Towards a new fundamental climate data record for microwave humidity sounders based on metrological best practice

Dissertation

zur Erlangung des Doktorgrades
an der Fakultät für Mathematik, Informatik und Naturwissenschaften
im Fachbereich Geowissenschaften der Universität Hamburg

vorgelegt von

Imke Hans

Hamburg, 2018

Als Dissertation angenommen vom Fachbereich Geowissenschaften

Tag des Vollzugs der Promotion: 02.11.2018

Gutachter: Prof. Dr. Stefan A. Buehler
                                                          Dr. Martin Burgdorf

Vorsitzender des Fachpromotionsausschusses   Prof. Dr. Dirk Gajewski
Geowissenschaften:

Dekan der Fakultät MIN: Prof. Dr. Heinrich Graener

Abstract

The goal of this work is to contribute to making historical remote sensing data usable for climate research. Carried out in the framework of the Horizon 2020 project "Fidelity and Uncertainty in Climate Data Records from Earth Observation" (FIDUCEO), this thesis focuses on the operational data of the Microwave humidity sounders Special Sensor Microwave Temperature (SSMT-2), Advanced Microwave Sounding Unit - B (AMSU-B) and Microwave Humidity Sounder (MHS) on board polar orbiting satellites. Since these instruments were not designed for climate monitoring, their data sets do not fulfil the requirements for this purpose. Rather, the data sets contain numerous data doubling and they do not provide detailed uncertainty information. Additionally, they show discontinuities from one instrumental mission to the successor mission. These discontinuities, or "inter-satellite biases", prevent the construction of long-term, stable data records for climate monitoring.

To obtain a data record usable for climate monitoring, these inter-satellite biases need to be understood and, ideally, removed. This would provide a harmonised, stable data record, free from instrumental issues. A data record for climate research should also be free from operational artefacts, such as doubled data. Also, a consistent format and a data frame containing a single orbit without any overlap with the preceding or succeeding orbit is desirable. Furthermore, the uncertainties within the measurement process need to be characterised and propagated to the fundamental measurand in the data set, which is the brightness temperature or radiance. This information would allow for constructing an uncertainty quantified fundamental climate data record (FCDR) of brightness temperatures as it is required for obtaining sound results in climate research.

The new FCDR for Microwave humidity sounders that I generate in this thesis handles the problems of the operational data: 1) The new FCDR provides extensive uncertainty information, considering also the correlation of underlying errors. 2) I improve the understanding of certain inter-satellite biases. In a harmonisation effort, the inter-satellite biases will then be reduced by a recalibration approach. First results of this harmonisation effort are presented in this thesis. 3) The new FCDR is based on consolidated data without data doubling. It is stored in an easy-to-use NetCDF format providing the brightness temperature, its uncertainty and detailed quality information within a single file per orbit.

To start, I investigate the behaviour of the instruments over their lifetime. The time evolution of the basic calibration quantities offers valuable information: it provides a useful overview of erratic periods for the later production of the FCDR and the analysis of inter-satellite biases. In a detailed study, the noise and the related noise equivalent differential temperature ($NE\Delta T$) are analysed. Apart from the stand-alone result of time evolution of $NE\Delta T$ as a quick overview for users, the study also introduces the recently proposed method of using the Allan deviation for noise computation. This Allan deviation is further used for the later production of the new FCDR.

As basis for calibration, uncertainty analysis and uncertainty propagation, a measurement equation is formulated. It models the impacts of the various physical effects on the measurement process. Within the uncertainty analysis and the generated FCDR files, three classes of effects

causing uncertainty on the brightness temperature are distinguished. The three classes reflect the different correlation behaviour of the underlying errors. I prepare an FCDR generator code that processes the raw level 1b data to calibrated level 1c brightness temperatures. The code executes the uncertainty propagation and prepares the FCDR files without data doubling from one equator crossing to the next crossing in the same flight direction. In this way, an unharmonised FCDR is produced for all considered instruments in a first step.

In a second step, I address the issue of the harmonised FCDR: I analyse the inter-satellite biases depending on time, scan angle and brightness temperature. For certain biases, I identify radio frequency interference (RFI) as a possible cause known for AMSU-B on NOAA-15 already, but not for MHS on NOAA-19. I also detect inconsistencies within the operational calibration and correct for these in my FCDR generator code. Moreover, the FIDUCEO harmonisation procedure is presented in a proof-of-concept study. Harmonisation is a recalibration approach that aims at reducing biases by optimising calibration parameters within the measurement equation. As inter-comparison method, simultaneous nadir overpasses (SNOs) are used. First test results show that the optimisation machine works in principle. In a first harmonised FCDR, I observe the expected reduction of bias in the limited range of brightness temperatures that the input data of SNOs cover. A promising method for further improvement on the method by extending the input data to warmer brightness temperatures is in discussion.

Zusammenfassung

Ziel dieser Arbeit ist es, zur Aufbereitung historischer Fernerkundungsdatensätze für die Klimaforschung beizutragen. Diese Dissertation wurde im Rahmen des Horizon 2020 Projekts "Fidelity and Uncertainty in Climate Data Records from Earth Observation" (FIDUCEO) durchgeführt und konzentriert sich auf die operationellen Daten der Mikrowellenradiometer für die Feuchtebestimmung Special Sensor Microwave Temperature (SSMT-2), Advanced Microwave Sounding Unit - B (AMSU-B) und Microwave Humidity Sounder (MHS), die an Bord von Satelliten in polaren Umlaufbahnen installiert sind. Da diese Instrumente nicht für die Beobachtung des Klimas entwickelt wurden, erfüllen ihre Datensätze nicht die Anforderungen für diesen Zweck. Vielmehr enthalten sie zahlreiche Datenduplizierungen, sie geben keine detaillierten Informationen über Unsicherheiten und zeigen darüber hinaus Diskontinuitäten zwischen aufeinanderfolgenden Missionen. Diese systematischen Messdifferenzen zwischen den Instrumenten auf verschiedenen Satelliten ("inter-satellite biases") verhindern bislang das Konstruieren langer, stabiler Datensätze für die Beobachtung des Klimas.

Um einen Datensatz zu erhalten, der für die Klimaforschung nutzbar ist, müssen diese systematischen Messdifferenzen verstanden sein und idealerweise beseitigt werden. Dies würde einen harmonisierten, stabilen Datensatz generieren, frei von instrumentellen Problemen. Ein Datensatz für die Klimaforschung sollte außerdem keine Artefakte des operationellen Betriebs aufweisen, wie zum Beispiel duplizierte Daten. Zudem sind ein konsistentes Format, sowie ein definierter Datenumfang von einem Orbit, ohne Überlapp mit dem vorherigen oder nachfolgenden Orbit, erstrebenswert. Außerdem müssen die Unsicherheiten im Messprozess charakterisiert und zu der fundamentalen Messgröße des Datensatzes, der Strahlungstemperatur oder Radianz, fortgepflanzt werden. Diese Informationen erlauben dann die Konstruktion eines unsicherheitsquantifizierten fundamentalen Klimadatensatzes der Strahlungstemperatur (fundamental climate data record, FCDR), wie er für fundierte Ergebnisse der Klimaforschung nötig ist.

Der neue FCDR aus den Daten der Mikrowellenradiometer für die Feuchtebestimmung, den ich in dieser Arbeit erstelle, bereinigt Probleme der operationellen Daten: 1) Der neue FCDR enthält ausführliche Informationen über Unsicherheiten, wobei auch Korrelationen der zugrundeliegenden Fehler betrachtet werden. 2) Ich verbessere das Verständnis bestimmter systematischer Messdifferenzen. Mit der Harmonisierung sollen die systematischen Messdifferenzen durch einen Rekalibrierungsansatz reduziert werden. Erste Ergebnisse dieser Harmonisierung werden in dieser Arbeit vorgestellt. 3) Der neue FCDR basiert auf konsolidierten Daten ohne Duplizierungen. Die Daten werden in ein benutzerfreundliches NetCDF Format geschrieben, welches die Strahlungstemperatur, ihre Unsicherheit sowie detaillierte Qualitätsinformation in einer einzigen Datei je Orbit enthält.

Zu Beginn untersuche ich das Verhalten der Instrumente über den Zeitraum ihrer Lebenszeit. Für die spätere Produktion des FCDRs und der Analyse der systematischen Messdifferenzen gibt die Zeitentwicklung der grundlegenden Kalibrierungsgrößen wertvolle Informationen und einen nützlichen Überblick über Zeitperioden anormalen Instrumentenverhaltens. In einer detaillierten Studie werden das Rauschen und die abgeleitete noise equivalent differential temperature (NEΔT) analysiert. Abgesehen von dem eigenständigen Ergebnis des Überblicks über die

zeitliche Entwicklung der NEΔT, führt die Studie auch die erst kürzlich vorgeschlagene Methode der Rauschbestimmung mittels der Allan deviation ein, welche in der späteren FCDR Produktion angewendet wird.

Als Basis für die Kalibrierung, die Analyse und Fortpflanzung der Unsicherheiten wird eine Messgleichung formuliert, welche die verschiedenen physikalischen Effekte modelliert, die auf den Messprozess einwirken. Bei der Analyse der Unsicherheiten und auch in den erzeugten FCDR Dateien werden drei Klassen von Effekten unterschieden, die Unsicherheiten in der Strahlungstemperatur hervorrufen. Die drei Klassen spiegeln das unterschiedliche Korrelationsverhalten der zugrundeliegenden Fehler wider. Ich schreibe einen FCDR-generierenden code, der die rohen Level-1b Daten zu kalibrierten Level-1c Strahlungstemperaturen prozessiert. Der code führt die Unsicherheitsfortpflanzung aus und erstellt die einzelnen FCDR Dateien ohne Datenduplizierung von einem Äquator-Überflug zum nächsten Überflug in derselben Flugrichtung. Auf diese Weise wird in einem ersten Schritt ein unharmonisierter FCDR für alle betrachteten Instrumente produziert.

Am harmonisierten FCDR wird in einem zweiten Schritt gearbeitet: Ich analysiere die systematischen Messdifferenzen, die eine Zeit-, Blickwinkel- und Strahlungstemperaturabhängigkeit aufweisen. Für einige Messdifferenzen identifiziere ich die elektromagnetische Interferenz (radio frequency interference, RFI) als mögliche Ursache, welche für AMSU-B auf NOAA-15 bereits bekannt war, beispielsweise jedoch nicht für MHS auf NOAA-19. Ich decke auch Unstimmigkeiten in der operationellen Kalibrierung auf und korrigiere sie in meinem FCDR code. Desweiteren stelle ich die FIDUCEO Harmonisierung im Rahmen einer Machbarkeitsstudie vor. Die Harmonisierung ist ein Rekalibrierungsansatz, der zum Ziel hat, die systematischen Messdifferenzen durch Optimierung von Kalibrierungsparametern in der Messgleichung zu reduzieren. Als Methode für den Vergleich von Instrumenten auf verschiedenen Satelliten werden simultaneous nadir overpasses (SNOs) verwendet. Erste Ergebnisse belegen das prinzipielle Funktionieren der Optimierungsmaschine. In einem ersten harmonisierten FCDR beobachte ich die erwartete Reduzierung der systematischen Messdifferenzen im begrenzten Strahlungstemperaturintervall, welches durch die SNOs in den Eingangsdaten abgedeckt wird. Ein vielversprechender Ansatz für eine weitere Verbesserung der Methode durch Ausweitung des Eingangsdatensatzes hin zu warmen Strahlungstemperaturen wird diskutiert.

Contents

Contents

1 Introduction and motivation for a new fundamental climate data record

The goal of this work is to generate a fundamental climate data record (FCDR) for satellite-borne passive microwave (MW) humidity sounders. The novelty of this data set is in three main aspects: 1) The data come with extensive uncertainty information for the recalibrated measurements. 2) Differences between sensors on different satellites ("inter-satellite biases") are reduced by a process called harmonisation. 3) The data set comes in an easy-to-use NetCDF format, the data set is based on consolidated data without data doubling and combines observables, their uncertainties and detailed quality information in a single file per orbit. These three aspects set apart the FCDR produced in this study from previous efforts in generating an FCDR for microwave radiometers.

An FCDR is a data set of sufficient length, consistency and stability to be of use for climate research. It is a level 1 data set. That means it contains the measured signal in radiance or brightness temperature. An FCDR forms the basis for a climate data record (CDR) which is a level 2 (or higher) product containing the geophysical quantities that are of direct interest in climate studies, such as sea surface temperature or upper tropospheric humidity. The brightness temperature measurements of the microwave radiometers considered in this study are used to retrieve upper tropospheric humidity (UTH). UTH is of special interest in climate research, as water vapour is a major greenhouse gas. Moreover, water vapour shows strong spatio-temporal variability that asks for global monitoring in order to assess its evolution. Only satellite-based measurements can provide this global coverage over long time periods. That is why an FCDR of microwave humidity sounder data is of special interest in climate research. But, why is it a major effort to create an uncertainty quantified, harmonised FCDR without inter-satellite biases? This question can be split into several aspects: What does harmonisation mean and how is it done? How do we include uncertainty information? And to start the investigation: What are the problems in the available data sets that need to be solved and how do we address these problems?

The data sets currently available are operational data sets that were not designed for climate monitoring but for assimilation in Numerical Weather Prediction (NWP) and reanalysis. These operational data sets exist since the early 1990s, when passive microwave radiometers on board polar orbiting satellites were first used for observation of atmospheric water vapour. Usually, the instruments have lifetimes of a couple of years and overlap with their predecessor and successor missions of similar instruments. Hence, the data sets now cover a time period of more than 20 years and therefore gain attention for the use in climate research. As pointed out however, these radiometers were not designed for recording long time series that are needed for climate monitoring. Climate research requires continuous time series, that means neither large gaps, nor jumps of the observed physical quantities should appear. The most difficult aspect, however, is stability: the time series of an observed physical quantity should only contain trends that reflect nature, in this case climate. But any instrumental effects causing time dependent trends should ideally be removed, otherwise they may mask the natural trend or feign a trend. In

their current shape, the available data records of the passive microwave radiometers do not meet these requirements as they show biases from one instrument to the other or strong degradation with time. Therefore, the data records cannot be used for climate monitoring, yet. For similar instrument families, efforts have been undertaken in the past to overcome the inter-satellite biases seen for those instruments. Zou et. al. have devised a recalibration scheme for the microwave sounding unit (MSU) instrument that is used for temperature profiling. All inter-satellite biases were assumed to originate from an uncorrected non-linearity of the sensor. This non-linearity and an additional offset were newly determined in Zou et al. (2006) to give a new set of calibration coefficients. However, this non-linearity is not necessarily the true physical non-linearity of the sensor, but a mathematical construct that is optimised such that it reduces the bias. A similar approach was used later by Zou and Wang (2011) for the AMSU-A instrument, also used for temperature profiling. For the microwave humidity sounders considered in this thesis, an FCDR generation effort has been carried out by Hanlon and Ingram (2016) from the UK Met Office as part of the CM SAF (Climate Monitoring Satellite Application Facility). From a detailed analysis of monthly zonal mean biases, they devised corrections in order to reduce the biases. These corrections assume that the diurnal cycle is symmetric around noon. This assumption is not fully true and it has to be considered that it may lead to masking of the actual inter-satellite bias. Yet, this bias correction by Hanlon and Ingram (2016) does not aim at finding the actual origin of the biases within the measurement process. This thesis, however, contributes to finding the actual origin of the biases and correcting for it, which is the major challenge of making historical data sets from passive microwave humidity sounders usable for climate research.

Apart from inter-satellite biases, the available data sets have a second deficit: the lack of uncertainty information. Any measured physical quantity should come with an uncertainty estimate, otherwise its scientific value is questionable. However, the only information on uncertainty that is provided in some of the data records is the noise equivalent differential temperature (NEΔT), indicating the overall sensitivity of the radiometer. For some instruments it is given as time varying quantity within the data sets, whereas for others only the pre-launch specifications for NEΔT are available. NEΔT, however, contributes to random uncertainty only. Truly random errors however, will average out when users investigate climate. Therefore, the knowledge of the NEΔT is not enough: the uncertainties that appear in some structured form and that will not average out, are particularly important and need to be considered by the users. A pixel level uncertainty covering all possible effects contributing to the (time dependent) uncertainty of the final measurand is not at all available. For climate research, however, this information is definitely needed.

The goal of this work is to contribute to the preparation of a data set for the passive microwave radiometers that fulfils the requirements for climate research, including comprehensive uncertainty information in the data records. The work to generate this new FCDR is part of the European project "Fidelity and Uncertainty in Climate data records from Earth Observation (FIDUCEO)". Within the scope of the project, FCDRs are also generated for instruments in the visible and infrared spectrum. The aim of this project is to understand the origins and quantify instrumental issues leading to random and systematic errors in the measurements. For this, the metrological perspective, i.e. the perspective of measurement science is adopted (Mittaz et al., 2018). This FIDUCEO philosophy makes use of a rigorous measurement equation approach. That means that the complete calibration chain of translating the initial raw signal to the output quantity of brightness temperature is modelled in a single function. This function, the measurement equation, depends on various physical quantities that have an impact on the cali-

bration. The effect of each of these quantities needs to be understood and uncertainty quantified in order to allow both understanding of inter-satellite biases and uncertainty propagation.

The uncertainty propagation is part of the FIDUCEO calibration of raw data to finally produce the FCDR. In this work I develop a new processing code adapting the described FIDUCEO philosophy of a rigorous measurement equation approach. A multitude of effects impacting on the measurement process are identified for each instrument and included in the measurement equation as well as in the uncertainty propagation. The new MATLAB code developed here enables the processing from raw counts (level 1b) to brightness temperature (level 1c), that means that the evaluation of the measurement equation including all required preparations is carried out. At this stage, the new processing imitates the operational one and reproduces the operational brightness temperatures in order to gain confidence in the functioning of the new processing code. As improvement on the operational calibration however, the code also allows for usage of corrected calibration parameters to overcome inconsistencies of the operational calibration, and above all, the code executes the uncertainty propagation and generates the FCDR files in the easy-to-use NetCDF format.

In a preparatory step towards these FCDR files, I analyse the very raw counts and basic quantities measured by the instruments over their lifetime. This helps to decide on usability of data for FCDR production as it gives a first overview of the evolution of the instrument to detect periods of erratic behaviour or strong degradation. Degradation can be related to a decreased sensitivity of the instrument (and increased NEΔT). I therefore carry out a detailed investigation of the noise and NEΔT evolution over the lifetime of the instruments. Periods of large noise are identified, and a set of usable data is suggested. From these suggested usable data, the FCDR files are produced. Moreover, the methods used for noise computation are further applied within the FCDR production.

The generated FCDR files combine calibrated brightness temperatures and pixel level uncertainty information. Hence, these files represent the intermediate unharmonised FCDR, that provides uncertainty information and reproduces the brightness temperatures of the operational calibration. However, these unharmonised FCDR files are still not comparable among the different individual radiometer instruments: The inter-satellite biases are not yet removed by the new processing. To reconcile the instruments, the biases themselves and possible causes are analysed then. In a sensitivity study, I analyse various identified effects affecting the measurement process regarding their influence on the final brightness temperature in a qualitative and quantitative sense. By comparing the results to existing biases and taking into account the identified problems and inconsistencies of the old operational calibration, I deduce possible approaches for harmonising the instruments.

As first approach to a harmonised data set, I correct the detected inconsistencies within the operational calibration procedure and adapt my processing code accordingly. For the actual reconciliation of all instruments of the entire MW data record, a full harmonisation procedure is required as it is designed within the FIDUCEO project. I present the FIDUCEO harmonisation process as a proof-of-concept, including the first results from this novel approach. The harmonisation itself is executed with a sophisticated mathematical process based on regression algorithms and developed within the FIDUCEO project by partners at the National Physical Laboratory and FastOpt GmbH. The regression tools "(fast) Orthogonal Distance Regression (ODR)" and "Error-in-Variables (EIV)" take into account the uncertainties for both the independent and dependent variables and are able to quickly process large data sets. The tools minimise the inter-satellite biases by optimising certain calibration parameters. I choose the calibration parameters to be optimised on the basis of the sensitivity study and from investigating

the instruments behaviour. The choice is no final statement, but it is an iterative process to find the combination of effects that is able to explain the observed biases. Within this process, I use the harmonisation output, namely the optimised calibration parameters, and include them in the FCDR processor. This results in a first version of harmonised MW FCDR.

The thesis is structured as follows:

Chapter 2 presents a short description of the considered MW instruments and their characteristics that are of interest for the calibration process. Here, I also present the various missions of the individual instruments as well as the data sets that are available so far.

Chapter 3 introduces the problem of inter-satellite biases for the considered MW instruments. I discuss known biases as they are detected with different inter-comparison methods.

Chapter 4 is dedicated to the analysis of instrumental behaviour and noise evolution over the lifetime of each individual instrument.

Chapter 5 encompasses a detailed description of the calibration process and uncertainty propagation (Section 5.1), the characterisation of the various effects that have an impact on the measurement process (Section 5.2) and the presentation of the abilities of the new FCDR processor code (Section 5.3). The chapter also presents intermediate results of the generated unharmonised FCDR version (Section 5.4).

Chapter 6 presents the concept for the generation of a harmonised FCDR and the FIDUCEO approach to this problem (Section 6.1). Starting from a sensitivity study on the various effects that have an impact on the measurement process (Section 6.2), the harmonisation procedure is explained (Section 6.4) and first results are presented (Section 6.5).

Chapter 7 concludes the thesis. Problems and achievements are identified and a brief outlook is given on further steps towards consistent data records for climate research.

> ***Note:*** Chapter 4 of this thesis contains a reprint of the publication
>
> > Hans, I., Burgdorf, M., John, V. O., Mittaz, J., and Buehler, S. A.: Noise performance of microwave humidity sounders over their lifetime, *Atmos. Meas. Tech.*, 10, 4927-4945, `https://doi.org/10.5194/amt-10-4927-2017`, 2017

Chapter 2 contains parts of the above publication. The appendix of the above publication is part of the appendices of this thesis. Moreover, the contents of following project reports are partly included in the thesis:

> Hans, I., Burgdorf, M., Woolliams, E.: Product User Guide - Microwave FCDR Release 0.2, Universität Hamburg, National Physical Laboratory, 2018

> Burgdorf, M., Hans, I., Prange, M., Mittaz, J., Woolliams, E.: D2_2 Microwave: Report on the MW FCDR uncertainty, Universität Hamburg, National Physical Laboratory, 2017

2 Considered instruments and data: MHS, AMSU-B and SSMT-2

This thesis focuses on the three instruments: Microwave Humidity Sounder (MHS), Advanced Microwave Sounding Unit - B (AMSU-B) and Special Sensor Microwave Temperature (SSMT-2). These three instruments are cross-track-scanning, passive microwave radiometers flown on polar orbiting satellites. The design is very similar for the three instruments: Within five frequency channels, they measure the radiation emitted by the atmosphere to space. The five channels capture radiation from several levels of altitude in the troposphere. The main purpose is humidity profile sounding, therefore three channels are placed around the strong absorption line of water vapour at 183 GHz and provide information on upper levels of the troposphere. The other two channels penetrate much deeper into the atmosphere, also down to the surface. Note however, that the exact central frequencies are not the same for all instruments, although they are very similar. The similarities between the instruments are also in the basic calibration procedure: A two point calibration on a known warm and a known cold target is used to translate the measured signal from Earth to radiation. Within one calibration or scan cycle, the instrument measures the radiation emitted by the warm calibration target (a black body on board the instrument) and the radiation emitted by the cold target (the signal from viewing to deep space). Then, the instrument measures the radiation from Earth through a multitude of Earth views arranged in a scan line perpendicular to the flight direction of the satellite. After applying several corrections to the simple two point calibration to account for the instruments impact on the measurement, one obtains the final radiation for the Earth view. Via Planck's formula this is transformed into brightness temperature, which is the common unit of radiation in the humidity sounding context.

The following sections present the detailed characteristics that are essential for the calibration of the AMSU-B, MHS and SSMT-2 instruments and the construction of a fundamental climate data record based on these microwave radiometers. I also briefly present the concept of a polar orbiting satellite and introduce the various missions for the individual instruments and the available data set.

Note that SSMT-2 instruments are included in the FCDR production, although large parts of the work are focussed on MHS and AMSU-B. This is because documentation and available data is reduced for SSMT-2 and hence it is obvious to first establish the new FCDR generation and harmonisation for the AMSU-B and MHS instruments.

This chapter on used data is based on the short data section in the following publication and slightly extended for this thesis

Hans, I., Burgdorf, M., John, V. O., Mittaz, J., and Buehler, S. A.: Noise performance of microwave humidity sounders over their lifetime, *Atmos. Meas. Tech.*, 10, 4927-4945, `https://doi.org/10.5194/amt-10-4927-2017`, 2017

	Channel (in this thesis)	Orig. channel	Centre frequency in GHz	Total bandwidth in GHz	Pre-launch NEΔT in K
	1	4	91.655±1.250	3	0.6
	2	5	150.00±1.25	3	0.6
SSMT-2	3	2	183.31±1.00	1	0.8
	4	1	183.31±3.00	2	0.6
	5	3	183.31±7.00	3	0.6
	1	16	89.0±0.9	2	0.37
	2	17	150.0±0.9	2	0.84
AMSU-B	3	18	183.31±1.0	1	1.06
	4	19	183.31±3.0	2	0.70
	5	20	183.31±7.0	4	0.60
	1	H1	89.0	2.8	0.22
	2	H2	157.0	2.8	0.34
MHS	3	H3	183.31±1.0	1	0.51
	4	H4	183.31±3.0	2	0.40
	5	H5	190.31	2.2	0.46

Table 2.1 The basic instrumental characteristics of SSMT-2, AMSU-B and MHS. Note that all channels have two side bands, although rarely indicated for MHS channels 1, 2 and 5 where the side bands are very close to each other. The values for the NEΔT stem from the specifications for SSMT-2, from NOAA-15 for AMSU-B and NOAA-18 for MHS.

2.1 Spectral channel characteristics

Table 2.1 (data taken from the NOAA KLM User Guide (Robel et al., 2009) and from Table 1 in Kobayashi et al. (2017)) shows the basic channel characteristics for the AMSU-B, MHS and SSMT-2 instrument. Note that the AMSU-B channels are usually numbered together with the channels of AMSU-A (which is not considered here), resulting in channels 16 to 20. In this study however, the AMSU-B instrument is considered independently and therefore the channels are labelled 1 to 5 as it is for the MHS instrument. This is also to keep a stable relation between the channels' names and their sounding altitude across all instruments: Channel 1 is the surface channel. Channel 2 detects radiation from lower levels of the troposphere whereas channel 3 to 5 are the main sounding channels for water vapour in the mid- to upper troposphere. For very dry conditions the surface may have an impact, too.

For AMSU-B and MHS, the water vapour sensitive channels are channels 3-5 with frequencies 183±1, 183±3 and 183±7 GHz (only 190 GHz for MHS) around the 183 GHz water vapour absorption line (see Table 2.1 for the precise values). They provide information on the tropospheric humidity. Channel 1 is at 89 GHz and channel 2 at 150 GHz (157 GHz for MHS); both offer a deeper view through the atmosphere down to the surface. See Fig. 2.1a for the position of the channels in the electromagnetic spectrum and corresponding opacity of the Earth's atmosphere. For SSMT-2, the order of the channels is different. The original channel 1 is at 183±3 GHz, channel 2 at 183±1 and channel 3 at 183±7 GHz. The original channels 4 and 5 are surface channels placed at 92 GHz and 150 GHz, respectively (NOAA-CLASS, 2016). However, again for simplicity and in order to fit the MHS-relation between the channels' names and their sounding altitude we will use the MHS naming of channels for SSMT-2 as well and refer to the water

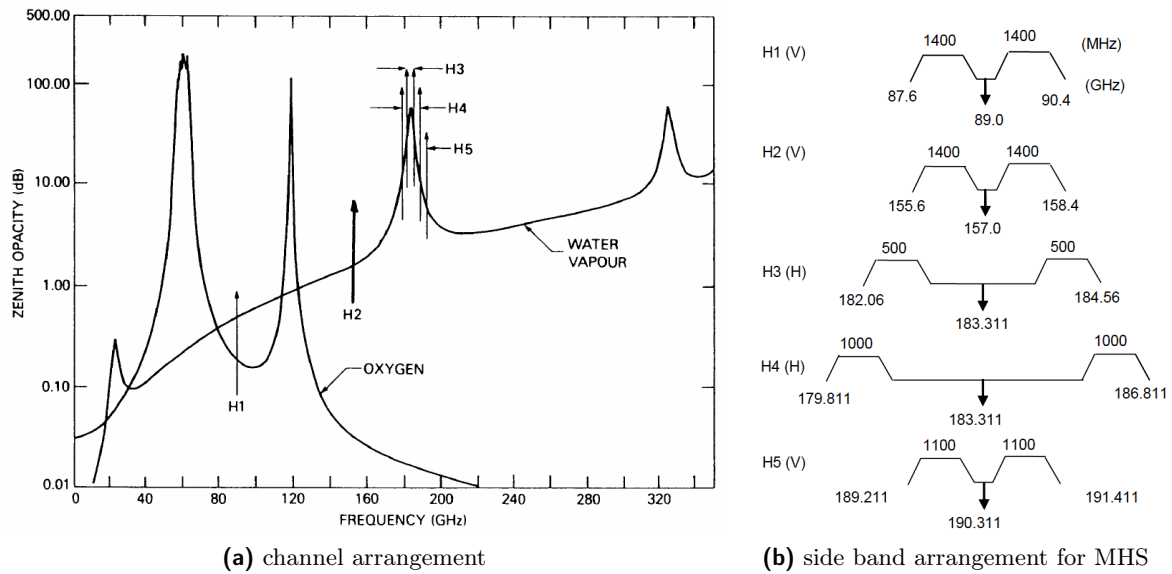**(a)** channel arrangement   **(b)** side band arrangement for MHS

Figure 2.1 The electromagnetic spectrum and opacity of the atmosphere for the spectral range of AMSU-B/ MHS/ SSMT-2 channels (indicated in the graphic (a) as channel H1-H5 in MHS numbering). (b) shows the side band arrangement for MHS. Figures from Robel et al. (2009), with friendly permission by NOAA.

vapour channels at 183±1, 183±3 and 183±7 GHz as channels 3, 4 and 5. The surface channels at 92 GHz and 150 GHz are labelled as channel 1 and 2. Note that the actual frequencies are not exactly the same for the different instruments, even though we will refer to them as one channel, e.g. the "89 GHz channels" encompass the 89 GHz channels of AMSU-B and MHS, but also the 92 GHz channel of SSMT-2. For the produced new MW-FCDR data sets however, we keep the original channel names. Information on the frequency and, for SSMT-2, on the corresponding channel in MHS/ AMSU-B is provided in the variables' attributes of the data sets.

Note that all channels have two side bands around the centre frequency. This is also valid for the channels 1, 2 and 5 of MHS (see Fig. 2.1b). Most often however, they are indicated as 89.0 GHz, 157.0 GHz and 190.31 GHz only, because the side bands are very close to each other. Only for the channels 3 and 4 (and 5 for AMSU-B and SSMT-2), the side bands are significantly separated such that they are placed above and below the 183.31 GHz water vapour absorption line.

For all channels, the recorded signal is averaged over the two side bands placed at slightly higher and lower frequencies around the centre frequency and further processed. The relative spectral response functions (RSRFs) for the two side bands is the same, mirrored at the centre frequency of the channel. For the processing of level 1b to level 1c, i.e. performing the photometric calibration and transformation to brightness temperature needed for the FCDR, the knowledge of the RSRF is not necessary. However, for the step towards a CDR, the knowledge of the RSRF becomes important. It is crucial to know on which frequencies the instrument actually operates in order to correctly simulate the output of the instrument due to the emission of the atmosphere by using radiative transfer models (RTMs). If the frequency of the instrument is not assumed correctly within the RTM, a systematic error occurs even if the instrumental calibration is perfect, simply because the instrument measures something different (i.e. at another frequency) than the RTM believes it to do. Therefore, the new FCDR provides information on the RSRF from pre-launch
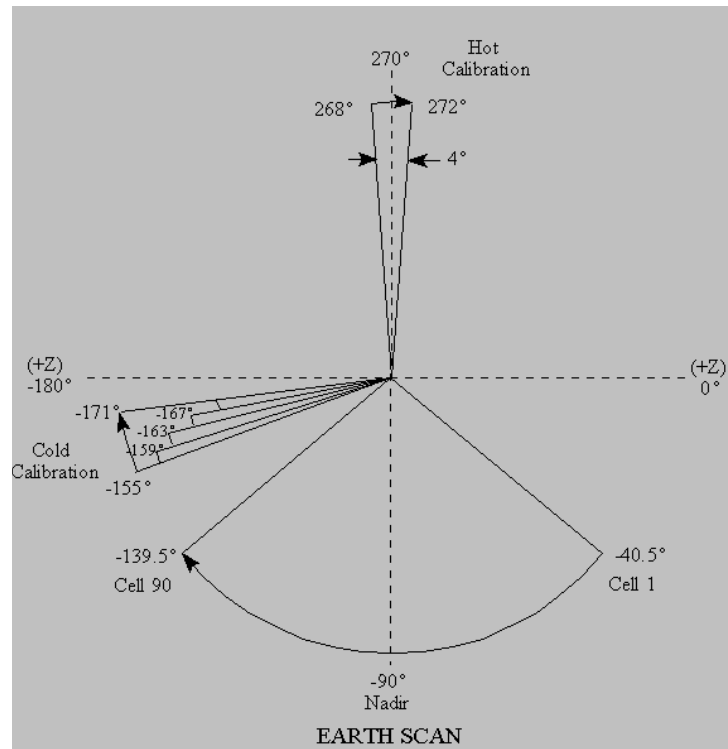
Figure 2.2 The scan motion of AMSU-B. Figure from Robel et al. (2009), with friendly permission by NOAA.

measurements, where available (so far only MHS on NOAA-18 and NOAA-19 satellites). For the other instruments, assuming a boxcar function in the RTM, with width and centre frequency as given in table 2.1, results in an uncertainty of below 0.1 K (Hans et al., 2018).

2.2 Scanning characteristics

MHS, AMSU-B and SSMT-2 are cross-track scanning instruments, capturing a line of 90 (28 for SSMT-2) contiguous Earth views per scan (45 (14) on each side of the sub-satellite point). For MHS, the antenna beamwidth is 1.1° for all channels and views and translates to a diameter of 16.3 km for the nadir fields of view (FOVs) at a nominal satellite altitude of 870 km (Robel et al. (2009), App.J3). See table 2.2 for other instruments. Thus, the instrument covers about 50° to each side of nadir resulting in a swath width of about 2180 km (WMO-OSCAR, 2016). For each calibration cycle of 8/3 s (8 s for SSMT-2), one scan line of Earth views is recorded. The calibration cycle for AMSU-B is displayed in Fig. 2.2. It is in principle the same for MHS. Four contiguous views of the internal warm calibration target (IWCT) are taken (for SSMT-2 the reflector stays at the position of 130.5° for the IWCT). Then the reflector's rotation is accelerated to reach the start position for the Earth views. At constant velocity the signal for the 90 (28) views is measured, before the reflector is accelerated again to reach the deep space views (DSVs). Four views of this cold target are recorded (again SSMT-2 stays at 229.5° for the DSV), before the reflector is accelerated to reach the warm calibration target again. One full rotation takes 8/3 s (8 s) and defines one calibration cycle (Robel et al., 2009).

The definition of the viewing angles in Fig. 2.2 is arbitrary and differs between the instruments. To later match the requirements for the polarization corrections in the calibration, the nadir field

	Channel	swath width in km	nominal beamwidth in °	FOV diameter (Nadir) in km	number of Earth views
	1		6.0	88	28
	2		3.7	54	28
SSMT-2	3	1500	3.3	48	28
	4		3.3	48	28
	5		3.3	48	28
AMSU-B		2259	1.1	16	90
MHS		2180	1.1	16	90

Table 2.2 The basic scanning characteristics of SSMT-2, AMSU-B and MHS (NOAA-CLASS, 2016; dms, 2004; Robel et al., 2009; WMO-OSCAR, 2016)

of views must be placed around 0° (see Sec. 5.3.2.5). Hence, the shaft positions from the raw data records are translated to this configuration in our processing.

2.3 Calibration targets

All three instruments use the deep space, i.e. the signal from the cosmic microwave background (CMB) as their cold calibration target. The corresponding brightness temperature of the CMB is a well known quantity: 2.72548 ± 0.00057 K (Fixsen, 2009). In the calibration procedure, a correction is made to this value to account for influences of the instrument and its surroundings on the measurement of the signal.

As internal warm calibration target, all instruments carry a black body, whose physical temperature is measured with five, seven or two platinum resistance thermometers (PRTs) for MHS, AMSU-B and SSMT-2 respectively; all PRTs are placed at the back of the black body. The black body of AMSU-B is built by a magnesium alloy substrate with an array of pyramidal tines having an aspect ratio of 4: 1. This is covered by a 1.3 mm layer of Eccosorb CR114. Thus, a very high emissivity of 0.9999 is achieved (Saunders et al., 1995). The temperature of the black bodies of the instruments is not actively controlled, but it is left free-floating. The black body is isolated against thermal instabilities by its large thermal inertia (Saunders et al., 1995), thus keeping the temperature relatively stable.

2.4 Polar orbiting satellites

The three instruments fly on polar orbiting satellites. This type of satellite has a low Earth orbit (LEO) of about 850 km altitude, in contrast to a geostationary orbit (GEO) which has an altitude of 35780 km (NOAA, 2017) and a constant position of the sub-satellite point on Earth. A polar orbiting satellite is a special case of the LEO satellites: its orbital plane has an inclination close to 90° against the equator. Thus, the satellite swath passes both poles. Moreover, the polar orbiting satellites that the AMSU-B, MHS and SSMT-2 instruments fly on, have a sun-synchronous orbit. For the sun-synchronous orbit, the orbital parameters are adjusted such that the angle between the orbital plane and the sun remains constant (EUMETSAT-webpage, 2018). As a consequence, the sub-satellite point always crosses the equator at the same local time, which is important for collecting consistent data time series. Note however, that the NOAA
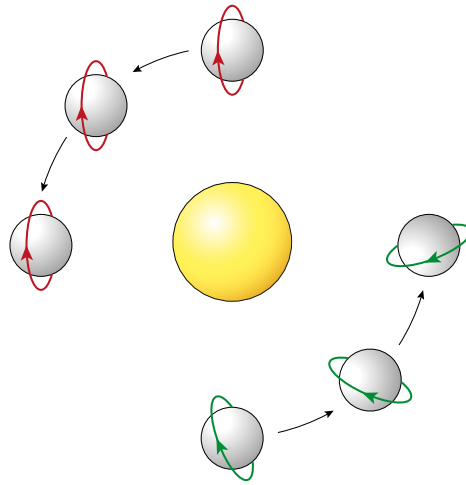
Figure 2.3 Fix orbit (red) and sun-synchronous orbit (green). Figure courtesy Oliver Lemke.
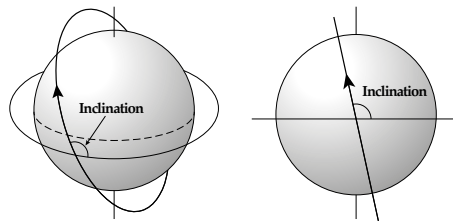

Figure 2.4 A satellite in a high inclination orbit around Earth. Figure courtesy Oliver Lemke.

satellites experienced an orbital drift over their lifetime so that their local equator crossing time (LECT) has slowly changed since launch (Ignatov et al., 2004). See Figures 2.3 and 2.4 for a sketch of a sun-synchronous polar orbit with the high inclination. The polar orbiting satellites enable a (almost) global coverage twice a day thanks to the broad swath of more than 2000 km (EUMETSAT-webpage, 2018). Figure 2.5 shows a typical picture of the global coverage. Another advantage of a polar orbiting satellite is a relatively high spatial resolution due to the low orbit, and a short revolution period that guarantees frequent coverage. Nonetheless, in contrast to a geostationary satellite, continuous measurements are impossible. Both types of satellites are therefore in wide use.

2.5 Missions

The three instruments have been used for a total of eleven missions of polar orbiting satellites, covering a time range of 25 years from 1991 to 2016: four for SSMT-2, three for AMSU-B and four for MHS. SSMT-2, being the oldest of this type of radiometer, was first launched in 1991 on the satellite F11 by the Defense Meteorological Satellite Program (DMSP). Missions on F12, F14 and F15 followed in 1994, 1997, and 1999. In 1998, the National Oceanic and Atmospheric Administration (NOAA) launched its satellite NOAA-15, being the first one of the 5th generation of polar orbiters. It carried the first AMSU-B instrument (see Fig. 2.6 for a drawing of the satellite and its payload). It was followed by NOAA-16 and NOAA-17
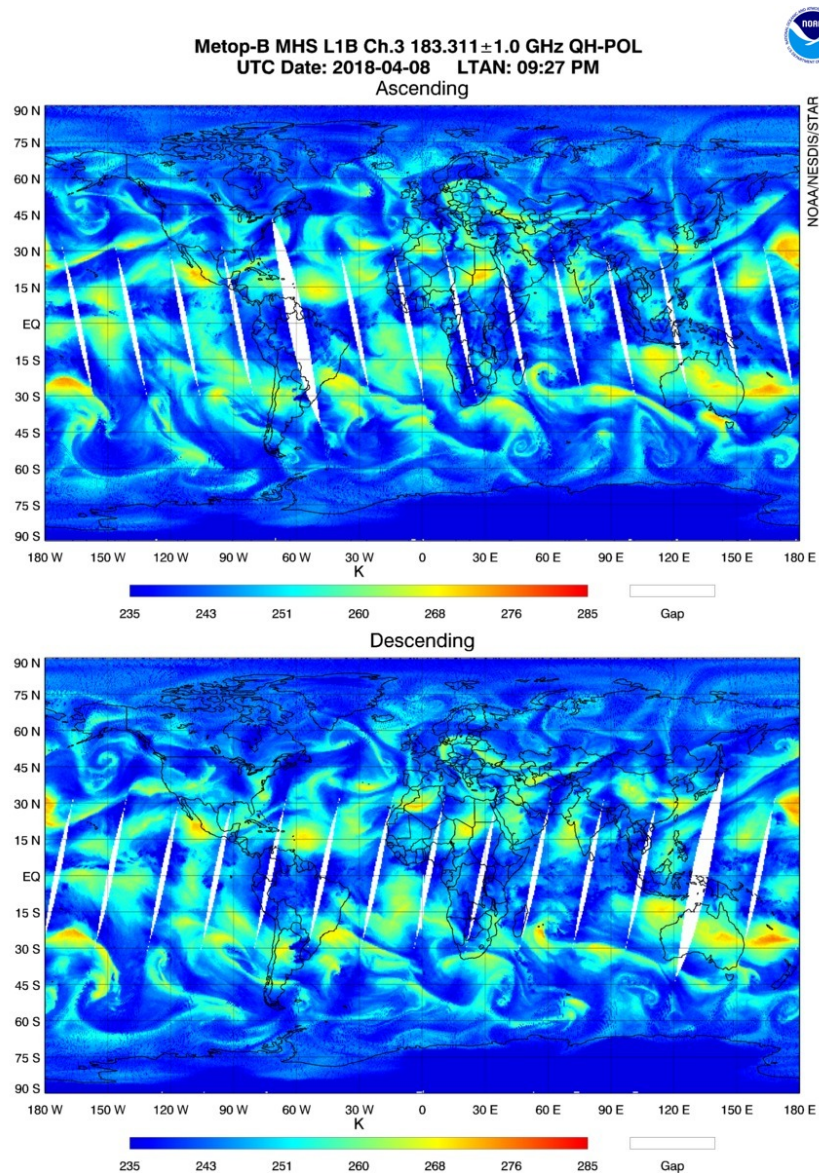
Figure 2.5 Global coverage of ascending and descending branch for MHS on Metop-B. Combined picture of several orbits of operational data. Figure from NOAA-STAR (2015). With friendly permission by NOAA/ NESDIS Center for Satellite Applications and Research.

satellites in 2000 and 2002. The successive satellites NOAA-18 and NOAA-19, launched in 2005 and 2009 already have the newer MHS instrument on board. The European satellites Metop-A and Metop-B (launched in 2006 and in 2013), controlled by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) also carry the MHS instrument. Figure 2.7 gives an overview of the missions and the equator crossing times of the satellites. Note that the Metop-satellites are non-drifting due to their controlled orbit.

Figure 2.6 Drawing of the configuration of the 5th generation of polar operational environmental satellite (POES). Figure from Kramer (2018), image credit: eoPortal, NASA.

2.6 Available Level-1b data

For the AMSU-B and MHS instruments both level 1b (containing raw counts) and level 1c (processed to brightness temperature) data files are available. The level 1b data records are downloaded from the NOAA CLASS (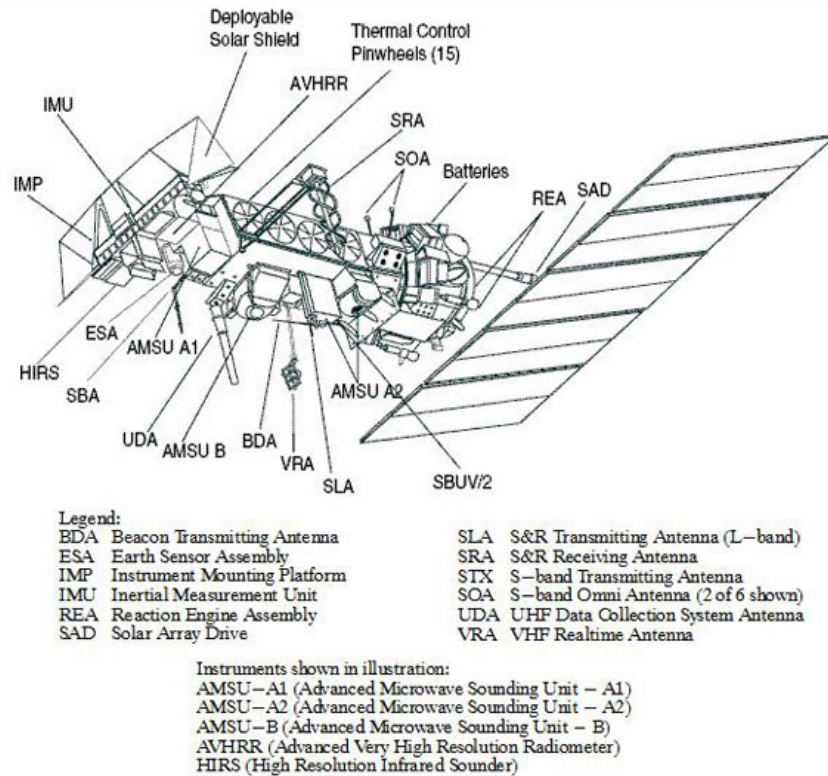Comprehensive Large Array-data Stewardship System) archive (NOAA, 2018). For SSMT-2 there are some inconsistencies regarding the time range of available data: On the NOAA NCEI (National Centers for Environmental Information, formerly NGDC, National Geophysical Data Center) data availability web page, there are longer time frames indicated for which SSMT-2 data should exist (reaching back to 1992) than on the NOAA CLASS page. This larger data set of SSMT-2 data has been reformatted to NetCDF by John and Chung (2014) and covers the range according to NCEI (shown in Fig. 2.7). But this is not the raw file providing all information that goes into the calibration and that we aim to look at and that we will use for FCDR production. For example, the NCEI-data file does not contain the temperature measured on the internal black body. Hence, to stay in line with the investigation of AMSU-B and MHS data obtained from NOAA CLASS, we only used the data for SSMT-2 that NOAA CLASS provides and that cover the time range indicated on the NOAA CLASS website (NOAA-CLASS, 2016). The data record format for the level 1b data that we use here is documented for AMSU-B and MHS in the NOAA KLM User Guide (Robel et al., 2009) and on the NOAA CLASS website for SSMT-2 (NOAA-CLASS, 2016).

Each data record file corresponds to approximately one orbit. However, there are also longer and shorter files covering only parts of one orbit. Moreover, there is always an overlap of some
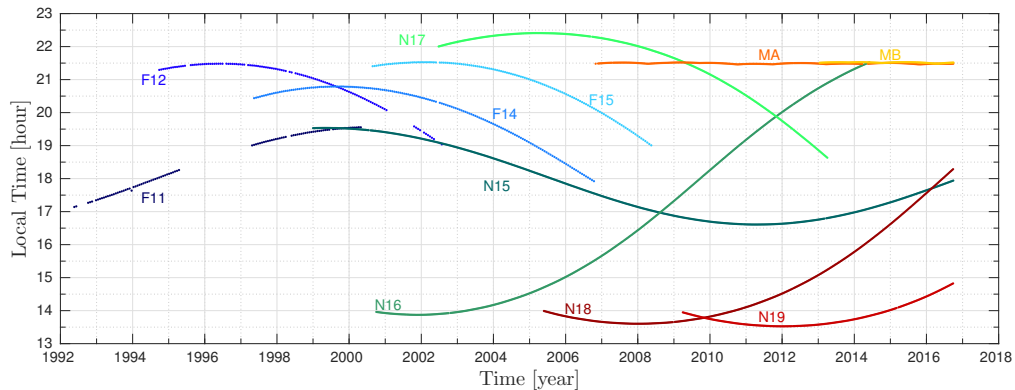
Figure 2.7 The local equator crossing times (ascending branch) for the satellites considered here. The graphic shows all times for which there is any data available from the instrument on board the respective satellite (regardless of quality issue). For the F11 to F15 satellites, we only look a part of the time series shown here, due to inconsistencies across the different sources of data (see text in section 2.6).

minutes between subsequent files. This makes any investigation on the evolution of the atmosphere complicated, since those double data need to be removed. A standardised format for ready-to-use FCDR files with equal length is one of the aims in the generation of a new MW FCDR within this thesis.

To read the level 1b data files, we have developed our own reader that is used in our processing. The level 1b files contain quality and status information and the information needed for the calibration, i.e. the raw counts from the signals of the calibration targets and Earth view as well as the temperature (or raw counts for AMSU-B) measured by the PRTs on the IWCT. Further calibration information is in the header or in separate files `clparams.dat` and `fdf.dat` delivered with the AAPP (ATOVS and AVHRR Pre-processing Package). The AAPP is used operationally to convert the level 1b data to level 1c. The level 1c data contain the brightness temperature for all channels, the geolocation information including satellite and solar angles and some quality information. Information on the noise equivalent differential temperature (NEΔT) is only provided with the AAPP 7-13 version for MHS. This gap will be closed with the FCDR generated in this work, since the FCDR will contain extensive uncertainty information.

For SSMT-2, the level 1b and level 1c data are not separated, but the level 1b file containing the raw counts already provides the computed brightness temperature, too.

In this work, the level 1b files for each instrument, i.e. the raw counts are used together with all required calibration information to process the data to a new level 1c data set.

3 The problem: Existing biases between MHS, AMSU-B and SSMT-2 instruments

Having eleven instruments of the same family that cover more than 20 years, is a good start for considering climate monitoring with these instruments' data sets. However, there is still a major drawback: So far, the output measurements of the individual instruments are not comparable. This is because there are biases between the instruments, i.e. the measured brightness temperatures of two instruments are not the same where they should be (e.g. because the instruments measure the same region of the atmosphere at the same time using the same frequency, or they should give the same mean brightness temperature for a region or time). Due to these biases, a comparison of the retrieved atmospheric state from the existing data of different instruments is not valid. Furthermore, even a comparison of measurements of the same instrument from different time periods or different regions is not valid in the case that the instrument experiences instabilities. These instabilities and the inter-satellite biases caused by them may mask natural trends or feign trends when comparing data from an individual instrument or between several instruments, respectively. In order to analyse real climate trends, stable and bias-free data sets are highly required for a valid comparison of data from different instruments, time periods, and regions.

Before setting out to investigate the causes for the biases, I discuss the existing biases between the individual instruments as they are revealed by different methods of inter-comparison. It is these existing biases that the harmonisation procedure later attempts to understand and reduce.

3.1 Biases revealed by different methods

Comparing two radiometers on board two polar orbiting satellites is no trivial task. Ideally, both instruments should measure the emitted radiation of a perfectly known target. This would allow for the true comparison of the instruments. In this case, one could optimise the calibration of one instrument to match the other. In this special case of a perfectly known target, even an absolute calibration and comparison would be possible. Unfortunately, there is no consolidated method for this absolute calibration, although there are attempts to use the moon as target (Burgdorf et al., 2016, 2018). Since the ideal case of a known target cannot be achieved yet, we have to use other methods to deduce at least the relative biases from one instrument to the other. A task, that is still not trivial: We have to make sure, that the instruments are measuring the same thing (although we do not know what the "thing" truly gives as signal). This is difficult, since the instruments are carried by different satellites in different orbits with different LECT, hence measuring any chosen region at different times of the day. To approximate this "same thing", there are different methods. The biases revealed by the different methods also contain slightly different information, since they are based on different sampling. In the following, I present the biases revealed by

1. Collocation methods, i.e. Simultaneous Nadir Overpasses (SNOs) and Simultaneous All Angle Collocations (SAACs)

2. averaging methods, i.e. global or zonal averages of monthly means

3. natural calibration targets, that are a special case of averaging methods

4. a method based on percentile-analysis that was recently developed in the Master's thesis by M. Prange (not yet submitted).

All methods make use of averaging at some point. Note however, that I choose the term "averaging methods" to summarise the methods that first compute averages of the data and afterwards compare these averages of two instruments to obtain a bias. This term "averaging methods" should therefore distinguish these methods from the collocation methods. The collocation methods compare two individual collocated measurements directly and afterwards deduce the overall bias from the mean of the differences of the individual measurements. The FIDUCEO harmonisation itself is based on the collocation methods. As an independent comparison method, I use the percentile analysis method to validate the harmonisation results.

3.1.1 Collocation methods

The idea of collocation methods reflects the most intuitive way of comparing the measurements from two instruments: Two instruments on board two different satellites look at the same time at the same target (latitude-longitude pair) with a similar viewing geometry (John et al., 2013a). The most important advantage of this method is, that any sampling issues due to natural variability like the diurnal cycle are minimised, if "the same time" and "same target" are defined sufficiently strict. Hence, the collocation method should provide a direct view onto the inter-satellite bias. This makes it interesting for inter-calibration efforts. A disadvantage of the collocation method is, that it is not well suited for surface channels due to the high variability of surface emissivity (Iacovazzi Jr. and Cao, 2008; John et al., 2012). This is even a problem for the sounding channels: in very dry atmospheres, as they often occur in high latitudes, even the sounding channels may receive a large contribution from the surface. Unfortunately, the significance of this issue is increased, since the collocations mostly occur at high latitudes due to the configuration of the orbits (John et al., 2012) (only for short periods of similar LECT, global collocations of two satellites may be obtained). This sampling at higher latitudes also results in a small range of sampled brightness temperatures. Hence, inter-satellite biases that vary over the full measured brightness temperature range are poorly characterised if comparing only collocations from high latitudes (Zou et al., 2006; John et al., 2012).

The collocations of instruments on two satellites are usually obtained with the constraint on comparing measurements with a similar viewing angle. The SNO method only uses the subset of collocations with near-nadir viewing angles for both instruments (John et al., 2012, 2013a), whereas the SAAC method uses also collocations of the off-nadir FOVs with similar viewing angle for both instruments (John et al., 2013b). In the following, I present the main results from those methods.

3.1.1.1 Simultaneous Nadir Overpasses (SNOs)

The method using SNOs for inter-calibration of radiometers on board sun-synchronous polar orbiters was developed by Cao et al. (2004). Efforts for inter-calibration of the High-resolution
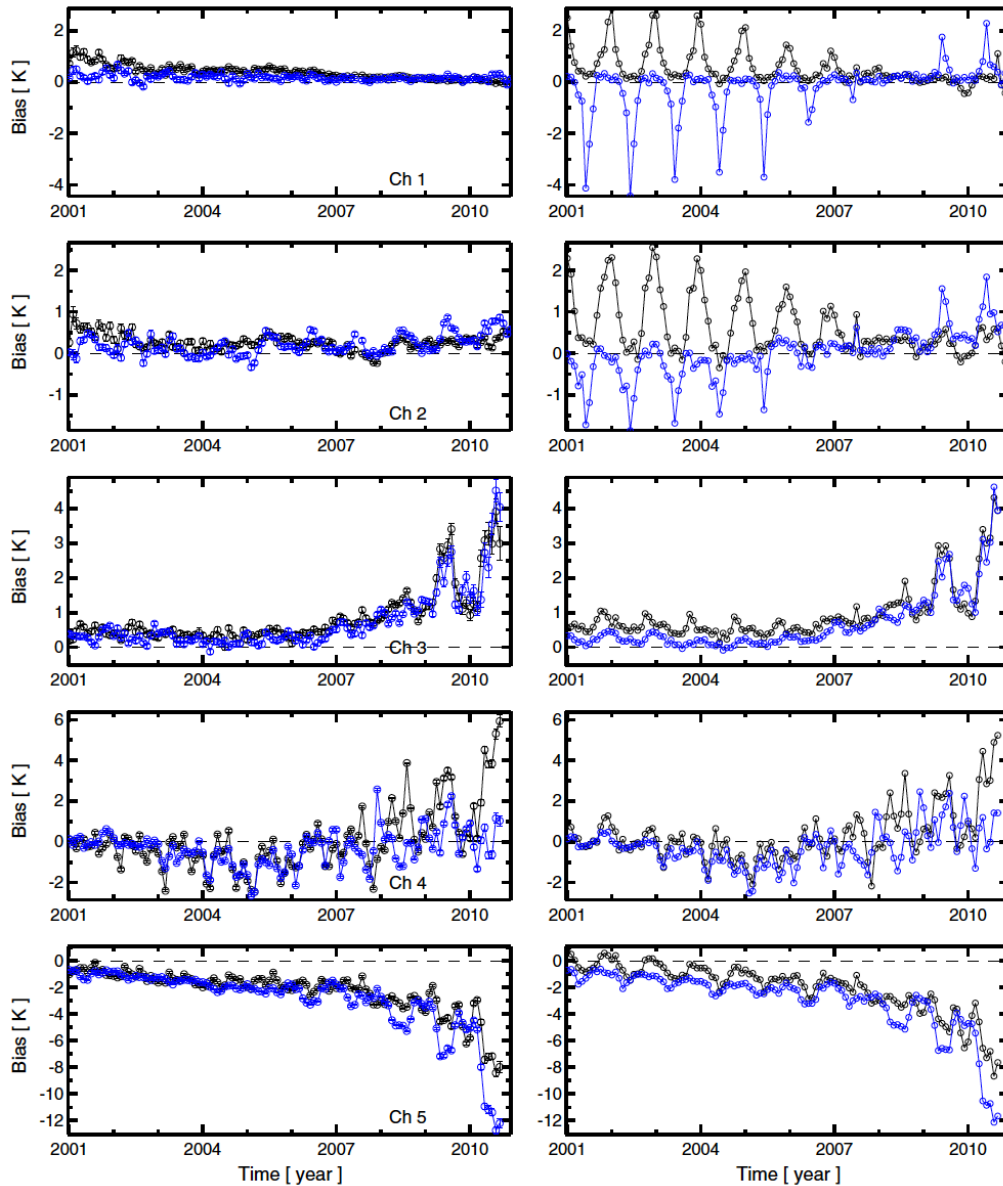
Figure 3.1 The inter-satellite biases (NOAA-16 – NOAA-15) estimated from SNOs (left panel) and zonal averages with ascending and descending passes combined (right panel). Black symbols represent data from Antarctica (SNOs, left panel) or from 70°S–80°S (zonal averages, right panel). Accordingly, blue symbols represent data from the Arctic or from 70°N–80°N. Figure from John et al. (2013a), with friendly permission by American Geophysical Union (AGU).

Infrared Radiation Sounder (HIRS) and of AMSU-A instruments were undertaken in e.g. Zou et al. (2006) and Shi et al. (2008). John et al. (2012) apply the SNO method to AMSU-B and MHS data for periods where global SNOs occurred. In John et al. (2013a), they present the SNO method in comparison to other inter-comparison methods by analysing biases of the NOAA-16 AMSU-B against NOAA-15 AMSU-B.

Figure 3.2 Time evolution of inter-satellite biases for channel 3 against MHS on NOAA-18, obtained from SNOs. Colours: yellow: NOAA-15, black: NOAA-16, magenta: NOAA-17, red: NOAA-18, cyan: Metop-A, green: Metop-B. Figure from Burgdorf and Hans (2016).

The allowed time difference is 5 minutes to avoid varying scene through cloud impact. The spatial distance threshold is 5 km, which is one third of the FOV-diameter at nadir. As "nadir", the four FOVs to either side of nadir are used (John et al., 2012).
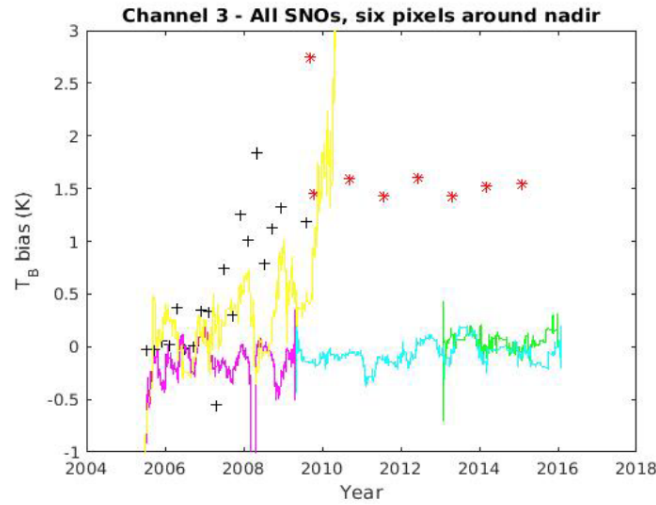
As pointed out above, the SNO events mostly occur at high latitudes. Due to orbital drift however, some satellite pairs approach each other in their LECT. In these periods of similar LECT, SNOs may occur over the whole globe. This was actually the case for the pairs NOAA-16 – NOAA-15 (in August 2008), Metop-A – NOAA-17 (in May/ June 2009), NOAA-18 – NOAA-19 (in September 2009). For these periods, the whole brightness temperature range could be sampled with SNOs and compared among the instruments (John et al., 2012). Biases varying with latitude and scene radiance (or brightness temperature) of either instrument were presented. These biases were strongest for channels 3, 4 and 5 for the NOAA-16 – NOAA-15 pair. It should be noted that at this period, both instruments had already severely degraded (see Sec. 4.2). Therefore, taking the brightness temperature of one instrument as the reference for the other does not shed light on individual instrumental problems and may even be misleading due to increased noise level (Sec. 3.1.4, 4.2). The biases for the Metop-A – NOAA-17 pair are rather small, usually below 1 K (John et al., 2012), but show some dependence of brightness temperature. Only channel 2 has a large bias, because AMSU-B and MHS channel 2 frequencies differ. For the sounding channels, biases tend to be higher at higher latitudes and lower temperatures. The NOAA-18 – NOAA-19 pair, however, should not be considered for the 2009-period at all, since the MHS on NOAA-19 experienced major problems in the sounding channels 3 and 4. The instrument was out of specifications and showed erratic behaviour from June to October 2009. In this time, an attempt was made to save the instrument by readjusting the gain, but it was not fully successful: the instrument stabilised in October 2009, but did not regain its specified abilities (see Sec. 4.2, 8.1, and in NOAA-OSPO (2015)). Unfortunately, two of the three considered periods of global SNOs are affected by major instrument degradation and noise increase, which lead to partly misleading results in deducing brightness temperature dependent biases (see also Sec. 3.1.4). Nonetheless, the study showed that biases varying with latitude and brightness temperature

exist (e.g. Metop-A – NOAA-17) and, as an important consequence, near-polar SNOs can only reveal part of the inter-satellite biases and should not be used alone for inter-calibration (John et al., 2012). This issue is addressed within FIDUCEO (see Chapter 6 on Harmonisation). In Sec. 3.1.4, I revisit the question of brightness temperature dependent biases.

Bias as function of time is investigated in John et al. (2013a): the NOAA-16 – NOAA-15 bias is computed over the period 2001-2010. The main result is the strong increase of the bias with time for channel 3, 4 and 5, the latter one showing a negative bias (see Fig. 3.1). This result was confirmed by using zonal averages for the high latitude bands as shown in the right panel of Fig. 3.1. This time dependent bias follows the evolution of the instrumental degradation (see Sec. 4.1).

In Burgdorf and Hans (2016), we present inter-satellite biases as function of time, found using the FIDUCEO collocations ("match-ups" generated according to Block (2015), $\Delta t < 5\,\text{min}$, $\Delta s < 5\,\text{km}$) for all AMSU-B and MHS instruments against MHS on NOAA-18. Figure 3.2 shows the time evolution for inter- satellite bias for channel 3 obtained from the SNO-match-ups (with "nadir" meaning 3 FOVS to either side of nadir) for global data, ascending and descending passes averaged. Note that the found biases are obtained from mostly near-polar SNOs and therefore do not sample the high end of the temperature range. Again, a strong increase of the bias of NOAA-16 and also NOAA-15 is visible against NOAA-18. NOAA-17 shows a small, slightly negative bias. NOAA-19 shows a strong increase in 2009 during its erratic period and afterwards stays at a constant high bias level. Metop-A and Metop-B roughly agree with NOAA-18, slightly shifted to a negative, but decreasing bias for Metop-A and a relatively stable positive bias for Metop-B.

3.1.1.2 Simultaneous All Angle Collocations (SAACs)

Using SNOs, any inter-satellite bias depending on the scan angle of the instrument cannot be sampled. However, scan-asymmetries have been observed for AMSU-B (Buehler et al., 2005), indicating that scan-dependent biases might exist.

