PhD Thesis

Remote sensing of ice clouds: synergistic measurements and radiative transfer simulations

Gerrit Holl

15 November 2013
To my beloved Catherine
Abstract

This thesis primarily considers the spaceborne remote sensing of ice clouds and frozen precipitation. Ice clouds are important for hydrology and for the Earth’s radiation budget, but many properties are difficult to measure, in particular using spaceborne instruments. A better quantification of ice clouds is needed to improve global climate models. This thesis presents steps toward such an improvement.

The first part of the thesis introduces topics related to the research presented in the second part, but presents no new scientific results. It gives a brief introduction to the history of atmospheric remote sensing and describes how the different parts of the electromagnetic spectrum can be used actively or passively. Then, it describes why ice clouds are important and what microphysical, optical, and macrophysical properties are used to describe atmospheric ice. Next, it briefly introduces the relevant topics in atmospheric radiative transfer. The first part concludes with a description of various approaches to retrievals, with a particular focus on those applied in this thesis.

The second part of the thesis describes new results. The bulk of the new results is described in five peer-reviewed publications, that are appended verbatim.

A major part of the work builds on the development of a toolkit to easily find co-incident measurements, or collocations, between any pair of satellite sensors. Four appended articles rely on this toolkit.

The first appended article uses the toolkit to obtain collocations between passive microwave and infrared on operational meteorological satellites with the Cloud Profiling Radar on CloudSat. It presents three examples. Firstly, from the collocated dataset and a dataset of synthetic profiles, the article compares the statistical relations between an official CloudSat Ice Water Path (IWP) product and microwave radiances. Secondly, it shows a point-by-point comparison between the same CloudSat IWP product, and an IWP product based on passive microwave. A more sophisticated set of systematic comparisons, including more satellites and sensors, is presented in a dedicated paper. Finally, the first paper provides a first preview of how the collocations can be used to train a new IWP retrieval from passive operational measurements. This too is the topic of a dedicated paper, where solar, terrestrial infrared, and microwave radiances are combined to obtain an improved IWP product from passive operational sensors, by training with an active combined radar-lidar product from CloudSat-CALIPSO.

The second appended article also relies on the collocations toolkit. Here, collocations between different copies of identical or very similar microwave sounders are used to assess how the inter-satellite bias depends on radiance and latitude.

The remaining two studies described in the thesis do not use existing measurements, but are based on radiative transfer modelling. One attached paper verifies that optimised frequency grids obtained in clear-sky simulations for terrestrial infrared instrument studies, can be applied directly for cloudy simulations. This result is relevant for future studies. Finally, the thesis includes a short study with retrieval simulations for a new sub-millimetre instrument concept.
Preface

This thesis concludes four years of research performed at Luleå University of Technology, Kiruna Campus. It consists of two parts:

- Part one is an introduction to the topic. Most of the content here does not describe new scientific results, but rather an overview of the knowledge required to understand the articles.
- Part two consists of new research, including a chapter describing a new instrument study and verbatim copies of peer reviewed publications, four published and one submitted.

This text can be considered a continuation of the so-called licentiate thesis (Holl 2011). In the Swedish academic system, the licentiate thesis is a thesis very similar in structure to the PhD thesis, that PhD candidates are encouraged to write approximately halfway through their PhD studies. A lot of the material published here was already published in the licentiate thesis. In case of questions, you can contact me at gerrit.holl@gmail.com.
Acknowledgements

Virtually no work is ever done alone. Human work is cooperative and the work resulting in this thesis is no exception. Without help and encouragement of dozens of people throughout my life, this thesis would not exist.

I would like to thank my primary supervisor, Stefan Buehler. From the very start when I joined the group part-time as a master student, your supervision has been dedicated and highly valuable. I appreciate this very much. Thanks!

Thanks to my secondary supervisors, Patrick Eriksson and Jana Mendrok. A big thanks to all who have contributed to the articles that I wrote and that resulted in this thesis: Stefan Buehler, Bengt Rydberg, Carlos Jiménez, Jana Mendrok, Ajil Kottayil, and Salomon Eliasson. Also thanks to Salomon Eliasson and Viju John, for initiating research projects to which I could provide a major contribution, thereby providing more material for my thesis. I am grateful for all the valuable input, in weekly group meetings or in smaller discussions, from all people who have worked in the research group during my time here. I have already mentioned Stefan Buehler, Jana Mendrok, Salomon Eliasson, and Ajil Kottayil. In addition, thanks to Oscar Isoz, Erik Johansson, Daniel Kreyling, Thomas Kuhn, Richard Larsson, Isaac Moradi, Mathias Milz, Madhuri Nukala, Anura Wickramanayake, and and Simon Östman.

I would also like to express my thanks to the other colleagues and former colleagues at the Division of Space Technology: Victoria Barabash, Johnny Ejemalm, Anita Enmark, Lars Jakobsson, Stefan Rova, Rikard Slapak, Anette Snällfot-Brändström, and Maria Winnebäck.

In the past years, I have had the opportunity to participate in various courses. I am grateful to the Swedish National Graduate School in Space Technology, in particular to director Marta-Lena Antti and the rest of the board members, for organising a good share of those courses. Additionally, I would like to thank Utrecht University for the summer school in Physics of the Climate System in 2009, the European Space Agency for the summer school in Earth Observation and Modeling in 2010, and the Université Joseph Fourier with all co-organisers and co-sponsors for my chance to participate in the European Research Course on Atmospheres in 2011. I would like to thank Luleå University of Technology for sponsoring my participation in a small, unconventional, but very useful course called Utveckling av Grupp och Ledare, where I learned surprisingly much about my own role in a group.

My thanks to the community for building and maintaining the Atmospheric Radiative Transfer Simulator along with the tools Atmlab and PyARTS. Additionally, a number of open source software packages not developed at our group have proven invaluable in my work. For the scientific data analysis I have used Python, numpy, scipy, matplotlib, and countless other, smaller packages that lie at the basis of the opensource software world that make using computers so much more pleasant and that have enhanced my productivity.
considerably. To write articles, to make presentations, and to design posters, \LaTeX{} and associated packages have been of great help. My gratitude to all people who have donated their time for developing open source software, for writing documentation, and for helping out on forums and mailing-lists.

I would like to thank lecturers and administration for my Bachelor in Applied Physics at the University of Twente, and for my Master in Space Science and Technology at the Julius-Maximilianus University of Würzburg and at Luleå University of Technology.

Finally, I wish to express thanks to a number of people who have been valuable for me on a personal level. Dank aan mijn ouders, Han en Tinelot, en aan mijn zussen, Elchien en Duveken.

De laatste jaren van mijn basisschooltijd heb ik doorgebracht op het Pædologisch Instituut in Duivendrecht. De hulp die ik hier heb gekregen van Peggy, Paul, Paula, Joke, Katrien, Lex, en anderen is niet te overschatten. Ik weet niet of ik zonder de zeer doelgerichte hulp in onder andere het aanleren van sociale vaardigheden de middelbare school had “overleefd”. Ik weet wel dat ik in mijn hele leven enorm veel aan de lessen heb gehad en zal hebben. Dank ook aan mijn vrienden uit die tijd.


Through easy and harder times, I have experienced valuable help from countless of people with whom I connected primarily or exclusively through the internet. I would especially like to thank Rik van Riel, who gave me wise advice in a period that I particularly needed it.

Shortly before I accepted Stefans offer to become his PhD student, I met you, Catherine. I do not possess the words to describe how fortunate I have been, how fortunate that you have come into my life. Therefore, I will just finish by saying a big thank you for everything we have shared in the past, do share in the present, and will share in the future.
Appended Papers


Related papers

The following papers are related to atmospheric remote sensing, but are not part of the PhD thesis.


Contents

Abstract v
Preface vii
Acknowledgements ix
Appended papers xi
Related papers xiii
Contents xv
List of figures xix
List of tables xix

1 Background 1

1 Introduction 3
  1.1 Why this thesis? .............................................. 3
  1.2 Aims ............................................................ 4
  1.3 Methods ....................................................... 4
    1.3.1 Collocations .............................................. 4
    1.3.2 Radiative transfer ....................................... 4
  1.4 Prior work ................................................... 5
  1.5 Presented research ......................................... 6
  1.6 This thesis .................................................. 6

2 Remote sensing 7
  2.1 Electromagnetic radiation .................................. 7
  2.2 Sensor considerations ...................................... 10
    2.2.1 Imaging .................................................. 12
    2.2.2 Sounding ............................................... 12
    2.2.3 Viewing geometry ...................................... 13
2.2.4 Orbits .......................................................... 13
2.3 Specific observation technologies ............................. 14
  2.3.1 Solar radiation .............................................. 14
  2.3.2 Terrestrial radiation ....................................... 15
    2.3.2.1 Infrared radiation ..................................... 15
    2.3.2.2 Passive microwave radiation ....................... 16
  2.3.3 Active ..................................................... 19
2.4 Collocations .................................................. 20
  2.4.1 Criteria .................................................... 21
  2.4.2 Method .................................................... 21
  2.4.3 Collocated datasets ...................................... 21

3 Ice clouds .......................................................... 23
  3.1 Radiative effects .............................................. 23
  3.2 Microphysical properties .................................... 24
  3.3 Optical properties .......................................... 26
  3.4 Macrophysical properties ................................... 26

4 Radiative transfer .................................................. 29
  4.1 Theory ......................................................... 29
    4.1.1 Clear-sky radiative transfer ............................ 30
    4.1.2 Scattering ................................................. 32
  4.2 Sensors ....................................................... 35
  4.3 Software ...................................................... 36
    4.3.1 Scattering ............................................... 36
    4.3.2 Optimisations ............................................. 36

5 Retrieval development ........................................... 39
  5.1 The inverse problem ......................................... 39
  5.2 Optimal estimation .......................................... 39
  5.3 Lookup tables / retrieval databases ....................... 40
    5.3.1 Sources .................................................... 40
    5.3.2 Applications — from simple lookups to multi layer perceptrons 40

II New research ..................................................... 43

6 Studying a new instrument concept ............................. 45
  6.1 Sub-millimetre radiometry .................................. 45
  6.2 Instrument study ............................................. 47
    6.2.1 Channel selection ....................................... 47
    6.2.2 Performance study ...................................... 50

7 Summary of appended articles .................................. 53
  7.1 Paper I ....................................................... 53
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2</td>
<td>Paper II</td>
<td>54</td>
</tr>
<tr>
<td>7.3</td>
<td>Paper III</td>
<td>55</td>
</tr>
<tr>
<td>7.4</td>
<td>Paper IV</td>
<td>56</td>
</tr>
<tr>
<td>7.5</td>
<td>Paper V</td>
<td>56</td>
</tr>
<tr>
<td>8</td>
<td>Conclusion</td>
<td>59</td>
</tr>
<tr>
<td></td>
<td>Bibliography</td>
<td>61</td>
</tr>
<tr>
<td></td>
<td>Acronyms</td>
<td>71</td>
</tr>
<tr>
<td>Paper I</td>
<td>Collocating satellite-based radar and radiometer measurements – methodology and usage examples</td>
<td>75</td>
</tr>
<tr>
<td>Paper II</td>
<td>Understanding inter-satellite biases of microwave humidity sounders using global simultaneous nadir overpasses</td>
<td>93</td>
</tr>
<tr>
<td>Paper III</td>
<td>Optimised frequency grids for infrared radiative transfer simulations in cloudy conditions</td>
<td>109</td>
</tr>
<tr>
<td>Paper IV</td>
<td>Systematic and random errors between collocated satellite ice water path observations</td>
<td>123</td>
</tr>
<tr>
<td>Paper V</td>
<td>SPARE-ICE: synergistic IWP from passive operational sensors</td>
<td>139</td>
</tr>
</tbody>
</table>
List of figures

2.1 First cyclone photograph ............................................. 7
2.2 Shortwave and longwave radiation ..................................... 8
2.3 AVHRR channels ........................................................... 11
2.4 AMSU-B Jacobians ......................................................... 12
2.5 AMSU-B channels ........................................................... 18
2.6 Footprints ................................................................. 20

3.1 Ice shape examples ......................................................... 25

4.1 Summary of radiative transfer properties ............................. 31
4.2 Phase functions ........................................................... 35

6.1 Sub-mm ice cloud measurement principle ............................. 46
6.2 Sub-mm particle-dependent sensitivity ................................. 46
6.3 IceMusic spectrum and channel positions ............................ 48
6.4 IceMusic Jacobians ......................................................... 49
6.5 IceMusic retrieval simulation performance .......................... 51

List of tables

4.1 Scattering regimes ......................................................... 33
4.2 Methods for scattering properties calculations ..................... 34

6.1 Proposed channels for IceMusic, centre frequencies and channel widths. 50
Part I.

Background
Chapter 1.
Introduction

1.1. Why this thesis?

Planet Earth is the home of all humanity, a population of more than seven billion people. We all live at the boundary between the solid Earth and the atmosphere that surrounds it. This atmosphere is a complex, highly dynamic system. Pressure, humidity, precipitation, and other variables, vary continuously in space and in time, in a phenomenon we experience as the weather. The statistics of the weather in a specific region over a long period of time is known as the climate.

The climate on planet Earth is itself dynamic, as boundary conditions such as orbital inclination and atmospheric composition change. During the last century, measurements and models show that temperatures are increasing and precipitation patterns are changing (Trenberth et al. 2007). It is very likely that the bulk of those changes are due to anthropogenic influences of atmospheric composition and surface vegetation (Hegerl et al. 2007). This climate change has the potential to significantly affect our natural environment (Rosenzweig et al. 2007). Therefore, a good understanding of the climate system is important to mankind.

Clouds have a strong effect on the Earth’s radiation budget (Rogers and Yau 1976). In the 4th Assessment Report (AR4), the Intergovernmental Panel on Climate Change (IPCC) wrote that ‘cloud feedbacks remain the largest source of uncertainty in climate sensitivity estimates’ (Randall et al. 2007, Section 8.6.3.2). In climate models, the magnitude and the sign of the feedback is almost entirely dictated by model assumptions (Stephens 2005). Therefore, there is a strong need for a better understanding of clouds.

Both modelling and remote sensing of clouds and precipitation are difficult, in particular if ice is present. The interaction of clouds with electromagnetic radiations depends on many factors, such as particle size, number density, mass density, phase, and shape. Spatial scales range from 1 μm to $10^3$ km, a range spanning 12 orders of magnitude. Remote measurements are hard to validate, because remote measurements are often column-integrated quantities, and column-integrated quantities are hard to measure in-situ. The lack of validated measurements means the models are poorly constrained and therefore hard to improve.

A good understanding of the climate system requires good observations. The World Meteorological Organisation (WMO) Global Climate Observing System (GCOS) has recommended a set of 50 Essential Climate Variables (ECVs) that should be systematically observed (WMO 2010). One can observe the atmosphere either through remote sensing or in-situ. Atmospheric remote sensing can be from the ground looking up, from the air looking either down or up, or from space looking down. From a climate monitoring
point of view, observations should ideally a) span a long time range, b) have a good time resolution (high sampling frequency), c) have global coverage, d) have a good horizontal spatial resolution, and e) have a small error.

Ground-based measurements have been around longest and therefore span the longest time range. However, instruments are typically stationary, and therefore have poor spatial coverage. The highest time resolution is obtained with either ground-based measurements or with spaceborne measurements from geostationary satellites. Global coverage can only be obtained from polar orbiting satellites. The horizontal spatial resolution for downlooking spaceborne sensors is exceeded only by downlooking airborne ones. The error on remote sensing measurements is usually higher than for in-situ observations. However, considering that global coverage can only be met with polar orbiting satellite observations, only with satellite observations we can ever hope to meet all 5 criteria.

1.2. Aims

This thesis shall be seen as a step toward an improvement of the space-based remote sensing of frozen clouds and precipitation. The focus is on the column density of ice, or Ice Water Path (IWP). IWP is a fundamental property for the characterisation of ice clouds, because it relates directly to mass continuity. In estimates of the atmospheric column density of ice, IWP, models and measurements vary by up to an order of magnitude (Waliser et al. 2009) and show different spatial distributions (Eliasson et al. 2011). Therefore, there is a need for improved IWP retrievals. Ultimately, the aim of an improved characterisation of atmospheric ice is to help constrain (and, eventually, improve) global climate models.

1.3. Methods

To study existing retrievals and work towards improved retrievals, I primarily use two methods. I briefly introduce them below, but I will describe them in more detail in section 2.4 and in chapter 4.

1.3.1. Collocations

A collocation occurs when different instruments remotely sense an overlapping area (or volume) at instants close in time. Collocations have advantages over using either instrument alone. For example, collocations between the CloudSat Cloud Profiling Radar (CPR) and National Oceanic and Atmospheric Administration (NOAA)-18 Microwave Humidity Sounder (MHS) can teach us how IWP affects radiances measured with MHS (as shown in Paper I), what the shared sensitivity range between different IWP products is (as studied in Paper IV), or can be used to develop an improved IWP product (such as in Paper V). Collocations are described in more detail in section 2.4 and in Paper I.

1.3.2. Radiative transfer

Radiative transfer describes the interaction of electromagnetic radiation with a medium it encounters. One example is the simulation of radiation emitted by the Earth’s surface and
1.4. Prior work

Spaceborne remote sensing of clouds dates back to the earliest meteorological satellites, where optical photography was used to identify spatial structures (see chapter 2). Quantitative measurements go back to around 1980, with the focus commonly on the interaction between clouds and radiation (Rossow and Schiffer 1991). Traditionally, downlooking remote sensing of clouds has been performed using solar or terrestrial infrared measurements (e.g. Rossow and Schiffer 1991; King et al. 1997). More recently, products have been generated from passive microwave (e.g. Ferraro et al. 2005; Boukabara et al. 2011) and from active sensors such as radar and lidar (e.g. Stephens et al. 2002, 2008; Delanoë and Hogan 2010; Deng et al. 2010). Cloud properties have also been retrieved from instruments in a limb geometry (e.g. Eriksson et al. 2007).

The A-Train (see section 2.3.3) is a constellation of satellites with the explicit aim to exploit collocations between different instruments (L’Ecuyer and Jiang 2010). The use of collocations, however, predates the launch of the A-Train satellites. The first publication to explicitly address the combination of measurements from different sensors appears to be a technical note written in Japanese (Aoki 1980). Collocations are used in many ways. For example, a combination of passive microwave and infrared has been used for rainfall retrievals (e.g. Kidd et al. 2003; Marzano et al. 2004; Rapp, Elsaesser and Kummerow 2009; Di Paola et al. 2012), comparable to how Paper V combines passive techniques to retrieve IWP. Another popular application is to collocate instruments that should in theory be identical, and use this to characterise biases and/or intercalibrate sensors. An example is Cao, Weinreb and Xu (2004). Paper II and John et al. (2013) apply collocations similarly. However, intercomparisons between sensors do not need to use collocations. Another way is to collect data over some region and time period, and then compare the statistics of the collected data (e.g. Wu et al. 2009; Eliasson et al. 2011). Paper I contains more references to articles focussed on the technical aspects of obtaining collocations.

A particular application in this thesis is to collocate operational and scientific sensors, in order to better understand or improve operational retrievals. Relatively little work has been performed in this direction. A recent study (under public review at the time of writing) is by Gong and Wu (2013), who use a statistical approach using the spaceborne cloud radar CloudSat to train a retrieval algorithm for ice cloud properties from passive microwave, similar to the method introduced in Paper I and further explored in Paper V. A more elaborate review of prior work can be found in the introductions to the five appended papers.

\footnote{Gong and Wu (2013) was published for open review two weeks after the submission of Paper V, and is therefore not mentioned in the latter manuscript.}
Chapter 1. Introduction

1.5. Presented research

In Paper I, I present three possible applications for the use of collocations. Two of those have resulted in dedicated publications: a thorough inter-comparison of IWP products in Paper IV, and the development of a new, synergistic IWP product in Paper V. In the years since Paper I, the collocation toolkit has vastly improved in flexibility, which has led to a number of spin-off projects.

In John et al. (2011) (one of the “related papers” to this thesis), we use the toolkit to collocate cloud-cleared data from the High resolution Infrared Radiation Sounder (HIRS) instrument with Advanced Microwave Sounding Unit (AMSU) measurements, to investigate biases introduced by cloud-clearing in the climatology of Upper Tropospheric Humidity (UTH). In Paper II, nadir collocations (there referred to as Simultaneous Nadir Overpasses (SNOs)) between different copies of AMSU-B and MHS are used to find the latitudinal dependence of inter-satellite biases. In John et al. (2013), another related paper, we look at Simultaneous All Angle Collocations (SAACs) to investigate scan asymmetries. Other projects using either the collocation toolkit or one of the resulting datasets are in various stages of development.

