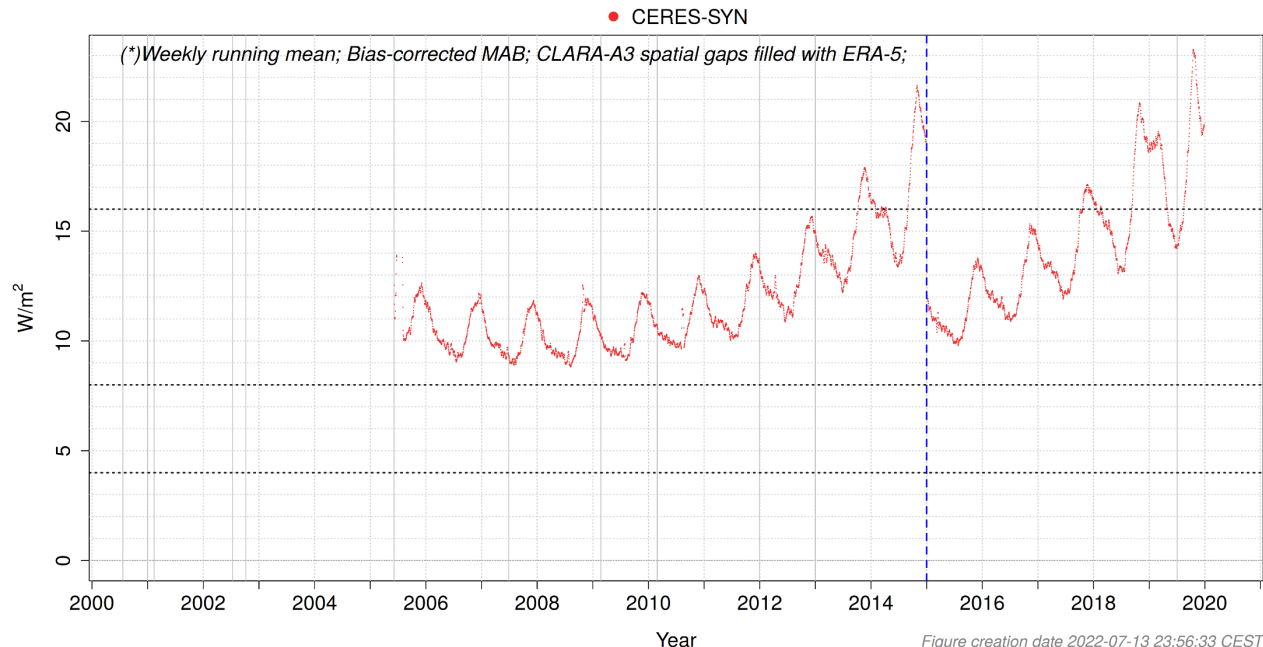
Figure 10-21: Global MAB between daily CLARA-A3 RSF and CERES-SYN, using only afternoon satellites NOAA-18 (2005-2014) or NOAA-19 (2015-2019)

Similar to the mean bias of this experimental afternoon-only CLARA-A3 time series (Figure 10-12), this suboptimal orbital configuration introduces seasonal patterns that were absent in the all-satellite configuration (Figure 5-5). Additionally, for the MAB we also notice a gradual increase of MAB (decrease of performance) due to orbital drift from afternoon towards evening (increasing solar zenith angle).

The increase in MAB as well as its seasonality can be explained by a degrading temporal coverage over regions characterized by large-scale regional climate phenomena with an asymmetric diurnal cycle (e.g. marine stratus thinning or land convection). This introduces strong regional biases, as documented in Section 5.3.1, which can be positive or negative, depending on the region and the kind of phenomena. Furthermore, a degrading temporal coverage also introduces strong biases with fast moving small-scale or heterogeneous weather systems (e.g., fronts), typically consisting of swirls with positive alongside negative bias, caused by temporal extrapolation. Globally averaged together, all these biases vary seasonally because of a hemispherical imbalance of the associated regional climate features' occurrence and strength, explaining the seasonal pattern of MAB. With any degradation of the temporal coverage (orbital configuration) these regional biases also grow accordingly, which directly leads to an increase of the global MAB regardless the sign of these regional biases. The sign of these regional biases is also the reason why the global mean bias during the CERES era remains relatively stable: negative and positive regional biases tend to balance (compensate) each other.

Now that the source of this increased MAB and its seasonality is established, the objective is to imitate the pre-CERES era processing error using CERES-era data. The experiment with NOAA-18-only and NOAA-19-only satellite is not suitable to imitate the afternoon-only periods in the early years of the record, as they are not representative for the local ECT (Equator Crossing Time) observed at those times. Therefore, each of the three typical pre-CERES orbital configurations (afternoon-only, morning-only, and morning+afternoon) are mimicked using carefully selected CERES-era equivalent orbital configurations in terms of ECT: an overview is provided in Table 10-2.

Table 10-2: Estimation of RSF uncertainty during pre-CERES era

| Orbital config.	Pre-CERES (1979-1999) orbital configuration		CERES-era (2000-2020) equivalent orbital configuration (similar local ECT**)			
	Satellite (*)	Duration (months)	Satellite*	Duration	MAB (W/m²)	
					MM	DM
(I) Afternoon satellite only	T-N N-7 N-7 N-9 N-9	1979/01-1980/01: 13 1981/09-1983/08: 24 1984/06-1985/01: 08 1985/02-1985/06: 05 1985/11-1986/10: 12 ***Total: 62 months***	N-16 N-18 N-19	2005/01-2006/12 2012/01-2013/12 2016/01-2017/12	Range: 4.1 - 6.3 Mean: **4.8** <8 (100%) <4 (0%) <2 (0%)	Range: 10.6 - 19.1 Mean: **13.5** <16 (88.7%) <8 (0%) <4 (0%)
(II) Morning satellite only	N-6 N-12	1980/02-1981/08: 19 1994/09-1994/12: 04 ***Total: 23 months***	N-15 N-15 N-16 N-18	2000/03-2000/07 2001/03-2001/06 2011/01-2011/12 2017/07-2018/06	Range: 5.2 - 8.6 Mean: **6.6** <8 (84.8%) <4 (0%) <2 (0%)	Range: 13.2 - 23.1 Mean: **17.4** <16 (35.3%) <8 (0%) <4 (0%)
(III) Afternoon and morning satellite only	N-7/-8 N-9/-8 N-9/-10 N-11/-10 N-11/-12 N-14/-12 N-14/-15	1983/09-1984/05: 09 1985/07-1985/10: 04 1986/11-1988/10: 24 1988/11-1991/09: 35 1991/10-1994/08: 35 1995/01-1998/11: 47 1998/12-2000/02: 15 ***Total: 169 months***	N-14/-15 N-15/-16 N-16/-18 N-18/-19	2001/03-2001/07 2004/01-2005/12 2011/01-2011/12 2017/07-2018/06	Range: 2.3 - 6.3 Mean: **3.6** <8 (100%) <4 (73.6%) <2 (0%)	Range: 6.7 – 19.0 Mean: **10.4** <16 (99.2%) <8 (15.7%) <4 (0%)
(*) T=TIROS, N=NOAA; (**) Equator Crossing Time						

The CERES-era equivalents for the afternoon-only orbital configuration are shown in Figure 10-22, while Figure 10-23 shows the resulting monthly MAB.
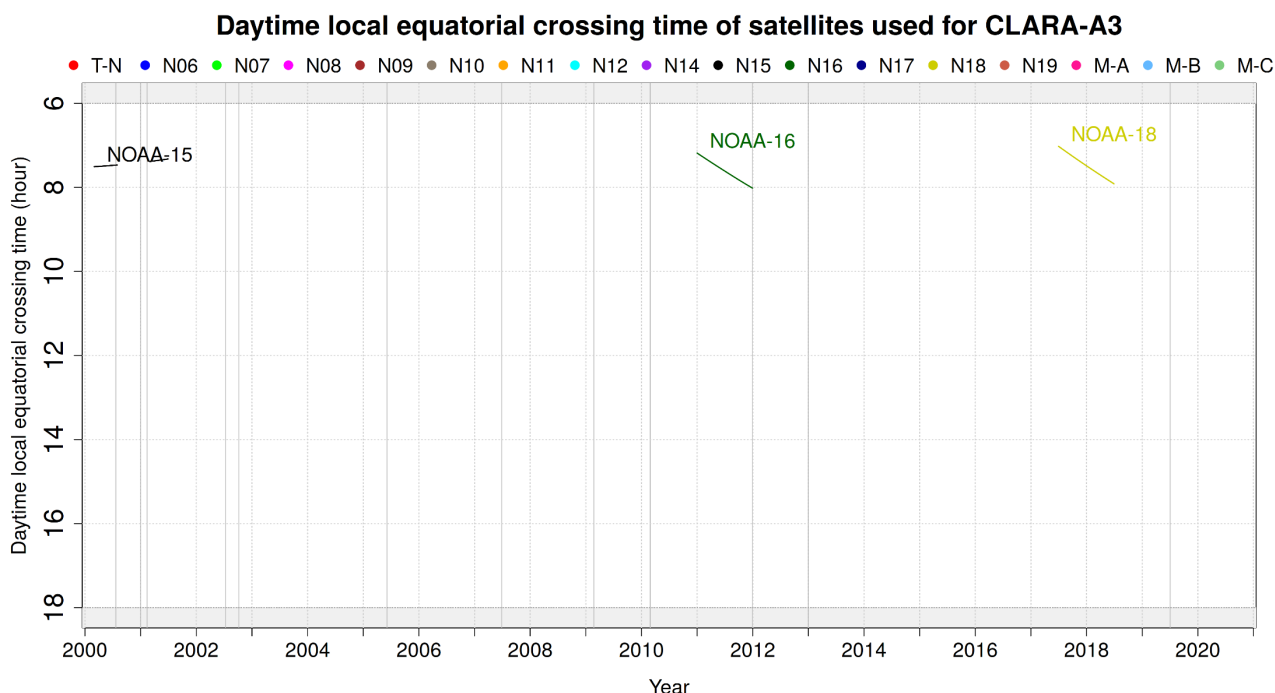
Daytime local equatorial crossing time of satellites used for CLARA-A3
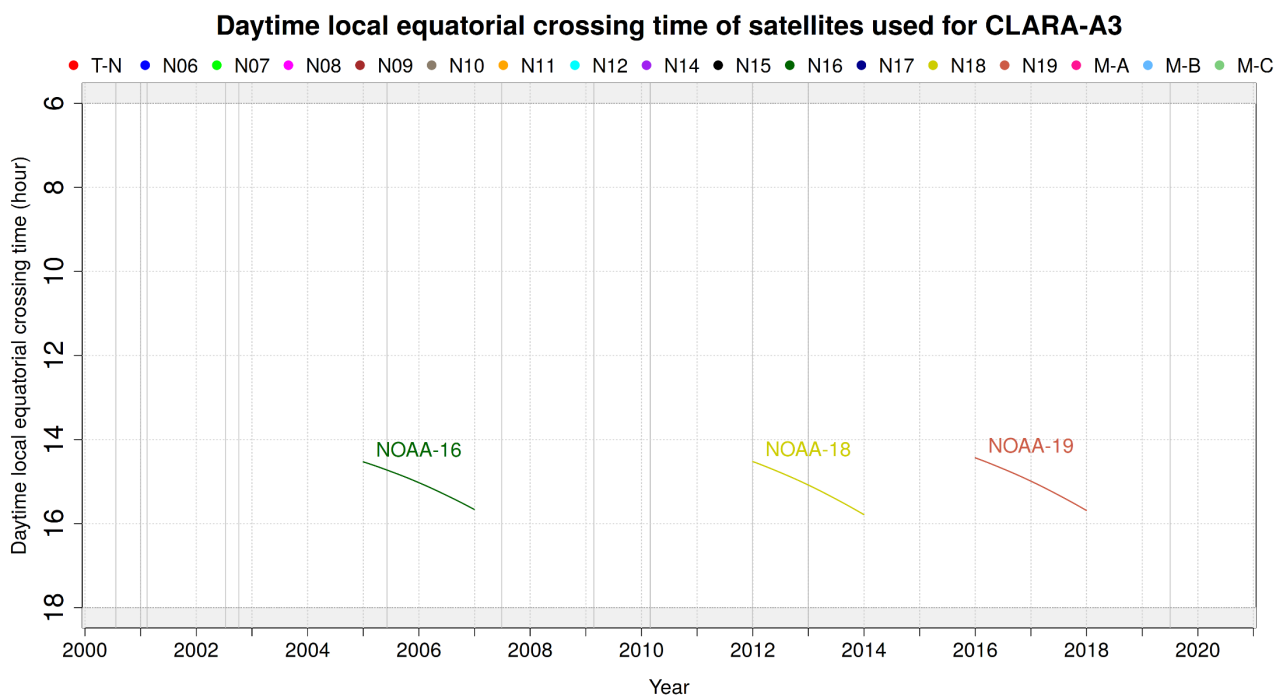
Figure 10-22: Daytime local equator crossing time of satellites used to mimic pre-CERES era afternoon-only orbital configuration