As modification of the SNO method, John et al. (2013b) developed the Simultaneous All Angle Collocations (SAACs), matching the corresponding FOVs for the two instruments (FOV 1 to FOV 1,...FOV 90 to FOV 90, hence no limb effect adds to the bias). They present scan-dependent biases averaged over years 2001-2005 and 2006-2010 for combinations of NOAA-15, NOAA-16, NOAA-17, NOAA-18 and Metop-A. The largest scan-dependent bias of about 15 K is observed for channel 4 of NOAA-15. In fact, all channels of NOAA-15 show scan-dependent biases, especially for the sounding channels in the later period of 2006-2010. The same is true for NOAA-16. NOAA-17 and NOAA-18 show smaller biases (see Fig. 3.3 for the scan-dependent biases for channel 3). The qualitative pattern of scan-dependent biases appears to be quite constant over time. They are also seen in zonal average procedures in Hanlon and Ingram (2016), presented in Sec. 3.1.2.1, where also data from NOAA-19 and Metop-B was analysed.

The origin of these scan-dependent biases needs to be investigated in the frame work of an inter-calibration effort like FIDUCEO. Some effects within the measurement process might cause scan-dependent biases (see Sec. 5.2), since they have scan dependence themselves. In fact, already during the in-orbit verification phase, large scan-dependent biases due to radio frequency interference (RFI) were observed for NOAA-15 (Atkinson and McLellan, 1998). These RFI biases could be corrected to a certain degree with a devised correction scheme (Atkinson, 2001) (see Sec. 5.2). Since the correction was not updated, its validity for later years of the instrument's life is questionable. Therefore, scan-dependent biases are likely to occur. In John et al. (2013b), the
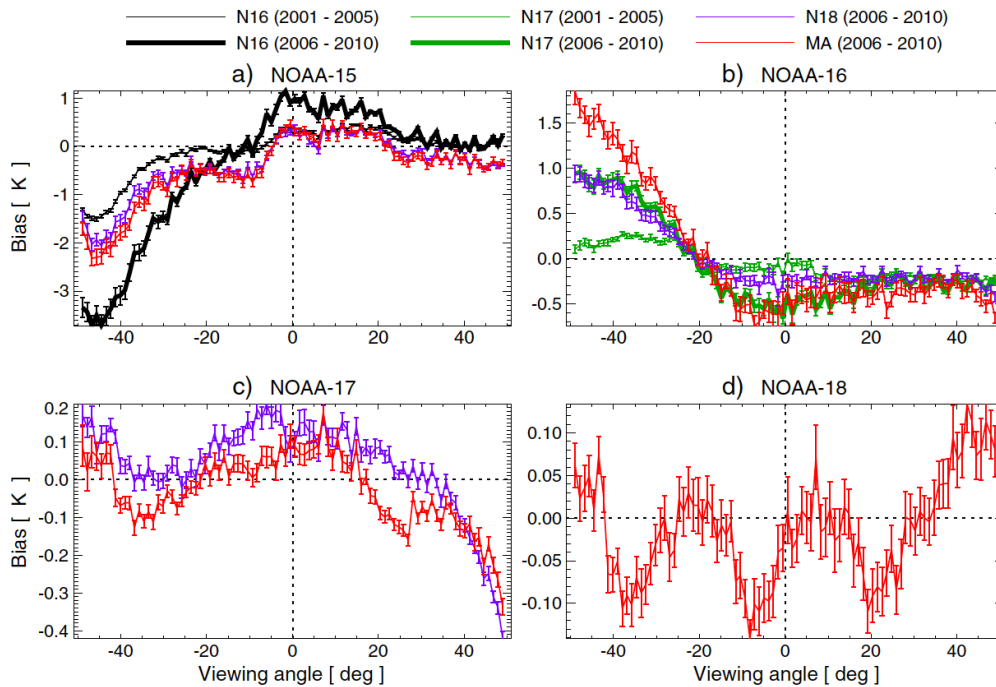
Figure 3.3 Scan-dependent biases of channel 3 (183±1 GHz) for different satellite combinations. For each sub-panel, (a) NOAA-15, (b) NOAA-16, (c) NOAA-17, and (d) NOAA-18, this particular satellite is subtracted from the satellites indicated in the legend. Mean biases are shown for the periods 2001–2005 and 2006–2010, when data is available for the individual satellites. Note that the y axis range of each panel is different. Figure from John et al. (2013b), with friendly permission by American Geophysical Union (AGU).

NOAA-15 scan-dependent biases are correctly identified as RFI issues. Also for NOAA-16, the possibility of RFI due to decreased gain in the sounding channels is suggested. However, this idea and its consequences are not discussed further in John et al. (2013b), nor investigated for other instruments. In Chapter 6, I will address this issue of RFI in the framework of harmonisation.

3.1.2 Averaging methods

A different approach than taking simultaneous measurements to compare two instruments, is to rely on a mean state of the atmosphere that both instruments should capture in the same way. Inter-satellite biases due to systematic differences should then shine through since small scale natural variability is reduced by the averaging process. This averaging might be executed over different spatial and temporal scales: confined targets can be used, meaning certain regions on Earth only. This was done in John et al. (2013a), see Sec. 3.1.3. Another way is to use latitude bins and compute zonal averages (Shi and Bates, 2011; John et al., 2013a; Hanlon and Ingram, 2016) or to investigate global means (Hanlon and Ingram, 2016). For temporal averaging, monthly means (Hanlon and Ingram, 2016) or even averages over longer periods (Hanlon and Ingram, 2016) can be used. For all types of averages, it should be considered that interesting features or dependences might be averaged out or that differences are misinterpreted as inter-satellite biases that in reality might be only due to different sampling times of the instruments and hence impact of diurnal cycle (John et al., 2013a). Looking at data from ascending and

descending passes separately may highlight some impact of the diurnal cycle. In order to see inter-satellite biases, both passes should be averaged however. This removes at least a part of the asymmetric diurnal cycle (Chung et al., 2007; John et al., 2012, 2013a). A clear advantage of comparison methods using averages is the availability of all time periods and global data since there is no restriction to polar regions or times of similar LECT as there is for the SNOs.

3.1.2.1 Zonal averages

Hanlon and Ingram (2016) give a report on the UK Met Office efforts to generate an FCDR for the Microwave humidity sounders by bias analysis and bias correction. Their investigation is based on averaging methods. Data are filtered for extreme cases (below 90 or above 320 K), high orography and clouds. However, the cloud filter may also exclude parts of Antarctica for channel 3, since the applied threshold is surpassed (comparing channel 5 and channel 3 may suffer from surface influence for dry atmospheres (Buehler et al., 2007) and wrongly classify data as cloudy).

For zonal avergages, Hanlon and Ingram (2016) arrange the data in 10° latitude bins. Monthly means are computed for each latitude bin, which are then further averaged over a long period of several years. Biases of each AMSU-B and MHS against MHS on NOAA-18 are then investigated for each latitude bin as function of FOV. This reveals scan-dependent biases. As in John et al. (2013b), strong scan-dependent biases are seen for both NOAA-15 and NOAA-16. In Fig. 3.4, the bias of NOAA-15 against NOAA-18, that means NOAA-15 – NOAA-18, is plotted as function of the FOVs 1-90. The scan dependence of the bias is very similar to the pattern seen in Fig. 3.3 from John et al. (2013b) who computed NOAA-18 – NOAA-15. Hence, note the opposite sign of the bias. This scan dependence of the bias , reproduced with the zonal averages by Hanlon and Ingram (2016), was assigned to RFI effects by John et al. (2013b). Quantitatively, the biases differ slightly: the bias obtained from the zonal averages is shifted up by about 0.5 K towards a
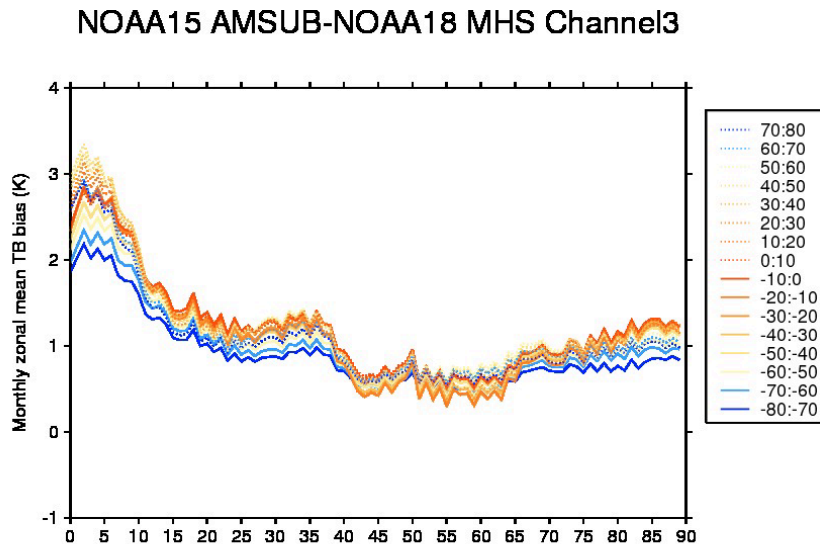
Figure 3.4 Scan dependence of NOAA-15 – NOAA-18 zonal mean bias for channel 3, averaged over all months June 2005 – Jan 2009. The monthly zonal biases are computed for each field of view separately and each 10° latitude band which is shown with a different coloured line, the northern hemisphere by dotted lines and the southern hemisphere by solid lines. Figure and caption from Fig. 18 in Hanlon and Ingram (2016), with friendly permission by Helen Hanlon and William Ingram.

"warmer" NOAA-15 relative to the bias obtained from the SAACs. Note that the zonal averages are subject to influences from the diurnal cycle, whereas the SAACs are not.

Other instruments than AMSU-B NOAA-15 or NOAA-16 show much weaker scan-dependent biases. However, it is an interesting fact that the bias patterns also for other instruments, e.g. NOAA-19, appear in each latitude bin over an averaged period of several years (Hanlon and Ingram, 2016). This suggests a quite stable underlying scan-dependent issue also for other instruments than NOAA-15 for which we know of scan-dependent issues due to RFI.

The time dependent bias for NOAA-16 – NOAA-15 is revealed also by zonal averages in John et al. (2013a), from which I showed the results for SNOs already in Fig. 3.1. The right panels of that figure show the results from zonal averages over polar regions. The strong time dependent bias of NOAA-16 – NOAA-15 is clearly visible and corresponds to the bias as seen from the SNOs. For channels 1 and 2, being influenced by surface, the biases show strong seasonal dependence that results from the diurnal cycle (John et al., 2013a). Note that the bias vanishes for 2008 where both satellites have the same LECT and sample the same phase of the day.

3.1.2.2 Global averages

In Hanlon and Ingram (2016), also global monthly means are investigated. Here, only near nadir FOVs (three FOVS to either side of nadir) are used. Again, the applied cloud filter may exclude parts of Antarctica for channel 3. Hence, the cold end of brightness temperatures will be undersampled for channel 3 compared to fully global data. The diurnal cycle impact is reduced in Hanlon and Ingram (2016) by averaging both the ascending and descending passes.

The time dependent bias for channel 3 for all AMSU-B and MHS instruments against MHS on NOAA-18 is displayed in Fig. 3.5. The observed biases agree quite well with the ones obtained with the FIDUCEO SNOs in Fig. 3.2: The strong increasing biases of NOAA-15 and NOAA-16 are again the dominant ones. Also NOAA-19 shows the almost constant bias. Metop-A and Metop-B only show very small deviations. NOAA-17 shows a slight positive decreasing bias, in contrast to Fig. 3.2, where the bias was more stable and slightly negative (note that both plots are based on near-nadir views only). This difference is probably due to different filtering (no cold end in 3.5). The observed change in time might be due to changes in the sampling of the diurnal cycle by orbital drift that affects the global averages, but not the SNOs.

In the Master's thesis by M. Prange, global, monthly means for all FOVs are used for investigating the NOAA-16 – NOAA-15 time dependent bias in order to compare the result to the bias observed in John et al. (2013a), where only near-nadir views were used for zonal averages, natural targets (see Sec. 3.1.3) and SNOs. To certain extent, the result matches the bias observed in John et al. (2013a). The differences can be explained by the scan-dependent bias observed by John et al. (2013b), presented in Fig. 3.3: in the Master's thesis, the absolute value of NOAA-16 – NOAA-15 bias for channel 3 increases strongly as observed in John et al. (2013a) (see Fig. 3.1). However, the NOAA-16 – NOAA-15 bias tends to negative values, whereas it becomes positive in John et al. (2013a). This is due to the strong scan-dependent biases of NOAA-15 (see Fig. 3.3) that does not influence the near-nadir FOVs, but has an impact on the overall bias if using all FOVs. The Master's thesis exemplarily shows by this comparison how results can differ between the used methods.

As overview, I compute the inter-satellite biases against NOAA-18 (all passes together) for all channels and all available times for which the instruments have overlap with NOAA-18. I use global averages, all FOVs and a high orography filter. The resulting time dependent biases are displayed in Fig. 3.6. All time dependent inter-satellite biases discussed so far are visible in this
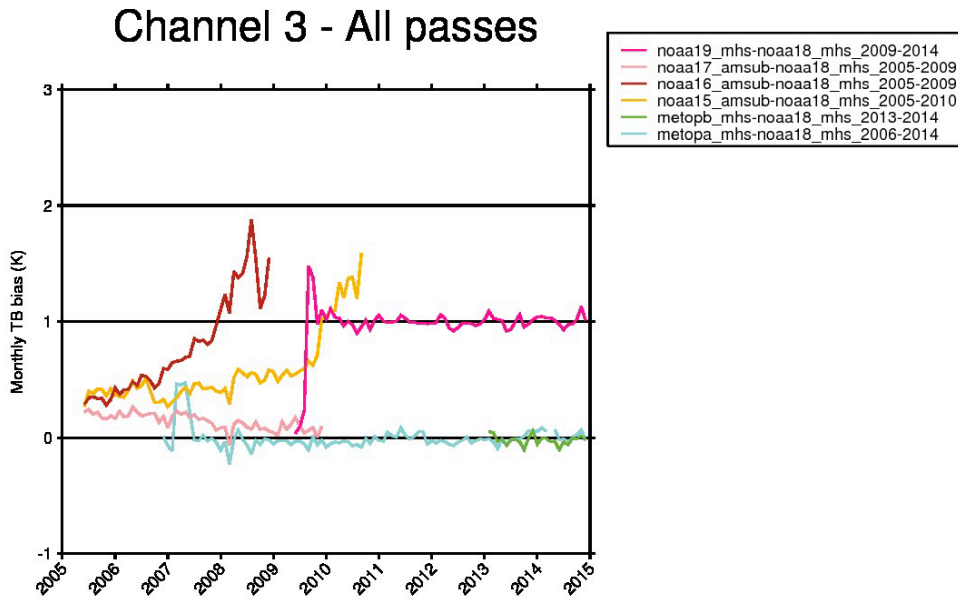
Figure 3.5 Global mean biases for nadir view only, computed with latitude binned data over all passes together. Figure and caption from Fig. 13 in Hanlon and Ingram (2016), with friendly permission by Helen Hanlon and William Ingram.

plot. Note that distinct differences may also exist among the sounding channels 3-5, which are very similar for AMSU-B.

For the SSMT-2 instruments, it is even more challenging to identify and analyse biases. This is, because the definition of a reference is difficult. Unfortunately, the SSMT-2 data sets from NOAA CLASS do not have overlap with the usual reference, which is MHS on NOAA-18. Even for the longer data sets from NGDC the overlap is hardly two years (see Sec. 2.6 for explanation of the different data sets). To overcome this problem of a lacking long-term reference, the SSMT-2 measurements can be compared to reanalysis data as in Kobayashi et al. (2017). Comparison to reanalysis data to gain insights and guidance is valid in the FIDUCEO philosophy. However, using reanalysis data as reference is not the FIDUCEO approach, since the traceability gets lost in the complexity of the underlying models and amount and variety of assimilated data within the reanalysis. Hence, a subsequent matching of instruments has to be done for the SSMT-2s. Figure 3.7 shows the time dependent bias computed from global monthly means, all FOVs, for the SSMT-2 on DMSP F14 and F15 (NGDC data, to have longer time series) against AMSU-B on NOAA-17. A strong seasonal pattern is visible in all channels. In addition, channel 1 of F15 shows a very peculiar bias change around July 2006, jumping from -5 K to +5 K against NOAA-17. Both instruments, on F14 and on F15 show a decreasing bias in channels 3-5. The bias changes in the range of -2 K to 0 K for channel 3, whereas channel 4 and 5 have larger biases up to -4 K. Against the reanalysis data, F14 shows a rather stable bias, whereas F15 shows a distinct increase in the bias (Kobayashi et al., 2017). Note however, that NOAA-17 also shows a slight time dependent bias against NOAA-18 (see Fig. 3.6). Hence, part of this F15 – NOAA-17 bias might be a real instrumental bias for F15 and part of this bias might be due to a changing NOAA-17 also. Looking further back into the past and matching F12 to F14, we see a nearly constant bias of about -0.5 K in channel 3 (see Fig. 3.8b), that was also observed in Kobayashi et al. (2017). Also Luo et al. (2017), who investigate biases among the different
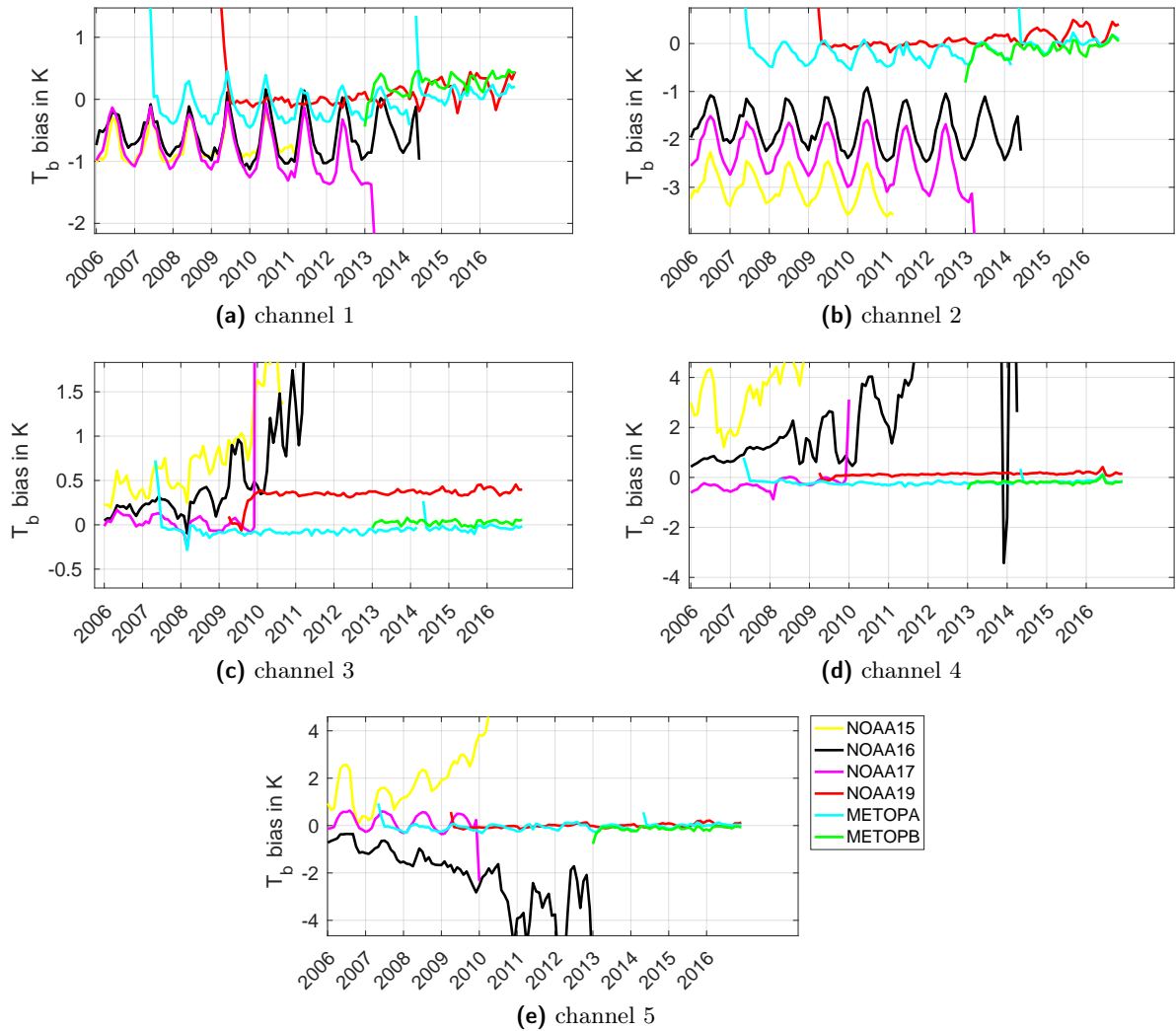
Figure 3.6 Inter-satellite biases against MHS on NOAA-18, based on global averages, all FOVs.

SSMT-2 instruments using SNOs, zonal averages and natural targets (tropical oceans) find a similar bias for F12 – F14. Only towards the end of F12, large biases occur which are also seen in Fig. 3.8b. Reaching even further back and matching F11 to F12, a constant offset of about 2 K of F11 against F12 in channel 3 is observed, superimposed with a seasonal pattern (see Fig. 3.8a). In Kobayashi et al. (2017), both F11 and F12 showed a slight decreasing bias against the reanalysis data, hence a constant bias between them is expected. A similar constant bias of about 1.5 K for F11 – F12 is observed by Luo et al. (2017). Compared to the F14 satellite, F12 also showed a relatively constant bias. However, against NOAA-17, the F14 satellite showed a time dependent one. Taking into account the time dependent bias of NOAA-17 against NOAA-18, it becomes clear that the bias of F12 against NOAA-18 cannot be identified as long as NOAA-17 and F14 are not fully characterised against NOAA-18.

This short discussion of SSMT-2 biases shows the difficulty especially for the older SSMT-2 instruments: the lack of a long-term reference hinders a bias analysis as it is possible for the other Microwave sounders AMSU-B and MHS.

(a) channel 1

(b) channel 2

(c) channel 3

(d) channel 4

(e) channel 5

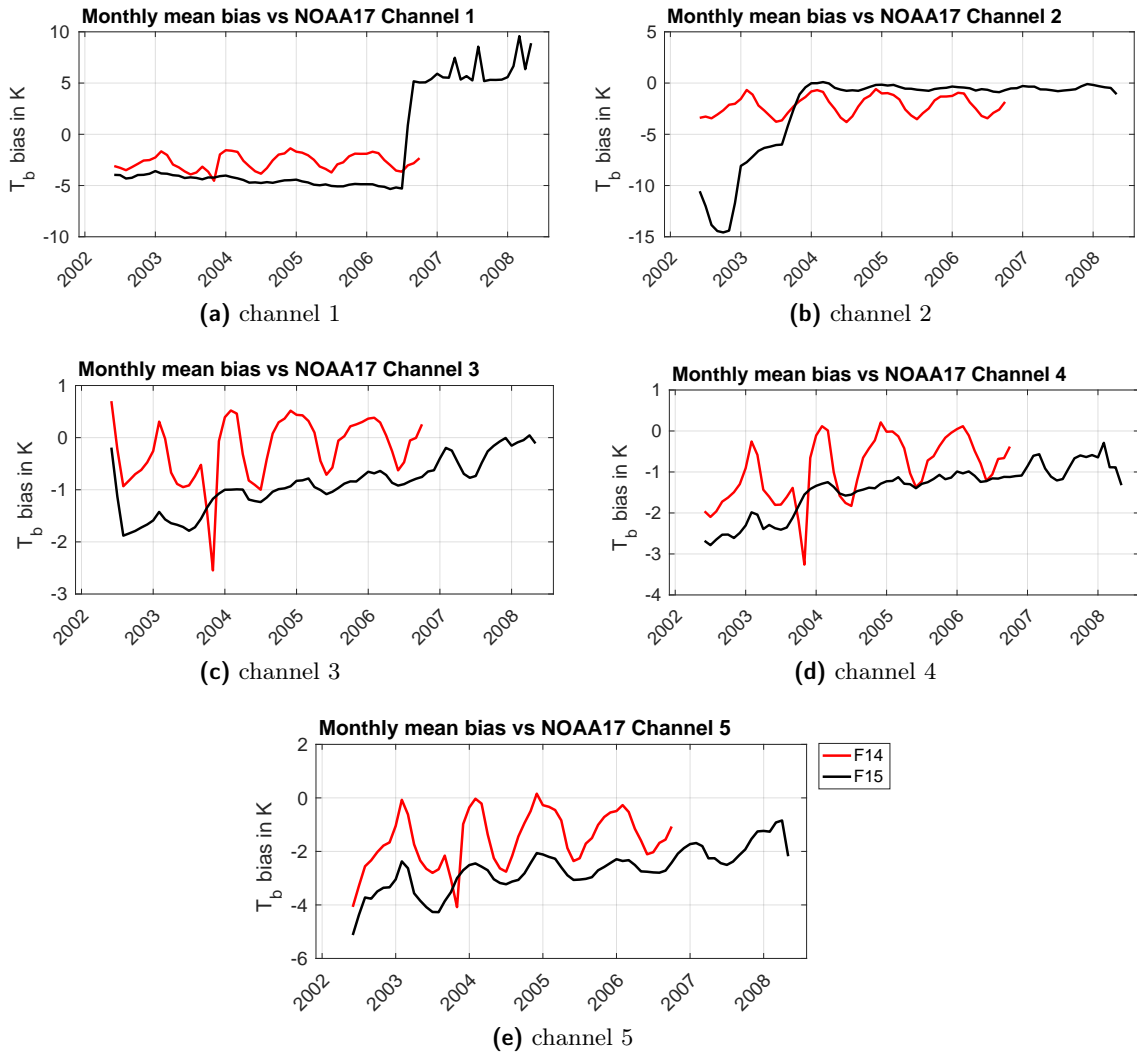**Figure 3.7** Inter-satellite biases of SSMT-2 on DMSP F14 and F15 against AMSU-B on NOAA-17, based on global averages, all FOVs.
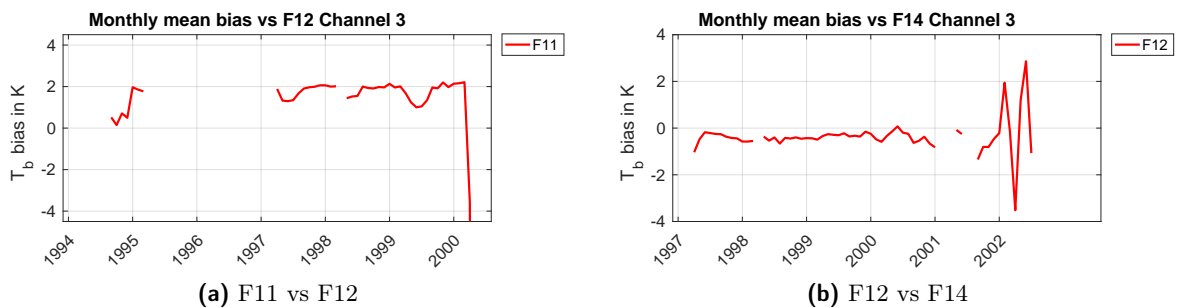

(a) F11 vs F12

(b) F12 vs F14

Figure 3.8 Inter-satellite biases of SSMT-2 on DMSP F11, F12 and F14, based on global averages, all FOVs. Channel 3 at $183 \pm 1\,\text{GHz}$.

3.1.3 Natural calibration targets

Natural calibration targets as used by John et al. (2013a) and Luo et al. (2017) are a special case of zonal averages. In John et al. (2013a), zonal averaged near-nadir brightness temperatures are used for the chosen region of Antarctica and tropical oceans. The idea of using confined regions is, that there is a small diurnal cycle for well chosen regions. This facilitates the identification of inter-satellite biases of two instruments that sample the region at different times of the day. However, John et al. (2013a) found that Antarctica was not well suited as target, since all channels may become surface channels and are therefore strongly affected by the diurnal cycle. The tropical oceans performed better, but were still subject to effects from the diurnal cycle. Nonetheless, the increasing positive bias of NOAA-16 – NOAA-15 was observed for the sounding channels. As pointed out above in Sec. 3.1.2.2, the NOAA-16 – NOAA-15 bias becomes negative if all FOVs are used.

3.1.4 Percentiles of brightness temperature data

Another statistics-based method uses percentiles of data instead of means. The method as described here was developed only recently in the Master's thesis by M. Prange. The advantage of the method is that the biases can be easily displayed as function of both time and percentiles which represent the covered brightness temperature range. For each month, I compute the 10th to 90th percentiles of the global data to cover the full brightness temperature range. The computation of percentiles is done separately for the ascending and descending branch. The inter-satellite biases against MHS on NOAA-18 are then computed as the difference of the corresponding percentiles for each branch. Finally, the mean bias as the average of biases from the ascending and descending branches is calculated.

Figure 3.9 displays the inter-satellite biases as function of percentiles and time, for ascending, descending and both together. The inter-satellite bias is colour-coded. Note that differences above and below $\pm 1\,\mathrm{K}$ are not displayed explicitly (except for NOAA-15 and NOAA-16). The figure shows the inter-satellite biases for the three AMSU-B and three MHS matched against MHS on NOAA-18. I show the results for channel 3, because it is the most interesting one from the perspective of an Upper Tropospheric Humidity CDR generator. Furthermore, the problems experienced with this channel for several instruments as shown by other inter-comparison methods do not allow for robust conclusions from the data. However, I also briefly discuss channel 4 for Metop-A and Metop-B, since it has a quite large bias compared to its other channels.

In the beginning, channel 3 on NOAA-15 shows a slight positive difference to NOAA-18 that increases towards higher percentiles (see Fig. 3.9a). At later times, the differences to NOAA-18 grow and it looks as if a distinct pattern of large negative bias in the lowest percentiles and large positive bias in the higher percentiles would build up. At this point it must be noted, that what looks like a bias pattern is *not* a real bias, but a consequence of increasing noise level (see Sec.4.2) for channel 3 and the analysis method using percentiles. The same evolution is seen in NOAA-16 channel 3 (Fig. 3.9b). Also for NOAA-19, this spurious "bias distribution" over the percentiles is visible, however, it builds up quite abruptly after its erratic period in 2009 (see Sec. 4.1, 4.2). Because of a severe gain decrease in this period, the noise level increased as for NOAA-15 and NOAA-16. Consequently, the covered range of brightness temperatures is extended. This is because any measured value at the "true" warm end of the temperature range will be measured sometimes as too low, but equally often as too large, thus extending the range towards higher temperatures. The same holds for the lower end. By computing percentiles, the measurements

(a) NOAA-15

(b) NOAA-16

(c) NOAA-17

(d) NOAA-19
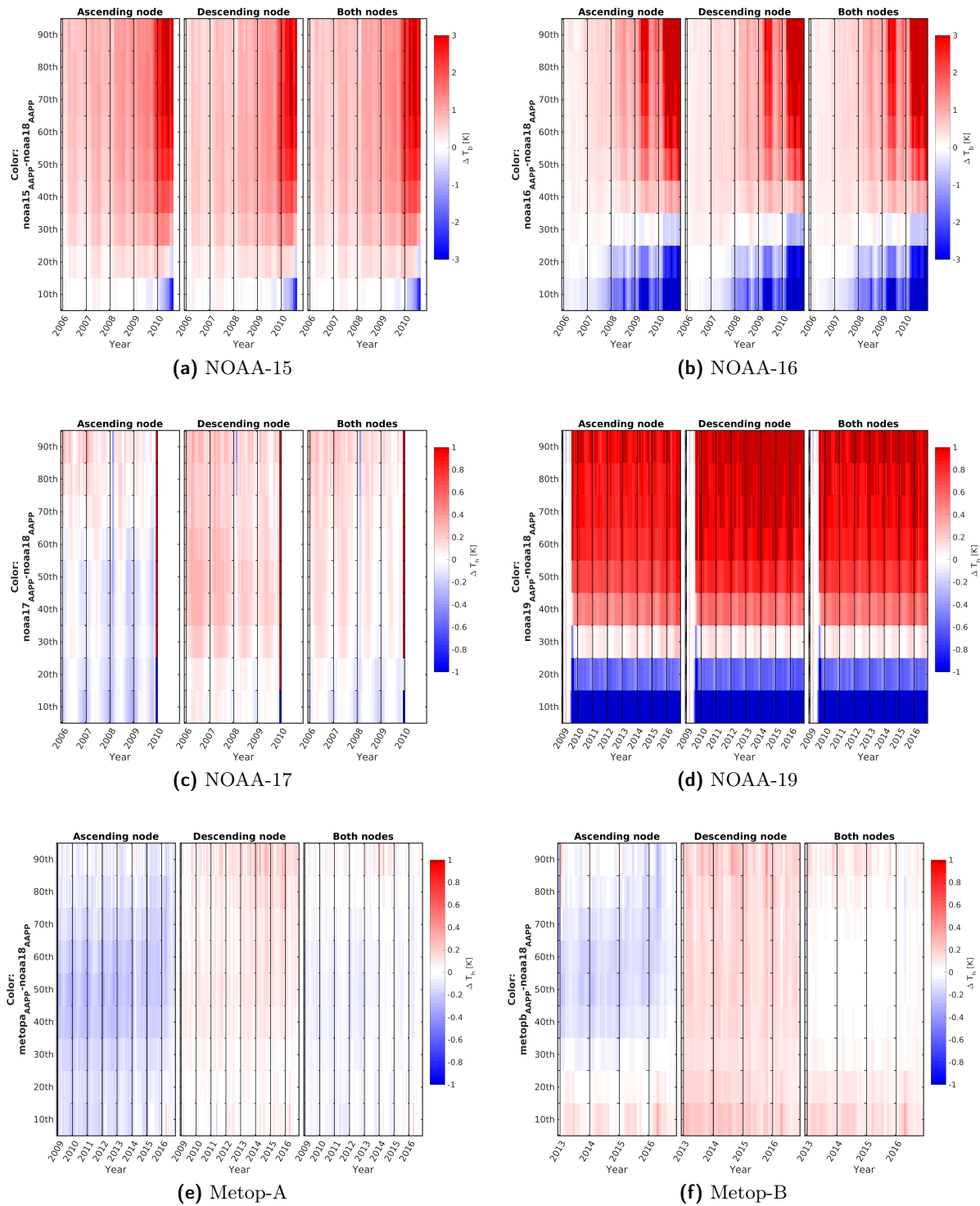
(e) Metop-A

(f) Metop-B

Figure 3.9 Channel 3, differences to MHS on NOAA-18, 10th to 90th monthly percentiles, all FOVs. Note the different colour scales for NOAA-15 and NOAA-16 compared to the others.

are sorted in ascending order and counted afterwards. Hence, the lower percentiles will be lower for the noisy instrument than for NOAA-18. At the higher percentiles, the noisy instrument will be warmer. This results in a spurious bias pattern that might be misinterpreted as brightness temperature dependent bias as in John et al. (2012). Of course, there is still an inter-satellite bias as can be seen in the time evolution of the biases for NOAA-15, NOAA-16 and NOAA-19. However, this bias is not caused by the noise effect revealed with the percentiles method: no bias can be caused by noise, as a systematic difference cannot be explained by random errors. To find the actual reason for the observed time dependent biases is part of the harmonisation.

For channel 3 of AMSU-B on NOAA-17 (Fig. 3.9c), a seasonal pattern, probably due to diurnal cycle effects is visible. Note that NOAA-17 and NOAA-18 have completely different LECTs (see Fig. 2.7) and therefore sample different times of the day. A weak brightness temperature dependent difference to NOAA-18 is visible, being slightly negative at the lower end and slightly positive at the higher end. At the very end of NOAA-17-AMSU-B lifetime, the instrument degrades abruptly (see also Sec. 4.2, 8.1) and the spurious "bias" pattern as seen for NOAA-15, NOAA-16 and NOAA-19 due to increased noise is visible. On top of the seasonal variations, a slight time evolution can be seen towards smaller biases. This is also visible in Fig. 3.6c. Note that also channel 4 shows a slight time dependence, however, into opposite direction (Fig. 3.6d).

For both nodes averaged, channel 3 of MHS on Metop-A (Fig. 3.9e) shows a slight negative bias for most times and percentiles. Only in the extreme percentiles of the warm and cold ends, the bias sometimes becomes positive. This seems to be more systematic at the warm end than at the cold end. Looking at the average of the ascending and descending node, channel 3 of MHS on Metop-B (Fig. 3.9f) is slightly warmer than NOAA-18 in the lower and higher percentiles, whereas no bias is visible around the 50th, 60th percentiles. This is also displayed in Fig. 3.10, where I plot the bias as function of the brightness temperature belonging to each percentile, for the six months from September 2014 to February 2015 for both Metop-A and Metop-B. A distinct dip of the bias towards zero bias can be seen for Metop-B. For Metop-A, the pattern is less pronounced and rather flat. The separation of both nodes in Fig. 3.9 (e) and (f) reveals that the descending node tends to be warmer in all percentiles, whereas the ascending node also shows distinct negative biases. Negative biases in the ascending, and positive biases in the descending node data also exist for Metop-A, which suggest some influence of the diurnal cycle that is not fully averaged out in the right panel. Note that Metop-A and Metop-B have the same LECT, whereas NOAA-18 has a quite different one (see Fig. 2.7). However, some instrumental issue may be present, since the bias patterns are different for Metop-A and Metop-B.

For the Metops, I also show the percentiles analysis for channel 4 (see Fig. 3.11), since they have an unexpected large negative bias in all percentiles against NOAA-18, bearing in mind that these are the newest instruments. This negative bias was also found by Hanlon and Ingram (2016), when looking at a global average over all times. It is also clearly visible in Fig. 3.6d.

3.2 Conclusion

In this chapter, I discussed the inter-satellite biases in the operational data of the microwave humidity sounders. Due to a lack of an absolute reference, one has to compare the instruments among themselves to obtain the inter-satellite biases. This is done by choosing one instrument as the reference. In the FIDUCEO project, the MHS on NOAA-18 is taken as reference. Using different inter-comparison methods, different aspects of biases can be revealed: dependence on time, on brightness temperature or on the scan angle. However, each of the methods is affected

(a) Metop-A   **(b)** Metop-B
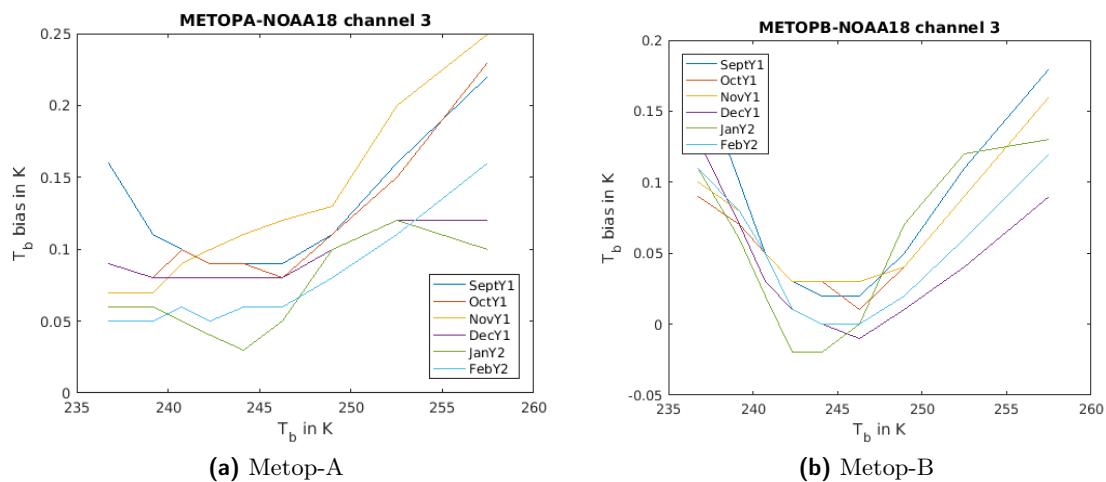
Figure 3.10 Channel 3, differences to MHS on NOAA-18 against brightness temperatures of 10th to 90th monthly percentiles, all FOVs. From September 2014 to February 2015.

(a) Metop-A   **(b)** Metop-B
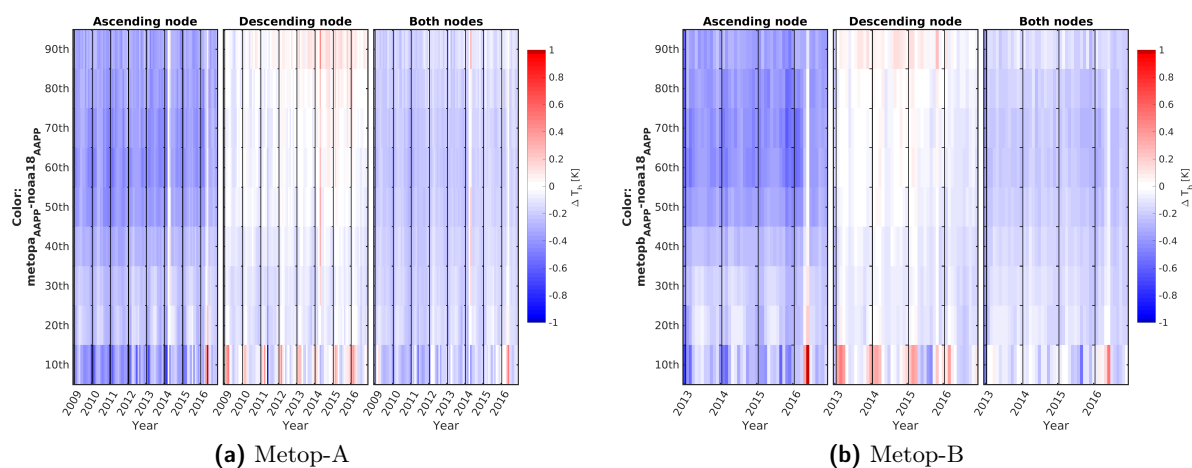
Figure 3.11 Channel 4, differences to MHS on NOAA-18, 10th to 90th monthly percentiles, all FOVs.

by different sampling issues, which makes it difficult to disentangle sampling effects from real inter-satellite biases.

The most important biases for the respective instruments on the listed satellites are:

- NOAA-15, NOAA-16: strong time and scan-dependent biases against NOAA-18 in channels 3-5

- NOAA-17: slight time and scan-dependent bias against NOAA-18 in channels 3-5

- NOAA-19: no bias before degradation around August 2009. Constant bias against NOAA-18 in channel 3 and to a lesser extent in channel 4.

- Metop-A, Metop-B: best agreement with NOAA-18. Small, brightness temperature dependent biases remain in channel 3. Channel 4 has stronger constant bias.

The SSMT-2 instruments cannot be compared directly to NOAA-18 due to the lack of an overlapping period. SSMT-2 on F14 and F15 can be compared to NOAA-17 and show a slight time dependent bias. However, this time dependent bias is not necessarily due to time dependent instrumental problems of the SSMT-2 on F14 and F15, but the bias may originate from time dependent instrumental changes of NOAA-17, too. Note that NOAA-17 shows a time dependent bias against NOAA-18. The trend of this bias change is compatible with the trend seen for the SSMT-2 against NOAA-17, if NOAA-17 brightness temperatures become cooler.

The harmonisation effort, as presented in this thesis, concentrates on the understanding and reduction of the most important inter-satellite biases listed above. SSMT-2 is not yet part of this procedure, because of the lack of a direct comparison with the reference NOAA-18. In Chapter 6 on Harmonisation, I revisit the existing, most important biases and investigate their possible causes as a starting point for the FIDUCEO harmonisation procedure.

4 Study of instrumental behaviour over time: Noise analysis and usable data

Before starting to work with data of the MW sounders, it is very useful to study the overall evolution of the basic calibration quantities and the noise performance in order to get an overview of the instruments behaviour over its life time. This overview is not only important for any users to judge the usability of the data for their purposes, but also the overview is crucial for the further production of the FCDR. Having this overview, we can later easily estimate for which periods we can expect good data. Also, we know where we have to expect bad data which we can test our FCDR production code on for the detection and flagging of bad or suspicious data.

In this chapter, we present the evolution of the calibration quantities to give an impression of the instrumental behaviour and we investigate in detail the noise performance of the instruments. From these studies, we deduce periods of usable data.

Major parts of this chapter are published in

> Hans, I., Burgdorf, M., John, V. O., Mittaz, J., and Buehler, S. A.: Noise performance of microwave humidity sounders over their lifetime, *Atmos. Meas. Tech.*, 10, 4927-4945, `https://doi.org/10.5194/amt-10-4927-2017`, 2017, Creative Commons Attribution 4.0 License

Note that in this thesis, the section on data of this publication has been moved to Chapter 2 on data above.

4.1 Evolution of basic calibration quantities

For studying the overall behaviour over the life time of the instruments, we read the level-1b data records and we create long time series of the basic calibration quantities. From this raw data record we read the counts for the black body views, i.e. the on board calibration target views (OBCT), the counts for the deep space views (DSVs) and the counts for the temperature sensors (Platinum Resistance Thermometers, PRTs) on the black body. The latter ones are transformed to temperature in Kelvin. We do not take into account any quality flags that might be set in the data record but only use raw unfiltered data in order to preserve the original recorded behaviour. For each channel and all scan lines of every orbit we calculate from those values the gain $G_n(i)$ for scan line n and channel i as

$$G_n(i) = \frac{\overline{C}_{\mathrm{OBCT}}(i) - \overline{C}_{\mathrm{DSV}}(i)}{\overline{T}_{\mathrm{PRT}} - 2.725K}, \tag{4.1}$$

where $\overline{C}_{\mathrm{OBCT}}$ and $\overline{C}_{\mathrm{DSV}}$ indicate the counts from the OBCT and DSVs, respectively, both averaged over the 4 views. $\overline{T}_{\mathrm{PRT}}$ denotes the average temperature measured by all temperature sensors (two for SSMT-2, seven for AMSU-B and five for MHS). The OBCT and DSV counts,
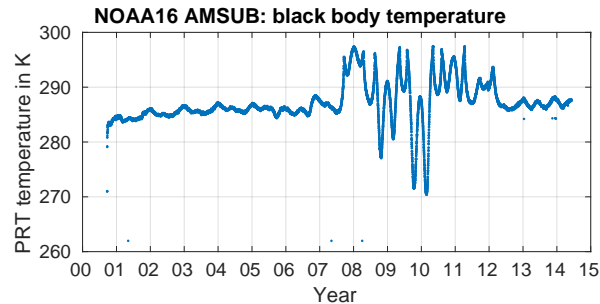
NOAA16 AMSUB: black body temperature

Figure 4.1 The temperature of the black body (average of 7 PRTs) on board AMSU-B on NOAA-16 over its life time from 2000 to 2014. Note the characteristic oscillating pattern that occurs during strong orbit drift towards an LECT of 18:00.

as well as the gain, are our input values for the noise estimation which is described in section 4.2 on noise performance.

From the level 1b readings we concatenate long time series and produce plots per instrument and calibration quantity. Here, I present some examples of these plots. Figure 4.5 at the end of this section shows the evolution of the gain for all AMSU-B and MHS instruments over their life time. This is a useful graphic to get an overview of the instrumental behaviour.

The AMSU-B instruments on NOAA-15 and NOAA-16 are known to have suffered from significant problems over their life time. Both have experienced a strong degradation over time, especially in channels 3-5, which was a superposition of two phenomena: Firstly, the satellites' orbital drift increased, leading to changes of the solar-beta-angle (Zou and Wang, 2011). This in turn produced strong changes in the temperature on the instrument, affecting all calibration quantities to varying degrees. Hence, a characteristic oscillating pattern appears in the temperature of the black body (see Fig. 4.1), and also on the gain. Secondly, a steady decrease of the gain (see Fig. 4.5a,4.5b) made the recorded signals for DSV and IWCT ever more similar. This is shown in Fig. 4.2 over the life time of AMSU-B on NOAA-16. This evolution decreased the overall sensitivity of the radiometer in the channels 3-5, which is discussed in Sec. 4.2. Finally, both targets, warm and cold, produce the same signal in the instrument. At this point, at the latest, the AMSU-B NOAA-16 data of channel 3 become totally useless. A very similar evolution takes place for channels 4 and 5. The time dependent biases reported in Chapter 3 show the same evolution as the gain degradation of both instruments. The reason for this is suspected to be RFI and further discussed in Chapter 6 on harmonisation. Channels 1 and 2 are mostly unaffected though and can be used safely.

The MHS instrument on NOAA-19 is another example of degradation. However, in this case, the degradation is not as steady as for the NOAA-15 and NOAA-16 AMSU-B: Shortly after launch, the MHS on NOAA-19 was "out of specifications" (NOAA-OSPO, 2015) in summer 2009. It showed erratic behaviour in channels 3 and 4 with periods of zero gain. After several gain adjustments were made, the instrument regained its stability, but at a degraded level for channel 3 and channel 4. In Figures 4.3, 4.4 and 4.5e, this problematic period is clearly visible. From 2010 on, both channel 3 and 4 remain stable, but with a significantly reduced gain, especially for channel 3. As for NOAA-15 and NOAA-16, the evolution of the gain in channels 3 and 4 matches well the evolution of the bias (Chapter 3). The reason is discussed in Chapter 6 on harmonisation.
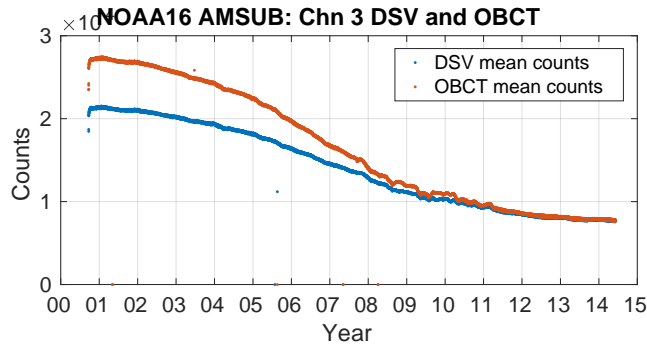
Figure 4.2 Evolution of the mean counts of the DSVs and IWCT views in channel 3 of AMSU-B on NOAA-16 over its life time from 2000 to 2014. Note how the recorded signal from DSV and IWCT approach each other.
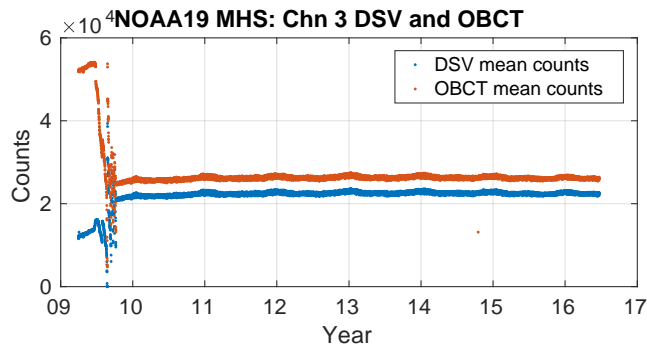
Figure 4.3 Evolution of the mean counts of the DSVs and IWCT views in channel 3 of MHS on NOAA-19 from 2009 to 2016. Note the problematic period in 2009 and the similarity of DSV and IWCT measurements afterwards.
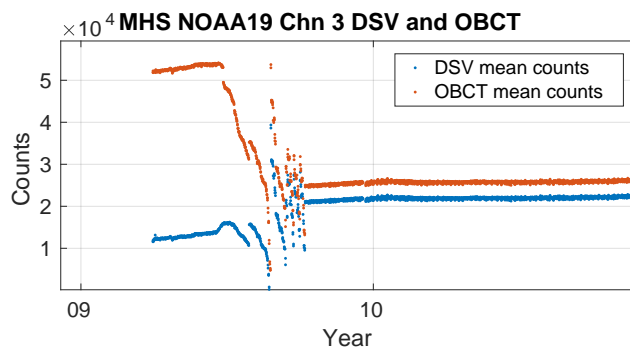
Figure 4.4 A zoom of Fig. 4.3 showing the problematic period in 2009 where several gain adjustments were made in order to bring the channel 3 back to specification. The channel stabilised at some point, but at a degraded level.

A detailed description of the evolution of the behaviour per instrument and quantity is provided in the appendix 8.1.

(a) NOAA-15

(b) NOAA-16

(c) NOAA-15

(d) NOAA-18

(e) NOAA-19

(f) Metop-A

(g) Metop-B
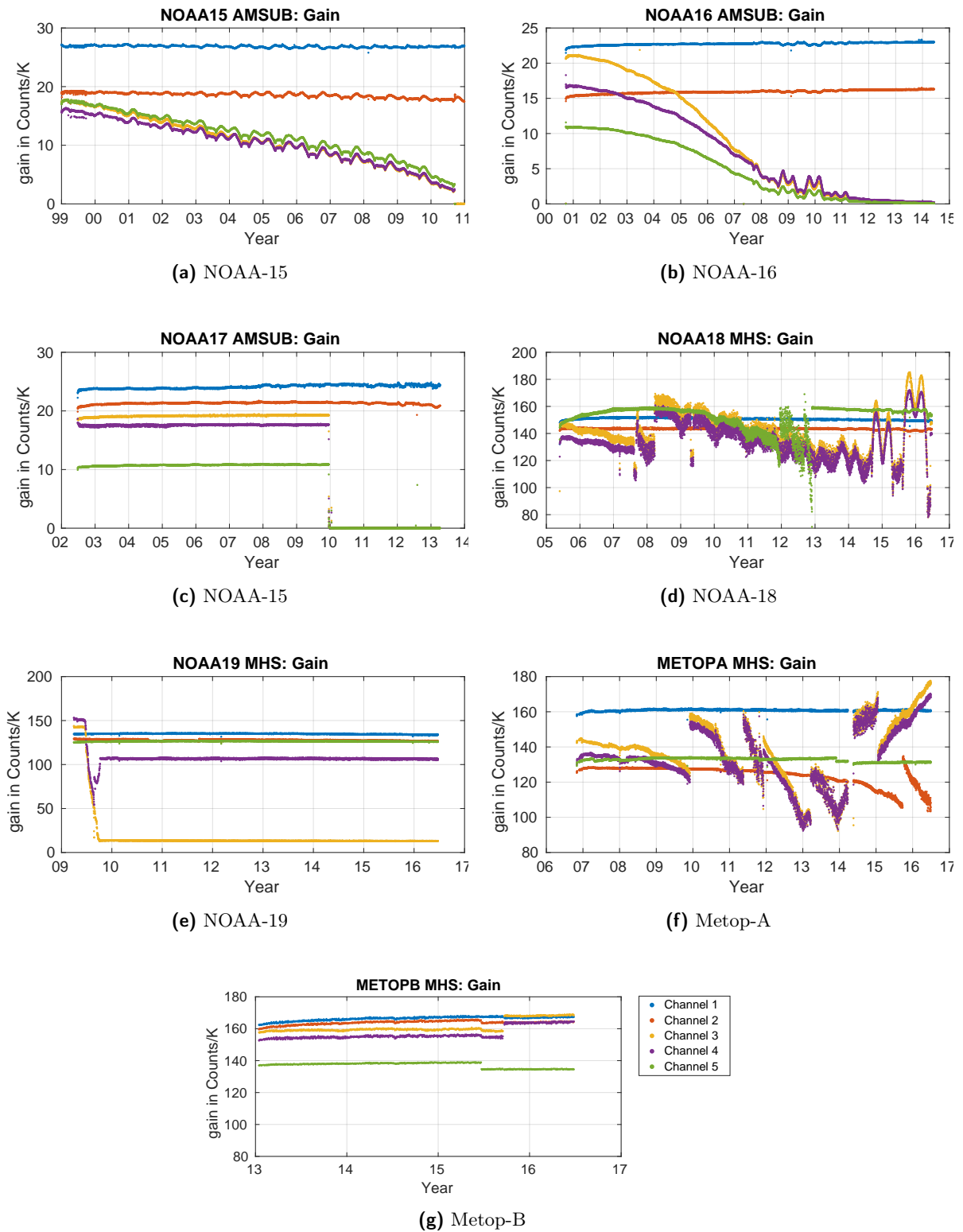
Figure 4.5 The evolution of the gain for all channels of the AMSU-B and MHS instruments over their life time. Note the dramatic decrease for some channels.

4.2 Noise performance of MW humidity sounders over their lifetime

This section is a reprint of large parts of the publication

> Hans, I., Burgdorf, M., John, V. O., Mittaz, J., and Buehler, S. A.: Noise performance of microwave humidity sounders over their lifetime, *Atmos. Meas. Tech.*, 10, 4927-4945, `https://doi.org/10.5194/amt-10-4927-2017`, 2017, Creative Commons Attribution 4.0 License

In this study, we calculate and assess the noise evolution over the lifetime of all individual instruments of the microwave sounders Special Sensor Microwave Water Vapor Profiler (SSMT-2), Advanced Microwave Sounding Unit-B (AMSU-B) and Microwave Humidity Sounder (MHS). So far, their data sets lack comprehensive information on uncertainty caused by noise: From the pre-launch measurements, one knows the specifications on the precision that the instruments had to meet. These values of noise equivalent differential temperature (NEΔT) are provided per instrument and channel in the NOAA KLM User Guide (Robel et al., 2009) and by their nature as specifications do not comprise any information on time evolution of noise. The ATOVS and AVHRR Pre-processing Package (AAPP) software used for the processing of raw level 1b data to level 1c data containing brightness temperatures, now provides with version 7.13 a measure of noise, namely a cold and a warm NEΔT, referring to the cold and warm calibration targets on board those microwave sounders. However, this information on noise in the AAPP-processed data sets is not available for all instruments. Graphical information on noise evolution is given on the NOAA-STAR-ICVS web page (NOAA-STAR, 2015), but this is also limited to a few periods and instruments. Comprehensive information on uncertainty caused by noise is not available for the end user interested in the measurements of the SSMT-2, AMSU-B or MHS instruments.

To close this gap, we determine and evaluate the time series of the noise for the SSMT-2 instruments on board the Defense Meteorological Satellite Program (DMSP) satellites F11, F12, F14 and F15, for the AMSU-B instruments on the satellites NOAA-15, NOAA-16 and NOAA-17 launched by the National Oceanic and Atmospheric Administration (NOAA) and for the MHS instruments on the satellites NOAA-18, NOAA-19 and the Metop-A and Metop-B satellites controlled by EUMETSAT. In the assessment of the noise evolution, we identify periods of low-quality data. To make this information easily accessible, we provide a graphical and descriptive overview over the whole lifetime of the instruments. From this overview, the users can estimate the uncertainty due to noise and can decide on the applicability of the data set for their purposes. Our method and tool to estimate count noise will be used in the evaluation of the uncertainty for the generation of new microwave sounder Fundamental Climate Data Records (FCDRs). Those are currently developed in the project "Fidelity and Uncertainty in Climate data records from Earth Observation (FIDUCEO)" in the framework of which this study has been carried out and that aims to adopt a rigorous metrological (measurement science) perspective to understanding the origins and quantifying various instrumental issues that lead to random and systematic errors (Mittaz et al., 2018).

Apart from the new comprehensive time series of noise evolution, our results also include the analysis of the spectrum of the noise in flight. This analysis is based on the statistical tool of the Allan deviation and its general form the *M*-sample deviation (Mittaz, 2016). We also use the Allan deviation for the calculation of the noise itself, in contrast to what has been done for the previously available noise estimates. The Allan deviation, well known in other disciplines (Tian

et al., 2015; Malkin, 2011; Allan, 1966), has been suggested only recently by Tian et al. (2015) for the estimation of noise in the measurements of microwave sounders in flight.

The noise in flight can be estimated with various methods. Atkinson (2015) reports on methods used and suggested by different agencies for the calculation of cold and warm NEΔT. The various methods include the standard deviation and also the Allan deviation as suggested by Tian et al. (2015). The disadvantage of the standard deviation is that it is sensitive to variations in the mean that naturally occur in the measurements of these kind of polar orbiting instruments (Tian et al., 2015). In this study, we follow the suggestion of Tian et al. and use the Allan deviation for the estimation of noise. To clarify the notion of noise at first, the next section is dedicated to the elaboration of a consistent noise terminology in the context of the microwave sounders.

This article is further structured as follows. After establishing the noise terminology used here, we explain our methods in detail. Later, our results on the analysis of the noise spectra and the time evolution of noise are presented. The discussion of these findings is followed by concluding thoughts. In the Appendix we provide a detailed description of the time series of the individual instruments.

4.2.1 Noise terminology

In theory, noise in the measurements of a radiometer such as the microwave sounders considered here can be related to the process of measuring and it can be calculated from instrumental quantities. This theory of noise in the measurements of a radiometer is explained by Ulaby and Long (2014) whom we follow here: The antenna delivers a power P_a to the receiver. In the Rayleigh-Jeans limit, this power is usually related to a temperature T_a as $P_a = kT_aB$, with k being Boltzmann's constant and B the bandwidth of the receiver. The precision with which the temperature T_a can be estimated by a measurement is referred to as the radiometer sensitivity dT. It is subject to any noise that may impact on the true signal and depends on the temperature of the whole system. So, the total system noise power $P_{sys} = P_a + P_{rec} = kT_{sys}B$ relates to the system temperature $T_{sys} = T_a + T_{rec}$ with T_a being the antenna temperature (which includes the true signal) and P_{rec} and T_{rec}, being the power and temperature of the receiver including the influence of the transmission line between antenna and receiver. Since the final measured output voltage is an integrated value from a receiver of bandwidth B and an integration time of t, the noise uncertainty to the radiometer sensitivity is

$$dT_N = \frac{T_{sys}}{\sqrt{Bt}} \tag{4.2}$$

However, one also has to consider fluctuations in the gain G on timescales shorter than one calibration cycle. These are not calibrated out, but impact on the recorded voltage and hence lead to fluctuations in the final measurement result. These short-term gain fluctuations lead to a term

$$dT_G = T_{sys}\frac{dG}{G} \tag{4.3}$$

Since both contributions are independent, the radiometer sensitivity finally reads

$$dT = \sqrt{dT_N^2 + dT_G^2} \tag{4.4}$$

$$\Rightarrow dT = T_{sys} \cdot \sqrt{\frac{1}{Bt} + \left(\frac{dG}{G}\right)^2} \tag{4.5}$$

where T_{sys} is the sum of antenna temperature and combined receiver-transmission line temperature. This radiometer sensitivity dT describes the smallest temperature difference that the radiometer can distinguish when looking at a target inducing an antenna temperature of T_a. It is therefore an uncertainty estimate on the measurement of T_a.

For in-flight monitoring of the radiometer sensitivity, eq. 4.5 is not well suited, since the receiver-transmission line temperature is not well accessible. Therefore, one does not usually use Eq. 4.5 to calculate the radiometer sensitivity, but one uses some kind of statistical estimation of the fluctuations in the measurements, e.g. in the counts that are the digitised output voltage. The counts may stem from the instrument's views of the cold or warm calibration target (deep space views, DSVs, and on-board calibration target, OBCT). This statistical estimation may be the standard deviation or the Allan deviation. This estimation of the fluctuations in the counts, referred to as count noise, comprises every noise that has contaminated the true signal from the antenna over the transmission line through amplifiers and mixers, including digitization noise. The count noise is therefore subject to both effects described in eq. 4.5 for the total radiometer sensitivity: noise from all electronic devices, that remains due to a finite integration time, and short-term gain fluctuations (on timescales shorter than one scan, i.e. one calibration cycle). This count noise estimate is in units of counts, i.e. one cannot compare the values directly among different sensors since the absolute count values are somewhat arbitrary. However, we can transform the count noise into a NEΔT, which then represents an estimate for the total radiometer sensitivity described by Eq. 4.5. This transformation includes the gain, i.e. NEΔT = noise-in-counts/ gain. The actual value of the gain taken for this estimate is the one corresponding to the scan lines from which the count noise has been calculated.

Hence, this transformation translates the fluctuations that we see in the counts (count noise) into a temperature difference that is equivalent to the noise by using the current gain. Any long-term changes in the gain will therefore impact on the time evolution of NEΔT. Altogether, this NEΔT includes the actual noise, i.e. short-term fluctuations of whatever origin, and the long-term variations of the gain. This is in contrast to the count noise estimate which reflects the pure short-term fluctuations.

This in-orbit analysis of noise can be carried out on both the counts of the DSV and OBCT views, whereas the Earth counts are not suited due to their natural variability when scanning over different scenes on Earth. The choice of target influences the antenna temperature in eq. 4.5 and consequently influences dT. Therefore, the dT calculated from the OBCT is expected to be larger than from the DSV. If choosing the DSV counts, one takes advantage of the fact that the brightness temperature of the DSV is very low. Therefore, the contribution of that signal to the antenna temperature is rather weak. The remaining contribution to the antenna temperature and of course the receiver temperature are of instrumental origin. Hence, analysing the DSV will give results (almost only) on the instrument itself. Converting the DSV count noise to a temperature we obtain the cold NEΔT, which corresponds to the radiometer sensitivity when looking at very cold scenes. The second choice, taking the count noise on OBCT will lead to the warm NEΔT after translating to a temperature. This warm NEΔT corresponds to the radiometer sensitivity when looking at a target of approximately 280 to 300 K (temperature of the OBCT, T_{BB}).