For radiative transfer simulations, this thesis uses primarily the Atmospheric Radiative Transfer Simulator (ARTS), a code that was available prior to the project or developed in parallel by others. However, small improvements to ARTS were made as part of the thesis project. Paper III describes a small but essential step for calculation cloudy radiative transfer simulations for infrared radiometers. These results are applicable in a variety of different contexts.

1.6. This thesis

Part I introduces theory and background information relevant to the research presented in Part II. You are currently reading chapter 1, the introduction. Chapter 2 provides the history of spaceborne remote sensing and introduces the basic physics that most remote sensing is based upon. The object of the remote sensing in this thesis, ice clouds, is introduced in chapter 3. Then, chapter 4 introduces radiative transfer and radiative transfer modelling, a tool of paramount importance for retrieval development. The latter is described in chapter 5. Part II describes new research results. A retrieval study for a new instrument concept is described in chapter 6. Then, in chapter 7, the articles appended at the end of the thesis are summarised, and their inclusion is justified by a brief description of my contribution, in particular where I am not the first author. Finally, before the backmatter and the included articles, chapter 8 concludes the thesis.
Chapter 2.

Remote sensing

As outlined in the introduction, only remote sensing from polar orbiting satellites can provide a global observation of Earth. This chapter briefly introduces aspects of atmospheric remote sensing relevant for this thesis work. A more complete introduction to (atmospheric) remote sensing can be found in Rees (2001) or other relevant textbooks.

Earth observation from space has been a major motivation for the launch of satellites since the dawn of the space age. The first successful weather satellite was TIROS-1, launched 1 April 1960, less than three years after the launch of Sputnik marked the beginning of the space age in October 1957. The ability to directly image the Earth from space (Figure 2.1) marked a revolution in meteorology and climatology. Now, synoptic-scale weather systems could be directly observed from space. Furthermore, measurements became available on a global scale, providing information in areas far away from any weather stations. It is therefore no surprise that after the launch of TIROS-1, hundreds of Earth observation satellites have been launched and are operated by organisations around the world, with always new satellites and sensors in planning. For a history of past Earth observation sensors, the review papers of Smith et al. (1986) and Kidd, Levizzani and Bauer (2009) are valuable.

2.1. Electromagnetic radiation

Almost all Earth observation satellites, and all satellites observing the lower atmosphere, measure electromagnetic radiation. They do so at various frequencies throughout the spectrum. The graphic in the margin on the next page shows frequencies and wavelengths at which atmospheric remote sensing operates (the various acronyms are expanded in a list in the back matter of the thesis). Electromagnetic radiation is emitted by any object with a non-zero absolute temperature, including the Sun, the Earth’s surface, and the Earth’s
atmosphere. For a blackbody, the energy emitted as a function of frequency is described by Planck’s law.

\[ L_{\nu, p} = \frac{2h\nu^3}{c^2(e^{h\nu/kT} - 1)}, \]  

(2.1)

or, with wavelength rather than frequency,

\[ L_{\lambda, p} = \frac{2hc^2}{\lambda^5(e^{hc/\lambda kT} - 1)}. \]  

(2.2)

Here, \( L_{\nu, p} \) is the spectral radiance in \( \text{Wsr}^{-1}\text{m}^{-2}\text{Hz}^{-1} \) or equivalent (where Hz may be replaced by another spectral unit, for example by m or \( \mu \)m for \( L_{\lambda, p} \)), \( h \) is the Planck constant with a value of \( 6.6261 \times 10^{-34} \) J s, \( c = 2.9979 \times 10^8 \) m/s is the propagation speed of electromagnetic radiation in vacuum (the speed of light), \( k = 1.3807 \times 10^{-23} \) J/K is the Boltzmann constant, \( T \) is the absolute temperature in K, and \( \lambda \) is the wavelength in m.

Figure 2.2 shows the blackbody curves for the Sun with an effective surface temperature of 5778 K and for the Earth at a surface temperature of 287 K, as well as Top Of Atmosphere (TOA) terrestrial radiation.

---

**Figure 2.2.** — Electromagnetic radiation as observed at the top of the Earth’s atmosphere: solar and terrestrial radiation. The lines show blackbody curves following Equation 2.2, whereas the filled area shows what part of terrestrial radiation is observed from a satellite sensor, as calculated with ARTS. The difference is absorbed by various gases. Infra-Red (IR) radiation is mostly absorbed by greenhouse gases such as water vapour and carbon dioxide. Note that the spectral irradiance curves are normalised to have the same peak (or, equivalently, area), so their magnitudes cannot be compared based on this figure.
2.1. Electromagnetic radiation

Real bodies are not blackbodies, but emit at a certain emissivity $\epsilon$,

\[ L_\nu = \epsilon(\nu)L_{\nu,p}, \]

or,

\[ L_\lambda = \epsilon(\lambda)L_{\lambda,p}, \]

where a blackbody, by definition, has $\epsilon = 1$. When $\epsilon$ is constant as a function of $\lambda$ or $\nu$, the source is, by definition, a grey body.

For the Earth, or indeed any object at a similar temperature, the spectral radiance in W sr$^{-1}$ m$^{-2}$ Hz$^{-1}$ emitted at the microwave wavelength of 20 GHz (1.54 cm) is approximately ten orders of magnitude less than the spectral radiance at the peak wavelength of 10.3 $\mu$m (Rees 2001, page 27). It is therefore not practical to express radiation in W sr$^{-1}$ m$^{-2}$ Hz$^{-1}$ when working with both infrared and microwave radiation. Terrestrial radiances are instead often expressed in brightness temperature, such as the Planck brightness temperature,

\[ T_{b,\text{Planck}} = \frac{hc}{\lambda k \ln \left(1 + \frac{1}{\epsilon} \left(\frac{h c}{\lambda k T} \right)^{-1}\right)}, \]

defined as the temperature that a blackbody emitting the observed radiation at a particular frequency would have. For downlooking sensors, the brightness temperature can be related to the temperature of the observed scene if the emissivity is known, although this becomes more complex if several scenes are in the field of view. The definition in Equation 2.4 is not unique. Other definitions of brightness temperature exist, such as the Rayleigh-Jeans brightness temperature, which instead uses the Rayleigh-Jeans approximation of the Planck function. Brightness temperature can also relate to the total received sensor power (in W) rather than the spectral radiance (in W sr$^{-1}$ m$^{-2}$ Hz$^{-1}$), and then there are even more variables in the precise definition. All of these are expressed in Kelvin, so care is needed when interpreting brightness temperatures.

Figure 2.2 also shows what part of terrestrial radiation is transmitted by the atmosphere and what part is not. Radiation emitted by the Sun and then reflected by Earth is referred to as solar or (less accurately) shortwave radiation, whereas radiation emitted by Earth is often referred to as terrestrial, longwave, or thermal radiation. Most Earth observation sensors measure at wavelengths where either solar or terrestrial radiation dominate, although some also measure where both have significant contributions. Terrestrial radiation at 8 to 13 $\mu$m is not significantly absorbed by greenhouse gases and reaches space if the sky is clear (without clouds or other particulate matter). This region of the electromagnetic spectrum is called the window region. Infrared radiation can be either terrestrial, solar, or a mixture of both. A major advantage of measuring terrestrial radiation is that it is available at all times, whereas a measurement of solar radiation requires sunlight and is only possible during the day (or perhaps in moonlight).

Although negligible for the Earth’s energy budget, microwave radiation as emitted from the Earth’s surface and atmosphere is very valuable for the purposes of remote sensing. Microwave radiation interacts with matter quite differently than visible or infrared radiation.

1Strictly speaking, radiation of either solar or terrestrial origin covers all wavelengths, and both are thermal. In this thesis, I characterise radiation based on its origin: solar or terrestrial.
does, and can therefore be used to obtain different information. For example, it is much less affected by clouds. More information on passive microwave sensors can be found in section 2.3.2.2 on page 16.

Note that, depending on context, scientific field, and region in the electromagnetic spectrum, different units are used to express the frequency or wavelength of electromagnetic radiation:

- **Wavelength.** Visible and infrared measurements are often referred to by wavelength in nanometres or micrometres, respectively. Wavelength in centimetres may be found in a radar context as well.

- **The wavenumber is the inverse of wavelength, and used for infrared radiation.** Wavenumbers are usually expressed in cm\(^{-1}\). 8 µm \(\triangleq\) 1250 cm\(^{-1}\) and 12 µm \(\triangleq\) 833 cm\(^{-1}\).

- **Frequency in GHz is used for microwave and sub-millimetre radiation.**

- **Others:** Energy in KeV is used for x-ray observations. Radar frequencies are commonly referred to by band letters (W-band, K-band, etc.) describing electromagnetic regions.

The graphic on page 8 illustrates the region of the electromagnetic radiation of interest for this thesis with frequencies in GHz and THz and wavelengths in µm and mm.

For a thorough introduction to the physics behind remote sensing, the reader is referred to Rees (2001) or one of the other textbooks on the subject.

### 2.2. Sensor considerations

This section discusses some sensor considerations relevant for this thesis. It is not meant to be complete; many aspects, such as polarisation or spectroscopy, are not considered at all, as they are not relevant for this work.

Fundamentally, a passive remote sensing instrument does not measure spectral radiance or radiant power directly, but some analogue or digital quantity such as noise power, antenna temperature, or electron count. The spectral response function (SRF) describes the sensor (channel) sensitivity to different frequencies. The total power received by a sensor,

\[
P = A\Delta\Omega \int_0^\infty f(\lambda)L(\lambda)\,d\lambda,
\]

relates to the brightness temperature (Rees 2001, section 6.3.4). In Equation 2.5, \(P\) is the power in W, \(A\) is area in m\(^2\), \(\Omega\) is the observed solid angle in sr, \(\lambda\) is wavelength in m, \(f(\lambda)\) is the sensor response function, and \(L(\lambda)\) is spectral radiance in W sr\(^{-1}\) m\(^{-2}\) Hz\(^{-1}\). Note that reality is often more complex than Equation 2.5 describes, as the equation assumes that the entire field of view contributes equally to the total sensor observed power. This is not true in general. For example, for microwave sensors, the antenna pattern is not at all constant, \(L\) is a function of angle, and therefore \(\Omega\) needs to be integrated over as well.
2.2. Sensor considerations

Figure 2.3. — The black lines show SRFs for AVHRR channels 4 and 5 on NOAA-19. Opacities (see Equation 4.4 on page 31) are simulated with ARTS (described in section 4.3). The vertical lines describe an optimised frequency grid as explained in Paper III. This is Figure 1 from Paper III.
As an example, Figure 2.3 shows the SRFs \( f(\lambda) \) for Advanced Very-High Resolution Radiometer (AVHRR) channels 4 and 5.

The details of sensor technology are beyond the scope of this text.

### 2.2.1. Imaging

As illustrated with the photograph in Figure 2.1 on page 7, the first remote sensing instruments in space were imagers. An image is a near-instantaneous two-dimensional view of a scene. Imagers are primarily used to determine spatial structures. Still, imagers such as AVHRR are also used to retrieve geophysical quantities. The pixels of a digital image are contiguous, whereas different soundings may be discontiguous. Imagers can have a very high spatial resolution, that may be in the order of metres rather than kilometres. Figure 2.6 on page 20 illustrates this. Examples of imagers on meteorological spacecraft are AVHRR, Visible Infrared Imager Radiometer Suite (VIIRS), and Spinning Enhanced Visible Infra-Red Imager (SEVIRI).

### 2.2.2. Sounding

This section briefly discusses some aspects of atmospheric sounding. For a thorough treatment, refer to Rodgers (2000), from where most of the material in this section is drawn.

Atmospheric sounding is the measurement of a vertical profile of some quantity in an atmosphere. Commonly measured quantities include temperature, humidity, or the concentration of trace gases. In remote sensing, sounding relies on radiation emitted by the atmosphere. For a down-looking sensor, height-dependency can be obtained by using multiple channels with frequencies at different distances from an absorption line. A sensor observing the limb can obtain a height resolution through scanning (see section 2.2.3).

When the atmosphere is opaque, any radiation emitted from the surface is absorbed by the lower atmosphere. Therefore, any radiation reaching a satellite sensor must be originating from higher layers in the atmosphere. Channels 18–20 of the AMSU-B instrument are placed at 1, 3, and 7 GHz from the 183 GHz absorption line (see also section 2.3.2.2, Figure 2.5, and section 4.1.1). The closer to the absorption line, the more absorption, the higher the altitude the radiation is coming from and the lower the brightness temperature (assuming absorption peaks in the troposphere, where temperature...
decreases with increasing altitude).

The jacobian is the derivative of the measurement vector with respect to the state,

$$\tilde{y} = I(\tilde{x}, \tilde{b}) \quad J = \frac{\partial I}{\partial \tilde{x}}. \quad (2.6)$$

a matrix that may contain many columns, because the measurement depends on many variables. In Equation 2.6, $\tilde{y}$ is the measurement vector (such as the three brightness temperatures for AMSU-B channels 18–20), $I$ relates the geophysical state to the measurement (such as described by the forward model, see chapter 4), $\tilde{x}$ is the state (describing geophysical parameters of interest), $\tilde{b}$ describes all other parameters influencing the measurement, and $J$ is the jacobian.

Figure 2.4 shows three rows of the jacobian for AMSU-B with respect to volume mixing ratio at different altitudes in the US Standard Atmosphere. Since water vapour concentrations vary in space and time, so does the jacobian, so this is just an example.

Sounding channels at frequencies of 183 GHz or higher may be affected by scattering from atmospheric ice. When retrieving profiles of temperature or humidity, this is undesirable. However, the signal from scattering can also be used to retrieve information related to atmospheric ice. This is the principle of cloud retrievals from microwave and sub-millimetre frequencies, which is further discussed in chapter 5 and section 6.1.

Unlike the observations by an imager, the observations by a sounder may not be contiguous. For example, the distance between soundings of the latest generation of HIRS instruments is much larger than the field of view of a single sounding. Therefore, in the case of sounders, it is not correct to speak of resolution or pixels, but rather of footprint size and distance between observations.

2.2.3. Viewing geometry

The viewing geometry of a sensor is an important design consideration.

Most satellite sensors used in Earth observation are down-looking. Regardless of whether the sensor can actually see all the way down or not, the sensor is oriented with the Earth’s surface within the field of view. This may be toward nadir (straight down) or at a particular off-nadir angle. Often, instruments scan across the track, thereby observing nadir and off-nadir, or conically, keeping the scan angle constant. A down-looking geometry allows for a good horizontal resolution, but the vertical resolution is poor.

A different geometry is to observe the limb. By looking at the horizon, a sensor on a satellite platform can observe the atmosphere against a background that is not the Earth: this may be cold space, the Sun, or a source of artificial radiation. Through scanning, the atmosphere is observed at different altitudes. This allows for a good vertical resolution and sensitivity to species with low concentrations, but since the line of sight through the atmosphere is long, the horizontal resolution is poor.

In this thesis, only down-looking sensors are used.

2.2.4. Orbits

Two particular orbit types are popular with Earth observation satellites. Many Earth observation satellites are either geostationary or sun-synchronous. The characteristics of
those special orbits are described below. Material for this section is drawn from Capderou (2005, sections 4.4 and 4.5).

For a geostationary orbit, a) the orbit angular velocity equals the Earth’s angular velocity, b) the orbit is circular, and c) the orbit lies in the equatorial plane. This results in a satellite that appears stationary for an observer on Earth. This orbit is popular because it allows a continuous observation of the same place. A disadvantage is that it implies a height of 35,788 km above the equator, so this orbit is not suitable if a high spatial resolution is needed, and observing high latitudes is difficult or impossible.

The plane of a sun-synchronous orbit is constant with respect to the Sun. Since the Earth is an oblate spheroid, an orbital plane is not constant with respect to the stars, but has a particular precession. If inclination and semi-major axis are chosen such that the precession is exactly one full circle per year, the orbital plane can be constant with respect to the Sun. Apart from technical advantages (such as the lack of a need to rotate solar panels), this offers advantages for Earth observation: the light conditions to observe a particular place are always the same. On the other hand, it also means that for long-term averages, observations occur only at one or two fixed local times. For atmospheric observations, this means that phenomena that occur primarily at other times might never be observed at all.

Sun-synchronous Earth observation satellites are usually in (near-)circular orbits.

A particular characteristic for a sun-synchronous orbit is the Local Time Ascending Node (LTAN). Since the orbit is constant with respect to the Sun, the mean local solar time at each ascending (northward) pass over the equator is constant too\(^2\). Due to perturbations from the gravitational influence of the Sun, the Moon, Jupiter, and irregularities in the geoid, a satellite does not naturally remain in an exact sun-synchronous orbit. For satellites without an actively controlled orbit, the LTAN drifts over time. Paper II exploits this drift for the intercalibration of passive microwave sensors.

All sensors used in this thesis are carried on satellites in polar-orbiting, sun-synchronous orbits.

\section*{2.3. Specific observation technologies}

Earth observation sensors operate throughout the electromagnetic spectrum, from 5.6 cm (5.3 GHz) for typical precipitation radars (Karlsson 1997) to 12 pm (300 eV) for ionospheric x-ray imagers, a range spanning nine orders of magnitude in wavelength. For spaceborne observations of weather and climate, wavelengths shorter than visible (390 to 700 nm) are rarely, if ever, used.

Technologies can be categorised in different ways. Here, I describe the technologies according to the radiation source: solar (section 2.3.1), terrestrial (section 2.3.2), or man-made (active, section 2.3.3) radiation.

\subsection*{2.3.1. Solar radiation}

Emission from the Sun can be approximated as coming from a blackbody at 5778 K, with intensity peaking at approximately 500 nm. As illustrated by Figure 2.2, the bulk of the

\(^2\text{The actual local solar time varies slightly due to the ellipticity of the Earth’s orbit around the Sun.}\)
2.3. Specific observation technologies

Radiation can be divided in Ultra-Violet (UV) (\(\lambda < 390\) nm), visible (390 nm < \(\lambda < 700\) nm), and near infrared (\(\lambda > 700\) nm). Most UV radiation is absorbed by the ozone layer, and near infrared radiation is partly absorbed by water vapour and other species. In the absence of clouds and aerosols, the bulk of visible radiation is transmitted through the atmosphere. Clouds and aerosols scatter and absorb solar radiation (although clouds do almost not absorb in the visible). In both cases, a significant part of visible radiation is still transmitted to the surface (or it would be pitch black on a cloudy day).

Optical imaging of Earth was the first application of space-based remote sensing, as illustrated by Figure 2.1. Even the crudest photo camera can be used for the identification of clouds. Still today, solar radiation is a major source of information on clouds — more on that will be discussed in chapter 5. Often, solar channels are on the same instrument as are infrared channels, such as with AVHRR, that is further discussed below.

2.3.2. Terrestrial radiation

Terrestrial radiation is radiation emitted by the Earth’s surface and atmosphere. As shown in Figure 2.2, this is significant at wavelengths longer than approximately 3 \(\mu\)m. At wavelengths shorter than 3 \(\mu\)m, reflected solar radiation dominates over emitted terrestrial radiation. At wavelengths longer than 3 \(\mu\)m, emission from the Earth’s surface and atmosphere is significant, and beyond roughly 5 \(\mu\)m, reflection can be neglected (Karlsson 1997). The bulk of the energy emitted by the Earth’s surface and atmosphere is emitted between 3 \(\mu\)m and 15 \(\mu\)m. For part of this region, radiation is absorbed by water vapour, carbon dioxide, or other gases (see Figure 2.2), which allows for atmospheric sounding. Between 15 \(\mu\)m and 150 \(\mu\)m (2 THz), carbon dioxide and water vapour in the Earth’s atmosphere absorb all radiation emitted from the surface and lower troposphere and nadir-looking observations have little to offer. Longer wavelengths are usually expressed by their frequency. The region between approximately 2 THz and 300 GHz (1 mm) is also known as sub-millimetre, and potentially useful for retrieval of cloud ice, even though here too atmospheric opacity is very high. See chapter 6 for an instrument study. Frequencies smaller than 300 GHz are known as microwave radiation, or, down to 30 GHz, as millimetre-wave.

