ERA report series

19 Pre-assimilation feedback on a Fundamental Climate Data Record of brightness temperatures from Special Sensor Microwave Imagers: A step towards MIPs4Obs?

Paul Poli[1], Carole Peubey[1], Karsten Fennig[2], Marc Schröder[2], Rob Roebeling[3], Alan Geer[1]

[1] ECMWF Research Department
[2] DWD Satellite based climate monitoring
[3] EUMETSAT Climate Services Division

Series: ERA Report Series

A full list of ECMWF Publications can be found on our web site under:
http://www.ecmwf.int/en/research/publications

European Centre for Medium Range Weather Forecasts
Shinfield Park, Reading, Berkshire RG2 9AX, England

Abstract

Responding to needs from the user community, the Climate Monitoring Satellite Application Facility (CM SAF) of EUMETSAT has produced a Fundamental Climate Data Record (FCDR) from brightness temperatures collected by the series of Special Sensor Microwave Imager (SSM/I) instruments. This FCDR is considered for use in future global reanalyses, and in particular ERA5, due to replace ERA-Interim. This report presents a first account of the information gained from the reanalysis user side, albeit before assimilation, using clear-sky radiative transfer computations. Earlier reanalyses (ERA-Interim and ERA-20C) are projected into observation space after collocation to the observation times and locations.

The assessment shows that the FCDR patches gaps in the SSM/I record at ECMWF and adds more than 13 satellite-years. A methodology to compare data records in details confirms the introduction of time offsets in the FCDR, consistently with leap seconds that were missing in the original record. The ancillary sea-ice information in the FCDR appears also of superior quality than in the original record. The ERA-20C reanalysis appears better than ERA-Interim at reproducing the inter-annual variability of the FCDR, but misses on the intra-monthly variability. The FCDR features a time consistency that is superior to the original record, thanks to the intercalibration. However, this intercalibration filters out some intra-month variability within the original data. Further conclusions about the data quality of the FCDR brightness temperatures will be best formed with assimilation experiments in reanalysis.

The comparison process generated a data record augmented by pre-assimilation feedback. This followed approach of augmenting an observational dataset with model-equivalents could in fact be generalized, and adopted to project not only reanalysis output but also climate model output into observational data space. This would complement the Obs4MIPs initiative, where each dataset matches one (or several) field(s) generated by a model integration such as from CMIP5. We propose to complement this alignment of observational datasets to climate model output datasets by a reverse procedure, *i.e.*, a MIPs4Obs concept, where one would match each model (or reanalysis) output dataset to observational data records. Where Obs4MIPs is from the onset designed to bring observations to the modellers, in the space and shape they are used to, MIPs4Obs would bring the model outputs to the community of observers, retaining the granularity of the original observation records. The interpretation of the "observation versus model" comparison results would be simplified, because the spatio-temporal coverage and sampling of the data considered would be *de facto* identical. Overall, such a pre-assimilation feedback approach could help assess and quality control Climate Data Records.

1 Introduction

One of the largest limitations of the ERA-Interim reanalysis (Dee *et al.*, 2011) is a set of discontinuities in the representation of the water cycle. These are caused by issues in the assimilation methodology employed then, causing jumps at each appearance or disappearance of each SSM/I sensor in the assimilation (Geer *et al.*, 2008).

Although the original SSM/I data record used in ERA-Interim has not been specifically questioned for its quality, it is known to be incomplete for some years and not using today's state-of-the-art calibration. This baseline data record, hereafter referred to as the **legacy data record (LDR)**, has two origins:

- Until mid-1999, the data come from Remote Sensing Systems (RSS, version 5, acquired for ERA-40); it includes inter-satellite calibration ('intercalibration' in the remainder of this report, for conciseness) that was state-of-the-art at the time. This data record includes only one satellite at a time, and
- Thereafter, the data are as acquired in Near-Real-Time (NRT) by ECMWF operations from the Global Telecommunications System (GTS), without intercalibration, but including usually several satellites.

The CM SAF has recovered as much as possible the original SSM/I data to construct a **Fundamental Climate Data Record** (Fennig *et al.*, 2013) spanning all the six satellites that carried SSM/I instruments. This FCDR is likely of greater length and fewer gaps than the LDR, or than that found in any other single global Numerical Weather Prediction (NWP) center's archive, because the FCDR includes data collected outside of the expected cut-off times allowed for NRT transmission.

Furthermore, characteristics of the SSM/I instruments are now better known than when these instruments were processed for NWP applications, so it has been possible to benefit from scientific advances in the creation of a FCDR. For example, the FCDR created by CM SAF uses a fixed set of antenna pattern corrections (CM SAF, 2013a), resulting in improved consistency across platforms than using sensor-dependent corrections. Both arguments suggest using such FCDR in future reanalyses, instead of the LDR.

Besides the CM SAF FCDR, two other state-of-the-art SSM/I FCDRs are now available, one from RSS (Wentz, 2013), the other from Colorado State University (CSU: Sapiano *et al.*, 2013). Curiosity dictates that one would look at all three of them, but under limited time and resources this report focuses on the CM SAF FCDR, while also revisiting the LDR used in ERA-Interim. A methodology on how to compare data records is outlined in this report, thereby proposing that future similar exercises can be carried out outside the complexity of a reanalysis system (ERA relies on the Integrated Forecasting System (IFS: ECMWF, 2007)).

The report is organized as follows. Section 2 explains how the FCDR is imported into the ECWMF Observation DataBase (ODB). Section 3 presents a methodology to compare the FCDR to reanalysis output, at the observation time and location, for clear-sky scenes. Section 4 proposes a method to compare the LDR and the FCDR. Section 5 validates the simulation tool by comparing simulations offline with the results of simulations carried out by the ERA-Interim reanalysis system, found in the ERA-Interim observation feedback archive. Section 6 discusses the results obtained by comparing simulations and the LDR and FCDR, the lessons learnt from the comparison with the LDR assimilated in ERA-Interim (considering here only the clear-sky assimilation), implications about future assimilation, and potential leads for improvement in a future iteration of the FCDR. Section 7 discusses the benefits of the approach followed here, namely augmenting an observational data record with reanalysis-equivalents, and how this approach could be generalized to systematically assess long-term data records from observations in a way that would complement Obs4MIPs. Section 8 draws conclusions.

2 Data sources

One source of data is the ERA-Interim observation feedback, which is being prepared for public release at time of writing. The observation feedback contains not only the SSM/I LDR observations (brightness temperatures), but also the differences observations minus ERA-Interim background (often abbreviated **O-B**, where **O** refers to observations and **B** refers to background), and the differences observations minus ERA-Interim analysis (often abbreviated **O-A**, where **A** refers to analysis). The feedback also includes in addition the variational bias correction estimates and data assimilation quality control flags, indicating whether each datum was assimilated or not, and reason(s) for rejection, when relevant.

A second source of data is the FCDR acquired from the CM SAF website, referenced by a digital object identifier (DOI:10.5676/EUM_SAF_CM/FCDR_SSMI/V001), retrieved from

http://dx.doi.org/10.5676/EUM_SAF_CM/FCDR_SSMI/V001

These observation data are converted from native NetCDF4 to the format used by the Observational Data Base Application Programming Interface (ODB API), using a Fortran program.

Each FCDR data file covers a full calendar day of data. During this conversion the data are assigned to the closest synoptic hour (0, 6, 12, or 18 UTC).

The following information is extracted from the NetCDF4 files into ODB files. The names of NetCDF elements are shown in italics in Table 1, between single quotes (‘ ’). In the conversion process, a quality control procedure flags as missing any observations at data locations for which any of the following is missing: latitude, longitude, zenith angle, azimuth, surface type, and spacecraft altitude. Similarly, data which have any of the following quality control indicators raised (non-zero) are flagged as missing: *‘qc_scan’*, *‘qc_channel’*, or *‘qc_fov_lo’*. Additionally, scan lines that only contain missing values (found with *‘qc_scan’* raised) are removed.

A particular attention is given to keep traceability to the original data, so the results can then be mapped back into the original data record to augment it with the reanalysis-added information. For each datum, only two records are required: identification of the original NetCDF file, and identification of the particular datum number within the NetCDF file.

ODB column name	Contents	Unit or format	Origin	Range
expver@desc	MARS attribute			NULL
type@desc	MARS attribute			NULL
class@desc	MARS attribute			NULL
stream@desc	MARS attribute			NULL
andate@desc	MARS attribute	YYYYMMDD	Calculated from *'time'*	19870708-20090101
antime@desc	MARS attribute	HH	Calculated from *'time'*	0, 6, 12, or 18
reportype@hdr	MARS attribute		Determined from satellite platform number	10002, 10004-10008
groupid@hdr	MARS attribute			11
source@hdr	Traceability attribute			'CMSAF001'

ODB column name	Contents	Unit or format	Origin	Range
collection_identifier@hdr	Traceability attribute	integer: NetCDF file date*100+ DMSP satellite number	Date and DMSP numbers indicated in the NetCDF filename	1987070808 - 2009010115
bufrtype@hdr	IFS attribute, for future assimilation			3
subtype@hdr	IFS attribute, for future assimilation			55
obstype@hdr	IFS attribute, for future assimilation			7
codetype@hdr	IFS attribute, for future assimilation			210
sensor@hdr	RTTOV attribute			6
satellite_identifier@sat	WMO identifier, common table C-5		*'wmo_satellite_ identifier'*	241, 243, 244, 246, 247, 248
satellite_instrument@sat	WMO identifier, common table C-8			905
seqno@hdr	ODB record number	Integer	Counter	>=1
date@hdr	Observation date	YYYYMMDD	Calculated from *'time'*	19870709-20081231
time@hdr	Observation time	HHMMSS	Calculated from *'time'*	0-235959
lat@hdr	Observation latitude	degreesNorth	*'lat'*	-87.8 to 87.7 for F8 -88.1 to 87.8 for F10 -87.8 to 87.6 for F11 -87.8 to 87.5 for F13 -88.0 to 87.8 for F14 -88.0 to 87.7 for F15
lon@hdr	Observation longitude	degreesEast	*'lon'*	-180.0 to 180.0 (exactly)
stalt@hdr	Spacecraft altitude	km	*'salt'*	834 to 889 for F8 728 to 876 for F10 837 to 886 for F11 839 to 884 for F13 837 to 884 for F14 833 to 881 for F15
zenith@sat	Satellite zenith angle	degrees	*'eia'*	52.96 to 53.68 for F8 52.07 to 53.81 for F10 53.07 to 53.59 for F11 52.64 to 53.29 for F13 52.80 to 53.42 for F14 52.68 to 53.19 for F15
azimuth@sat	Satellite azimuth angle	degrees	*'laz'*	-180 to 180 (exactly)

ODB column name	Contents	Unit or format	Origin	Range
surface_class@obssurf	Type of surface		*'sft_lo'*	0 for water mapped into 5 for IFS, 1 for land mapped into 0 for IFS, 2 for coast mapped into 6 for IFS, 11 for sea-ice mapped into 3 for IFS, 12 for sea-ice edge mapped into 4 for IFS
orbit@radiance	Orbit number			
scanline@radiance	Scan line number		Same as seqno	>=1
scanpos@radiance	Scan position number			1 to 64, '*nAcrossTrackLoRes*'
entryno@body	ODB entry number within record	integer	counter	1 to 7
unique_identifier@body	Traceability attribute	integer	record number within the original NetCDF file, *iscanposition* + *ichannel*nAcrossTrack* + *itime*nAcrossTrack*nchannel*	Lowest value cannot be below. Highest value depends on *ntime*.
vertco_reference_1@body	Channel number	integer	*'channel'*	1 to 7
vertco_type@body	IFS attribute			3
varno@body	IFS attribute			119
obsvalue@body	Observed brightness temperature	K	*'tb'*	See Table 2
obsvalue_cal@body	Intercalibrated brightness temperature	K	*'tb'* plus *'ical'*	See Table 2
rainy@qc	Rain contamination flag		*Computed using observed BTs from channels 3, 4, and 5*	Clear-sky if 0

Table 1: List of parameters read or computed from the CM SAF FCDR NetCDF4 files and written into ODB files

Table 1 indicates that longitudes can be exactly -180.0 or +180.0. Yet, these two values point to the same meridian on the surface of the Earth. The reason for this is the rounding applied after the range check. This suggests that modifications be applied to future SSM/I releases as well as other CDRs currently being developed at the CM SAF, so that longitudes are uniquely defined. Additionally, the maximum values of spacecraft altitudes found for F8, F10, F11, and F15 are larger than reported in the ATBD Appendix A Table 1. This discrepancy is simply caused by the ATBD giving a broad range and not exactly reflecting the full contents of the data files.

The column '*rainy@qc'* is added to discriminate pixels believed to be contaminated by rain (in which case '*rainy@qc*' differs from 0). The rain detection algorithm is as in the IFS. It considers observed brightness temperatures from channels 3 (22 GHz vertically polarized, noted T^{b}_{3}), 4 (37 GHz vertically polarized, noted

T^b_4) and 5 (37 GHz horizontally polarized, noted T^b_5). A scene is identified as clear ('*rainy@qc=0*') unless any of the following conditions is met:

- T^b_3, T^b_4, or T^b_5 is outside reasonable bounds, namely either below 50 K or above 350 K
- the difference (T^b_4-T^b_5) is below 40 K
- over sea, when T^b_3 and T^b_4 are both below 280 K, the liquid water path, estimated using the formula [4.2993 + 0.399635 *log* (280-T^b_3) - 1.406920 *log* (280-T^b_4)], exceeds 0.01

min/ max	F8 w/o inter-calib.	F8	F10 w/o inter-calib.	F10	F11 w/o inter-calib.	F11	F13 w/o inter-calib.	F13	F14 w/o inter-calib.	F14	F15 w/o inter-calib.	F15
chan. 1 19V	131/ 310	132/ 309	130/ 310	130/ 309	131/ 310	131/ 310	134/ 310	134/ 311	135/ 310	135/ 311	132/ 310	133/ 310
chan. 2 19H	81/ 299	81/ 298	80/ 300	80/ 300	80/ 300	80/ 300	80/ 300	80/ 300	88/ 300	88/ 300	81/ 300	81/ 300
chan. 3 22V	135/ 310	137/ 309	131/ 310	131/ 310	130/ 309	130/ 309	131/ 310	131/ 311	133/ 309	133/ 310	134/ 309	134/ 309
chan. 4 37V	132/ 310	134/ 309	133/ 310	133/ 309	134/ 310	134/ 310	132/ 310	132/ 310	133/ 310	132/ 309	134/ 310	134/ 310
chan. 5 37H	110/ 296	111/ 296	112/ 300	112/ 300	112/ 300	111/ 300	112/ 300	111/ 300	111/ 300	111/ 300	111/ 300	111/ 300
chan. 6 85V	134/ 307	135/ 308	132/ 311	133/ 313	133/ 310	132/ 310	132/ 310	132/ 311	132/ 310	132/ 311	133/ 310	133/ 311
chan. 7 85H	87/ 296	88/ 296	122/ 302	122/ 303	122/ 305	122/ 300	122/ 300	122/ 302	122/ 300	122/ 301	121/ 301	122/ 303

Table 2: Minima and maxima of brightness temperatures, sampling 1/15th of the CM SAF SSM/I FCDR (blue columns, without application of the intercalibration). Note that all values appear within physical range, because unphysical value have effectively been removed by application of the FCDR quality control flags

Table 2 indicates the minima and maxima found by inspecting the FCDR every 15th day. These values are not robust metrics, as would be the first or last percentile. However, percentile tests would not necessarily catch extreme, seldom, outliers, which is the purpose of this table: to identify if any unphysical value passed the various quality controls. The results indicate that these various quality controls work satisfactorily. A first analysis of the FCDR in the form of Table 2 had enabled to identify a number of issues, such as a few negative brightness temperatures and apparently incorrect data from F8 channels 6 and 7. After information exchange and a code review, it was found that some quality controls in the conversion step from NetCDF4 to ODB had not been propagated (this is now rectified). In particular, a quality flag characterizing the synthetic data for F8 channels 6 and 7 was applied incorrectly. As they stood, these erroneous data were easy to identify and would have been flagged for removal in a possible assimilation. Although, such a detailed interaction, although it may only occur rarely between a data provider and its users, proved to be extremely beneficial.

Table 2 suggests that the effect of the intercalibration is rather small. This is confirmed by Figure 1, which shows that most monthly estimates of intercalibration are within 1 K.

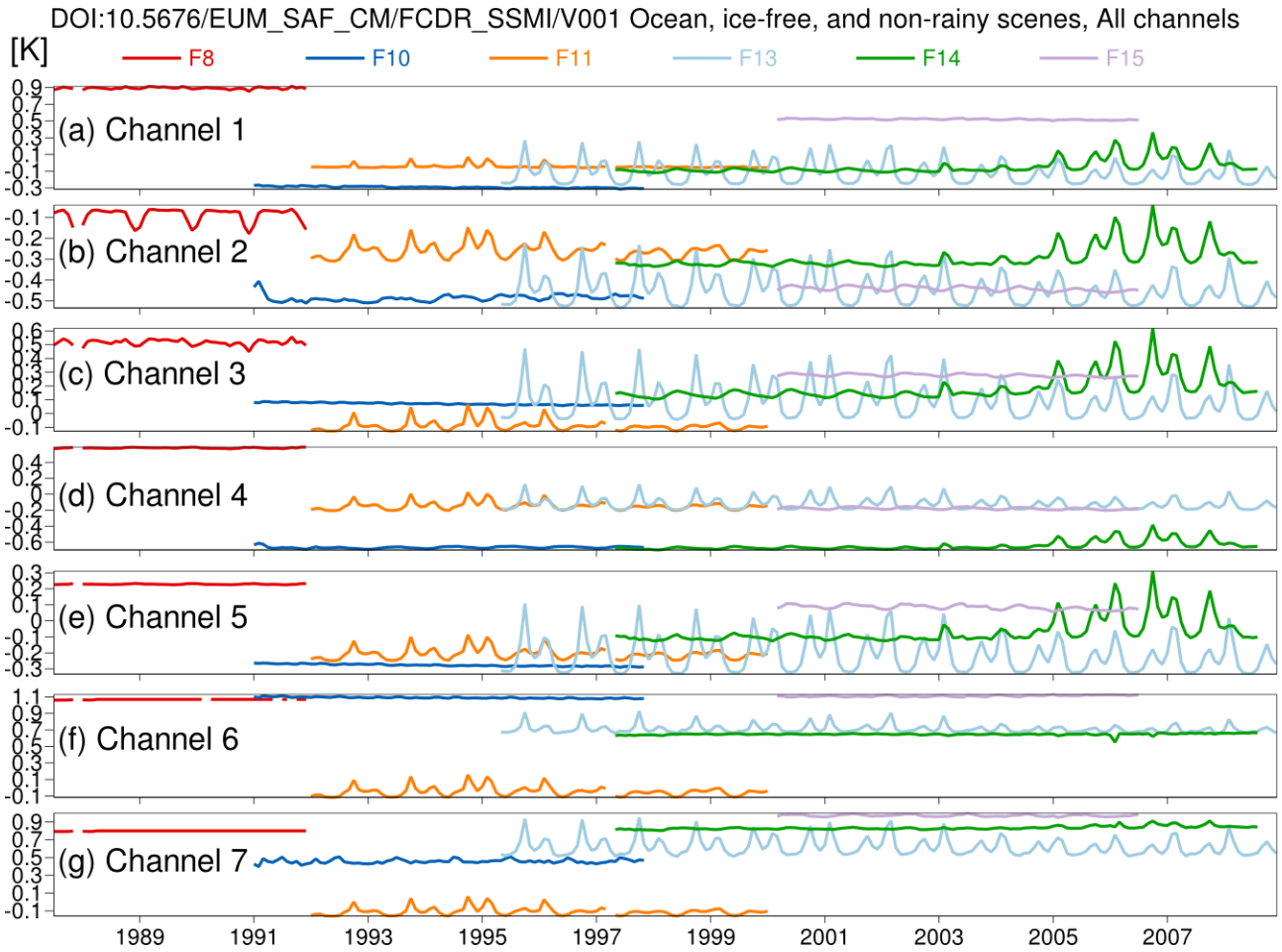

Figure 1: Monthly average intercalibration, sampling 1/15th of the CM SAF SSM/I FCDR

3 Methodology to compare FCDR with gridded reanalysis datasets

In this report we compare the CM SAF SSM/I FCDR with two reanalyses. The first reanalysis, ERA-Interim (Dee *et al.*, 2011), assimilated SSM/I brightness temperature observations, as explained earlier. Using this reanalysis enables to assess the discontinuities in the representation of the water cycle in ERA-Interim by comparison to a state-of-the-art FCDR that tracks water vapor.