The end users, however, will be interested in the (scene) NEΔT that they have to expect for a certain Earth pixel in their data sets. However, the NEΔT cannot be calculated directly from the Earth counts as explained above. But, it can be estimated from the cold and warm NEΔT. As eq. 4.5 expresses, the NEΔT or radiometer sensitivity depends on the antenna temperature. Therefore, the NEΔT when looking at an Earth scene of 240 K will be close to but slightly

smaller than the warm NEΔT. Knowing both the cold and the warm NEΔT, users can calculate their scene NEΔT as

$$\text{NE}\Delta\text{T} = \text{NE}\Delta\text{T}_{cold} + (T_{\text{a,scene}} - T_{\text{a,DSV}}) \cdot m \tag{4.6}$$

$$\text{with } m = \frac{\text{NE}\Delta\text{T}_{warm} - \text{NE}\Delta\text{T}_{cold}}{T_{\text{a,OBCT}} - T_{\text{a,DSV}}}$$

$$\text{and } T_{\text{a,OBCT}} - T_{\text{a,DSV}} = T_{\text{BB}} - 2.725\,K$$

This equation is obtained from combining eq. 4.5 applied for warm and cold NEΔT (i.e. using $T_{\text{a,OBCT}}$ and $T_{\text{a,DSV}}$ as antenna temperatures in T_{sys}). Using the resulting two equations for eq. 4.5 with scene NEΔT will yield the above equation 4.6.

To obtain an estimate of the current scene NEΔT, users need the estimate for the cold and warm NEΔT corresponding to the time window where their current earth pixel belongs. Moreover, they need the corresponding temperature of the OBCT (measured temperature of the black body, T_{BB}) and the $T_{a,scene}$ of their data set to finally calculate scene NEΔT with Eq. 4.6.

4.2.2 Methods

The subsection on data of this publication has been moved to Chapter 2 on instruments and data used throughout this thesis.

4.2.2.1 Noise estimation

The standard deviation can be used to estimate the noise on the counts as explained in the above section on noise terminology. This estimation has been used before (Tian et al., 2015; Atkinson, 2015), but the standard deviation has a disadvantage in the context of noise monitoring of instruments on polar orbiting satellites - since the standard deviation is based on measuring the difference of the values from the sample's mean, the standard deviation will only provide a sensible representation of the precision of the sample if the sample truly has a constant mean. However, due to the orbiting movement of the instrument around the Earth, all measured quantities show orbital variations, i.e. they have a non-stationary mean over one orbit. For such cases, the standard deviation is biased as it expects a stationary mean and would measure the deviation of a single measured value from the overall (erroneously stationary) mean over the full orbit. To reduce this bias, one has to define sub-samples along the orbit, for which the real mean is approximately stationary. Hence, the standard deviation becomes highly dependent on these chosen sample sizes and is therefore less suited for consistent in-orbit monitoring of different instruments.

The Allan deviation does not show this bias and is less dependent on choices (Tian et al., 2015; Mittaz, 2016). Therefore, we use the Allan deviation as statistical tool to estimate the noise on the counts. The Allan deviation, or its square, the Allan variance, is a special case of the more general M-sample variance (Allan, 1966), which is defined as

$$\sigma_M^2(M) = \frac{1}{M-1} \left(\sum_{i=0}^{M-1} y_i^2 - \frac{1}{M} \left[\sum_{i=0}^{M-1} y_i \right]^2 \right), \tag{4.7}$$

where y_i is a measured value from the sample and M denotes the number of values of the sample that are used for the calculation. In other words, M adjacent measurements yield one value $\sigma_{M_j}^2$.

The associated total M-sample variance for a total sample of N measurements is then calculated as the average over all $\sigma^2_{M_j}$ with $j \in 1, 2, \ldots, N/(M-1) - 1$. With $M = 2$ in eq. 4.7, one obtains the Allan variance that effectively uses two adjacent measurements of the data series:

$$\sigma^2_{\text{Allan}} = \sigma^2_M(2) = \frac{1}{2}(y_1 - y_0)^2 \tag{4.8}$$

The total Allan deviation for N measurements is then written as

$$\sigma_{\text{Allan, tot}} = \sqrt{<\sigma^2_{\text{Allan}}>_N}$$

$$= \sqrt{\frac{1}{2(N-1)} \sum_{n=1}^{N-1} (y_{n+1} - y_n)^2} \tag{4.9}$$

It is not biased due to longer-term trends such as orbital variations, because the Allan deviation averages over a sample size N the deviation of directly adjacent measurements. For a sample size of N values, one computes $N-1$ Allan deviations and averages these. The question of an appropriate value of N has been investigated in Tian et al. (2015) for the new instrument Advanced Technology Microwave Sounder (ATMS). Since it has the same scanning routine as the SSMT-2, AMSU-B and MHS, we follow the suggestions of Tian et al. - the lower limit of N is set by the stability of the Allan deviation with changing N. For small sample sizes of $N < 300$, the Allan deviation fluctuates. From $N = 300$ on, it takes a stable value. Following this study, we therefore use a sample size of $N = 300$ scan lines, providing us with about eight total Allan deviations per orbit. As expected, by comparing the standard deviation with the Allan deviation we found up to 40 times larger variations in the noise estimate over one orbit for the standard deviation. Also, increasing the sample size for the Allan deviation does not significantly change our results. This agrees with the results in Tian et al. (2015) concerning the stabilization of the Allan deviation above $N = 300$ scan lines for those instruments.

For defining the number of N, one could also use a different approach. This relates to the question of what a single measurement is and what adjacent means in the context of the investigated instruments. As explained above, the instruments have a scanning and calibration cycle of 8/3 seconds during which they record the signal from four warm calibration target views, from 90 Earth views and from four DSVs. Between the different targets they record nothing. Having in mind this scanning and calibration cycle, there are two approaches of noise estimation using the Allan deviation. On the on hand, one could use the Allan deviation between the individual adjacent four calibration views i.e. in each cycle one gets three Allan deviations. Over one orbit one averages these $3 \cdot k$ Allan deviations (with $k =$ number of scan lines in the orbit) to get the final total Allan deviation. We will call this the inter-pixel method. On the other hand, one can act on the scale of scan lines, as usually done for noise investigations so far (Tian et al., 2015; Atkinson, 2015; EUMETSAT, 2013). For this, one calculates the Allan deviation between two adjacent scan lines for all four views separately and then averages over the four obtained Allan deviations before applying the average over N scan lines. We chose this inter-scan-line method with $N = 300$ in our study. The reason for this is that the results of our analysis of the noise spectrum speak in favour of this inter-scan-line-method for noise estimation (see section 4.2.3.1): the inter-scan-line method will give a better estimate of the uncertainty in the data due to noise - compared to the inter-pixel method, which underestimates the uncertainty for non-white noise spectra.

To analyse the noise spectrum, we make use of the Allan deviation and the general M-sample variance again. Together, they make an interesting tool to determine the noise spectrum in a

simple way (Allan, 1966, 1987; Barnes and Allan, 1990). The quotient of the M-sample variance and the Allan variance, each averaged over the same sample size, is the so-called bias function (Barnes and Allan, 1990)

$$B_1(M) = \frac{< \sigma_M^2 >_N}{< \sigma_{\text{Allan}}^2 >_N}. \tag{4.10}$$

The behaviour of $B_1(M)$ for varying M is characteristic for different noise spectra. We let M vary from 2 to 20. We simulate white noise (constant power spectral density) and pink noise (or $1/f$ noise - noise with power spectral density proportional to inverse of frequency, i.e. $1/f$) in MATLAB and determine their bias functions over the indicated range of M. This serves as a comparison tool for the bias functions obtained from real data to estimate the nature of the noise spectrum of the data. This spectral analysis is carried out on the counts of the DSVs (DSV counts).

In this study, we investigate three estimates of noise. We calculate the Allan deviation on the deep space view counts to obtain the DSV count noise:

$$\Delta C_{\text{DSV}} = \sqrt{\frac{1}{2(N-1)} \sum_{n=1}^{N-1} \sum_{k=1}^{K} (C_{\text{DSV}_{k,n+1}} - C_{\text{DSV}_{k,n}})^2} \tag{4.11}$$

First, the difference in the counts C_{DSV} from scan line n to scan line $n+1$ is calculated for each view k separately. Then, the average for all $K=4$ views is taken. Then the total Allan deviation is computed as the average over all $N-1$ values obtained for the window of $N=300$ scan lines. This estimate of count noise is then translated into a temperature. We deduce the cold NEΔT by dividing by the gain corresponding to the first of the two adjacent scan lines (equally one could take the gain corresponding to the second one):

$$\text{NE}\Delta\text{T}_{\text{cold}} = \sqrt{\frac{1}{2(N-1)} \sum_{n=1}^{N-1} \sum_{k=1}^{K} \left(\frac{C_{\text{DSV}_{k,n+1}} - C_{\text{DSV}_{k,n}}}{G_n} \right)^2}. \tag{4.12}$$

Similarly, we calculate the warm NEΔT by replacing the DSV counts by the counts of the on-board calibration target (OBCT counts) in eq. 4.12. These three measures, i.e. the DSV count noise, and the cold and the warm NEΔT, are monitored over the lifetime of the instruments for each channel. The long time series that are displayed in this study contain data only for every 50th orbit (Fig. 4.10, 4.11, 4.12) in order to avoid a stronger overlapping of symbols and to maintain readability.

4.2.3 Results

4.2.3.1 Analysis of noise spectrum

The noise spectrum for the different channels has a non-white component that is more or less strongly pronounced for the different instruments and years. We present the effects of this mixed spectrum on the calculation of the noise time series. As examples, we pick two orbits from different years of MHS on NOAA-18 (2005: orbit 505, 2007: orbit 4500). The spectrum is calculated for these orbits with the bias function introduced in Eq. 4.10. The bias functions for each 300-scan-line window are further averaged over the orbit and the four DSVs. In this way,

we obtain for each of the two orbits an averaged bias function as shown for each channel in Fig. 4.6, together with the simulated bias functions for white and pink noise. For channels 1, 3 and 4, the bias function is close to that of pure white noise and therefore indicates for these channels a strong white noise component that is dominant over the pink one in the count noise of the DSV. The spectra for channels 2 and 5 look different, though. Both channels show a strong deviation from the pure white noise case, indicating a mixture of white and pink noise for both years.

How far this affects our noise estimates can be deduced from looking at the corresponding periods in time for the actual calculated count noise. In Fig. 4.7 the time evolution of the DSV count noise for the five channels is shown. In addition to the count noise calculated with the inter-scan-line method, we also provide the estimates obtained from the inter-pixel method for the two investigated orbits (red dots). The comparison of both methods' results together with the spectra in Fig. 4.6, indicate that both methods agree as long as there is a strong white noise component only (channels 1,3 and 4). Hence, the jump in DSV count noise in late 2007 in channels 3 and 4 is captured by both methods. At this time one can observe sudden jumps in the mean counts as well as a suddenly increased spread of the recorded counts around the mean, not only for the counts in the DSV but also for the OBCT counts. This is probably due to a gain adjustment for channels 3 and 4 in September 2007 (NOAA-OSPO, 2015).

If the noise spectrum is a mixture with a strong pink component, however, as is the case for channels 2 and 5, the inter-scan-line method gives a higher value than the inter-pixel method. This difference in the results of the two methods seems reasonable, since a pink noise (1/f noise) contaminated signal, having larger noise power at smaller frequencies, has variations due to noise on a longer timescale than the inter-pixel timescale. Thinking of the calibration cycle of the instruments, one can imagine the following scenario and consequences for the uncertainty estimation. At the beginning of the Earth scan, the signal suffers from a certain unknown portion of noise. Later in the scan, when looking at DSV, the signal from the target (deep space) itself is smaller of course. The portion of noise that contaminates the signal will have changed in the meantime, too. In the case of pure white noise, we know the range of that longer-term change since it will be defined by the standard deviation of the underlying distribution. This standard deviation is described by the value of the inter-pixel count noise. However, in the case of pink noise (or a mixture of white and pink), that longer-term change may have a different magnitude because of the stronger contribution of smaller frequencies to the noise. The inter-pixel noise value cannot capture this larger, longer-term change. Executing all the processing of the measured signal, one obtains the final brightness temperature $T_{a,scene}$. Naturally, this $T_{a,scene}$ is not the real value, but $T_{a,scene}$ is an estimate that will have an uncertainty. If we took the inter-pixel noise value as the uncertainty due to noise, we would underestimate the uncertainty. These longer-term variations between different targets within one calibration cycle are captured in the inter-scan line method (as far as they do not exceed the timescale of two scan lines) and therefore yield a higher value as noise estimate. This possibly significant change in the amount of noise that can happen between the measurements of the Earth views and the calibration views due to pink noise should be included in an estimate of uncertainty of the final brightness temperature measurement. Therefore, avoiding underestimation of the uncertainty, we use the inter-scan-line method for the calculation of noise.

Exemplarily we investigate the noise spectrum for the different instruments and channels in some chosen orbits and years across their lifetime. Naturally, this investigation cannot fully resolve the evolution of changes in the spectrum, but our analysis provides snapshots of the overall evolution of the spectrum. The AMSU-B and MHS instruments show in their channels either pure white noise or a mixture of white and pink noise. The distribution of this characteristic
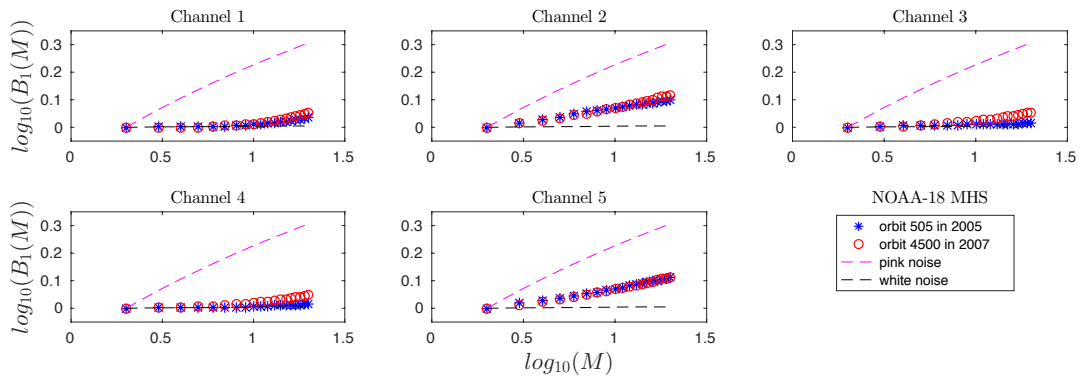
Figure 4.6 The bias functions $B_1(M)$ per channel for the orbits 505 of the year 2005 and 4500 of the year 2007 for MHS on NOAA-18.

among the channels is not fixed, however - a certain channel, for example the central water vapour channel 3, does not necessarily exhibit the same noise characteristic in all AMSU-B and MHS instruments. Furthermore, the characteristic may change as well in time. Looking at AMSU-B on NOAA-17 in Fig. 4.8 for example, channel 3 shows a strong pink component in the year 2006, whereas 2 years before in 2003 the pink component was less pronounced. This change in spectrum, adding some pink component to the noise, is also captured in our noise estimation by the inter-scan-line method: We detect a higher noise value accounting for the increased level of uncertainty that is due to the increased pink component. This is visible in the corresponding DSV count noise shown in Fig. 4.9.

For the SSMT-2 instruments, the bias function method as we use it here for analysing the noise spectrum does not work properly for all times and channels. The reason for that lies in the absolute count values that are so small for the SSMT-2 that the digitisation noise may impact and distort the picture. To improve this bias function method for the usage on data affected by digitization noise, one should simulate the digitization as well as the white and pink noise as has been presented by Mittaz (2016). Another aspect that impacts the noise analysis even more is the multitude of outliers in the measurements of the SSMT-2 instruments that often disturb the noise estimation. As mentioned above, we applied no filtering in order to get the whole picture of the instruments' behaviour: the instability in the performance of PRT, OBCT and DSV measurements of SSMT-2 is clearly visible in comparison to the other instruments. In the processing of the data to level 1c FIDUCEO FCDRs, those outliers are filtered out and do not contribute to the noise estimation executed on the fly.

4.2.3.2 Evolution of noise

We provide an overview over the evolution of noise in the different channels over the lifetime of the instruments (a detailed description of the instruments' performance is given in the Appendix). The three measures of noise, i.e. the DSV count noise, and the cold and the warm NEΔT, are displayed for all instruments and channels in Fig. 4.10, 4.11 and 4.12. The DSV count noise (see Fig. 4.10) is given in absolute counts and is therefore not suited for a comparison of noise levels of different instruments. The individual instrumental stability of the noise level can be observed very well, however. Looking at channels 3 and 4 of SSMT-2 on F14, one can observe a significant increase of the DSV count noise from 2001 on. A strong degradation of the DSV count noise is
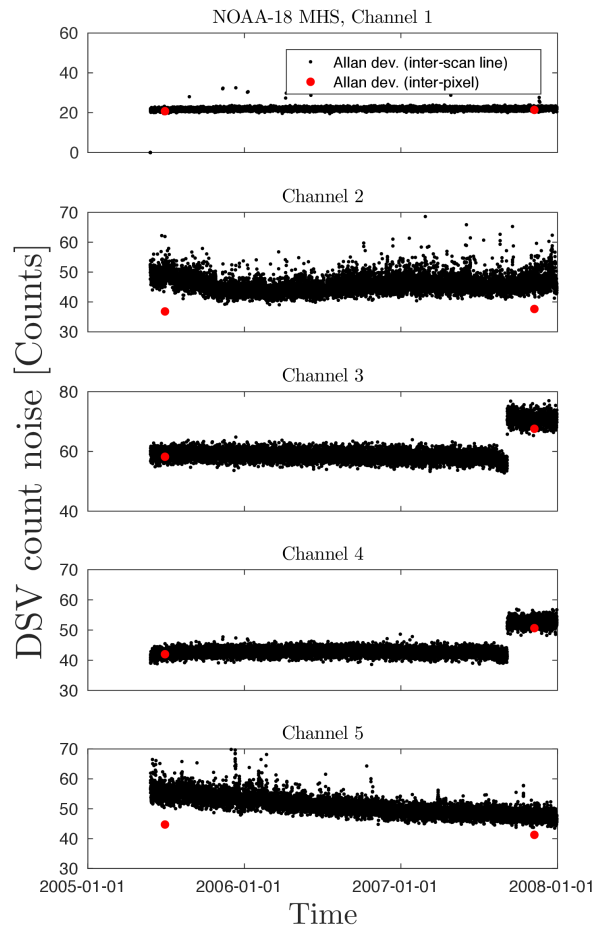
Figure 4.7 The DSV count noise per channel for the years 2005 to 2008 for MHS on NOAA-18. The red dots indicate the noise in DSV counts calculated with the inter-pixel method. For comparison with the inter-scan-line method, we applied this method exemplarily only for the two orbits 505 of 2005 and 4500 of 2007, for which we investigated the spectrum as well (see Fig. 4.6).

visible also for channel 1 of AMSU-B on NOAA-17: from 2007 on, the noise often peaks at almost 10 times higher values than its original one. Channels 3 and 4 of MHS on Metop-A show a rather smooth change over several years: from 2009 to 2012 the DSV count noise smoothly increased, then it abruptly jumped back to its initial value before increasing smoothly again. During the years 2014 to 2016 it then decreased again. The DSV count noise of AMSU-B on NOAA-15 and NOAA-16 varies only very slightly and smoothly over the lifetime of the instruments.

Both instruments, however, show a very different picture for the warm and cold NEΔT. Its evolution is displayed in Fig. 4.11 and 4.12. The NEΔT is influenced by the underlying count noise and the gain used for the conversion to temperature. Therefore, the evolution reflects the interplay of both quantities. The overall increase of NEΔT therefore relates to an increase of the count noise or a decrease of the gain. The increases in DSV count noise discussed above are quite visible in the cold NEΔT as well, e.g. for the channel 1 of AMSU-B on NOAA-17 or
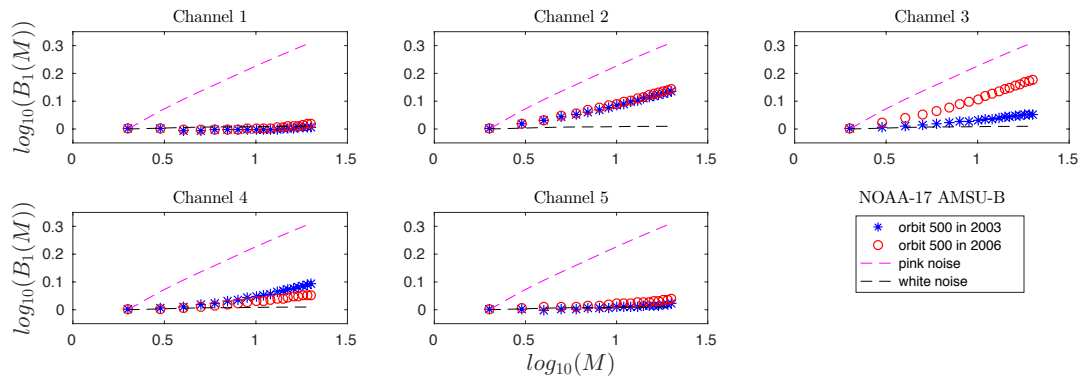
Figure 4.8 The bias functions $B_1(M)$ per channel for the orbits 500 of the year 2003 and the year 2006 for AMSU-B on NOAA-17.

for channels 3 and 4 of MHS on Metop-A. The same is valid for the count noise of the internal calibration target views and the warm NEΔT, which is usually about $0.1\,\mathrm{K}$ higher than the cold NEΔT. For channels 3 to 5 of AMSU-B on NOAA-15 and NOAA-16, which showed an almost stable count noise, the cold and warm NEΔT show a strong increase over the lifetime reaching e.g. $5\,\mathrm{K}$ in channel 3 in 2010, superimposed with an oscillating pattern. This increase is due to a strong degradation and decrease of the gain that was observed by John et al. (2013b),too. The oscillating pattern is also observed in many other measured quantities for these periods and is probably related to the change of the solar beta angle as the orbit of the satellite drifts - see the Appendix and Zou and Wang (2011). This changing pattern is also visible for cold NEΔT of channels 3 and 4 of MHS on NOAA-18 from late 2014 on. However, there is no steady degradation of the gain as for NOAA-15 and NOAA-16, such that the cold NEΔT remains at rather low values. The cold NEΔT also reflects erratic behaviour of the instrument when the smooth evolution of the quantities is interrupted by sudden jumps. For example, channels 3 and 4 of MHS on NOAA-19 suffer from an incident in late 2009 where NEΔT suddenly rises and falls again, but stays at an increased level.

4.2.4 Discussion

In this study we used the Allan deviation to calculate the evolution of the noise as well as the noise spectrum for the microwave sounders SSMT-2, AMSU-B and MHS in order to assess the quality of the data with respect to uncertainty due to noise.

The analysis of the noise spectrum showed that in some channels there is a significant non-white component that may change during the lifetime of the instruments. Together with the corresponding periods of count noise evolution in time, the analysis of the spectrum revealed that the inter-scan-line method for computing the Allan deviation is better suited for the purpose of uncertainty estimation than the inter-pixel method that underestimates the uncertainty if a pink noise component is present. Although the analysis of the noise spectrum was carried out on some orbits only, it definitely shows important aspects of the spectrum and its possible evolutions. Nonetheless, a full analysis of the noise spectrum would require a study on all orbits to track the evolution of the spectrum over time.

For the quality assessment of the microwave sounder data, we investigated the evolution of noise (count noise and NEΔT) over the lifetime of the instruments. The graphical overview we
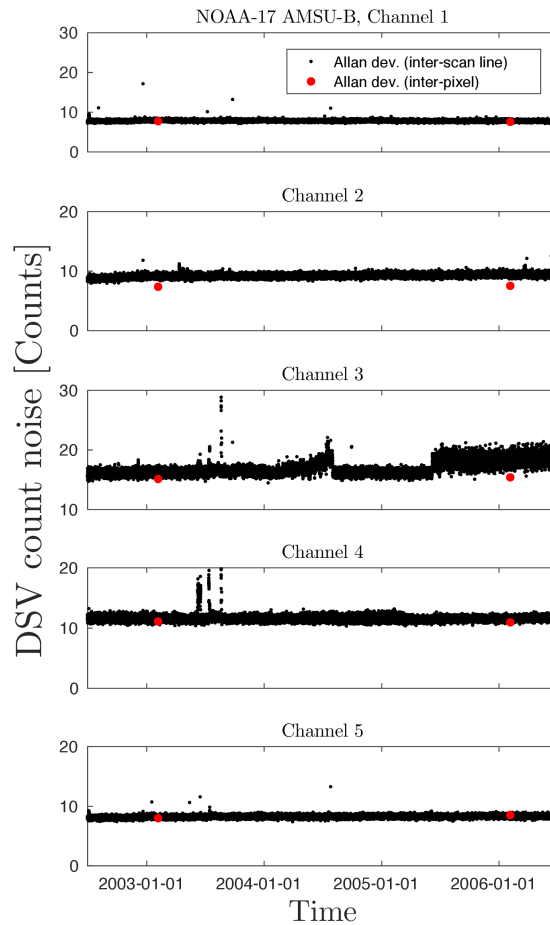
Figure 4.9 The DSV count noise per channel for the years 2003 to 2006 for AMSU-B on NOAA-17. The red dots indicate the noise in DSV counts calculated with the inter-pixel method. For comparison with the inter-scan-line method, we applied this method exemplarily only for these two orbits 500 of 2003 and 500 of 2006, for which we investigated the spectrum as well (see Fig. 4.8).

provided with Fig. 4.10 - 4.12 on the evolution of the noise gives a first impression of the quality of the data. The various outliers that we did not filter out on purpose indicate problematic periods of the instruments. The actual reasons for the various kinds of outliers are unclear.

Degradation in quality also manifests itself in an increasing cold NEΔT. This degradation can have two causes: First, the actual noise level measured in the count noise may have increased. This effect is hardly visible on mission timescales as the count noise is rather stable for most instruments. But on monthly timescales, the effect of increasing and subsequent decreasing of count noise shines through in the changes of cold NEΔT. Yet, the count noise does not cause an overall steady degradation for the investigated instruments. The second possible reason for degradation, however, has a strong impact on NEΔT in the observed cases: if the gain decreases and therefore the measured counts of DSV and OBCT converge, the NEΔT increases strongly. This reflects that the radiometer sensitivity, which NEΔT is a measure of, strongly degrades

Figure 4.10 Time evolution of the DSV count noise for the five frequency channels.
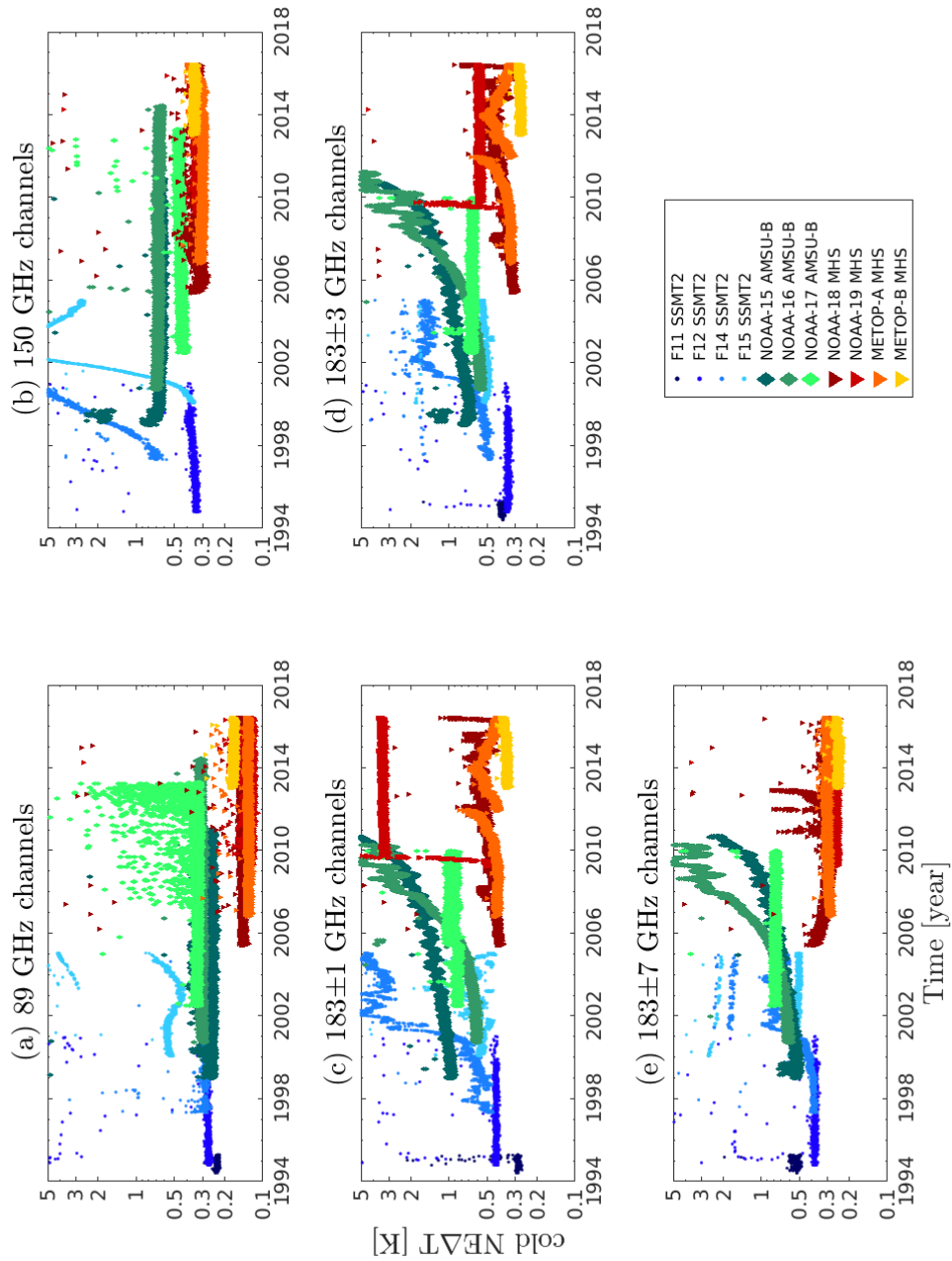
Figure 4.11 The time evolution of the cold NEΔT for the five frequency channels.
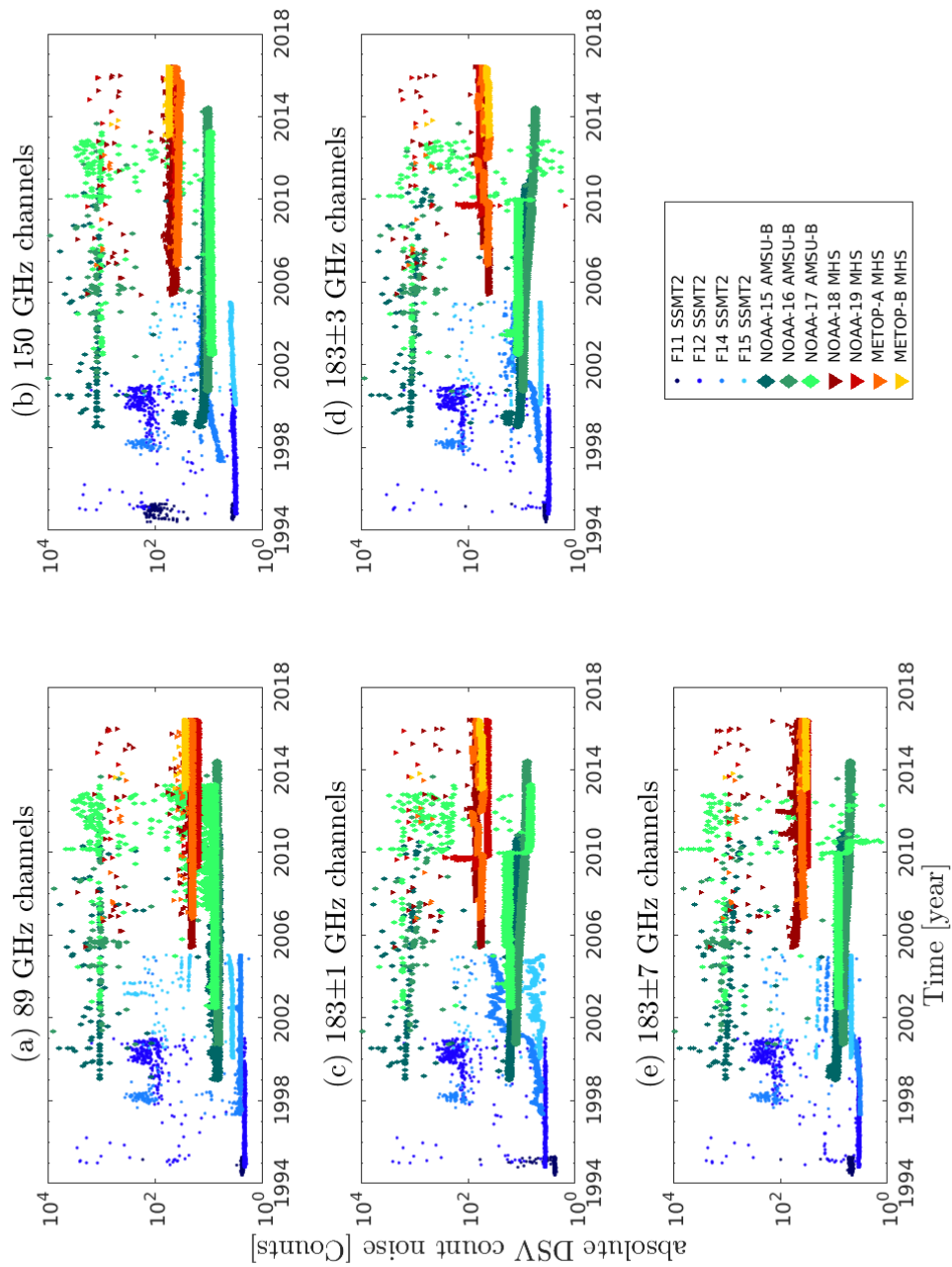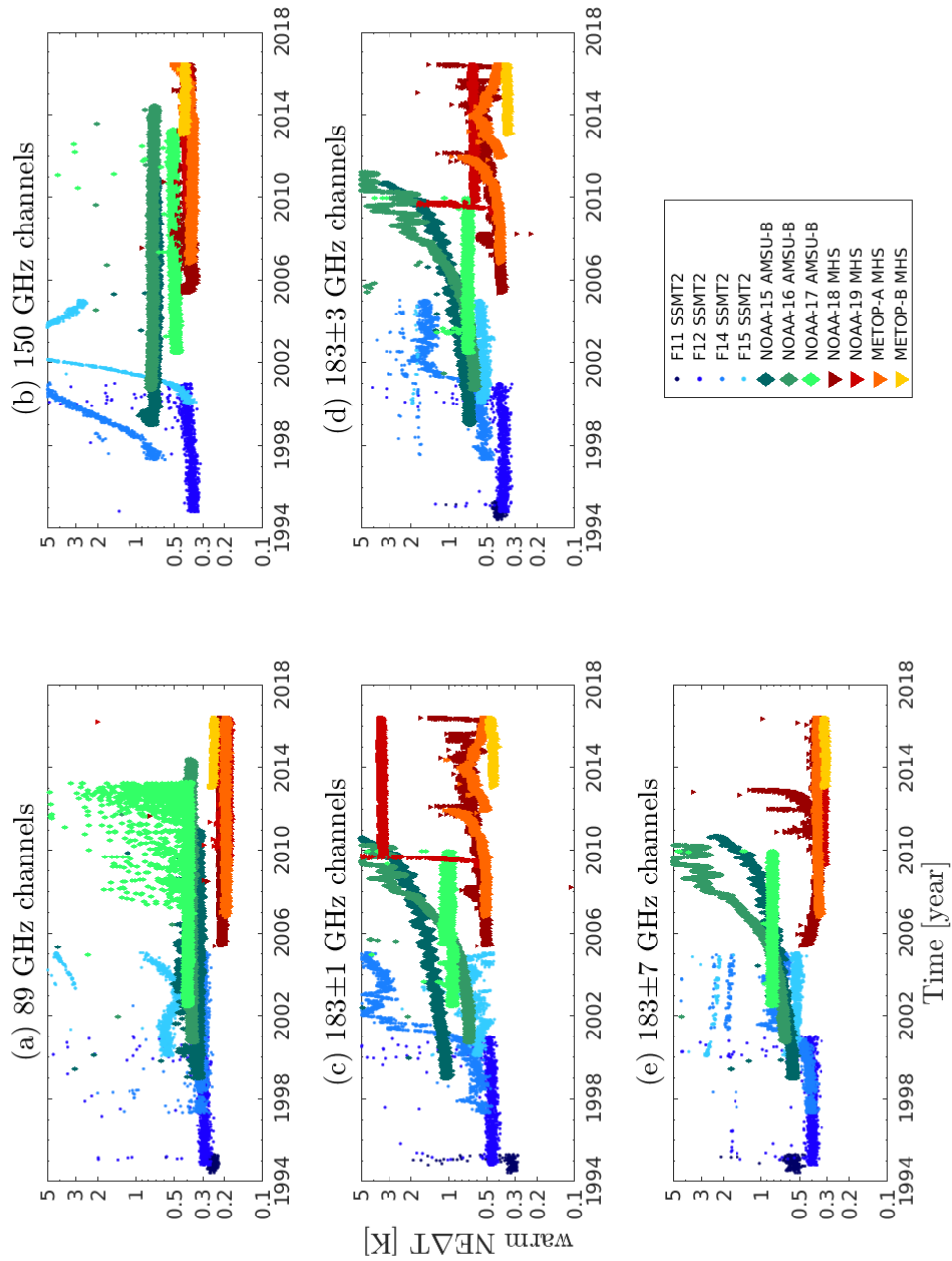
Figure 4.12 The time evolution of the warm NEΔT for the five frequency channels.

and the instrument is no longer able to distinguish temperatures properly. For example, it can only determine a temperature with an uncertainty of about $5\,\mathrm{K}$, as is the case for channel 3 of AMSU-B on NOAA-16 in 2010. This effect of gain degradation and increase of NEΔT is visible on both short and long timescales. The pattern induced by the change of the solar beta angle modifies NEΔT on monthly timescales and an overall continuous degradation of the gain causes a steady increase of cold and warm NEΔT, as seen for NOAA-15 and NOAA-16.

As intuitively obvious, an ageing satellite or sensor may degrade since its components have a limited lifetime. Accordingly, one can observe this degradation for many of the considered instruments. An interesting fact here is the different evolution for the different channels: when the three water vapour sounding channels severely degrade, the lower peaking channels may be unchanged, i.e. they may show no sign of ageing. Or, there are events that are visible in all channels, but only have long-lasting impact on certain channels. For the newer satellites, some adjustments were made during operation and this protected the instruments from degradation and kept them at an acceptable noise level. The lowest and most stable noise, but also the shortest data record so far, was from the MHS instrument on board the Metop-B satellite.

As an easy-to-use tool for information on noise we provided plots of the time evolution for all individual instruments of this microwave sounder family. These plots may help to decide on the usability of the data for a certain application. They were given for the DSV count noise, the warm and the cold NEΔT. Users of the data have to decide which level of uncertainty their product generation might still bear and which threshold of NEΔT they would set to limit the uncertainty. As a further result, we provide a chart in Fig. 4.13, which shows the periods of data for a threshold of cold NEΔT $< 1\,\mathrm{K}$.
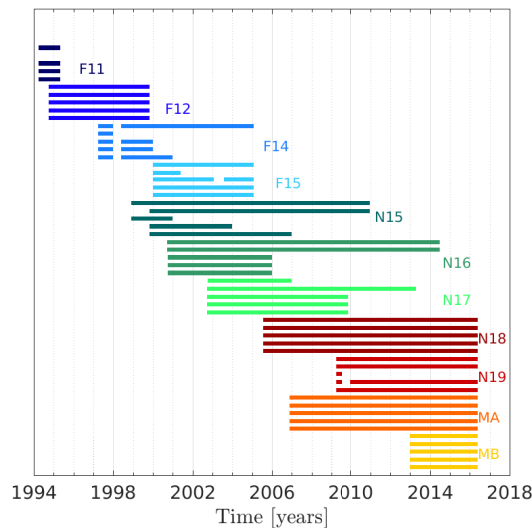
Figure 4.13 Usable microwave data records with cold NEΔT $<$1K. The five bars per satellite correspond to the channels 1 to 5 (from top to bottom).

For atmospheric product retrieval, Fig. 4.11 and 4.12 together with eq. 4.6 can be used to estimate the correct scene NEΔT. Since warm and cold NEΔT typically differ by only approximately $0.1\,\mathrm{K}$, a reasonable approximation would be also to simply use the warm NEΔT as an estimate for the scene NEΔT.

4.2.5 Conclusion

The results of our study provide users with information on the uncertainty due to noise that they should expect when using the data sets of the microwave sounders SSMT-2, AMSU-B and MHS.

The chart in Fig. 4.13 reveals the possibility to concatenate the available data for constructing gap-less long time series since 1994 at a noise level below 1 K for all frequency intervals that the instruments cover. This is of major interest for climate researchers who need long time series with low noise levels in order to investigate possible trends.

Apart from the stand-alone results as information content for users of these microwave sounders' data, our analysis is of direct use for the FIDUCEO project: the method for estimating the count noise for the DSV and OBCT will be used in the processing of level 1b to level 1c FIDUCEO FCDR in order to provide on-the-fly input values for the uncertainty propagation. This FCDR will provide a field-of-view-wise estimate of uncertainty in brightness temperature due to count noise for every scan line and orbit. Additionally, the FCDR will contain extensive information that will further close the gap of lacking information on uncertainty.

Acknowledgements

5 The production of the microwave fundamental climate data record (MW-FCDR)

This chapter describes the actual step from raw level 1b data to calibrated level 1c brightness temperature. This step is a complex calibration process in itself. However, the FIDUCEO FCDR reaches even further and adds many aspects so far lacking in the available level 1c data sets. Most importantly, FIDUCEO adds the aspect of uncertainty. Within the FIDUCEO vocabulary (FIDUCEO, 2018), an uncertainty-quantified fundamental climate data record (FCDR) is defined as

> A record of calibrated, geolocated, directly measured satellite observations in geophysical units (such as radiance) in which estimates of total uncertainty (or error covariance) and/or dominant components of uncertainty (or error covariance) are provided or characterised at pixel-level (and potentially larger) scales. The FCDR should be provided with all relevant auxiliary information for the data to be meaningful, including, e.g. time of acquisition, longitude and latitude, solar and viewing angles, sensor spectral response.

To meet these requirements, the new FCDR generation process described here includes a reshaping of the orbital data to files reaching from one equator crossing in a chosen flight direction to the next crossing (see Sec. 5.3.1 for the choice of direction). This removes all doubled data resulting from overlapping subsequent orbits. Moreover, a detailed but user-friendly set of quality information flags is provided based on quality checks during the calibration process. The most important aspect in the new level 1b to the level 1c processing is the propagation of uncertainty that finally provides uncertainty information on pixel level. For the propagation of uncertainty, I first have to identify the various sources of uncertainty by a rigorous measurement equation approach: the measurement equation is defined with all effects that disturb the measurement of the actual signal. For all these effects, the physical origin, the impact on the measurement equation and the propagation of uncertainty is explained in the first part of this chapter. In the second part of this chapter, I describe the algorithms used for the FCDR production and point to the corresponding scripts of the FCDR production code. The format of the FCDR files is also explained. This chapter concludes with example contents from the new (but still unharmonised) FCDR.

Parts of this chapter have already been published in FIDUCEO project reports:

> M. Burgdorf, I. Hans, M. Prange, J. Mittaz, E. Woolliams: D2_2 Microwave: Report on the MW FCDR uncertainty, Universität Hamburg, National Physical Laboratory, 2017

> I. Hans, M. Burgdorf, Emma Woolliams: Product User Guide - Microwave FCDR Release 0.2, Universität Hamburg, National Physical Laboratory, 2018

5.1 The MW measurement equation and propagation of uncertainty

This section explains the measurement equation approach and the uncertainty propagation and uncertainty classification with respect to correlation behaviour. This classification is of interest for the generation of higher level products in a CDR where regridding and averaging is used.

5.1.1 Measurement equation

The measurement equation is a mathematical expression connecting all quantities involved in a measurement (Woolliams et al., 2017). For the instruments considered here, the measurement equation is an analytical expression that relates the measured raw counts (which are digitised voltages) to the desired output of radiance or brightness temperature. The measurement model underlying the measurement equation for the considered MW sounders is a two-point calibration model: from a hot and a cold reference measurement the desired output quantity is estimated from a linear interpolation between those references. In its simplest form, the measurement equation would read

$$L_{\mathrm{E}} = L_{\mathrm{IWCT}} + \frac{L_{\mathrm{IWCT}} - L_{\mathrm{S}}}{C_{\mathrm{IWCT}} - C_{\mathrm{S}}} \cdot (C_{\mathrm{E}} - C_{\mathrm{IWCT}}) \qquad (5.1)$$

where L stands for radiance and C for count. The indices IWCT, S and E refer to internal warm calibration target, space and Earth respectively. Hence, the equation translates the counts obtained for a measurement of Earth (in practice one field of view, i.e. pixel) to radiance. This equation would need to be evaluated for every field of view, every scan line and every channel. However, for the actual processing from counts to radiance, a much more complex equation needs to be used. The complexity builds up with adding various effects that affect the measurement process. In the following, I build up the full measurement equation used in the processing. The details on the effects named here, are discussed in the section on individual effects (see section 5.2).

First of all, the calibration measurements for the internal warm calibration target and space are weighted mean values. The same holds for the measured temperature of the internal warm calibration target used to estimate L_{IWCT}. The averaged values are marked as $\overline{x}$ with x being place holder for the different quantities.

The linear interpolation is a first order, but reasonable, approximation of the instruments behaviour. Nonetheless, a non-linearity correction in shape of a second order term in C_E is applied: ΔL_{nl}.

Hence, the measurement equation has changed its form to

$$L_{\mathrm{E}} = L_{\mathrm{IWCT}} + \frac{L_{\mathrm{IWCT}} - L_{\mathrm{S}}}{\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}}} \cdot (C_{\mathrm{E}} - \overline{C}_{\mathrm{IWCT}}) + \Delta L_{\mathrm{nl}}. \qquad (5.2)$$

Further corrections are applied to the calculation of L_{IWCT} and L_{S}, so that they now enter as $\tilde{L}_{\mathrm{MIWCT}}$ and $\tilde{L}_{\mathrm{MS}}$, taking account of the fact that they are measured "representatives" of the radiance of the internal warm calibration target L_{IWCT} and of the radiance seen in the deep space view L_{S} rather than L_{IWCT} and L_{S} themselves.

A similar argument holds for the radiation emitted by Earth, L_{E}: the instrument does not measure the pure signal coming from Earth, but because of stray radiation entering through the side-lobes of the antenna, it measures a contaminated signal L_{ME} composed of Earth radiation,

cosmic microwave background radiation (CMB) and platform radiation (Pl). To account for this, one applies an antenna pattern correction (coefficients g_E, g_S, g_{Pl}). A final correction is added to account for polarisation effects: ΔL_{pol}. Therefore, we obtain as measurement equation for obtaining L_E

$$L_E = \frac{1}{g_E}(L_{ME} - g_S L_{CMB} - g_{Pl} L_{Pl}) + \Delta L_{pol} \tag{5.3}$$

$$\text{with} \quad L_{ME} = \tilde{L}_{IWCT} + \frac{\tilde{L}_{IWCT} - \tilde{L}_S}{\overline{C}_{IWCT} - \overline{C}_S} \cdot (C_E - \overline{C}_{IWCT}) + \Delta L_{nl} \tag{5.4}$$

Including all these corrections with their corresponding equations we obtain the full measurement equation:

$$\begin{aligned}
L_E = \frac{1}{1 - g_S - g_{Pl}} & \left[B(\nu_{ch}, A + b \cdot (\overline{T}_{IWCT} + \delta T_{ch})) \right. \\
& + \frac{B(\nu_{ch}, A + b \cdot (\overline{T}_{IWCT} + \delta T_{ch})) - B(\nu_{ch}, A_s + b_s \cdot (T_{CMB0} + \Delta T_c))}{\overline{C}_{IWCT} - \overline{C}_S} \cdot (C_E - \overline{C}_{IWCT}) \\
& + q_{nl} \frac{(C_E - \overline{C}_S) \cdot (C_E - \overline{C}_{IWCT})}{(\overline{C}_{IWCT} - \overline{C}_S)^2} \\
& \quad \cdot \left[B(\nu_{ch}, A + b \cdot (\overline{T}_{IWCT} + \delta T_{ch})) - B(\nu_{ch}, A_s + b_s \cdot (T_{CMB0} + \Delta T_c)) \right]^2 \\
& \left. - g_S \cdot B(\nu_{ch}, A_s + b_s \cdot (\overline{T}_{CMB0})) - g_{Pl} L_{Pl} \right] \\
& + \alpha[B(\nu_{ch}, A + b \cdot (\overline{T}_{IWCT} + \delta T_{ch})) - L'_E] \cdot \frac{1}{2}(\cos 2\theta_E - \cos 2\theta_S)
\end{aligned} \tag{5.5}$$

The "measurement diagram" visualises the dependences of the different terms of the equation also reaching to deeper levels. Figure 5.1 shows the measurement diagram for the MHS, AMSU-B and SSMT-2 instruments. The final radiance for the measured Earth pixel is represented in the violet box as L_E. It is obtained in a last corrective step from the core measurement equation: the blue box in the middle represents the core measurement equation including the linear interpolation (two-point calibration) and the additional term ΔL_{nl} accounting for non-linearity. The blue box is linked to surrounding boxes by "twigs" representing the sensitivity of the core measurement equation to certain quantities. These quantities (having their own colour) again depend on certain quantities from deeper levels of the calibration process. At the very end of each twig, the uncertainties associated with the quantity are indicated.

At several points and levels within this measurement diagram, a term "+0" is added. This term represents the fact that we know that the measurement equation as it is written is only an approximation to the truth. This equation is a simplified model of the measurement process. For example, the non-linearity correction assumes a quadratic non-linearity only. This is a strong simplification as the pre-launch tests do not necessarily confirm this assumption. Nonetheless, this non-quadratic non-linearity is small compared to other effects and therefore included in the "+0". Note that unknown effects may impact on the measurement also. As they are unknown, they cannot be modelled or corrected for. Hence, the measurement equation indicates its imperfection by the "+0".
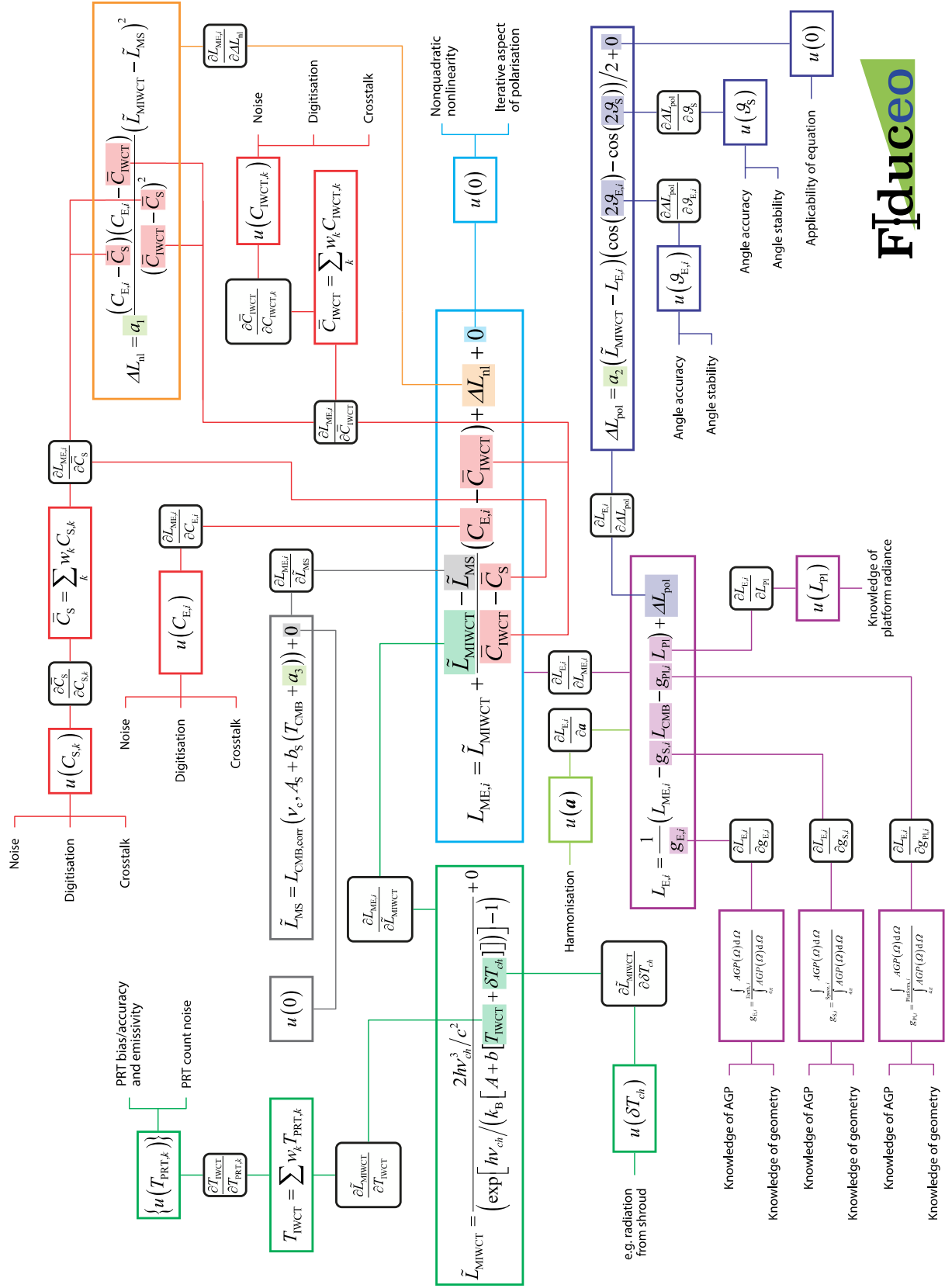
Figure 5.1 The full MW measurement equation. Graphical design: Jonathan Cherry and Emma Woolliams.

5.1.2 Uncertainty information and propagation

The FIDUCEO philosophy of uncertainty information and propagation, presented in project reports by Woolliams et al. (2017) and Merchant et al. (2017), has been developed within the project consortium to achieve harmony between different instrument types and provide a most coherent FCDR product family. As the uncertainty propagation is part of the MW-FCDR generation effort carried out in the scope of this thesis, I briefly present in this section the main aspects of this metrology-inspired approach explained in detail by Woolliams et al. (2017) and Merchant et al. (2017).

Uncertainty is defined in the International Vocabulary of Metrology (VIM, 2008) as

> a non-negative parameter characterizing the dispersion of the quantity values being attributed to a measurand, based on the information used.

In other words, the measured value in a measurement along with its unknown error (i.e. deviation from the truth) is understood as a draw from a probability distribution associated with the measurement. The uncertainty of that measured value is in turn a measure of the underlying probability distribution.

Of course, one tries to correct for every known error within a measurement process. However, even after correction, it remains an unknown deviation from the true value. Random errors cannot be corrected for even in principle, and hence, a second unknown deviation from the truth remains.

Nonetheless, one can estimate the uncertainty associated with the measurement. For this, a detailed uncertainty analysis combining uncertainties from all effects is required. The uncertainty analysis starts from the measurement equation that defines the calibration process disturbed by certain effects. The disturbing effects are partly corrected for - nonetheless, the remaining unknown error requires the estimation of uncertainty that this effect produces for the final measurand, the brightness temperature. The Guide to the Expression on Uncertainty in Measurement (GUM, 1995) suggests two methods to determine uncertainty: first, analytically with the Law of Propagation of Uncertainty and second, numerically with Monte Carlo Methods. The latter one has the advantages of being more representative with highly non-linear measurement models and complex probability distributions. However, it is numerically expensive due to numerous iterations required. For the FIDUCEO FCDRs, we use the Law of Propagation of Uncertainty. This only asks for the sensitivity coefficients c_j, i.e. the partial derivatives, associated with each effect and additionally the uncertainty $u(x_j)$ on the calibration parameters for the effect itself:

$$u^2(L) = \sum_{j=1}^{n_j} c_j^2 u^2(x_j) + 2 \sum_{j=1}^{n_j-1} \sum_{j'=j+1}^{n_j} c_j c_{j'} u(x_j, x_{j'}) \tag{5.6}$$

If the individual parameters x_j are uncorrelated, i.e. they do not share a common error effect, the second term is zero. This metrologically independent case is assumed for the MW, since it is valid for at least most of the identified effects. Hence, the uncertainty on the final radiance simplifies to

$$u^2(L) = \sum_{j=1}^{n_j} c_j^2 u^2(x_j) \tag{5.7}$$

The input uncertainty $u(x_j)$ needs to be estimated. According to metrological standards (GUM, 1995), an uncertainty estimate may be obtained from statistical analysis and from expert judgement. Both methods are used within the FIDUCEO FCDR generation.

The uncertainty information in the FCDR is the key element for the later derivation of any level 2 (or higher) products' uncertainty information. This higher level uncertainty information is required for numerical weather prediction reanalysis, climate modelling, climate prediction and eventually political decisions. It must allow for the discrimination of observations with lesser or greater uncertainty and also quantify uncertainty across the applied averaging scales of data (Merchant et al., 2017). To meet these requirements in higher-level derived products, detailed, but user-friendly uncertainty information needs to be provided on the FCDR level. The FIDUCEO approach for the "easy-FCDR" handles the trade-off between "detailed and adequate for the complexity" on one side and "easy-to-use and therefore probable to be used" on the other side by the following scheme: apart from the brightness temperature, geolocation and quality information, there is uncertainty information provided as

- per-pixel, per-channel magnitude of brightness temperature uncertainty

- per-orbit, per-channel length scales of cross-element and cross-line brightness temperature error correlation

- per orbit cross-channel brightness temperature error correlation

Note that pixel is understood here as "point" with coordinates (element, line) with element referring to the field of view (FOV) and line to the scan line of a certain orbit.

This set of uncertainty-related variables is necessary to represent the variability of the uncertainty in brightness temperature between pixels and channels and to provide information on the spatial and spectral correlation behaviour. This information is of interest for CDR producers as they will usually apply averaging procedures.

Within the provided uncertainty information, we distinguish three classes of effects (i.e. error sources) that have different correlation behaviour:

- **Independent errors**: There are effects sufficiently close to white noise causing deviations from truth that vary independently between pixels in a channel. Hence the term "independent error". Errors of this class are random in the sense that their origin is stochastic and that they cannot be corrected for even in principle.

- **Structured errors**: This class refers to effects that have a certain spatial structure within one orbit. The structure emerges from calibration procedures that result in errors being closely related from one pixel to the next. In this sense, originally random effects might result in a structured one on the level of brightness temperature. Also systematic effects (that could be corrected for in principle if better information was available) may contribute to structured errors.

- **Common errors**: This class refers to effects that have correlation scales exceeding one orbit; often they are correlated across the whole mission. Systematic effects will enter this class. This also holds for any calibration parameters optimised within the harmonisation procedure that retrospectively re-calibrates the instruments.

The total uncertainty on the brightness temperature is composed of all three contributions: the uncertainty due to independent effects, due to structured effects and due to common effects.

But, by construction, the three components behave differently when data is aggregated. For independent effects, the uncertainty reduces differently in an averaging process than for the structured effects. Also, the uncertainty due to common effects cannot be reduced at all. This is crucial knowledge for the producer of gridded and averaged products, and hence the FIDUCEO easy MW-FCDR provides separately, for each channel

- a variable of the same dimension as the brightness temperature containing uncertainty due to the combination of all independent effects,

- a variable of the same dimension as the brightness temperature containing uncertainty due to the combination of all structured effects,

- and a variable of the same dimension as the brightness temperature containing uncertainty due to the combination of all common effects.

Apart from the explained uncertainty variables, the easy FCDR also provides information on correlation, as indicated above. This information comes as typical length scales over which correlation is expected and as cross-channel correlation matrix. The full cross-channel correlation matrices (for the independent, structured and common effects) are provided in the FCDR files since they are of tractable size for all instruments investigated by FIDUCEO (Merchant et al., 2017). In the MW case, these matrices are three 5 x 5 matrices that store the cross-channel correlation information for the three effect classes. These three classes are necessary for the following reason: the classes were defined with respect to the spatio-temporal correlation structures - however, these do not necessarily coincide with the spectral correlation. This means, that the independent effects may still show correlation across the channels.

To compute the cross-channel correlation matrix for all effects of a certain effect class, a scheme has been developed by Merchant et al. (2017). It first requires the cross-channel (index c) parameter error correlation matrix $\mathbf{R}_c^{p,k}$ for each effect k for all pixels p. In this context, a pixel is again understood as a tupel (*line, element*) with *element* spanning the available fields of view (FOVs). This cross-channel parameter error correlation matrix is a 5 x 5 matrix for the MW case and has the form

$$\mathbf{R}_c^{p,k} = \begin{pmatrix} 1 & r_k(x_j(1), x_j(2)) & \dots \\ r_k(x_j(2), x_j(1)) & 1 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{5.8}$$

where $r_k(x_j(1), x_j(2))$ is the error correlation coefficient between the measurement equation term x_j evaluated for channel 1 and channel 2. Strictly speaking, this matrix would have to be evaluated for all pixels, which would include a huge computational effort. In this thesis, I provide rough estimates, valid for all pixels, to capture the overall nature of the individual effects. The individual $\mathbf{R}_c^{p,k}$ per effect k are given in the sections on the effects.

The second matrix needed per effect and pixel is the cross-channel parameter uncertainty matrix $\mathbf{U}_c^{p,k}$, having the same dimension as $\mathbf{R}_c^{p,k}$. The uncertainty matrix $\mathbf{U}_c^{p,k}$ is diagonal with the elements of the per-channel uncertainty associated with the effect and pixel:

$$\mathbf{U}_c^{p,k} = \begin{pmatrix} u_k(x_j(1, l, e)) & 0 & \dots \\ 0 & u_k(x_j(2, l, e)) & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{5.9}$$

$u_k(x_j(1, l, e))$ is the uncertainty in pixel $p = (l, e)$ in channel 1 due to the term x_j in the measurement equation associated with the effect k. The $u_k(x_j(ch, l, e))$ are computed as part of the uncertainty propagation in the FCDR production.

Having these matrices for all effects, the cross-channel error correlation matrix for the three effects classes can be computed as explained in the following paragraphs.

Independent effects The cross-channel error correlation matrix for the independent effects i is computed as follows: At first, the covariance matrix is obtained from

$$\mathbf{S}_{c,i} = \langle \sum_i \mathbf{U}_c^{p,i} \mathbf{R}_c^{p,i} \mathbf{U}_c^{p,i^T} \rangle_p \tag{5.10}$$

where an average over pixels p is executed. This average must be computed over a useful range of pixels (if not all are taken). The pixel range must be chosen such that it samples both lines and elements in a representative manner. In the MW case, I choose the full set of elements, i.e. all FOVs meaning 90 for the MHS and AMSU-B and 28 for the SSMT-2. For the lines, a reasonable thinning must be applied since one orbit file contains about 2300 lines, which is beyond reasonable computational effort for the correlation estimation. I sample every 100th scan line in order to preserve the overall evolution of the uncertainty of one FOV and channel across the orbit. In Fig. 5.2(a), the uncertainty due to antenna position in Earth views for FOV 45 in channel 2 of an MHS NOAA-18 orbit is displayed. The pattern of varying uncertainty over the orbit mainly originates from the polarisation correction. The antenna position in Earth views influences the brightness temperature only through the polarisation correction term. Hence, the uncertainty in brightness temperature due to the antenna position in Earth views behaves qualitatively similarly to the uncertainty due the polarisation correction parameter: larger uncertainty in colder regions (e.g. at the poles) and smaller uncertainty in warmer regions. This is reproduced in the pattern visible in Fig. 5.2(a), showing the expected change of warm and cold regions with low and high uncertainty, respectively, over the course of the orbit starting and ending at the equator. Figure 5.2(b) shows the computed covariances for channel 1 and 2 at every 100th scan line. The most important features are captured by this sampling. These covariances (per channel) are then averaged over for all FOVs to obtain the 5 x 5 matrix $\mathbf{S}_{c,i}$. The final step computes the cross-channel error correlation matrix $\mathbf{R}_{c,i}$ as

$$\mathbf{R}_{c,i} = \mathbf{U}_{c,i}^{-1} \mathbf{S}_{c,i} \mathbf{U}_{c,i}^{-1^T} \tag{5.11}$$

with the matrix $\mathbf{U}_{c,i}$ being

$$\mathbf{U}_{c,i} = \begin{pmatrix} \sqrt{[S_{c,i}]_{1,1}} & 0 & \cdots \\ 0 & \sqrt{[S_{c,i}]_{2,2}} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{5.12}$$

If all independent effects are also spectrally independent, then $\mathbf{R}_{c,i}$ is the identity matrix.

