

D4_6 Metrological assessment of consistency, stability and uncertainty of FIDUCEO FCDRs

J. Mittaz, G. Holl, I. Hans, M. Burgdorf, F. Ruethrich, Y. Govaerts, C. Merchant, R. Phipps, R. Roebeling,
V. O. John, M. Taylor, J. Mollard

FIDUCEO Consortium

30 August 2019



FIDUCEO has received funding from the European Union's Horizon 2020 Programme for Research and Innovation, under Grant Agreement no. 638822

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

Contents

1	Introduction	6
1.1	Scope.....	6
1.2	Version Control	6
1.3	Applicable and Reference Documents.....	6
1.4	Glossary.....	7
1.5	Executive Summary.....	8
2	MVIRI:.....	10
2.1	The FIDUCEO MVIRI FCDR.....	10
2.2	Data	10
2.2.1	MVIRI VIS Observations	10
2.2.2	SEVIRI Measurements.....	11
2.2.3	SCIAMACHY Measurements.....	12
2.3	Assessment Methodology.....	12
2.3.1	Expected performance of the FCDR.....	12
2.3.2	Harmonised time-series of reflectance and uncertainty	13
2.3.3	Homogenised time series and stability analysis	13
2.3.4	Comparison against SEVIRI	15
2.4	Assessment Results and Discussion	16
2.4.1	Expected performance of the FCDR.....	16
2.4.2	Harmonised time series of reflectance and uncertainty	17
2.4.3	Homogenised time series and stability analysis	18
2.4.4	Comparison against SEVIRI	20
2.5	Summary and Conclusions.....	22
3	Microwave sounders.....	23
3.1	Brief intro on what the FIDUCEO FCDR is	23
3.2	Method of validation/ reference sensors used/ new methods used	23
3.2.1	Attempts to harmonise MW data with the FIDUCEO harmonisation method.....	24
3.3	Mutual/ relative consistency within sensor series	32
3.3.1	Assessment of uncertainties	32
3.3.2	Assessment of stability and consistency	35

D4_6 Metrological assessment of consistency, **V1.1** stability and uncertainty of FIDUCEO FCDRs

3.4	Positives of FIDUCEO methods/ FCDR	41
3.5	Summary	41
4	HIRS	42
4.1	Introduction	42
4.2	Harmonisation of HIRS FCDR	42
4.3	Summary	46
5	AVHRR	47
5.1	Introduction	47
5.2	New methods used	47
5.3	Consistency of the FIDUCEO AVHRR FCDR	48
5.4	Overall stability of the FIDUCEO AVHRR FCDR	51
5.4.1	The 11µm and 12µm channels.....	52
5.4.2	The 3.7µm channel.....	53
5.5	Summary	54
5.6	Positives of FIDUCEO methods and FCDR.....	55

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

Table of Figures

Figure 1: Spectral Response functions of Meteosat-2-7 plotted together with SCanning Imaging Absorption spectroMeter for Atmospheric CartographY (SCIAMACHY) spectra acquired during 2002 at three target sites. a) shows the spectra measured by SCIAMACHY at the Algeria site, b) shows the spectra measured at the Nile-delta and the c) shows those measured at the Atlantic-1 site. Spectra are plotted in transparent black to better illustrate their spread. Spectra with strong cloud contamination were removed before plotting (see Section 3.9.1). Note that for band-adjustment/homogenisation more sophisticated filtering regarding clouds and scene heterogeneity is applied.	12
Figure 2: Convolved clear-sky SCIAMACHY spectra above Algeria-3 using Meteosat-2 and Meteosat-5 SRFs that were valid at each satellites' launch date.	14
Figure 3: Mean clear sky SCIAMACHY spectrum above the Algeria-3 site with standard deviation and with examples of the MVIRI and SEVIRI spectral response functions that are used for convolution.	15
Figure 4: Theoretically expected time series of the TOA reflectance derived from 6 Meteosat instruments at two reference sites using the (old) pre-launch measured SRFs and the (new) reconstructed, spectrally degrading SRFs. Grey shading indicates the combined standard uncertainty introduced by the SRFs.	16
Figure 5: Recalibrated MVIRI clear sky reflectance time series and uncertainties for the Algeria-3 site, calibrated with the reconstructed, in-flight characterised SRF.	17
Figure 6: Recalibrated MVIRI clear sky reflectance time series and uncertainties for the Atlantic-1 site, calibrated with the reconstructed, in-flight characterised SRF.	17
Figure 7: Homogenised time series of recalibrated broadband reflectances at Algeria-3, Nile-delta and Atlantic-1.....	19
Figure 8: Anomalies of the homogenised reflectance time series and their trends. Anomalies are the deviation from the mean annual reflectance cycles. Additional filtering for cloud contamination was applied. The filter computes the 25th percentile from a rolling kernel of 30 days around each reflectance measurement. Measurements that are brighter than this value are rejected in order to display only measurements that are certainly cloud free. Note that for the Atlantic site the periods with globally elevated aerosol loads due to volcanic eruptions are excluded from the stability analysis.	19
Figure 9: Histogram of SEVIRI HRVIS reflectance plotted as reference together with the histograms of the operational MVIRI reflectance and the recalibrated MVIRI reflectance, as obtained from cloud-free Algeria-3 pixels at 12:00 UTC slots during March 2005. The MVIRI datasets are band adjusted to the SRF of the SEVIRI HRVIS band according to the given Spectral Band Adjustment Factors (SBAFs). The SBAFs are computed using the same set of SCIAMACHY spectra and the difference between both SBAFs is entirely due to the different shapes of the MVIRI SRFs in the operational and the recalibrated datasets.....	20
Figure 10: Histogram of SEVIRI HRVIS reflectance plotted as reference together with the histograms of the operational MVIRI reflectance and the recalibrated MVIRI reflectance, as obtained from entirely cloud-covered pixels at 12:00 UTC slots during March 2005. Clouds are classified into high clouds (a), middle clouds (b) and low clouds (c). The MVIRI datasets were band adjusted to the SRF of the SEVIRI HRVIS band according to the given SBAFs. For each cloud-class the SBAFs for the operational dataset and for the recalibrated dataset are computed using the same set of SCIAMACHY spectra and the difference between both SBAFs is entirely due to the different shapes of the SRFs in the operational and recalibrated datasets.	21
Figure 11: Histogram of SEVIRI HRVIS reflectance plotted as reference together with the histograms of the operational MVIRI reflectance and the recalibrated MVIRI reflectance, as obtained from cloud-free Atlantic-	

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

1 pixels at 12:00 UTC slots during March 2005. The MVIRI datasets were band adjusted to the SRF of the SEVIRI HRVIS band according to the given SBAFs. The SBAFs are computed using the same set of SCIAMACHY spectra and the difference between both SBAFs is entirely due to the different shapes of the SRFs in the operational and the recalibrated datasets.	22
Figure 12 Difference in brightness temperature (channel 3) of FIDUCEO and AAPP processing for the first half of the example orbit of Metop-B.	23
Figure 13 Histogram of brightness temperature differences (K-residual) between MHS on MetOp-A and on NOAA-18, channel 3, before (left panel) and after (right panel) harmonisation.	25
Figure 14 Histogram of brightness temperature differences (K-residual) between AMSU-B on NOAA-16 and MHS on NOAA-18, channel 18, or 3 resp., before (left panel) and after (right panel) harmonisation.	26
Figure 15 Comparison of resulting biases for AAPP and harmonised FCDR processing for MHS on MetOp-A, channel 3 against MHS on NOAA-18 (for September 2014 - February 2015). Harmonisation parameters are non-linearity and warm target correction. The uncertainty bars represent the standard deviation of the bias over the six months of used data.	26
Figure 16 Abundance pattern of brightness temperatures for FIDUCEO match-ups of MHS on Metop-A and on NOAA-18.	27
Figure 17 As Figure 15, but the harmonisation parameters are non-linearity, warm target correction, polarisation correction and cold target correction. The uncertainty bars represent the standard deviation of the bias over the six months of used data.	28
Figure 18 Comparison of 10 th to 90 th percentile of brightness temperature in the matchup data before harmonisation (AAPP) and after harmonisation. The matchup data contain only near-nadir views. The four harmonisation parameters are non-linearity, warm target correction, polarisation correction and cold target correction.	29
Figure 19 As Figure 15, but for near-nadir views only. Note the difference to and the similarity to Figure...30	
Figure 20 Before and after harmonization: The differences X – NOAA18 with X = NOAA17, METOPA, and METOPB, are plotted as a function of time. Mean values for all matchups in intervals of 1.5 months were calculated, and then the time series was smoothed.	32
Figure 21 Channel 1: Brightness temperature and its uncertainties (for an example orbit of MHS on Metop-B in July 2015). Upper left: brightness temperature. Upper right: uncertainty of brightness temperature due to independent effects. Lower left: uncertainty of brightness temperature due to structured effects. Lower right: uncertainty of brightness temperature due to common effects.	34
Figure 22 As for Figure 21, but for channel 3.	34
Figure 23 Histogram of the three classes of uncertainties for September 2014, MHS on Metop-B, all channels.	35
Figure 24 Comparing inter-satellite biases of MHS and AMSU-B instruments against MHS on NOAA-18 in operational AAPP data (left) and FCDR data (right) as function of time. The FCDR provides more consistent and stable time series due to improved calibration. The instruments agree within the uncertainties (only uncertainties due to common effects displayed here). Note that channel 2 differs slightly in frequency for MHS and AMSU-B and therefore are not expected to match completely.	37
Figure 25 Comparing the time series of brightness temperature in channel 1 for all MW instruments in operational data (NGDC for SSMT-2 and AAPP processed data for AMSU-B and MHS) and in FCDR data as function of time. The FCDR provides more consistent and stable time series due to improved calibration.	

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

For the FCDR, the shaded regions denote the uncertainty due to common effects (hardly visible for channel 1 on this scale). There is no corresponding uncertainty estimate on the operational data.	38
Figure 26 As figure 25, but for channel 2. Note the uncertainty for NOAA-15 due to uncorrected RFI.	38
Figure 27 As Figure 25, but for channel 3. The instruments agree within the uncertainties. A distinct improvement is also achieved for SSMT2.....	39
Figure 28 As Figure 25, but for channel 4. Note that NOAA-15 still shows differences with respect to the other instruments. This is due to the imperfect RFI correction.	39
Figure 29 As Figure 25, but for channel 5.	40
Figure 30 Comparison of resulting biases for AAPP with and without harmonization for MHS on MetOp-A, channel 3, against MHS on NOAA-18 (for November 2006 - September 2010). 121,830 matchups between these instruments were used to calculate the bias, so in most cases the uncertainties due to random scatter of the differences between MetOp-A and NOAA-18 are smaller than the thickness of the lines.	40

Table of Tables

Table 1: Spatial and spectral characteristics of Meteosat Visible Infra-Red Imager (MVIRI) visible (VIS), thermal infrared (TIR) and water vapour (WV) channels.	11
Table 2: Spatial and spectral characteristics of the four Spinning Enhanced Visible Infra-Red Imager (SEVIRI) visible (VIS) and near infrared (NIR) channels. HRVIS: high-resolution visible.	11
Table 3: Evaluation sites where recalibrated (harmonised) and band-adjusted (homogenised) time series were generated along with spectral characteristics and thresholds for the filtering of SCIAMACHY spectra.	13
Table 4: Characterisation of the sites used for the comparison of the SEVIRI HRVIS dataset with the operational MVIRI dataset and the harmonised MVIRI FCDR. The sites and criteria are used for the extraction of suited SCIAMACHY spectra for the homogenisation and for the extraction of SEVIRI/MVIRI observations.....	15
Table 5: Stability assessment of the harmonised and homogenised FCDR at three reference sites.	19

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

1 Introduction

This report addresses the assessment of the Fundamental Climate Data Records (FCDRs) using the methodologies developed under WP2 for mutual and relative consistency and stability (trend artefacts and step changes) both within and across sensor-series. Where possible the FCDR uncertainty information will be validated. Each sensor team has found it necessary to assess stability in different ways, because of the different context of different sensor series.

1.1 Scope

FCDRs addressed in this document: Advanced Very High Resolution Radiometer (AVHRR), High resolution Infrared Sounder (HIRS), Meteosat Visible and Infrared Imager (MVIRI) and microwave (MW) sounders.

1.2 Version Control

Version	Reason	Reviewer	Date of Issue
1.0	Release version	Authors	30 August 2019
1.1	Minor update after Final Review	RP	15/10/19

1.3 Applicable and Reference Documents

Dec2013	Decoster, I.; Govaerts, Y.M.; Baudrez, E.; Dewitte, S.; Velazquez, A.; Blazquez, A.; Cornelis, J. Evidence of pre-launch characterization problem of Meteosat-7 visible spectral response. <i>Remote Sensing Letters</i> 2013 , 4, 1008–1017. doi:10.1080/2150704X.2013.828181.
Gov1999	Govarts, Y. M. Correction of the Meteosat-5 and -6 radiometer solar channel spectral response with the Meteosat-7 sensor spectral characteristics. <i>International Journal on Remote Sensing</i> 1999 , 10 (18), 3677–3682. doi: 10.1080/014311699211273.
Dec2014	Decoster, I.; Clerbaux, N.; Baudrez, E.; Dewitte, S.; Ipe, A.; Nevens, S.; Blazquez, A.; Cornelis, J. Spectral Aging Model Applied to Meteosat First Generation Visible Band. <i>Remote Sensing</i> 2014 , 6, 2534–2571. doi: 10.3390/rs6032534.
Gov2018	Govaerts, Yves; Rüthrich, Frank, John, Viju; Quast, Ralf, Climate Data Records from Meteosat First Generation – Part I: Simulation of Accurate Top-Of-Atmosphere Spectral Radiance over Pseudo-Invariant Calibration Sites for the Retrieval of the In-Flight MVIRI/VIS Spectral Response, <i>Remote Sensing</i> 2018 , 10(12), 1959; doi:10.3390/rs10121959.
Qua2019	Ralf Quast, Ralf Giering, Yves Govaerts, Frank Rüthrich and Rob Roebeling, Climate Data Records from Meteosat First Generation Part II: Retrieval of the In-Flight Visible Spectral Response, <i>Remote Sens.</i> 2019 , 11(5), 480; https://doi.org/10.3390/rs11050480
Joh2019	Viju O. John, Tasuku Tabata, Frank Rüthrich, Rob Roebeling, Tim Hewison, Reto Stöckli and Jörg Schulz, On the Methods for Recalibrating Geostationary Longwave Channels Using Polar Orbiting Infrared Sounders, <i>Remote Sens.</i> , 2019 , 11, 1171; https://doi.org/10.3390/rs11101171
Wol1985	Wolff, T. An image geometry model for METEOSAT. <i>Int. J. Remote Sens.</i> 1985 , 6, 1599–1606, doi:10.1080/01431168508948308.
Key1981	Keys, R.G. Cubic Convolution Interpolation for Digital Image Processing. <i>IEEE Trans. Acoust. Speech Signal Process.</i> 1981 , 29, 1153–1160, doi:10.1109/TASSP.1981.1163711.
Rue2019	Frank Rüthrich, Viju O. John, Rob A. Roebeling, Ralf Quast, Yves Govaerts, Emma Wooliams and Jörg Schulz, Climate Data Records from Meteosat First Generation Part III: Recalibration and Uncertainty Tracing of the Visible Channel on Meteosat-2–7 Using Reconstructed, Spectrally Changing Response Functions, <i>Remote Sens.</i> 2019 , 11(10), 1165; https://doi.org/10.3390/rs11101165

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

Gov2004	Govaerts, Y.; Clerici, M. MSG-1/SEVIRI Solar Channels Calibration Commissioning Activity Report. 2004. Available online: https://www.eumetsat.int/website/home/Data/Products/Calibration/MSGCalibration (accessed on 31 July 2019).
Lic2006	Lichtenberg, G.; Kleipool, Q.; Krijger, J.M.; van Soest, G.; van Hees, R.; Tilstra, L.G.; Acarreta, J.R.; Aben, I.; Ahlers, B.; Bovensmann, H.; et al. SCIAMACHY Level 1 data: Calibration concept and in-flight calibration, <i>Atmos. Chem. Phys.</i> 2006 , <i>6</i> , 5347–5367, doi:10.5194/acp-6-5347-2006.
SPH2013	SCIAMACHY Product Handbook. 2013. Available online: https://earth.esa.int/web/guest/missions/esa-operational-eo-missions/envisat/instruments/sciamachy-handbook/wiki (accessed on 31 January 2019).
Til2017	Tilstra, L.G.; Tuinder, O.N.E.; Wang, P.; Stammes, P. Surface reflectivity climatologies from UV to NIR determined from Earth observations by GOME-2 and SCIAMACHY. <i>J. Geophys. Res. Atmos.</i> 2017 , <i>122</i> , 4084–4111, doi:10.1002/2016JD025940.
Bur2011	Burrows, J.P.; Goede, A.P.H.; Muller, C.; Bovensmann, H. SCIAMACHY—The Need for Atmospheric Research from Space. In <i>SCIAMACHY-Exploring the Changing Earth's Atmosphere</i> ; Gottwald and, M., Bovensmann, H., Eds.; Springer, Dordrecht 2011 ; doi: 10.1007/978-90-481-9896-2_1.
Sne2011	Snel, R.; Lichtenberg, G.; Noël, S.; Krijger, M.; Slijkhuis, K.; Bramstedt, K. Calibration and Monitoring. In <i>SCIAMACHY-Exploring the Changing Earth's Atmosphere</i> ; Gottwald and, M., Bovensmann, H., Eds.; Springer, Dordrecht 2011 ; doi: 10.1007/978-90-481-9896-2_5.
NOAA KLM User Guide	Robel, J.; others. NOAA KLM User's Guide with NOAA-N, -N' Supplement. Technical report, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, National Climatic Data Center, Remote Sensing and Applications Division, 2009
Atkin1998	Atkinson, N.C.; McLellan, S. Initial evaluation of AMSU-B in-orbit data. <i>Proceedings of SPIE – The International Society for Optical Engineering</i> 1998 , <i>3503</i> , 276–287. doi:10.1117/12.319506.
Atkin2001	Atkinson, N.C. Calibration, monitoring and validation of AMSU-B. <i>Advances in Space Research</i> 2001 , <i>28</i> , 117–126. doi:10.1016/S0273-1177(01)00312-X
John2012	John, V.O.; Holl, G.; Buehler, S.A.; Candy, B.; Saunders, R.W.; Parker, D.E. Understanding inter-satellite biases of microwave humidity sounders using global simultaneous nadir overpasses. <i>Journal of Geophysical Research</i> 2012 , <i>117</i> . doi:10.1029/2011JD016349.
John2013	John, V.O.; Holl, G.; Atkinson, N.; Buehler, S.A. Monitoring scan asymmetry of microwave humidity sounding channels using simultaneous all angle collocations (SAACs). <i>Journal of Geophysical Research</i> 2013 , <i>118</i> , 1536–1545. doi:10.1002/jgrd.50154
PUG	Product User Guide, Microwave version 4.1
Hans2019a	Hans, I. Burgdorf, M., Buehler, S.A., Lang, T., Prange, M., John, V.O.: An uncertainty quantified fundamental climate data record for microwave humidity sounders. <i>Remote Sens.</i> 2019 , <i>11</i> (5). Doi: 10.3390/rs11050548.
Hans2019b	Hans, I. Burgdorf, M., Buehler, S.A.: On-board radio frequency interference as origin of inter-satellite biases for microwave humidity sounders, in preparation for special issue on Radio Frequency Interference in Remote Sensing
Prange2019	Prange, M.: Application and Comparison of Inter-Calibration Methods for Satellite Microwave Humidity Sounders, Master's Thesis, Universitaet Hamburg, 2019

1.4 Glossary

AAPP	ATOVS and AVHRR Pre-processing Package
AMSU-B	Advanced Microwave Sounding Unit -B

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

ATOVS	Advanced TIROS-N Operational Vertical Sounder
AVHRR	Advanced Very High Resolution Radiometer
CDR	climate data record
DMSP	Defense Meteorological Satellite Program
DSV	deep space view
ECV	essential climate variable
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FCDR	fundamental climate data record
IWCT	internal warm calibration target
LECT	local equator crossing time
MHS	Microwave Humidity Sounder
MW	microwave
NEdT	Noise equivalent differential temperature
NetCDF	Network Common Data Format
NOAA CLASS	Comprehensive Large-Array Stewardship System
NWP	numerical weather prediction
RFI	radio frequency interference
SNO	simultaneous nadir overpass
SSMT-2	Special Sensor MicrowaveWater Vapor Profiler
UTH	upper tropospheric humidity

1.5 Executive Summary

Fundamental Climate Data Records (FCDRs) have been generated for four series/categories of sensors. These FIDUCEO FCDRs have some common characteristics: not merely convenient formatting standards etc, but that fact that they are built on common, carefully debated approaches to uncertainty characterisation and multi-sensor calibration – so called harmonisation. Harmonisation is recalibration of data to bring consistency while respecting real differences between sensors, such as in spectral response.

Metrological assessment as reported here is designed to gauge how successful the work of FIDUCEO has been in producing consistent and stable FCDRs. Each sensor team has faced different challenges and has developed different approaches to assessment, while retaining the same metrological framework and approach.

The Meteosat Visible and Infrared Imager (MVIRI) FCDR is a dataset of a broad visible-wavelength reflectance channel on sensors in geostationary orbit. The key issue for the MVIRI FCDR is the continuous degradation/change of spectral response in orbit, becoming less responsive in the blue part of the spectrum over time. The work done involved estimating this degradation and harmonising the series given those estimates. The FCDR generated is shown by several comparisons to have high stability, which is demonstrated above invariant sites where stability is expected (Algeria). The consistent recalibration, the better characterisation of the SRFs and the improved quality flagging led to significant improvements of the reflectance measurements, particularly for high clouds. Problems likely related to the eruption of ‘El Chichon’ in 1982 should be addressed in future work on better characterisation of the stratospheric aerosols, so that this small shortcoming can be mitigated in the future.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

The microwave sounders (MW) FCDR stores the brightness temperature measured by the microwave humidity sounders Special Sensor Microwave Water Vapor Profiler (SSMT-2), Advanced Microwave Sounding Unit-B (AMSU-B) and Microwave Humidity Sounder (MHS) installed on board polar orbiting satellites. This FCDR overcomes certain problems of the available historical operational data of these instruments, which were designed for assimilation in numerical weather prediction (NWP), but whose length of timeseries makes the timeseries relevant for climate, if suitably stabilized. The microwave brightness temperatures records are addressed for the FCDR in three regards: reduction of bad data problems, uncertainty estimation and harmonisation. Useful progress is recorded against these three aims. Improvements in the parameters of the measurement equation and the correction for RFI contamination have increased the inter-satellite consistency. The instruments agree within the uncertainties due to common effects. The harmonization procedure shows potential for and need for further improvements. First, as understanding of data problems increases and the data become “cleaner” with each reprocessing, the harmonisation procedure becomes more effective. Second, further developments on the selection of the observation-pairs used for harmonisation are needed, to ensure optimization is effective across a full range of conditions such as all scan positions, scene temperature ranges and instrumental state. Strategies so far give the hope-for benefits of harmonisation for partial ranges of conditions. Future developments building on the developed techniques promise a more comprehensive outcome, by developing methods to provide a broader range of constraints to the harmonisation process.