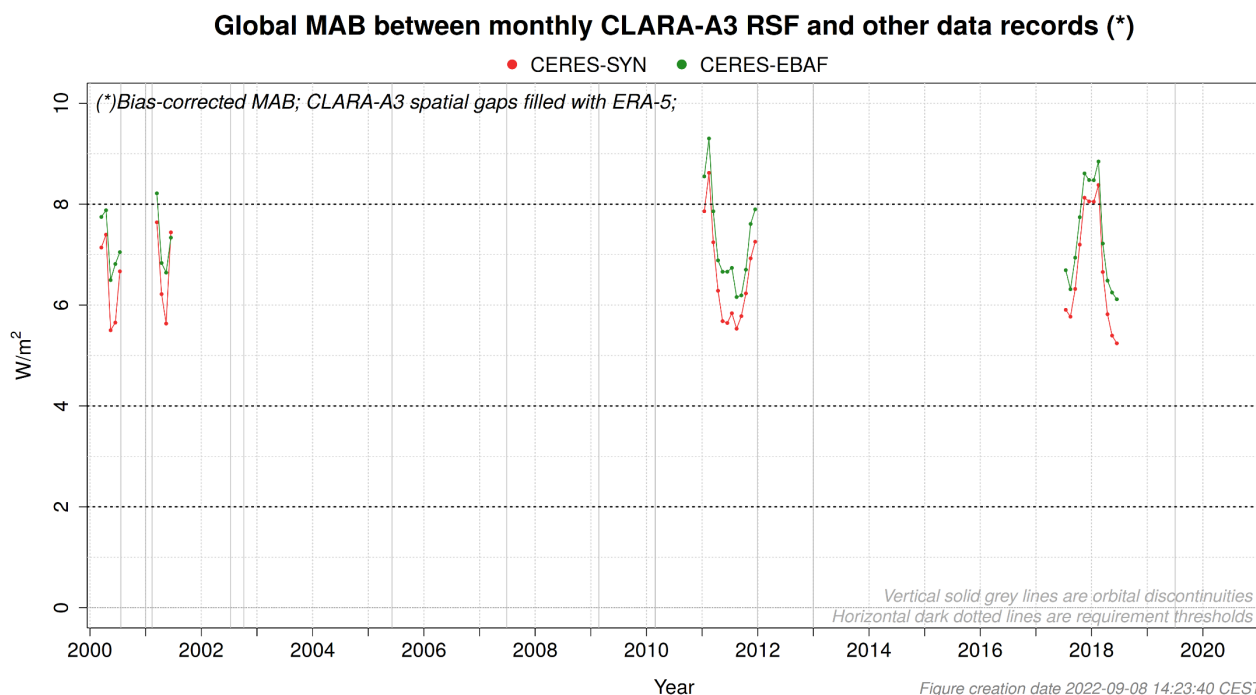**Global MAB between monthly CLARA-A3 RSF and other data records (*)**
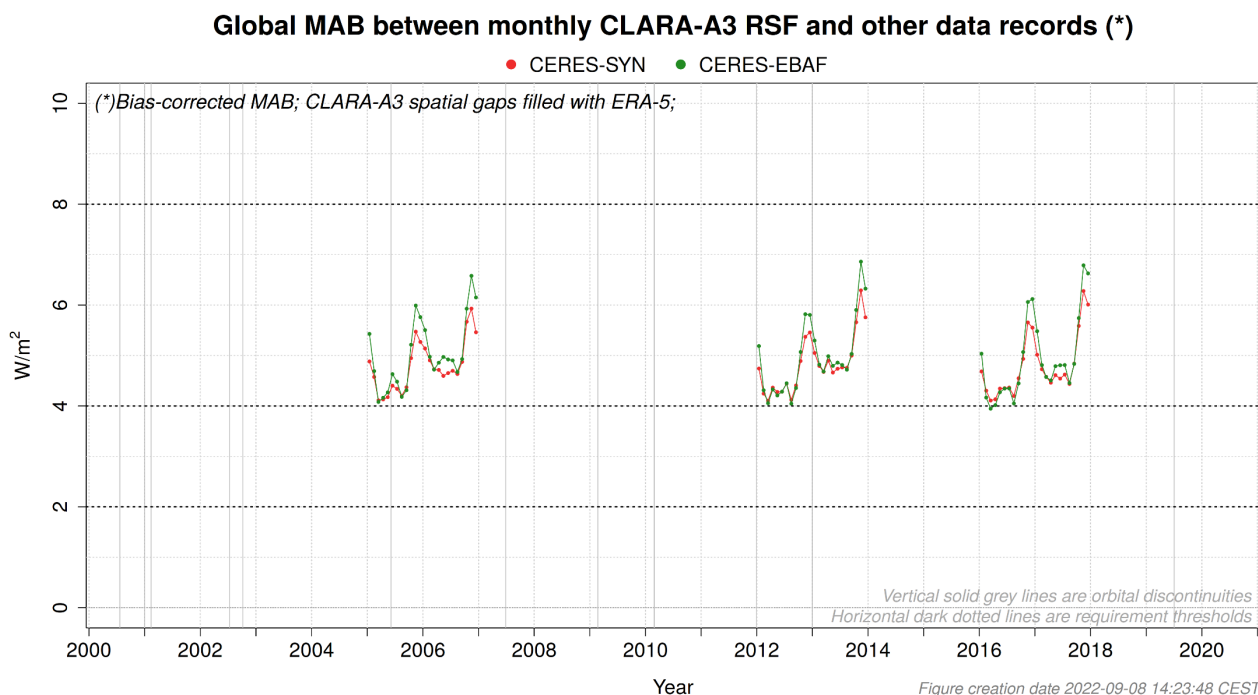
Figure 10-23: Global MAB between monthly CLARA-A3 RSF and CERES, using only afternoon satellites with ECT given by Figure 10-22.

The CERES-era equivalents for the morning-only orbital configuration are shown in Figure 10-24, while Figure 10-25 shows the resulting monthly MAB.
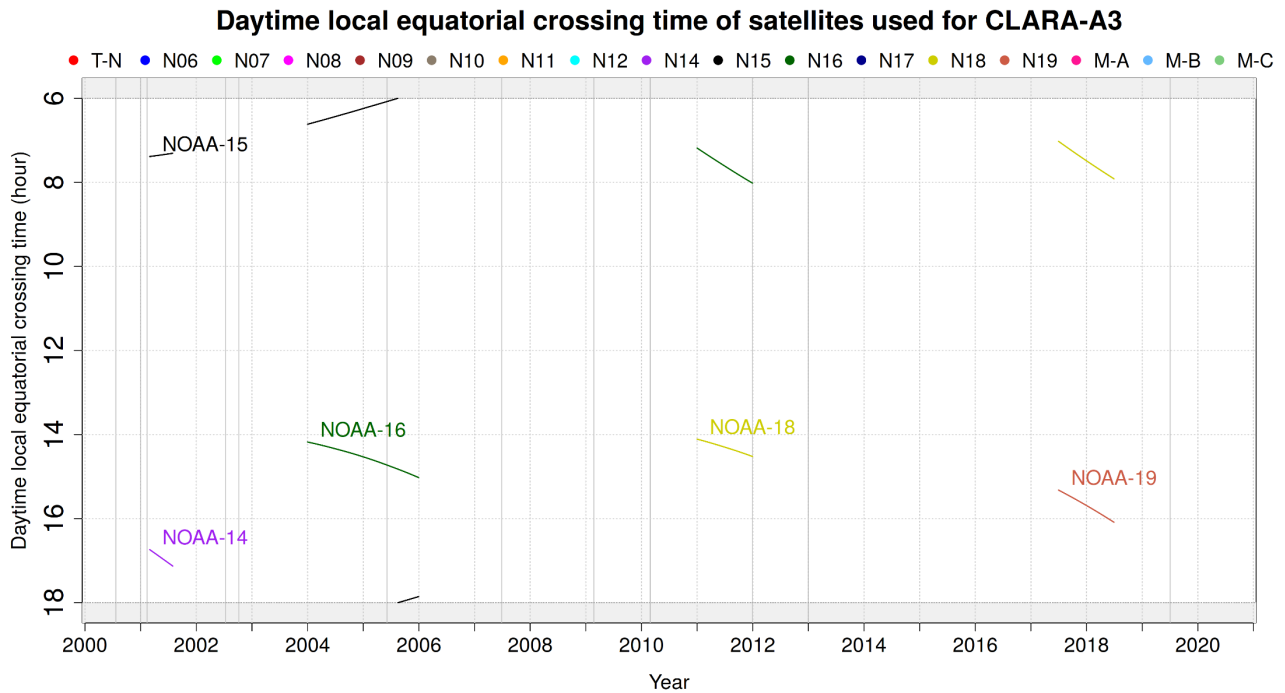
Daytime local equatorial crossing time of satellites used for CLARA-A3

Figure 10-24: Daytime local equator crossing time of satellites used to mimic pre-CERES era morning-only orbital configuration

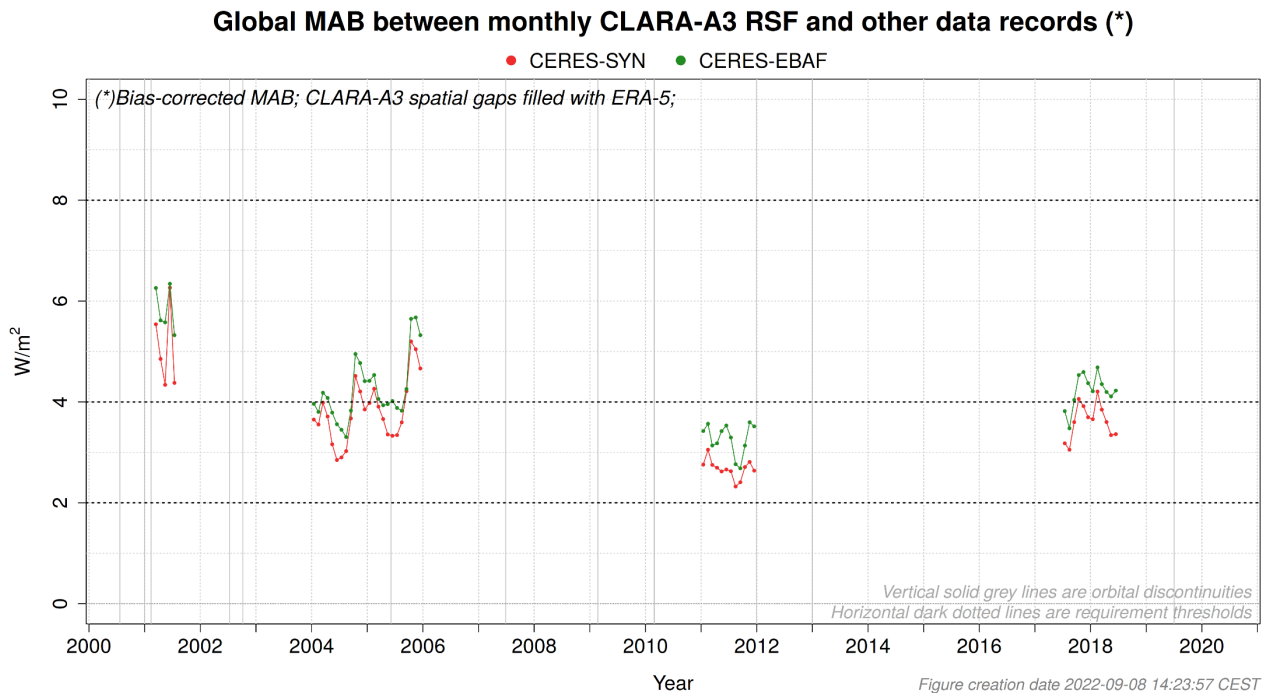### Global MAB between monthly CLARA-A3 RSF and other data records (*)

Figure 10-25: Global MAB between monthly CLARA-A3 RSF and CERES, using only morning satellites with ECT given by Figure 10-24.

The CERES-era equivalents for the afternoon-only orbital configuration are shown in Figure 10-26, while Figure 10-27 shows the resulting monthly MAB.

Figure 10-26: Daytime local equator crossing time of satellites used to mimic pre-CERES era morning+afternoon orbital configuration

Figure 10-27: Global MAB between monthly CLARA-A3 RSF and CERES, using morning+afternoon satellites with ECT given by Figure 10-26.

Using the average MAB for each of the three mimicked suboptimal (pre-CERES) orbital configurations, it is now possible to "fill" the gaps in the entire data record's accuracy time series (i.e. extend Figure 5-4), as given in Figure 10-28.

Global MAB between monthly CLARA-A3 RSF and other data records (*)
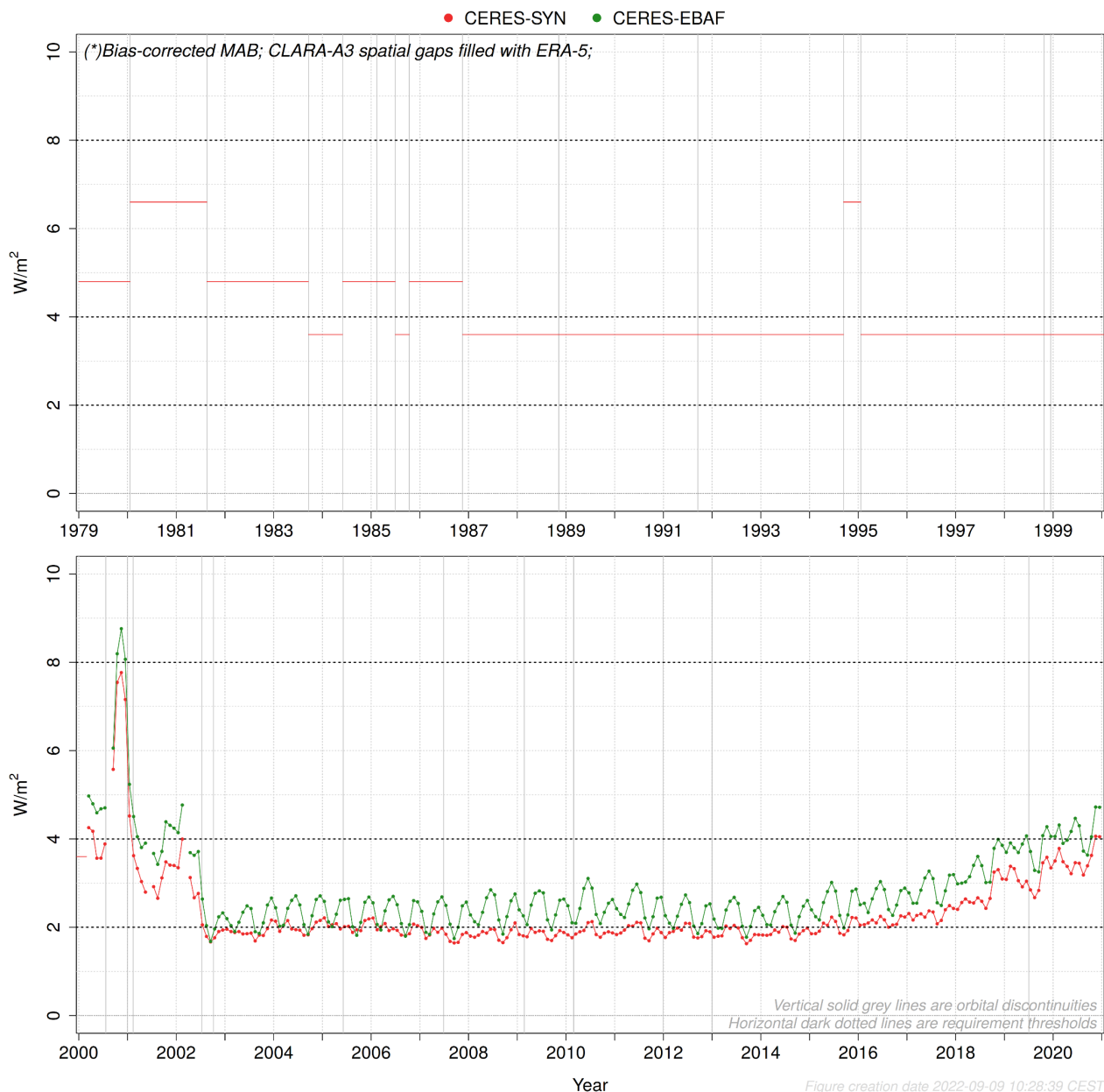
Figure 10-28: Global MAB between monthly CLARA-A3 RSF and other data records; horizontal red lines indicate estimations of MAB during pre-CERES era, based on actual CERES-era MAB using equivalent orbital configurations.