2.3.2.1. Infrared radiation\(^3\)

Observing terrestrial infrared radiation was one of the first applications of Earth observation from space. Already on the second TIROS satellite, an instrument measuring terrestrial infrared radiation was included, the Medium Resolution Infrared Radiometer (MRIR), with five channels and 55 km horizontal resolution. Since then, numerous imagers and sounders have been launched on polar-orbiting, geostationary, and other satellites. The first sounders were launched on the Nimbus-3 satellite in 1969 (Rodgers 2000), and HIRS was first put in orbit on-board Nimbus-6 on 12 June 1975. Subsequent editions of HIRS were flown on TIROS-N, NOAA-6, and all later NOAA and MetOp polar orbiting satellites to date. Operational imaging dates back to the launch of the Very High Resolution Radiometer (VHRR) on-board NOAA-2 on 15 October 1972. VHRR was succeeded by

\(^{3}\)Information for this and the following paragraph is partly obtained from http://nssdc.gsfc.nasa.gov/nmc, http://www.wmo-sat.info/oscar, and Cracknell (1997)
AVHRR with the launch of TIROS-N on 13 October 1978. Since then, AVHRR has been present on 17 operational polar-orbiting satellites by NOAA and MetOp, with 1 more planned (MetOp-C), although it is now being superseded by the more sophisticated VIIRS with the launch of Suomi National Partnership Program (NPP) in October 2011. The Fengyun-3 and MetOp-SG programmes (will) carry sensors comparable to VIIRS.

The NOAA KLM User’s Guide contains detailed technical information about the operational meteorological sensors on the NOAA satellites (Robel et al. 2009). Most of the information in the paragraphs below is retrieved from this User’s Guide.

The High resolution Infrared Radiation Sounder (HIRS) is a sounder with 20 channels. It has one channel measuring solar radiation, 12 channels measuring purely terrestrial infrared radiation, and 7 channels measuring a mixture of both. It obtains 56 Earth view samples across track at a maximum scan angle of 49.5°, corresponding to a swath width of 2179 km. The distance between the centres of neighbouring samples is 39 km. HIRS/3 (on NOAA-15, -16 and -17) has a footprint diameter at nadir of 20 km, with 19 km between footprint edges. With HIRS/4 (on NOAA-18, -19, MetOp-A, and MetOp-B), the footprint diameter is reduced to 10 km, so that the gap between neighbouring footprints increases to 29 km. HIRS/4 footprints are illustrated in Figure 2.6.

The Advanced Very-High Resolution Radiometer (AVHRR) is an imager with 6 channels (5 are in operation simultaneously) in the visible and infrared. As briefly discussed before, AVHRR has a long history, dating back to 1978 or even 1972. The latest generation, AVHRR/3, flies on NOAA-15 to 19 and on MetOp-A to C, which are all sun-synchronous satellites. Channels 1, 2, and 3A measure in the visible and near-infrared range and are calibrated in terms of albedo. Channels 4 and 5 measure terrestrial radiation. Figure 2.3 shows the spectral response functions for channels 4 and 5. Channel 3B is located in a spectral region where both reflected solar radiation and emitted terrestrial radiation are significant. Channels 3B, 4, and 5 are all calibrated in terms of brightness temperature.

A single AVHRR scanline consists of 2048 contiguous pixels ranging from 1.1 × 1.1 km² at nadir to 6.24 × 2.3 km² at the scan edge. Due to storage and bandwidth limitations, not all data are downlinked globally. Global, so-called Full Resolution Area Coverage (FRAC) data are available only for MetOp satellites. For other satellites, full-resolution data are available only for particular regions of the world, at pre-order, or by direct downlink for users operating their own ground station. Global data are available only at a reduced resolution in a product known as Global Area Coverage (GAC). In GAC data, four adjacent samples are averaged and only every third scanline is downlinked, thus reducing the data more than tenfold. Figure 2.6 shows the footprints of GAC data. In this thesis, all AVHRR measurements used are GAC.

2.3.2.2. Passive microwave radiation

Karlsson (1997) defines the microwave region as 0.1 to 10 cm. This corresponds to a frequency range of 3 to 300 GHz. Lower frequencies are generally not useful for atmospheric remote sensing. Higher frequencies are within the domain of sub-millimetre radiation, described in this thesis in chapter 6.

Spaceborne microwave observations of the atmosphere began with the launch of Nimbus-5 on 11 December 1972, carrying the single-channel Electrically Scanning Microwave
Radiometer (ESMR) for sea-ice observations. Since then, microwave radiometers have gradually improved. The 4-channel Microwave Sounding Unit (MSU) was first launched on TIROS-N on 13 October 1978. MSU has channels from 50.3 to 57.95 GHz and a footprint at nadir with a diameter of 124 km. Nine copies were launched between TIROS-N up to and including the launch of NOAA-14 on 30 December 1994. The similar Special Sensor Microwave/Temperature (SSM/T)-1 has seven channels from 50.5 to 59.4 GHz and a footprint at nadir with a diameter of 174 km. SSM/T-1 is part of the payload of several satellites from the Defense Meteorological Satellite Program (DMSP) series. Both MSU and SSM/T-1 were temperature sounders. Special Sensor Microwave/Imager (SSM/I), first launched on 20 June 1987 on DMSP 5D-2/F8, has seven channels from 19.35 to 85.5 GHz and a footprint between 15 km and 69 km depending on frequency. Higher frequency channels became available with SSM/T-2 (from 28 November 1991 on DMSP 5D-2/F11), AMSU (from 13 May 1998 on NOAA-15), and MHS (from 20 May 2005 on NOAA-18), instruments including channels around the strong water vapour absorption line at 183.310 GHz. With the launch of Suomi NPP in October 2011, this has further improved with the Advanced Technology Microwave Sounder (ATMS). The Chinese polar meteorological satellite series, Fengyun-3, includes the Micro-Wave Humidity Sounder (MWHS)-1 with the same channels as MHS (Zou, Ma and Qin 2012), and will include MWHS-2 with channels similar to ATMS. MWHS and ATMS have not been used in this thesis.

Apart from these operational sensors, numerous special-purpose, scientific sensors exist. For example, in late 2011, the Indian-French satellite Megha-Tropiques was launched. This carries the Sondeur Atmosphérique du Profil d’Humidité Intertropicale par Radiométrie (SAPHIR), with six channels around the 183.310 GHz absorption line (versus only three for MHS, AMSU-B and MWHS). The reader is referred to the review papers by Kidd, Levizzani and Bauer (2009) and Thies and Bendix (2011), as well as the Observing Systems Capability Analysis and Review (OSCAR) database at http://www.wmo-sat.info/oscar, for more information on recent, current and future sensors.

The Advanced Microwave Sounding Unit (AMSU)-B (Saunders et al. 1995) and the newer Microwave Humidity Sounder (MHS) (Kleespies and Watts 2007) are microwave radiometers designed for atmospheric sounding of humidity. AMSU-B is carried on NOAA-15, -16, and -17. MHS is carried on NOAA-18, -19, and MetOp-A and -B, and will be carried on MetOp-C.

Figure 2.5 shows the positions of the channels in AMSU-B. MHS is similar, but channel 2 is at 157 GHz (instead of 150 GHz) and channel 5 has only the band at 190 GHz (instead of 183 ± 7 GHz). There are also differences in the polarisation. These bands allow for sounding of atmospheric water vapour, as illustrated in Figure 2.4. The dual sideband characteristic of AMSU-B channels 3 to 5 are typical for microwave radiometers.

According to the NOAA KLM User’s Guide (Robel et al. 2009), the antenna spatial response function and the scan characteristics of AMSU-B and MHS are quite similar. Both scan across-track at 90 different viewing angles symmetrically around nadir, with no actual nadir observation. The maximum viewing angle for AMSU-B is 48.95°, for MHS it is 49.44°. For a nominal spacecraft altitude of 833 km, the instantaneous field of view has a diameter of 16 km at nadir, corresponding to the half-power level of an assumed Gaussian antenna pattern. This increases to roughly 52 × 27 km² at the outermost footprint. The
Figure 2.5. — Total opacity as a function of frequency for the AMSU-B frequency range. The opacity (see Equation 4.4) is shown for seven different values of Precipitable Water Vapour (PWV). The grey shaded areas show the locations of the AMSU-B channels. Channel positions from the NOAA KLM User’s Guide (Robel et al. 2009).

This is Figure 1 from Paper II.
2.3. Specific observation technologies

The effective field of view is slightly larger, because during the integration time, the spacecraft is moving with respect to the Earth’s surface and the instrument scan angle is changing with respect to the spacecraft. Variations in satellite altitude above the surface cause slight changes in the size of the instrument footprint. Figure 2.6 illustrates the effective field of view of MHS.

2.3.3. Active

Active sensors rely on artificial radiation. The source and the detector may either be on the same platform, such as with radar or lidar, or on different platforms, such as in the use of Global Navigation Satellite System (GNSS) for atmospheric soundings.

A radar transmits pulses of radiation at microwave frequencies and measures the backscattered signal as a function of time, thus obtaining information about the distance at which the radiation was backscattered. Radars are widely used for ground-based remote sensing of precipitation (e.g. Karlsson 1997), and spaceborne imaging radars have a significant history. However, only two radars for measuring clouds and precipitation are currently in orbit: one on the Tropical Rainfall Measuring Mission (TRMM), for liquid clouds and precipitation, and one on CloudSat, for solid clouds and precipitation.

The Cloud Profiling Radar (CPR) is a 94 GHz (W band) radar carried on-board CloudSat (Stephens et al. 2002). It has a footprint of approximately 1.4 km. Its primary purpose is to observe the vertical structure of clouds and precipitation. CloudSat and CPR were launched 28 April 2006.

A lidar operates by a similar principle to a radar, but uses radiation emitted by a laser, at wavelengths at or near the visible part of the spectrum. Lidars can observe smaller particles than radars, but the signal gets attenuated more rapidly.

The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) is a dual-wavelength polarisation lidar carried on-board Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) (Winker, Pelon and McCormick 2003; Winker et al. 2009) and is the only Earth-observing spaceborne lidar operational at the time of writing (future lidars for altimetry, wind profiling, or aerosol profiling are not planned until after 2015). It operates at 523 nm and 1064 nm and has a footprint of approximately 70 m. CALIOP observes the vertical structure of aerosol layers and for clouds with a visible optical depth of \( \tau < 5 \).

CloudSat, CALIPSO, and a number of other satellites, are part of the A-Train, or “afternoon train”, a constellation consisting of research satellites dedicated to atmospheric monitoring⁴. They fly in a sun-synchronous orbit with an altitude of 705 km, a LTAN of 13:30, and an inclination of 98°. An overview of the A-Train is given by L’Ecuyer and Jiang (2010).

The nearly equal ground tracks between the A-Train satellites allow for numerous synergies, such as combined active-passive ice cloud retrievals used in Paper IV and Paper V. From launch until April 2011, when CloudSat suffered a battery anomaly, CloudSat and CALIPSO were only 12.5 seconds or 93.8 km apart. They had coincident measurements for more than 90% of the time (L’Ecuyer and Jiang 2010). Additionally, ⁴According to L’Ecuyer and Jiang (2010), Aura project scientist Mark Schoeberl coined the name A-Train after a jazz song written by Billy Strayhorn and Duke Ellington.
NOAA-18 and NOAA-19 are sufficiently close to the A-Train to have a large number of collocations, as shown for NOAA-18 in Paper I (NOAA-19 was not launched at the time the work for Paper I was performed). Collocations are further described in section 2.4.

At the time of writing (summer 2013), CloudSat is operational in a daylight only mode, after suffering from a battery anomaly in 2011. It is back in the A-Train, but coincident measurements with CALIPSO occur much less frequently than before, and synergistic applications between radar and lidar are limited.

2.4. Collocations

Collocations form a major tool for the research in this thesis. They are introduced in Paper I (including a review of earlier work), and also of major importance for the work presented in Paper II, Paper IV, and Paper V, as well as in the studies by John et al. (2011), John et al. (2013), and Eliasson et al. (submitted 2013). A review of prior research using collocations is given in Paper V.

Illustration of sensor footprints over the Kiruna region

Figure 2.6. — Illustration of footprint sizes for CloudSat, MHS, and HIRS and pixel size for AVHRR. The data for the background map are ©OpenStreetMap contributors.
2.4. Collocations

2.4.1. Criteria
A collocation occurs if different instruments observe an overlapping area or volume at approximately the same time. Typically, the sensor footprint is large compared to the geophysical quantity of interest, whereas the duration of a measurement is negligibly short. Therefore, the spatial criterion compared to the footprint size is more strict, than the temporal criterion compared to the measurement duration. In the context of this thesis, collocated instruments are spaceborne remote sensing instruments, usually on different satellites. However, a collocation can also be considered between a satellite-based and a ground-based measurement. Figure 2.6 illustrates the sizes of footprints and pixels for various instruments.

The interpretation of “the same place” and “the same time” depends on the instruments and the usage of the collocations. For example, Seiz, Tjemkes and Watts (2007) match visible imagery on adjacent geostationary satellites to retrieve cloud top height by stereophotogrammetry. For such a high-resolution observation of a rapidly changing environment, the match should ideally occur within a minute. A quite different situation occurs when comparing radiosonde-measured UTH with satellite-retrieved UTH, as done by Moradi et al. (2010). This quantity is changing relatively slowly, and the satellites have a larger field of view. Here, a maximum time interval of two hours is considered acceptable. Such conditions are a tradeoff between the amount of data available and the required precision for the collocation. One can also attempt to correct the location for winds and considered altitude (geolocation data refer to the ground), but this is not done for the present study.

2.4.2. Method
I developed a toolkit to easily find collocations between any pair of sensors. This toolkit has been used to produce several collocated datasets. Over the years, I invested considerable time and effort to make this toolkit as generic as possible. This has led to a wide variety of applications not foreseen when preparing the first collocated dataset.

The first version of the algorithm is described in Paper I. The algorithm was significantly revised shortly after this paper was published, because it was not suitable for new collocated datasets. The revised algorithm is described in Paper II and Paper V.

The code is implemented in MATLAB® and publicly distributed as part of the Atmlab software package. The collocation toolkit can be downloaded through http://www.sat.ltu.se/projects/collocations/toolkit.php.

2.4.3. Collocated datasets
The collocation toolkit has been used in a large number of studies.

The first dataset generated with the collocation toolkit is a dataset matching the passive sensors AMSU, MHS, and HIRS on operational meteorological satellites, with the active CPR carried by CloudSat. It also includes IWP from Microwave Surface and Precipitation Products System (MSPPS). CPR measurements, in particular IWP from product 2B-CWC-RO, are averaged over the MHS footprint. Paper I presents three examples of possible applications, two of which are expanded in Paper IV and Paper V, respectively.
The work by John et al. (2011) relies on a collocated dataset between AMSU-B and cloud-cleared HIRS. Here, the sensors are on the same satellite, so the collocation is relatively straightforward. The flexibility of the collocation toolkit makes it easy to obtain the collocations in practice.

Another major class of collocations is between different copies of the same sensor, or between very similar sensors. For example, section 2.3.2.2 describes several instruments with channels near the 183 GHz water vapour absorption line: MHS, AMSU-B, SSM/T-2, SAPHIR, and MWHS. All those sensors are carried on satellites in low earth orbit, and all but SAPHIR are on sun-synchronous satellites. Some of those satellites have actively controlled orbits, but others do not, and have a LTAN that is slowly drifting as a function of time. Such drift leads to pairs of satellites incidentally having the same LTAN, which is a necessary requirement for global collocations between sun-synchronous satellites.\footnote{It is not sufficient, however: MetOp-A and MetOp-B have exactly the same orbit, but their position within the orbit is 180° apart, so they never collocate.}

**Paper II** uses a subset of such global collocations, SNOs, to investigate the latitudinal dependence of inter-instrument biases. John et al. (2013) use collocations between pairs of MHS and AMSU-B to investigate scan asymmetries. Projects involving collocations between SSM/T-2 and AMSU-B as well as between SAPHIR and MHS are ongoing.

For the work presented in **Paper IV**, collocations are similar to the ones in **Paper I**, but on a level-2 basis: rather than collocating radiances, **Paper IV** uses collocations between different IWP products. Apart from the aforementioned CPR 2B-CWC-RO IWP, this includes IWP from combined radar-lidar (CloudSat and CALIPSO), from solar reflectences (AVHRR and MODerate resolution Imaging Spectroradiometer (MODIS)), and from passive microwave (MHS).

Another ongoing project uses collocations between GNSS Receiver for Atmospheric Sounding (GRAS) and AMSU-A.

Indeed, the collocation toolkit allows for a great variety of applications.
Chapter 3.

Ice clouds

Cloud properties have been identified by the WMO Global Climate Observing System (GCOS) as Essential Climate Variables (ECVs) (WMO 2010, p. 61). One of the aims of this thesis is to improve quantitative measurements of cloud properties from space (see also section 1.2). This chapter describes in more detail why clouds and in particular ice clouds are important and what properties we can use to quantify clouds and precipitation. Those properties may be related to the Earth’s radiation budget, to hydrology, or to parameters required to perform radiative transfer simulations (see also chapter 4).

Such properties can be categorised in different ways. In this thesis, microphysical properties (section 3.2) are intensive (i.e. their value does not depend on the size of the system) and their value can vary as a function of location. Examples are the effective radius and the Ice Water Content (IWC). Optical properties (section 3.3) are similar, but relate to interaction with electromagnetic radiation. Macrophysical properties (section 3.4) relate to a property of an entire cloud or atmospheric column, and can be either intensive or extensive. Examples are visible optical depth, column mass, or cloud top height.

3.1. Radiative effects

Like greenhouse gases, clouds absorb and re-emit terrestrial radiation, and therefore heat the Earth’s surface. Unlike greenhouse gases, clouds reflect solar radiation, and therefore cool the Earth’s surface. The magnitude of the feedback is much smaller than the magnitude of either contribution, so a relatively small difference in either solar or terrestrial fluxes can effect a relatively large difference in net flux. The global net cloud feedback in the climate system depends on surface temperature, atmospheric temperature, and cloud properties. A thorough review on cloud radiative feedback in the climate system is presented by Stephens (2005), and most of the material in this section is drawn from there.

The magnitudes of the solar and terrestrial fluxes through a cloudy atmosphere depend on many macrophysical and microphysical cloud properties. Macroscopic changes in cloud cover, temperature, or geographic distribution, affect mean global fluxes, as does a change in the diurnal cycle of any of those properties. Cloud albedo depends on cloud optical depth, which in turn depends on geometric cloud thickness and physical properties such as particle sizes, total water or ice content, particle shapes (for ice clouds), and others. An increase in global-mean temperature may change the total solar and terrestrial fluxes and therefore the energy budget of the Earth. The effect of climate change on the sign of the energy budget is not clear. Quoting Stephens (2005): “we have no clear theory that suggests the accumulated effects of cloud feedbacks are in any way a function of
Chapter 3. Ice clouds

global-mean temperature”. More recently, Dessler (2010) find that the net feedback is likely positive (0.54 ± 0.74 W m⁻² K⁻¹), confirmed by both models and observations. The subject remains an area of active research.

Without considering clouds, at most low latitudes, the radiation budget is positive: the total power of solar radiation absorbed exceeds the total power of terrestrial radiation emitted to space. At high latitudes, this budget is negative (Peixoto and Oort 1992, section 6.8.2).

At low latitudes, the surface is warm and the surface emits large quantities of radiation. Clouds absorb this. Depending on their altitude, cloud tops are colder or much colder than the surface, and hence emit much less radiation to space than they absorb from the surface, thereby reducing outgoing terrestrial radiation. Meanwhile, a significant part of incoming solar radiation is transmitted, in particular if the clouds are thin, such as cirrus. Therefore, clouds over warm surfaces have a surface heating effect, in particular if they are thin.

At high latitudes, the surface is cold and the surface does not emit much radiation. The temperature difference between the surface and cloud tops is much smaller than above warm surfaces (cloud tops may even be warmer than the surface), so clouds do not affect outgoing terrestrial radiation much. As they still reflect solar radiation, they have a surface cooling effect.

Thus, clouds enhance the meridional gradient of the net radiation budget and, consequently, atmospheric circulation (Stephens 2005).

Both cloud optical thickness \( \tau \) and effective radius \( r_e \) have direct importance for the cloud radiative feedback in the climate system. Optical thickness \( \tau \) affects cloud reflectivity for solar radiation, as described by Nakajima and King (1990). Cloud particle effective radius \( r_e \) affects cloud emissivity for terrestrial radiation for thin clouds. Water, snow, and ice all have an emissivity at 8 to 12 \( \mu \)m of more than 0.98 (Rees 2001, Figure 3.25). Hence, the emissivity of a non-transparent cloud at those wavelengths is close to unity. However, for thin clouds, this emissivity is less than one, and terrestrial radiation may be transmitted\(^1\). In conclusion, both \( \tau \) and \( r_e \) have direct importance for cloud radiative feedback in the climate system.