The second reanalysis, ERA-20C, didn't assimilate SSM/I observations. The ERA-20C dataset resembles very much that generated by the ERA-20C ensemble production described by Poli *et al.* (2013). There are minor differences, addressing in particular the issues listed in the table presented in the conclusion section of the aforementioned reference. Using this reanalysis enables to carry out an independent comparison to the FCDR.

Both reanalysis datasets are available for download from the ECMWF public data server (http://apps.ecmwf.int/datasets/). All these data are retrieved from ECMWF MARS archive, in the form of GRIB files, with horizontal reduced Gaussian grids, from the analysis fields. The following parameters are used:

- from all model level fields: temperature and humidity
- from the surface fields: surface pressure, land-sea mask, skin temperature, 10-metre zonal and meridional wind components, surface geopotential, 2-metre temperature, 2-metre dewpoint temperature, and sea-ice cover.

The following information, collocated from the reanalysis fields to the locations and times of the observations, is then added to each observational record entry:

- surface skin temperature
- two-metre temperature
- ten-metre zonal wind
- ten-metre meridional wind
- sea-ice fraction
- surface elevation
- land-sea mask

This additional information is useful to discriminate the data, for example by scene type (such as over sea-ice). The ERA-Interim analysis fields are available every 6 hours; the ERA-20C analysis fields are available every 3 hours. The closest neighbor in time is used, there is no time interpolation, and all analysis field parameters are interpolated with a bilinear horizontal interpolation to the observation location.

The EUMETSAT NWP-SAF fast radiative transfer model RTTOV version 11 (Saunders *et al.*, 2011) is used to compute, at each observation pixel location and time, the brightness temperature estimates from the three-dimensional information found in the reanalysis GRIB output.

The program that runs these interpolations followed by calls to RTTOV11 expects the following arguments:

- ***-i*** input ODB file name
- ***-o*** output ODB file name
- ***-n*** number of model levels (60 for ERA-Interim, 91 for ERA-20C)
- ***-g*** path where input GRIB files are located
- ***-a*** time resolution of analysis fields (6 for ERA-Interim, 3 for ERA-20C)

- ***-f*** time resolution of forecast fields, and ***-d*** analysis hour offset from 00 UTC for forecast initial times (both pieces of information are useful if comparing with forecasts, but this functionality is not used here)
- ***-t*** **<FB>** name of the ODB table to append to column names, for example '*@ei*' for labelling all added columns as containing information from ERA-Interim, or '*@e2oper*' for labelling all added columns as containing information from the ERA-20C deterministic production
- ***-s*** WMO satellite identifier (this number is verified against the data found in the file, but is needed in order to determine the radiative transfer coefficient file to be loaded)
- ***-e*** **1** if the emissivity found in the ODB is to be used (this option was only used for validation purposes in section 5, to verify matching computations with the ERA-Interim LDR feedback, which contains the emissivity computed by ERA-Interim), otherwise the following two options are used:
- ***-v*** version of the FASTEM model to use for emissivity computations (version 5 by default), and
- ***-y*** **<EMIS>** name of the column which contains the emissivity calculated by RTTOV.

The parameters listed in Table 3 are then added to the ODB records.

ODB column name	Contents
orography**<FB>**	Surface elevation (in meters)
lsm**<FB>**	Land-sea mask (from 0.0 for sea-only to 1.0 for land-only)
tsfc**<FB>**	Skin temperature (in K)
t2m**<FB>**	Two-metre temperature (in K)
seaice**<FB>**	Sea-ice cover (0. no ice, 1. fully covered by sea-ice)
u10m**<FB>**	Ten-metre zonal wind (in m/s)
v10m**<FB>**	Ten-metre meridional wind (in m/s)
<EMIS><FB>	Emissivity computed by RTTOV
fgvalue**<FB>**	Brightness temperature calculated by RTTOV (in K)
fg_depar**<FB>**	Difference observation minus the brightness temperature calculated by RTTOV (in K)
fg_depar_cal**<FB>**	Difference intercalibrated observation minus the brightness temperature calculated by RTTOV (in K)

Table 3: List of added feedback columns

The processing suite follows the data flow shown by the flowchart in Figure 2. This flowchart omits that the FCDR is compared to several gridded datasets. The last step, feeding-back the reanalysis-added values into the native format of the FCDR (NetCDF in the present case), isn't implemented yet, but would help data records producers to collect and assemble feedback from various users (this prospect is further discussed in section 8). Except for this step, the implementing processing suite is shown in Figure 3. This suite, run here for only 1/15th of the SSM/I FCDR (considering only 1 day of data every 15), takes a couple of days on the ECMWF Linux-based computing cluster.

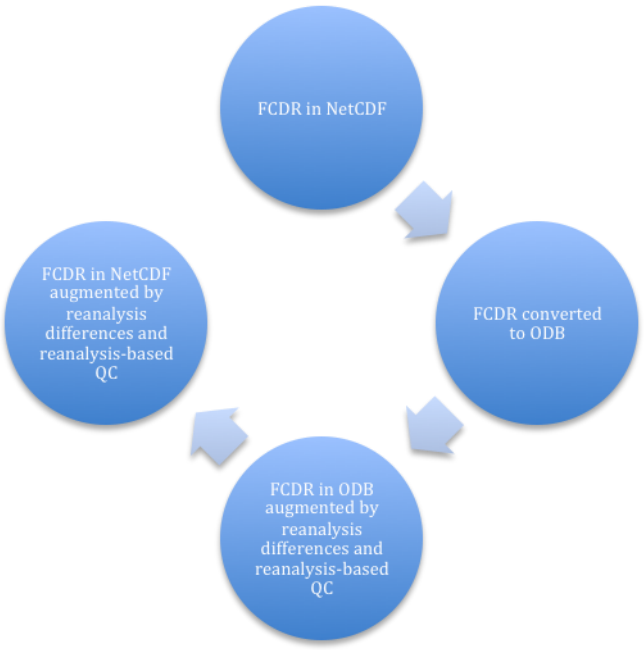

Figure 2: Data flowchart of the processing suite that compares the FCDR with reanalysis

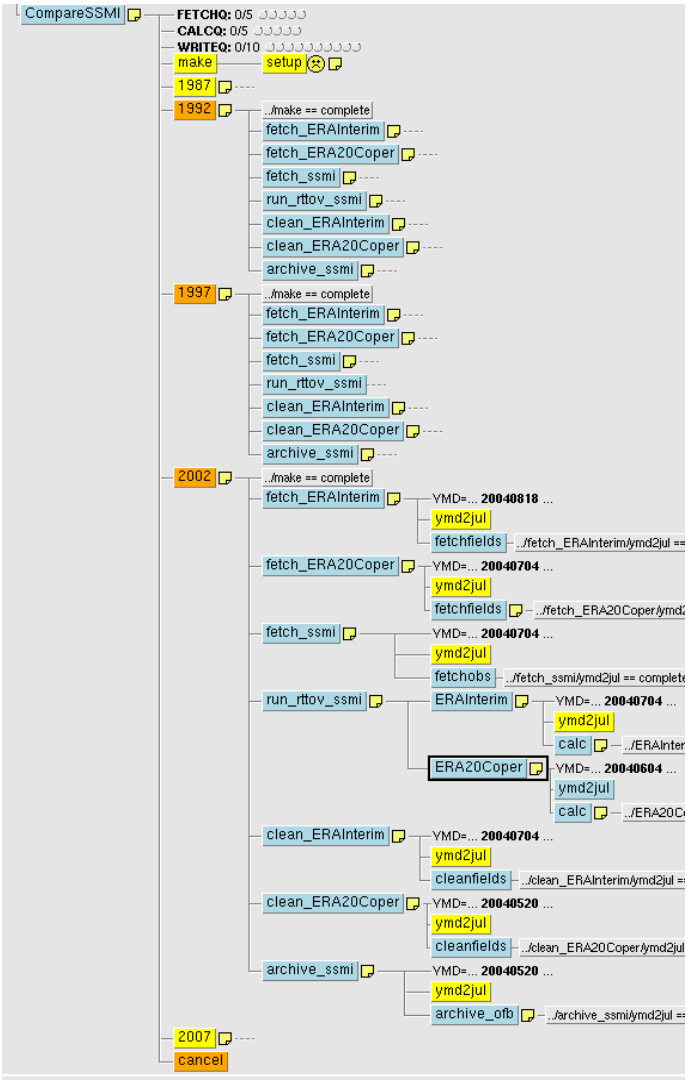

Figure 3: Snapshot of the simulator suite, here run in parallel 5-year batches (1987-1991, 1992-1994 etc…). Each batch sequentially acquires ERA-Interim and ERA-20C fields in GRIB format, SSM/I FCDR observations in NetCDF4 format, converts to ODB, runs the interpolation and RTTOV11 from each reanalysis at the observations locations. After cleaning-up the GRIB files, the suite archives the ODB files, augmented with pre-assimilation feedback, to the Observation Feedback Archive (OFA).

At the last step ("archive_ofb"), the ODB data augmented by model equivalents are archived on temporary MARS storage (MARSSCRATCH), which has a retaining time sufficient to allow for investigations. For the present work, the data were archived under class=e2, expver=1900, type=ofb, stream=oper. This facility (Observation Feedback Archive, OFA) was developed with help of the EU FP7 ERA-CLIM project (http://www.era-clim.eu).

The ODB files thus produced can be exploited with the ODB API software (https://software.ecmwf.int/wiki/display/ODB/ODB+API). This library supports SQL queries including search, sort, and aggregate statistical functions. These functions are used extensively in the present report, for example to compute statistics, to discriminate between channels, to isolate sea versus land ('*lsm@ei<0.01'* for ocean-only according to the ERA-Interim land-sea mask), to select specific latitude bands, to retain only regions with sea-ice concentration below 1% ('*ice@ei<0.01*' according to the ERA-Interim sea-ice), or to retain ocean regions believed to be free of rain contamination ('*rainy@qc=0*'). There are also more advanced features allowed by using a query language, for example to find where the surface classes reported in the FCDR agree with ancillary information interpolated from the reanalyses regarding sea-ice.

4 Comparison between LDR and FCDR

4.1 Spatio-temporal coverage

As mentioned in the introduction, a motivation for this study is to evaluate, before assimilation into a full-fledged reanalysis, whether the CM SAF SSM/I FCDR improves on the SSM/I LDR. Figure 4 compares the data found in the SSM/I LDR for channel 1, in panel (a1), with those found in the CM SAF SSM/I FCDR, in panel (b1), for the satellite F13 and date 19970116, and only over sea-ice-free ocean points. The LDR coverage as found in the assimilation feedback appears less dense because of thinning carried out during pre-processing in ERA-Interim. That thinning retained one scan position out of five, and one scan line out of five. This thinning does not show too much here, because of the resolution employed for plotting (1^o latitude x 1^o longitude). Generally the geographical match between the LDR and the FCDR appears reasonable. Possible rain bands show at the same locations in both data records. Over the Western Pacific, one notices an additional quarter of orbit of data in the FCDR than in the LDR. The ERA-Interim assimilation logs confirm for this day that the F13 SSM/I data file for 06 UTC is smaller than for the other synoptic hours. This confirms one of the motivations for using the FCDR as containing more data than the LDR, to patch the gaps in the record.

Another feature in these maps is the larger extent towards the poles in the FCDR, compared to the LDR. Both are subject to filtering by sea-ice cover from ERA-Interim below 1%, and land-cover below 1%. Removing this criterion (not shown) does yield a complete coverage map for the FCDR, including land and Polar Regions. The LDR map, regardless of this criterion, always only shows ocean points, because a pre-screening procedure had removed observations over land before ERA-Interim assimilation. Inspecting the ERA-Interim code for reasons for data rejection prior to data assimilation also indicates that the surface type found in the SSM/I data was used to remove any observation that was not above open water. It is hence possible that the surface type reported in the SSM/I data received in NRT did not contain the most realistic sea-ice map. Figure 5 shows the surface classes found in the FCDR. The sea-ice area appears to be much smaller than the area that is missing in the LDR, suggesting the LDR contained probably too many points reported as sea-ice.

In order to assess whether this reduced sea-ice coverage in the FCDR is realistic, it is now compared with the sea-ice coverage found in ERA-Interim. For the date considered here, the sea-ice data in ERA-Interim come from the NOAA/NCEP 2D-Var weekly product (Reynolds *et al.*, 2002). Given that the sea-ice fraction from ERA-Interim is interpolated to the FCDR data, it is then straightforward to isolate with the SQL query the locations where the FCDR surface class indicates sea-ice (or sea-ice edge) and ERA-Interim sea-ice fraction also indicates sea-ice (fraction reaching or exceeding 1%), and to do similarly for the other 3 possible combinations, *i.e.*, no sea-ice in both, or sea-ice in one but not the other. This produces a contingency table of 4 maps shown in Figure 6. The maps are restricted to sea points only, according to the ERA-Interim land-sea mask (‘*lsm@ei<0.01*’). Panels (1a) and (2b) indicate that FCDR sea-ice and NCEP sea-ice agree most often. Panel (2a) shows that the FCDR features ice shelves that extend further away from the poles; as indicated by Figure 5c, this is caused by including the sea-ice margin in the FCDR selection criteria. Panel (1b) shows Arctic coastlines as ice-free in the FCDR but covered by ice according to NCEP. The NCEP sea-ice did not use SSM/I data in input, so this comparison is between independent sea-ice estimates. It is hence not clear which one of the two is correct. Another source of comparison is the sea-ice fraction from ERA-20C, actually from HadISST2 (Titchner and Rayner, 2014): it uses as input retrievals from passive microwave instruments, including from SSM/I.

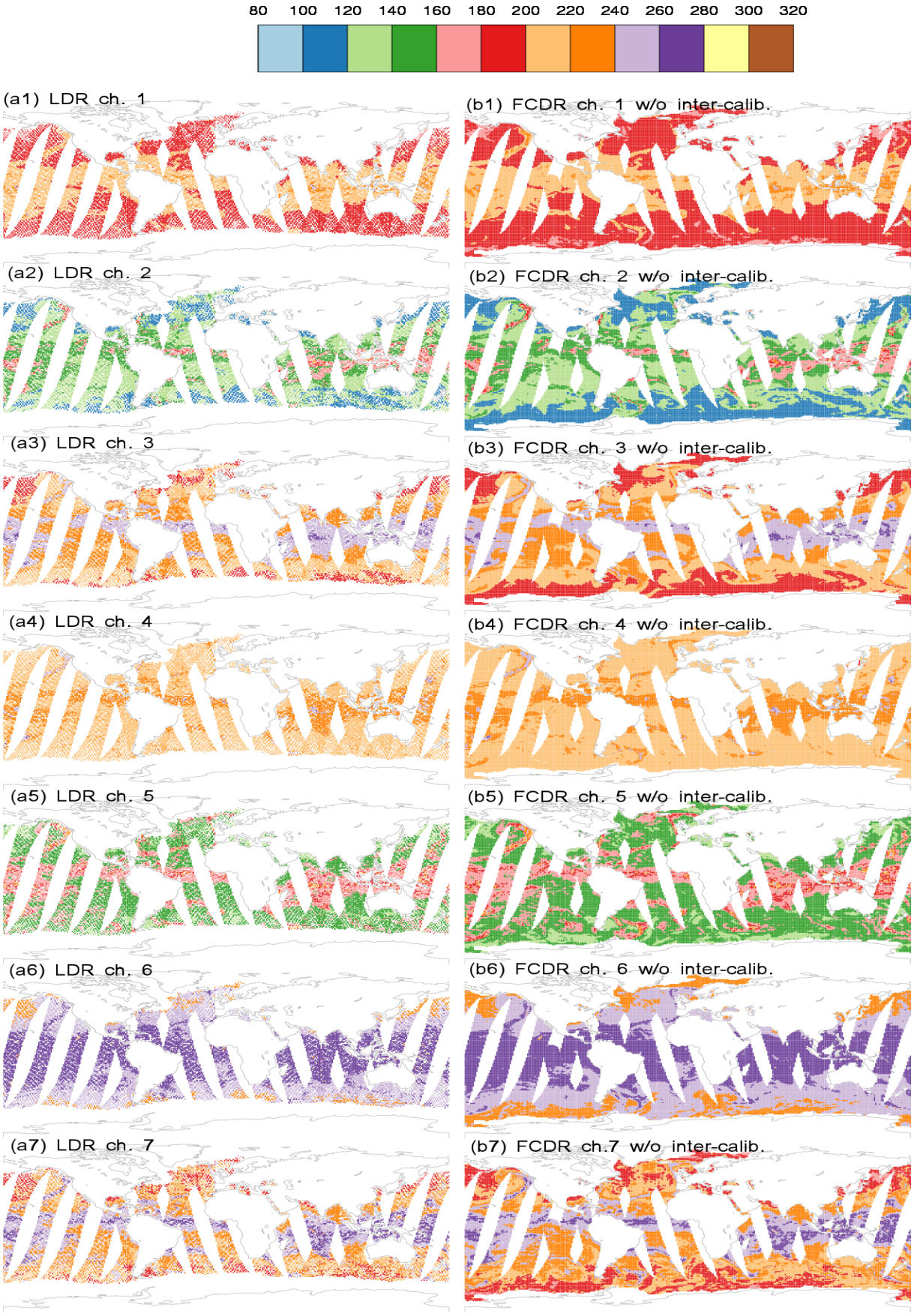

Figure 4: Brightness temperatures (K) from (a) LDR, which had been thinned spatially in ERA-Interim, and (b) FCDR without intercalibration, for satellite F13 on 19970116, considering only ice-free ocean points, binned to 1° x 1°

Restricting again the maps to only sea points according to the ERA-Interim land-sea mask (‘*lsm@ei<0.01’*) to avoid discrepancy between differing land-sea masks between ERA-Interim and ERA-20C (at different horizontal resolutions), Figure 7 shows findings similar to those above: many coastal locations are ice-free in the FCDR, but covered by sea-ice according to ERA-20C.