Taking into account the temporal coverage of each of these three typical pre-CERES orbital configurations, it is now possible to construct an estimation of the total monthly RSF uncertainty, by taking a temporally weighted average over the mean MAB's in each period (Table 10-3).
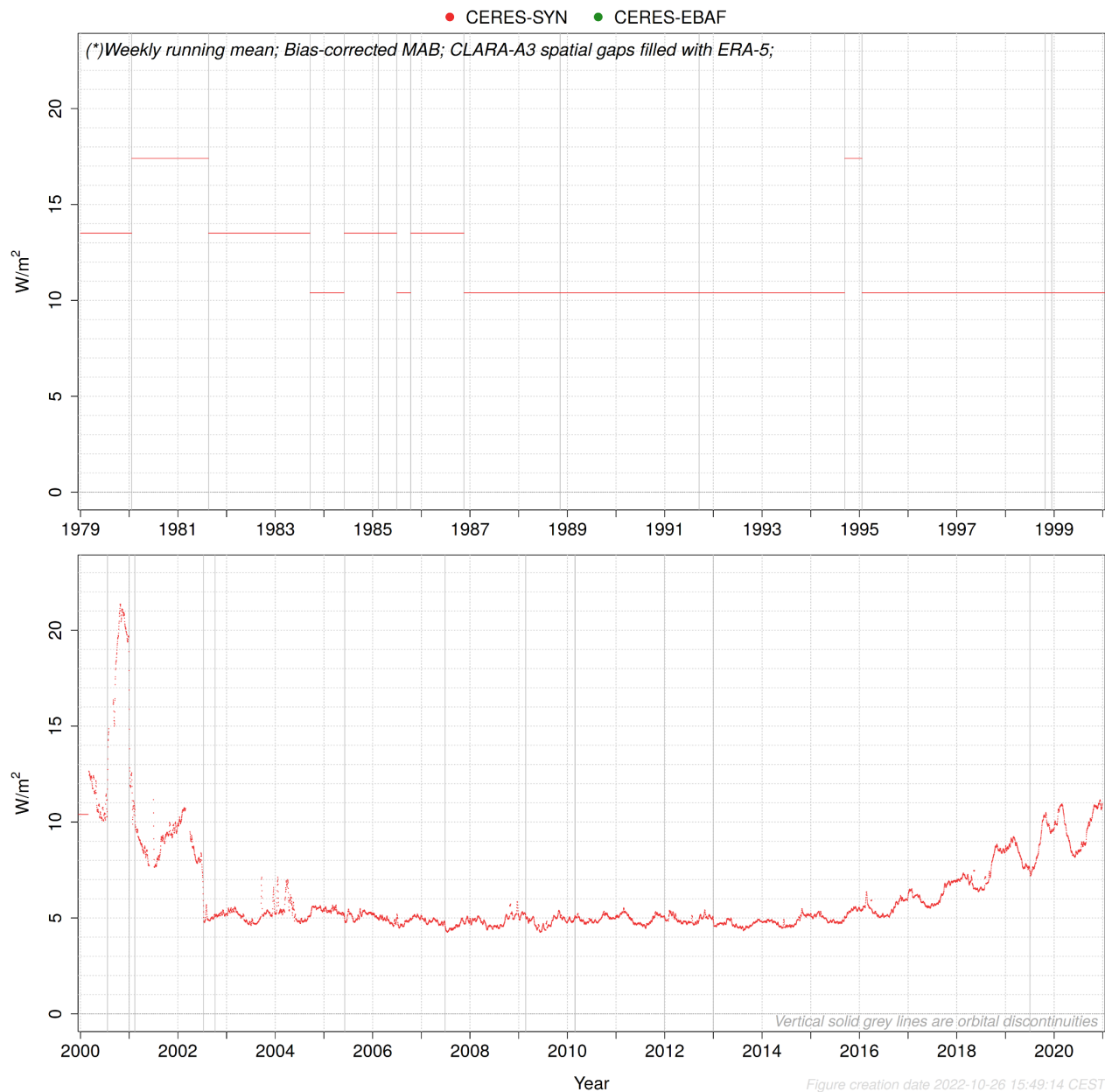
Table 10-3: Estimation of Monthly RSF uncertainty for the entire CLARA-A3 data record

Time span	Orbital configuration	Frequency of meeting the requirement (%)			Average MAB (W/m²)	Temporal share (%)
		Threshold 8 W/m²	Target 4 W/m²	Optimal 2 W/m²		
Pre-CERES era (1979/01-2000/02)	(I) Afternoon satellite only	100%	0.0%	0.0%	4.8 W/m²	62 months (12.4%)
	(II) Morning satellite only	84.8%	0.0%	0.0%	6.6 W/m²	23 months (4.6%)
	(III) Afternoon and morning satellite only	100%	73.6%	0.0%	3.6 W/m²	169 months (33.7%)
CERES era (2000/03-2020/12)	All available satellites	100%	96.4%	49.8%	2.3 W/m²	247 months (49.3%)*
Total record	**All available satellites**	**99.3%**	**72.4%**	**24.6%**	**3.2 W/m²**	**501 months (100%)***
(*) Three months are not taken into account: August 2000, June 2001 and March 2002, since CERES products contain data gaps in those months;						

The same is done for the daily means (Figure 10-29 and Table 10-4).

Global MAB between daily CLARA-A3 RSF and other data records (*)

Figure 10-29: Global MAB between daily CLARA-A3 and CERES-SYN RSF; horizontal red lines indicate estimations of MAB during pre-CERES era, based on actual CERES-era MAB using equivalent orbital configurations.

Table 10-4: Estimation of Daily RSF uncertainty for the entire CLARA-A3 data record

Time span	Orbital configuration	Frequency of meeting the requirement (%)			Average MAB (W/m²)	Temporal share (%)
		Threshold 16 W/m²	Target 8 W/m²	Optimal 4 W/m²		
Pre-CERES era (1979/01-2000/02)	(I) Afternoon satellite only	88.7%	0.0%	0.0%	13.5 W/m²	62 months (12.4%)
	(II) Morning satellite only	35.3%	0.0%	0.0%	17.4 W/m²	23 months (4.6%)
	(III) Afternoon and morning satellite only	99.2%	15.7%	0.0%	10.4 W/m²	169 months (33.7%)
CERES era (2000/03-2020/12)	All available satellites	98.4%	81.2%	0.0%	6.2 W/m²	247 months (49.3%)*
Total record	**All available satellites**	**94.5%**	**44.9%**	**0.0%**	**9.0 W/m²**	**501 months (100%)***

(*) Three months are not taken into account: August 2000, June 2001 and March 2002, since CERES products contain data gaps in those months;

10.5 RSF bias regional patterns investigation and assessment

10.5.1 Asymmetric diurnal cycle: example marine stratocumulus thinning

The orbital configuration affects the frequency and timing of the observations throughout the day, leading to different kinds of biases (e.g. bias due to absence of afternoon satellite is not necessarily the same as bias due to absence of morning satellite). To obtain the daily mean RSF, a SZA-dependent albedo model (a.k.a. shortwave diurnal cycle model) extrapolates the observations symmetrically w.r.t. noon (cfr [RD 7]). Hence, the sensitivity of the daily mean RSF bias is related to the (a)symmetry of the actual observed diurnal cycle (*w.r.t. noon*), which, in turn, depends on specific regional climatological phenomena.

An important example are the typical **marine stratocumulus** regions located over the eastern parts of tropical ocean basins (Eastman and Warren, 2014), where clouds persist in the morning and dissolve (i.e. become thinner, 'burn off') during the afternoon (Figure 10-30). In these areas the large-scale subsiding atmosphere flow brings dry and warm air downward and ocean upwellings bring cold water upward (Brient et al., 2019).
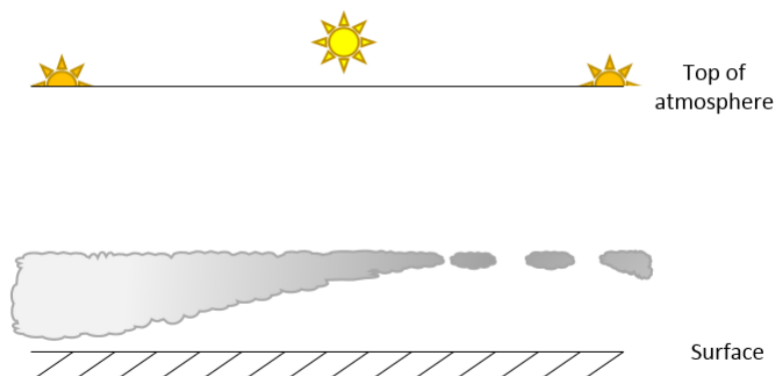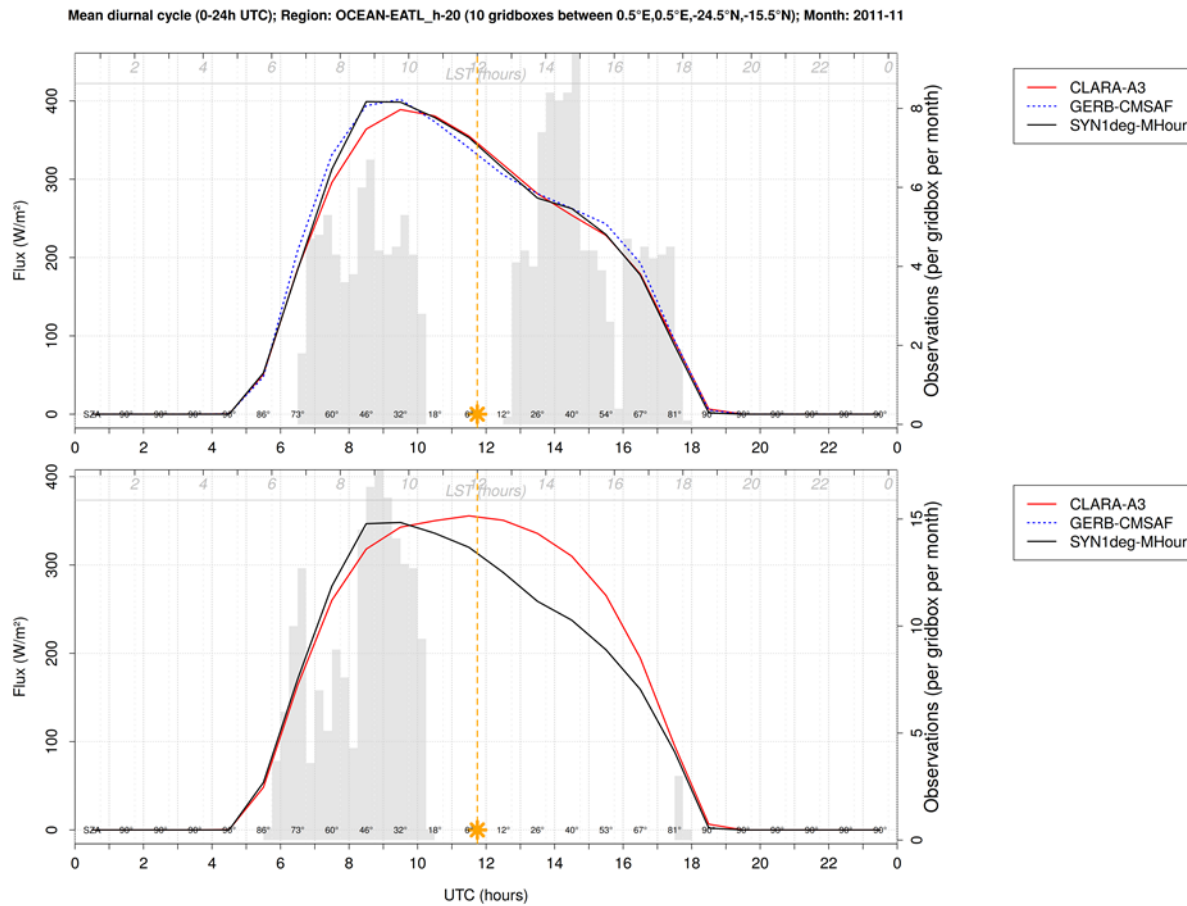
**Figure 10-30:** Thinning of marine stratocumulus (Gristey et al., 2018)

The importance of temporal coverage is demonstrated by zooming in on a representative region located in the eastern Atlantic Ocean (Figure 10-31, calculated from region '30' in Figure 10-32): as shown by the good correspondence with both CERES[5] and GERB[6], the RSF monthly mean diurnal cycle for October 2011 is well captured in CLARA-A3, which at that time makes use of the NOAA-18 and NOAA-19 afternoon observations (cfr Figure 4-10). The daily mean bias of CLARA-A3 w.r.t. CERES is relatively small (-1.9W/m²). However, the RSF for the month October 2020, in the same region and during the same season, suffers from a large positive bias, which is caused by the absence of available afternoon satellites (cfr Figure 4-10: orbital drift of NOAA-19 after 2016): the SZA-dependent albedo model (shortwave diurnal cycle model) extrapolates the morning observations symmetrically w.r.t. noon, ignoring the reality of dissolving ('thinning' or 'burning') of marine stratus. This results in a daily mean bias of +12.7W/m². Such biases are strong enough to be visible in the annual mean bias (Figure 5-6: red ocean areas west of African and South American land masses during 2019-2020). The opposite would be true for an absence of (mid-) morning observations (cfr Figure 4-10: no midmorning orbit before 2002), whereby the afternoon observation would be symmetrically extrapolated into the morning, leading to an underestimation of daily mean RSF (Figure 5-6: blue ocean areas west of African and South American land masses during 2000-2001).