3.2. Microphysical properties

Cloud microphysical properties are important for understanding cloud radiative properties, cloud optical properties, and cloud hydrology. Cloud microphysics is discussed in detail by e.g. Rogers and Yau (1976). Here, only a few cloud properties relevant for this thesis will be discussed.

Hydrometeors are atmospheric particles consisting of condensed (possibly frozen) water, such as cloud droplets, raindrops, or snowflakes. Hydrometeors can be liquid or frozen (or a mixture of the two) and come in sizes ranging from particles smaller than a few micrometer to several hundred micrometer (Wallace and Hobbs 2006), or even larger in the case of snowflakes or hailstones. Additionally, frozen particles come in a large variety

\(^1\) This can be exploited for remote sensing purposes. For example, Rädel et al. (2003) describe the dependence of cloud emissivities at 8 \( \mu \)m and 11 \( \mu \)m on particle effective radius \( r_e \).
Figure 3.1. — Examples of ice shapes observed during several flight campaigns. Figure from Heymsfield et al. (2002).
of shapes. The shape of a cloud ice particle or of a (precipitating) snowflake depends on the history of temperature, humidity, and collisions with other particles, that the particle experienced (Libbrecht 2005). Figure 3.1 shows examples of ice particles as collected by Heymsfield et al. (2002) during several tropical field campaigns.

Retrievals of properties such as IWP need ice cloud particle size and shape information. Since those cannot be retrieved directly and independently, they need to be parametrised. Size distributions are often parametrised based on in situ measurements. The size distribution gives a quantity describing the number of particles (such as particle number or mass density) as a function of the particle size. Since particles may be aspherical, particle size is not uniquely defined. Frequently chosen sizes include the volume equivalent sphere diameter and the maximum diameter. One size distribution (used, for example, by ARTS) is given by McFarquhar and Heymsfield (1997). Theirs is the sum of a first-order gamma distribution and a log-normal distribution. The distribution depends on IWC and temperature.

Particle shapes are (even) harder to parametrise, because there is no single, unique, numerical property to describe the shape.

### 3.3. Optical properties

Optical properties describe the interaction of electromagnetic radiation with cloud or precipitation particles. Note that “optical” does not imply visible radiation, but also applies to other wavelengths, such as infrared or microwave.

Optical properties, such as the extinction coefficient, can be measured with reasonable accuracy in a laboratory environment. These optical properties are a function primarily of three microphysical properties: particle size, particle phase, and particle shape. Optical properties also depend on temperature, because the refractive index does. In the context of this thesis, optical properties are relevant in the context of radiative transfer simulations and their discussion will be deferred to section 4.1.2.

### 3.4. Macrophysical properties

Macrophysical physical properties are important for the radiative balance and optical properties, but also for hydrological applications and interesting in their own right.

Ice Water Content (IWC) (a microphysical property) is the mass density of ice in an air parcel, commonly expressed in g/m$^3$. The macrophysical property Ice Water Path (IWP) is the IWC integrated along a vertical column,

\[
IWP = \int_{z_b}^{z_t} IWC(z)dz \quad [g/m^2],
\]

where $z$ is the position along the column and IWP is considered between altitudes $z_b$ and $z_t$. Usually, $z_b$ and $z_t$ span the entire atmosphere. If $z_b$ and $z_t$ do not span the entire atmosphere, the IWP is a partial-column IWP or pIWP.
Cloud Top Height (CTH), Cloud Top Pressure (CTP), and Cloud Top Temperature (CTT) describe the highest physical extent of the cloud. The top of a cloud may not be sharply defined, so some care needs to be taken when considering these quantities. Cloud-emitted infrared radiation reaching a satellite sensor originates from the top layers of a (sufficiently thick) cloud. Therefore, parameters describing this region, such as cloud top height, pressure, and temperature, are important to consider in satellite remote sensing of clouds.
Chapter 4.

Radiative transfer

Radiative transfer describes the transfer of radiation through a medium, from emission to absorption. In the context of this thesis, radiation is emitted by the Earth’s surface or atmosphere, transferred through the atmosphere (including clouds), and finally either escaping into space, or absorbed by the Earth’s surface, the Earth’s atmosphere, or by a spaceborne detector. In other words, in this thesis, radiation is of terrestrial origin and considered in the context of remote sensing. Other aspects, such as solar radiation or broadband fluxes relevant for energy flux calculations, are not considered. For a comprehensive introduction, the reader is referred to one of the many textbooks on the subject.

4.1. Theory

Radiative transfer is most concisely summarised by the radiative transfer equation. The radiative transfer equation is valid under the following conditions:

- Local Thermodynamic Equilibrium (LTE). Under LTE conditions, temperature is well-defined from collisions between molecules. A necessary and sufficient condition for LTE is that the rate of collisions between molecules is much higher than the rate of spontaneous energy level changes (i.e. emission and absorption) within the molecules (Thomas and Stamnes 2002). LTE is fulfilled if the medium density is sufficiently high, the frequency is sufficiently low, and the intensity is not too high. This is true for terrestrial radiation in the Earth’s troposphere and stratosphere, but breaks down in the upper atmosphere and for radiation at UV or higher frequencies. Under LTE, blackbody emission is described by Equation 2.1 on page 8.

- Fully elastic scattering. In case of elastic scattering, the frequency of radiation remains constant before and after a scattering event. For the radiative transfer of terrestrial radiation, virtually all scattering can be considered as elastic, and inelastic scattering can be ignored for most practical purposes.

If those conditions are fulfilled, then the scalar radiative transfer equation (where scalar means that polarisation is not considered) can be formulated as

$$\frac{dI(\nu, \mathbf{r}, \mathbf{\hat{n}})}{ds} = -\beta_e(\nu, \mathbf{r}, \mathbf{\hat{n}})I(\nu, \mathbf{r}, \mathbf{\hat{n}}) + \beta_a(\nu, \mathbf{r}, \mathbf{\hat{n}})B(\nu, T(\mathbf{r}))$$

$$+ \int Z(\nu, \mathbf{r}, \mathbf{\hat{n}}, \mathbf{\hat{n}}') I(\nu, \mathbf{r}, \mathbf{\hat{n}}') d\mathbf{\hat{n}}', \quad (4.1)$$
where $I$ refers to spectral radiance in e.g. W sr$^{-1}$ m$^{-2}$ Hz$^{-1}$, $\nu$ is frequency, $r$ is the position, $\hat{n}$ and $\hat{n}'$ are unit vectors characterising outgoing and incoming direction, respectively, $s$ is the position along the line of sight (in units of length), $\beta_e$ is the extinction coefficient, $\beta_a$ is the absorption coefficient, $B$ is blackbody radiation as given by Equation 2.1, $Z$ is the scattering phase function, and $T$ is the local temperature.

Equation 4.1 describes the change of radiation along the line of sight. Three terms contribute to this change:

- The first term, $\beta_e(\nu, r, \hat{n})I(\nu, r, \hat{n})$, is preceded by a negative sign and describes losses through extinction. Radiation losses can occur either through absorption or by the radiation being scattered away from the line of sight, i.e. $\beta_e = \beta_a + \beta_s$, where $\beta_s$ is the scattering coefficient.

- The second term, $\beta_a(\nu, r, \hat{n})B(\nu, T(r))$, is the emission source. Under LTE conditions, Kirchhoff’s law states that absorptivity $\alpha$ equals emissivity $\epsilon$. If this were not the case, an object could reach infinite temperature by systematically absorbing more energy than it emits.

- The third term, $\int_{4\pi} Z(\nu, r, \hat{n}, \hat{n}')I(\nu, r, \hat{n}') d\Omega'$, is the scattering source term. It describes radiation originating from any direction scattered into the line of sight.

The absorption coefficient $\beta_a$ (in m$^{-1}$) is the absorption cross-section $\sigma_a$ per unit volume. $\sigma_a$ (in m$^2$) for a particle or molecule describes how much radiation is absorbed. The absorption efficiency $Q_a$ is the ratio between $\sigma_a$ and the physical cross-section, and describes what fraction of radiation passing through the surface is absorbed. The scattering coefficient $\beta_s$, cross-section $\sigma_s$, and efficiency $Q_s$ are defined accordingly, but relate to scattering. The extinction coefficient $\beta_e$, cross-section $\sigma_e$, and efficiency $Q_e$ are the sum of the respective absorption and scattering quantities.

The scattering phase function $Z$ describes the distribution of radiation after a beam is scattered, or, equivalently, the probability density function for the direction in which a single photon is scattered. It may be normalised in numerous ways, such as to 1, to $4\pi$ (the solid angle covering the full sphere), to $\beta_s$, to $\sigma_s$, or to $Q_s$. In Equation 4.1, the scattering phase function is normalised with respect to $\beta_s$, e.g. $\int_{4\pi} Zd\Omega = \beta_s$.

Figure 4.1 illustrates the relation between various properties occurring in radiative transfer.

### 4.1.1 Clear-sky radiative transfer

In the context of this chapter, “clear-sky” radiative transfer means radiative transfer in a purely gaseous atmosphere, whereas “cloudy” radiative transfer means radiative transfer with scattering particles. In reality, gases and aerosols also scatter radiation. In this thesis, aerosols are not considered, and for terrestrial radiation, scattering from gases can be neglected. Physical processes important in clear-sky radiative transfer are still important in cloudy radiative transfer and the solution for clear-sky radiative transfer can be considered as a subset of the solution for cloudy radiative transfer. However, the solution approach
is different, because the cloudy problem is considerably more complex (more factors are involved) and more complicated (a more difficult problem to solve).

Since we excluded scattering by definition, the only relevant processes for clear-sky radiative transfer are absorption and emission, so that

$$\frac{dI(\nu, \mathbf{r}, \mathbf{n})}{ds} = \beta_a(\nu, \mathbf{r}, \mathbf{n}) \left[ B(\nu, T(\mathbf{r})) - I(\nu, \mathbf{r}, \mathbf{n}) \right],$$

which is Equation 4.1 with \( \beta_e = \beta_a \) and \( Z = 0 \). Equation 4.2 can be solved analytically. If emission can be neglected (e.g., because there is a very strong background), it becomes the Lambert-Beer law,

$$I = I_0 e^{-\tau},$$

where \( I_0 \) is the background spectral radiance, and \( \tau \) is the optical depth or the opacity,

$$\tau \equiv \int_0^\tau \beta_a \, dl.$$

As illustrated in Figure 4.1, the determination of \( \beta_a \) starts with fundamental quantum-mechanical considerations. Electromagnetic radiation interacts with the various gases present in the atmosphere. Each constituent gas absorbs at particular regions in the electromagnetic spectrum. From quantum-mechanical considerations, multi-atom molecules (as almost all gases constituting the atmosphere are) can vibrate and rotate according to
discrete states. Only photons with an energy equal to the difference between a currently held energy level and a higher energy level can be absorbed. Those frequencies correspond to absorption lines. Equivalently, higher energy levels are reached due to thermal excitation, and photons with corresponding energies are emitted by anything with non-zero temperature.

From just considering the transitions, absorption lines would have no width and absorption would occur only if the photon frequency exactly matches the transition frequency. However, several processes cause so-called line broadening. Due to the Heisenberg uncertainty principle, the frequency associated with an absorption line is not exactly determined, but has a certain \( a \) natural broadening. Molecules do not exist alone, but are part of an ensemble of many molecules forming a gas, a liquid, or a solid. Consequently, nonzero temperature and pressure cause \( b \) thermal and \( c \) pressure broadening, respectively. Natural broadening is very small compared to thermal and pressure broadening, and in the troposphere, pressure broadening is the dominating mechanism.

For the absorption coefficient \( \beta_a \), all lines for all constituent gases need to be considered. Those lines and their dependences are catalogued in databases such as the HIgh resolution TRANsmission (HITRAN) database (Rothman et al. 2009).

More information on gaseous absorption can be found in textbooks such as Goody and Yung (1995), Liou (2002), or in the ARTS Theory Guide (Eriksson et al. 2013a) and references therein.

### 4.1.2. Scattering

Like molecules, particles emit and absorb electromagnetic radiation. Unlike molecules at infrared and microwave wavelengths, particles additionally scatter incoming radiation. For remote sensing of clouds, scattering, absorption, and emission by clouds can all be relevant, depending on cloud phase, particle size, and wavelength.

The physics of scattering by particles is an entire field in itself. This thesis only scratches the surface. For details, the reader is referred to textbooks such as the ones by Mishchenko, Travis and Lacis (2002) or Liou (2002), or to the ARTS Theory Guide (Eriksson et al. 2013a) and references therein. Material in this section is drawn from these sources and from Wallace and Hobbs (2006, section 4.4.1).

In clear-sky radiative transfer, the absorption coefficient \( \beta_a \) can (in theory) ultimately be determined from quantum-mechanical considerations along with knowledge of the atmosphere, as described above and illustrated in Figure 4.1. For radiative transfer in a cloudy atmosphere, those considerations are not sufficient for the determination of \( \beta_a \) and the determination of \( \beta_a \) is not sufficient to solve the radiative transfer equation. To solve Equation 4.1, additionally the extinction coefficient and the scattering phase function are needed. As illustrated in Figure 4.1, those properties can be obtained starting from the fundamental material properties for the scattering object (such as water in liquid or frozen form).

Interaction between radiation and particles depends on the refractive index of the material (which in turn is a function of phase and temperature), on particle size in relation to wavelength, and on particle shape. The size parameter relates to the order of magnitude
4.1. Theory

of the ratio between particle size and radiation wavelength,

$$x \propto \frac{a}{\lambda},$$

(4.5)

and gives an indication on how the radiation interacts with the particle. Here, $\lambda$ is wavelength and $a$ is a characteristic particle size, which may be the diameter for a spherical particle or the maximum diameter for a non-spherical particle. The order of magnitude for $x$ determines what model best describes the interaction. The wavelength of relevant terrestrial electromagnetic radiation varies by four orders of magnitude (see page 8) and so do the sizes of atmospheric ice particles (see section 3.2). Since those are independent, the size parameter $x$ can vary over a range of 7 to 8 orders of magnitude. Table 4.1 lists a number of example size parameters for hydrometeors that may appear in the atmosphere.

<table>
<thead>
<tr>
<th>Type</th>
<th>Characteristic length (mm)</th>
<th>$\lambda = 10 \mu m$</th>
<th>$\lambda = 1.6 \text{ mm } \approx 183 \text{ GHz}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Snowflake</td>
<td>10.0</td>
<td>$6 \times 10^2$</td>
<td>$4 \times 10^1$</td>
</tr>
<tr>
<td>Raindrop</td>
<td>1.0</td>
<td>$6 \times 10^1$</td>
<td>4</td>
</tr>
<tr>
<td>Ice crystal</td>
<td>0.1</td>
<td>6</td>
<td>$4 \times 10^{-1}$</td>
</tr>
<tr>
<td>Cloud droplet</td>
<td>0.01</td>
<td>$6 \times 10^{-1}$</td>
<td>$4 \times 10^{-2}$</td>
</tr>
</tbody>
</table>

Table 4.1. — Size parameters for typically occurring hydrometeors in the atmosphere, observed at infrared and microwave wavelengths. Based on Liou (2002, Table 5.1).

Different models exist to calculate particle optical properties in either exact or approximate ways. Many of those models are described in Mishchenko, Travis and Lacis (2002) and in Mishchenko, Hovenier and Travis (2000). Such models take as an input the refractive index $n$ and particle physical properties (size and shape), and calculate the scattering phase function and two parameters uniquely characterising $Q_a$, $Q_s$, and $Q_e$, as summarised in Figure 4.1.

For spherical particles, Lorenz-Mie theory describes exactly and analytically the interaction between electromagnetic radiation and scattering particles, by expressing the scattering efficiency and the scattering phase function in terms of refractive index and size parameter. A derivation can be found in Liou (2002, section 5.2).

When $x \ll 1$, the Rayleigh scattering approximation applies and the scattering efficiency follows $Q_s \propto \lambda^{-4}$. In the Rayleigh regime, scattering is weak and gets weaker rapidly with increasing wavelengths. Rayleigh scattering is relevant for visible light scattering from molecules and causes the blueness of the sky, because blue light (which has a short wavelength) undergoes (multiple) scattering and arrives at the observer from all directions, whereas yellow light (which has a longer wavelength) arrives mainly straight from the Sun. Rayleigh scattering also occurs for microwave scattering from small cloud particles, but the number density of cloud particles is too small for this to be of any significance for passive remote sensing purposes. The Rayleigh approximation is valid for scatterers of any shape. Rayleigh scattering phase functions are quite smooth, like the one shown in panel b in Figure 4.2.
For scattering from atmospheric ice particles, $x \not\ll 1$, $Q_s \not\ll \lambda^{-4}$, and the Rayleigh approximation is no longer valid. In this regime, the shape of the scattering particle is important.

For non-spherical particles, Lorenz-Mie theory does not apply and no analytical solution exists (Liou 2002, section 5.4). The T-Matrix method is applicable to rotationally-symmetric particles. Generic methods that can be applied to particles of arbitrary shape are computationally expensive. Mishchenko, Hovenier and Travis (2000) provide an overview of numerical methods and laboratory measurements to determine optical properties of particles of arbitrary shape. Such methods include, but are not limited to, Discrete Dipole Approximation (DDA) and Finite Difference Time Domain (FDTD). Table 4.2 summarises

<table>
<thead>
<tr>
<th>Shape</th>
<th>Method</th>
<th>Remark</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spherical</td>
<td>Lorenz-Mie</td>
<td>exact</td>
</tr>
<tr>
<td>Rotationally symmetric</td>
<td>T-Matrix</td>
<td></td>
</tr>
<tr>
<td>Other</td>
<td>DDA, FDTD, ...</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2. — Summary of scattering calculation methods that can be applied to ice particles (Mishchenko, Hovenier and Travis 2000).

For particles of any shape, when $x \not\ll 1$, the Rayleigh approximation is invalid. For spherical non-absorbing particles, for example, $Q_s(x)$ oscillates as a function of $x$ and $\lim_{x \to \infty} Q_s = 2$. In a cloud, many different particle sizes are present, and $x$ is sufficiently large so that $dx/d\lambda$ is small. Therefore, the total scattering of visible light does not depend on wavelength, and the colour (spectral composition) before and after scattering are the same. This is why clouds generally appear white (or grey), as only in a cloudy sky the Sun reveals its true white colour.

Figure 4.2 shows two scattering phase functions, here normalised to the scattering cross-section in $\text{m}^2$. In Figure 4.2a, $x = 58$, and in Figure 4.2b, $x = 0.28$, so neither scattering phase function is in the Rayleigh regime, although the microwave one is quite close. Several differences are apparent from the figure:

- The scattering cross-section in the infrared example is much larger than the one in the microwave example. This makes intuitive sense, because the particle is so much larger relative to the wavelength, as quantified by the size parameter. Indeed, thin clouds consisting of small particles are transparent at microwave frequencies, whereas a cloud does not need to be very thick for it to be opaque at infrared wavelengths.

- The scattering phase function for the infrared example has a very strong forward peak, whereas the scattering phase function in the microwave example is very smooth. Solution methods that discretise the scattering phase function, such as the Discrete Ordinate ITerative (DOIT) method, would need a very fine angular grid to fully characterise the forward peak. A very fine angular grid is computationally expensive. Therefore, methods that do not discretise the scattering phase function, such as
Monte Carlo (MC) methods, are advantageous. Both DOIT and MC are briefly discussed in section 4.3.1.

- The scattering phase function for the infrared exhibits an oscillatory behaviour as a function of the scattering angle. Like the strong forward peak, an accurate characterisation of these oscillations with discrete ordinate methods like DOIT requires a very fine angular grid. Again, MC methods do not suffer from this problem.

Although Figure 4.2 is for illustration purposes only, the conclusions can be generalised for other shapes.

4.2. Sensors

All prior considerations in this chapter consider monochromatic simulations without considering either the spectral response function or the antenna pattern. The observed channel radiance is given by Equation 2.5 (page 10). To simulate Equation 2.5, a number of monochromatic simulations need to be carried out in order to characterise the total channel power. The naive way to carry this out would be to use a full and densely spaced grid for both the frequency and the sensor field of view. However, this might result in very many radiative transfer calculations, which can be computationally very expensive, in particular for cloudy simulations. Fortunately, other methods exist; see, for example, section 4.3.2 for a method, implemented in ARTS, to significantly reduce the number of monochromatic calculations.
4.3. Software

In this thesis, I used two radiative transfer models. One is a version of SHDOMPPDA (Evans 2007), used in the project described in chapter 6. The other, ARTS, is described in this section.