Near the Arctic coasts, the agreement of two independent sea-ice datasets (NCEP and HadISST2) versus a disagreeing FCDR does not necessarily mean that the latter is incorrect. One possible explanation is the finer horizontal resolution of the FCDR compared to the two sea-ice datasets, which were both interpolated from the reanalysis (*i.e.,* low) resolution, 80 km in the case of ERA-Interim and 125 km in the case of ERA-20C. Reduced sea-ice near the coasts is actually more plausible because of possibly enhanced mixing there for several reasons. For example, Wang *et al.* (2003) have indicated that accounting for tidal mixing in a high-resolution sea-ice model leads to reduced sea-ice fraction near the coasts, with more polynyas (areas of open water surrounded by sea-ice) found when this effect is taken into account.

Another finding, comparing Figure 6 and Figure 7, is that the HadISST2 sea-ice extent appears larger than NCEP. One possibility may be that the former has a sharper sea-ice transition than the other.

Overall, should one apply the same pre-screening procedure as ERA-Interim did prior to assimilation, it seems that a more realistic selection would be done from the FCDR than from LDR. This is yet another reason to prefer one to the other: the FCDR contains improved contextual information, which influences the pre-screening quality controls.

In terms of temporal coverage, Table 4 indicates that the FCDR presents a more complete record than the LDR, adding more than 13 satellite-years.

4.2 Brightness temperatures

Comparing the brightness temperatures found in the LDR and the FCDR can be done by matching the two records. To this end we consider a time period when the LDR is made up of data received from the GTS, the same satellite (F13) and same day (19970116) as before. The LDR and FDCR observations are matched for this day, considering only non-missing observation values. First we take the intersection of both data records for the minutes of the day that are common to both, allowing for up to +/- 10 seconds difference. Then, for each minute of data present in both data records, we look for a match to each LDR observation, picking the FCDR observation which is the closest (in terms of non-intercalibrated brightness temperature), among the set of FCDR observations that satisfy all the following criteria:

- scan position number (1-64) is identical to the LDR scan position number,
- time is within 10 seconds of the LDR reported observation time,
- location is within 100 km distance of the LDR reported position, and
- channel number (1-7) is identical to the LDR channel number.

Figure 8 indicates that most matches are found at +1 second. The FCDR ATBD explains that leap seconds had been introduced with considerable delay in the operational processing chain of SSM/I, and so the FCDR processing corrects this omission. In fact, between the day F13 was launched (March 1995) and 16 January 1997, only 1 leap second was introduced, on 1 January 1996 (Source: BIPM, http://www.bipm.org/en/bipm-services/timescales/time-ftp/publication.html#nohref). It is then without surprise that this leap second, introduced properly during the FCDR processing, reappears as giving the best match to the LDR. In terms of distance, this leap second difference introduces an average distance of approximately 7 km, which is also consistent with a satellite velocity of 7.5 km/s. For the observations that are matched between FCDR and LDR (at +1 second), Figure 9 shows that the both data records agree within 0.2 K standard deviation for

channels 1 to 5, with mean differences upwards of 0.3 K, and differences are the largest for channels 6 and 7, with standard deviations slightly above 1.1 K. Here we probably see the effect of the FCDR processing applying Antenna Pattern Correction (APC) different to that applied to data received from the GTS.

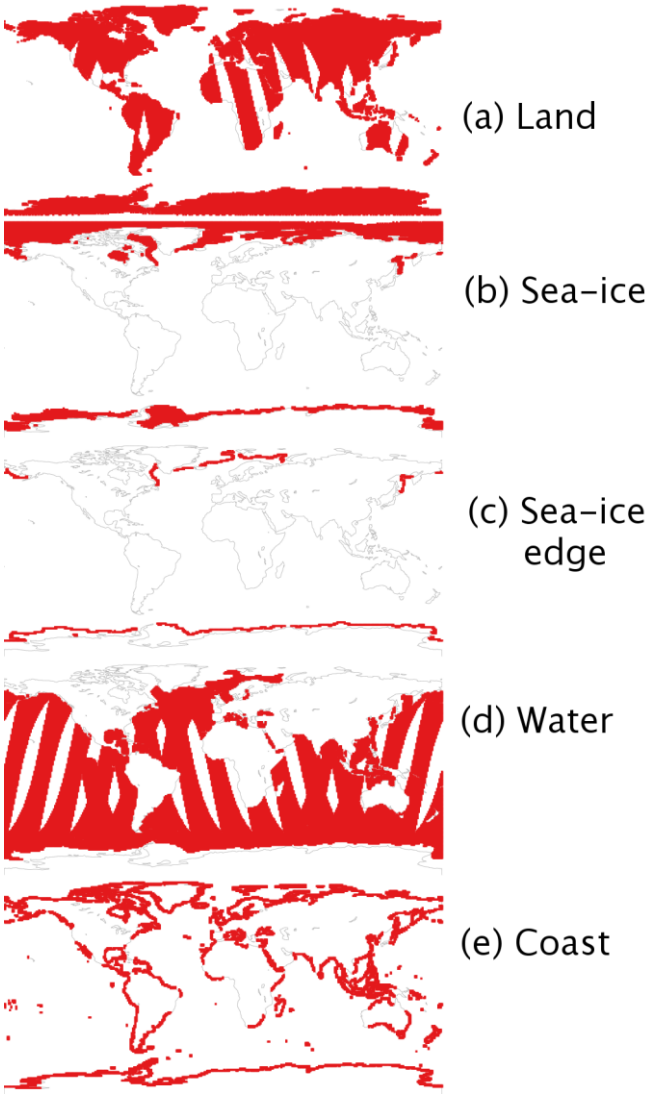

Figure 5: Surface classes in the FCDR for 19970116, satellite F13

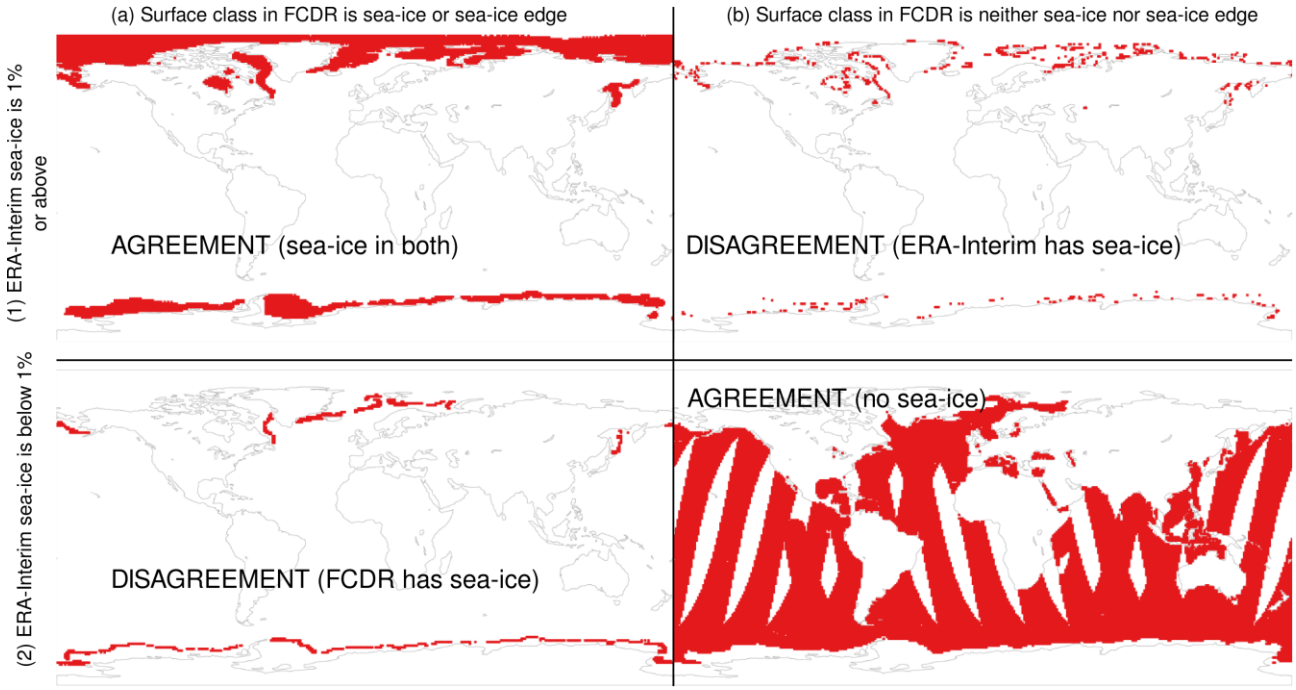

Figure 6: Contingency table for the sea-ice reported in the surface class in the FCDR (columns a, b) and matching sea-ice fraction from ERA-Interim at 1% or above (rows 1, 2), for 19970116, satellite F13

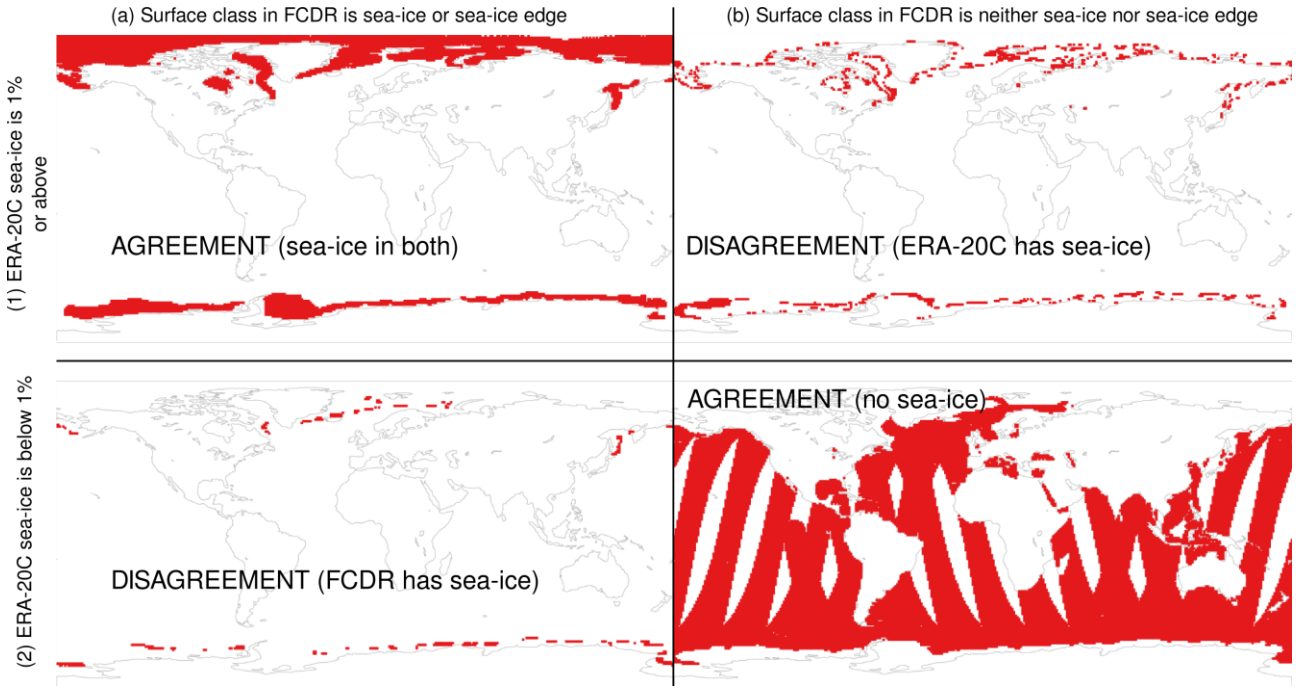

Figure 7: Same as Figure 6, but comparing with ERA-20C sea-ice

Satellite	Record	Start year/month	End year/month	Net gain from LDR to FCDR
F8	LDR	198708	199201	None major
	FCDR	198707	199112	
F10	LDR	199201	199212	+6 years
	FCDR	199101	199711	
F11	LDR	199301	199512	+5 years
	FCDR	199201	200001	
F13	LDR	199601	200812	+7 months
	FCDR	199505	200812	
F14	LDR	199906	200808	+2 years
	FCDR	199705	200808	
F15	LDR	200006	200608	None major
	FCDR	200003	200607	

Table 4: Start and end months of LDR and FCDR,

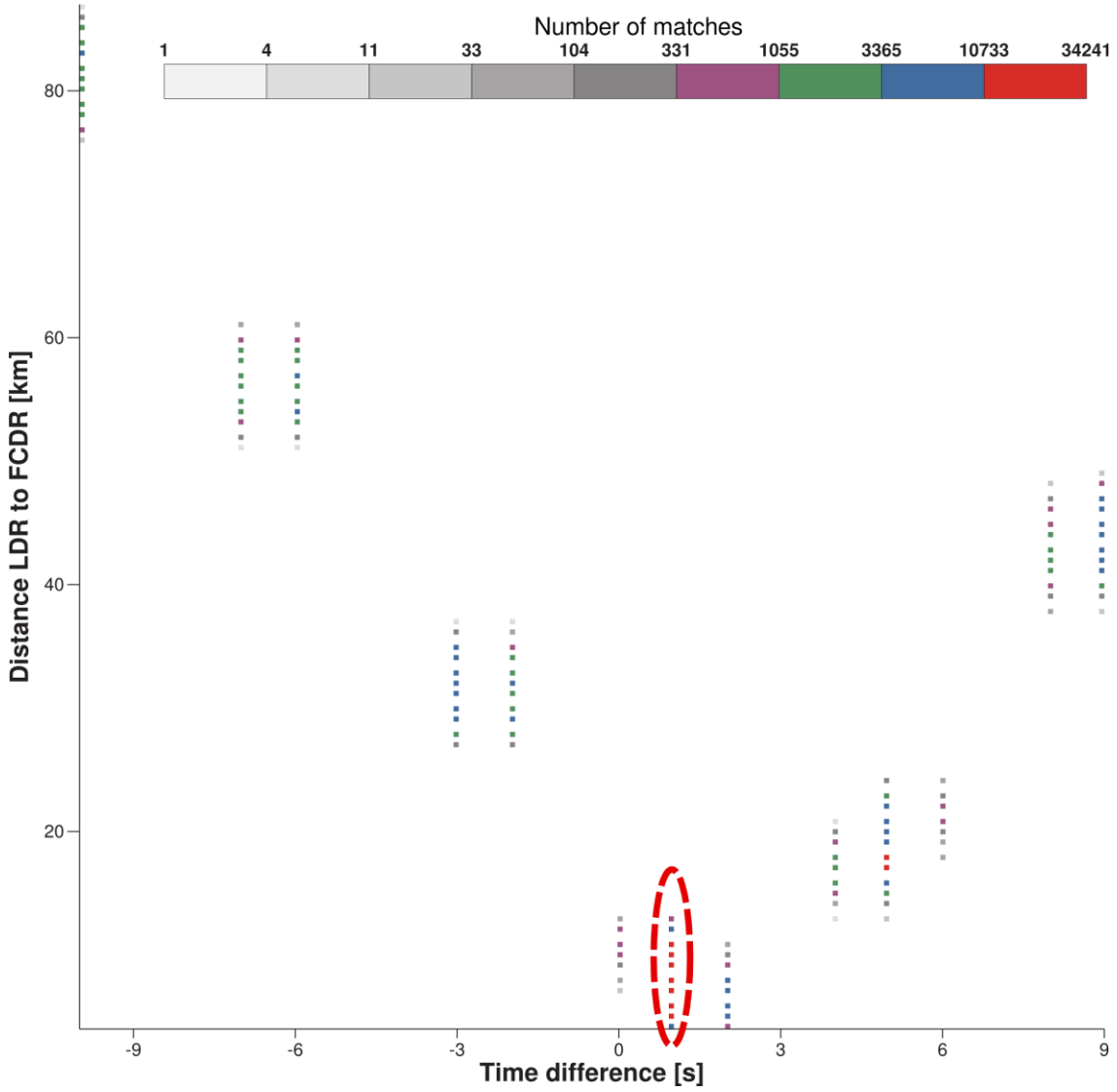

Figure 8: Number of FCDR matches to the LDR observations for satellite F13 on 19970116. The red ellipse highlights the greatest density of matches, when FCDR time = LDR time +1 second

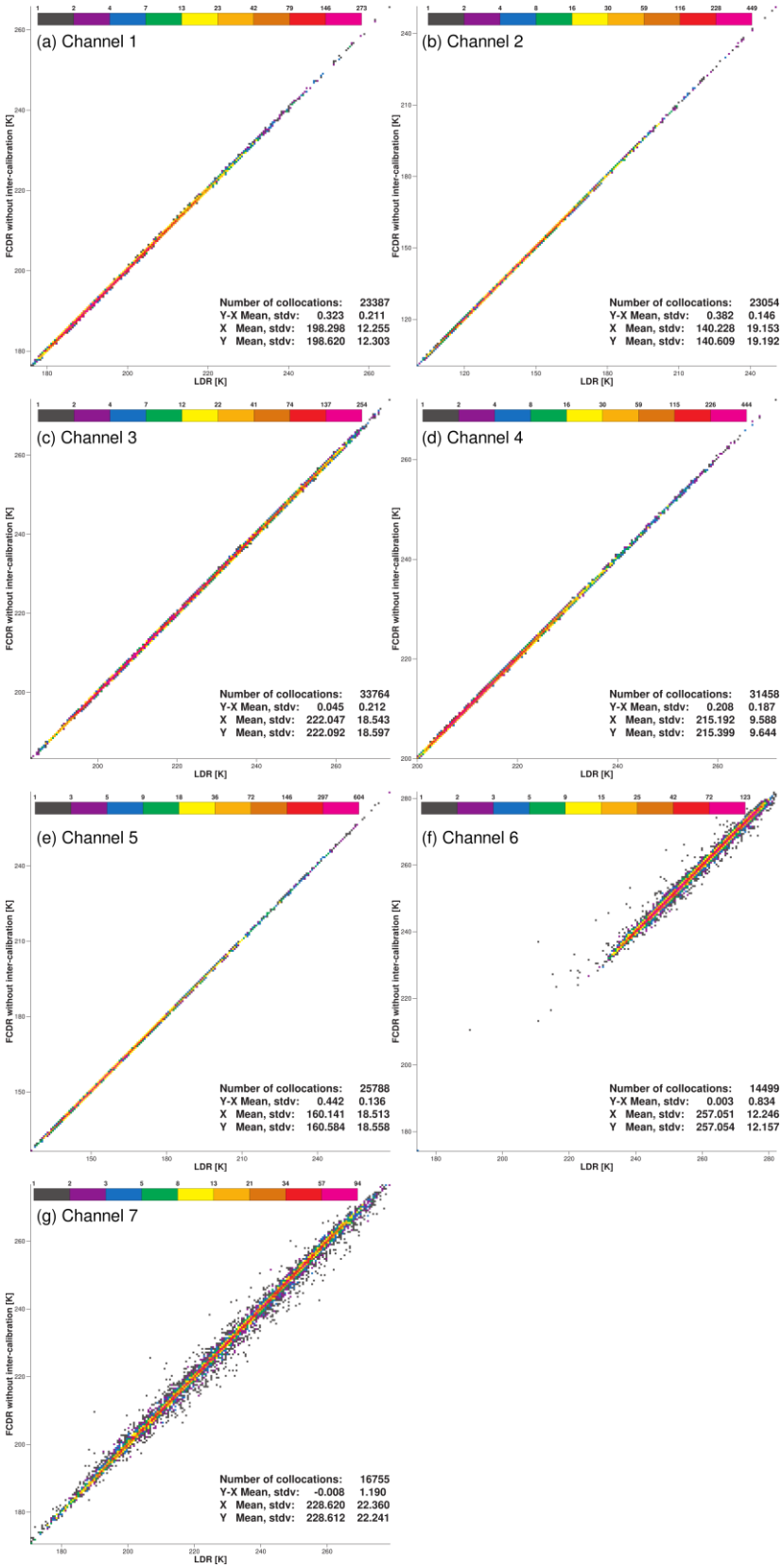

Figure 9: Comparison between the LDR observations (horizontal axis) and the FCDR, without intercalibration (vertical axis), for observations matched with the procedure described in the text on 19970116 for satellite F13 and with FCDR time = LDR time +1 second

5 Validation of the simulation tool

Before analysing results related to the comparison of the FCDR or LDR to the two reanalyses, a validation of the simulation tool is required. This can be done by comparing its results with earlier computations from ERA-Interim, assuming the latter result from a technically valid implementation of RTTOV in the IFS. However, Table 5 indicates that there are several differences between the two sets of computations, which make such validation not straightforward.

