ECMWF COPERNICUS | REPORT

Copernicus Climate Change Service

Quality Evaluation Report

Fundamental Climate Data Record of Microwave Humidity Sounder

Issued by: EUMETSAT

Date: 11/10/2019

REF.: C3S_311b Task 1.1 D1.1, D1.2, D1.3

IMPLEMENTED BY

Quality Evaluation Report

Fundamental Climate Data Record of Microwave Humidity Sounder

Lizzie Good (EUMETSAT CM SAF*)
Freya Aldred (EUMETSAT CM SAF*)
Mike Grant (EUMETSAT)
Timo Hanschmann (EUMETSAT)
Viju John (EUMETSAT)
Rob King (EUMETSAT, CM SAF*)
Rob Roebeling (EUMETSAT)
Jörg Schulz (EUMETSAT)
Christoforos Tsamalis (EUMETSAT CM SAF*)

*Met Office

Issued by: EUMETSAT

Date: 11/11/2019

REF.: C3S_311b-1.1_D1.1, D1.2, D1.3

Service contract: 2016/C3S_311b _EUMETSAT

Contents

List of Figures

List of Tables

1 Introduction

This Quality Evaluation Report has been compiled together with staff from the UK Met Office acting as partner of the EUMETSAT Climate Monitoring Satellite Application Facility (CM SAF).

The Quality Evaluation Report describes the validation of the 183 GHz fundamental climate data records (FCDR) produced for the Copernicus Climate Change Service (C3S) [AD1]. The FCDRs include the 183 GHz channels on the Microwave Humidity Sounder (MHS), the Advanced Technology Microwave Sounder (ATMS), the Micro-Wave Humidity Sounder 1 (MWHS/1), and MWHS/2. It is complementary to the efforts in the FIDUCEO (Fidelity and uncertainty in climate data records from Earth Observations) Horizon 2020 project, where FCDRs have been produced for the microwave humidity sounders SSM/T2 (Special Sensor Microwave/Temperature-2), AMSU-B (Advanced Microwave Sounding Unit), and MHS. A key objective of the FIDUCEO project has been the derivation and provision of fully-characterised measurement uncertainties, following the fundamental principles of metrology. Thus, the FIDUCEO FCDRs for SSM/T2, AMSU-B and MHS, which cover the period 1994-2017, are provided with detailed quality and uncertainty data that offer users additional information compared with the conventional brightness temperature (BT) datasets provided by other data set developers (0).

The C3S microwave humidity FCDRs are generated using a modified version of software developed by the University of Hamburg within the FIDUCEO project. The modifications include converting the software from Matlab to Python and adding the capability to process MWHS/1, MWHS/2, and ATMS data. The objective for this modified software is to produce FCDRs for MHS on board Metop A and B, MWHS/1 and 2 on board FY-3A, -3B and -3C, and ATMS on board S-NPP that is complementary to the FIDUCEO FCDRs, such that all the FCDRs can be used together in applications, including the generation of Climate Data Records (CDRs). The FCDRs comprise of calibrated brightness temperatures with uncertainty components, geolocation information, as well as detailed quality flags. They are generated directly from the satellite Level 1a (L1a) products, which contain the digital counts at the full spatial and temporal resolution of the instrument, together with georeferencing data and all the calibration data needed to convert these observations to BTs.

This report is concerned only with the validation of the 183 GHz FCDRs. Table 1 lists the sensors that are included in the FCDRs, together with the period of data availability in each case. The channels evaluated in this study are shown in Table 2.

Table 1: Summary information for the MW sounders used to produce the FCDRs indicating the satellite platform, the temporal period for each instrument, and the local equator crossing time (LECT) of the descending node. The LECT is indicative at the launch time of each satellite. LECT source: https://www.wmo-sat.info/oscar/ . *Denotes approximate LECT times

Instrument	Satellite	Period	LECT descending
MHS	**Metop-A**	**05/2007-12/2018**	**09:30**
MHS	**Metop-B**	**04/2013-12/2018**	**09:30**
MWHS/1	**FY-3A**	**07/2008-05/2014**	**09:05***
MWHS/1	**FY-3B**	**12/2010-12/2018**	**01:38***
MWHS/2	**FY-3C**	**09/2013-12/2018**	**10:15***
ATMS	**SUOMI NPP**	**12/2011-12/2018**	**01:25**

Table 2: 183 GHz channels present on each instrument used in the beta and final microwave sounder FCDR. The channel number for each instrument is shown where these are present.

Sensor	183.31±1	183.31±1.8	183.31±3	183.31±4.5	183.31±7	190.31
MHS	3		4			5
MWHS/1	3		4		5	
MWHS/2	11	12	13	14	15	
ATMS	22	21	20	19	18	

For ease of reporting the suite of FCDRs – for MWHS/1 (FY-3A/B), MWHS/2 (FY-3C), MHS/Metop A/B, and ATMS/S-NPP - are hereafter referred to as 'C3S microwave humidity FCDR` or just 'FCDR' for simplicity in this document.

The instruments comprising the C3S microwave humidity FCDR are referred to only by their platform name in plot annotation, and in general, in the text, for example the MWHS/1 on FY-3B is referred to as 'FY-3B'. For a full list of the platforms, see Table 1.

1.1 Purpose of the document

Within Task 1.1 in C3S_311b [AD1] EUMETSAT provides FCDRs for observations from microwave sounder instruments on board polar orbiting operational weather satellites, the C3S microwave humidity FCDR. This C3S microwave humidity FCDR is defined by a set of three documents. These documents are essential for users to use and understand the products. The documents are:

- Algorithm Theoretical Baseline Document (ATBD) (methods and algorithm description) [0];
- Quality Evaluation Report (informs on the accuracy, precision, and stability of the product);
- User Guide (provides essential information for the user on the product definition, how to access and work with the data, and contains information on major limitations on the usage) [0].

This Quality Evaluation Report aims to provide information on the product quality for the C3S microwave humidity FCDR to detect and mitigate remaining issues in the data before their delivery to C3S. This information is derived by means of data analysis including comparison to other microwave sensor data and reference data. This includes:

- Comparisons to operationally processed data;
- Comparisons to simulated observations, based on the ERA5 reanalysis dataset;
- Simultaneous nadir overpass analysis.

1.2 Structure

This document is structured as follows:

Section 1	Introduction describing the purpose and scope of this quality evaluation report
Section 2	Background provides an overview on the data used in the study
Section 3	Validation strategy summarises the approaches for validating the data records and the reference data
Section 4	Validation results
Section 5	Conclusions
Sections 6-8	Appendices for additional results

1.3 List of Abbreviations

Abbreviation	Description
AAPP	ATOVS and AVHRR Pre-processing Package
ATMS	Advanced Technology Microwave Sounder
AMSU-B	Advanced Microwave Sounding Unit - B
BT	Brightness Temperature
CDR	Climate Data Record
CMA	China Meteorological Administration
CM SAF	Climate Monitoring Satellite Application Facility
ECMWF	European Centre for Medium-Range Weather Forecasts
ECV	Essential Climate Variable
ERA5	ECMWF/C3S Reanalysis version 5
FIDUCEO	Fidelity and uncertainty in climate data records from Earth Observations
FCDR	Fundamental Climate Data Record
FY-3	FengYun-3
FOV	Field of View
EPS	EUMETSAT Polar System
GCOS	Global Climate Observing System
ITCZ	Inter Tropical Convergence Zone
LECT	Local Equator Crossing Time
Metop	Meteorological Operational Satellite Program of Europe
MHS	Microwave Humidity Sounder
MWHS	Micro-Wave Humidity Sounder
MW	Microwave
OBCT	on-board calibration target
QC	Quality control
NOAA	National Oceanic and Atmospheric Administration
NRT	Near-Real-Time
NWP	Numerical Weather Prediction

1.4 Applicable Documents

#	Title	Version
AD1.	C3S Implementation Plan 2019 EUMETSAT-D9.2	1.0

1.5 Reference Documents

RD1.	Buehler, S. A., M. Kuvatov, T. R. Sreerekha, V. O. John, B. Rydberg, P. Eriksson, and J. Notholt, 2007, A cloud filtering method for microwave upper tropospheric humidity measurements, Atmos. Chem. Phys., 7, 5531-5542, doi: https://doi.org/10.5194/acp-7-5531-2007
RD2.	M. Dowell, P. Lecomte, R. Husband, J. Schulz, T. Mohr, Y. Tahara, R. Eckman, E. Lindstrom, C. Wooldridge, S. Hilding, J.Bates, B. Ryan, J. Lafeuille, and S. Bojinski, 2013: Strategy Towards an Architecture for Climate Monitoring from Space. Pp. 39
RD3.	Gu S., Y. Guo, Z. Wang and N. Lu, 2012, "Calibration Analyses for Sounding Channels of MWHS Onboard FY-3A," in IEEE Transactions on Geoscience and Remote Sensing, vol. 50, no. 12, pp. 4885-4891, doi: 10.1109/TGRS.2012.2214391

RD4.	Hans, I.; Burgdorf, M.; Buehler, S.A.; Prange, M.; Lang, T.; John, V.O., 2019, An Uncertainty Quantified Fundamental Climate Data Record for Microwave Humidity Sounders. *Remote Sens.*, *11*, 548.
RD5.	Hanschmann, T., John, V., Roebeling, R. Grant, M. And Schulz, J., 2019, Algorithm Theoretical Basis Document: Fundamental Climate Data Record of Microwave Humidity Sounder, EUMETSAT report to C3S
RD6.	Hanschmann, T., John, V., Roebeling, R. Grant, M. And Schulz, J., 2019, Product User Guide: Fundamental Climate Data Record of Microwave Humidity Sounder, EUMETSAT report to C3S
RD7.	He, J., Zhang, S., and Wang, Z., 2015, "Advanced Microwave Atmospheric Sounder (AMAS) Channel Specifications and T/V Calibration Results on FY-3C Satellite," in IEEE Transactions on Geoscience and Remote Sensing, vol. 53, no. 1, pp. 481-493, doi: 10.1109/TGRS.2014.2324173
RD8.	Hocking J., P. Rayer, D. Rundle, R. Saunders, M. Matricardi, A. Geer, P. Brunel and J. Vidot, 2017, RTTOV v12 Users Guide, NWPSAF-MO-UD-037, https://nwpsaf.eu/site/software/rttov/.
RD9.	GCOS-154, 2011: Systematic Observation Requirements for Satellite-Based Products for Climate,
RD10.	Kazumori, M. and English, S. J., 2015, Use of the ocean surface wind direction signal in microwave radiance assimilation. Q.J.R. Meteorol. Soc., 141: 1354–1375. doi:10.1002/qj.2445
RD11. .	Kim, E., C.-H. J. Lyu, K. Anderson, R. V. Leslie, and W. J. Blackwell, 2014, S-NPP ATMS instrument prelaunch and on-orbit performance evaluation, J. Geophys. Res. Atmos., 119, 5653–5670, doi:10.1002/2013JD020483
RD12.	Kleespies, T. J. and Watts, P., 2006, Comparison of simulated radiances, Jacobians and linear error analysis for the Microwave Humidity Sounder and the Advanced Microwave Sounding Unit-B. Q.J.R. Meteorol. Soc., 132: 3001-3010. doi:10.1256/qj.05.03
RD13.	Lawrence H., N. Bormann, A. J. Geer, Q. Lu and S. J. English, "Evaluation and Assimilation of the Microwave Sounder MWHS-2 Onboard FY-3C in the ECMWF Numerical Weather Prediction System," in IEEE Transactions on Geoscience and Remote Sensing, vol. 56, no. 6, pp. 3333-3349, 2018, doi: 10.1109/TGRS.2018.2798292
RD14.	Moradi, I., S. Buehler, V. John, A. Reale, and R. Ferraro, 2013, Evaluating instrumental inhomogeneities in global radiosonde upper tropospheric humidity data using microwave satellite data, IEEE T. Geosci. Remote, 51, 3615–3624, doi:10.1109/TGRS.2012.2220551.
RD15.	Roca, R., Brogniez, H., Chambon, P., Chomette, O., Cloché, S., Gosset, M. E., Mahfouf, J.-F., Raberanto, P. and Viltard, N, 2015, The Megha-Tropiques mission: a review after three years in orbit. Front. Earth Sci. 3:17. doi: 10.3389/feart.2015.00017
RD16.	Smith A., 2017, Radiance Simulator v2.0 User Guide, NWPSAF-MO-UD-040, https://nwpsaf.eu/site/software/radiance-simulator/.

RD17.	Wang, D., C. Prigent, L. Kilic, S. Fox, C. Harlow, C. Jimenez, F. Aires, C. Grassotti, and F. Karbou, 2017, Surface Emissivity at Microwaves to Millimeter Waves over Polar Regions: Parameterization and Evaluation with Aircraft Experiments. J. Atmos. Oceanic Technol., 34, 1039–1059, https://doi.org/10.1175/JTECH-D-16-0188.1.
RD18.	Yang, W., John, V. O., Zhao, X., Hui Lu, H. and Kenneth R. Knapp, K. R., 2016, Satellite Climate Data Records: Development, Applications, and Societal Benefits, Remote Sens. 2016, 8, 331; doi:10.3390/rs8040331
RD19.	John, V. O., Holl, G., Buehler, S. A., Candy, B., Saunders, R. W., and Parker, D. E. (2012), Understanding intersatellite biases of microwave humidity sounders using global simultaneous nadir overpasses, J. Geophys. Res., 117, D02305, doi:10.1029/2011JD016349.

1.6 Document Change Record

Version	Date	Document Number	Changes
1	16.04.2019	C3S_311b Task 1.1 D1.1, D1.2, D1.3	Draft version
2	15.08.2019	C3S_311b Task 1.1 D1.1, D1.2, D1.3	Final version

1.7 Definitions

The following definitions are used throughout the document.

Data levels (http://www.wmo.int/pages/prog/sat/dataandproducts_en.php):

- Level 1A Instrument counts with geolocation and calibration information attached but not applied
- Level 1B Geolocation and calibration information applied to the instrument counts
- Level 1C Instrument specific. For example, 1b data converted to Brightness Temperature

Products types:

- Operational data - are Level 1 data provided by an institution who has operational responsibility for the sensor. In this study, this is the China Meteorological Administration (CMA) for MWHS/1 and MWHS/2, the National Oceanic and Atmospheric Administration (NOAA) for ATMS and MHS, and Centre National d'Etudes Spatiales (CNES) for SAPHIR.
- Fundamental Climate Data Record [0] – The term FCDR denotes a well-characterised, long-term data record, usually involving a series of instruments, with potentially changing measurement approaches, but with overlaps and calibrations sufficient to allow the generation of products that are accurate and stable, in both space and time, to support climate applications. FCDRs are typically calibrated radiances, backscatter of active instruments, or radio occultation bending angles. FCDRs also include the ancillary data used to calibrate them. The term FCDR has been adopted by GCOS and can be considered as an international consensus definition.

Statistics:

- Observations minus simulated BTs – is the difference between the observations that comprise the C3S microwave humidity FCDR and equivalent data that have been simulated using a radiative transfer model with ERA5 (ECMWF/C3S 5th reanalysis) atmospheric and surface parameters.
- Independent uncertainty – is the standard uncertainty that describes the errors associated with an observation that are uncorrelated in time and space. Independent errors arise from random effects causing errors that manifest independence between pixels, such that the error in one pixel is in no way predictable from knowledge of the error in another pixel, were that knowledge available. Independent errors therefore arise from random effects operating on a pixel level, the classic example being detector noise. The Independent uncertainty is mostly referred to as the random uncertainty.
- Structured uncertainty – is the standard uncertainty that describes the errors associated with an observation that are correlated over local time and space scales (e.g. within an orbit). The correlation length scales should be defined either in the data product or in its documentation.
- Common uncertainty – is the standard uncertainty that describes the errors associated with an observation that are fully correlated in time and space (e.g. during the lifetime of an instrument). The common uncertainty is also referred to as the systematic uncertainty.
- Total uncertainty – is the standard uncertainty that describes the total uncertainty budget, where all the uncertainty components (U_a, U_b, U_c, …) have been added in quadrature, i.e.:

$$U_A = \sqrt{{U_a}^2 + {U_b}^2 + {U_c}^2 + \cdots} \qquad 1\text{-}1$$

  In this document, the independent, the structured, and the common uncertainty components are added to the total uncertainty.
- The mean difference μ (also frequently reported as 'bias') results from the arithmetic mean of the difference over the members of two data records. This measure should indicate how close the observation, y, is on average to a reference observation, o (representing the best estimation of the truth). It indicates whether the data record on average over- or underestimates the reference data record and is defined as:

$$\mu = \frac{1}{n}\sum_{k=1}^{n}(y_k - o_k) = \bar{y} - \bar{o} \qquad 1\text{-}2$$

  This study considers SAPHIR brightness temperatures as common reference, because it can be compared to all FCDRs but is not part of the data record. Further, SAPHIR observations cover the whole dynamic range due to its low inclination angle.
- The sample standard deviation $\boldsymbol{\sigma}$ (also frequently reported in literature as 'precision' or 'bias-corrected root-mean-squared error') is a measure of the spread around the mean value of the distribution formed by the differences between the test and the reference data records. This measure should tell how individual parameter estimations are distributed relative to the mean difference defined as:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}((y_k - o_k) - (\bar{y} - \bar{o}))^2} \qquad 1\text{-}3$$

- Double-difference – is the difference between two differences with respect to a reference data set:

$$DD_{y_a - y_b} = (y_a - o) - (y_b - o) \qquad 1\text{-}4$$

Where y_a is the brightness temperature from sensor *a*, y_b is the brightness temperature from sensor *b*, and o is the reference brightness temperature. The advantage of this method over simply calculating the $y_a - y_b$ difference is that any differences in observation time or space are removed by calculating the difference with respect to the reference BT, which should be close to the observation time and space of y_a or y_b.

- Uncertainty overlap analysis – is the analysis of the agreement between data points considering the uncertainties on the data points. The total uncertainty is assumed here to represent the 1-sigma total uncertainty (see above definition). Therefore, if two equivalent data sets are compared, it is assumed that ~ 68% of data should agree within their uncertainty ranges if the uncertainties for both data sets are approximately correct. This is assessed through the calculation of *Z* for each pair of data points:

$$Z = \frac{(y_b - y_a)}{\sqrt{(U_a{}^2 + U_b{}^2)}} \qquad \text{1-5}$$

Where y_a is the brightness temperature from sensor *a*, y_b is the brightness temperature from sensor *b*, U_a is the total uncertainty on y_a and U_b is the total uncertainty on y_b. If the distribution of *Z* for all data points has a standard deviation of one, then the data sets are considered to be in agreement. If the standard deviation of this distribution is more than one, then the uncertainties for one or both data sets are too small. If the standard deviation of the distribution is less than one, then the uncertainties for one or both data sets are too large. It may be possible to establish which data set has uncertainties that are too large or small, through the inter-comparison of several pairs of data sets. Any uncertainty arising due to the collocation or 'matchup' of the observations, e.g. uncertainty in scene inhomogeneity, is assumed to be zero in this experiment.

Other:

- Simultaneous nadir overpass (SNO) – is when two different satellite instruments have approximately the same near-nadir earth view over space at the same time, within some agreed limits in time, space, and view-angle difference. In this study, satellite observations from two different sensors are considered SNOs if the observations are within 5 km and 5 minutes to each other, and where the satellite zenith angle differences are within 5°.

2 Background

This section provides a description of the C3S microwave humidity FCDR evaluated in this study.

2.1 MHS on board Metop A and B

Metop A is Europe's first operational meteorological polar satellite launched on the 19 October 2006. Metop is the acronym for the Meteorological Operational Satellite Program of Europe. Metop B, the second satellite of the EUMETSAT Polar System (EPS), was launched on the 17 September 2012. Both satellites are on a sun-synchronous orbit with local equator crossing time (LECT) of 21:30 stable over time for the ascending node. For Metop A the orbit was kept stable until August 2017, after which the orbit started to drift to an earlier LECT. Metop A and B carry the same suite of instruments, which includes the MHS.

MHS is a five-channel across-track scanning radiometer with 90 field of views (FOV) separated by 10/9 degrees (0). It has a scan range of ±49.44° with respect to the nadir direction leading to a swath of 2310 km, while the FOV size is about 16 km at nadir. Each scan takes 8/3 s to complete and it consists of four views of the on-board calibration target (OBCT), followed by the 90 Earth views and the four views of the deep space. There are three water vapour sensitive channels around the 183 GHz absorption line with frequencies 183.31±1.0 GHz, 183.31±3.0 GHz and 190.31 GHz (Table 2). The first two channels are double-sideband symmetric about the water vapour line, while the 190.31 GHz has a single band-pass.

2.2 ATMS on board S-NPP

On 28 October 2011 NASA launched the Suomi- National Polar-Orbiting Partnership (S-NPP) satellite for NOAA. It was placed into an 824 km, polar Sun-synchronous, 16 day-repeat-cycle orbit with an ascending node LECT of 13:30 (RD11). S-NPP carries the first ATMS. ATMS is an "integrate-while-scan" total-power-radiometer, with 96 FOVs per scan and 22 channels. The angular range of the ATMS scan is ±52.725° relative to nadir, inducing a swath of 2580 km. Like MHS the scan period of ATMS is 8/3 s and it uses four views of the deep space and the OBCT to perform the calibration. Of the 22 channels, five are around the 183 GHz water vapour absorption line with a FOV size of about 16 km at nadir. More specifically these moisture sensitive channels have the following frequencies: 183.31±1.0 GHz, 183.31±1.8, 183.31±3.0 GHz, 183.31±4.5 and 183.31±7.00 GHz (Table 2).

2.3 MWHS/1 and 2 on board FY-3A, -3B and -3C

FengYun-3 (FY-3) is the second generation of Chinese polar sun-synchronous satellites. Two development phases were considered for the FY-3 satellites, with FY-3A and FY-3B considered experimental, while FY-3C and the subsequent platforms were considered operational. FY-3A was launched on 27th May 2008 on an orbit with a LECT of 22:00 for the ascending node. It began its operational phase on 12 January 2009 and reached its end of life on 5 January 2015[1]. FY-3B was launched the 4 November 2010 on an orbit with a LECT of 14:00 for the ascending node. It became operational 2 June 2011 and is still operating at the time of writing[2]. The first fully operational satellite of the FY-3 series, FY-3C, was launched on the 23 September 2013 on an orbit with a LECT of 22:00 for the

[1] https://directory.eoportal.org/web/eoportal/satellite-missions/f/fy-3
[2] https://directory.eoportal.org/web/eoportal/satellite-missions/f/fy-3

ascending node. For all FY-3 satellites, the orbit drifts with time and thus these LECTs, which are also provided in Table 1, are approximately correct at the time of writing.

Both FY-3A and FY-3B carry the same suite of instruments, which includes the Microwave Humidity Sounder (MWHS/1). MWHS/1 is a total power microwave radiometer with five channels in the range of 150-191 GHz. Three of them are sounding channels around the 183 GHz water vapour absorption line with frequencies 183.31±1.0 GHz, 183.31±3.0 GHz and 183.31±7.00 GHz, each with two passbands and only sensitive to vertical polarization at nadir (0). MWHS/1 has a nominal field of view of 1.1° full-width at half-maximum, with pointing accuracy of 0.1° and 16 km FOV size at nadir. It has a scan range of ±53.35°, which is translated to a swath of 2660 km. Every 8/3 s, MWHS/1 scans through 98 Earth views, three space views, and three OBCT views. MWHS/2 on board FY-3C is an evolution of MWHS/1 with 15 channels in the range 89 to 191 GHz (0). More specifically, it has two additional channels in the 183 GHz water vapour absorption line with frequencies 183.31±1.8 and 183.31±4.5, thus it is similar to ATMS, while the rest of the technical elements are the same as for MWHS/1. Operations were suspended on 31 May 2015 for an anomaly investigation and resumed on 15 July 2015[3].

2.4 FCDR production

A brief description of the processing software and steps that are used for the creation of the C3S microwave humidity FCDR is provided below. More detailed information can be found the Algorithm Theoretical Basis Document (ATBD).

The software for processing the microwave humidity sounder data to create the C3S microwave humidity FCDR is developed in Python3. It is based on Matlab code that was used to produce the FIDUCEO MW humidity sounders FCDRs (0), thus the output from this system for the MHS 183 GHz channels is equivalent to FIDUCEO. The software does not follow single-orbit-processing logic for each sounder, but instead it interprets the data as continuous measurements over time. This is achieved by processing over the specific time range within slots of 8 hours of data moving forwards in time by applying a 4 h step. The 8 hours of data processing involves data quality control, pre-processing steps, and application of the measurement model including the uncertainty propagation. Then the central 6 h of the data are passed to the writing routine, which creates equator-to-equator output netCDF4 files following established data format conventions (C3S, CF-convention, ACDD - GCMD Science Keywords).

Microwave sounders measure Earth radiation in terms of counts by scanning across-track. For the conversion of counts to radiance or brightness temperature two calibration targets are used, the deep space and the on-board warm calibration target (OBCT). The space temperature, also known as cosmic background temperature, is constant at about 2.73 K. The on-board calibration target temperature is stabilised and measured by platinum resistance thermometers (PRT). With the known calibration target temperatures and their respective measured counts, a two-point calibration equation is applied that converts the earth view counts into radiance or brightness temperature.

Before applying the calibration equation, the counts from the space, the OBCT and the Earth views are quality controlled by using a number of tests in order to prevent the use of suspicious/corrupted data further along the processing chain. The same applies to the

[3] https://www.wmo-sat.info/oscar/satellites/view/115

PRT temperatures. Dubious data are flagged by setting one of the four quality flags and are excluded from further processing.

Pre-processing steps include the computation of the solar azimuth angle, the calculation of the Allan deviation of the counts over 300 scanlines as a measure of the uncertainty, and the estimation of possible impact from the moon on space counts. Another pre-processing step is the averaging of space and OBCT counts over 7 lines for MHS, MWHS/1 and 2 using a triangular function, while for ATMS the averaging is performed over 5 or 9 lines, depending on the channel, using a boxcar function to reduce the noise. In addition, the instrument temperature is calculated to estimate the possible corrections for any warm or cold bias, together with the non-linearity coefficient and the associated uncertainties. Finally, the antenna viewing position reported in the input file is compared with the computed position and if the difference is larger than a threshold, the respective pixel is flagged.

Calibration is performed using the quality controlled and averaged counts to radiance for MHS, MWHS/1 and MWHS/2 or to brightness temperature for ATMS. Firstly, the linear calibration is applied and then the non-linear correction term is calculated and added. This provides the antenna temperature, which is further corrected by applying the antenna pattern correction to compensate for side lobe effects on the antenna or the instrument mirror.

The major advantage of the C3S microwave humidity FCDR is that the brightness temperatures at 183 GHz channels are accompanied by their uncertainties. Three different kinds of uncertainties are provided per field of view: independent (or random), structured and common (or systematic). The independent uncertainty includes effects that affect a specific FOV, while the structured uncertainty describes effects that induce correlation among adjacent FOVs. The common uncertainty is applicable on much longer timescales (e.g. the lifetime of the instrument). The most important contributions to the independent uncertainty are due to earth counts and the earth view antenna position. On the other hand, the structured uncertainty is mainly affected by uncertainties in the space view counts, OBCT counts and OBCT temperatures, and in the case of MHS by the antenna position of the space view. Finally, the common uncertainty includes contributions from the OBCT temperature, the non-linearity coefficient, the warm and cold temperature bias correction, the polarisation correction coefficient, the antenna pattern correction coefficient, and the antenna position.

As mentioned above, in addition to uncertainties the C3S microwave humidity FCDR is also characterised by four quality flags, mainly assigned in the pre-processing steps. The first is the 'data_quality_bitmask', which summarises any issue related to the calibration target, such as moon intrusion or bad PRT measurements. The 'quality_issue_pixel_bitmask' is the second quality flag and this indicates any issue related to the calibration target measurement or earth view measurement, such as an insufficient number of valid calibration counts. The 'quality_pixel_bitmask' indicates issues related to the input data and the geolocation, such as invalid geolocation, invalid time variable, or if the L1a data are flagged. Finally, the 'quality_scanline_bitmask' indicates where there are issues related to the transmission of the data. In this study, all quality flags are considered, and only data where none of these flags are present are used in the analysis (Section 3.2).

3 Validation strategy

This section summarises the validation strategy and the data sets use to evaluate the C3S microwave humidity FCDR. The evaluation is performed using three separate approaches:

- Comparison of the FCDRs with equivalent operational BTs
- Comparison of the FCDRs with simulated top of atmosphere (TOA) BTs
- Analysis of simultaneous nadir overpasses (SNOs)

Further details for each evaluation strand are provided in Section 3.2. This study also includes an assessment of the uncertainties and quality control (QC) flags in the C3S microwave humidity FCDR, which is described in Section 3.2.4.

3.1 Validation Datasets

3.1.1 Operational MHS L1c data

MHS data for Metop A and B have been obtained from the NOAA Comprehensive Large Array-Data Stewardship System (CLASS). As NOAA CLASS provides files with raw counts (Level 1a) for MHS, the NWP SAF ATOVS and AVHRR Pre-processing Package (AAPP) software was used to obtain level 1c BTs. More specifically version 8.4 of AAPP has been used to process the MHS data on board Metop A and B into HDF5 L1c files. Metop A was declared operational the 15 May 2007, while for Metop B this occurred on 24 April 2013. In 2014 Metop A MHS unexpectedly entered fault mode on 27 March, a situation which lasted until the 20 May when the sounder recovered.

3.1.2 Operational ATMS L1c data

The ATMS Sensor Data Record (SDR) HDF5 files for S-NPP, which contain the brightness temperatures, have been obtained from the NOAA CLASS for the period 9 March 2017 to the end of 2018. For the period up to 8 March 2017, SDR reprocessed data have been kindly provided by NOAA. It is expected that there will be no change between the reprocessed and the operational datasets. The SDR dataset includes the antenna pattern correction. It is useful to note that the geolocation information is not included in the SDR file, but it is provided in a separate file. The commissioning phase for S-NPP ended in March 2012.

3.1.3 Operational MWHS/1 and 2 L1c data

The operational MWHS/1 and 2 L1c data with calibrated and geolocated BTs are sourced from HDF5 files provided to EUMETSAT by CMA, who operate the FY (FengYun) satellite programme. In the case of the MWHS/1 instruments, the files contain also the raw counts and much of the needed information to perform the two point calibration. This information has been used for the creation of the respective FCDR that has been evaluated in this study. FY-3A was declared operational on 12 January 2009, FY-3B the 2 June 2011, and FY-3C the 5 May 2014. The FY-3C satellite suffered some operational problems between May and August 2015 and for this reason, there is a data gap between 31 May and 11 July 2015. CMA had to make changes, which affected the environment temperature of MWHS/2 inducing biases especially for channels 183.31±3.0 GHz and 183.31±4.5 GHz (0).

3.1.4 SAPHIR L1b data

The Sounder for Atmospheric Profiling of Humidity in the Intertropics by Radiometry (SAPHIR) is on board the Megha-Tropiques satellite, which was launched on 12 October 2011. The orbit of Megha-Tropiques is circular and is characterised by a 20° inclination at the Equator and an altitude of 866 km (0). Thus, the orbit is asynchronous with a variable local equator crossing time. SAPHIR is a cross-track scanning radiometer observing the Earth's atmosphere with a scan angle of ±42.96° around nadir inducing a swath of 1700 km. Each scan line is composed of 130 non-overlapping FOVs with size of 10 km at nadir. SAPHIR has six channels located in the water vapour absorption band with frequencies 183.31 ± 0.2 GHz, 183.31 ± 1.1 GHz, 183.31 ± 2.8 GHz, 183.31 ± 4.2 GHz, 183.31 ± 6.8 GHz and 183.31 ± 11 GHz. It should be noted here that the SAPHIR frequencies are slightly different from the other MW humidity sounders (Table 2). SAPHIR is not included in the C3S microwave humidity FCDR, and as such, it constitutes an independent source used for validation and more specifically through the analysis of simultaneous nadir overpasses (SNOs; Section 3.2.3). For this purpose, L1A2 orbit-wise HDF5 SAPHIR data have been downloaded from ICARE[4]. The files include geolocated BTs without uncertainties per pixel. In this study, all the SAPHIR data have been used except those with pixel quality flag for the bit number 15 with value equal to 1 (meaning invalid brightness temperature).

3.1.5 Simulated TOA BTs

TOA BTs have been simulated in this study using ERA5 profiles and the NWP SAF (Numerical Weather Prediction Satellite Application Facility) Radiance Simulator (RadSim). RadSim (0) constitutes an interface for the radiative transfer model RTTOV (0), also from the NWP SAF. In this study, RTTOV is used to simulate the BT that a MW humidity sounder would see at a particular satellite zenith angle for the 183 GHz channels (Table 2). Surface emissivities for the simulations were computed using the FASTEM-6 model (0) over sea, and the emissivity atlases TELSEM2 (0) based on Special Sensor Microwave Imager/Sounder (SSMIS) observations over land and sea ice. The emissivity atlases are provided at 0.25 degrees resolution through the NWP SAF RTTOV webpage[5]. In some locations, mainly over coastlines and close to sea ice – open ocean boundaries, there is inconsistent information between the ERA5 and the SSMIS atlases concerning the surface type (sea, land, sea ice) given the slightly different resolution of the datasets (~0.28° vs. 0.25°) and the fact that the atlases are a climatological product (one atlas for each calendar month). In those cases, a constant value for emissivity is used, assuming the surface type indicated by ERA5, which is 0.95 for land and 0.92 for sea ice (the default values in RadSim). It is useful to note that ERA5 assimilates observations from all the instruments in Table 1, thus the comparison between ERA5 and the C3S microwave humidity FCDR cannot be considered independent.

Simulations are performed for each satellite field of view (FOV), considering the respective longitude, latitude, satellite zenith and azimuth angle. Given that the ERA5 temporal resolution is 1 hour, no temporal interpolation is attempted and the closest hourly ERA5 file to satellite time observation is used for the simulation. By contrast, RadSim interpolates the ERA5 fields into the spatial location of the satellite observations. The simulations only include liquid clouds, so scattering effects are not considered.

---

[4] http://www.icare.univ-lille1.fr/mt
[5] https://nwpsaf.eu/site/software/rttov/

3.2 Validation methods

For all evaluation strands, only C3S microwave humidity FCDR data with a QC flag setting of zero are considered in the analyses. The FCDR uncertainty data are only considered in the SNO analysis, as uncertainties are not available for either the operational or simulated BT data sets.

3.2.1 Comparison between FCDRs and operational BTs

This evaluation strand is concerned with the comparison between the C3S microwave humidity FCDR (Section 2) and the equivalent operational BTs (Section 3). Ideally, the two data sets should agree very closely as they have been derived from the same raw satellite observations (i.e. digital counts). However, differences may arise due data artefacts or anomalies that have been corrected for in one data set but not in the other; it is expected that the C3S microwave humidity FCDR should be more consistent over time as the calibration does not change and should be optimal. The following diagnostics have been used to assess the agreement between the two BT products for each instrument and for each 183 GHz channel:

- Time series of daily, globally-averaged BTs and their standard deviations for each data set
- Time series of the daily, zonally-averaged BT differences (FCDR minus operational) and their standard deviations within latitude bands of ±30°, ±60°, ±90° and the mid-latitudes between 60°S-30°S and 30°N-60°N
- Hovmöller diagrams showing the variation in daily mean differences for 1°-latitude bands with time

Both the C3S microwave humidity FCDR and operational data sets are screened using the FCDR QC flags. This approach was chosen partly to ensure consistency between the different sensor comparisons and partly because some of the QC flags provided with the operational data sets are very poorly understood. For example, no clear user documentation exists in English for the MWHS QC flags. The data are also screened for unrealistic outliers using the range 100-400 K.

The C3S microwave humidity FCDR data are matched with the equivalent operational data using the latitudes and longitudes of the observations. Any duplicate scanlines are discarded and only the first occurrence is included in the analysis, whatever the value of the QC flag. The results from this evaluation strand are presented in Section 4.1.

3.2.2 Comparison between the FCDRs and simulated TOA BTs

This analysis compares the C3S microwave humidity FCDR BTs (Section 2) with equivalent BTs that have been simulated using RadSim and ERA5 (Section 3.1.5). The following diagnostics have been used to assess the agreement between the data sets for each instrument and for each 183 GHz channel:

- Global maps of BT differences (FCDR minus simulation) for a specific date showing the differences between the two data sets on a per FOV basis. Two global maps are considered, separating the day into two parts for visualization, with the differences that occur between a) 00:00 and 11:00 UTC, and b) 12:00 and 23:00 UTC.
- Time series of zonally-averaged daily BT differences and the standard deviation of the differences (FCDR minus simulation)

- Time series of percentiles for the distributions of daily BT differences
- Time series of double-differences (Section 1.7), using both Metop A and S-NPP as the reference data sets
- Hovmöller diagrams showing variation in daily mean differences by latitude with time

For the time series and Hovmöller diagrams, the C3S microwave humidity FCDR observations contaminated by clouds (mostly convective or precipitating) are excluded from the analysis. This is to ensure the observations are consistent with the simulations, where scattering effects are not included (Section 3.1.5). Cloud-contaminated pixels are removed from the observations using two tests following the work of Buehler et al. (0). Firstly, the observed BTs are discarded if the difference between the channels 183.31±7 GHz (or 190.31 GHz for MHS) and 183.31±1 GHz is negative. The second test discards the observations if the values of channel 183.31±1 GHz are lower than a minimum view-angle dependent value based on cloud-free simulations. It is important to note that these two tests have been developed mainly for Upper Tropospheric Humidity (UTH) retrievals and are not optimized (i.e. masking clouds correctly) for the lower peaking channels (e.g. 183.31±7 GHz or 190.31 GHz). In addition, as the second test was developed for MHS, the threshold values used for the larger view angles of ATMS and MWHS/1 and 2 have been extrapolated. To minimize the impact of emissivity uncertainties over land and ice, which are based on climatological values, the satellite-simulation differences are only calculated over the ocean and large lakes within ±60 degrees. Analysis of the inter-comparisons between the C3S microwave humidity FCDR and the simulated data is presented in Section 4.2.

3.2.3 Analysis of Simultaneous Nadir Overpasses

The final evaluation strand consists of direct comparisons between the FCDRs. This is performed through the analysis of SNOs (Section 1.7). In this study, satellite observations from two different sensors are considered SNOs if the observations are within 5 km and 5 minutes, and for where the satellite zenith angles are within 5°. The agreement between the different satellites indicates the relative difference between them and is therefore indicative of the accuracy. The limitation of this approach is that the SNOs for the satellites listed in Table 1 nearly always occur near the poles as they are in polar orbit. To increase the number and geographical coverage of this analysis, SNOs with SAPHIR (Section 3.1.4) are also included. For the comparison, channels are matched by their central frequency as shown in Table 2. For example, MWHS/1 channel 3 is compared to ATMS channel 22. For SAPHIR, the 183.31±1.1 GHz channel is matched with the 183.31±1.0 GHz on MHS/MWHS/ATMS, the 183.31±2.7 GHz channel is matched with the 183.31±3.0 GHz channel, and the 183.31±6.6 GHz channel is matched with the 183.31±7.0 GHz channel on MWHS/ATMS and 190.31 GHz channel on MHS. Only channels present on all instruments are considered in this analysis, i.e. 183.31±1.0 GHz, 183.31±3.0 GHz and 183.31±7.0 GHz (or 190.31 GHz for MHS). The following diagnostics are assessed for each sensor pairing:

- Scatter plot of BTs from sensor 1 vs BTs from sensor 2
- Scatter plot of BT differences vs BTs from a reference sensor (where the reference sensor is selected from the sensor pair).
- Time series of BT differences presented as box plots.

Both scatter plots are plotted as two-dimensional histograms with bin-size of 1 K and data density are normalised to power 0.3. Only data that are not masked by the QC flag checks are included in the comparisons.

For the box-plot time series, each box represents one month of data. As for the scatter plots only data that are not masked by the QC flags are included. The central line is the median (or 50th percentile), the box represents the inter-quartile range, and the whiskers are the 10th and 90th percentiles. The y-range of each plot represents the 1st and 99th percentile from the distribution of all the BT differences, with the exception of the FY-3C and SAPHIR pairing, where the 4th and 99th percentiles are used for aesthetic reasons.

All available data are analysed, so the SNOs presented encompass the full range of data available (Table 1), with the exception of the SNOs involving SAPHIR, which are restricted to the eastern hemisphere only due to the computational time required to process these SNOs. The analysis of SNOs from the FY-3A and FY-3B sensor pairing are also excluded from this study as there are many unexplained, unrealistic anomalies that require much further investigation. All other available SNO pairs are included and the analysis of these SNOs is presented in Section 4.3.

3.2.4 Assessment of uncertainties and quality flags

One of the key benefits of the FIDUCEO project for generating FCDRs is in the provision of well-characterised uncertainties and quality flags (Section 2.4). For the analysis of uncertainties, two diagnostics are considered:

- Time series of percentiles describing the global distribution of values for each uncertainty component and the total uncertainty (Section 1.7)
- Uncertainty overlap analysis using C3S microwave humidity FCDR data from the SNOs (note SAPHIR is excluded from this analysis as no uncertainties are provided with these data)

The aim of the first diagnostic is to examine the values of the uncertainties and assess whether they are reasonable with respect to the other findings this study. The percentiles shown are for the 1st, 25th, 50th, 75th, and 99th percentiles. The daily distributions are calculated from using the total uncertainty for each valid pixel on that calendar day. The second diagnostic, the uncertainty overlap analysis, should indicate whether the uncertainties for different instrument pairs are collectively too large or too small.

For the assessment of the quality flags, the following diagnostics are considered:
- Global maps showing where data are flagged in each QC test
- The mean differences and their standard deviation of the FCDR minus operational BT differences with the quality flags on and quality flags off (i.e. used and not used)
- Time series of the percentage of pixels rejected by quality flags for each instrument

The aim of the first diagnostic is to assess whether the QC information appears to be reasonable. The second diagnostic should demonstrate whether outliers - including those that may not be visible to the eye – are in general correctly flagged by the QC tests. The final diagnostic is there simply to indicate what fraction of the data are removed by the QC tests and to check that this fraction is neither too high, nor too low.

4 Validation Results

4.1 Comparison between FCDRs and operational BTs

4.1.1 MHS on board of Metop A

4.1.1.1 Time series analysis

The time series of mean differences for Metop A is shown in Figure 12. In general, the C3S microwave humidity FCDR is warmer than the operational BTs by up to 0.1 K. There is a notable period of approximately 3 months in the first half of 2007 where the differences between the FCDR and operational data sets are about 0.4 K to 0.7 K. This is due to an error in the Metop A data files from NOAA CLASS, where there are incorrect calibration corrections in the file header for the cold space and for the warm target temperature for the period 27th February to 19 June 2007. This error is not present in the FCDR. There is a weak negative trend in the 183.31±3 GHz channel, which induces a change in the mean BT of <0.05 K and is therefore considered negligible. The time series of mean BT differences also exhibits a small annual cycle with amplitude of ~0.025 K, which is again considered negligible. For the 183.31±3 GHz and 190.31 GHz channels, there is some stratification of the differences by latitude band, but this is <0.05 K, which again is considered negligible. The global average is closer to zero, which suggests that the differences at higher latitudes must be ~0 K.

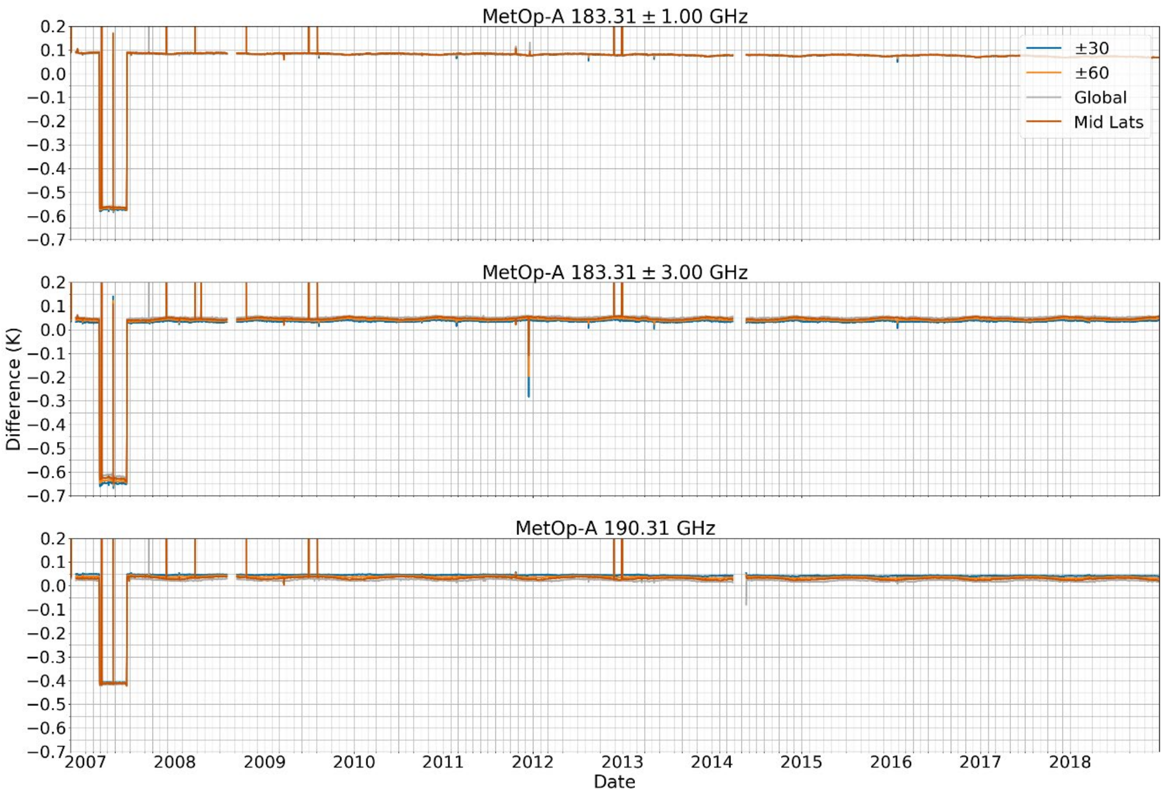

Figure 1: Time series of mean differences for Metop A (FCDR minus NOAA/CLASS L1b) for each of the 183 GHz channels and the 190.31 GHz channel (equivalent to a 183.31±7 GHz channel)

There are a few spikes in the agreement, particularly before 2010, most of which correspond to spikes in the standard deviations (not shown). The time series of the number of available data are generally quite stable, although noisy, and offer no insight into these anomalies. Apart from a few spikes, the time series of standard deviations of the daily mean differences is very stable and close to ~0.1 K for all channels (not shown). Notably, the 2007 anomaly is not visible in this time series.

4.1.1.2 Hovmöller diagrams

Figure 13 shows the Hovmöller diagram for the mean daily differences for the Metop A 183.31±1 GHz channel. The 3-month discontinuity in 2007 with negative mean differences is clearly visible. As for the FY-3 analyses, annual cycles are present at high latitudes and in the tropics, and there are linear features at about 75°N/S that correspond to the sharp increase in the number of observations at this latitude (not shown). However, as for FY-3B and -C the amplitudes of these annual cycles are <<0.1 K and are therefore considered negligible. Similar features are also apparent in the Hovmöller diagrams for the standard deviation of the mean daily differences although as for the standard deviation of the differenced time series, the 2007 discontinuity is not visible (Figure 14). In general, both the daily mean differences and their standard deviations appear very stable. Notably, the linear features that were seen in the FY-3 Hovmöller diagrams at ~0° and ~17°N are not present, which further confirms that the source of these features is in the operational FY-3 data, rather than in the C3S microwave humidity FCDR.

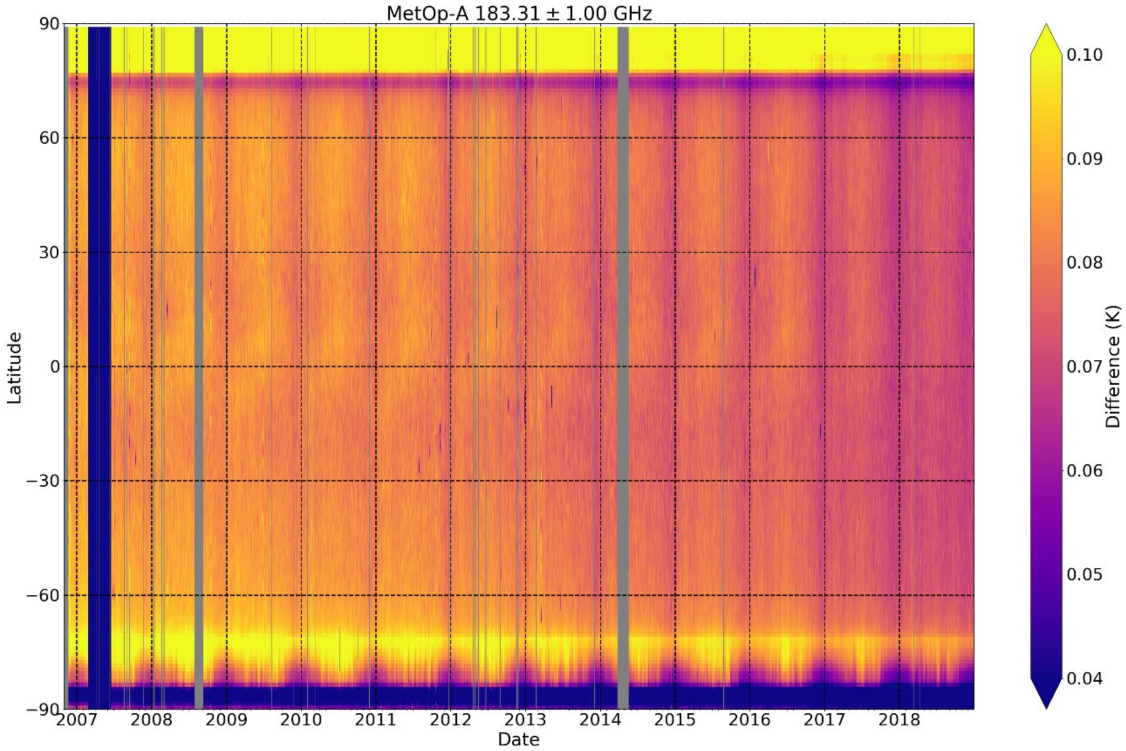

Figure 2: Hovmöller diagram showing the change in daily mean difference by latitude with time for Metop A for the 183.31±1 GHz channel. The grey bands correspond to data missing from the analysis.