The Atmospheric Radiative Transfer Simulator (ARTS) is a radiative transfer model for the simulation of radiative transfer of terrestrial radiation. ARTS can calculate polarised radiances in up to three spatial dimensions in any geometry above an ellipsoidal planet. It can obtain either (monochromatic) spectral radiances or calculate passive instrument radiances, taking into account the spectral response function and the antenna pattern. An overview of the capabilities of ARTS is given by Eriksson et al. (2011), with more details in the references therein and in the ARTS User Guide (Eriksson et al. 2013b).

4.3.1. Scattering

In a radiative transfer simulation, ARTS treats the clear-sky and the cloudy parts separately. For the cloudy part, two different solvers exist:

- The DOIT method is a polarised three-dimensional discrete ordinate solver for cloudy radiative transfer. It was developed by Emde et al. (2004), primarily for microwave and sub-mm radiances. DOIT relies on a discretisation of the scattering field. This causes a problem for strongly non-linear scattering phase functions as occur for typical ice crystals in the infrared (see Figure 4.2), because it would require a very fine angular grid to fully resolve the scattering phase function and calculations would become very slow. Although methods exist to alleviate this (e.g. Wiscombe 1977), those have not been implemented in ARTS.

- In the context of radiative transfer, the Monte Carlo (MC) approach involves simulating the radiative transfer statistically. A general overview is given by Mayer (2009). The implementation of MC in ARTS is described by Davis, Emde and Harwood (2005). ARTS implements a reverse MC approach, which means photons are “generated” at the sensor and then traced backward. The photon is transmitted through the three-dimensional model atmosphere step by step. At each step, local optical properties determine the chance that the photon was emitted or scattered at this point. If it is scattered, a random direction is drawn according to the scattering phase function. Therefore, the performance of a MC algorithm does not suffer from a strongly peaked scattering phase function. For this reason, ARTS-MC is the tool of choice in this thesis.

4.3.2. Optimisations

ARTS is a physical model in the sense that it solves the radiative transfer equation starting from first principles as illustrated in Figure 4.1. This makes ARTS slow compared to many other models. In particular, many radiative transfer models use parametrisations for the characterisation of clouds. Although parametrisations speed up the calculation considerably,
they do so at the cost of accuracy. ARTS implements a number of optimisations to make simulations faster, without compromising much on accuracy.

A full characterisation of gaseous absorption for a typical channel on an infrared radiometer (Figure 2.3, page 11) requires a much larger number of monochromatic simulations than a full characterisation of a channel radiance on a microwave radiometer (Figure 2.5, page 18). To fully resolve all absorption lines on an infrared radiometer, thousands of monochromatic simulations need to be simulated. Buehler et al. (2010) describe a method to derive a small set of frequencies that are representative for the entire channel, even without resolving all absorption lines. For HIRS, Buehler et al. (2010) show that a relative error in channel radiance of less than $10^{-4}$ can be obtained by simulating spectral radiances for less than twenty frequencies. Buehler et al. (2010) apply this to clear-sky simulations, but Paper III demonstrates that this is also applicable to cloudy simulations.

Other optimisations implemented in ARTS include absorption lookup tables (Buehler, Eriksson and Lemke 2011) and a sensor response matrix (Eriksson et al. 2006).
Chapter 5.

Retrieval development

5.1. The inverse problem

In the context of (atmospheric) remote sensing, a retrieval is the process of obtaining a geophysical quantity of interest from a calibrated measurement. For sensing the atmosphere, measurements can be reflectances, radiances, reflectances (in case of an active instrument), or other quantities (see chapter 2). A data set of calibrated measurements is also known as a ‘level 1’ data set or as a Sensor Data Record (SDR). The steps that need to be taken in going from the measurement to the geophysical quantity depend on what is being measured and what quantity is sought.

A radiative transfer model, such as described in chapter 4, can calculate the calibrated measurement $y$ as a function of the geophysical state $x$ (Eriksson et al. 2013a),

$$y = F(x, b) + \varepsilon(x, b),$$

where $y$ is a vector representing the calibrated measurement, $F$ is the forward model, $x$ is the vector containing the geophysical quantities to be retrieved, $b$ is the vector contains everything else affecting the measurement (geophysical quantities, instrument properties, spectroscopic properties, etc.), and $\varepsilon$ is the measurement error. The forward model is by definition the model that calculates sensor observations from the quantities in Equation 5.1, and in the present context, the radiative transfer model is the forward model. The task of a retrieval, then, is to find a value for $x$ for a given $y$ for which Equation 5.1 holds or holds as well as possible. The problem is also known as the inverse problem, and described in detail by Rodgers (2000). Depending on the field, the difficulty of this problem ranges from trivial to impossible.

5.2. Optimal estimation

There exist several different approaches for atmospheric retrievals. More often than not, there is no unique solution to Equation 5.1, and a framework is needed to choose the ‘best’ solution according to a sensible quantification of ‘best’. A classical approach is Bayesian Optimal Estimation in the framework treated in detail by Rodgers (2000). Optimal estimation is a mathematical approach, and can be taken either analytically or numerically. A derivation is beyond the scope of this thesis, but the so-called maximum likelihood solution takes the set of measurements with associated errors (error covariance matrix) and the set of a priori information with associated errors (a priori covariance matrix), in order to choose a geophysical state consistent with both. Each retrieval involves minimising a
cost function, which requires iteratively evaluating the forward model. When evaluating
the forward model is computationally expensive, retrieving large numbers of measurements
is not feasible. Moreover, optimal estimation assumes Gaussian statistics, which limits its
applicability for retrieving highly non-Gaussian quantities such as IWP (although retrieving
in log-space alleviates this problem somewhat).

5.3. Lookup tables / retrieval databases

An alternative to optimal estimation outlined above, is to retrieve using a lookup table or a
retrieval database. Here, the solution approach is not mathematical, but rather statistical
or empirical. Rather than performing optimal estimation for every single retrieval, a lookup
table relates the desired quantity \( x \) to measurements \( y \), taking into account other
factors \( b \) either explicitly or implicitly. In the case of IWP retrievals, such a table may
either contain IWC or IWP directly, or it may contain quantities from which IWP can be
calculated through an explicit parameterisation, such as the visible optical depth \( \tau \) or the
effective radius \( r \).

5.3.1. Sources

A lookup table is often built by performing forward model simulations on a set of predefined
physical states. For example, the MODIS retrieval algorithm is based on a lookup table
which uses microphysical properties based on in-situ measurements (Baum et al. 2005a),
and reflectances calculated with a broadband radiative transfer model (Baum et al. 2005b),
relating \( \tau \) and \( r \) to reflectances in six solar channels (King et al. 2003). An explicit
parametrisation then relates IWP to \( \tau \) and \( r \). Evans et al. (2012) stochastically generate
profiles of humidity, temperature, IWC, and two parameters indicating particle sizes, then
use a radiative transfer model to simulate radiances and radar reflectivity at microwave
and sub-millimetre frequencies.

A different source of a retrieval database is explored in this thesis. Paper I introduces
a retrieval database based on collocations (see section 2.4) between passive, operational
instruments and IWP from CloudSat CPR. This is further developed in Paper V, where
this retrieval database is used to develop a new IWP product.

5.3.2. Applications — from simple lookups to multi layer perceptrons

A lookup table or retrieval database can be used in several ways to perform retrievals. The
simplest way is a direct lookup with some form of interpolation. If the database has many
dimensions and is large, this becomes memory-intensive and computationally expensive,
and other methods may be preferred. Bayesian Monte Carlo Integration (BMCI) (Evans
et al. 2002) uses a smaller retrieval database where the entries are randomly spaced (hence
the ‘Monte Carlo’ aspect) and calculates \( x \) and its error based on entries in the vicinity of
\( y \). However, the number of entries needed in the retrieval database still increases rapidly
with the size of \( y \) (i.e. with the number of measurements per entry).

The methods outlined above require to relate every single measurement to the retrieval
database. Inevitably, performing retrievals with a larger retrieval database will require
more computer resources. An alternative method is to use the retrieval database to obtain a parametrisation. A parameterisation is a simplified mathematical model — an equation or a set of equations — relating the desired quantity \( x \) to measurements \( y \), taking into account other factors \( b \) either explicitly or implicitly. A major advantage is that, regardless of the process of obtaining the parameterisation, the actual retrieval requires only the evaluation of a handful of equations. This makes the retrieval faster than any other. One method to obtain a parameterisation from a retrieval database is by training an Artificial Neural Network (ANN).

An ANN is an interconnected assembly of processing units known as neurons. A particular type of ANN is the Multi Layer Perceptron (MLP), which can be used for fitting a complex function, for example, for geophysical retrievals (Krasnopolsky 2007). In an MLP, the nodes in an ANN are divided into layers, where each layer is fully connected to the next layer, i.e. all nodes in layer \( n \) are connected to all nodes in layer \( n + 1 \). Each connection has a nonlinear associated function with some parameters, for example, a weight \( w \) and a bias \( b \). The simplest MLP consists of three layers. When used for a geophysical retrieval, those are the input layer \( y \), the hidden layer, and the output layer \( x \). The input layer has one node for each input quantity (such as channel radiance), while the output layer has one node for each quantity to be retrieved (such as IWP). The hidden layer can have any number of nodes.

Before an ANN can be used for retrievals, each connection needs to be assigned values for the weight \( w \) and the bias \( b \). Weights and biases are assigned in an iterative process known as training, which involves minimising a cost function. In training, the neural network is shown a database of inputs and targets. The training will cause the ANN to learn the relation between \( y \) and \( x \), although there is a risk that the network learns not only the relation between \( y \) and \( x \), but also of the noise. There are several approaches to prevent this. The approach used in this thesis is called the early stopping criterion. In this approach, the data are divided in training, validation, and testing data. Training ends if the cost function is still being reduced with the training data, but no longer with the validation data. Then, the testing data can then be used to identify the quality of the retrieval.

In this thesis, ANNs are used in Paper I, in Paper V, and in the study presented in chapter 6. In particular MLPs for geophysical retrievals are further discussed by Krasnopolsky (2007).
Part II.

New research
Chapter 6.

Studying a new instrument concept

Chapter 2 introduced various techniques used in atmospheric remote sensing, while chapter 5 described how these techniques can be used for retrievals. The usage of these techniques is well consolidated. This chapter introduces sub-millimetre radiometry and its application to remote sensing of ice clouds. Although not nearly as well-established as microwave or infrared, sub-millimetre atmospheric remote sensing is a promising technology that has been demonstrated in simulations and through the use of aircraft prototypes. In the context of this chapter, sub-millimetre covers the frequency range from 183 GHz (1.64 mm) to 2 THz (0.15 mm).

In this chapter, section 6.1 describes the background and history of sub-millimetre radiometry for remote sensing of ice clouds. Then, section 6.2 describes a study of an instrument concept, and compares it to a previous instrument concept study. The results presented in section 6.2 are not yet published elsewhere.

6.1. Sub-millimetre radiometry

The European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) operates meteorological satellites in polar sun-synchronous orbits in the EUMETSAT Polar System (EPS). From the year 2021 onward, a new generation of satellites will be launched as part of the EPS Second Generation (SG) series. Three of those satellites, called MetOp-SG-B1, -B2, and -B3, will carry (among other instruments) the Ice Cloud Imager (ICI), a sub-millimetre radiometer with channels up to 664.0 ± 4.2 GHz.

EUMETSAT’s decision to include ICI on the MetOp-SG-B satellites is the latest news in the history of proposed sub-millimetre radiometers. The first study to investigate frequencies higher than for the 183 GHz channels for spaceborne observations of clouds was by Gasiewski (1992). Since then, there have been many studies and proposals improving the concept (Evans and Stephens 1995; Evans et al. 1998; Kunzi 2001; Ackerman et al. 2005; Buehler et al. 2007, 2012a). Aircraft prototypes exist, such as the Compact Scanning Submillimeter Imaging Radiometer (CoSSIR) (Evans et al. 2005) and the International Sub-Millimetre Airborne Radiometer (ISMAR) (Pritchard et al. 2012), and sub-millimetre radiometry is used both for astronomy and atmospheric limb sounding (Murtagh et al. 2002; Waters et al. 2006). However, no down-looking instrument is currently in space.

Figure 6.1a, from Buehler et al. (2007), shows the measurement principle for sub-millimetre remote sensing of ice clouds. The instrument observes an opaque background of water vapour. Upwelling radiation from this water vapour gets scattered by the presence of atmospheric ice. This causes a reduction in the observed radiances, as illustrated by
Chapter 6. Studying a new instrument concept

Figure 6.1. — Physical principle of ice cloud measurements at sub-millimetre frequencies. Panel (a) shows an illustration of the viewing geometry and measurement principle for a downlooking sub-millimetre radiometer. Panel (b) shows simulated radiances for a synthetic ice cloud with varying IWP for various sub-millimetre frequencies. Both figures are from Buehler et al. (2007).

Figure 6.2. — Panel (a) shows the brightness temperature difference between a clear-sky and a cloudy simulation, where a constant amount of cloud ice is distributed in particles of varying sizes and for different frequencies. Panel (b) shows a selection of size distributions. Both figures are from Buehler et al. (2007).
6.2. Instrument study

IceMusic is a new instrument concept for a sub-millimetre radiometer based on a different technology compared to other sub-millimetre radiometers. The proposed concept uses cryogenic lumped element kinetic inductance detectors (Doyle et al. 2008), which could potentially allow for thousands of detectors and achieve a noise level down to 9 mK (Pete Hargrave, personal communication). It is proposed to have single-sideband channels with a width of at least 1% of the centre frequency. This section presents a retrieval simulation study for the IceMusic concept.

The primary aim of the study is to select IceMusic channels and to study the performance of the proposed selection for retrieving IWP. A secondary aim is to set up a software system that makes it easy to study the performance of new instrument concepts. Such a system would consist of several components, such as a method to generate a retrieval database (including profile sources and a radiative transfer model) and a method to use this database for developing a retrieval algorithm or for performing retrievals. Note that this section does not present a complete study, but is intended to be inserted into a paper describing the IceMusic instrument more completely. This text in this thesis considers only the retrieval performance.

6.2.1. Channel selection

Choosing the channels for a new instrument concept involves selecting channel locations, widths, and shapes, in order to optimize the instruments’ use for scientific and/or operational objectives. In the case of IceMusic, the primary objective is the quantification of the total column ice mass (IWP). For optimally probing particles of different sizes, it is desirable to have channels at varying wavelengths (see Figure 6.1 and Figure 6.2). Additionally, by placing channels such that opacity varies, they receive their signals from different parts of the atmosphere, allowing for some vertical resolution (although the vertical resolution can never be as good as for active instruments). This can be quantified by the Jacobian as described in section 2.2.2 ‘Sounding’.

Figure 6.3 shows a clear-sky spectrum for a midlatitude winter scenario as defined by Anderson et al. (1986), simulated with ARTS as seen looking straight down from a point above the atmosphere. The simulation uses a blackbody surface and spectroscopic properties of absorbing gases from the HITRAN 2004 edition (Rothman et al. 2005). Even though the primary objective for IceMusic is to measure IWP, the channel selection takes place using clear-sky radiative transfer studies, because single scattering properties
Chapter 6. Studying a new instrument concept

Figure 6.3. — Microwave and sub-millimetre spectrum with IceMusic channel positions. Clear-sky spectrum calculated with ARTS. The blue line shows monochromatic radiances, while the grey shaded region shows the mean and standard deviation of the brightness temperature for a hypothetical rectangular channel with a width of 1% of the centre frequency. For example, at 100 GHz it shows the radiance for a rectangular channel from 99.5 to 100.5 GHz, whereas at 1 THz it shows the radiance for a rectangular channel from 995 to 1005 GHz. The green rectangles show the selected channels.
vary much more slowly with frequency, than opacity does. The figure also shows channel positions, selected by optimising the Jacobians as described below.

Figure 6.4. — Jacobians for IceMusic channel selection. See also Table 6.1 for a more complete definition of each channel.

Figure 6.4 shows the temperature Jacobians for the IceMusic instrument concept, using the same midlatitude winter scenario as was used for Figure 6.3. The channel selection starts with a suggestion with 12 channels positioned between 89 GHz and 2 THz, where the initial guess is mostly based on prior experience (Stefan Buehler, personal communication). The channel selection is then fine-tuned based on two criteria: to sample the size distribution, and to get vertically resolved information. To meet the first criterion, it is sufficient that the channels have different wavelengths/frequencies. To meet the second criterion, the channels must have different Jacobians (see section 2.2.2). In order to meet both criteria independently, it is desirable to have pairs of channels that have different frequencies, but are sensitive to the same altitude. Because the instrument concept was thought to fly in a midlatitude winter flight campaign, the channel positions were optimised to get pairs
of matching Jacobians for the midlatitude winter scenario in particular. For example, Figure 6.4 shows that the channel at 321.5 GHz has a Jacobian similar to the channel at 186.31 GHz. Therefore, these channels obtain information from the same layers of the atmosphere, but are sensitive to different particle sizes.

<table>
<thead>
<tr>
<th>no.</th>
<th>freq. (GHz)</th>
<th>width (GHz)</th>
<th>peaks (km)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>89</td>
<td>10</td>
<td>surface</td>
</tr>
<tr>
<td>2</td>
<td>150</td>
<td>10</td>
<td>surface</td>
</tr>
<tr>
<td>3</td>
<td>183.31</td>
<td>1.8</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>186.31</td>
<td>1.8</td>
<td>3</td>
</tr>
<tr>
<td>5</td>
<td>190.31</td>
<td>1.8</td>
<td>1.5</td>
</tr>
<tr>
<td>6</td>
<td>205</td>
<td>10</td>
<td>surface</td>
</tr>
<tr>
<td>7</td>
<td>312.5</td>
<td>10</td>
<td>1.5</td>
</tr>
<tr>
<td>8</td>
<td>321.5</td>
<td>3.2</td>
<td>3</td>
</tr>
<tr>
<td>9</td>
<td>505</td>
<td>10</td>
<td>4.5</td>
</tr>
<tr>
<td>10</td>
<td>530</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>11</td>
<td>875</td>
<td>20</td>
<td>4.5</td>
</tr>
<tr>
<td>12</td>
<td>1500</td>
<td>20</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 6.1. — Proposed channels for IceMusic, centre frequencies and channel widths.

The final selection of channels proposed for IceMusic is shown in Table 6.1.

6.2.2. Performance study

After the channel selection, instrument performance for cloudy retrievals is studied. In the present study, this builds on the work of Evans et al. (2012). Evans et al. (2012) present a sophisticated method to generate artificial one-dimensional cloudy atmospheric states, using microphysical assumptions consistent with in-situ measurements from the Tropical Composition, Cloud and Climate Coupling (TC4) experiment (Toon et al. 2010) and vertical correlations consistent with CloudSat reflectivities. By applying radiative transfer simulations (Evans 2007) to these states, Evans et al. (2012) construct a database that can be used for retrieval studies (see also section 5.3).

Evans et al. (2012) use this database with Bayesian Monte Carlo Integration (BMCI) and a method known as Markov-Chain Monte Carlo, or optimal estimation when these fail. The IceMusic study uses the retrieval database from Evans et al. (2012), but with an altered selection of channels, and with a neural network based retrieval.

For the present study, Frank Evans provided the code used in Evans et al. (2012), as well as code for DDA calculations (following Evans 2007) to calculate single scattering properties (see section 4.1.2) and the results of such calculations for the channels he used for simulating the CoSSIR instrument. IceMusic has a different channel selection, which necessitated additional DDA calculations. The model considers four shapes. In the present study, the size ranges considered for the different shapes were: plate aggregates (6.31 to 355 µm), spherical aggregates (5.01 to 497 µm), snow aggregates (63.1 to 891 µm), and
6.2. Instrument study

hail (398 to 3162 µm). In the original system by Evans et al. (2012), spherical and snow aggregates were considered up to 1584 µm, and plate aggregated up to 398 µm. However, as DDA calculations get more time-consuming for larger particles and shorter wavelengths, in this study, the largest particles were not considered, so that calculation times became manageable. This may be permissible, because at high frequencies, a sensor will not look far enough down into the atmosphere to observe the region where those large particles reside (see Figure 6.4). This hypothesis still needs further testing. For details on the particle shapes and the DDA calculates, refer to Evans (2007); Evans et al. (2012).