Component	ERA-Interim observation feedback	Simulation tool	Mitigation measure to align the simulation tool for validation purposes
Observation input	LDR	FCDR	Convert the LDR to the same format as FCDR, and use the result as input
Model analysis fields input	Semi-Lagrangian (SL) buffers in IFS, at T255 horizontal resolution and 60 vertical levels, with 30-minute time-step model integration	Retrieved from MARS at T255 horizontal resolution and 60 vertical levels, only available 6-hourly	Consider only observations within 15 minutes of the 6-hourly analysis fields. Small differences because SL buffers contain more information than the archive.
Vertical interpolation	IFS interpolation to the 43 pressure levels of radiative transfer coefficients (RTCOEF)	RTTOV version 11 interpolation to the 43 pressure levels of RTCOEF	Unchanged (source of difference)
Horizontal interpolation	IFS SL interpolation routines	Simulation tool interpolation (bilinear in latitude, longitude)	Unchanged (source of difference)
Radiative transfer model	RTTOV version 8	RTTOV version 11, but same RTCOEF (except for format)	Unchanged (source of difference)
Emissivity model over ocean	FASTEM version 2	FASTEM version 5	Use FASTEM version 2, or the emissivity in the ERA-Interim feedback

Table 5: Differences between the simulation tool and the RTTOV computations carried out within the data assimilation, and mitigation measures applied in this section of the report, for validation purposes

The third column indicates the steps taken to minimize the differences and align the simulation tool as much as possible with the radiative transfer simulations carried out by the ERA-Interim data assimilation system. Applying all these measures to observations above ice-free ocean regions (‘*lsm@ei<0.01 and seaice@ei<0.01’*) shown in Figure 10a, for 16 January 1997, retains a total of 1369 SSM/I soundings.

The first direct comparison between the brightness temperatures simulated by the tool and those found in ERA-Interim feedback shows a bias of 2.1 K and a standard deviation of 0.5 K (Figure 10a). However, ERA-Interim employed a different version of the radiative transfer model (RTTOV version 8) than used in the tool (RTTOV version 11). Although both use the same radiative transfer coefficients, the emissivity models differ: ERA-Interim used FASTEM version 2 whereas the tool uses FASTEM version 5, to the produce state-of-the-art comparisons to the FCDR. Choosing FASTEM version 2 (by running the simulation tool with the option "-v 2") reduces the standard deviation to 0.3 K (Figure 10b), but increases the bias to 3.5 K. There are several reasons while the emissivity could still differ from that used by ERA-Interim, so another approach is to use the emissivity that was saved in the ERA-Interim feedback (by running the simulation tool with the option "-e 1"). Figure 10c shows that this gives a bias and a standard deviation that are both 0.2 K.

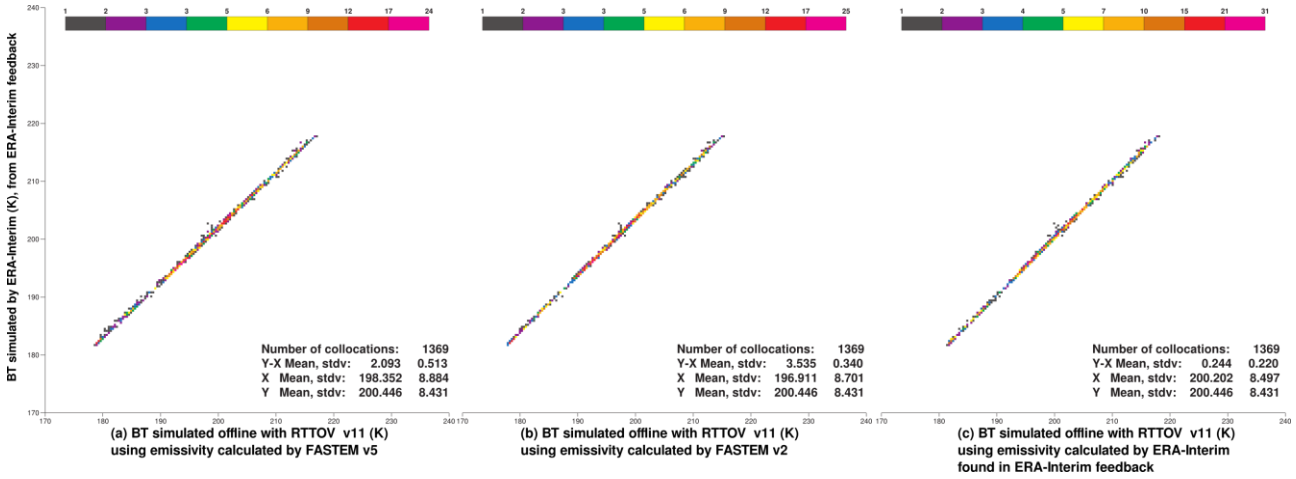

Figure 10: Comparison between (horizontal axis) simulations carried out offline and (vertical axis) brightness temperatures computed within the ERA-Interim data assimilation system, retrieved from the ERA-Interim observation feedback archive, for SSM/I channel 1. Panels (a)-(c) show the impact of the emissivity model. (a) FASTEM version 5 (used in the remainder of the present report), (b) FASTEM version 2 (version that was employed in ERA-Interim), (c) emissivity found in the ERA-Interim observation feedback archive

These results are only for one channel, and differ somewhat for the other channels. Table 6 shows a summary for all channels. The table suggests that the choice of the emissivity model has a large influence on the ability to reproduce ERA-Interim results. Highlighting in green the results that are closest to ERA-Interim calculations, we find that the emissivity that was computed within ERA-Interim gives the best match. Although it does not allow the best reproduction of ERA-Interim results, the choice of FASTEM v5 induces a mean offset which is channel- and region-dependent, but has no significant influence on the standard deviation fit to observations. A similar investigation carried out for another satellite (F8) showed very similar results, suggesting the robustness of the offsets generated by application of the offline tool (as compared to the ERA-Interim RTTOV simulations during the assimilation).

Bias, and standard deviation in K	LDR *minus* ERA-Interim (O-A, for reference, without any bias correction applied)		LDR *minus* Simulation with Emissivity found in ERA-Interim feedback		LDR *minus* Simulation with Emissivity calculated by FASTEM v2		LDR *minus* Simulation with Emissivity calculated by FASTEM v5		ERA-Interim *minus* Simulation with Emissivity found in ERA-Interim feedback		ERA-Interim *minus* Simulation with Emissivity calculated by FASTEM v2		ERA-Interim *minus* Simulation with Emissivity calculated by FASTEM v5	
chan. 1 19V	-0.2	3.3	0.0	3.3	3.3	3.2	1.9	3.2	0.2	0.2	3.5	0.3	2.1	0.5
chan. 2 19H	1.5	6.0	5.0	6.1	3.5	6.0	3.7	6.1	3.6	0.8	2.0	0.5	2.3	0.6
chan. 3 22V	0.8	2.5	0.8	2.5	3.2	2.4	1.7	2.5	0.0	0.4	2.3	0.6	0.9	0.6
chan. 4 37V	-1.1	4.6	-0.8	4.6	2.0	4.5	0.3	4.5	0.3	0.2	3.0	0.3	1.4	0.3
chan. 5 37H	3.4	9.4	7.9	9.5	5.3	9.4	3.4	9.5	4.5	0.9	1.9	0.5	0.1	0.7
chan. 6 85V	0.5	3.7	0.3	3.8	1.8	3.8	1.3	3.8	-0.2	0.3	1.3	0.5	0.8	0.5
chan. 7 85H	7.0	7.3	10.0	7.7	7.8	7.5	5.6	7.1	3.0	1.3	0.8	0.7	-1.4	1.3

Table 6: Validation of the simulation tool by comparison with the calculations found in the ERA-Interim feedback, using the same 1369 SSM/I soundings from F13 shown in Figure 10, but showing all channels in the present table. Dark (light) green cells show the closest match to O-A (respectively: ERA-Interim calculations).

6 Analysis of the differences between data records (FCDR, LDR) and gridded reanalysis datasets (ERA-Interim, ERA-20C)

6.1 Spatial differences

We first assess maps of mean differences between the FCDR and one reanalysis. Figure 11 shows maps of mean differences between the FCDR brightness temperatures and simulations from ERA-Interim, for channel 7 (85H) around Australia. We consider for each satellite at least 6 days of data, spread over at least 3 months, and averaged in 0.1 degree x 0.1 degree latitude, longitude bins. The figure separates satellites (rows 1 to 6), and ascending orbits (column a) from descending orbits (column b). The maps show positive haloes along the coastlines. These are caused by (improperly) simulating coastal observations as if they were completely over ocean. This issue is thus an artefact of the simulation procedure. Another element to consider is that the mean of differences between observations and reanalysis contains reanalysis biases, which can be regional and slowly varying.

Such maps cannot be used alone to identify geolocation problems. Berg *et al.* (2013) proposed a method to address the issue of geolocation with SSM/I data, caused by instrument mounting angle, satellite attitude, and satellite position differing slightly from expectations (even if all may be within the original mission specifications). The CM SAF devoted particular attention to this issue (section 4.2 of the ATBD). We consider here the difference between mean statistics for ascending orbits and those for descending orbits, as done by Berg *et al.* (2013). Such approach removes any inconsistency in the simulation suite near the coast, and mitigates the impact of large, systematic reanalysis biases. However, diurnal reanalysis biases would manifest themselves differently between ascending versus descending orbits, and remain visible on such maps. Even so, such bias should not necessarily be larger at the coast (unless it is proven that reanalysis diurnal biases are most pronounced in coastal areas). Furthermore, rainy scenes, which may not be sampled equally by ascending and descending orbits, can dominate the observational signal, and therein could be a diurnal cycle dependence, with more convection in late afternoon/early night than in the morning. We thus apply the rain detection scheme (though this only works over oceans, we retain all scenes over land). Maps in Figure 12b show differences between statistics considered earlier for ascending orbits minus descending orbits. These contain the signal from observations as well as from reanalyses, bearing in mind the possible presence of diurnal reanalysis biases in coastal areas. Over the coastlines, the haloes shown earlier cancel out for all satellites, except for satellite F14, where one can clearly see the coastlines delineated, with signs that differ between the Northern and Southern sides of the continent. Maps in Figure 12a show differences (ascending orbits) minus (descending orbits) in ERA-Interim simulations. For F14, one finds differing shades for land versus sea, but the coastlines do not stand out, other than being an imaginary line where colours shift from positive to negative. The origin of the problem for F14 is confirmed as being in the observations, with Figure 12c showing differences (ascending minus descending) in the FCDR. This would tend to suggest that there may be room for further improvement in the SSM/I geolocation, at least for the F14 satellite, for a future FCDR version.

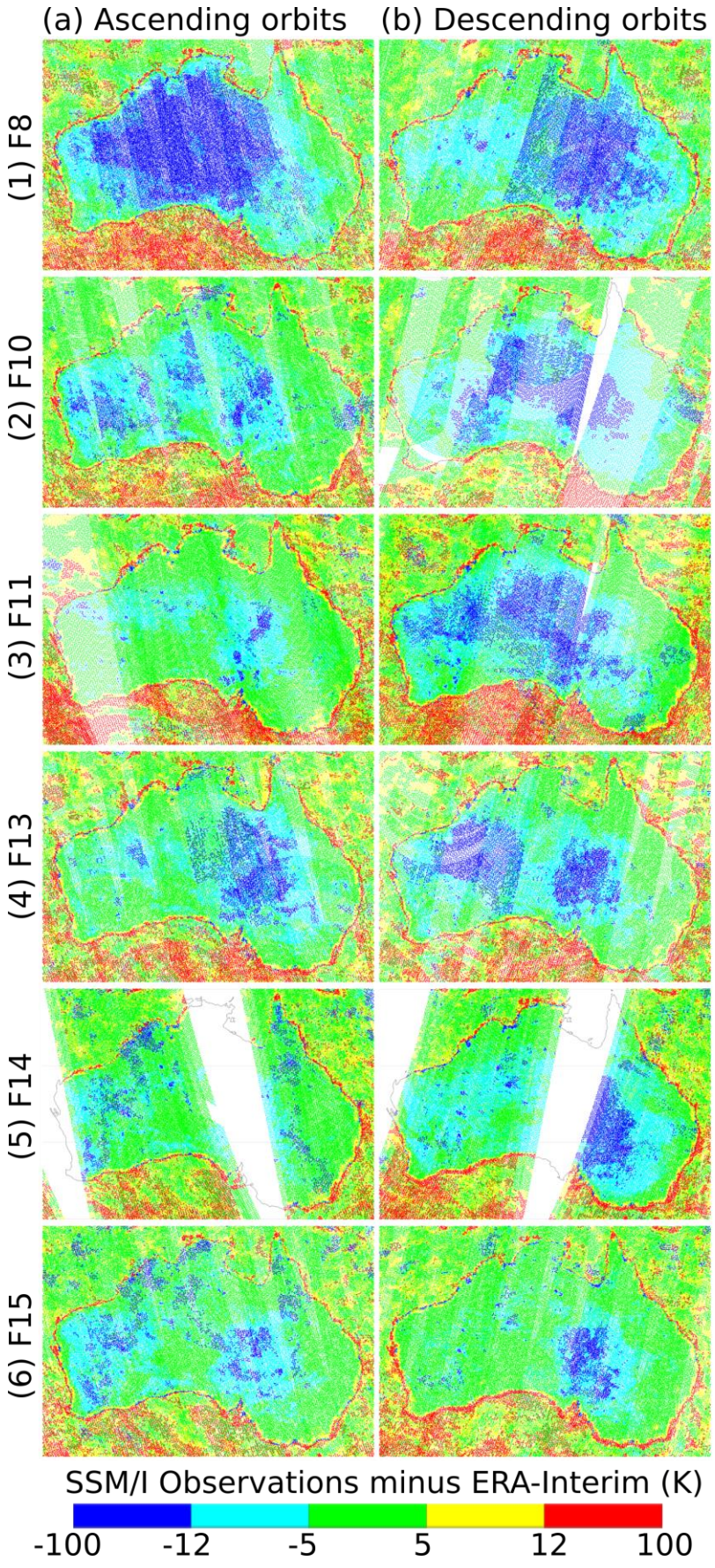

Figure 11: Maps of mean differences between the FCDR brightness temperatures and computations from ERA-Interim (using the simulation tool), for channel 7 (85H). Dates considered are as follows: F8, 19870716, 19870731, 19870815, 19870830, 19870914 19870929; F10, 19920116, 19920131, 19920215, 19920301, 19920316, 19920331; F11, 19950415, 19950430, 19950515, 19950530, 19950614, 19950629, 19960409, 19960424, 19960509, 19960524, 19960608, 19960623; F13, 19960409, 19960424, 19960509, 19960524, 19960608, 19960623; F14 and F15, 20020116, 20020131, 20020215, 20020302, 20020317, 20020401.

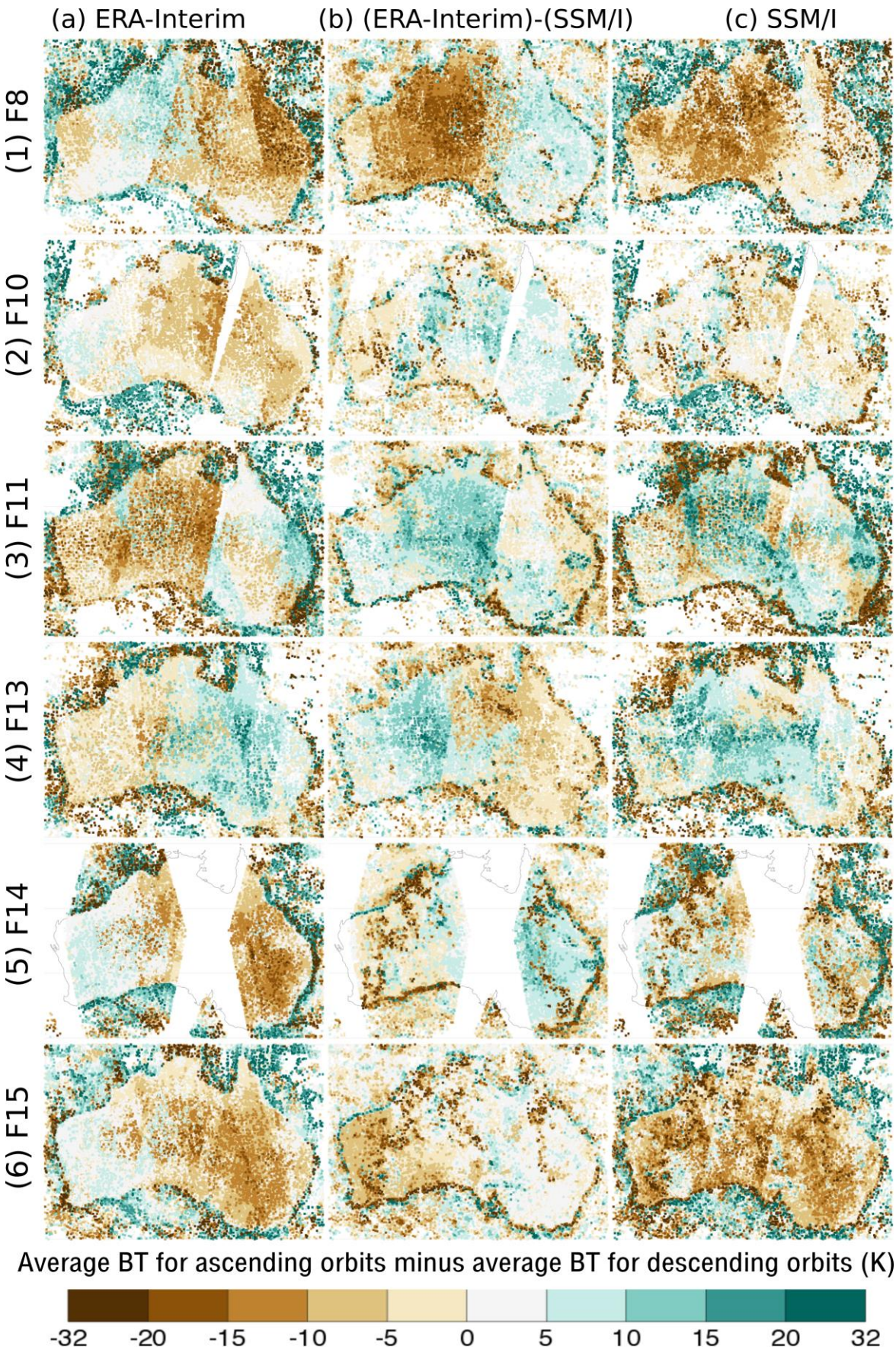

Figure 12: Differences between ascending and descending orbits for channel 7 (85H), for the same scenes as shown in Figure 11

6.2 Temporal differences

Figure 13 to Figure 19 show, for channels 1 to 7, several timeseries at monthly resolutions. In order to remain within the validity domain of the simulation tool, all statistics are generated from observations that are only over sea, with sea-ice concentration below 1, and believed to be free of rain contamination.