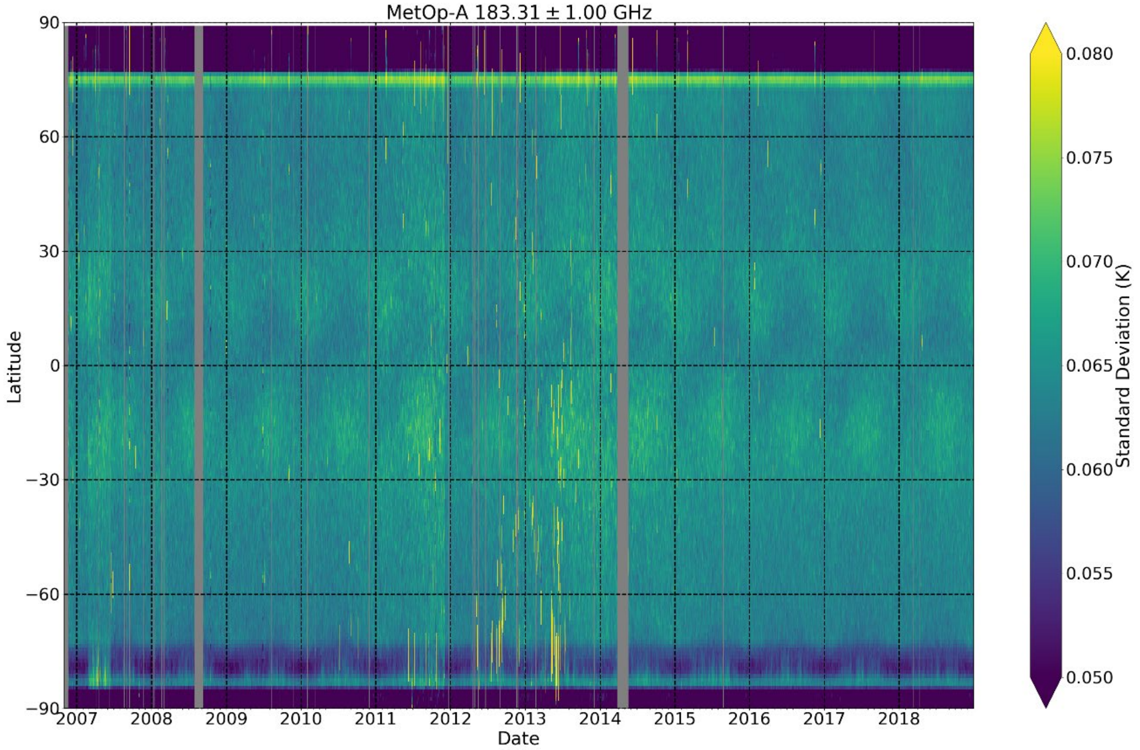

Figure 3: Hovmöller diagram showing the change in standard deviation of the daily mean difference by latitude with time for Metop A for the 183.31±1 GHz channel.

4.1.1.3 Key points

- The C3S microwave humidity FCDR is generally up to ~0.1 K warmer than the operational Metop A BTs
- The standard deviation of the daily mean differences is very stable and close to ~0.1 K for all channels.
- There is a 3-month discontinuity in the mean differences in 2007 due to a calibration error in the operational data files from NOAA CLASS. This error is not present in the FCDR
- The linear features in the standard deviations at ~0° and ~17°N that were present for FY-3 are not present for Metop A, which further confirms that the source of the linear features is in the operational FY-3 data rather than the C3S microwave humidity FCDR.

4.1.2 MHS on board of Metop B

4.1.2.1 Time series analysis

Figure 15 shows the time series of daily mean differences for all channels on Metop B. In general, the differences between the C3S microwave humidity FCDR and operation data are negligible at <0.05 K for all channels. As for Metop A, the time series show stratification of the differences by latitude band and annual cycles, but given the scale of the differences, these features are also considered to be negligible. There is a general upward trend in the BT differences for the 183.31±1 GHz channel, but the overall change is very small (<0.05 K). The time series of standard deviations is also very stable and very close to zero and apart from a few spikes, shows no notable discontinuities or anomalies (not shown).

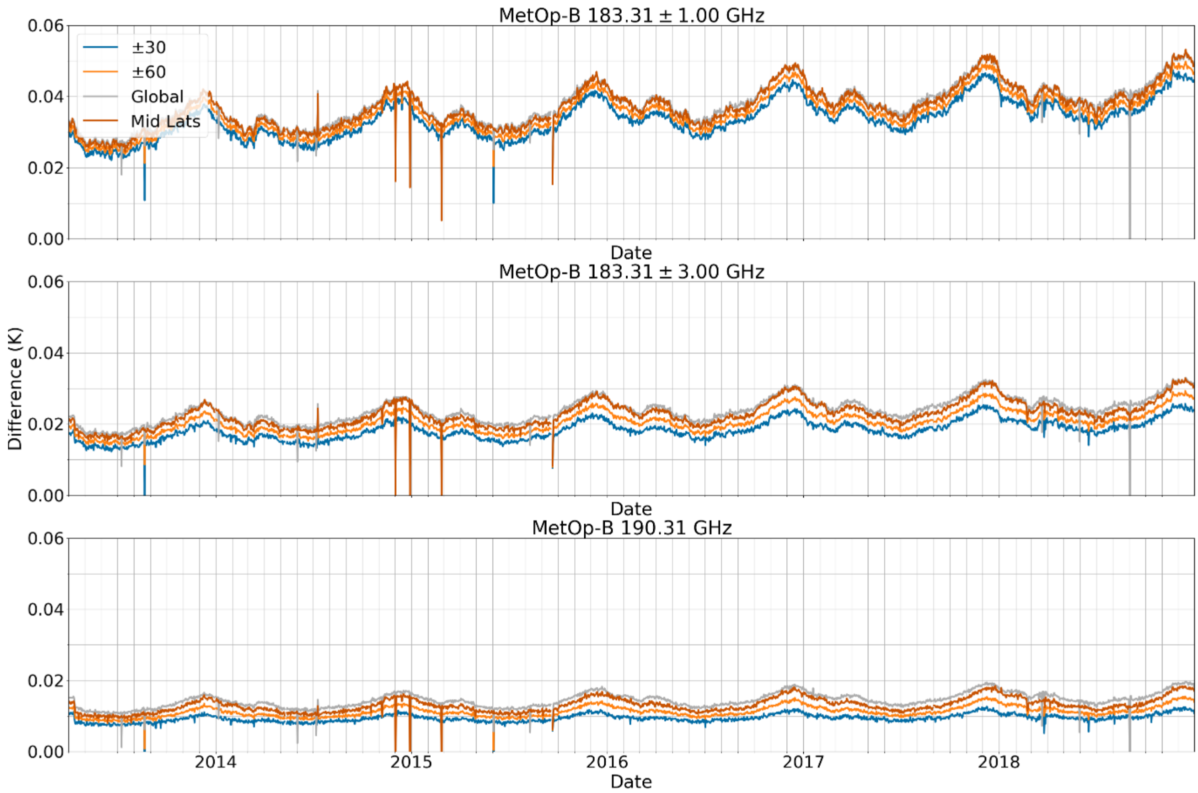

Figure 4: Time series of mean differences for Metop B (FCDR minus operational) for each of the 183 GHz channels and the 190.31 GHz channel (equivalent to a 183.31±7 GHz channel)

4.1.2.2 Hovmöller diagrams

Figure 16 shows the Hovmöller diagram for the Metop B daily mean differences for the 190.31 GHz channel. The diagram confirms the stability of the differenced time series with no visible anomalies or discontinuities. As seen for the FY-3 analysis and for Metop A, there are annual cycles present in the tropics and at higher latitudes, but as the amplitude of these changes is only a few hundredths of a K it is therefore negligible. The results for the 183 GHz channel standard deviations for all channels are similar, with no significant features (not shown).

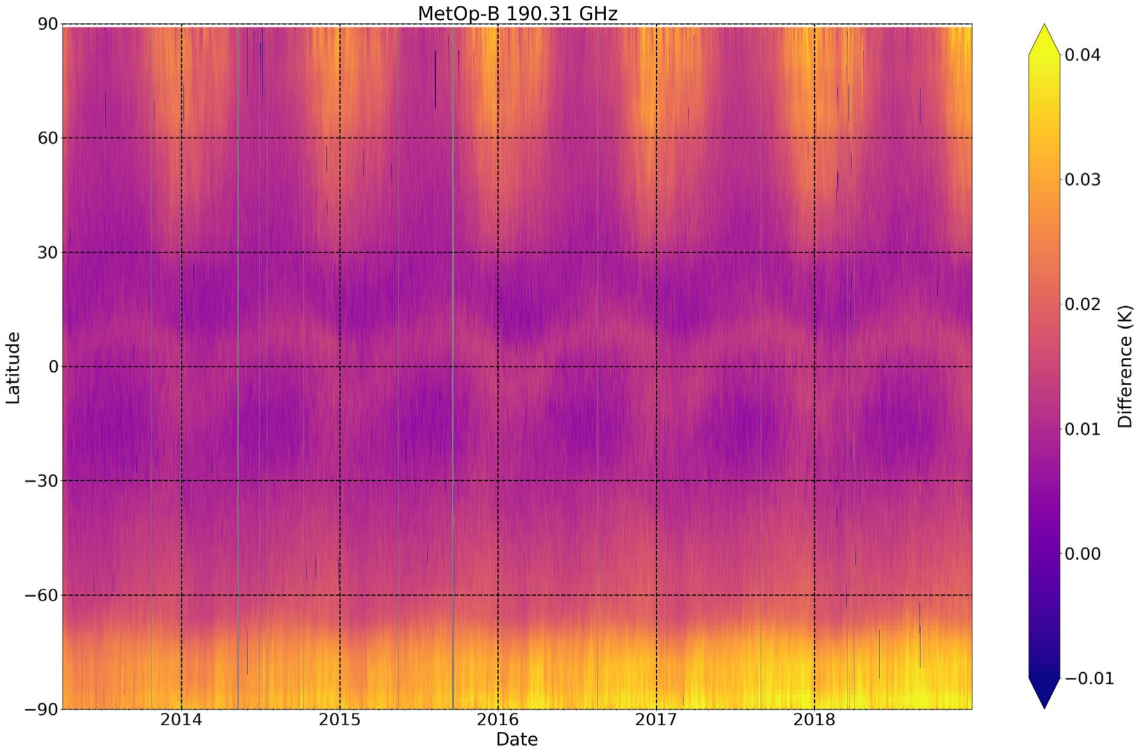

Figure 5: Hovmöller diagram showing the change in daily mean difference by latitude with time for Metop B for the 190.31 GHz channel. The grey bands correspond to data missing from the analysis

4.1.2.3 Key points

- The differences between the Metop B FCDR and operational BTs are typically a few hundredths of a K and are generally very stable.
- There are no significant features in the differences between the data sets. (Any features show amplitudes of a few hundredths of a K and are therefore considered negligible.)

4.1.3 ATMS on board of S-NPP

4.1.3.1 Time series analysis

Figure 17 shows the time series of differences between the C3S microwave humidity FCDR and operational data for S-NPP. In general, the time series are very stable, although the differences are much larger than for the other sensors analysed in this study. The smallest difference of ~1.4 K is observed for the 183.31±7 GHz channel, followed by the 183.31±1 GHz channel (~1.5 K), 183.31±1.8 GHz channel (~1.7 K), 183.31±4.5 GHz channel (~1.8 K), and finally the 183.31±3 GHz channel (~1.9 K). In general, there is very little difference between the different zonal bands until mid-2016, when the global time series is ~0.1 K lower than any of the other zonal bands. Given the other zonal bands are for the tropics and mid-latitudes, this negative change in the differences must occur at high latitudes. This change is also apparent in the time series of standard deviations (Figure 18).

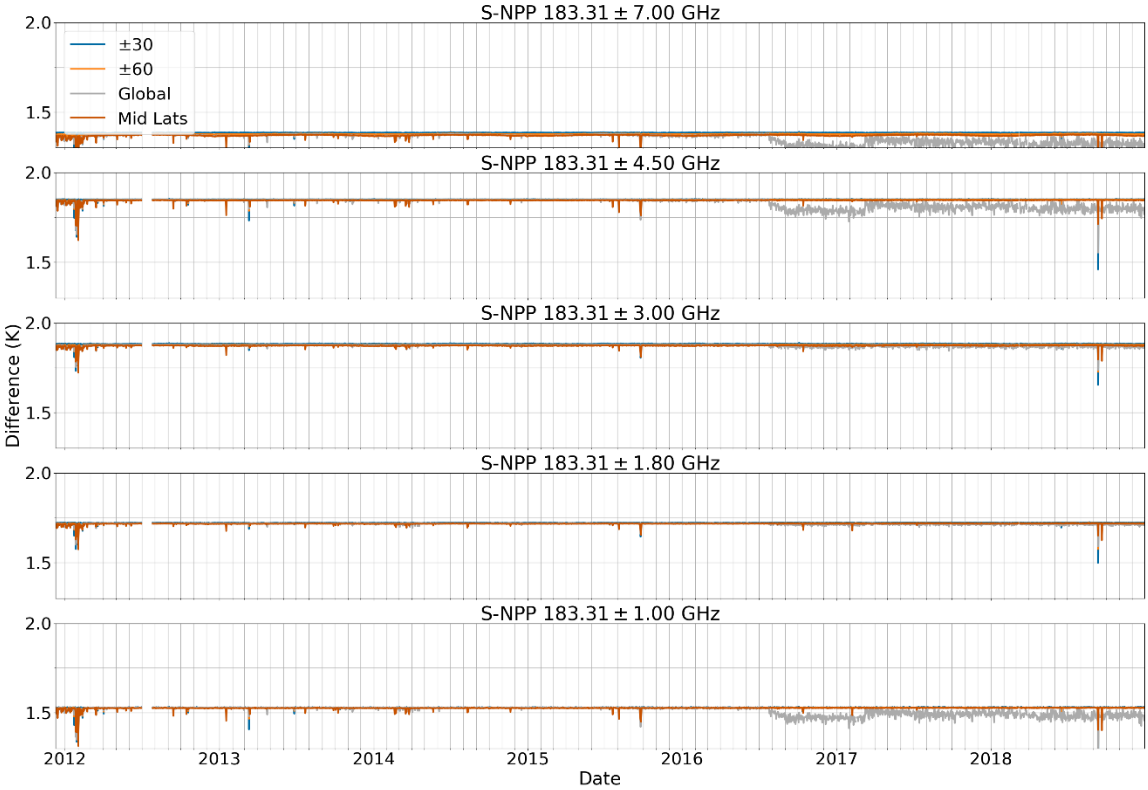

Figure 6: Time series of mean differences for S-NPP (FCDR minus operational) for each of the 183 GHz channels

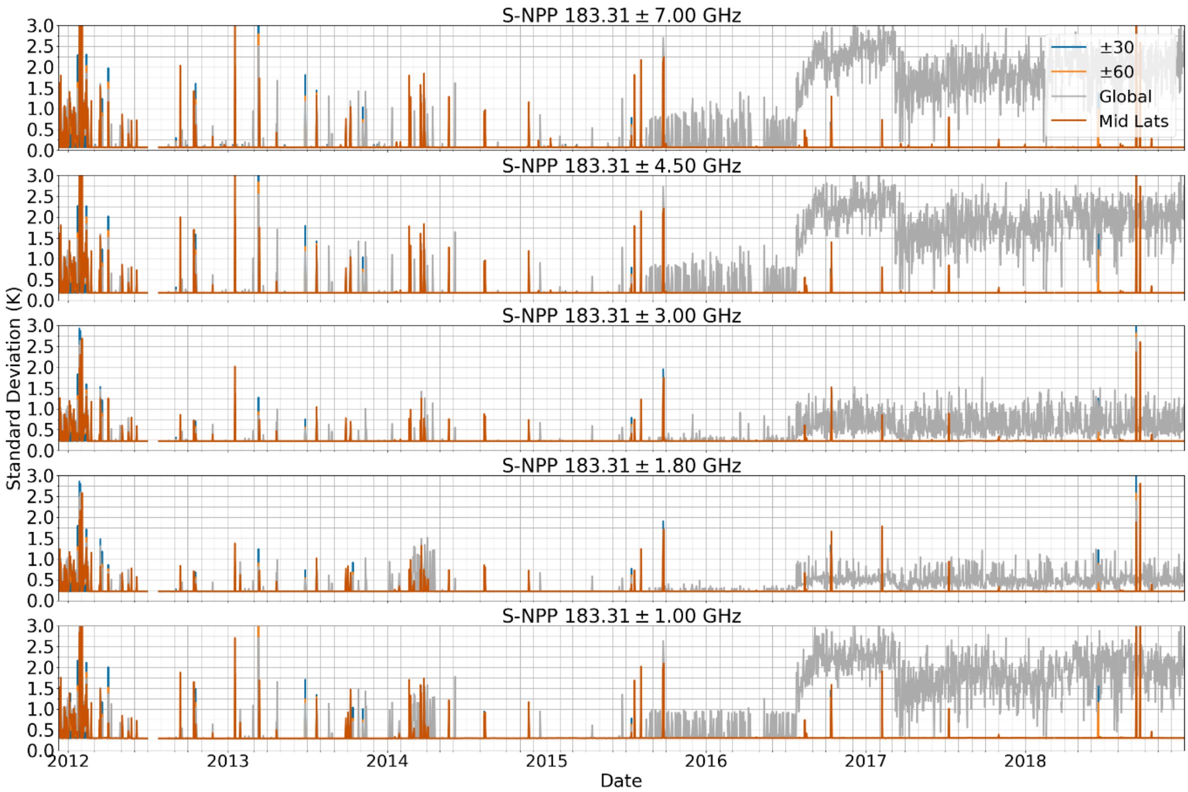

Figure 7: Time series of the standard deviation of differences for S-NPP (FCDR minus operational) for each of the 183 GHz channels.

As for the mean differences, the standard deviations are generally quite stable and <0.5 K, apart from a few spikes, until mid-2016 when the global standard deviation increases by up to 3 K. However, this anomaly is preceded by a high-number of spikes in the global time series from mid-2015, which suggests that the high-latitude anomaly may begin earlier. Other than the onset of this high-latitude anomaly, there are no other visible discontinuities in the time series.

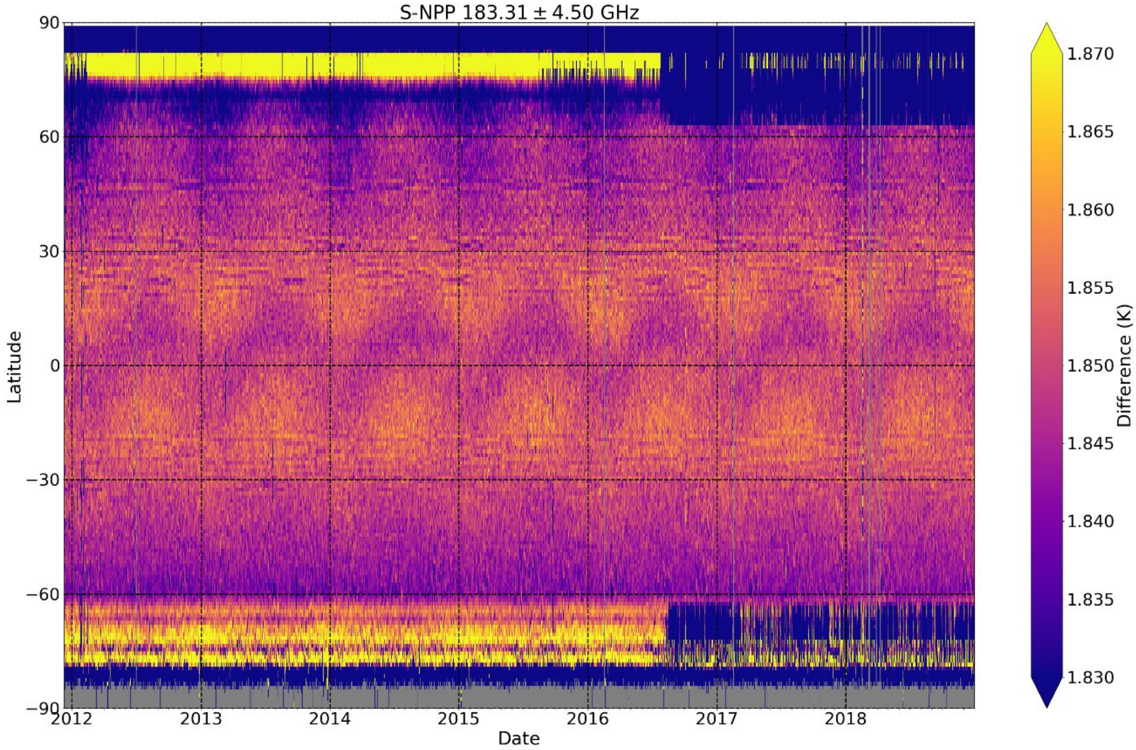

Figure 8: Hovmöller diagram showing the change in daily mean difference by latitude with time for S-NPP for the 183.31±4.5 GHz channel. The grey bands correspond to data missing from the analysis

4.1.3.2 Hovmöller diagrams

Figure 19 shows the Hovmöller diagram for the daily mean differences for the 183.31±1.8 GHz channel. The diagram shows a clear zonal structure with much larger mean differences above 60°N/S until mid-2016, after which the differences reduce significantly. In the southern hemisphere, the onset of the high-latitude change is abrupt. However, in the northern hemisphere, the onset appears more gradual and begins in mid-2015. This may explain the earlier change in the global standard deviations compared with the global mean difference time series noted above. This change in the mean differences at high-latitudes is seen in the Hovmöller diagrams for all channels, although there is some variability between channels. For example, there is little evidence of any phased onset of the reduction in mean differences in either hemisphere for the for the 183.31±1.8 GHz channel (Appendix A). The Hovmöller diagrams for the standard deviation of the daily mean differences also show the high-latitude discontinuity (Appendix A: 183.31±7 GHz channel only).

As seen for the FY-3 and Metop analyses, there are annual cycles present in the tropics and at higher latitudes. However, it should be noted that in Figure 19, and for the three higher-peaking channels, the amplitude of these changes is only a few hundredths of a

K and is therefore negligible. However, for the 183.31±7 GHz channel these annual cycles involve changes of ~0.1 K (not shown), which may be significant. The Hovmöller diagram for the standard deviation of daily mean differences shown in Appendix A also shows annual cycles, but this is also considered to be negligible given that the changes are only a few hundredths of a K at most.

Although the mean differences between the S-NPP FCDR and operational data sets are large, the Hovmöller diagrams demonstrate that for most of the globe, the variation in mean difference in space and time is actually very small. This is also supported by the low standard deviations, and low variability of the standard deviations in space and time for most of the globe (Appendix A). This suggests that the difference between the two data sets could be systematic.

4.1.3.3 Key points

- Mean differences between the C3S microwave humidity FCDR and operational data sets for S-NPP range between ~1.4 K and ~1.9 K (FCDR is warm).
- The standard deviations of the global mean differences are generally <0.5 K. However, for most of the globe the variation in the mean and standard deviation is of the order of hundredths of a K, which suggests the large mean difference between the C3S microwave humidity FCDR and operational data sets may be systematic.
- Both the mean differences and standard deviations are quite stable until mid-2016, when the standard deviation of the global mean differences increases by up to ~3 K, and the global mean difference reduces and becomes noisier.
- This change in mid-2016 is attributed to a change in the mean difference above 60°N/S, which is higher than for the rest of the globe until this time, after which begins to reduce. For some areas above 60°N/S, the mean difference begins to reduce from mid-2015, but for most areas, the change is abrupt.
- As seen for the FCDR vs operational data comparisons for other sensors, there is an annual cycle in daily mean differences and their standard deviations in the tropics and at high latitudes. However, the amplitude of these annual cycles is insignificant for all channels, other than perhaps the 183.31±7 GHz channel where the amplitude reaches ~0.1 K.

4.1.4 MWHS/1 on board of FY-3A

4.1.4.1 Time series analysis

Figure 91 presents the C3S microwave humidity FCDR and operational time series for FY-3A. There are significant differences between the two data sets, which appear to reduce with decreasing channel sounding height. This can be seen more clearly in the differenced time series shown in Figure 102, which confirm the that the C3S microwave humidity FCDR is cooler than the operational BTs by up to 1.5 K. There is a general improvement in the agreement between the C3S microwave humidity FCDR and the operational BTs with time for the 183.31±1 GHz channel. However, it should be noted that this does not necessarily imply that the accuracy of the C3S microwave humidity FCDR is improving over time for this channel. By contrast, the mean differences for the183.31±3 and 183.31±7 GHz channels are reasonably stable after the end of the commissioning phase (12 January 2009) and before mid-2012. The 183.31±7 GHz channel shows the best agreement between the datasets (approx. -0.5 K mean difference), followed by the 183.31±3 GHz (approx. -1.0 K) and finally the 183.31±1

GHz (approx. -1.0 to -1.5 K). There is some stratification of the differences by latitude band, only visible with this plotting range for the 183.31±7 GHz channel, with slightly smaller differences observed at tropical latitudes. All channels show a significant discontinuity in September 2013, which is clear in the differenced time series but only just visible in both the individual data set time series (Figure 91). The cause of this discontinuity is unknown, but as it appears to be in both data sets, this is likely an instrumental issue for which no correction has been applied. A smaller discontinuity is also apparent earlier in May 2013, which is most significant in the 183.31±7 GHz channel.

The time series of the standard deviations of the differences for FY-3A (Appendix A) indicates typical standard deviations that range between <0.1 K (183.31±1 GHz) and ~1.7 K (183.31±7 GHz). While the standard deviation for the 183.31±1 GHz and 183.31±3 GHz appear quite stable in time, there is a slight decrease in standard deviation with time for the 183.31±7 GHz channel. The May and September 2013 discontinuities are also present in the time series of standard deviations, although the May event is barely visible. The number of data available for each day (Appendix A) is quite stable after the commissioning phase (January 2009) apart from a large drop in available data around the end of 2010/early 2011, the cause of which is unknown.

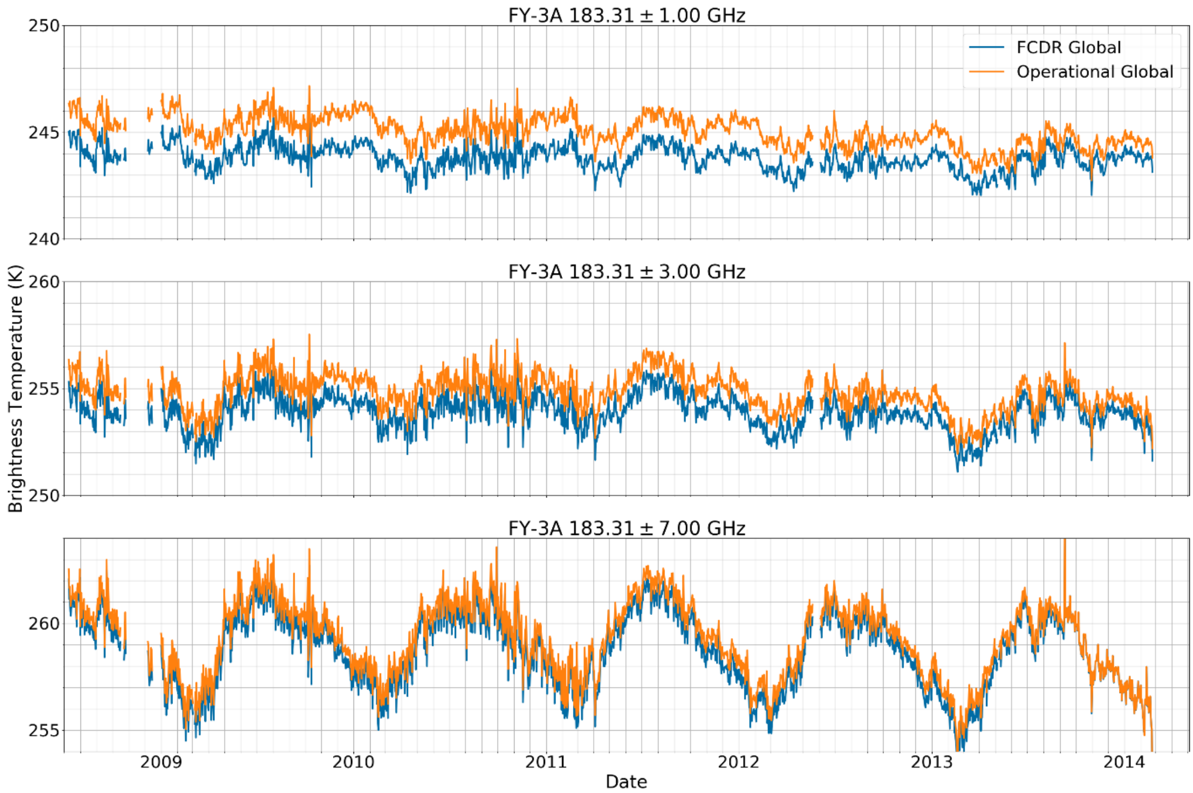

Figure 9: Time series of FCDR (blue) and operational (orange) FY-3A data.

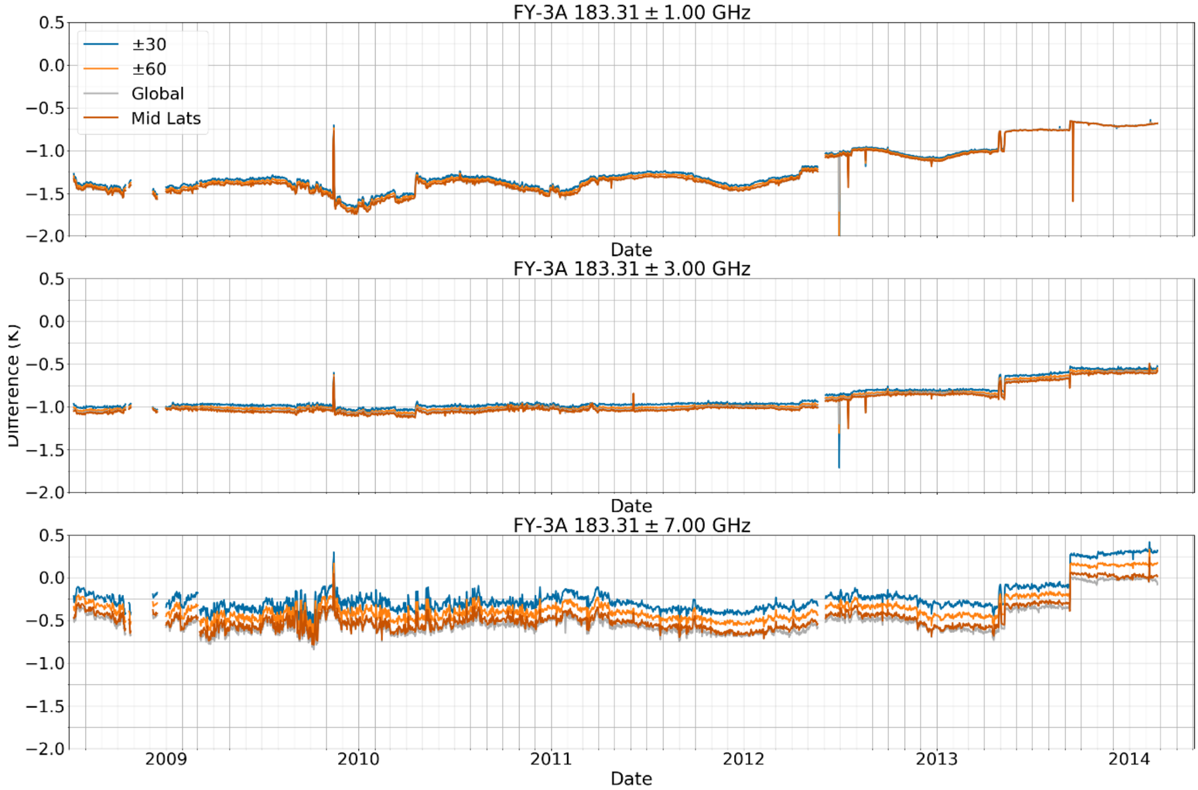

Figure 10: Time series of mean differences for FY-3A (FCDR minus operational) for each channel. The data have been averaged over latitude bands ±30° (blue), ±60° (orange), ±90° (grey), and mid-latitude bands 60°S-30°S and 30°N-60°N (red).

4.1.4.2 Hovmöller diagrams

Figure 113 shows the Hovmöller diagram for the 183.31±1.0 GHz channel. The data show little zonal structure, which is expected given the results shown in Figure 102 (time series of mean difference stratified by latitude bands). The variation in agreement with time apparent in Figure 102 is also evident here: the differences are strongly negative and quite stable until May 2013, when they become less negative, followed by another positive step change in September 2013. These discontinuities are also visible in the results for the other two channels (183.31±7 GHz channel shown in Appendix A). A weak annual cycle in the differences is evident in the tropics and at mid-to-high latitudes. The annual cycles may be a result of a BT-dependent difference between the two data sets that is apparent because of the colder BTs that occur when cloud is present. Thus, when there is increased cloud, there will be an increased occurrence of negative BT differences. Over the tropics, this is likely to be related to the Inter Tropical Convergence Zone (ITCZ), which tends to have a stronger presence in the northern hemisphere (NH) due to the larger land masses. The annual cycle at mid-to-high latitudes, for example, could be due to the increased cloud present during the mid-latitude storms that occur mainly during the winter months. Some linear features at high latitudes are evident in the mean differences for the 183.31±7 GHz; these are also present in the equivalent standard deviation plots (Figure 124; 183.31±1 GHz and 183.31±3 GHz channels not shown). These linear features appear to correspond to the change in the number of data due to the higher number of overpasses at these latitudes (Figure 135) so may be some artefact of the averaging process. There are also linear features visible in the standard deviations for all channels at ~0° and ~17°N. These linear features are attributed to the calibration process that is applied to the data in the CMA orbit files, which start at 17°N and end at 0.5°N. In both the operational data and FCDR processing the calibration counts are averaged over seven scanlines, which causes a discontinuity for the first and last six scan lines in the operational orbit files where the full seven scanlines required are not available. This discontinuity is avoided in the FCDR processing by calibrating the data in ~10000 scanlines 'blocks' that are about 5 orbits and include the preceding and subsequent scanlines required for the full seven-line averaged calibration (0). Thus, the linear features are an artefact in the operational data and not the C3S microwave humidity FCDR.

The September 2013 discontinuity is evident in the standard deviations for the 183.31±3 GHz and 183.31±7 GHz channels, where the standard deviation is reduced, but not in the 183.31±1 GHz channel. By contrast, the May 2013 discontinuity is present in the 183.31±1 GHz and 183.31±3 GHz channels, but not in the 183.31±7 GHz. These observations are consistent with the time series analysis presented in Section 4.1.4.1.

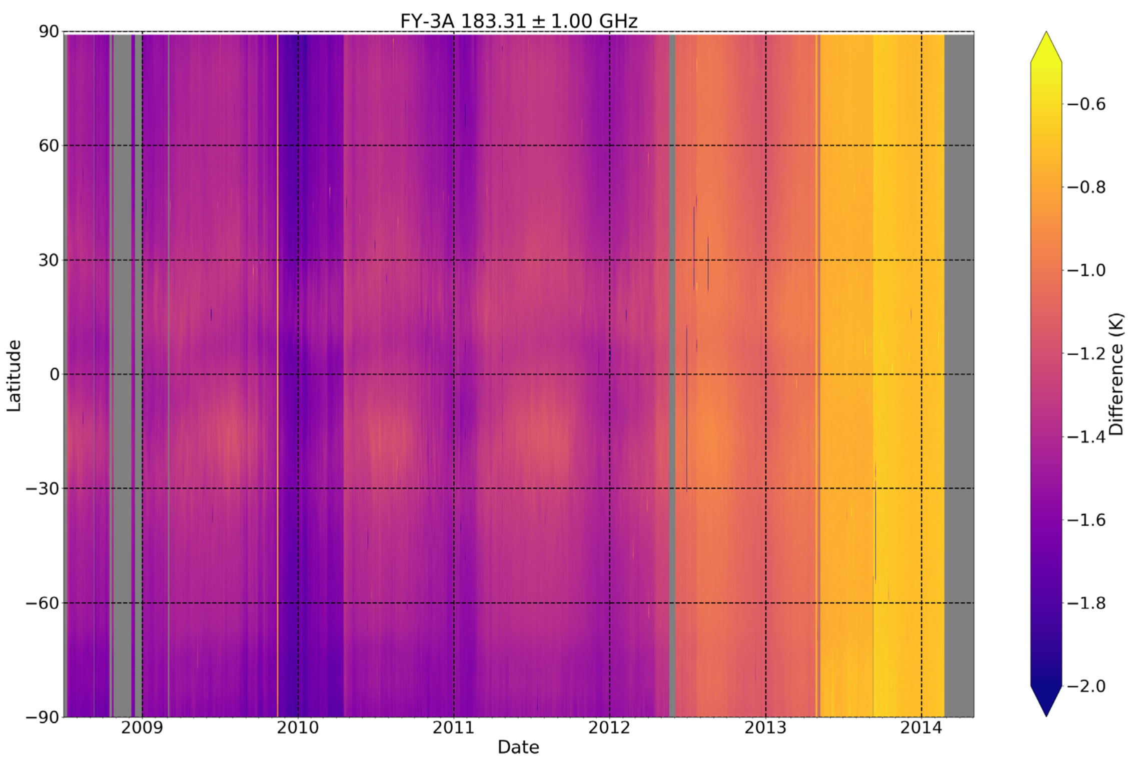

Figure 11: Hovmöller diagram showing the change in daily mean difference by latitude with time for FY-3A for the 183.31±1 GHz channel. The grey bands correspond to missing data.

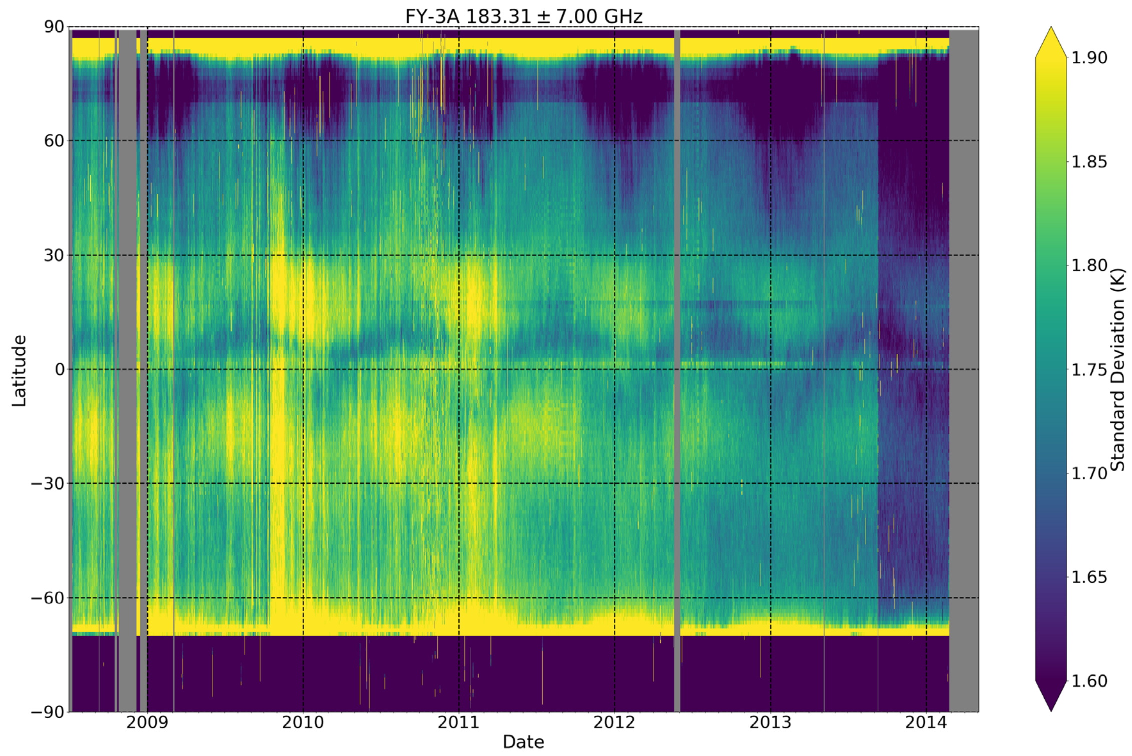

Figure 12 Hovmöller diagram showing the change in the standard deviation of the daily mean differences by latitude with time for FY-3A for the 183.31±7 GHz channel.

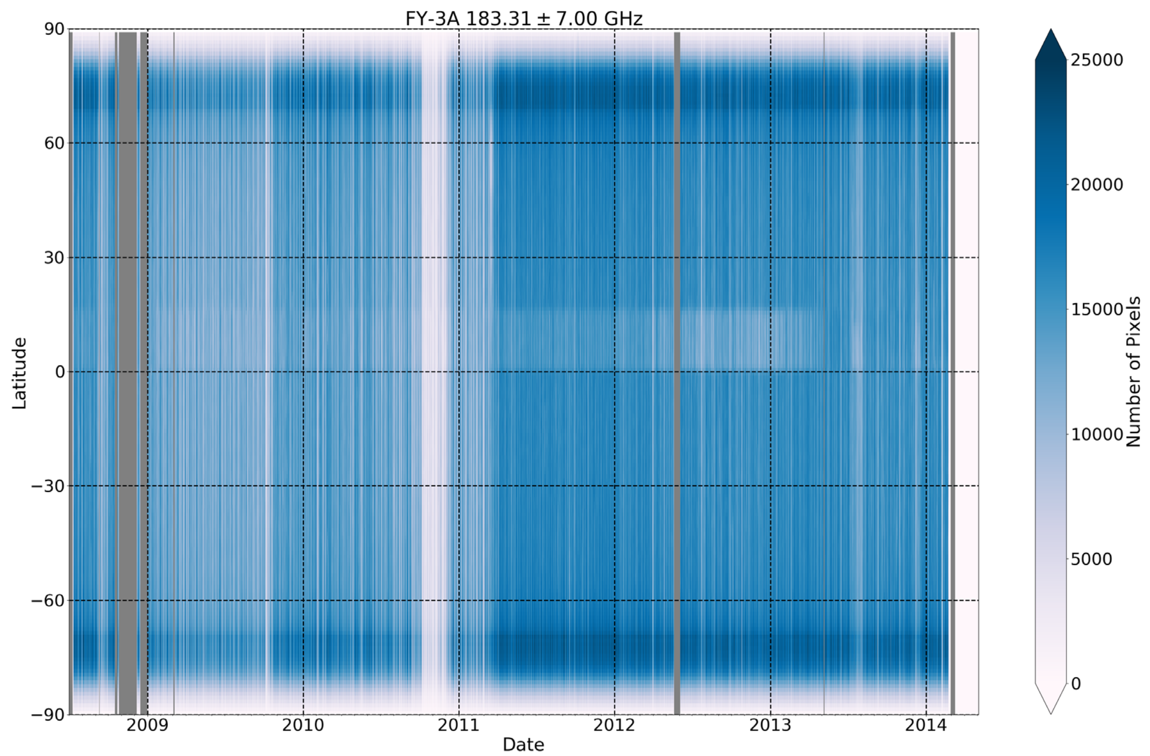

Figure 13: Hovmöller diagram showing the change in the number of daily data by latitude with time for FY-3A for the 183.31±7 GHz channel.

4.1.4.3 Key points

- The FY-3A FCDR is cooler than the operational data set
- The 183.31±7 GHz channel shows the best agreement between the datasets (approx. -0.5 K), followed by the 183.31±3 GHz (approx. -1.0 K) and finally the 183.31±1 GHz (approx. -1.0 to -1.5 K)
- Standard deviations of differences range between <0.1 K (183.31±1 GHz) and ~1.7 K (183.31±7 GHz)
- All channels show significant discontinuities in May and September 2013
- There is a large drop in the number of available data around the end of 2010/early 2011
- An annual cycle in the FCDR minus operational differences is evident in the tropics and at high latitudes, which is likely to be due to a BT-dependent difference between the two data sets that is apparent because of the cloud-related BT variability that results from the ITCZ and mid-to-high latitude storms.
- Linear features are present in the Hovmöller diagrams at high latitudes that appear correspond to the change in the number of data due to the higher number of overpasses at these latitudes.
- Linear features are visible in the standard deviations for all channels at ~0° and ~17°N. These linear features are attributed to the calibration process that is applied to the data in the CMA orbit files, which start at 17°N and end at 0.5°N and are not present in the FCDR data.

4.1.5 MWHS/1 on board of FY-3B

4.1.5.1 Time series analysis

Figure 146 shows the time series of mean differences for FY-3B. The agreement for FY-3B is substantially better than for FY-3A, with mean differences of ~0.1 K for the 183.31±1 GHz and 183.31±3 GHz, and approximately -0.2 K for the 183.31±7 GHz channel. The time series are also much less noisy. The better agreement for FY-3B compared with FY-3A is likely to be due to the different calibration approaches used for each instrument. With the exception of FY-3A and S-NPP, a non-linear correction is applied to the other sensors that comprise the suite of FCDRs at the radiance level, which follows the approach adopted in the NWP SAF AAPP (ATOVS and AVHRR Pre-processing Package; https://www.nwpsaf.eu/site/software/aapp/). For FY-3A and S-NPP, this non-linear correction is applied in BT space, which may explain the different results obtained for these two sensors (0). The mean differences for the 183.31±1 GHz and 183.31±3 GHz channels are generally stable with time, apart from a very small (<0.1 K) discontinuity in early 2014 that is present in all channels. All channels also show anomalies at the end of 2016. In the 183.31±1 GHz and 183.31±3 GHz channels this is characterised by missing data, but in the 183.31±7 GHz channel, there is a negative step in the time series of almost 0.1 K. There is also a small discontinuity (<0.1 K) about half-way through 2012 present in the 183.31±7 GHz channel following a very small positive trend in the differences since April 2011, when FY-3B became operational, that is not visible in the other 183 GHz channels. The reason for this discontinuity is currently unknown, but may, for example, be due to a sensor drift correction applied to this channel by CMA in their operational data set that has not been applied to the C3S microwave humidity FCDR. Examination of the single data set time series does not yield any insight into whether the drift is present in one or both data sets (Appendix A).

The better agreement between the FCDR and operational data for FY-3B compared with FY-3A is also apparent in the time series of the standard deviation of the differences, which is typically <0.05 for the 183.31±1 GHz and 183.31±3 GHz channels, and ~0.1 K for the 183.31±7 GHz channel (not shown). As observed for the mean differences, the time series of standard deviations are generally stable. The time series for the number of available data are also very stable, with no notable changes over time (not shown).

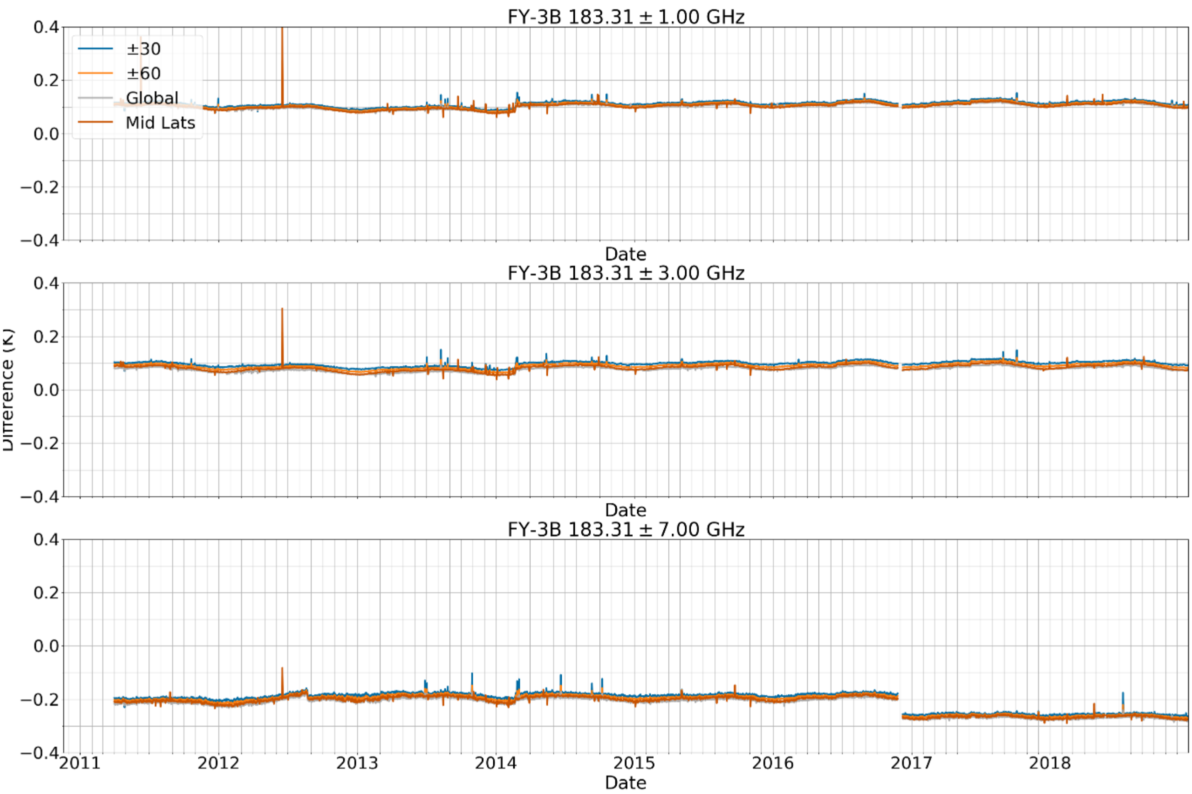

Figure 14: Time series of mean differences for FY-3B (FCDR minus operational) for each of the 183 GHz channels.