---

[5] CERES-SYN1deg-MHour Ed.4.1 (https://ceres-tool.larc.nasa.gov/ord-tool/)
[6] TOA Radiation from GERB/SEVIRI ed.2.0: TRS Mean diurnal cycle (https://wui.cmsaf.eu/)

Mean diurnal cycle (0-24h UTC); Region: OCEAN-EATL_h-20 (10 gridboxes between 0.5°E,0.5°E,-24.5°N,-15.5°N); Month: 2011-11

Figure 10-31: Monthly mean diurnal cycle of CLARA-A3 RSF for October 2011 and 2020, in 10 Eastern Atlantic grid boxes (region '30' in Figure 10-32). Grey bars indicate observation density. (top) year 2011; (bottom) year 2020, without afternoon observations.
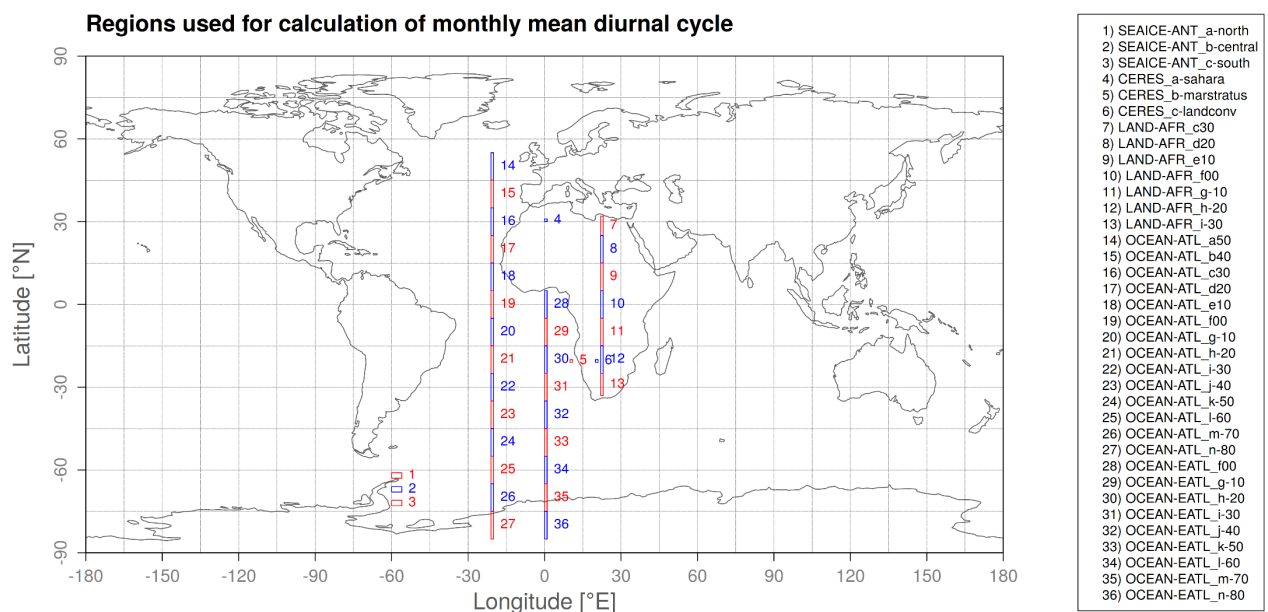

Figure 10-32: regions used for calculation of monthly mean diurnal cycle

10.5.2 Underestimation over Antarctica with late afternoon orbit (months DJF)

This section shows some monthly mean RSF diurnal cycles for a region in Antarctica, denoted by number '36' in the map of Figure 10-32. All diurnal cycles are created for the month November, and only differ in year and orbital configuration.

Figure 10-33 and Figure 10-34 show respectively the absolute flux and the bias of monthly mean RSF diurnal cycle for November **2000**, when only the (drifted) late afternoon satellite **NOAA-14** is available. Every hour, the cloud cover (CLD) and cloud optical thickness (COT) of both CLARA-A3 and CERES are shown below in the graph. Between 3h-18h UTC there are very large negative biases (up to -100W/m² at noon (!)); these are accompanied with fully overcast CLARA-A3 cloud cover (100%) whereas CERES keeps more or less the same cloud cover as the hours before and after (between 30-40%). The resulting monthly mean bias for this region amounts -51.3W/m². This type of bias is, however, not seen in Figure 10-35, showing the bias' diurnal cycle for the same region and month but now for the year **2011**, i.e. when many **more satellites** are available, resulting in a modest monthly mean bias of -7.6 W/m².
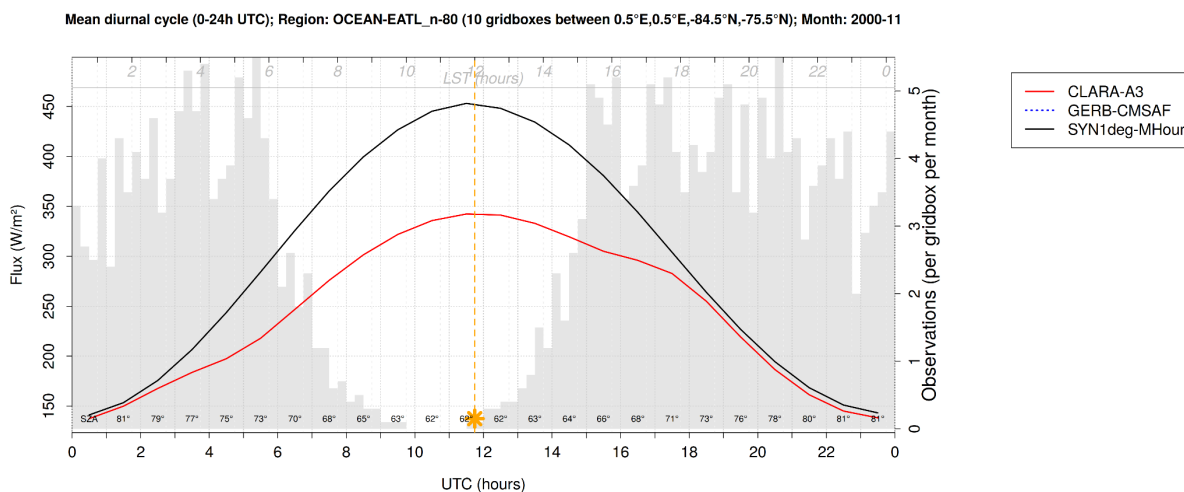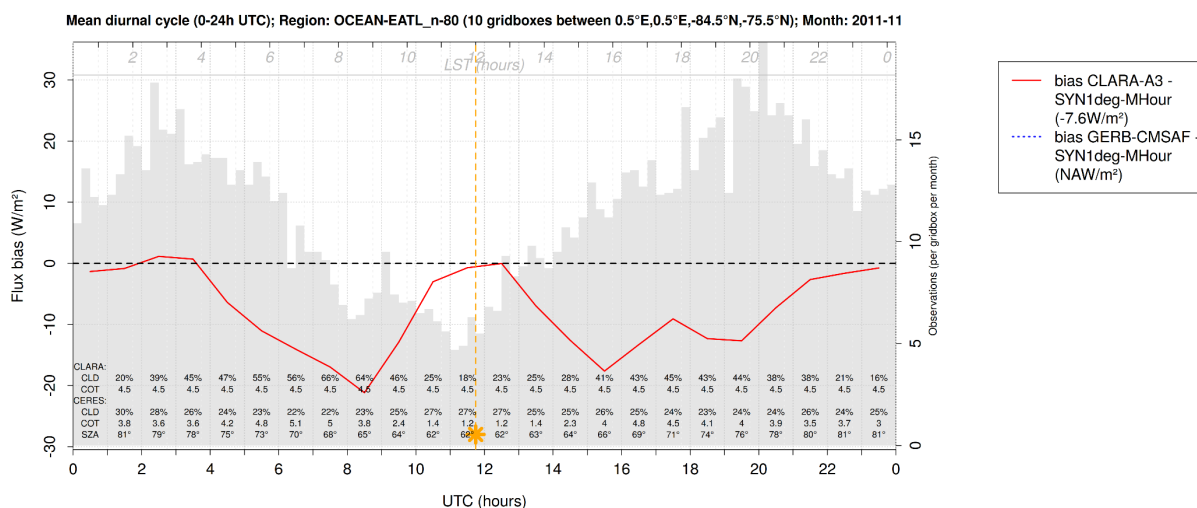

Figure 10-33: CLARA-A3 and CERES-SYN RSF monthly mean diurnal cycle for November <u>2000</u>, in 10 Antarctic grid boxes (region '36' in Figure 10-32). Grey bars indicate observation density.
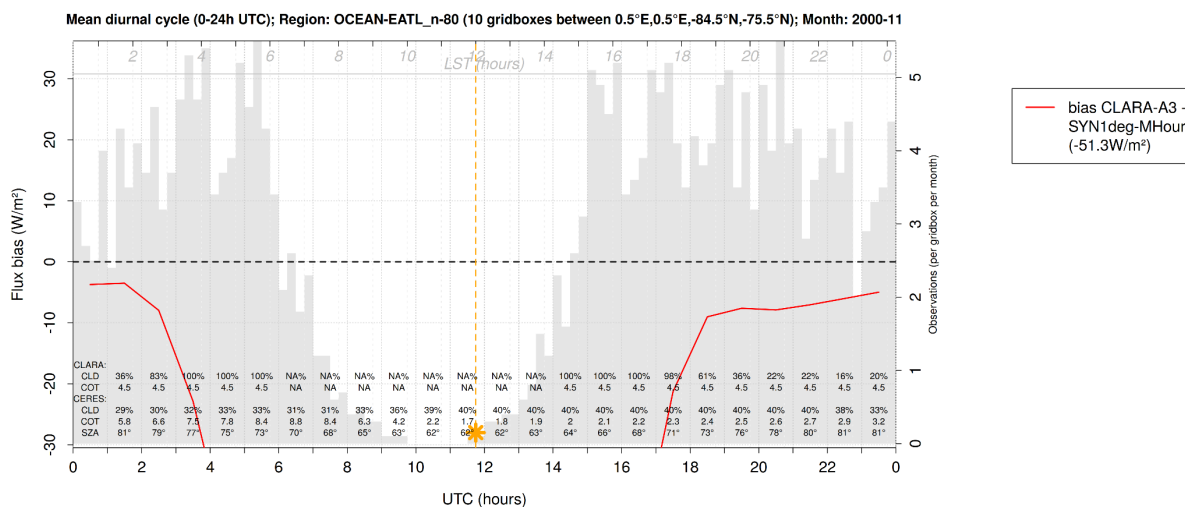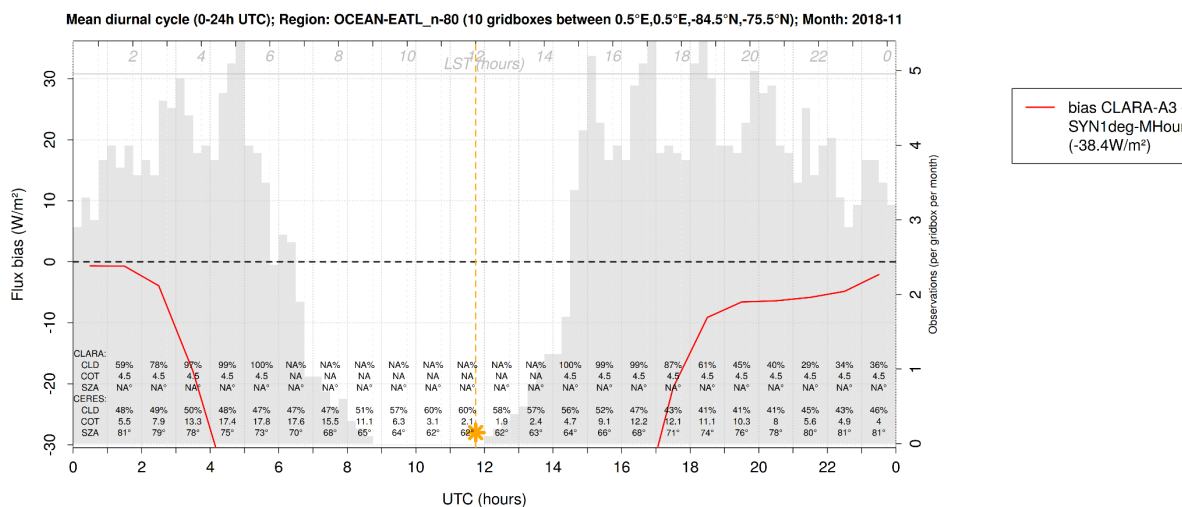
Figure 10-34: Bias between CLARA-A3 and CERES-SYN RSF monthly mean diurnal cycle for November 2000, in 10 Antarctic grid boxes (region '36' in Figure 10-32). Grey bars indicate observation density.