A similar conclusion holds for the assessment of FCDRs for High Resolution Infrared Sounders (HIRS) and Advanced Very High Resolution Radiometers (AVHRRs). The FCDR for AVHRR consists of infrared and visible imagery over a complex series of 9 satellites, with the three IR (3.7, 11 and 12 μm) channels addressed within FIDUCEO. The variability of uncertainty over time, both between satellites and within their lifetimes, is demonstrated in the assessment, and is valuable new information for users. Harmonisation has been more successful for the longer-wavelength channels (11 and 12) because inter-sensor pairs were more numerous and better distributed than for the 3.7 μm channel. Harmonisation is effective for the AVHRR/3 design instruments in the split window region. AVHRR/3s have improved design and are also closest in time to (usually overlapping with) the reference sensor (AATSR). The matches used for harmonisation span the colder range of scene temperature for most sensor pairs, and results are best for observations that reflect that sampling distribution. As with MW, the key to future improvement of results across the whole data range is to develop methods to provide a broader range of constraints to the harmonisation process.

The FCDR for HIRS addresses 19 channels on a series of 12 sensors. With 228 sensor-channel combinations, this is the most stringent test for the general uncertainty estimation and harmonisation methodologies developed in FIDUCEO, since it was well beyond the resource of the project to attempt bespoke solutions for the different challenges seen with different sensors and channels. In general, uncertainty characterization has delivered important new information to users of the FCDR, although the assessment reveals that instabilities in one aspect of the uncertainty estimation causes intermittent artefacts. Significant success in harmonisation has been achieved across some of sensor-channel combinations. For the eleven HIRS channels at wavelengths longer than 5 μm (and whose spectral definition remains stable across the series), the HIRS BTs are effectively harmonised. Best results are obtained for channels where the available inter-sensor matchup distributions thoroughly spanned the observation range, and in other cases it will be necessary to develop methods to provide a broader range of constraints to the harmonisation process.

D4_6 Metrological assessment of consistency, **V1.1** stability and uncertainty of FIDUCEO FCDRs

2 MVIRI:

2.1 The FIDUCEO MVIRI FCDR

The FIDUCEO MVIRI Fundamental Climate Data Record (FCDR) contains data from 35 years of geostationary satellite observations from Meteosat First Generation (MFG) Satellites. The instrument onboard those satellites, the Meteosat Visible and Infrared Imager (MVIRI), has been developed to support forecasters at the national weather centers with visually interpretable information about the state of the atmosphere. It acquires one image of the earth disk below the satellite every 30 minutes with nominal pixel size at the sub satellite point of 5km in the IR/WV bands and 2.5 km in the visible band. While originally being designed in the 70s, the last of the seven MVIRI instruments was launched in 1997. The specifications for the instrument performance as well as the requirements for pre-launch tests have evolved in between the launch dates, hindering the exploitation of the 35 years of observations for climate studies up to now. Among other difficulties, previous studies have particularly pointed out problems with the pre-launch characterization of the sensor spectral response functions (Dec2013 and Gov1999) as well as signs of spectral degradation (Dec2014). Furthermore, users of satellite data increasingly demand a thorough characterization of the measurement uncertainty, both in terms of magnitude and correlation in the domains of space, time and wavelength. The FIDUCEO MVIRI FCDR now contains reconstructed spectral response functions (SRFs) that account for the spectral degradation of the sensors (Gov2018 and Qua2019). Those SRFs are used for a consistent recalibration of all observations, spanning MVIRI on Meteosat 2-7, and for the computation of the top of atmosphere bidirectional reflectance factors that are included in the dataset. Independent and structured uncertainties of each measurement are traced from physical effects and correlations between effects as well as correlations in time and space are considered and provided along with the data. While the FIDUCEO project has been targeting on the visible band of MVIRI, the FCDR also contains the recalibrated IR and WV bands as obtained from EUMETSATs other activities (Joh2019).

2.2 Data

2.2.1 MVIRI VIS Observations

The Meteosat satellites were operated in a geostationary orbit and were designed to provide one image of the earth every 30 min. Continuous data from the sub-satellite position of 0° were collected since Meteosat-2, comprising now more than 24 years of data. In 2006, the 0° service was taken over by Meteosat Second Generation (MSG) satellites. During 1998–2017, Meteosat-5 and 7 were re-located to the east to provide Indian Ocean data coverage (IODC). The visible band of MVIRI essentially consisted of four silicon photodiodes, of which two were active and two were backup. The active silicon detectors responded to the light from the place where the telescope was pointing. The position of the two active sensors on the radiometer's focal plane was shifted relative to each other in north-south direction, so that they measured adjacent lines.

The MFG satellites were spin-stabilised satellites and therefore, unlike other three-axis stabilised satellites, they did not need a rotating scan-mirror. The visible images resulted from the interplay of the detectors measuring, the satellite rotating at a defined speed and the telescope tilting to a defined angle. During one revolution of the satellite, each of the two active sensors acquired one scan-line across the earth. The two scan-lines were transmitted to the ground station during that part of the revolution, during which the radiometer was not directed towards the earth. Before transmission, the original sensor voltage was

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

truncated into digital count values C . This A/D conversion was done on 8 bits (256 levels) since Meteosat-4, Meteosat-2 and -3 were still encoded to 6-bit values (64 levels). Upon reception in the ground segment, those latter were inflated to the 256-level range of the 8-bit data. For the VIS (IR) detectors each line consisted of 5000 (2500) pixels corresponding to a nadir resolution of 2.25km (4.5km) (Table 1) In the ground segment the raw image lines were combined with metadata into the so-called Level 1.0 format. The processing into the Level 1.5 data format includes the navigation of the images into a georectified grid (Wol1985). For this purpose, information about the orbit attitude and inclination were used to identify the best matching Level 1.0 pixel for each cell in the rectified Level 1.5 grid. The 4×4 surrounding Level 1.0 pixels were then averaged using distance-weighted cubic spline interpolation (Key1981). The accuracy of the georectification was then accessed using a set of Landmarks. The FIDUCEO FCDR represents the most recent Version of the MVIRI data record. The generation of the FCDR is comprehensively described in 3 peer reviewed papers (Gov2018, Qua2019 and Rue2019).

Table 1: Spatial and spectral characteristics of Meteosat Visible Infra-Red Imager (MVIRI) visible (VIS), thermal infrared (TIR) and water vapour (WV) channels.

Channel	sampling nadir (km)	Nominal spectral band (μm)
VIS 0.7	2.25	0.40–1.10
WV 6.4	4.5	5.70–7.10
TIR 11.5	4.5	10.5–12.5

2.2.2 SEVIRI Measurements

SEVIRI is a radiometer onboard a series of four Meteosat Second Generation (MSG) satellites that are operated by EUMETSAT in a geostationary orbit. In 2002, the first MSG satellite (Meteosat-8) was launched. Similar to the MFG satellites, the MSG satellites are spin stabilised. The SEVIRI instrument operates 12 channels simultaneously. Three of its channels are at visible and near infrared wavelengths between 0.6 and 1.6 μm , eight channels are at infrared wavelengths between 3.8 and 14 μm and one channel is a high-resolution visible (HRVIS) channel. For this study, particularly the HRVIS channel is of relevance due to its broad spectral coverage being comparable to the MVIRI VIS channel. As for MVIRI VIS, the detectors employed are silicon photodiodes. In contrast to the latter, however, not only two detectors are operated simultaneously, but an array of 9 detectors. With this setup, the HRVIS channel acquires 9 scanlines during each revolution of the satellite. In this way a fulldisk earth scan can be performed every 15 min, much faster than with MVIRI. During the period 2004–2006, Meteosat-8 was operated at a sub-satellite longitude close to Meteosat-7. This period therefore is predestined for useful comparisons between SEVIRI and MVIRI. The characteristic of four solar channels are summarised in Table 2 according to the Meteosat-8 (MSG1) commissioning report (Gov2004).

Table 2: Spatial and spectral characteristics of the four Spinning Enhanced Visible Infra-Red Imager (SEVIRI) visible (VIS) and near infrared (NIR) channels. HRVIS: high-resolution visible.

Channel	Sampling Nadir (km)	Nominal Spectral Band (μm)	Mission requirement on SNR @1% albedo
VIS 0.6	3	0.56–0.71	10.1
VIS 0.8	3	0.74–0.88	7.28

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

NIR1.6	3	1.50–1.78	3.0
HRVIS	1	0.37–1.25	4.3

2.2.3 SCIAMACHY Measurements

SCIAMACHY was a scanning nadir and limb spectrometer covering the ultraviolet (UV) through visible to shortwave infrared (SWIR) spectral range. SCIAMACHY was a joint development of Germany, the Netherlands and Belgium and was launched in February 2002 onboard the ENVISAT platform operated by the European Space Agency (ESA) (Lic2006). About 10 years after launch, on April the 8th 2012, ESA lost contact with the ENVISAT satellite. ENVISAT was a sun-synchronised polar orbiting satellite with a local equator crossing time of 10:00 AM and an orbital period of about 100 minutes. SCIAMACHY performed nadir and limb measurements. In limb mode, the instrument observed a certain volume of the atmosphere about 7 min before it was observed in nadir mode. The orbit swath is 960 km wide. The wavelength range covered by SCIAMACHY is 240–2380 nm in eight spectral channels with a spectral resolution between 0.2–1.5 nm. Light that entered the instrument was dispersed using an assembly of prisms and holographic diffraction gratings onto the arrays of 1024 detectors per channel. While reticon photodiodes were used for the five UV-VIS channels, the three SWIR channels were equipped with Indium Gallium Arsenide detectors (SPH2013). The 1024 detectors were sub-divided into clusters that are useful for trace-gas retrieval. As each detector, after the dispersion and bending of the incoming light beam, represented a unique wavelength, the clusters corresponded to wavelength regions. For each of the 56 clusters the integration time could be varied, resulting in various spatial resolutions. This allowed a higher spatial resolution for the most important spectral regions and longer integration times where needed. In order to cover the entire measured spectrum, the measurements of all clusters have to be integrated into the broadest pixel size. Global coverage (in nadir mode) is achieved in six days (Til2017). The instrument was originally designed for studying atmospheric chemistry and aerosols (Bur2011). The SCIAMACHY Level 1 dataset has undergone several recalibration campaigns. The latest version (V8) of the Level 1b dataset contains all calibration parameters described in [Sne2011].

2.3 Assessment Methodology

2.3.1 Expected performance of the FCDR

The different characteristics of the 6 MVIRI instruments are expected to introduce jumps even in the final, recalibrated FCDR. Moreover, those jumps are expected to depend on the spectral characteristics of the surface in a pixel (Figure 1).

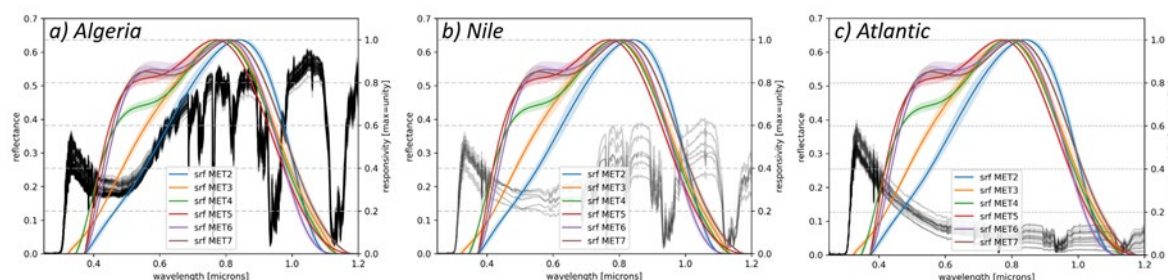


Figure 1: Spectral Response functions of Meteosat-2-7 plotted together with SCanning Imaging Absorption spectroMeter for Atmospheric Cartography (SCIAMACHY) spectra acquired during 2002 at three target sites. a) shows the spectra measured by

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

SCIAMACHY at the Algeria site, b) shows the spectra measured at the Nile-delta and the c) shows those measured at the Atlantic-1 site. Spectra are plotted in transparent black to better illustrate their spread. Spectra with strong cloud contamination were removed before plotting (see Section 3.9.1). Note that for band-adjustment/homogenisation more sophisticated filtering regarding clouds and scene heterogeneity is applied.

To assess the expected performance of the FCDR against the operational version, reflectance spectra measured by the SCIAMACHY instrument are exploited. The cloud-filtered spectra acquired at two sites (Algeria and Atlantic) during 2002 are used to derive two characteristic mean spectra. Convolved with the SRFs of the instruments, they yield a band-integrated reflectance that is theoretically expected to be measured by each instrument at the moment where the SRF is valid. Two sets of SRFs are used: The first set are the operational SRFs that were derived pre-launch for each instrument. Those SRFs were considered to be constant over the lifetime of a satellite. Consequently only one SRF is available per satellite. The second set of SRFs are the reconstructed versions (Qua2019), valid for 45 days each.

2.3.2 Harmonised time-series of reflectance and uncertainty

The recalibrated MVIRI FCDR represents a harmonised dataset because the same pseudo-invariant calibration sites (PICS) are used for all 6 instruments. For an assessment of the actual performance of the FCDR against the expected performance, reflectances at two reference sites are extracted from the fulldisk: Algeria and Atlantic. In order to avoid impacts from variable observation geometries, only image time slots between 11:00 and 13:00 UTC are considered. The reflectance of n images from each day within a $m=3 \times 3$ pixel box around the central latitude/longitude (Table 3) are averaged. The uncertainties of those measurements are combined according to Equation 1 and Equation 2, where $u_i(\hat{R})$ is the independent uncertainty of the averaged reflectance and $u_s(\hat{R})$ is the structured uncertainty of the averaged reflectance. Note that for the structured uncertainty it is assumed that the errors within a 3×3 pixel box are fully correlated while the errors for subsequent images are independent.

Equation 1

$$u_i(\hat{R}) = \frac{1}{\sqrt{n * m}} \sqrt{\sum_{t=1}^n \sum_{k=1}^m u_i(R_{t,k})^2}$$

Equation 2

$$u_s(\hat{R}) = \frac{1}{\sqrt{n}} \sqrt{\sum_{t=1}^n \left(\sqrt{\sum_{k=1}^m u_s(R_{t,k})^2} \right)^2}$$

Table 3: Evaluation sites where recalibrated (harmonised) and band-adjusted (homogenised) time series were generated along with spectral characteristics and thresholds for the filtering of SCIAMACHY spectra.

Site	Land Cover Type	Dominant Spectral Contribution	Central Latitude	Central Longitude	Size of Box
Algeria-3	Desert	Red	30.32	7.66	2° × 2°
Nile	Agricultural land	Green	30.5	31.25	0.5° × 0.5°
Atlantic-1	Sea	Blue	-22.5	9.5	2° × 2°

2.3.3 Homogenised time series and stability analysis

In order to test the assumption that the jumps and trends in the harmonized time series are explained by the different SRFs and their spectral degradation, the time series have to be homogenised. This is also an important precondition for an assessment of the stability. Homogenisation here describes the process of

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

adjusting all sensors of a series to a common reference sensor. This process is also called band-adjustment. The reference sensor is Meteosat-5 at start of lifetime. Band-adjustment is only possible if the reflectance spectrum of the target site is well known. For this purpose SCIAMACHY spectra are used and extracted above the sites in Table 3. As the sites are assumed to have an invariant reflectance spectrum with time, the spectra acquired by SCIAMACHY during its lifetime can also be used for the homogenisation of older satellites, such as Meteosat-2.

In a first step the SCIAMACHY spectra are cloud filtered using three tests. Those tests are well described in Rue2019. Each cloud free spectrum is then convoluted with the SRF of Meteosat-5 as reference and with the SRF of the sensor that has to be adjusted. A linear fit to the convolved reflectance pairs then yields the band adjustment parameters (Figure 2)

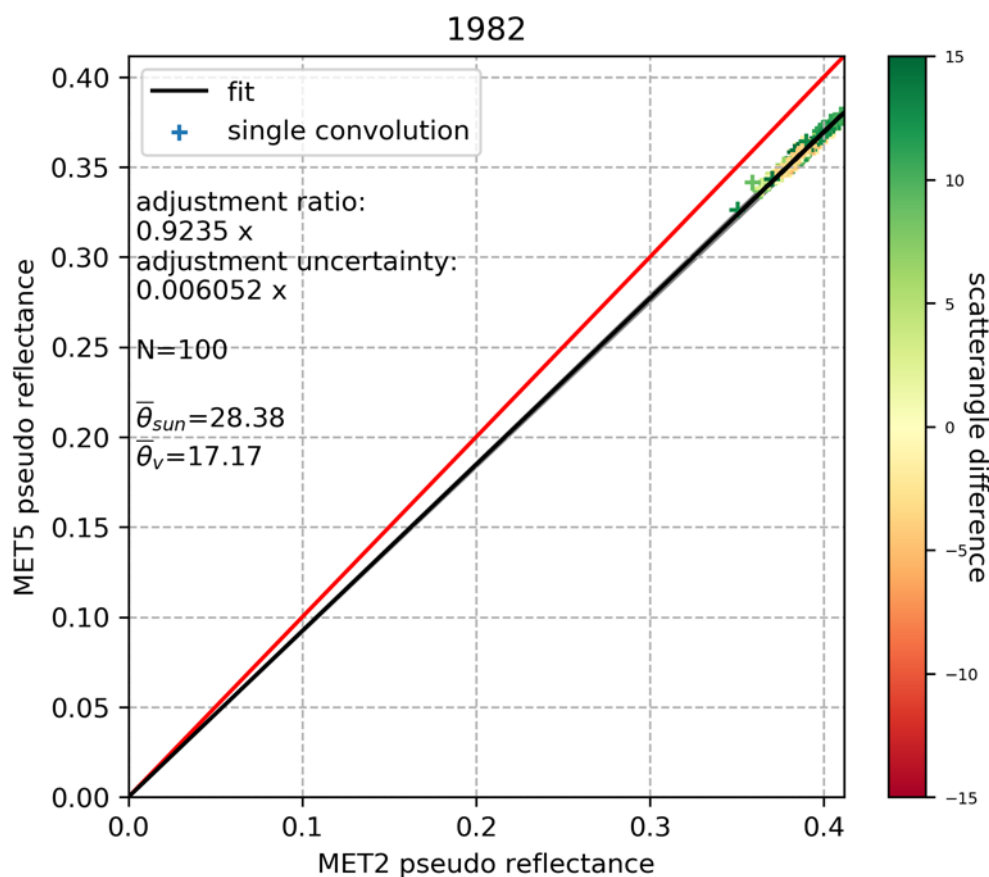


Figure 2: Convolved clear-sky SCIAMACHY spectra above Algeria-3 using Meteosat-2 and Meteosat-5 SRFs that were valid at each satellites' launch date.

When applied to the entire Meteosat time series, the spectral band adjustment parameters have to be updated for every change of the SRFs due to the spectral degradation.

The assessment of the decadal stability is applied on the homogenised time series above the three validation sites. In a first step an additional cloud filter is applied. The filter uses the distribution statistics from a rolling kernel of 30 days around each reflectance measurement. Measurements that are brighter than the 25% percentile are rejected in order to analyse only measurements that are certainly cloud free. For the Atlantic-

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

1 site also periods with globally elevated aerosol loads due to volcanic eruptions (El Chichon and Pinatubo) are excluded from the stability analysis. Those additional filters are meant to ensure that the computed stability is not corrupted by changes in the performance of the operational cloud mask or by increased Rayleigh-scattering from aerosols. In a second step, the smoothed mean annual cycle is subtracted from the time series in order to get rid of any seasonality. In a third step the measurements are then aggregated to daily means and a generalised linear model (GLM) model is fitted to the deseasonalised time series. The regression slope is evaluated as a measure of the stability in relation to the mean reflectance at each target site.

2.3.4 Comparison against SEVIRI

The validation of the presented FCDR is performed using superior observations from the SEVIRI sensor onboard the first Meteosat Second Generation satellite. This satellite was operated on a sub-satellite point (SSP) that was close to MET7 during the years 2004-2006. For validation purposes particularly the high-resolution visible (HRVIS) band of SEVIRI is useful, due to its similarity with MVIRI VIS. For the validation the operational MVIRI product and the new FIDUCEO FCDR are spectrally adjusted to the HRVIS band using the same methodology as described above for the time series homogenisation. An example of the difference between the SRFs of the dataset versions is provided in Figure 3. Histograms of the band adjusted observations over spectrally different target regions and clouds (Table 4) were then extracted.

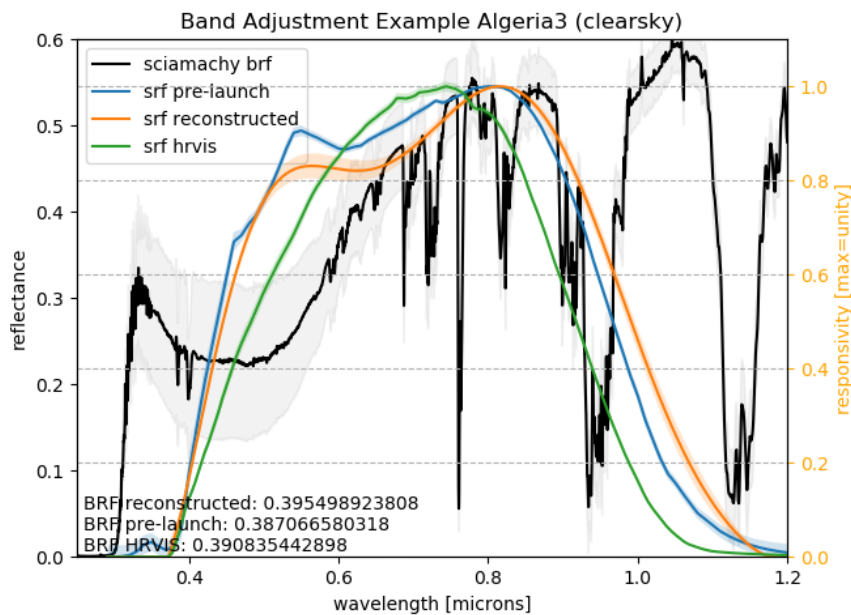


Figure 3: Mean clear sky SCIAMACHY spectrum above the Algeria-3 site with standard deviation and with examples of the MVIRI and SEVIRI spectral response functions that are used for convolution.

Table 4: Characterisation of the sites used for the comparison of the SEVIRI HRVIS dataset with the operational MVIRI dataset and the harmonised MVIRI FCDR. The sites and criteria are used for the extraction of suited SCIAMACHY spectra for the homogenisation and for the extraction of SEVIRI/MVIRI observations.