4.1.5.2 Hovmöller diagrams

Figure 157 is a Hovmöller diagram showing the change in daily mean difference by latitude with time for the FY-3B 183.31±1 GHz channel. The discontinuity in early 2014 is clearly visible in the Hovmöller diagrams for all channels (183.31±3 and 183.31±7 GHz not shown). As seen for FY-3A, there is an annual cycle evident in the tropics and at mid-to-high latitudes in all channels. As noted previously, these annual cycles probably occur because there is a BT-dependent difference between the two data sets that is apparent because of the colder BTs that occur when cloud is present. However, the amplitude of these cycles is much smaller here than for FY-3A, and do not exceed 0.1 K (<0.1 K for FY-3B vs ~0.5 – 1.0 K for FY-3A) and so are negligible and not considered further here.

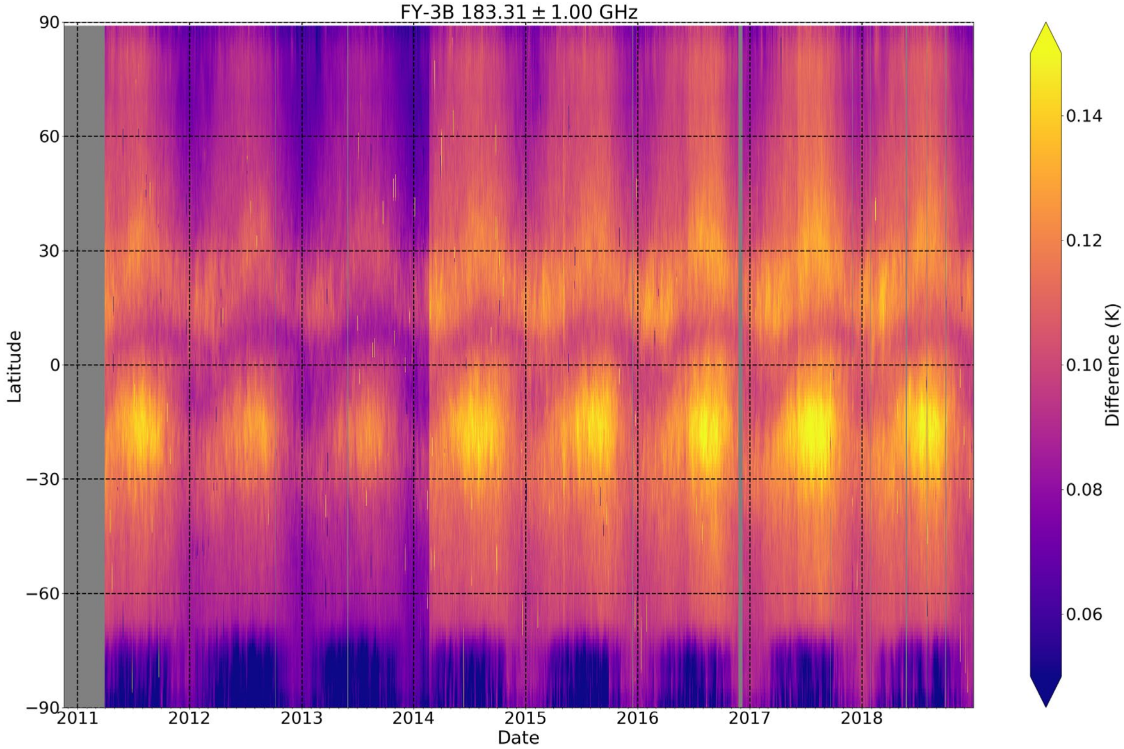

Figure 15: Hovmöller diagram showing the change in daily mean difference by latitude with time for FY-3B for the 183.31±1 GHz channel. The grey bands correspond to data missing from the analysis.

Figure 168 shows the variation in the standard deviation of the daily mean differences as a function of latitude and time for the 183.31±7 GHz channel (183.31±3 and 183.31±3 GHz not shown). The discontinuity present in early 2014 is clearly visible. As for the mean differences, there is an annual cycle at mid-to-high latitudes in all channels. However, the amplitudes of these annual cycles are again <0.1 K thus are not considered further here. The linear features observed for FY-3A at ~0° and ~17°N are also present for FY-3B (Section 4.1.4).

4.1.5.3 Key points

- The agreement for FY-3B is substantially better than for FY-3A, with mean differences of ~0.1 K for the 183.31±1 183.31±7 GHz channel GHz and 183.31±3 GHz, and approximately -0.2 K for the 183.31±7 GHz channel. The standard deviations are typically <0.1 K.
- There is a discontinuity in the BT differences (FCDR minus operational) at the end of 2016, which induces a negative step of ~0.1 K in the 183.31±7 GHz channel. . Another discontinuity appears in early 2014 affecting all channels, although of smaller magnitude ~0.05 K.
- As for FY-3A, there are linear features visible in the standard deviations for all channels at ~0° and ~17°N. These linear features are attributed to the calibration process that is applied to the data in the CMA orbit files, which start at 17°N and end at 0.5°N and are not present in the FCDR data.

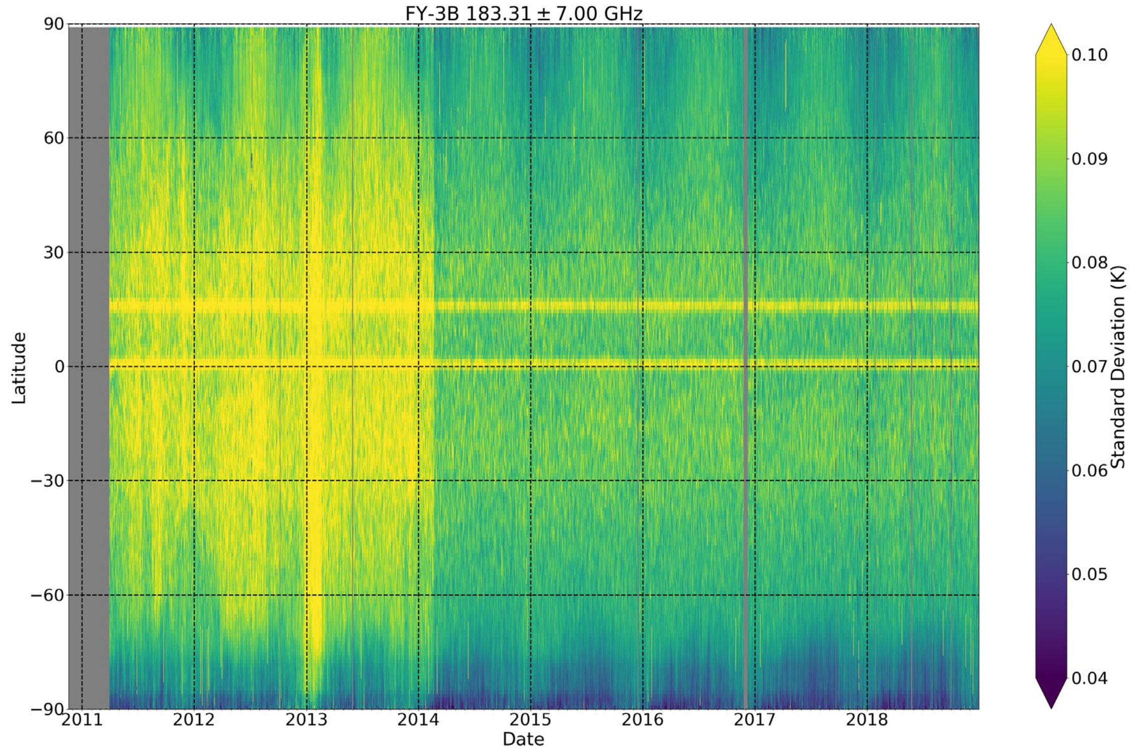

Figure 16: Hovmöller diagram showing the change in the standard deviation of the daily mean differences by latitude with time for FY-3B for the 183.31±7 GHz channel. The grey bands correspond to data missing from the analysis.

4.1.6 MWHS/2 on board of FY-3C

4.1.6.1 Time series analysis

Figure 179 shows the time series of C3S microwave humidity FCDR and operational data from FY-3C. In general, the agreement between the two data sets is good and stable, apart from discontinuity in the agreement for the 183.31±1 GHz channel in early 2015. This can be seen clearly in the time series of differences (Figure 1810), which shows a positive step of about 0.5 K for this event. The time series of differences also shows a smaller discontinuity in mid-2014, which is characterised by an increase in the noise between this discontinuity and the early 2015 event, more obvious in channel 183.31±7 GHz. Mean BT differences are generally very close to zero for the 183.31±1 GHz, 183.31±1.8 GHz and 183.31±3 GHz channels. For the 183.31±4.5 GHz and 183.31±7 GHz the mean differences are approximately -0.3 and -0.7, respectively.

The daily standard deviations are typically very close to zero for all channels although there are many spikes, particularly before 2016 (Appendix A). The spike density is particularly high around the time of the two discontinuities in mid-2014 and early 2015.

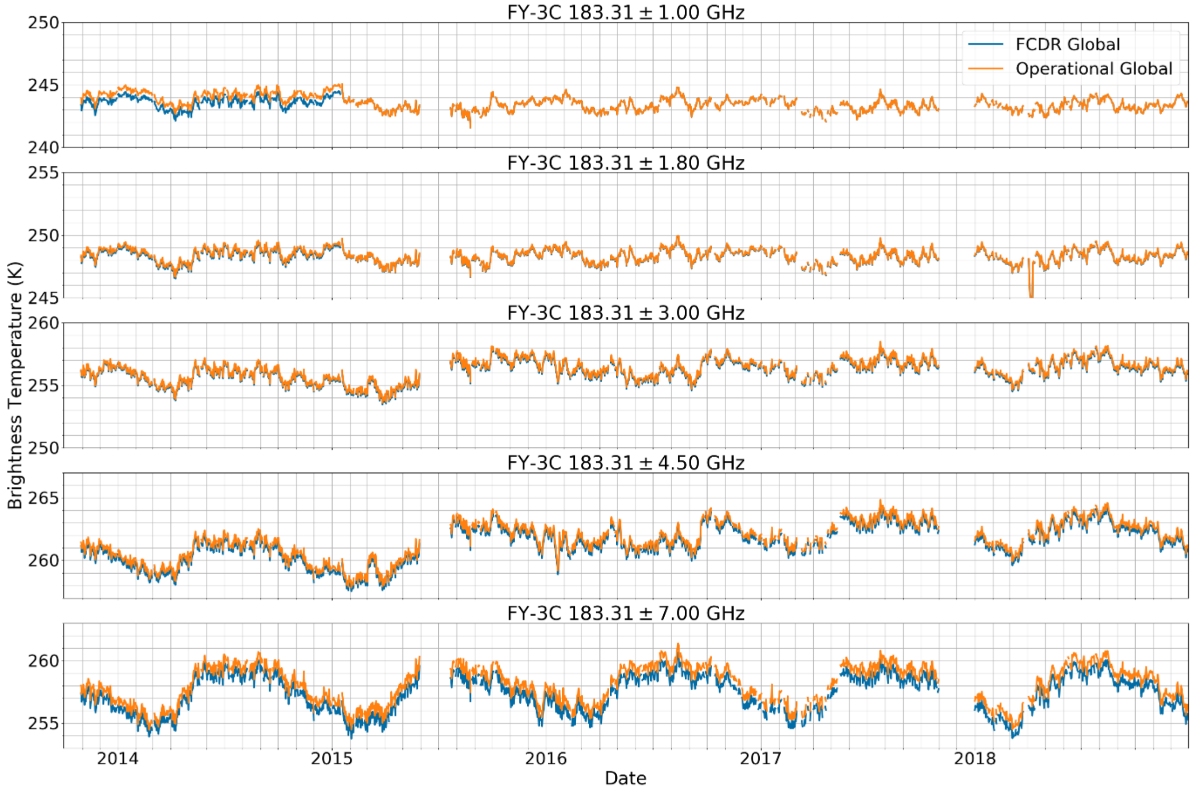

Figure 17: Time series of FCDR (blue) and operational data for FY-3C

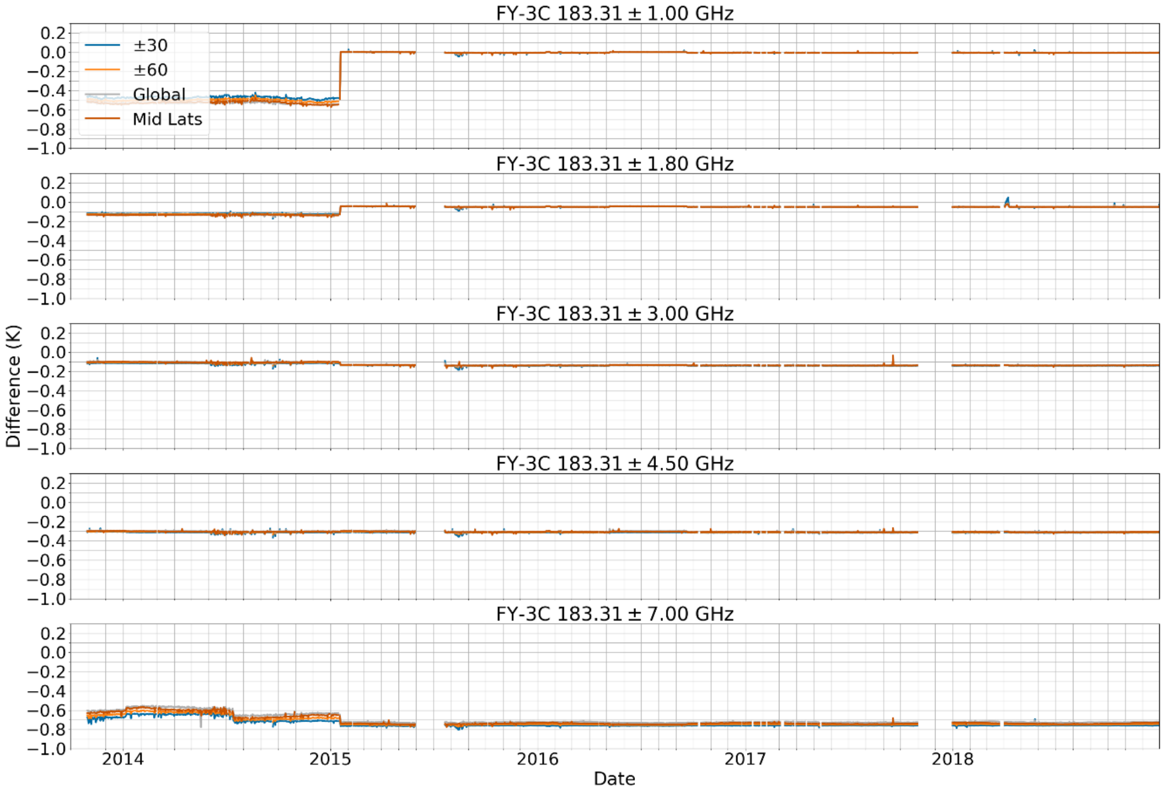

Figure 18: Time series of BT differences (FCDR minus operational) for FY-3C.

4.1.6.2 Hovmöller diagrams

Figure 1911 shows the Hovmöller diagram for the mean differences for the 183.31±7 GHz channel. The mid-2014 and early 2015 discontinuities are clearly visible in the data. Annual cycles in the tropics and mid-to-high latitudes are also visible and are similar to those seen previously for FY-3A and FY-3B. As for FY-3B, the amplitudes of these annual cycles are <0.1 K and are thus considered negligible. Further discontinuities are also apparent in mid-2015, early 2016, and late 2016 in the 183.31±1 GHz channel Hovmöller diagram. However, it should be noted that these further discontinuities are of the order of a few hundredths of a K at most so are also considered negligible. The Hovmöller diagrams for the standard deviation of mean differences show no further remarkable features other than those already noted previously. As for FY-3A and FY-3B, these data show annual cycles at high latitudes and in the tropics, and the two linear bands in all channels at ~0° and ~17°N (Appendix A; only shown for the 183.31±7 GHz channel).

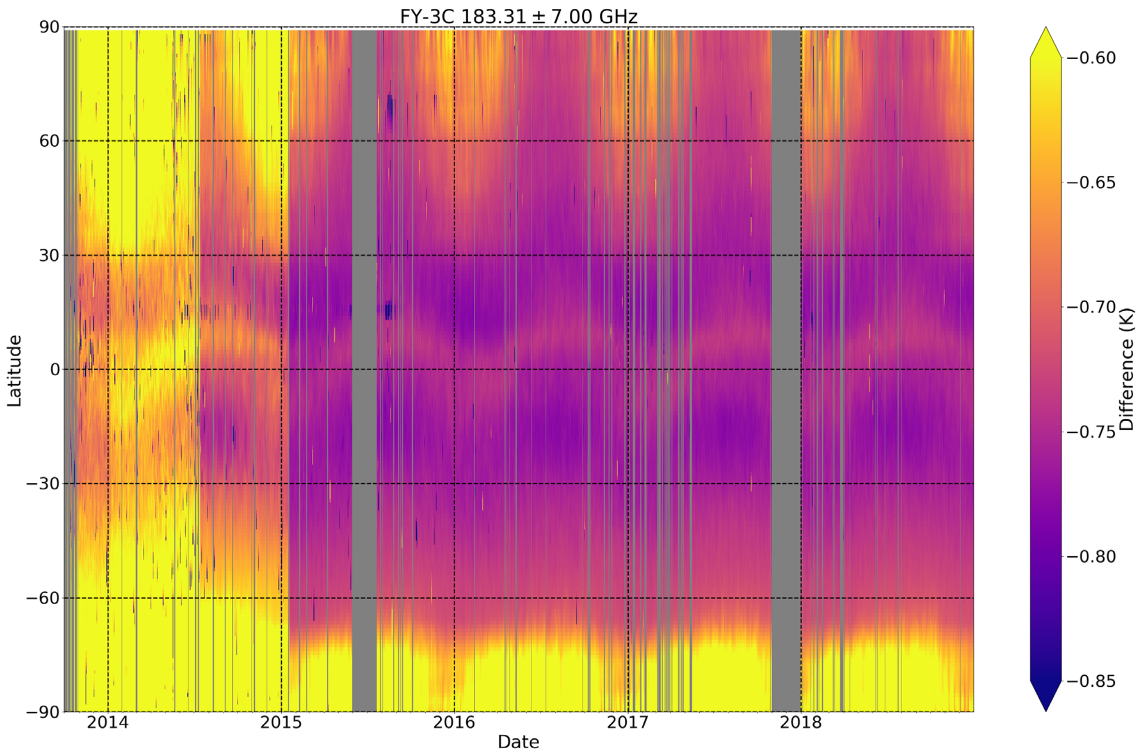

Figure 19: Hovmöller diagram showing the change in daily mean difference by latitude with time for FY-3C for the 183.31±1 GHz channel. The grey bands correspond to data missing from the analysis.

4.1.6.3 Key points

- Mean BT differences are generally very close to zero for the 183.31±1 GHz, 183.31±1.8 GHz and 183.31±3 GHz channels. For the 183.31±4.5 GHz and 183.31±7 GHz the mean differences are approximately -0.3 and -0.7 K, respectively.
- Standard deviations are typically very close to zero for all channels
- There are discontinuities in the times series in mid-2014 and early 2015. The 2015 event is strongest in the 183.31±1 GHz, where the FCDR minus operational difference exhibits a positive step of ~0.5 K.
- In general, the agreement between the FCDR and operational data is more stable and less noisy from 2016.

- As for FY-3A and FY-3B, there are linear features in the standard deviations at ~0° and ~17°N.

4.2 Comparison between FCDRs and simulated TOA BTs

4.2.1 Global Difference Maps

Figure 20 shows maps of the global differences between the S-NPP FCDR and equivalent simulated BTs for the 183.31±7 GHz channel on 26 May 2012 for all the available data. The FCDR BTs are generally colder than the simulated BTs, although there are regions, mostly over the land, where they are warmer. As noted in Section 3.1.5, only the absorption from liquid water clouds has been included in the simulations. Thus, there are areas where clouds are found in the observations but are missing from the simulations, which are generally characterized by swathes or patches of more negative differences (FCDR colder indicated by dark blue colours). Large positive and negative differences are also seen at high latitudes, more obvious over Antarctica, mainly due inadequate surface emissivity values used in the simulations (emissivities are based on a climatology). The same is true over other landmasses, although the differences there are smaller in magnitude.

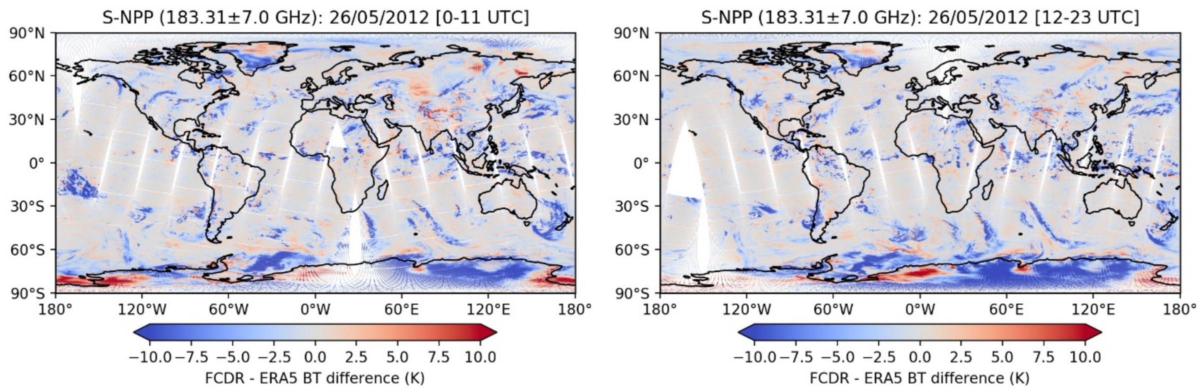

Figure 20: Global maps showing the FCDR ATMS/S-NPP minus equivalent simulated BTs for 183.31±7.0 GHz channel differences for 00:00-11:00 UTC (left) and 12:00-23:00 UTC (right) on 26 May 2012.

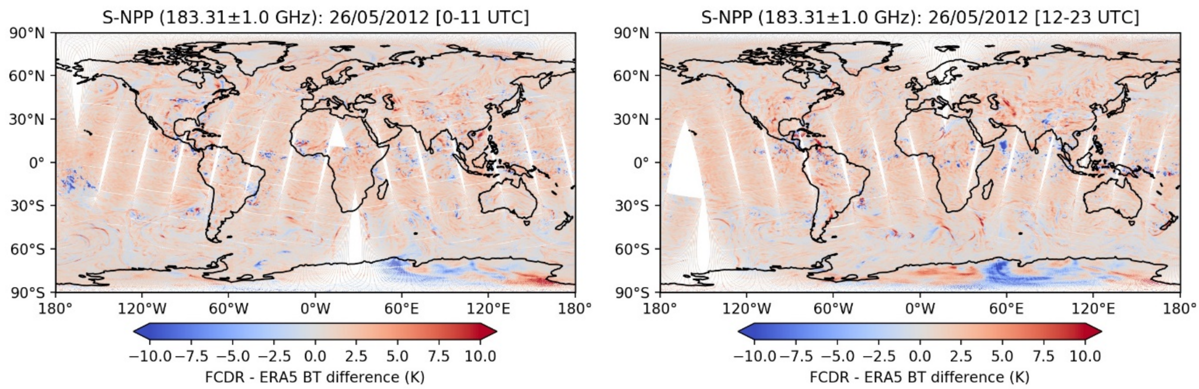

Figure 21: Global maps showing the S-NPP FCDR minus equivalent simulated BTs for 183.31±1.0 GHz channel differences for 00:00-11:00 UTC (left) and 12:00-23:00 UTC (right) on 26 May 2012.

Figure 21 shows the same maps but for the 183.31±1.0 GHz channel. For this channel, the differences are mostly slightly positive, although there are some localised regions where the differences are negative. These coincide with the large negative differences seen in Figure 20. As for the 183.31±7.0 GHz channel, the largest differences are seen at high latitudes. The global difference maps for the other S-NPP channels are similar, and generally suggest that the differences become more positive, with a reduction in the number and intensity of cloud-related features, as the channel sounding height increases (not shown).

Figure 22 and Figure 23 show the same results presented in Figure 20 and Figure 21, but only for oceans and large lakes, and restricted to between ±60° latitude. The observations have also been cloud-cleared using the Buehler (0) tests (see Section 3.1.5), which removes many of the negative-difference structures noted in the unmasked data shown above. However, it does not remove all the negative patches, which is particularly apparent for the low peaking channels. Based on these results, only data over oceans and large lakes between ±60° latitude that have been cloud-cleared using the Buehler tests are considered for the rest of this evaluation strand (as stated in Section 3.2.2). However, the results shown below suggest that there may be some cloud-related artefacts in the data that should be considered when interpreting the results.

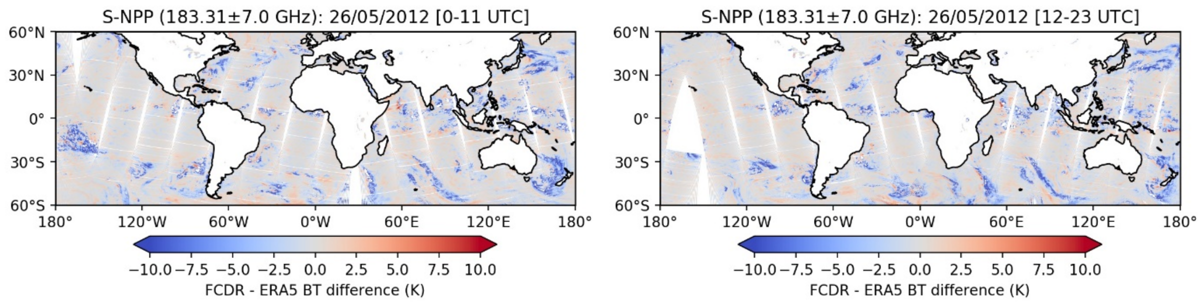

Figure 22: Global maps showing the S-NPP FCDR minus equivalent simulated BTs for 183.31±7.0 GHz channel differences for 00:00-11:00 UTC (left) and 12:00-23:00 UTC (right) on 26 May 2012. The data shown are cloud-cleared (observations only), over oceans and lakes only, and restricted to ±60° latitude.

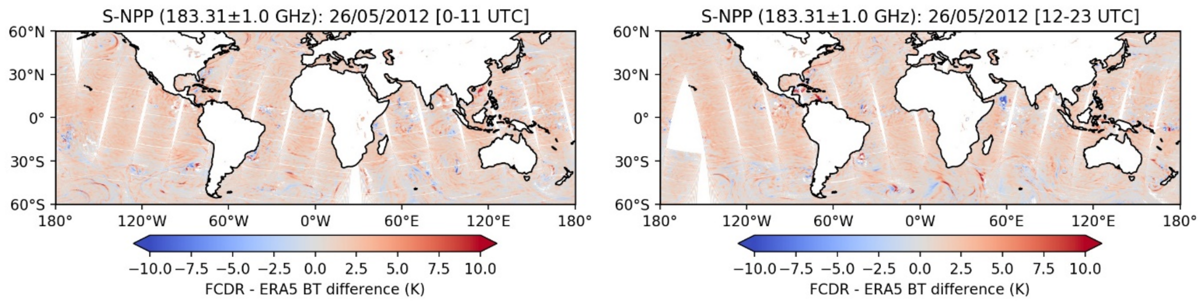

Figure 23: Global maps showing the S-NPP FCDR minus equivalent simulated BTs for 183.31±1.0 GHz channel differences for 00:00-11:00 UTC (left) and 12:00-23:00 UTC (right) on 26 May 2012. The data shown are cloud-cleared (observations only), over oceans and lakes only, and restricted to ±60° latitude.

4.2.2 Time Series Analysis

4.2.2.1 Time series of FCDR minus simulated BTs

Figure 24 shows the global mean time series of cloud-cleared C3S microwave humidity FCDR BTs minus simulated BTs for the 183.31±1 GHz channels for all instruments over oceans and large lakes between ±60° latitude. The time series for Metop A and B, and S-NPP are all stable with time. Notably, the discontinuity in the Metop A data near the beginning of the time series in 2007 seen in the comparison with operational BTs (Section 0) is not present, which is confirmation that this anomaly exists only in the operational BTs. Collectively these three sensors are ~1 K warmer than the simulations. There is almost no distinction between the time series for Metop A and B, while S-NPP is ~0.3 K warmer. By contrast, the time series for the FY-3 sensors are around 1-2.3 K colder than the simulations. The time series for FY-3A is the least stable and shows a pronounced negative trend with time. However, the discontinuities identified earlier in this study in May and September 2013 (Section 4.1.4) are not clearly visible. There is also a weak negative trend in the FY-3B and -3C data. The time series for FY-3C is very variable and the annual cycle, which is expected, is barely visible. The early 2015 discontinuity in the FY-3C FCDR/operational data analysis is also visible in Figure 24 (Section 4.1.6). However, the mid-2014 discontinuity also identified in this analysis is not obvious in the comparison between the FCDR and the simulations. However, there is a ~0.2 K step change obvious in mid-2016.

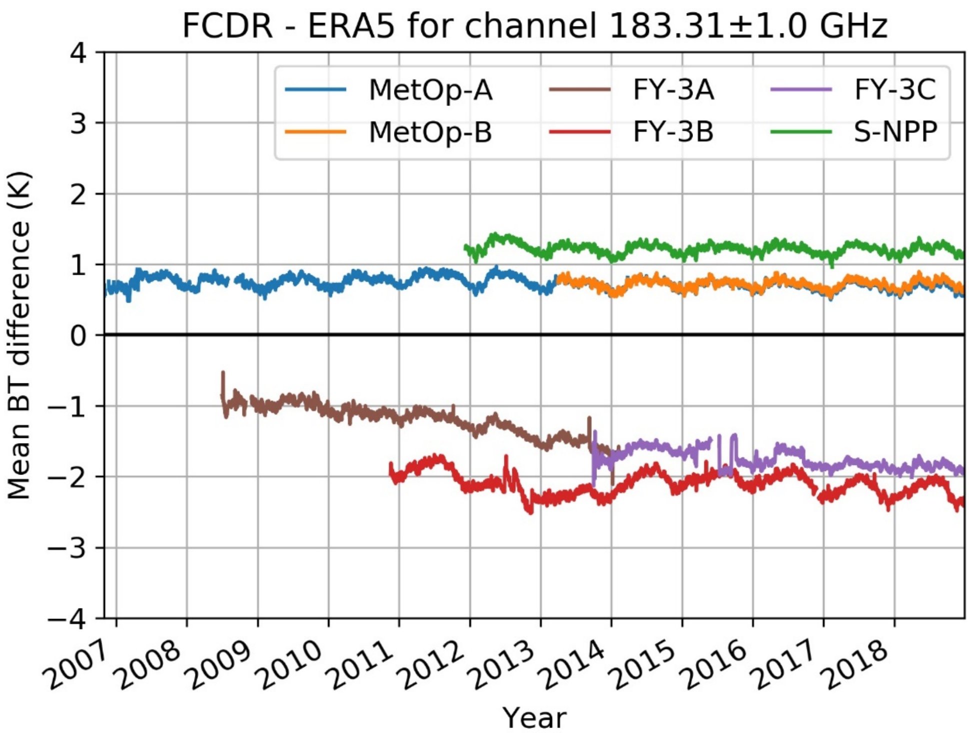

Figure 24: Time series of FCDR BTs minus simulated BTs for all sensors for the 183.31±1.0 channel.

The results for the other channels are generally quite similar (Appendix B). For the 183.31±1.8 GHz (S-NPP and FY-3C only, not shown), the S-NPP observations are ~0.5 K warm, while the FY-3C time series is very similar to that for the 183.31±1 GHz channel but is slightly more stable. The general pattern for the 183.31±3 GHz channel is broadly the same, although Metop A and B are now ~0.3 K cooler than the simulations, while S-NPP is ~0.5 K warmer. The FY-3A and -B data are very similar to the other high-peaking channels, although the shape of the FY-3A time series is now more parabolic, rather than linear with a negative slope. The FY-3C time series differs the most from the higher-sounding channels, and is mostly positive compared with the simulations, but highly variable, ranging between -0.5 and +2 K, with some strong discontinuities (Appendix B). For the 183.31±4.5 GHz channel, the S-NPP time series is slightly warmer at ~0.7 K, while the FY-3C data demonstrate an even higher variability compared with the 183.31±3 GHz channel, which ranges between 0.5 and ~4 K.

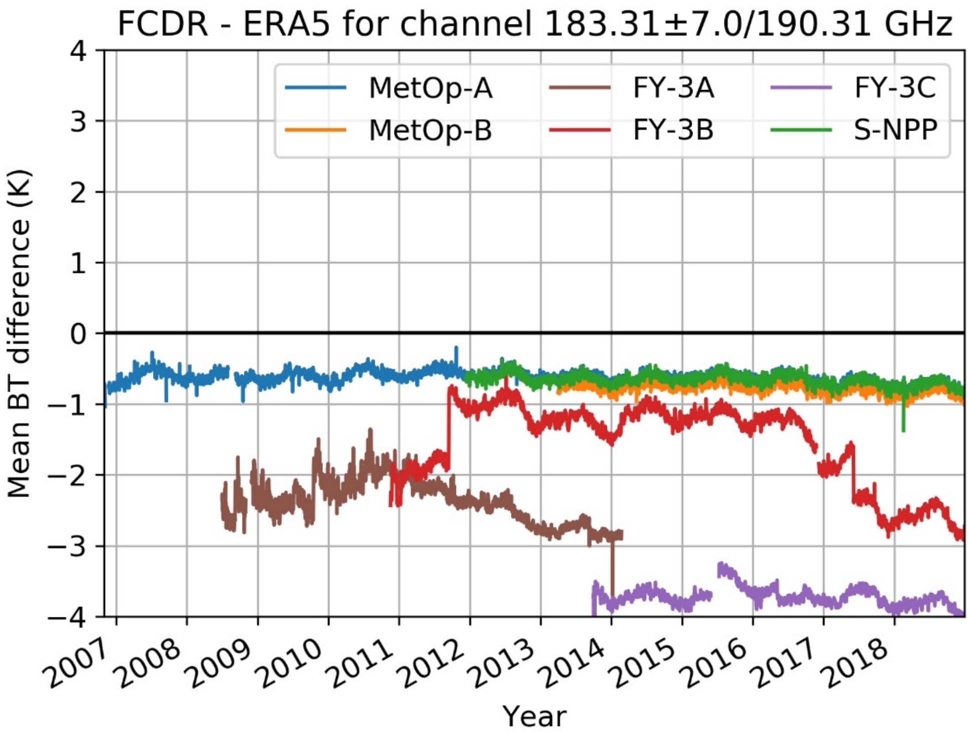

Figure 25: Time series of FCDR BTs minus simulated BTs for all sensors for the 183.31±7.0 channel.

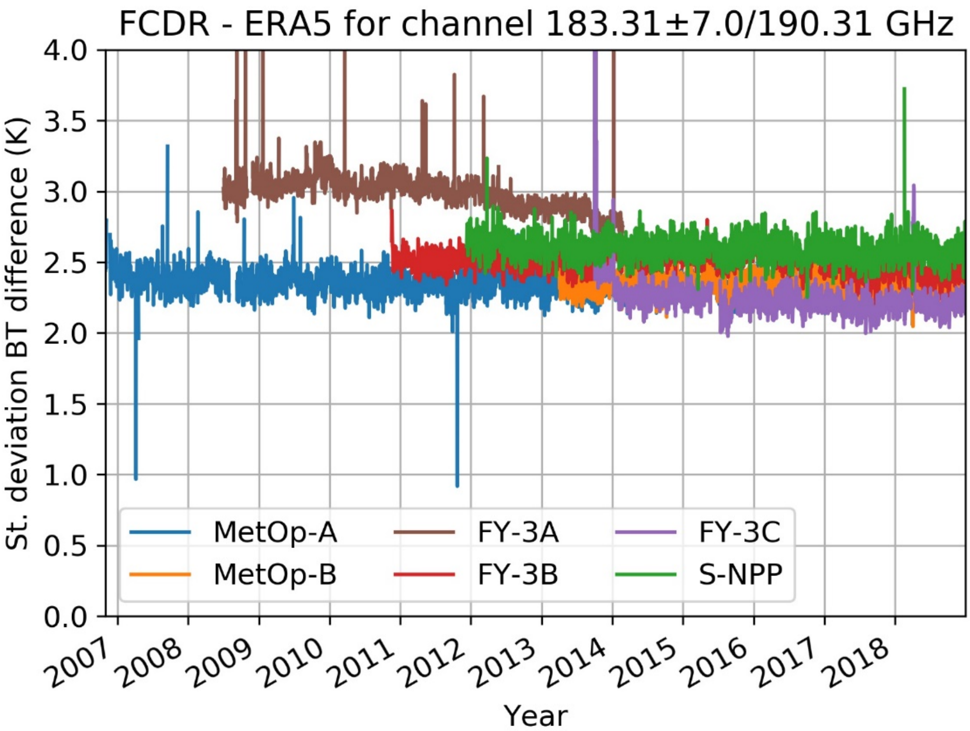

Figure 26: Time series of the standard deviation of the FCDR BTs minus simulated BTs for all sensors for the 183.31±7.0 channel.

For the lowest sounding channels at 183.31±7 GHz, or 190.31 GHz for Metop A and B, all the time series are colder compared to the simulations (Figure 25). S-NPP, Metop A and B are now all very similar at around -0.8 K. This is also confirmed in the SNO analysis described in Section 4.3.15. The FY-3C time series is quite stable at around -3.7 K compared with the simulations, while FY-3A and -B are more variable with mean differences that range between -0.9 K and -3.0 K. There is a significant discontinuity in the FY-3B data in the second half of 2011 where there is a positive step of ~1 K, and a later one in the first half of 2017 where there is a negative step of ~0.5 K. Neither of these discontinuities are apparent in the FCDR/operational data inter-comparison (Section 4.1.5), which suggests that these discontinuities must be present in both data sets. By contrast, the discontinuity identified in this earlier analysis at the end of 2016 are not obvious in Figure 25, although there are missing data that correspond to this anomaly.

The time series of the standard deviations of the daily differences between the FCDRs and equivalent simulated BTs have also been analysed. The standard deviations for the FY-3 FCDRs are higher than those for either Metop A, B or S-NPP for all channels except the 183.31±7 GHz / 190.31 GHz channels (Figure 26). The time series for all instruments is reasonably stable for all channels and typically range between 1.3 and 2.5 K for all channels (not shown) except the 183.31±7 GHz / 190.31 GHz channels, where they are higher and noisier. Except for FY-3A, the percentiles for each sensor indicate that the distribution of BT differences is negatively skewed for the 183.31±7.0 GHz / 190.31 GHz channels, and that the lower two percentiles (1% and 2.5%) are considerably noisier than the others. An example for FY-3B is shown in Figure 27. This negative skew and higher variance are a result of the negative BT differences that occur when clouds that scatter radiation at 183 GHz are present in the observed BTs but are not in the simulations, where only liquid clouds are included (Section 3.1.5). This indicates that the screening applied to the observations in this study to remove these clouds is not completely effective in all channels (Section 3.2.2). The negative skew and high variance in the lower percentiles clearly decreases as the channel sounding height increases (not shown), such that for the 183.31±1.0 GHz, the distribution is quite symmetrical and the variance in the upper two percentiles (97.5 and 99 %) is similar to the lowest percentiles (Figure 28). This is not surprising given that the cloud mask applied here has been derived for the Upper Tropospheric Humidity (UTH) retrievals using only the 183.31±1.0 GHz channel. Some non-zero and varying agreement between the FCDRs and the simulations is expected as differences in the atmospheric states between the ERA5 reanalysis and the true atmosphere observed by the satellites will also influence the agreement between the two data sets. In order to minimise this dependence on the simulations, the next section investigates the double differences between sensors.

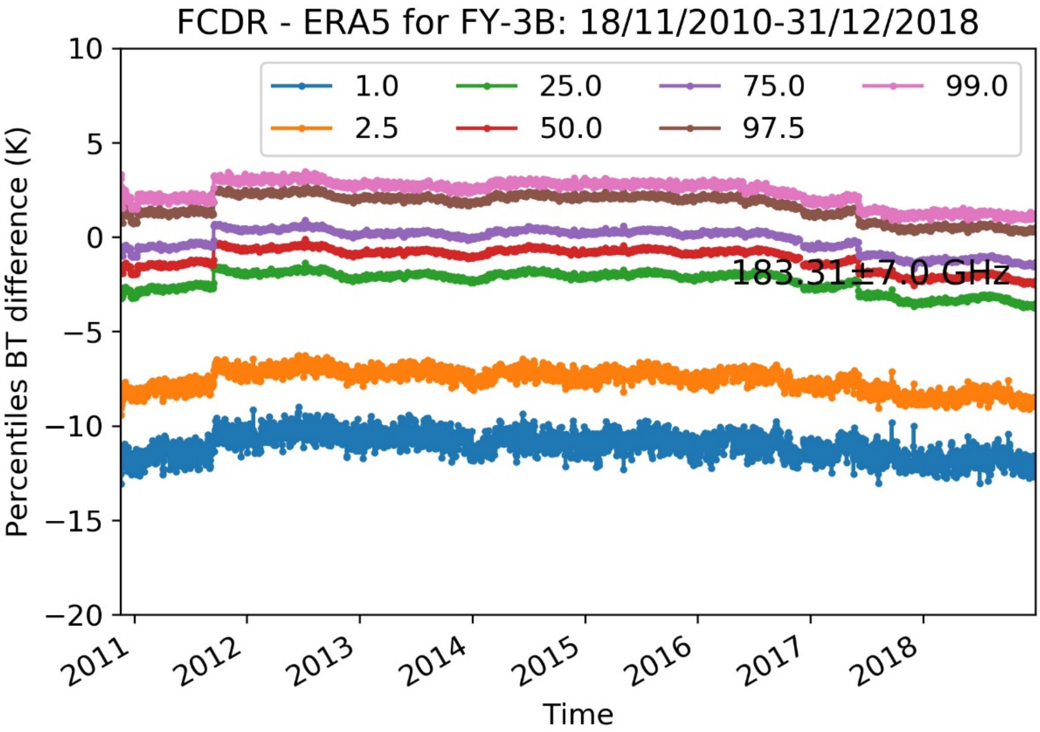

Figure 27: Time series of percentiles for the 183.31±7 GHz channel for FY-3B. The legend indicating which line corresponds to each percentile is provided in the legend at the top of the figure.

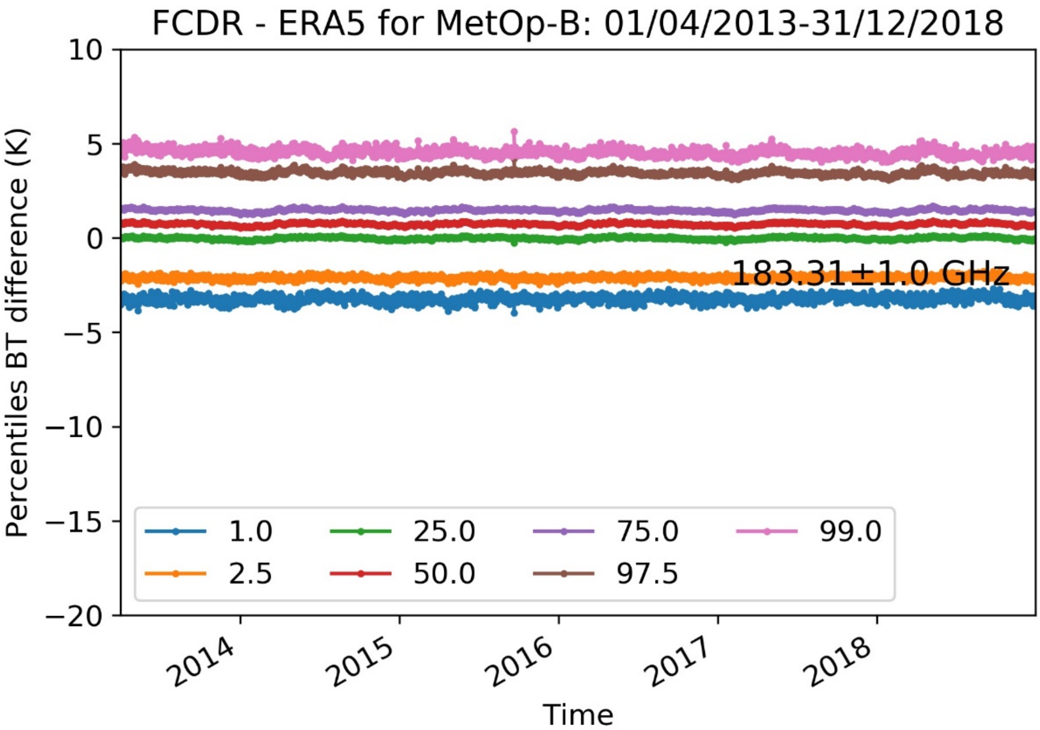

Figure 28: Time series of percentiles for the 183.31±1 GHz channel for Metop B. The legend at the top of the figure indicates which line corresponds to each percentile.

4.2.2.2 Time series of double differences

Figure 29 shows the time series of the double difference of the daily mean differences for the 183.31±1 GHz channel for each sensor referenced to S-NPP (see Section 1.7 for the explanation of the double difference method). For Metop A and B, the annual cycle has been removed, and the data show that the time series are very stable with a constant difference of ~0.4 K from S-NPP. This demonstrates the excellent agreement between Metop A and B, and that the difference with S-NPP is simply an offset. For the FY-3 sensors, most of the annual cycle has been removed and the slight negative trends in the FY-3A and FY-3C data are still evident (compare with Figure 24, which shows the single-sensor time series). FY-3C has some significant discontinuities in 2013 and 2015. There is still significant variability in the FY-3B time series, and discontinuities are apparent in the second half of 2012. None of these discontinuities are apparent in the comparisons between the FCDRs and operational data reported in Sections 4.1.5 and 4.1.6.

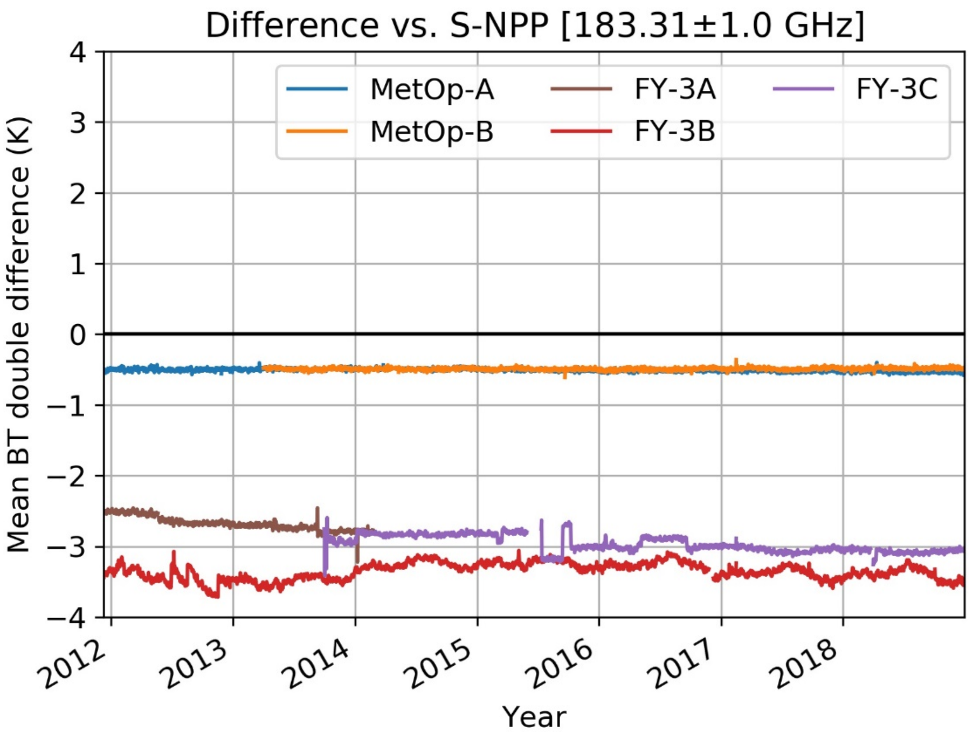

Figure 29: Time series of double differences for each sensor referenced to S-NPP for the 183.31±1 GHz channel

The FY-3C time series also shows a fairly stable, but slightly negative trend in the 183.31±1.8 GHz channel referenced to S-NPP (Appendix B). However, the time series for this instrument is highly variable with many discontinuities in the 183.31±3 GHz (Figure 30). Notably discontinuities occur in mid-2015, for a six-month period in the middle of 2016, and in early 2018. None of these discontinuities are apparent in the comparisons with the operational FY-3C data (Section 4.1.6), which suggests that these discontinuities are present in both the FCDR and the operational data for this instrument. Significant discontinuities are also visible in the FY-3A and FY-3B time series in mid-2012. None of these discontinuities are apparent in the comparisons between the FCDRs and operational data reported in Sections Section 4.1.4 and 4.1.5, again suggesting the discontinuities could be present in both the FCDR and the operational data.