With the Evans et al. (2012) system, a set of $1 \times 10^6$ cases (profiles with associated measurements) were generated. Following the early stopping criterion (described in section 5.3.2), profiles were divided in 70% training, 15% validation, and 15% testing. Gaussian random noise with a standard deviation of 0.1 K was added. After the training was finished, where validation data were used to prevent overfitting, testing data were used to assess the median relative retrieval error.

![Comparison IceMusic/Ciwsir](image)

**Figure 6.5.** — Retrieval simulations based on the database by Evans et al. (2012) (panel (a)) or as performed by Jiménez et al. (2007) (panel (b)). Note that errors are much lower in the Jiménez et al. (2007) study.

Figure 6.5a shows an estimate of the retrieval error for an IWP retrieval, using either the IceMusic channel configuration, or a channel configuration identical to Cloud Ice Water path Submillimetre Imaging Radiometer (CIWSIR) as studied by Jiménez et al. (2007). For comparison, Figure 6.5b shows performance that Jiménez et al. (2007) found using similar methodology, but a less sophisticated retrieval database.

Regardless of the retrieval approach, any error estimates are determined by the complexity of the forward model or the variability contained in the retrieval database. It is possible to
construct a trivial retrieval database, in which all IWP values can be retrieved with 0 error with a very simple instrument, simply by constructing a unique and perfectly invertible $y = F(x)$ mapping. However, such a database would have no connection to the real world, and such a retrieval algorithm would be useless. The reliability of an error estimate, then, is determined by the representativeness of forward model or retrieval database. In the present example, the retrieval database by Evans et al. (2012) has a higher variability than the one used by Jiménez et al. (2007). Therefore, the artificial retrieval problem is more difficult, which may explain the larger error. Hence, one should be careful in interpreting error estimates from retrieval studies, and particularly careful in comparing error estimates between different studies. This hypothesis needs further testing.

To obtain a real error estimate, then, requires an independent source of true (IWP). In particular for an integrated quantity like an IWP measurements, in-situ measurements are hard to obtain. Some studies, such as Deng et al. (2013), attempt to compare remote and in-situ measurements for IWC — but even before vertical integration, this is challenging. Certainly, retrieval error characterisation remains a challenge.
Chapter 7.

Summary of appended articles

The bulk of the new research presented in this thesis is presented in five papers, that are appended verbatim. All of the papers have multiple authors, so none is fully my work. This chapter presents summaries of the appended papers, and briefly characterises what part of each paper was done by me.

7.1. Paper I


Included at page 73.

Paper I describes collocations between CloudSat Cloud Profiling Radar and NOAA-18 MHS. The paper is based on the work I did for my Master Thesis (Holl 2009). It consists of three parts.

Firstly, the paper describes an algorithm to find collocations. This algorithm is optimised for collocations between two instruments where at least one instrument observes at a fixed viewing angle (typically the nadir), so that observations lie on a line. All footprints are treated as point measurements, and collocations depend only on distance in space and time between different footprints. Collocation criteria are split in a spatial and temporal part. If there is any temporal overlap (± the time criterion) between two considered granules (a granule being one file containing satellite data, typically one orbit), the spatial criterion is tested.

Secondly, Paper I applies the algorithm to find collocations between CloudSat CPR and the suite of AMSU-A, AMSU-B, MHS, and HIRS on the NOAA and MetOp operational meteorological satellites. The paper shows that only NOAA-18 has global collocations with CloudSat CPR\(^1\). It presents patterns in MHS viewing angles and latitudes at which collocations occur, and shows that those collocations do not occur all the time, but intermittently. Since MHS footprints are much larger than CPR footprints, the paper presents two datasets: one dataset where each collocation matches one MHS measurement to one CPR measurement (so that several collocations share the same MHS measurement, but have distinct CPR measurements), and one dataset where each collocation matches

\(^{1}\)NOAA-19 was not launched at the time this investigation was performed.
one MHS to a collection of CPR footprints, presenting statistics (mean, standard deviation, etc.) of the latter. Those collocations are available for public use.

Finally, Paper I explores several applications for the datasets derived in the second part. Three applications are presented. The first application is a direct comparison between MHS-averaged CPR IWP and an IWP product from NOAA National Environment Satellite, Data and Information Service (NESDIS) MSPPS. This comparison shows that MSPPS IWP is much smaller than CPR IWP, even for clouds that are detected at MHS frequencies. This application is explored in more detail in Paper IV.

The second application looks at statistics between brightness temperatures and IWP. The statistics between the two are compared between a dataset of artificial states with simulated brightness temperatures developed by Rydberg et al. (2009) and a dataset based on collocations. Although small differences exist between the two datasets, the features are largely the same.

For the third application, the collocated dataset is used as a training database for the development of an IWP retrieval. A subset of nadir-looking, tropical, MHS-averaged, cloudy collocations is selected. This subset is then used to train an ANN, mapping MHS channels 3–5 against CPR IWP. The resulting network can then be used to retrieve IWP from MHS measurements. Paper I then presents the performance of this retrieval and studies how it is affected by adding HIRS channels. Since clouds with IWP smaller than 100 gm$^{-2}$ are invisible to passive microwave radiation, addition of infrared HIRS should significantly increase the retrieval. However, the effect is small, likely because of the poor horizontal sampling by HIRS. A more sophisticated retrieval based on the same principle is derived in Paper V.

In conclusion, Paper I introduces a method to obtain collocations and shows that those are a highly useful tool with a large variety of applications.

I did almost all work for Paper I, except the description of the dataset developed by Rydberg et al. (2009).

### 7.2. Paper II


Included at page 93.

**Paper II** uses a newly developed collocated dataset to globally compare nadir observations from different copies of the MHS and AMSU-B sensors on sun-synchronous satellites, in order to investigate how inter-instrument biases vary as a function of radiance and latitude. Here, a bias is a systematic (as opposed to random) difference between the measurements of different copies of the same instrument. In the ideal case, the bias is 0. Traditionally, biases are derived from polar Simultaneous Nadir Overpasses (SNOs), but this may not be globally representative.
**Paper II** exploits the drift in LTAN (see section 2.2.4) for operational satellites from NOAA and MetOp. This drift, illustrated in Figure 3 in **Paper II**, causes some pairs of satellites to temporarily have the same LTAN. During these periods, collocations between such a pair of satellites occur globally.

The algorithm I developed for **Paper I** is not optimal for the purposes in **Paper II**, because it is optimised for the situation where at most one satellite is scanning and includes off-nadir observations. In **Paper I**, one of the considered instruments has observations only along the ground track, but for **Paper II**, both are scanning instruments. Therefore, I improved the collocation algorithm as described in **Paper II**.

**Paper II** considers only near-nadir collocations, here referred to as SNOs. It presents collocations between NOAA-16 and NOAA-15 during August 2008, between MetOp-A and NOAA-17 during April and May 2009, and between NOAA-19 and NOAA-18 during September 2009. The collocations are then used to explore the latitudinal dependence of inter-satellite biases. The paper shows that the latitudinal dependence is significant, and that an inter-calibration of humidity sensors needs to take this into account.

My contribution to **Paper II** was the development and application of this algorithm, and the description in the appendix of **Paper II**. I developed the algorithm not exclusively for the purposes of **Paper II**, but also for other research.

### 7.3. Paper III


**Paper III** shifts the focus to radiative transfer. The paper builds on the work by Buehler et al. (2010). Buehler et al. (2010) present a method to derive an optimised frequency grid for the simulation of radiances observed by broadband sensors. Buehler et al. (2010) apply this to the infrared radiometer HIRS and show how the number of frequencies can be reduced from several thousand to less than twenty, but show this only for clear-sky simulations (without clouds or other scatterers). In **Paper III**, we show that a frequency grid derived under clear-sky conditions can be used for simulations of cloudy atmospheres.

In the paper, we select 50 profiles from a dataset by Chevallier, Di Michele and McNally (2006). For all profiles, we use the ARTS-Monte Carlo (MC) model to perform simulations for HIRS channel 11 (7.33 μm) with both the reference grid, and the optimised grid from Buehler et al. (2010). We perform each simulation ten times, because the Monte Carlo method is stochastic, and the difference between subsequent simulations can be used to calculate an error estimate. We also apply the method from Buehler et al. (2010) to

---

2 Even though **Paper II** uses only (near)-nadir collocations, I wanted to optimise the collocation algorithm in any case.
Chapter 7. Summary of appended articles

AVHRR channel 5 (10.8 µm), and perform the same comparison as for HIRS channel 11. In both cases, we show that the magnitude of the bias for cloudy simulations is less than 0.04 K and the magnitude of the root mean square error is less than 0.4 K.

We also investigate the tradeoff between runtime and MC error for different numbers of photons per monochromatic pencil-beam, for both the reference grid and the fast grid. For AVHRR-5, the reference grid contains 5461 frequencies, while the fast grid contains only 5 frequencies. A reference grid simulation with 1 photon per frequency (5461 for the entire channel) has roughly the same accuracy as an optimised grid simulation with 1000 photons per frequency (5000 for the entire channel). However, the simulation with the optimised grid runs much faster, likely due to reduced overhead associated with each monochromatic pencil beam simulation.

For Paper III, I have done almost all work and written all text.

7.4. Paper IV


Included at page 122.

Paper IV expands on the first example application presented in Paper I. It presents a systematic comparison between IWP products from different datasets, using the joint radar-lidar retrieval raDAR/liDAR (DARDAR) as a reference. In total, eight datasets — from radar, solar, or passive microwave — are compared to DARDAR.

In the comparison, Paper IV separates systematic and random errors and characterises the IWP range where each dataset shares sensitivity with DARDAR. The sensitivity range of DARDAR probably exceeds that of any considered passive measurement, although that can not be proven with the results in Paper IV. If we do make this assumption, however, we can find the sensitivity range for each passive dataset.

Paper IV finds that datasets from solar sources have the largest IWP sensitivity range, and that MODIS has the lowest random error with DARDAR (this is further explored by Eliasson et al. (submitted 2013)). The paper also compares a new IWP product based on a retrieval algorithm first presented in Paper I and further explored in Paper V.

For Paper IV, I have developed the underlying collocations, contributed a new IWP product, and have been highly involved in the analysis, having numerous discussions on science and data analysis with first author Salomon Eliasson. Additionally, I contributed by getting into contact with authors of some of the involved datasets, thereby expanding the volume of the comparisons.

7.5. Paper V

56
7.5. **Paper V**


Included at page 139.

The final paper included in the PhD thesis, **Paper V**, describes a retrieval algorithm for a synergistic IWP dataset, explores synergies between different passive techniques, and presents the SPARE-ICE IWP dataset.

The paper starts with collocating NOAA-18 MHS and AVHRR with CloudSat CPR. IWP from the CloudSat 2C-ICE product, which is based on combined radar and lidar measurements, is averaged over the MHS footprint, as is AVHRR. This results in a dataset matching solar reflectances, terrestrial infrared radiances, passive microwave radiances, and auxiliary information, against 2C-ICE IWP. This dataset is used as a retrieval database (see section 5.3) to train artificial neural networks for an IWP retrieval (see section 5.3.2).

**Paper V** describes a systematic exploration of synergies between an IWP retrieval based on any combination of solar, terrestrial infrared, and passive microwave. Combinations of two techniques mostly outperform any individual retrieval, and a combination of three techniques in turn performs better than any combination of two. Finally, a set of inputs is decided for the use in the final product, that is called SPARE-ICE.

The last part of the paper explores SPARE-ICE with MODIS and MSPPS for a set of case studies and a 2007 gridded mean. The comparisons show that SPARE-ICE appears to perform well, even in difficult conditions. SPARE-ICE is available for public use.

I did almost all the work for this paper.
Chapter 8.

Conclusion

This thesis consists of four published articles, one submitted article, one work in progress, and an introduction to the field.

A large part of the work builds on the collocation toolkit, described in section 2.4. Collocations between passive, operational sensors and active, scientific sensors, are highly useful for a variety of applications. Some of those have been shown in this thesis (Paper I, Paper II, Paper IV, and Paper V), while others have been published as side projects (John et al. 2011; Buehler et al. 2012b; John et al. 2013; Eliasson et al. submitted 2013). In particular, the proximity of NOAA-18 to the A-Train is very beneficial for product development from the former, and it would be good if future scientific satellites are also placed in orbits close to operational ones, so that the operational ones can benefit.

Other parts of the PhD work were based on radiative transfer modelling, a more classical tool for retrieval development. Paper III shows that a clear-sky derived optimised frequency grid can be used in cloudy simulations. This will be useful in future work. Finally, chapter 6 describes a project involving retrieval simulations for a new sub-millimetre instrument concept.

Science is never finished, and many of the projects presented call for further work. The work presented in Paper II can relatively easily be extended to other sensors, such as for SAPHIR on the French-Indian Megha-Tropiques satellite, or MWHS on the Chinese Fèngyú”n-3 series. Comparisons such as presented in Paper I and Paper IV can be expanded with more IWP products, extended to other products, or be made more sophisticated or more focussed along the lines of Eliasson et al. (submitted 2013). Certainly, the work presented in Paper V could be extended or improved upon in several ways, that are mentioned in the outlook of the manuscript.

The study presented in chapter 6 is not really finished. A very interesting aspect to look at would be to explore synergies between sub-millimetre and active radar. This may be a topic of future study.

Overall, this thesis should present a modest but real contribution to the problem of spaceborne remote sensing of atmospheric ice.
Bibliography


Bibliography


Bibliography


Bibliography


Bibliography

United Kingdom and New York, NY, USA: Cambridge University Press, pp. 79–131 (cit. on p. 3).


Acronyms

AMSU Advanced Microwave Sounding Unit. 6, 12, 13, 17, 18, 21, 22, 53, 54
ANN Artificial Neural Network. 41, 54
AR4 4th Assessment Report. 3
ARTS Atmospheric Radiative Transfer Simulator. 5, 6, 8, 11, 26, 32, 35–37, 47, 48, 55
ATMS Advanced Technology Microwave Sounder. 17
AVHRR Advanced Very-High Resolution Radiometer. 11, 12, 15, 16, 20, 22, 56, 57
BMCI Bayesian Monte Carlo Integration. 40, 50

CALIOP Cloud-Aerosol Lidar with Orthogonal Polarization. 19
CALIPSO Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation. 19, 20, 22
CIWSIR Cloud Ice Water path Submillimetre Imaging Radiometer. 51
CoSSIR Compact Scanning Submillimeter Imaging Radiometer. 45, 50
CPR Cloud Profiling Radar. 4, 19, 21, 22, 40, 53, 54, 57
CTH Cloud Top Height. 27
CTP Cloud Top Pressure. 27
CTT Cloud Top Temperature. 27

DARDAR raDAR/liDAR. 56
DDA Discrete Dipole Approximation. 34, 50, 51
DMSP Defense Meteorological Satellite Program. 17
DOIT Discrete Ordinate ITerative. 34–36

ECV Essential Climate Variable. 3, 23
EPS EUMETSAT Polar System. 45
ESMR Electrically Scanning Microwave Radiometer. 16
EUMETSAT European Organisation for the Exploitation of Meteorological Satellites. 45, 71

FDTD Finite Difference Time Domain. 34
FRAC Full Resolution Area Coverage. 16

GAC Global Area Coverage. 16
GCOS Global Climate Observing System. 3, 23
GRAS GNSS Receiver for Atmospheric Sounding. 22

HIRS High resolution Infrared Radiation Sounder. 6, 13, 15, 16, 20–22, 37, 53–56
HITRAN HHigh resolution TRANsmision. 32, 47
Acronyms

**ICI** Ice Cloud Imager. 45  
**IPCC** Intergovernmental Panel on Climate Change. 3  
**IR** Infra-Red. 8  
**ISMAR** International Sub-Millimetre Airborne Radiometer. 45  
**IWC** Ice Water Content. 23, 26, 40, 52  
**IWP** Ice Water Path. 4–6, 21, 22, 26, 40, 41, 46, 47, 51, 52, 54, 56, 57, 59  
**LTAN** Local Time Ascending Node. 14, 19, 22, 55  
**LTE** Local Thermodynamic Equilibrium. 29, 30  
**MC** Monte Carlo. 35, 36, 55, 56  
**MHS** Microwave Humidity Sounder. 4, 6, 17, 19–22, 53, 54, 57  
**MLP** Multi Layer Perceptron. 41  
**MODIS** MODe rate resolution Imaging Spectroradiometer. 22, 40, 56, 57  
**MRIR** Medium Resolution Infrared Radiometer. 15  
**MSPPS** Microwave Surface and Precipitation Products System. 21, 54, 57  
**MSU** Microwave Sounding Unit. 17  
**MWHS** Micro-Wave Humidity Sounder. 17, 22, 59  
**NESDIS** National Environment Satellite, Data and Information Service. 54  
**NOAA** National Oceanic and Atmospheric Administration. 4, 11, 15–17, 20, 53–55, 57, 59  
**NPP** National Partnership Program. 16, 17  
**OSCAR** Observing Systems Capability Analysis and Review. 17  
**PWV** Precipitable Water Vapour. 18  
**SAAC** Simultaneous All Angle Collocation. 6  
**SAPHIR** Sonde Atmospherique du Profil d’Humidité Intertropicale par Radiometrie. 17, 22, 59  
**SDR** Sensor Data Record. 39  
**SEVIRI** Spinning Enhanced Visible Infra-Red Imager. 12  
**SNO** Simultaneous Nadir Overpass. 6, 22, 54, 55  
**SRF** spectral response function. 10–12  
**SSM/I** Special Sensor Microwave/Imager. 17  
**SSM/T** Special Sensor Microwave/Temperature. 17, 22  
**TC4** Tropical Composition, Cloud and Climate Coupling. 50  
**TIROS** Television and InfraRed Observation Satellite. 7, 15–17  
**TOA** Top Of Atmosphere. 8  
**TRMM** Tropical Rainfall Measuring Mission. 19  
**UTH** Upper Tropospheric Humidity. 6, 21  
**UV** Ultra-Violet. 15, 29  
**VHRR** Very High Resolution Radiometer. 15
**VIIRS** Visible Infrared Imager Radiometer Suite. 12, 16

**WMO** World Meteorological Organisation. 3, 23
Collocating satellite-based radar and radiometer measurements – methodology and usage examples

Collocating satellite-based radar and radiometer measurements – methodology and usage examples

G. Holl¹, S. A. Buehler¹, B. Rydberg², and C. Jiménez³

¹Department of Space Science, Luleå University of Technology, Kiruna, Sweden
²Department of Radio and Space Science, Chalmers University of Technology, Göteborg, Sweden
³Laboratoire d’Etude du Rayonnement et de la Matière en Astrophysique, Centre National de la Recherche Scientifique, Observatoire de Paris, Paris, France

Abstract. Collocations between two satellite sensors are occasions where both sensors observe the same place at roughly the same time. We study collocations between the Microwave Humidity Sounder (MHS) on-board NOAA-18 and the Cloud Profiling Radar (CPR) on-board CloudSat. First, a simple method is presented to obtain those collocations and this method is compared with a more complicated approach found in literature. We present the statistical properties of the collocations, with particular attention to the effects of the differences in footprint size. For 2007, we find approximately two and a half million MHS measurements with CPR pixels close to their centrepoints. Most of those collocations contain at least ten CloudSat pixels and image relatively homogeneous scenes. In the second part, we present three possible applications for the collocations. Firstly, we use the collocations to validate an operational Ice Water Path (IWP) product from MHS measurements, produced by the National Environment Satellite, Data and Information System (NESDIS) in the Microwave Surface and Precipitation Products System (MSPPS). IWP values from the CloudSat CPR are found to be significantly larger than those from the MSPPS. Secondly, we compare the relation between IWP and MHS channel 5 (190.311 GHz) brightness temperature for two datasets: the collocated dataset, and an artificial dataset. We find a larger variability in the collocated dataset. Finally, we use the collocations to train an Artificial Neural Network and describe how we can use it to develop a new MHS-based IWP product. We also study the effect of adding measurements from the High Resolution Infrared Radiation Sounder (HIRS), channels 8 (11.11 μm) and 11 (8.33 μm). This shows a small improvement in the retrieval quality. The collocations described in the article are available for public use.

1 Introduction

Atmospheric remote sensing from satellites is a major source of data for the atmospheric sciences and for operational weather forecasting (Kidd et al., 2009). Measurements from Earth observation satellites have a global or near-global coverage. However, the accuracy of products derived from such measurements is often poor (Wielicki et al., 1995; Wu et al., 2009). A combination of observations from different instruments enables applications that are impossible with single-instrument measurements. One way to combine measurements is through collocations. A collocation is an event where different (satellite) sensors observe the same location at roughly the same time. The collocations considered here are mainly between active measurements from the Cloud Profiling Radar on-board CloudSat, and passive measurements from microwave and infrared sensors on-board National Oceanic and Atmospheric Administration (NOAA)-18.