Structured effects The computation of the cross-channel error correlation matrix for structured effects s is the same as for the independent effects:

$$\mathbf{S}_{c,s} = \langle \sum_s \mathbf{U}_c^{p,s} \mathbf{R}_c^{p,s} \mathbf{U}_c^{p,s^T} \rangle_p \tag{5.13}$$

$$\mathbf{R}_{c,s} = \mathbf{U}_{c,s}^{-1} \mathbf{S}_{c,s} \mathbf{U}_{c,s}^{-1^T} \tag{5.14}$$
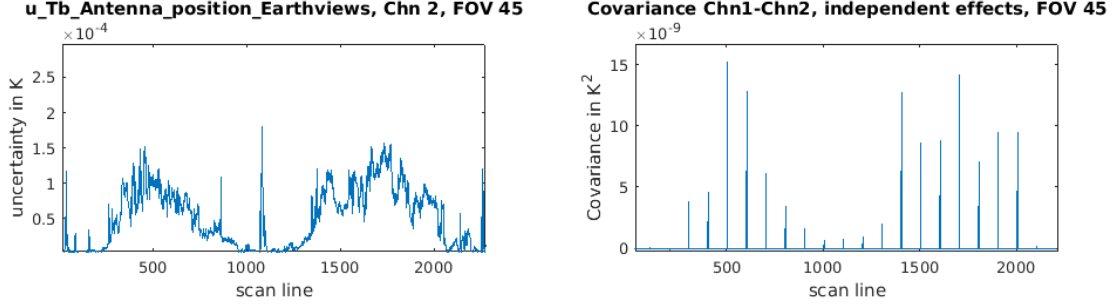
(a) Uncertainty of T_b due to Antenna position in Earth views

(b) Covariance of channel 1 and 2 for independent effects, computed for every 100th scan line

Figure 5.2 Justifying the chosen sampling at every 100th scan line for the cross-channel correlation computation: the main features of change in uncertainty (in (a)) are captured in (b). Uncertainty due to noise on Earth counts (second contributor to independent effects) not shown, since similar.

with the matrix $\mathbf{U}_{c,s}$ being

$$\mathbf{U}_{c,s} = \begin{pmatrix} \sqrt{[S_{c,s}]_{1,1}} & 0 & \cdots \\ 0 & \sqrt{[S_{c,s}]_{2,2}} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{5.15}$$

Common effects The computation of the cross-channel error correlation matrix for structured effects *co* is the same as for the independent effects:

$$\mathbf{S}_{c,co} = \langle \sum_s \mathbf{U}_c^{p,co} \mathbf{R}_c^{p,co} \mathbf{U}_c^{p,co^T} \rangle_p \tag{5.16}$$

$$\mathbf{R}_{c,co} = \mathbf{U}_{c,co}^{-1} \mathbf{S}_{c,co} \mathbf{U}_{c,co}^{-1^T} \tag{5.17}$$

with the matrix $\mathbf{U}_{c,co}$ being

$$\mathbf{U}_{c,co} = \begin{pmatrix} \sqrt{[S_{c,co}]_{1,1}} & 0 & \cdots \\ 0 & \sqrt{[S_{c,co}]_{2,2}} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{5.18}$$

These three cross-channel error correlation matrices are provided in the easy FCDR.

Further correlation information is given for the spatio-temporal correlations, mostly concerning the structured effects: The typical length scales of cross-element and cross-line brightness temperature error correlation. The latter one is only affected by the structured effects, per definition (independent effects have zero correlation across lines, common effects are correlated even over more than one orbit). The correlation across lines is introduced by the 7-scan line averaging procedure for temperature measurements and calibration target measurements in the calibration (see Sec. 5.3.2.4). Two scan lines separated by a distance larger than 7 scan lines, do not share any calibration measurement and are therefore uncorrelated. Hence, the length scale over which the correlation vanishes is 7 scan lines. The cross-element brightness temperature correlation length scale, I estimate as the full range of FOVs (90 and 28 for AMSU-B, MHS and SSMT-2 respectively): every FOV within a scan line is calibrated with the same underlying calibration

measurements, hence there is full correlation for the structured effects across all FOVs. Some common effects, as the antenna pattern correction, will only correlate over a shorter range of about 5-10 FOVs (see Sec. 5.2.1), but this is not taken into account here as the structured effects already correlate over the whole range of FOVs. For the independent effects, the length scale is zero FOV by construction.

With the per pixel uncertainty variables for all effects classes, the cross-channel correlation matrices and the typical correlation length scales, the easy FCDR users are provided with extensive uncertainty and correlation information.

5.2 The individual effects represented in the measurement equation

The complexity of the measurement diagram emerges from the various effects that need to be accounted for in order to retrieve the actual signal one is interested in. "Effect" means here any disturbing influence on the calibration that distorts the true signal. The equations that model each individual effect are incorporated in the measurement equation and represented within the diagram by coloured boxes. The branches trace the effects down to lower levels of the calibration procedure, indicated also by the partial derivatives, and finally relate the quantity that is causing the effect to the uncertainty of the quantity.

In this section I present the known effects that have an impact on the calibration procedure: I explain their origin and the models used to quantify them in the calibration process. From the full measurement equation, I deduce the sensitivity coefficient and also give information on its correlation structure. The assumptions made on the input uncertainty are based on statistical analysis and expert judgement (see Sec. 5.1.2).

This section is the physical basis for the translation of the calibration process into algorithms that finally build the FCDR production code.

5.2.1 The antenna pattern correction

Origin of the effect and its model in the measurement equation

The microwave radiometers are designed such that their main beam efficiency is larger than 95%, the rest of the detected radiation comes from side lobes of the antenna. Pre-launch tests analysed the antenna gain pattern (AGP) to investigate the effect of the side lobes on the final measurement of antenna temperature (Hewison and Saunders, 1996).

The antenna temperature T_A, can be obtained from the convolution of the antenna gain pattern $F(\theta, \phi)$ with the far-field radiance distribution $B(\nu, T_{AP}(\theta, \phi))$ of a source radiating at a temperature T_{AP} and frequency ν (Hewison and Saunders, 1996):

$$T_A = B^{-1} \left\{ \frac{\int_{4\pi} B(\nu, T_{AP}(\theta, \phi)) F(\theta, \phi) d\Omega}{\int_{4\pi} F(\theta, \phi) d\Omega} \right\} \tag{5.19}$$

with $d\Omega = \cos\theta d\theta d\phi$ being the elemental solid angle and B and B^{-1} representing the Planck Function and its inverse. During the swath, the side lobes are exposed to different radiation. This effect is strongest for the pixels at the edge of the scan: Here, the side lobes detect radiation from space or from the platform. Figure 5.3a shows a coarse map of the radiance distribution due to the geometry of the instrument itself and other instruments on board the satellite (a
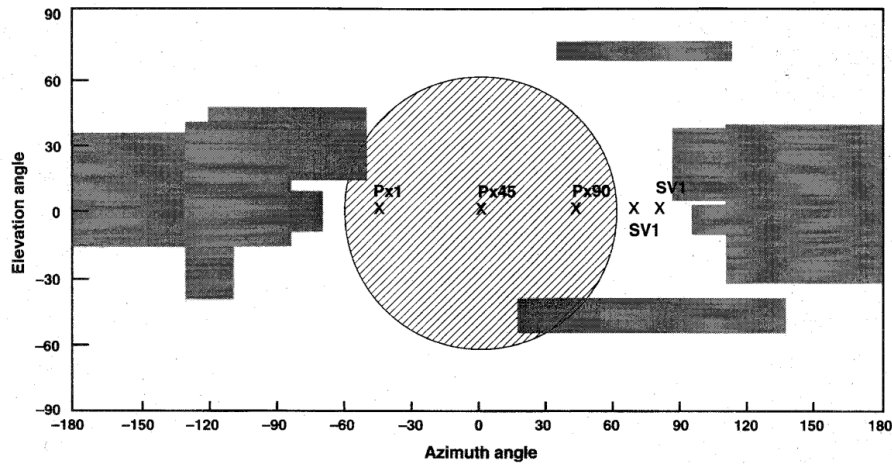
(a) AMSU-B scan environment for NOAA platforms
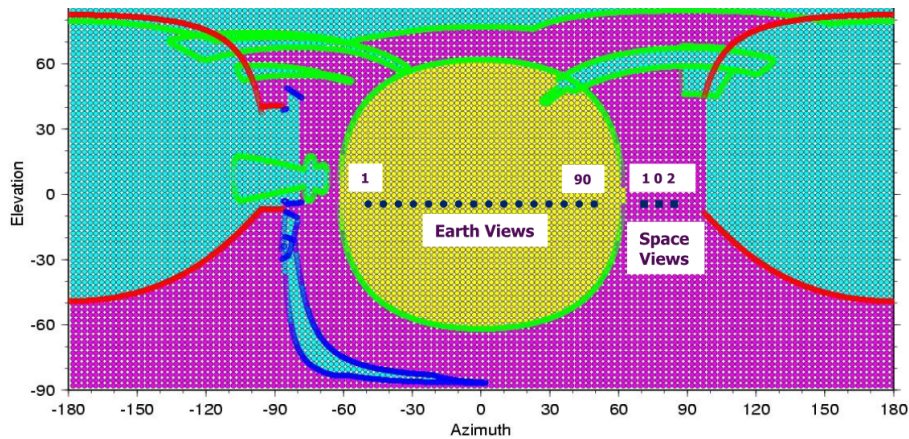
(b) MHS scan environment on Metop satellites

Figure 5.3 Geometrical map of the environment of the instrument on board the satellite. Figure 5.3a from Hewison and Saunders (1996), ©1996, IEEE. Figure 5.3b from Klaes and Ackermann (2014), with friendly permission by Jörg Ackermann.

corresponding map for the MHS instruments on board the Metop satellites is shown in Fig. 5.3b). Together with the measured antenna pattern and Eq. 5.19, this map was used in Hewison and Saunders (1996) to predict the required correction for the antenna temperature for each FOV.

This correction, presented in Hewison and Saunders (1996), can also be expressed as the overall efficiencies g_E, g_S, g_{Pl} of the antenna with which it detects radiation from Earth (E), space (S) and platform (Pl). Values for these efficiencies are stored in the calibration file `fdf.dat` of the operational processor AAPP. Their origin is not explicitly documented in AAPP, but they must have been obtained from a convolution of the geometry map with the measured antenna pattern as indicated for AMSU-B by Hewison and Saunders (1996) and explained for MHS in Klaes and Ackermann (2014). See Fig. 5.4 for their values for the individual instruments. The efficiencies g vary for the pixels i along the swath. The efficiencies can be expressed with $AGP(\Omega) = F(\theta, \phi)$ as

73

$$g_E(i) = \frac{\int_{\text{Earth},i} AGP(\Omega)d\Omega}{\int_{4\pi} AGP(\Omega)d\Omega} \tag{5.20}$$

$$g_S(i) = \frac{\int_{\text{Space},i} AGP(\Omega)d\Omega}{\int_{4\pi} AGP(\Omega)d\Omega} \tag{5.21}$$

$$g_{Pl}(i) = \frac{\int_{\text{Platform},i} AGP(\Omega)d\Omega}{\int_{4\pi} AGP(\Omega)d\Omega} \tag{5.22}$$

which are indicated as the quantities associated with the antenna pattern correction effect within the measurement diagram Fig. 5.1.

The sum of the three radiance contributions gives the antenna temperature

$$T_{\text{A}} = B^{-1}\left\{g_{\text{E}}B(T_{\text{E}}) + g_{\text{S}}B(T_{\text{S}}) + g_{\text{Pl}}B(T_{\text{Pl}})\right\}. \tag{5.23}$$
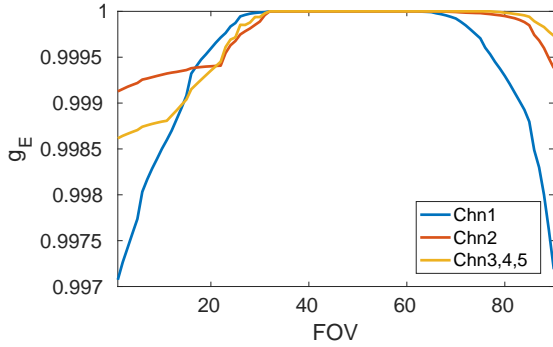
(a) all AMSU-Bs and the MHS on NOAA-18, correct channel assignment
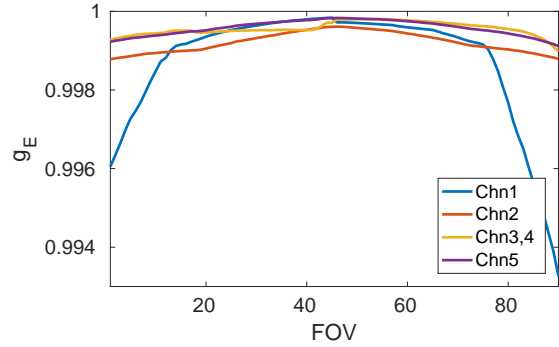
(b) MHS on NOAA-19
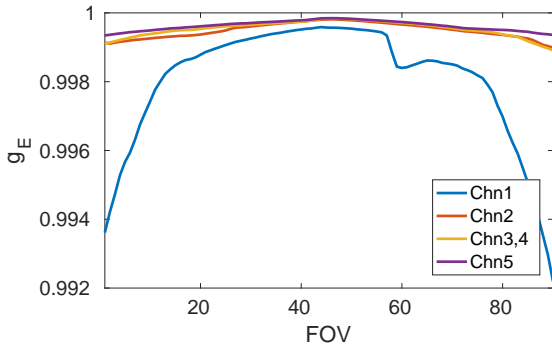
(c) MHS on Metop-A
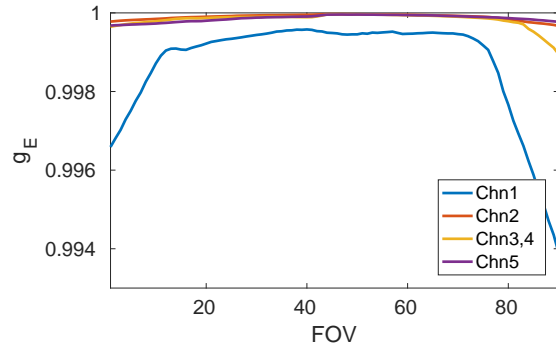
(d) MHS on Metop-B

Figure 5.4 Antenna efficiencies g_E for Earth views of the individual AMSU-B and MHS instruments. Channel assignment was wrong for all AMSU-Bs and the MHS on NOAA-18 (see paragraph on sensitivity coefficients and uncertainties below). Note that I corrected the assignment of values to channels already for this figure.

Within the calibration process, this antenna correction has to be applied. It is represented in the measurement diagram as

$$L_E = \frac{1}{g_{\mathrm{E}}}(L_{\mathrm{ME}} - g_{\mathrm{S}}L_{\mathrm{CMB}} - g_{\mathrm{Pl}}L_{\mathrm{Pl}}) \tag{5.24}$$

with L_{E} being the actual radiance of the Earth scene we are interested in, L_{ME} being the measured radiance (corresponding to the antenna temperature), L_{CMB} being the radiance of the cosmic microwave background (i.e. the true signal from deep space) and L_{Pl} being the radiance of the platform. In the actual performed calibration, it is assumed that $L_{\mathrm{Pl}} = L_{\mathrm{E}}$, i.e. $g_{E'} = g_{\mathrm{Pl}} + g_{\mathrm{E}}$. See Sec. 5.2.2 for further explanation. This transforms Eq. 5.24 to

$$L_E = \frac{1}{g_{E'}}(L_{\mathrm{ME}} - g_S L_{\mathrm{CMB}}) \tag{5.25}$$

Note that for SSMT-2, no antenna pattern correction is provided. The antenna efficiency for all FOVS of the Earth views is set to 1 in the header of the level 1b files.

Sensitivity coefficients and uncertainty

The sensitivity coefficients $c(i)$ (i varying for channel and pixel) associated with this effect are the partial derivatives of L_{E} with respect to the efficiencies $g(i)$. Note that the final measurement equation for L_{E} includes the term ΔL_{pol} which in turn depends on L_{E}:

$$c_{E',i} = \frac{\partial L_{\mathrm{E},i}}{\partial g_{E',i}} = -\frac{1}{g_{E'}^2}(L_{\mathrm{ME}} - g_S L_{\mathrm{CMB}})\left(1 - \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S))\right) \tag{5.26}$$

$$c_{S,i} = \frac{\partial L_{\mathrm{E},i}}{\partial g_{S,i}} = -\frac{L_{\mathrm{CMB}}}{g_{E'}}\left(1 - \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S))\right) \tag{5.27}$$

The uncertainty on the efficiencies g is not documented from pre-launch tests. Therefore I have to estimate the uncertainty based on the values given. The actual values given in the AAPP calibration files (`fdf.dat`) for the efficiencies vary for the different channels and instruments. However, it is not clear why some values are associated with certain channels or instruments: E.g. all AMSU-B instruments have the same g. Within this set however, channel 16 and channel 19, and channel 17 and channel 20 have the same efficiency values. The explanation for this confuse assignment is not documented and might be simply a coding error: the assignment does not make any sense, since it is the sounding channels 18-20 that share the same quasi-optical path (this is why channel 19 was measured as representative in Hewison and Saunders (1996) for the other sounding channels). Proceeding with the MHS instruments, MHS on NOAA-18 has again the same values for g as the AMSU-B instruments (see Fig. 5.4a). For Metop-A, there is an old set of values that is again the one from the AMSU-Bs and that is used per default value in the AAPP (Atkinson, 2016). Yet, a set of newer values has been provided by Jörg Ackermann (EUMETSAT) in 2007 (see comments in `fdf.dat` and Fig. 5.4c). This newer set is based on antenna pattern measurements from MHS for Metop-A and the geometrical map for the Metop platforms, see Fig. 5.3b (Klaes and Ackermann, 2014). The MHS on NOAA-19 and Metop-B also got their own values (Figures 5.4 (d) and (b)). From the observed spread of values among the channels, I deduced a rough estimate of uncertainty as follows.

The applied correction to the Earth views, i.e. $1 - g_{E'}$, is computed for all channels. To cover the variation between the channels, a deviation of 50% of the correction would have to be added/
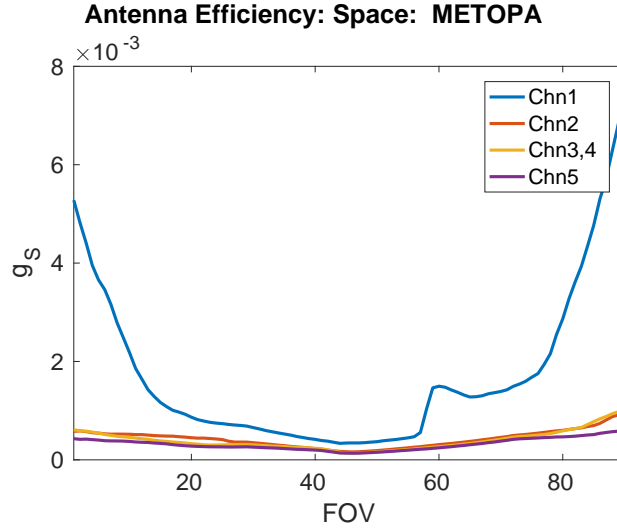
Figure 5.5 Contribution from space g_S in antenna efficiency of Earth views of MHS on Metop-A for each FOV. Note the unexpected increase of space contribution around FOV 60 in channel 1 (see text).

subtracted. Hence, as coarse estimate of uncertainty from the spread of values over the channels, I use 50% of the correction on Earth views. The uncertainty on the efficiency value for space are 50% of g_S. Hence, input uncertainties for the FCDR production are

$$u(g_{E'}) = 0.5 \cdot (1 - g_{E'}) \tag{5.28}$$
$$u(g_S) = 0.5 \cdot g_S \tag{5.29}$$

Another indicator for rather large uncertainty is an unexplained issue within the antenna efficiency: it is unclear why the contribution from space g_S of MHS on Metop-A (second version in `fdf.dat`) shows a distinct feature in channel 1 around FOV 60 (see Fig. 5.5). This feature should rather be expected in g_{Pl}, since the geometrical map shows a part of the platform reaching close to FOV 60 (azimuth around 25°, see Fig. 5.3b).

Hence, this rough and rather large estimation of 50% uncertainty on the correction reflects the miss-assignment of values to channels for AMSU-B, the unexplained re-usage of AMSU-B antenna efficiencies for MHS on NOAA-18 (and even Metop-A per default in AAPP; it can be changed manually) and unexplained issues within the efficiency values.

Since the efficiencies g are valid for the whole mission, the effect for the antenna pattern correction is counted among the common effects that have a correlation length scale larger than one orbit: e.g. pixel 90 will have the same uncertainty due to the antenna pattern correction for all orbits of the mission.

The effect of antenna pattern correction is fully correlated between channels 3 and 4. For AMSU-B, channel 5 is also fully correlated to 3 and 4. The cross-channel parameter correlation matrix for MHS is

$$\mathbf{R}_c^{p,g} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.30}$$

5.2.2 Radiation of the platform

Origin of the effect and its model in the measurement equation

As pointed out in Sec. 5.2.1 on the antenna pattern correction, the radiation of the platform contributes to the overall radiation detected within the side lobes of the antenna. This is due to the geometry of the spacecraft and the gain pattern of the antenna. To fully compensate this effect of contamination one would have to know the actual radiation of the platform. This is difficult to achieve: one would not only need a coarse map of the geometry of the spacecraft as in Fig. 5.3, but one would need knowledge on the exact geometry and optical characteristics of all surfaces from other instruments and the spacecraft itself. Since this information is not available, also the FIDUCEO processing, just as the operational AAPP processor, has to use the approximation presented already in the pre-launch tests in Hewison and Saunders (1996): it is assumed that the platform has a reflectivity of one, and furthermore, the platform only reflects the Earth radiation — even more, only the Earth radiation detected in the pixel that the correction should be applied to (this assumption is translated into the calibration procedure by considering the sum of the Earth and platform efficiencies $g_E + g_{Pl}$). As pointed out in Hewison and Saunders (1996), no reflection of sunlight is assumed. Moreover, the surroundings of the instruments on board the spacecraft are rather located in the near-field than in the far-field. The far-field assumption is used however, in order to derive the platform influence from the geometry and antenna pattern. This approximation is considered to be useful to estimate the influence of the platform (Hewison and Saunders, 1996).

The further simplification of $L_{Pl} = L_E$ is a strong one. Nonetheless, it is a useful approximation due to lack of detailed information. Also, the effect of changing the assumed radiation (or brightness temperature) from typical Earth temperatures to cosmic microwave background temperature, i.e. about 99% only results in a small change of final brightness temperature for a corrected Earth pixel: I carried out a short investigation (documented also in project report by Burgdorf et al. (2017)) by only using either the cosmic microwave background for the radiance of platform or the Earth radiance of the observed pixel, i.e. comparing the effect of both extreme cases. For MHS on Metop-B, channel 1 has the largest antenna pattern correction. The change in brightness temperature was -0.22 K at the very edge of the scan. For channel 3, which has a much smaller antenna pattern correction, the change was only -0.03 K (see Fig. 5.6).

This investigation shows that the effect of the radiance of the platform is weak for the sounding channels with small antenna pattern correction. For channel 1 it is more important. However, since the extreme cases of possible assumptions still only lead to a change of 0.08%, assuming an Earth temperature of 280 K for channel 1, it seems a fair solution to first stick to the AAPP assumption for the calibration. Of course, in the context of inter-satellite biases, this small effect should be kept in mind.

The radiance of the platform enters the measurement equation as part of the antenna pattern correction: see Eq. 5.24. Using the discussed approximation the equation turns into Eq. 5.25.

Sensitivity coefficients and uncertainty

In order to propagate the uncertainty due to this radiance of platform and its assumed value, the uncertainty propagation uses the sensitivity coefficient based on the full model of Eq. 5.24. Hence, the sensitivity coefficient for the radiance of the platform reads

$$c_{L_{Pl},i} = \frac{\partial L_{E,i}}{\partial L_{Pl}} = -\frac{g_{Pl,i}}{g_{E,i}} \left(1 - \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S)) \right). \qquad (5.31)$$

Channel1 Difference: Tb(Rad_Pl=Rad_Earth)−Tb(Rad_Pl=Rad_CMB)

(a) channel 1

Channel3 Difference: Tb(Rad_Pl=Rad_Earth)−Tb(Rad_Pl=Rad_CMB)

(b) channel 3
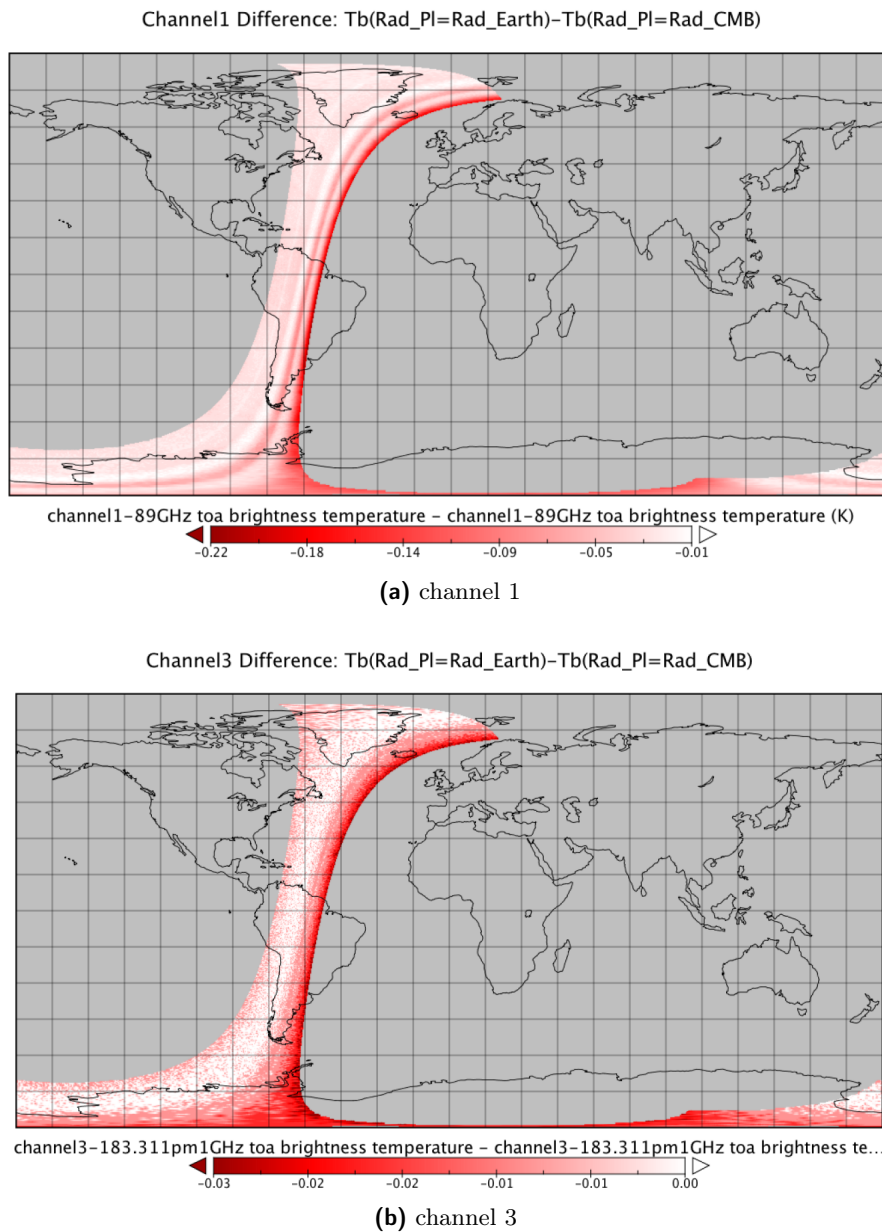
Figure 5.6 Difference in Earth view Brightness Temperature when using $L_{Pl} = L_E$ or $L_{Pl} = L_S$ for one example half-orbit of MHS on Metop-A.

The estimated uncertainty on the radiance of the platform, i.e. the assumed brightness temperature of the platform is 10% of the mean Earth brightness temperature, which corresponds to about 25 K.

The radiation of the platform is classified as common effect. Its uncertainty has no structure within a single orbit, but the assumption of $L_{Pl} = L_E$ is made for the whole mission.
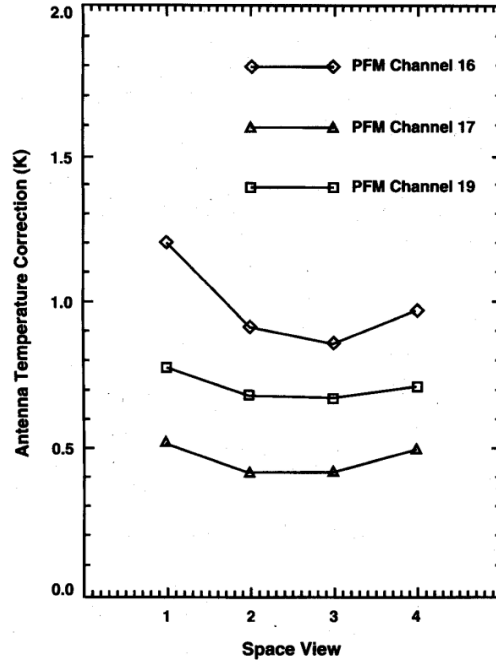
Figure 5.7 Cold bias correction for AMSU-B. From Hewison and Saunders (1996), ©1996, IEEE. Note the different naming of space view configurations starting at 1, compared to the text and (Atkinson and McLellan, 1998; Atkinson, 2000, 2002).

The effect of assumed radiance of platform is fully correlated between all channels. The cross-channel parameter correlation matrix is

$$\mathbf{R}_c^{p,L_{Pl}} = \begin{pmatrix} 1 & 1 & \dots \\ 1 & 1 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \qquad (5.32)$$

5.2.3 Cold space bias correction

Origin of the effect and its model in the measurement equation

This effect could be referred to as antenna pattern correction for the DSV, since it relates to the same effect presented in Sec. 5.2.1: Through the side lobes, radiation from platform and Earth limb enters the beam so that the overall measured signal from deep space view consists of the radiation from cosmic microwave background, from Earth limb and platform (Hewison and Saunders, 1996). There are four space view configurations SPV0 to SPV3 corresponding to different positions available for the instruments. For AMSU-B, SPV0 is the one closest to the Earth views and hence mostly influenced by Earth limb. SPV3 is closer to the z-axis and could therefore be more exposed to radiation from the platform (Atkinson, 2000, 2002). During pre-launch tests, a correction for the measured brightness temperature for each space view configuration has been determined in Hewison and Saunders (1996), based on the measured antenna pattern and the spacecraft geometry (see Fig. 5.7).

This correction ΔT_c accounts for the contamination through the side lobes. The optimal space view configuration was chosen during the in-orbit verification phase. For NOAA-15 and NOAA-

16, SPV3 was chosen (Atkinson and McLellan, 1998; Atkinson, 2000) due to the lowest count readings, whereas SPV2 was chosen for NOAA-17 (Atkinson, 2002).

The correction ΔT_c is given as value in Kelvin and is added to the temperature of the cosmic microwave background T_{CMB} to yield

$$T_S = T_{\mathrm{CMB}} + \Delta T_c. \tag{5.33}$$

This cold space bias correction enters the measurement equation through the radiance $\tilde{L}_{\mathrm{MS}}$ in the core equation for L_{ME} (blue box in the measurement diagram 5.1). $\tilde{L}_{\mathrm{MS}}$ represents the measured radiance in the DSV and hence incorporates the corrected signal as in Eq.5.33:

$$\tilde{L}_{\mathrm{MS}} = B(\nu, A_s + b_s(T_{\mathrm{CMB}} + \Delta T_c)) \tag{5.34}$$

where B is the Planck function and ν is the channel frequency. Note that the correction ΔT_c has been labelled as a_3 in the measurement equation diagram since it had been considered as possible harmonisation parameter (see Sec. 6.4.2, S harmonisation parameters are named a_1 to a_S). A_s and b_s are band correction factors (see Sec. 5.2.10).

Sensitivity coefficients and uncertainty

The sensitivity coefficient is the partial derivative of L_{E} with respect to ΔT_c.

$$
\begin{aligned}
c_{\Delta T_c} = \frac{\partial L_{\mathrm{E}}}{\partial \Delta T_c} = & -\frac{1}{g_{E\prime}^2} \left(1 - \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S)) \right) \\
& \cdot \left(-\frac{C_{\mathrm{E}} - \overline{C}_{\mathrm{IWCT}}}{\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}}} - 2q_{nl}\frac{(C_{\mathrm{E}} - \overline{C}_{\mathrm{S}})(C_{\mathrm{E}} - \overline{C}_{\mathrm{IWCT}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}})^2}(\tilde{L}_{\mathrm{MIWCT}} - \tilde{L}_{\mathrm{MS}}) \right) \\
& \cdot \frac{\partial B(\nu, A_s + b_s \cdot (T_{\mathrm{CMB}} + \Delta T_c))}{\partial \Delta T_c}
\end{aligned}
\tag{5.35}
$$

where B is again the Planck function of which we need the derivative with respect to temperature in this case (ΔT_c behaves as the temperature). A_s and b_s are band correction factors that I have determined for the deep space view (see Sec. 5.2.10).

The uncertainty on ΔT_c for the unharmonised case must be estimated from the provided values for ΔT_c in the `mhs_clparams.dat`/ `amsub_clparams.dat` files or in the in-orbit verification tests (Atkinson and McLellan, 1998; Atkinson, 2000, 2002) since no uncertainty is given. From the values given for the four space view configurations, I estimate the uncertainty as the standard deviation of the set of four values given for each channel, per instrument. Since MHS on NOAA-18, for unknown reasons, has the same ΔT_c for all channels and space view configurations according to the `mhs_clparams.dat` file, I use 100% of the given value as uncertainty estimate in this case. For SSMT-2, a "cold path temperature correction" is provided in the header of the level 1b files per channel. I use the standard deviation over the channels as uncertainty estimate. This effect of cold space bias correction is counted among the common effects, since the value for ΔT_c is valid for the whole mission.
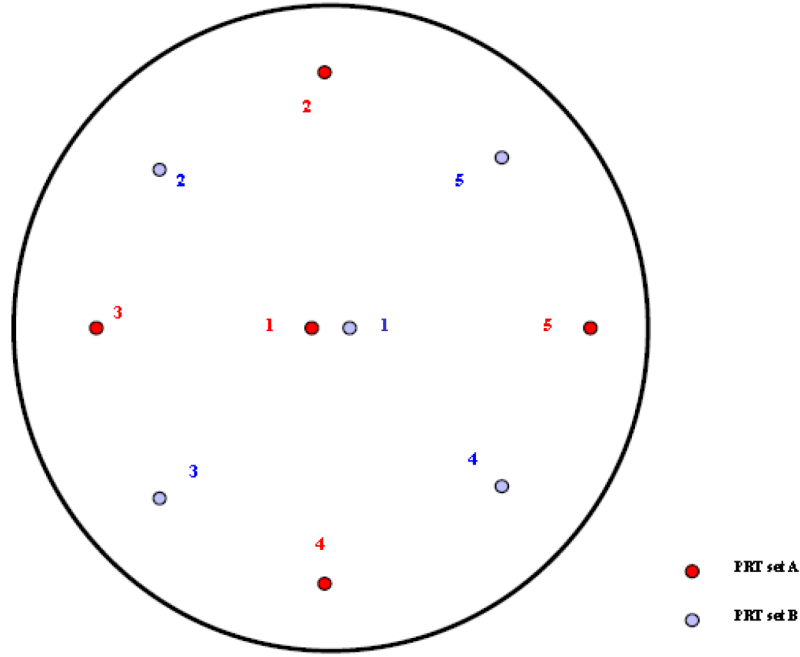
Figure 5.8 Arrangement of PRT sensors for MHS on the black body as seen from bottom of instrument. AMSU-B has a similar arrangement with two more PRT sensors. Figure from EUMETSAT (2013), with friendly permission by EUMETSAT. Note that PRT readings are only given for one set (A or B) within the level 1b file.

The effect of cold space bias correction is fully correlated between channels 3 and 4. For AMSU-B and SSMT-2, channel 5 is also fully correlated to 3 and 4. The cross-channel parameter correlation matrix for MHS is

$$\mathbf{R}_c^{p,\Delta T_c} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \tag{5.36}$$

5.2.4 Temperature measurement of the black body

Origin of the effect and its model in the measurement equation

The temperature of the IWCT (black body) is measured at its back by 7, 5 or 2 platinum resistance thermometers (PRT) for AMSU-B, MHS and SSMT-2 respectively. The PRTs for MHS and AMSU-B are arranged in a circle with one PRT placed in the centre, see Fig. 5.8. For the temperature measurement for one individual scan cycle, the PRTs are averaged. For all AMSU-B instruments, PRT 6 was always excluded since a small bias was observed against the others (Saunders et al., 1995). This is changed in my FCDR processor: Since this bias is small compared to the accuracy of $0.1\,\mathrm{K}$ of the PRTs, it does not seem appropriate to exclude PRT 6 for all cases. Hence, PRTs are only excluded based on the quality checks in the FCDR processor (see Sec. 5.3.2.2).

The averaged PRT-temperature, again averaged over 7 scan lines then enters the measurement equation as T_{IWCT} for the scan line to be calibrated (details on the average procedure are given in Sec. 5.3.2.4). This temperature is used to determine the radiance emitted by the IWCT:

$$\tilde{L}_{\text{MIWCT}} = B(\nu, A + b \cdot (T_{\text{IWCT}} + \delta T_{ch})) \tag{5.37}$$

Sensitivity coefficients and uncertainty

The IWCT temperature also enters the polarisation correction, apart from the core measurement equation. Hence, the sensitivity coefficient for the IWCT temperature is

$$
\begin{aligned}
c_{\delta T_{\text{IWCT}}} = \frac{\partial L_E}{\partial \delta T_{\text{IWCT}}} = &-\frac{1}{g_{E'}} \left(1 - \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S)) \right) \\
&\cdot \left(1 + \frac{C_E - \overline{C}_{\text{IWCT}}}{\overline{C}_{\text{IWCT}} - \overline{C}_S} + 2q_{nl}\frac{(C_E - \overline{C}_S)(C_E - \overline{C}_{\text{IWCT}})}{(\overline{C}_{\text{IWCT}} - \overline{C}_S)^2}(\tilde{L}_{\text{MIWCT}} - \tilde{L}_{\text{MS}}) \right) \\
&\cdot \frac{\partial B(\nu, A + b \cdot (T_{\text{IWCT}} + \delta T_{ch}))}{\partial \delta T_{ch}} \cdot b \\
&+ \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S))\frac{\partial B(\nu, A + b \cdot (T_{\text{IWCT}} + \delta T_{ch}))}{\partial \delta T_{ch}} \cdot b
\end{aligned} \tag{5.38}
$$

The uncertainty on the measurement of the black body temperature T_{IWCT} has two components: the systematic uncertainty on T_{IWCT} which relates to the accuracy of the PRTs, that is estimated as $0.1\,\text{K}$ (expert judgement from the National Physical Laboratory, NPL). The second component is noise on T_{IWCT} due to noise on the individual PRTs. This noise is calculated in the FCDR production as has been explained in Section 4.2. for the IWCT and DSV counts. Instead of "views", the individual measurements are now composed of individual PRTs. The overall computation using the Allan deviation is the same. One obtains quite small noise values of $0.002\,\text{K}$ (AMSU-B) to $0.02\,\text{K}$ (MHS).

The possible systematic error of T_{IWCT} is included in the common effects. The noise effect is part of the structured errors since it varies over one orbit, inheriting the 7-scan-line average structure.

The T_{IWCT} is the same for all channels. Therefore, the effect is fully correlated between the channels. The cross-channel parameter correlation matrix is

$$\mathbf{R}_c^{p,T_{\text{IWCT}}} = \begin{pmatrix} 1 & 1 & \cdots \\ 1 & 1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{5.39}$$

5.2.5 Warm target bias correction

Origin of the effect and its model in the measurement equation

As for the deep space view, also for the view of the IWCT, a correction term δT_{ch} is included in the overall temperature of the black body (Saunders et al., 1995). However, from the pre-launch test in Saunders et al. (1995) it was deduced that this δT_{ch} is probably zero for all channels of the investigated proto-flight model (i.e. AMSU-B on NOAA-15). Within the scope of these pre-launch tests, it was also investigated how the shroud of the Earth target that was used for testing, influenced the measured brightness temperature. This was carried out to investigate the

testing facilities and their impact on the actual tests on the instrument, but it is still interesting to know that a change of $5\,\mathrm{K}$ of the shroud compared to the Earth target resulted in a change of $0.2\,\mathrm{K}$ for all channels. Keeping the shroud at most $1\,\mathrm{K}$ off compared to the Earth target, was identified as sufficient to exclude influence from the shroud. For the black body of the instrument, the situation might be similar: The black body and its shroud are probably close to thermal equilibrium and hence, the effect of the shroud would be small. The influence of stray radiation from the further environment of the black body, meaning inner parts of the instruments that might have different temperatures, should also be small, since this radiation only enters through the side lobes of the beam. Further effects on the radiation of the IWCT to be included in δT_{ch} are horizontal and vertical temperature gradients. The vertical ones are negligible according to the study in Hewison (1991). Horizontal gradients may occur and investigation on that was carried out within the FIDUCEO project: gradients of about $0.2\,\mathrm{K}$ were seen (Burgdorf et al., 2017) when comparing the PRT measurements to each other for certain time periods of NOAA-16. Given the systematic uncertainty on the PRT sensors of $0.1\,\mathrm{K}$, the effect is not large though. Any reflection on the black body of radiation of surrounding elements is extremely small since the emissivity of the black body is very close to one: values of 0.9999 (Hewison, 1991; Saunders et al., 1995) for the IWCT of AMSU-B are observed.

Given these possible effects on the radiation of the IWCT, one can understand δT_{ch} as an approximate combination of all these effects. As pointed out in Saunders et al. (1995), this correction term might be used for in-orbit calibration adjustments. This is picked up in the harmonisation process, where δT_{ch} will enter.

So far, the values used in the operational calibration for δT_{ch} are zero for all MHS and AMSU-B instruments and channels (as indicated in Saunders et al. (1995)), except for AMSU-B on NOAA-17: channel 20 has a non-zero, negative value (see `clparams.dat` files). The reason is not clear, since the frequency does not differ much from the other sounding channels, hence the sounding channels should all get the same correction. Different values, hence the possible channel dependence indicated by index ch, would only be expected for channels 16 and 17 due to frequency dependent emission from the surroundings of the black body. For the SSMT-2 instruments, all instruments and channels are provided with a non-zero, positive "warm path temperature correction" in the file header.

The correction enters the measurement equation as additive term to the measured temperature of the black body T_{IWCT}. As such, it appears in the radiance of the IWCT:

$$\tilde{L}_{\mathrm{MIWCT}} = B(\nu, A + b \cdot (T_{\mathrm{IWCT}} + \delta T_{ch})) \tag{5.40}$$

B is the Planck function and A and b are the band correction factors, accounting for the breakdown of the monochromatic assumption for the outer sounding channels (see Sec. 5.2.10).

Sensitivity coefficients and uncertainty

The sensitivity coefficient for δT_{ch} is obtained as partial derivative from the measurement equation. Note that the radiance of the black body enters not only the core measurement equation and the non-linearity term but also the polarisation correction term. The coefficient is

$$
\begin{aligned}
c_{\delta T_{ch}} = \frac{\partial L_E}{\partial \delta T_{ch}} = -\frac{1}{g_{E\prime}} & \left(1 - \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S)) \right) \\
& \cdot \left(1 + \frac{C_{\mathrm{E}} - \overline{C}_{\mathrm{IWCT}}}{\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}}} + 2q_{nl}\frac{(C_{\mathrm{E}} - \overline{C}_{\mathrm{S}})(C_{\mathrm{E}} - \overline{C}_{\mathrm{IWCT}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}})^2}(\tilde{L}_{\mathrm{MIWCT}} - \tilde{L}_{\mathrm{MS}}) \right) \\
& \cdot \frac{\partial B(\nu, A + b \cdot (T_{\mathrm{IWCT}} + \delta T_{ch}))}{\partial \delta T_{ch}} \cdot b \\
& + \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S))\frac{\partial B(\nu, A + b \cdot (T_{\mathrm{IWCT}} + \delta T_{ch}))}{\partial \delta T_{ch}} \cdot b
\end{aligned}
\tag{5.41}
$$

As first estimate of the uncertainty, I use 100% of the value given for channel 20 on NOAA-17, which corresponds to 0.16 K. Given that some gradients of 0.2 K were observed (at least from what the PRT sensors yield), and that there is no justification for channel 20 on NOAA-17 to be different from the others, this assumed uncertainty seems an acceptable choice. For SSMT-2, I use the standard deviation over the five channels as uncertainty estimate for δT_{ch}.

The warm target bias correction as it is implemented in the measurement equation is a common effect, since it introduces errors on larger scales than one orbit. Its values are valid for the whole mission.

The underlying effect of distribution of physical temperature is the same for all channels. Hence, the effect is assumed as fully correlated between the channels. The cross-channel parameter correlation matrix is

$$
\mathbf{R}_c^{p,\delta T_{ch}} = \begin{pmatrix} 1 & 1 & \cdots \\ 1 & 1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix}
\tag{5.42}
$$

5.2.6 Non-linearity

Origin of the effect and its model in the measurement equation

The calibration of the microwave radiometers is based on a two-point linear interpolation. However, the instrument is not necessarily perfectly linear in its response and translation of brightness temperature. Pre-launch tests on the AMSU-B instruments verified the linear response of the receiver (Saunders et al., 1995) over the tested temperature range of 85 - 330 K. Also, the deviations in the AMSU-B measured brightness temperature and the actual reference target temperature in the testing facilities were determined. Looking at the departure from a straight line fit of the brightness temperature as function of the target temperature, the deviation from linearity was observed over the tested temperature range. No strong non-linearity could be observed (all below 0.1 K). Only for channel 16, a second order fit was suggested to be appropriate. Ground tests for requirements checks by Matra Marconi also showed no distinct non-linearity in shape of a second order polynomial (Khan and Shaw, 1999). For the operational processor AAPP, non-linearity coefficients (for a second-order term in C_E) are provided nonetheless as look-up table in the `mhs_clparams.dat` and `amsub_clparams.dat` files. For all channels, non-linearity

coefficients are given for three reference instrument temperatures. The non-linearity coefficient to use for a certain scan line is then obtained from linear interpolation between the three temperatures for the current instrument temperature (which the temperature of the local oscillator for channel 5 is used as proxy for (EUMETSAT, 2013)). For AMSU-B, channel 18-20 have zero as non-linearity coefficients. For MHS, all channels are provided with non-zero non-linearity coefficients. Channel 1 or 16, respectively, has the strongest non-linearity coefficient. Interestingly, almost all coefficients are negative. Only for MHS on NOAA-19, all coefficients for every channel and reference temperature are positive. It is not clear where these values originate from. Probably, they were determined in pre-launch test, but it is not documented in AAPP. Note that SSMT-2 has no non-linearity correction.

In the FIDUCEO measurement equation, we include the non-linearity correction as a second order term in the counts. This correction enters the core measurement equation as

$$\Delta L_{\mathrm{nl}} = q_{nl} \frac{(C_{\mathrm{E}} - \overline{C_{\mathrm{S}}})(C_{\mathrm{E}} - \overline{C}_{\mathrm{IWCT}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}})^2}(\tilde{L}_{\mathrm{MIWCT}} - \tilde{L}_{\mathrm{MS}}) \tag{5.43}$$

The coefficient q_{nl} is the non-linearity coefficient. It is counted among the possible parameters to be harmonised in order to obtain a better estimation, hence it is named a_1 in the measurement diagram.

Sensitivity coefficients and uncertainty

The partial derivative of L_E with respect to the coefficient q_{nl} yields the sensitivity coefficient for the effect of non-linearity:

$$\begin{aligned} c_{\delta q_{nl}} = \frac{\partial L_E}{\partial \delta q_{nl}} = &-\frac{1}{g_{E'}}\left(1 - \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S))\right) \\ &\cdot \frac{(C_{\mathrm{E}} - \overline{C_{\mathrm{S}}})(C_{\mathrm{E}} - \overline{C}_{\mathrm{IWCT}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}})^2}(\tilde{L}_{\mathrm{MIWCT}} - \tilde{L}_{\mathrm{MS}}) \end{aligned} \tag{5.44}$$

Because of the unknown origin and uncertainty of the non-linearity coefficients for the reference temperatures, I use an uncertainty of 100% of the values for q_{nl} obtained from interpolation. Within this uncertainty estimate, the true uncertainty on the given reference values, as well as the uncertainty on the instrument temperature and the interpolation are included.

The effect of non-linearity is counted among the common effects, since its correlation length scale is larger than one orbit (the underlying non-linearity coefficients are assumed to be constant even over the whole mission).

Each channel showed a different (very small) non-linearity in the pre-launch tests by Matra Marconi (Khan and Shaw, 1999). The effect is assumed to be uncorrelated between the channels. Therefore, the cross-channel parameter correlation matrix is the identity matrix

$$\mathbf{R}_c^{p,q_{nl}} = \mathbf{I}_{5x5} \tag{5.45}$$

5.2.7 The polarisation correction

Origin of the effect and its model in the measurement equation

A polarisation correction depending on scan angle and Earth radiance has been devised in order to account for varying reflectivity of the antenna with respect to different polarisation directions

and for the defined polarisation of the individual channels (Labrot et al., 2011). The suggested form is

$$L_{\text{pol}} = \frac{1}{2}\alpha(\tilde{L}_{\text{MIWCT}} - L_E)(\cos(2\theta_E) - \cos(2\theta_S)) \tag{5.46}$$

with α being a channel dependent coefficient representing $1 - R_\parallel/R_\perp$, where $R_\parallel$ and $R_\perp$ denote the reflectivity for electric fields parallel and perpendicular to the plane of incidence, respectively. The angles θ_E and θ_S are the scan angles for Earth view and space view respectively, measured from nadir at $0°$. In this form, the correction enters the measurement equation. Note that in principle an iteration would be necessary, since the correction itself depends on L_E. Since it is a small correction anyway, we follow the AAPP strategy and only execute one step of the iteration by using the uncorrected L_E in Eq. 5.46 to determine the corrected L_E. This approximation is indicated in the "+0" term in the measurement diagram.

For the AMSU-B instruments, the α values are all zero in the AAPP calibration files, although a scan-dependent cold bias was observed pre-launch and a correction scheme very similar to Eq. 5.46 was developed in Saunders et al. (1995). The best fit obtained as correction was

$$\Delta L = \gamma \frac{(\tilde{L}_{\text{MIWCT}} - L_E)}{(\tilde{L}_{\text{MIWCT}} - L_S)} \cdot \sin(75° - |\theta_E|) \tag{5.47}$$

I investigated the similarity between Eq. 5.46 and 5.47 by comparing the trigonometric terms

$$\sin(75° - |\theta_E|) \qquad \text{and} \qquad \frac{1}{2}(\cos(2\theta_E) - \cos(2\theta_S)) \tag{5.48}$$

in both correction schemes ($75°$ is the approximate position of the DSV). The remainder of the formulae is equal, or close to constant and might be absorbed in γ. Figure 5.9 shows the difference of the trigonometric expressions as function of θ_E, for two values of θ_S ($75°$ and $69°$), representing different DSV configurations. For the range of θ_E that is relevant for the calibration ($|\theta_E| \in [0, 50°]$), both curves are very close to zero. Hence, one can deduce that the trigonometric expressions should correct the same. That means, the cold biases seen in Saunders et al. (1995) were probably due to uncorrected polarisation effects.

Unfortunately, no coefficients for the $\gamma(\nu)$ from Saunders et al. (1995) are reported. For the NOAA-18 MHS, non-zero α are provided in the AAPP files, also the old Metop-A version has non-zero values since it is simply the same as the NOAA-18 version (default value for AAPP is to use the old Metop-A version, it can be switched by the user (Atkinson, 2016)). It is not clear, however, how these values have been determined. Also, the given values for α vary considerably between the channels from 0.0002 in channel 1 to -0.0022 in channels 3 and 4 (the negative sign accounts for the different polarisation direction). This large spread of values is suspicious, since the reflectivity of the mirror, which α is related to, is not likely to vary significantly over the frequency range covered by the channels. Moreover, since AMSU-B showed a scan-dependent bias pre-launch, it is unclear why it only has zero α. Hence, it is desirable to obtain optimised values, for all instruments. Therefore, the coefficient α is a possible candidate for harmonisation and is hence labelled as a_2 in the measurement equation diagram.

Note that no polarisation correction at all is provided for SSMT-2. Hence, α is assumed to be zero.
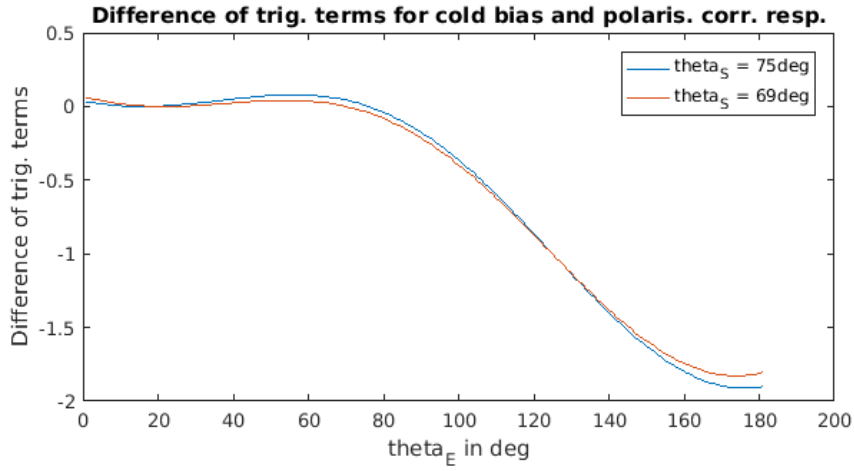
(a) for angles from 0 to 180°
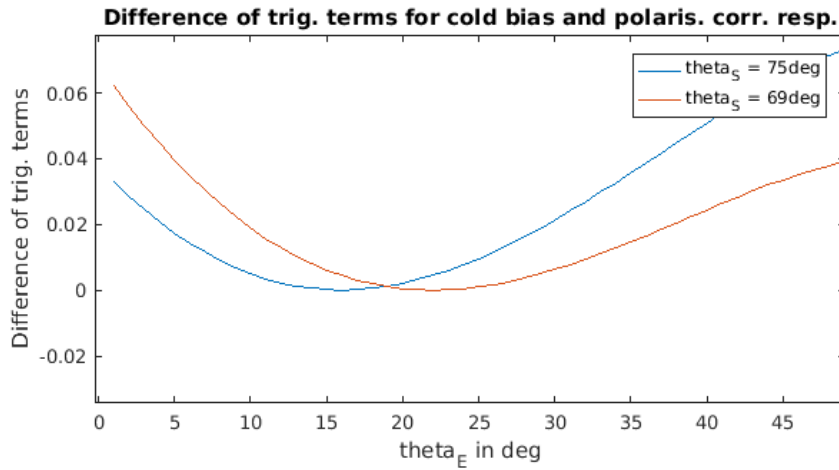
(b) zoom-in of (a) to relevant range for θ_E

Figure 5.9 The difference of the trigonometric expressions in Eq. 5.46 and 5.47 as function of θ_E. See text for further explanation.

Sensitivity coefficients and uncertainty

The sensitivity coefficient for α is the partial derivative

$$c_\alpha = \frac{\partial L_{\mathrm{E}}}{\partial \alpha} = \frac{1}{2}(\tilde{L}_{\mathrm{MIWCT}} - L_{\mathrm{E}})(cos(2\theta_E) - cos(2\theta_S)) \tag{5.49}$$

As uncertainty estimate I use 100% of the values provided for MHS on NOAA-18, since no indication of uncertainty or origin of the values is found. This polarisation correction is included in the common effects.

Since the coefficient α relates to the reflectivity of the mirror that should not vary significantly within the covered frequency range, I assume as first approximation that the effect is fully correlated between the channels. Hence, the cross-channel parameter correlation matrix is

$$\mathbf{R}_c^{p,\alpha} = \begin{pmatrix} 1 & 1 & \cdots \\ 1 & 1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{5.50}$$

Further quantities in the polarisation correction that may cause small errors are the angles for Earth and space view. Their sensitivity coefficients are

$$c_{\theta_E} = \frac{\partial L_{\mathrm{E}}}{\partial \theta_E} = -\alpha(\tilde{L}_{\mathrm{MIWCT}} - L_{\mathrm{E}})\sin(2\theta_E) \tag{5.51}$$

$$c_{\theta_S} = \frac{\partial L_{\mathrm{E}}}{\partial \theta_E} = \alpha(\tilde{L}_{\mathrm{MIWCT}} - L_{\mathrm{E}})\sin(2\theta_S) \tag{5.52}$$

The angles have two uncertainty components: There is a random variation in the position (angle) that can be seen for a single FOV over several scan lines. I used the standard deviation of these position measurements over one orbit to estimate the uncertainty due to these random variations for the Earth (0.04°) and space view angles (0.02°). On top of that, there is a systematic uncertainty on these angles: It has been observed that there is a systematic deviation for the measured pointing from the nominal pointing in the range of 0.1° for some of the instruments, channels and views (Ackermann, 2017). Therefore, I assume a systematic uncertainty of 0.1° for Earth and space views, accounting for the slight miss-alignments.

For SSMT-2, no separate information on the angles is provided. Hence, the nominal positions are used in the FCDR code. As rough estimate on random uncertainty, I use the provided precision of 0.5°. For the systematic uncertainty, the estimate for MHS is used. Note that these uncertainties only enter, since I use 100% of the NOAA-18 MHS value as uncertainty estimate for α of SSMT-2 as well.

The random variations of the angles for the Earth views are counted among the independent effects. Spectrally, these variations are not considered to be independent, but fully correlated as the variations probably result from small changes in the positioning of the rotating reflector by the motor. This is equal for all channels. The cross-channel parameter error correlation matrix is therefore estimated as

$$\mathbf{R}_c^{p,\theta_E} = \begin{pmatrix} 1 & 1 & \cdots \\ 1 & 1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{5.53}$$

The random variations of the angles for the space view are counted among the structured effects (since they affect all FOVs in one line equally). As for the Earth views, the cross-channel correlation matrix is

$$\mathbf{R}_c^{p,\theta_S} = \begin{pmatrix} 1 & 1 & \cdots \\ 1 & 1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{5.54}$$

The systematic uncertainties in the angles for Earth and space views are counted among the common effects.

5.2.8 Noise effects on Earth views, space views and IWCT views

Origin of the effect and its model in the measurement equation

The noise on DSV, IWCT and Earth counts has been discussed in detail in Section 4.2. The computation of the noise on counts by the Allan deviation for each view is part of the FCDR processor. Also, the 7-scan line average procedure is taken into account for the DSV and IWCT noise as estimate for the final uncertainty on the respective counts that enters the uncertainty propagation. The noise on Earth counts is estimated by rescaling the IWCT and DSV counts noise (without 7-scan line average) from the IWCT and DSV temperature to the Earth temperature of the pixel (described in 4.2).

The counts for IWCT, DSV and Earth enter the core measurement equation as well as the non-linearity correction term:

$$L_{\mathrm{E}} = L_{\mathrm{IWCT}} + \frac{L_{\mathrm{IWCT}} - L_{\mathrm{S}}}{\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}}} \cdot (C_{\mathrm{E}} - \overline{C}_{\mathrm{IWCT}}) + \Delta L_{\mathrm{nl}} \qquad (5.55)$$

with

$$\Delta L_{\mathrm{nl}} = q_{nl} \frac{(C_{\mathrm{E}} - \overline{C}_{\mathrm{S}})(C_{\mathrm{E}} - \overline{C}_{\mathrm{IWCT}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}})^2} (\tilde{L}_{\mathrm{MIWCT}} - \tilde{L}_{\mathrm{MS}}). \qquad (5.56)$$

Sensitivity coefficients and uncertainty

The sensitivity coefficients read for the counts on DSV

$$
\begin{aligned}
c_{\overline{C}_{\mathrm{DSV}}} = \frac{\partial L_{\mathrm{E}}}{\partial \overline{C}_{\mathrm{DSV}}} = & -\frac{1}{g_{E'}} \left(1 - \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S)) \right) \\
& \cdot \frac{(C_{\mathrm{E}} - \overline{C}_{\mathrm{IWCT}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}})^2} (\tilde{L}_{\mathrm{MIWCT}} - \tilde{L}_{\mathrm{MS}}) \\
& \cdot \left[1 + q_{nl} \frac{(\tilde{L}_{\mathrm{MIWCT}} - \tilde{L}_{\mathrm{MS}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}})} \cdot (-2C_{\mathrm{E}} + \overline{C}_{\mathrm{IWCT}} + \overline{C}_{\mathrm{S}}) \right]
\end{aligned} \qquad (5.57)
$$

on IWCT view

$$
\begin{aligned}
c_{\overline{C}_{\mathrm{IWCT}}} = \frac{\partial L_{\mathrm{E}}}{\partial \overline{C}_{\mathrm{IWCT}}} = & \frac{1}{g_{E'}} \left(1 - \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S)) \right) \\
& \cdot \frac{(C_{\mathrm{E}} - \overline{C}_{\mathrm{S}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}})^2} (\tilde{L}_{\mathrm{MIWCT}} - \tilde{L}_{\mathrm{MS}}) \\
& \cdot \left[-1 + q_{nl} \frac{(\tilde{L}_{\mathrm{MIWCT}} - \tilde{L}_{\mathrm{MS}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}})} \cdot (-2C_{\mathrm{E}} + \overline{C}_{\mathrm{IWCT}} + \overline{C}_{\mathrm{S}}) \right]
\end{aligned} \qquad (5.58)
$$

and on Earth view

$$
\begin{aligned}
c_{\overline{C}_{\mathrm{E}}} = \frac{\partial L_{\mathrm{E}}}{\partial \overline{C}_{\mathrm{E}}} = & \frac{1}{g_{E'}} \left(1 - \frac{1}{2}\alpha(\cos(2\theta_E) - \cos(2\theta_S)) \right) \\
& \cdot \frac{(\tilde{L}_{\mathrm{MIWCT}} - \tilde{L}_{\mathrm{MS}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}})} \\
& \cdot \left[1 - q_{nl} \frac{(\tilde{L}_{\mathrm{MIWCT}} - \tilde{L}_{\mathrm{MS}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_{\mathrm{S}})} \cdot (-2C_{\mathrm{E}} + \overline{C}_{\mathrm{IWCT}} + \overline{C}_{\mathrm{S}}) \right]
\end{aligned} \qquad (5.59)
$$

The corresponding uncertainties are obtained from the on-the-fly computation of the noise by the evaluation of the Allan deviation during the FCDR processing.

The DSV and IWCT count noise is part of the structured effects since the DSV and IWCT calibration counts enter the 7-scan line average procedure and therefore exhibit a correlation structure on smaller length scales than one orbit. The noise on Earth counts belongs to the independent effects. All three noise components are uncorrelated between the channels. I investigated this by computing the noise over certain months and calculating the cross-channel correlation. In Fig. 5.10, the spectral correlation matrix is shown exemplarily for the DSV count noise for one month of MHS on NOAA-18. The used correlation coefficients are those from Pearson and Spearman, measuring a linear relation and the ranking, respectively. No strong off-diagonal elements are present for this month. Other randomly picked months gave a similar result. Hence, the cross-channel parameter correlation matrices are the 5x5 identity matrix

$$\mathbf{R}_c^{p,\overline{C}_{\mathrm{DSV}}} = \mathbf{R}_c^{p,\overline{C}_{\mathrm{IWCT}}} = \mathbf{R}_c^{p,C_{\mathrm{E}}} = \mathbf{I}_{5x5}. \tag{5.60}$$

### 5.2.9 Radio frequency interference

Origin of the effect and its model in the measurement equation

When the first calibrated AMSU-B data were available after launch of NOAA-15, strong scan-dependent biases that reached up to 40 K were observed for four of five channels (Atkinson and McLellan, 1998; Atkinson, 2001). The suspicion arose that radio frequency interference (RFI) from transmitters was the reason. AMSU-B (and afterwards MHS, too) may be influenced by six transmitters on board the platform, of which some are active at all times, some are not used and some are used for data dumping to ground stations and hence only active at some times. The STX-1 transmitter used for High Resolution Picture Transmission (HRPT) and the SARR (Search and Rescue Repeater) were first identified to cause the largest bias for AMSU-B on NOAA-15. The detailed characterisation of which transmitter causes what bias for which channels was obtained by dedicated tests: The transmitters were switched on an off in turn over subsequent orbits to quantify the biases. As the biases appeared to be stable, a correction scheme was devised (Atkinson and McLellan, 1998; Atkinson, 2001). This correction scheme was implemented in the level 1b data header, so that the user could correct for the biases (biases on DSV and IWCT were already corrected within the level 1b data). However, the biases did not remain stable, as new bias modes occurred (Atkinson, 2001). Again, new correction schemes were defined for the five channels being affected as
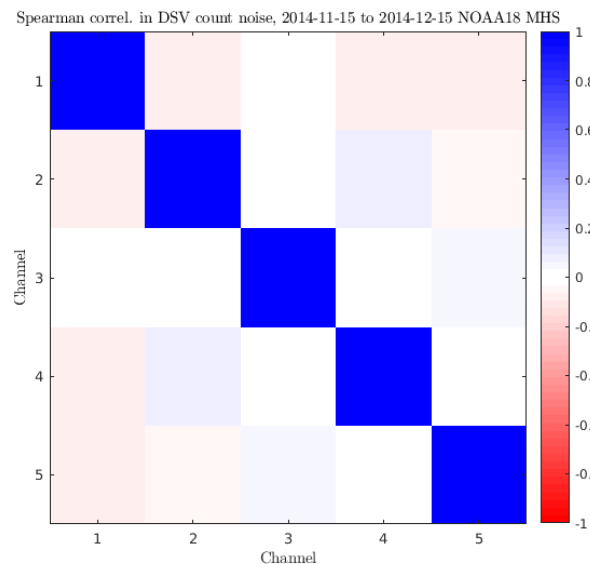
- channel 16: < 1 K

- channel 17: < 15 K

- channel 18: < 2 K

- channel 19: about 35 K

- channel 20: about 8 K.