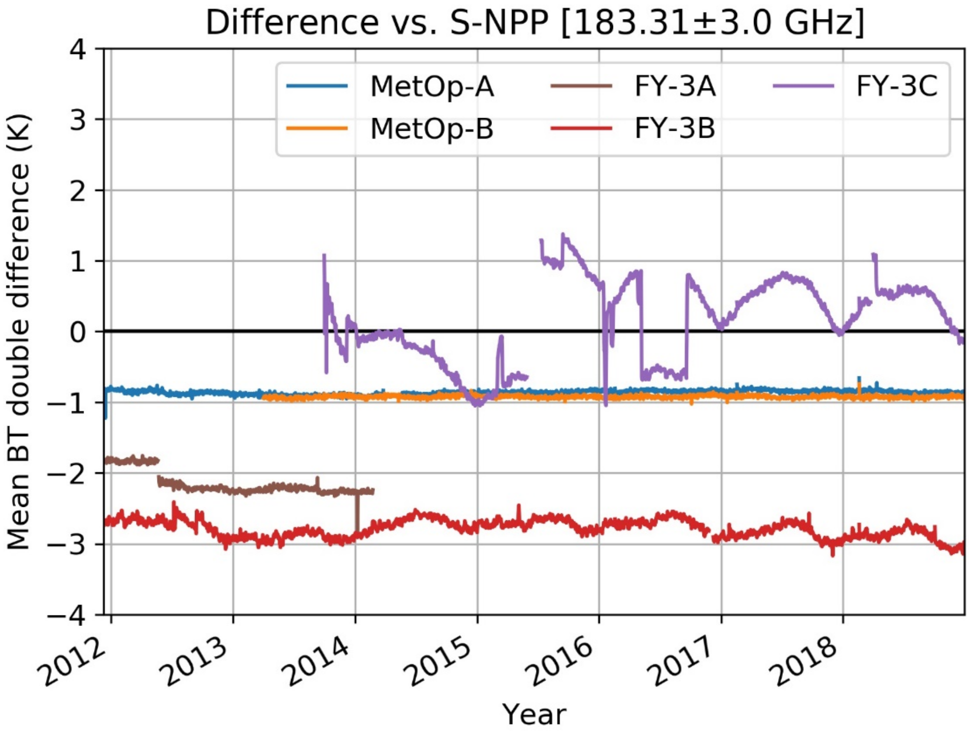

Figure 30: Time series of double differences for each sensor referenced to S-NPP for the 183.31±3 GHz channel

For the 183.31±4.5 GHz channel, the FY-3C time series is erratic and warm compared with the S-NPP data varying between 0 K and 4 K (Appendix B); the variability is similar to that observed for the 183.31±3.0 GHz channel (Figure 30) and for the single sensor time series analysis (Appendix B). However, the FY-3C double-differenced time series is much more stable for the 183.31±7 GHz, with only one significant discontinuity occurring in mid-2015. This discontinuity is not present in the comparison between the FCDR and the operational FY-3C data, suggesting it may be present in both data sets. The double-differenced time series for the FY-3A has a negative trend varying between about -1.8 K and -2.1 K but is otherwise quite stable. The FY-3B time series is the most unstable for this channel and varies by ~0.5 K during the first 5 years, after which the negative step in mid-2017 that was seen in the previous section is present. The Metop A and B time series for this channel are very close to zero difference from S-NPP with no significant discontinuities. It should be noted, however, that the frequency for this channel is different between the two sensors, i,e. 190.31 GHz vs. 183.31±7 GHz. There is also a small offset between Metop A and B of ~0.1 K that is more obvious in the double-differenced time series than in the single-sensor time series for this channel presented in Figure 25.

The double-differenced time series referenced to Metop B give broadly the same results (Appendix B). However, they highlight additional discontinuities for FY-3A at the very beginning of the time series in 2009 and towards the end of 2009 in the 183.31±1 and 183.31±3 GHz channels. Additional discontinuities are also visible in the FY-3B data at the beginning of the time series at the end of 2010, and a major positive step towards the end of 2011 in the 183.31±7 GHz channel (Appendix B).

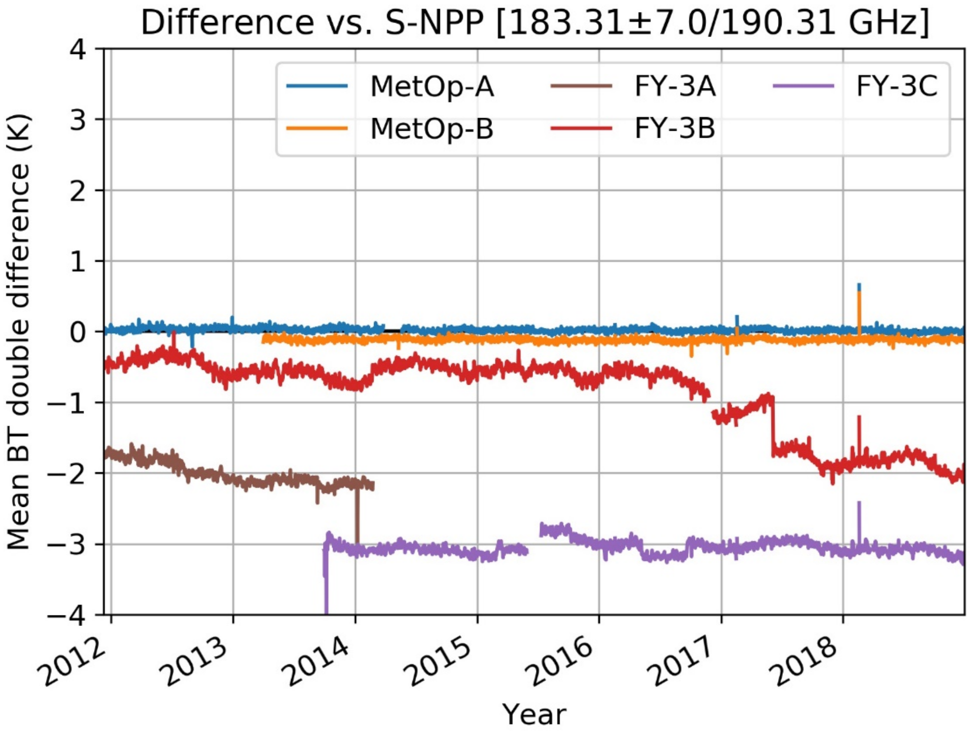

Figure 31: Time series of double differences for each sensor referenced to S-NPP for the 183.31±7 GHz channel

4.2.2.3 Key points

- The time series for Metop A and B, and S-NPP are all very stable with time and are generally in close agreement. This suggests that the differences observed between the S-NPP FCDR and operational data (Section 0) may be due to the operational data.
- There is virtually no distinction between the Metop A and B time series, except for in the 183.31±7 GHz channel where there is a small offset (~0.1 K).
- The time series for FY-3A, -B and -C are often highly variable. The time series for FY-3C in particular is highly erratic for the 183.31±3 GHz and 183.31±4.5 GHz channels.
- There is a negative trend for FY-3A in the 183.31±1 GHz channel, the 183.31±3 GHz channel and from 2011 in the 183.31±7 GHz channel.
- All the FCDR data for the 183.31±7 GHz (or 190.31 GHz for Metop A and B) are cold compared to the simulations.
- The standard deviations for the FY-3 FCDRs are higher than those for either Metop A, B or S-NPP for all channels except the 183.31±7 GHz / 190.31 GHz channels
- For all sensors, the standard deviations are highest for the 183.31±7 GHz channel. In general, the distributions of FCDR minus simulations are negatively skewed for the 183.31±7.0 GHz / 190.31 GHz channels and the lower percentiles of the distributions are noisy. These features are attributed to cloud contamination in the observations.
- The discontinuities identified previously in FY-3A data in May and September 2013 are not visible. New discontinuities are identified here in 2009 and mid-2012.
- For FY-3B, the discontinuity identified in the FCDR vs operational data analysis at the end of 2016 is not obvious. New discontinuities, not seen previously, are identified

at the end of 2010, in the second half of 2011, the second half of 2012, and in the first half of 2017.
- For FY-3C, the discontinuities identified previously in this study in 2013 and 2015 in are visible, but the mid-2014 discontinuity is not. A new discontinuity, not identified previously, is also apparent for ~6 months in the middle of 2016, and in early 2018.

4.2.3 Hovmöller diagrams

Figure 32 shows the variation in the FCDR minus simulations daily mean difference as a function of latitude and time for FY-3B 183.31±7.0 GHz channel. The results for this channel indicate a general difference of approximately -2 K, as noted already in the time series analysis, but there is variation in the magnitude with latitude that also shows an annual cycle. There is a strong annual cycle in the northern-hemisphere tropics, with the negative differences peaking at higher latitudes during the NH summer. There are also annual cycles in the mid latitudes. For example, the annual cycle between ~35°N and 60°N shows the more negative differences during the NH winter moving towards higher latitudes during the NH summer months. These annual cycles are quite similar to those apparent in the Hovmöller diagrams presented in the analysis of the differences between the FCDR and equivalent operational data sets. However, they have much larger amplitude here. As noted in this earlier analysis, the annual cycle may be a result of a BT-dependent difference between the two data sets that is apparent because of the colder BTs that occur when cloud is present. Over the tropics, this is likely to be related to the ITCZ, which tends to have a stronger presence in the NH due to the larger land masses. The annual cycle between ~35°N and 60°N, for example, could be due to the increased cloud present during the mid-latitude storms that occur mainly during the NH winter months. However, in this analysis the annual cycles may also be related to failures in the cloud screening. Evidence presented in the previous section suggests that the cloud tests are not always successful at removing cloud that scatters radiation at 183 GHz, particularly in the lower-peaking channels (e.g. 183.31±7 GHz). Thus, when there is increased cloud, there will be an increased occurrence of negative BT differences due to cloud contamination and any annual cycle in cloud occurrence will manifest itself as an annual cycle in the BT differences between the observations and the simulations. The annual cycles in the tropics and mid-latitudes become less apparent as the channel sounding height increases (i.e. higher in the atmosphere) such that they are barely visible in the 183.31±1 GHz channels (Figure 33).

The discontinuities seen in the time series presented above are also evident in the Hovmöller diagrams. For example, the FY-3B mean differences are more positive between the end of 2011 and early 2017 (Figure 32). Finally, the erratic nature of the FY-3C time series very apparent at all latitudes, with a clear negative trend in the BT differences during 2014 and the first part of 2015, followed by a period of missing data, after which the differences are positive. The period where FY-3C has a negative step change for a few months in the middle of 2016 is also present in the Hovmöller diagram in Figure 34.

Figure 35 shows the standard deviation of the daily mean differences for the FY-3A 183.31±7.0 GHz channel. Annual cycles are also apparent in these data where higher standard deviations correspond to where the FCDR minus simulated BT difference is more negative. This behaviour is consistent with an increased prevalence of cloud in the observations with respect to the simulations as cloud is highly variable and therefore likely to result in a higher observed BT variance. As observed for the daily mean differences, the variability in the standard deviations is higher for the 183.31±7.0 GHz channel compared with the 183.31±1.0 GHz channel (Figure 36). Results for the intermediate 183 GHz channels show the same tendency noted for the daily mean differences, with a general reduction in standard deviation, and its variability with latitude and time, with increasing channel sounding height. The discontinuities noted earlier in this section, in both the time series and Hovmöller diagrams, are also present in the standard deviations.

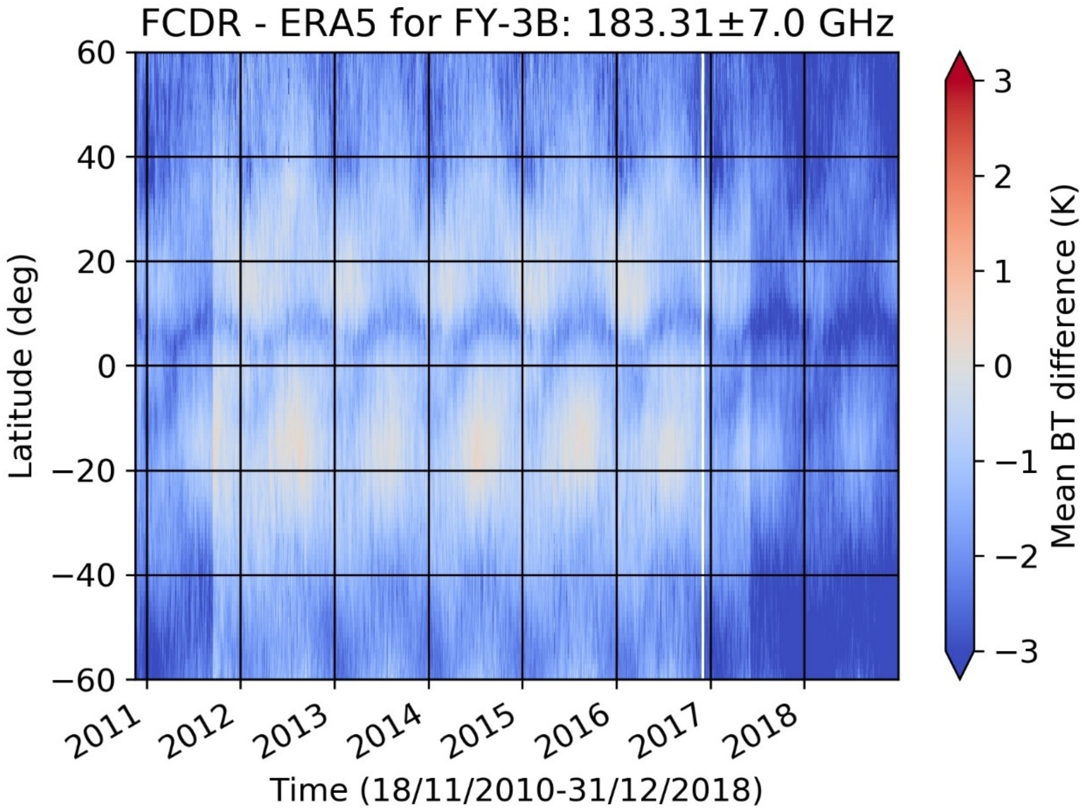

Figure 32: Hovmöller plot showing the change in daily mean difference by latitude with time for FY-3B for the 183.31±7 GHz channel.

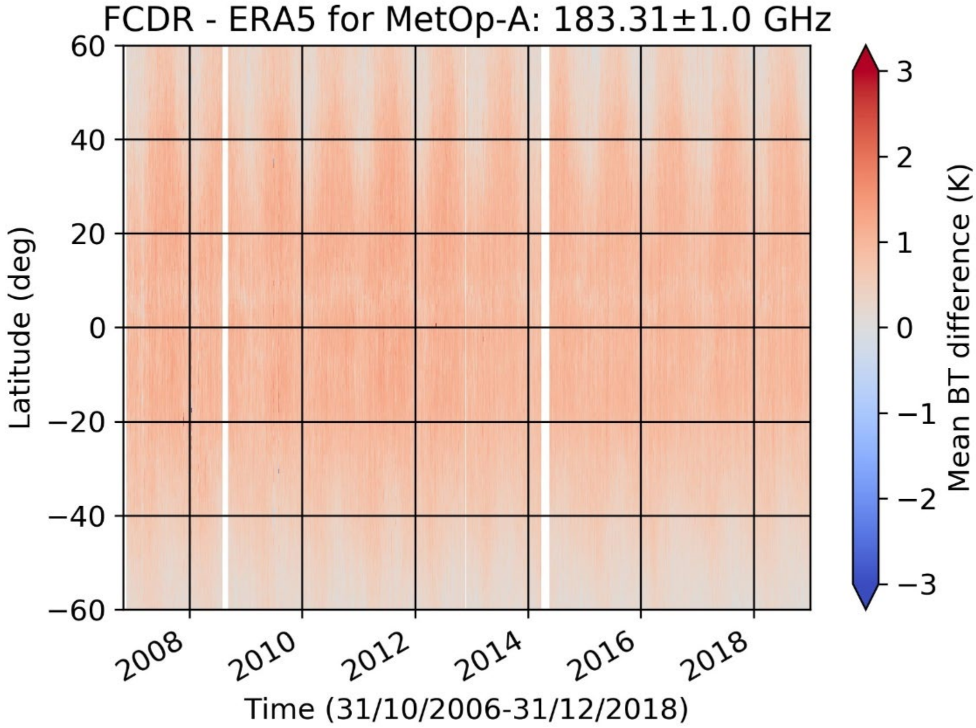

Figure 33: As for Figure 32 but for the Metop A 183.31±1 GHz channel. Note that the time axis intervals are two years rather than one, as shown for the other sensors.

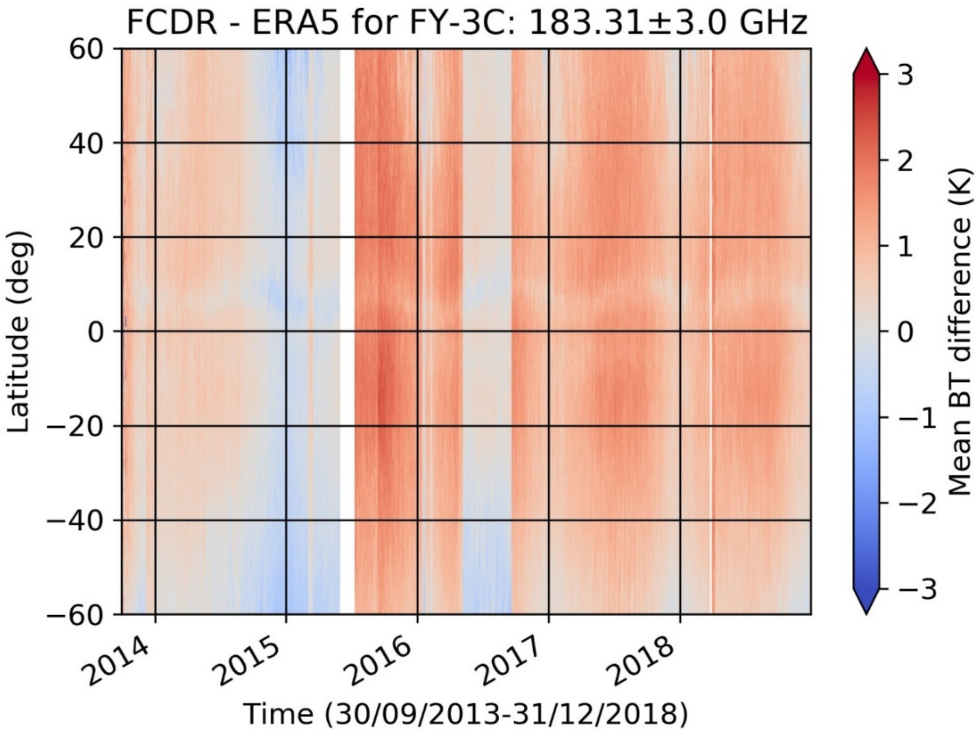

Figure 34: Hovmöller plot showing the change in daily mean differences by latitude with time for the FY-3C 183.31±3 GHz channel.

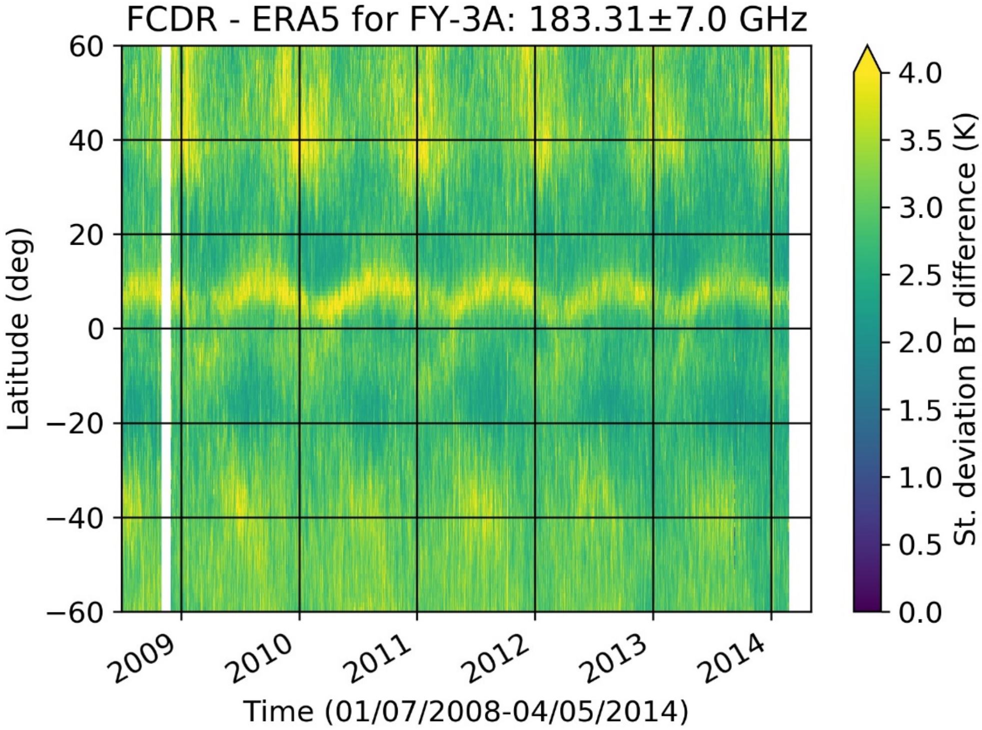

Figure 35: Time series of the standard deviation of the daily mean differences for the FY-3A 183.31±7 GHz channel

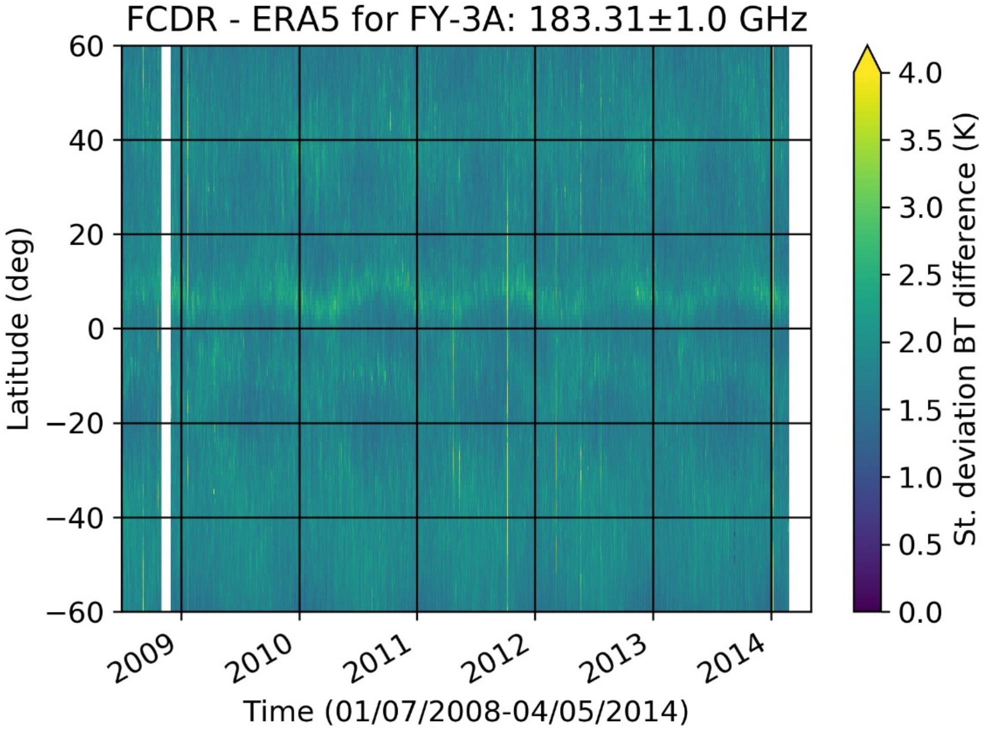

Figure 36: Time series of the standard deviation of the daily mean differences for the FY-3A 183.31±1 GHz channel

4.2.3.1 Key points

- The Hovmöller diagrams show annual cycles in the northern-hemisphere tropics, with the negative differences peaking at higher latitudes during the NH summer, and annual cycles in the mid latitudes (~35°N/S to 60°N/S).
- The annual cycles may be a result of a BT-dependent difference between the two data sets that is apparent because of the colder BTs that occur when cloud is present. This could also be because of failures in the cloud screening. In the tropics, the cycles could be linked to the ITCZ and at mid-to-high latitudes, they could be linked to mid-latitude storms. The hypothesis that the annual cycles are linked to cloud is strengthened by the stronger annual cycles and higher standard deviations seen in the 183.31±7.0 GHz channel, which is more sensitive to the lower atmosphere and thus more affected by the presence of clouds.
- The annual cycles are also visible in the standard deviation Hovmöller diagrams and higher standard deviations appear to be linked to more negative BT differences (FCDR colder than the simulations).

4.3 Analysis of Simultaneous Nadir Overpasses

4.3.1 FY-3A vs Metop A

The SNOs for FY-3A and Metop A achieve near-global coverage (Appendix C). In total there are more than 650,000 SNOs. The correlation between the sensors is >0.99, although there are some outliers. Figure 37 shows the BT differences vs Metop A BTs for this set of SNOs. For all three channels, the BT difference has a non-linear dependency on the Metop A BT but this is particularly strong for the 183.31±3 GHz channel. The scatter is higher for the 183.31±7 GHz, reflected in the standard deviation of the differences, which is 3.2 K for this channel compared to 2.3 and 1.5 for the 183.31±3 GHz and 183.31±1 GHz channels, respectively. The mean differences are 0.6 K, 1.9 K and 1.7 K for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

Figure 38 shows the box-plot time series for this sensor pair. There is a weak annual cycle in these data, where the most positive differences are seen in ~June. There is also perhaps a weak positive trend in all channels.

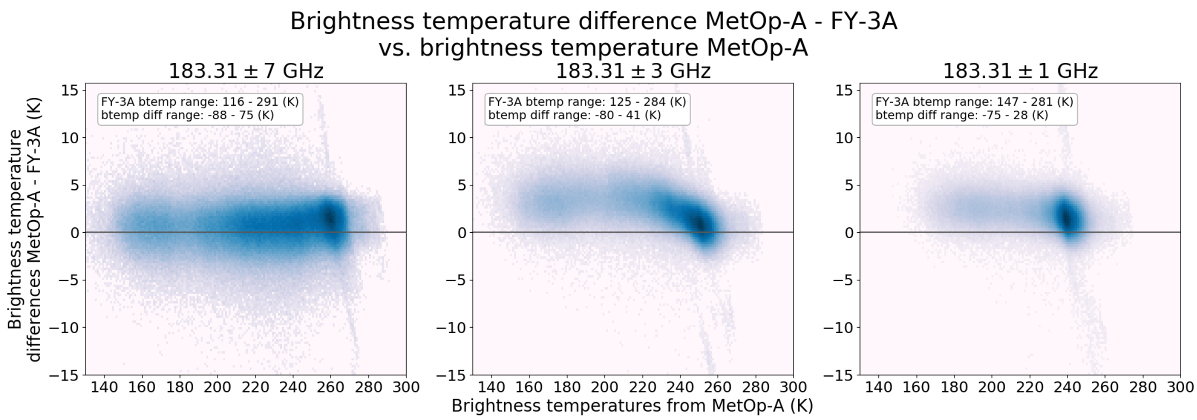

Figure 37: Metop A minus FY-3A BT differences vs Metop A BTs.

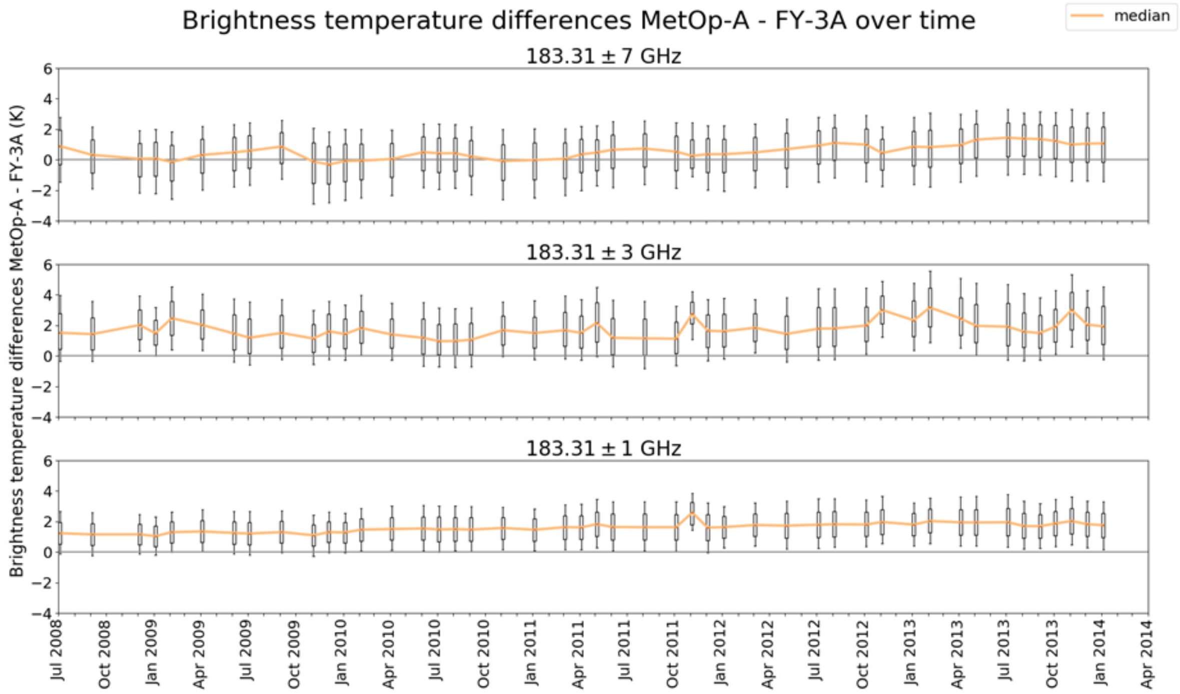

Figure 38: Box-plot time series of monthly BT differences for FY-3A / Metop A SNOs.

4.3.2 FY-3A vs Metop B

As for FY-3A and Metop A, the SNOs for FY-3A and Metop B also achieve near-global coverage (Appendix C). However, these SNOs only occur over a window of less than one year (n≈460,000). In general, the results for this sensor pair are very similar to the FY-3A and Metop A pairing, with a non-linear dependency of the BT differences on Metop B BTs (Figure 39). Mean differences (and standard deviations) for each channel are 1.1 K (1.9 K), 1.8 K (1.8 K) and 1.8 K (1.1 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively. The box-plot time series shown in Figure 40 suggests there is little variability with time in the SNO statistics (although note the limited <1 year time period).

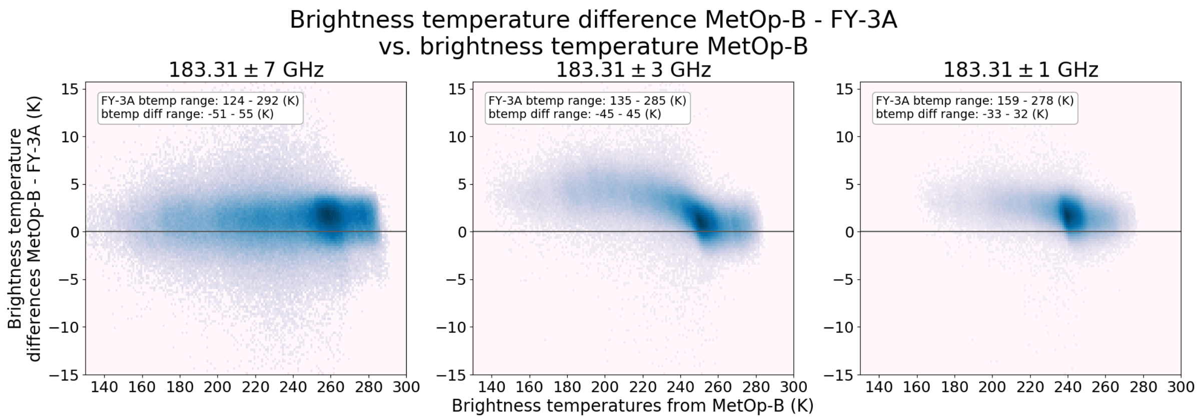

Figure 39: Metop B minus FY-3A BT differences vs Metop B BTs.

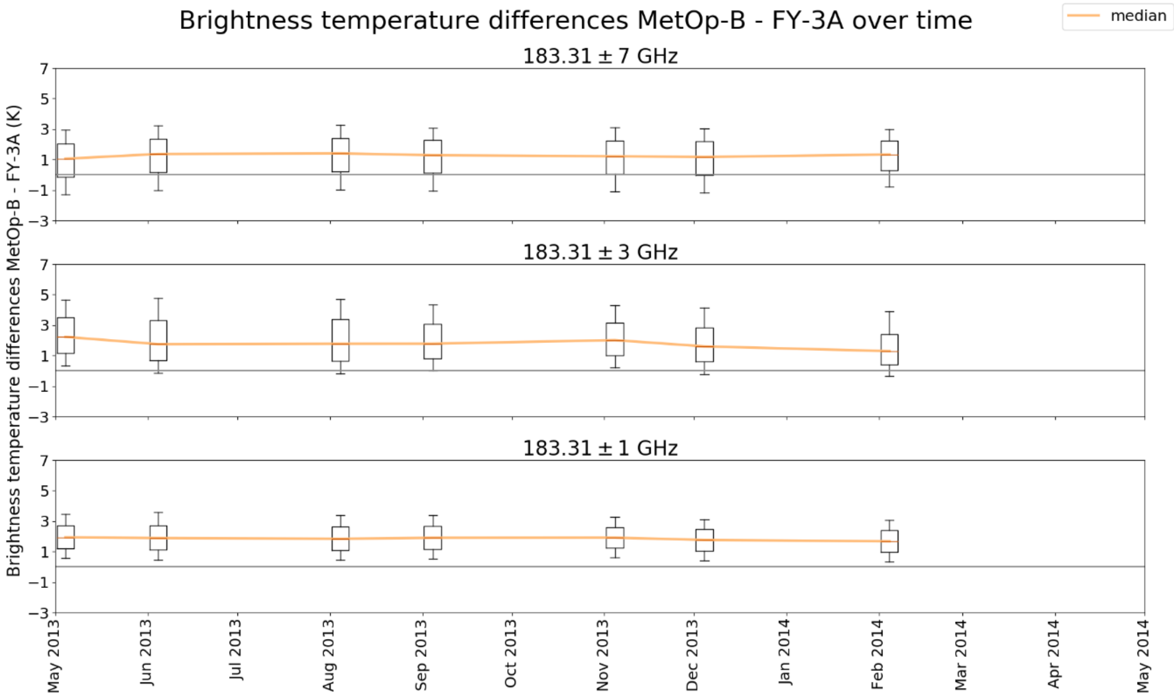

Figure 40: Box-plot time series of monthly BT differences for FY-3A / Metop B SNOs.

4.3.3 FY-3A vs SAPHIR

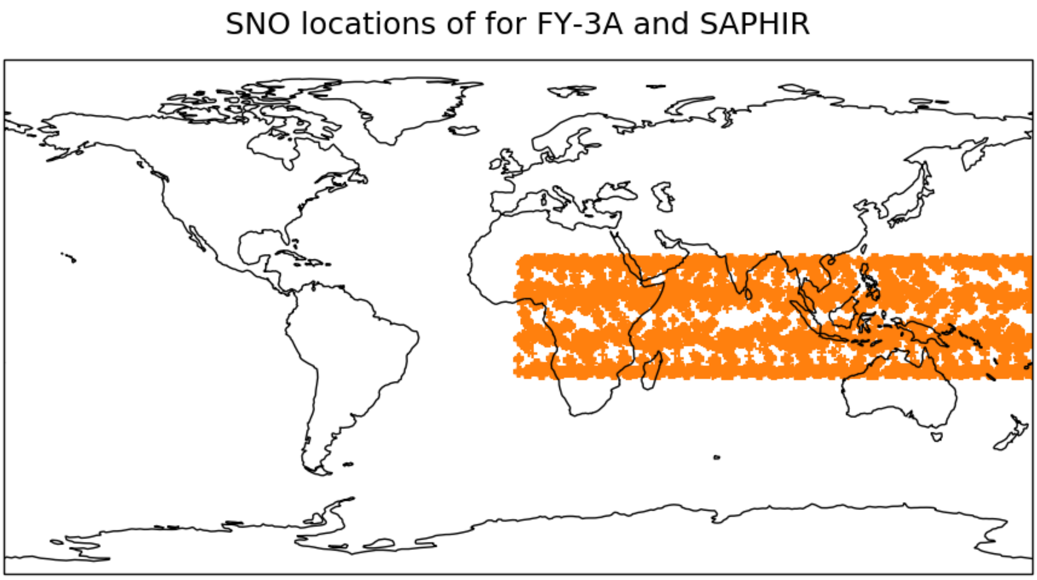

Figure 41: Map showing the locations of the FY-3A and SAPHIR SNOs.

Figure 41 shows the locations of the FY-3A/SAPHIR SNOs, which occur only at tropical latitudes owing to the orbit of SAPHIR (Section 3.1.4). There are around 47,000 SNOs for this comparison. Despite the limited checks on the SAPHIR QC flags, the correlation between the two data sets is >0.94. However, there appears to be a weak dependency of the BT difference on SAPHIR BT (Figure 42), particularly for the 183.31±1 GHz channel pair, which shows an approximately negative linear relationship (i.e. BT differences become more negative with increasing SAPHIR BT).

The box-plot time series shown in Figure 43 suggests that the differences between the sensors are stable in time. There appears to be a wider distribution of differences in September 2013, but otherwise the month-to-month variability in the distributions is quite consistent. The mean differences (and standard deviations) are 0.9 K (3.1 K), -0.4 K (2.3 K), 2.0 K (1.8 K), for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

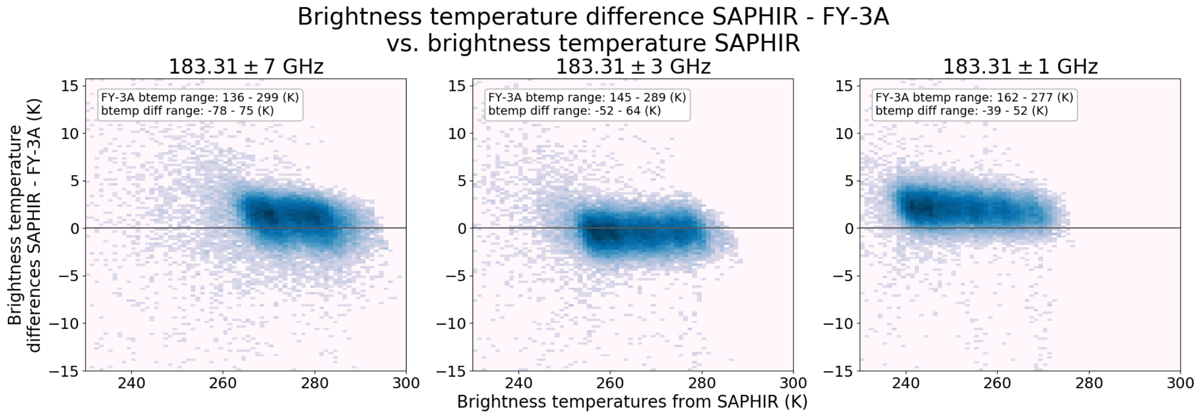

Figure 42: SAPHIR minus FY-3A BT differences vs SAPHIR BTs

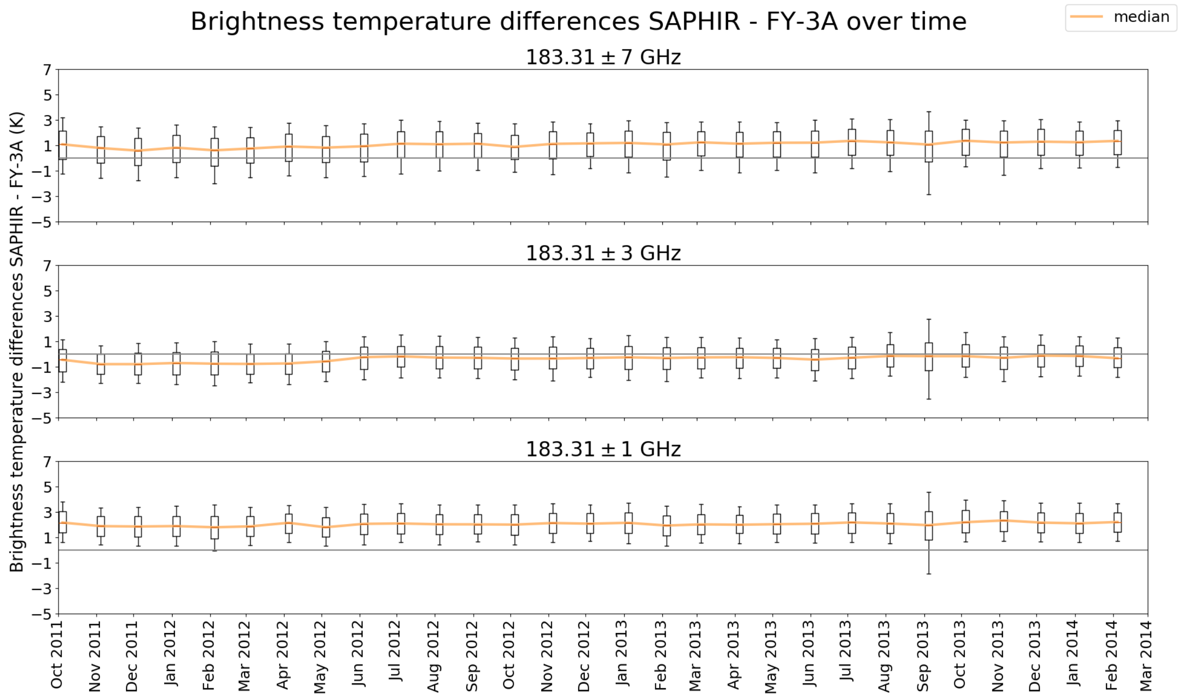

Figure 43: Box-plot time series of monthly BT differences for FY-3A / SAPHIR SNOs

4.3.4 FY-3A vs S-NPP

All the ~75,000 SNOs between FY-3A and S-NPP occur near the poles (Appendix C). As for the comparisons between FY-3A and the Metop data, there appears to be a non-linear dependency of the BT differences on the S-NPP BT (Figure 44). However, it should be noted that this is less conclusive given that most of the BTs occur within a very limited range due to the high-latitude locations of the SNOs. In addition, the SNOs only occur in

April – June 2012 and September - November 2013 (Figure 45). The mean (and standard deviation) of the BT differences are 1.1 K (2.0 K), 2.5 K (1.5 K), 1.4 K (1.3 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

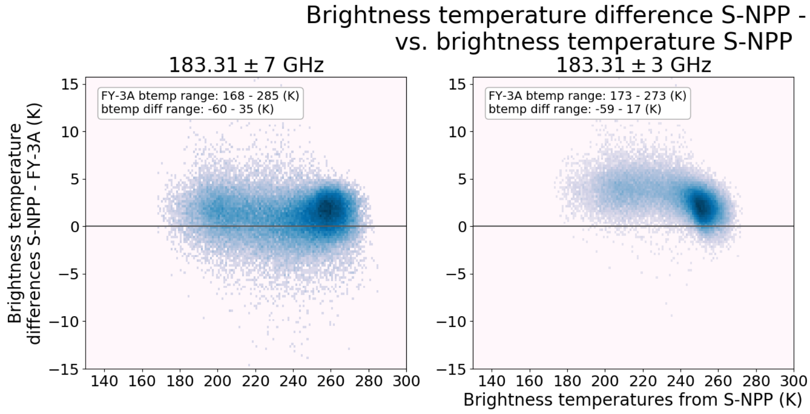

Figure 44: S-NPP minus FY-3A BT differences vs S-NPP BTs

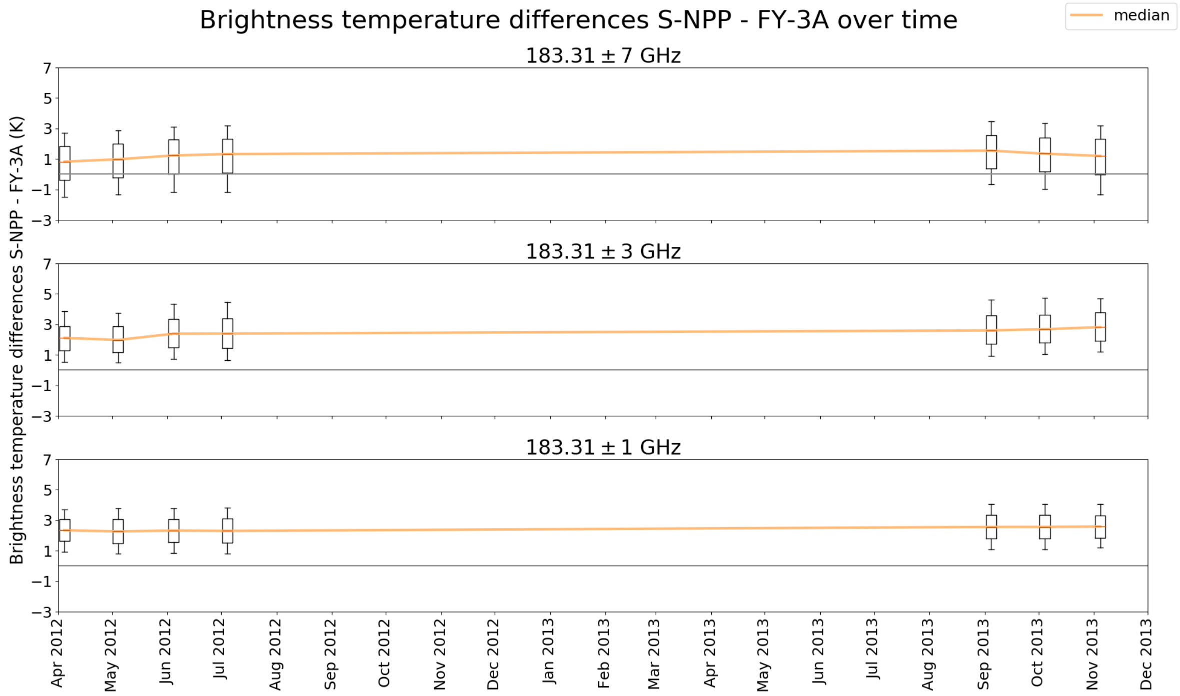

Figure 45: Box-plot time series of monthly BT differences for FY-3A / S-NPP SNOs

4.3.5 FY-3B vs FY-3C

The locations of the ~70,000 FY-3B / FY-3C SNOs are very similar to the locations of the FY-3A / S-NPP SNOs and occur only near the poles (Appendix C – shown only for FY-3A / S-NPP). As seen for the FY-3A SNO pairs, there is a non-linear dependency of the BT differences on the FY-3C BTs (Figure 46) with a general tendency of more negative BT differences at higher FY-3C BTs. The box-plot time series (Figure 47) suggest the difference is reasonably stable for the 183.31±1 GHz channel, but there is an annual cycle in the differences for the 183.31±7 GHz and 183.31±3 GHz channels that appears to peak in January/February. The mean (and standard deviation) of the BT differences are 0.8 K (2.6 K), 3.9 K (1.4 K), 0.8 K (1.2 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