Site	Criterion	Central Latitude	Central Longitude	Size of Box (lat × lon)
Algeria-3	Cloud fraction = 0	30.32	7.66	4° × 4°
High-cloud	Cloud-top-pressure < 200hPa	0.0	0.0	10° × 10°

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

Mid-cloud	Cloud-top-pressure 200-700hPa	0.0	0.0	$10^\circ \times 10^\circ$
Low-cloud	Cloud-top-pressure >700hPa	0.0	0.0	$10^\circ \times 10^\circ$
Atlantic-1	Cloud fraction = 0	-22.5	9.5	$10^\circ \times 5^\circ$

2.4 Assessment Results and Discussion

2.4.1 Expected performance of the FCDR

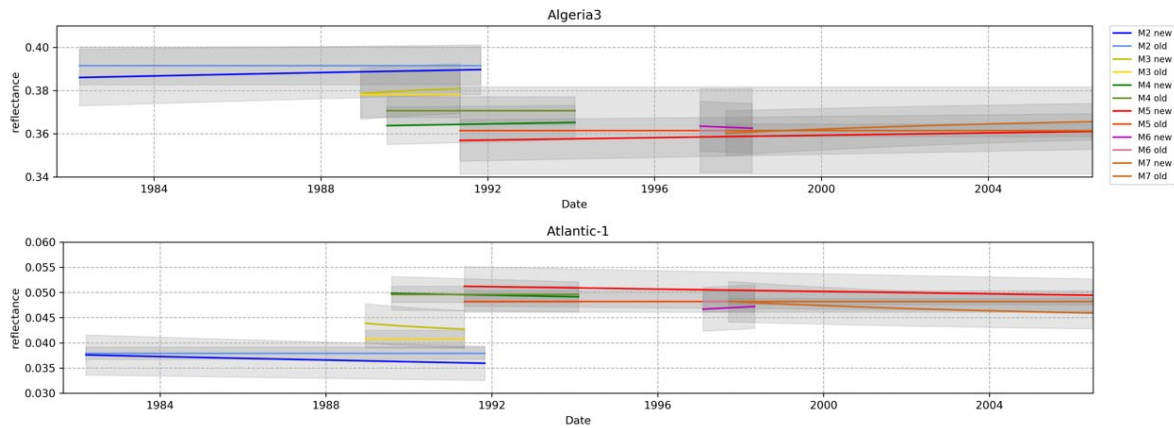


Figure 4: Theoretically expected time series of the TOA reflectance derived from 6 Meteosat instruments at two reference sites using the (old) pre-launch measured SRFs and the (new) reconstructed, spectrally degrading SRFs. Grey shading indicates the combined standard uncertainty introduced by the SRFs.

The artificial time-series of theoretically expected reflectance at the Algeria and Atlantic site (Figure 4) confirm that the recalibrated FCDR will have remaining jumps. Those jumps are caused by the different shapes of the SRFs. As such, they are not remainders of an incomplete recalibration process, but rather reflect the fact that the measurements are indeed different. Depending on the surface type of a pixel and its spectral BRDF, the jumps can be larger or smaller, positive or negative. Additionally, as a unique feature of the new FIDUCEO FCDR, the SRF of the MVIRI visible band is changing with time. This means that after some months in orbit even observations from the same MVIRI have to be considered spectrally different. In the dataset this becomes apparent as long-term trends over certain surfaces. As for the jumps, those trends are not a sign of an invalid recalibration process and reflect expected differences of the observations. Any interpretation or validation of the dataset has to consider the different spectral representation of the measurements e.g. by using appropriate spectral band adjustment functions.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

2.4.2 Harmonised time series of reflectance and uncertainty

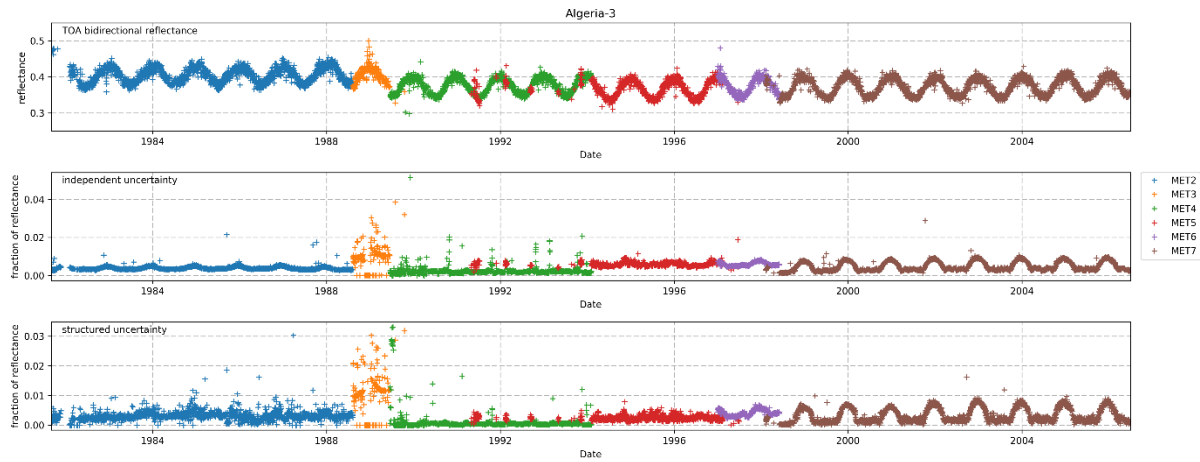


Figure 5: Recalibrated MVIRI clear sky reflectance time series and uncertainties for the Algeria-3 site, calibrated with the reconstructed, in-flight characterised SRF.

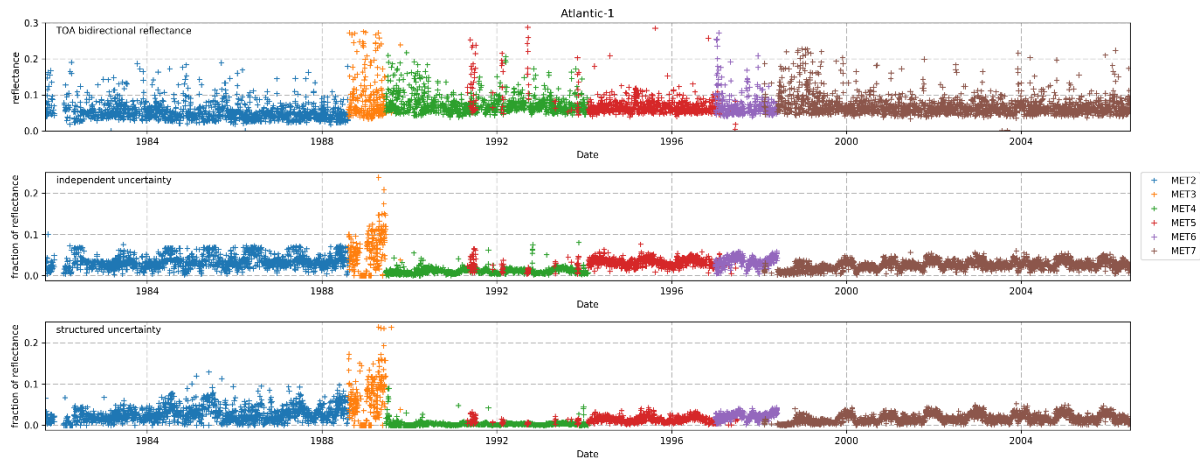


Figure 6: Recalibrated MVIRI clear sky reflectance time series and uncertainties for the Atlantic-1 site, calibrated with the reconstructed, in-flight characterised SRF.

The actually observed time series of clear-sky reflectance extracted from the harmonised MVIRI VIS FCDR match well with the above-described expectations. The most striking features are the jumps between the different sensors. Particularly Meteosat-2 and -3 deviate from the other satellites. At the Algeria-3 target site, with its dominant spectral contribution in the red, the clear-sky reflectance values from Meteosat-2 and -3 are brighter than those of Meteosat-4,-5,-6, and -7 (Figure 5). At the Atlantic-1 target site, with its dominant spectral contribution in the blue, the clear-sky reflectance values from Meteosat-2 and -3 are darker than those from the other satellites (Figure 6). The observed differences can be explained by the differences between the SRFs of Meteosat-2 and -3 and the other satellites (see Figure 1 and Figure 4). From Figure 1 it can be seen that the spectral response of Meteosat-2 and -3 is much weaker between 0.4 and 0.6 μm compared to the spectral response of Meteosat-4,-5,-6,-7. Therefore, the clear-sky reflectance of Meteosat-2 and -3 measured at Atlantic-1, which reflects most at wavelengths smaller than 0.6 μm must be lower than from the other satellites. The opposite applies for Algeria-3, which reflects most at wavelengths larger than 0.6 μm . For the green Nile delta (not shown) no noticeable jumps are present.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

Figure 5 and Figure 6 also show time series of the independent and structured uncertainties determined with the methods laid out in (Rue2019). The uncertainty for data from the Meteosat-2 satellite is small because of the coarse digitisation. While the coarse digitisation increases the digitisation noise (see equation 10 in Rue2019), this is compensated by the fact that the noise of the sensor voltage before reaching the analogue-to-digital converter has to be much larger in order to trigger any noise in the digitised signal. The seesaw-like pattern of the Meteosat-2 uncertainties may be attributed to the annual cycle of the instrument temperature. This cycle is a result of the satellite being heated by the sun from different angles over the year and it is recorded in the instrument telemetry data. Highest instrument temperatures occur during the winter months and highest diurnal amplitudes of the temperatures occur during the eclipse seasons in spring and autumn. Higher instrument temperatures lead to an increase of the noise of the onboard electronics and increase the differences of the sensitivities of the two detectors. In this way, the variations in instrument temperature affect the independent and the structured uncertainty. The uncertainty of Meteosat-2 data has been reduced as compared to what was reported in Rue2019. This was achieved by an improved flagging of lines containing no data. For Meteosat-3 the improved flagging has not been applied.

The influence of instrument temperature variations is smaller for the newer satellites (Meteosat-4,-5,-6,-7) because of enhanced sun protection. The bulging pattern of the independent uncertainty of Meteosat-7 can be attributed to a very different noise level of the two detectors on this satellite. This increases the uncertainty of the signal when combined for the two detectors. During June–August, the difference in noise level between the two detectors is mitigated by the temperature dependent switches of the Analog-to-Digital converters of the two detection chains (Rue2017). The bulging pattern is only visible at the bright Algeria site, while it is not visible at the dark Atlantic site that shows higher uncertainty because the measurement of the much smaller reflectance is very sensitive to instrument noise.

2.4.3 Homogenised time series and stability analysis

The homogenised MVIRI VIS FCDR time series of clear-sky reflectance do not exhibit any prominent jumps between sensors, proving that the FCDR correctly accounts for the real shape of each sensors SRF (Figure 7). Outliers at the Algeria site only occur for the short duration of the Meteosat-3 coverage during winter, pointing to problems of the operational cloud mask. Cloud contamination at the Nile delta is generally higher. Furthermore, impacts of human activities may occur at this site. This leads to a generally higher probability for variability, and that is confirmed by data. The clear-sky reflectance above the Atlantic are, as expected, around 0.05. The high spread of the observed values at this site again points to observations that were not flagged cloudy by the operational cloud mask. In order to exclude the variable performance of the cloud mask from the stability evaluation of the dataset, we implemented a consistent filtering across all satellites as an intermediate step (see section on methods). After applying this filtering the anomalies of the clear-sky reflectance time series (Figure 8) have been evaluated on their decadal stability. The results reveal very stable behaviour for Algeria-3 and for the Nile delta sites with trends around -0.6% and 1.7 % decadal change in reflectance (Table 5). Results for the Atlantic-1 site are worse with around 7.5% decadal change in reflectance. The main contribution to this change comes from the too dark values of Meteosat-2 (Figure 8). Since the homogenisation has removed the effects of the different SRFs, this finding indicates a potential overestimation of the Meteosat-2 SRF in the blue region of the spectrum.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

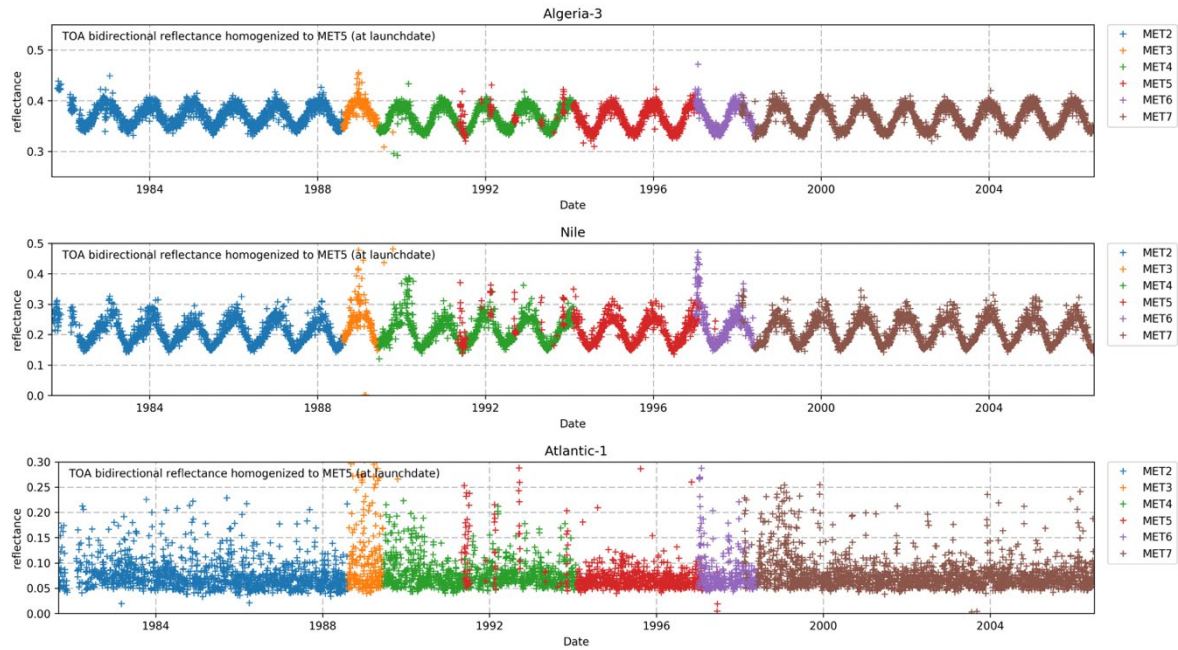


Figure 7: Homogenised time series of recalibrated broadband reflectances at Algeria-3, Nile-delta and Atlantic-1.

Table 5: Stability assessment of the harmonised and homogenised FCDR at three reference sites.

Site	Mean Reflectance	Stability of Reflectance [reflectance decade ⁻¹]	Relative Stability [% decade ⁻¹]
Algeria-3	0.36 ± 0.02	-0.0024 ± 0.000	-0.65 ± 0.02
Nile	0.18 ± 0.03	0.0031 ± 0.000	1.78 ± 0.03
Atlantic-1	0.05 ± 0.00	0.0038 ± 0.000	7.46 ± 0.01

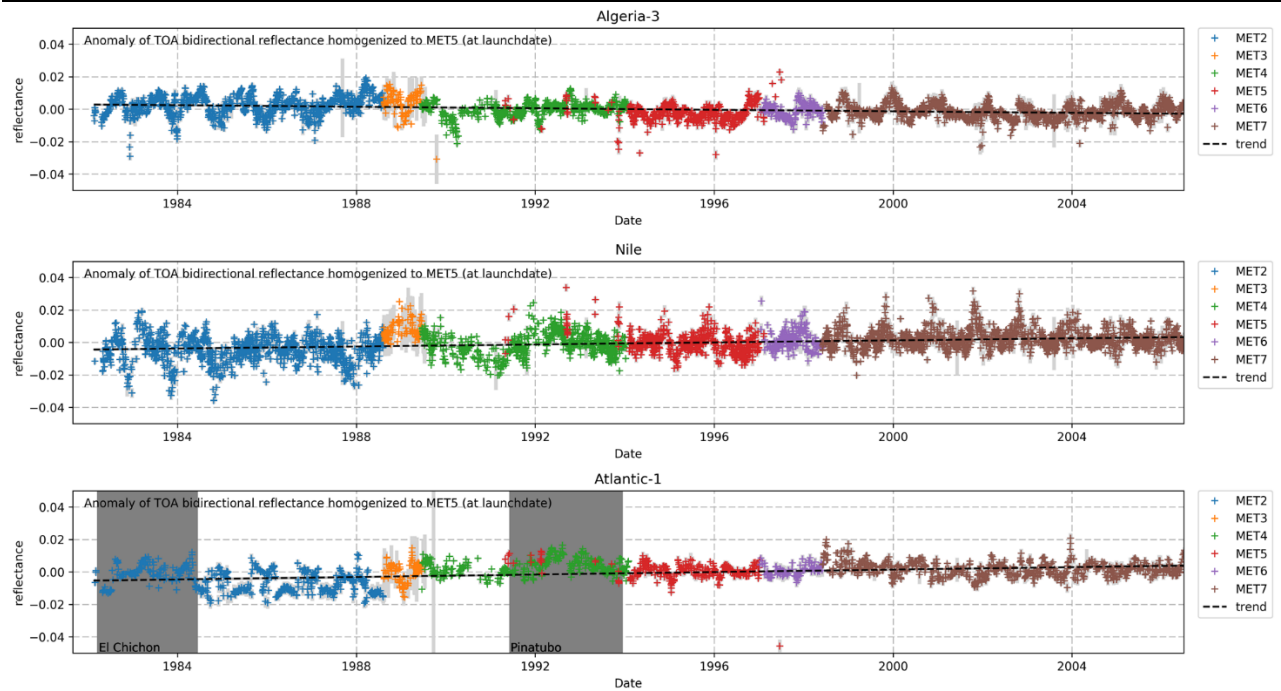


Figure 8: Anomalies of the homogenised reflectance time series and their trends. Anomalies are the deviation from the mean annual reflectance cycles. Additional filtering for cloud contamination was applied. The filter computes the 25th percentile from a rolling kernel of 30 days around each reflectance measurement. Measurements that are brighter than this value are rejected in

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

order to display only measurements that are certainly cloud free. Note that for the Atlantic site the periods with globally elevated aerosol loads due to volcanic eruptions are excluded from the stability analysis. Figure 12.

2.4.4 Comparison against SEVIRI

For the Algeria-3 desert target, the histograms of clear-sky reflectance from both MVIRI datasets do not deviate much from the histogram of the SEVIRI dataset (Figure 9). Only a subtle bright bias against SEVIRI is observed for both.. The good fit at this site is attributed to the fact that the three datasets were all calibrated using desert sites with comparable, red spectral characteristics. While this forces all instruments to measure the same over desert-like sites, a bias is introduced as soon as objects in the Field of View (FoV) have different spectral characteristics, such as clouds and ocean surfaces. This is apparent in the reflectance histograms for cloudy scenes shown in Figure 10. This Figure shows the histograms of fully cloudy pixels with cloud-tops at three different pressure levels (high clouds (above 200 hPa), middle clouds (between 200 hPa and 700 hPa) and low clouds (below 700 hPa)). From the figure it can be seen that reflectance values across all cloud levels are lower for the operational MVIRI dataset than in the SEVIRI HRVIS dataset. In the histograms of the recalibrated MVIRI reflectance, this dark bias is slightly reduced for low and middle clouds and strongly reduced for high clouds. For high clouds, the observed shift towards brighter cloud reflectance values would affect the top of the atmosphere outgoing shortwave radiation by roughly 8 W/m². Similar behaviour is observed over the Atlantic-1 target. Here the histogram from the operational dataset is also darker than that from the SEVIRI dataset (Figure 11) and the histogram from the recalibrated MVIRI VIS reflectance matches much better with the histogram of the SEVIRI dataset. The better match between the recalibrated data and the SEVIRI dataset is encouraging and confirms that the clear-sky and cloudy reflectance at target sites with different spectral characteristics are better (assuming SEVIRI is the better characterised instrument) represented when using reconstructed SRFs compared to the use of original (pre-launch) SRFs.

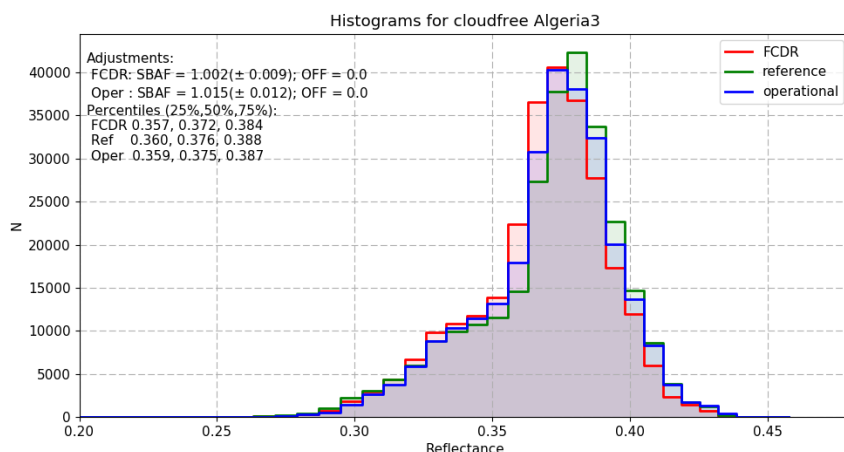


Figure 9: Histogram of SEVIRI HRVIS reflectance plotted as reference together with the histograms of the operational MVIRI reflectance and the recalibrated MVIRI reflectance, as obtained from cloud-free Algeria-3 pixels at 12:00 UTC slots during March 2005. The MVIRI datasets are band adjusted to the SRF of the SEVIRI HRVIS band according to the given Spectral Band Adjustment Factors (SBAFs). The SBAFs are computed using the same set of SCIAMACHY spectra and the difference between both SBAFs is entirely due to the different shapes of the MVIRI SRFs in the operational and the recalibrated datasets.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

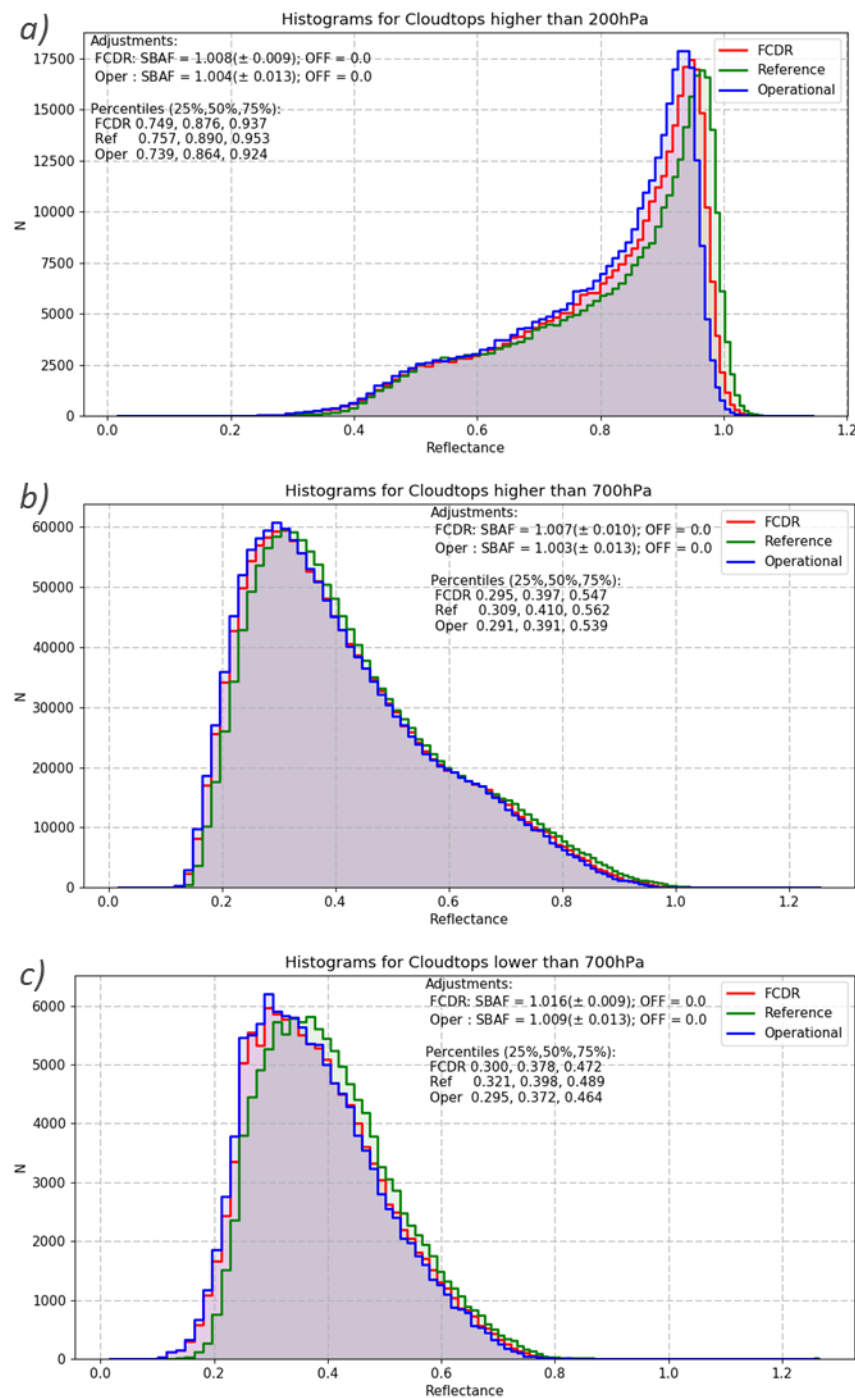


Figure 10: Histogram of SEVIRI HRVIS reflectance plotted as reference together with the histograms of the operational MVIRI reflectance and the recalibrated MVIRI reflectance, as obtained from entirely cloud-covered pixels at 12:00 UTC slots during March 2005. Clouds are classified into high clouds (a), middle clouds (b) and low clouds (c). The MVIRI datasets were band adjusted to the SRF of the SEVIRI HRVIS band according to the given SBAFs. For each cloud-class the SBAFs for the operational dataset and for the recalibrated dataset are computed using the same set of SCIAMACHY spectra and the difference between both SBAFs is entirely due to the different shapes of the SRFs in the operational and recalibrated datasets.