The effect of RFI is particularly difficult to characterise and handle: it is only possible by tests such as those that were carried out in the early life-time of NOAA-15. Only these tests can give
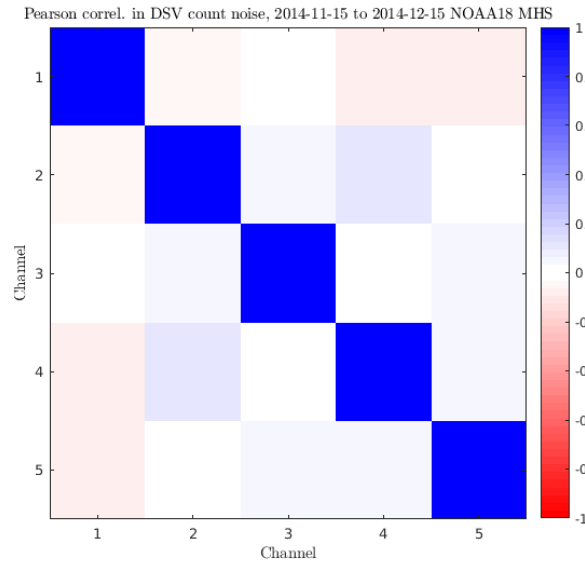
(a) Pearson correlation coefficient

(b) Spearman correlation coefficient

Figure 5.10 Cross-channel correlation coefficients computed for the DSV noise for one month.

the actual biases or even prove the influence of RFI. However, since RFI has shown to be highly variable over time, regular tests would be necessary to keep track of the required bias correction schemes. The corrections supplied in the level 1b header were not updated after 2001 (Robel et al., 2009; Atkinson, 2018). The consequence is that AMSU-B NOAA-15 data cannot be fully corrected and hence, cannot be used for climate monitoring. The problem of RFI causing (scan-dependent) biases increases with time for NOAA-15, since the gain decreases strongly and the radiometer sensitivity degrades accordingly (see Chapter 4) and therefore, the relative impact of RFI increases. This is corroborated by the findings of pre-launch tests on the remaining AMSU-B instruments after the NOAA-15 launch: it was found that the radiation interfering with the

Earth signal enters the receiver by the main reflector. Since the main reflector is rotating to cover all scan angles, the scan dependence of RFI bias is introduced. The pickup occurred only in the video amplifier and detector (Atkinson, 2001). Hence, the actual impact of RFI happens at the back-end of the receiver, whereas gain reduction is a front-end effect (John et al., 2013b). Therefore, a reduced gain produces an increased impact of RFI. The same decrease in gain as for NOAA-15 happened for channels 18-20 of AMSU-B on NOAA-16 (see Chapter 4). Hence, it might be possible that RFI also influences this instrument in its degraded gain status and makes its data unusable for climate studies, although the shielding of the instrument to prevent RFI had been improved (Atkinson, 2001), and no RFI was observed during the in-orbit verification test phase for the channels 18-20. However, a weak RFI impact was observed in the DSV and IWCT view for channel 17 (Atkinson, 2000), indicating that NOAA-16 is not completely free of this RFI effect. Hence, also channels 18-20 could have been affected already by RFI influence, which was too weak to be detected during the in-orbit verification phase several years before the strong gain degradation. The possible increasing impact of RFI on NOAA-16 was also suspected in John et al. (2013b), but no consequences were formulated. For the NOAA-17 AMSU-B, some RFI related biases were observed in channel 19 (Atkinson, 2002) and a correction scheme was devised and stored in the level 1b data header. To which extent this correction scheme is trustworthy at later times remains unclear. Because of the experienced non-predictability of RFI on NOAA-15, it cannot be ruled out that RFI on NOAA-17 also changed.

For the MHS instruments, RFI was not considered as an issue (all bias correction fields in the level 1b data are zero filled). Nonetheless, for degraded gain periods as seen for channel 3 and 4 of NOAA-19 MHS after highly problematic periods (see Chapter 4), it may be possible that RFI affects the signal to a significant amount. Since no dedicated in-orbit tests are possible from a retrospective perspective, it is difficult to shed light on this issue. Investigating biases against other instruments, by using only data where certain transmitters were switched on and off, does not fully enlighten the problem. With the monthly-percentiles method, I investigate the data of MHS on NOAA-19 for September 2010 to February 2011. As the bias against NOAA-18 was stable (see Sec. 3.1.2.2, 3.1.4), a short period is sufficient for a first impression). Allowing all transmitters to be present in the data shows the known strong bias against NOAA-18. Allowing no transmitters to be switched on leaves almost no data (reduction to 0.03%). These extreme cases do not help to characterise the influence of RFI. However, also allowing for certain transmitters to be switched on, does not yield a clear result either. Only, since the SARR and STX-1 transmitters are active at all times (except for very few instances of a few scan lines), it cannot be ruled out that they cause the bias.

In Chapter 6, Sec. 6.3 on Harmonisation, I will revisit this RFI issue in the context of understanding and reducing inter-satellite biases.

As explained above, the correction schemes devised for the NOAA-15 and NOAA-17 AMSU-B instrument, might not be fully reliable. Therefore, they are not applied within the FCDR processor. The production of FCDR automatically excludes the periods of degradation using quality checks.

The correction schemes work on the level of counts, i.e. RFI is no effect that would get an extra term in the measurement equation, but it already distorts the very input signal, namely the Earth view, DSV and IWCT counts.

Sensitivity coefficients and uncertainty

As RFI directly affects the counts, the sensitivity coefficients for the effect of RFI equal those of the DSV, IWCT and Earth view counts (see Eq. 5.57, 5.58 and 5.59).

To include an uncertainty on the brightness temperature due to RFI, the following conservative uncertainty estimate on the counts is made: the observed biases in the verification phase are used as uncertainty estimate (where they are only given in temperature, I translate them back into counts by using the gain of the according period). This uncertainty, propagated through the sensitivity coefficients to the uncertainty in brightness temperature, is included in the uncertainty budget as soon as any transmitter is switched on. For the other instruments than those on NOAA-15 and NOAA-17, no biases are documented and hence, no uncertainty related to RFI can be assigned or estimated, since RFI could cause any bias. The information on transmitter status, however, is provided in the FCDR for all instruments, except for SSMT-2 where no such information is available.

The RFI is part of the common effects. The cross-channel parameter correlation matrix is diagonal, since the effect of RFI is completely uncorrelated between the channels. All channels can be affected in a different way. Hence, we have the identity matrix

$$\mathbf{R}_c^{p,\text{RFI}} = \mathbf{I}_{5x5} \tag{5.61}$$

5.2.10 Band correction factors A and b

Origin of the effect and its model in the measurement equation

The five channels for SSMT-2, AMSU-B and MHS all measure radiance at two side bands around the central frequency of the channel. For the sounding channels at 183±3 and 183±7 GHz, the side bands are placed quite far away from the central frequency. For these channels, the monochromatic assumption for using only the central frequency in the Planck function breaks down and band corrections need to be applied (Saunders et al., 1995). Hence, band correction coefficients are applied to the temperature of the IWCT when it is converted to radiance $\tilde{L}_{\text{MIWCT}}$. These coefficients are designed such as to model the effect that the instrument actually measures an averaged radiance from both pass bands.

In the measurement equation, the coefficients A and b are applied to the corrected temperature of the IWCT:

$$\begin{aligned}
\tilde{L}_{\text{MIWCT}} &= B(\nu_{ch}, A + b \cdot (T_{\text{IWCT}} + \delta T_{ch})) \\
&= \frac{2h\nu_{ch}^3}{c^2} \frac{1}{\exp\left[\dfrac{h\nu_{ch}}{(k_b[A + b \cdot (T_{\text{IWCT}} + \delta T_{ch})])}\right] - 1}
\end{aligned} \tag{5.62}$$

The values for A and b are listed in the AAPP `mhs-` and `amsub_clparams.dat` files. They are also given in the pre-launch report in Saunders et al. (1995).

Improving on the operational values

The values provided in the calibration files are useful, but they can be easily improved (if there is still potential for improving) by reducing the deviation of calculating the Planck function at the central frequency and the corrected temperature compared to the mean of the Planck function

evaluated at the upper and lower side band frequency and the uncorrected temperature. This can be expressed as minimising the expression $|Q - 1|$ with the quotient Q being

$$Q = \frac{B(183.31\,\text{GHz}, A + b \cdot 280\,\text{K})}{0.5 \cdot [B(183.31 - 3\,\text{GHz}, 280\,\text{K}) + B(183.31 + 3\,\text{GHz}, 280\,\text{K})]} \tag{5.63}$$

for the channel at 183±3 GHz. For the channel at 183±7 GHz, the quotient is adapted accordingly. I chose the temperature of 280 K as typical temperature for the IWCT. By minimising the expression $|Q - 1|$, I obtain new band correction coefficients:

$$A = 0.0015\,\text{K} \qquad b = 1.00025 \tag{5.64}$$

and for the channel at 183±7 GHz

$$A = 0.00289\,\text{K} \qquad b = 1.00138. \tag{5.65}$$

Compared to the old coefficients ($A = -0.0031\,\text{K}, b = 1.00027$ and $A = -0.0167\,\text{K}, b = 1.00145$ for channels at 183±3 GHz and 183±7 GHz, respectively), I reduced the difference of Q to 1 from $3.61690 \cdot 10^{-6}$ to now $9.99 \cdot 10^{-9}$ for the channel at 183±3 GHz and from $2.62 \cdot 10^{-6}$ to now $1.01 \cdot 10^{-8}$ for the channel at 183±7 GHz. I.e. a small improvement could still be achieved.

No such band correction coefficients are given for the DSV and its temperature, although the same argument of break down of monochromatic assumption holds there. In the EUMETSAT product generation specification for MHS (EUMETSAT, 2013), it is written that the temperature of the DSV needs to be band corrected. However, it is referred to the correction for the IWCT, as if indicating to use the same coefficients. This is not the correct way: The Planck curve has a different shape for the temperatures occurring in DSV compared to IWCT temperatures. Hence, one cannot apply the same coefficients. To overcome this discrepancy, I have determined A_s and b_s to be the band correction coefficients for the DSV. For this, I find the coefficients A_s and b_s that minimise $|Q - 1|$ with

$$Q = \frac{B(183.31\text{GHz}, A_s + b_s \cdot 3\,\text{K})}{0.5 \cdot [B(183.31 - 3\,\text{GHz}, 3\,\text{K}) + B(183.31 + 3\,\text{GHz}, 3\,\text{K})]} \tag{5.66}$$

I do the same with the corresponding frequencies for the channel at 183±7 GHz. As temperature I chose 3 K as it is around the typical DSV temperature corrected with ΔT_c for contamination. I obtain as values for the channel at 183±3 GHz

$$A_s = 0.00397\,\text{K} \qquad b_s = 0.99857 \tag{5.67}$$

and for the channel at 183±7 GHz

$$A_s = 0.00392\,\text{K} \qquad b_s = 0.99811. \tag{5.68}$$

These values enter the measurement equation in

$$\tilde{L}_{MS} = B(\nu, A_s + b_s(T_{\text{CMB}} + \Delta T_c)) \tag{5.69}$$

Using these band corrections reduces the deviation from 1 of the quotient compared to using no correction (i.e. $A_s = 0\,\text{K}$ and $b_s = 1$). For the channel at 183±3 GHz, $|Q - 1|$ is reduced from $3.3214 \cdot 10^{-4}$ to $1.67 \cdot 10^{-6}$. For the channel at 183±7 GHz, $|Q - 1|$ is reduced from $1.81066 \cdot 10^{-3}$ to $1.19 \cdot 10^{-6}$. The impact on the final brightness temperature is quite small (see below). Since

these new band coefficients provide a conceptual improvement, they enter the FIDUCEO MW calibration.

For the SSMT-2, no band correction coefficients were provided for the data sets so far. It is fair to assume that no band correction has been applied. Yet, a simple improvement on this can be made: since SSMT-2 has the same channels as AMSU-B, I use the same new band correction coefficients for SSMT-2 as for AMSU-B. Also for MHS channel 4, I use the new coefficients.

The effect on the final brightness temperature is quite small. When using both the new coefficients A, b as well as the new coefficients for the DSV A_s and b_s the final brightness temperature changes by about $-0.0009\,\mathrm{K}$ for channel 4 and by about $-0.0036\,\mathrm{K}$ for channel 5 compared to using the AAPP-values for the IWCT and simply zero and one for the DSV. Therefore, no significant improvement in terms of measured signal can be achieved, but it is the idea of FIDUCEO to improve the calibration process and its understanding wherever potential is discovered.

The uncertainty on the band correction coefficients is neglected.

5.3 FCDR production

Translating the concepts for the FCDR production into code requires handling of large data sets of different quality. Apart from implementing the measurement equation in all its subtleties along with the uncertainty propagation, it is required to perform quality checks, pre-processing steps and the reshaping of the data to equator-to-equator files. Finally, the FIDUCEO easy FCDR format is implemented to provide easily accessible NetCDF-4 files. Within this section, I present the overall abilities and an algorithm based documentation of the MW-FCDR generator that I implemented in MATLAB-2016. Some more detailed explanations on certain processing steps can be found in the Appendix 8.2.

The MW-FCDR generator is composed of two main parts: first, the definition of the data set to be processed as a single orbit (Sec. 5.3.1), and second, the actual preparation of that single orbit and its processing to level 1c and writing to NetCDF (Sec. 5.3.2).

After execution of the first block, the correct starting and ending scan lines for all orbits of one month (for a certain instrument) are defined. The equator-to-equator frame as well as data gaps are taken into account to find the correct scan lines. Having defined the single orbits of one month, the second block starts with the actual preparation of each single orbit, its processing to level 1c and the writing to NetCDF.

5.3.1 Preparatory steps: Defining the consolidated data set

A key aspect for the FIDUCEO FCDR being more convenient and easy-to-use is the framing of the data from one ascending/descending equator crossing to the next ascending/descending equator crossing. This eliminates all the usual overlaps between adjacent files and all random repetitions of fractions of orbits that are present in the NOAA CLASS data archive. In the following I present the main steps within the FCDR generator to achieve a cleaned-up consolidated data set for processing to level 1c. For some more detailed information on the steps and involved functions, see Appendix 8.2.2.

The FCDR generator works on monthly data. That means, the whole available data set for one instrument for a certain month is read and processed further. Before any processing actually starts, duplicated data from overlapping adjacent files or repetitions of fractions of orbits are removed. Only this allows the further steps to be executable in a meaningful manner. The next step represents the key aspect of the FIDUCEO FCDR files: it defines the start and end lines of the new orbital file within the equator-to-equator frame. Depending on the Local Equator Crossing Times (LECT) of the satellites, the FIDUCEO consortium has defined for which satellites the ascending node or descending node should be used as start and end point for the equator-to-equator frame. This differentiation between the satellites is necessary since the AVHRR instrument (flying on the same satellites as AMSU-B and MHS) requires a certain correction for disturbing sunlight that depends on the relative position to the sun and cannot be executed at the edges of a defined orbit. Hence, for each satellite that AVHRR is flown on, the ascending or descending node are chosen as start point depending on the LECT. In order to define the corresponding start and end lines, the equator crossings belonging to the chosen flight direction are determined in the processor by using the sign of the latitude's gradient over the scan lines as indicator.

Having found all significant equator crossings for defining start and end lines, a check needs to be made for data gaps: in the case that a large data gap occurs that reaches further than the next significant equator crossing, the FCDR processor detects this and ends the file at the start

of the gap. The next file then only starts at the first line after the gap. This strategy is chosen to not merge data from two equator-to-equator-frames into one file. Finally, a set of start and end scan lines is available for all equator-to-equator frames of the month, accounting for data gaps. This set is the starting point for the second block of the FCDR processor: the processing to level 1c.

5.3.2 Processing to level 1c and writing the Easy FCDR

From the first block of the FCDR generator, the start and end lines are defined for all files to be processed and written. Using this information, the monthly data set is chopped into the equator-to-equator files and processed further.

5.3.2.1 Set-up: collecting raw calibration data

The processing starts with a set-up of all quantities required for the evaluation of the measurement equation and uncertainty propagation for a single equator-to-equator file. For all per-scan line quantities that are collected for the calibration from the equator-frame data set, missing scan lines, i.e. small data gaps within the equator-frame are identified and filled up with fill-values. In the beginning of the set-up, general data are collected such as time, geolocation and satellite and solar angles. Afterwards, the raw counts for all views (i.e. samples) of the DSV and IWCT are collected as well as the PRT temperatures (5, 7 or 2 for MHS, AMSU-B or SSMT-2 respectively). For the AMSU-B instrument, the PRT temperatures are not directly provided in the level 1b files but need to be computed from the given counts $C_{k,l}$ and conversion coefficients $a_{i,k}$ (stored in the level 1b header for PRT sensors k). The conversion is a third order polynomial (Robel et al., 2009)

$$T_{k,l} = a_{0,k} + a_{1,k}C_{k,l} + a_{2,k}C_{k,l}^2 + a_{3,k}C_{k,l}^3 \tag{5.70}$$

yielding the temperature of PRT sensor k ($k = 1 \ldots 7$) for scan line l. For the DSV, IWCT counts and PRT temperatures, a preliminary average is calculated over all views/PRT-sensors per scan line. Also, a preliminary noise estimate is computed with the Allan deviation for each channel. Both the preliminary average and the preliminary noise estimate are required for the quality checks that follow the collection of raw calibration data.

5.3.2.2 Quality Checking

An essential step before the calibration is quality checking. This ensures that only data of good quality enter the processing towards a calibrated brightness temperature. The quality checking excludes bad data, e.g. outliers, and flags the affected views/ PRT-sensors and scan lines. The flags are used further in the processing to identify the bad lines and finally also provide the input for the construction of quality information bit masks for the FCDR users.

5.3.2.2.1 Moon intrusion   The first quality check is executed on the DSV only: the moon intrusion checking. It is only available for AMSU-B and MHS, since an auxiliary file from AAPP is required that does not exist for SSMT-2. In any case, moon intrusions in SSMT-2 are hard to detect because of the noise and the weak signal of the moon. Since the FOV is larger for SSMT-2, the moon fills a smaller fraction in the DSV and hence results in a weaker signal compared to the MHS and AMSU-B. Yet, this also implies that the error introduced by a moon intrusion is of much less importance for SSMT-2.

For AMSU-B and MHS, the AAPP provides subroutines `amsubcl.F` and `mhscl.F` creating per level 1b file a log-file ("cl-file") that, among other things, provides information on the minimum and maximum moon-angle occurring in each level 1-b file. Further, if the minimum moon-angle is smaller than 2°, it contains a list of all contaminated DSVs and scan lines. This information is used in the moon-checking code to flag and exclude contaminated measurements. At first, the scan line numbers listed in the cl-files are related to the scan line numbers of the new equator-to-equator frame so that the flagging and excluding of lines is later executed correctly. After identifying the contaminated lines, the number of contaminated views per scan line is counted. The criterion for good/ bad lines are

$$\text{\#contaminated views} < 4 \qquad \rightarrow \qquad \text{good} \tag{5.71}$$

$$\text{\#contaminated views} = 4 \qquad \rightarrow \qquad \text{bad} \tag{5.72}$$

Therefore, if at least one view is not contaminated, then I use this view as measurement for the scan line. If all four views are contaminated, the DSV measurements for this line is not used. As an intermediate step, the average over the good views is computed to give the mean DSV signal for the usable lines. This step is required in order to store the correctly computed mean based on the moon-checking. In the normal quality checks, a moon intrusion is not necessarily detected and therefore, this pre-calculated DSV average must be included after the other quality checks have provided their intermediate average per line. This pre-calculated DSV average for the partly moon-contaminated lines is later included in the DSV mean per scan line that actually enters the 7-scan line average.

It may occur that the AAPP-routine creating the cl-files partly failed and no contaminated scan lines are listed, although the minimum moon-angle is smaller than 2°. In this unusual event, the moon-checking of the FCDR generator throws an error. Hence, the moon-checking fails for this orbit. This case is flagged in the quality bit mask as invalid data for the whole orbit, since it cannot be known then from the cl-file which data of the orbit might be contaminated.

5.3.2.2.2 Quality checks on DSV, IWCT and PRTs

For all three quantities, the raw DSV, IWCT counts and the PRT temperatures, four quality checks are executed in order to filter out outliers and bad data. Some tests are closely related to the ones proposed in EUMETSAT (2013) and AAPP. However, the applied thresholds differ and are better adapted to the instruments behaviour analysed in Chapter 4.

The first check ensures that the values (i.e. counts from DSV or IWCT or temperatures from PRT) are in overall reasonable limits. This threshold test per view/ sensor and scan line asks whether

$$x_{\min} < x < x_{\max} \tag{5.73}$$

for a value x (with $x \in \{C_{\text{DSV}}, C_{\text{IWCT}}, T_{\text{PRT}}\}$) compared to channel and instrument dependent thresholds $x_{\min}$ and $x_{\max}$ that are pre-defined in the parameter files `coeffs_SATSEN...mat` that I prepared for the FCDR processing (see Appendix 8.2.1). If the statement is true, the view/ sensor passes this test.

The second test is the median test. It only accepts values x that are close to the median $\hat{x}$ over the views or sensors:

$$|\hat{x} - x| \leq 3\sigma \tag{5.74}$$

with σ being the preliminary noise estimate from the Allan deviation on the raw counts before the quality checking. This preliminary estimate is slightly larger than the true noise since outliers

still have an impact on the data. However, it is a useful estimate for application within the quality checks.

The third test is only executed per line and checks whether the minimum and maximum value for this line are too different:

$$\max(X_l) - \min(X_l) \leq 5\sigma \tag{5.75}$$

with X_l being the set of values for all views/ sensors of the scan line.

As an additional test I filter out jumps in the calibration data. It is necessary for events where all views/ sensors collectively jump to higher/ lower values. This event could not be detected by the previous tests, but could still result in a significant error if the Earth counts of that scan cycle do not jump accordingly. Hence, any jump in the intermediate mean (averaged over all views before quality checking), are flagged and the corresponding scan lines are excluded. The thresholds for these jumps are also predefined in the parameter files `coeffs_SATSEN...mat`. Despite of analysis of typical instrumental behaviour, the thresholds might still miss a jump or erroneously detect a jump, since not all events for all times and orbits can be captured in an analysis.

A view or sensor is accepted as good, if

- it passed the threshold test
- and it passed the median test

The classification of the views/ sensors is required to compute the mean values for the scan line using good views only. The classification of the scan line itself is derived as follows: A scan line is classified as bad, if

- only one view/sensor is accepted as good
- or a jump was detected
- or the scan line failed the minimum-maximum test

For a bad scan line, the calibration cannot be executed because of unusable calibration data from either DSV, IWCT, or PRT data. A bad scan line is flagged and excluded from further processing. If less than 300 scan lines of the equator-to-equator file are good, the affected channel is excluded from further processing, not calibrated for the orbit and filled with fill-values.

For each good scan line l, the mean value $\overline{x}_l$ is calculated from

$$\overline{x}_l = \frac{1}{\sum_k w_k} \sum_k w_k x_{k,l} \qquad \text{with} \qquad w_k = 0, 1 \tag{5.76}$$

where w_k is the weight, derived from the flag associated with each view/ sensor k: 1 is usable, 0 is not usable. All scan lines, for which less than the full number of views/ sensors could be used are flagged in order to inform the user. The mean value for each good scan line then enters the 7-scan line average procedure.

For the DSV an intermediate step is necessary in order to include the pre-calculated mean for the moon-contaminated lines where at least one view could be used. Since these lines and views might be classified erroneously as good/ bad, the actual moon-checking based mean for those lines is assigned to $\overline{x}_l$ at this stage.

The set of lines for the three quantities DSV, IWCT counts and PRT temperatures that are usable for the 7-scan line average and for further processing is composed of good lines, and lines

for which at least one DSV is uncontaminated from a Moon intrusion. These are used for the final noise estimation, the 7-scan line average and the calibration.

One could consider filling up the scattered bad scan lines with a mean value of the adjacent good lines, after the latter received their 7-scan line average. This is not yet implemented in the FCDR processor.

5.3.2.2.3 Quality checks on Earth views Bad data in Earth views also need to be excluded from processing, since they may even cause the processing to fail: outliers, for which the signal is far too high or too low, might result in negative values within the processing that prevent functions like *log* to be evaluated. Therefore two tests are executed in order to filter out bad data: As for the calibration data, I also check for jumps in the Earth view data. The thresholds used to identify jumps must be less strict for Earth views, but still strict enough to filter out most of the bad data. This is more prone to wrong classification of good/ bad data than it is for the calibration data, since the Earth view data reflect the changes in the scene which can be strong at cloud edges for example. The second test detects all negative and zero Earth counts. All pixels identified as unusable data are then excluded from the calibration.

5.3.2.3 Estimating noise

Having defined the good views/ sensors and lines, the noise is estimated by using the Allan deviation as explained in Section 4.2, Eq. 4.9, but only using the measurements classified as good. For the FCDR processing, I apply a rolling average-window of 300 scan lines to get a smooth varying estimate over the course of the orbit. Hence, each scan line to be calibrated has a noise estimate assigned to it that enters the 7-scan line average like the underlying calibration quantity itself.

5.3.2.4 Weighted rolling average over 7 scan lines

After averaging the usable DSV and IWCT count measurements as well as the PRT temperatures for a certain scan line, a weighted rolling average over 7 scan lines is applied to these three quantities and their respective noise estimates. The precise definition of usable lines from the quality checking is important in order to correctly calculate the 7-scan line average. Because the 7-scan line average is a weighted rolling average, its weights will need adjustment in case of unusable lines within the 7-scan line frame. This adaptation of weights is important for data gaps, that means truly missing scan lines, or scan lines with bad calibration data that do not enter further processing. The weighted average $\tilde{x}_0$ for a single line 0 is computed as

$$\tilde{x}_0 = \sum_{i=-n}^{n} w_i \overline{x}_i \tag{5.77}$$

$$\text{with} \quad \sum_{i=-n}^{n} w_i = 1 \tag{5.78}$$

$$\text{and} \quad w_i = \frac{1}{n+1}\left(1 - \frac{|i|}{n+1}\right). \tag{5.79}$$

This average is computed for all usable lines only. For all three instruments, I use $n = 3$ to achieve the 7 scan line average. This is also done in the operational calibration. The number of 7 scan lines has been chosen in Saunders et al. (1995) based on tests of noise impact. The
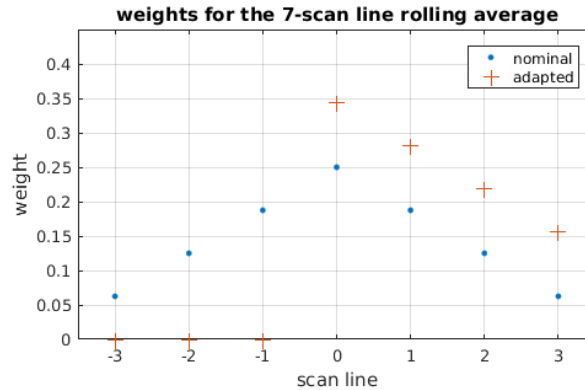
Figure 5.11 Weights for the scan lines used in the weighted rolling average for scan line 0.

choice of 7 scan lines was a good trade-off between smoothing the signal by reduction of noise through averaging and increasing the influence of noise by its low frequency $1/f$ noise again. The nominal values for the weights of the 7 scan lines are displayed in Fig. 5.11 as blue dots.

If one or more scan lines in this ± 3 scan line frame are unusable, their weight is set to zero. Hence, the weights of the remaining scan lines within this frame are adjusted to again fulfil Eq. 5.78. This is achieved by partitioning the missing weight equally among the remaining weights. The orange crosses in Fig. 5.11 show the adjusted weights for the case of a data gap at scan lines -3, -2 and -1, which cannot be used to compute the 7-scan line average for scan line 0. With the adapted, increased weights, the average over the remaining scan lines is computed to obtain the average for scan line 0. Similarly, the weights are adapted for the case of a single bad scan line within the 7-scan line frame.

The beginning and the end of the equator-to-equator file, however, are designed such that the actual line of the equator-crossing is only the fourth line of the file. Hence, it can be calibrated correctly with the full 7-scan line average. The first three lines (and last three lines) of the file are only stored for this purpose. They do not get calibrated themselves. Hence, for the start and end of the file the weight adjustment is not required.

The procedure of the weighted rolling average over 7 scan lines is executed for all basic calibration quantities to provide the averaged value for the DSV and IWCT counts and the PRT temperatures that enter the measurement equation, as well as the respective noise estimates entering the uncertainty propagation.

5.3.2.5 Set-up: collecting calibration parameters

Apart from the actual calibration data from the calibration targets, the full calibration procedure using the measurement equation also requires the various calibration parameters in order to take into account the effects discussed above in Sec. 5.2. During the set-up, all parameters are read from the header of the level 1b files, are set as hard coded values or read from the `coeffs_SATSEN...mat` files.

For the non-linearity coefficient q_{nl} and the warm target bias correction δT_{ch}, only preliminary values are read from the header: the actual value to be used in the measurement equation is obtained from an interpolation. For both parameters, three values are given per channel for three reference instrument temperatures T_{LOrefMin}, T_{LOrefNom} and T_{LOrefMax}. As proxy for the instrument temperature, the temperature of the local oscillator for channel 5 is used, following

EUMETSAT (2013). Depending on the current instrument temperature, the correct value for q_{nl} and δT_{ch} is calculated by linear interpolation between the three reference temperatures. The linear interpolation is done piecewise, i.e. between T_{LOrefMin} and T_{LOrefNom} (let this interval have index A) and between T_{LOrefNom} and T_{LOrefMax} (let this interval have index B). The reference temperatures T_{LOrefMin} etc. are given in the level 1b header. The slopes for the interpolation are m_A and m_B, the axis intercept is n_A and n_B. p acts as place holder in the formulae below and stands either for q_{nl} (i.e. non-linearity coefficient, q_{nl}) or δT_{ch}. The slopes and intercepts are computed using

$$m_A = \frac{(p_{\text{Min}} - p_{\text{Nom}})}{(T_{\text{LOrefMin}} - T_{\text{LOrefNom}})} \tag{5.80}$$

$$n_A = p_{\text{Nom}} - m_A \cdot T_{\text{LOrefNom}} \tag{5.81}$$

and for second temperature interval

$$m_B = \frac{(p_{\text{Nom}} - p_{\text{Max}})}{(T_{\text{LOrefNom}} - T_{\text{LOrefMax}})} \tag{5.82}$$

$$n_B = p_{\text{Max}} - m_B \cdot T_{\text{LOrefMax}} \tag{5.83}$$

The measured local oscillator temperature $T_{\text{LO},l}$ for line l is obtained from the count readings of a temperature sensor and its count-to-temperature conversion coefficients $t_0 \ldots t_4$ with a fourth order polynomial.

$$T_{\text{LO},l} = t_0 + t_1 C_l + t_2 C_l^2 + t_3 C_l^3 + t_4 C_l^4 \tag{5.84}$$

Before interpolation, it is checked whether the local oscillator temperature $T_{\text{LO},l}$ for line l is between T_{LOrefMin} and T_{LOrefNom} (use m_A and n_A) or between T_{LOrefNom} and T_{LOrefMax} (use m_B and n_B). The interpolation for δT_{ch} and q_{nl} is then

$$p_l = m_i \cdot T_{\text{LO},l} + n_i \tag{5.85}$$

with i being either A or B, depending on $T_{\text{LO},l}$ and p being either q_{nl} or δT_{ch}. Unlike the reference values for the non-linearity, the reference values for δT_{ch} are zero for all channels except channel 20 for AMSU-B on NOAA-17 (see Sec. 5.2.5). Hence, the interpolation for δT_{ch} has no effect, if using the reference values for δT_{ch} given in AAPP.

Further steps are also required for the measured position angles for the Earth views and DSVs . The angles θ are converted from the position counts to degree with conversion factor 0.007031250 (from AAPP, `clparams.dat`-file). Moreover, the angles for Earth views and DSVs enter the polarisation correction and therefore must come in a certain shape, having nadir at 0°. This is achieved for MHS with (i is the index for the space view i):

$$\theta_E = 0.007031250° \cdot C_{\text{posE}} - 180° \tag{5.86}$$

$$\theta_{S_i} = 0.007031250° \cdot C_{\text{posS}_i} - 180° \tag{5.87}$$

For AMSU-B, a different handling is needed since the angular positions are defined differently:

$$\theta'_E = 0.007031250° \cdot C_{\text{posE}} + 131.061° \tag{5.88}$$

$$\theta_E = \theta'_E - \texttt{floor}(\theta'_E/360°) \cdot 360° - 180° \tag{5.89}$$

$$\theta_{S_i} = 0.007031250° \cdot C_{\text{posS}_i} + 131.061° - 180° \tag{5.90}$$

As θ_S in the polarisation correction, I finally use the mean of the four provided values for the DSVs.

All position angles undergo a quality check, since a deviation from the nominal positions could result in a wrong geolocation. The condition for a good position for the 90 Earth views is

$$|\theta_{\mathrm{nom},\mathrm{E}}(i) - \theta_{\mathrm{E}}(i)| \leq \varepsilon_E \tag{5.91}$$

ε_E is 0.11 for AMSU-B and 0.25 for MHS, from AAPP `clparams.dat`-files (for SSMT-2, none of the angles is actually measured and stored in the level 1b files, hence, no checking is possible). If one view does not meet the condition, the scan line gets flagged as having problematic geolocation.

For the DSV, the condition is

$$|\theta_{\mathrm{ideal},\mathrm{S}}(i) - \theta_{\mathrm{S}}(i)| \leq \varepsilon_{\mathrm{S}} \tag{5.92}$$

ε_{S} is 0.5 for AMSU-B and 0.8 for MHS, from `clparams.dat`-files.

Having set all calibration parameters and computed q_{nl} and δT_{ch}, their corresponding uncertainties are set as hard coded values as described in Sec. 5.2.

In the case that the calibration is not possible for any channel, since there are insufficient (less than 300) scan lines of usable PRT temperature data, the nominal processing stops and writes an empty file with quality information only.

Usually however, the set-up is followed by the evaluation of the measurement equation.

5.3.2.6 Evaluation of the Measurement Equation

All calibration quantities have been preprocessed (quality-checked, averaged) to have the required shape to enter the measurement equation. The full measurement equation, evaluated for every channel, FOV and scan line reads

$$
\begin{aligned}
L_E = \frac{1}{g'} \Bigg[& B(\nu_{ch}, A + b \cdot (\overline{T}_{\mathrm{IWCT}} + \delta T_{ch})) \\
& + \frac{B(\nu_{ch}, A + b \cdot (\overline{T}_{\mathrm{IWCT}} + \delta T_{ch})) - B(\nu_{ch}, A_s + b_s \cdot (T_{\mathrm{CMB0}} + \Delta T_c))}{\overline{C}_{\mathrm{IWCT}} - \overline{C}_S} \cdot (C_E - \overline{C}_{\mathrm{IWCT}}) \\
& + q_{nl} \frac{(C_E - \overline{C}_S) \cdot (C_E - \overline{C}_{\mathrm{IWCT}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_S)^2} \\
& \quad \cdot [B(\nu_{ch}, A + b \cdot (\overline{T}_{\mathrm{IWCT}} + \delta T_{ch})) - B(\nu_{ch}, A_s + b_s \cdot (T_{\mathrm{CMB0}} + \Delta T_c))]^2 \\
& -(1 - g') \cdot B(\nu_{ch}, A_s + b_s \cdot (\overline{T}_{\mathrm{CMB0}})) \Bigg] \\
& + \alpha [B(\nu_{ch}, A + b \cdot (\overline{T}_{\mathrm{IWCT}} + \delta T_{ch})) - L'_E] \cdot \frac{1}{2}(\cos 2\theta_E - \cos 2\theta_S)
\end{aligned}
\tag{5.93}
$$

It uses the AAPP assumption $L_{Pl} = L_E$, and hence using $1 = g_E + g_S + g_{Pl}$ and $g' = 1 - g_S = g_E + g_{Pl}$. B represents the Planck function.

The last term, proportional to α is the polarisation correction. As pointed out in Sec. 5.2.7, an iterative procedure would be necessary. Since the correction is small compared to L_E, only

one step is actually executed, i.e. L'_E to be inserted in the correction is simply L_E evaluated from the measurement equation *without* the polarisation correction:

$$
\begin{aligned}
L'_E = \frac{1}{g'} \Big[& B(\nu_{ch}, A + b \cdot (\overline{T}_{\mathrm{IWCT}} + \delta T_{ch})) \\
& + \frac{B(\nu_{ch}, A + b \cdot (\overline{T}_{\mathrm{IWCT}} + \delta T_{ch})) - B(\nu_{ch}, A_s + b_s \cdot (T_{\mathrm{CMB0}} + \Delta T_c))}{\overline{C}_{\mathrm{IWCT}} - \overline{C}_S} \cdot (C_E - \overline{C}_{\mathrm{IWCT}}) \\
& + q_{nl} \frac{(C_E - \overline{C}_S) \cdot (C_E - \overline{C}_{\mathrm{IWCT}})}{(\overline{C}_{\mathrm{IWCT}} - \overline{C}_S)^2} \\
& \quad \cdot [B(\nu_{ch}, A + b \cdot (\overline{T}_{\mathrm{IWCT}} + \delta T_{ch})) - B(\nu_{ch}, A_s + b_s \cdot (T_{\mathrm{CMB0}} + \Delta T_c))]^2 \\
& - (1 - g') \cdot B(\nu_{ch}, A_s + b_s \cdot (\overline{T}_{\mathrm{CMB0}})) \Big]
\end{aligned}
\tag{5.94}
$$

In principle, this means that the sensitivity coefficient for any effect x should get another contribution, namely

$$
-\alpha \cdot \frac{\partial L'_E}{\partial x} \approx -\alpha \cdot \frac{\partial L_E}{\partial x}
\tag{5.95}
$$

The largest absolute value for α is 0.0022 for channel 3 and 4. Hence, the uncertainty on L_E would be

$$
u_{L_E}(x) = \frac{\partial L_E}{\partial x}(1 + 0.0022) \cdot u_x
\tag{5.96}
$$

$$
= 1.0022 \cdot u'_{L_E}(x)
\tag{5.97}
$$

for $u'_{L_E}(x)$ being the uncertainty on L_E due to effect x if not including the reappearance of L_E in the sensitivity coefficient. The deviation in the uncertainty is only 0.2%. Even if Eq. 5.95 is an approximation only, the difference between taking the second appearance of L_E (and hence all parameters) into account and ignoring it, is negligible for the final uncertainty value.

Note: If needed, the FCDR processor is also able to handle the general case without the AAPP-assumption of $L_{Pl} = L_E$. In this case, a temperature of the platform must be assumed and set as a hard coded value.

As a final step of the calibration, to obtain the final brightness temperature for the Earth scene, the radiance is converted with the inverse Planck function B^{-1}:

$$
T_{\mathrm{B,E}} = \frac{1}{b} \cdot [B^{-1}(\nu_{ch}, L_E) - A]
\tag{5.98}
$$

Strictly speaking, the usage of the band correction factors A and b is not necessary here. They were determined to account for the fact that the spectrum of a black body was measured at two separated frequencies rather than at a central one. However, the Earth signal is affected by the absorption lines and therefore it is no black body. The measured Earth signal is an average of the radiances at two frequencies, e.g. placed at the slopes of the 183 GHz water vapour absorption line. The brightness temperature is only another unit for this mean radiance, defined as the temperature that a black body would have if it gave this radiance signal. Hence, it is a question of convention, whether to include the band corrections in the brightness temperature of the Earth or not. For consistency, I follow the AAPP method using A and b.

5.3.2.7 Uncertainty propagation

The uncertainty propagation uses the sensitivity coefficients as defined in the respective sections on effects and evaluates them for all channels, FOVs and scan lines. The final uncertainty on brightness temperature due to a single effect x is computed as

$$u_x(T_{\mathrm{B,E}}) = \frac{\partial T_{\mathrm{B,E}}}{\partial L_{\mathrm{E}}} \frac{\partial L_{\mathrm{E}}}{\partial x} u(x). \tag{5.99}$$

With c_i being the sensitivity coefficient $\partial L_{\mathrm{E}}/\partial x_i$, we have for the set of independent effects i

$$u_i^2(T_{\mathrm{B,E}}) = \sum_i \left(\frac{\partial T_{\mathrm{B,E}}}{\partial L_{\mathrm{E}}} c_i u(x_i) \right)^2. \tag{5.100}$$

For the structured effects s we have

$$u_s^2(T_{\mathrm{B,E}}) = \sum_s \left(\frac{\partial T_{\mathrm{B,E}}}{\partial L_{\mathrm{E}}} c_s u(x_s) \right)^2 \tag{5.101}$$

and for the common effects co we have

$$u_{co}^2(T_{\mathrm{B,E}}) = \sum_{co} \left(\frac{\partial T_{\mathrm{B,E}}}{\partial L_{\mathrm{E}}} c_{co} u(x_{co}) \right)^2. \tag{5.102}$$

The uncertainty due to RFI is included in $u_{co}^2(T_{\mathrm{B,E}})$ only after the compiling of quality flags (Sec. 5.3.2.8). This is justified by the fact that the uncertainty induced by RFI is only added when any transmitter is on. This information is contained in the quality flag settings.

5.3.2.8 Compiling Quality Flags

The FIDUCEO easy FCDR is designed to have extensive, but easy to use quality information. To meet this requirement, a set of meaningful flags must be compiled from the newly processed flags from the quality checks during processing and from the multitude of available level 1b flags. Only part of the quality information from the level 1b is of interest for the FCDR users, cannot be inferred from the data during processing and therefore has to be copied over from the level 1b file. This concerns general and detailed information on time and geolocation problems. The corresponding flags are located in the quality information blocks of the record (Robel et al., 2009). All quality flags provided in the level 1b files on these issues are compiled into one flag each for time and for geolocation. The flag for time or for geolocation problems is set, if any of the corresponding level 1b flags was set for the scan line.

Further information that is copied over from level 1b files is the information on the status of the transmitters (STX, SARR). This is stored in the level 1b files in the analog telemetry block (Robel et al., 2009).

Quality information on the calibration data and Earth view data are collected as flags from the quality checks performed during FCDR processing. These quality flags include information on missing scan lines and bad scan lines that could not be calibrated. The quality checks also keep track of which line was calibrated using less than the nominal number of calibration views/ PRT sensors, or whether the 7-scan line average was executed on less than 7 lines for a certain line. Also, moon intrusions are flagged according to their impact on calibration. Bad Earth pixels are also flagged.

This quality information, combining level 1b and FCDR processing information is compiled to give three bit masks, one of which exists per channel yielding seven bit masks in total. In Sec. 5.4, I present a table displaying the complete set of quality flags.

The final uncertainty estimate for RFI is made after all quality bit masks are set: From the quality bit masks given per scan line, the information on transmitter status is read. If any transmitter is on for a certain scan line, then the uncertainty on brightness temperature due to RFI is included in the uncertainty due to common effects. This is only the case for the AMSU-B instruments on NOAA-15 and NOAA-17 that are known to be affected by RFI to some extent and that some bias measurements from the in-orbit verification phase tests are available for.

5.3.2.9 Estimation of cross-channel correlation

As a further post-processing step, the cross-channel correlation matrix is computed: per effect, the estimated cross-channel parameter correlation matrix is hard coded. Afterwards, the corresponding uncertainty matrix is built from the uncertainty information on that effect obtained from the uncertainty propagation. The final cross-channel correlation matrices per effects class are computed using Eq. 5.10 to 5.18. This includes the averaging procedure over a chosen sample of scan lines and FOVs as explained in Sec. 5.1.2. Any channel that was not calibrated for the current orbit receives fill values in the corresponding row and column of the correlation matrix.

5.3.2.10 Writing FCDR files

The FIDUCEO easy FCDR files are written as NetCDF-4 format. Each equator-to-equator frame corresponds to one NetCDF file, without overlap with adjacent ones. For each of the three MW instruments, I implemented an individual writer adapted to the features of each instrument. The overall set of variables is the same however. The FCDR files are provided with global attributes and individual attributes per variable in order to make the FCDR self-explaining to a certain degree.

The processed level 1c data are scaled to be stored as integer in order to reduce the required disk storage. The scale factor is part of the attributes of each variable and allows for automatic re-scaling in e.g. MATLAB and *python* NetCDF-reading functions or the software *panoply*.

Data entries for uncalibrated or missing scan lines get filled with fill values. This is also valid for the three lines before and after the starting and ending equator-crossing, respectively: these lines were used only for proper calibration of the starting/ ending line at the equator-crossing, but were not calibrated themselves.

Apart from the level 1c variables, I also include identifiers in order to maintain traceability to the original level 1b file that each scan line originates from.

Instrument and Satellite	Start date	End date
SSMT-2 F11	1994-06-01	1995-04-17
SSMT-2 F12	1994-10-13	2001-01-08
SSMT-2 F14	1997-04-28	2005-01-18
SSMT-2 F15	2000-01-24	2005-01-18
AMSU-B NOAA-15	1999-04-01	2011-03-28
AMSU-B NOAA-16	2001-03-20	2011-12-31
AMSU-B NOAA-17	2002-10-15	2013-04-10
MHS NOAA-18	2005-08-30	2017-11-30
MHS NOAA-19	2009-04-14	2017-11-30
MHS Metop-A	2007-05-21	2017-11-30
MHS Metop-B	2013-01-29	2017-11-30

Table 5.1 Periods of available FCDR data

5.4 The new MW-FCDR: example contents

The new FIDUCEO processor is run over all available NOAA-CLASS level 1b data of AMSU-B, MHS and SSMT-2 instruments. The time ranges covered are therefore the same as for the study of noise performance in Section 4.2. Table 5.1 summarises the available periods. Note that for certain periods, there are uncalibrated channels due to instrumental degradation. Also, single orbits may be missing because of lack of data or impossible calibration. Each FCDR file has a size of about 6.8 MB for AMSU-B and MHS or 0.6 MB for SSMT-2. Note that SSMT-2 only has 28 pixels instead of 90 pixels, which reduces the file size significantly. In total, we have 2.2 TB for the whole MW-FCDR.

Note that the FCDR at this stage is an *un*harmonised FCDR. Nonetheless, it is already a big leap forward compared to existing operational data, since the FCDR provides the data in an easy-to-use NetCDF format, on the basis of consolidated data. Furthermore, the FCDR provides extensive uncertainty information. I will present some details of these two novel aspects in the following.

Across all instruments of the FIDUCEO project, a common (basic) format has been developed in the project consortium in order to guarantee homogeneity among the various FIDUCEO FCDRs, as far as the different characteristics of the instruments allow. Each file of the FIDUCEO FCDR is a NetCDF-4 file containing the data for a full orbit from one equator-crossing to the next one in the same direction. The naming convention for the FIDUCEO FCDR is

FIDUCEO_FCDR_L1C_{SEN}_{SAT}_{START}_{END}_EASY_vX.X_fvX.X.nc

For the microwave FCDR, SEN stands for SSMT2, AMSUB or MHS; SAT is replaced by the corresponding satellite F11, F12, F14, F15, NOAA15, NOAA16, NOAA17, NOAA18, NOAA19, METOPA or METOPB. START and END are the start an end date and time of the orbit in the format YYYYMMDDHHMMSS. vX.X and fvX.X. indicate the version number for the data processing and the format respectively.

The header of each file provides global attributes for general information. An example of the global attributes section of the header is given in Fig. 5.12. The global attributes inform again about sensor, satellite, start and end times as easily accessible string variables. Also, the data file names of the underlying original level 1b data used for this specific orbit are listed.

```
// global attributes:
    :Conventions = "CF-1.6";
    :institution = "Universitaet Hamburg";
    :source = "NSS.MHSX.M1.D15187.S1522.E1551.B1451919.MM.gz
NSS.MHSX.M1.D15187.S1553.E1614.B1451919.MM.gz
NSS.MHSX.M1.D15187.S1608.E1702.B1451920.SV.gz
NSS.MHSX.M1.D15187.S1702.E1753.B1452020.MM.gz ";
    :title = "Microwave humidity sounder Easy-Fundamental Climate Data
Record (MW-Easy-FCDR)";
    :history = ; // double
    :references = ; // double
    :id = "product doi will be placed here";
    :naming_authority = "Institution that published the doi";
    :licence = "This dataset is released for use under CC-BY licence
(https://creativecommons.org/licenses/by/4.0/) and was developed in
the EC \\nFIDUCEO project Fidelity and Uncertainty in Climate Data
Records from Earth Observations. Grant Agreement: 638822.";
    :writer_version = "MATLAB script write_easyFCDR_orbitfile_MHS.m";
    :satellite = "metopb";
    :instrument = "mhs";
    :comment = "WARNING: This is an early pre-beta version. ";
    :StartTimeOfOrbit = "06-Jul-2015 15:47:58";
    :EndTimeOfOrbit = "06-Jul-2015 17:29:18";
```

Figure 5.12 Example contents of global attributes in the header of NetCDF FCDR file for the example orbit of MHS on Metop-B, start: 2015-07-06 15:47:58 end: 2015-07-06 17:29:18.

Information about the original level 1b files is valuable for maintaining traceability back to the original data. It is complemented by both a file identifier variable `scanline_map_to_origl1bfile` that links every scan line to a level 1b file listed in the global attributes, and a scan line identifier variable `scanline_origl1b` that relates every scan line of the FCDR to the original scan line number in the corresponding level 1b file. Figure 5.13 shows the scan line identifier and the file identifier for an exemplary orbit.

Further auxiliary variables included in the FCDR are the satellite and solar zenith and azimuth angles for the AMSU-B and MHS instruments.

Taking an exemplary orbit of MHS on Metop-B from 06-Jul-2015 15:47:58 to 06-Jul-2015 17:29:18, I present the contents of the FCDR files in the following.

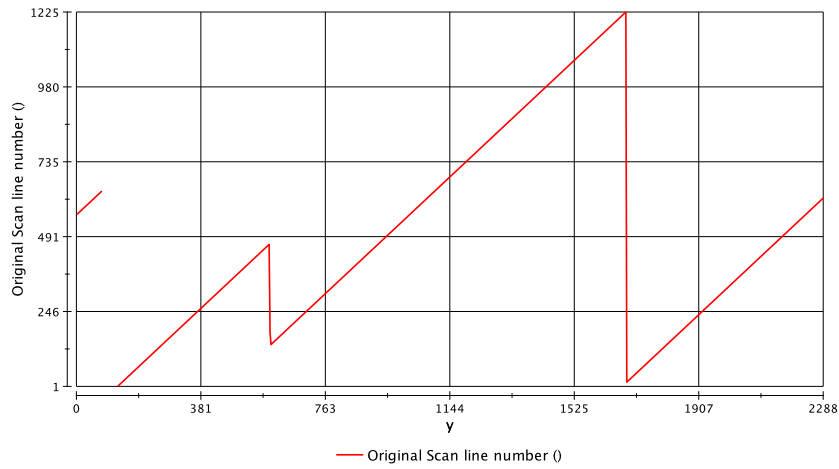### 5.4.1 Quality information

To judge the overall usability of the data contained in a single FCDR file, the variable

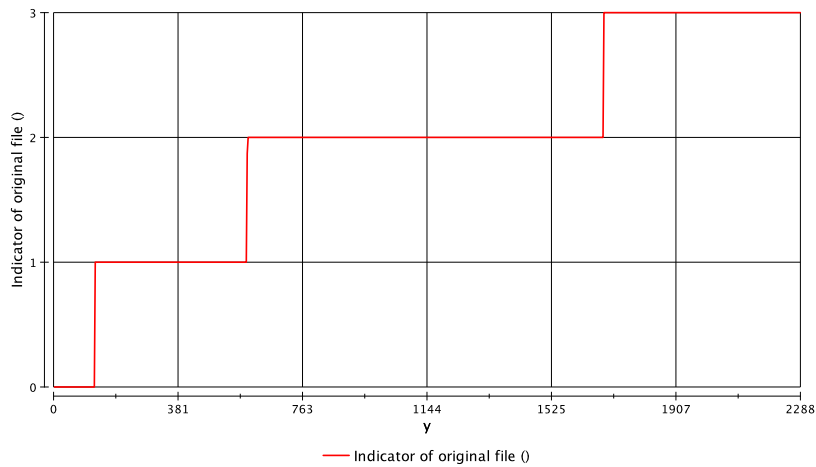<div align="center">

`quality_pixel_bitmask`

</div>

provides general, per pixel quality information. This quality flags variable is available for all sensors considered in the FIDUCEO project and therefore has a common format. A second quality bit mask

<div align="center">

`data_quality_bitmask`

</div>

(a) level 1b - scan line identifier

(b) level 1b - file identifier

Figure 5.13 Example contents of identifier variables for maintaining traceability back to original level 1b data records. Orbit of MHS on Metop-B start: 2015-07-06 15:47:58 end: 2015-07-06 17:29:18.

gives sensor specific quality information. The meaning of the bit positions is explained in the variables' attributes. See Tables 5.2 and 5.3 at the end of this subsection on quality information for the full explanation of the flag meanings for these quality bit masks.

Bit 0 corresponds to the least significant bit, i.e. $2^0 = 1$. The general bit mask

```
quality_pixel_bitmask
```

is compiled as

$$\text{flag} = b_7 \cdot 2^7 + b_6 \cdot 2^6 + \ldots + b_1 \cdot 2^1 + b_0 \cdot 2^0, \tag{5.103}$$

using the bits b_7 to b_0 that can take either 0 or 1 to communicate the information presented in Table 5.2. Any other bit mask is compiled accordingly.

To get a first overview of the quality of the example orbit of MHS on Metop-B, the general bit mask is displayed in Fig. 5.14a on a latitude – longitude grid. Note the equator-to-equator frame of the orbit. Obviously there are two issues within this orbit: 1. There is a data gap close

(a) `quality_pixel_bitmask` : General information

(b) `data_quality_bitmask` : Sensor specific information
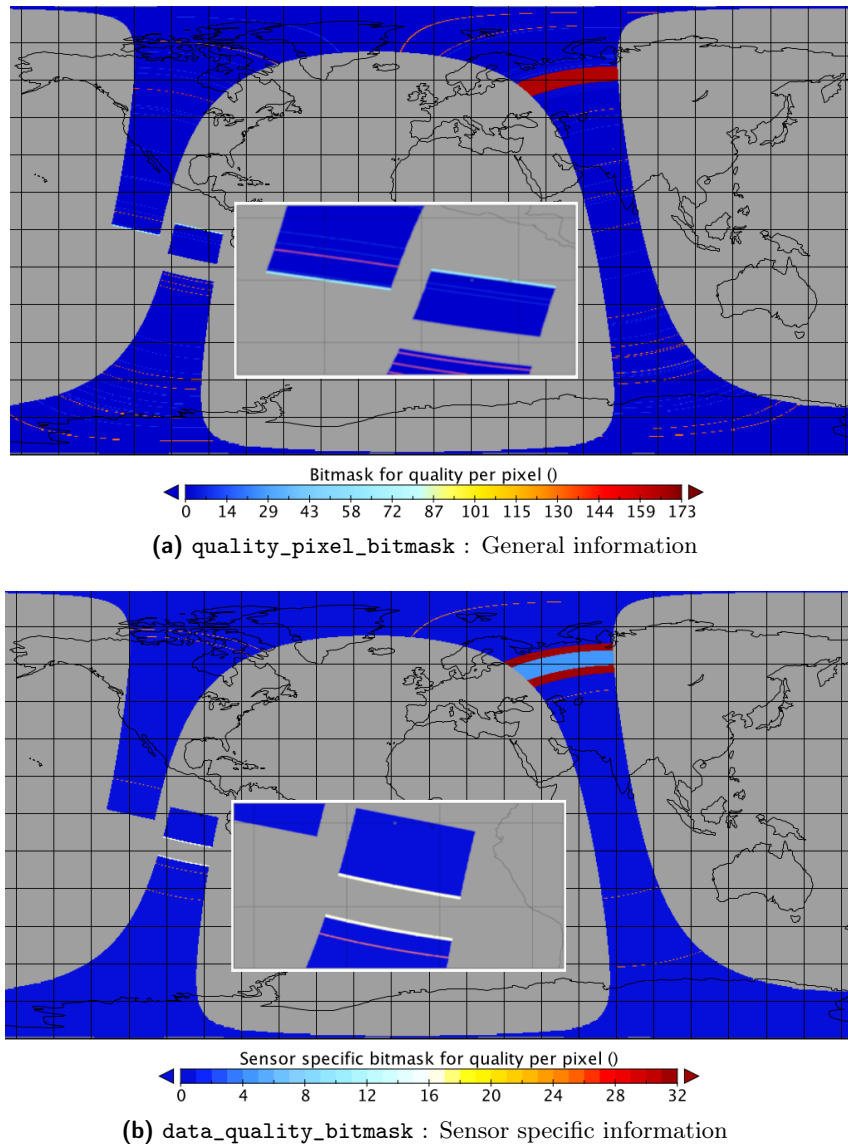
Figure 5.14 Quality bit masks plotted for the full orbit file of the example orbit of MHS on Metop-B.

to the start of the file. 2. Some further quality problem when the satellite flies over Russia. In the following I analyse the quality information communicated through the different bit masks.

The first issue is flagged in the `quality_pixel_bitmask` as "padded data" since fill values have been inserted for this gap. This flag value of 65 (bit 6 is set for "padded data" and bit 0 set for "invalid") does not show up in the graphic since no geolocation values are available for the data gap either and hence no lat-lon plotting is possible. The "padded data" and "invalid" flag is also raised for the first three and last three lines stored in the orbit file. Note the light blue shade at the start and end of the data, on which is zoomed in (see the white-framed picture in Fig. 5.14a). This flag is caused by the lack of calibration for these lines: They are only stored for correct calibration of the 4th and 4th to last lines that actually mark the true equator crossing.

The second issue is mostly flagged as 165 in Fig. 5.14a (at some lines an "invalid geolocation" adds to it to yield the 173 indicated on the colour scale). Since 165 is an odd number, obviously

bit 0 has been set to indicate "invalid" pixels which must not be used. The full bit mask reads 10100101 (from most significant bit at left to least significant at right). Hence, bit 2 is set to flag "invalid input", bit 5 is set to relate this to a "sensor error". Bit 7 is set to indicate that there are "uncalibrated channels" for these pixels.

Furthermore, there are some scattered scan lines during the orbit that are flagged as 130 which means that bit 7 and 2 are set. This indicates "use with caution", because there are "uncalibrated channels" for that line. The channel specific information can be inferred from the bit masks `quality_issue_pixel_ChX_bitmask` that I explain further below.

The `quality_pixel_bitmask` communicates quality issues of a general nature and might occur for any instrument on board a polar orbiter. Hence, this quality bit mask is available for any FIDUCEO FCDR as mentioned above. The sensor specific ones (valid for all channels of the instrument) are stored in

<div align="center">

`data_quality_bitmask`,

</div>

which is explained in Table 5.3 and displayed in Fig. 5.14b for the example orbit. The two mentioned issues within this example orbit are further described through this `data_quality_bitmask`. The white-framed picture in Fig. 5.14b zooms in on the very edges of the data gap that are coloured in faint yellow. The bit mask has the value 16, indicating that the calibration is suspected to be non-nominal due to issues with PRT sensors. In fact, due to the data gap, the 7-scan line average for the PRT sensors could only be executed over fewer lines which triggered the setting of this flag. Some further scan lines in the orbit are flagged with 24, indicating that both flags for "suspect PRT" and "suspect black body temperature" are set. This combination of flags indicates that less than 5 PRT sensors were used for the single line and that a possible gradient might be overlooked by that. This is not critical however, since the adjacent scan lines still build up the 7 scan line average that is finally used for the calibration of that line.

The second issue, over Russia, is marked by `data_quality_bitmask` values of 32 and 4. Both of them indicate a moon intrusion. The flag 32 indicates that the calibration is suspected to be non-nominal due to a moon intrusion, i.e. that some DSV are moon-contaminated and only the uncontaminated views have entered the calibration. Note also, that this event raises the "use with caution" flag in the general `quality_pixel_bitmask` (not visible in the plot of Fig. 5.14a since 2 is too close to zero to be resolved in the colour scale). This quality flag of 32 is set at the beginning and the end of the moon intrusion, where the contamination only partially affects the DSVs. In between, the flag 4 is set to indicate that the lines were not calibrated due to full moon contamination in all four DSVs.

The channel-specific bit masks (see Table 5.4)

<div align="center">

`quality_issue_pixel_ChX_bitmask`

</div>

complement the general one and indicate for every channel individually, which scan line was calibrated, but with fewer lines or views than nominal and which lines were not calibrated. See Fig. 5.15a and 5.15b for a zoomed-in picture of the channel-specific bit mask for Channel 1 for this orbit. Note the suspicious calibration at the edges of the data gap in 5.15a (flag 3: "suspect calib DSV" and "suspect calib IWCT") and the increasing and decreasing impact of the moon intrusion in 5.15b (flag 1 and 4, "suspect calib DSV" and "no calib bad DSV", respectively). Also note the uncalibrated scan line (flag 4: "no calib bad DSV") directly after the data gap in 5.15a. Since there is bad data for this scan line in the DSV, it is not calibrated. Consequently, the three preceding and the three succeeding lines are flagged as "suspect calib DSV" to indicate

that the 7-scan line average for those lines was executed with only 6 lines. Over the orbit, some more of such scattered bad scan lines are found due to bad DSV or IWCT view data in the different channels.
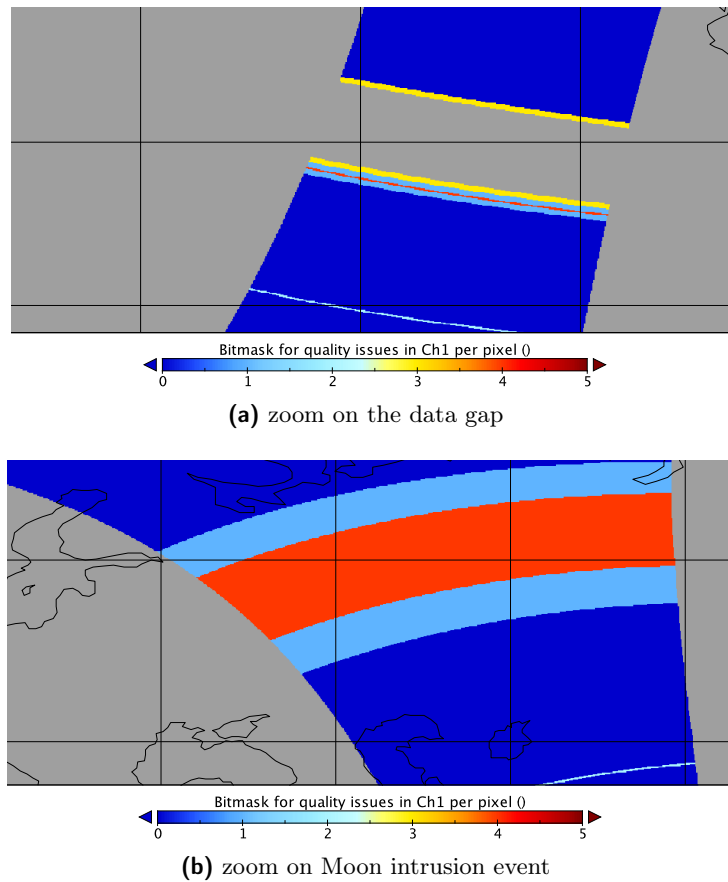

(a) zoom on the data gap

(b) zoom on Moon intrusion event

Figure 5.15 Variable `quality_issue_pixel_Ch1_bitmask`: channel specific information for the example orbit of MHS on Metop-B.

GENERAL QUALITY INFORMATION		
Flag name	**Bit**	**Description**
invalid	0	General flag for invalid data. Set to TRUE if any of the following is set: invalid_input, invalid_geoloc, invalid_time, sensor_error, padded_data or any sensor specific flag that indicates invalid data.
use_with_caution	1	Input data flags set that indicate potential errors. Set to TRUE if one or more of the original sensor data flags indicate possible (but usually not critical) problems or if data in a single channel is not usable. Definition of this flag combination in sensor specific section.
invalid_input	2	Input data invalid flag. Set to TRUE if a combination of the original sensor data flags indicates unuseable data. Definition of this flag combination in sensor specific section.
invalid_geoloc	3	Flag is raised if the geolocation or viewing-geometry data of this pixel is not valid.
invalid_time	4	Flag is raised if the acquisition time data of the pixel is not valid.
sensor_error	5	Flag is raised if the measurement data or sensor status data is not valid.
padded_data	6	Pixel contains fill value or repeated data; the corresponding measurement data is stored in the previous/next orbit file. Usually this data originates from correlation-calculations overlapping orbit-file boundaries.
incomplete_channel_data	7	Flag is raised if data for one or more channels is incomplete.

Table 5.2 Explanation of `quality_pixel_bitmask`

SENSOR SPECIFIC QUALITY INFORMATION		
Flag name	**Bit**	**Description**
moon_check_fails	0	The check for Moon intrusion failed. Hence no valid DSV data. If set, "invalid_input" is also set.
no_calib_bad_prt	1	All PRT measurements are bad. Calibration impossible. If set, "sensor_error" is also set.
no_calib_moon_intrusion	2	Moon intrusion detected. Moon contaminates all four DSV. If set, "sensor_error" and "invalid_input" is also set.
susp_calib_bb_temp	3	Less than the full number of PRT sensors has been used for calibration. An unaccounted for temperature gradient might be missed. If set, "use_with_caution" is also set.
susp_calib_prt	4	Less than the full number of PRT sensors has been used for calibration. OR: Fewer scan lines have been used to get the weighted average of the current one. None of those issues impacts the final calibration significantly.
susp_calib_moon_intrusion	5	Moon intrusion detected. At least one DSV could be used for calibration. If set, "use_with_caution" is also set.

Table 5.3 Explanation of `data_quality_bitmask`

CHANNEL SPECIFIC QUALITY INFORMATION		
Flag name	**Bit**	**Description**
susp_calib_DSV	0	Less than 4 DSV could be used for calibration. This includes the case of partial Moon contamination. OR: Less than 7 scan lines have been used to get the weighted average of the current one. None of those issues impacts the final calibration significantly.
susp_calib_IWCT	1	Less than 4 IWCT views could be used for calibration. OR: Less than 7 scan lines have been used to get the weighted average of the current one. None of those issues impacts the final calibration significantly.
no_calib_bad_DSV	2	Bad DSV data for this scan line. Calibration impossible.
no_calib_bad_IWCT	3	Bad IWCT data for this scan line. Calibration impossible.
bad_data_earthview	4	Bad data from Earth view.