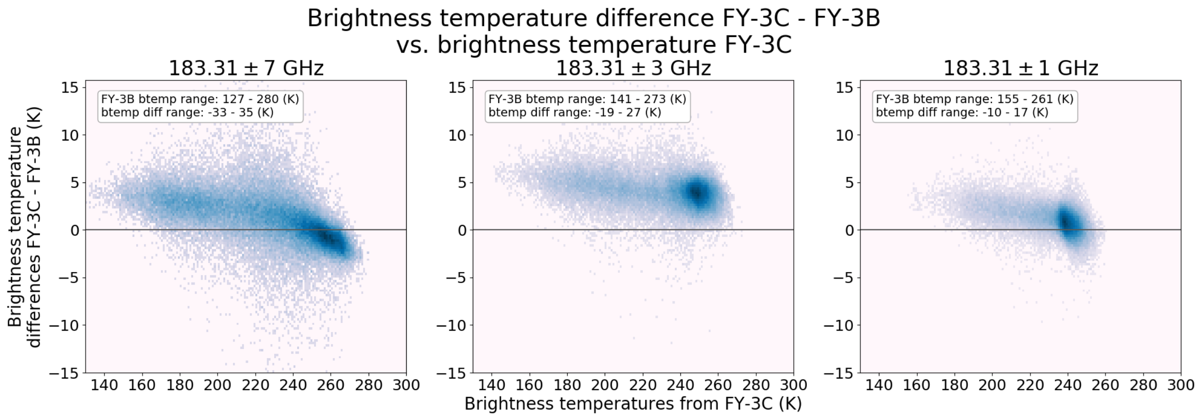

Figure 46: FY-3C minus FY-3B BT differences vs FY-3C BTs

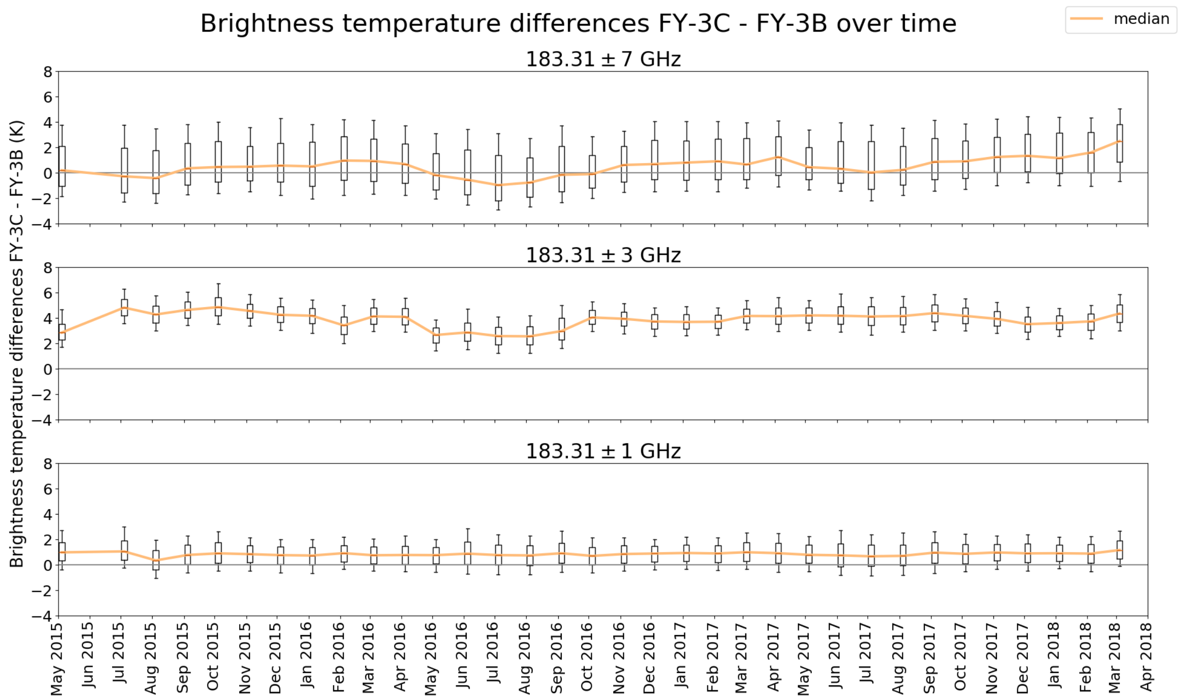

Figure 47: Box-plot time series of monthly BT differences for FY-3B / FY-3C SNOs

4.3.6 FY-3B vs Metop A

The locations of the ~140,000 FY-3B / Metop A SNOs are very similar to the locations of the FY-3A / S-NPP SNOs and occur only near the poles (Appendix C – shown only for FY-3A / S-NPP). Figure 48 shows that as seen for previous sensor pairs, there is a strong non-linear dependency of the BT difference on the Metop A BTs, which has a negative slope (BT differences become more negative with increasing BT). As for the FY-3B / FY-3C SNOs, there appears to be an annual cycle in these BT differences that appears to peak in January/February (Figure 49). In general, this peak also tends to coincide with a smaller interquartile range of the BT differences. The mean (and standard deviation) of the BT differences are 3.6 K (3.0 K), 3.1 K (1.9 K), and 2.7 K (1.3 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

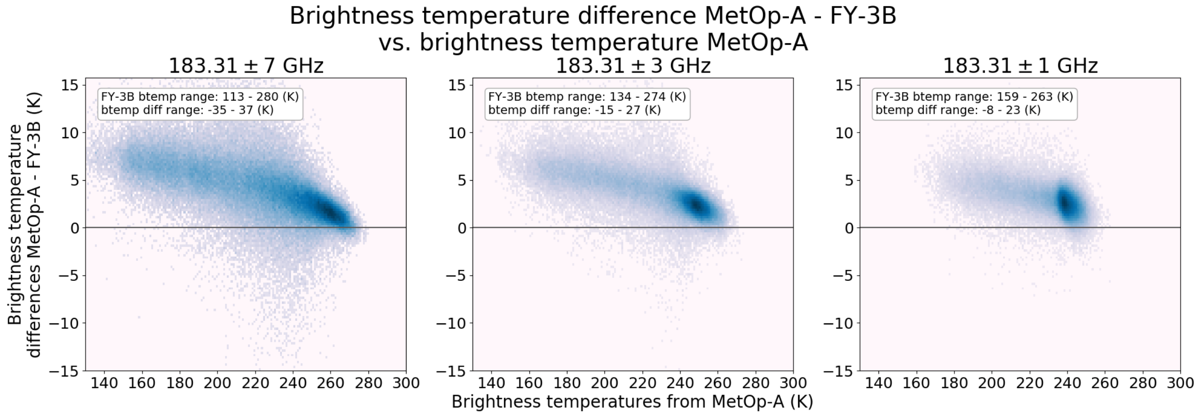

Figure 48: Metop A minus FY-3B BT differences vs Metop A BTs

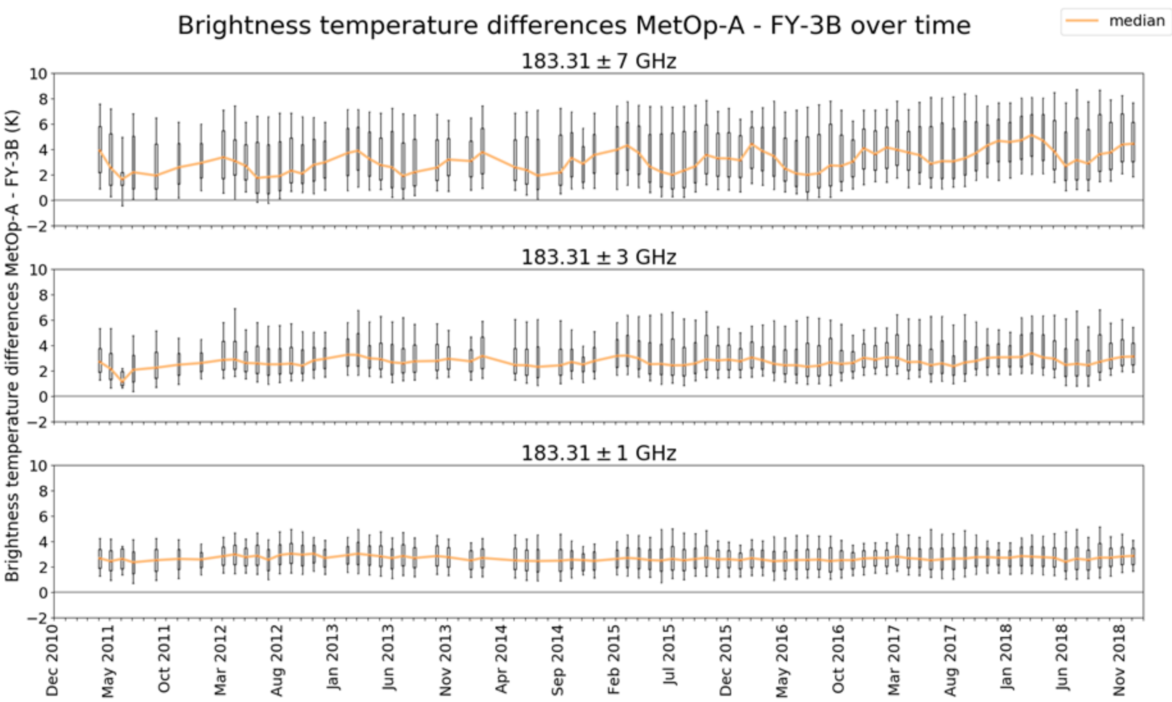

Figure 49: Box-plot time series of monthly BT differences for FY-3B / Metop A SNOs

4.3.7 FY-3B vs Metop B

The locations of the ~105,000 FY-3B / Metop B SNOs are very similar to the locations of the FY-3A / S-NPP SNOs and occur only near the poles (Appendix C – shown only for FY-3A / S-NPP). As seen in previous sensor pairs, there is a non-linear response of the BT difference to the Metop B BTs, which is particularly strong for the 183.31±7 GHz channel. There is also an annual cycle in the BT differences that appears to peak in January/February together with a narrower distribution in the BT differences (Figure 51). The mean (and standard deviation) of the BT differences are 3.7 K (2.9 K), 3.2 K (1.9 K), and 2.8 K (1.3 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

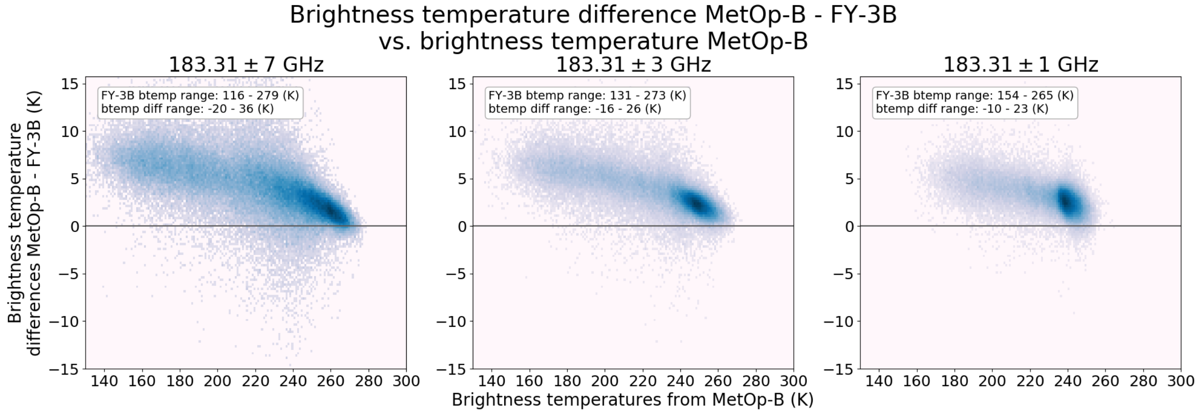

Figure 50: Metop B minus FY-3B BT differences vs Metop B BTs

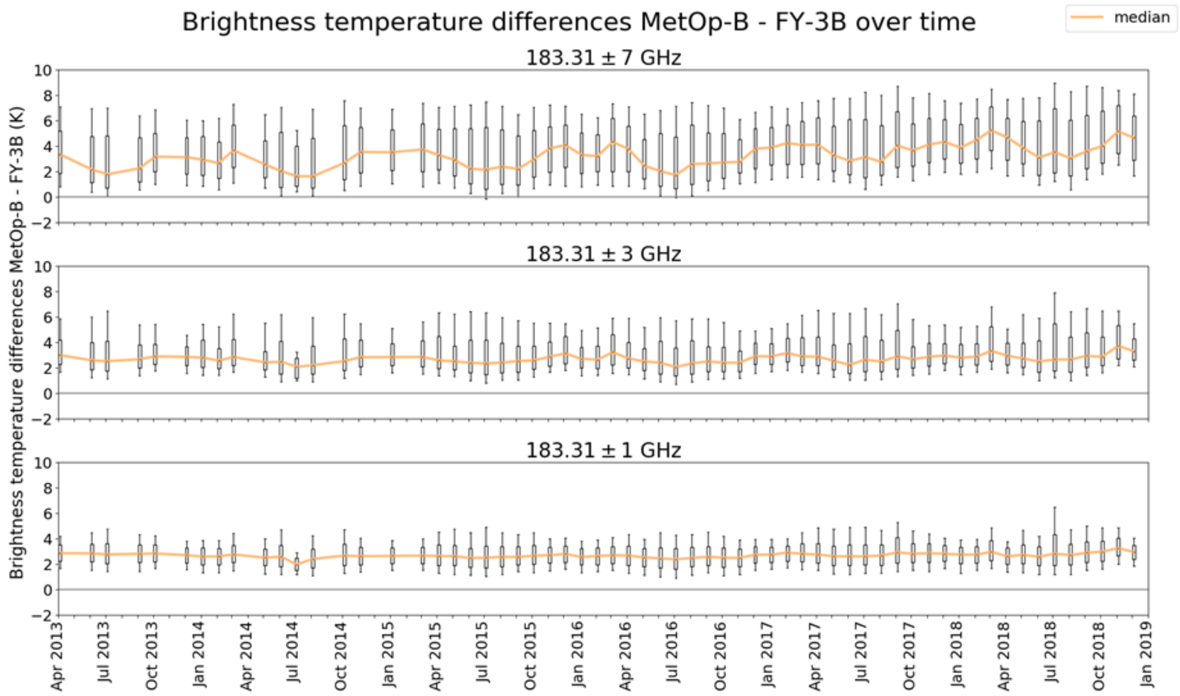

Figure 51: Box-plot time series of monthly BT differences for FY-3B / Metop B SNOs

4.3.8 FY-3B vs SAPHIR

As for the FY-3A / SAPHIR SNO analysis, all ~81,000 FY-3B / SAPHIR SNOs also occur in the tropics (Figure 41). Figure 52 shows the BT differences plotted against the SAPHIR BTs, which show a clear negative relationship for each channel (i.e. BT differences become more negative with increasing SAPHIR BT). The box plot time series is generally quite stable both in terms of the median difference for each month, and also the interquartile range (Figure 53). There is an anomaly in September 2013 where the distribution is uncharacteristically wide; this is similar to the anomaly seen in the FY-3A / SAPHIR SNO analysis, suggesting an anomaly in the SAPHIR data rather than the FCDR. The mean (and standard deviation) of the BT differences are -0.78 K (4.3 K), 0.0 K (3.1 K), and 2.6 K (2.3 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

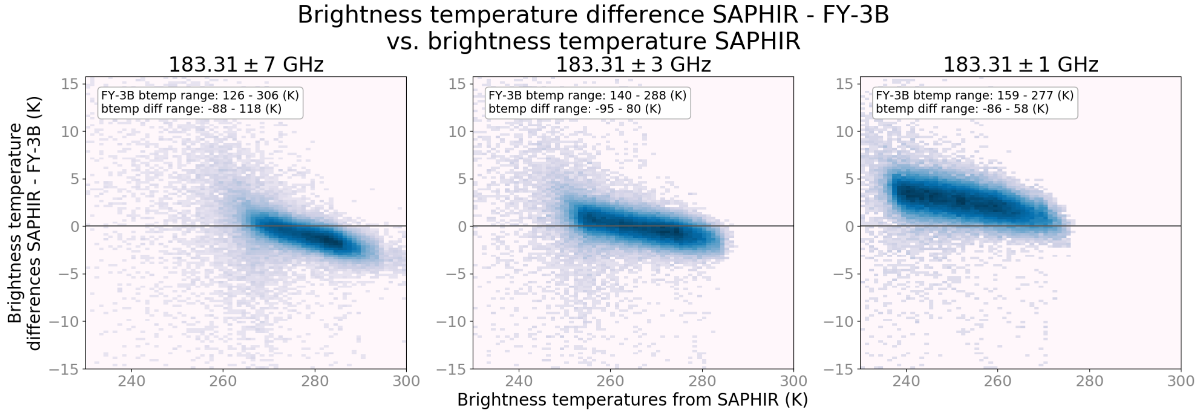

Figure 52: SAPHIR minus FY-3B BT differences vs SAPHIR BTs

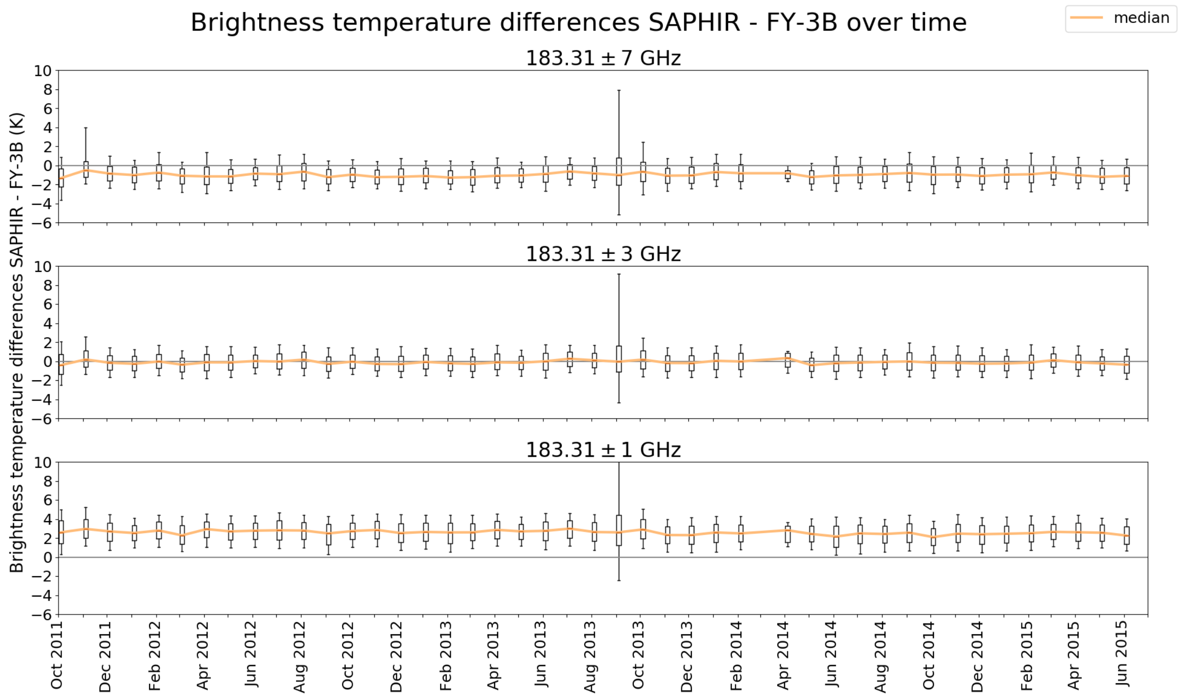

Figure 53: Box-plot time series of monthly BT differences for FY-3B / SAPHIR SNOs

4.3.9 FY-3B vs S-NPP

All ~450,000 SNOs for the FY-3B / S-NPP sensor pair occur near the poles (Appendix C – shown only for FY-3A / S-NPP). As observed for previous sensor pairings involving the FY-3 instruments, there is a strong dependency of the BT difference on S-NPP BTs Figure 54). The box-plot time series shown in Figure 55 shows a very clear example of the annual cycle seen for previous sensor pairings, that peaks in Jan/Feb together with a reduction in the interquartile range of BT differences. The mean (and standard deviation) of the BT differences are 4.1 K (3.7 K), 4.4 K (2.5 K), and 3.6 K (2.5 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

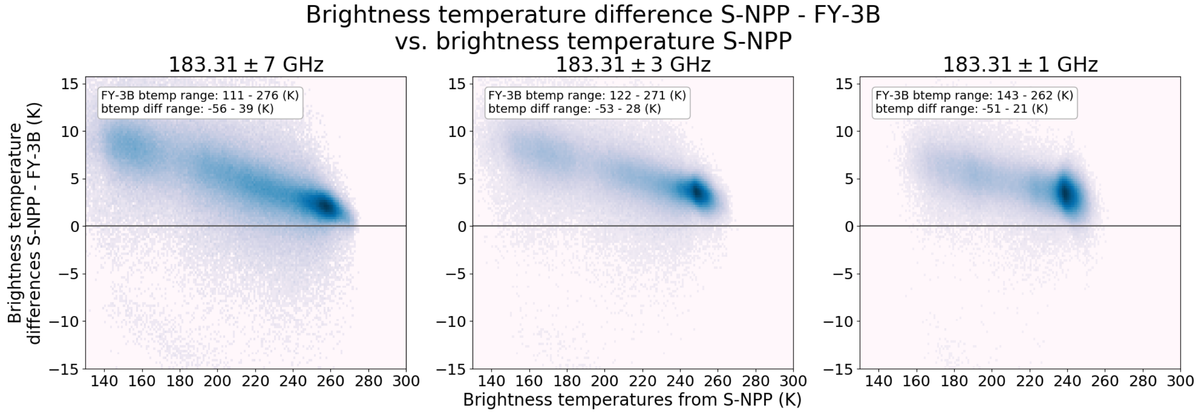

Figure 54: S-NPP minus FY-3B BT differences vs S-NPP BTs

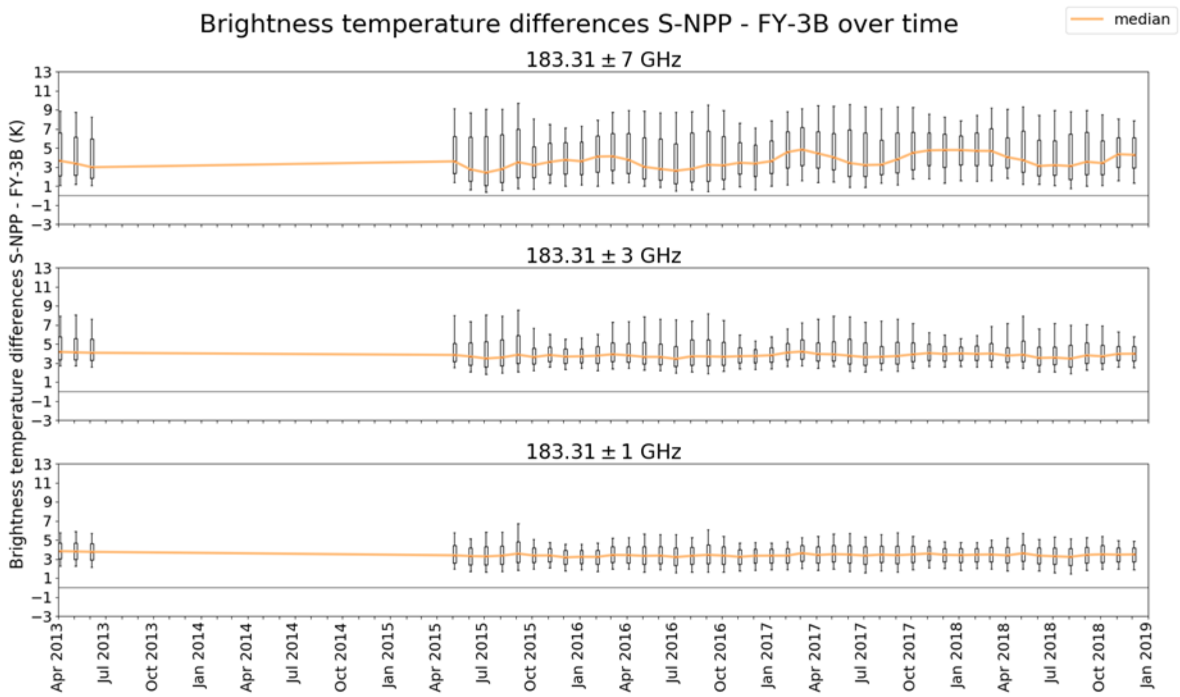

Figure 55: Box-plot time series of monthly BT differences for FY-3B / S-NPP SNOs

4.3.10 FY-3C vs Metop A

All ~440,000 SNOs for the FY-3C / Metop A sensor pair occur near the poles (Appendix C – shown only for FY-3A / S-NPP). As observed for previous sensor pairings involving the FY-3 instruments, there is a dependency of the BT difference on Metop A BTs but this is weak for the 183.31±7 GHz and 183.31±1 GHz channels (Figure 56). The box-plot time series shown in Figure 57 shows some variability in the agreement with time, but any annual cycle is unclear. The mean (and standard deviation) of the BT differences are 2.7 (1.4 K), -0.7 K (1.4 K), and 1.7 K (1.0 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively. It is notable that the 183.31±7 GHz and 183.31±1 GHz channels are more stable over time compared with the 183.31±3 GHz channel that has some jumps (e.g. mid 2015) in agreement with the results from the comparisons between the FY-3C FCDR and the equivalent simulations (Figure 30).

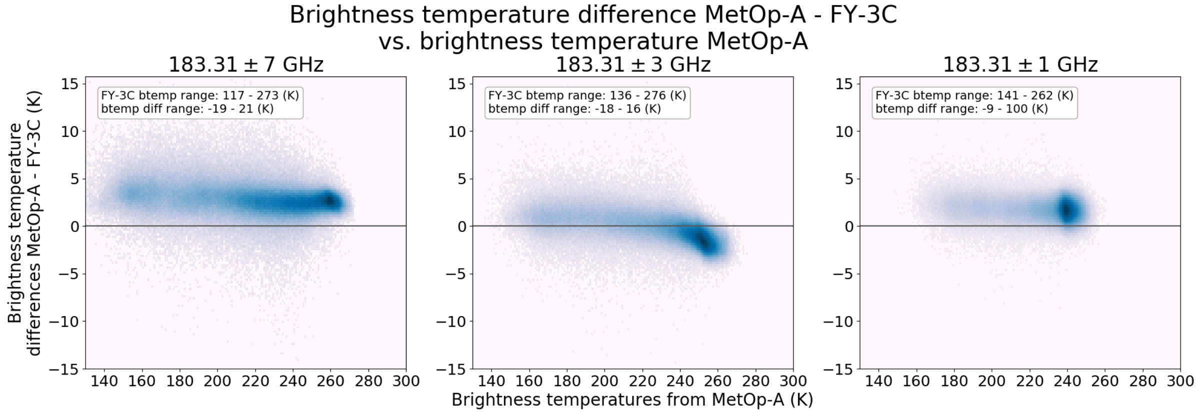

Figure 56: Metop A minus FY-3C BT differences vs Metop A BTs

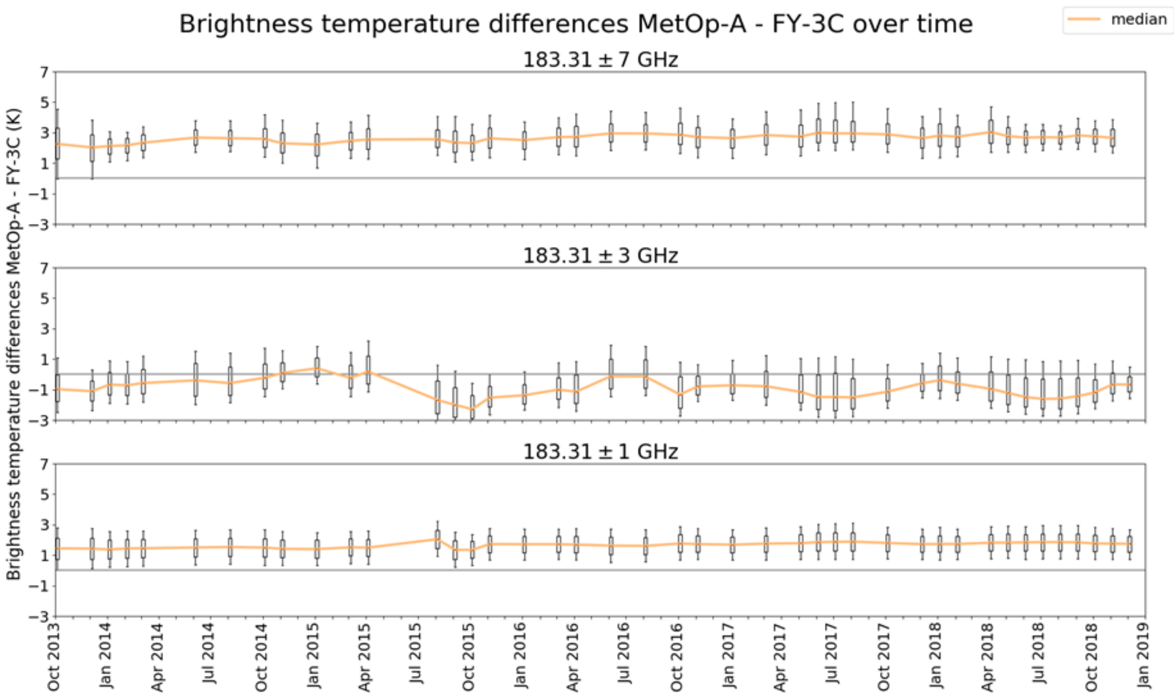

Figure 57: Box-plot time series of monthly BT differences for FY-3C / Metop A SNOs

4.3.11 FY-3C vs MetopB

All ~500,000 FY-3C / Metop B SNOs occur at high latitudes but extend further south compared with the previous sensor pairs with SNOs near the poles (Appendix C). The dependency of the BT differences on Metop B BTs is very similar to the results for the FY-3C / Metop A SNO analysis. The time series analysis is also very similar and shows no annual cycle or other systematic variability (Figure 59). The variability in BT differences with time in the 183.31±3 GHz channel is particularly high. The distribution of BT differences for all channels appear to be particularly high in April 2018. The mean (and standard deviation) of the BT differences are very similar to the comparison between FY-3C and Metop A: 2.7 (2.0 K), -0.6 K (1.7 K), and 1.8 K (1.0 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

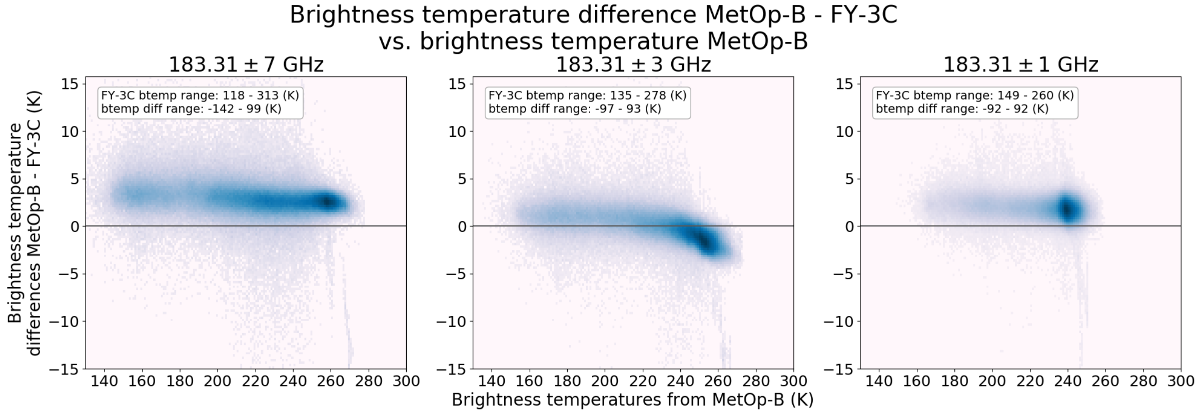

Figure 58: Metop B minus FY-3C BT differences vs Metop B BTs

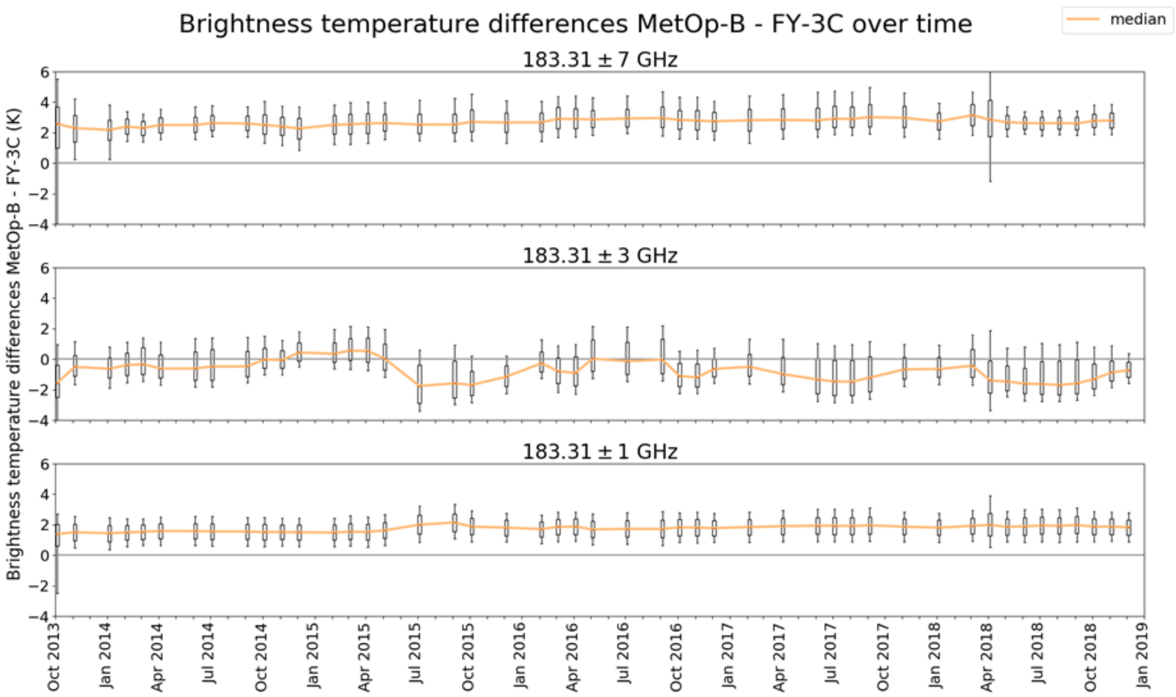

Figure 59: Box-plot time series of monthly BT differences for FY-3C / Metop B SNOs

4.3.12 FY-3C vs SAPHIR

As for previous SNO analyses using SAPHIR data, all comparisons for this sensor pairing are restricted to tropical latitudes in the western hemisphere (Figure 41). Figure 60 shows the dependency of the BT differences on SAPHIR BT, which has a negative slope (i.e. BT differences become more negative with increasing BT), particularly for the 183.31±7 GHz channel. Figure 61 shows the box-plot time series for the ~33,000 FY-3C / SAPHIR SNOs grouped by month of the year, which appear quite stable for the 183.31±7 GHz and 183.31±1 GHz channels, but less so for the 183.31±3 GHz channel, which shows variability in the median of ~1 K. The mean (and standard deviation) of the BT differences are 2.1 (3.5 K), -1.9 K (2.5 K), and 2.3 K (1.7 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

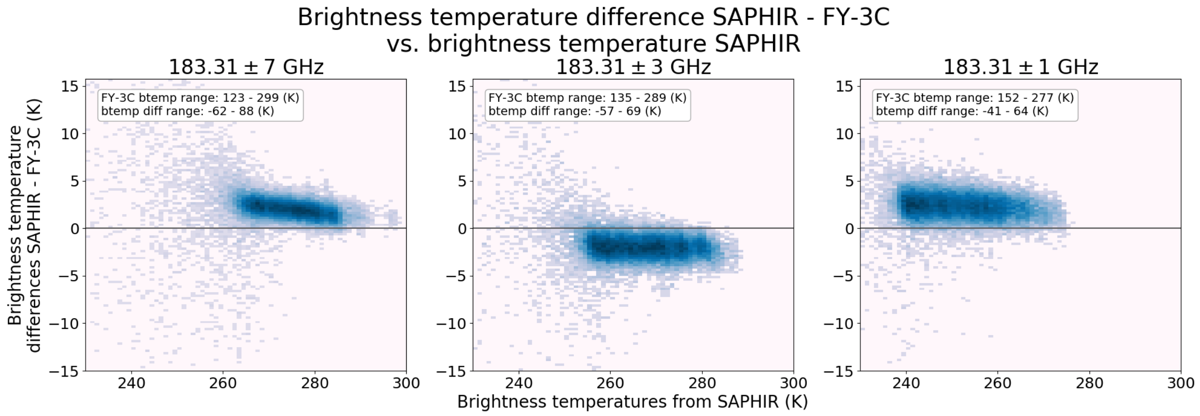

Figure 60: SAPHIR minus FY-3C BT differences vs SAPHIR BTs

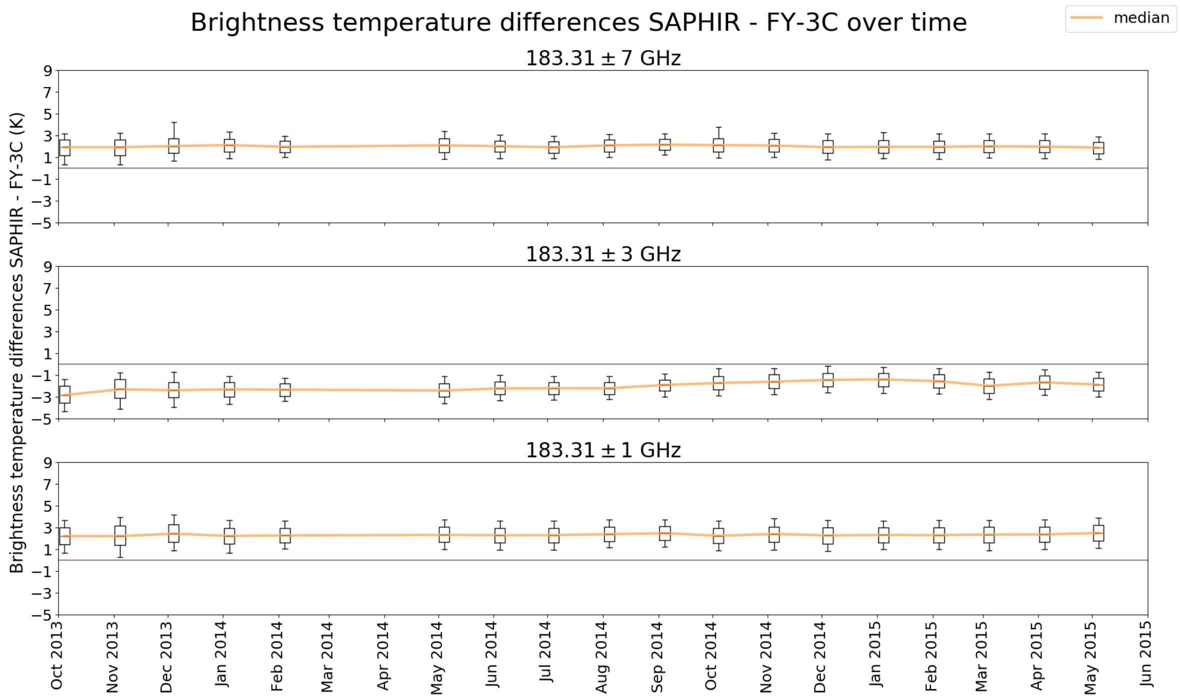

Figure 61: Box-plot time series of monthly BT differences for FY-3C / SAPHIR SNOs

4.3.13 FY-3C vs S-NPP

The ~180,000 SNOs for the FY-3C / S-NPP sensor pair all occur near the poles (Appendix C – see the map for FY-3A and S-NPP). As for all other SNO pairs including an FY-3 sensor, the BT differences exhibit some dependency on BT, particularly for the 183.31±7 GHz channel (Figure 62). Figure 63 shows the box-plot time series, which shows that the BT differences for the two lower sounding channels at 183.31±3 GHz and 183.31±7 GHz vary with time, although there are too few data to determine whether there is an annual cycle present. By contrast, there is very little variability with time in the 183.31±1 GHz channel. The mean (and standard deviation) of the BT differences are 2.1 (3.7 K), -0.5 K (2.3 K), and 1.9 K (3.5 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

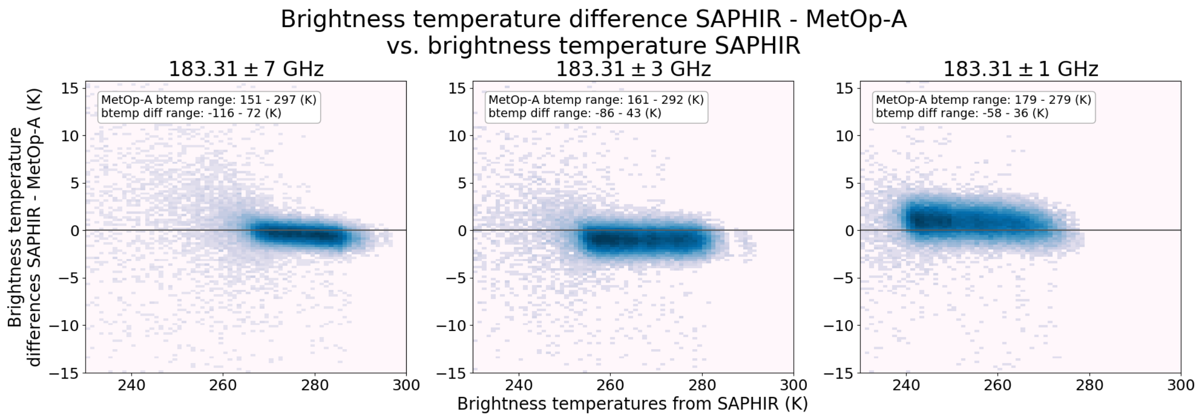

Figure 62: S-NPP minus FY-3C BT differences vs S-NPP BTs

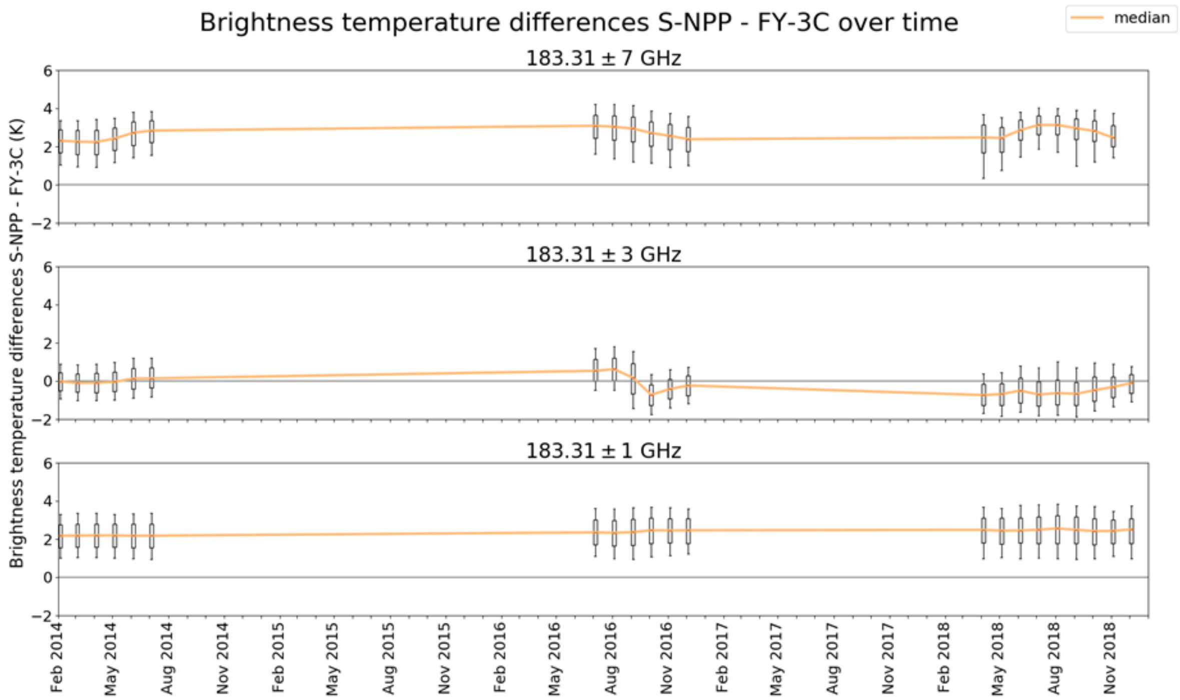

Figure 63: Box-plot time series of monthly BT differences for FY-3C / S-NPP SNOs

4.3.14 Metop A vs SAPHIR

As for previous SNO analyses using SAPHIR observations, all ~60,000 SNOs for Metop A / SAPHIR occur in the tropics. Figure 64 shows a weak dependency of the BT difference on BT. Figure 65 shows the box-plot time series for these BT differences grouped by month of the year, which is remarkably stable in time in both the median and the interquartile range with the exception of September 2013, also seen in previous SNO analyses that include SAPHIR. The mean (and standard deviation) of the BT differences are -0.3 (2.4 K), -0.9 K (1.7 K), and 0.9 K (1.5 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

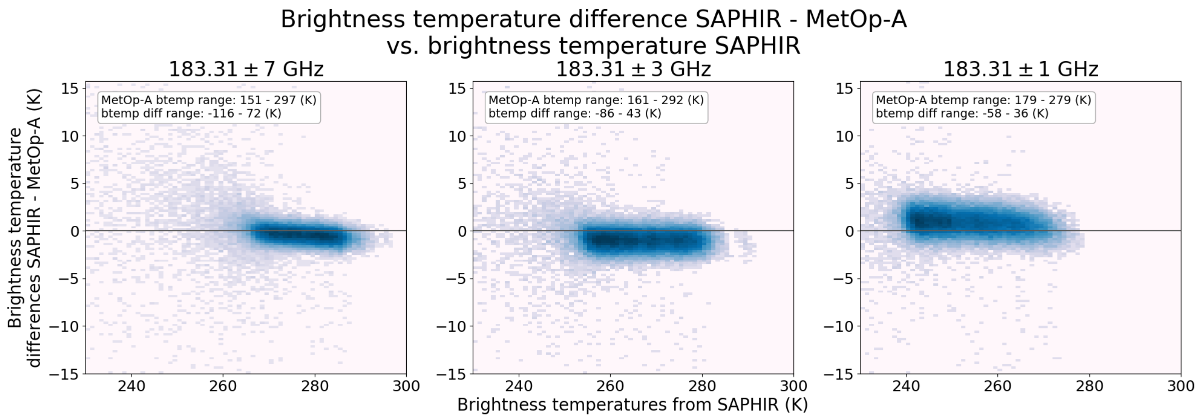

Figure 64: SAPHIR minus Metop A BT differences vs SAPHIR BTs

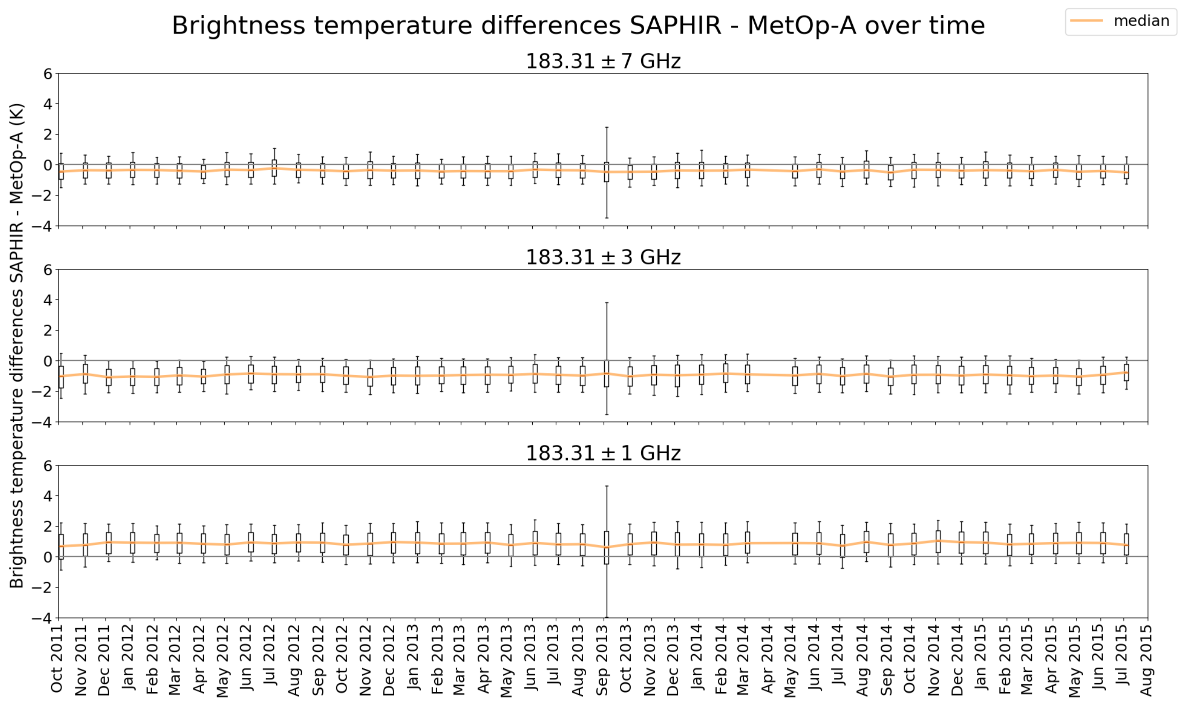

Figure 65: Box-plot time series of monthly BT differences for Metop A / SAPHIR SNOs

4.3.15 Metop A vs S-NPP

All ~114,000 Metop A / S-NPP SNOs occur near the poles. Figure 66 shows there is a very weak dependency of the S-NPP minus Metop A BT difference on S-NPP BTs, particularly in the two upper sounding channels at 183.31±3 GHz and 183.31±1 GHz. The box-plot time series in Figure 67 shows that there is some variability in the BT difference with time for the 183.31±7 GHz channel and to a lesser degree in the 183.31±3 GHz. This variability in time is consistent with what we should expect based on the frequency difference between the 2 instruments as shown in Figure 2 of RD19. By contrast, the 183.31±1 GHz channel is very stable, both in terms of the variability in the monthly median and in the inter-quartile range. The mean (and standard deviation) of the BT differences are 0.0 (1.2 K), 0.8 K (1.0 K), and 0.7 K (1.2 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively

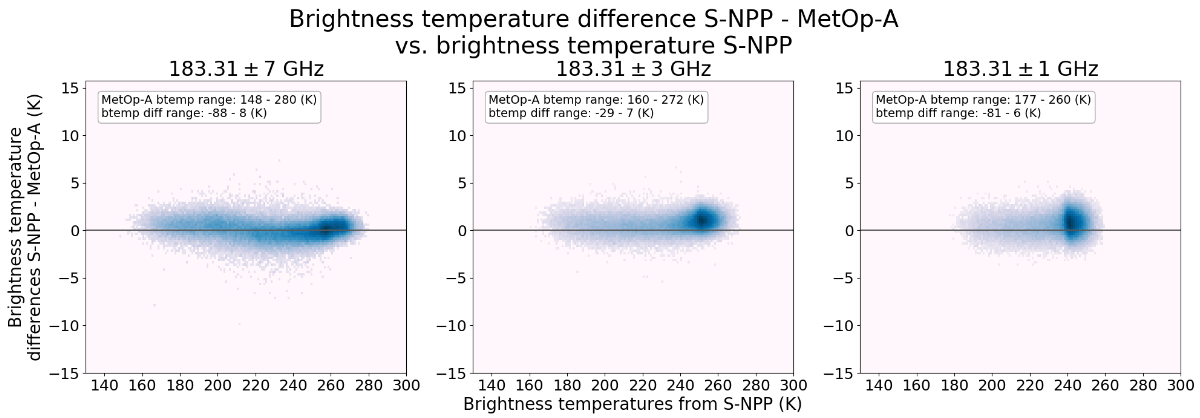

Figure 66: S-NPP minus Metop A BT differences vs S-NPP BTs

.