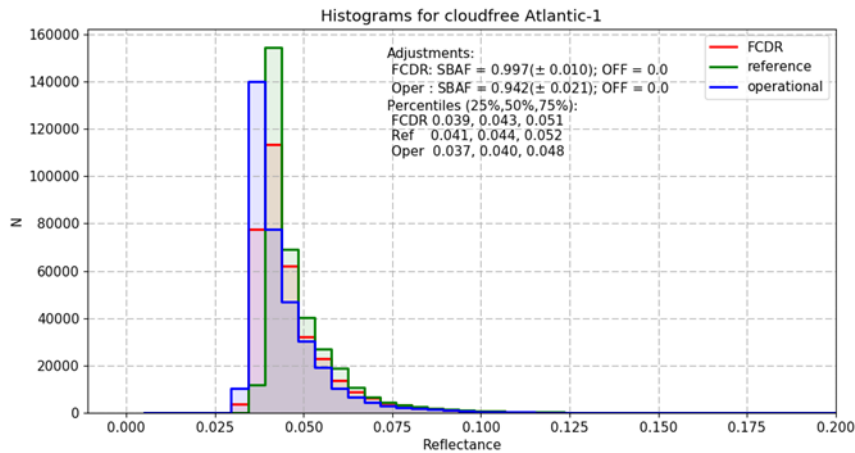


Figure 11: Histogram of SEVIRI HRVIS reflectance plotted as reference together with the histograms of the operational MVIRI reflectance and the recalibrated MVIRI reflectance, as obtained from cloud-free Atlantic-1 pixels at 12:00 UTC slots during March 2005. The MVIRI datasets were band adjusted to the SRF of the SEVIRI HRVIS band according to the given SBAFs. The SBAFs are computed using the same set of SCIAMACHY spectra and the difference between both SBAFs is entirely due to the different shapes of the SRFs in the operational and the recalibrated datasets.

2.5 Summary and Conclusions

The FIDUCEO efforts on the reconstruction of the VIS Band spectral response functions (SRFs) and the recalibration of the MVIRI VIS measurements have yielded an improved data record of the broadband visible reflectance at the top of atmosphere spanning more than 35 years. The reconstructed SRFs correctly account for the spectral degradation of the silicon photodiodes and the related change of the responsivity particularly in the blue part of the spectrum. The FCDR has high stability which is demonstrated above invariant sites where stability is expected (Algeria). The consistent recalibration, the better characterisation of the SRFs and the improved quality flagging led to significant improvements of the reflectance measurements, particularly for high clouds. However, slight issues were found with the characterisation of the SRF of Meteosat-2 above oceans, which may be caused by heavy contaminations with volcanic aerosols after the eruption of 'El Chichon' in 1982. With a better characterisation of the stratospheric aerosols this small shortcoming could be mitigated in the future.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

3 Microwave sounders

Much of this work has been published in the peer reviewed journal article, Reference Document Hans2019a.

3.1 Brief intro on what the FIDUCEO FCDR is

The FIDUCEO fundamental climate data record (FCDR) is a new FCDR that stores the brightness temperature measured by the microwave humidity sounders Special Sensor Microwave Water Vapor Profiler (SSMT-2), Advanced Microwave Sounding Unit-B (AMSU-B) and Microwave Humidity Sounder (MHS) installed on board polar orbiting satellites. This FCDR overcomes certain problems of the available historical operational data of these instruments, which were not designed for climate monitoring, but for assimilation in numerical weather prediction (NWP). As these data sets now cover more than 20 years, they become increasingly relevant for climate science. Climate research requires long time, consistent and uncertainty quantified data records that the available operational data sets do not provide. New aspects of the FIDUCEO FCDR are 1) easy handling and removal of sampling artefacts, 2) extensive pixel-level uncertainty information considering correlation behavior of the underlying errors - a strategy based on metrological principles - and 3) understanding and reducing inter-satellite biases by a recalibration approach. The new FCDR shall serve as a starting point for the construction of a climate data record (CDR). The CDR will be a gridded data set containing the essential climate variable (ECV) upper tropospheric humidity (UTH). UTH is an important ECV to be monitored: water vapor is a major greenhouse gas that is highly variable in space and time. Hence, global continuous observations are required for which satellite borne measurements are of high interest.

3.2 Method of validation/ reference sensors used/ new methods used

In a first step, we validate our newly set-up and coded processing chain for MW sensors. This is achieved by using the measurement equation with all parameters defined as in AAPP. This serves as test of our implementation, since we expect to reproduce AAPP output. Indeed, we are able to reproduce the operational AAPP results. This is shown in Figure 12 for the first half of an example orbit of MHS on Metop-B (start: 2015-07-06 15:47:58, end: 2015-07-06 17:29:18). The small visible differences are due to slightly different quality checks and no significant systematic differences are visible.

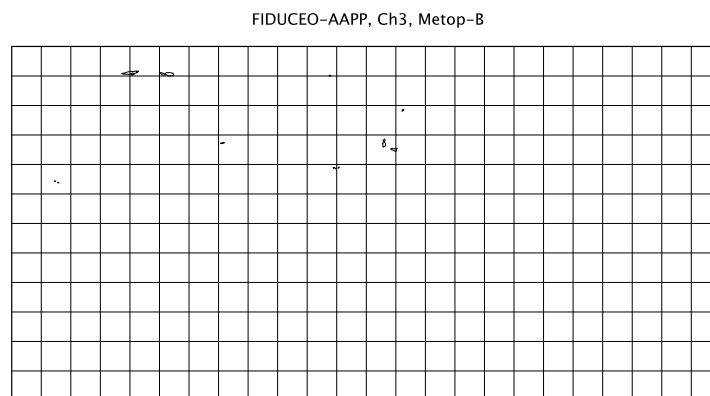


Figure 12 Difference in brightness temperature (channel 3) of FIDUCEO and AAPP processing for the first half of the example orbit of Metop-B.

Having validated our implementation of the measurement equation, we apply improvements to the measurement equation as explained in Hans2019a. The improvements consist in correction of assignment

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

errors of certain parameters (antenna pattern correction), improvement of the band correction and most of all, application of a newly developed correction scheme for sensors affected by radio frequency interference (RFI) [Hans2019b].

Using this improved measurement equation, we produce the new, consistent MW FCDR that is validated in Section 3.3.2.

For validating the new FCDR, we compare the time series of the five channels to the operational data processed with AAPP. We use global monthly means, for all viewing angles and ascending and descending passes averaged. A second approach is to compare the two data sets on the basis of matchup data.

The uncertainties stored in the FIDUCEO FCDR have no direct equivalent in existing data sets. A comparison is hence only possible to a limited degree. See next section for results.

3.2.1 Attempts to harmonise MW data with the FIDUCEO harmonisation method

The FIDUCEO harmonization is a novel approach for the inter-calibration of remote-sensing instruments, using recalibration instead of bias-correction on the brightness temperature. The harmonization uses SNOs as inter-comparison method. The data set of FIDUCEO matchups (SNOs and also off-nadir matches) between the sensors and the reference is generated by Brockmann Consult GmbH in WP3. For the found match-ups, the required information for calibration is collected in a second step in order to form the harmonization input data.

Instead of bias-correcting the measured brightness temperatures, the harmonization seeks to correctly account for the underlying physical effects that cause the bias in the calibration procedure. This recalibration approach uses the measurement equation and attempts to find optimized values for certain parameters in the measurement equation such that the inter-satellite bias is reduced. In an iterative procedure of harmonization and analysis of its results, the correct combination of calibration parameters that explains and reduces the biases needs to be found. Afterwards, reprocessing the data to level 1c using the optimized calibration parameters in the measurement equation generates the harmonized FCDR.

3.2.1.1 Harmonisation results using two parameters

For the first harmonization runs, the set of calibration parameters to be optimized is composed of the non-linearity correction and the warm target correction (only channel 3 is harmonized in the first attempts as it is the most important one for the derivation of the UTH CDR). The non-linearity is a relatively weak effect, but it allows by construction for a quadratic deviation from the linear calibration model. The warm target correction is a relatively strong effect. Its most important impact is on the warm temperature end, but it also affects the cold end of brightness temperatures to a small degree. These parameters are chosen as optimization parameters in the first harmonization run because of their potential impact on the brightness temperature, their easy technical optimization (single constants) and because of the undocumented origin of the operational values for the parameters.

Directly after the harmonization attempt, the harmonization output statistics on remaining inter-satellite differences provide information on the successful completion of the harmonization process in principle. The harmonization output provides the K-residual, i.e. the remaining inter-satellite difference. In case of successful harmonization, the K-residual is normally distributed around zero. If this is not the case, the

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

harmonization could not randomize the inter-satellite differences, i.e. unexplained biases remain. This can be due to bad choice of harmonization parameters or due to instrumental effects that were not considered.

In the case of successful harmonization according to the harmonization output statistics, we validate the reprocessed data. To validate this obtained first version harmonized FCDR, we carry out the same analysis of statistics as for the operational data: we compute monthly statistics of the operational and of the harmonized FCDR data. These statistics encompass the 10th to 90th percentile per month. Hence, the investigation of the differences in the percentiles between the sensors provides an impression of the inter-satellite bias across the covered brightness temperature range. Results are discussed in the following.

The harmonization diagnostics indicate that the bias reduction was successful for the MHS on Metop-A and on Metop-B as well as for the AMSU-B on NOAA-17. This is conveyed for example by the resulting distribution of the K-residual for these instruments matched against NOAA-18. The K-residual measures the remaining difference between the instrument and the reference in the input data set of the used match-ups after harmonization. After successful harmonization, the K-residual should be normally distributed around 0 K, indicating that a bias reduction in the input data was achieved and only random differences remain. This is the case for the MHS on Metop-A and on Metop-B as well as for the AMSU-B on NOAA-17. Figure 13 shows the difference in brightness temperature for Metop-A - NOAA-18 before and after harmonization (note that there is little change as the bias of Metop-A and NOAA-18 is very small already before harmonization).

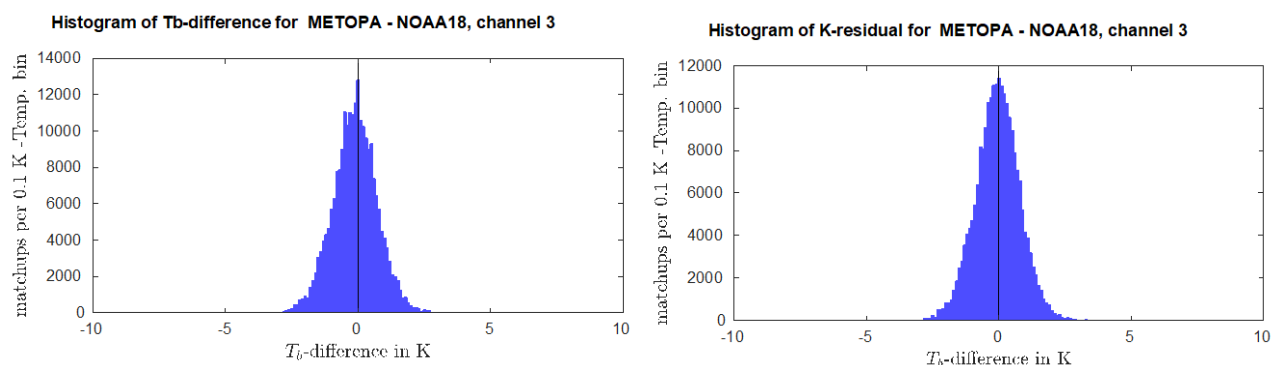


Figure 13 Histogram of brightness temperature differences (K-residual) between MHS on MetOp-A and on NOAA-18, channel 3, before (left panel) and after (right panel) harmonisation.

Yet, a normal distribution around zero is not achieved for NOAA-16 and NOAA-19 (for NOAA-15, the asymmetry of the K-residual histogram is only very weak). For these instruments, a bias remains which was not reduced by optimizing the non-linearity correction and the warm target correction. Figure 14 shows the difference in brightness temperature before and after harmonization for NOAA-16 - NOAA-18. Note that the K-residual is shifted to -0.5 K. Hence, there is an asymmetry indicating that the differences between the sensors were not randomized. This remaining bias is discussed at the end of this section.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

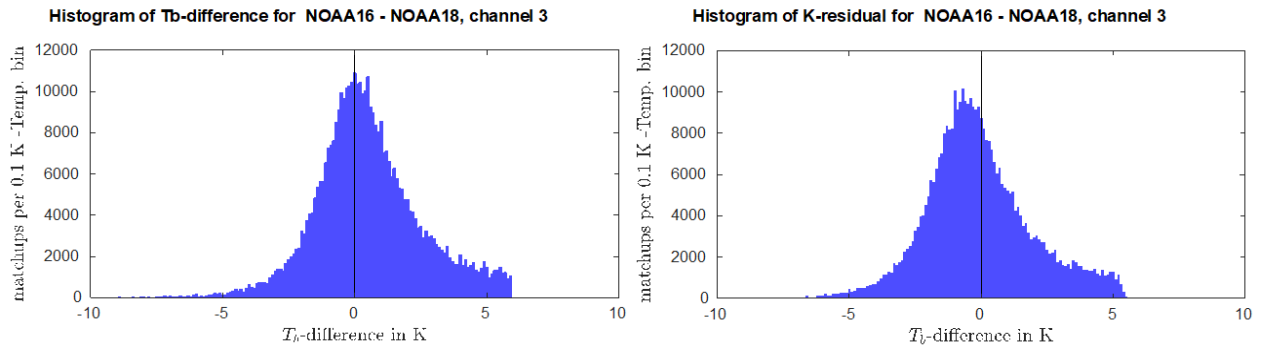


Figure 14 Histogram of brightness temperature differences (K-residual) between AMSU-B on NOAA-16 and MHS on NOAA-18, channel 18, or 3 resp., before (left panel) and after (right panel) harmonisation.

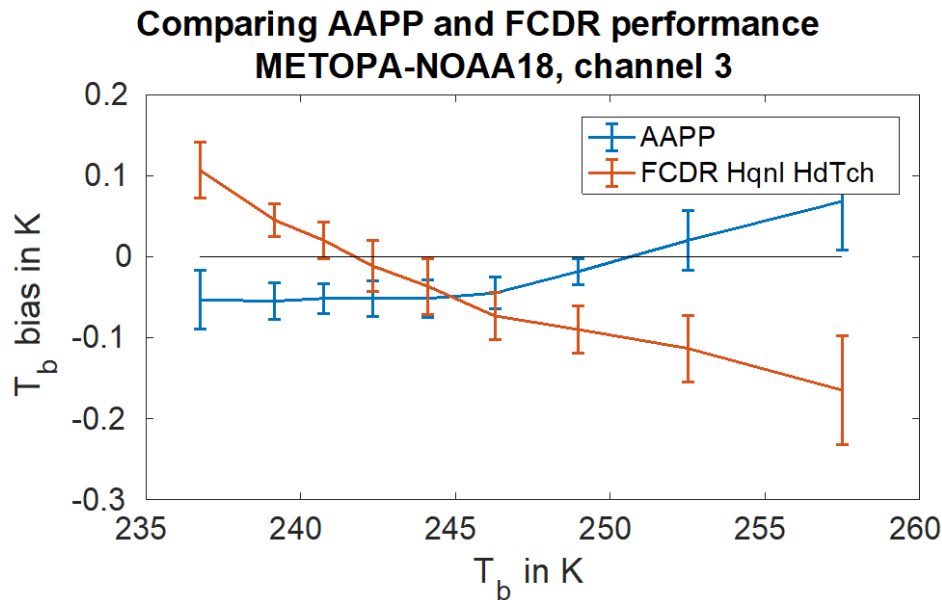


Figure 15 Comparison of resulting biases for AAPP and harmonised FCDR processing for MHS on MetOp-A, channel 3 against MHS on NOAA-18 (for September 2014 - February 2015). Harmonisation parameters are non-linearity and warm target correction. The uncertainty bars represent the standard deviation of the bias over the six months of used data.

For the successful cases of Metop-A, Metop-B and NOAA-17, the optimized parameters for the non-linearity and the warm target correction can be used in the measurement equation to reprocess all data from level 1b to level 1c in order to obtain a harmonized FCDR. In order to test the first harmonization results, we process half a year of Metop-A data to a harmonized FCDR, providing us with a first impression of the overall success of the harmonization. In Figure 15, the bias in the monthly percentiles (10th to 90th percentile) of MHS on Metop-A - NOAA-18 is displayed as function of the respective brightness temperature of the percentiles. The bias is shown before harmonization (operational AAPP processed data, blue) and after harmonization (using optimized parameters in the FCDR generation, orange). A reduction of the bias is achieved at 239 to 245 K, while it is even increased in other ranges of brightness temperature. The reduction is mainly achieved in the temperature range where most match-ups occur (note that SNOs usually occur at high latitudes). The remaining temperature ranges are poorly sampled by the input data based on SNOs (see Figure 16 for the abundance pattern of SNOs). Hence, the harmonization cannot optimize the calibration for these ranges. Another reason for the failure of the bias reduction over wide intervals of the temperature

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

range is that the chosen combination of optimization parameters may not be the correct choice to account for the true cause of the bias.

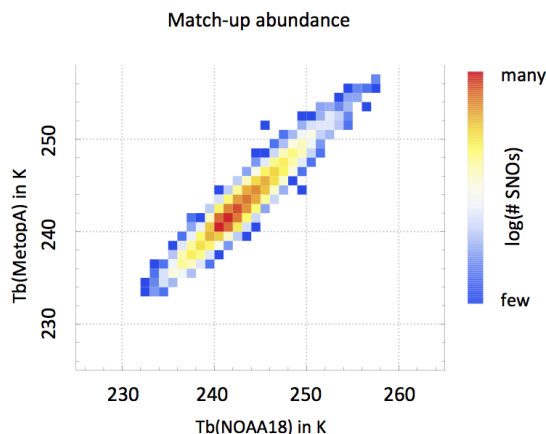


Figure 16 Abundance pattern of brightness temperatures for FIDUCEO match-ups of MHS on Metop-A and on NOAA-18.

The first results of the harmonization showed two problems that prevent us from generating the harmonized MW FCDR at this stage. First, the chosen parameter combination of non-linearity and warm target correction may not be the correct choice to account for the cause of the biases. This deficiency of the current harmonization is not unexpected, since the harmonization procedure is thought of as an iterative process. This iteration of re-running the harmonization with different optimization parameters is expected to finally discover the best combination of calibration parameters that required optimization in order to reduce the inter-satellite bias. Hence, executing more harmonization runs with different optimization parameters is expected to improve our understanding of the origin of the biases. Another modification of the harmonization can be achieved by using a different reference instrument. Because of the possible deficiencies of MHS on NOAA-18 (for unknown reasons it received the same antenna pattern correction as the AMSU-Bs), it is interesting to execute the harmonization with another reference instrument, such as MHS on Metop-A, for example. The second problem of the harmonization is the poor sampling of the brightness temperature range by the used match-ups. Tropical SNOs that allow extending the temperature range to warmer temperatures [John2012] are scarce. This is because there are too few periods of similar LECT for two satellites, and moreover, a similar LECT is a necessary, but not sufficient condition for global SNOs. Hence, tropical SNOs are rare which induces the poor sampling of the warm temperature end. As long as the vast majority of match-ups covers only a very small range of a few Kelvin around 242 K, the harmonization has no knowledge of what the optimization of certain parameters may produce at the warm temperature end. Therefore, it is highly required to sample the warm temperature end adequately. A promising method for the collection of “warm match-ups” was developed in parallel to the harmonization efforts and is expected to improve the harmonization procedure.

The failure of the harmonization for the AMSU-Bs on NOAA-15 and NOAA-16 as well as of the MHS on NOAA-19 indicates that it was not possible to randomize the differences between these instruments and the reference NOAA-18, even in the restricted range of brightness temperatures covered by the match-ups. The underlying problem may be related to issues with radio frequency interference (RFI), which introduces biases. This problem is known for NOAA-15 from the in-orbit verification phase [Atkin1998,2001]. And for

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

NOAA-16 as well, a growing influence of RFI due to a decreased gain was suggested in [John2013]. In Hans2019b, we analyze in detail this issue for AMSU-B on NOAA-16 and MHS on NOAA-19 and we find high evidence that RFI causes the strong inter-satellite biases observed for NOAA-16, NOAA-19 and NOAA-15. We also developed a new RFI-correction scheme that is applied in the calibration process for the final FIDUCEO FCDR validated in Section 3.3 and presented in Hans2019a.

3.2.1.2 Harmonisation results using four parameters

Any further harmonization attempts were mostly executed for Metop-A, in order to investigate further the applied optimization procedure and its dependence on the used match-up data and chosen harmonization parameters.

In further harmonization runs, we extended the set of “cold matchups” (obtained from Brockmann Consulting) by “warm matchups”. The warm matchups were obtained with the newly developed method of using the geostationary IR-sensor SEVIRI as filter for constant brightness temperature (in the water vapor channel) over a certain time window at a certain location [Prange2019]. The temperature range of these matchups is around 260 - 270 K, which is significantly warmer than most of the cold matchups (around 240 K). With this extended temperature range by combining the two matchup data sets, the dynamical range of brightness temperatures is better sampled and should allow for better constraints on the harmonization problem.