Figure 10-35: Bias between CLARA-A3 and CERES-SYN RSF monthly mean diurnal cycle for November 2011, in 10 Antarctic grid boxes (region '36' in Figure 10-32). Grey bars indicate observation density.

Finally, a test is performed to verify whether the satellite itself (NOAA-14) might be the problem. This is done by mimicking the same orbital configuration using (only!) the – also drifting – satellite **NOAA19** during November **2018**, which then has about the same ECT as NOAA-14 had in November 2000. The result is shown in Figure 10-36, which shows a similar bias pattern; it can be concluded that this bias is not related to a specific satellite (e.g. NOAA-14) but to a specific set of conditions (snow/ice surface type, southern hemisphere summertime, drifted satellite with late afternoon ECT, large biases in retrieved cloud cover).

Figure 10-36: Bias between CLARA-A3 and CERES-SYN RSF monthly mean diurnal cycle for November 2018, obtained by only making use of the drifted satellite NOAA-19, in 10 Antarctic grid boxes (region '36' in Figure 6 9). Grey bars indicate observation density.
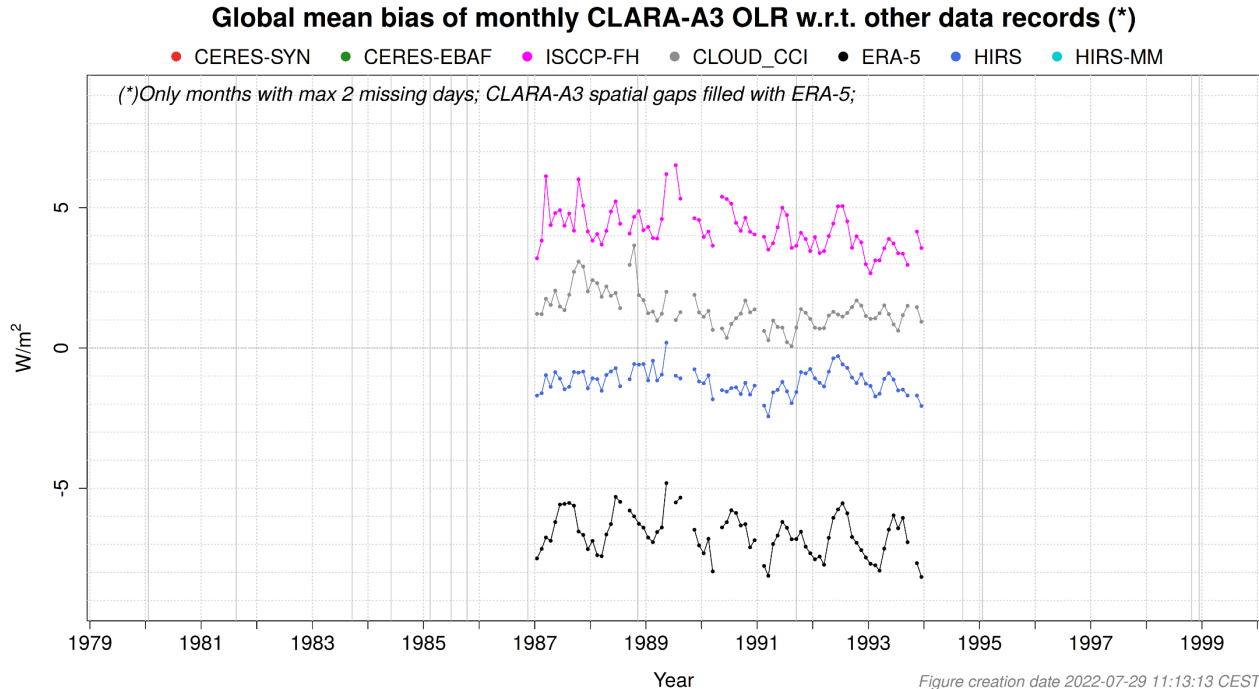
10.6 OLR mean bias investigation and assessment

In this section an experiment is done to investigate the negative bias anomaly of OLR in the first few years of the record (based on TIROS-N and NOAA-6). This is done with a rudimentary test: the daily and monthly mean OLR is calculated using **only** the morning satellites NOAA-10 or NOAA-12 (Figure 10-37). The comparison between NOAA-10-based and NOAA-6-based OLR would give us more certainty about the calibration hypothesis because it excludes other potential reasons (like instrument, since they are both AVHRR/1 instruments lacking the 10.8μm thermal channel, and like orbital configuration, since they are both morning orbits with nearly the same ECT, i.e. SZA).

**Daytime local equatorial crossing time of satellites used for CLARA-A3**
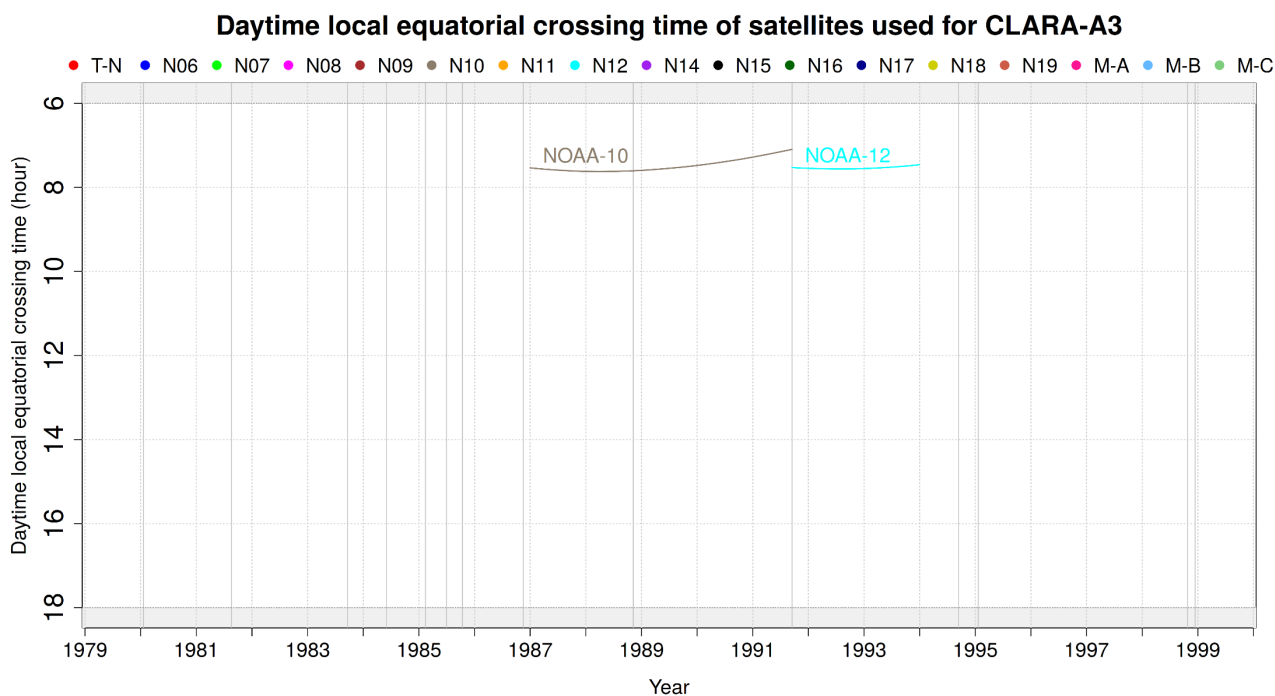
Figure 10-37: Daytime local equator crossing time of morning satellites NOAA-10 (1987-1991) or NOAA-12 (1991-1993)

The resulting monthly mean bias is shown in Figure 10-38.

Global mean bias of monthly CLARA-A3 OLR w.r.t. other data records (*)

Figure creation date 2022-07-29 11:13:13 CEST

Figure 10-38: Global mean bias of monthly CLARA-A3 OLR w.r.t. other data records, using only morning satellites NOAA-10 (1987-1991) or NOAA-12 (1991-1993)

For comparison, the all-satellite (default) orbital configuration is shown in Figure 10-39.

Global mean bias of monthly CLARA-A3 OLR w.r.t. other data records (*)

● CERES-SYN   ● CERES-EBAF   ● ISCCP-FH   ● CLOUD_CCI   ● ERA-5   ● HIRS   ● HIRS-MM
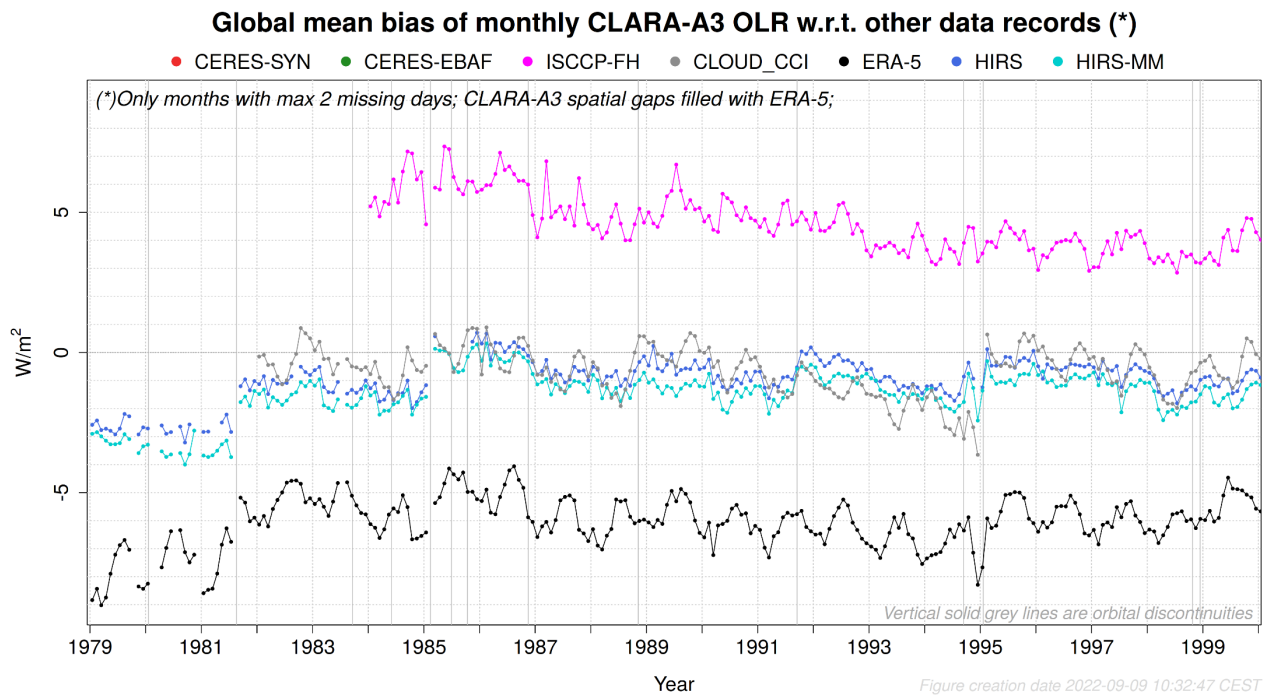
Figure 10-39: Global mean bias of monthly CLARA-A3 OLR w.r.t. other data records, using all available satellite (default configuration).

A comparison of Figure 10-38 and Figure 10-39 (blue curve) indicates that the use of morning-only satellites causes on average a decrease in monthly mean OLR of about -0.5W/m², which is relatively small and not enough to explain the much more negative bias during the morning-only configuration during 1981-1982 (from NOAA-6). The latter can be assumed to be caused by the underlying L2b data record which in turn is likely impacted by an issue with the calibration of the FDR, or an issue with the spectral response correction factors.

Furthermore, the relatively 'flat' bias curve of CLARA-A3 OLR w.r.t. HIRS (blue curve) in Figure 10-38 also indicates that the use of a different instrument (i.e. with or without the 10.8µm thermal channel) does not have a significant impact on the global mean bias; NOAA-10 (1987-1991) has the oldest instrument on board called AVHRR/1, without the 10.8µ channel, whereas NOAA-12 (1991-1993) has a newer instrument on board called AVHRR/2, with the 10.8µ channel.