Table 5.4 Explanation of `quality_issue_pixel_ChX_bitmask` for Channel X, where X indicates the channel name.

5.4.2 Brightness temperatures, their uncertainty and correlation

The brightness temperature for each channel is the actual observable stored in the FCDR. Its value is based on the calibration executed with the FCDR processor, taking into account quality issues. Hence, no calibrated brightness temperature is available in the data gap, during the moon intrusion where all four DSVs were affected, and for the scattered scan lines with bad calibration data. Note that the latter cannot be resolved at the scale of the figures 5.16 to 5.20. Figures 5.16 to 5.20 show the brightness temperature (a), the uncertainty of the brightness temperature due to independent effects (b), due to structured effects (c) and due to common effects (d) during the orbit for all channels.

The **independent effects** are composed of

- noise on Earth counts
- and random variations of the antenna position in Earth views.

The latter one enters only by the polarisation correction proportional to α. In the current example, α is zero for MHS on Metop-B, using the AAPP defined calibration parameters. Hence, in this example, the independent effects only include the noise on Earth counts. The noise on Earth counts often dominates the uncertainty budget, as can be seen in a histogram also (see Fig. 5.21a). Unlike the DSV and IWCT signal used in the calibration, the Earth counts originate from a single measurement only and undergo no averaging. Hence, their noise contribution is significantly higher than for the DSV and IWCT views. The noise in the DSV and IWCT views enters the structured effects. The list of **structured effects** is

- DSV count noise
- IWCT count noise
- PRT noise
- random variations of the antenna position in DSVs

The list of **common effects** is

- PRT accuracy
- warm target bias correction δT_{ch}
- cold target bias correction ΔT_c
- non-linearity q_{nl}
- polarisation correction α
- antenna pattern correction g_i
- radiance of platform
- antenna position accuracy in Earth views
- antenna position accuracy in DSVs
- radio frequency interference

The uncertainty in brightness temperature due to common effects is shown in Figures (d) of Fig. 5.16 to 5.20. For channel 1, a very clear scan-dependent pattern is seen that is related to the antenna pattern correction, which is strongest in channel 1 compared to the other channels. At

the very edge of the scan, the applied correction is largest and the trust in it is lowest. Hence, the propagated uncertainty in brightness temperature reflects this change of trust over the scan. This is also true for the other channels, but much less pronounced. Especially for the sounding channels, the uncertainty is increased at lower temperatures (e.g. Antarctica). This is due to the polarisation correction that has higher impact at low temperatures. Therefore, uncertainty on the polarisation correction translates to uncertain brightness temperature at the lower end.

Figure 5.21a shows a histogram of uncertainties of the whole month of September 2014 for MHS on Metop-B. For all channels but channel 1, the independent effects dominate the uncertainty budget as mentioned earlier. This is an important point for climate monitoring: this uncertainty contribution can be reduced by averaging procedures, and hence a large part of uncertainty can be overcome by sufficient amount of data. Only channel 1 has overlapping distributions of common and independent effects and even a long tail of higher uncertainty due to common effects that cannot be reduced by any averaging. Compared to the persistent uncertainty due to common effects, the uncertainty due to structured effects is mostly small, only for channel 2 there is overlap of both uncertainty contributions.

The overall appearance of the histogram of uncertainty is dramatically changed, if RFI possibly affects the data. Figure 5.21b shows the histogram for AMSU-B on NOAA-15 for September 2005. For all channels, the uncertainty due to common effects dominates the uncertainty due to structured and independent effects which are hardly resolved at this scale (they are very similar to the ones shown for MHS on Metop-B, though). The uncertainty due to common effects dominates because the uncertainty due to RFI is included in the common effects. Note that RFI easily introduces biases of many Kelvin (see Sec. 5.2.9) which is represented here by high uncertainty. This uncertainty contribution is only present for NOAA-15 and 17, which we have actual characterisation of RFI for, at least from the beginning of the mission. The uncertainty due to RFI is added for any time that any transmitter is switched on. The transmitter status is communicated via the `quality_scanline_bitmask`.

Information on correlation is provided in further variables: correlation length scales and cross-channel correlation matrices for the different effect classes. The correlation length scales cross-element and cross-line give a coarse estimate of length scales over which correlation is present (see Sec. 5.1.2). They are given each as single number valid for all scan lines.

The cross-channel correlation matrices are obtained from initial estimation of cross-channel correlation for each effect and propagated through to give the final cross-channel correlation matrix for brightness temperatures for the different effect classes. The procedure was described in Sec. 5.1.2. Figure 5.22 shows the cross-channel correlation matrices for the independent, structured and common effects for the exemplary Metop-B orbit. The independent effects are also spectrally independent, and therefore, the correlation matrix is the identity matrix. The structured effects include spectrally independent effects (DSV, IWCT noise), as well as a highly correlated effect: the noise on the PRT sensors influences all channels. Thus, the final cross-channel correlation matrix from structured effects is close to the identity matrix, but still has small non-zero off-diagonals. Among the common effects, there are some that are highly correlated between the sounding channels, or even all channels (PRT accuracy). Hence, the brightness temperatures of all channels are correlated to some degree (all elements of the matrix are equal to or larger than 0.46).

The spectral correlation information needs to be taken into account if more than one channel is used for a retrieval of a geophysical quantity. Spatio-temporal correlation must be accounted for in averaging procedures by including correlation terms in the uncertainty computation of an averaged value.
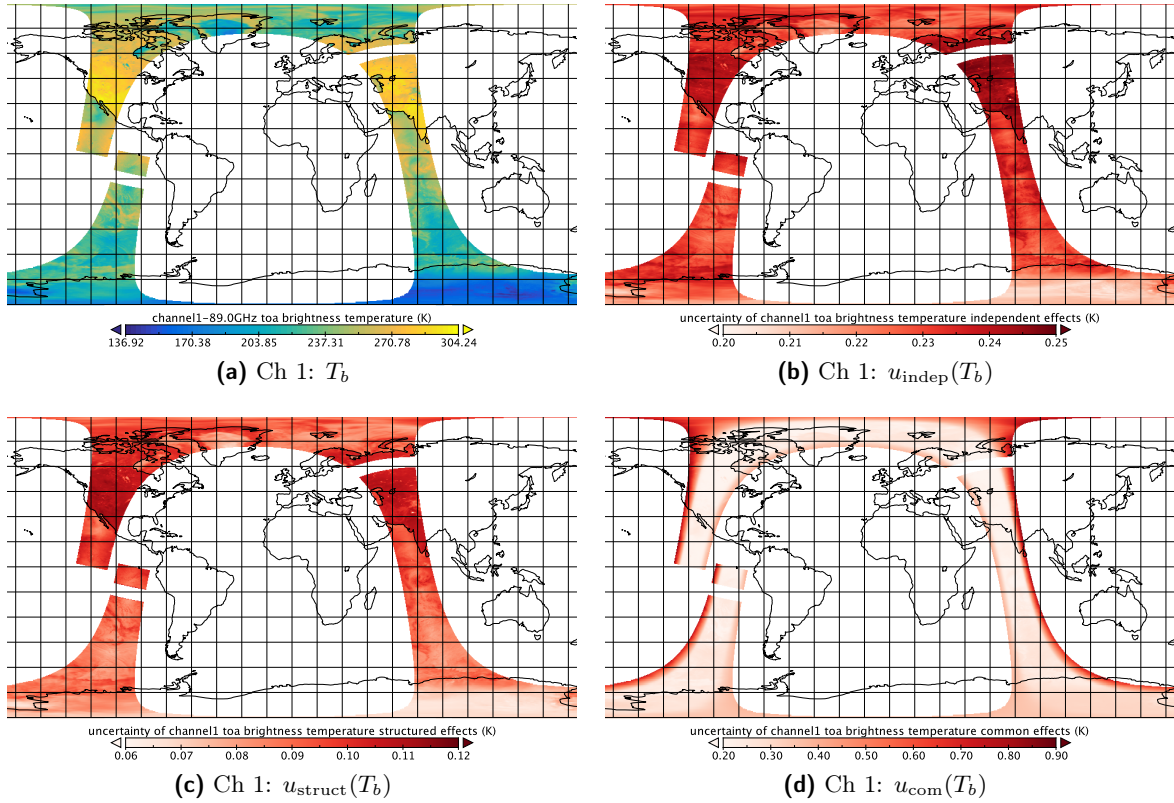
(a) Ch 1: T_b

(b) Ch 1: $u_{\mathrm{indep}}(T_b)$

(c) Ch 1: $u_{\mathrm{struct}}(T_b)$

(d) Ch 1: $u_{\mathrm{com}}(T_b)$

Figure 5.16 Channel 1: Brightness temperature and its uncertainties (example orbit of MHS on Metop-B).

(a) Ch 2: T_b

(b) Ch 2: $u_{\mathrm{indep}}(T_b)$

(c) Ch 2: $u_{\mathrm{struct}}(T_b)$

(d) Ch 2: $u_{\mathrm{com}}(T_b)$
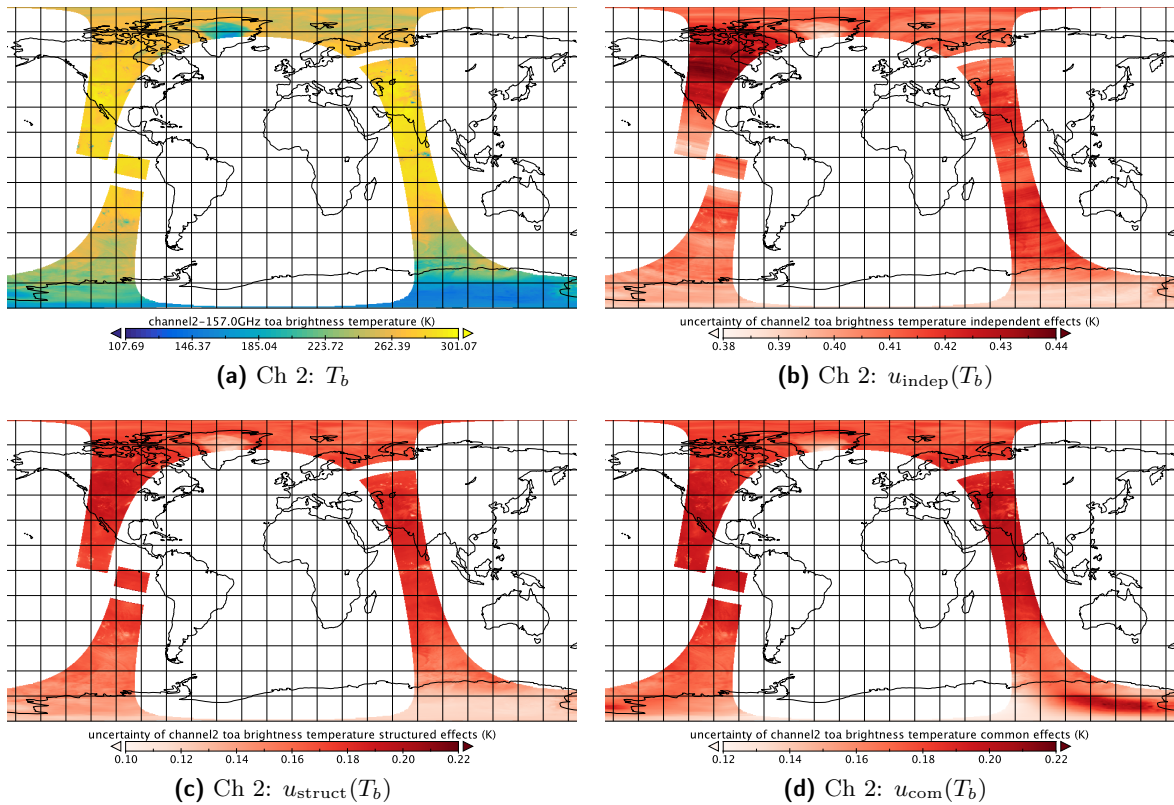
Figure 5.17 Channel 2: Brightness temperature and its uncertainties (example orbit of MHS on Metop-B).

(a) Ch 3: T_b

(b) Ch 3: $u_{indep}(T_b)$

(c) Ch 3: $u_{struct}(T_b)$

(d) Ch 3: $u_{com}(T_b)$
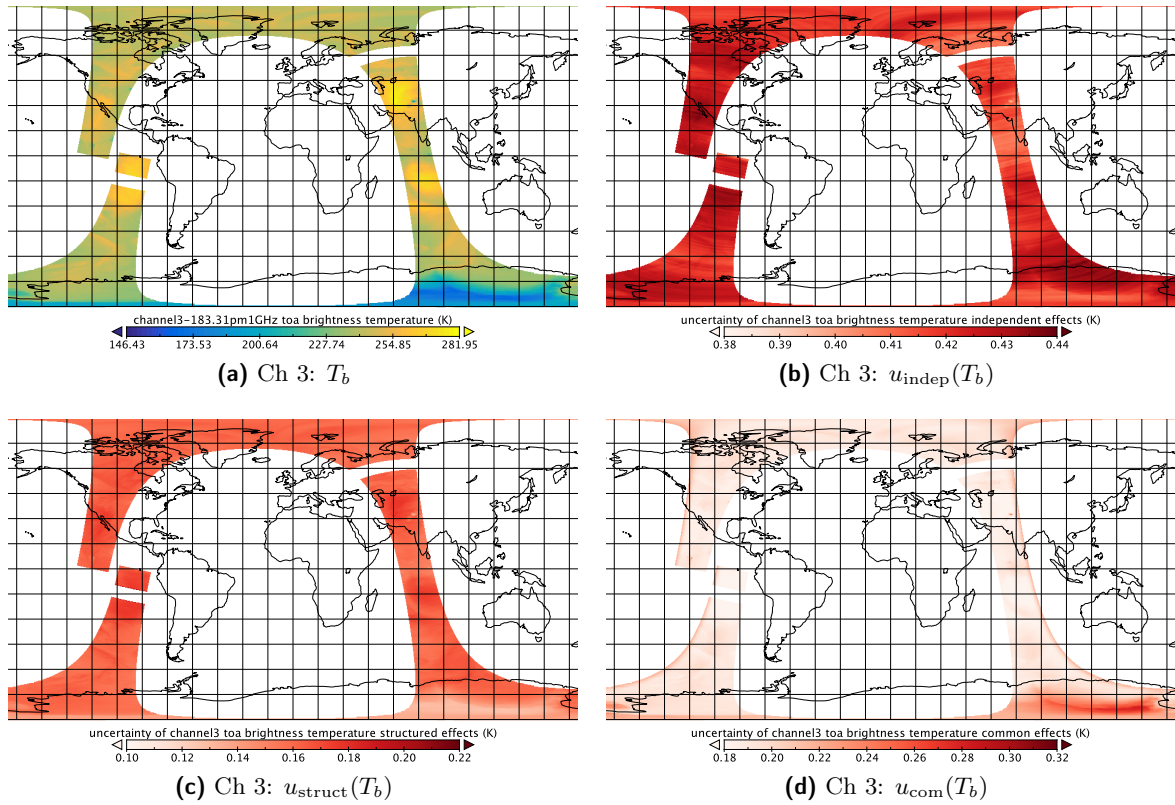
Figure 5.18 Channel 3: Brightness temperature and its uncertainties (example orbit of MHS on Metop-B).

(a) Ch 4: T_b

(b) Ch 4: $u_{indep}(T_b)$

(c) Ch 4: $u_{struct}(T_b)$

(d) Ch 4: $u_{com}(T_b)$
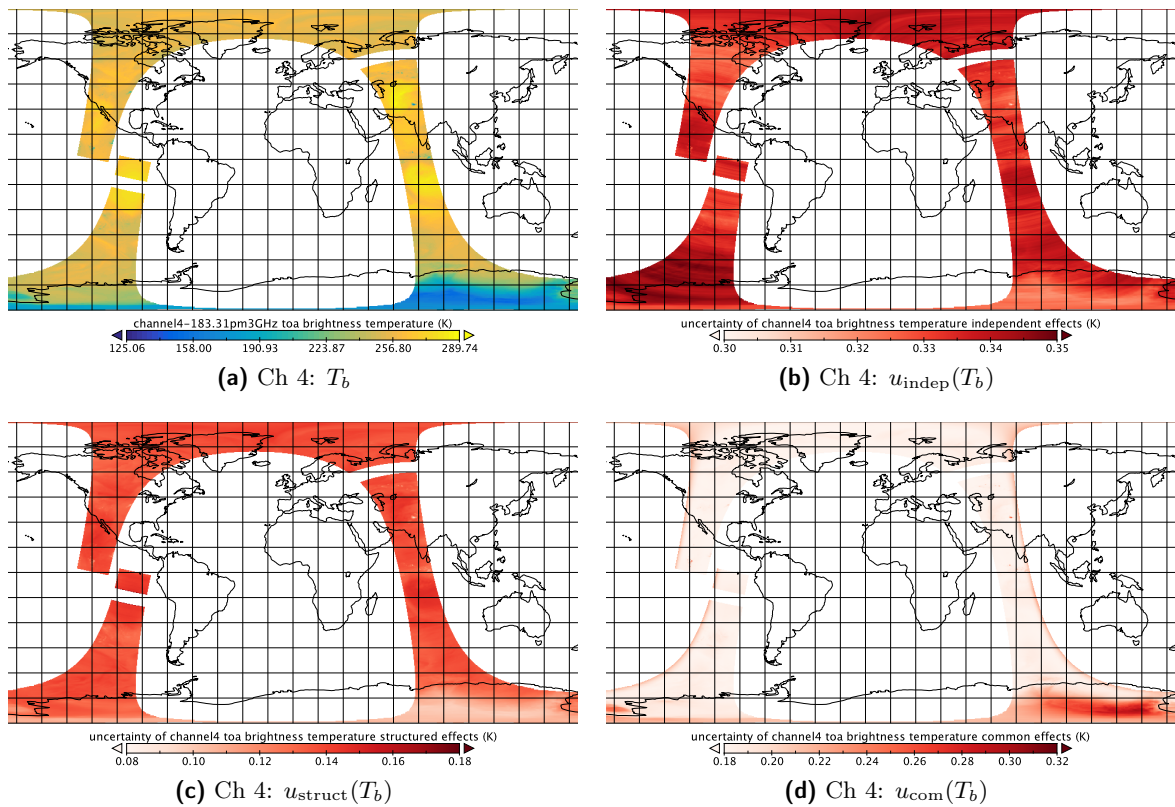
Figure 5.19 Channel 4: Brightness temperature and its uncertainties (example orbit of MHS on Metop-B).

(a) Ch 5: T_b

(b) Ch 5: $u_{\mathrm{indep}}(T_b)$

(c) Ch 5: $u_{\mathrm{struct}}(T_b)$

(d) Ch 5: $u_{\mathrm{com}}(T_b)$
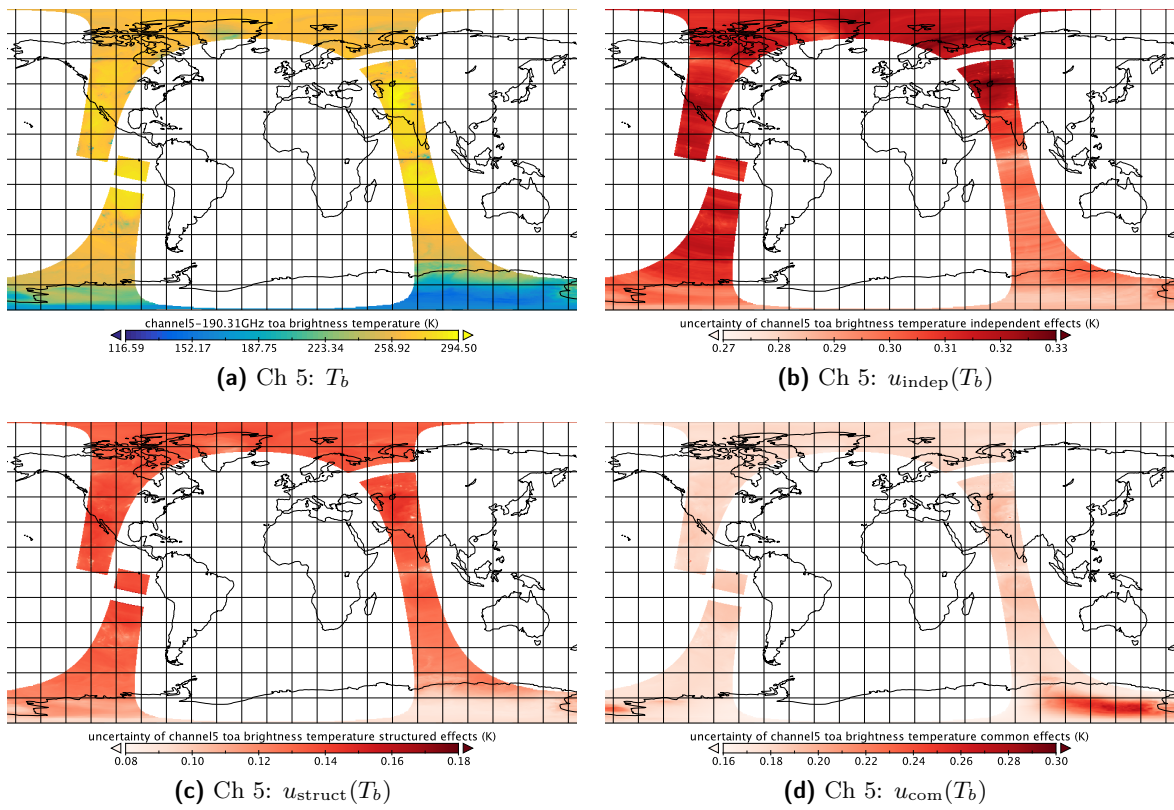
Figure 5.20 Channel 5: Brightness temperature and its uncertainties (example orbit of MHS on Metop-B).

(a) September 2014, MHS on Metop-B

(b) September 2005, AMSU-B on NOAA-15
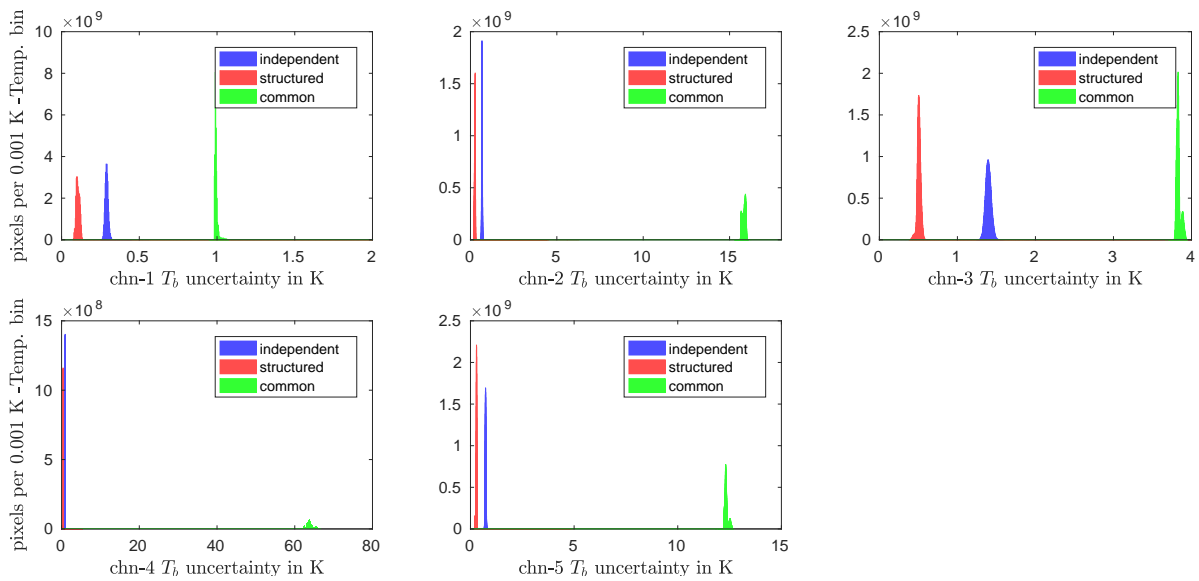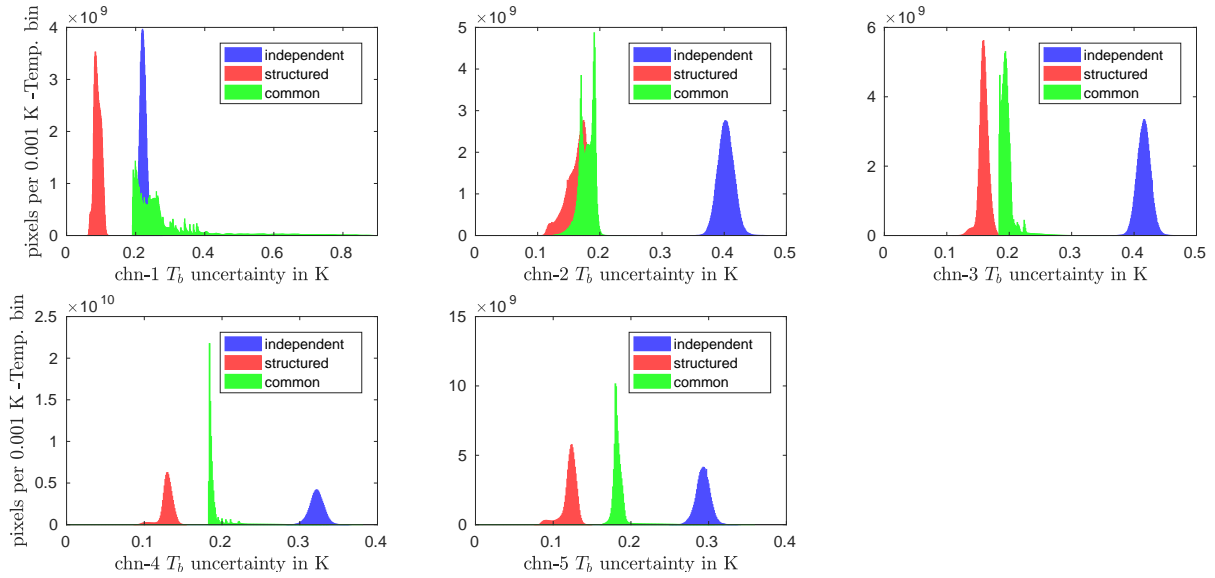
Figure 5.21 Histogram of the three classes of uncertainties for Metop-B and NOAA-15, all channels.

(a) independent

(b) structured

(c) common
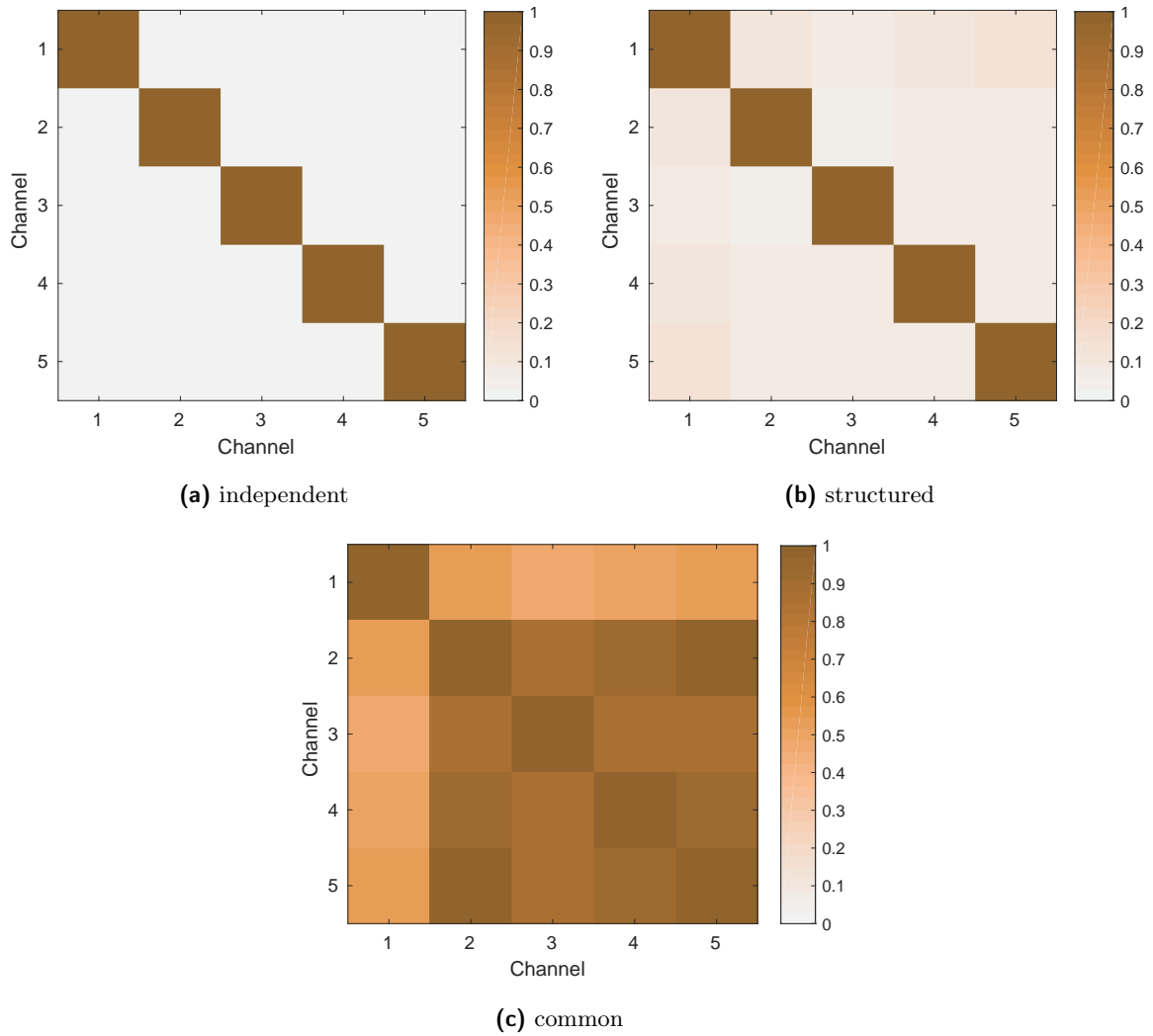
Figure 5.22 Cross-channel correlation matrices for the three effect classes. Data from example orbit of MHS on Metop-B.

5.5 Conclusion: FCDR achievements

A new MW-FCDR was generated for the eleven missions of MHS, AMSU-B and SSMT-2 instruments. The FCDR is based on the measurement equation approach and it includes the propagation of uncertainties. As mentioned above, the FCDR is not yet harmonised at this stage. This intermediate step of producing an unharmonised FCDR is important however, since it is now possible to compare the results obtained from the new FCDR processor with those from the operational AAPP processing. The comparison of the two processing procedures is required to gain confidence in the correct functioning of the new FCDR processor. This is also necessary to have a defined starting point for harmonisation: The FCDR processor must be able to reproduce the AAPP results at first in order to improve on them later.

Using the inter-comparison method of monthly percentiles described in Section 3.1.4, I compare inter-satellite biases according to AAPP-processed data with those biases obtained after FCDR processing. I use a period of six months for each satellite combination, using NOAA-18 as reference. Ideally, the FCDR processor must reproduce the same biases as seen in the operational AAPP processed data. Indeed, imitating AAPP within the FCDR processor by applying all calibration parameters as defined in the AAPP auxiliary files (`clparams.dat`, `fdf.dat`) leads to the same inter-satellite biases as they are seen in AAPP processed data. This is shown in Fig. 5.23 for the instruments on NOAA-15, NOAA-19 and Metop-A.

Note however, that the processing procedures disagree for NOAA-15 (and NOAA-17, not shown here) for channels that are affected by RFI. For these channels, namely channel 3 for AMSU-B on NOAA-15 displayed in Fig. 5.23b, an RFI correction had been devised that is applied within AAPP. The FCDR processor does not apply this correction. Therefore, the two processing methods deviate in the resulting bias. I decided to not include the RFI correction, because the RFI-correction is only valid for the very first years of the missions, but is not correct for later times any more. This is because the correction would have required updates that were not given in the past (Robel et al., 2009; Atkinson, 2018). Hence, the data do not get (fully) RFI-corrected if applying the outdated correction. I address this problem of RFI in Sec. 6.3 again, in the context of harmonisation.

Summarizing the achievements of the FCDR production so far, we have: Two of the three novel aspects of the FIDUCEO FCDR compared to available operational data or previous efforts to create an FCDR are achieved already:

1. The FCDR provides extensive uncertainty information on pixel level, required for derivation of high level products including respective uncertainty information.

2. It comes in a user-friendly format combining brightness temperature and its uncertainties, quality information bit masks and their meanings all in a single NetCDF-file per equator-to-equator frame of a whole orbit.

In addition, I applied some minor corrections in the calibration, namely the improvement on the band correction factors (Sec. 5.2.10). These do not significantly influence the final brightness temperature, but they are conceptually improved in the sense of the FIDUCEO philosophy.

Moreover, being able to reproduce the biases seen in AAPP processed data, I am now confident about the abilities of the FCDR generator and the unharmonised FCDR as starting point for the harmonisation. From this point on, actual recalibration of data by improved calibration and harmonisation is possible.

(a) AMSU-B NOAA-15, Channel 1

(b) AMSU-B NOAA-15, Channel 3

(c) MHS NOAA-19, Channel 1

(d) MHS NOAA-19, Channel 3

(e) MHS Metop-A, Channel 1
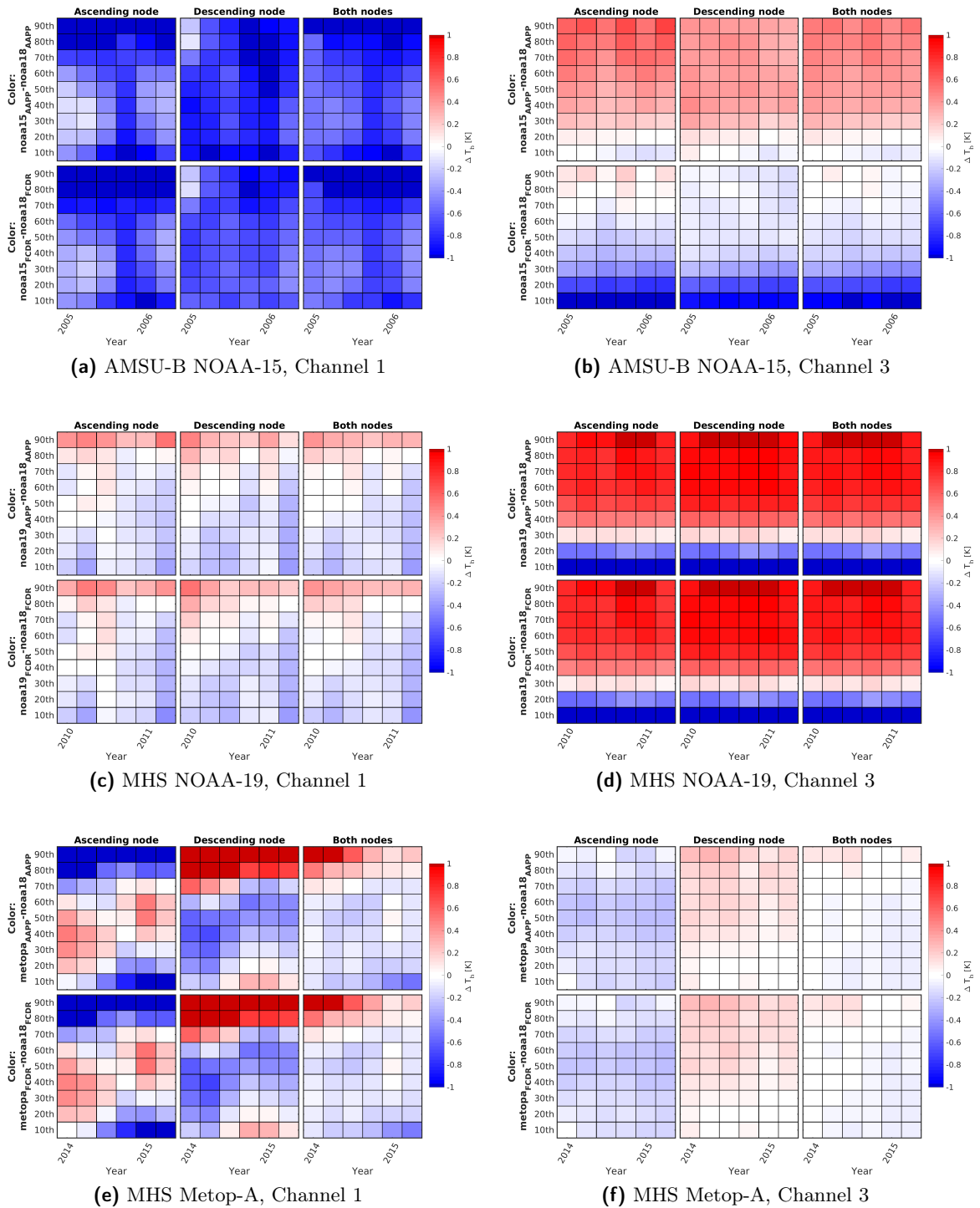
(f) MHS Metop-A, Channel 3

Figure 5.23 Inter-satellite biases against MHS on NOAA-18 computed with percentile-method. The upper row in each panel displays the AAPP-processed data, the lower one shows data processed by the FCDR generator. Data period from September to February of the indicated years. Note that the FCDR generator imitates the AAPP almost perfectly. Only (b) shows a discrepancy of the AAPP and FCDR generator. This is due to the application of the old RFI-correction scheme in AAPP. The FCDR generator does not apply the correction scheme.

6 Understanding and reducing biases: The harmonised MW FCDR

After introducing the problem of inter-satellite biases in Chapter 3 and discussing in detail the calibration procedure and instrumental effects in Chapter 5, I now address in this Chapter 6 on harmonisation the question of possible instrumental issues as causes for the biases. Moreover, I present the FIDUCEO harmonisation in a proof-of-concept way for generating a consistent, harmonised MW humidity sounder FCDR.

First, I explain the ideas and concepts behind harmonisation that form the frame, in which the FIDUCEO harmonisation is carried out. As a preparatory step towards this harmonisation, I then investigate the actual impact of the various effects onto the final brightness temperature in a sensitivity study. Further, I analyse inter-satellite biases, suggesting possible causes and deriving improvements on the calibration. The obtained information from the sensitivity study and bias analyses then enters the FIDUCEO harmonisation procedure. I use first results from this harmonisation procedure to produce an early version of a harmonised FCDR and discuss the successes, problems and possible improvements.

6.1 The idea behind harmonisation

Existing inter-satellite biases are the reason to consider harmonisation in the first place: the historical data sets from instruments of the same family, focusing on the MW humidity sounders in this thesis, do not form a stable long-term record. This means that the measurements from the instruments on board different satellites are not comparable. Comparable data for different space and time is required however, to investigate climate trends. Hence, it is required to reconcile the instruments to obtain a consistent long-term record from their combined data sets.

Within FIDUCEO, this goal is approached by harmonisation. Harmonisation is about reducing inter-satellite biases. But it only attempts to reduce biases that cannot be explained so far: for example, harmonisation accounts for the expected difference in the measurements of two sensors that have a slightly different spectral response function. That means, harmonisation does not homogenise or force the sensors to look as if they were the same, but it only reduces the bias up to the remaining difference due to that different spectral response function. Also, FIDUCEO harmonisation is no bias correction. Bias correction, as it is done for example by the NOAA (F)CDR-program described in Ferraro (2016) for AMSU-B and MHS, computes offsets or factors to be applied to the brightness temperatures to reconcile the instruments. FIDUCEO harmonisation, however, is a recalibration approach which is novel in the efforts of inter-calibration of satellite borne instruments. Recalibration means optimisation of calibration parameters within the measurement equation to obtain new recalibrated brightness temperatures that are reconciled with a chosen reference.

This optimisation of calibration parameters tries to find the best values for the parameters that reduce the bias in brightness temperature against a reference. An independent bias analysis can improve the understanding of the instrument and hence provide information to the harmonisation

process. The harmonisation process, in turn, provides information about which biases might be caused by which combination of effects represented by the optimised calibration parameters.

Of course, it should be noted that the harmonisation will be a *relative* recalibration only, since no SI-reference is available and instead one has to choose an instrument as reference. For the MW humidity sounders, MHS on NOAA-18 has been chosen as reference in FIDUCEO and in other inter-calibration efforts (Hanlon and Ingram, 2016; Ferraro, 2016) as it is relatively stable and it has the longest data record and hence overlapping periods with most instruments.

I discussed another problem in Chapter 3: There are different methods to compare the instruments, and these methods are affected by different sampling issues and therefore reveal only different parts of the true instrumental inter-satellite bias. The harmonisation uses an SNO-like method, based on "match-ups", because the result of this inter-comparison method is close to the true instrumental bias (impact of atmospheric variability is minimised by appropriate matching criteria). Moreover, the harmonisation optimises calibration parameters within the measurement equation and hence uses actual measurements for inter-comparison rather than averages of many measurements.

The FIDUCEO harmonisation procedure itself, namely the optimisation, is carried out by FIDUCEO project partners from FastOpt GmbH and National Physical Laboratory (NPL). Figure 6.1 shows the production flow of the FIDUCEO FCDRs. The intermediate result of unharmonised FCDR solely relies on the input level 1b data, the uncertainty estimates and the FCDR production code. The harmonised FCDR is produced similarly, however, some calibration parameters will be optimised and hence enter the measurement equation with a new value. For input, the harmonisation process requires the match-up data which it should compare. Further it needs to know the calibration mechanism, namely the measurement equation, and lastly, the optimisation parameters need to be chosen from the set of calibration parameters in the measurement equation. The choice of optimisation parameters requires ideas from the analysis of the instruments characteristics, the bias and possible causes. I will discuss this in Sec. 6.2 and 6.3. As output, the harmonisation procedure provides a set of optimised calibration parameters. These I put back into the measurement equation in the FCDR generator to produce a harmonised FCDR.

6.2 Sensitivity study on the individual effects

In order to obtain information on the actual impact of the various effects on the final brightness temperature, I analyse the sensitivity of the brightness temperature to changes in the corresponding parameters of certain effects that are modelled in the measurement equation. Some of these parameters may be possible candidates for optimisation parameters in the harmonisation. Therefore, it is interesting to gain an impression of their potential impact onto the brightness temperature and to "simulate biases" that the parameters can introduce. This is also important for validation of the harmonisation procedure: knowing the impact of changed parameters can hint at whether a suggested change by the harmonisation seems reasonable. This is especially important for temperature ranges that are not well represented in the harmonisation input data based on the SNO-like match-up data set.

The effects analysed in this sensitivity study are the antenna pattern correction, the polarisation correction, the non-linearity, the cold space bias correction and the warm target bias correction. For each of these effects, I use the operational value for the corresponding parameter X as reference value X_{ref}. Then, I vary the parameter X in the scheme $k \cdot X_{\mathrm{ref}}$ with
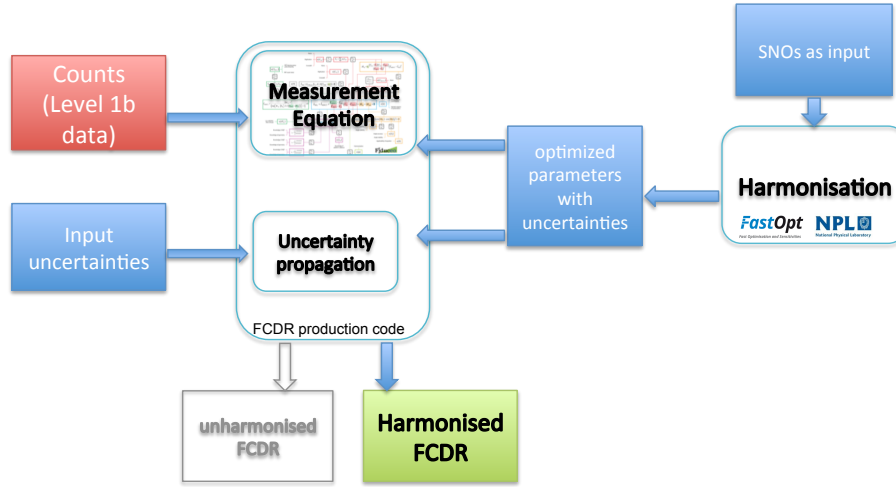
Figure 6.1 Production flow of the FIDUCEO FCDR. For the unharmonised FCDR (grey), only the left hand side input to the FCDR generation code is needed (data set presented in Sec. 5.4). The production of the harmonised FCDR additionally uses the optimised calibration parameters obtained from the harmonisation process carried out by FastOpt GmbH and NPL.

$k \in [0, 0.25, 0.5, ...2, 3, 4]$ corresponding to eleven values for the parameter ("value1 ... value11" in the plots). For the antenna pattern correction, that actually consists of three parameters $g_E, g_S, g_{\mathrm{Pl}}$ (only two are independent), I use the following scheme: $g_E = 1 - k \cdot g_S - k \cdot g_{\mathrm{Pl}}$, with $k \in [0, 0.25, 0.5, ...2, 3, 4]$. For effects that have a reference value of zero (for certain channels), I use the reference value from another instrument/ channel, to vary a non-zero value. For each realisation of the parameter X for a certain effect, I compute the brightness temperatures $T_{\mathrm{b,X}}$ in an example orbit. Then, I compute the differences in the brightness temperature compared to the brightness temperature for the reference value $T_{\mathrm{b,ref}}$. This difference ΔT_b is plotted as function of the reference brightness temperature for FOV 1, 15, 30, 45, 60, 75 and 90. As example, I used one orbit of AMSU-B on NOAA-17 and one orbit of MHS on Metop-A. Figures 6.2 to 6.6 show the results of the sensitivity study for channel 3 of Metop-A.

The only effects depending on FOV are the antenna pattern correction and the polarisation correction (Fig. 6.2, 6.6). However, their dependence is very different. The antenna pattern correction has stronger impact on the outer FOVs than on the inner FOVs. However, the effect is not symmetric around nadir. For example, FOV 90 can be more affected than FOV 1, although they have the same absolute value of the viewing angle. This is due to the asymmetric environment of the instrument that affects the radiation in the side lobes. The polarisation correction is completely symmetric around nadir, but it has the opposite dependence: the outer FOVs are less affected than the inner FOVs. This is because the inner FOVs require a larger correction to account for the different reflectivity of the mirror for horizontal and vertical polarisation and for the polarisation sensitivity of the receiver.

The impact of the effects also differs in its dependence on brightness temperature. The non-linearity shows a quadratic dependence on brightness temperature (Fig. 6.4). This is of course by construction, since the non-linearity was modelled as a second order polynomial. All other effects show a linear dependence on the brightness temperature. However, the effects differ in their maximum/ minimum impact range. Strong impact on the cold end and weak impact on the warm end is shown by the cold space correction and the polarisation correction (Fig. 6.3,
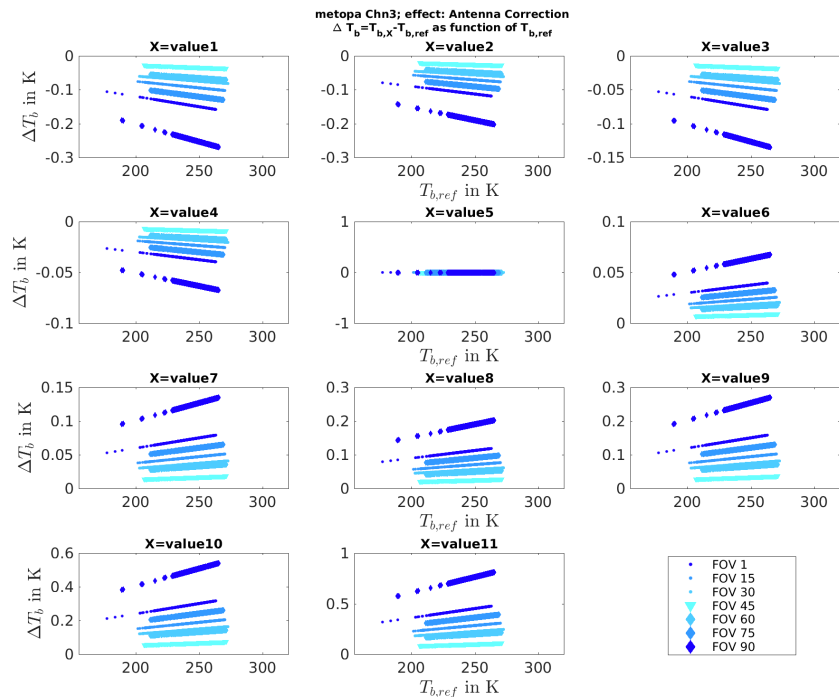
Figure 6.2 Sensitivity study on antenna pattern correction for MHS on Metop-A. Value1 to Value11 indicate the increase of factor k (see text). Value5 corresponds to $k = 1$, i.e. the operational reference value. Accordingly $\Delta T_b = 0$ in this case.
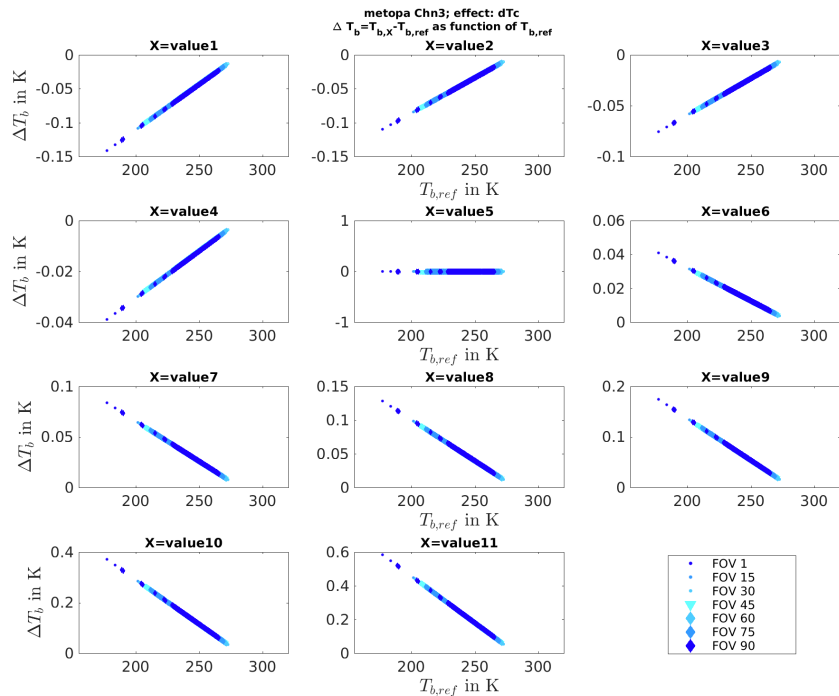
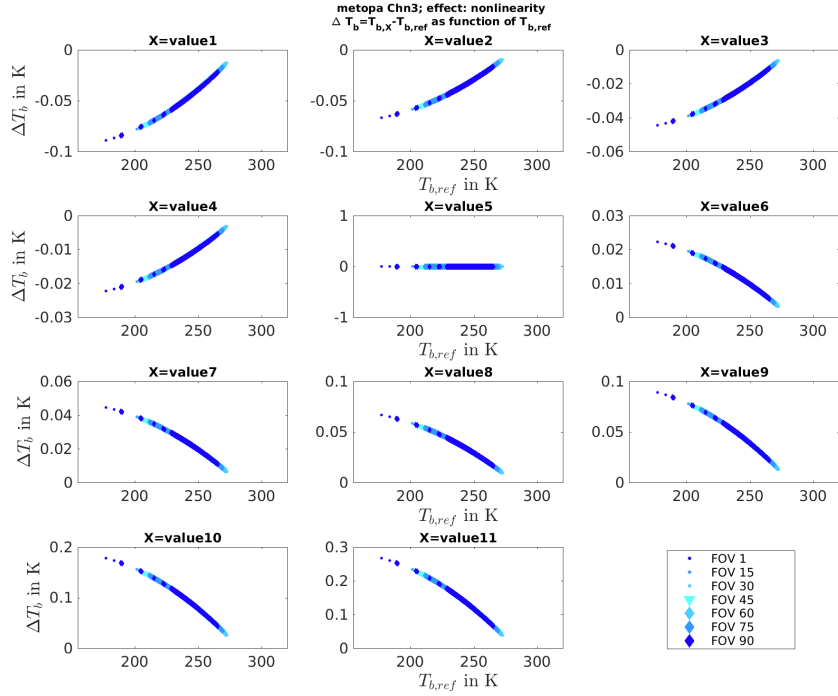

Figure 6.3 As Figure 6.2, but for the cold space correction.

Figure 6.4 As Figure 6.2, but for the non-linearity correction.

6.6). The opposite behaviour is observed for the antenna pattern correction and the warm target correction shown in Fig. 6.2 and Fig. 6.5, respectively. Of course, quantitative differences make some effects stronger than others for the same relative change of $k \cdot X$. Most effects show a significant impact at either the cold or the warm end of the temperature range. The warm target correction, however, has a significant impact over the whole covered brightness temperature range such that it also affects the rather low brightness temperatures (see Fig. 6.5).

The linear dependence that is seen for most of the effects is quite weak. That means, if one looks at a relatively small range of brightness temperatures in a bias analysis (which often is the case due to the distribution of brightness temperatures), one does not necessarily detect a linear bias, because the range is simply too small to make the linearity visible. Moreover, a true bias may result from a combination of wrongly corrected effects. This might result in a complicated bias dependence on the brightness temperature. Furthermore, the FOVs may be affected differently and sampling issues may also add to the biases, thus making the biases more complex. Hence, the results from the sensitivity study cannot easily point out reasons for biases by comparing the sensitivity-plots to plots of brightness temperature dependent biases. However, the results from the sensitivity study help assessing the impacts of the effects and advise the harmonisation procedure.

In principle, all effects can be possible candidates for optimisation parameters. Yet, not all effects are equally easy to treat as an optimisation parameter. Moreover, not all effects have enough potential to account for biases without changing their parameter's value dramatically.

The antenna pattern correction is a valid candidate for optimisation, since there are some inconsistencies in the operational values. The sensitivity to possible changes as I show them here, however, only cover the aspect of overall change of the correction by a factor, as I did not change the qualitative pattern of the correction. Nonetheless, it is possible that the geometrical maps and
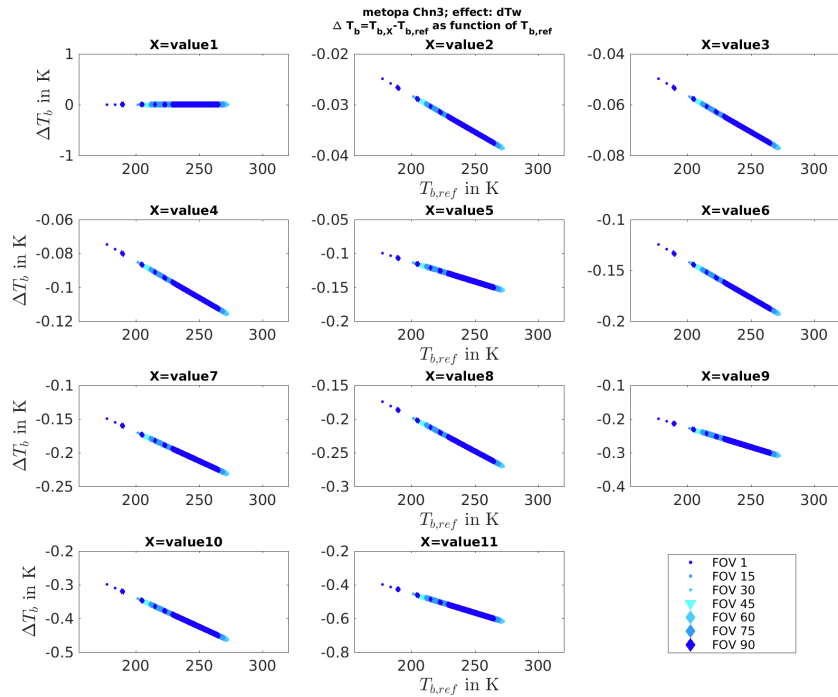
Figure 6.5 As Figure 6.2, but for the warm target correction. Note that the operational reference value is zero for this effect (value1). To vary the parameter in multiples of k, I used the only non-zero value given for this effect, taken from AMSU-B on NOAA-17 for channel 5 ($\delta T_{ch} = -0.16$ K). This value is reached for value5, i.e. $k = 1$.

the antenna patterns from which the correction was deduced are not correct for all instruments. A major drawback of using the antenna pattern correction for harmonisation is its complexity. It would be necessary to optimise each contribution from of Earth, space and platform g_E, g_S, g_{Pl} for each FOV, adding up to 270 optimisation parameters (only 180 are independent though, since $g_E + g_S + g_{Pl} = 1$). To avoid this large number of optimisation parameters, it is possible to use polynomial fits to approximate the antenna pattern correction. The coefficients of these polynomial fits would serve as optimisation parameters in this approach. However, this approach will introduce errors inherent to itself. Hence, in the very first harmonisation runs as I present them in this thesis, the antenna pattern correction is not yet used.

A parameter that is very easily accessible is the cold space correction ΔT_c (labelled as a_3 in its status as harmonisation parameter in the measurement equation diagram, Fig. 5.1). However, the sensitivity to this correction is quite small in the temperature range where most observations are made (around 245 K for channel 3). Therefore, to account for biases of only 0.2 K in this temperature range, changes of four times the operational value are needed (see Fig. 6.3). This corresponds to a correction of several Kelvin instead of a fraction of a Kelvin. Changes like this do not seem very likely to be the correct modification of the calibration. First tries in optimising a_3 with the harmonisation procedure did not produce any useful results. Harmonisation suggested huge changes of the cold space correction parameter. Hence, for further harmonisation attempts it is not considered at first.

The non-linearity is also a very small effect that cannot account for large biases, unless its value is changed dramatically. Nonetheless, it is used as an optimisation parameter already in
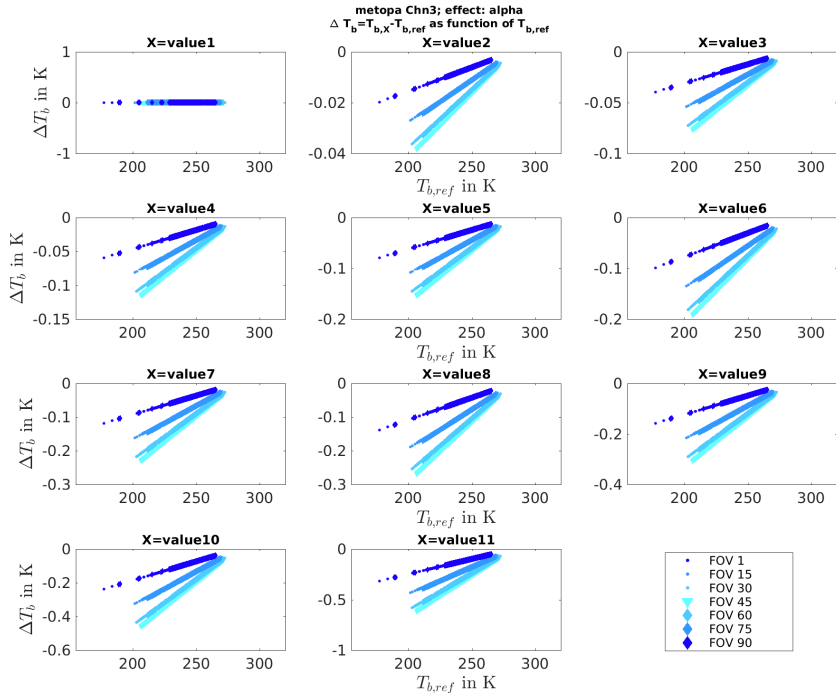
Figure 6.6 As Figure 6.2, but for the polarisation correction. Note that the operational reference value is zero for this effect (value1). To vary the parameter in multiples of k, I used the only non-zero value given for this effect, taken from MHS on NOAA-18 ($\alpha = -0.0022$). This value is reached for value5, i.e. $k = 1$.

the first harmonisation attempts to allow for non-linear corrections at all. Higher order terms are not considered here.

The warm target correction has a large effect in the usual temperature range (changes of $0.2\,\mathrm{K}$ are easily realised by $k = 1.5$, see Fig. 6.5). For unknown reasons, only channel 5 of AMSU-B on NOAA-17 has an operational non-zero value of -0.16 K. All other instruments and channels get zero (this is why I used the NOAA-17 value as X in all $k \cdot X$). Because of this unexplained fact, together with the possible impact of the correction, the warm target correction parameter δT_{ch} is used as optimisation parameter. It is mainly interesting for the harmonisation of the MHS on the Metop satellites and the AMSU-B on NOAA-17. For the AMSU-B on NOAA-15 and NOAA-16, as well as for the MHS on NOAA-19, the correction is still too small to account for such large biases with erratic time dependence as seen for these instruments in Chapter 3 (see Sec.6.3 for further analysis of these biases). Since the warm target bias correction encompasses all possible influences of the surroundings of the black body, it can be argued that the correction should be time dependent to allow for heating and cooling of the black body shroud and environment with the orbital movement. It is not clear however, which of the provided temperature measurements from the inner parts of the instrument are important for such an additional heating effect in the surroundings of the black body. The temperatures provided in the level 1b file cover only the local oscillator, the mixer, the amplifier, the sub-reflector and the "PRT board" temperature (it is not documented where these temperature sensors are actually located). None of these temperatures appears as a useful proxy for the temperature of the shroud for example. Hence, one can assume a combination of several temperatures and optimise coefficients for each temperature to find a

time dependent δT_{ch}. This is not yet part of the harmonisation attempts. So far, the warm target correction is treated as constant for the mission.

The calibration parameter α for the polarisation correction is non-zero only for MHS on NOAA-18 within the operational parameter set of AAPP. Hence, this value is used in the sensitivity study as X in all $k \cdot X$. The polarisation correction impact varies with the polarisation (horizontal, vertical) of the channel's receiver. For vertical polarisation channels (all channels in AMSU-B, channels 1, 2 and 5 in MHS), the polarisation correction parameter α (or a_2, in its role as harmonisation parameter) would be positive and produces a warm correction (for Earth temperatures smaller than the black body temperature, see Eq. 5.46). For horizontal polarisation channels (channel 3 and 4 of MHS), the situation is reversed (see Fig. 6.6): a_2 is negative and produces a cold correction. The impact increases for cooler brightness temperatures and close to nadir. For channel 3, around $245\,\mathrm{K}$, changes of $0.1\,\mathrm{K}$ at nadir are already reached for $k = 1$ (i.e. using $a_2 = -0.0022$ from NOAA-18). Hence, this effect is strong and can generate significant biases in the case of a wrong value for α. Wrong values for α are likely for some instruments, since only MHS on NOAA-18 actually receives a non-zero correction for this effect, although a characteristic bias indicating a required correction was observed for AMSU-B in pre-launch tests already (see Sec.5.2.7). Hence, it is highly likely that further instruments aside from MHS on NOAA-18 require a polarisation correction. Therefore, the polarisation correction a_2 is an optimisation parameter that is determined within the second attempts of harmonisation.

6.3 Bias analysis and improvement of calibration

6.3.1 RFI related biases

The strongest biases against MHS on NOAA-18 in the most relevant water vapour channel 3 at 183 ± 1 GHz are observed for the AMSU-B on NOAA-15 and NOAA-16 and for the MHS on NOAA-19. In all three cases, an increased noise appears at the same time. Due to this high noise ($>1\,\mathrm{K}$), most of the periods may be excluded, if only low noise is acceptable for users. However, climate studies use averaged data, and because noise belongs to independent uncertainties that are reduced in averaging procedures, noise is not critical for climate research as long as the time series is long enough. Therefore, it is not necessarily required to exclude data with increased noise from climate studies. Consequently, these periods of increased noise must not be left aside, but they have to be included in an analysis to understand and possibly reduce the strong biases occurring in these periods.

I previously indicated the relation of gain decrease and bias increase for NOAA-15, NOAA-16 and NOAA-19 and identified a possible cause to be RFI (see Sec. 3.1.1, 4.1, 5.2.9). Here, I investigate this issue in more detail.

As explained in Sec. 5.2.9, the RFI is a scan-dependent effect that impacts the back-end of the receiver, adding an extra amount of signal to the recorded signal. Hence, if the gain is low, this extra amount has a relatively stronger impact onto the overall signal than if the gain was high. This mechanism can lead to biases caused by RFI at decreased gain for instruments that showed no RFI at the beginning. Because their gain was at sufficiently high level, any RFI impact was not detectable. This was already suggested by John et al. (2013b), however, without reporting consequences or any further investigation. In the following, I discuss this RFI issue for all instruments and I draw conclusions for the harmonisation of the affected instruments.

I start with a closer look on the AMSU-B on NOAA-15, which has known issues with RFI. Apart from the gain decrease – bias increase relation, there is a further indication that RFI

(a) NOAA-15 **(b)** NOAA-16

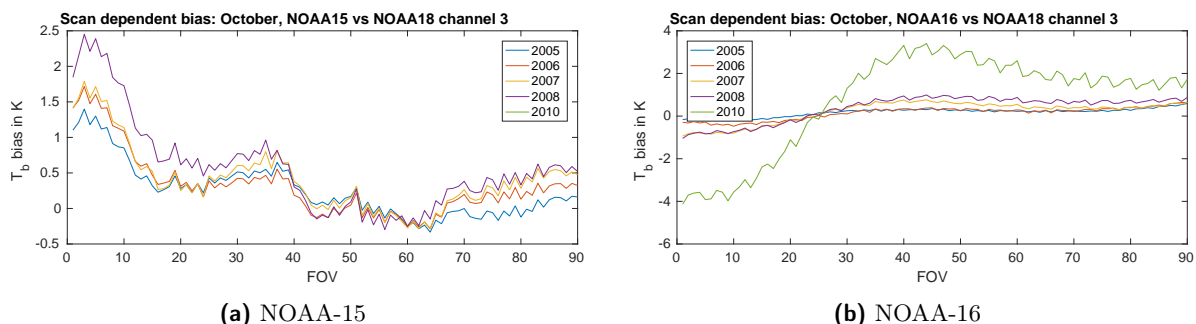**Figure 6.7** Scan-dependent bias against MHS on NOAA-18 for channel 3. Monthly means, all latitudes.