6.2.1 Impact of the intercalibration

The first row in Figure 13 shows the impact of the intercalibration in the FCDR, using the ERA-20C analysis as a comparison basis. The mean differences with ERA-20C are more stable over time, for all satellites and channels (see Figure 14 to Figure 19), when the intercalibration is applied, than when it isn't applied. This result validates the approach followed with the intercalibration to improve the stability and consistency within the data record.

It is interesting to note that the satellite employed as reference for the intercalibration can nearly be inferred from the plot: the grey lines are mostly clustered when the F11 satellite is available. This is most visible in Figure 18, for channel 6, for which one notes that, without intercalibration, all satellites other than F11 are consistent between one another. The F11 satellite stands out from the others for that channel. After application of the intercalibration (to F11), all the other satellites are brought in line. In such instance, it is questionable whether F11 is really the best reference to intercalibrate channel 6 data. The choice of F11 as the reference in the CM SAF FCDR was motivated by a combination of factors (see ATBD): F11 is the satellite that has the most overlaps with other satellites, is stable compared to *in situ* measurements, and has low non-linearity response (compared F13). However, Figure 18 illustrates that choosing "the best reference" to use to align data records in a FCDR remains a difficult problem. Given this uncertainty, namely that several references are possible, assuming some distribution of their likelihood (all equally likely, some more likely than others, etc…), an ensemble of FCDRs obtained with an ensemble of "best references" could be another way forward.

Another element of interest is that the intercalibration appears less efficient for F14 than for the other satellites. In the FCDR, channel 4 (37V) shows a mean difference between F14 and the other satellites approximately 0.5 K. This difference is nearly as much as 1K for channel 5 (37H). A comparison of F13 and F14 against collocated observations from the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI) is shown in Figure 20, for matching channels. Timeseries of differences (F13 minus TMI), and (F14 minus TMI), are conceptually similar to the top rows of in the earlier 7 figures, except that TMI instead of ERA-20C is used as the comparison baseline. The double differences (differences between these two sets of differences) are all within +/-0.2K, and are not as large as in the comparison to ERA-20C, especially for channels 4 and 5 (37V and 37H), where F13 and F14 differ by 0.5 to 1 K in spite of the intercalibration. Figure 21 shows density plots of the differences with respect to TMI for channel 5 (37H), with the horizontal axis showing differences between SSM/I channels 4 and 5 (37V minus 37H, useful to detect rain contamination). Differences F13 minus TMI and F14 minus TMI both depict a nearly identical behaviour and offset. As Figure 20, this is also in disagreement with Figure 17 which shows a difference between F13 and F14 of approximately 0.5K for channel 5. Furthermore, the intercalibration in the FCDR from CSU also shows a correction value of zero for this channel (see Table III shown by Sapiano *et al.*, 2013), suggesting that non-intercalibrated lines should also show no differences with respect to ERA-20C. Overall the differences seen between F14 and the reanalyses cannot to be explained at this point, but the samples used for all these comparisons are of course different, and therein may lay some of the explanation. However, applying similar rain checks to filter out the rain-contaminated collocation situations considered with TMI does not change the results. Another possibility could be an incorrect use of RTTOV or incorrect coefficients

in the offline tool to simulate the F14 satellite. Attention will have to be given to monitor closely the data assimilation of F14 from the FCDR.

6.2.2 Improvement in temporal consistency from LDR to FCDR

The second row in Figure 13 shows the mean differences between the two SSM/I observation records (FCDR, LDR) and the two reanalyses (ERA-20C, ERA-Interim). Differences between non-intercalibrated FCDR and ERA-20C resemble very much the differences between non-intercalibrated FCDR and ERA-Interim. Mean differences for channel 1 (as well as channel 4, Figure 16, and channel 7, Figure 19) feature an offset of approximately 1K between F8 and F10. Mean differences for channel 6 (Figure 18) are similar across satellites, except for F11.

However, for channel 3 (22V, sensitive to water vapor), the differences with ERA-20C are more stable over time than with ERA-Interim. The reasons behind this are not straightforward and can be understood from the combination of timeseries shown here. They have to do with ERA-Interim having assimilated SSM/I data. The first element to note is a jump in the timeseries of differences between the non-intercalibrated FCDR and ERA-Interim on 1 January 1992. This coincides with the introduction of F10 in ERA-Interim. ERA-Interim only assimilated one satellite a time until mid-1999. Few profiles in rainy situations were assimilated from F8, but with F10 this situation changed. The number of profiles in the so-called 'rain assimilation' jumped from near-zero to about ten thousand per day. The observation operator in this assimilation method produced about twice the observed amount of rain, and the subsequent assimilation corrected this by drying the analysis accordingly (Geer *et al.*, 2008). The net result was that in January 1992 there was a rapid, sudden drying of the ERA-Interim analyses. A trace of this problem is visible in the timeseries of differences between the clear-sky data from the LDR considered for assimilation and the ERA-Interim background: there is a spike of about -1K during the first month of 1992, during which the analyses dried up, and the variational bias correction also got adjusted. Because ERA-20C is completely independent from the SSM/I data record, it does not feature any such discontinuity for channel 3 in 1992 with respect to the FCDR, whether intercalibrated or not.

It is important to remind that the statistics related to the LDR are over a set of observations different from those in the FCDR, and the validation carried out in section 5 reported systematic offsets between calculations from ERA-Interim and offline simulations. After mid-1999, the difference (LDR minus ERA-Interim) features the same temporal evolution as the difference (FCDR minus ERA-Interim), with a downward trend of about -0.5 K/decade over the course of the following 10 years. Before mid-1999, there is an offset between these two timeseries, which has to do with a change in the origin of SSM/I data in the LDR around that time (before, intercalibrated record from RSS; after, as received from the GTS; see section 1). This explains why the green curves agree well with the grey after 1999 (and do not, before). For surface-sensitive channels, the agreements are not as good as for channel 3, and differences (offsets up to ~2K) were explained in section 5.

The quality of the time continuity of the RSS data record before 1999 in the LDR and assimilated in ERA-Interim is noted, for there are generally smoother transitions in the mean between satellites before 1999 than after.

For completeness, the figures also shows mean differences between the LDR and ERA-Interim analysis after application of the bias correction. These are, as expected, aligned on the zero line throughout the period.

6.2.3 Comparison of the variability

The third row in Figure 13 shows the standard deviations of differences between the two SSM/I observation records (FCDR, LDR) and the two reanalyses (ERA-20C, ERA-Interim). For all channels (Figure 14 to Figure 19), these are largest with respect to ERA-20C, which did not assimilate SSM/I or any other humidity observations. For the water vapor channel (channel 3), the standard deviations of differences between the non-intercalibrated FCDR and ERA-20C is approximately 7 K. This number, much larger than with comparison to ERA-Interim background or analyses, at approximately 3 or 2 K, is only a third of the intra-month variability within the observations shown in the fourth row, at approximately 21 K. Consequently, the ERA-20C reanalysis explains approximately 90% of the variance in the SSM/I observations. Assuming that this variance corresponds to natural variability in total column water vapor, this indicates that the ERA-20C reanalysis explains 90% of the natural intra-month variability in total column water vapor in the tropics. Accounting for the SSM/I instrument noise to the 21 K standard deviation within the observations, the percentage of total column variance explained by ERA-20C in the tropics may be higher than 90%. The remaining 10% or so of variance that are missing may be caused by the use of monthly Sea-Surface Temperature (SST) forcing data and the absence of humidity observations in the assimilation. Using the same method, it would be interesting to assess how much additional variance can be explained by a daily SST, or is found in model-only integrations such as ERA-20CM (Hersbach *et al.*, 2015).

The standard deviations of differences between the non-intercalibrated FCDR and ERA-Interim analysis is, for channel 3, approximately 2.5 K, which is larger than for LDR minus ERA-Interim analysis but smaller than LDR minus ERA-Interim background. One possible explanation, for not reaching such values as in the LDR statistics, is the temporal mismatch; the population of observations considered here includes all the data within a day: during the assimilation the proper time of the observations was taken into account with a model integration (time-step of 30 minutes) whereas the simulation tool picks the nearest neighbor in time from analysis fields with a resolution of 6 hours.

One notes, for all satellites available, a decrease in the standard deviation of differences with respect to ERA-20C towards the end of the time period. This may be due to the sharp increase in the quality of ERA-20C SST in the Southern Oceans during the 2000s.

One remarks for channel 3 that the standard deviations of differences LDR minus ERA-Interim analysis after bias correction (in blue) decrease by approximately 10% after June 1999. This indicates that the analysis suddenly fitted more the observations. This is confirmed by inspecting the width of the shading which shows differences between background and analysis departures from the LDR before bias correction for that channel. This beam becomes narrower after mid-1999, consistent with observations more affected by random noise than by structural, correlated noise, before that. This illustrates *a priori* the double-edged aspect of using a FCDR (with intercalibration) for assimilation into reanalysis. The benefit is to present the assimilation with more continuous timeseries; the disadvantage is that data may contain larger structural errors. As shown by the timeseries of LDR minus analysis and with bias correction, which improve approximately 1999, such structural errors cannot be dealt with by the assimilation as well as raw, non-intercalibrated, observations, for which the variational bias correction and the assimilation may be able to find a better agreement between all the data sources. However, a proper answer to this question requires a data assimilation experiment.

The fourth row in Figure 13 shows the intra-month variability within the FCDR by considering the standard deviation within the domain and within the month. Generally the FCDR features slightly lower standard deviation when the intercalibration is applied, indicating that some the intra-month variability found in the original data was removed during the intercalibration. However, the difference between the two is marginal.

As expected however, one does not find episodes when the intercalibrated FCDR features systematic larger variability than the raw FCDR; an additional spurious signal (not found here) would have suggested the introduction of noise or gross errors in the intercalibration process.

The water vapor sensitive channel 3 (22V, Figure 15) shows evidence of the El Nino events. During 1991/2 and 1997/8 the intra-month standard deviations within the data are larger.

Overall, statistics from channels 6 and 7 from F8 in the FCDR stand out from the others. The data from the two 85 GHz channels on F8 were affected by sensor problems, and the data in the FCDR (as explained by the ATBD and the PUM) are synthetic, and contain no additional information compared to channels 1 to 5. For assimilation purposes, it may be reasonable to consider rejecting those data.

Figure 13: (next page) Monthly timeseries for channel 1 (19V). First row: mean differences between the FCDR and the ERA-20C reanalysis. Second and third rows: mean and standard deviation of various differences between the FCDR, the LDR, and the ERA-Interim and ERA-20C reanalyses. Fourth row: intra-month variability within the FCDR. Note that statistics related to the LDR are from the ERA-Interim observation feedback, for observations that passed all the assimilation quality controls; they thus correspond to a smaller observational sample than the FCDR

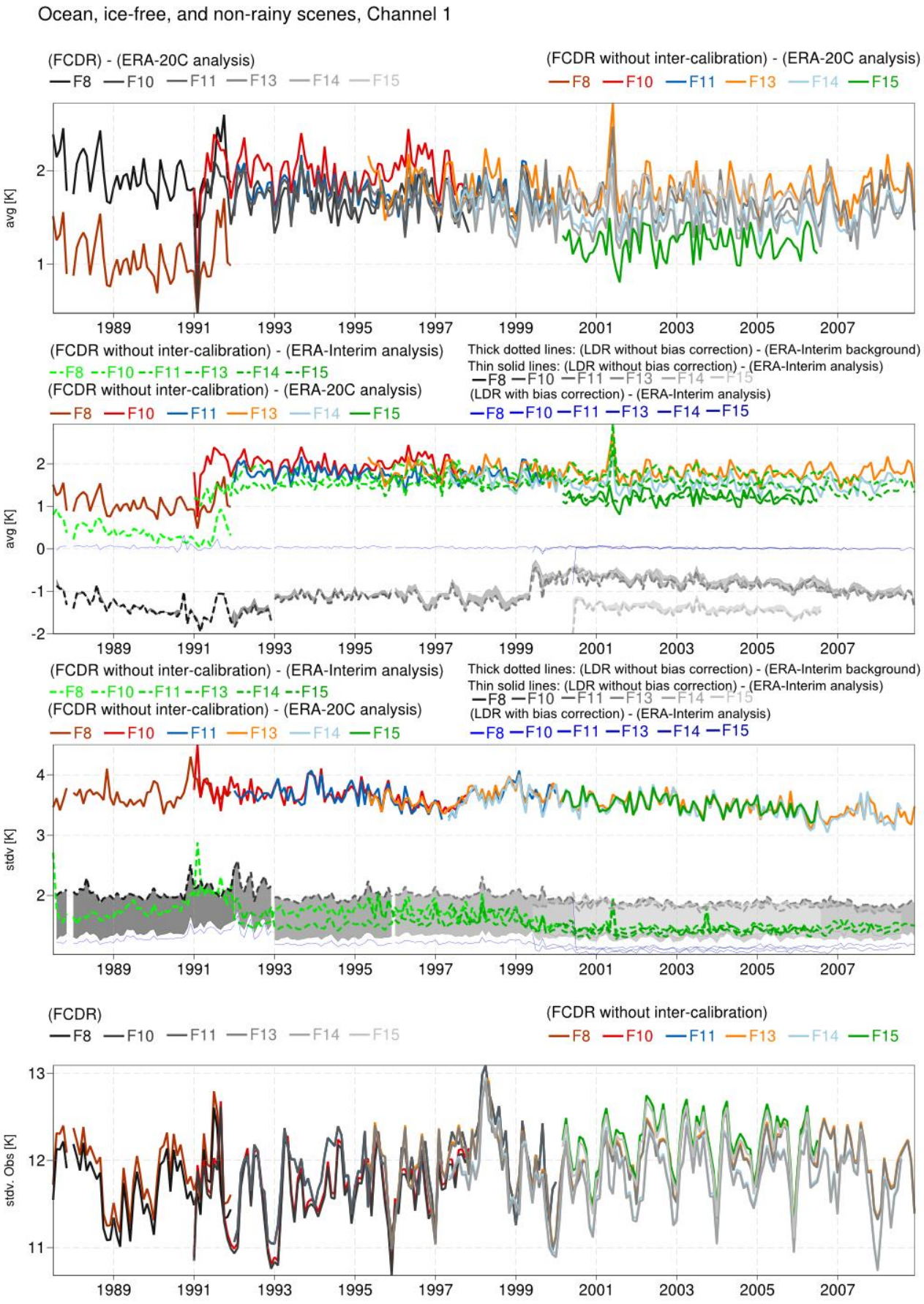

Figure 13: see caption on previous page

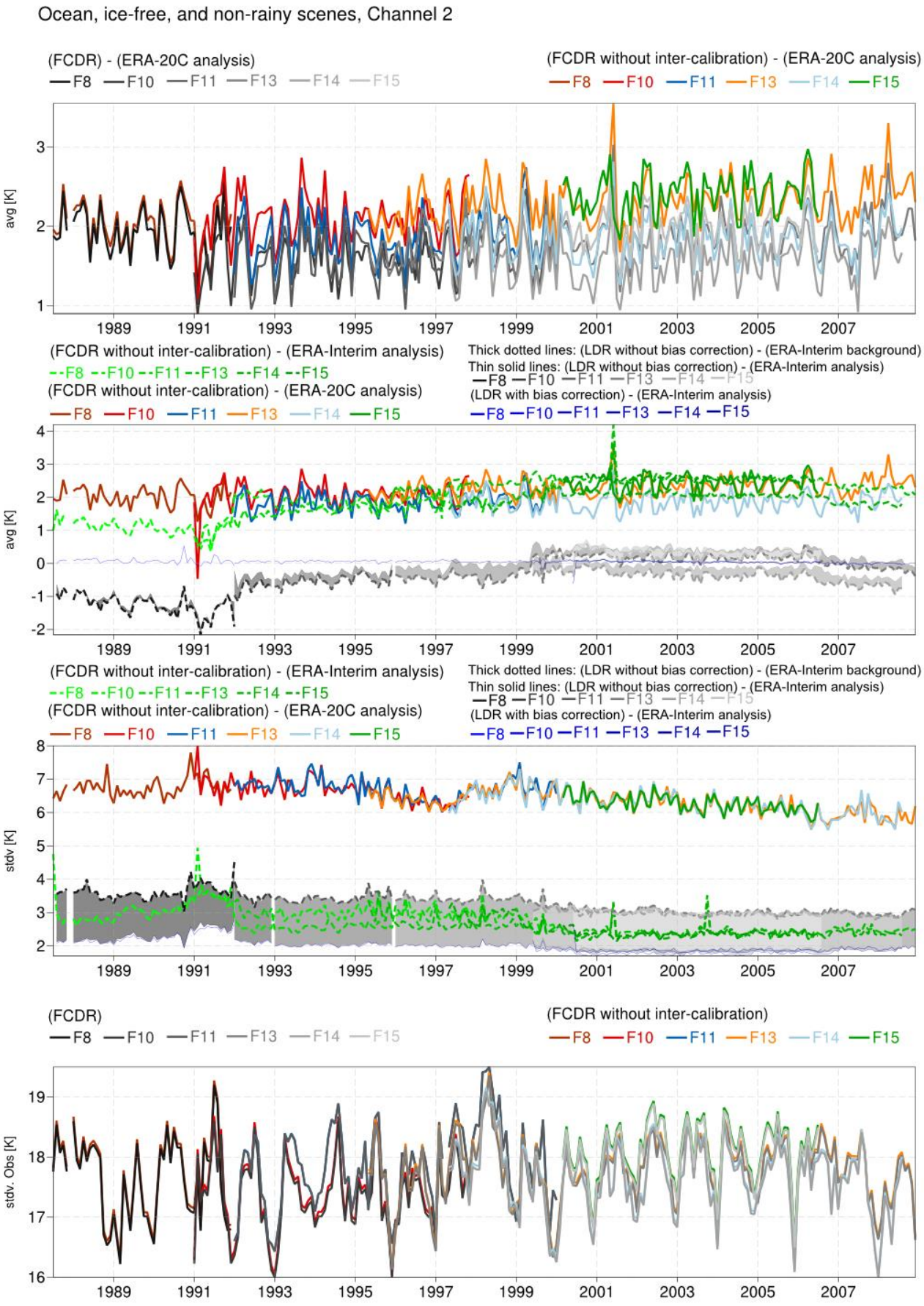

Figure 14: Same as previous figure, but for channel 2 (19H)