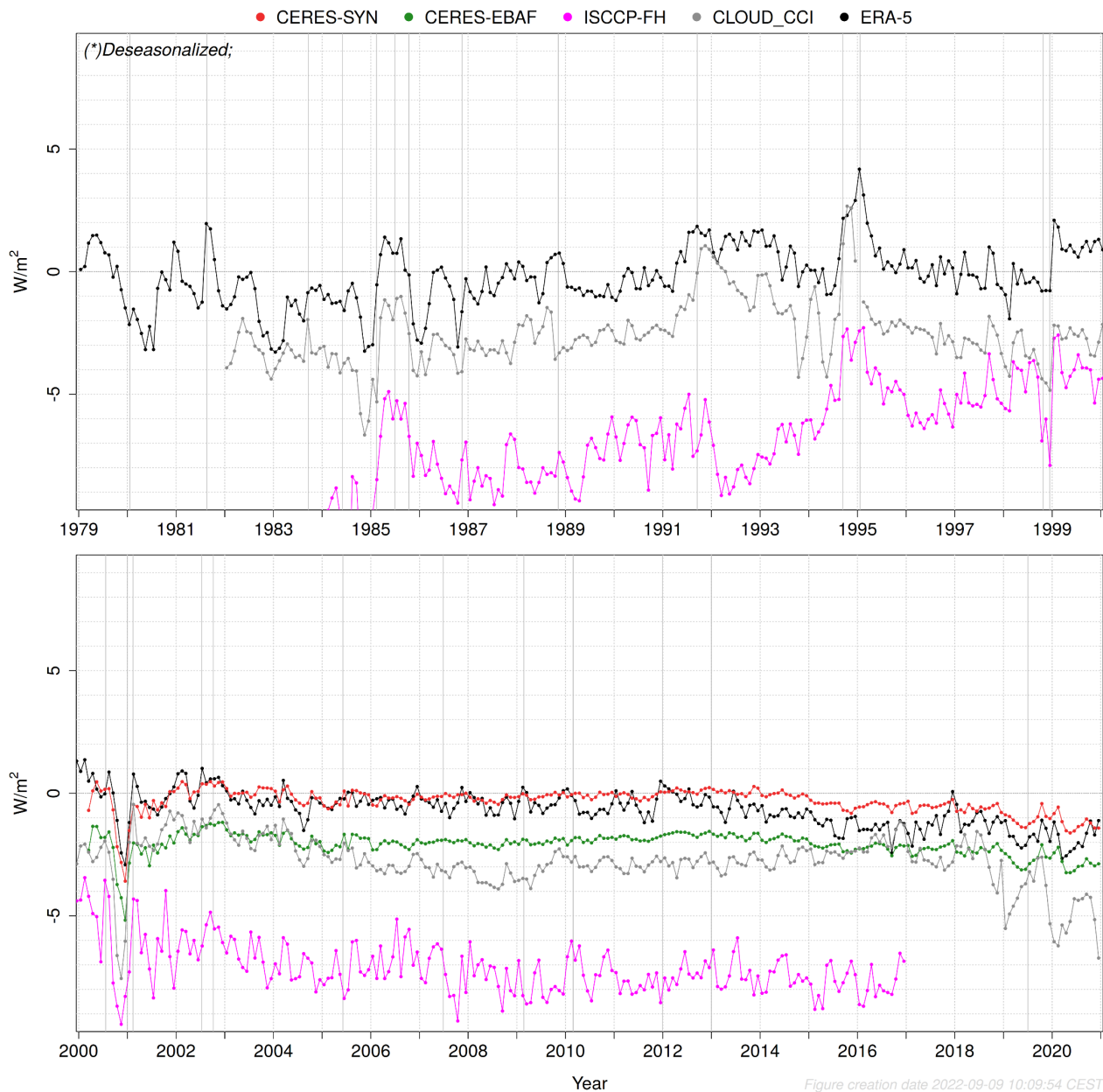## 10.7 Results for RSF without spatial gap-filling of CLARA-A3

10.7.1 Mean bias and stability
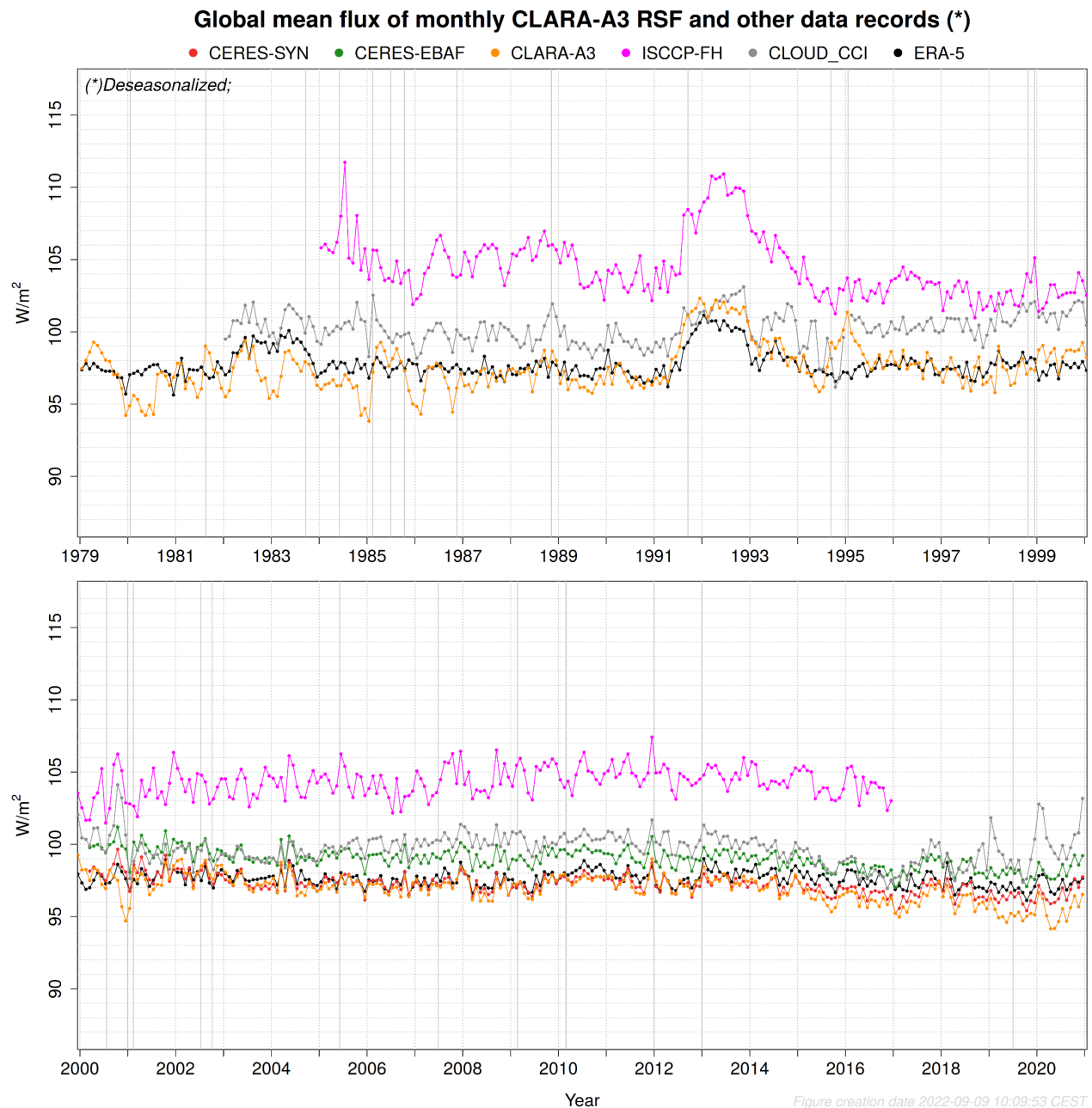
Figure 10-40: Global mean flux of monthly CLARA-A3 RSF and other data records

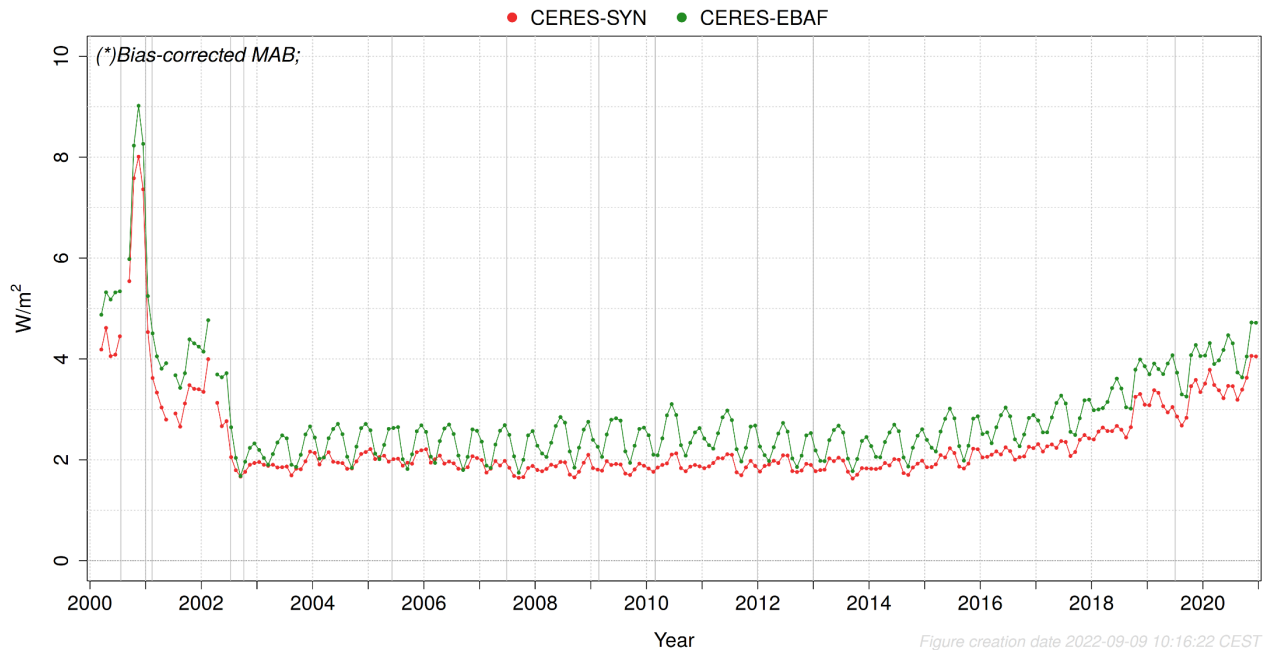**Global mean bias of monthly CLARA-A3 RSF w.r.t. other data records (*)**

Figure 10-41: Global mean bias of monthly CLARA-A3 RSF w.r.t. other data records.

10.7.2 Processing error (regional uncertainty)
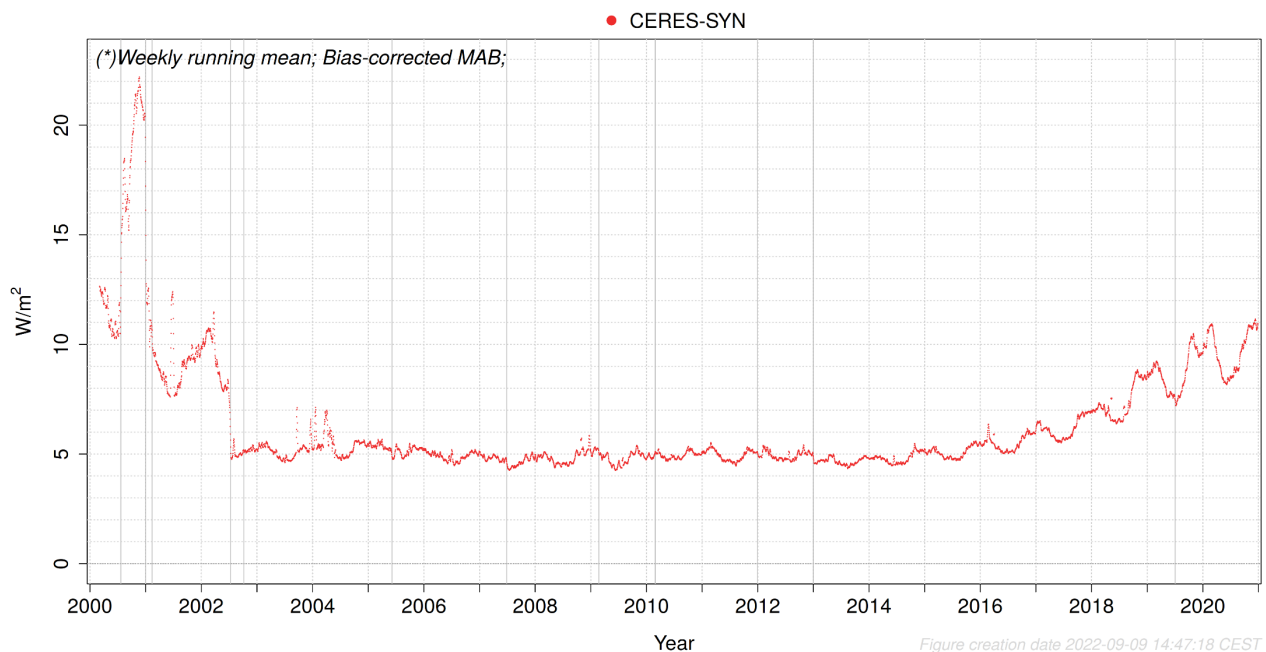
10.7.2.1 Monthly

Global MAB between monthly CLARA-A3 RSF and other data records (*)

Figure 10-42: Global MAB between monthly CLARA-A3 RSF and other data records

10.7.2.2 Daily

Global MAB between daily CLARA-A3 RSF and other data records (*)

Figure 10-43: Global MAB between daily CLARA-A3 RSF and CERES-SYN1deg-Day

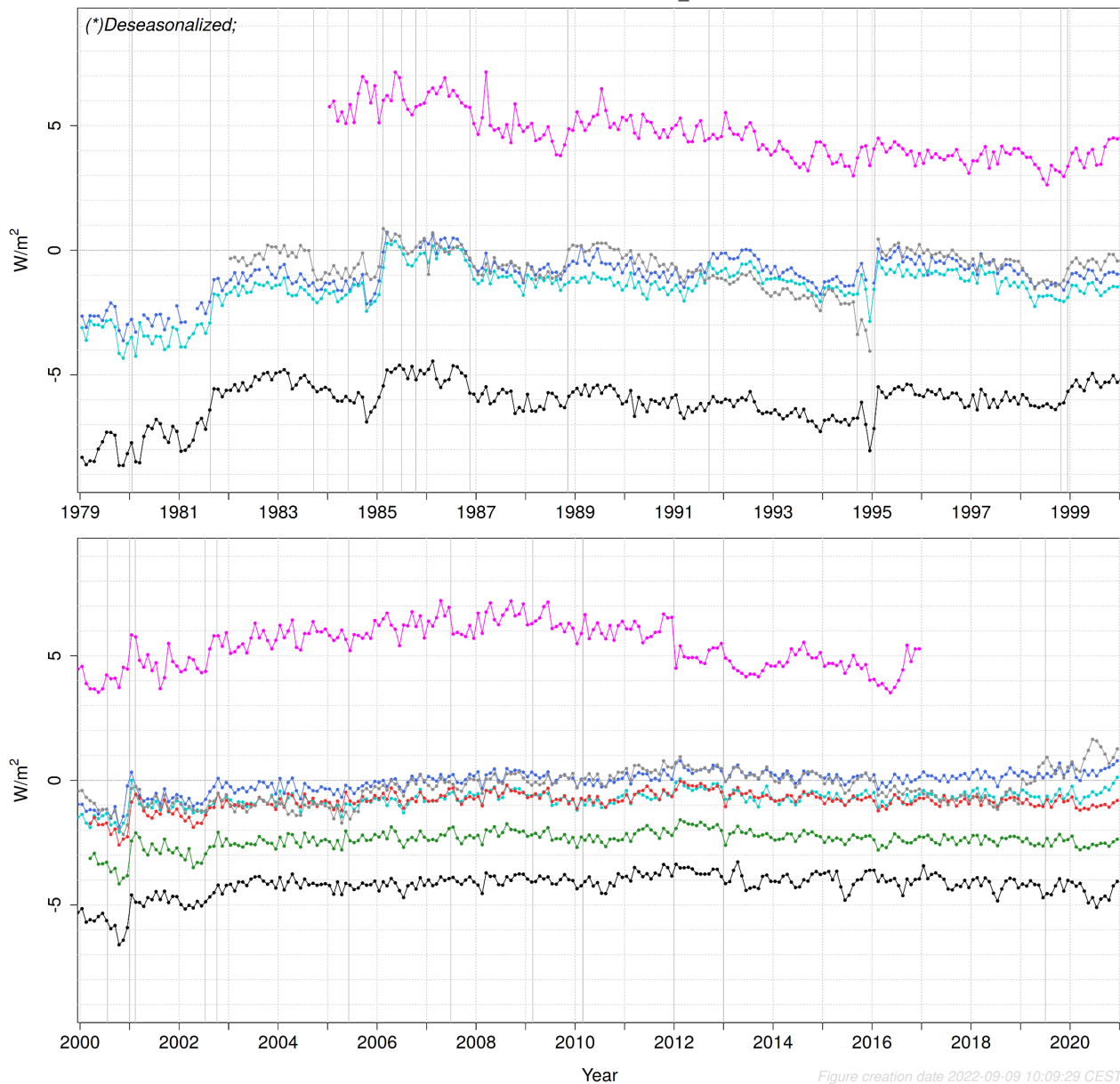## 10.8 Results for OLR without spatial gap-filling of CLARA-A3