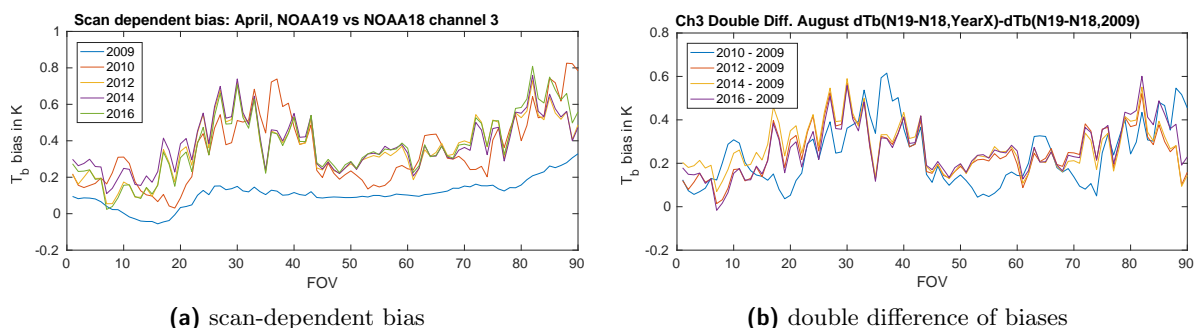**(a)** scan-dependent bias **(b)** double difference of biases

Figure 6.8 Scan-dependent bias of MHS on NOAA-19 against MHS on NOAA-18 for channel 3. Monthly means, all latitudes. The figure shows the scan-dependent bias for April in several years (a) and the difference of the biases in later years compared to 2009 (b).

causes the bias, namely the strong scan dependence of RFI. At the beginning of the NOAA-15 mission, strong scan-dependent biases were seen and attributed to RFI (see Chapter 3). These scan-dependent biases were also observed by John et al. (2013b). Taking a closer look at the scan-dependent biases as they were observed by Hanlon and Ingram (2016) using zonal averages, a zigzag pattern, varying from one FOV to the next can be seen, superimposed with a long-range pattern varying slowly over the FOVs. What is important about the zigzag pattern, is its stability. The zigzag pattern occurs for all latitude bins and was obviously not averaged out completely over the long time period the plots were made for (several years). Hence, there must be an underlying systematic effect. Producing similar plots, but for monthly means of all latitudes, I observe the same patterns. Figure 6.7a shows biases of NOAA-15 channel 3 against MHS on NOAA-18, that means NOAA-15 – NOAA-18, from monthly means for October of the years 2005 to 2008. The stability of the pattern is obvious. This cannot be noise or any natural variability effect. Note also that the intensity of the peaks and the smooth long-range pattern grows over the years. This evolution corresponds to the decreasing gain (see Fig. 4.5a).

A similar picture is obtained for NOAA-16 (see Fig. 6.7b). A stable pattern of peaks superimposed with a smoother pattern is visible for all years, intensifying with time as the gain decreases (see Fig. 4.5b).

The third instrument with significant gain decrease and bias increase is MHS on NOAA-19. The scan-dependent bias shown for April of several years of the period 2009 to 2016 in Fig. 6.8a supports the theory of an impact of RFI on the instrument. In 2009, before the problems and

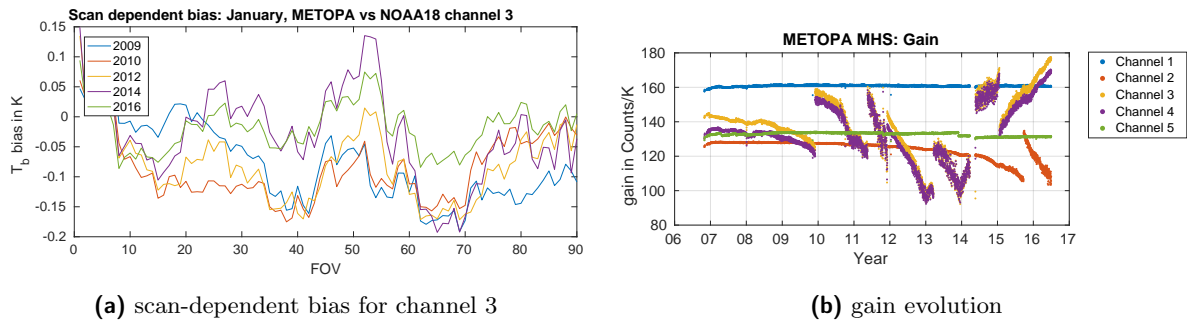**(a)** scan-dependent bias for channel 3

(b) gain evolution

Figure 6.9 Scan-dependent bias of MHS on Metop-A against MHS on NOAA-18 for channel 3, monthly means for January, all latitudes shown in (a). Gain evolution for all channels shown in (b).
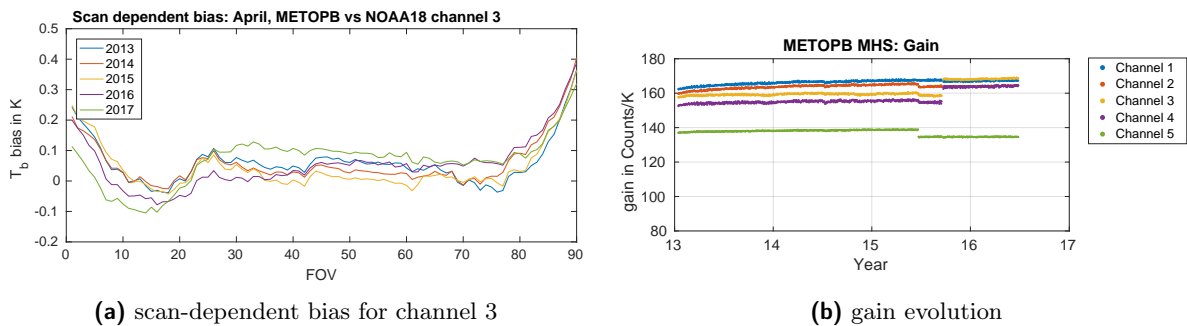

(a) scan-dependent bias for channel 3

(b) gain evolution

Figure 6.10 As Figure 6.9, but for Metop-B and month April.

the associated gain decrease started for NOAA-19 (see Fig. 4.5e), there is only a faint zigzag pattern. Note that the smooth pattern with a dip around FOV 15 may be related to a different issue on the reference instrument on NOAA-18. This issue introducing the smooth pattern is constant over the years and therefore also affects the later years shown in 6.8a (see Sec. 6.3.3 for the discussion of the issue). From 2010 on, when the gain has decreased significantly, a strong zigzag pattern is visible. Between 2010 and 2012, the pattern changes slightly for several FOVs, but it remains stable afterwards. A clear picture of the pattern can be obtained when subtracting the scan-dependent biases of NOAA-19 vs NOAA-18 in the year 2009 (showing a faint zigzag pattern only) from the scan-dependent biases in the years from 2010 and later. This double difference is displayed in Fig. 6.8b. The figure shows the remaining zigzag pattern which emerges by removing the background bias, i.e. the constant smooth pattern. To some degree, this remaining zigzag pattern even shows a regularity resembling an interference pattern that the rotating mirror scans through.

Even for the MHS on Metop-A, a slight zigzag pattern can be observed in this analysis of monthly means. Figure 6.9a shows the scan-dependent bias for January of 2009 to 2016. Note that the peaks only reach up to 0.15 K and down to -0.2 K, which is significantly less than the peaks observed for NOAA-15, NOAA-16 or NOAA-19. Nonetheless, the pattern changes marginally with the gain: the differences for adjacent FOVs are smaller in January 2016 (relatively high gain) compared to those in January 2014 (relatively low gain, see Fig. 6.9b). Since this coincides again with the gain evolution, it is possible that the data of MHS on METOP-A may be affected by RFI, if the gain decreases strongly in the future.
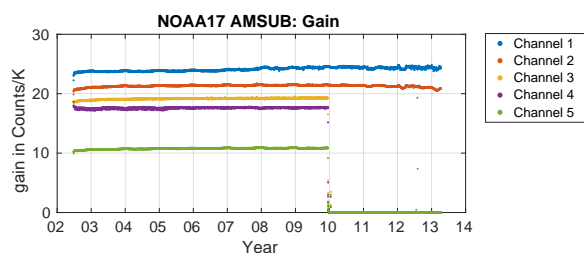
Figure 6.11 Gain evolution for AMSU-B on NOAA-17.

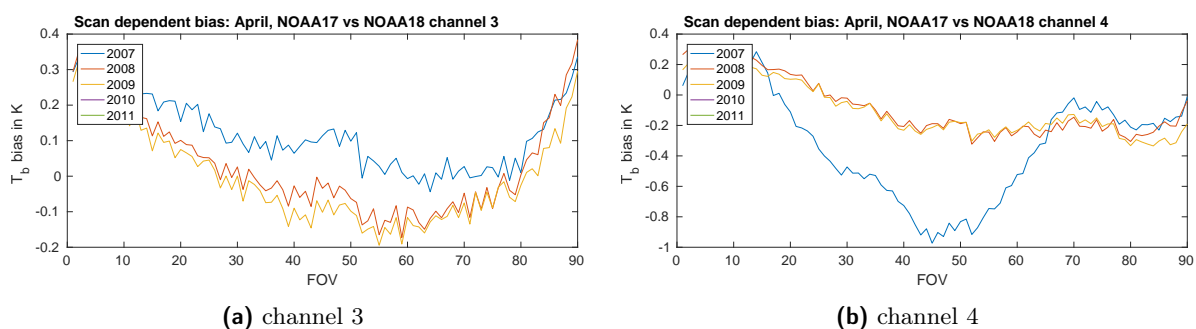**(a)** channel 3

(b) channel 4

Figure 6.12 Scan-dependent bias of AMSU-B on NOAA-17 against MHS on NOAA-18. Monthly means, all latitudes. No data exist for AMSU-B on NOAA-17 in 2010 and after due to instrument failure in the sounding channels (see Fig. 6.11).

The MHS on Metop-B also shows a persistent slight peak pattern that may be assigned to RFI (see Fig. 6.10a). Over the years, the intensity of the pattern changes marginally. As the gain is also relatively constant, with a small increase for channel 3 in late 2015 only (see Fig. 6.10b), a correlation between the intensity of the pattern and the gain evolution can be neither confirmed nor excluded. It should be noted that the smoothly varying scan-dependent bias pattern with a dip around FOV 15 may be related to the issue on NOAA-18, that I discuss in Sec. 6.3.3.

Finally, I investigate the scan-dependent bias for AMSU-B on NOAA-17 and possible relations to RFI. It should be noted that channel 4 has known issues with RFI and that a correction scheme was devised for that channel (Atkinson, 2002). As mentioned earlier, this correction scheme is not necessarily valid for later years, since the absolute impact of RFI might change over time. This was observed in the early years of AMSU-B on NOAA-15 when different RFI modes appeared which changed the bias pattern (Atkinson, 2001). Consequently, a scan-dependent bias induced by RFI may vary over time due to two reasons. First, the absolute impact of RFI may change, making a previously devised correction scheme invalid. This may lead to a qualitative change of the scan-dependent bias pattern. Second, the relative impact of RFI may change due to a varying gain. This may lead to an intensification of a qualitatively unchanged pattern. While the latter reason may be true for NOAA-15, NOAA-16 and NOAA-19, it does not hold for NOAA-17. NOAA-17 has a very stable gain in channels 3-5 (until complete failure in 2010, see Fig.6.11). An intensification of the pattern is thus not expected here. A qualitative change of RFI, however, is suggested by Figure 6.12. Figure 6.12a shows the scan-dependent bias in channel 3 for April in 2007, 2008 and 2009. While the pattern is very similar for 2008 and 2009, it is qualitatively different in 2007. A similar situation is observed for channel 4 (see Fig. 6.12b).

Here, the qualitative change from 2007 to 2008 is even more pronounced. The mean bias over all FOVs changes from a negative value to a mean bias close to zero. This change of bias against NOAA-18 is also observed in the time dependent bias analysed in the global averages in Fig. 3.6d: while the bias is negative in April 2007, it is close to zero in April 2008. In summary, a qualitative change of absolute RFI impact on channel 4 (and to less extent in channel 3 and 5), may have first resulted in scan-dependent biases as seen in 2007. The biases may emerge as they cannot be corrected for by the old correction scheme which was constructed for bias patterns as they occurred in the early years of the instrument. Further changes of RFI after 2007 may have caused the change of scan-dependent bias as seen in Fig. 6.12b from the year 2007 to 2008. This discussion shows the particular difficulty in correcting RFI related biases which may change erratically in time.

Concluding the above analysis of the time evolution of scan-dependent biases, it is highly probable that channel 4 of AMSU-B on NOAA-17 is affected by qualitative absolute changes of RFI (the channels 3 and 5 are affected to less extent), whereas relative changes of RFI due to gain variations are more dominant for channel 3 of the AMSU-B on NOAA-15, and on NOAA-16 and for the MHS on NOAA-19 (also for channel 4 and 5 for the AMSU-B on NOAA-15 and on NOAA-16 and channel 4 for the MHS on NOAA-19, showing similar behaviour). This theory of RFI related biases for NOAA-15, NOAA-16 and NOAA-19 is supported by the facts that firstly, RFI was confirmed for NOAA-15 by Atkinson and McLellan (1998); Atkinson (2001), and secondly, that other instruments show persistent zigzag bias patterns over the scan, correlated with a degradation of the gain – a behaviour which is very similar to the behaviour of NOAA-15. This bias pattern, highly variable from FOV to FOV, but stable at least over certain time periods, cannot be produced by noise or any natural variability effect, nor is there any known systematic effect that could produce such a scan-dependent bias pattern. Hence, I suggest the interference pattern due to RFI as most convincing explanation for these scan-dependent biases. Also, it seems likely that all microwave humidity sounders on board the NOAA and Metop satellites are affected by RFI, but the bias only becomes distinctly visible if the gain decreases.

In order to reduce the strong biases for the NOAA-15, NOAA-16 and NOAA-19 instruments, the harmonisation procedure cannot be applied easily. Of course, an extra optimisation parameter can be introduced in the measurement equation to account for RFI, but this parameter would merely be an additive number ΔC_E to the Earth/ IWCT/ DSV counts to account for the extra amount of signal. This ΔC_E however, cannot be optimised easily by harmonisation to reduce the inter-satellite bias. This is because it is completely different for all FOVs and it can even change erratically in time (change of absolute RFI impact as for NOAA-17 and early NOAA-15). Allowing for this behaviour means allowing ΔC_E to have a different value for all FOVs and times, which results in the optimisation solver being allowed to do anything to make the measurements fit. This would effectively work as "optimising the counts themselves" to reduce the bias. This is actual manipulation of the very measurement and no sound recalibration. To properly use harmonisation on this effect requires the definition of strict constraints on this ΔC_E for each FOV and each time period where the RFI is believed to be constant. This can prevent the optimisation machine from over-correcting the actual measurements.

In order to reduce the bias caused by RFI even if the current harmonisation cannot, more sophisticated methods for analysing the bias and finally devising possible corrections need to be applied. An in-depth analysis of the time and scan-dependent biases to filter out the interference pattern could be a strategy to obtain a correction scheme for each period where RFI seems stable. Gain changes should not be a problem, since they automatically scale the correction, if a devised correction ΔC_E in terms of counts would be implemented in the measurement equation.

(a) Metop-A channel 3

(b) Metop-A channel 4
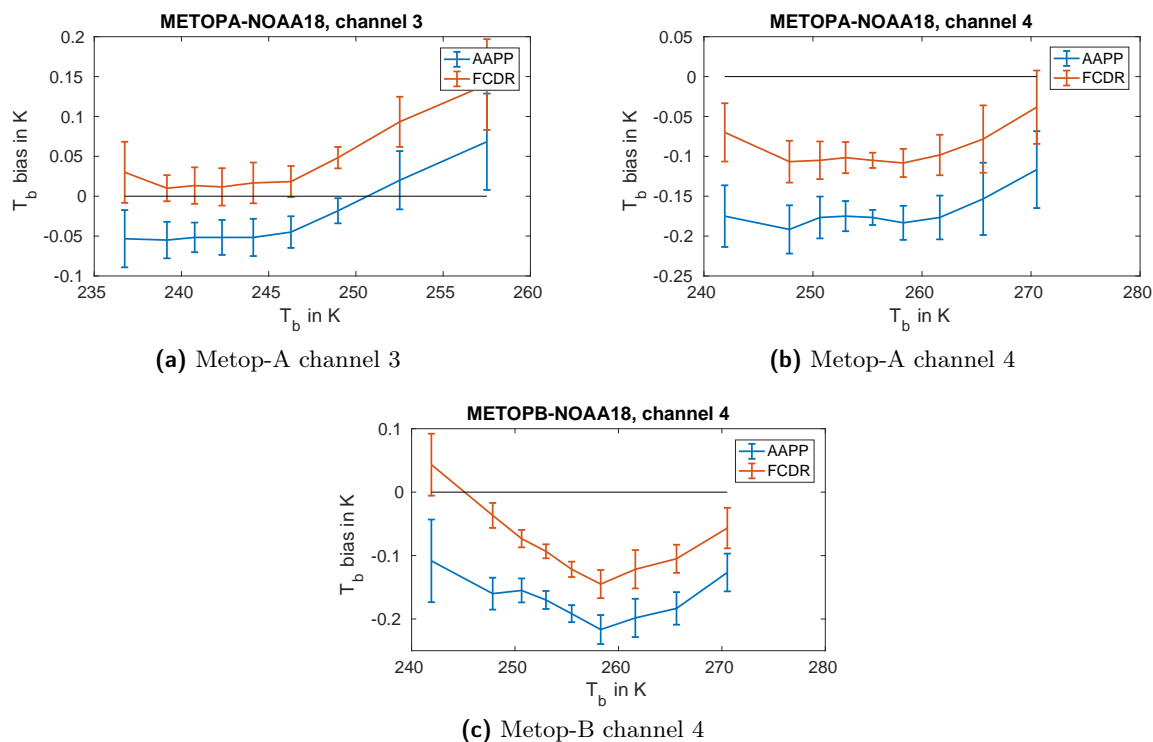
(c) Metop-B channel 4

Figure 6.13 Comparison of resulting biases for AAPP and FCDR processing, using the correct antenna pattern correction in the FCDR processing.

Having in place this correction for each FOV and period of constant RFI, possible ranges of allowed change could be formulated. This would provide the required constraints for applying a subsequent harmonisation procedure, which could then further refine the RFI correction if necessary. This is interesting work, but beyond the scope of this thesis.

6.3.2 Improvement of antenna pattern correction

An improvement of the operational calibration by AAPP can be achieved by correcting for obvious mistakes. The assignment of antenna pattern correction parameters to the different channels is incorrect for AMSU-B (see Sec. 5.2.1). As a consequence, it is also wrong for MHS on NOAA-18, which only receives the antenna pattern correction of AMSU-B. If AAPP is used in its default configuration, the antenna pattern correction is wrong even for Metop-A. This is because the default version of antenna pattern correction for Metop-A is copied from NOAA-18. As pointed out in Sec. 5.2.1, channel 4 and 5 receive the same correction as channel 1 and 2. However, they should get the same correction as channel 3. This is the first change that I apply for improvement in my FCDR generator code. This affects all AMSU-Bs and the MHS on NOAA-18. For Metop-A, I use the new version of correction parameters as a second change for improvement. To demonstrate the improvement only due to correcting for this wrong assignment of antenna pattern correction, I do not change the polarisation correction for Metop-A, but still use the default values copied over from NOAA-18 (see Sec. 5.2.7). For September 2014 to February 2015, I reprocess the NOAA-18 and Metop-A measurements to obtain new brightness temperatures, based on the corrected assignment of antenna pattern correction. I compute the 10th to 90th

percentile and the average for each percentile over these six months. The corresponding standard deviation is used as uncertainty estimate. Figure 6.13 shows these averages and demonstrates the performance of the AAPP compared to the FCDR with corrected assignments. For channel 3 (Fig. 6.13a), a slight improvement can be achieved in the lower percentiles up to 246 K. The uncertainty bars include the zero-bias line for this brightness temperature range. For the 70th to 90th percentile however, AAPP provides a smaller bias. Note that the bias has no simple linear dependence on the brightness temperature. This situation is different for the "simulated biases" seen in the sensitivity study that were induced by varying a calibration parameter of a single effect (see Sec. 6.2). These "simulated biases" were potential changes of the measured brightness temperature due to a certain effect that showed a distinct dependence on the brightness temperature, characteristic for this particular effect. Real biases as observed in Fig. 6.13a do not show such a characteristic dependence. Therefore, the dependence of the bias cannot be easily assigned to a certain effect. Looking at the inter-satellite biases in Fig. 6.13, it should be noted that all FOVs are considered together. Furthermore, the temperature range is quite small compared to the temperature range in the sensitivity study that was not based on percentiles, but on all measurements of a single orbit. Thus, any dependence on brightness temperature becomes much clearer in the sensitivity study than on the reduced temperature range of the shown biases. Moreover, Fig. 6.13 shows a bias with two instruments involved such that other instrumental issues, sampling issues and diurnal cycle effects also have an impact. For channel 4 (Fig. 6.13b), the large bias of about -0.2 K (see Fig. 3.11 also) can be reduced by applying the antenna pattern correction with correct assignment to channel 4 for NOAA-18 and the newer version antenna pattern correction for Metop-A. A similar result is obtained for Metop-B (Fig. 6.13c) that also showed a large bias in channel 4. The remaining bias may be due to other instrumental effects or different sampling of the asymmetric diurnal cycle. Note that the plotted data here are averages from all data, i.e. both ascending and descending branch. Another reason could be also instrumental deficiencies of NOAA-18 (see Sec. 6.3.3).

An improvement due to correct assignment of antenna pattern corrections would also be expected for all AMSU-B instruments. This is especially expected for channel 4, because channel 4 erroneously receives the same correction as channel 1, which is most different from the correct correction values for channel 4. The correct correction values for channel 4 are those assigned to channel 3, since all three sounding channels 3-5 of AMSU-B have the same quasi-optical path and should therefore receive the same antenna pattern correction. No significant improvement can be observed, however, if correcting for the erroneous assignment. This is due to the impact of RFI on channel 4 for all AMSU-Bs that dominates over the bias introduced by the wrong antenna pattern correction.

6.3.3 Deficiencies of MHS on NOAA-18

Apart from the wrong assignment to channels, it is likely that MHS on NOAA-18 has more deficiencies in its antenna pattern correction. Since the correction was simply copied over from the AMSU-B instruments, it may not correct adequately for the contamination of the signal in the side lobes of MHS on NOAA-18. Evidence for these additional deficiencies is obtained from looking at scan-dependent biases as I did for the investigation of RFI related biases above. Comparing the scan-dependent biases in channel 3 for NOAA-19 – NOAA-18 and for Metop-B – NOAA-18 (see Fig. 6.8a, 6.10a), the same smoothly varying bias pattern with a dip around FOV 15 and an increase towards FOV 90 can be observed. I suggest that this bias pattern emerges from a wrong antenna pattern correction for NOAA-18. This is supported by what is visible

Antenna Efficiency: Space: MHS NOAA19 - NOAA18

(a) channel 3

Antenna Efficiency: Space: MHS METOPB - NOAA18
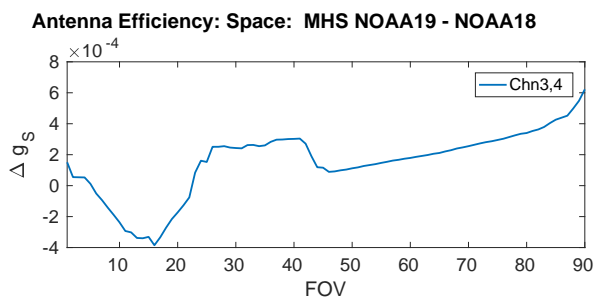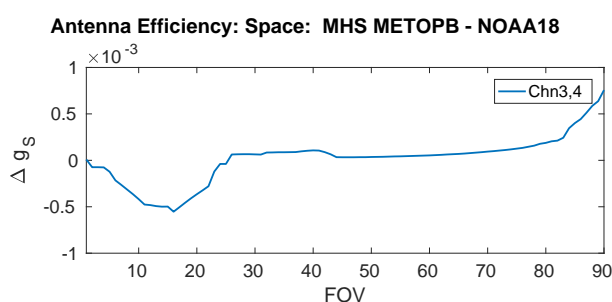
(b) channel 4

Figure 6.14 Differences in the antenna pattern correction for NOAA-19, Metop-A compared to NOAA-18. The difference in the contribution from space g_S for each Earth view is plotted. Compare the pattern to Fig. 6.8a, 6.10a.

in Fig. 6.14. Figure 6.14 displays the antenna pattern correction difference $g_S(\text{NOAA-19}) - g_S(\text{NOAA-18})$ and $g_S(\text{Metop-B}) - g_S(\text{NOAA-18})$. This is the difference of the space contribution efficiency to the overall signal in the Earth view between MHS on NOAA-19 and on NOAA-18, and between MHS on Metop-B and on NOAA-18, respectively. For both instruments on NOAA-19 and on Metop-B, the difference in the antenna pattern correction compared to NOAA-18 looks very similar and matches the bias pattern seen in Fig. 6.8a and Fig. 6.10a. This hints at a wrong antenna pattern correction for NOAA-18. For the Metop-A instrument, this bias pattern is not visible, since both instruments, on Metop-A and on NOAA-18, receive the same antenna pattern correction in AAPP, and therefore make a similar mistake in correcting the signal for antenna pattern effects. Hence, this bias pattern is not visible.

Any instrumental issue on NOAA-18 is not correctable by the harmonisation (at the moment), since it relies on NOAA-18 as reference in the current setting. Also, an impact of RFI on MHS on NOAA-18 cannot be ruled out. An RFI impact may distort the signal measured by MHS on NOAA-18. In this case, RFI would have an impact on all inter-satellite biases that use MHS on NOAA-18 as the reference. As pointed out in Sec. 6.1, the harmonisation can only produce a relative improvement, since there is no absolute, external reference. Executing harmonisation with another reference is an interesting effort for further understanding the causes for the biases, but it is beyond the scope of this thesis.

6.4 The harmonisation procedure

In the following I briefly describe the harmonisation procedure with its input and output data as designed and set up within the FIDUCEO project. Early results using this method are described in Sec. 6.5. So far, only channel 3 is considered in the harmonisation, because it will be used for the CDR generation.

6.4.1 Input data

The input data set that the FIDUCEO harmonisation is based on, is produced with an SNO-like collocation method on the AAPP processed level 1c data by the project partner Brockmann Consulting GmbH (Block, 2015). This method generates so called match-ups for each sensor pair. A match-up is an event, where both sensors measure the same target at the same time, using the thresholds $\Delta t < 5$ min, $\Delta s < 5$ km. Moreover, a cloud-screening as in Buehler et al. (2007) is applied to allow for clear-sky cases only. Unlike SNOs, the FIDUCEO match-ups so far allow for all FOVs. Hence, the match-up system finds events where FOV 1 of sensor A is matched with FOV 45 of sensor B. Consequently, the limb effect will have a different impact on the corresponding measurements, thus always displaying a natural bias. However, there will also be a match for FOV 45 of sensor A with FOV 90 of sensor B. This match-up will show the same natural bias due to the limb effect, but with a different sign. Hence, the bias is cancelled out in a large data set.

The match-ups are collected for all sensor pairs and stored in NetCDF files (per week) with information on original level 1c data file name, time, scan line and FOV of the match-ups and the corresponding measured brightness temperature. These match-up files form the basis for the harmonisation input files.

The harmonisation input files, designed in a common format for all instruments considered in FIDUCEO, are constructed by Martin Burgdorf at Universität Hamburg per microwave humidity sounder sensor-pair for all their match-ups over their mission. Within these files, all information is stored that is required for the harmonisation procedure: from the level 1b data corresponding to each match-up, the required information for calibration is collected. This includes the counts measured for the DSV, IWCT and Earth view, as well as the temperature of the black body for all seven scan lines used for the scan line of the match-up. For the reference sensor, the AAPP processed brightness temperature is stored. Uncertainty estimates are stored for each quantity, based on typical values obtained from the production of the unharmonised FCDR. Additionally, an estimate for the expected random difference of the two sensors $K_r = 0.5$ K is provided, based on an estimate of natural variability of the atmosphere within the space and time thresholds, obtained from a standard deviation of the surrounding FOVs. The harmonisation input file is cleaned from bad data by excluding any flagged data and any data outside reasonable ranges. A matrix containing correlation information of adjacent match-ups is added to the file by project partners at NPL.

6.4.2 The optimisation solver

The prepared harmonisation input file per sensor-pair is then handed over to the optimisation solver developed and run by project partners at FastOpt GmbH and NPL (Giering et al., 2017; Quast, 2018). The optimisation solver is a complex code that solves the harmonisation problem for all sensor pairs at once. The harmonisation problem itself is an optimisation problem of

minimising a certain quantity. This quantity is the deviation of the expected difference K_{ij} and the observed difference in radiance of two sensors $L_i - L_j$

$$(L_i - L_j) - K_{ij}. \tag{6.1}$$

The full cost function J to be minimised accounts for uncertainties of all involved quantities and prior knowledge. For the harmonisation parameters $\mathbf{a}_1, \ldots, \mathbf{a}_S$ for the sensors $1 \ldots S$ it reads (Giering et al., 2017):

$$
\begin{aligned}
J(\mathbf{a}_1, \ldots, \mathbf{a}_S) = &\frac{1}{2} \sum_{i<j} (\mathbf{d}(\mathbf{a}_i, \mathbf{a}_j) - \mathbf{K}_{ij})^T C_{ij}^{-1} (\mathbf{d}(\mathbf{a}_i, \mathbf{a}_j) - \mathbf{K}_{ij}) \\
&+ \frac{1}{2} \sum_{i<j} \sum_{k=1}^{m_{ij}} (\mathbf{q}_{ijk} - \overline{\mathbf{q}}(t_{ijk}))^T C_{\mathbf{q}_i}^{-1} (\mathbf{q}_{ijk} - \overline{\mathbf{q}}(t_{ijk})) \\
&+ \frac{1}{2} \sum_i (\mathbf{a}_i - \overline{\mathbf{a}}_i)^T \overline{C}_{\mathbf{a}_i}^{-1} (\mathbf{a}_i - \overline{\mathbf{a}}_i)
\end{aligned}
\tag{6.2}
$$

with $\mathbf{K}_{ij}$ being the vector that contains the expected measurement differences between sensor i and j for all match-ups. The expected difference can be due to spectral response function differences. For channel 3 of the MW sounders, there is no such expected systematic difference; only a random component $\mathbf{K}_r$ is considered. m_{ij} counts the match-ups, $\mathbf{q}_{ijk}$ denote the sensor state variables, i.e. the counts of the DSV, IWCT and the temperature measured by the PRTs. C_{ij}, $C_{\mathbf{q}_i}$ and $\overline{C}_{\mathbf{a}_i}$ are the error covariance matrices for the measurements, for the sensor state variables, and for the harmonisation parameters, respectively. $\mathbf{d}(\mathbf{a}_i, \mathbf{a}_j)$ is the difference between the measurements of the two sensors i and j

$$
\mathbf{d}(\mathbf{a}_i, \mathbf{a}_j) = \begin{pmatrix} f_i(t_{ij1}, \mathbf{a}_i, \mathbf{q}_{ij1}) \\ f_i(t_{ij2}, \mathbf{a}_i, \mathbf{q}_{ij2}) \\ \vdots \\ f_i(t_{ijm_{ij}}, \mathbf{a}_i, \mathbf{q}_{ijm_{ij}}) \end{pmatrix} - \begin{pmatrix} f_j(t_{ji1}, \mathbf{a}_j, \mathbf{q}_{ji1}) \\ f_j(t_{ji2}, \mathbf{a}_j, \mathbf{q}_{ji2}) \\ \vdots \\ f_j(t_{jim_{ij}}, \mathbf{a}_j, \mathbf{q}_{jim_{ij}}) \end{pmatrix}
\tag{6.3}
$$

where t_{ijk} is the time and f the measurement equation of the sensors. The terms in the cost function

1. minimise the K-residual, i.e. the deviation from the expected difference K_{ij}, by optimising harmonisation parameters $\mathbf{a}$.

2. take into account uncertainties in all variables to optimise the value of the sensor state variables within their uncertainties.

3. use prior information on the harmonisation parameters $\mathbf{a}$, which is optional (not yet included in the first harmonisation runs).

NPL and FastOpt GmbH have developed two methods to minimise the cost-function to be used for the harmonisation problem (Quast, 2018): Errors-in-Variables (EIV) and Orthogonal Distance Regression (ODR). Both methods handle uncertainties in both the independent and dependent variables, unlike ordinary least squares. In the specific harmonisation case, the EIV method can handle all uncertainties, whereas the ODR does not handle structured uncertainties. For both methods, FastOpt GmbH developed fast versions that do not optimise the sensor state variables, but the calibration parameters only.

Using the input file to populate the vectors and matrices, the optimisation solver then minimises J by finding new values for the harmonisation parameters $\mathbf{a}$.

6.4.3 Output data

The most important output of the harmonisation procedure is a set of optimised harmonisation parameters **a**. I include these parameters in my FCDR generator code to then produce the harmonised FCDR. Also, the harmonisation procedure outputs the covariance matrices for all parameters of all sensors. These matrices contain the uncertainty information on the new optimised **a**. Moreover, the K-residual is provided in the output data and can be used for diagnostics of the harmonisation procedure. For an ideal harmonisation, the K-residual shows a random distribution consistent with the underlying errors. If the K-residual is not randomly distributed, this is a hint at a problem from the sensor or the match-up data that could not be solved (Quast, 2018).

Having a new set of optimised harmonisation parameters, I produce a new version of an FCDR. This harmonised FCDR can be harmonised with respect to different effects, depending on the chosen set of harmonisation parameters.

6.5 Harmonising the instruments regarding certain effects: Results

The first test results from the FIDUCEO harmonisation procedure are based on an fast-ODR run by FastOpt GmbH. In this run, channel 3 for all AMSU-B and all MHS instruments is harmonised simultaneously against the reference MHS on NOAA-18. The input data set includes match-ups over the whole lifetime of the satellites, except for MHS on NOAA-19, where the highly problematic period around September 2009 was excluded. The input data are not yet fully quality checked, but cleaned with a rough filter for strong outliers only, therefore, the following results can only be considered preliminary. The fully quality checked data will be used in the next runs. As harmonisation parameters, the non-linearity correction q_{nl} and the warm-target correction δT_{ch} are used. Hence, the vector of harmonisation parameters for this run is $\mathbf{a} = (q_{nl}, \delta T_{ch})$.

To verify the correct functioning of the harmonisation procedure, FastOpt GmbH check whether the K-residual distribution is random. Figure 6.15a shows the K-residual for the harmonisation of Metop-A. To guide the eye, a fit of a normal distribution (solid line) and a Poisson distribution (dashed lines) are plotted as well. The K-residual is purely random, following the normal distribution and indicating a successful harmonisation in the sense that the parameters were adjusted such that the inter-satellite bias is reduced for the input data set. This is not the case for the MHS on NOAA-19 for example (see Fig. 6.15b). In this case, the K-residual distribution is not symmetric around zero, indicating remaining problems and biases. It must be kept in mind that the outliers have not been removed from the data and may distort the results. Nonetheless, the asymmetry of the K-residual histogram substantiates the hypothesis I suggested in Sec. 6.3, that MHS on NOAA-19 is affected by RFI. According to this idea, no adjustment of q_{nl} or δT_{ch} can reduce the bias and a harmonisation attempt on this would produce non-random K-residuals as seen in Fig. 6.15b. A similar picture can be seen for NOAA-15 and NOAA-16, while NOAA-17 and Metop-B also follow the normal distribution.

To assess the actual success of the harmonisation in terms of reduction of inter-satellite bias in the whole FCDR, I include the optimised values for the harmonisation parameters of the non-linearity q_{nl} and of the warm target correction δT_{ch} in the FCDR generator to reprocess the level 1b data (see Tab. 6.1). Hence, I produce an early version of a harmonised FCDR for six months (September 2014 - February 2015) for MHS on Metop-A. Figure 6.16 shows the average
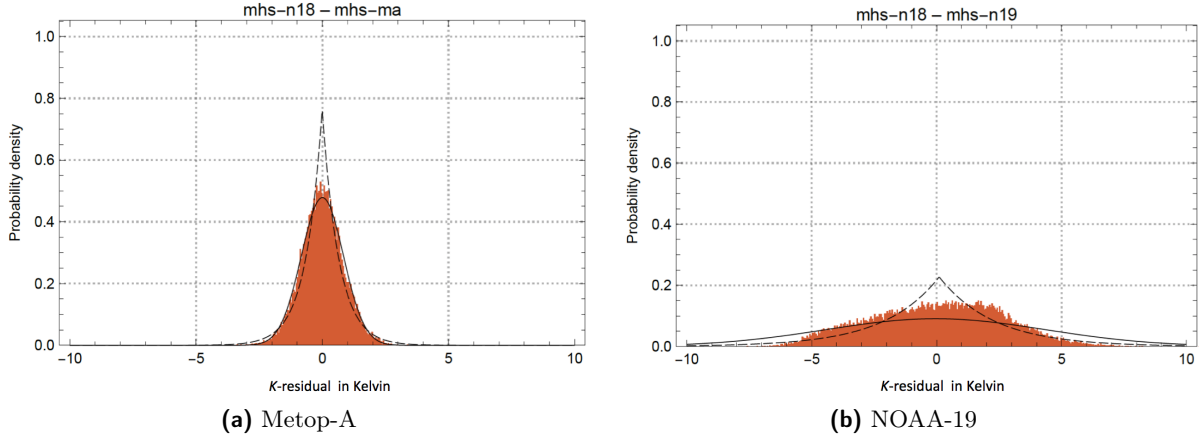
(a) Metop-A **(b)** NOAA-19

Figure 6.15 Histogram of K-residual after first harmonisation run. Figure courtesy Ralf Quast.

over the six months for the 10th to 90th percentile of the brightness temperatures. It is the same as Fig. 6.13a, but the FCDR is now based on the harmonised parameters q_{nl} and δT_{ch}. This figure shows the "before" (blue, AAPP) and the "after" (orange, FCDR) of harmonisation. An improvement, meaning a smaller absolute value of the bias, is only achieved around 240 to 244 K. Beyond that, the FCDR shows a larger bias than the operational AAPP processed level 1c data. The overall pattern of the bias over the range of brightness temperatures changed also. Instead of an increase to a positive bias at high temperatures, the bias now tends to large negative values. At the low temperature end, a negative bias is turned into a positive bias. Clearly, the bias was overcompensated by adjusting the non-linearity from about $-25\,\mathrm{W}^{-1}$ to $-256\,\mathrm{W}^{-1}$ and the warm target correction from $0\,\mathrm{K}$ to $-0.9\,\mathrm{K}$. The former produces warmer temperatures for Metop-A at low temperatures, the latter produces colder temperatures for Metop-A at high temperatures (see Sec. 6.2), resulting in the changed bias pattern.

The chosen harmonisation parameters cannot explain the inter-satellite bias against NOAA-18, as their adjustment does not lead to a real improvement. Also, the very strong increase of non-linearity appears improbable. However, another important reason for the limited success of this harmonisation run is the following. The input data set that the harmonisation uses contains match-up data only. These match-ups are obtained as SNO-like collocations and therefore only provide brightness temperatures measured at high latitudes, hence sampling a small range of temperatures. Figure 6.17 shows the abundance of matched brightness temperatures for MHS on NOAA-18 and Metop-A. Most of the matches are around 240 to 244 K. Consequently, the

effect	parameter	operational value	estimated uncertainty	optimised value (1st run)	uncertainty (1st run)
non-linearity	q_{nl}	-25 W^{-1}	100%	-256 W^{-1}	45 W^{-1}
warm target corr.	δT_{ch}	0 K	100% of 0.16 K	-0.9 K	0.2 K

Table 6.1 Optimised parameters for the non-linearity q_{nl} and warm target δT_{ch} correction from the first harmonisation run. Note that the listed operational value is an average of the non-linearity coefficient varying between -35 and -15 W^{-1} with instrument temperature. The estimated uncertainty on the operational value is explained in Sec. 5.2.5 and 5.2.6. It does not enter harmonisation. The uncertainty of the optimised values is a result of the harmonisation procedure.
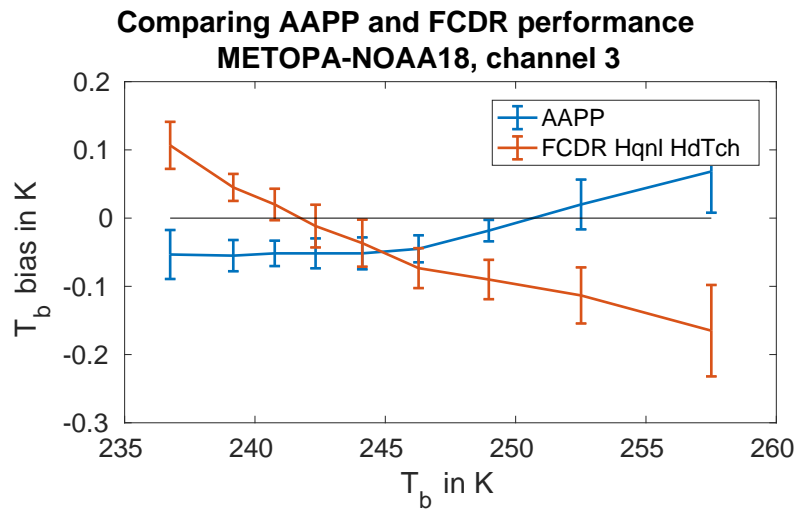
143

**Comparing AAPP and FCDR performance
METOPA-NOAA18, channel 3**

Figure 6.16 Comparison of resulting biases for AAPP and harmonised FCDR processing for MHS on Metop-A, channel 3 against MHS on NOAA-18 (for September 2014 - February 2015). Harmonisation parameters are non-linearity q_{nl} and warm target correction δT_{ch}. The uncertainty bars represent the standard deviation of the bias over the six months of used data.
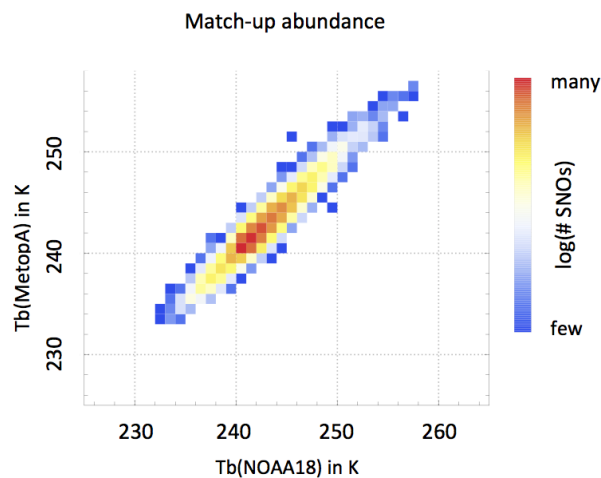

Figure 6.17 Abundance pattern of brightness temperatures for FIDUCEO match-ups of MHS on Metop-A and on NOAA-18. Figure courtesy Ralf Quast.

optimisation solver optimises the calibration parameters to harmonise these brightness temperatures. If I then use the new calibration parameters to produce an FCDR including the full temperature range, the extrapolation fails as the parameters were not optimised to reduce biases at other temperatures. This is an inherent problem of the harmonisation based on SNO-like collocations only. However, since this approach is the way to access actual measurements for recalibration, the procedure is pursued, but requires improvement. With an improved procedure, it is then possible to iterate on finding the correct combination of parameters to reduce the bias.

A second and third harmonisation run were executed on fully quality checked data (affecting mostly AMSU-B). The second harmonisation run shows the same results for the calibration

parameters as the first run that used the rough filter. Hence, the deduced harmonised FCDR for Metop-A does not change compared to the results from the first run presented above and does not improve the results. Also, NOAA-15, NOAA-16 and NOAA-19 still cannot be harmonised as the remaining K-residual indicates. A third run allowing for harmonisation using reference-to-sensor and also sensor-to-sensor match-ups at the same time, shows that instruments which cannot be harmonised with the chosen parameters may deteriorate the result for others. Using matches of Metop-A and NOAA-16 for the harmonisation produces calibration parameters for Metop-A that increase the bias (non-random K-residual). The harmonisation tries to reduce the overall occurring biases, at the cost of increasing individual biases. Hence, the NOAA-15, NOAA-16 and NOAA-19 instruments should be excluded at first from harmonisation using sensor-to-sensor match-ups also, due to their possible RFI-induced bias. Executing the same harmonisation runs with the polarisation correction as third parameter still does not improve the overall results.

To significantly improve the functioning of the FIDUCEO harmonisation, an extension of the input data is necessary to also cover the warm end of the brightness temperatures with adequate sampling. It is impossible to achieve this by SNOs, since global SNOs only happen for periods of similar LECT, thus covering only a small period of the instrument's lifetime. A promising method could be the collocation via virtual match-ups - an idea that arose by Stefan A. Buehler after the first harmonisation results and within the frame work of the Master's theses by J. Mrziglod and M. Prange (ongoing work in both cases). The idea of virtual match-ups is that a UTH-sensitive sensor (i.e. corresponding to channel 3 of MHS) on board a geostationary satellite should serve as "mediator" between two sensors on board polar orbiting satellites. That means, the geostationary instrument serves as "filter" on match-ups with a large time window between the two polar orbiting satellites. For example, the Spinning Enhanced Visible and Infrared Imager (SEVIRI) on board the geostationary METEOSAT-10 satellite is collocated with both MHS on NOAA-18 and on Metop-A separately (all collocations are found by J. Mrziglod and M. Prange using the collocation tool by J. Mrziglod). In a second run, those matches are collected ("filtered") where the SEVIRI brightness temperature in the 6.2μm and 7.3μm channels remains constant (e.g. $\Delta T = 1$K) for several hours over a confined small region. In this way, virtual match-ups of NOAA-18 and Metop-A are generated. The virtual match-ups are separated in time (and slightly in space), but should still sample a very similar state of the upper troposphere (to which extent this approximation holds, is still under investigation). This method allows for adequate sampling of the high temperature end, since the constant SEVIRI temperatures are mostly found in subsidence regions as expected. Figure 6.18 shows the abundance of those virtual match-ups for only ten days of NOAA-18 and Metop-A in December 2013, as well as the traditional SNOs. The cold temperature end is sampled by the traditional SNOs, whereas the warm temperature end is sampled by the virtual match-ups. The standard deviation is similar for both methods. Hence, this method is promising for the extension of the harmonisation input data for a warm temperature calibration point.

An integration of this method in the production of the harmonisation input files and a new harmonisation run based on the combined set of FIDUCEO match-ups and virtual match-ups, is planned, but beyond the scope of this thesis.

6.6 Conclusion

Before starting actual harmonisation of MW data in this chapter, I analysed the sensitivity of the final brightness temperature to the various effects affecting the measurement process. In this
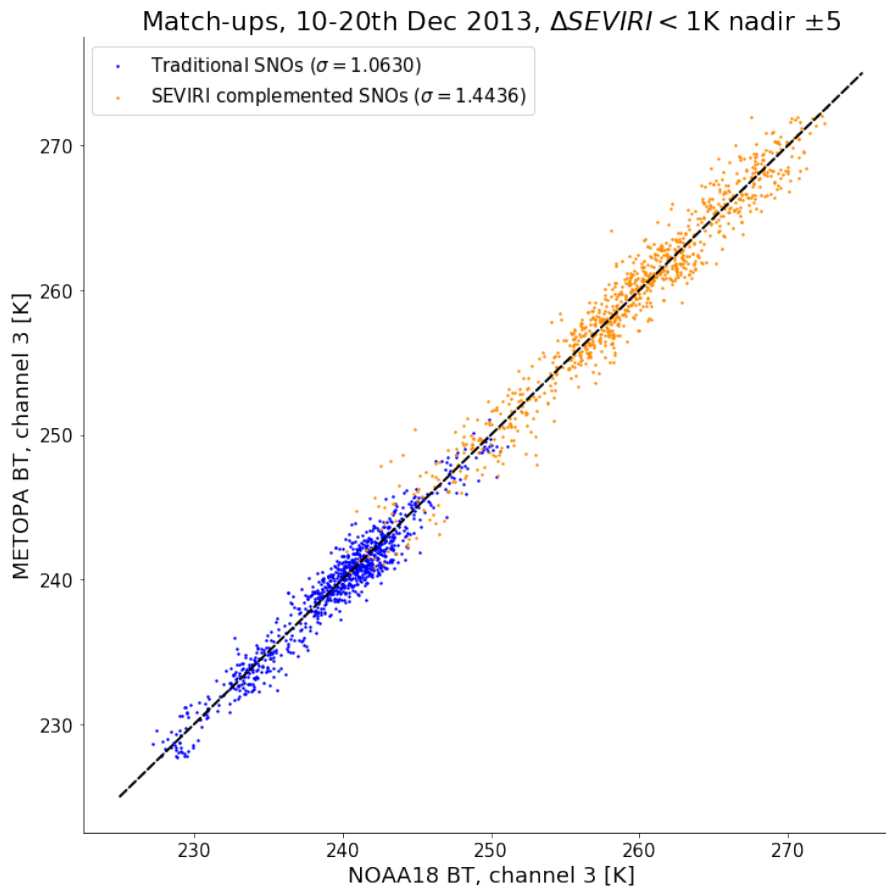
Figure 6.18 Scatter plot of brightness temperatures for SNOs (blue) and virtual match-ups (orange) of MHS on Metop-A and on NOAA-18. Five FOVs to either side of nadir only. Figure courtesy J. Mrziglod and M. Prange.

study, I showed the dependence on measured brightness temperature and on FOV of the various effects, and I identified weak effects such as the non-linearity correction or strong effects such as the warm target correction. Both effects are used in the first test harmonisation run.

As an intermediate step, I analysed in more detail the strong biases seen in channel 3 (partly in channel 4 and 5 also) for the AMSU-B on NOAA-15, NOAA-16 and the MHS on NOAA-19 against MHS on NOAA-18. From this analysis I suggest RFI as a possible explanation for the biases. Consequently, harmonising these instruments is a particularly difficult undertaking. Any straight-forward harmonisation attempt on RFI effects is impossible since an additive correction parameter would have to be applied to the raw counts. Furthermore, free optimisation of this new parameter would be necessary in order to account for erratic RFI behaviour. This would result in random manipulation of the raw data and is thus not feasible. Strict constraints would have to be applied to this correction parameter. These strict constraints would need to be obtained from a detailed study of the observed bias patterns. This study may already yield a new devised correction scheme as was given for the first years of NOAA-15. This correction scheme could then be further optimised by harmonisation, if necessary.

By correcting for a wrong assignment of antenna pattern corrections to channels for AMSU-Bs and MHS on NOAA-18, I was able to reduce some biases especially for channel 4 that erroneously received the correction for channel 1 in the operational AAPP processing.

The analysis of scan-dependent biases revealed that the antenna pattern correction for MHS on NOAA-18 is probably wrong and causes visible biases for NOAA-19 and Metop-B when compared to NOAA-18. This is plausible, since the same correction is assigned to MHS on NOAA-18 and the AMSU-Bs. This implies that a different reference instrument is necessary for harmonisation in the long run.

The FIDUCEO harmonisation procedure was presented with its first results, problems and possible improvements. The idea of harmonisation is to reduce and explain inter-satellite biases in a recalibration approach. The recalibration is an optimisation procedure, where certain parameters of the measurement equation are changed such that the overall inter-satellite biases are reduced. The underlying data set for inter-comparison is a collection of match-ups, i.e. SNO-like events. The first harmonisation results reveal the expected problem that the SNOs cause since they only cover a small range of brightness temperatures. A bias reduction may be achieved in this small range, but it causes large biases outside the range. An improvement on the input data set to overcome this problem, is currently being developed in the frame work of Master's theses. The geostationary instrument SEVIRI is used as a filter to generate match-ups of two polar-orbiting instruments for warm temperatures as well. This method would extend the harmonisation input data and allow for better harmonisation. Then, the correct combination of parameters to reduce the bias can be obtained in an iterative manner by optimising different combinations of possible harmonisation parameters in several harmonisation runs.

7 Summary, conclusions and outlook

Climate research requires long-term, stable data records of geophysical observables that enable scientists to detect trends of climate. These climate data records (CDRs) can be derived from the underlying fundamental climate data records (FCDRs) that store the radiance or brightness temperature of satellite based measurements.

The generation of uncertainty quantified FCDRs and derived CDRs is the overall aim of the Horizon 2020 project FIDUCEO, which this thesis is part of. The aim of this thesis is to contribute to the generation of a new FCDR for microwave humidity sounders as basis for a CDR of upper tropospheric humidity (UTH). The three new aspects within this MW FCDR generated in the framework of this thesis are:

1. extensive uncertainty information on pixel level, including information on correlation behaviour,

2. the understanding and the reduction of inter-satellite biases, to be refined with the novel method of harmonisation,

3. a user-friendly NetCDF data format of the FCDR files, based on consolidated data.

These aspects address three main problems of the currently available data from operational processing. The data lack uncertainty information, contain inter-satellite biases and require tedious handling of doubled data that need cleaning before usage. This cleaning is part of the generation of FCDRs based on consolidated data of the MHS, AMSU-B and SSMT-2 instruments.

So far, the inter-satellite biases prevent the construction of long-term, stable data records. This problem was described in Chapter 3. Depending on the inter-comparison method used, the actual value of observed biases differ because of different sampling. Nonetheless, strong time and scan-dependent biases against the chosen reference satellite NOAA-18 are seen for all methods, for channel 3-5 of AMSU-B on NOAA-15 and NOAA-16 and for channel 3 and 4 of MHS on NOAA-19. Smaller biases are also visible for AMSU-B on NOAA-17, and MHS on both Metop-A and Metop-B (John et al., 2013a; Hanlon and Ingram, 2016). I also showed these biases using a method based on the analysis of monthly percentiles, developed recently in a Master's thesis. The understanding of the strong biases of these three instruments and also of the smaller biases for the Metops, for example, was later treated in Chapter 6, introducing also the FIDUCEO harmonisation.

Addressing the problem of missing uncertainty information in the available data sets is an important part of the thesis. Uncertainty information should be included in the FCDR files as a pixel-wise standard uncertainty of the brightness temperature. This uncertainty estimate should incorporate all known sources of uncertainty. To achieve this, an uncertainty analysis must be based on the understanding of the instrument's calibration, on its behaviour and on the evolution of performance over its lifetime.

In Chapter 4, I studied the behaviour of the basic calibration quantities over the lifetime of the instrument. The evolution of the temperature of the black body, the internal warm calibration target, is strongly affected by orbital drift. The latter induced a change of the solar-beta-angle and hence produced a strong heating and cooling when the satellite approaches an LECT of 18:00. This was the case already for NOAA-15, NOAA-16 and recently for NOAA-18. The strong temperature change of the black body itself does not affect the calibration, since the latter is carried out on much shorter time scales than the temperature change happens. Nonetheless, other parts of the instrument may be affected by this temperature change in a negative way and cause operational issues.

For AMSU-B on NOAA-15 and on NOAA-16, the impact of orbital drift is preceded and accompanied by a strong decrease in gain. This decrease in gain has consequences for the sensitivity of the instrument, leading to an increased NEΔT. Also, the accuracy is reduced. The decreased gain increases the impact of RFI, which may induce a bias. I observe the very same phenomenon for channel 3 (and to a lesser extent for channel 4) of MHS on NOAA-19. In this case the decrease is sudden instead of continuous.

For the MW humidity sounders on NOAA-17 and Metop-B, the gain is stable throughout the mission. Whereas for MHS on Metop-A and NOAA-18, the gain in channels 3 and 4 shows strong changes, but never decreases as strongly as for NOAA-15 and NOAA-16. Hence, no critical increase of NEΔT is seen.

The evolution of noise and NEΔT is studied in detail in section 4.2. I started with clarifying the noise terminology, differentiating between noise on counts and its translation to NEΔT using the gain. Hence, NEΔT is not pure noise, but represents the sensitivity that is determined inclusively by the gain. I used the Allan deviation for noise estimation. This method has recently been suggested by Tian et al. (2015) for use with polar orbiting instruments. The Allan variance is the special case for $M = 2$ of the M-sample variance using M adjacent samples. This method is based on differences of adjacent samples and is hence insensitive to orbital variations, whose effects become apparent on longer time scales. This is an advantage for noise estimation in flight compared to the standard deviation.

The method also allows for easy determination of the noise spectrum by so-called bias-functions that make use of a quotient of the general M-sample variance and the Allan variance. The variation of this quotient with varying M is characteristic for white or pink noise ($1/f$-noise), for example. Computing the bias-functions for the DSV counts of the sensors and comparing to the bias function of white noise varying with M reveals that the sensors have different noise spectra in different channels, also varying with time. Often, however, a mixture of white and pink noise is observed. A more detailed study may be carried out in the future for more time periods, focussing on the changes of noise spectra.

The main outcome of my noise analysis is the overview of the time evolution of the NEΔT over the lifetime of the instruments. The evolution shows strong increases of NEΔT due to increases of noise (for some SSMT-2), or due to strong decreases of the gain as for NOAA-15, NOAA-16 and NOAA-19. Covering the whole lifetime of the instruments, the produced plots provide a helpful overview of NEΔT evolution for all MW humidity sounders.

Hence, the study on noise is a stand-alone result. However, it also serves as an important part of the uncertainty analysis regarding the generation of an FCDR. The analysis of the noise and the basic calibration quantities, such as the calibration counts and the derived gain, gave insight to the overall evolution of the instrument. Using the Allan deviation in this study, I tested and established the method to calculate the noise as part of the FCDR uncertainty analysis and later FCDR production.

For the FCDR uncertainty analysis, an understanding of the instruments' calibration is required in order to assess the relevant sources of errors and estimate their induced uncertainty. In Chapter 5, I introduced the metrological measurement equation approach as used within FIDUCEO. The calibration procedure is expressed in one function translating raw counts to brightness temperature. This function models all considered "effects". The "effects" are physical processes or measurement conditions that have an impact on the measurement and that need to be accounted for in order to remove the introduced distortion from the measured signal.

The modelled effects are grouped into three classes of resulting uncertainty, distinguishing the different correlation behaviour of the underlying unknown errors. The three classes are *independent*, *structured* and *common* uncertainties. The class of independent uncertainty incorporates the uncertainties due to all effects that produce truly random errors on pixel level. Hence, the noise on Earth counts enters this class. The class of structured uncertainty is composed of effects that produce random errors by their nature, but that become correlated at pixel level through the calibration procedure. The noise on the temperature measurement of the black body and the noise on the calibration counts from DSVs and IWCT views are part of this class, since their resulting uncertainty on the brightness temperature is correlated due to the applied 7-scan line weighted rolling average in the calibration. The class of common uncertainty is composed of all contributing systematic effects that induce correlation on length scales larger than one orbit, usually over the whole mission. This class consists of corrections on the calibration targets, the non-linearity correction, the antenna pattern correction and the polarisation correction.

For each effect, an input uncertainty estimate must be obtained, since no uncertainty information is provided for the operational parameters modelling the effects. This estimate of input uncertainty per effect can be based on an in-depth analysis, similar to the noise contribution analysis. Each noise component is computed with the Allan deviation and serves as uncertainty estimate of the respective noise effect on Earth view, DSV, IWCT or black body temperature measurement. Other estimates are based on the spread of calibration parameters for different instrument configurations given in the operational AAPP calibration, or on very coarse estimates of 50-100% due to implausible assignments of calibration values in AAPP. For each effect, I also provide estimates on the cross-channel correlation and on the correlation length scales along and across scan lines.

Using these input uncertainties, I computed the final uncertainty on the brightness temperature with the law of propagation of uncertainty, making use of the partial derivatives of the measurement equation. For each class of uncertainty, the corresponding final uncertainty of the brightness temperature is given as a combination of the uncertainties of all members of the class. Hence, the FCDR files provide pixel level uncertainty of the brightness temperature due to independent, structured and common effects. The propagation of uncertainty together with the actual processing and calibration is done by my FCDR generator code implemented in MATLAB.

The FCDR generator code, a major result of the work in Chapter 5, executes the required cleaning and consolidation of the data. It produces orbit-wise FCDR files in the NetCDF-4 format from one equator crossing to the next in the same flight direction. This removes all data doubling from the overlap of subsequent files and hence prevents the appearance of sampling artefacts. Moreover, concise quality flags are set during the processing and provided as pixel-level information along with the explanation of each flag within each of the NetCDF FCDR files.

The FCDR generator can imitate operational AAPP results, which increases my confidence in the functioning of the implemented processing chain and measurement equation. Because of the lack of detailed uncertainty information in the previously available data, the uncertainties now

contained in the FCDR cannot be compared to other data. However, the estimates that exist for NEΔT from instrument specifications or other test-wise applied methods of NEΔT-computation (Atkinson, 2015) correspond well to my results for the uncertainties of brightness temperature due to noise effects. Looking at the three classes of uncertainties, histograms reveal that the independent effects mostly dominate the overall uncertainty budget. This is encouraging, since this uncertainty can be reduced by averaging processes applied in the later CDR production, whereas the common uncertainties cannot be reduced.

The FCDR code allows for the adjustment of the calibration parameters related to all effects. This enables the easy change of the calibration based on new insights on better calibration parameters for certain effects. In this regard, I corrected the assignment of antenna pattern corrections to channels for AMSU-B, which was wrong in AAPP.

Concluding, the production of the unharmonised FCDR works and a complete beta version is processed. This unharmonised FCDR already answers two problems of operational data, since it provides extensive uncertainty information and stores FCDR files in a user-friendly format, based on consolidated data.

The third problem, namely existing inter-satellite biases and their reduction, is addressed in Chapter 6 on the harmonised FCDR. Harmonisation aims at reducing inter-satellite biases by finding the best calibration parameters for certain physical effects. "Best" means that the optimised values of these calibration parameters lead to changes of the brightness temperature that minimise the bias. This is a recalibration approach that sets apart the FIDUCEO harmonisation from other inter-calibration attempts.

The intended change of calibration parameters for certain effects induces an intended change of brightness temperature that should ideally reduce the existing bias. To analyse possible induced changes in brightness temperature for the different effects, I carried out a sensitivity study on the various effects. In this study, I discussed different dependencies of the changes on the scan angle of the field of view (FOV), and on the brightness temperature. Non-linear dependence on the brightness temperature is only introduced by the non-linearity correction by construction. However, differently orientated linear dependencies exist with a strong effect at the warm temperature end and a weak effect at the cold end, or vice versa. I identified relatively weak effects, such as the non-linearity correction, or relatively strong effects, such as the warm target correction and the polarisation correction. This sensitivity study provides an overview of the potential of the different effects to change the brightness temperature. This helps to provide information for the harmonisation process and later understand and judge harmonisation results.

In another preparatory step for the harmonisation, I analysed existing inter-satellite biases. This does not only help to understand harmonisation results, but also prevents inappropriate optimisation of parameters.

I analysed in detail the strong biases observed for channel 3 of the AMSU-B on NOAA-15 and NOAA-16 and of the MHS on NOAA-19 against MHS on NOAA-18. These biases show a strong dependence on time and on scan angle. The increase of bias over time is strongly correlated with the decrease of the gain. The scan dependence in turn shows distinct, stable patterns with strong differences in the bias from one FOV to the next similar to an interference pattern. These stable patterns intensify when the gain decreases. From these analyses, and the proven radio frequency interference (RFI) on NOAA-15 during the in-orbit verification phase, I suggest RFI as a cause for the strong inter-satellite biases for NOAA-15, NOAA-16 and NOAA-19. The

analysis was carried out for channel 3 so far, but channel 4 (and 5 for AMSU-B) show the same time-dependent bias and gain correlation and are likely also affected by RFI.

Another unexpected large bias is observed for channel 4 of the MHS on Metop-A and Metop-B against NOAA-18. I was able to identify the cause for this in the wrong assignment of antenna pattern corrections to the channels in the operational processing AAPP and in the usage of the old version of the antenna pattern correction for Metop-A by default (which was copied over from NOAA-18 in AAPP). Both issues are corrected in my FCDR generator code.

However, another identified issue is not solved so far: analysing scan-dependent biases for NOAA-19 and Metop-B against NOAA-18, I revealed that the operational antenna pattern correction for NOAA-18 is probably wrong. This is suggested by the distinct scan-dependent bias patterns. The bias patterns resemble the patterns of the differences in the antenna pattern correction of the respective instruments for NOAA-19 – NOAA-18, and for Metop-B – NOAA-18. An erroneous antenna pattern correction for MHS on NOAA-18 is probable, since the correction used in AAPP for MHS on NOAA-18 was only copied from AMSU-B on NOAA-15. Hence, it would be desirable to construct an adequate antenna pattern correction for the MHS instrument on NOAA-18 as it was obtained for NOAA-19 and the Metops.

Finally, I presented the harmonisation procedure as designed in FIDUCEO and discussed its first results. The harmonisation is based on SNOs as the inter-comparison method and uses an ODR (orthogonal distance regression) optimisation solver, implemented by project partners at NPL and FastOpt GmbH, to obtain new values for certain chosen calibration parameters. Using these new parameters in the calibration, the inter-satellite bias is expected to be reduced.