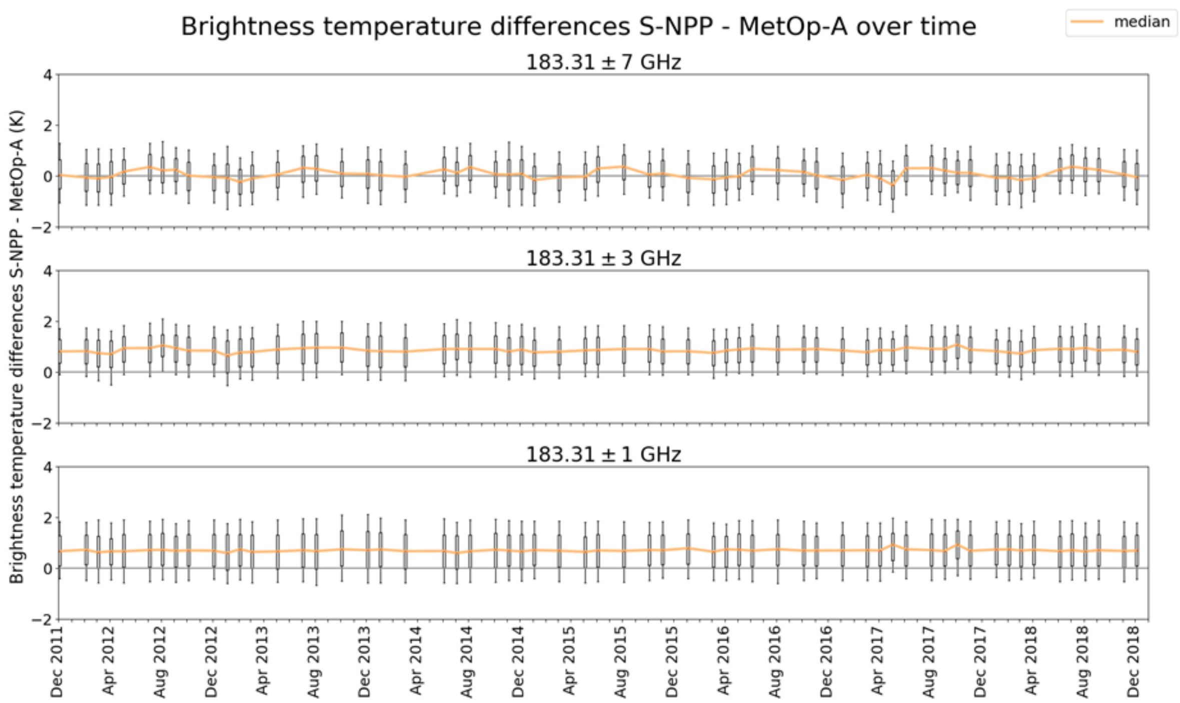

Figure 67: Box-plot time series of monthly BT differences for Metop A / S-NPP SNOs

4.3.16 Metop B vs SAPHIR

As for previous SNO analyses using SAPHIR data, all ~35,000 Metop B / SAPHIR SNOs occur in the tropics in the western hemisphere. Figure 68 shows the BT difference has a very weak dependency on the SAPHIR BTs, where the slope is slightly negative (i.e. BT differences become more negative with increasing SAPHIR BT). Figure 69 shows that the monthly BT differences are remarkably stable, both in terms of the median and the interquartile range. As seen in the previous SNO analyses with SAPHIR data, there are a particularly high number of outliers in September 2013. The mean (and standard deviation) of the BT differences are -0.2 (2.6 K),- 0.8 K (2.0 K), and 0.9 K (1.8 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

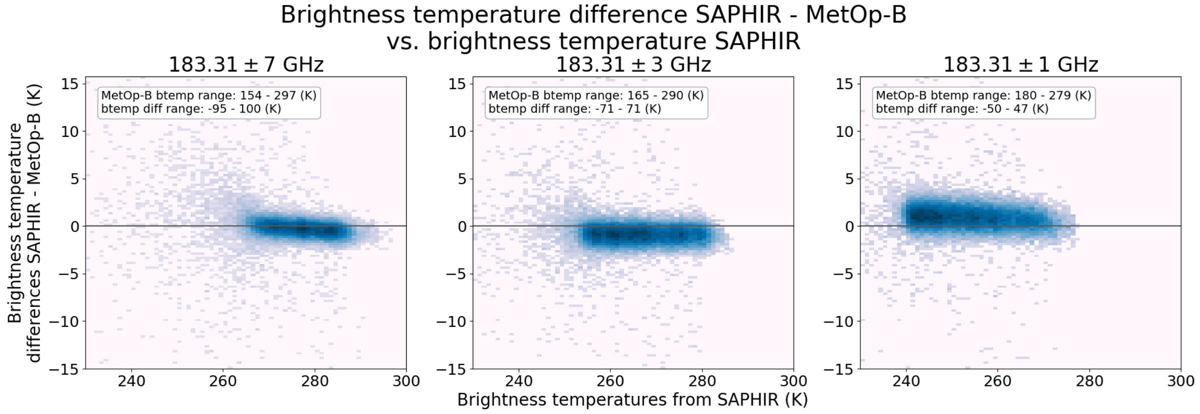

Figure 68: SAPHIR minus Metop B BT differences vs SAPHIR BTs

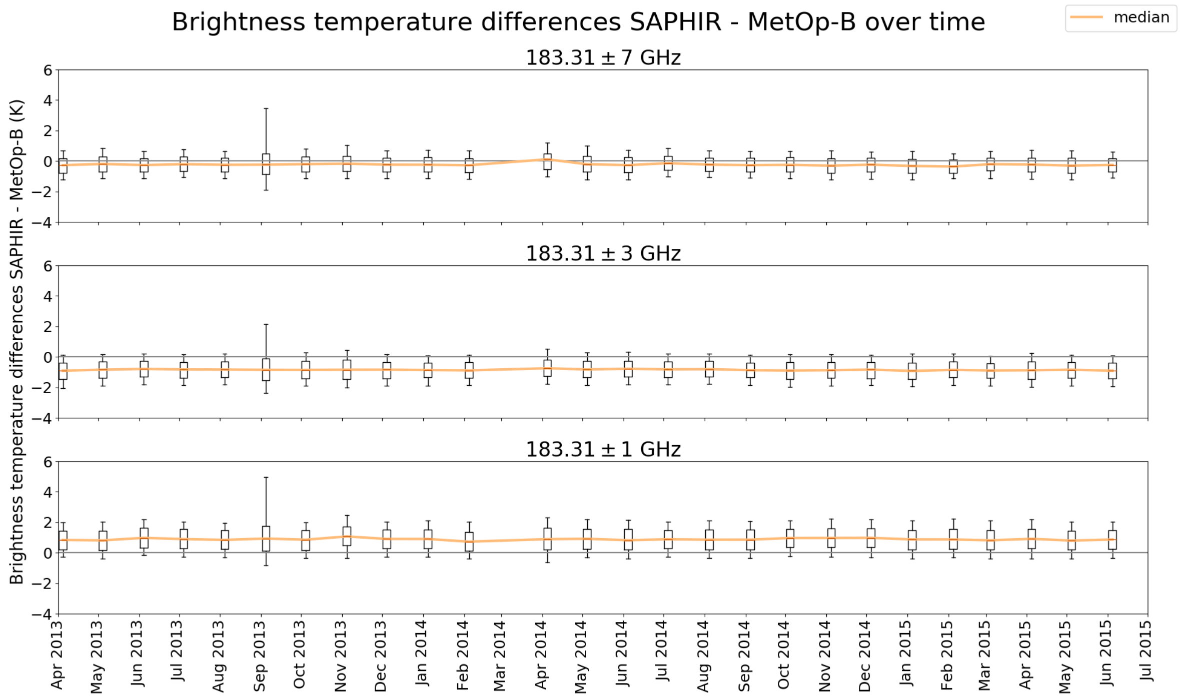

Figure 69: Box-plot time series of monthly BT differences for Metop B / SAPHIR SNOs

4.3.17 Metop B vs S-NPP

All ~95,000 Metop B / S-NPP SNOs occur close to the poles (Appendix C: shown only for FY-3A and S-NPP). Figure 70 shows that the BT differences are generally quite stable with S-NPP BTs, which is in contrast to most of the SNO pairs presented in previous sections. Figure 71 shows the box-plot time series. These results suggest that the monthly differences are generally stable for the 183.31±3 GHz and 183.31±1 GHz channels, but variable for the 183.31±7 GHz channel. The time series for this channel exhibits a weak annual cycle that peaks in August. The mean (and standard deviation) of the BT differences are 0.0 K (1.2 K), 0.8 K (1.1 K), and 0.6 K (1.2 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

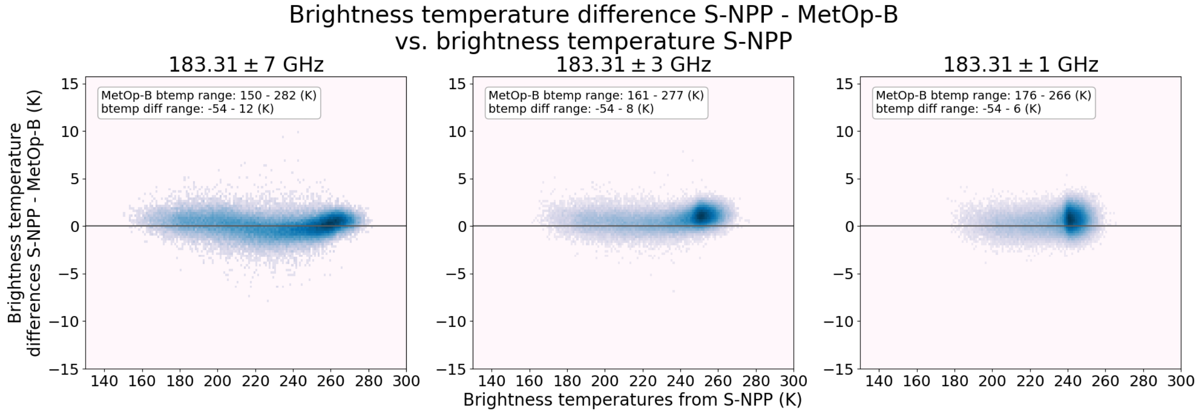

Figure 70: S-NPP minus Metop B BT differences vs S-NPP BTs

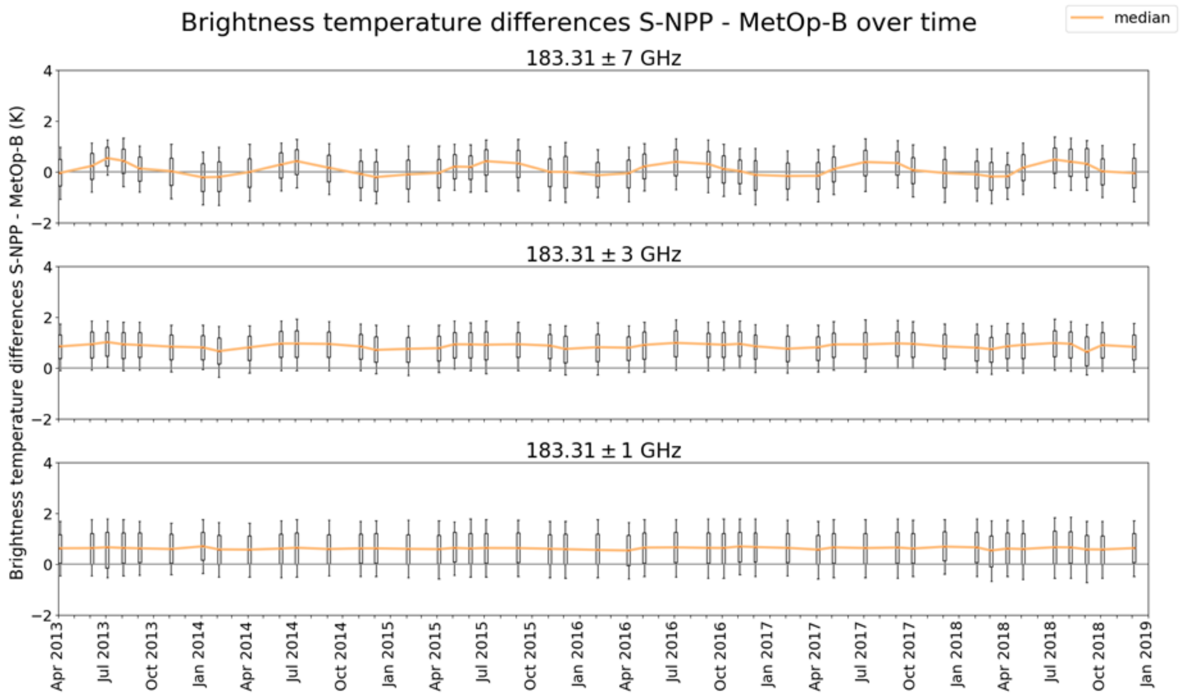

Figure 71: Box-plot time series of monthly BT differences for S-NPP / Metop B SNOs

4.3.18 S-NPP vs SAPHIR

As for previous SNO analyses using SAPHIR data, all ~60,000 S-NPP / SAPHIR SNOs occur in the tropics in the western hemisphere. Figure 72 shows there is no dependency of the BT differences on SAPHIR BTs. Figure 73 shows the box-plot time series for the monthly differences, which is very stable for all channels both in the median, and the interquartile range, with the exception of the September 2013 anomaly, which has been apparent in all other SNO analyses in this study that have used SAPHIR. The mean (and standard deviation) of the BT differences are -1.0 K (2.7 K), -2.1 K (2.1 K), and -0.1 K (1.7 K) for the 183.31±7 GHz, 183.31±3 GHz and 183.31±1 GHz channels, respectively.

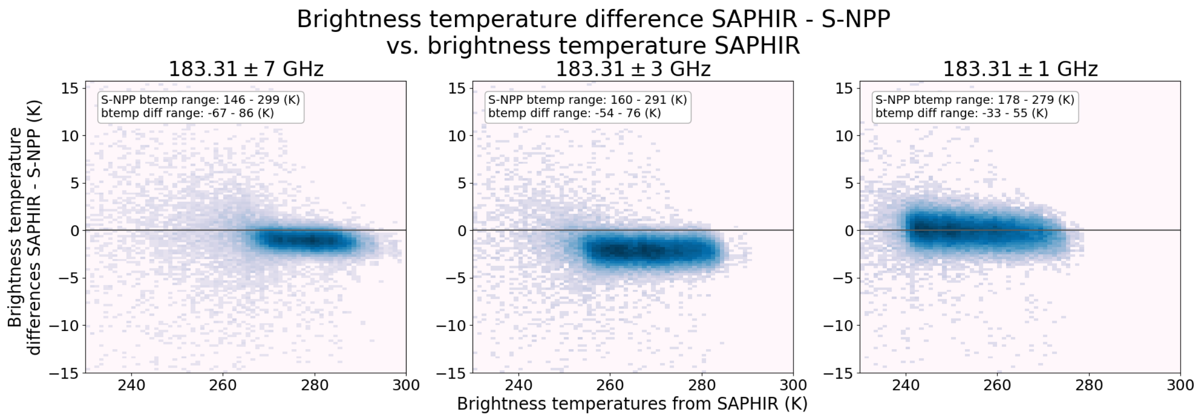

Figure 72: SAPHIR minus S-NPP BT differences vs SAPHIR BTs

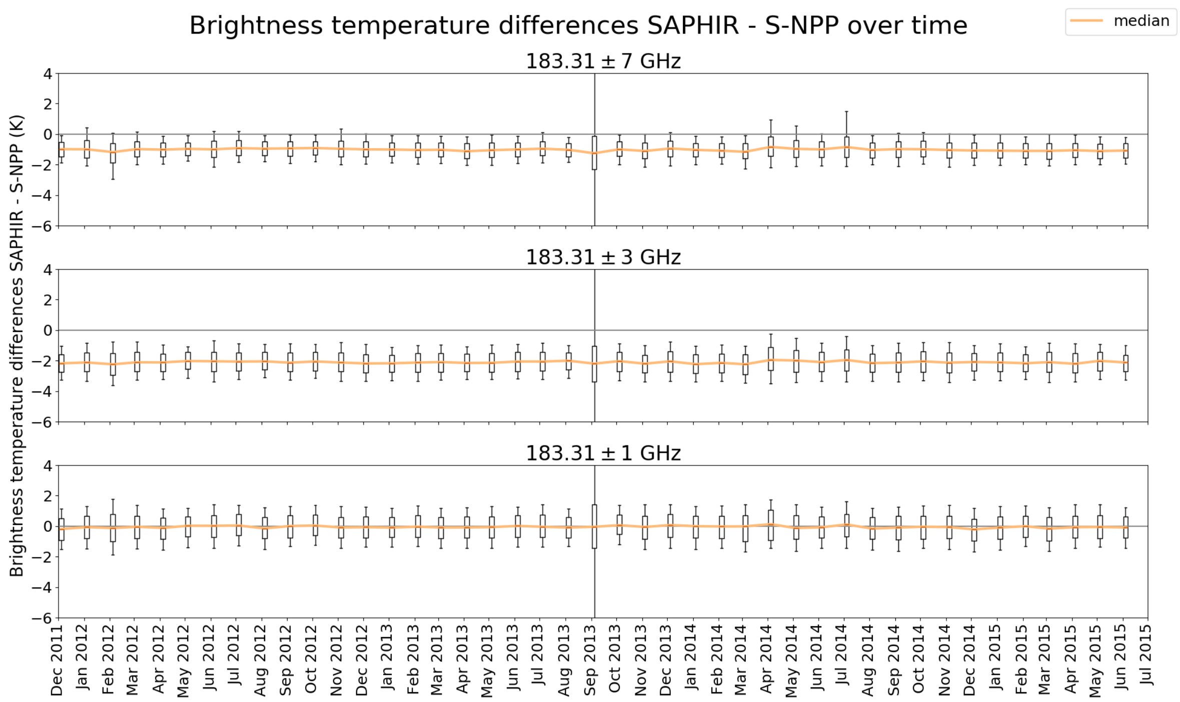

Figure 73: Box-plot time series of monthly BT differences for S-NPP / SAPHIR SNOs

4.3.19 Summary

Table 3 shows a summary of the SNO analysis results. The summary includes a subjective assessment of the response of the BT differences to a reference BT, and whether there is variability in the BT differences with time. The objective of this summary is to provide a quick reference to the main results of the SNO analysis performed in this study, particularly when considering the overall conclusions of the work. See the text in the preceding sections for context.

Table 3: Summary of the SNO analysis results. Column 'StDev' is the standard deviation of the BT differences. Column 'BT Res.' Indicates whether there is a response in the BT differences to a reference BT with 'W' indicating weak, 'M' indicating moderate and 'S' indicating strong. Column 'Time Var.' Indicates whether there is variability in the BT differences with time with 'W' indicating weak, 'M' indicating moderate and 'S' indicating strong. Column 'Loc' indicates the geographical location of the SNOs: global (G), tropical (T), or polar (P). The mean differences are calculated

according to the formula: sensor_column1-sensor_column2. 'N' stands for None and means there is no dependence either with BT nor with time.

Channel Sensor Pair		183.31±7 GHz Mean (K)	183.31±7 GHz StDev (K)	183.31±3 GHz Mean (K)	183.31±3 GHz StDev (K)	183.31±1 GHz Mean (K)	183.31±1 GHz StDev (K)	BT Res.	Time Var.	Loc
Metop A	FY-3A	0.6	3.2	1.9	2.3	1.7	1.5	M	W	G
Metop B	FY-3A	1.1	1.9	1.8	1.8	1.8	1.1	M	-	G
SAPHIR	FY-3A	0.9	3.1	-0.4	2.3	2.0	1.8	W	W	T
S-NPP	FY-3A	1.1	2.0	2.5	1.5	1.4	1.3	M	-	P
FY-3C	FY-3B	0.8	2.6	3.9	1.4	0.8	1.2	M	M	P
Metop A	FY-3B	3.6	3.0	3.1	1.9	2.7	1.3	S	M	P
Metop B	FY-3B	3.7	2.9	3.2	1.9	2.8	1.3	S	M	P
SAPHIR	FY-3B	-0.8	4.3	0.0	3.1	2.6	2.3	M	N	T
S-NPP	FY-3B	4.1	3.7	4.4	2.5	3.6	2.5	S	M	P
Metop A	FY-3C	2.7	1.4	-0.7	1.4	1.7	1.0	W	W	P
Metop B	FY-3C	2.7	2.0	-0.6	1.7	1.8	1.0	W	W	P
SAPHIR	FY-3C	2.1	3.5	-1.9	2.5	2.3	1.7	W	N	T
S-NPP	FY-3C	2.1	3.7	-0.5	2.3	1.9	3.5	W	M	P
SAPHIR	Metop A	-0.3	2.4	-0.9	1.7	0.9	1.5	W	N	T
S-NPP	Metop A	0.0	1.2	0.8	1.0	0.7	1.2	W	N	P
SAPHIR	Metop B	-0.2	2.6	-0.8	2.0	0.9	1.8	W	N	T
S-NPP	Metop B	0.0	1.2	0.8	1.1	0.6	1.2	N	W	P
SAPHIR	S-NPP	-1.0	2.7	-2.1	2.1	-0.1	1.7	N	N	T

4.4 Assessment of uncertainties and quality flags

4.4.1 Uncertainties

4.4.1.1 Time series analysis

Figure 74 shows the time series of percentiles of the global daily uncertainties for the 183.31±1 GHz channel on FY-3A. The plot shows that the common uncertainty is ~0.1 K for the whole period and for the whole globe, as the individual percentiles are not visible. By contrast, the other uncertainty components show some variability, which is as expected. For the independent uncertainty, the percentiles vary between ~0.7 and ~0.8 K, while the structural uncertainty varies between ~0.2 and ~0.3 K. The total uncertainty ranges between ~0.8 and ~0.9 K. The equivalent results for the 183.31±3 GHz channel are very similar, but the results for the 183.31±7 show some differences (Figure 75). Before mid-2011 in particular, there is a lot of variability in the independent and structured uncertainty components, which vary between about 0.6 and 1.6 K and 0.2 and 0.8 K, respectively. After 2011, the time series becomes more stable, and the ranges for this uncertainty components are ~0.6 to ~0.8 K and ~0.2 to ~0.3 K, respectively. Consequently, the total uncertainty ranges between ~0.6 and ~1.6 K before mid-2011, and ~0.6 and ~0.9 K after mid-2011. In summary, the values of the uncertainty components show very little variability, apart from the 183.31±7 GHz channel before mid-2011, and suggest that the total uncertainty on the data from any channel is typically <1 K.

Similar results are obtained for the other sensors. For FY-3B and -3C (not shown) the common uncertainty is also ~0.1 K. The structured and independent uncertainties are slightly smaller than for FY-3A, ranging between about 0.3 to 0.6 K for the independent uncertainties, 0.1 to 0.2 K for the structural uncertainties, and 0.3 to 0.7 K for the total uncertainties. In this case, the 183.31±1 GHz channel has the higher uncertainties, while the 183.31±3 GHz and 183.31±7 GHz channels, and the intermediate channels for FY-3C, are at the lower end. It is worth noting that the FY-3C 183.31±3 GHz, 183.31±4.5 GHz and 183.31±7 GHz channels exhibit higher noise in the independent and structured uncertainties prior to 2016. This is consistent with the earlier findings in this study that note a more stable and less noisy agreement between the FCDR and operational data from 2016 (Section 4.1.6).

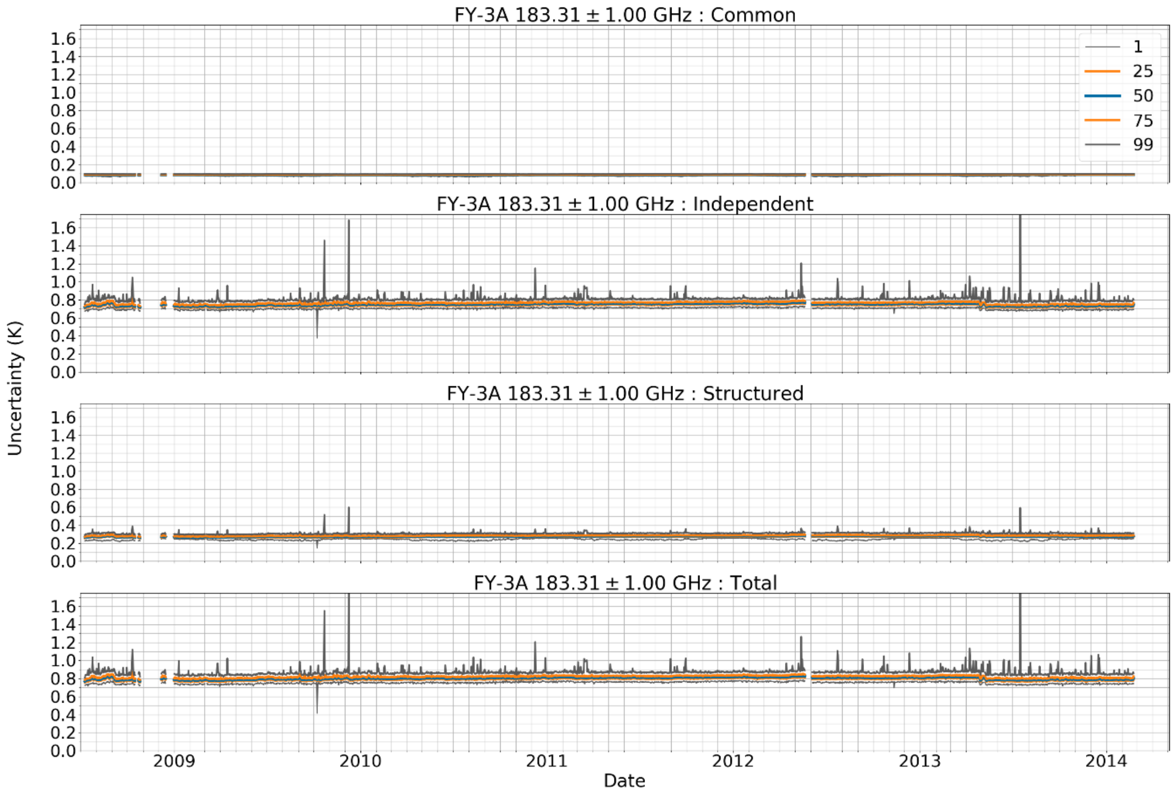

Figure 74: Time series of percentiles for global distribution of uncertainties for the FY-3A 183.31±1 GHz channel

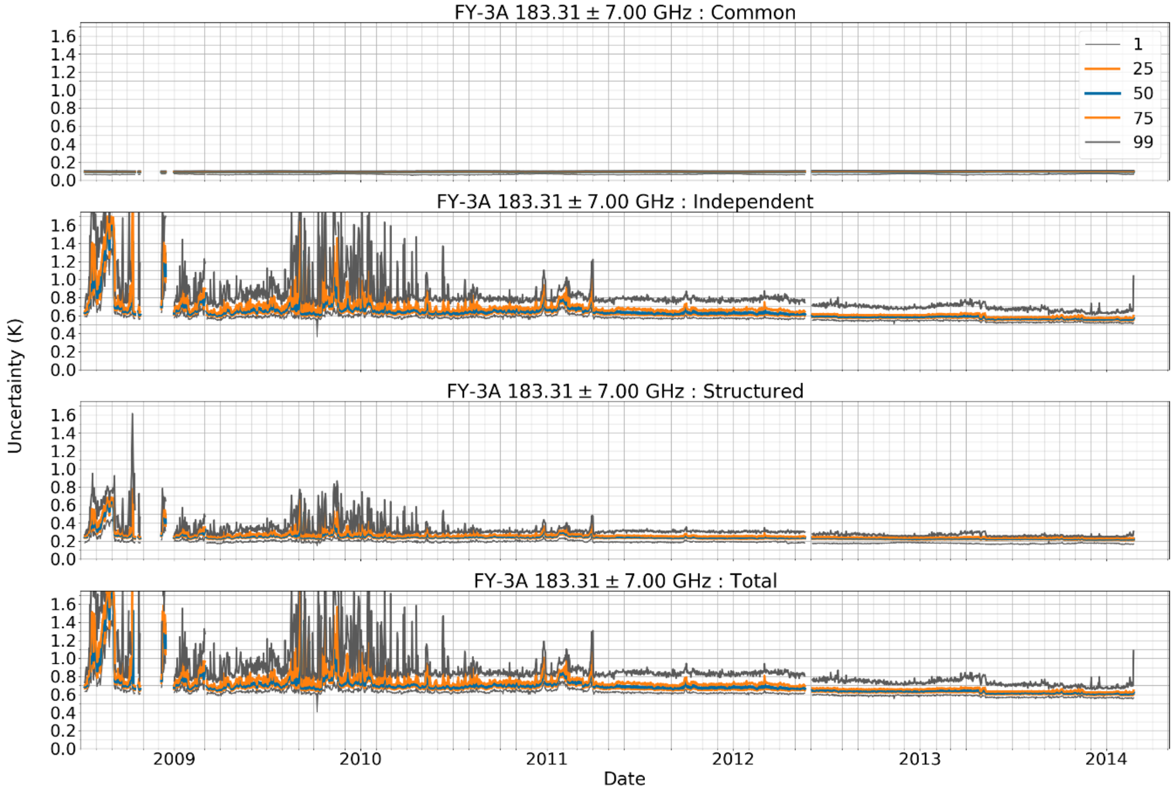

Figure 75: Time series of percentiles for global distribution of uncertainties for the FY-3A 183.31±7 GHz channel

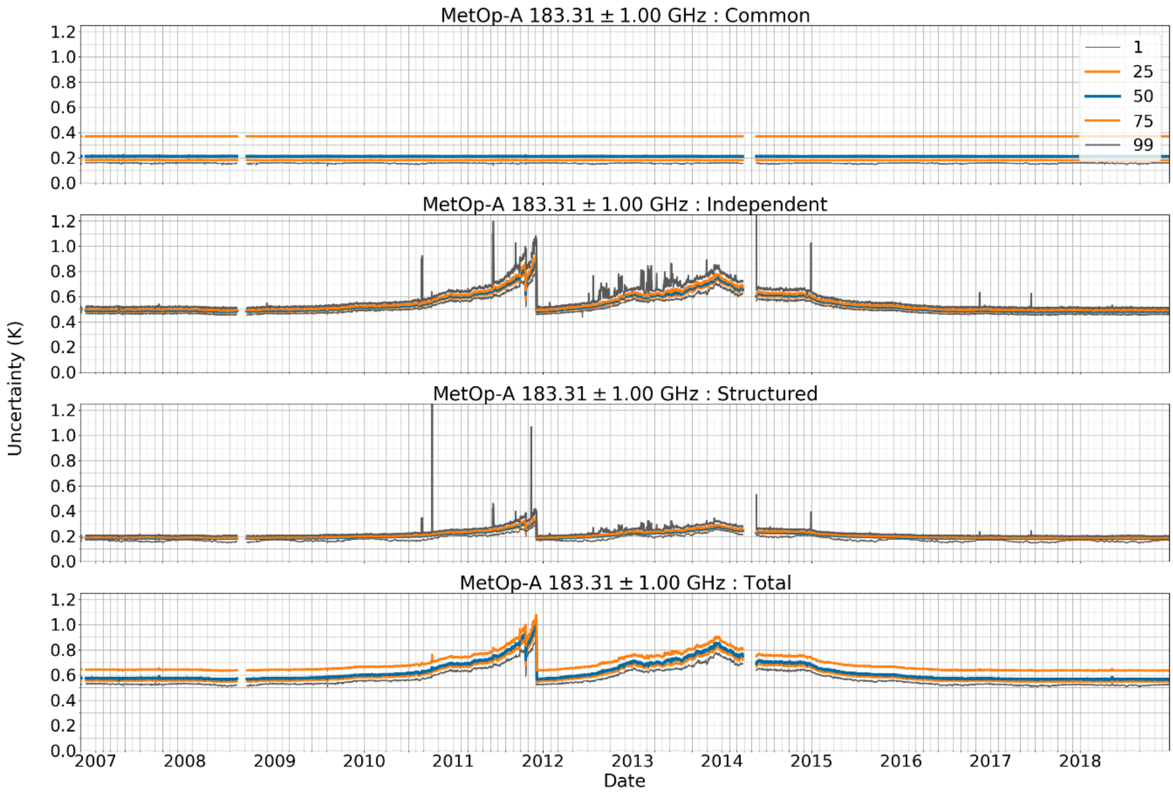

Figure 76: Time series of percentiles for global distribution of uncertainties for the Metop A 183.31±1 GHz channel

For Metop A, there appears to be quite a wide range of uncertainty values applied to the data as and the individual percentiles are quite distinct (Figure 76). The common uncertainty ranges between about 0.2 and 0.5 K, which is different from the FY-3 instruments that had a fairly constant value of ~0.1 K. For the 183.31±1 GHz channel, the independent uncertainty is mostly ~0.5 K, but there is a period of significant variability between mid-2011 and 2015, where values range up to ~1.0 K. This period of variability is also present in the structured uncertainty, which has a typical value of ~0.2 K that rises to ~0.4 K in this period. This results in a total uncertainty that has the same pattern; values range between about 0.6 K and 1 K, with some variation around these values corresponding to the different percentiles. The results are similar for the 183.31±3 GHz channel, although the independent uncertainty values are approximately 0.1 K lower. The results for the 183.31±7 GHz channel do not show this anomalous period and have very stable independent (~0.3 K) and structured (~0.1 K) uncertainties. It is thought that this anomaly is due to the change in the local oscillator, that was switched to the back-up in 2011. A change in the uncertainties is also apparent in 2014, when the Metop A gain was adjusted.

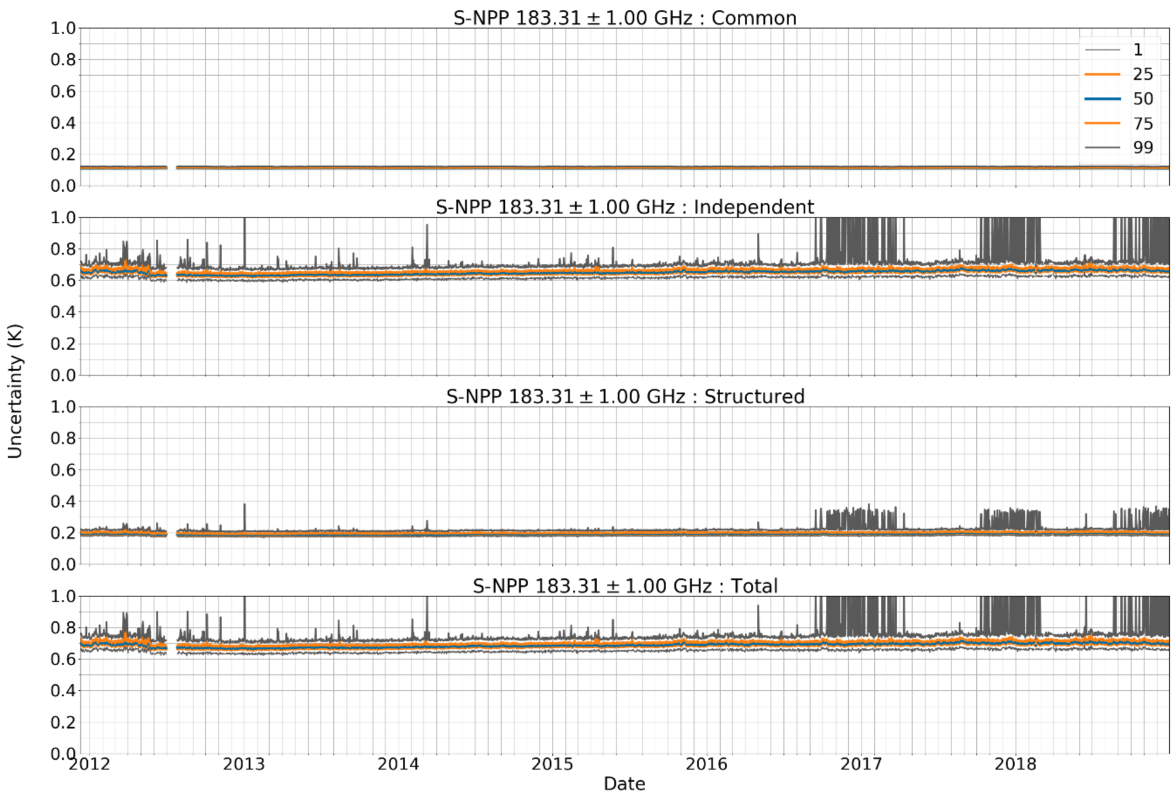

Figure 77: Time series of percentiles for global distribution of uncertainties for the S-NPP 183.31±1 GHz channel

The results for Metop B show no significant features in the time series of uncertainties and there is little distinction between the percentiles (not shown). The common uncertainty is ~0.2 K for all channels, the independent uncertainty ranges between ~0.3 and 0.4 K, while the structured uncertainty is ~0.1 to ~0.2 K. The total uncertainty budget is therefore ~0.4 to ~0.5 K.

The results for S-NPP are also quite stable, with only a small distinction between the percentiles for the upper-sounding channels (Figure 77). The common uncertainty is ~0.1 K for all channels, the independent uncertainty ranges between ~0.3 K and 0.7 K, and the structured uncertainty between ~0.1 K and ~0.2 K. This results in a total uncertainty of ~0.4 to ~0.7 K. The spread of the percentiles decreases with channel sounding height; the lowest sound channel (183.31±7 GHz), the percentile range is ~0.1 K. The time series exhibits increased noise during the first half of 2012, and then towards the end of 2016 there are three 'spikey' periods.

In general, the total uncertainties are <1 K, although 'baseline' values (i.e. outside of some of the anomalous periods discussed above) are typically <0.7 K. At this point, it is useful to reflect on results presented earlier in the study, particularly those from the SNO analysis. The average 'mean difference' ('standard deviation') between SNOs is 1.8 K (1.5 K), 1.7 K (1.7 K) and 1.9 K (2.4 K), for the 183.31±1 GHz, 183.31±3 GHz, and 183.31±7 GHz channels, respectively. The results of the SNO analysis therefore suggest that the FCDR uncertainties may be underestimated, particularly for the 183.31±7 GHz channel, which tends to have a much higher standard deviation. Additionally, for FY-3A, the differences between the FCDR and operational data are typically -0.5 to -1.5 K (FCDR minus operational) with standard deviations of up to ~1.7 K, which supports this conclusion. Similarly, the mean differences between the S-NPP and operational FCDR are ~1.5 K. The double differences for the different FY-3 FCDRs reported in Section 4.2.2.2, which are typically differ from Metop A / S-NPP by 1-4 K, also suggest the uncertainties for these instruments may be significantly underestimated.

4.4.1.2 Uncertainty overlap analysis

The uncertainty overlap analysis, defined in Section 1.7, has been performed for all SNOs, excluding those with SAPHIR, as the SAPHIR data set does not include any uncertainties. Only total uncertainties in the FCDR are considered (Section 3.2.4). Figure 78 shows the histograms for the SNO uncertainty analysis for FY-3B and FY-3C. The distributions for the 183.31±3 GHz and 183.31±1 GHz channels are closer to Laplacian than Gaussian distributions, while the distribution for the 183.31±7 GHz channel is quite asymmetrical. All the distributions have positive skew and have standard deviations that are much greater than one, which implies that one or both of the FY-3B and FY-3C uncertainties are underestimated. The 183.31±1 GHz channel has the smallest standard deviation (2.2 K compared to ~3.5 K for the other two channels), suggesting that the uncertainties for this channel are better defined.

Figure 79 shows the distributions for another SNO pair, Metop A and S-NPP. These distributions are also approximately Laplacian but are generally more symmetrical and narrower than those for FY-3B and FY-3C. For all three channels, the standard deviations are greater than one, again implying that the uncertainties are underestimated, but they are smaller than those for the FY-3B and FY-3C pairing. The standard deviation for the 183.31±1 GHz and 183.31±3 GHz channels is 1.4 K, while the 183.31±7 GHz channel again has the largest standard deviation of 2.2 K.

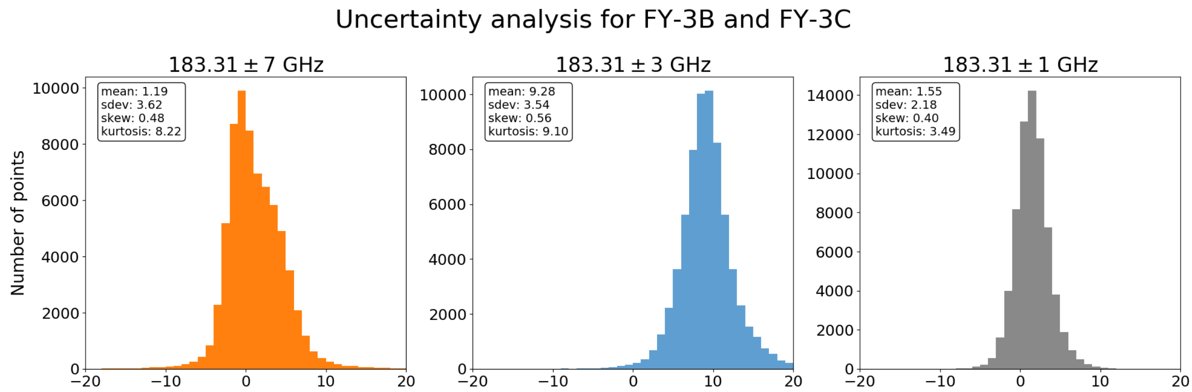

Figure 78: Distribution for the uncertainty overlap analysis for FY-3B and FY-3C

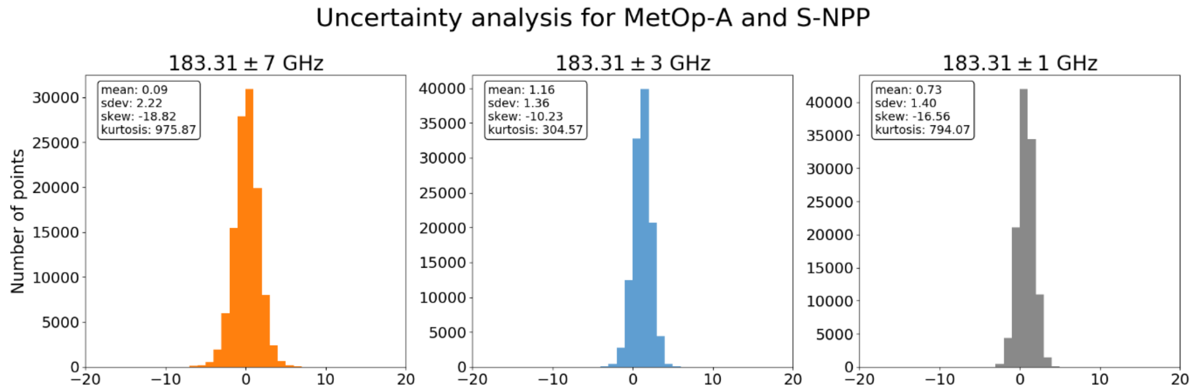

Figure 79: Distribution for the uncertainty

This pattern, where the 183.31±7 GHz channel has the highest standard deviation, is typical for all SNOs, except for those pairs that include FY-3A (Table 4). For the FY-3A pairing, it is always the 183.31±3 GHz channel that has the highest standard deviation. For SNO pairs that do not include FY-3A, the 183.31±1 GHz channel has the lowest standard deviation. All SNO pairs have standard deviations that are greater than one, which suggests that all the FCDRs have uncertainties that are underestimated for at least the geographical regions where the SNO analysis has been performed. For all SNO pairs, except Metop A/FY-3A and Metop B/FY-3A which have global coverage, these regions are near the poles.

Table 4: Standard deviations for the uncertainty overlap analysis. Standard deviations shown in bold represent the channel with the highest standard deviation for a particular SNO pair. Standard deviations that are underlined represent the channel with the lowest standard deviation. The 'Coverage' column indicates whether the SNO analysis is global (G) or near the poles (P). The 'Mean' column represents the mean standard deviation for each sensor pair across all channels.

Sensor Pair		Coverage	183.31±7 GHz (K)	183.31±3 GHz (K)	183.31±1 GHz (K)	Mean (K)
Metop A	FY-3A	G	2.5	3.1	2.2	2.6
Metop B	FY-3A	G	1.5	3.1	2.2	2.3
S-NPP	FY-3A	P	1.5	2.4	2.0	2.0
FY-3C	FY-3B	P	3.6	3.5	2.2	3.1
Metop A	FY-3B	P	3.9	3.8	2.3	3.3
Metop B	FY-3B	P	4.0	4.2	2.5	3.6
S-NPP	FY-3B	P	5.1	4.9	3.7	4.6

Sensor Pair		Coverage	183.31±7 GHz (K)	183.31±3 GHz (K)	183.31±1 GHz (K)	Mean (K)
Metop A	FY-3C	P	2.7	2.4	1.2	2.1
Metop B	FY-3C	P	4.4	3.4	1.5	3.1
S-NPP	FY-3C	P	7.0	3.8	4.0	4.9
S-NPP	Metop A	P	2.2	1.4	1.4	1.7
S-NPP	Metop B	P	2.4	1.7	1.4	1.8

Notably, the mean standard deviations are best for SNO pairs that include only S-NPP and Metop A or B. This is consistent with the general findings in this study that suggest that the FCDRs for these instruments are the most stable and have the best agreement with the reference data sets used in this study. Although analysis indicates that the uncertainties are underestimated for all instruments, the higher standard deviations for the FY-3 pairings suggests that the uncertainties are particularly underestimated for these sensors. The worst mean standard deviations generally include FY-3B in the SNO pairing, which suggests that the underestimate may be strongest for FY-3B, although the S-NPP and FY-3C pair also perform poorly. Note, that the match-up uncertainty arising from scene inhomogeneity is considered to be zero in this study.

4.4.2 Quality flags

4.4.2.1 Maps

Figure 80 shows an example map that summarises the QC flags, i.e. where any of the four QC flags are set (see Section 2.4), from the Metop A FCDR on 15 October 2009. The map shows many scan lines are flagged and that these occur in all regions. Figure 81 shows the same map, but just for the data quality QC flag, which shows a much smaller number of pixels or scanlines that are flagged by this QC information. Figure 82 and Figure 83 show the equivalent maps for the quality issue and quality pixel QC flags, which show a greater number of data that have been flagged. By contrast, no data are flagged for this example in the quality scanline QC information (Figure 84). These maps show that many of the Metop A data are flagged with issues on this day, most of which are from the quality issue and quality pixel QC information.