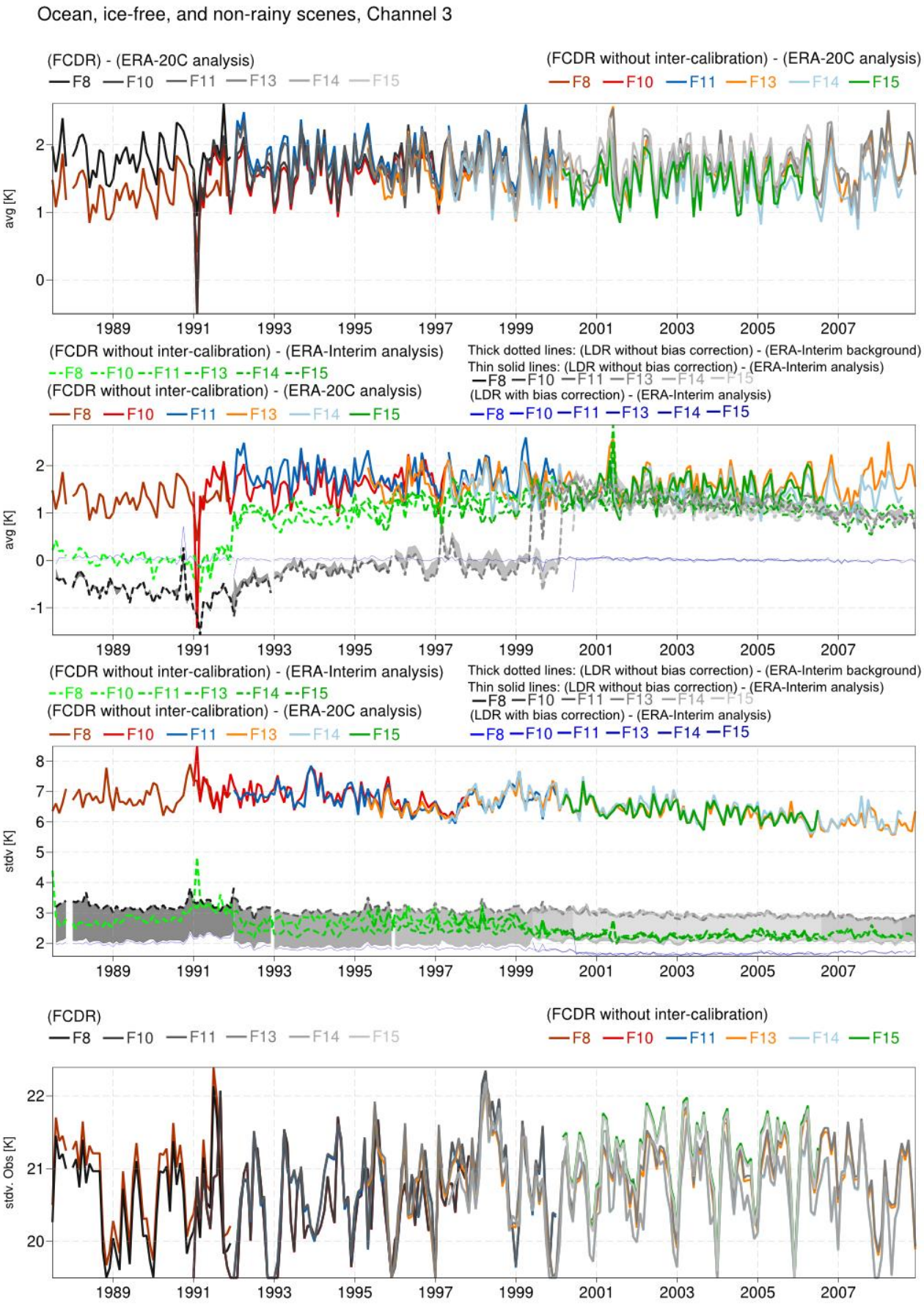

Figure 15: Same as previous figure, but for channel 3 (22V)

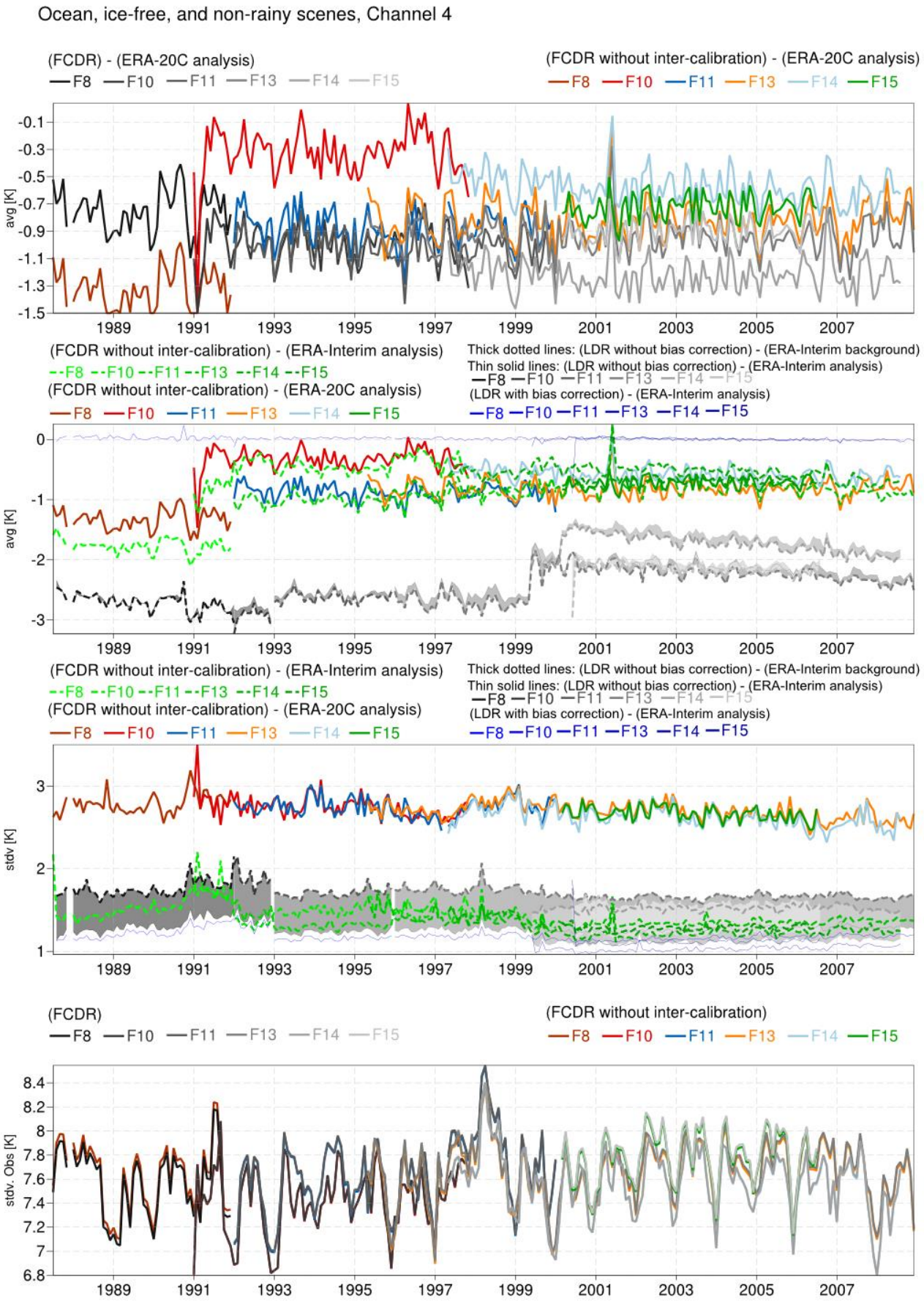

Figure 16: Same as previous figure, but for channel 4 (37V)

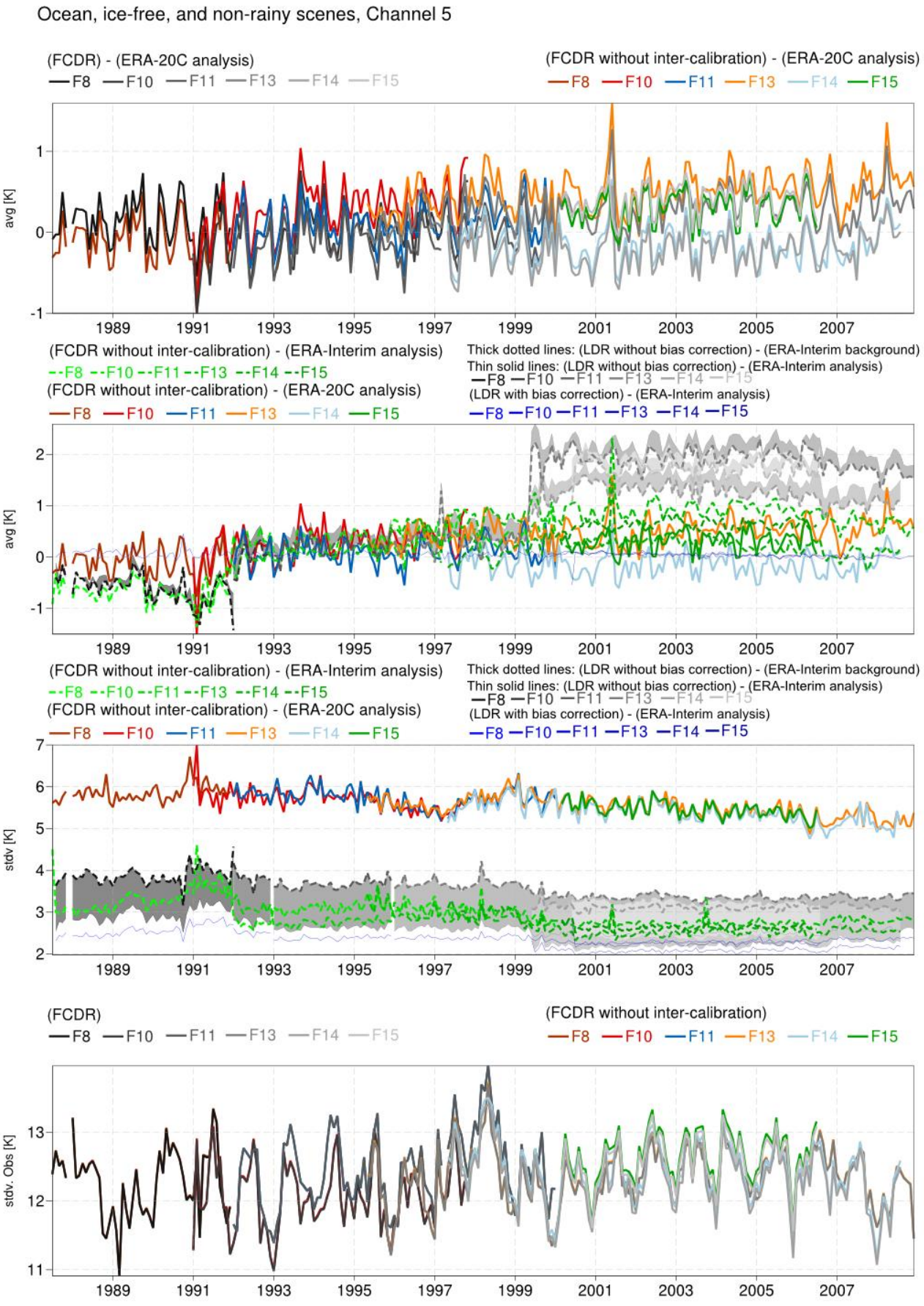

Figure 17: Same as previous figure, but for channel 5 (37H)

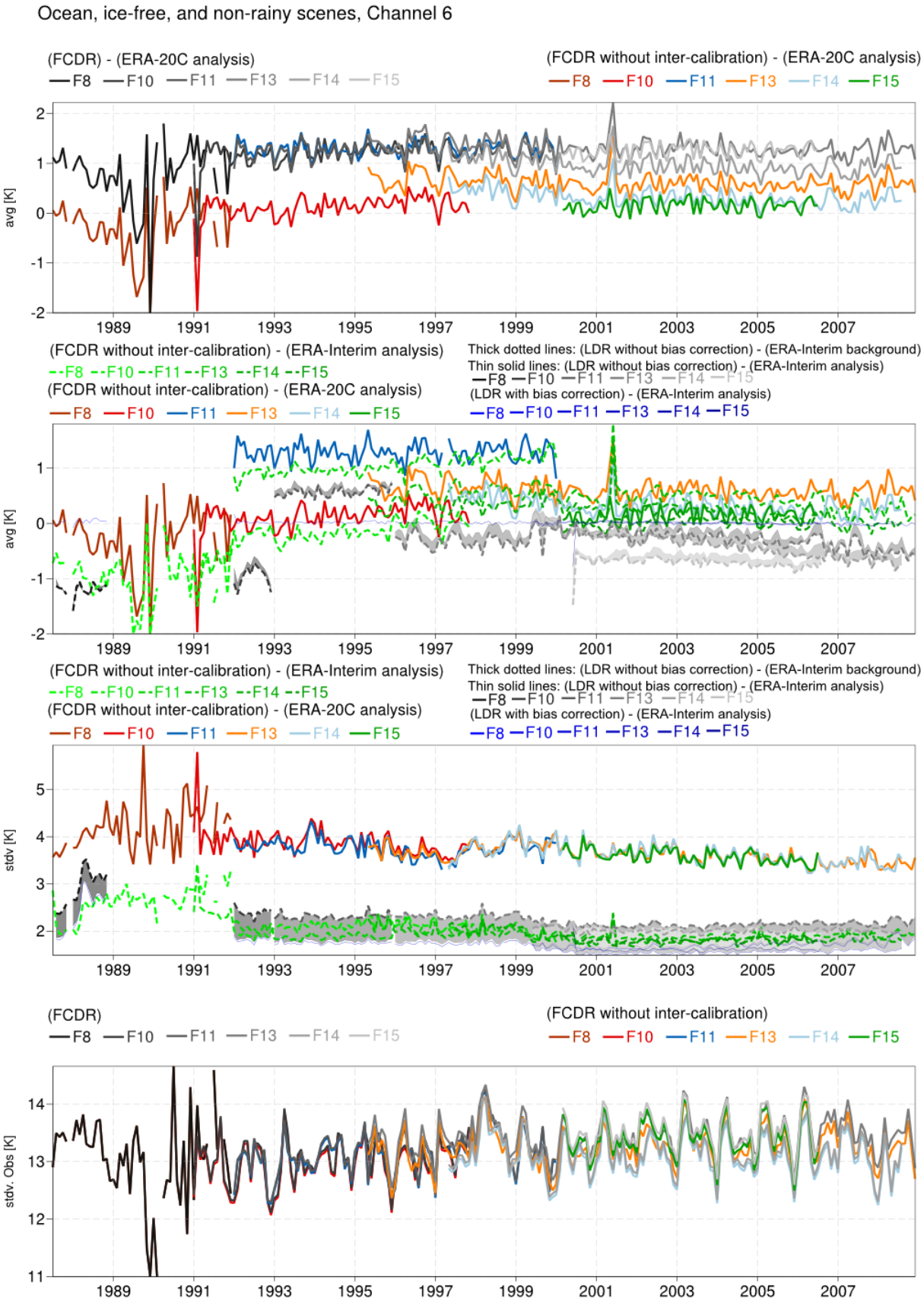

Figure 18: Same as previous figure, but for channel 6 (85V)

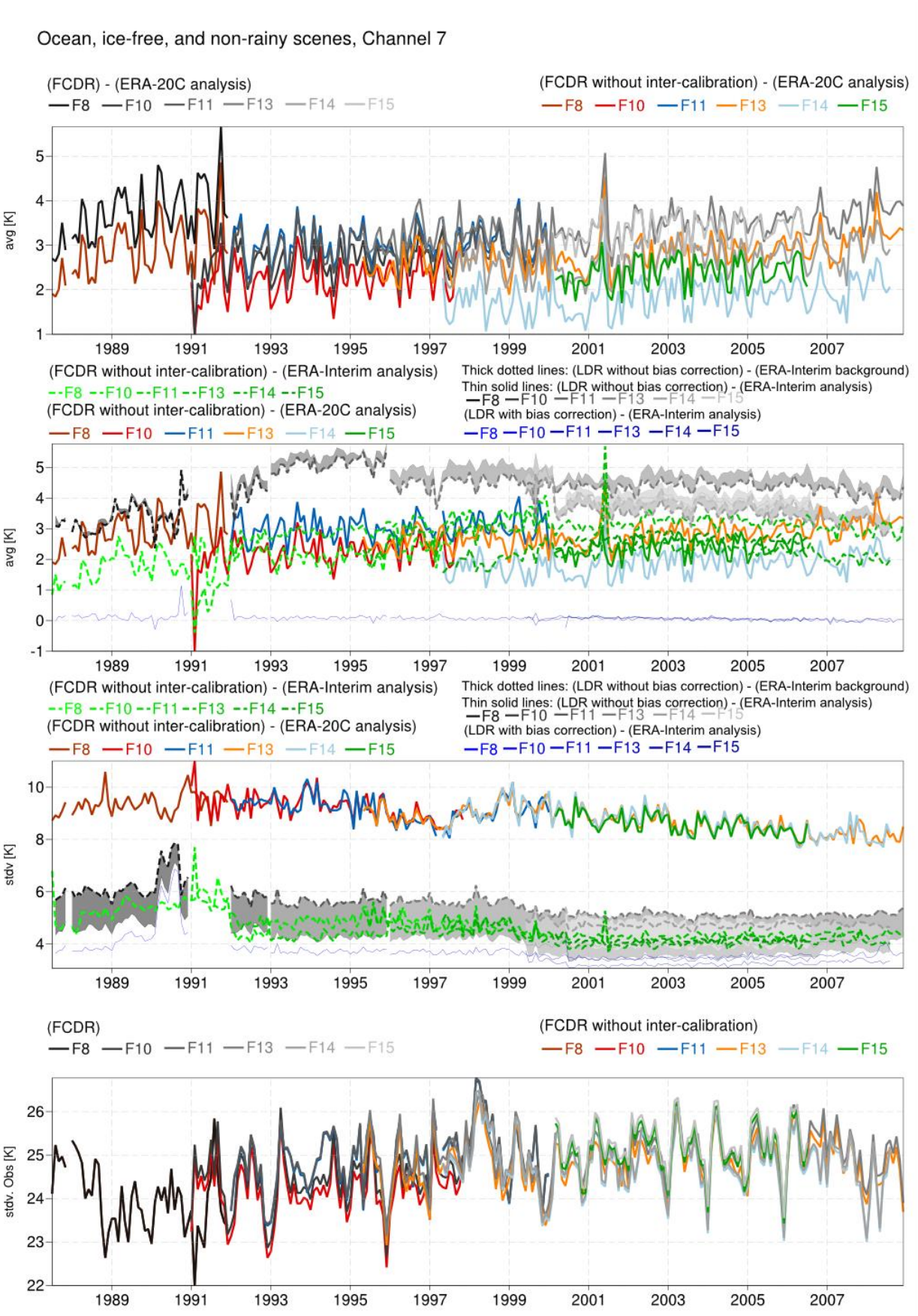

Figure 19: Same as previous figure, but for channel 7 (85H)

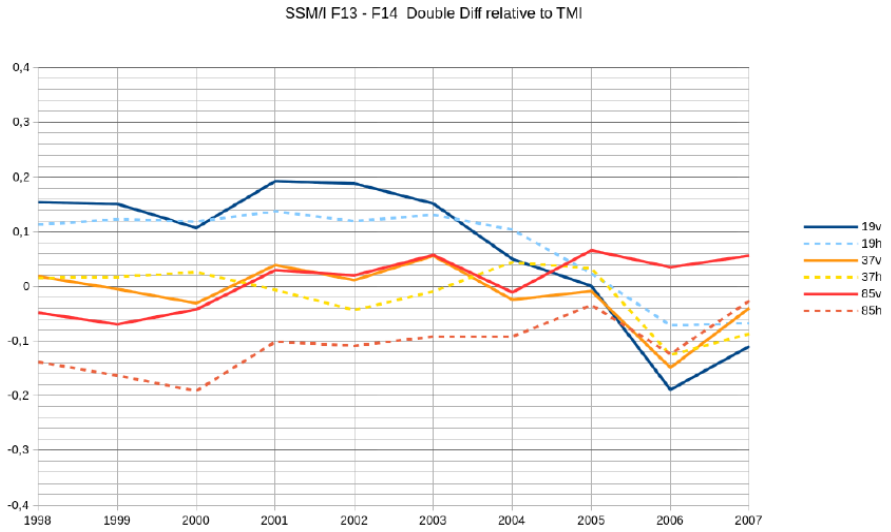

Figure 20: Timeseries of mean double differences, (F13-TMI) minus (F14-TMI), using collocated TMI observations as a reference.