In a first test run of harmonisation presented in this thesis, two calibration parameters were optimised: the non-linearity correction and the warm target correction. The harmonisation diagnostics plots show that the parameters for NOAA-15, NOAA-16 and NOAA-19 could not be optimised to only give a remaining random deviation from the reference NOAA-18 in the used input data based on SNOs. A bias remains, showing that an optimisation is not possible using these chosen parameters.

As I suspect RFI to be the cause for the large biases of these instruments, I do not expect the current harmonisation procedure to successfully reduce these biases. The RFI impact is not an effect such as the other effects considered, as it cannot be cast into a formula in the measurement equation with a constant parameter to be optimised. RFI requires a scan-dependent and possibly time dependent additive correction on the raw counts, which forbids the simple application of an optimisation procedure as it would optimise the measurement signal. To use an automatic optimisation procedure for the harmonisation of instruments that are affected by RFI, strict constraints for the correction of each individual FOV for each period of constant RFI impact must be applied. These constraints on the possible range of an applied correction would need to be devised beforehand in a detailed study.

For the other sensors, on NOAA-17 and the Metops, however, no apparent failure of the optimisation procedure can be seen. Hence, optimised parameters were obtained that reduce the inter-satellite bias in the input data. To generate a first part of a harmonised FCDR and validate the harmonisation result, I used the optimised parameters in my FCDR generator code to obtain newly calibrated brightness temperatures for a period of six months. The bias was successfully reduced in the brightness temperature range that is well represented in the input data. However, outside this range, at other brightness temperatures, the optimised parameters induced a strong change, as it is expected based on the sensitivity study. This strong change increased the bias in these temperature ranges. Hence, the harmonisation procedure needs refinement to reduce the inter-satellite biases over the whole covered brightness temperature range.

Two problems in the harmonisation procedure can be identified so far. First, in order to use also off-nadir data, all combinations of FOVs were allowed in the SNO-like match-ups. However, this mixes strongly limb-affected data with non-limb-affected data, which is problematic in case the effect does not average out over the input data set. The alternative to only harmonise groups of FOV-combinations raises the question of justified thresholds in order to form groups. Moreover, it will be hard to harmonise the off-nadir groups due to significantly reduced amount of data. Consequently, as a first approach, one can concentrate on near-nadir data only.

The second, fundamental problem concerns the confined range of covered brightness temperatures in the input data based on SNOs. SNOs usually occur at high latitudes and therefore cover only the lower end of brightness temperatures. Hence, the harmonisation only operates at one calibration point without knowledge of the warm end of the brightness temperature range.

To overcome this problem, a promising method is being developed in Master's theses. By virtually matching two polar orbiting MW humidity sounders, with the geostationary instrument SEVIRI acting as a filter, the accessible brightness temperature range can be extended to warm temperatures, since these virtual match-ups occur in low-latitude subsidence regions. It is intended to include these virtual match-ups in the input for harmonisation. It is expected to yield better results, since the harmonisation then also operates on the warm temperature end.

The harmonisation is not a one-way procedure. Apart from the above mentioned important refinements, the procedure itself is considered to be iterative. Several combinations of possible harmonisation parameters need to be tested in the optimisation procedure in order to find the correct combination that reduces and explains the bias. This is the overall goal of the recalibration approach. Also, a different choice of reference may be considered for further harmonisation runs. This approach is reasonable because of the probable deficiencies of the reference instrument on NOAA-18 concerning its antenna pattern correction.

As pointed out above, I do not expect the current harmonisation procedure to work for the NOAA-15, NOAA-16 and NOAA-19 due to the impact of RFI. Nonetheless, by further investigation of the pattern seen in the scan-dependent bias analysis, it could be possible to devise an RFI correction scheme for each sensor to be applied to the counts. This would be an update and improvement of the RFI corrections devised during the in-orbit verification phase of NOAA-15. When formulating the required constraints, the new correction could be further improved by harmonisation. However, it should be kept in mind that the RFI-contamination of the calibration views is not accessible to the scan-dependent bias analysis.

The possible impact of RFI should also be kept in mind for the remaining MHS on NOAA-18 and the Metops, which are currently not affected by large RFI-induced biases. Note however, that I identified characteristic patterns of RFI. In case the gain reduces strongly, the impact of RFI may grow also for these instruments and consequently cause significant biases.

As an overall conclusion, I point out that the effort to generate an uncertainty quantified, harmonised FCDR of MW humidity sounder data is on a promising path. This thesis contributed to this major effort by the preparation of an uncertainty quantified FCDR in a user-friendly format combining measurand, uncertainty and quality information in a single file per equator-to-equator orbital frame. Also, the FCDR generator code is in place for easy adjustment for the generation of a harmonised FCDR. Even if the harmonised FCDR is not yet produced, detailed analyses of the instruments and the identification of a potential improvement in the calibration were achieved in this thesis. The harmonisation set-up by project partners has matured to a

point where it can generate results. For this harmonisation procedure, a significant improvement is expected by the expansion of the input data.

Building and testing the FIDUCEO harmonisation procedure to date has been focussed on the MHS and AMSU-B instruments. The inclusion of the more complicated case of the SSMT-2 instruments is expected in the future and would extend the available long-term data record.

A successful harmonisation and the subsequent production of a harmonised FCDR is a necessary condition for the generation of long-term, stable CDRs. From the MW humidity sounder FCDR, a UTH CDR will be derived. There is already ongoing work in the framework of a Master's thesis to derive a gridded UTH product with propagated uncertainties. This study is based on the unharmonised FCDR produced in this thesis to establish the processing chain while the harmonisation is still being improved. Finally, the UTH CDR will be created based on the harmonised FCDR from MW humidity sounders to obtain stable, long-term series of UTH data usable for climate research.

8 Appendices

8.1 Sensor time series: Evolution of basic calibration quantities

This section is a reprint of the appendix of the publication

> Hans, I., Burgdorf, M., John, V. O., Mittaz, J., and Buehler, S. A.: Noise performance of microwave humidity sounders over their lifetime, *Atmos. Meas. Tech.*, 10, 4927-4945, `https://doi.org/10.5194/amt-10-4927-2017`, 2017, Creative Commons Attribution 4.0 License

In the following we investigate the stability of the individual instruments flying on different satellites by looking at the long time series of the above-mentioned observables, mostly at the cold NEΔT as indicator of the overall noise. For every channel, we display the cold NEΔT over all considered missions in Fig. 4.11 from 0.1 to 5 K. We state which data we would definitely suggest to exclude, based on the rather high threshold of 1 K. The remaining useful periods are displayed in Fig. 4.13. We are interested in long-term evolutions in the sensor or sudden incidents impacting the instrument. Hence, the normal orbital variations are not investigated further, since their effect on the cold NEΔT, even in the case of stronger changes, is only very small by construction.

8.1.1 DMSP-F11 (SSMT-2)

The DMSP-F11 was launched in 1991. The NOAA CLASS data set starts at 01 April 1994 and ends on 24 April 1995 with some data gaps of several days or weeks. The time record exhibits some issues: Sometimes the time stamp indicating the seconds of the day is zero (without a change of day) or has values larger than 86400 s. The corresponding scan lines are excluded in our processing and do not enter the time series.

8.1.1.1 OBCT temperature (2 PRT)

Both PRT sensors show normal behaviour throughout the time range. The temperature on the black body changes in an interval of about 5K around 290 K.

8.1.1.2 Channels 1 and 2

Channel 1 has a stable gain and a low cold NEΔT of 0.2 K over the whole time range (see Fig. 4.11a, black line). Channel 2, however, is damaged from the start: the gain is constantly zero as the signal for the OBCT and DSV counts is the same. Hence, it is of no use for research. The cold NEΔT has infinite values and therefore does not appear in Fig. 4.11b.

8.1.1.3 Channels 3, 4 and 5

The gain is stable at about 10 Counts K^{-1}, except for some erroneous outliers between -5 and 10 Counts K^{-1} (very similar values for all three channels). In November 1994, there is a complete orbit of bad outlier data spreading between -5 and 10 Counts K^{-1}. The cold NEΔT is quite stable at around 0.3, 0.4 and 0.5 for channels 3, 4 and 5 respectively - except for the corresponding outliers of the gain (Fig. 4.11c-e, black line). From late 1994 and early 1995 on, cold NEΔT of channels 3 and 5 shows more frequently higher values around 1.3 K. This is due to a most peculiar aspect: there are jumps up and down in the OBCT and DSV counts within an orbit from the year 2001 on (it already appears before, but rather seldom): The orbital change has the expected smooth shape before it suddenly jumps to a higher or lower level and there continues its course. The origin of these jumps is unclear.

8.1.2 DMSP-F12 (SSMT-2)

The second SSMT-2 instrument was brought to its orbit on 28 April 1994. The NOAA CLASS data set runs from 13 October 1994 to 08 January 2001 with some data gaps of several weeks.

Beside the time record issues mentioned for F11, the instrument on F12 revealed more wrong time stamps for many data points: the time stamp goes back to some hours before the actual time and therefore produces artificial abrupt rises and drops in the time evolution of the observables. Hence, additionally to the filter used for F11, we use a second one excluding all data whose time stamp is smaller than the previous one.

8.1.2.1 OBCT temperature (2 PRT)

The PRT sensors do not show any peculiarity, except for several groups of outliers in 1994 (around 288 K) and more widely distributed outliers in 1999. Both PRTs show slight oscillatory changes in the black body temperature of about 4 K around an increasing mean of 300 to 304 K. In 1994 and later in 1999 there are several groups of outliers.

8.1.2.2 Channels 1 and 2

The lower peaking channels show the same behaviour as the water vapour channels described below, with similar values. They cannot be used for research purposes after 1999 either (see Fig. 4.11a,b, violet line).

8.1.2.3 Channels 3, 4 and 5

Apart from outliers, the gain is stable until 1999 at around 10 or 9 Counts K^{-1} for channels 4 and 5 and 3 respectively. The same holds for the DSV count noise and the cold NEΔT (0.43, 0.34, 0.38 K for channel 3,4,5). From later 1999 on, there are very many outliers that are rather widely spread such that cold NEΔT also reaches above 5 K for those data points and the remaining line of cold NEΔT around 0.4 K appears quite thin (see Fig. 4.11c-e, violet line). This makes the water vapour channels less suited for research purposes.

8.1.3 DMSP-F14 (SSMT-2)

The third SSMT-2 instrument was only launched on board the DMSP-F14 on 10 April 1997. The NOAA CLASS data set starts on 28 April 1997 and ends on 18 January 2005

8.1.3.1 OBCT temperature (2 PRT)

The black body temperature slightly oscillates with a period of about 6 months around a decreasing mean from 294 K to 292 K in 2005. Both PRT sensors agree within the random uncertainties throughout the investigated time frame.

8.1.3.2 Channels 1 and 2

The low-peaking channels show a similar behaviour as the water-vapor channels until 1998, when the instrument suffers from several issues, described in detail for channels 3 to 5. Channel 1 recovers from the critical 1998 phase and has a very low cold NEΔT at the level of the pre-1998 value of 0.3 K (Fig. 4.11a, blue line). Channel 2, however, does not recover after May 1998 - instead, the signal of the OBCT and DSV approach each other resulting in a strongly increasing cold NEΔT surpassing 1.5 K at the end of 1998. Afterwards, it even reaches 8 K before decreasing slightly again, but always staying above 6 K (Fig. 4.11b, blue line). Channel 2 is therefore only usable for 1997. Both channels also show the jumps of unclear origin, already mentioned for F11, but to a lesser extent than the water vapour channels described below.

8.1.3.3 Channels 3, 4 and 5

The gain remains stable at 10 and 9 Counts for channels 4 and 3, and 5, respectively, throughout the lifetime. However, during the first half of 1998 the instrument suffers from some incidents: several additional levels of gain emerge and the OBCT and DSV counts show extensive jumps. Thus it appears as if the levels of OBCT and DSV count signals split into two branches each. Then, the branches for DSV count signals approach those for OBCT. The resulting gain levels are lower than the original stable value or even close to zero which leads to many high peaks of cold NEΔT of even >1000 K (not visible in Fig. 4.11c-e). Data from this period, i.e. January - May 1998, should not be used. Apart from this period the cold NEΔT is quite stable with slight changes around 0.5 K for channels 3 and 4, or 0.4 K for channel 5. In 2001, however, cold NEΔT increases above 1.5 K for channel 4 (4.5 K for channel 3 and 0.7 K for channel 5) and stays at this high level (see Fig. 4.11d, blue line). This corresponds to the development of the DSV count noise: after 1998, the DSV count noise also increases from initially 4 Counts to 15 Counts, slightly at first, then more strongly in 2001 (even to 30 Counts in channel 3, and 8 Counts for channel 5). Then, the values fluctuate around this increased level. To some extent, this correlates with the more frequent appearance of the jumps in the OBCT and DSV counts within an orbit as mentioned already for F11. Due to the described increase of the cold NEΔT, channels 3 and 4 should not be used from year 2001 on. Channel 5 might be used with caution due to higher uncertainty resulting from the jumps.

8.1.4 DMSP-F15 (SSMT-2)

On 12.12.1999 the DMSP-F15 satellite was launched carrying the last SSMT-2 instrument. The NOAA CLASS data set encompasses the measurements from 24 January 2000 to 18 January 2005.

8.1.4.1 OBCT temperature (2 PRT)

Throughout the considered time frame, both PRT sensors indicate a stable, only slightly oscillating black body temperature around an increasing mean of 295 to 298 K.

8.1.4.2 Channels 1 and 2

After a stable phase at the beginning of the mission, the gain gets slightly unstable for channel 1 and smoothly increases from $7\,\mathrm{Counts\,K^{-1}}$ to $9\,\mathrm{Counts\,K^{-1}}$ before decreasing to $5\,\mathrm{Counts\,K^{-1}}$. Accordingly, cold NE$\Delta$T increases from $0.6\,\mathrm{K}$ to $0.8\,\mathrm{K}$. In February, March and September 2003, channel 1 suffers from very large noise $>5\,\mathrm{K}$. These periods should be excluded. Furthermore, in 2003, there is a second level of cold NEΔT values which the measurement jumps to and off, increasing from 3 to $4\,\mathrm{K}$ (Fig. 4.11a, light blue line). This pattern can be seen in the DSV count noise as well and relates to the same jumps of unclear origin as those mentioned below for the channels 3 to 5. These are also visible in the OBCT and DSV counts of channel 2. However, the gain for channel 2 already decreases from 2001 on, when the OBCT and DSV signals become similar. Accordingly, cold NEΔT rises and even reaches $5\,\mathrm{K}$. It does not decrease below $2.3\,\mathrm{K}$ afterwards (Fig. 4.11b light blue line). Hence, channel 1 could be used with caution due to some higher uncertainty, whereas channel 2 is of no use due to its large noise.

8.1.4.3 Channels 3, 4 and 5

The gain is quite stable at a constant value of $7\,\mathrm{Counts\,K^{-1}}$ for channel 4 ($8\,\mathrm{Counts\,K^{-1}}$ for channels 3 and 5), but has many outliers even down to a negative gain of $-3\,\mathrm{Counts\,K^{-1}}$. Cold NE$\Delta$T is mostly stable at $0.5\,\mathrm{K}$ ($0.6\,\mathrm{K}$ for channels 3 and 5). In 2003, cold NEΔT temporarily increases in channel 3 to $1.5\,\mathrm{K}$, but decreases again to $0.8\,\mathrm{K}$ (see Fig. 4.11c, light blue line). Channels 4 and 5 remain quite stable (Fig. 4.11d-e, light blue line). However, from the start, the jumps of unclear origin, mentioned for the surface channels above and for F11 and F14, appear in channels 4 and 5 and make the DSV count noise as well as the cold NEΔT change suddenly between two courses.

8.1.5 NOAA-15 (AMSU-B)

On 13 May 1998 the NOAA-15 satellite was launched having the first AMSU-B sensor on board as a subunit of the AMSU instrument. The operational data start on 15 December 1998. The instrument was turned off on 28 March 2011 (NOAA-OSPO, 2015), but already in late 2010 the data are too noisy to be used. Here, we investigate the NOAA CLASS data set from the start of operational data until the end of 2010. AMSU-B was turned off due to problems with the scan motor making measurements impossible. However, there are still data records being sent to Earth which cannot be used, of course, since these contain no measurement data but random numbers.

The NOAA-15 satellite started with an LECT of about 19:30, reached about 16:30 in 2010 and drifted back to 18:00. Its quick orbital drift over its lifetime impacted on the AMSU-B instrument: a characteristic pattern of peaks and drops becomes visible in the time evolution of many observables from 2002 on (see also NEΔT in Fig. 4.11 and 4.12). The same pattern can also be seen for the Microwave Sounding Unit (MSU) instrument on the earlier NOAA-14 satellite (Grody et al., 2004), for the AMSU-A on NOAA-15 (Zou and Wang, 2011) as well as on the AMSU-A and AMSU-B on board the successor satellite NOAA-16 (see below) which has already experienced the same strong orbital drift as other NOAA satellites. In Zou and Wang (2011), focusing on AMSU-A, a connection of this pattern to a changing solar beta angle due to orbital drift is seen. This angle is defined as the angle between the vector from Earth to Sun and the orbital plane of the satellite. Hence, a changing angle will influence the exposure of the

instrument to the Sun and may therefore impact its performance. An investigation of this is beyond the scope of this overview of microwave data.

The AMSU-B on NOAA-15 also suffered from the radio frequency interference (RFI), with channels 2 and 4 being impacted most. It introduced a scan-dependent bias that also affected the DSVs as well as the 90 Earth views. The impact was not constant in time, however. For example, in the period of October 1998 to September 1999, the measurements are biased for half the orbit before returning to normal behaviour for the rest of the orbit (Atkinson, 2001). This is also visible in the cold $NE\Delta T$ of channels 2, 4 and 5 (see Fig. 4.11b,d,e, dark green line).

8.1.5.1 OBCT temperature (7 PRT)

From the start, the black body shows strong variations of temperature (5 to 8 K) on monthly scale. Moreover, there are many drops to 262 K which are probably related to the PRT sensors. All seven sensors mostly agree throughout the lifetime, apart from some events where they drop or jump to different temperature levels. There are also many randomly distributed outlier values of the different PRT sensors. From 2002 on, the orbital drift induced the changing pattern mentioned above, which becomes clearly visible and remains until the end of the data set.

8.1.5.2 Channels 1 and 2

The counts for the OBCT and DSV are quite stable, except for small changes on the monthly scale. However, the counts often drop to zero (either for both targets or for one of them) which results in constant levels of outliers in the gain at -60, 0 or 100 Counts K^{-1}. Yet, apart from some random outliers the gain is mostly stable at its initial value of 30 and 20 Counts K^{-1} for channels 1 and 2 respectively. The changing pattern mentioned above becomes more pronounced in the course of time, but as the OBCT and DSV counts almost change accordingly there are only very small changes in the gain (~ 1 Count K^{-1}) and no decline. The cold $NE\Delta T$ remains quite stable at 0.25 K (channel 1) and 0.6 K (channel 2), see the dark green line in Fig. 4.11a and b, respectively. Filtering out the scan lines of outlier values, and excluding channel 2 from start until November 2000, when a phase of unstable cold $NE\Delta T$ ends, will provide a useful data set.

8.1.5.3 Channels 3, 4 and 5

The water vapour channels are subject to more quality issues. From the start, one can observe slowly decreasing counts for the DSV signal and quicker decreasing for the OBCT counts. For the first years up to the end of 2001, the resulting gain still has acceptable values and cold $NE\Delta T$ is about 1 K for channel 3 or 0.8 K and 0.6 K for channels 4 and 5, respectively. From 2002 on, however, the changing pattern as seen in the black body temperature shines through also to the cold $NE\Delta T$ and the degradation gets stronger. The recorded signals for OBCT and DSV approach each other until the gain becomes very small (below 6 Counts K^{-1} for an initial value of 20 Counts K^{-1}) and, consequently, the cold $NE\Delta T$ rises above 2.5 K. Finally in the middle of September 2010, the gain drops to zero resulting in NAN values for cold $NE\Delta T$. Data should not be used for channel 3 from 2001 on, for channel 4 from 2004 on or for channel 5 from 2007 on, as the cold $NE\Delta T$ increases beyond 1 K (see Fig. 4.11c-e, dark green line).

8.1.6 NOAA-16 (AMSU-B)

The second AMSU-B instrument was sent to space on board the NOAA-16 satellite on 21 September 2000. The operational data started on 20.03.2001. Finally, NOAA-16 was decommissioned on 09 June 2014. Compared to its predecessor, NOAA-16 was exposed to an even stronger orbital drift from about 14:00 to 22:00 LECT (see Fig. 2.7). In 2007, the earlier mentioned changing pattern for the observables emerges, probably related to the solar beta angle (see above, Zou and Wang (2011)). This is visible in NEΔT, too (see Fig. 4.11 and 4.12).

8.1.6.1 OBCT temperature (7 PRT)

The black body temperature only shows small oscillations on a monthly scale, reaching about 4 K in late 2006, though. As for NOAA-15, the PRT sensors also often drop to 262 K. There are also periods of months, when the PRT sensors differ by about 10 K for several orbits. Then, from October 2007 on, the variations in the overall evolution become more severe as the strong changing pattern becomes visible with an amplitude of 5 to 10 K. In 2012, the pattern ceases and only small changes around 288 K can be seen.

8.1.6.2 Channels 1 and 2

The low peaking channels show quite acceptable data having a cold NEΔT of 0.3 K. Nonetheless, over the whole lifetime, the OBCT and DSV counts often drop to zero or jump to other quite stationary levels (especially from 2004 on). This is transported to the gain and also causes outliers of up to 2 K in cold NEΔT. In channels 1 and 2 the changing pattern is very faint and only changes the gain by about $\pm 1\%$. Therefore, the cold NEΔT also appears stable at the scale of Fig. 4.11a,b, green line).

8.1.6.3 Channels 3, 4 and 5

Initially, the gain is rather stable for the three water vapour channels. A slight decreasing starts in early 2001 after higher orbit-to-orbit variations that can be seen in OBCT and DSV counts, as well. In 2002, the OBCT counts start to decrease morw quickly than the DSV counts, and hence the gain decreases continuously. Four years later, in 2006, the gain has decreased from initially 22 Counts K^{-1} down to 9 Counts K^{-1} in channel 3 (the other channels show a similar evolution) and cold NEΔT has risen from 0.6 K to 1.4 K. The degradation for the three channels continues further as the gain decreases (OBCT and DSV counts getting close to one another) and cold NEΔT increases. From late 2007 on, the changing pattern shines through in the counts and the cold NEΔT (see Fig. 4.11c-e, green line) reaches 18 K in 2011, when the gain approaches zero, and increases beyond 50 K in 2014 as the signal recorded for the OBCT and DSV is basically the same. Doing a two-point calibration is not sensible at this stage and produces completely useless data due to absurdly high noise with cold NEΔT > 10 K. One should stop using NOAA-16 data from the end of 2005 when cold NEΔT surpasses 1 K and degradation keeps advancing in channels 3-5.

8.1.7 NOAA-17 (AMSU-B)

On board NOAA-17 the last AMSU-B instrument was launched on 24 June 2002. Its operational data set starts on 15.10.2002 and ends on 10 April 2013. NOAA-17 drifted from about 22:00 to 19:00 LECT over its mission (Fig. 2.7).

8.1.7.1 OBCT temperature (7 PRT)

The seven PRT sensors indicate a stable black body temperature softly oscillating on the yearly scale around 285 K (slightly increasing to 287 K). As for the other AMSU-B instruments, the PRT measurements also often drop to 262 K. In 2010, the overall evolution remains, but the measured values of the seven sensors jump between discrete levels and follow the overall evolution with different constant offsets. There also appear strong peaks from 2011 on, a sharp drop to 275 K in early 2013 and then an increase again.

8.1.7.2 Channels 1 and 2

Until 2007, channel 1 has a stable gain, cold NEΔT and DSV count noise. Then, sharp peaks (of factor > 4 to stable noise value) appear in the DSV count noise. Later the peaks reach even a value of factor of 10 times the stable noise value and outliers even a factor of > 50. Moreover, the peaks become more frequent such that the underlying constant DSV count noise of initially 8 Counts becomes less visible. Hence, channel 1 gets very noisy (cold NEΔT peaks reach up to 5 K) due to the DSV count noise that transfers to the overall cold NEΔT, see Fig. 4.11a, light green line. The gain is also impacted from the high DSV count noise peaks, since the DSV counts apparently have a larger variation that becomes visible in jumps and drops of the gain to certain levels whilst keeping the overall initial value of 24 Counts K^{-1}. Channel 2 shows a similar behaviour, though less pronounced, i.e. the frequency of the appearing peaks is smaller (see Fig. 4.11b light green line). Filtering out the scan lines of outlier values will lead to a usable data set for channel 2. Channel 1 also needs filtering, but from 2007 on, one should not use the data at all, since they get too noisy, as described at the beginning of the paragraph.

8.1.7.3 Channels 3, 4 and 5

Apart from small jumps and drops in channels 3 and 4 in 2003 and 2004, all three channels have stable cold NEΔT values of 0.85 K, 0.7 K and 0.8 K, respectively. In December 2009 however, a sharp drop of both OBCT and DSV counts results in a gain of almost zero and a huge cold NEΔT of 2000 K or infinite (NAN) values (see Fig. 4.11c-e, light green line). From December 2009 on, the NOAA-17 AMSU-B data for the sounding channels cannot be used for any research questions.

8.1.8 NOAA-18 (MHS)

The first MHS instrument was installed on board the NOAA-18 satellite launched on 20 May 2005. The operational data set starts on 30 August 2005. The mission is still ongoing; however, our data set for investigation ends in May 2016. From its start until May 2016 it drifted from 14:00 LECT to 18:00.

8.1.8.1 OBCT temperature (5 PRT)

The five PRT sensors agree in the slight oscillations onthe yearly scale of the black body temperature around 284 to 287 K. Apart from a few outlier values of several PRTs, the measurements are quite stable and show a stable black body temperature. However, in August 2014, the strong changing pattern as seen for the NOAA-15 and 16 satellites emerges and leads to maximum (minimum) temperature of 298 K (270 K). This pattern is still visible at the end of the used data set in May 2016.

8.1.8.2 Channels 1 and 2

Apart from outlying values, both channels 1 and 2 have a stable gain and cold NEΔT around 0.14 K and 0.36 K, respectively, over the lifetime (see Fig. 4.11a,b, dark red line). The changing pattern visible in the black body temperature is only prominent in the OBCT and DSV counts that change accordingly, thus resulting in a stable gain.

8.1.8.3 Channels 3, 4 and 5

A first, all three channels show a stable gain (in the range of 140 Counts K^{-1}), with small discrete jumps and drops. The orbital variation around the mean is larger than for channels 1 and 2, often about ± 10 Counts K^{-1}, and also shows changes over the years. Channel 5 has very large orbital variation in 2011 and 2012 and also significant changes in the DSV count noise for these periods, but then it is suddenly reduced by a factor of 20 by controlled gain adjustment (Bonsignori, 2007). Thus, channel 5 is less variable from 2013 on. The changing pattern is apparent in the gain in 2014: its strongest impact is on channel 3 (up to 90 Counts K^{-1} within a month), then on channel 4 and finally on channel, 5 where it is hardly visible. Cold NEΔT is also stable at first (0.5 K, 0.4 K, 0.3 K, for channels 3 to 5), but also shows the jumps in the gain and increases slightly until, in 2014, the changing pattern becomes visible and increases or decreases cold NEΔT (Fig. 4.11c-e, dark red line). Temporarily, cold NEΔT reaches 0.95 K in channel 3 (0.8 K for channel 4, whereas channel 5 remains stable since 2013 at 0.3 K). It is a usable data set, but one should be aware of the temporarily increased noise and therefore larger uncertainty for all three channels. Channel 5 has the fewest problems from 2013 on.

8.1.9 NOAA-19 (MHS)

On 06 February 2009 the NOAA-19 satellite was launched carrying the second MHS instrument. The operational data start on 02 June 2009. So far, NOAA-19 has drifted from 14:00 to 15:00 LECT. It is still operational.

8.1.9.1 OBCT temperature (5 PRT)

All five PRT sensors measure the same stable temperature of the black body, oscillating slightly on the yearly scale around 285 K.

8.1.9.2 Channels 1 and 2

Throughout the lifetime both channels are stable and have a constant cold NEΔT of 0.13 K and 0.33 K respectively. In Fig. 4.11a,b, the corresponding red line is directly behind the orange one of Metop-A.

8.1.9.3 Channels 3, 4 and 5

Channels 3 and 4 begin stable, but show erratic behaviour in July 2009. The OBCT and DSV signal suffer from major incidents, resulting in a strongly diminished gain. Following the drop in the gain, cold NEΔT increases from 0.5 K to 3.4 K in channel 3 (Fig. 4.11c, red line). Yet, channel 4 recovers from the incidents in 2009 and then remains stable at 0.58 K (Fig. 4.11d, red line). Channel 5 is stable throughout the mission having a low cold NEΔT of 0.27 K (Fig. 4.11e, red line). From the data set of NOAA-19, channel 3 should not be used.

8.1.10 Metop-A (MHS)

The third MHS instrument was carried to orbit on board the Metop-A satellite launched on 19 October 2006. The operational data start on 15 May 2007. The instrument is still active. Unlike the NOAA satellites, the Metop satellites do not exhibit orbital drift. Their local equator crossing time remains stable at 21:30.

8.1.10.1 OBCT temperature (5 PRT)

The temperature of the black body is quite stable over the mission so far and shows small variations on a 3-monthly scale around 283 K. There are a few orbits with outlier values, mainly in the first years of the mission and there is a larger data gap in spring 2014.

8.1.10.2 Channels 1 and 2

Both channels do not show any anomalies and remain stable at their initial cold NEΔT values of 0.13 K and 0.31 K respectively (see Fig. 4.11a,b, orange line). The latter one increases slightly to 0.34 K in 2016.

8.1.10.3 Channels 3, 4 and 5

The gain is constantly adjusted during operation to correct for decreases and increases and to keep it within certain limits (Bonsignori, 2007). Overall, the resulting cold NEΔT is quite stable around 0.5 K or 0.6 K for channel 3, peaking at 0.7 K in late 2011. For channel 4 there is a slightly lower noise of 0.3 K, peaking at 0.5 K in late 2011. Channel 5 is stable throughout the mission with low cold NEΔT of 0.27 K (see Fig. 4.11c-e, orange line). As for channel 5 of the MHS instrument on NOAA-18, the DSV count noise changes over the mission in channels 3 and 4. This is visible in the cold NEΔT as well.

8.1.11 Metop-B (MHS)

On 17 September 2012 the Metop-B satellite was launched with the fourth MHS instrument on board. The first operational data are available for 29 January 2013 when it replaced the Metop-A for operational purposes (WMO-OSCAR, 2016). The mission is envisaged to end after 2018. As Metop-A, Metop-B has no orbital drift either.

8.1.11.1 OBCT temperature (5 PRT)

Until the end of the considered time frame (May 2016), the temperature of the black body varies with an amplitude of about 2 K on a 3-monthly scale around 281 K. There are only four events of outlier values so far.

8.1.11.2 Channels 1 and 2

A small decrease of the gain can be observed for channel 2. However, this degradation is always corrected for by adjusting the gain and resetting it to higher values. The cold NEΔT is stable at 0.18 K for channel 1 and 0.36 K for channel 2 (see Fig. 4.11a,b, yellow line).

8.1.11.3 Channels 3, 4 and 5

The adjustment of the gain to keep it at a quasi-constant level is also prominent for channels 3 to 5 (with the smallest adjustments for channel 5). The cold NEΔT is stable at $0.35\,\text{K}$, $0.27\,\text{K}$ and $0.25\,\text{K}$ for channels 3, 4 and 5 respectively (see Fig. 4.11c-e, yellow line).

8.2 The FCDR production code

8.2.1 Overview of functions

In the following, I present the dependencies of the used scripts and functions. Note that the MW-FCDR generation code requires the open-source tool *atmlab* (atm, 2018), a collection of remote-sensing related helpful functions implemented for MATLAB. Note that some *atmlab* functions such as `invcm2hz.m` are not explicitly listed here.

The script `set_coeffs.m` is not part of the FCDR generator itself, but it is required to set certain instrumental calibration parameters before the generation code is executed. The constants are then stored in the files `coeffs_SATELLITESENSOR_antcorr_alpha.mat` that are read during the calibration (replace SATELLITE and SENSOR by the name of the satellite and sensor, respectively).

The script `run_batch_mission.m` is used to send several jobs to the cluster. It calls the function `FCDR_generator(satellite, sensor, year, month)`. This function is the enveloping function to produce FCDR files for a certain instrument and month. Within this function, the two main blocks of orbit-definition and orbit-processing are executed. The first block for consolidating the data set consists of (note: an indented script (S) or function (F) is called by the previous script/function with smaller indent):

MW_l1b_read_monthly (F)

    find_granules_for_period (F)

    satreaders.poes_radiometer_level1b (F)

        read_MW_level1b (F)

            read_MHS_allvar(F)

                read_MHS_header(F)

                read_MHS_record(F)

            read_AMSUB_allvar(F)

                read_AMSUB_header(F)

                read_AMSUB_record(F)

            read_SSMT2_allvar(F)

                ssmt2_read(F)

                read_SSMT2_record(F)

            doy2date(F) (by A. Booth)

            mills2hmsmill(F)

remove_duplicated_l1bdata (S)

RunLength_M(F)

determine_EQcrossings (F)
determine_data_gaps (F)

After this first block, having defined all equator crossings and data gaps, the function `FCDR_generator` continues with the second block for processing and writing:

process_FCDR (S)

    chop_data_EQ2EQ (F)

    setup_MHS_processing2l1c (S)/ ... AMSUB.../...SSMT2...

        fill_missing_scanlines (F)/ fill_missing_scanlines_SSMT2 (F)

        doy2date(F) (by A. Booth)

        calculate_solarAndsatellite_azimuth_angle (S)

        calculate_AllanDeviation_DSV_IWCT_withoutQualflags (S)

        mooncheck_processing (S)

        qualitychecksDSV_allchn (S)

            filter_plateausANDpeaks(F)

        qualitychecksIWCT_allchn (S)

            filter_plateausANDpeaks(F)

        qualitychecksPRT_allsensors (S)

            filter_plateausANDpeaks(F)

        qualitychecksEarthview (S)

            filter_plateausANDpeaks_earth(F)

        calculate_AllanDeviation_DSV_IWCT_PRT (S)

        setup_roll_average (S)

            unify_qualflags(F)

        calculate_roll_average (S)

        qualitychecksEarthLocation (S)

        qualitychecksSpaceViewLocation (S)

        qualitychecksIWCTViewLocation (S)

    prepare_empty_file (S)

    write_easyFCDR_orbitfile_MHS (S)/ ...AMSUB/ ...SSMT2

        change_type (F)

        change_type_zero_nan(F)

    measurement_equation (S)

        planck(F)

> invplanck(F)

uncertainty_propagation_optimized (S)

> DplanckDf(F)

> DplanckDT(F)

> DinvplanckDrad(F)

quality_flags_fromlevel1b (S)

> fill_missing_scanlines(F)/ fill_missing_scanlines_SSMT2(F)

quality_flags_setting_easyFCDR (S)

add_RFI_uncertainty (S)

estimate_correlation_cross_channel (S)

write_easyFCDR_orbitfile_MHS (S) / ...AMSUB/ ...SSMT2

> change_type (F)

> change_type_zero_nan(F)

All processed orbit files are stored in directories `easy/vXfvY/satellite/year/month/day`, together forming the MW-FCDR for data version vX and format version fvY.

8.2.2 Description of preparatory steps

8.2.2.1 Collecting and reading data

The FCDR generator is designed to work on a full month of level 1b data for a certain instrument on a certain satellite. Hence, the first step for the reading of that whole month is to collect all data paths for this instrument and month. From the chosen instrument, satellite, year and month that are specified as arguments of `FCDR_generator` and passed on to `MW_l1b_read_monthly`, the atmlab function `find_granules_for_period` collects all corresponding data paths in a list. Then, this list is looped over: The each element of this list is taken by the atmlab function `satreaders.poes_radiometer_level1b` (adapted from its original form `poes_radiometer` for the calling of the full level 1b data-reader). The child-function `read_MW_level1b` allows for flexible reading of level 1b microwave data files: Depending on the choice of instrument, the function calls the corresponding level 1b reader `read_MHS_allvar`, `read_AMSUB_allvar` or `read_SSMT2_allvar`. Within these readers, internal functions for the reading of the header and data record are called. The data read from each file is stored in a structure. Before it is appended to the previous file's structure in the parent function `MW_l1b_read_monthly`, a field "time" is created within the data structure to take the converted time in seconds since 1970-01-01 00:00. This time variable is required for unique identification of each recorded line over the month.

8.2.2.2 Removing duplicated data

Having the concatenated data sets now building up the full monthly data record, the cleaning for duplicated data is executed in `remove_duplicated_data`. Within the monthly data record, I introduced an indicator per scan line to identify which original file it belongs to. Using this information, the algorithm looks for duplicates of the end time of one file and the start time of the next ten following files and marks the detected duplicated scan lines. This is necessary

to both identify the normal overlap between subsequent files and to mark the start lines of all fractional orbit files: it often happens that there exist some data files containing only fractions of a corresponding complete data file. These fractional orbit files have the same start time or end time as their corresponding full file, and hence they are listed before or after the complete file. The algorithm therefore has to identify the longest file corresponding to a certain start time and mark it as to be used, whereas all fractional files after the complete one are marked for deletion. Fractional files before the complete file are marked for deletion since they are in between the end line of one file and the new start line of the complete file and are therefore interpreted as the normal overlap. Iterating over all orbital files gives a list of scan lines and files to be removed from the monthly data record. This deletion is executed after the list has been completed with the last iteration. The result is a monthly data set without overlap between files.

8.2.2.3 Finding equator crossings

The choice on which flight node (ascending/descending) should be used for each instrument (see Sec. 5.3.1) is implemented in `determine_EQcrossings` based on the argument for the current satellite. Within `determine_EQcrossings`, the equator crossings are determined from the changes of sign of the latitude. As indicating latitude, I use the latitude of the virtual centre pixel, i.e. the mean value of the latitude of the two innermost pixels (45 and 46 for AMSU-B and MHS and 14 and 15 for SSMT-2). The flight direction is obtained from the sign of the change of the latitude: greater than zero for northbound, i.e. ascending, and smaller than zero for southbound, i.e. descending (the flight direction is already stored in AMSU-B and MHS level 1b files and could be used from there, but not in SSMT-2 level 1b files. To keep the code more general, it uses the explained criteria for all three instruments). The scan line numbers for each equator crossing within the month are stored for the next steps.

8.2.2.4 Finding critical data gaps

As critical data gaps I consider those gaps that start in one equator-to-equator frame, but end in another. In this case, the algorithm would concatenate two data sets that might be even several orbits apart. To prevent this, the code must be able to detect critical data gaps and set the end of one file to the start of the gap, and the start of the next file to the end of the gap.

The function `determine_data_gaps` handles this as follows: From calculating the time differences (in seconds since 1970-01-01 00:00) of adjacent scan lines, the code detects jumps in time. It then checks which of the scan lines associated with these jumps have been collected among the equator crossings already (this wrong-classification as equator crossing happens, if the satellite actually crossed the equator during the gap). Hence, two groups of scan line candidates for data gaps are identified: those that have been classified erroneously as equator crossings, and those that were not interpreted as equator crossing because no change of sign of latitude has happened over the gap. For the first group, the found candidates are removed from the list of equator crossings and they are kept as start lines of data gaps. For the second group, further checks are made: first, it needs to be checked whether the data gap ends within the time scale of one orbit, or whether it only ends later. In the former case, the scan line candidate is deleted from the list: no new file will be started since the gap is fully contained within one orbit. In the latter case, the scan line is a true candidate for a critical data gap. A last check ensures, that there was no scan line shortly before the one in question that was in the list of candidates. If there is one, then there is a cluster of data gaps after a critical one. In this case, only the first line detected

from this cluster is saved as scan line before a critical data gap. Finally the scan lines that had been classified erroneously as equator crossings and the scan line candidates having passed the tests are stored in a common set of scan lines before a critical data gap.

Having found all equator crossings and start lines of large data gaps, this information needs to be merged to give one set of scan lines to be used as start and end points for the equator-to-equator files. The merging is done within the main function `FCDR_generator`: A matrix C (Eq.(8.1)) is created that stores the sorted scan lines of the equator crossings and the start lines of the gaps. In the second and third row, the matrix stores the corresponding start-offset values (the line is used as start of file) and end-offset values (the line is used as end of file) for each line, respectively, that get subtracted from or added to the scan line number when the data is chopped into equator-to-equator pieces.

$$C = \begin{bmatrix} EQcr1 & EQcr2 & ... & EQcr42 & Gap1 & EQcr43 & ... \\ 3 & 3 & ... & 3 & -1 & 3 & ... \\ 3 & 3 & ... & 3 & 0 & 3 & ... \end{bmatrix} \tag{8.1}$$

For the equator crossings the end- and start offset is 3. This accounts for the fact that the calibration requires the three lines before and after a certain scan line and hence, three lines before the start and after the end of one file must be stored so that the full calibration data is available also for the first and last line of the file. For the start lines of a data gap it is different. The actual end of a file must be precisely at the start line of the gap. Hence, the end-offset is 0 which will be added to the scan line value. The actual start of a file after a gap must be the start line of the gap +1 to securely overcome the gap. Hence, the start-offset which will be subtracted from the scan line value is −1, to yield the required +1. This matrix C containing all information for the start/ end lines of each equator-to-equator file to be created within this month is then used in the next block for the actual preparation of the orbital files and the processing (see Alg. 1).

> **for** *EQfile=EQcr1 to EQcrEND* **do**
> > startline = C(1,EQfile)
> > start_offset = C(2,EQfile)
> > startoffile = startline - start_offset
> > endline = C(1,EQfile+1)-1
> > end_offset = C(3,EQfile+1)
> > endoffile = endline + end_offset
> > data(EQfile) = data(startoffile to endoffile)
>
> **end**

Algorithm 1: Chop data into equator-to-equator frames taking into account significant data gaps. See Eq. 8.1 for matrix C. This separation into the equator-to-equator files is done in the second block in `process_FCDR /chop_data_EQ2EQ`.

8.2.3 Description of processing steps to l1c

8.2.3.1 Weighted rolling 7-scan line average

For the evaluation of the measurement equation, the calibration data from DSV, IWCT and PRTs need to undergo the weighted rolling 7-scan line average on usable scan lines only. Excluding certain scan lines means, that the weights need to be adapted so that the sum of the weights still equals 1. Hence, for every scan line to be calibrated, the three lines before and after need to be checked for usability. If any of these lines is not usable, the weight that should

have been assigned to it must be distributed among the other lines to give new weights. Having collected all usable lines and computed their weights, the scan line in question can receive its 7-scan line averaged value for the three quantities (two of them per channel even). This procedure would need to be applied for all scan lines of the orbit file. In order to implement this efficiently in MATLAB, I have developed a strategy using (sparse) matrices. In the following I explain this strategy. The explanations are also included as comments in the documented code `calculate_roll_average.m`.

The overall aim is to compute the 7-scan line average as a matrix-vector multiplication

$$\overline{\mathbf{x}} = \mathbf{G}\mathbf{x} \tag{8.2}$$

of a matrix $\mathbf{G}$ containing the adapted weights for the individual lines and a vector $\mathbf{x}$. $\mathbf{x}$ is of length n, with n being the total number of scan lines of the current orbit file. $\mathbf{x}$ contains the pre-calculated mean over the usable view/ sensors for the DSV, IWCT and PRT measurements, respectively, i.e. $\mathbf{x} \in \{C_{\mathrm{DSV}}, C_{\mathrm{IWCT}}, T_{\mathrm{PRT}}\}$ for a single channel.

To achieve this form of equation, several intermediate steps are required for efficient computation:

First, I create a matrix $\mathbf{A}$ with dimensions n x n. It contains only ones for each line that is theoretically used to compute the 7-scan line average for a certain line. Hence, $\mathbf{A}$ is a banded matrix with ones on the three upper and lower diagonals and the main diagonal:

$$\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & \ddots \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ \vdots & & & & & & \ddots & & \ddots \end{pmatrix} \tag{8.3}$$

$\mathbf{A}$ is stored as sparse matrix. As second matrix, I create $\mathbf{W}$, having the same structure as $\mathbf{A}$, but containing the nominal weights for the 7 scan lines:

$$\mathbf{W} = \begin{pmatrix} 0.2500 & 0.1875 & 0.1250 & 0.0625 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\ 0.1875 & 0.2500 & 0.1875 & 0.1250 & 0.0625 & 0 & 0 & 0 & 0 & 0 \\ 0.1250 & 0.1875 & 0.2500 & 0.1875 & 0.1250 & 0.0625 & 0 & 0 & 0 & 0 \\ 0.0625 & 0.1250 & 0.1875 & 0.2500 & 0.1875 & 0.1250 & 0.0625 & 0 & 0 & 0 \\ 0 & 0.0625 & 0.1250 & 0.1875 & 0.2500 & 0.1875 & 0.1250 & 0.0625 & 0 & 0 \\ 0 & 0 & 0.0625 & 0.1250 & 0.1875 & 0.2500 & 0.1875 & 0.1250 & 0.0625 & 0 \\ 0 & 0 & 0 & 0.0625 & 0.1250 & 0.1875 & 0.2500 & 0.1875 & 0.1250 & 0.0625 & \ddots \\ 0 & 0 & 0 & 0 & 0.0625 & 0.1250 & 0.1875 & 0.2500 & 0.1875 & 0.1250 \\ 0 & 0 & 0 & 0 & 0 & 0.0625 & 0.1250 & 0.1875 & 0.2500 & 0.1875 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.0625 & 0.1250 & 0.1875 & 0.2500 \\ \vdots & & & & & & \ddots & & & \ddots \end{pmatrix} \tag{8.4}$$

W is also stored as sparse matrix.

Third, I prepare a matrix **U** (one for each channel), again of dimension n x n, indicating which scan lines can be used (based on missing lines and the quality checks identifying bad lines). This matrix **U** is not sparse. Exemplarily, I show **U** for the case that scan line 3 and 9 cannot be used:

$$\mathbf{U} = \begin{pmatrix} 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & \cdots \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & \ddots \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \\ 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 1 & \\ \vdots & & & & & & \ddots & & & \ddots \end{pmatrix} \tag{8.5}$$

In the next step a matrix **D** is calculated, being the same as **W**, but unusable lines are removed:

$$\mathbf{D} = \mathbf{W} * \mathbf{U} \tag{8.6}$$

The asterisk * stands for element-wise multiplication:

$$D_{ij} = W_{ij} \cdot U_{ij} \tag{8.7}$$

Hence, **U** removes all weights for unusable lines.

The weights of the unused lines now need to be distributed evenly among the remaining lines that are actually used in the ± 3 scan line frame around the line in question. To achieve this, I collect per scan line the weights of the *un*usable lines. This is stored in matrix **K**:

$$\mathbf{K} = \mathbf{W} - \mathbf{D} \tag{8.8}$$

To evenly distribute the weights of the unused lines, I need to calculate the sum of these unused weights per scan line, which I store in the vector **s**, and I need to count the number of used lines, which I store in the vector **m**.

$$\mathbf{s} = \text{sum of unused weights, per scan line} = \text{sum of elements of } \mathbf{K}\text{'s rows} \tag{8.9}$$

$$\mathbf{m} = \text{number of used lines, per scan line} = \text{non-zero elements of } \mathbf{D}\text{'s rows} \tag{8.10}$$

Note that I have to set **m** to 1 artificially, where **m** would be zero. This is because I have to divide by **m** in the next step. The extra contribution that I add wrongly here, is removed later on by the element-wise multiplication with the matrix of used lines. This operation destroys all non-needed values.

Now, the share is calculated that needs to be added to the used lines' weights:

$$p_i = \frac{s_i}{m_i} \tag{8.11}$$

or

$$\mathbf{p} = \mathbf{s} * /\mathbf{m} \tag{8.12}$$

where I indicate element-wise division with */ to yield a vector $\mathbf{p}$ of the same length as $\mathbf{s}$ and $\mathbf{m}$.

As intermediate step, I create a matrix that contains all lines that are actually used for the average, i.e. they are usable from a quality check perspective *and* they are needed in a certain ± 3 scan lines frame:

$$U'_{ij} = \begin{cases} 1, & \text{for } U_{ij} = 1 \land A_{ij} = 1, \\ 0, & \text{else.} \end{cases} \tag{8.13}$$

This is achieved by using the logical operator AND on matrices $\mathbf{U}$ and $\mathbf{A}$:

$$\mathbf{U}' = \mathbf{U}\&\mathbf{A} \tag{8.14}$$

Now, I can obtain the new weights from the old weights plus the computed share $\mathbf{p}$. The vector $\mathbf{p}$, of length n is expanded artificially to a n x n matrix $\mathbf{P}$ to enable element-wise multiplication (symbol *) with the matrix $\mathbf{U}'$ of used lines. Thus, the additional shares are added correctly and any wrongly introduced contributions from unusable lines (where I set $\mathbf{m}$ to 1) are removed in this step:

$$\mathbf{G} = \mathbf{D} + \mathbf{P} * \mathbf{U}' \tag{8.15}$$

Finally the weighted rolling average is computed in the desired form:

$$\overline{\mathbf{x}} = \mathbf{G}\mathbf{x} \tag{8.16}$$

This procedure is applied for $\mathbf{x}$ being the counts on DSV or IWCT (each per channel), and being the temperature of the PRTs. In Fig. 5.11, the orange crosses display the adjusted weights for the case of three unusable lines before the scan line in question.

Bibliography

(1995). *Evaluation of measurement data – Guide to the expression of uncertainty in measurement (GUM) JCGM 100:2008*. Joint Committee for Guides in Metrology (JCGM). `http://www.bipm.org/en/publications/guides`, last access: 15.07.2018.

(2004). Defense Meteorological Satellite Program (DMSP) Appendix C.2. `https://eosweb.larc.nasa.gov/project/fire/fire3_ace/ACEDOCS/data/appen.c.2.pdf`. last access: 21.05.2018.

(2008). *International Vocabulary of Metrology – Basic and General Concepts and Associated Terms (VIM) JCGM 200:2008*. Joint Committee for Guides in Metrology (JCGM), 3rd edition edition. `http://www.bipm.org/en/publications/guides`, last access: 15.07.2018.

(2018). atmlab tool and documentation. `http://www.radiativetransfer.org/tools/`.

Ackermann, J. (2017). Dataset Description. MHS Instrument Pre-Launch Characteristics. `https://www.eumetsat.int/website/home/Satellites/CurrentSatellites/Metop/MetopDesign/MHS/index.html`. Last access: 15.07.2018.

Allan, D. W. (1966). Statistics of atomic frequency standards. *Proceedings of the IEEE*, 54:221–230.

Allan, D. W. (1987). Should classical variance be used as a basic measure in standards metrology? *IEEE Trans. Instrum. Meas.*, 36:646–654.

Atkinson, N. C. (2000). Performance of AMSU-B Flight Model 2 (FM2) during NOAA-L Post Launch Orbital Verification Tests, AMB112. Technical report, Remote Sensing Instrumentation, Met Office.

Atkinson, N. C. (2001). Calibration, monitoring and validation of AMSU-B. *Advances in Space Research*, 28(1):117–126.

Atkinson, N. C. (2002). Performance of AMSU-B Flight Model 3 (FM3) during NOAA-M Post Launch Orbital Verification Tests, AMB113. Technical report, Remote Sensing Instrumentation, Met Office.

Atkinson, N. C. (2015). NEDT specification and monitoring for microwave sounders. Technical report, MetOffice.

Atkinson, N. C. (2016). AAPP User Guide, ATOVS and AVHRR Pre-processing Package (AAPP), NWPSAF-MO-UD-036. `https://www.nwpsaf.eu/site/software/aapp/documentation/userguide/#Antenna_pattern_correction_for_microwave_sounders`. Version 1.3, last access: 13.05.2018.

Atkinson, N. C. (2018). (personal communication).

Bibliography

Atkinson, N. C. and McLellan, S. (1998). Initial evaluation of AMSU-B in-orbit data. *Proceedings of SPIE - The International Society for Optical Engineering*, 3503:276–287.

Barnes, J. A. and Allan, D. W. (1990). Variances based on data with dead time between the measurements (Technical Note 1318). Technical report, National Institute of Standards and Technology (NIST).

Block, T. (2015). FIDUCEO Multi-sensor Match up System Implementation Plan, V1.0. Technical report, Brockmann Consult GmbH.

Bonsignori, R. (2007). The Microwave Humidity Sounder (MHS) In-Orbit Performance Assessment. *Proc. of SPIE*, 6744(67440A):1–12.

Buehler, S. A., Kuvatov, M., and John, V. O. (2005). Scan asymmetries in AMSU-B data. *Geophysical Research Letters*, 32.

Buehler, S. A., Kuvatov, M., Sreerekha, T. R., John, V. O., Rydberg, B., Eriksson, P., and Notholt, J. (2007). A cloud filtering method for microwave upper tropospheric humidity measurements. *Atmospheric Chemistry and Physics*, 7(21):5531–5542.

Burgdorf, M., Buehler, S. A., Lang, T., Michel, S., and Hans, I. (2016). The moon as a photometric calibration standard for microwave sensors. *Atmospheric Measurement Techniques*, 9:3467–3475.

Burgdorf, M. and Hans, I. (2016). Evaluation Report on Microwave Humidity Sounder Version One FCDR, V1.0. Technical report, Universität Hamburg, FIDUCEO.

Burgdorf, M., Hans, I., Prange, M., Lang, T., and Buehler, S. A. (2018). Inter-channel uniformity of a microwave sounder in space. *Atmospheric Measurement Techniques*, 11(7).

Burgdorf, M., Hans, I., Prange, M., Mittaz, J., and Woolliams, E. (2017). D2.2 Microwave: Report on the MW FCDR uncertainty. Technical report, Universität Hamburg, National Physical Laboratory.

Cao, C., Weinreb, M., and Xu, H. (2004). Predicting simultaneous nadir overpasses among polar-orbiting meteorological satellites for the intersatellite calibration of radiometers. *Journal of Atmospheric and Oceanic Technology*, 21:537–542.

Chung, E. S., Sohn, B. J., Schmetz, J., and Koenig, M. (2007). Diurnal variation of upper tropospheric humidity and its relations to convective activities over tropical africa. *Atmospheric Chemistry and Physics*, 7(10):2489–2502.

EUMETSAT (2013). MHS Level 1 Product Generation Specification, EUM.EPS.SYS.SPE.990006. Technical report, EUMETSAT.

EUMETSAT-webpage (2018). Satellite orbits. `https://www.eumetsat.int/website/home/Satellites/LaunchesandOrbits/SatelliteOrbits/index.html`. last access: 28. April 2018.

Ferraro, R. (2016). AMSU-B/MHS Brightness Temperature - Climate Algorithm Theoretical Basis Document, NOAA Climate Data Record Program CDRP-ATBD-0801 Rev. 1. Technical report, NOAA STAR.

FIDUCEO (2018). Vocabulary - Uncertainty Quantified Fundamental Climate Data Record. `http://www.fiduceo.eu/content/uncertainty-quantified-fundamental-climate-data-record-fcdr`.

Fixsen, D. J. (2009). The temperature of the cosmic microwave background. *The Astrophysical Journal*, 707(2).

Giering, R., Quast, R., Mittaz, J., Hunt, S., Harris, P., Woolliams, E., Dilo, A., Cox, M., and Merchant, C. (2017). A novel framework to harmonise satellite data series for climate applications. poster, EUMETSAT Conference Rome.

Grody, N. C., Vinnikov, K. Y., Goldberg, M. D., T.Sullivan, J., and Tarpley, J. D. (2004). Calibration of multisatellite observations for climatic studies: Microwave Sounding Unit (MSU). *Journal of Geophysical Research: Atmospheres*, 109.

Hanlon, H. and Ingram, W. (2016). Algorithm Theoretical Basis Document, Fundamental Climate Data Record of Microwave Brightness Temperatures, CF-26/27/28, version 1.0. Technical Report SAF/CM/UKMO/ATBD/FCDR_MWAVE, Satellite Application Facility on Climate Monitoring (CM SAF).

Hans, I., Burgdorf, M., and Woolliams, E. (2018). Product user guide - Microwave FCDR release 0.2. techreport, Universität Hamburg and National Physical Laboratory.

Hewison, T. J. (1991). A Thermal Model of Black Body Targets. Met Office (RS) Working Paper No.29, Aug 1991, Available from Met O(RS) Y70 DERA, Farnborough, Hants, GU14 0LX, UK. 12 pp.

Hewison, T. J. and Saunders, R. (1996). Measurements of the AMSU-B Antenna Pattern. *IEEE Transactions on Geoscience and Remote Sensing*, 34(2).

Iacovazzi Jr., R. A. and Cao, C. (2008). Reducing Uncertainties of SNO-Estimated Intersatellite AMSU-A Brightness Temperature Biases for Surface-Sensitive Channels. *Journal of Atmospheric and Oceanic Technology*, 25(6):1048–1054.

Ignatov, A., Laszlo, I., Harrod, E. D., Kidwell, K. B., and Goodrum, G. P. (2004). Equator crossing times for NOAA, ERS and EOS sun-synchronous satellites. *International Journal of Remote Sensing*, 25(23):5255–5266.

John, V. O., Allan, R. P., Bell, B., Buehler, S. A., and Kottayil, A. (2013a). Assessment of inter-calibration methods for satellite microwave humidity sounders. *Journal of Geophysical Research*, 118:4906–4918.

John, V. O. and Chung, E. S. (2014). Creating a microwave based FCDR for tropospheric humidity: Initial assessment of SSM/T-2 radiances. *Global Space-based Inter-Calibration System (GSICS) Quarterly, Special Issue on Microwave*.

John, V. O., Holl, G., Atkinson, N., and Buehler, S. A. (2013b). Monitoring scan asymmetry of microwave humidity sounding channels using simultaneous all angle collocations (SAACs). *Journal of Geophysical Research*, 118:1536–1545.

Bibliography

John, V. O., Holl, G., Buehler, S. A., Candy, B., Saunders, R. W., and Parker, D. E. (2012). Understanding inter-satellite biases of microwave humidity sounders using global simultaneous nadir overpasses. *Journal of Geophysical Research*, 117(D2).

Khan, S. and Shaw, G. (1999). PFM Radiometric Calibration Test Report, MHS-TR-JA281-MMP. Technical report, MATRA MARCONI SPACE UK limited.

Klaes, D. and Ackermann, J. (2014). Techniques and processes for pre-launch characterisation of new instruments. Technical report, EUMETSAT.

Kobayashi, S., Poli, P., and John, V. O. (2017). Characterisation of Special Sensor Microwave Water Vapor Profiler (SSM/T-2) radiances using radiative transfer simulations from global atmospheric reanalyses. *Advances in Space Research*, 59(4):917–935.

Kramer, H. J. (2018). NOAA POES Series - 5th Generation Satellites. `https://directory.eoportal.org/web/eoportal/satellite-missions/n/noaa-poes-series-5th-generation`. last access: 21.05.2018.

Labrot, T., Lavanant, L., Whyte, K., Atkinson, N., and Brunel, P. (2011). AAPP Documentation Scientific Description, version 7.0, Document NWPSAF-MF-UD-001. Technical report, Numerical Weather Prediction Satellite Application Facility (NWP SAF).

Luo, Z. J., Jeyaratnam, J., Shah, N., Rossow, W. B., and Ferraro, R. (2017). Inter-calibration of UTH-related radiances from SSM/T-2 and AMSU-B. GSICS, MW sub-group meeting, presentation slides, `http://gsics.atmos.umd.edu/bin/view/Development/20170111`, last access: 19.07.2018.

Malkin, Z. M. (2011). Study of astronomical and geodetic series using the Allan variance. *Kinematics and Physics of Celestial Bodies*, 27:42–49.

Merchant, C. J., Woolliams, E., and Mittaz, J. (2017). Uncertainty and Error Correlation Quantification for FIDUCEO easyFCDR Products: Mathematical Recipes, v0.9. Technical report, University of Reading, National Physical Laboratory.

Mittaz, J. P. D. (2016). Instrument Noise characterization and the Allan/M-sample variance. `http://www.fiduceo.eu/sites/default/files/publications/%noise_and_allan_variance_report.pdf`. last access: 14. Dec 2017.

Mittaz, J. P. D., Wolliams, E., and Merchant, C. J. (2018). Applying Principles of Metrology to Historical Earth Observations from Satellites. *Metrologia*. in preparation.

NOAA (2017). NOAA's Geostationary and Polar-Orbiting Weather Satellites. `http://noaasis.noaa.gov/NOAASIS/ml/genlsatl.html`. last access: 28. April 2018.

NOAA (2018). CLASS - Comprehensive Large Array-data Stewardship System. `https://www.class.ncdc.noaa.gov/saa/products/welcome`. data downloaded in 2015-2017.

NOAA-CLASS (2016). DMSP SSM webpage. `https://www.class.ngdc.noaa.gov/data_available/dmsp/index.htm`. last access: 14. Dec 2017.

NOAA-OSPO (2015). POES Operational Status webpage. `http://www.ospo.noaa.gov/Operations/POES/status.html`. last access: 14. Dec 2017.

NOAA-STAR (2015). ICVS webpage. `https://www.star.nesdis.noaa.gov/icvs/index.php`. last access: 14. Dec 2017.

Quast, R. (2018). Radiance Harmonisation for Stability. presentation slides. FIDUCEO User Workshop, Lisbon.

Robel, J. et al. (2009). NOAA KLM user's guide with NOAA-N, -N' supplement. Technical report, National Oceanic and Atmospheric Administration, National Environmental Satellite, Data, and Information Service, National Climatic Data Center, Remote Sensing and Applications Division.

Saunders, R. W., Hewison, T. J., Stringer, S. J., and Atkinson, N. C. (1995). The radiometric characterization of AMSU-B. *IEEE Transactions on Microwave Theory and Techniques*, 43(4):760–771.

Shi, L. and Bates, J. J. (2011). Three decades of intersatellite-calibrated High-Resolution Infrared Radiation Sounder upper tropospheric water vapor. *Journal of Geophysical Research*, 116.

Shi, L., Bates, J. J., and Cao, C. (2008). Scene Radiance–Dependent Intersatellite Biases of HIRS Longwave Channels. *Journal of Atmospheric and Oceanic Technology*, 25(12):2219–2229.

Tian, M., Zou, X., and Weng, F. (2015). Use of Allan Deviation for Characterizing Satellite Microwave Sounder Noise Equivalent Differential Temperature (NEDT). *IEEE Geoscience and Remote Sensing Letters*, 12(12):2477–2480.

Ulaby, F. T. and Long, D. G. (2014). *Microwave Radar and Radiometric Remote Sensing.* University of Michigan Press.

WMO-OSCAR (2016). WMO OSCAR Observing Systems Capability Analysis and Review Tool webpage. `http://www.wmo-sat.info/oscar/satellites/view/307`. last access: 14. Dec 2017.

Woolliams, E., Mittaz, J., Merchant, C. J., and Harris, P. (2017). D2.2a: Principles behind the FCDR effects tables. Technical report, National Physical Laboratory and University of Reading.

Zou, C.-Z., Goldberg, M. D., Cheng, Z., Grody, N. C., Sullivan, J. T., and Tarpley, D. (2006). Recalibration of microwave sounding unit for climate studies using simultaneous nadir overpasses. *Journal of Geophysical Research*, 111.

Zou, C.-Z. and Wang, W. (2011). Intersatellite calibration of AMSU-A observations for weather and climate applications. *Journal of Geophysical Research*, 116.

List of Figures

List of Tables

List of Publications and Project Reports

Hans, I., Burgdorf, M., John, V. O., Mittaz, J., and Buehler, S. A.: Noise performance of microwave humidity sounders over their lifetime, *Atmos. Meas. Tech.*, 10, 4927-4945, `https://doi.org/10.5194/amt-10-4927-2017`, 2017

Hans, I., Burgdorf, M., Woolliams, E.: Product User Guide - Microwave FCDR Release 0.2, Universität Hamburg, National Physical Laboratory, 2018

Burgdorf, M., Hans, I., Prange, M., Mittaz, J., Woolliams, E.: D2_2 Microwave: Report on the MW FCDR uncertainty, Universität Hamburg, National Physical Laboratory, 2017

Burgdorf, M., Buehler, S. A., Lang, T., Michel, S., and Hans, I.: The Moon as a photometric calibration standard for microwave sensors, *Atmos. Meas. Tech.*, 9, 3467-3475, `https://doi.org/10.5194/amt-9-3467-2016`, 2016

Burgdorf, M., Hans, I., Prange, M., Lang, T., and Buehler, S. A.: Inter-channel uniformity of a microwave sounder in space, *Atmos. Meas. Tech.*, 11, 4005-4014, `https://doi.org/10.5194/amt-11-4005-2018`, 2018

Acknowledgements

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Dissertation mit dem Titel "Towards a new fundamental climate data record for microwave humidity sounders based on metrological best practice" selbstständig verfasst und keine anderen als die angegebenen Hilfsmittel – insbesondere keine im Quellenverzeichnis nicht benannten Internet-Quellen – benutzt habe. Alle Stellen, die wörtlich oder sinngemäß aus Veröffentlichungen entnommen wurden, sind als solche kenntlich gemacht. Ich versichere weiterhin, dass ich die Dissertation oder Teile davon vorher weder im In- noch im Ausland in einem anderen Prüfungsverfahren eingereicht habe und die eingereichte schriftliche Fassung der auf dem elektronischen Speichermedium entspricht.

Hamburg, den 31.07.2018   Imke Hans