10.8.1 Mean bias and stability
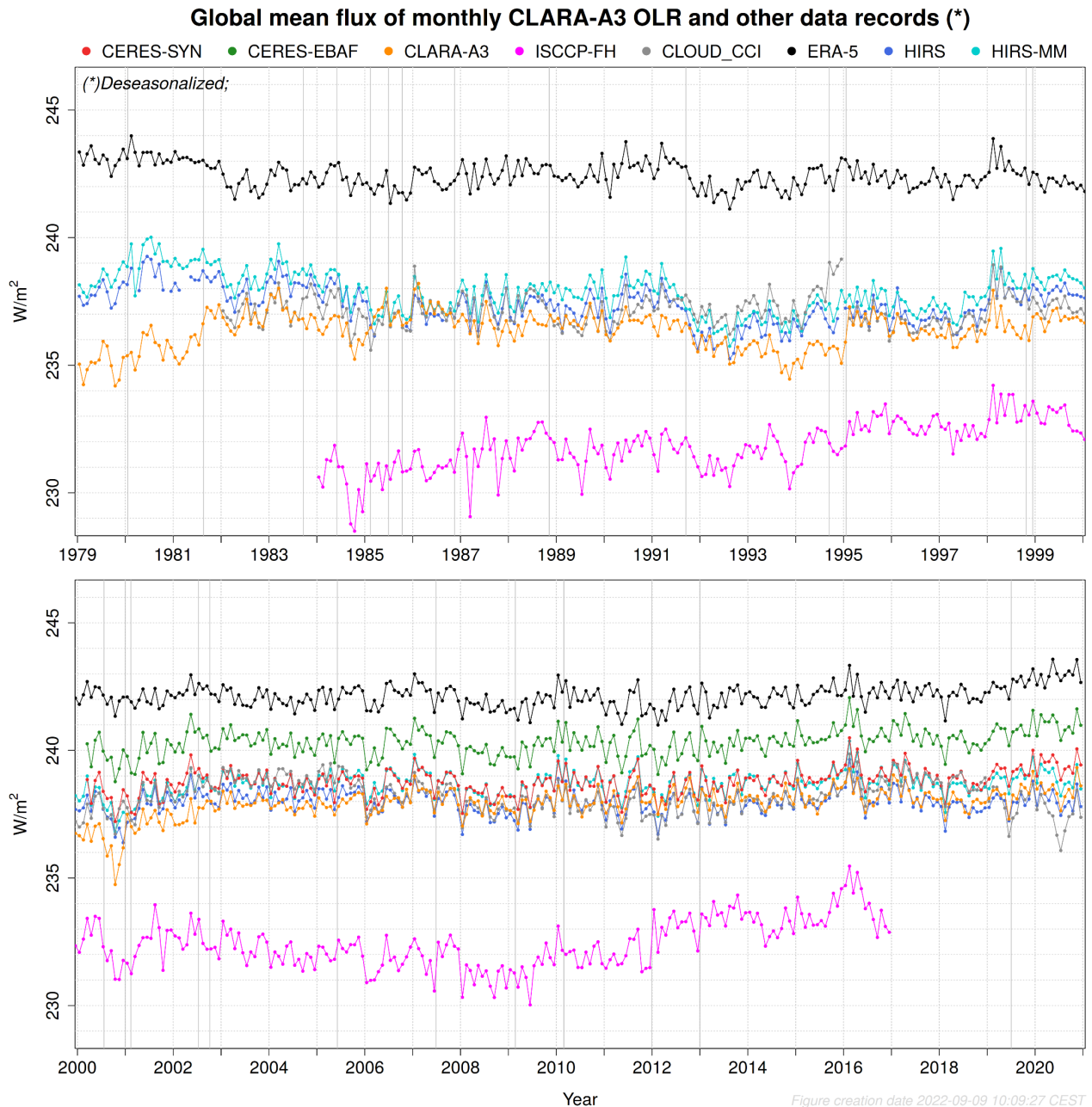
Figure 10-44: Global mean flux of monthly CLARA-A3 OLR and other data records

Global mean bias of monthly CLARA-A3 OLR w.r.t. other data records (*)

Figure 10-45: Global mean bias of monthly CLARA-A3 OLR w.r.t. other data records.

10.8.2 Processing error (regional uncertainty)

10.8.2.1 Monthly

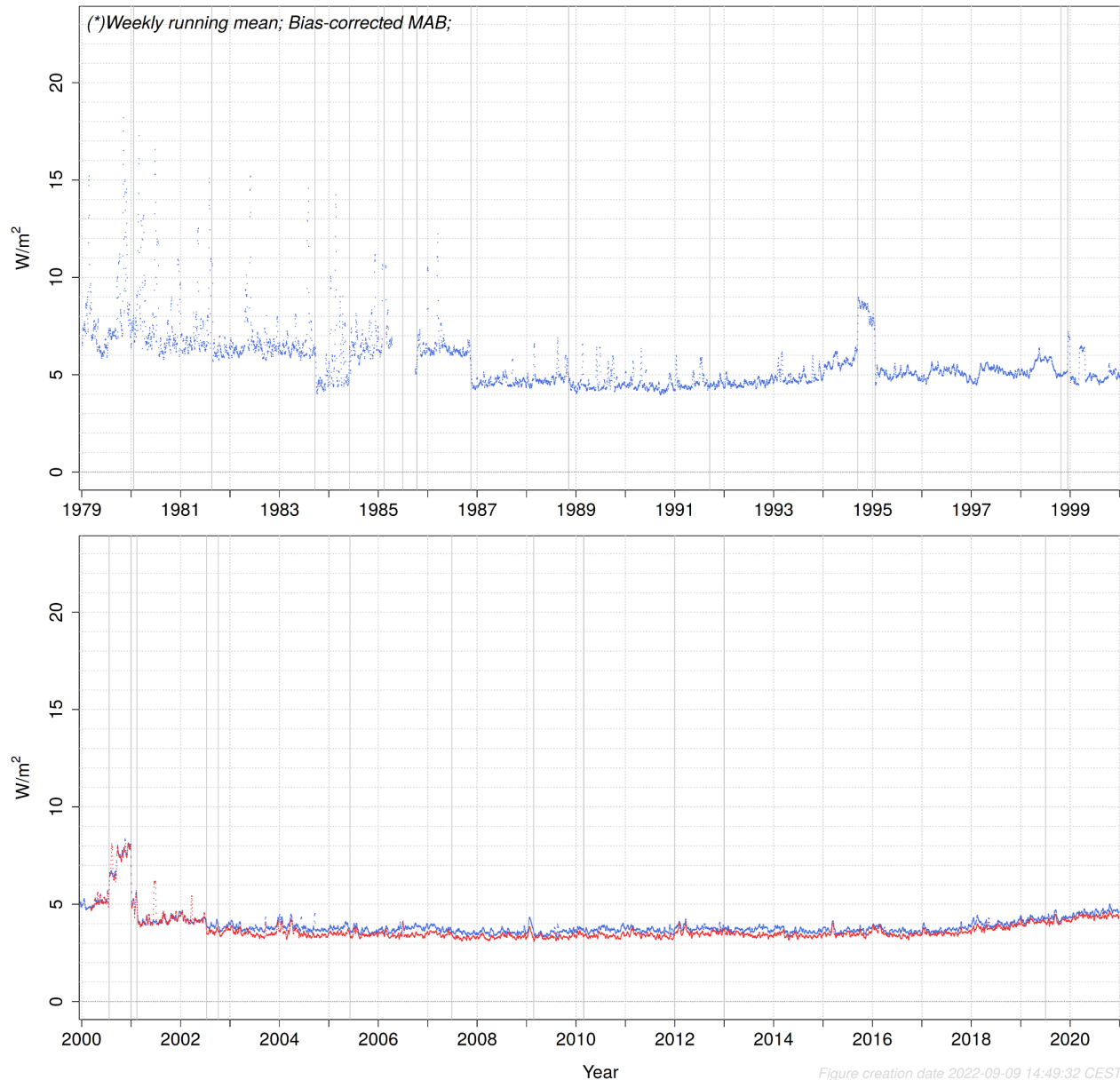**Global MAB between monthly CLARA-A3 OLR and other data records (*)**
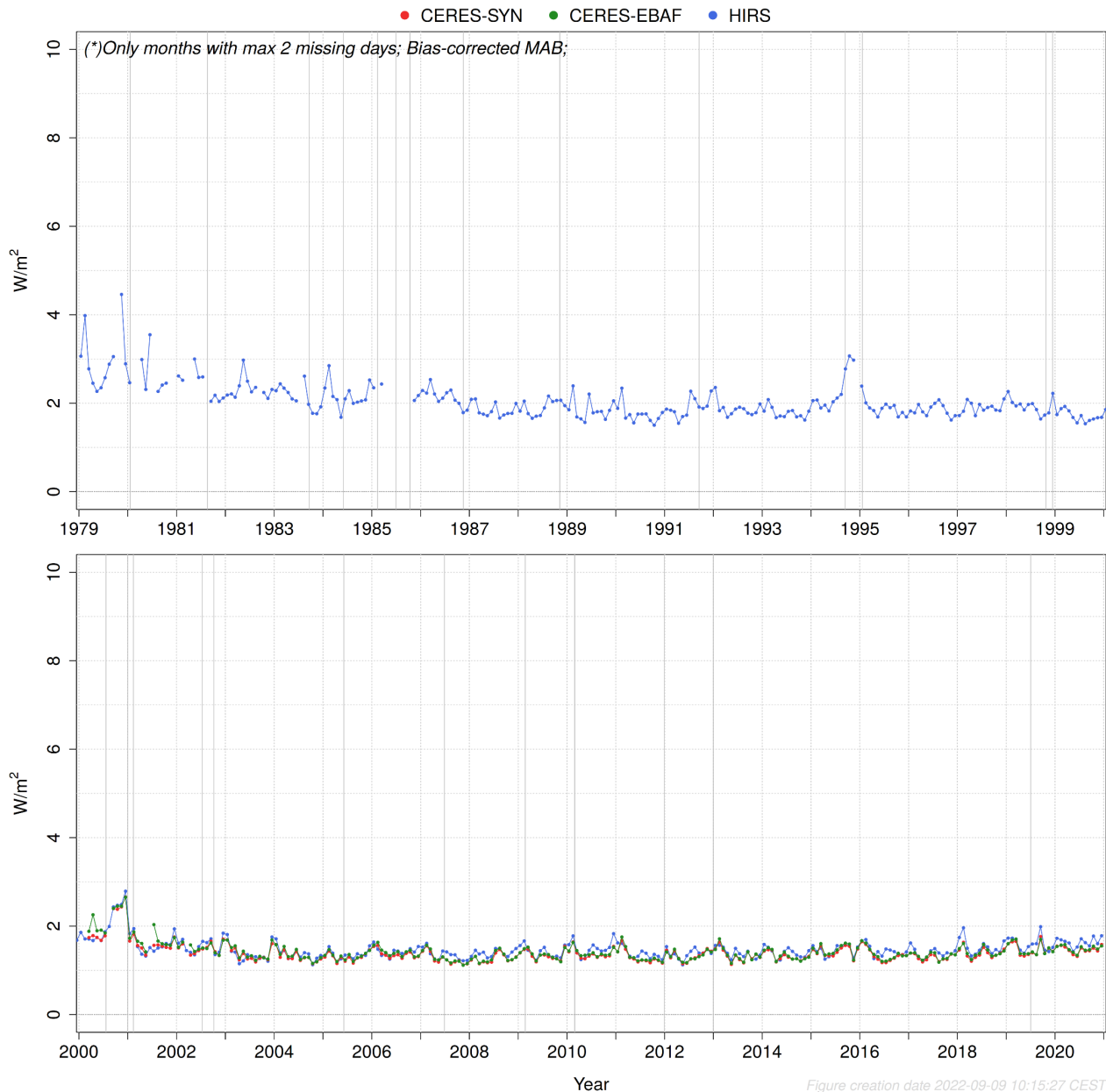
Figure 10-46: Global MAB between monthly CLARA-A3 OLR and other data records