However, the better sampling of the brightness temperatures alone does not lead to a successful harmonization in the sense that the biases are reduced consistently. We have to allow the harmonization to optimize more parameters at once. In principle, we should optimize all calibration parameters that we do not fully trust. This would include the antenna pattern correction, which is extremely difficult to access as optimization parameter. Instead, we use the four parameters of non-linearity coefficient, warm target correction, cold target correction and polarization correction in the next step. In the following, we present the attempts to harmonize the MW data using these four harmonization parameters, and the cold and warm matchups as training data.

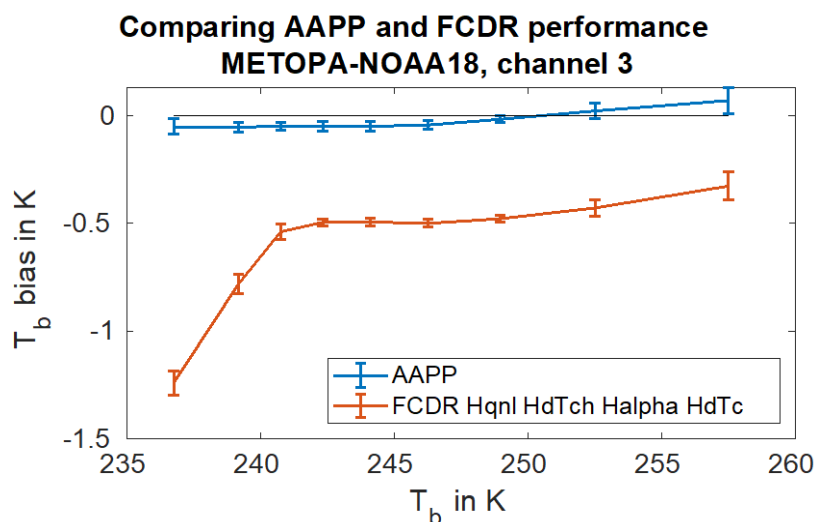


Figure 17 As Figure 15, but the harmonisation parameters are non-linearity, warm target correction, polarisation correction and cold target correction. The uncertainty bars represent the standard deviation of the bias over the six months of used data.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

The four-parameter harmonization run with warm and cold matchups for Metop-A – NOAA-18 is successful on the matchup data themselves. However, when applying the harmonized coefficients for the nonlinearity, the warm target correction, the cold target correction and the polarization correction within the FCDR processing, the overall biases are not reduced but rather increased. This can be seen in Figure 18 showing the differences of the monthly percentiles for Metop-A – NOAA-18 when processed with AAPP or with the harmonized FCDR code. The bias is about -0.5K for the harmonized FCDR over wide ranges of temperature, although the harmonization procedure itself was successful in reducing the bias on the matchup data set. Having excluded the possibility of coding errors in either the FCDR processing or the harmonization code, we have investigated this issue further.

The optimized parameters for nonlinearity and polarization correction are optimized such that they reduce the bias on the matchup data that contain only near-nadir matches. The FCDR however, is processed for all angles. For the off-nadir angles, the polarization correction has only little influence and hence, the optimized parameter only has a small impact. However, the optimized nonlinearity coefficient has the same impact over all FOVs. Hence, its influence is not compensated for the off-nadir angles as it was for the nadir ones (optimized in harmonization). Therefore, the large new value for the non-linearity may now introduce artificial biases in the off-nadir data and hence distort the overall performance of the FCDR, which leads to the increased bias when looking at monthly percentiles from all scan angles. To check this suspicion of near-nadir/ off-nadir differences due to the polarization correction, we display the percentiles of the matchups data in Figure 18 and compare it to the FCDR monthly percentiles using near-nadir data only, shown in Figure 19.

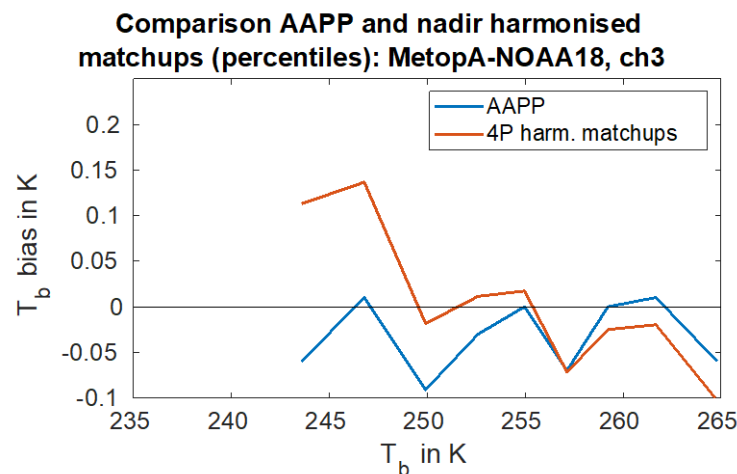


Figure 18 Comparison of 10th to 90th percentile of brightness temperature in the matchup data before harmonisation (AAPP) and after harmonisation. The matchup data contain only near-nadir views. The four harmonisation parameters are non-linearity, warm target correction, polarisation correction and cold target correction.

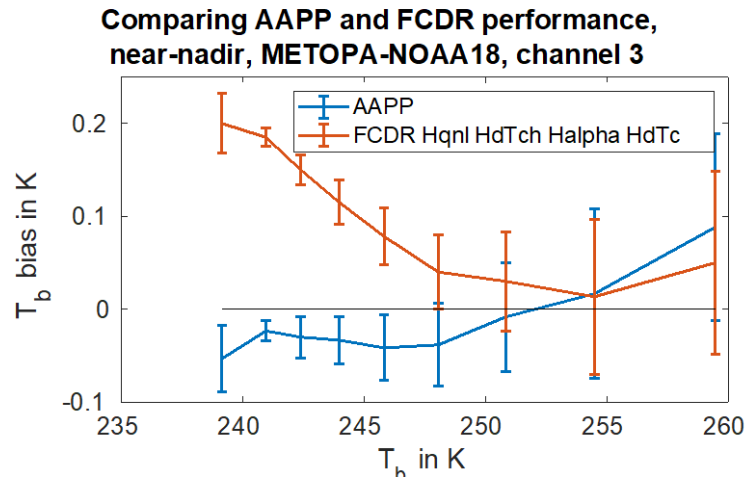


Figure 19 As Figure 15, but for near-nadir views only. Note the difference to and the similarity to Figure

The difference Metop-A – NOAA-18 is very similar in both plots. Note that the upper plot on percentiles of the matchups corresponds to a slightly shifted T_b -range with respect to the lower plot. Note also that the increase seen in the monthly percentiles for $T_b > 255$ K in Fig. 19 is probably due to diurnal cycle effects that do not show up in the match up data (the same increase of bias can be seen for percentiles of Metop-B, which samples the same time of the day as Metop-A).

Summarizing, we see an increased bias in the harmonized FCDR for monthly percentiles using all viewing angles. Doing the same exercise for near-nadir views, we obtain a small bias for the harmonized FCDR for monthly percentiles that resembles the bias that is observed in the percentiles from the matchups themselves (by construction only near-nadir). This corroborates the suspicion that the off-nadir views are wrongly calibrated when using the harmonized parameters that were obtained for near-nadir views only.

Concluding, we suspect that using the polarization correction as harmonization parameter while only using near-nadir matches, will not result in consistent harmonization. Moreover, having added the warm matchups to the cold ones, the idea is to further extend the range of temperatures by also adding the very cold matchups including very dry scenes over Antarctica (this comes at the cost of including cloudy pixels as well, but this only increases the noise). A first harmonization test on Metop-A, performed only shortly before the end of the project, using this further extended matchup data set (warm, cold, very cold matchups) shows promising results. Hence, further improvement of the sampling of temperatures and scan angles may yield a more stable harmonization in the future.

As the harmonization did not yet reach the required level of reliability, and the improvement of the sampling for the harmonization is still in progress, we did not process the whole FCDR time series with trial harmonization parameters (as the harmonized parameters so far distort the overall calibration if including all data, it does not make sense to process the whole archive).

From the above analysis, we deduct that the harmonization problem needs to be constrained adequately in both temperature range and viewing angles. To achieve this, a homogeneous sampling of temperatures and

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

viewing angles is required in the harmonization input data, i.e. the matchup data. Efforts to extend the data set were undertaken and show promising first results. Moreover, the harmonization parameters should reflect the complete set of uncertain calibration parameters. Consequently, the antenna pattern correction should enter the harmonization procedure. This is a complex undertaking because 270 values (180 independent values) would need to be optimized. However, they are constraint by the (uncertain) geometry of the instruments' surroundings and of the (uncertain) antenna pattern. To approximate the antenna pattern correction with a polynomial and optimize its coefficients, is a first approach, but it cannot reflect the actual complexity of the true, required antenna pattern correction.

An important argument that made us investigating alternatives to harmonization, in order to achieve a consistent calibration, is the strong and time dependent bias seen for NOAA-15, NOAA-16 and NOAA-19. These biases are the most important ones that actually prevent the construction of consistent time series, and these biases cannot be reduced by the harmonization method as it is designed. The reason for this is the origin of the bias, which is most probably RFI. Since this effect influences the very raw counts, it has an impact on the measurement even before the calibration takes place. Consequently, it is not part of the measurement equation in the form of a simple constant parameter. Rather, the RFI varies arbitrarily from viewing angle to viewing angle, and also over time. To correct this disturbing effect, we developed a scheme to be applied on the raw counts before entering the measurement equation. This reduces the biases significantly [Hans2019b]. In principle, it is now possible to use this scheme on the counts in the measurement equation also during harmonization and allow for small changes (a few counts) of this scheme. A good sampling of all viewing angles with exact matches of angles is highly required for such an undertaking. This is not possible in the remaining time frame, however. Consequently, the FCDR is based on an improved measurement equation. The most important aspects in this improvement are the correction of assignments of parameters to channels/ instruments, and the application of the newly developed RFI-correction scheme.

3.2.1.3 *Harmonisation results using three parameters*

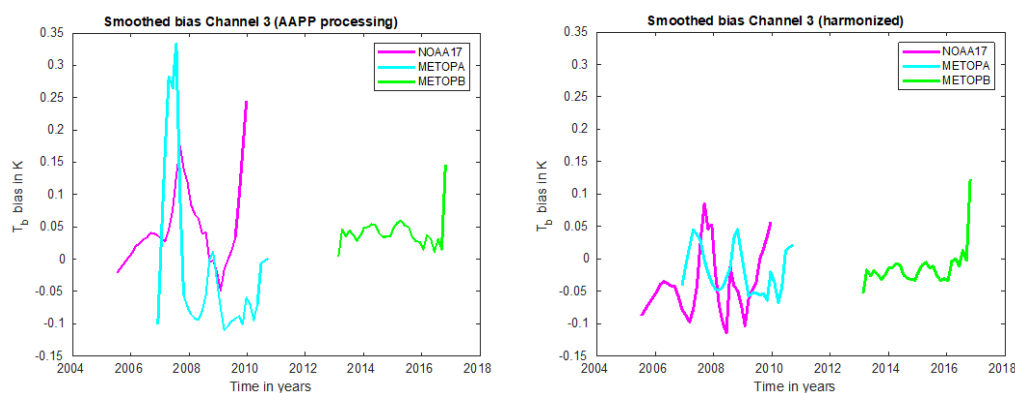
In the next step, harmonisation tests with the extended matchup sets mentioned in the previous section were carried out with AMSU-B on NOAA-17 and MHS on Metop-B as well. In each case MHS on NOAA-18 was used as reference, and very cold matchups with scene temperatures far below 240 K were included. In addition to this extension of the range of temperatures, we included matchups of any pixel numbers between 40 and 49 as well as matchups of pixels with the same distance from nadir, e. g. no. 1 with no. 90, no. 2 with no. 89, etc. This way we made sure that the geometry of the viewing direction relative to the atmospheric layers was almost the same for both observations of a matchup.

FastOpt calculated harmonisation parameters for AMSU-B on NOAA-17 and MHS on MetOp-A and MetOp-B, i. e. those instruments that were only little or not at all affected by RFI. The new values differed much less from the corresponding default values in AAPP (non-linearity coefficient, antenna reflectivity, and cold space correction factor) than those obtained from the first harmonization runs. When they are employed in the measurement equation, they improve considerably the accuracy of the measured brightness temperature (see Figure 20).

The mean value for the differences in all match-ups from a given instrument pair is listed in Table 1. As expected, the bias against NOAA18 is in each case significantly reduced after harmonization. The precision did not improve, because the random scatter, which is due to noise of the receivers and mismatch

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

uncertainties, cannot be removed. The numbers in the table were calculated without any averaging or smoothing.



20 Before and after harmonization: The differences $X - \text{NOAA18}$ with $X = \text{NOAA17}$, METOPA, and METOPB, are plotted as a function of time. Mean values for all matchups in intervals of 1.5 months were calculated, and then the time series was smoothed.

	Bias [K]	Standard Deviation [K]
NOAA-17 (harmony) – NOAA-18	-0.021	1.47
NOAA-17 (AAPP) – NOAA-18	+0.058	1.45
Metop-A (harmony) – NOAA-18	-0.004	1.33
Metop-A (AAPP) – NOAA18	-0.039	1.33
Metop-B (harmony) – NOAA-18	-0.025	1.05
Metop-B (AAPP) – NOAA-18	+0.039	1.05

Table 4-1: Change of bias of different instruments relative to MHS on NOAA-18 due to harmonization

It would be interesting to apply the harmonization method of FastOpt to the raw data of NOAA-15, NOAA-16, and NOAA-19 after RFI correction and with the correct APC for both NOAA-18 and Metop-A. Basically harmonization only makes sense once the main calibration errors in the processing pipeline have been eliminated.

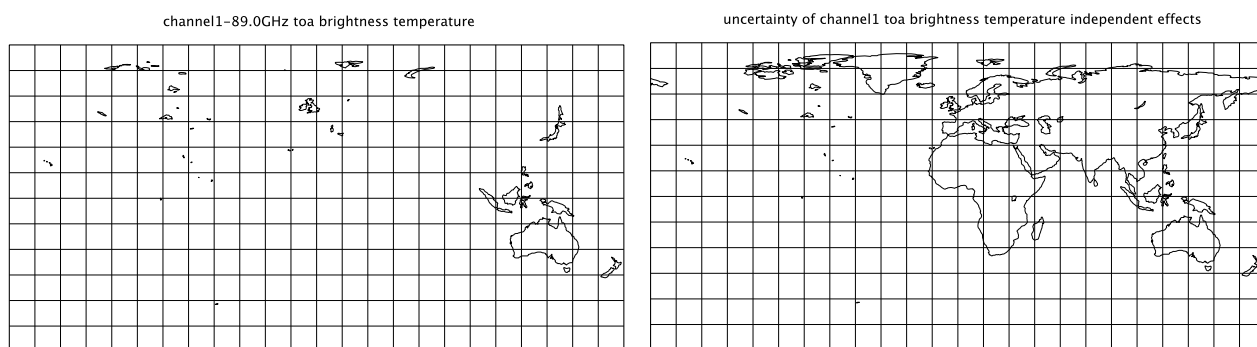
3.3 Mutual/ relative consistency within sensor series

3.3.1 Assessment of uncertainties

Because of the lack of uncertainty information in the operational data, a comparison of the uncertainties stored in the new FCDR to other uncertainties is only possible to a limited degree. A comparison of the uncertainties due to independent and structured effects to the instrumental specification for NEdT is possible, since both classes of uncertainty are dominated by either the noise on the Earth counts or the noise on the calibration counts, which is also the important aspect in the estimate of NEdT. The comparison reveals a fair agreement of deduced uncertainties from the FCDR processing and the reported specified NEdT [NOAA KLM User Guide]. During periods of erratic instrumental behavior, such as the later years of AMSU-B on NOAA-15 and on NOAA-16 as well as of MHS on NOAA-19 when the instruments failed the NEdT specifications, the uncertainties of the brightness temperature stored in the FCDR reflect well this erratic

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

behavior by an increased uncertainty due to independent and structured effects. The uncertainty due to common effects has no equivalent in the available data sets. Hence, a first analysis of its characteristic distribution over an orbit is especially interesting. Figure 21 and Figure 22 show the uncertainty due to common effects for channel 1 and 3, respectively. Note that the uncertainty introduced by the antenna pattern correction is the strongest contribution for channel 1; the picture of a high uncertainty on the edges of the swath and a low uncertainty around nadir clearly shows the dominance of the antenna pattern correction over the other effects. This dominance arises because of the large operational correction value and the input uncertainty estimates (50% of the operational correction value). Note that channel 3 has a much smaller operational correction. Hence, using the same estimate of 50% leads to a smaller uncertainty than for channel 1. For channel 3, other effects gain influence as can be deduced from the more prominent temperature dependence (compare the brightness temperature in 10 to the common uncertainties). This temperature dependence of the uncertainty with large uncertainties in cold regions and small uncertainties in warm regions is due to the polarization correction determined by the coefficient α . This correction also introduces a scan dependence, which leads to smaller uncertainties at the edge of the swath and to larger uncertainties near nadir. This is by construction of the correction formula, which accounts for the polarization sensitivity of the receiver in connection with the different reflectivity of the rotating mirror for horizontally and vertically polarized radiation. Note that we used 100% of the operational values for the coefficient α as input uncertainty estimate (operational value from NOAA-18 as no other instrument receives a non-zero correction, see above). Nonetheless, it is still the antenna pattern correction that dominates the scan dependence of the uncertainty. At the right edge, the uncertainty is increased compared to nadir. At the left edge, the antenna pattern correction has less influence. A comparison of the three classes of uncertainties is visible in Figure 23, showing histograms of the respective uncertainties for each class for one month of MHS data (MHS on Metop-B, September 2014). The histograms for the three sounding channels resemble each other, although the exact positions of the distributions vary slightly. For channels 1 and 2, the picture is changed as either the independent and common uncertainties, or the structured and common uncertainties overlap. The long tail for the common effects in channel 1 reaching 0.8 K is due to the antenna pattern correction affecting the outermost pixels most strongly. Most of the data, however, have common uncertainties below 0.3 K also for channel 1. Except for channel 1, the independent uncertainties dominate the other classes. From the perspective of a climate scientist, this is encouraging, as averaging procedures used for climate studies can reduce the uncertainty due to independent effects. Averaging cannot reduce the smaller, common uncertainty, however, and hence, its relative importance increases on longer times scales.



D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

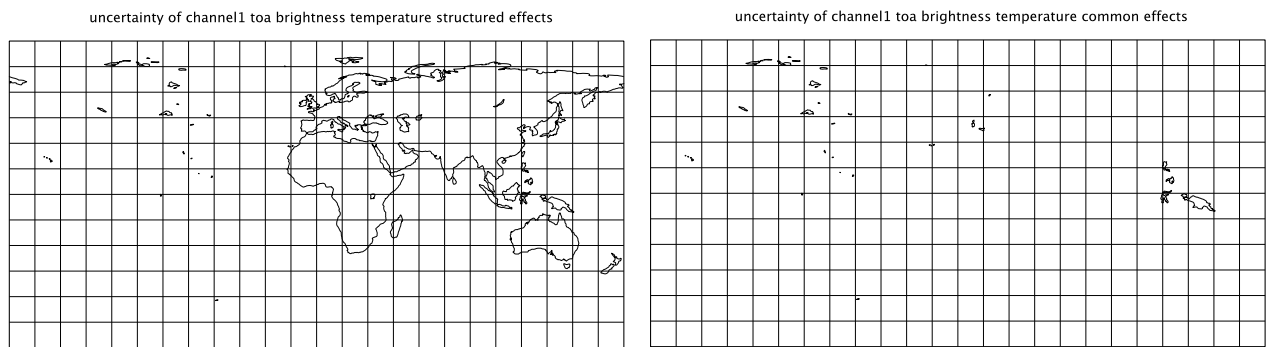


Figure 21 Channel 1: Brightness temperature and its uncertainties (for an example orbit of MHS on Metop-B in July 2015). Upper left: brightness temperature. Upper right: uncertainty of brightness temperature due to independent effects. Lower left: uncertainty of brightness temperature due to structured effects. Lower right: uncertainty of brightness temperature due to common effects.

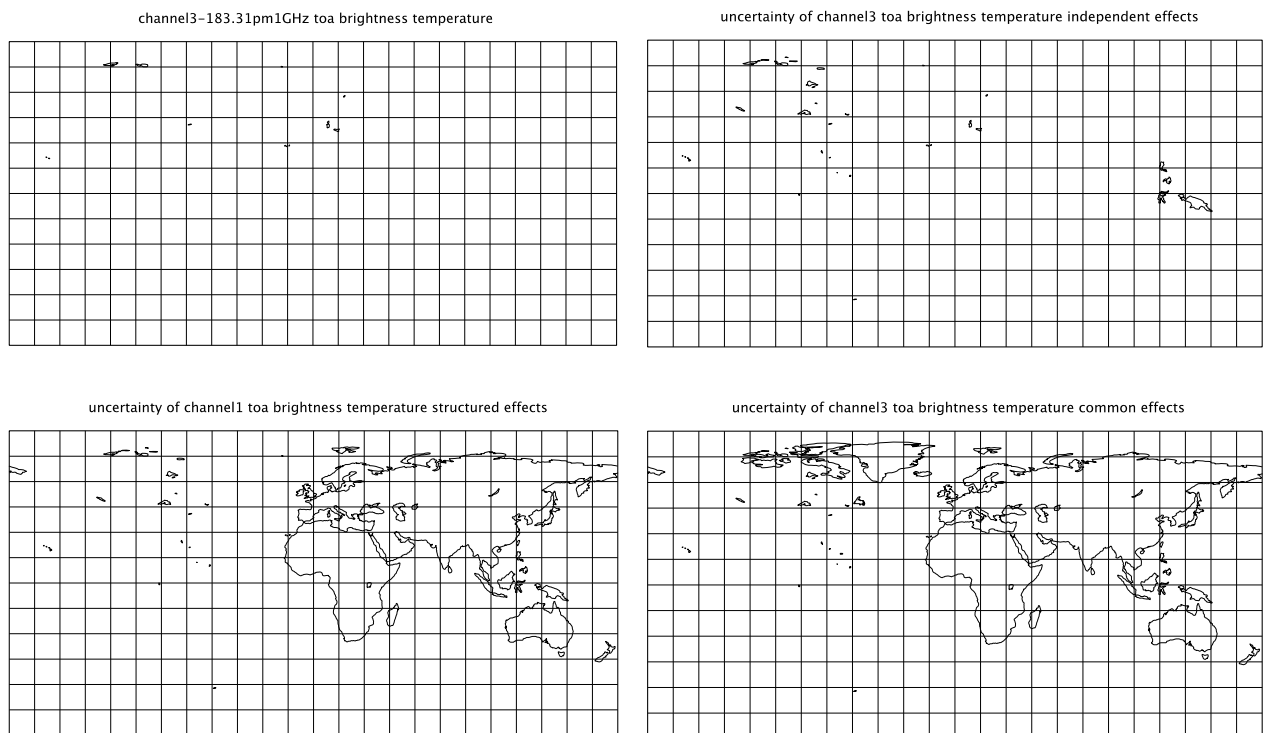


Figure 22 As for Figure 21, but for channel 3.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

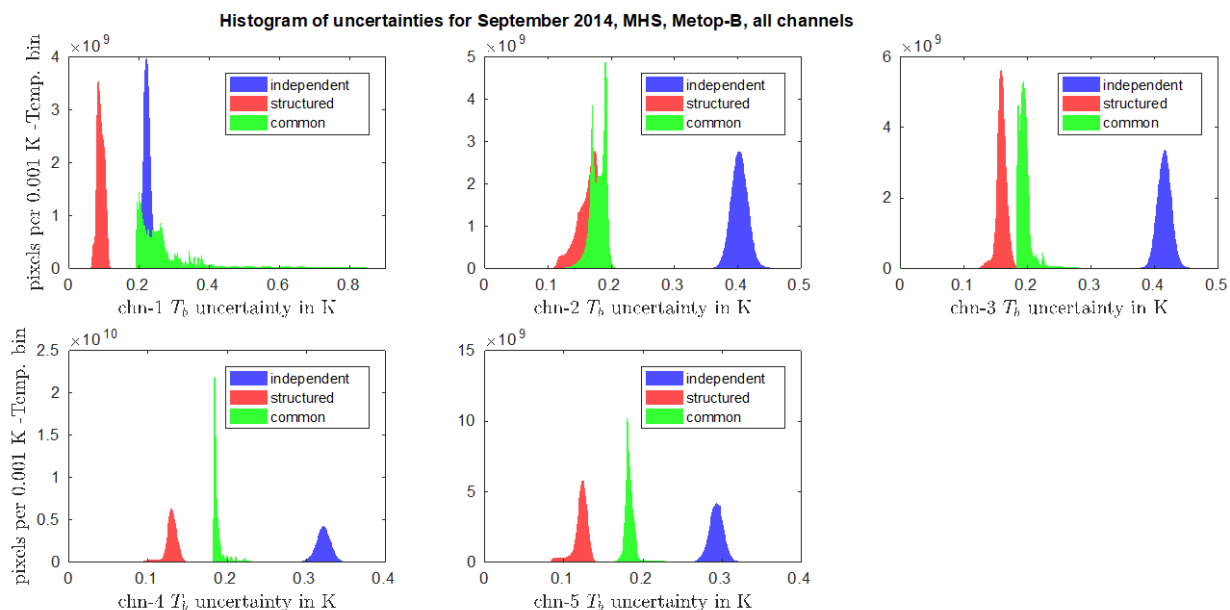


Figure 23 Histogram of the three classes of uncertainties for September 2014, MHS on Metop-B, all channels.