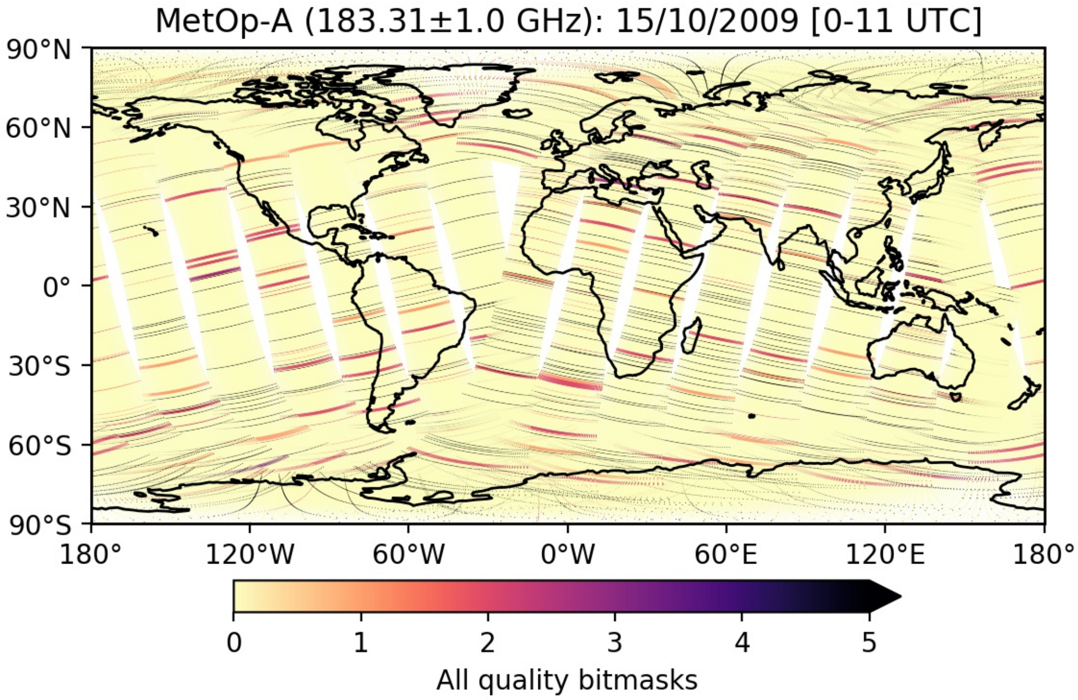

Figure 80: Map of Metop A QC summary flags for the 183.31±1 GHz channel

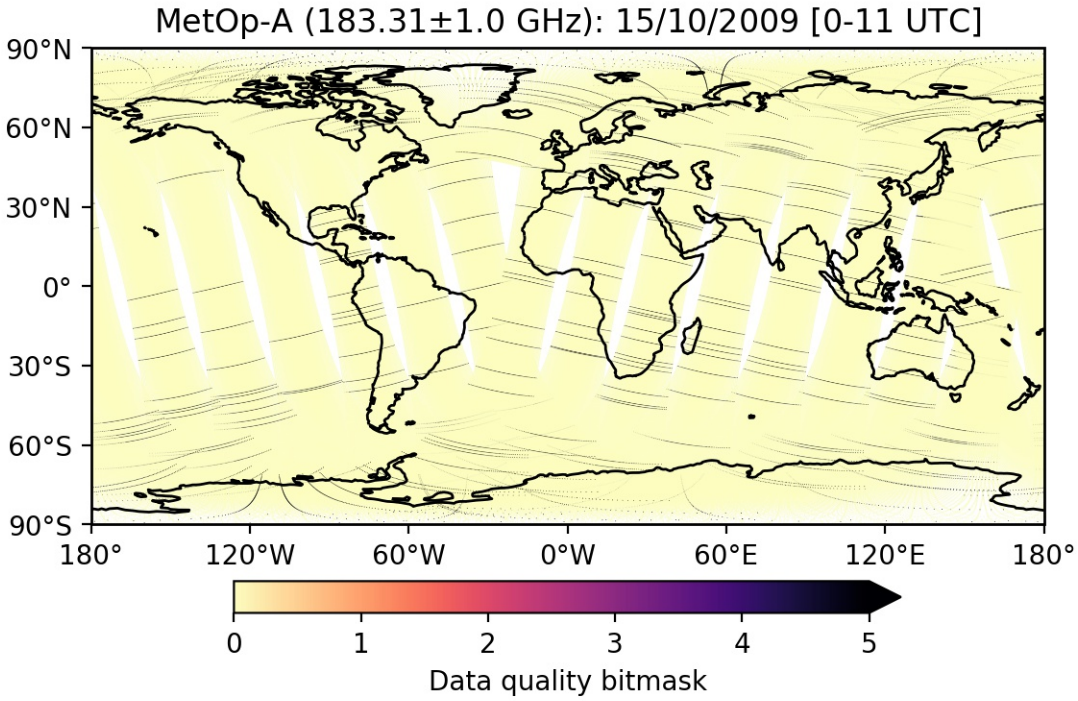

Figure 81: Map of Metop A data quality QC flag for the 183.31±1 GHz channel

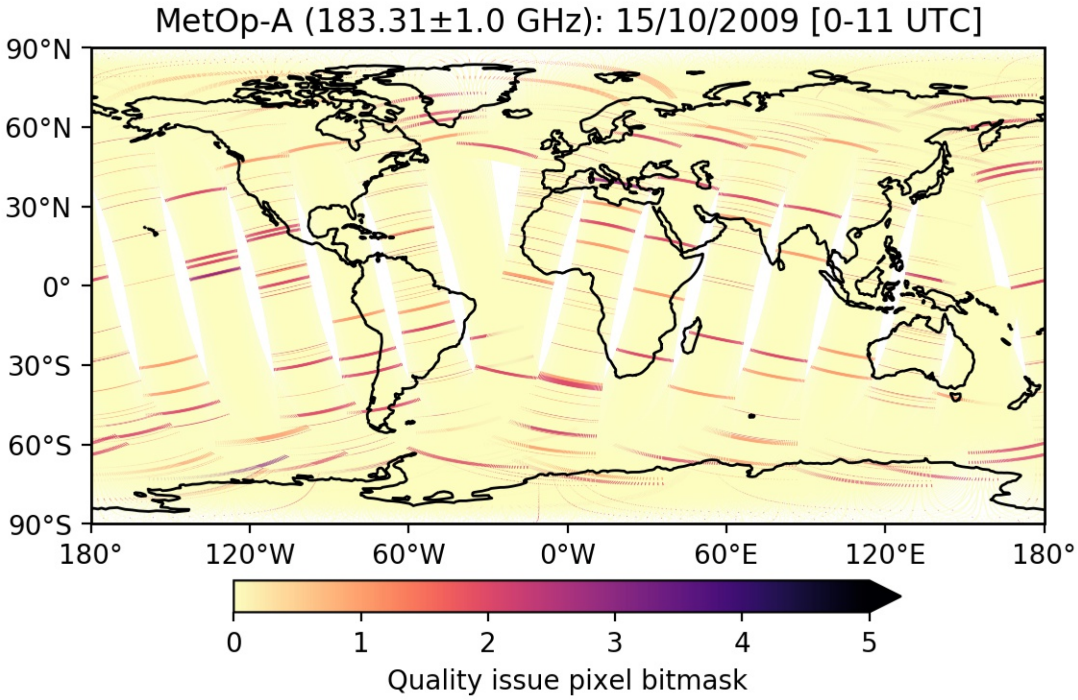

Figure 82: Map of Metop A quality issue pixel QC flag for the 183.31±1 GHz channel

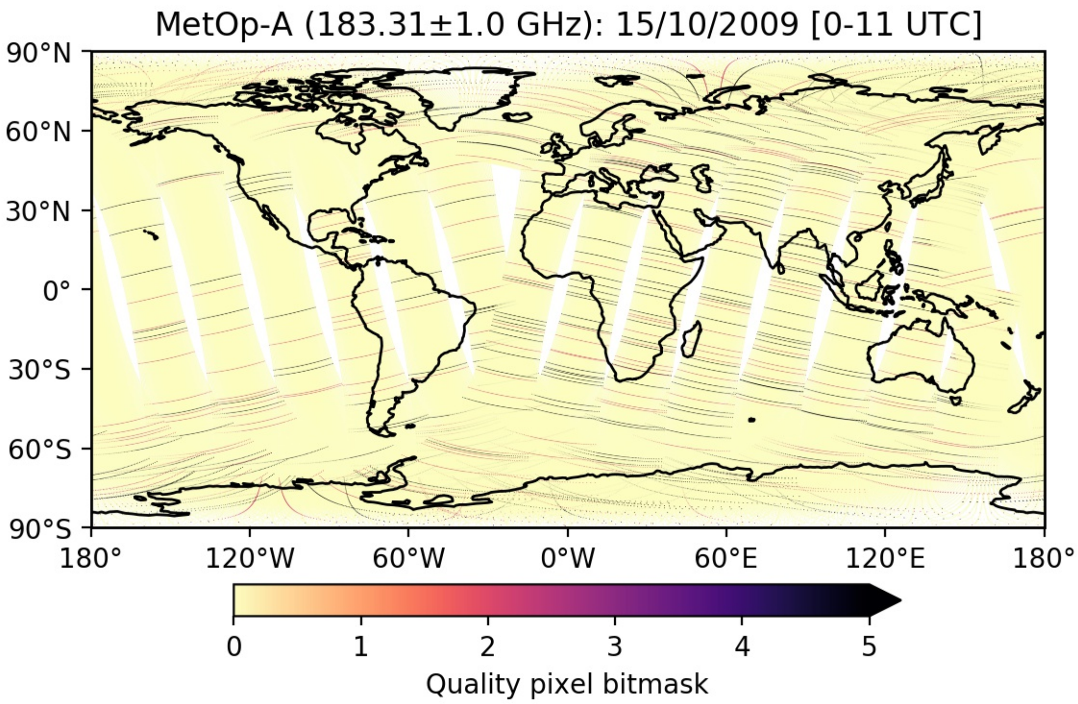

Figure 83: Map of Metop A quality issue QC flag for the 183.31±1 GHz channel

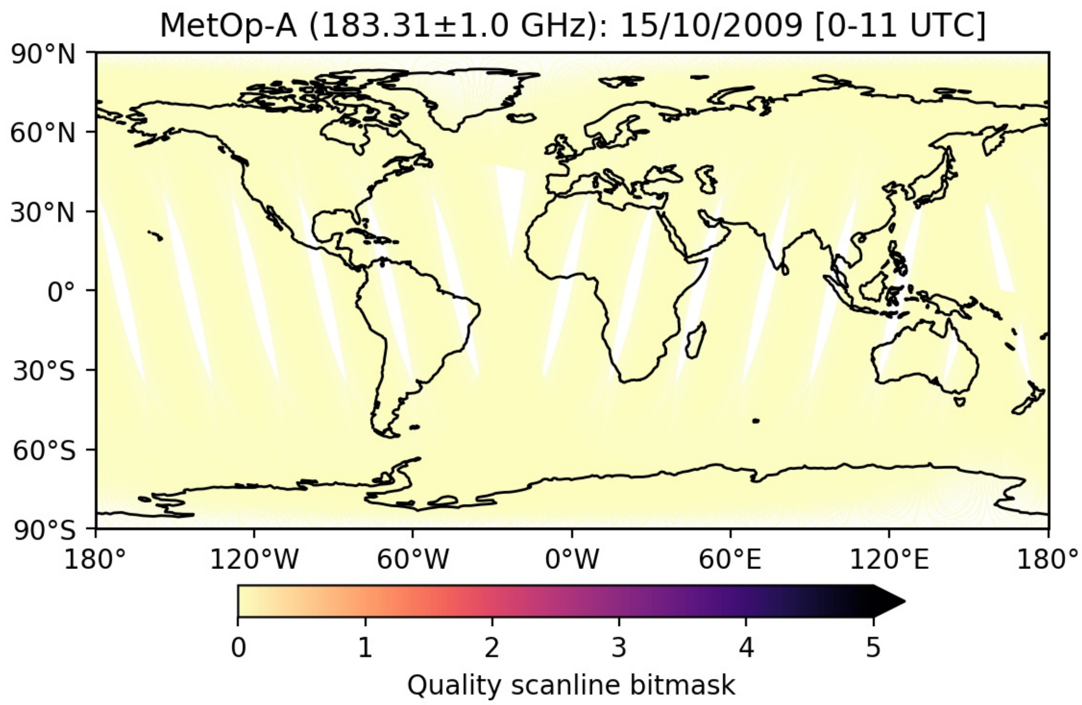

Figure 84: Map of Metop A quality scanline QC flag for the 183.31±1 GHz channel

4.4.2.2 Time series analysis

Evidence presented in Section 4.4.2.1 suggests that the QC tests are flagging a significant number of data and that flagged data occur in all regions. Figure 85 shows the FCDR minus operational time series for FY-3B where the QC information has not been applied to the data. There are several spikes present, including a large spike at the end of 2016, and a period of higher noise from mid-2013 to 2015. Figure 86 shows the same time series, but with the QC information applied. The large spike at the end of 2016 has been removed, and the noise in the period from mid-2013 to 2015 is reduced. The noisy data at the start of the record has also been removed.

The positive effects from applying the QC information can also be seen in the time series of the standard deviations of the mean differences between the FCDR and operational data sets. Figure 87 shows the time series of standard deviations for FY-3B without QC information applied to the data. The time series is full of high-standard deviation spikes and is generally very noisy. Figure 88 shows the same time series, but with the QC information applied. There are still many spikes present in the data, but the number and value (i.e. standard deviation that they represent) of the spikes is significantly reduced, making the time series much less noisy. Results are similar for the other sensors comprising the suite of FCDRs (not shown).

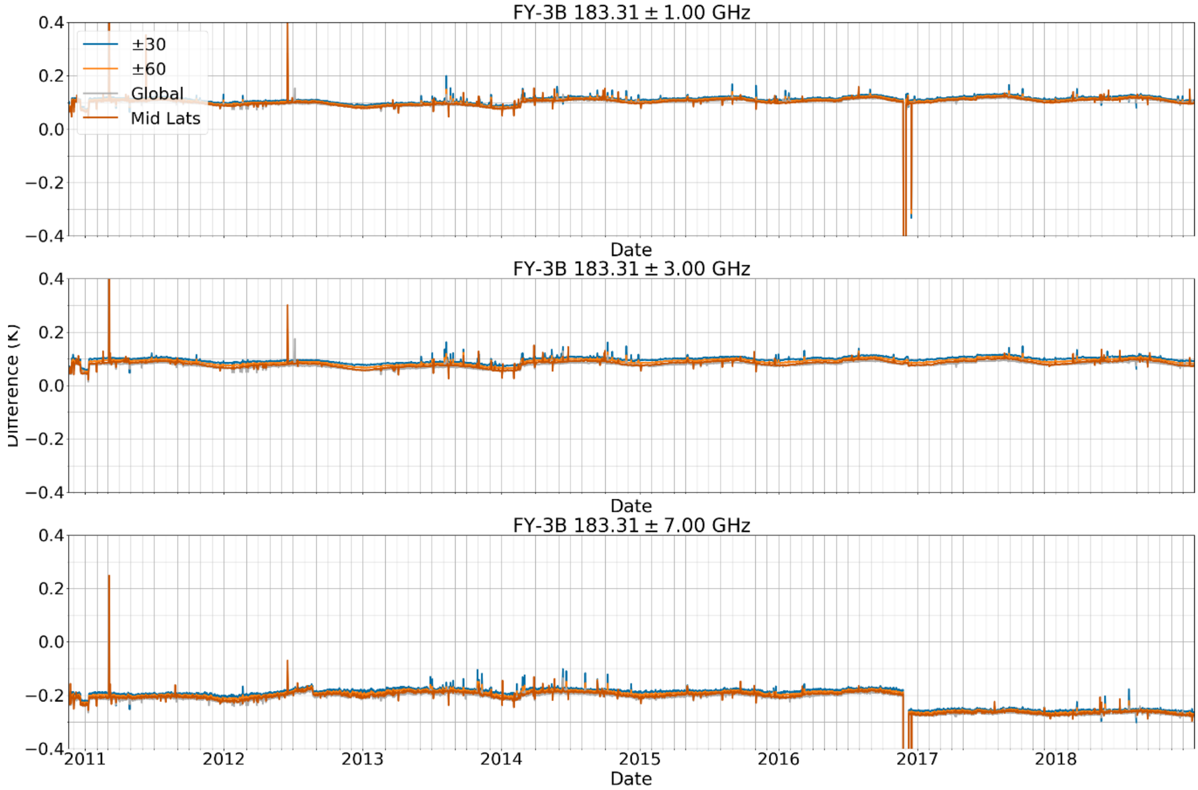

Figure 85: Time series of daily mean differences (FCDR minus operational) for FY-3B with no QC applied.

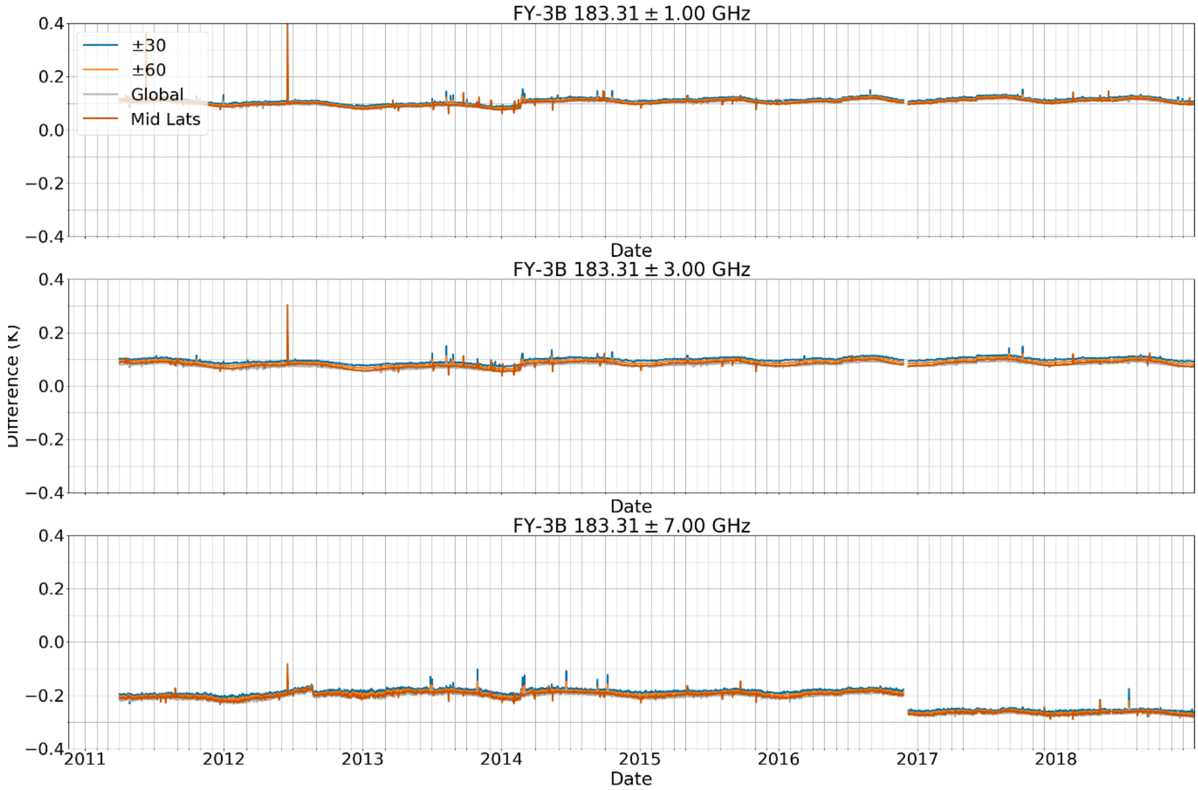

Figure 86: Time series of daily mean differences (FCDR minus operational) for FY-3B *with* QC applied.

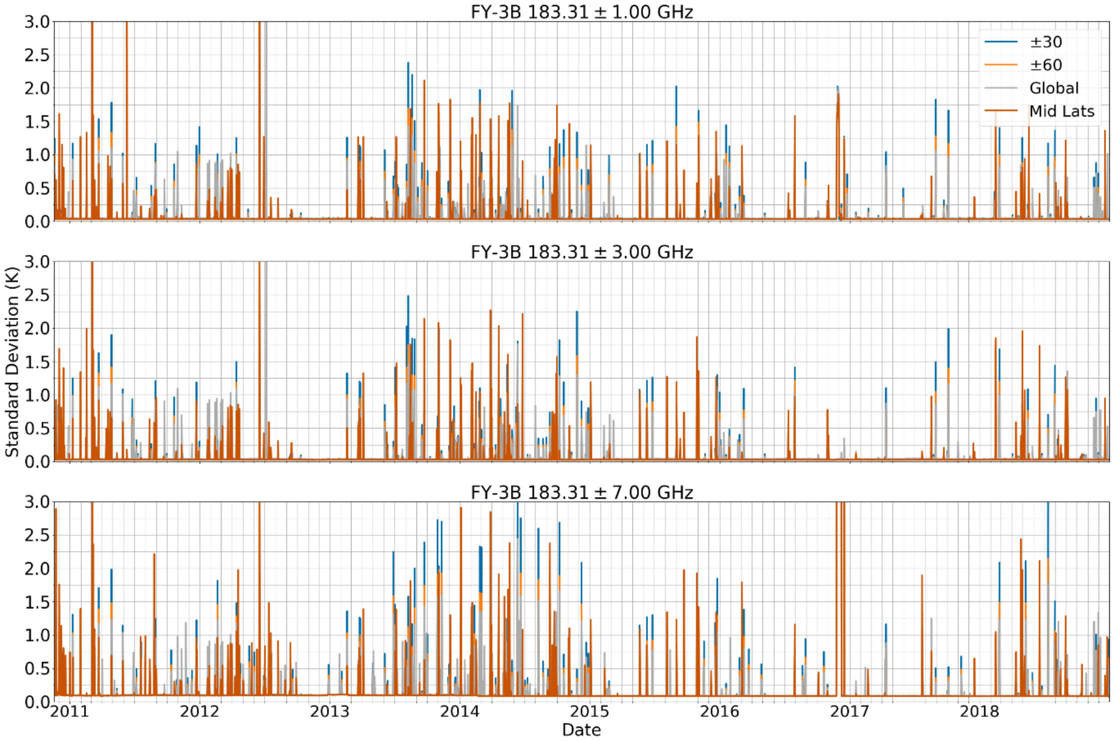

Figure 87: Time series of the standard deviations of the mean differences (FCDR minus operational) without QC information applied to the data for FY-3B.

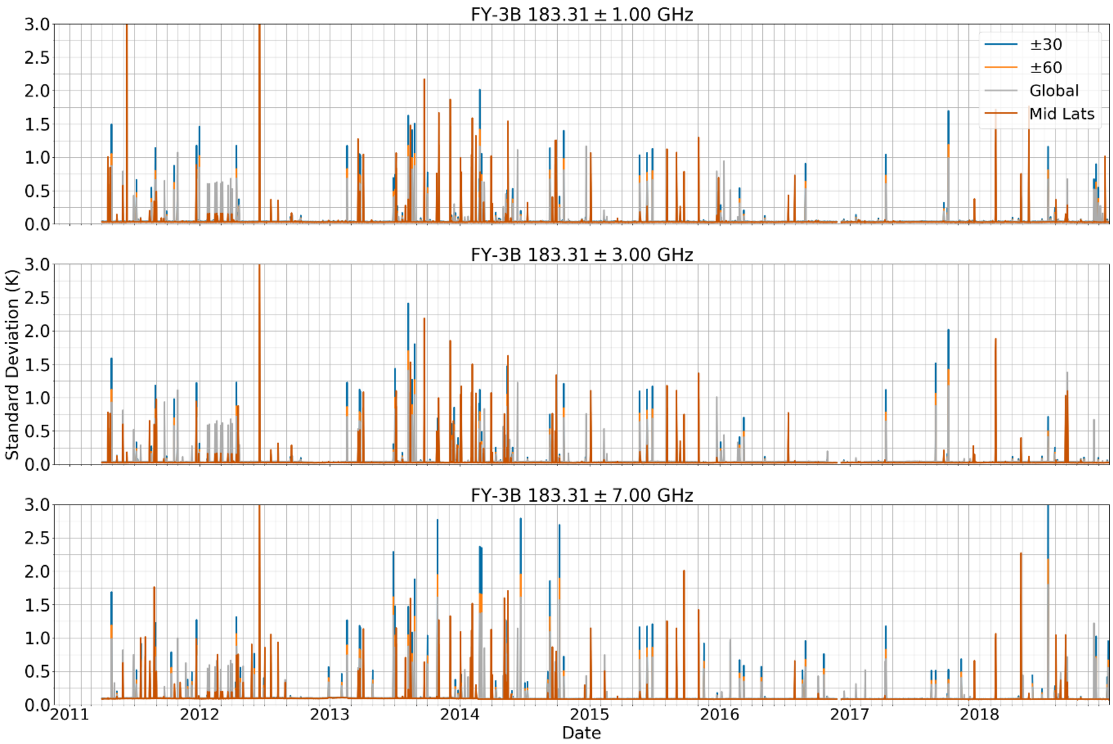

Figure 88: Time series of the standard deviations of the mean differences (FCDR minus operational) *with* QC information applied to the data for FY-3B.

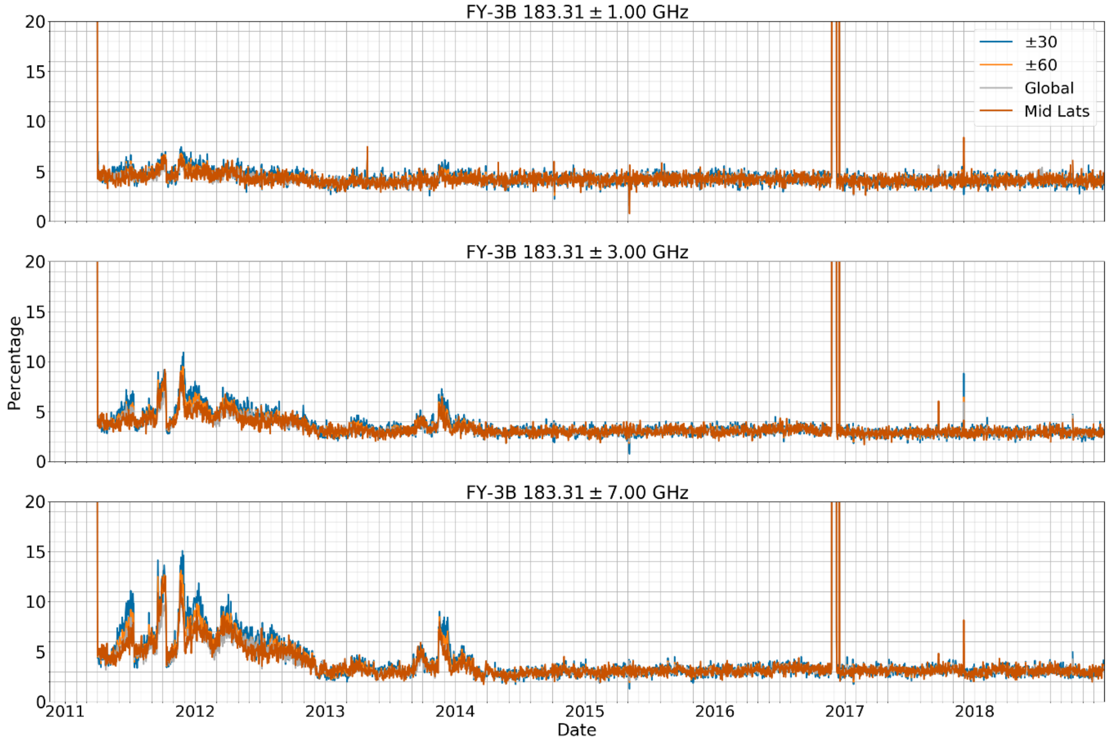

Figure 89: Time series of the % of pixels removed each day using the QC information for FY-3B.

4.4.2.3 Time series of percentage of data removed using QC information

Figure 89 shows the FY-3B time series of the % of the total data removed each day through application of the QC information. The plot suggests that this screening removes <5% of the data, which is considered to be a reasonable proportion of the data.

The QC % time series for the other instruments comprising the FCDR suite present similar results. For FY-3A, the percentage of pixels removed by applying the QC information is typically between 5 and 10% for the 183.31±1 GHz and 183.31±3 GHz channels, but often <5% for the 183.31±7 GHz channel (not shown). Interestingly, the September 2013 discontinuity is evident in these time series (see Section 4.1.4), where the % increases gradually just before this event, then there is a negative step change in the % during the event. For FY-3C, the percentage of pixels removed by applying the QC information is very stable and ~5% for the whole time series. The results for Metop A are also quite stable and are ~8% for all channels. Notably, there are discontinuities in the QC % time series towards the end of 2011 and in early 2018 that are not present in the comparisons with the operational data (Section 0) or simulations (Section 4.2). For Metop B, the time series of the percentage of data removed by the QC process has a slightly positive trend, ranging between about 6 and 8 % (Figure 90). There are also discontinuities that are present in mid-2014 and late 2015, which are again not present in the comparisons with the operational data (Section 0) or simulations (Section 4.2). The results for S-NPP suggest that this FCDR has the fewest number of data that are rejected by applying the QC information. This is a very stable ~3% for all channels, apart from the first few months of 2016, when the number of data rejected rises to >20%. This is coincidental with the change between the FCDR and operational data that occurs at high-latitudes noted previously (Section 0).

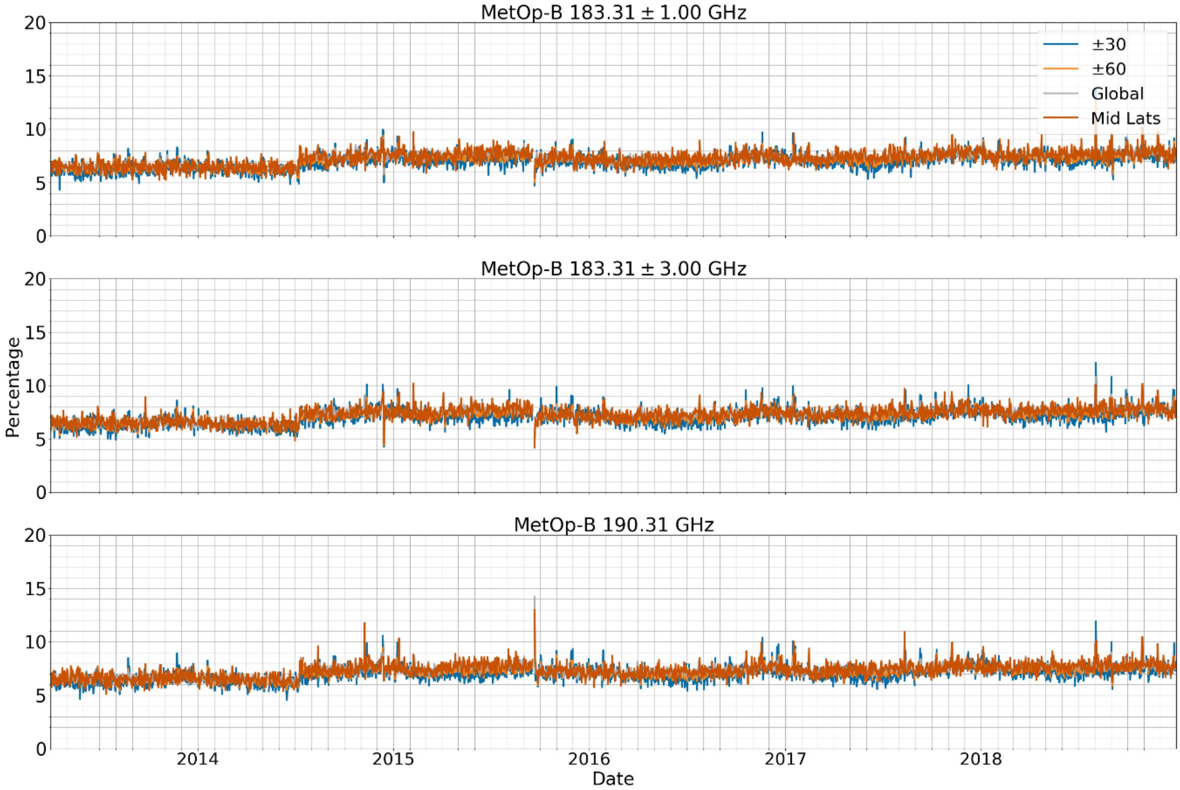

Figure 90: Time series of the % of pixels removed each day using the QC information for Metop B.

5 Conclusions

The results of the quality evaluation suggest that the FCDRs for Metop A and B are of very good quality. Through the analysis of double-differences calculated using simulated TOA BTs based on ERA-5 atmospheric profiles, Metop A and B are found to agree almost exactly, except for the 183.31±7 GHz channel where there is a small offset (~0.1 K). Virtually no difference is found between the Metop B FCDR and the operational Metop B L1c product with differences amounting to just a few hundredths of a K, which are negligible. The agreement between the Metop A FCDR and operational L1c product is also very good, with both mean differences and standard deviations of ~0.1 K (FCDR warm). The exception is an approximately 3-month period in 2007 where the FCDR exhibits a negative step-change with respect to the operational data. This negative step is due to wrong cold and warm target correction values assigned in the level 1b files archived at NOAA CLASS, but not in the FCDR, which is further confirmed through the comparison between the FCDR and simulated BTs where the step-change is absent. Both Metop A and B show remarkable stability with respect to the simulated BTs and show no discontinuities.

The results for S-NPP are also very good. However, there appears to be a very stable but systematic difference between the FCDR and the operational BTs of ~1.4 to ~1.9 K (FCDR warm), depending on the channel. There is also a significant discontinuity at high latitudes in 2016, where the mean difference between the FCDR and operational data reduces, such that the global standard deviation increases from <0.5 K to nearly 3 K. However, analysis of the FCDR with respect to the simulations and operational data suggest that the FCDR is very stable with time at all latitudes between ±60°.

Both Metop A and B, and S-NPP demonstrate good agreement with the simulated data. Some non-zero and varying agreement between the FCDRs and the simulations is expected as differences in the atmospheric states between the ERA5 reanalysis and the true atmosphere observed by the satellites will also influence the agreement between the two data sets. However, mean differences between the FCDRs and the simulations between ±60° latitude (higher latitudes are excluded from this particular analysis) are generally within ±1 K with standard deviations of around 1.5 – 2.5 K. The analysis of double differences suggests that the zonal mean differences between Metop A and B, and S-NPP, do not exceed ±1 K. This apparently good agreement is also confirmed through the analysis of SNOs. There are no SNOs between Metop A and B, but the analysis of SNO pairs for Metop A or B with S-NPP indicates that the mean differences and their standard deviations are consistently within ~1 K. However, it should be noted that the SNO analysis for these sensor pairs is restricted to very high latitudes and do not include assessment of the S-NPP 183.31±1.8 GHz and 183.31±4.5 GHz channels. Collectively, these results suggest that the systematic difference between the FCDR and operational data for S-NPP may be due to the operational data, rather than the FCDR. An independent assessment of these FCDRs using SAPHIR further confirms the quality of the Metop A, B and S-NPP data. It should be noted that the frequencies of the SAPHIR channels do not match those of Metop A, B and S-NPP exactly, so some differences are expected. Mean differences are within ±1 K, with the exception of the SAPHIR/S-NPP 183.31±7 GHz channel (183.31±6.6 GHz for SAPHIR), which is -2.1 K, while the standard deviations range between 1.5 and 2.7 K.

The results of the quality evaluation suggest that the FY-3A, -B, and -C FCDRs are of lower quality than those for Metop A, B and S-NPP. However, this comparative low quality is likely to originate from the L1a data rather than from the processing of the FCDRs as there is good agreement between the FCDRs and operational data sets, particularly for FY-3B and FY-3C. For these sensors, the agreement between the FCDRs and operational data is generally within a few tenths of a K. However, for FY-3A, the FCDR is 0.5 to 1 K colder than the operational data, with standard deviations of up to ~1.7 K depending on the channel. Comparisons between the FY-3 FCDRs and simulations also present some large and variable differences. In particular, FY-3C is highly erratic for the 183.31±3 GHz and 183.31±4.5 GHz channels, with the mean differences varying by several K, and sometimes abruptly. The FY-3A and -3B time series are consistently cold compared with the simulated TOA BTs and have a tendency for negative trends with time. Despite these larger mean differences, it is notable that the standard deviations of the differences are quite stable and similar to those obtained for Metop A, B and S-NPP. All the FY-3 FCDRs have significant discontinuities. For FY-3A, these occur in 2009 and mid-2012. The FY-3A FCDR also exhibits discontinuities with respect to the operational data in May and September 2013. For FY-3B, there are discontinuities at the end of 2010, in the second half of 2011, the second half of 2012, and in the first half of 2017. There is also a discontinuity present in the FCDR with respect to the operational data at the end of 2016. For FY-3C, there are discontinuities in 2013, 2015 and in early 2018, in addition to a 6-month anomaly in the middle of 2016. A discontinuity with respect to the operational FY-3C data is also identified in mid-2014.

Results for the SNO analyses confirm the relatively poor performance of the FY-3 FCDRs compared with Metop A, B and S-NPP. SNO pairs that include one of the FY-3 FCDRs tend to have the largest mean differences and standard deviations. With exception of the 183.31±3 GHz channel on FY-3C, the SNO analysis indicates that FY-3A, -B and -C is generally warmer than Metop A, B and S-NPP (note that the 183.31±1.8 GHz and 183.31±4.5 GHz channels are not assessed in the SNO analysis). The FY-3 sensors also appear to have a BT-dependent response, i.e. for a 1 K change in true BT, the observed FY-3 BTs do not necessarily change by 1K; this issue seems to be particularly bad for the FY-3B. This response is frequently non-linear and appears in all three channels assessed. The SNO analysis also suggests some variation in the BT differences with time, which may, at least in part, be a result of this BT-dependent response. In some cases, this variation has an annual cycle. In general, Metop A, B and S-NPP do not appear to have a significant BT-dependent response and/or show a variation in BT differences with time, but there are exceptions. For example, there is a weak non-linear dependency of the S-NPP minus Metop B BT differences on S-NPP BT for both the 183.31±3 GHz and 183.31±7 GHz channels and the time series for these channels has an annual cycle with amplitude of 0.5-1 K.

A BT-dependent response in the FY-3A observations may also be responsible for the weak annual cycle in the agreement between the FY-3A FCDR and operational BT data. There are annual cycles in this comparison both in the NH tropics and at high latitudes that appear to be linked to changes in cloud frequency and intensity that are associated with the ITCZ and mid-to-high latitude storms. Similar annual cycles are also present in the comparisons between the FCDRs and simulated BTs, although these are much more intense with larger BT variations. In this case, any annual cycles may also be due to failures in the cloud screening applied to the observations to make them consistent with the simulated BTs. A decreasing intensity of these annual cycles with channel sounding height and a negative skew in the lowest-sounding channels at 183.31±7.0 GHz and 190.31 GHz is further evidence to support this theory. (BTs over clouds that are not transparent at 183 GHz are cold and will therefore result in a cold tail in the distribution of differences where cloud contamination occurs.)

One of the main benefits of the FCDRs over the operational L1c data is in the provision of well-characterised uncertainties and detailed QC information. The results of the quality evaluation suggest that the uncertainties on the FCDRs are underestimated, particularly for the FY-3 sensors. The total uncertainties for all the FCDRs are typically <1 K but based on the results presented here, it seems likely that the errors frequently exceed this limit. There are four separate QC flags provided with the FCDRs, each of which indicate different issues within the data. Using the QC flags together generally results in up to 10% of the data being rejected, although this varies slightly with each FCDR. Use of the QC flags appears to remove bad data points successfully, resulting in a data set with lower noise.

Finally, this report has highlighted anomalies in the FY-3 operational L1c data at 17°N and 0.5°N. These latitudes are where the operational data start and finish and where is also a disruption in the operational calibration, which is averaged over a sequence of scan lines. These anomalies are not present in the FCDR due to a more consistent approach to the averaging required for the calibration.

6 Appendix A: Supplementary figures for the FCDR BTs vs operational BTs

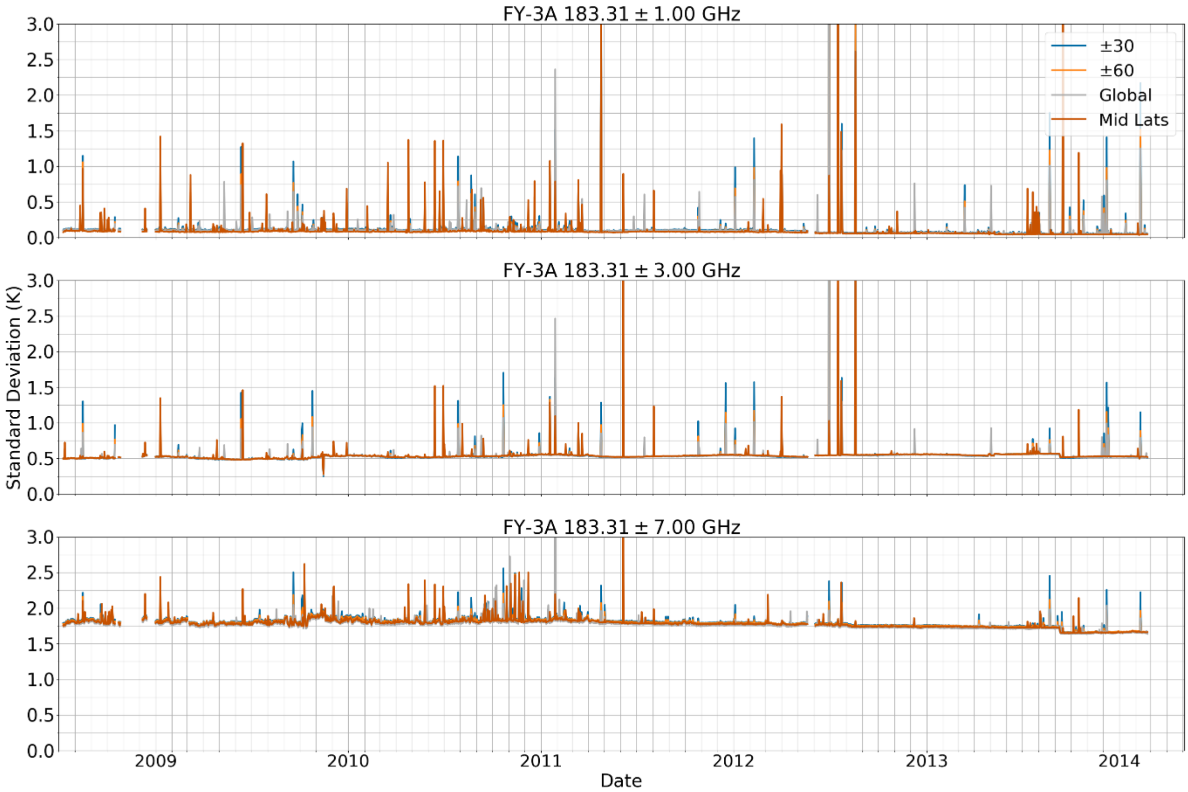

Figure A1: Time series of the standard deviation of differences for FY-3A (FCDR minus operational) for each of the 183 GHz channels. The data have been averaged over latitude bands ±30° (blue), ±60° (orange), ±90° (grey), and mid-latitude bands 60°S-30°S and 30°N-60°N (red).

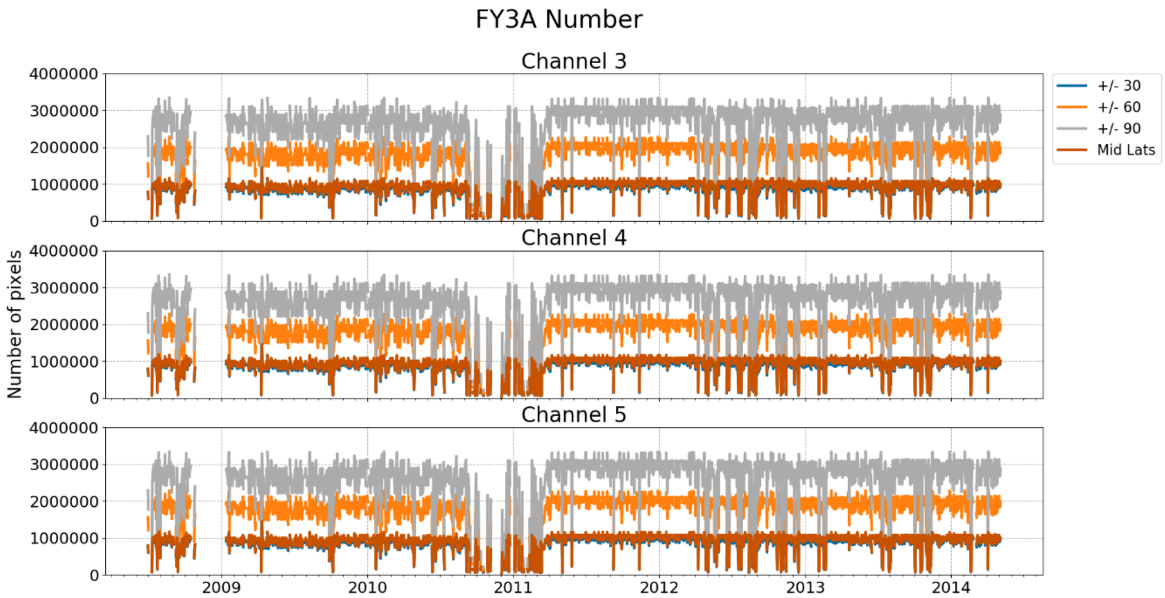

Figure A2: Time series of the number of available comparisons for FY-3A for each of the 183 GHz channels. The data have been averaged over latitude bands ±30° (blue), ±60° (orange), ±90° (grey), and mid-latitude bands 60°S-30°S and 30°N-60°N (red).

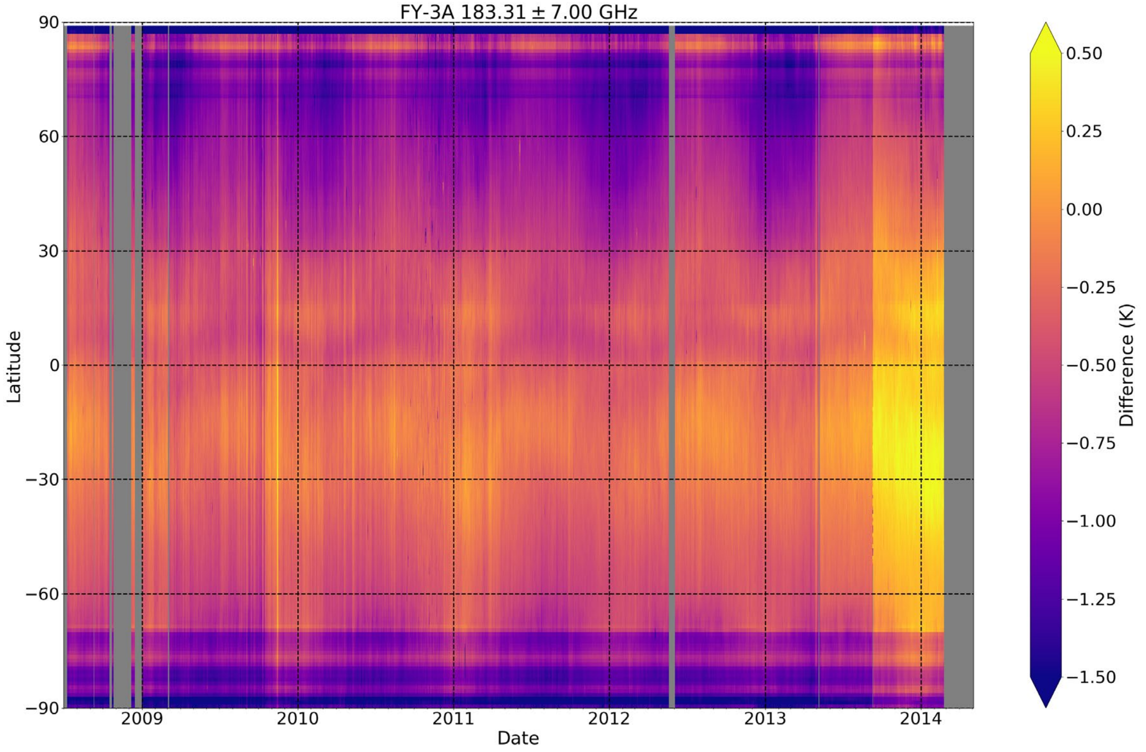

Figure A3: Hovmöller diagram showing the change in daily mean difference by latitude with time for FY-3A for the 183.31±7 GHz channel. The grey bands correspond to missing data.

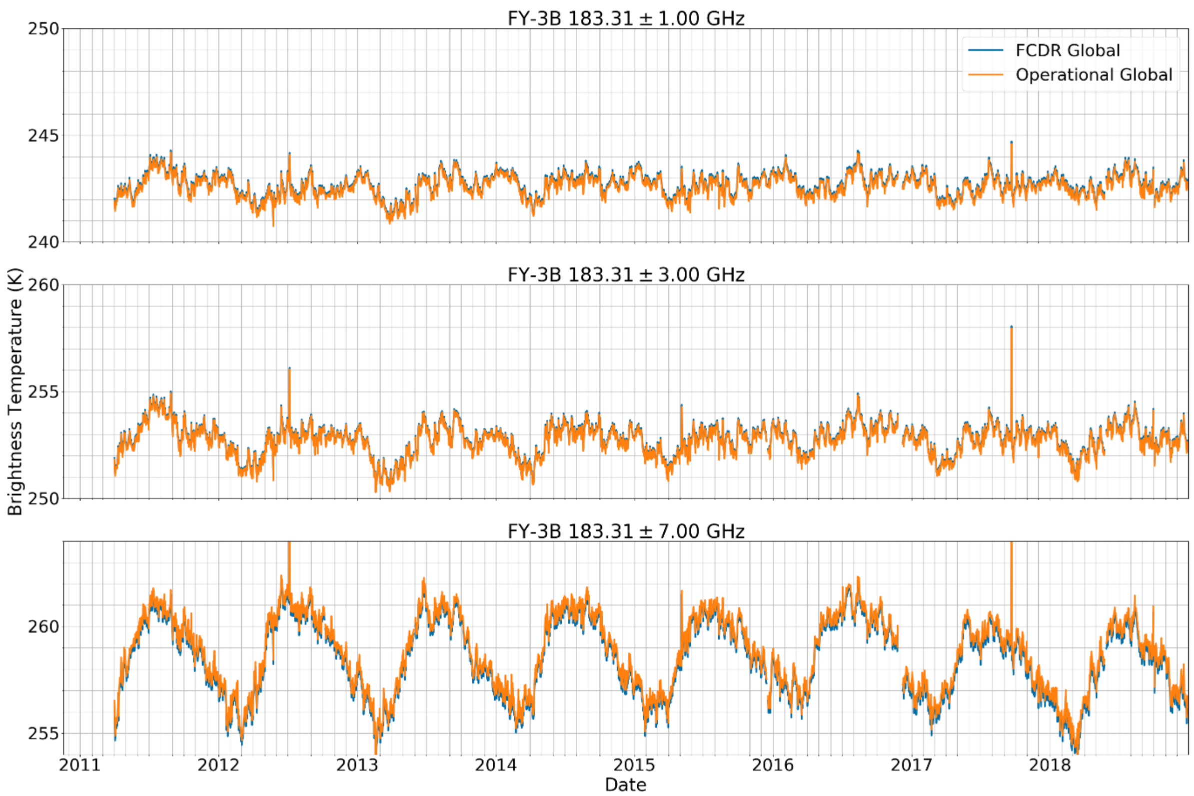

Figure A4: Time series of FCDR (blue) and operational (orange) data for FY-3B.