One product obtained by remote sensing measurements is the Ice Water Path (IWP), the vertically integrated Ice Water Content (IWC) or the column mass density of ice in the atmosphere. Ice clouds are important for the climate, because they absorb and scatter thermal radiation and reflect solar radiation, and thus influence the radiation budget of the Earth (Stephens, 2005). As shown by John and Soden (2006), different General Circulation Models (GCMs) disagree by an order of magnitude about the climatology of IWP. Also IWP values from remote sensing measurements differ considerably (Wu et al., 2009). Therefore, it is important to improve the quality of ice cloud retrievals. A good understanding of the cloud signal in microwave radiometer measurements is an important step in the development of retrieval algorithms for possible future missions, such as the Cloud Ice Water Submillimetre Imaging Radiometer (CIWSIR), proposed by Buehler et al. (2007).
Collocations between sensors on the same platform are commonly used (for example, see Frey et al., 1996; Bennartz, 2000). The idea to collocate data from different satellite platforms is not new either. Wielicki and Parker (1992) compare the cloud cover obtained with sensors of different spatial resolution. The A-Train constellation was motivated by the advantages of using a combination of measurements (Stephens et al., 2002). Already before CloudSat's launch, Miller et al. (2000) described how to use active sensor data as a priori information for passive sensor retrievals, anticipating “a considerable overlap of CloudSat with the Earth Observing System (EOS) PM and Geostationary Operational Environmental (GOES) satellites”. Several recent studies use the new possibilities from the A-Train (for example, Holz et al., 2008; Kahn et al., 2008). However, not much work has been published on actual collocation methods. The first publication on the subject appears to be a technical note written in Japanese (Aoki, 1980). Judging from the abstract, Aoki (1980) describes how to match AVHRR and HIRS/2 if the instruments are on the same satellite. Other conference papers on the subject are Nagle (1998) and Sun et al. (2006). The first peer-reviewed publication on the subject appears to be Nagle and Holz (2009), discussed in more detail in Sect. 3.1.1.

No literature exists that focusses on collocations between an active instrument such as the Cloud Profiling Radar (CPR) on-board CloudSat and passive, operational instruments on Polar Orbiting Environmental Satellites (POES) such as the MHS on the NOAA-18. However, such collocations have relevant applications. Although a satellite like CloudSat has high quality products, the coverage is small compared to operational satellites, and it will have a limited lifetime. If we can use collocations between CloudSat CPR and NOAA-18 MHS to improve the operational microwave IWP retrieval, the advantages will last much beyond the lifetime of the A-Train satellites and have a much higher spatial coverage. Even passive microwave data from before CloudSat could be reprocessed with an improved algorithm. Whereas Miller et al. (2000) describe a retrieval that requires collocated data for each individual retrieval, we show that collocations can be used to develop new retrievals, that can then be used for non-collocated passive radiometer measurements.

The main purpose here is to study collocations between CloudSat CPR and NOAA-18 MHS. Collocations with MHS and AMSU-B on other POES were also located, but due to the large distances between the satellites, few useful collocations were found. Hence, the study focuses on NOAA-18 MHS. The collocation procedure is described in Sect. 3. The secondary purpose of the study is to look at possible uses of the collocations. Three applications are described in Sect. 4. Firstly, the NOAA National Environmental Satellite, Data and Information Service (NESDIS) Microwave Surface and Precipitation Products System (MSPPS) IWP product is compared with the IWP product from the CPR on-board CloudSat (Sect. 4.1). Simulated radiances from generated clouds are used to study the relation between brightness temperature and IWP, and compare this with the statistics of the collocated dataset (Sect. 4.2). Finally, in Sect. 4.3, we use microwave radiances, with and without infrared measurements, to train an Artificial Neural Network with the CloudSat IWP as a target. Such a network can then be used to develop a new IWP product from microwave (and IR) measurements. Such applications were not found in peer-reviewed literature.

2 Instruments

The Cloud Profiling Radar (CPR) is a radar instrument on-board the sun-synchronous CloudSat satellite (Stephens et al., 2002), launched 28 April 2006. It has an operating frequency of 94 GHz and measures profiles of backscattering ratio at a 0.16° off-nadir angle. CloudSat generates a profile every 1.1 km along-track. A profile footprint is 1.3 km across-track and 1.7 km along-track. A profile is taken every 0.16 s. CloudSat is part of the A-Train constellation. It has an inclination of 98.26° and a Local Time Ascending Node (LTAN) varying between 13:30 and 13:45 local solar time. We use the ROIWP (Radar-Only Ice Water Path) field from the 2B-CWC-RO (level 2b, Cloud Water Content, Radar Only) product, version 008. Austin et al. (2009) describe the algorithm to calculate IWC from radar reflectivity profiles. They report an upper limit of the uncertainty of 40%. However, throughout this article, we assume CloudSat CPR to represent the truth since it is supposed to provide the most accurate measurements of IWP. The data originate from the CloudSat Data Processing Center. All measurements are geolocated and time-associated.

The Advanced Microwave Sounding Unit-B (AMSU-B) and its successor the Microwave Humidity Sounder (MHS) are microwave radiometers (Saunders et al., 1995; Kleespies and Watts, 2007). MHS channels 3–5 correspond to AMSU-B channels 18–20. We use the MHS channel numbers. Channel 3 has a centre frequency of 183.31±1.00 GHz with a bandwidth of 500 MHz, channel 4 has a centre frequency of 183.31±3.00 GHz with a bandwidth of 1000 MHz, and channel 5 has a centre frequency of 183.31±7.00 GHz (AMSU-B) or 190.31 GHz (MHS) with a bandwidth of 2000 MHz (AMSU-B) or 2200 MHz (MHS). We use channels 3–5 because of the prominent water vapour spectral line at 183.31 GHz. In this article, we neglect the differences between AMSU-B and MHS. Although they are not the same, the standard deviation of the difference is much larger than the mean difference, so that a simple correction is not possible (Kleespies and Watts, 2007). Because of its proximity to CloudSat, we focus on NOAA-18 and MHS for the collocations. However, we have also looked for collocations with MetOp-A (a satellite operated by the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT)), NOAA-15, NOAA-16 and NOAA-17, so with a total of five satellites. The MHS field of view is


www.atmos-meas-tech.net/3/693/2010/
around 1.1°, and the footprint size at nadir is around 15 km in diameter. It scans across-track in angles from $-49.44°$ to $49.44°$ with 90 measurements per scan line. A scan takes 8/3 s. MHS is currently present on NOAA-18, NOAA-19 and MetOp-A, whereas AMSU-B is present on NOAA-15 through NOAA-17. All those satellites are sun-synchronous satellites. NOAA-18 has an inclination of 98.74° and a LTAN of 13:39\textsuperscript{1}. This is close to CloudSat, which leads to a large number of collocations, as described later in the article.

MHS measures the antenna temperature, which can be calibrated to obtain a brightness temperature in units of Kelvin. We use the ATOVS and AVHRR Pre-processing Package (AAPP) software package to apply this calibration, described by Labrot et al. (2006) (ATOVS stands for Advanced TIROS Operational Vertical Sounder, where TIROS stands for Television InfraRed Observation Satellite). We obtain the radiances from the NOAA CLASS archive.

All those satellites also carry the infrared radiometer High Resolution Infrared Radiation Sounder (HIRS), either HIRS/3 or HIRS/4. HIRS measures in 20 channels, one visible and nineteen infrared. We use channels 8 ($\lambda = 11.1$ µm, a window channel) and 11 ($\lambda = 7.33$ µm, a humidity channel) because ice clouds are clearly visible at those wavelengths. HIRS/3 is present on NOAA-15 through NOAA-17 and HIRS/4 is present on NOAA-18, NOAA-19 and MetOp-A. HIRS scans the atmosphere in 56 angles between $-49.5°$ and $49.5°$. Those measurements are not on the same grid as the MHS measurements (see Fig. 1). A HIRS scan takes 6.4 s.

3 Finding collocations

The footprint size of the considered sensors is in the order of kilometres, whereas the measurement duration is in the order of milliseconds. The spatial extent of a measurement is of the same order as the physical extent of a cloud or larger (kilometers), but the time order of a measurement (fraction of a second) is much smaller than a typical cloud lifetime (minutes to hours) (Rogers and Yau, 1979).

Thus, to have a meaningful collocation, the footprints need to have a physical overlap. However, the time between the measurements can be much larger than the duration of a measurement. Hence, a collocation occurs when the sensors observe exactly the same place at approximately the same time.

As shown in Fig. 1, an MHS footprint is an order of magnitude larger than a CPR footprint and HIRS measurements are not on the same grid as MHS measurements.

We create two collocated datasets. In the first dataset, there is an entry for each CloudSat measurement collocating with an MHS measurement, so that there can be many collocations for the same MHS pixel. In the second dataset, each collocation has a unique MHS measurement and CPR pixels are averaged. For each MHS measurement, we note the number of CPR pixels inside the MHS pixel, the average CPR IWP value, the standard deviation of the CPR IWP and the fraction of cloudy CPR pixels. For the averaging, we consider the CPR pixels as point measurements and the MHS pixels as circular measurements with a radius of 7.5 km and a constant sensor spatial response function inside this area. In reality, the sensor spatial response function of MHS is better approximated by a Gaussian shape. Although this might reduce the representativeness, this effect is small compared to other error sources. The total area covered by the CPR pixels is still much smaller than the MHS footprint area. This leads to a sampling error, as discussed in Sect. 3.3 below.

Both datasets are available for public use.

3.1 Collocation finding procedure

The collocation finding procedure consists of four steps. The steps are described in detail in the following text.

1. Orbits (granules) with time overlap are selected.
2. Orbit sections are selected according to a rough temporal criterion.
3. Measurements possibly fulfilling the spatial criterion are selected.
4. The temporal criterion is applied to the selected measurements.

The measurement data as obtained from the data providers is stored as one file for each orbit. Those files, known as granules, contain geolocated, time-referenced measurements. The geolocation refers to the actual measurement; the position of the satellite is not available and not required for the procedure (in contrast to Nagle and Holz (2009) discussed further down). The filenames contain information about the starting and ending time of the data contained by the granule.

For each CPR granule, we locate all NOAA and MetOp granules that have a time overlap with the CPR granule. Those are two granules for each POES for each CPR granule, or a total of ten files for each CPR granule to search for collocations (MetOp-A and NOAA-15 through NOAA-18).

We read the CPR file along with each of the associated POES files. The start and end times of the files are different. The segment with time overlap is selected, plus the segment where the time difference is less than the maximum time interval for a collocation to be considered. For example, if the CPR granule covers 10:00–11:30 UT, and a POES granule covers 11:00–12:30 UT, and our maximum time difference is 15 min, we consider the data in the interval 10:45–11:45, or more precisely 10:45–11:30 in the CPR granule and 11:00–11:45 in the POES granule.

As defined above, a collocation has a spatial and a temporal criterion. We use a two-step approach: first we look for any collocations that might meet the spatial criterion, and then whether those also meet the time criterion.

Starting from the orbit data screened according to the first temporal criterion as explained above, we find the measurements that meet the spatial criterion. In the first step, we do not consider the true pixel size or the sensor spatial response function of either sensor. Instead, we treat the measurements as points and define a maximum distance to select the measurement pair for further consideration. The sensor spatial response function and the effective field of view can be used later to select a subset of those collocations or a weighting of them to consider the MHS spatial response function.

We consider the ground track of each scan angle of the MHS (track A) and compare it to the single scan in the CPR (track B), but the following procedure works as well if both instruments are scanning.

If two ground tracks are plotted, a human observer can see immediately whether there is any spatial overlap or not. Computers can not, so the following algorithm is used to identify points where the spatial overlap condition is met.

1. The distance in km between successive points of the ground track is computed for both ground tracks, considering only the segments screened according to the temporal criterion discussed above. The maximum speed of the ground tracks is assumed to be the maximum distance.

2. Start with $n = 1$.

3. Find close points to $A_n$ in $B$. Here, $A_n$ is the $n$-th measurement in track $A$. Figure 2 shows the distance from a CloudSat CPR pixel to all pixels in a MHS track for a fixed viewing angle. If any collocations exist, they will be close to the global minimum. Find points meeting the distance criterion by the following method.

   (a) Choose $N$ equidistant points (henceforth samples) from $B$ as shown in Fig. 2. Combined with the first and the last point of the track, the samples define the edges for $N + 1$ intervals. All intervals contain the same number of points, with the exception of the last interval, that may contain less points than the others.

   (b) Find which sample is closest to $A_n$. Call this sample $B_m$.

   (c) Consider $B_{m+1}, B_{m+2}, \ldots, B_{m+r}$ where $B_{m+r}$ is the first sample that does not meet the spatial condition or the last measurement point of the granule. Consider $B_{m-1}, B_{m-2}, \ldots, B_{m-t}$ where $B_{m-t}$ is the first sample that does not meet the spatial condition or the first point of the granule. If $N$ is large enough, all points that meet the spatial criterion are contained by the super-interval $(B_{m-t}, B_{m+r})$, because the minimum of the distance from $A_n$ to $B$ will be contained by it (if $N$ is too small, this interval
may contain only a local minimum). An example of such a super-interval is shown by the thick line in Fig. 2. Consider this super-interval.

(d) Calculate the distance between \( A_n \) and every point in the super-interval.

(e) Note all points for which the spatial condition is met. If there are no such points, remember the distance of the closest point.

As shown in Fig. 2, \( N = 20 \) is already sufficiently large to guarantee that any points in \( B \) meeting the spatial criterion are contained in the super-interval. However, with \( N = 20 \) the number of points in the super-interval for which the distance to \( A_n \) will be calculated is still quite large. A larger \( N \) means the super-interval will be smaller, but the number of samples for which the distance will be calculated will be larger. The choice of \( N \) is thus an optimisation problem to reduce the number of distance calculations. We have chosen \( N = 200 \).

4. If there were any points for which the spatial condition was met, increase \( n \) by 1 and start again from 3.

5. If there were no points for which the spatial condition was met, calculate the least number of points remaining before it could be met: increase \( n \) by

\[
\text{smallest} \frac{\text{distance} - \text{spatial condition}}{\text{max speed}}
\]

and start again from 3. For example, if the shortest distance is 120 km, the spatial condition distance 20 km, and the max speed 10 km/point, \( n \) will be increased by \( \frac{120-20}{10} = 10 \).

This works, because if the minimum distance from \( A_n \) to \( B \) is 120 km and the distance between \( A_n \) and \( A_{n+10} \) is 100 km, the maximum distance between \( A_{n+10} \) and \( B \) cannot be less than 20 km.

The procedure described above is not the fastest possible (for example, point (d) could be optimised further) but with this algorithm, the bulk of the time running the code searching for collocations was spent on downloading files from a local server and decompressing them.

From all points obtained with the procedure described above, those for which the time difference is less than 15 min are selected. Even though many of those CPR measurements are outside the MHS pixel, all are stored in the collocated dataset, because the MHS pixel size is a function of the scan angle, and some applications may allow for the CPR pixel to be (just) outside the MHS measurement. Also, it is cheap to select a subset of collocations, but to find pixels slightly further away than the initial criterion, the algorithm would need to be rerun.

For each collocation and for each sensor (CPR, MHS, HIRS and AMSU-A), we store the location (lat/lon), the measurement time, the time of the first measurement in the file (to help finding the file containing the measurement) and the location of the point inside the datafile (row/column). We also store the distance of each centerpoint to the CPR centerpoint, and the time difference (MHS time minus CPR time). With this information, one can find exactly which of the CPR pixels fall inside the MHS pixels, possibly considering the sensor spatial response function.

For the second dataset, we collect the CPR pixels in an MHS pixel and calculate the number of CPR measurements, the average, the standard deviation and the coefficient of variation (standard deviation divided by mean) of the IWP product. Here, we choose a circular MHS pixel area with a radius of 7.5 km, so we are certain that the CPR pixels are contained by the MHS pixel independently of the scan angle. We also note the cloud fraction, defined as the number of CPR pixels with at least 1 g m\(^{-2}\) of ice divided by the total number of CPR pixels inside the MHS measurement.

### 3.1.1 Comparison with Nagle and Holz (2009)

The method described above is quite different from the method described by Nagle and Holz (2009), henceforth referred to as “NH”.

NH divide the two instruments to be collocated into a master and a slave, where the small slave observations are projected on the large master footprint. They find the location of the satellites as a function of time (forward navigation) and “estimate the time at which a slave satellite passes abeam of a master FOV on the surface” (inverse navigation). They then calculate simultaneous nadir observations (SNO), when two satellites pass over any point on the ground within a certain time window. For this calculation, NH use an orbital prediction model. They search the scan lines around the SNO for overlap with the master FOV. NH assign weights to each of the slave observations based on the sensor spatial response function of the master.

NH claim that their method works for any combination of satellite, aircraft and ground observations. However, a scanning instrument might very well collocate with a ground observation without any SNO if the measurement is strongly off-nadir. For (near)-parallel orbits, this can be the case between different satellites as well. In fact, at one point NH “presuppose that the two orbital planes are not nearly coincident”.

NH use the satellite position to calculate the projected sensor spatial response function on the Earth surface. We use an expression from Bennartz (2000) to calculate the size of the pixel, and we do not presently consider the sensor spatial response function.

NH was designed to be computationally efficient and may very well be faster than our method. However, our method is conceptually simpler than NH. Our method does not require any forward or inverse navigation. It finds collocations regardless of the presence of simultaneous nadir observations. For some applications, only simultaneous nadir observations are of interest; in this case, NH and our method should give the same result.

The processing of slightly more than two years of data from CloudSat and five AMSU/MHS sensors with our methods took around two weeks of computer time on a powerful workstation (Intel Xeon Dual Quadcore 3.20 GHz, 16 Gigabyte Random Access Memory (RAM)). Most of this time was due to transferring files over the network and decompressing them. We did not carry out a comparison of speed and results using a common set of source data.

3.2 Collocation statistics

We have located collocations for the period between 15 June 2006 13:12 and 4 October 2008 10:34. For the year 2007, we have found 124 822 977 collocations between the NOAA-18 MHS and the CloudSat CPR, where the maximum distance between MHS and CPR centre points did not exceed 15 km and the time difference between MHS and CPR measurements was limited to 15 min. With a maximum distance of 7.5 km and counting the MHS pixels, the number of collocations reduces to 2 669 135. If only tropical nadir points are selected (within 30 degrees of the equator, within 1 degree of nadir), around 1% or 26 410 MHS pixels remain.

Figure 3 shows the latitudes at which collocations occur between the CloudSat CPR and the MHS/AMSU-B on different satellites. It shows that only the NOAA-18 MHS has collocations with the CPR globally. This is due to the fact that the LTAN of the NOAA-18 (13:39) is always similar to the CloudSat LTAN (13:30–13:45). NOAA-18 is near the A-Train constellation and thus near CloudSat, because CloudSat is part of the A-Train. All other POES considered in this study have collocations with CloudSat CPR only near the poles.

Figure 4 shows at which angles and latitudes the collocations occur. At the equator, no nadir collocations with a time difference of less than one minute occur. Rather, the viewing angle is slightly off-nadir. If two satellites pass through the same place in space\(^2\) with one minute in between, the Earth rotates so their subsatellite points are roughly 1 m/24 h·40 075 km≈27.8 km apart. For a NOAA-18 altitude of 850 km, the viewing angle then needs to be \(\tan^{-1}(27.8/850) = 1.9^\circ\). In reality, the satellites do not pass through the exact same point, and the viewing angles for collocations within one minute are slightly larger. The CloudSat has a slightly lower inclination than NOAA-18, so for a collocation to occur, NOAA-18 has to look to the left when it reaches its northernmost point and to the right when it reaches its southernmost point.

CloudSat and NOAA-18 are in some sort of “orbital resonance”, as shown in Fig. 5, showing the collocations in January 2007. Figure 5 shows a time series of the number of collocations per hour, where the upper left is 1 January, 00:00–00:59 and the lower right is 31 January, 23:00–23:59 (inclusive). The figure shows a collocation pattern with a 56-h period: 16 h with collocations, 40 h without.