3.3.2 Assessment of stability and consistency

The reduction of inter-satellite biases and the corresponding improved consistency is visible in Figure 24, which shows the time series of inter-satellite biases vs NOAA-18 of global monthly means, averaged over all scan angles and both ascending and descending passes. Both, the operational and the FCDR data are displayed.

Note that the peaks that suddenly appear in the time series (Metop-A: 2007, 2012, 2014, NOAA-19: 2009, 2012) are due to different amount of data usable for the specific month for the different instruments. The reference is MHS on NOAA-18 because of its long time series and stability. Note also that only AMSU-B and MHS instruments are shown. SSMT-2 cannot be compared to NOAA-18 because of missing overlap. Hence, SSMT-2 is not displayed in Figure 24. The improvement of the SSMT-2 calibration by the strict application of the measurement equation is visible in Figure 25 to Figure 29 showing the time series of the brightness temperature for all instruments and channels for the operational data and the new FCDR data.

Channel 1 is affected by the correction of APC for NOAA-18 and Metop-A. No RFI correction is applied (since the method for deriving the correction does not work for channels strongly affected by the diurnal cycle). There is reasonable agreement between the instruments. Note for example the increasing amplitude of the seasonal cycle for MHS on NOAA-19 as the satellites drift apart. At the same time, the amplitude decreases for Metop-A as NOAA-18 drifts closer to it. The AMSU-B instruments show an offset with respect to the MHS instruments. Note that the channel does not have exactly the same characteristics for AMSU-B and MHS (they differ in bandwidth and location of the two pass bands).

Channel 2 shows very similar results as channel 1. However, AMSU-B on NOAA-15 differs from the other AMSU-Bs. This is because the channel is not RFI-corrected in the FCDR (correction scheme cannot be derived with the applied method). The matching of NOAA-15 to NOAA-18 is not real, since channel 2 differs for MHS and AMSU-B (note the off-set for the other AMSU-Bs).

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

Channel 3 is affected by the correction of APC for NOAA-18 and Metop-A. Moreover, the RFI-correction scheme for NOAA-15, NOAA-16, NOAA-17 and NOAA-19 affects this channel. A clear improvement in the reduction of the inter-satellite bias is visible. And as well the individual instrumental stability has improved, especially for NOAA-16. For NOAA-15, the derived RFI-correction scheme is imperfect, because of its strong RFI contamination already in the early years that can be compared to NOAA-18. This imperfection is represented by an increased uncertainty due to common effects.

For channel 4 and 5, similar improvements as for channel 3 can be seen.

Validating the harmonization of Metop-A with matchups over a large range of scene temperatures in a similar way as we did for the first harmonization runs, we obtain the following result. We plot the average inter-satellite bias (Metop-A – NOAA18, from 121,830 match-ups) as function of brightness temperature. This is shown in Figure 30 for the operational AAPP processed data (blue) and the harmonized AAPP data (orange) of channel 3 from November 2006 to September 2010. This validation shows that the improved harmonization using three parameters does not distort the calibration as the application of the trial harmonization parameters did (compare Figure 15 and Figure 30, bearing in mind that the bias in the figures was calculated for quite different time periods and with quite different methods). Quite the opposite, it results in an overall improved consistency of Metop-A with respect to NOAA-18. This is the case in particular for the lowest scene temperatures, which can only be found over Antarctica without cloud filtering.

Overall, the applied improvements in the parameters of the measurement equation and the correction for RFI contamination have increased the inter-satellite consistency. The instruments agree within the uncertainties due to common effects. The harmonization procedure shows potential for further improvements.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

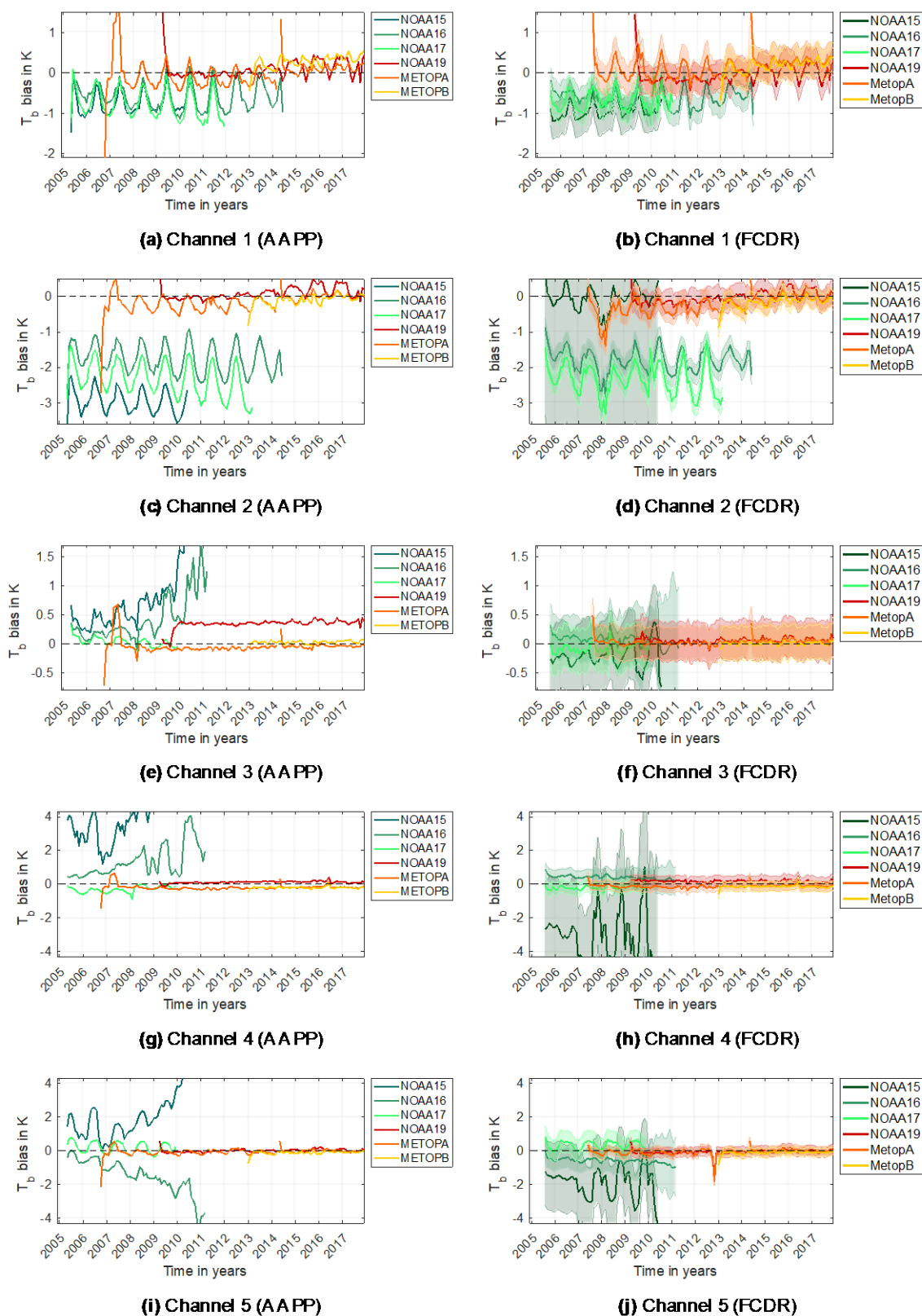


Figure 24 Comparing inter-satellite biases of MHS and AMSU-B instruments against MHS on NOAA-18 in operational AAPP data (left) and FCDR data (right) as function of time. The FCDR provides more consistent and stable time series due to improved calibration. The instruments agree within the uncertainties (only uncertainties due to common effects displayed here). Note that channel 2 differs slightly in frequency for MHS and AMSU-B and therefore are not expected to match completely.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

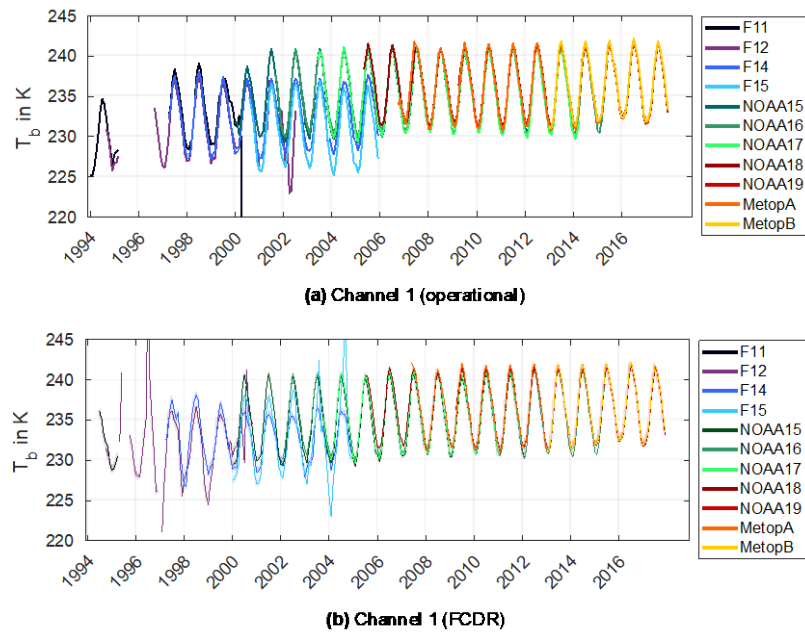


Figure 25 Comparing the time series of brightness temperature in channel 1 for all MW instruments in operational data (NGDC for SSMT-2 and AAPP processed data for AMSU-B and MHS) and in FCDR data as function of time. The FCDR provides more consistent and stable time series due to improved calibration. For the FCDR, the shaded regions denote the uncertainty due to common effects (hardly visible for channel 1 on this scale). There is no corresponding uncertainty estimate on the operational data.

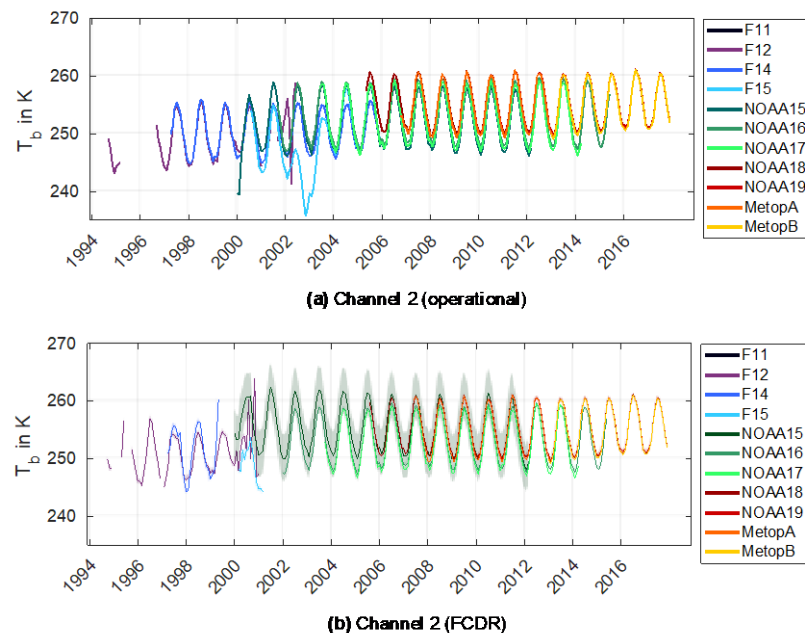


Figure 26 As figure 25, but for channel 2. Note the uncertainty for NOAA-15 due to uncorrected RFI.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

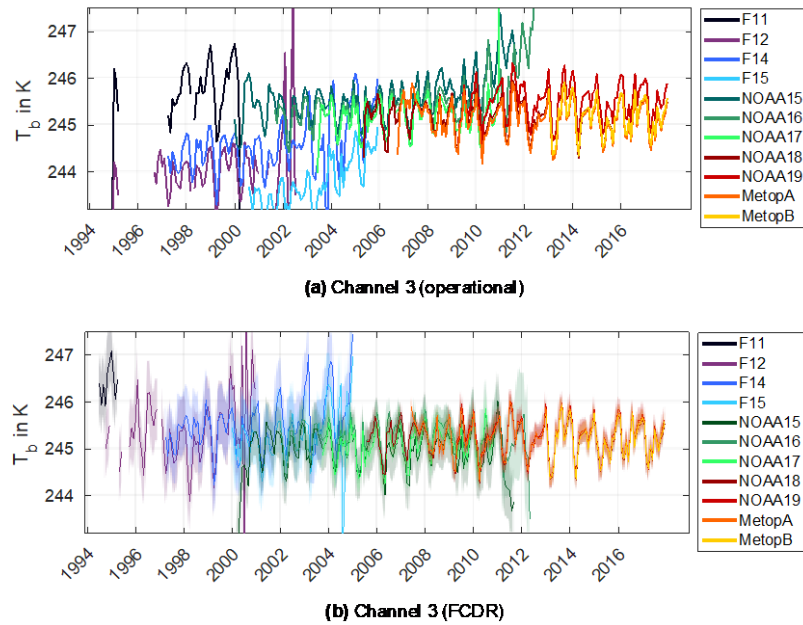


Figure 27 As Figure 25, but for channel 3. The instruments agree within the uncertainties. A distinct improvement is also achieved for SSMT2.

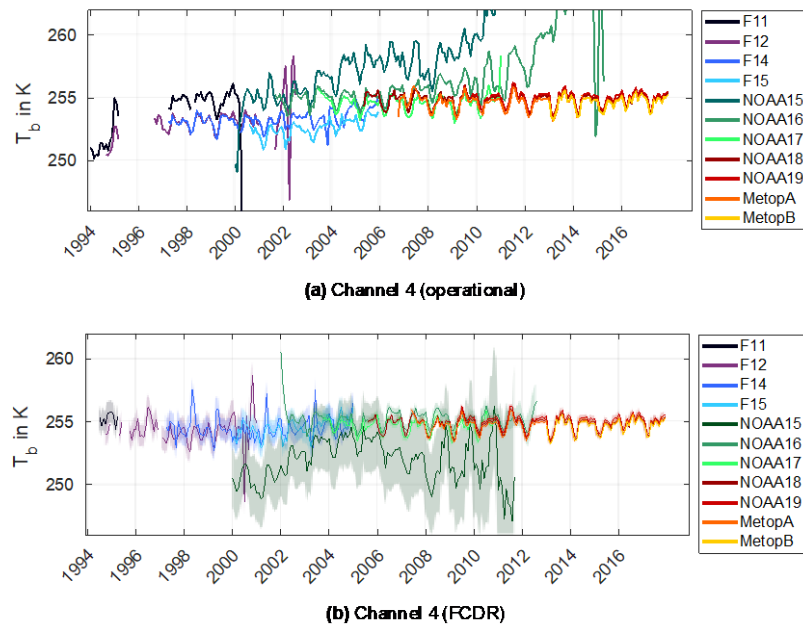


Figure 28 As Figure 25, but for channel 4. Note that NOAA-15 still shows differences with respect to the other instruments. This is due to the imperfect RFI correction.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

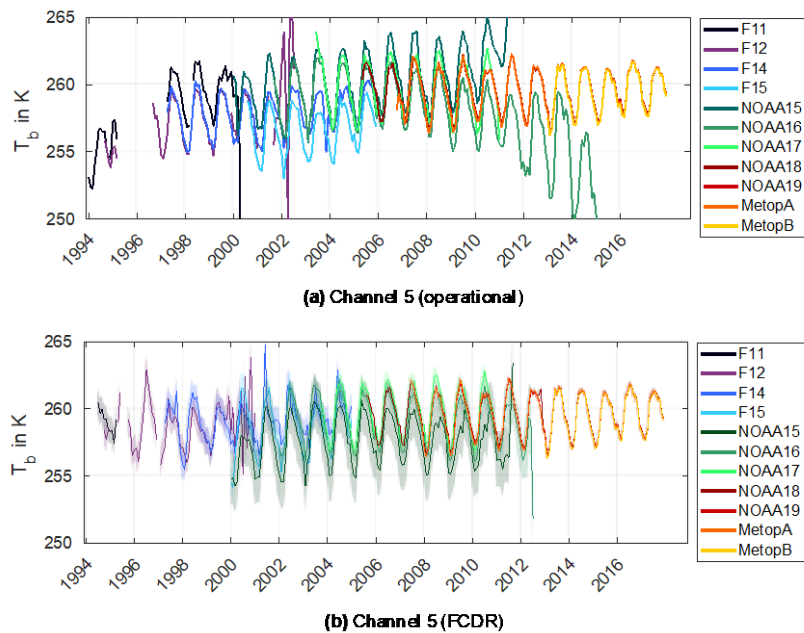


Figure 29 As Figure 25, but for channel 5.

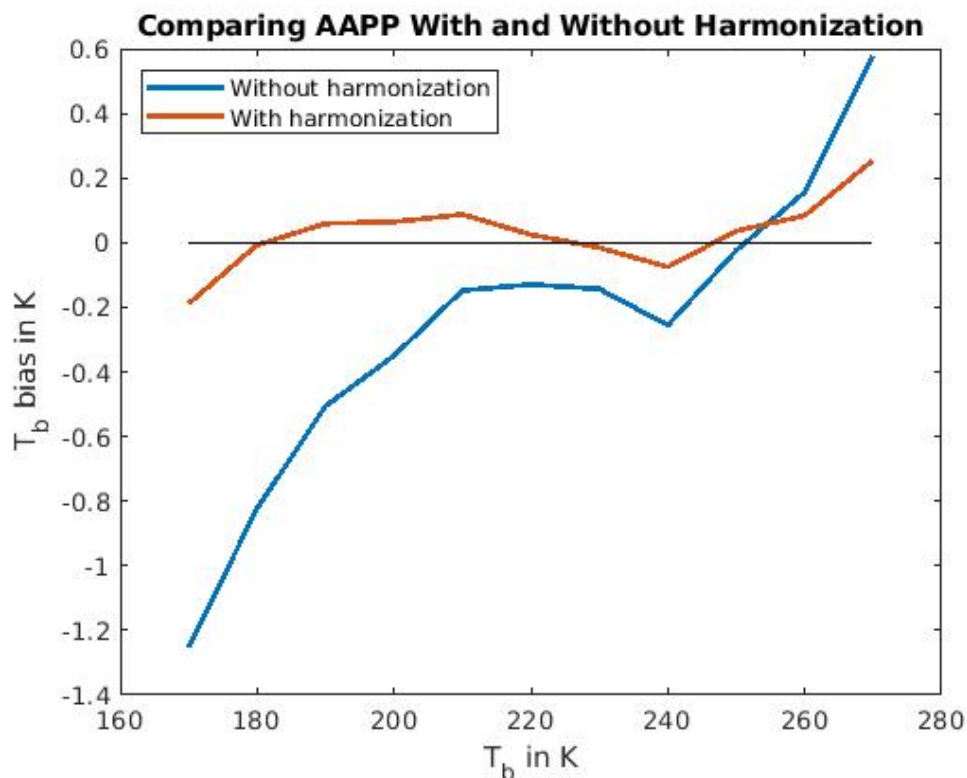


Figure 30 Comparison of resulting biases for AAPP with and without harmonization for MHS on MetOp-A, channel 3, against MHS on NOAA-18 (for November 2006 - September 2010). 121,830 matchups between these instruments were used to calculate the bias, so in most cases the uncertainties due to random scatter of the differences between MetOp-A and NOAA-18 are smaller than the thickness of the lines.

D4_6 Metrological assessment of consistency, **V1.1** stability and uncertainty of FIDUCEO FCDRs

3.4 Positives of FIDUCEO methods/ FCDR

First, we facilitate the handling of the data by providing the FCDR in ready-to-use NetCDF files. Each of these files covers one orbit reaching from one equator-crossing to the next in the same flight direction. This equator-to-equator frame removes all doubled data and hence prevents the appearance of corresponding sampling artefacts. The FCDR contains concise quality information on pixel level and maintains the traceability back to the original level 1b data.

Second, the FCDR provides extensive pixel-level uncertainty information, considering the correlation behavior of the underlying errors. This is communicated by three uncertainty variables (see PUG, D2.2) in the FCDR files.

Third, the large inter-satellite biases and instabilities were understood, and reduced by an improved calibration based on correcting wrong calibration parameter assignments, on a newly developed RFI-correction scheme and on strict application of the measurement equation. Remaining differences are covered by the common uncertainties.

3.5 Summary

The FIDUCEO FCDR for microwave humidity sounders addresses the three problems of operational data that we intended to solve in order to make the data usable for climate research. The third problem of inter-satellite biases may be further investigated with the FIDUCEO harmonization method to further reduce the biases. This is an on-going effort, giving a future perspective for further improvement. Overall, the generation of a consistent uncertainty quantified MW FCDR has been achieved.