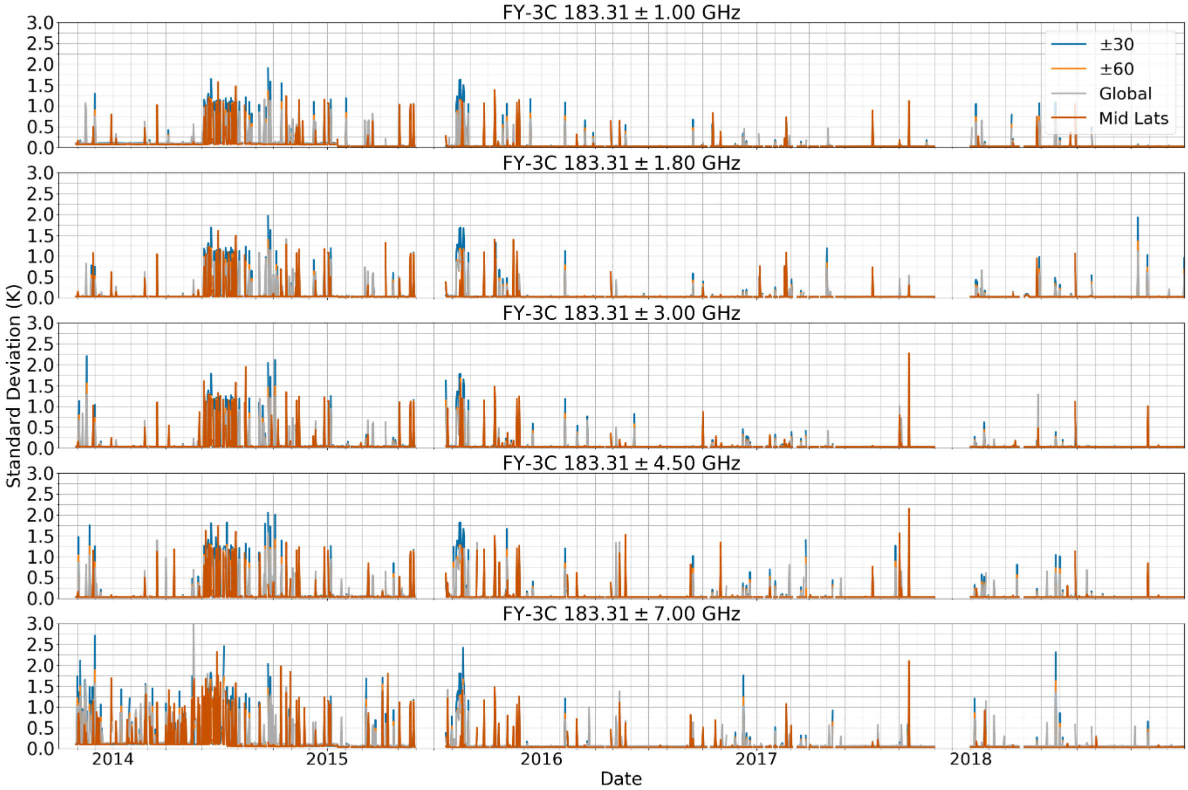

Figure A5: Time series of the standard deviation of differences for FY-3C (FCDR minus operational) for each of the 183 GHz channels. The data have been averaged over latitude bands ±30° (blue), ±60° (orange), ±90° (grey), and mid-latitude bands 60°S-30°S and 30°N-60°N (red).

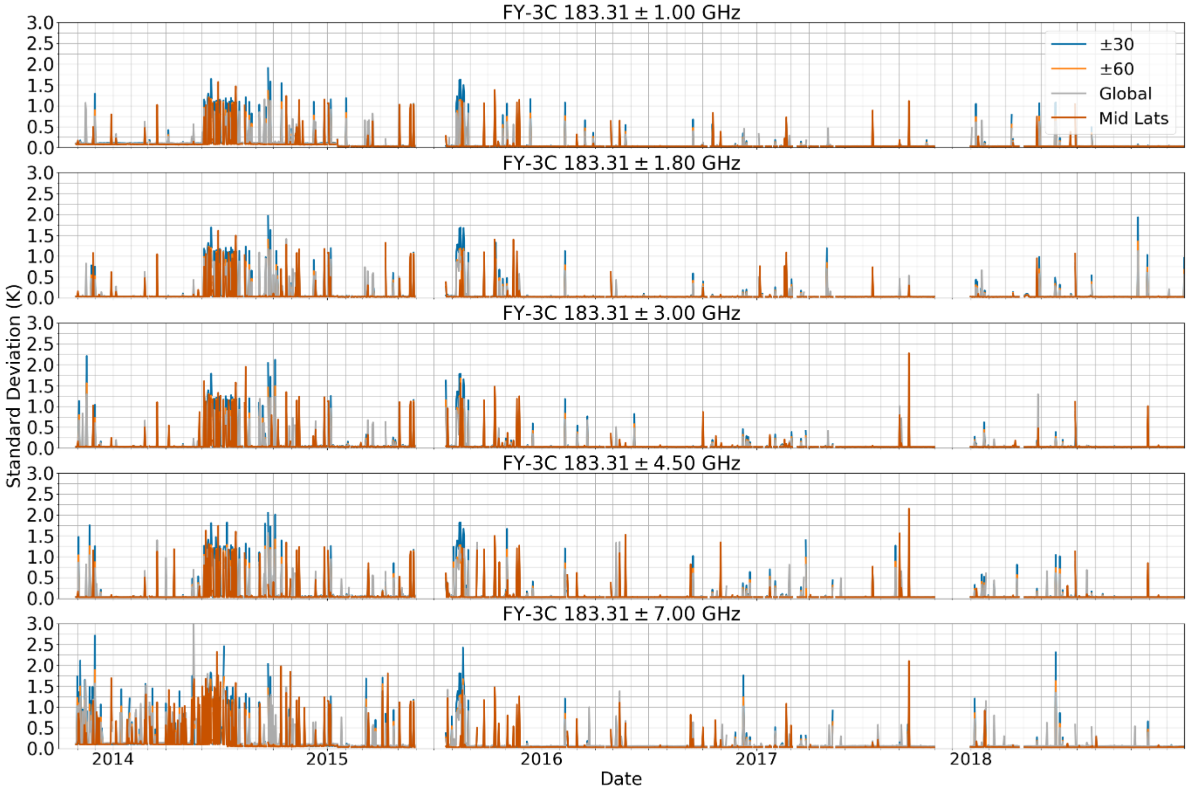

Figure A6: Time series of the standard deviation of differences for FY-3C (FCDR minus operational) for each of the 183 GHz channels.

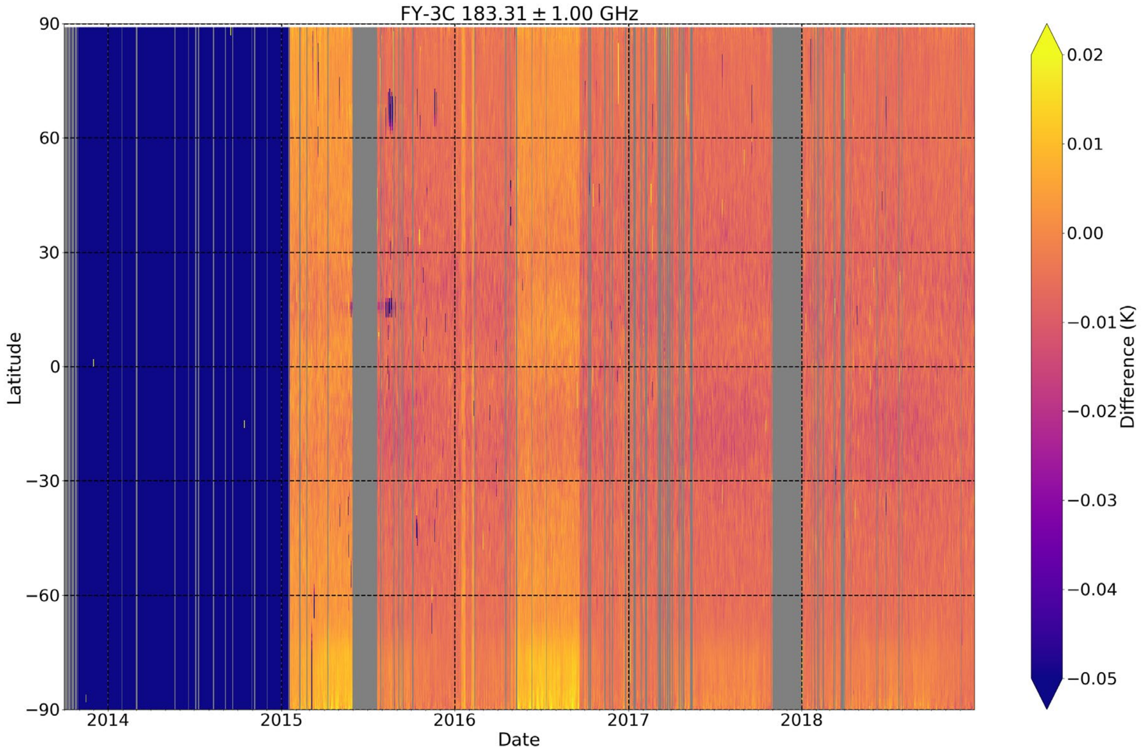

Figure A7: Hovmöller diagram showing the change in daily mean difference by latitude with time for FY-3C for the 183.31±1 GHz channel. The grey bands correspond to data missing from the analysis.

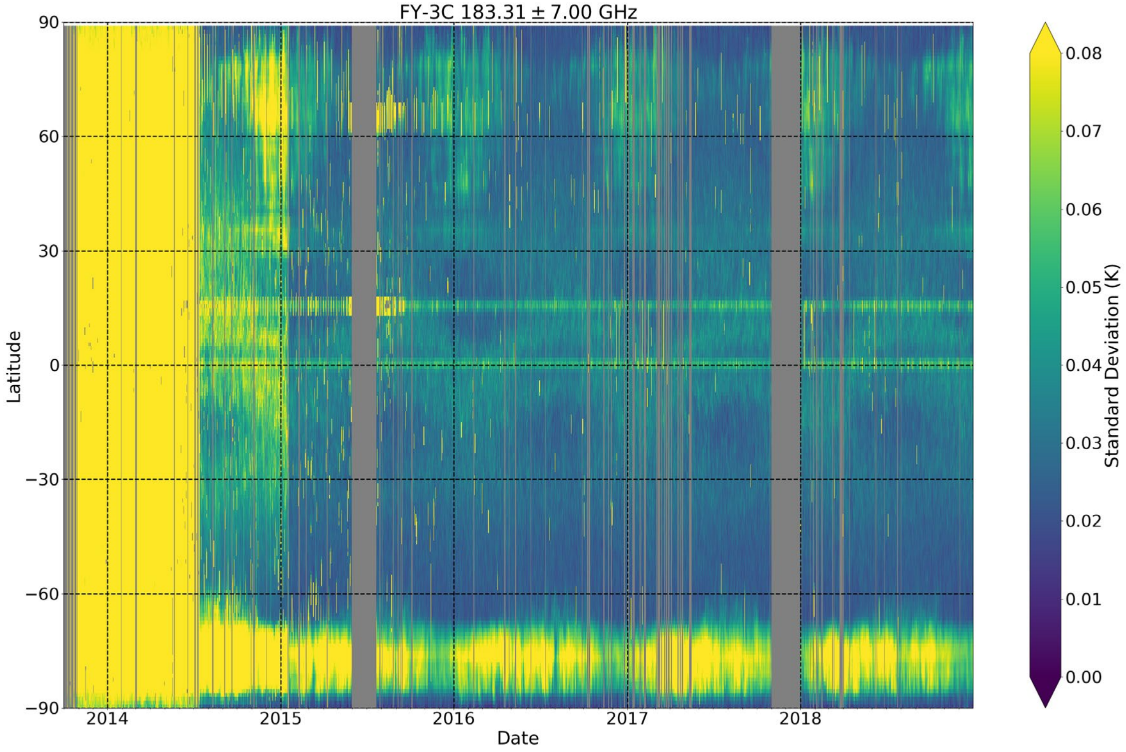

Figure A8: Hovmöller diagram showing the change in standard deviation of daily mean differences by latitude with time for FY-3C for the 183.31±7 GHz channel. The grey bands correspond to data missing from the analysis.

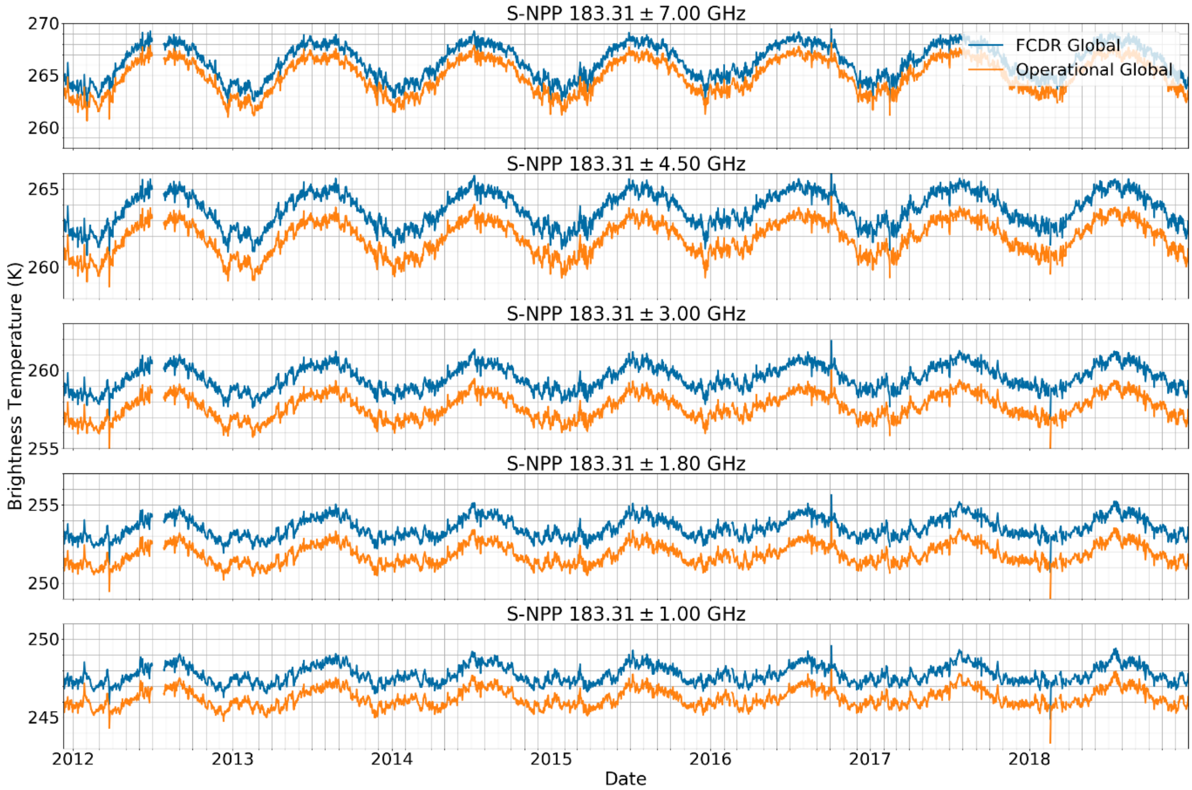

Figure A9: Time series of FCDR (blue) and operational (orange) data for S-NPP

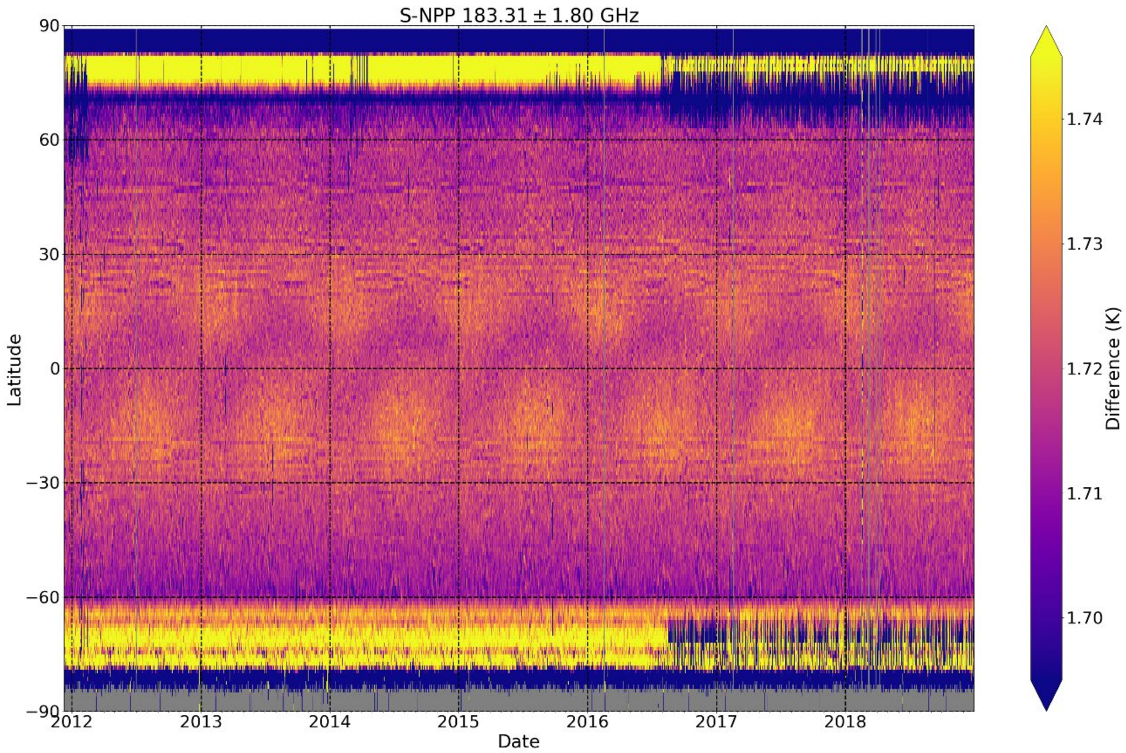

Figure A10: Hovmöller diagram showing the change in daily mean difference by latitude with time for S-NPP for the 183.31±1.8 GHz channel. The grey bands correspond to data missing from the analysis

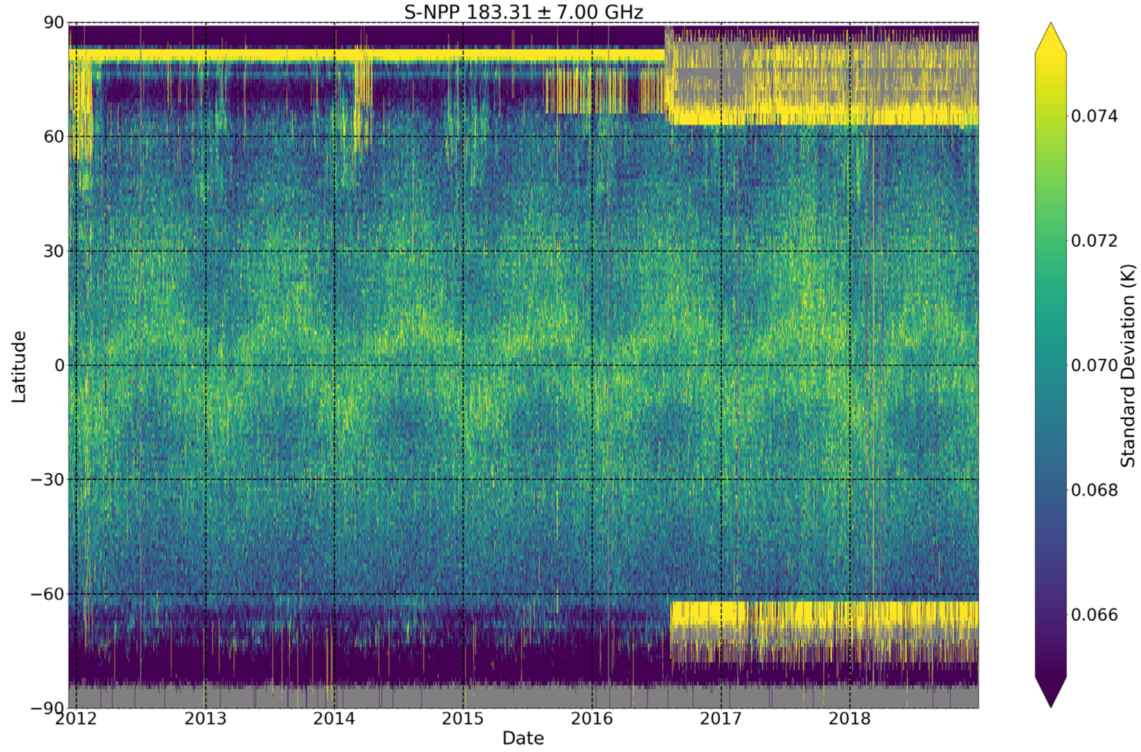

Figure A11: Hovmöller diagram showing the change in standard deviation of daily mean differences by latitude with time for S-NPP for the 183.31±7 GHz channel. The grey bands correspond to data missing from the analysis.

7 Appendix B: Supplementary Figures for the FCDR vs Simulated BTs

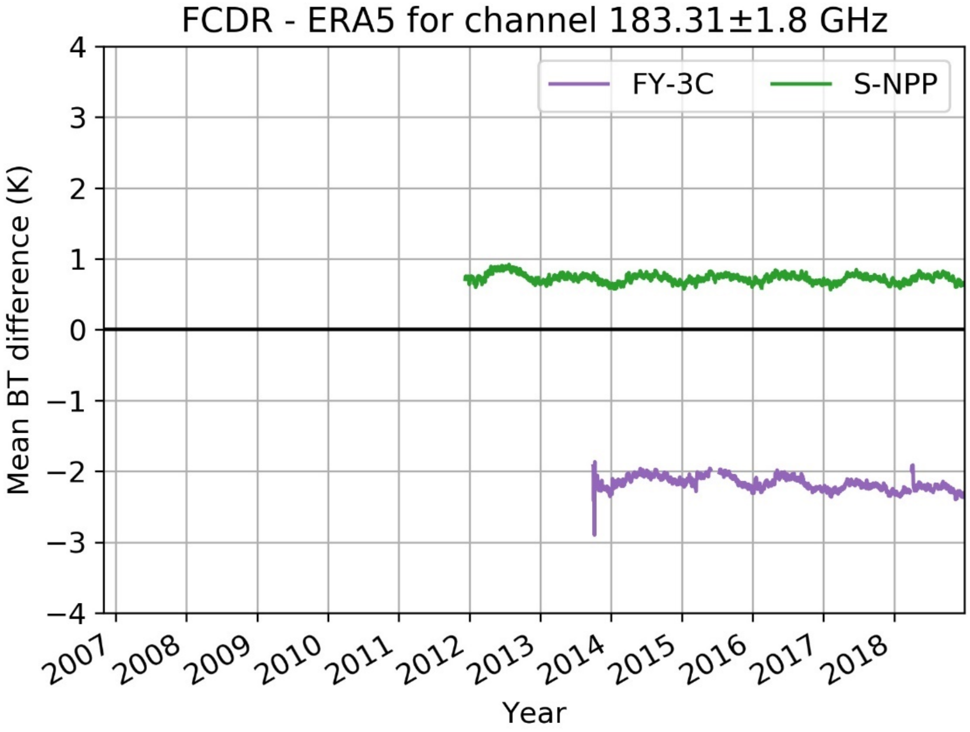

Figure B1: Time series for the FCDR minus operational data for all sensors for the 183.31±1.8 GHz channel.

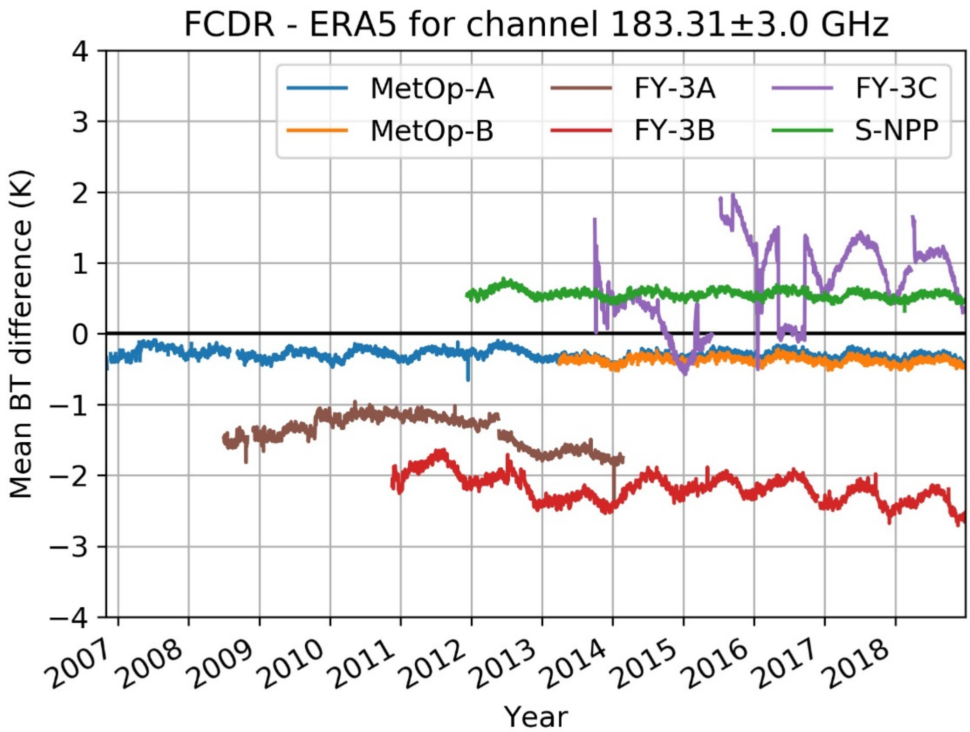

Figure B2: Time series for the FCDR minus operational data for all sensors for the 183.31±3 GHz channel.

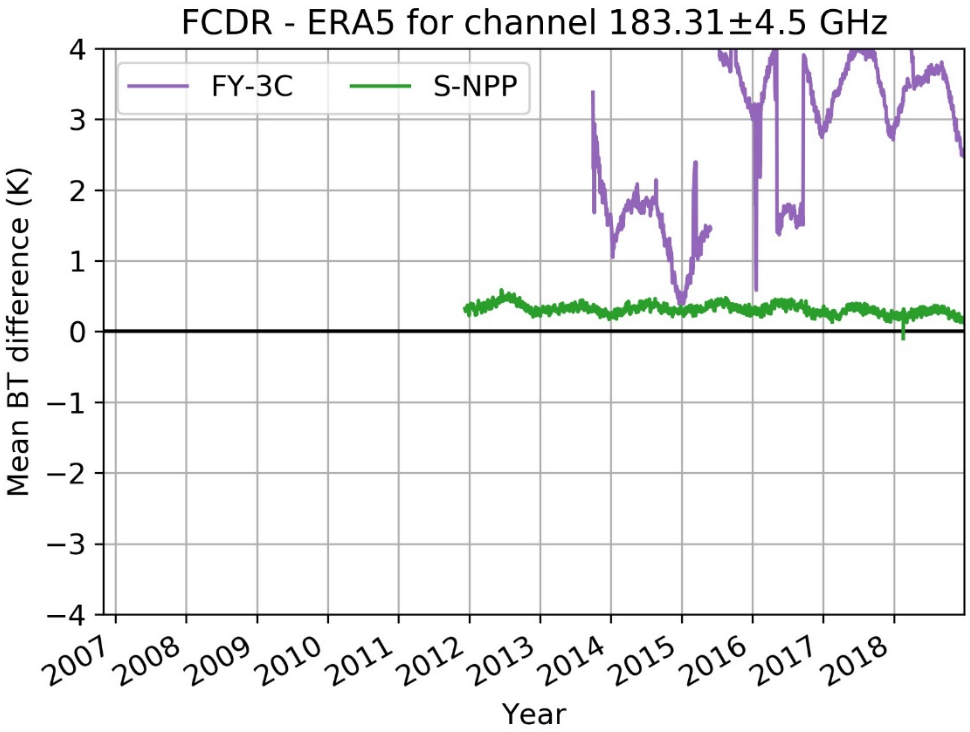

Figure B3: Time series for the FCDR minus operational data for all sensors for the 183.31±4.5 GHz channel.

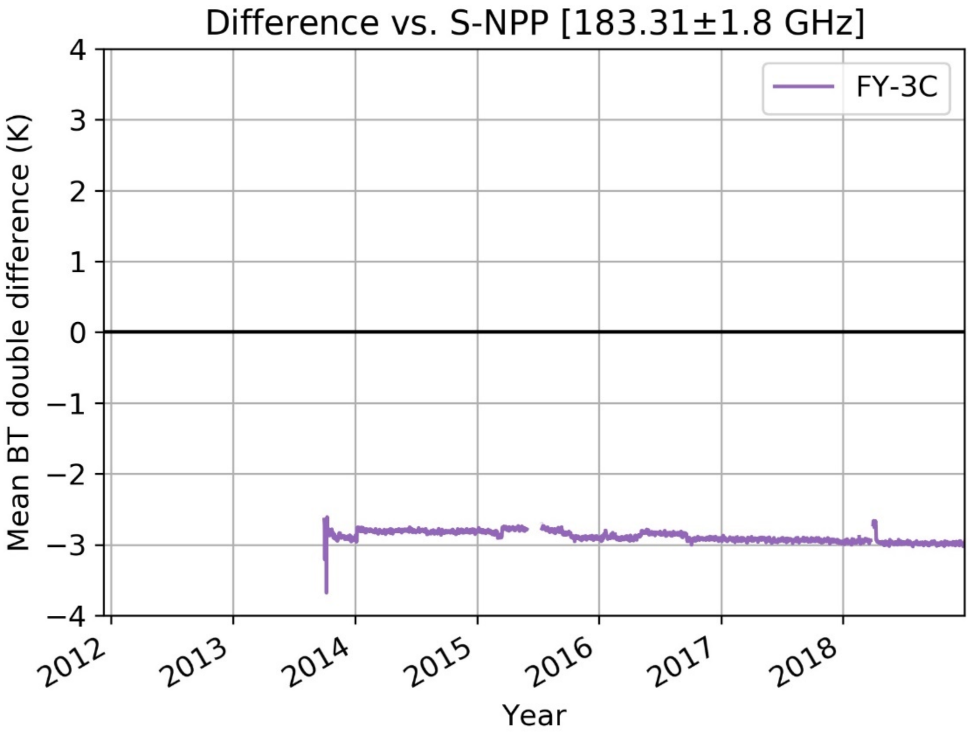

Figure B4: Time series of double differences referenced to S-NPP for the 183.31±1.8 GHz channel.

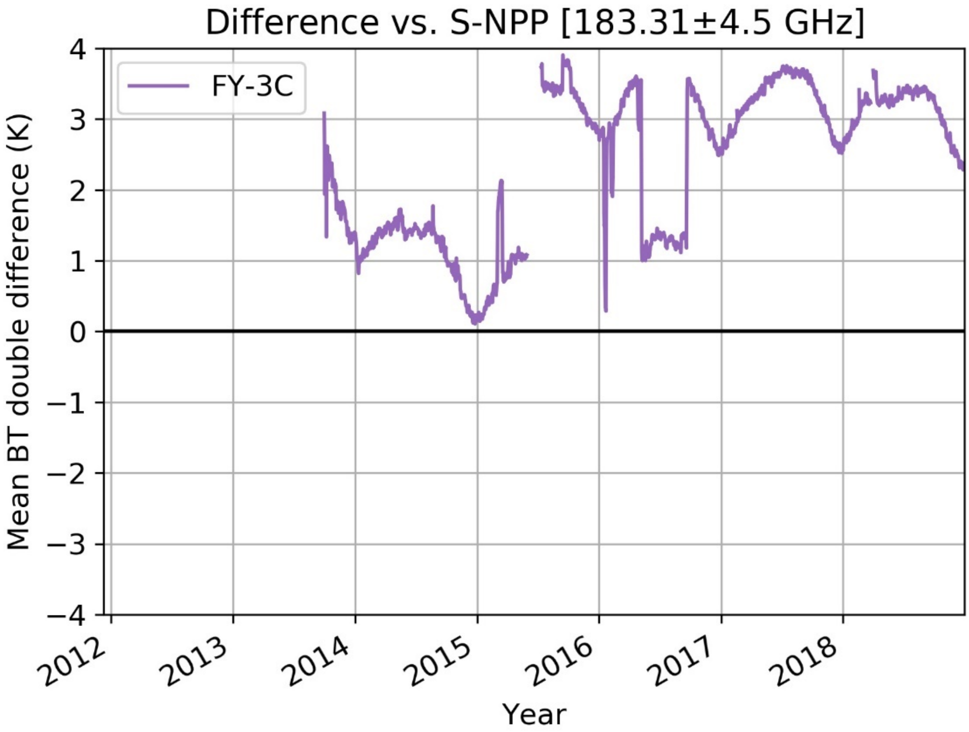

Figure B5: Time series of double differences referenced to S-NPP for the 183.31±4.5 GHz channel.

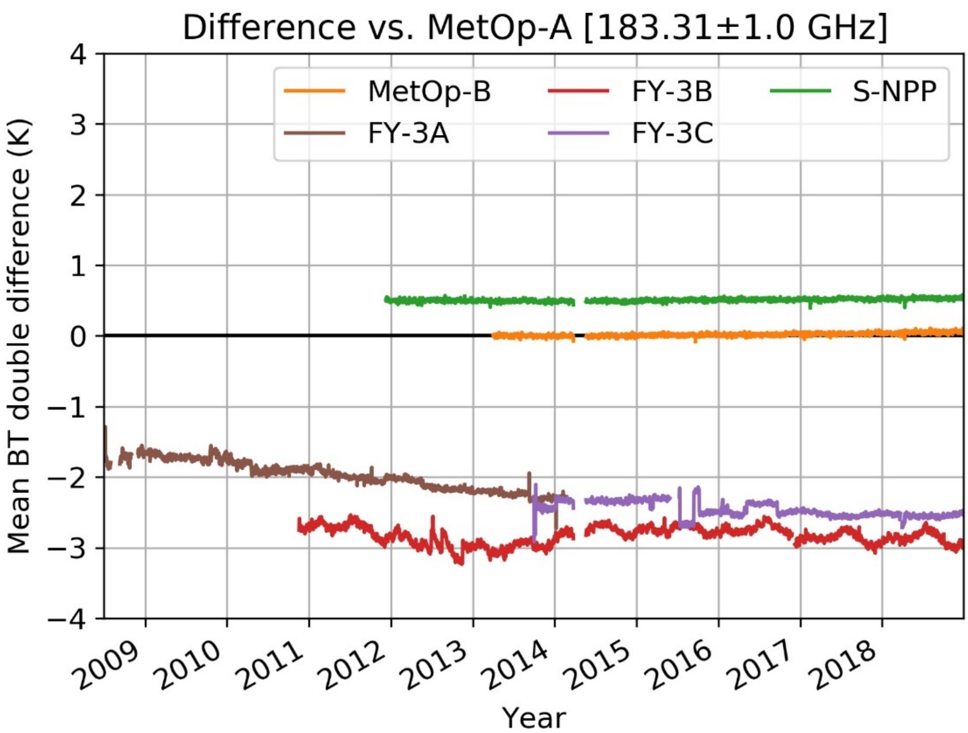

Figure B6: Time series of double differences referenced to Metop A for the 183.31±1 GHz channel.

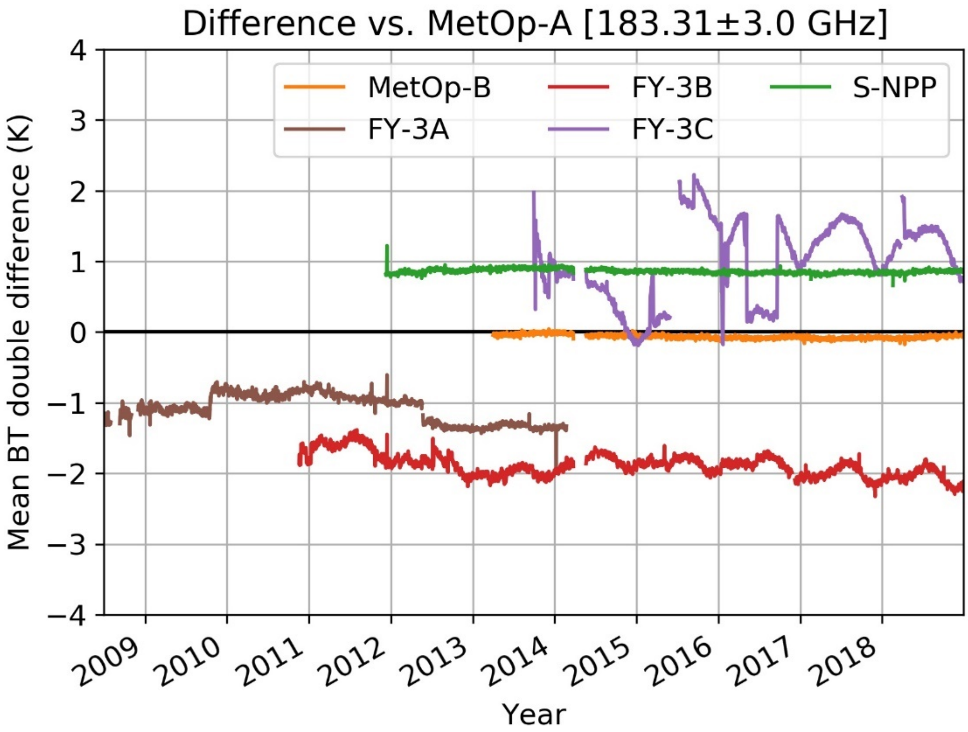

Figure B7: Time series of double differences referenced to Metop A for the 183.31±3 GHz channel.

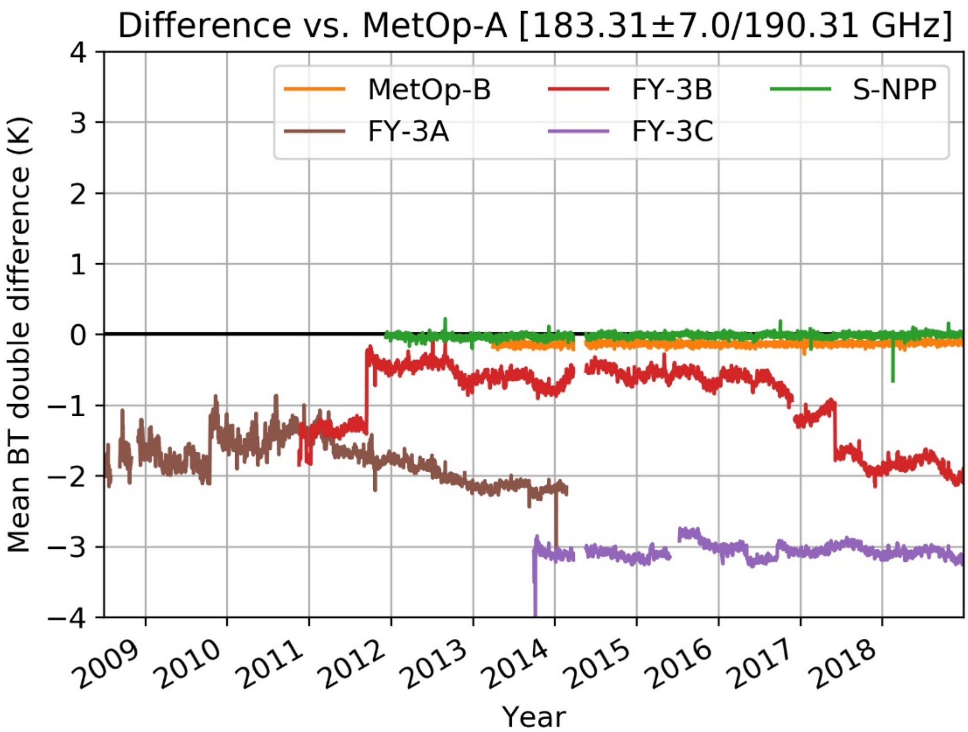

Figure B8: Time series of double differences referenced to Metop A for the 183.31±7 GHz channel.

8 Appendix C: Supplementary figures for the analysis of simultaneous nadir overpasses

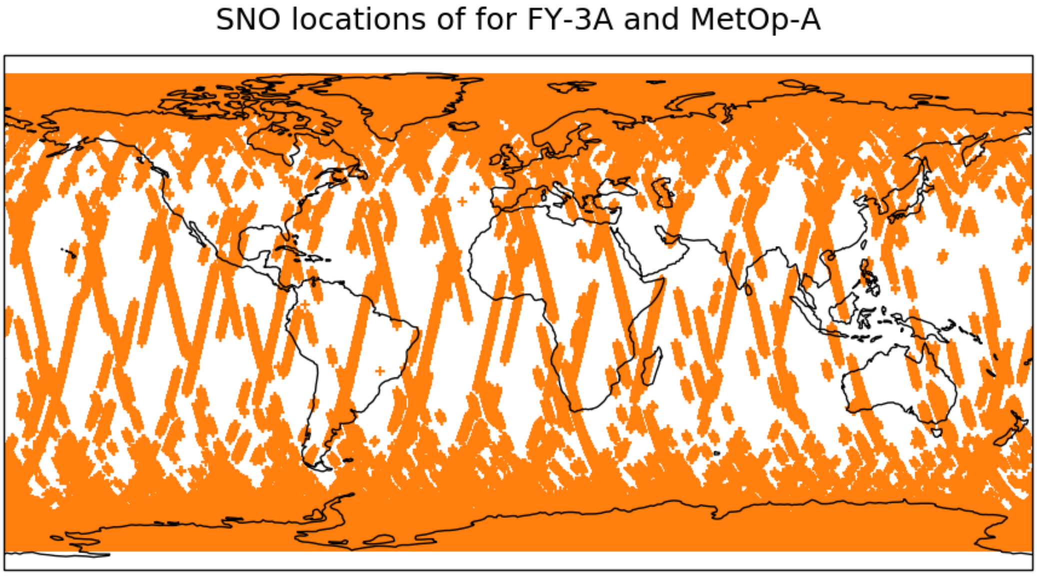

Figure C1: Map showing the locations of the FY-3A and Metop A SNOs.

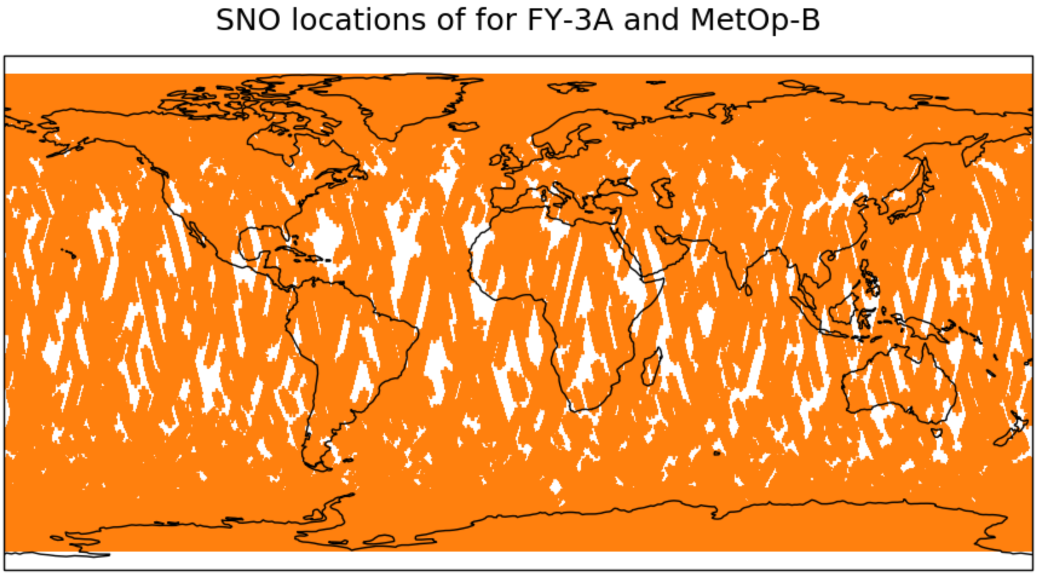

Figure C2: Map showing the locations of the FY-3A and Metop B SNOs.

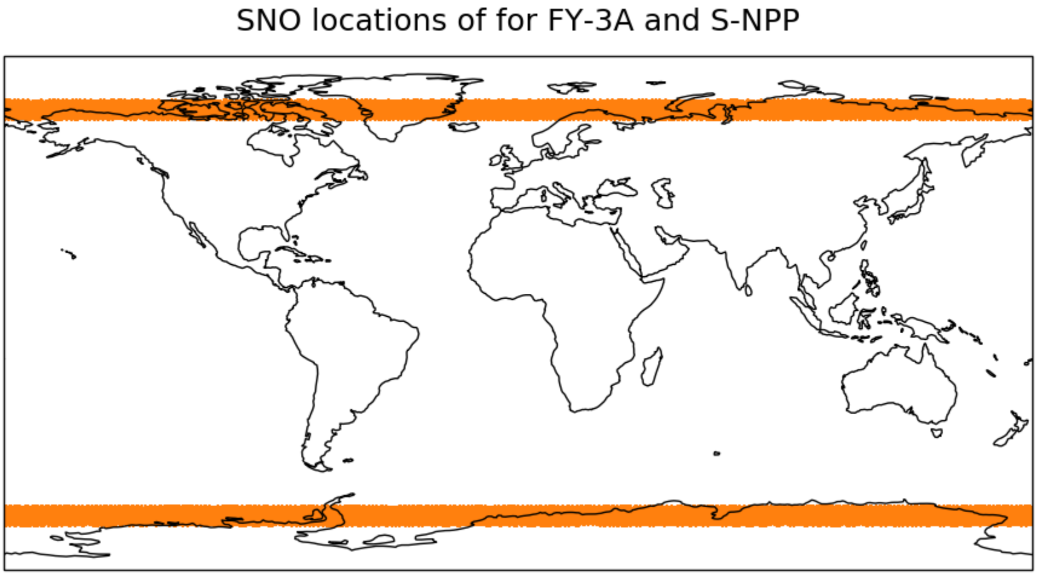

Figure C3: Map showing the locations of the FY-3A and S-NPP SNOs

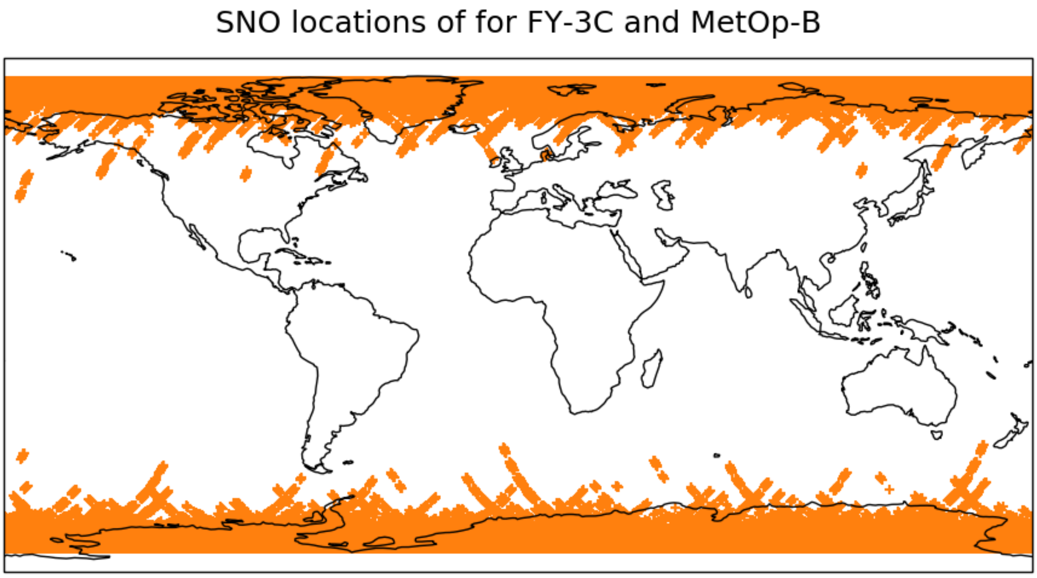

Figure C4: Map showing the locations of the FY-3C and Metop B SNOs

ECMWF Shinfield Park Reading RG2 9AX UK