10.8.2.2 Daily

Global MAB between daily CLARA-A3 OLR and other data records (*)
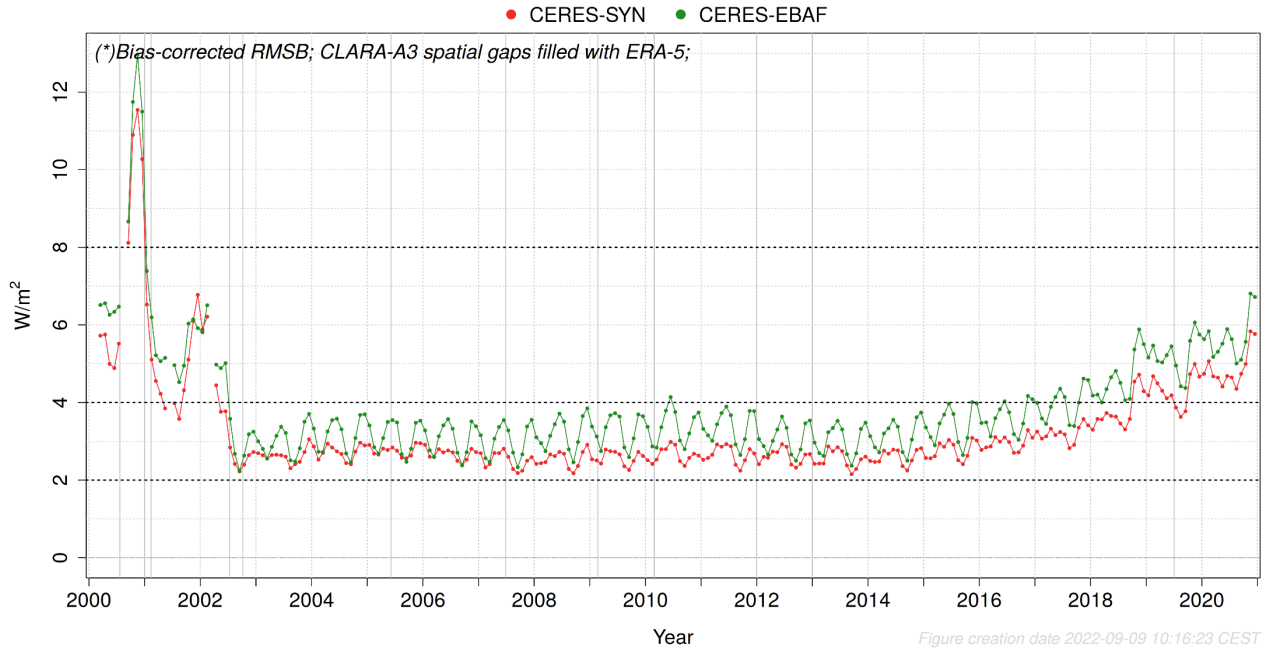

Figure 10-47: Global MAB between daily CLARA-A3 OLR and other data records

10.9 Processing error (regional uncertainty) with RMSB

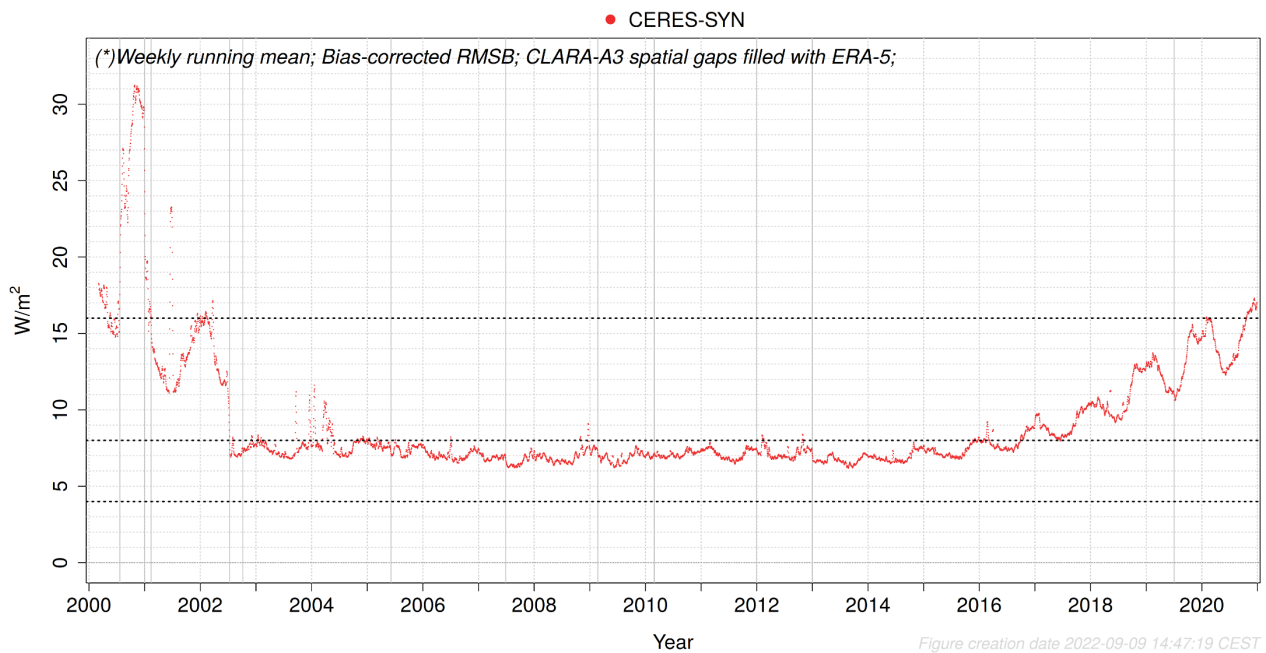### 10.9.1 RSF

Global RMS Bias between monthly CLARA-A3 RSF and other data records (*)

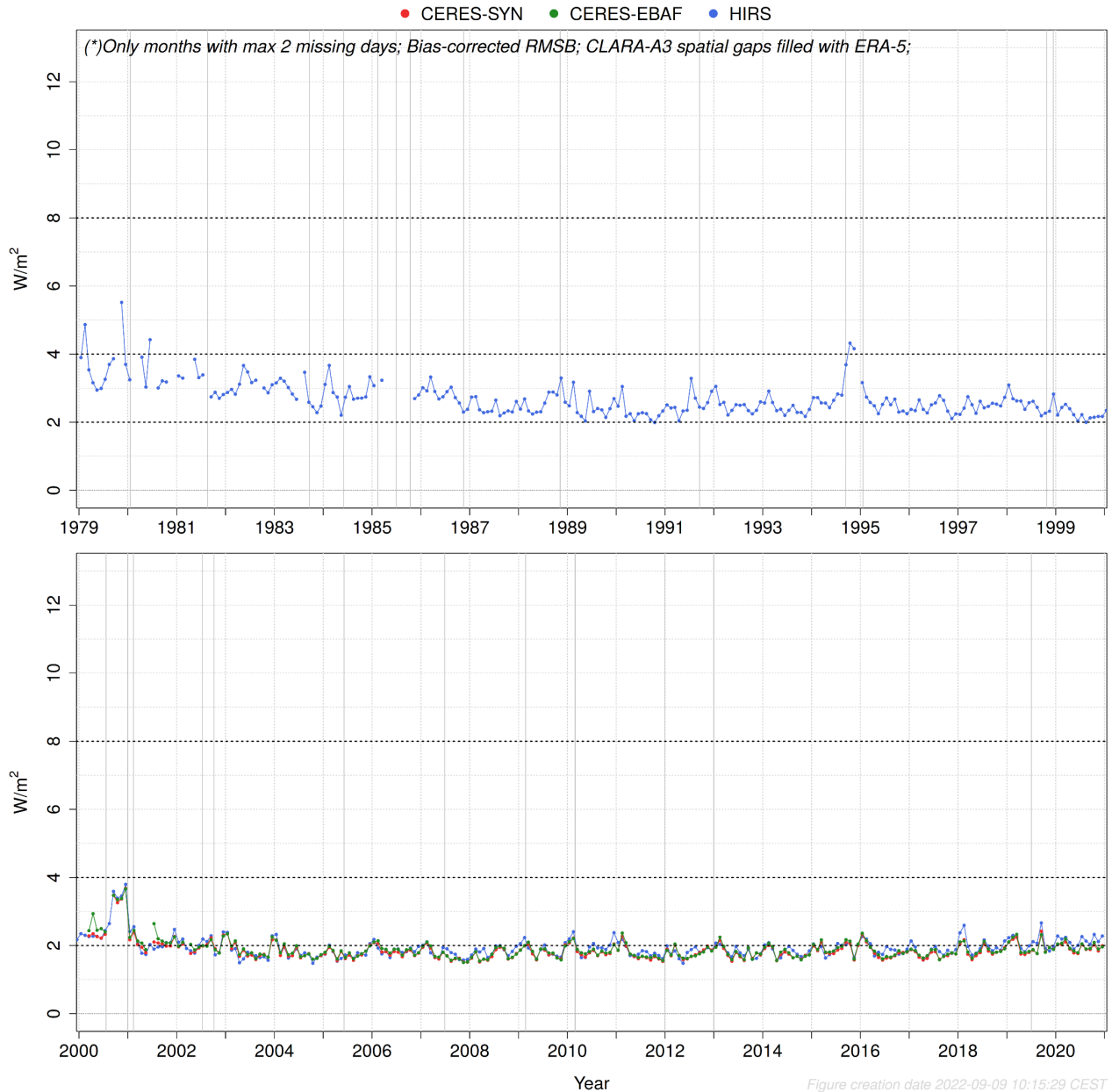**Figure 10-48:** Global RMSB between monthly CLARA-A3 RSF and other data records

Global RMS Bias between daily CLARA-A3 RSF and other data records (*)

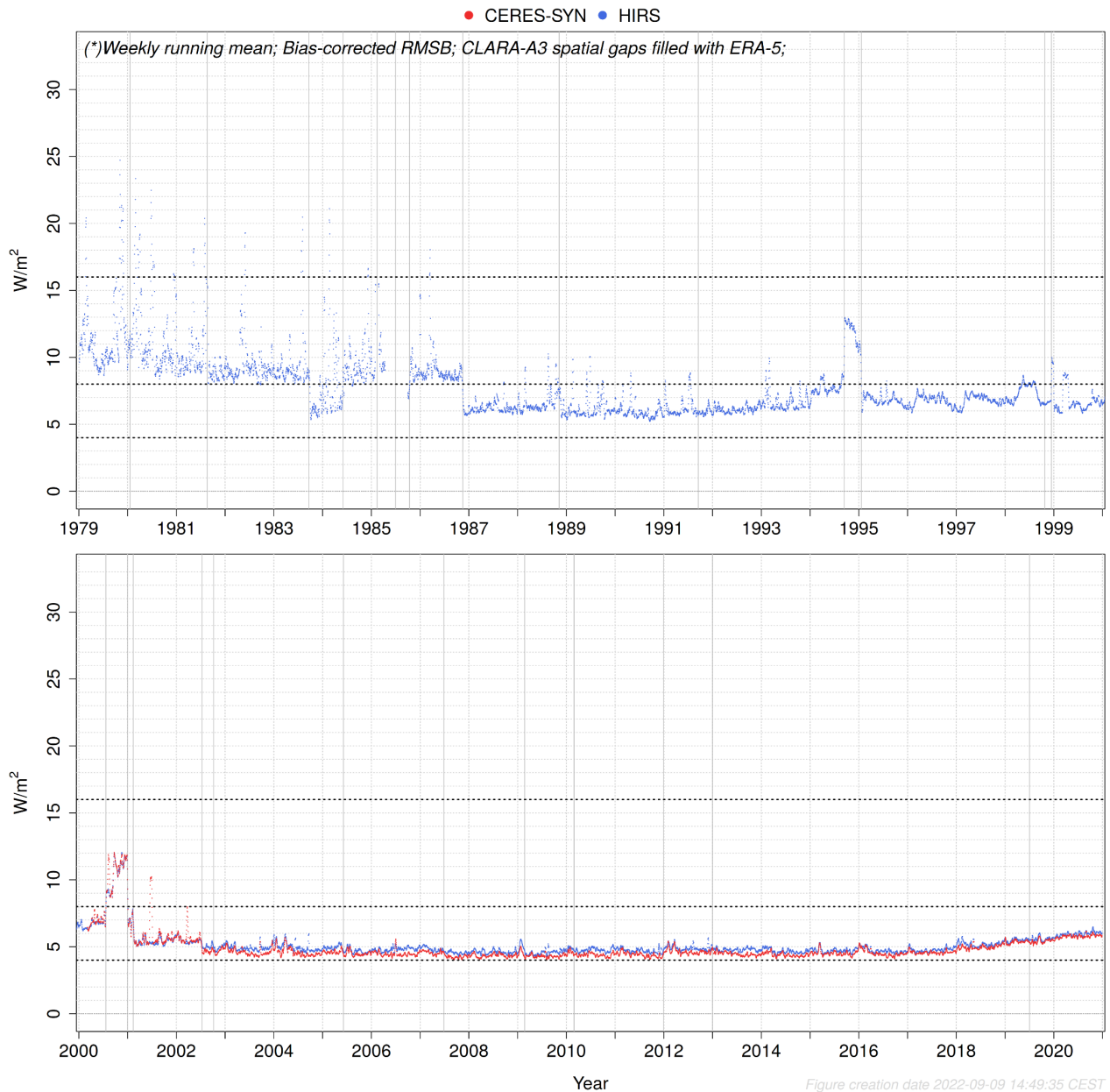**Figure 10-49:** Global RMSB between daily CLARA-A3 RSF and CERES-SYN1deg-Day

10.9.2 OLR

Global RMS Bias between monthly CLARA-A3 OLR and other data records (*)

● CERES-SYN ● CERES-EBAF ● HIRS

(*)Only months with max 2 missing days; Bias-corrected RMSB; CLARA-A3 spatial gaps filled with ERA-5;

Figure creation date 2022-09-09 10:15:29 CEST

Figure 10-50: Global RMSB between monthly CLARA-A3 OLR and other data records

Global RMS Bias between daily CLARA-A3 OLR and other data records (*)

Figure 10-51: Global RMSB between daily CLARA-A3 OLR and other data records