D4_6 Metrological assessment of consistency, **V1.1** stability and uncertainty of FIDUCEO FCDRs

4 HIRS

4.1 Introduction

The FIDUCEO FCDR for HIRS contains a 37 year long record of satellite retrievals of radiance and brightness temperatures for wavelengths ranging from 0.69 μm to 14.95 μm . The FCDR has been produced using a metrological approach to errors and uncertainty. This leads to a much more detailed expressions of uncertainty provided in three distinct components (independent, structured and common) at a pixel-by-pixel level. Improvements over the operational data have also been added, the two most significant being a new approach to the self emission modelling, required because the calibration data is only taken every 40 scanlines so changes in self emission between calibration observations must be taken into account, together with a Harmonisation processes intended to remove any error between sensors. The FCDR also has full traceability, including links to the original files that were used initially as a basis of the FCDR, allowing users to understand how uncertainty was calculated and what processes and sampling have occurred in the FCDR production.

Long-term datasets of atmospheric variables are important to understanding weather, climate, chemistry and change, and the role of satellites, with extensive global coverage and large swath areas are a valuable constraint on the knowledge of the past. From a climate study perspective the stability of the HIRS FCDR is crucial and will be discussed further below.

4.2 Harmonisation of HIRS FCDR

The harmonisation process attempts to stabilize the continuous data record across multiple satellites, allowing a consistently calibrated, continuous record to be established without jumps that are not understood (e.g., because of instrument differences). The Harmonisation process itself uses sensor-to-sensor matchups, both HIRS to HIRS as well as HIRS to a reference which in this case is IASI, and recovers improved calibration parameters for all the sensors in a way which should leave true sensor-to-sensors differences (such as differences in the spectral responses) while removing spurious trends/errors (full details of the Harmonisation process can be found elsewhere in the FIDUCEO documents/papers).

In order to see how stable the HIRS data is we have looked for trends in the median temperature of the Earth over the complete time period covered by the FCDR. While the Earth radiance is not a metrologically constant reference, using it will provide insights into possible trends in the data record. As an example, Figure 5-1 shows the output from channel 5 (13.9 μm). The top panel, showing the retrieved median brightness temperature, appears as a line with variations on seasonal timescales. The line is very consistent despite the data being made up of many satellites with differing sensors (different coloured points). Differences from sensor-to-sensor are small and are of a size that can be accounted for from small differences between individual sensors' characteristics. This plot is also shows a downward trend in the observed brightness temperature of order 1 K/decade. This potential climate signal would otherwise have been impossible to detect without harmonisation.

Uncertainty characteristics are not as consistent between sensors as the harmonised calibration (lower panels of same figure). It is expected that the errors in individual satellites will be different, and are likely to change over time, because of the variability of components and their degradation. In this figure, the common uncertainty, shown in the bottom panel, is a good example of this, with well-defined uncertainties that are

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

small, stable and continuous for each individual satellite. There is a clear distinction between the uncertainty in pre NOAA-14 satellites, which used HIRS 1 & 2, and post NOAA-15 satellites, which used HIRS 3. The structured uncertainty also shows a split in characteristics around NOAA 14 and 15. Later sensors have a lower average structural uncertainty, but are less stable, with increases frequency of very high uncertainties. These instances of large uncertainty require further investigation, as they may be erroneous (e.g., due to ill-conditioning of how uncertainties are calculated). (Note that because of the thickness of the plotting line, the apparent incidence of such excursions is exaggerated.) The independent uncertainty both increases and becomes more time-dependent with the recent sensors, with the exception of NOAA-12. Such variability in uncertainties are related to variations in individual error sources and as shown can be quite complex. The time variability of uncertainty is important information for users.

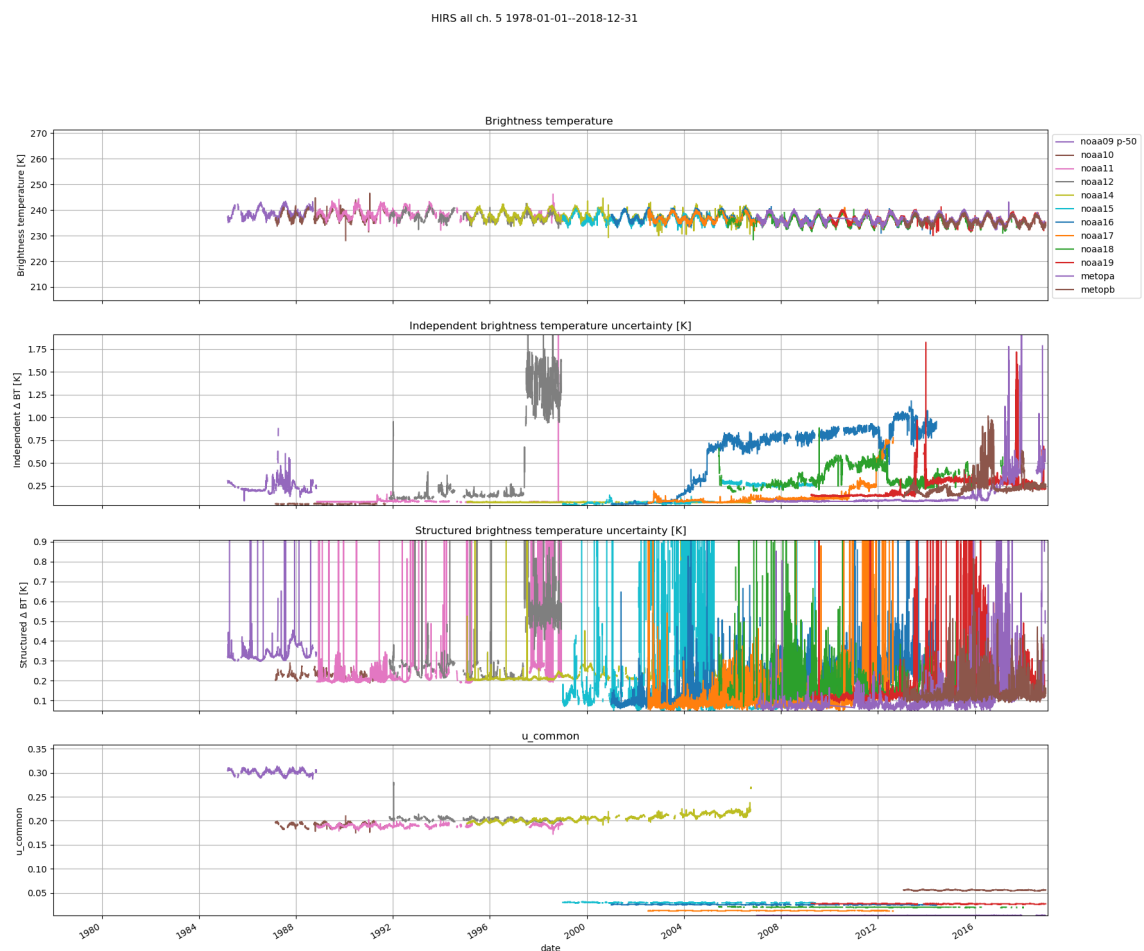


Figure 5-1 - Brightness temperature (top panel) and three uncertainty components (independent, structured and common respectively from second top to bottom) associated with Ch 5 (13.9 μm) from HIRS, after harmonisation.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

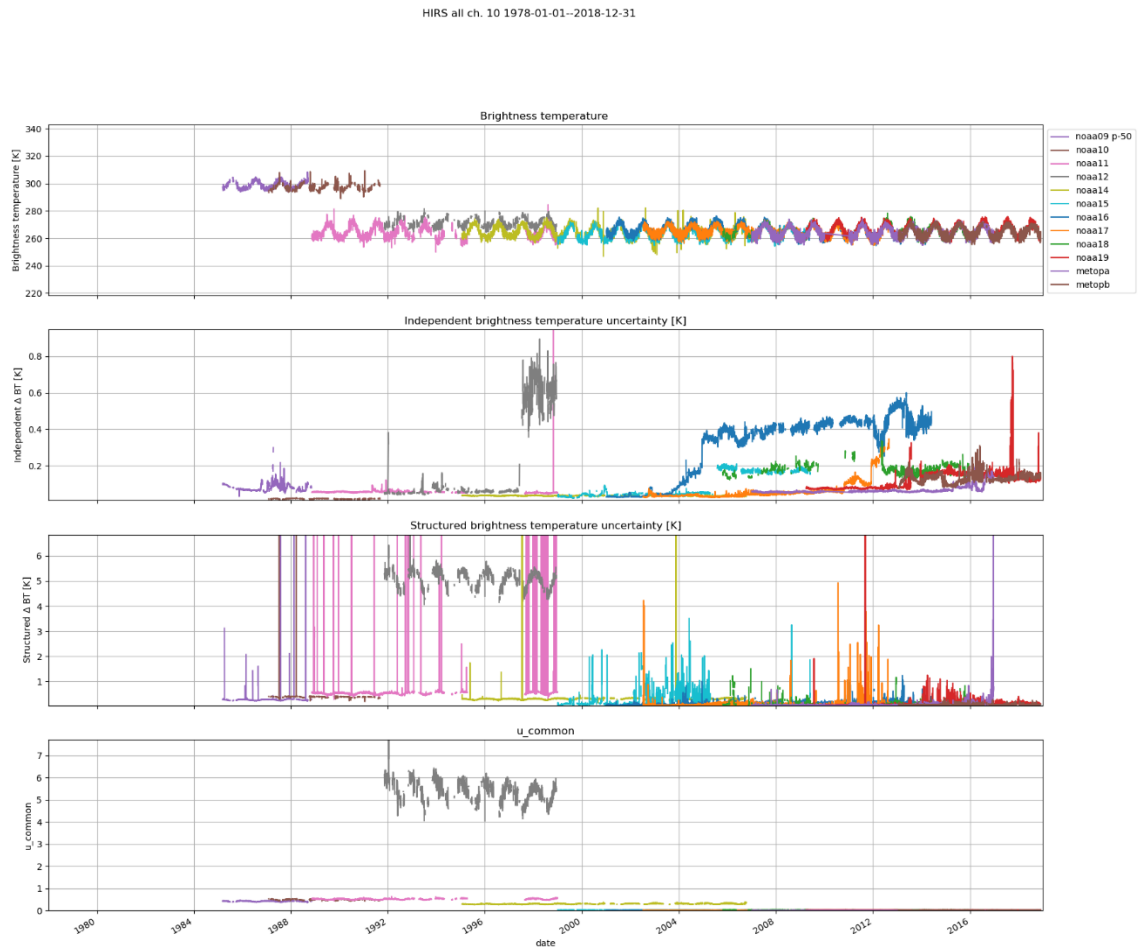


Figure 5-2 – Brightness temperature and errors associated with Ch 10 (12.5 & 8.1 μm) from HIRS, after harmonisation. Panels are as figure 5-1.

Figure 5-2 shows channel 10 output. The brightness temperature has 3 clear separate relations, one in which sensors back to NOAA 11 (with the exception of NOAA 12) agree, one in which NOAA 09 and 10 agree, and one in which NOAA 12 is different again. From NOAA 11 onwards (excluding NOAA 12) all sensors use a 12 μm channel and all have a crossover with at least another satellite with a similar spectral response. When harmonisation takes place, these channels are synced up and produce a continuous 12 μm channel set. NOAA 09, 10 and 12 all have the old HIRS 2 sensor, in which channel 10 has a different spectral response at 8.1 μm . The difference between the 8.1 μm and 12 μm spectral response functions is taken into account in the harmonisation process but while NOAA 09 and 10 have an overlap, NOAA 12 does not overlap with another sensor at 8.1 μm so cannot be harmonised. This explains why NOAA 09 and 10 have good agreement, but NOAA 12 does not. In the interest of scientific usage as a continuous record, channel 10 should be considered as three separate sets: NOAA 11 onwards (excluding NOAA 12); NOAA-12; and NOAA-9 and 10.

Several other channels also show similar patterns in discrepancy across the boundary between HIRS 2 and HIRS 3. Whilst in some cases this is due to changing of the channel wavelengths (Channel 12 also moves from 6.7 to 6.5 μm between HIRS 2 and HIRS 3 for example and channel 4 has a different spectral response function), characteristics of the uncertainties in other channels can also be split in this way. Similarly, the latter part of the NOAA-12 retrievals also have higher uncertainties associated with it compared to other

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

satellites. This does not, however, appear to have an effect on the harmonisation process on the brightness temperatures.

In total, channels 2 – 9 and 11 show continuous data in which harmonisation has ensured a smooth transition between sensors. As shown above, the inconsistency in channel 10 and 12 is due to changes in the wavelength measured and these should ideally be treated separately by the user.



Figure 5-3 – As figure 5-1, but for channel 15 (4.45 μm).

Figure 5-3 shows the time series of brightness temperature and uncertainties for channel 15 (4.45 μm). The retrieved brightness temperatures show that at most satellite overlaps, there is not a consistent agreement, at times up to 10 K in difference. There are also erroneous data which can be seen as spikes upwards and downwards outside the normal range of temperatures recorded. Some of “observations” on NOAA 15 and 19 report a temperature to 0 K; clearly, the (extensive) steps taken to filter bad data are not 100% effective for this channel and require further improvement. As seen in previous figures, the consistency of the common uncertainty is good across all satellites, but is very time dependent for structured uncertainty. The sporadic nature of the structured error arises from the self-emission model used in the FCDR generation model, which in its current form can give unstable results, and can lead to occasional spikes in the uncertainty of up to a few K. Further work on the self-emission model will be able to reduce the spikes obtain a more stable and consistent uncertainty estimate.

4.3 Summary

For HIRS channels at wavelengths longer than 5 μm , the HIRS BTs are relatively stable and there is consistency from sensor to sensor. The two cases where a direct comparison between sensors of a given “channel” gives a large difference are expected because of significant changes in the channels’ spectral response functions. There is considerable time variability in the different uncertainty components, especially for the structured case. Such variability is expected, because instruments do change behaviour over time, but some of the results are also suspected to arise from intermittent instability in calculations related to the self-emission model, which therefore requires additional development.

For the channels around 4 μm , the results are more varied, and larger than expected differences from sensor to sensor persist even though the harmonisation process has been applied. Recommended steps for improvement in the $\sim 4 \mu\text{m}$ channels are: to consider harmonising using a different measurement equation than at longer wavelengths, as embedding the better linearity of the shorter-wavelength detectors into the measurement equation give more stable harmonisation results; re-appraise the filtering of the matchup data used in harmonisation, since for these channels scattered solar radiation (which is not accounted for in the harmonisation process) may be important.

Users therefore must note that the distinction in the degree of stability achieved between the longer ($>5 \mu\text{m}$) and shorter ($<5 \mu\text{m}$) channels.

5 AVHRR

5.1 Introduction

The FIDUCEO FCDR contains reflectance and brightness temperatures from the Advanced Very High Resolution Radiometer (AVHRR) over a 28 year period, with the intended product intended for use as a long term climate dataset, rather than for use with operational forecasting or numerical weather predictions. Versions of the sensors placed onboard the NOAA satellite systems did evolve over time, although the sensor series has been recording near-continuously since the 1980s.

The FCDR has been produced using a metrological approach to errors and uncertainty. This includes considering more accurate representations of errors associated with the underlying data errors differences of sensor versions. This leads to much more detailed expressions of uncertainty, which is provided at a pixel-by-pixel level in the FCDR, per channel. Quality flags and masks are provided to allow the user control over the level of data they wish to include in any future work. The FCDR also has full traceability, including links to the original L1B files that were used initially as a basis of the FCDR, allowing users to understand how uncertainty was calculated and what processes and sampling have occurred in the FCDR production.

5.2 New methods used

The operational calibration for AVHRR is based upon the work by Walton et al. (1998), but has been shown to have multiple issues, as highlighted by Mittaz et al. (2009) and Mittaz & Harris (2011). The new methods used in the FIDUCEO project are comprehensively covered in Mittaz et al. (2019), and so a summary only is provided here:

This new method is based upon an improved radiance measurement function, which is highlighted in this equation:

$$L = a_0 + \frac{(\epsilon + a_1)L_{\text{ICT}} - a_2(\bar{C}_S - \bar{C}_{\text{ICT}})^2}{\bar{C}_S - \bar{C}_{\text{ICT}}} (\bar{C}_S - C_E) + a_2(\bar{C}_S - C_E)^2 + a_3f(T_{\text{inst}}) + 0$$

Where L is the Earth radiance calculated from the Earth Count (C_E). Corrections are made for the instrument self-emission radiance in Earth view, L_{ICT} , which is a function of the instrument temperature, T_{inst} and an offset correction is determined from the averaged space view counts during the calibration cycle, $\bar{C}_S$. T_{inst} is estimated for every scanline. The calibration terms of a_0 , a_1 , a_2 and a_3 are calculated during the harmonisation process, with respect to the reference sensor.

The full FCDR development in FIDUCEO has included an analysis of the error correlation structure across spectral bands and across space (from pixel to pixel within a scanline and from scanline to scanline within an orbit/image).

The FIDUCEO harmonisation is a novel approach for the inter-calibration of remote-sensing instruments, using recalibration instead of bias-correction on the brightness temperature. The harmonisation uses SNOs

D4_6 Metrological assessment of consistency, **V1.1** stability and uncertainty of FIDUCEO FCDRs

as inter-comparison method. The data set of FIDUCEO matchups (SNOs and also off-nadir matches) between the sensors and the reference is generated by Brockmann Consult GmbH in WP3. For the found match-ups, the required information for calibration is collected in a second step in order to form the harmonisation input data.

Instead of bias-correcting the measured brightness temperatures, the harmonisation seeks to correctly account for the underlying physical effects that cause the bias in the calibration procedure. This recalibration approach uses the measurement equation and attempts to find optimised values for certain parameters in the measurement equation such that the inter-satellite bias is reduced. In an iterative procedure of harmonisation and analysis of its results, the correct combination of calibration parameters that explains and reduces the biases needs to be found. Afterwards, reprocessing the data to level 1c using the optimised calibration parameters in the measurement equation generates the harmonised FCDR.

5.3 Consistency of the FIDUCEO AVHRR FCDR

The AVHRR FCDR product is designed to give the user a long, uninterrupted time series of observations which can be used to produce a variety of products. This is normally difficult, as one satellite generally lasts only about 10 years, and deteriorates throughout its lifetime. To get a significant length in time, multiple satellites with similar sensors must be used. One of the key requirements when using multiple satellites to create a single dataset is a smooth transition from one satellite to the next, with consistency of calibration.

The FIDUCEO AVHRR FCDR is harmonised to prevent unexplained jumps (indicated inconsistent calibration) and to attempt to metrologically ensure continuity throughout the time series. Brightness temperatures need a reference, which in the case of AVHRR is provided by AATSR. Figure 6-1 shows two time series of the “K residual” between two satellites for the 11 μm channel: this K residual represents a double difference: it is the expected (i.e., explainable) difference between the two satellites (due to spectral response and matchup differences) minus the observed difference of the measurements. The K residual is therefore a measure of the degree to which the calibration between sensor pairs may be inconsistent. The top figure shows the K residual when observations use the operational calibration, which shows a wide range of values depending on the satellite pairing based on the SNO matchups. For example, NOAA 17 to NOAA16 has a consistently high K residual, whilst NOAA 16 to NOAA 15 has a K residual which averages that are negative for the same time period. In these cases, it is clear that there are unaccounted for calibration-related differences between the sensors. The bottom panel shows the harmonised FIDUCEO product. The reduction in K residual across all satellite pairings is clear to see, showing that the harmonisation process has greatly reduced the differences in the satellite matchups and highlights how the harmonisation in the FIDUCEO FCDR has improved the continuity of the dataset across the multiple sensors at the locations/temperatures sample by the SNO data. A perfect harmonisation would show a mean K residual of zero per satellite pairing, and a Gaussian distribution in the ratio of observed radiance error to the associated uncertainty with a standard deviation of one. Figure 6-2 shows the mean and distribution of the K residuals for the 11 micron channel. Modern sensor pairs apart from the n19-n18 parings show near 0 means and have Gaussian distributions centred around or near zero and a standard deviation of the delta radiance/uncertainty around one, indicating the quality of the continuity of the dataset after harmonisation for the AVHRR/3 sensors. Things are more complex for the older (AVHRR/2) sensors with an apparent underestimate of the total uncertainty (Gaussian distribution larger than one standard deviation).