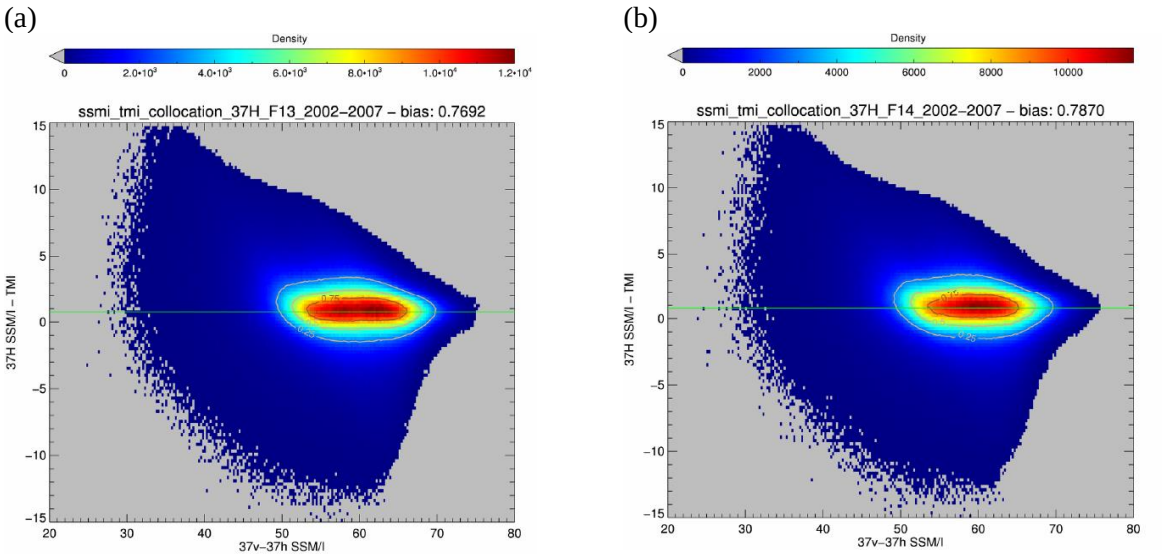

Figure 21: Density plots of single differences of collocated observations for channel 5 (37H). (a) F13 minus TMI, (b) F14 minus TMI

Another set of data found in the FCDR documents the health status of the sensors. This information is also read from the NetCDF4 files and saved in small metadata files, as documented in Table 7.

ODB column name	Contents	Unit or format	Origin	Range
date	Observation date	YYYYMMDD	Calculated from *'time'*	19870709-20081231
satellite_identifier@sat	WMO attribute		*'wmo_satellite_identifier'*	241, 243, 244, 246, 247, 248
channel	Channel number	integer	*'channel'*	1 to 7
rotation	instrument rotational speed	rpm	from '*rotation*'	F8: 31.595085/31.595875 F10: 31.595529/31.595985 F11: 31.595682/31.596106 F13: 31.595477/31.595991 F14: 31.595666/31.596164 F15: 31.595159/31.595319
trhl_var	variance of hot load temperature	K**2	from '*trhl_var*'	F8: 0.000009/0.000293 F10: 0.000894/0.001847 F11: 0.000083/0.000457 F13: 0.000066/0.002271 F14: 0.000319/0.000614 F15: 0.000765/0.001048
hotc_var	variance of hot load reading	count**2	from '*hotc_var*'	See Table 8
colc_var	variance of cold load reading	count**2	from '*colc_var*'	See Table 8
nedt	noise equivalent temperature	YYYYMMDD	from *'nedt'*	See Table 8

Table 7: List of parameter metadata read from the CM SAF FCDR NetCDF4 files and written into small ODB files

Channel	Satellite	Variance of hot load reading		Variance of cold load reading		Noise equivalent temperature	
		Min	Max	Min	Max	Min	Max
1	0.13	0.21	0.04	0.09	0.37	0.51	0.13
	0.17	0.22	0.06	0.08	0.52	0.57	0.17
	0.10	0.26	0.04	0.12	0.34	0.54	0.10
	0.11	0.21	0.03	0.15	0.37	0.59	0.11
	0.11	0.24	0.05	0.10	0.35	0.55	0.11
	0.18	0.25	0.06	0.09	0.48	0.54	0.18
2	0.12	0.50	0.04	0.08	0.36	0.73	0.12
	0.15	0.20	0.05	0.08	0.50	0.56	0.15
	0.10	0.20	0.04	0.08	0.29	0.45	0.10
	0.10	0.22	0.03	0.09	0.29	0.50	0.10
	0.16	0.25	0.07	0.10	0.42	0.59	0.16
	0.19	0.22	0.06	0.08	0.50	0.56	0.19
3	0.20	0.30	0.11	0.18	0.63	0.78	0.20
	0.32	0.39	0.13	0.16	0.57	0.63	0.32
	0.22	0.35	0.07	0.16	0.49	0.62	0.22
	0.23	0.32	0.11	0.16	0.50	0.65	0.23
	0.21	0.37	0.12	0.16	0.52	0.72	0.21
	0.23	0.26	0.10	0.12	0.60	0.65	0.23
4	0.06	0.11	0.03	0.07	0.28	0.44	0.06
	0.09	0.11	0.04	0.05	0.37	0.42	0.09
	0.07	0.10	0.02	0.05	0.28	0.39	0.07
	0.05	0.11	0.02	0.10	0.23	0.43	0.05
	0.08	0.11	0.04	0.07	0.31	0.43	0.08
	0.08	0.11	0.03	0.05	0.32	0.41	0.08
5	0.09	0.12	0.04	0.08	0.33	0.55	0.09
	0.09	0.11	0.04	0.05	0.38	0.45	0.09
	0.08	0.14	0.04	0.07	0.31	0.42	0.08
	0.06	0.13	0.01	0.06	0.28	0.44	0.06
	0.08	0.11	0.04	0.07	0.34	0.50	0.08
	0.08	0.10	0.03	0.04	0.37	0.41	0.08
6	0.00	0.69	0.00	0.61	-899999987309 02931225156687 84119808.0	38808264.00	0.00
	0.12	0.14	0.07	0.09	0.56	0.61	0.12
	0.11	0.15	0.07	0.11	0.56	0.69	0.11
	0.11	0.16	0.07	0.11	0.45	0.99	0.11
	0.12	0.16	0.08	0.11	0.50	0.65	0.12
	0.14	0.15	0.09	0.10	0.55	0.60	0.14
7	0.14	0.39	0.11	0.45	0.67	18476090.00	0.14
	0.12	0.17	0.08	0.11	0.57	0.65	0.12
	0.12	0.16	0.07	0.10	0.43	0.52	0.12
	0.12	0.17	0.07	0.12	0.46	1.59	0.12
	0.12	0.16	0.07	0.10	0.47	0.59	0.12
	0.12	0.15	0.08	0.10	0.61	0.68	0.12

Table 8: Range of metadata, sampling 1/15th of the FCDR

Timeseries of the noise equivalent temperatures are shown in Figure 22. Information shown for the 85GHz channels on F8 comes from the real data and does not represent an estimate of the noise in the synthetic data. This information is useful metadata for users to find out when the channel degraded (large numbers for these two channels on F8 also appears in Table 8).

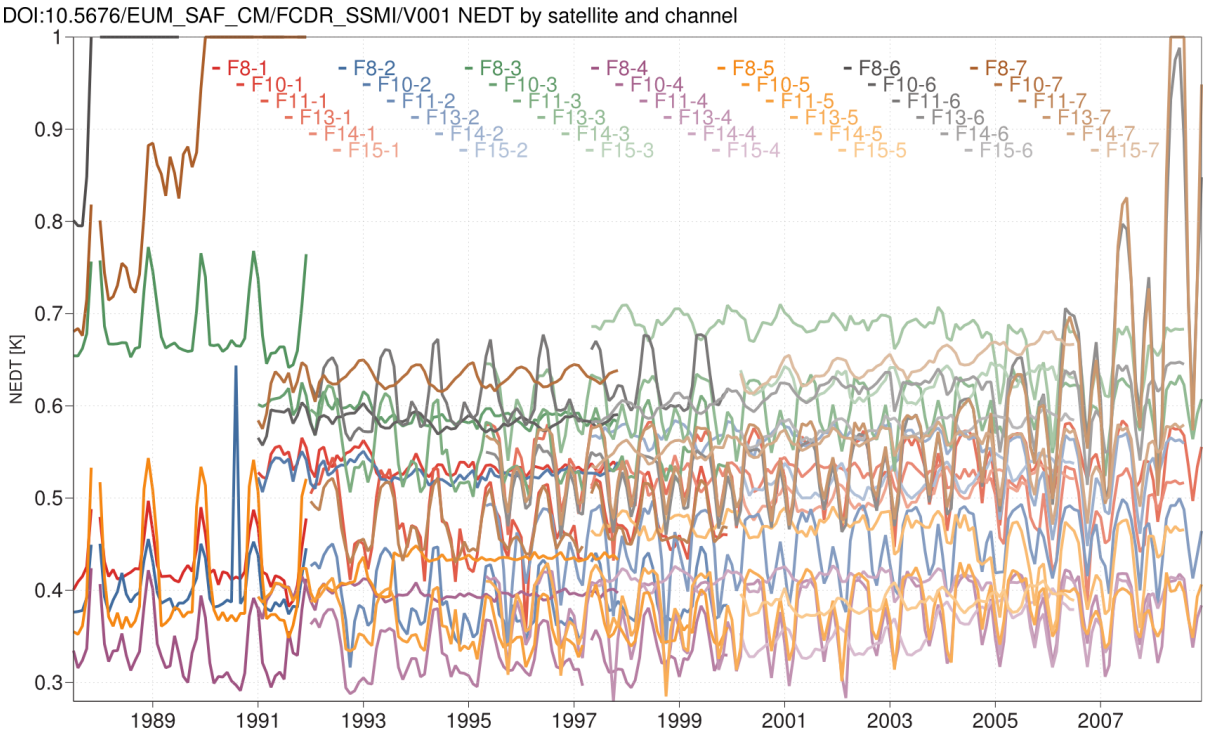

Figure 22: Timeseries of noise equivalent temperature for all channels and satellites

Also of interest is an apparent general tendency, for several channels, to feature increasing noise as the satellites age. This is most apparent for channels 6 and 7 on F13, the longest lasting satellite, but can be seen for other channels and satellites. This increase in noise, if confirmed, could mean that particular attention needs to be paid when the FCDR data are later exploited by noise filtering or data assimilation schemes. These typically assume some error characteristics. However, in IFS, the NEDT only contributes a small part of the observation error assumed by the data assimilation (on the order of 2 to 3 K for SSM/I), the rest being contributed by representativeness and observation operator errors.

7 Generalizing the comparisons of observational data records to gridded datasets: MIPs4Obs?

7.1 A shortcoming of Obs4MIPs, addressed with COSP

Since 1995, the World Climate Research Programme's (WCRP) Coupled Model Intercomparison Project (CMIP) has coordinated climate model experiments involving multiple international modelling teams. Through these experiments, climate modellers and scientists from around the world gained insights into the processes, mechanisms, and consequences of climate variability and climate change (Meehl *et al.*, 2014). The CMIP model experiments have routinely been the basis for future climate change assessments made by the Intergovernmental Panel on Climate Change (e.g., IPCC, 2013).

In support of the 5th phase of the CMIP (CMIP-5), the Observations for Model Intercomparison Projects (Obs4MIPs) initiative was launched, to support the analysis of the CMIPs with the provision of observational products. The observational products published under the Obs4MIPS need to match the CMIP protocol strictly "in terms of variables, temporal and spatial frequency, time period spanned, data format, and error or uncertainty estimates", as explained on the Earth System Commodity Governance website (https://www.earthsystemcog.org/projects/obs4mips/). This means that observation geometry and sampling are generally ignored. Beyond the technical requirements it is of utmost importance to verify that proposed observational products are useful for specific CMIP experiments. An overview of currently planned CMIP experiments is given by Meehl *et al.* (2014). The list of observational products that can be submitted to Obs4MIPS comprises monthly products that are used for classical assessments, but may also comprises several higher space-time resolution products that are intended to serve process studies. Obs4MIPs also facilitates the matching of satellite observations to model simulations through the use of simulated observational products. In the framework of CMIP, modelling groups emulate a selection of observational products online within their models or offline from the model outputs. The emulation process requires that the modellers implement an Observational Simulator Package, which simulates observational data and converts them to the mandatory CMIP format, regardless of the availability of corresponding observations at a particular date/time and location.

One approach that does take care of this particular sampling problem is the CF-MIP Observational Simulator Package (COSP; Bodas-Salcedo *et al.*, 2011), which can be operated in two modes. The 2D mode is typically used to compare model outputs and ISCCP, without space-time collocation, as in Obs4MIPs. There is in addition a 1D mode, where so-called curtain files are used to achieve a space-time collocation between model and observational data. In both cases the comparison can be conducted in observation (*e.g.*, radiance) or in geophysical parameter space. Although the latter is an effective manner of matching model simulations with satellite observations, the process of each of the 30 CMIP modelling groups implementing COSP is resource-consuming. Additionally, in the 2D mode, most of the data produced by these simulators do not have, and will never have, an observational equivalent to compare with, for the very reason that observations always have limited spatio-temporal coverage.

7.2 An example of Obs4MIPs-like comparison

Starting from the Obs4MIPs concept of projecting observations into analogues "in terms of variables, temporal and spatial frequency, and periods" to the climate model outputs (Teixeira *et al.*, 2014), we give now an example of such an approach. We apply here what Teixeira *et al.* call a 'Stage II' comparison: we compare total column water vapor (TCWV) retrieved by various groups (RSS and HOAPS) with model integration and reanalysis datasets. We consider four observational products, one climate model integration (ERA-20CM, described by Hersbach *et al.*, 2015), and three reanalysis products (ERA-Interim and ERA-

20C as described earlier, and JRA-55 described by Kobayashi *et al.*, 2015). Figure 23a shows monthly means of TCWV over the tropical oceans for January 1988-June 2014. allows to spot offsets or systematic differences between the products, and to appreciate that some products have more pronounced annual cycles than others. We show for completeness ERA-20C analyses and 3-hour forecasts (or backgrounds) but there is essentially no difference between them for the scales shown here. ERA-20C and ERA-20CM stand out as missing approximately 2 mm of TCWV, according to the observational products. This is consistent with a long-standing, known bias of the IFS model, on which much progress has been achieved in the most recent release of the ECMWF model (Peter Bechtold, *personal communication*). JRA-55 is also on the lower side of the various observational product estimates. Differences between the various observational products are larger for the earlier-generation products shown here (RSSv6 and HOAPSv3.1) than for the more recent ones (RSSv7 and HOAPSv3.2), which agree better between one another. One notes, however, that RSSv7 features a slightly higher trend than HOAPSv3.2, although still within variability. The annual cycle appears too weak in ERA-20C (and ERA-20CM). This may be another manifestation of the model bias, but could reflect the choice of SST forcing: both ERA-20CM and ERA-20C use monthly SSTs, which may thus miss out on the yearly minima and maxima as well as intra-month variability as noted earlier when comparing to the observation variability, respecting the observation sampling.

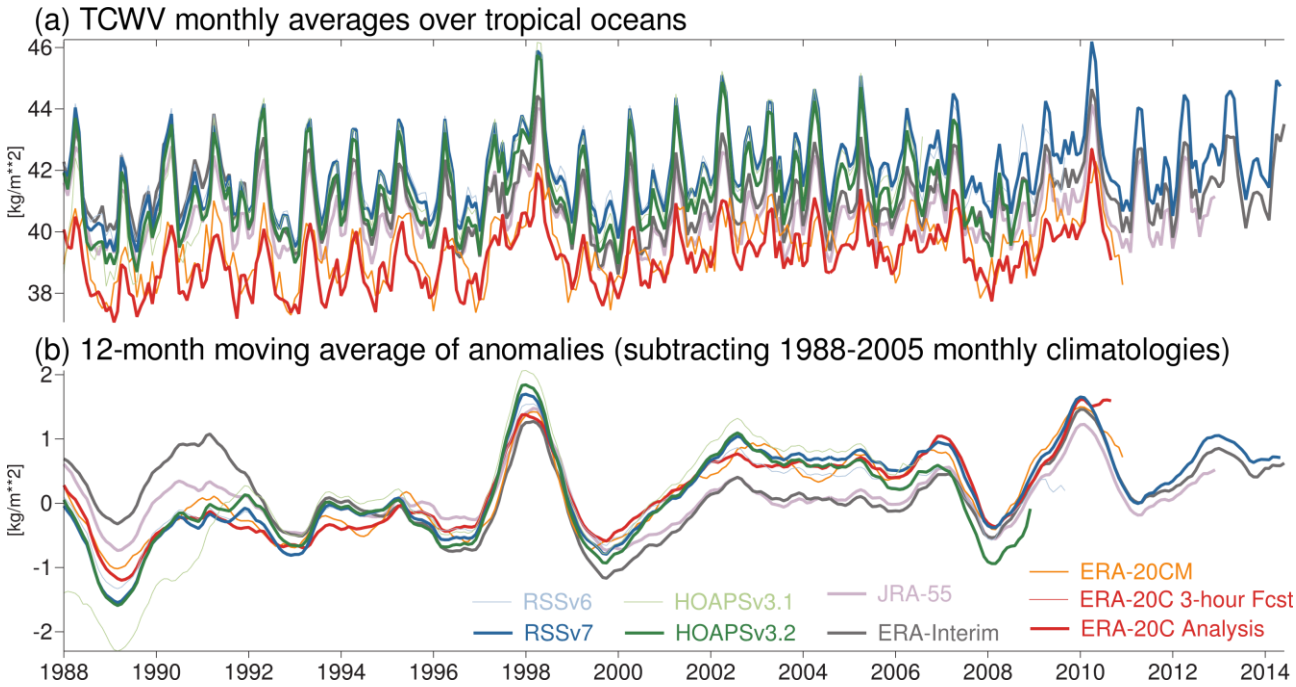

Figure 23: Time series of total column water vapor monthly anomalies over tropical ocean (latitudes 20S-20N) from several reanalyses (ERA-Interim, JRA-55, ERA-20C), one model-only integration (ERA-20CM, i.e., AMIP-type), and 4 observational products: Remote Sensing Systems (RSS) v6 and v7, and Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data (HOAPS) v3.1 and v3.2

Figure 23b shows anomalies, and indicate discrepancies in the low-frequency variability between the various estimates. The known problems in the TCWV continuity of ERA-Interim are somewhat visible with a sudden drop in 1992 (drying was introduced whenever additional SSM/I data were assimilated using the 1D+4D rain assimilation, Geer *et al.*, 2008), going from more moist than RSS and HOAPS pre-1992 to drier than them in the 2000s, recovering approximately 2009 to the level of RSS, as SSM/I 1D+4D rain assimilation ceases in ERA-Interim. These problems are absent from ERA-20C(M). JRA-55 also appears to have some issues in the late 1980s (more moist than RSS and HOAPS) and the 2000s (possibly too dry).