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

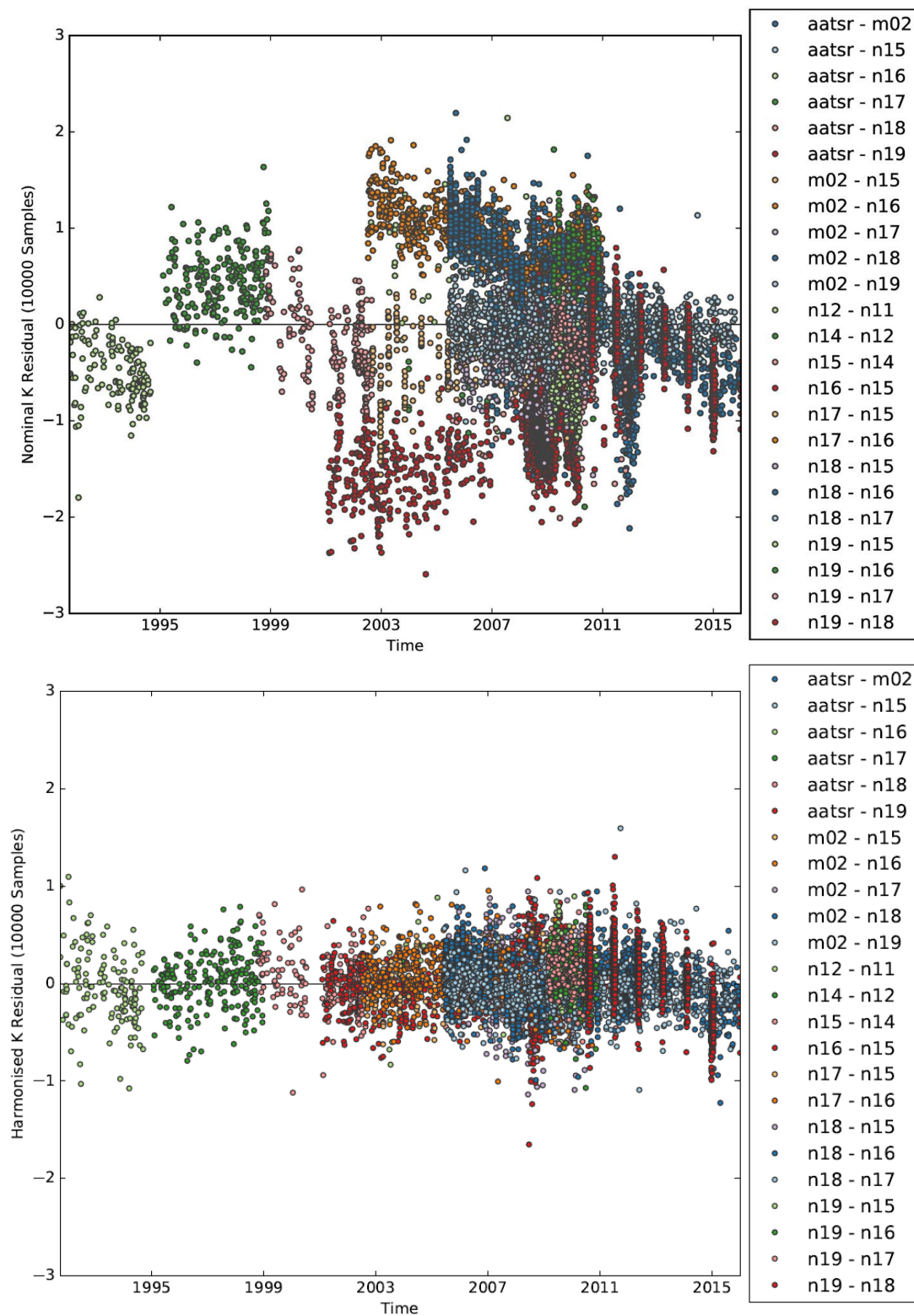


Figure 6-1 – K residual from FIDUCEO MMD satellite comparisons using Operational calibration (top) and FIDUCEO Harmonised calibration (bottom) for the $11\mu\text{m}$ channel.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

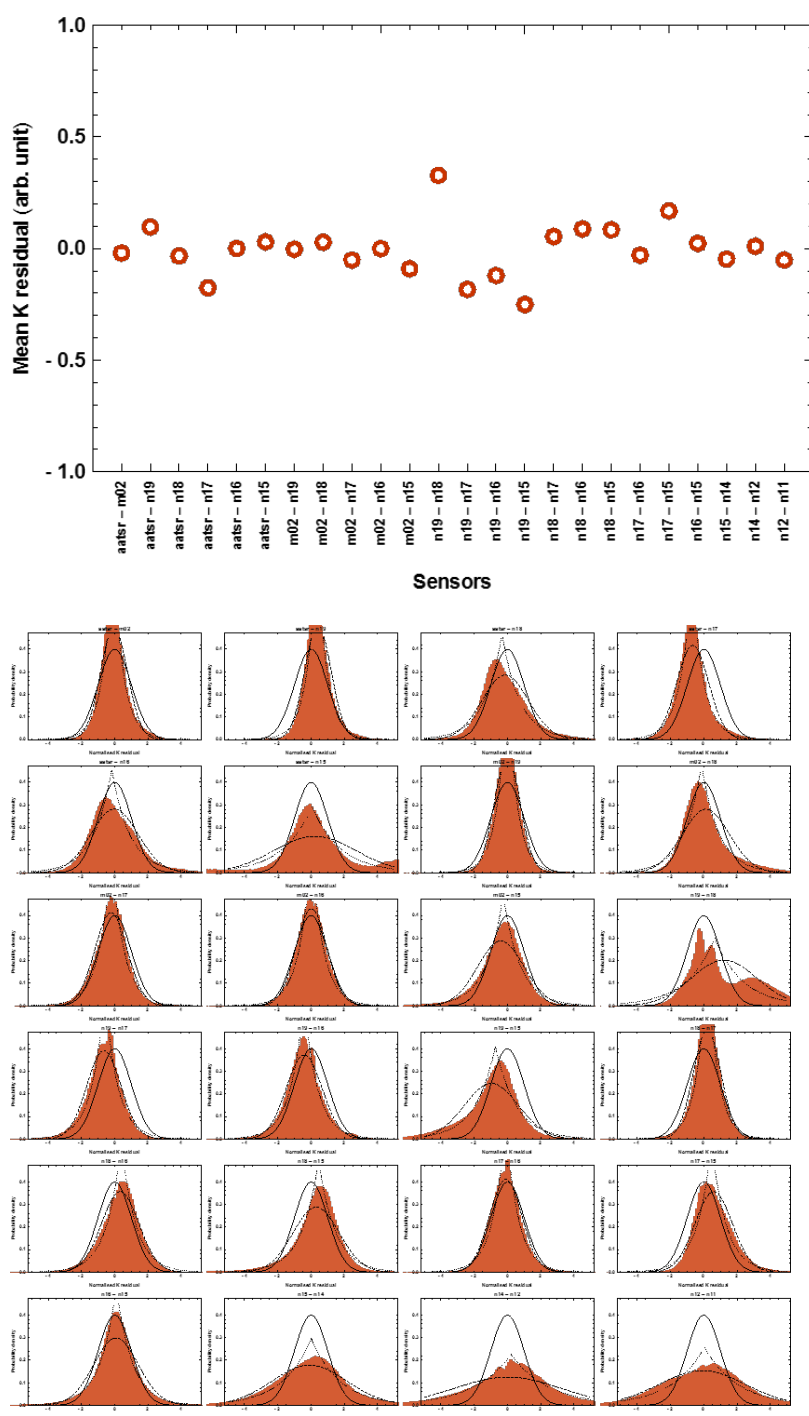


Figure 6-2 – (Top) Mean K residual of harmonised 11 μm channel from AVHRR by sensor pairing. (Bottom) Distribution of K residuals from harmonised 11 μm channel per overlapping sensor pair. The plots show the most recent pairs at top left and run to the oldest pair bottom right.

Figure 6-3 shows the 3.7 μm channel. The harmonisation has shown significant improvements at the matchup locations and has led to fairly unskewed Gaussian distributions centred around zero. Here. However, Figure 6-3 shows clear evidence of the impact of digitisation on the matchup data. This is caused by the predominantly cold (240K-250K) matches used from the SNO locations. Overall, though, from the

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

harmonisation and analysis of the SNO matchup data the harmonisation process has minimised the K residuals and has, in general, shown that the uncertainties used are mostly consistent with the expected statistics.

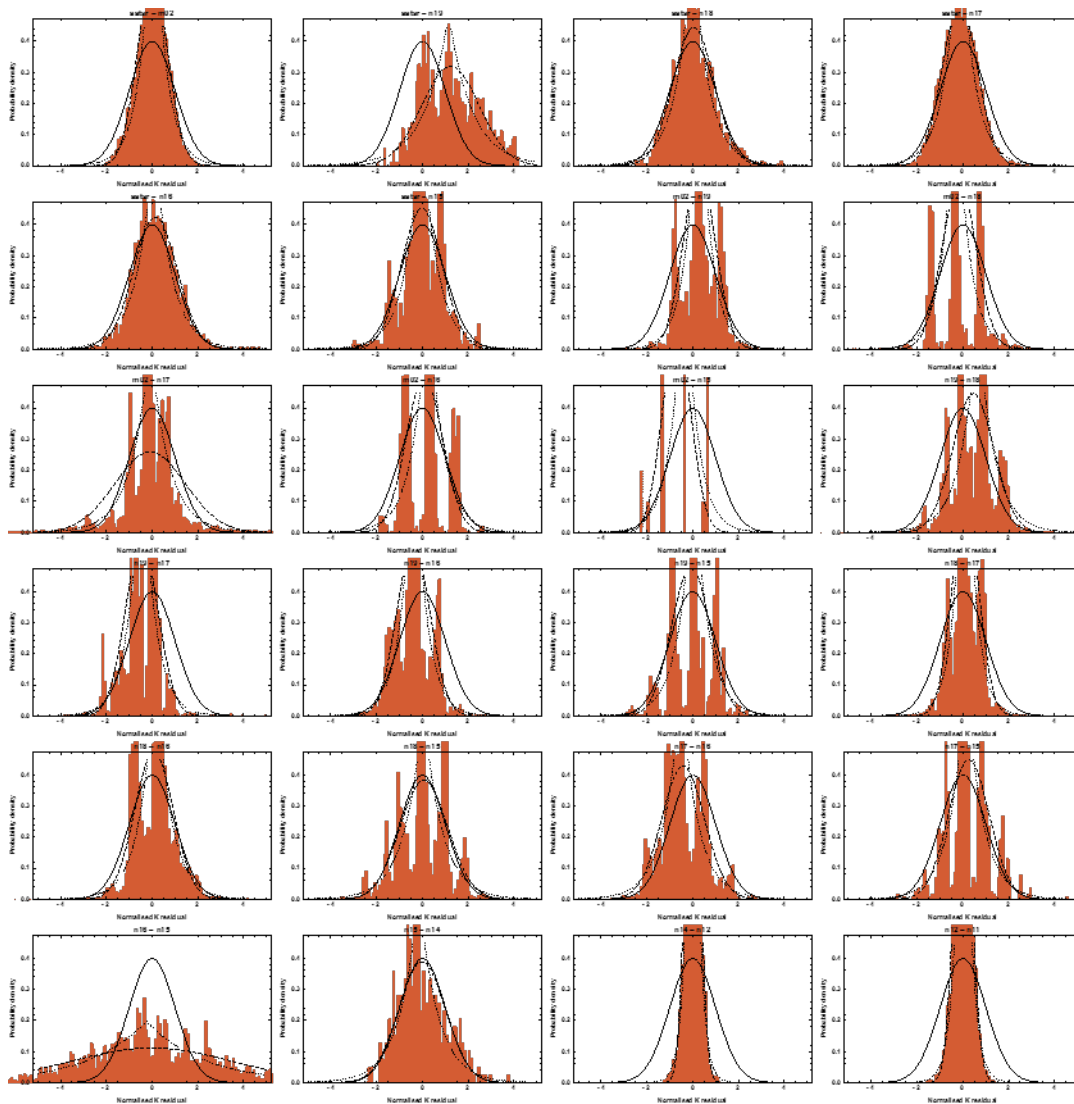


Figure 6-3 – Distribution of the K residual from 3.7 um channel of AVHRR, one plot per sensor pairing.

5.4 Overall stability of the FIDUCEO AVHRR FCDR

As a further check on the stability of the AVHRR radiances, average brightness temperature integrated over a day is considered, on the assumption that on long timescales (multiple years) the total radiance of the Earth in a given channel will be relatively stable (except for climatic scale changes). The Earth is a metrological reference, and many processes influence such an average. For example, all the satellites have different and evolving equator crossing times (apart from MetOp-A) which mean they sample different and evolving portions of the Earth's diurnal cycle as well as having evolving thermal environments for the satellite itself.

D4_6 Metrological assessment of consistency, V1.1 stability and uncertainty of FIDUCEO FCDRs

But using the daily average brightness temperature is a check capable of detecting significant problems with the FIDUCEO calibration.

5.4.1 The 11 μ m and 12 μ m channels

Starting with the two windows channels, Figure 6-4 and Figure 6-5 show the daily averaged data for the 11 μ m and 12 μ m channels respectively for all the sensors in the current FIDUCEO FCDR (NOAA11 to MetOp-A). Both plots seem to show that the AVHRR/3 sensors are more stable than the earlier AVHRRs (AVHRR/2) and there is an apparent trend for both channels before the year 2000 (the approximate date for the switch from AVHRR/2 to AVHRR/3 sensors). It is also apparent that there are differences in the means for the sensors which become most apparent from NOAA-16 (purple) back. NOAA-11 (blue) also seems to be consistently high compared to the other AVHRRs though it should be born in mind that the references for it (NOAA-12) are themselves very problematic with large instrument temperature variations and solar contamination effects. When looking at an average over all scene temperatures it is clear that the harmonisation process for the 11 μ m and 12 μ m channels needs further improvement, concentrated in the earlier years (those most distant from the reference sensor period).

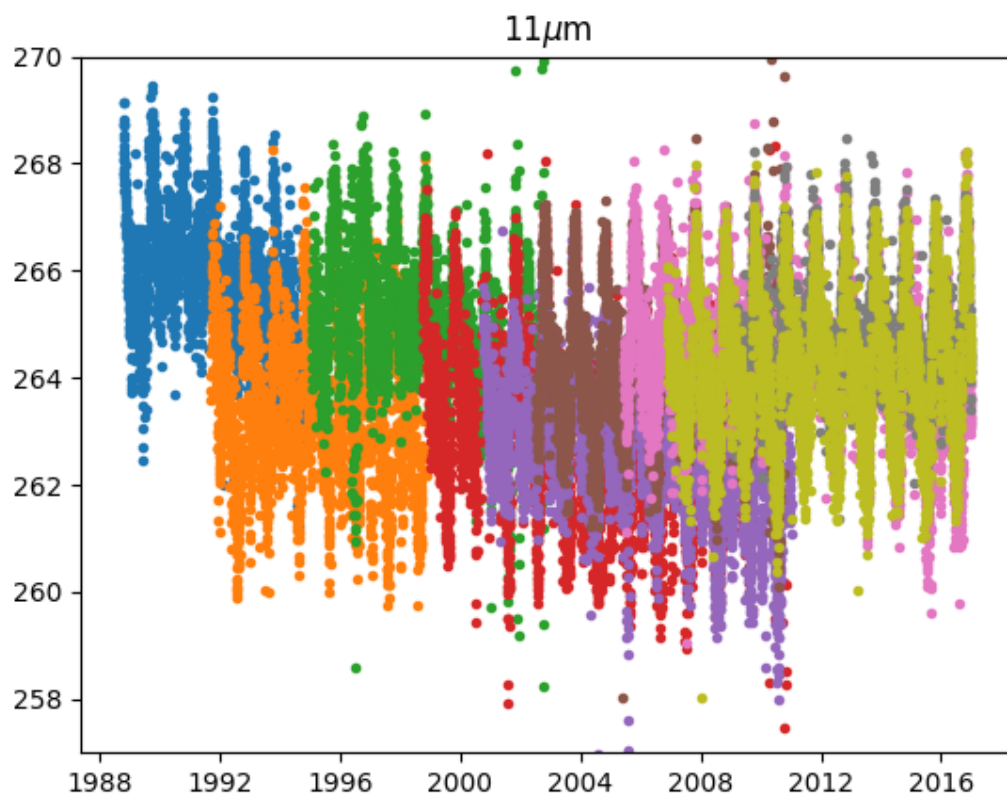


Figure 6-4- Daily average brightness temperatures for the 11 μ m channel AVHRR data for all sensors from NOAA11 through to Metop-A. The AVHRR/3 data from NOAA-16 (2001) to MetOp-A looks fairly stable but the earlier data (NOAA-11 to NOAA-15) shows some evidence of a trend. The different AVHRRs are coloured as NOAA11:Blue, NOAA12:Orange, NOAA14: Green, NOAA-15:Red, NOAA-16: Purple, NOAA-17:Brown, NOAA18: Pink, NOAA-19:Grey, MetOp-A:Olive.

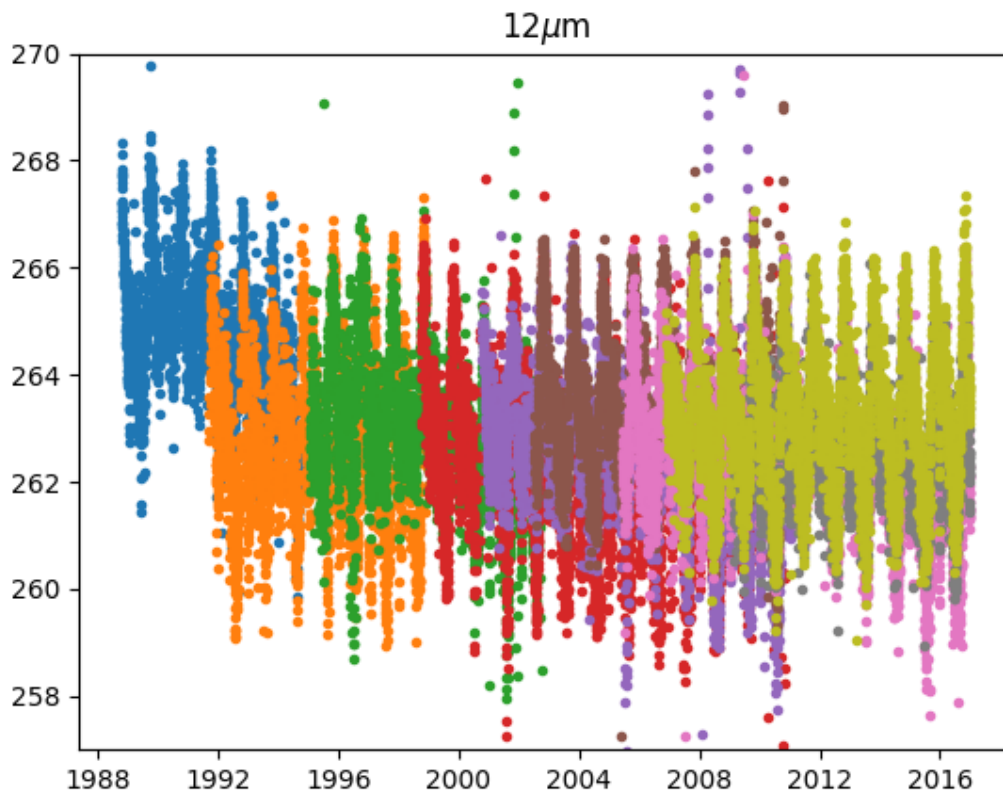


Figure 6-5 Daily average brightness temperatures for the 12μm channel. The 12μm channel looks slightly more stable than the 11μm channel but there is still a trend for the early AVHRRs.

5.4.2 The 3.7μm channel

The case of the 3.7μm channel is somewhat different from that of the 11μm and 12μm channels because the SNO matchups are much more limited. This is because SNOs can't be used in the daytime portions of the orbit because of atmospheric solar scattering. So for some sensor-to-sensor SNO matches the number of matchups can be very limited. On top of this there is also the problem that for the 3.7μm channel there are extra sources of error that do not exist for the 11μm and 12μm channel, namely solar contamination of the calibration data as well as direct solar radiance which is seen in the Earth view data itself. Both of these error sources occur at night and will potentially further limit high quality matchups.

The impact of the solar contamination error source is highlighted in Figure 6-6. For the current harmonisation it was decided to remove all solar contaminated matches and only keep night time matches with no solar contamination as this is the most error free case. It did, however, severely limit the number of matchups particularly for the AVHRR/2 sensors which did not have a sun shield fitted and therefore are much more prone to solar contamination issues. What is apparent in Figure 6-6 is that those sensors most impacted by solar contamination (the AVHRR/2 sensors, NOAA-11 to NOAA-14) have a very poor harmonised calibration compared to the AVHRR/3 cases. This is almost certainly due to the much smaller number of matchups available here.

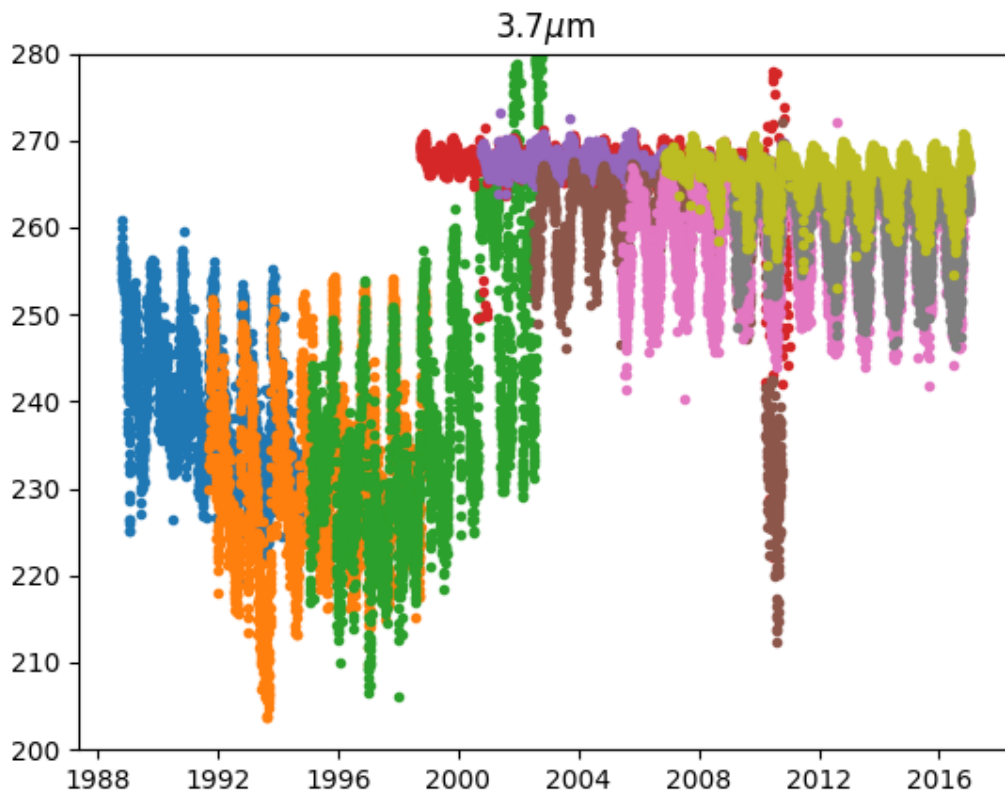


Figure 6-6 Daily average brightness temperatures for the 3.7μm channel. For many sensors the 3.7μm channel shows more seasonal variability than was seen in the 11μm and 12μm channels and the range of variability varies from sensor to sensor. All the AVHRR/2 sensors (NOAA-11:Blue, NOAA-12:Orange, NOAA-14:Green) are also significantly offset from the AVHRR/3 sensors. The brown spike (NOAA-17) around 2010 is due to the sensor cross into the terminator and becoming very unstable after March 2010.

5.5 Summary

Results of the daily average Earth temperature in the three IR channels of the AVHRR show that the harmonisation applied has not achieved the hoped-for stability for the earlier part of the record, farthest from the reference sensor. It is to be expected that errors accumulate with increasing temporal separation from the reference, but the discrepancies are larger than hoped for. The overlaps available and used for matches between sensor pairs are one limitation that has hampered the harmonisation approach: some overlap periods are very short and geometrically restricted. More matches covering a larger spread of scene temperature, instrument temperature and orbital position are needed to obtain a better harmonisation. The SNO matches currently used are concentrated at the poles, and provide a limited snapshot of all the scene temperatures/satellite environment temperatures that occurred. It has turned out that these are inadequate to span all the processes that need to be harmonised.

The mathematical aspects of the attempted harmonisation remain valid, but new strategies for matching that expand the range of conditions sampled are needed to give the harmonisation process statistically adequate inputs. New forms of matches need to be developed with associated uncertainty characterisation,

D4_6 Metrological assessment of consistency, **V1.1** stability and uncertainty of FIDUCEO FCDRs

which will require significant future effort (particularly the uncertainty aspect). Within the FIDUCEO project we were not able to develop the matchup data beyond simple SNOs originally planned for.

Another limitation may arise from the AVHRR measurement equation used in the harmonisation attempted, which (although more realistic than the operational calibration equation) may be insufficiently flexible to capture all the relevant behaviour. For example, instrument temperature model term, which parameterises the (previously unaddressed) effect of instrument temperature on calibration behaviour a linear model as a function of instrument temperature for each sensor. Evidence in the AVHRR Uncertainty document D2.2, suggests the true dependence is more complex, both non-linear and a function of time. An element of the remaining trends seen in the whole Earth results may arise from the constraints of the simple parameterisation.

5.6 Positives of FIDUCEO methods and FCDR

The FIDUCEO AVHRR FCDR facilitates the handling of the data by providing data in ready-to-use NetCDF files. Each of these files covers one orbit reaching from one equator-crossing to the next in the same flight direction. This equator-to-equator frame removes all doubled data and hence prevents the appearance of corresponding sampling artefacts. The FCDR contains concise quality information on pixel level and maintains the traceability back to the original level 1b data. Secondly, the FCDR provides extensive pixel-level uncertainty information, considering the correlation behaviour of the underlying errors. This is communicated by three uncertainty variables (see PUG, D2.2) in the FCDR files. Thirdly, the large inter-satellite biases and instabilities have been largely reduced due to both the harmonization and correction of calibrations. While not fully optimal for all sensors/instrument types this provides a significant improvement when using the data over long time periods, and ensures good levels of confidence in any future uses of the dataset. Finally, the work and methodologies provided by the FIDUCEO project, based on metrological reasoning, have been proven to generally improve both the accuracy and the quantification of uncertainty of AVHRR retrievals.