Overall these comparisons are quite telling, with ERA-20C(M) featuring the low frequency closest to the observational products and one may expect that Obs4MIPs will facilitate more detailed studies of this type, drilling down into weather types and events and modes of variability, to improve understanding of climate model shortcomings.

7.3 What could we learn from a comparison in observation space?

We present now an example of assessment in the observation space of SSM/I. Figure 24a shows the mean brightness temperatures collected by 3 satellites over the Eastern Tropical Pacific (latitudes 1.5 to 2.5 degrees North, longitudes 135 to 90 degrees West), after applying a high-pass filtering in the longitudinal direction (subtracting the longitudinal moving average with a window of 12 degrees), and a temporal filtering to retain the 20 to 50 day signal. The anomalies extracted in this fashion suggest waves in the longitude and time domain (the vertical axis shows day of the year 1996). These are possibly manifestations of tropical waves, propagating westward. The calculations from ERA-Interim, all at the same sample as the observations, and filtered in the same way, show in Figure 24b much reduced traces of such oscillations. However, ERA-20C, in Figure 24c, completely misses such waves. A likely explanation is the monthly SST forcing used as input. Carrying out the same type of investigation by applying RTTOV to climate model outputs could help better understand whether this variability is properly represented in those models.

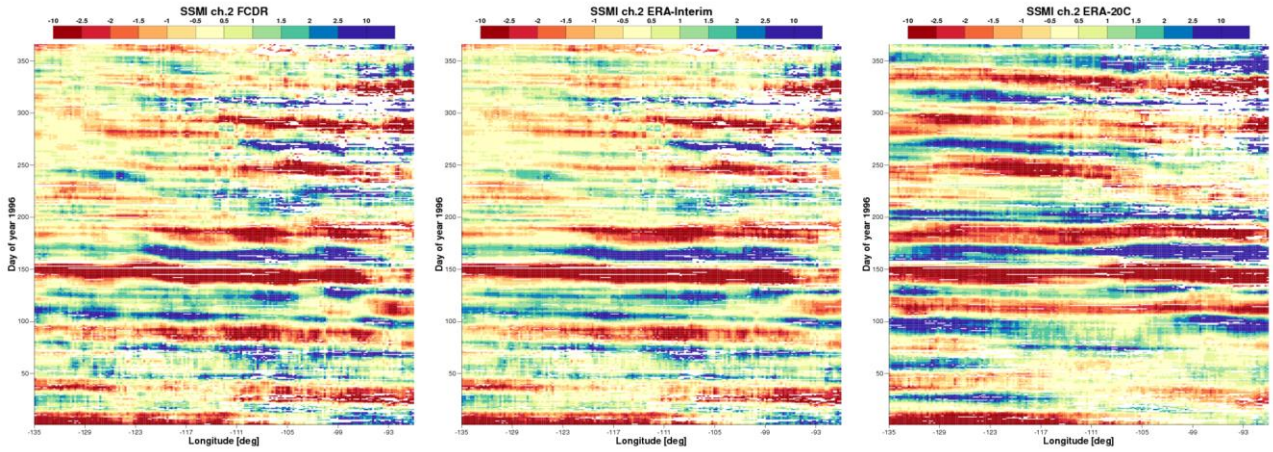

Figure 24: (left) Average SSM/I channel 2 brightness temperature from the FCDR for F10, F11, and F13, at 0.25 degree longitude resolution (horizontal axis) and per day of the year 1996 (vertical axis, from 1 January at the bottom to 31 December at the top) in the Eastern Tropical Pacific, subject to spatial and temporal filtering described in the text; (middle) Simulations from ERA-Interim at the same times and locations as the observations, using the methodology described in the present report; (right) Simulations from ERA-20C

7.4 Considerations for a generalized MIP4Obs

The fact that so far mainly satellite datasets and only sample *in situ* data have been covered by Obs4MIPs stems from the deceptively apparent complete spatial coverage offered by satellite data. Yet, even polar-orbiting satellite data are not temporally complete for they only partially sample the diurnal cycle. This focus on seemingly complete datasets sets from the beginning a limit to the level of details that one can exploit when the comparisons are restricted to the use of geophysical variables provided on a regular grid at a daily frequency. Extending the Obs4MIP principle to observation space would in fact go against several foundations of Obs4MIPs, such as the typology of allowed variables, all being geophysical ones, not

instrument-sensor dependent such as brightness temperatures or atmospheric delays, and the reference to gridded products throughout layers of Obs4MIPs documentation, file formats, and data model.

The strength of the Obs4MIPs protocol is its focus on defining consistently the metadata for the model outputs. However, a limitation is that the Obs4MIPS protocol only succinctly defines the metadata for the observations, since these are supposed to resemble the model outputs. Since the debut of the World Meteorological Organization (WMO) the definition of observation metadata has been the focus of working groups, for example to define standards for reporting and exchanging observations. Such prior investment in the definition of observation data attributes should indeed form the basis for any comparison into observation space. In short, it would make equal sense to also map model output comparisons into observational data records by simply augmenting them with model equivalents.

In fact, mapping the model simulations and reanalysis gridded data to the observation space, location, and time, presents a complementary approach to Obs4MIPs. The prospect of augmenting existing FCDRs (which are quite limited in numbers for a given geophysical variable) by equivalents coming from the largely superior number of model runs, should enable direct comparison between all models at similar times and locations, along with an observational quantity that is unaltered or of a higher standard (*e.g.*, if reprocessed as an FCDR). Such approach deserves being considered before all modelling groups invest into generating gridded datasets containing observation-like quantities.

For all the reasons mentioned above, it would make sense to consider a reverse protocol, tentatively named ‘MIPs4Obs’, that could offer a complementary framework of comparisons to Obs4MIPS. Within such a MIPS4Obs, the model outputs would simply be made into analogues ‘in terms of variables, temporal and spatial frequency, and periods’ to observational products, respecting the observation sampling and geometry.

The approach presented in this report can be seen as a step into this direction, by projecting gridded data into analogues of a FCDR.

8 Conclusions

The present report assesses the added value of the CM SAF SSM/I of brightness temperatures over the existing data record currently available at ECMWF, hereafter called legacy data record (LDR). This assessment is carried out outside the assimilation system, thereby proposing methodologies to compare observational data records (LDR and FCDR), and to compare a data record with reanalysis datasets (ERA-Interim and ERA-20C). Such methodologies could be similarly applied to other data records of the same variable, such as the SSM/I brightness temperature FCDRs from CSU and from RSS, but also to other observational data records, and to other gridded datasets such as climate model outputs.

Regarding the comparison of the LDR to the FCDR, the geographical coverage of the FCDR is found to match the coverage of the legacy SSM/I data available for assimilation in ERA-Interim for a selected date. However, as expected, the temporal coverage is much improved, patching gaps in the LDR, and adding more than 13 satellite-years (mostly because in the LDR only one DMSP satellite is available at any given time until 1999). The FCDR also appears to contain better ancillary information regarding sea-ice, which is an important factor affecting pre-screening quality control. The detailed matching procedure illustrated in the report is also able to infer that the FCDR for a given date and a given satellite is shifted by +1 second in time and 7 km in distance compared to the original record (LDR); this information appears valid given the dates of introduction of leap seconds and the velocity of DMSP satellites.

The tools developed for mapping the gridded datasets into observation space build on earlier efforts such as a comparison of prior reanalyses to Nimbus-4 interferometer observation data (Poli, 2013). The principles are generic enough that they could be applied to any observational data record, provided an observation operator exists. The simulation tool relies largely on RTTOV and is validated by comparing its results in a controlled simulation environment with those found in the ERA-Interim observation feedback (assuming the latter result from a correct implementation of RTTOV in IFS). The validation notes that robust offsets exist between the offline and online simulation results, owing to the use of different emissivity models (ERA-Interim used an older version of FASTEM); other factors such as internal interpolations are assumed to be causing the rest of the differences.

Problems related to inaccurate geolocation of the SSM/I pixels have been addressed in the CM SAF FCDR. Maps shown in this report suggest that this geolocation may be further improved in a future release of the FCDR. Mapping reanalysis fields to the observations (space, date, time, and location) allows to compare data records with identical spatio-temporal coverage, and removes all discussion on how different sampling affects the comparison. For the 22 GHz channel (channel 3) sensitive to total column water vapor, the SSM/I FCDR is found to conform to the independent reanalysis (ERA-20C), explaining 90% of the variability in the data. This result is verified by confirmed by comparing means of TCWV with observational products (RSS and HOAPS), where ERA-20C matches better than ERA-Interim these datasets. One missing element of variability in ERA-20C water cycle is suggested to be in tropical waves, possibly because of the monthly SST employed as forcing.

Comparing the FCDR to ERA-Interim yields a lower standard deviation than when comparing to ERA-20C (approximately 3 K instead of from 6-7 K). This is found to be consistent with the assimilation statistics from the ERA-Interim observation feedback, which assimilated the LDR. In the mean, the departures between the FCDR and ERA-20C are more stable over time than those between the FCDR and ERA-Interim, which suffers from discontinuities in its representation of the water cycle, as also verified here by considering inter-annual variability of TCWV. This indicates that ERA-20C could be considered for acting as a possible transfer standard for observation reprocessing work.

The FCDR intercalibration reduces largely mean satellite-to-satellite differences and low-frequency variability when compared to reanalyses. However, this has the cost of decreasing slightly the intra-month variability within the data.

The choice of the reference satellite for intercalibration, F11, is visible from most timeseries. For one channel (85V, channel 6) this choice is interesting for it essentially increases (in the mean) the brightness temperatures from all other satellites. This brings up the difficult question of how to optimally choose a reference for intercalibration.

Close inspection of the residual statistics from the ERA-Interim assimilation illustrates the dilemma in reanalysis for choosing an observation input. Using an intercalibrated data record as input presents the reanalysis with observations which are more consistent over time. However, the data thus brought forward for assimilation contain structural errors (correlated, slow-varying errors) which pose problem to the current variational bias correction schemes. When the LDR shifts from the RSS SSM/I record to the GTS record, one observes a sharp reduction in the standard deviation of the assimilation residuals, especially after bias correction. In the light of this, as well as the difficult choice of the reference satellite and the intra-month variability lost during intercalibration, it seems advisable, before ERA-Interim replacement (ERA5) enters production, to run dedicated assimilation experiments of the FCDR with and without the intercalibration applied. Key metrics to monitor should include the assimilation bias estimates inferred by the variational bias correction scheme, in the context of all other observations and forcings. On that note, the two 85 GHz channels (channels 6 and 7) from the F8 satellite should probably not be assimilated into global reanalyses.

The present pre-assimilation feedback is meant to be followed up by assimilation experiments in reanalysis. Yet, the lessons learnt with a comparison outside such a system, but using the tools and methods presented here, indicate that such comparison approach could assist in quality control and assessment of CDRs.

The report proposes that the approach presented here could be seen as a stepping stone to a MIPs4Obs initiative, to complement Obs4MIPs. The novelty would be to build upon the existing standards and metadata definitions of the observing community and then provide to observational data records analogues computed from model outputs. As shown in this report, this could enable to spot data problems and to better understand the observational record, as well as note deficiencies in the various gridded datasets (climate model outputs or reanalyses). By virtue of a projection that would mirror Obs4MIPs, the results could then further feed modelling developments and planning of observation campaigns. In fact, such an approach has started in one particular observational community: ICOADS will adopt a value-added database framework in the near future (Smith *et al.*, 2011), though it is intended for the 'improved datasets' that have used ICOADS, and not really for CMIP comparisons. The more general MIPs4Obs approach proposed here would be inclusive of CMIP comparisons. Whereas the concept of Obs4MIPs mostly limits it to observational data offering low-frequency variability information, MIP4Obs would be inclusive of measurement campaign experiments, or similarly short observational data records.

Acknowledgements

The authors thank the CM SAF for provision of the SSM/I FCDR and the NWP-SAF for provision of RTTOV11. Work conducted in this report was partially funded by the ECMWF reanalysis project, EU FP7 ERA-CLIM2 Grant Agreement no. 607029, EU FP7 CORE-CLIMAX Grant Agreement no. 313085, and benefited from fruitful exchanges at the 6th EUMETSAT's Working Group on Data Record Generation. We thank Dick Dee (ECMWF) for comments that helped improve this document.

References

Berg W, Sapiano MRP, Horsman J, Kummerow C. 2013: Improved Geolocation and Earth Incidence Angle Information for a Fundamental Climate Data Record of the SSM/I Sensors. *IEEE Trans. Geosci. Remote Sens.*, **51**(3):1504--1513. DOI: 10.1109/TGRS.2012.2199761

Bodas-Salcedo A, and Co-authors, 2011: COSP: Satellite simulation software for model assessment, *Bull. Amer. Meteor. Soc.*, **92**: 1023--1043. DOI:10.1175/2011BAMS2856.1

CM SAF (EUMETSAT Satellite Application Facility on Climate Monitoring). 2013: Algorithm Theoretical Basis Document for Fundamental Climate Data Record of SSM/I Brightness Temperatures. DOI: 10.5676/EUM_SAF_CM/FCDR_SSMI/V001.

CM SAF (EUMETSAT Satellite Application Facility on Climate Monitoring). 2013: Product User Manual for Fundamental Climate Data Record of SSM/I Brightness Temperatures. DOI: 10.5676/EUM_SAF_CM/FCDR_SSMI/V001

Dee DP, and Co-authors. 2011: The ERA-Interim reanalysis: configuration and performance of the data assimilation system. *Q.J.R. Meteorol. Soc.*, **137**: 553--597. DOI: 10.1002/qj.828

ECMWF. 2007. IFS documentation cy31r1. Available from ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom. Online at http://old.ecmwf.int/research/ifsdocs/CY31r1/index.html

Fennig K, Andersson A, Schröder M. 2013: Fundamental Climate Data Record of SSM/I Brightness Temperatures. Satellite Application Facility on Climate Monitoring, DOI:10.5676/EUM_SAF_CM/FCDR_SSMI/V001

Geer AJ, Bauer P, Lopez P. 2008: Lessons learnt from the operational 1D + 4D-Var assimilation of rain- and cloud-affected SSM/I observations at ECMWF. *Q.J.R. Meteorol. Soc.*, **134**: 1513--1525. DOI: 10.1002/qj.304

Hersbach H, Peubey C, Simmons A, Berrisford P, Poli, P, Dee D. 2015: ERA-20CM: a twentieth-century atmospheric model ensemble. *Q.J.R. Meteorol. Soc.* doi: 10.1002/qj.2528

Intergovernmental Panel on Climate Change (IPCC). 2013: Climate Change 2013: The Physical Science Basis, Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change. Edited by T. F. Stocker et al., Cambridge University Press, Cambridge, U.K.

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

Meehl GA, Moss R, Taylor KE, Eyring V, Stouffer RJ, Bony S, Stevens B. 2014: Climate model intercomparisons: preparing for the next phase. *Eos Trans Am Geophys Union*, **95(9):**77–78. DOI: 10.1002/2014EO090001

Poli P, and Co-authors. 2013: The data assimilation system and initial performance evaluation of the ECMWF pilot reanalysis of the 20th-century assimilating surface observations only (ERA-20C). ERA Report Series 14, September 2013, 59 pp., available from ECMWF, Shinfield Park, Reading. Online at http://old.ecmwf.int/publications/library/do/references/list/782009

Poli P. 2013: Use of hyperspectral infrared satellite observations for climate reanalyses. ECMWF / EUMETSAT NWP-SAF workshop on efficient representation of hyperspectral infrared satellite observations, ECMWF, Shinfield Park, Reading RG2 9AX, United Kingdom, 5-7 November 2013. Presentation available from http://old.ecmwf.int/newsevents/meetings/workshops/2013/NWP-SAF_satellite_observations/presentations/index.html

Reynolds RW, Rayner NA, Smith TM, Stokes DC, Wang W. 2002: An improved in situ and satellite SST analysis for climate. *J. Clim.* **15**(13): 1609--1625

Sapiano M, Berg W, McKague D, Kummerow C. 2013: Towards an Intercalibrated Fundamental Climate Data Record of the SSM/I Sensors, *IEEE Trans. Geosci. Rem. Sens.*, **51**, 1492-1503 DOI: 10.1109/TGRS.2012.2206601

Saunders R, Hocking J, Rundle D, Rayer P, Matricardi M, Geer A, Lupu C, Brunel P, Vidot J. 2013: RTTOV-11 science and validation report. EUMETSAT NWP SAF, 62pp. available from http://nwpsaf.eu/deliverables/rtm/rtm_rttov11.html

Smith SR, Bourassa MA, Woodruff S, Worley S, Kent E, Josey S, Rayner N, Freeman E, Reynolds R. 2011: Developing an ICOADS Value-Added Database (IVAD) to Support Climate Research. Available from http://icoads.noaa.gov/ivad/IVAD_concept_plan_final.pdf

Teixeira J, Waliser D, Ferraro R, Gleckler P, Lee T, Potter G. 2014: Satellite Observations for CMIP5: The Genesis of Obs4MIPs. *Bull. Amer. Meteor. Soc.*, early online release DOI: 10.1175/BAMS-D-12-00204.1

Titchner HA, Rayner NA. 2014: The Met Office Hadley Centre sea ice and sea-surface temperature data set, version 2: 1. sea ice concentrations. *J. Geophys. Res.* **119**:2864--2889, DOI: 10.1002/2013JD020316

Wang J, Kwok R, Saucier FJ, Hutchings J, Ikeda M, Hibler W Haapala J, Coon MD, Meier HEM, Eicken H, Tanaka N, Prentki D, Johnson W. 2015: Working Toward Improved Small-scale Sea Ice-Ocean Modeling in the Arctic Seas. *Eos* **84** (34)**:** 325–336

Wentz FJ. 2013: SSM/I Version-7 Calibration Report. Rep no 011012, Remote Sensing Systems, Santa Rosa, California, USA, 46pp. available from http://images.remss.com/papers/rsstech/2012_011012_Wentz_Version-7_SSMI_Calibration.pdf