3.3 Sampling effects

As shown in Fig. 1, an MHS footprint is an order of magnitude larger than a CPR footprint. The smallest MHS pixel is the nadir-viewing pixel, which has a diameter of 16 km. The CPR pixel can be approximated by an ellipse of 1.3 by 1.7 km\(^2\). It covers at most 0.65% of the area an MHS pixel:

\[
\frac{A_{\text{CPR}}}{A_{\text{MHS}}} = \frac{\pi \cdot 1.3 \cdot 1.7}{\pi \left(\frac{16}{2}\right)^2} = 0.0065 = 0.65\%
\]

Many CPR measurements fit in one MHS measurement. Since the CPR is not a scanning instrument, CPR pixels never fill an MHS pixel completely. In the best case, a nadir MHS pixel contains around 15 CPR pixels (or only 13 when we limit the collocations to CPR pixels within 7.5 km of the MHS centrepoint). The total area is less than 15·0.65%≈9.75% because of the overlap between subsequent CPR pixels. Usually, the CloudSat ground track does

\(^{2}\)The same place in space in an Earth-centered inertial reference system.
A collocation is considered representative, or good, if the CPR IWP statistics for the area covered by CPR are the same as the statistics of a hypothetical CPR IWP covering the full MHS pixel.

CPR pixels inside the MHS pixel have the same statistics as they would if they would fill the entire MHS pixel. Whether the collocation is representative cannot be known exactly, because high-resolution information on the part of the MHS pixel not covered by CPR pixels is not available in this approach. However, we can look at some indicators to make an educated guess as to how well the CPR pixels represent the MHS pixel.

Figure 6 shows three graphs that give some insight in the sampling error. The MHS pixel is assumed to be circular with a radius of 7.5 km.

In Fig. 6a we can see that most collocations contain a relatively large number of CPR pixels, but many do not. When the number of CPR pixels inside the collocation is small, the CPR pixels are close to the MHS footprint edge and poorly represent the MHS pixel. The highest number of CPR pixels inside a MHS pixel occurs when the CPR groundtrack passes close to the centre of the MHS footprint. This is the optimal case.

Figure 6b shows a histogram of the coefficient of variation of the CPR IWP product for the CPR pixels within 7.5 km of the MHS centrepoint. A small coefficient of variation corresponds to a homogeneous cloud. The more homogeneous the cloud, the more representative the CPR pixels are for the complete MHS footprint area. We use the coefficient of variation rather than the standard deviation because the standard deviation is likely to be much larger for clouds with a high IWP than for clouds with a low IWP. Selecting collocations based on the standard deviation would throw away many of the measurements with high IWP. The coefficient of variation is largest when some CPR pixels measure a strong cloud and others do not measure any cloud at all. This indicates the presence of a strong, localised cloud, which significantly reduces our trust in the representativeness of the CPR pixels.

In Fig. 6c, the distribution of CPR inside MHS is shown for three cases. The red dots show a case with an extremely high coefficient of variation (2.106; note in panel (b) that a coefficient of variation larger than 2 is so rare that it is not visible in the histogram). Since a thick cloud that is only 1 km in diameter is unlikely, this happens usually when the cloud is just on the edge of the MHS pixel. In either case, the CPR pixels do probably not share the same statistics as the MHS footprint and the collocation is not useful. The green dots show a case with a very low coefficient of variation (0.017; cases where all CPR pixels have the same nonzero measurement and the coefficient of variation is 0 occur as well, but the IWP value tends to be 1 g m\(^{-2}\) so it would not be visible in this graph). The portion of the cloud imaged by CPR has a roughly constant IWP of around 70 g m\(^{-2}\). It is quite likely that the rest of the MHS pixel looks similar. The example in blue shows a collocation with a coefficient of variation of 0.354.

When the criteria discussed above are applied, sampling effects are reduced and a large number of collocations remain.
4 Applications

Collocations can be used in many different ways. This section presents some possible applications of collocations between CloudSat CPR and NOAA-18 MHS. Three examples are explored in the following subsections. This section is meant to show what can be done with such a collocated data set and does not provide a comprehensive study of the different applications.

4.1 Comparison with NESDIS IWP

Various algorithms exist to determine IWP from microwave radiometer measurements (Liu and Curry, 2000; Zhao and Weng, 2002; Weng et al., 2003). The National Environment Satellite, Data and Information Service (NESDIS) publishes an operational IWP product from MHS measurements in the Microwave Surface and Precipitation Products System (MSPPS). Zhao and Weng (2002) assume spherical ice particles and calculate the effective particle diameter from the ratio between the scattering at 89 GHz and 150 GHz. They assume a constant bulk volume density and calculate the IWP from this. They also discuss how errors propagate in the retrieval algorithm, but no discussion of systematic error and no validation for the NESDIS MSPPS IWP was found in this paper, nor elsewhere in the literature. Waliser et al. (2009) find a dry bias in the NESDIS IWP product. They explain this from the Zhao and Weng (2002) screening criteria and the MHS insensitivity for ice particles smaller than 0.4 mm.

CloudSat IWP has a systematic uncertainty of up to 40% (Austin et al., 2009). Judging from the available data, the detection limit for CloudSat IWP is 1 g m⁻².

Figure 7 shows a comparison of the NESDIS MSPPS IWP with the CloudSat IWP. It shows that the NESDIS IWP is systematically smaller than the CPR IWP. For many nonzero CloudSat measurements, the NESDIS IWP is zero. This is because thin clouds are (almost) transparent for microwave radiation in the frequencies at which MHS operates (Greenwald and Christopher, 2002). For some NESDIS IWP measurements, the CloudSat IWP is zero. This happens due to the different footprint sizes. The MHS footprint is much larger than the CPR footprint. A cloud that does not cover a complete MHS pixel may be missed by the CPR (see Sect. 3.3).

MSPPS IWP is systematically lower than CPR IWP by approximately 70–90%. Austin et al. (2009) estimate the CPR accuracy to 40%, based on a retrieval blind comparison study by Heymsfield et al. (2008), which was based on simulated radar observations for ice particle data from aircraft in-situ measurements. While the profiles considered in that study may not be representative for all atmospheric cases, we can
still consider the CPR data to be considerably better validated than the MSPPS data. It is therefore likely that the difference reflects a real low bias in the MSPPS data. This is partly a fundamental problem, because of the transparency of thin clouds to radiation at MHS frequencies. However, MSPPS underestimates the IWP for thick clouds as well. A more accurate IWP product based on microwave measurements is probably possible. One way to obtain such a product is by using a neural network, described later in the article.

4.2 Comparison of BT-IWP relations

As a second application example, we investigate the relation between the MHS channel 5 brightness temperature and the associated Ice Water Path for two different datasets. The first dataset consists of the collocations, providing a mapping between brightness temperatures and independent IWP. The second dataset consists of a mapping generated from 30,000 synthetic atmospheres as described below. Note that this mapping predates the collocated measurements. Rydberg et al. (2009) use this method to derive IWC from the Sub-Millimetre Radiometer (SMR) on the Odin satellite. It can potentially be used to derive IWP from MHS.

Atmospheric states, including clouds, are generated following the procedure described by Rydberg et al. (2009), and a brief overview is given here. Cloud states are generated in a series of steps, where two-dimensional (2-D) radar reflectivity fields from the Cloud Profiling Radar on-board CloudSat serve as the basis for obtaining realistic cloud structures. Orbit sections of CloudSat data (with a resolution of ~ 250 m in vertical by 2 km along the scan line) are transformed to 3-D by inputting those into a stochastic iterative amplitude adjusted Fourier transform algorithm (Venema et al., 2006). This algorithm generates surrogate 3-D radar measurement fields with the same spatial resolution as the original fields.

Cloud microphysical fields are generated in such a way that the surrogate 3-D radar reflectivity fields are conserved. This is done by assuming that spherical ice particles can be used to represent the single scattering properties of natural occurring ice particle populations. We lack information about the true shape of the ice particles, which is different for different cloud types, and the most generic assumption is to assume spheres. This is also the assumption made by Austin et al. (2009) for the CloudSat CPR IWP retrieval. The accuracy of this approximation is difficult to assess, because the true microphysical parameters are unknown. Furthermore, the cloud ice particle size distribution (PSD) parameterisation derived by McFarquhar and Heymsfield (1997) (hereafter MH97) is assumed to be the best representation of the tropical mean PSD. MH97 depends on temperature and ice water content (IWC), and is used to map radar reflectivity fields to IWC and PSD fields. However, it should be clear that local PSD may deviate significantly from MH97. For temperatures above 273 K, clouds are assumed to consist entirely of spherical water particles and the PSD of stratus cloud derived by Deirmendjian (1963) is used.

Weather data (temperature, humidity, and pressure) and ozone information, originating from the European Centre for Medium-Range Weather Forecasts (ECMWF), are obtained from the CloudSat auxiliary data archive (ECMWF-AUX). ECMWF-AUX contains ECMWF state variable data interpolated to each CPR bin. These fields are handled as described by Rydberg et al. (2009) in order to have a realistic variability that accounts for variations on scales not resolved by ECMWF.

Radiative transfer simulations of nadir viewing AMSU-B channel 20 (corresponding to MHS channel 5) are performed using version 1.1 of the Atmospheric Radiative Transfer Simulator (ARTS). This is a development of the first version, ARTS-1 (Buehler et al., 2005), where two scattering modules, a discrete ordinate iterative method (Emde et al., 2004) and a reverse Monte Carlo algorithm (Davis et al., 2005) have been implemented to solve the polarised radiative transfer equation. The Monte Carlo module is used and the 3-D variability of the atmosphere is fully considered in the radiance simulations. The lower and upper sidebands of AMSU-B channel 20 are represented by single frequencies of 176.01 and 189.91 GHz, respectively. For a diverse set of atmospheric profiles, the root mean square error between this approximation and a setup with a finer frequency grid is 0.020 K. The instrument antenna spatial response function is assumed to be a 2-D Gaussian with a full-width half-power beamwidth of 1° in both dimensions. Pencil beam simulations with a grid spacing matching the
atmospheric states horizontal resolution are performed. After the antenna weighting the precision of the simulations is better than 0.5 K. The IWP is extracted along each pencil beam where radiative transfer simulations are performed. The atmospheric scenario has a higher spatial resolution than AMSU-B, so the simulated IWP are weighted according to the antenna pattern to obtain the AMSU-B IWP.

Figure 8 shows a comparison between the two relations. We average the CPR IWP over the MHS pixel, and select a subset of collocations. For the collocations, only measurements that are within 20 degrees of the equator are used, in order to prevent a signal from the surface (Buehler and John, 2005). Only collocations where the MHS measurement is within 5 degrees of nadir are used, so that no significant limb effect occurs. Finally, collocations are selected where all CPR pixels are cloudy and the coefficient of variation is smaller than one, for reasons discussed in Sect. 3.3 above.

The figure shows AMSU-B channel 20 or MHS channel 5 brightness temperature as a function of the IWP (logarithmic) for the two different datasets. In blue are the collocated measurements (MHS channel 5 and CPR IWP). The red boxes show simulated radiances for generated atmospheric states (AMSU-B channel 20 and generated IWP).

The figure shows that both datasets have largely the same statistical features. For IWP up to around 100 g m\(^{-2}\), the effect on the brightness temperature is minimal, because thin clouds are not resolved at MHS channels 3–5 frequencies (Greenwald and Christopher, 2002). For higher values of IWP, the brightness temperature decreases logarithmically as a function of IWP. For IWP >100 g m\(^{-2}\), the simulated brightness temperatures are slightly higher than the observed ones.

The microphysical assumptions for the generated atmospheric states are based on MH97, which differ from the assumptions in the CloudSat retrieval. This might contribute to the observed differences.

Overall, the variability in the simulated brightness temperatures is smaller than the variability in the observed brightness temperatures. This effect is stronger for higher values of the IWP. Several factors may contribute to this discrepancy. The CPR pixels are much smaller than the MHS pixels, so the measured value is averaged over a smaller area. If a small, concentrated cloud exists inside a MHS pixel, the CPR might either see it, in which case it measures a higher IWP than the MHS, or it might miss it, so it measures a lower IWP. This adds to the variability. Additionally, the generated atmospheric states might not fully resolve the natural variability of cloud microphysical parameters and of atmospheric temperature and humidity.

4.3 Developing a retrieval using neural nets

An artificial neural network (ANN) is an interconnected assembly of processing units called neurons (e.g. Jiménez et al., 2003). Neural nets are widely used to statistically characterise the mapping between radiometric measurements and related geophysical variables (e.g. Krasnopolsky, 2007). We use an ANN to characterise the mapping between MHS radiances and the CPR IWP, and then use the trained ANN to retrieve IWP from the MHS measurements. We call this retrieval MHS-CPR IWP.

MHS-CPR IWP has both advantages and disadvantages compared to other retrieval approaches. One can use a neural network with simulated rather than measured radiances, or one can use a more classical retrieval method. As we use the collocated measurements, an advantage is the relative simplicity; there is no need for a potentially complicated radiative transfer model with many possible sources of error. On the other hand, the collocations approach may introduce a number of errors, as discussed in Sect. 4.3.1. However, an MHS-CPR IWP can complement the other existing retrieval methods. The retrieval quality can never become as good as CloudSat, but the spatial and temporal coverage will be much larger.

The neural network approach described below is in the exploration phase and will be developed further.

We select a subset of collocations that provide a relatively homogeneous dataset. The subset is restricted to pixels over ocean within 20 degrees of the equator, because a warm (and humid) atmosphere prevents the MHS from getting a signal from the surface (Buehler and John, 2005). Due to these restrictions, the neural network is only applicable to the tropics. A strongly off-nadir measurement is colder due to the limb effect (Buehler et al., 2004). For the training, we restrict ourselves to measurements within 5 degrees of nadir. This avoids the need to compensate for this effect (described below). The neural network works for nadir measurements or measurements where the limb effect is compensated.

As discussed in Sect. 3.3, the MHS measurement compromises a larger area than the CloudSat measurement, even when we average the CPR pixels inside an MHS pixel. If a small, strong event is present inside an MHS pixel, the CloudSat might miss it completely or measure exactly this event. In both cases, the observed MHS radiance is the same, but the CPR IWP can vary considerably. For that reason, we select only homogeneous measurements: the collocation shall contain at least ten CPR pixels, all measuring at least 1 g m\(^{-2}\), and the standard deviation shall not exceed the mean value. The selection of only “cloudy pixels” for the training leads to a wet bias, because the neural network tends to the mean state if it has insufficient information from the input. We want to explore the effect of adding HIRS channels on the neural network retrieval. Hence, we choose collocations where at least five CPR pixels are within 10 km of the nearest HIRS pixel.

Finally, only collocations where the time interval is at most ten minutes are selected.

For the year 2007, we find 2627 collocations that meet the criteria described above.
For the neural network calculations, we use the MATLAB Neural Network toolbox V6.0.1 (R2008b). The collocations are divided in 60% training, 15% testing and 25% validation. MHS channels 3, 4 and 5 are the inputs. As a target, we choose the log IWP which was found to work better than the ordinary IWP. The transformation is reversed after the application of the neural network. Throughout the process, CPR IWP is assumed to be the truth. The training is considered to be finished if the error with the testing data increases for fifteen consecutive iterations. After training, we store a neural network that we can then use for our retrieval.

To compensate for the limb effect, we correct the brightness temperatures before we input them to the network. For each viewing angle and channel, the mean brightness temperature is calculated. We use only tropical measurements (within 30 degrees of the equator) to prevent an angle-dependent signal from Antarctica, which is mainly seen by one side of the scan. The limb effect is minimal for the two viewing angles closest to nadir. The average brightness temperature for those angles is the reference. The limb effect can be quantified by the difference between the reference brightness temperature and the mean brightness temperature for a certain viewing angle. We compensate for the limb effect by adding this difference to all measurements for this viewing angle.

In Fig. 9 we show an example of how a NN IWP product might look like. The data is for 1 January 2008. The left panels show the MHS brightness temperatures between 08:56 and 19:02 UTC, the right panel shows the IWP retrieved by the neural network.

4.3.1 Error analysis

Four sources of error can be identified: (a) The CPR IWP uncertainty is up to 40% (Austin et al., 2009). This propagates directly into the MHS-CPR IWP. (b) Collocation mismatches add noise to the training data, as discussed in Sect. 3.3. This may or may not result in an error in the MHS-CPR IWP (noise in the input data need not change the best fit). (c) The inversion from MHS data inherently has a limited accuracy, leading to a significant uncertainty in the MHS-CPR IWP. (d) The MHS has a radiometric noise of up to 0.55 K and might suffer from calibration errors.

Figure 10 shows a scatter plot between CPR IWP and collocated MHS-CPR IWP. Both axes are logarithmic. (a) and (d) do not contribute to the variability seen here. MHS-CPR IWP could still perfectly reproduce MHS-CPR IWP even considering the MHS radiometric noise, because this noise is part of the training data. If it would do so, CPR IWP might still be off by 40% compared with the true atmospheric IWP, but Fig. 10 would not show variability.

The variability is consistent with simulations similar to the ones described in (Jiménez et al., 2007). Since those simulations did not use collocations, the dominant source of the variability in Fig. 10 is likely to be the inversion error (c).

For low IWP, the network exhibits a wet bias. Thin clouds are (almost) completely transparent at MHS frequencies (Buehler et al., 2007), so with only those measurements, there is no information for thin clouds. With no information, the neural network tends towards the mean state. Since only cloudy CPR pixels were used for the training, this explains the wet bias.

Figure 11 shows the neural network sensitivity to MHS radiometric noise. A subset of tropical nadir measurements for 2007 are selected. For practical reasons, this subset consists of the MHS measurements for where collocations could be found; however, as the CloudSat values are not used for this figure, those measurements are effectively a sample of all MHS measurements for 2007. The figure shows the mean fractional IWP error as a function of IWP and input noise. For this figure, the neural network is applied twice. First, the unperturbed input data (MHS brightness temperatures for channels 3, 4 and 5) are fed into the ANN. This gives an unperturbed IWP for each measurement. Then, we add gaussian noise, starting with $\sigma=0.1$ K, to the input data, and feed this perturbed data to the ANN. This results in a perturbed IWP denoted by $\tilde{\text{IWP}}$. For each collocation, the fractional error is calculated as $\left|\frac{\tilde{\text{IWP}}}{\text{IWP}} - 1\right|$. Those fractional errors are divided into bins according to the unperturbed IWP value.
For each bin, we calculate the mean fractional error. This process is repeated for higher values of $\sigma$, up to $\sigma=2.0$ K, taking steps of $\sigma=0.1$ K.

Unsurprisingly, Fig. 11 shows that a higher input noise results in a higher error in the output. This effect is linear. The mean fractional error as function of IWP is less straightforward. The error is largest for IWP values of around $100 \text{ g m}^{-2}$ and smaller for values that are either larger or smaller. This can be explained as follows. For small IWP, a small perturbation in the brightness temperatures has little influence on the IWP. The ANN does not interpret the brightness temperature noise as IWP. This is in line with the observation that thin clouds are transparent to the frequencies at which MHS operates (Greenwald and Christopher, 2002), and can also be seen in Fig. 8. For large IWP, MHS channels 3–5 will observe large depressions in brightness temperature, and a 2 K noise is much smaller than the signal, so its effect on the output is also small. However, for intermediate values of IWP, around $100 \text{ g m}^{-2}$, the noise is of a similar order of magnitude as the signal, and the ANN is quite sensitive to input noise. The actual radiometric noise for MHS depends on the channel, but is always below 0.55 K (Kleespies and Watts, 2007). This means that radiometric noise is unlikely to be a dominant error source for this kind of IWP retrieval method.

4.3.2 Adding HIRS

Thin clouds are not visible by MHS channels 3–5 because the effect of ice clouds on microwave radiation at those frequencies is relatively small. In the infrared, the situation is different: even a small cloud has an observable effect, but an infrared sensor does not see the difference between a medium cloud and a thick cloud, because the sensor is saturated quickly (Jiménez et al., 2007). Hence, we can expect the retrieval quality to improve if we combine infrared and microwave measurements.

Figure 12 shows a scatter plot similar to Fig. 10, but with additional HIRS channels 8 and 11 (chosen for their clear cloud signal). The number of collocations used for the neural net remains the same, because we already preselected the

Fig. 9. The neural network (see text) can be used to retrieve IWP from radiances. The figure shows observations by NOAA-18 in the descending node on 1 January 2008 between 10:54 and 17:20 UTC (local time during the night). The left panels show the brightness temperatures as observed by the MHS channels 3–5. The right panel shows the IWP as generated with the neural network as described in the text. Cold areas in the left panel correspond to wet areas in the right panel.
collocations so that at least five CPR pixels are less than 10 km from the nearest HIRS pixel centerpoint. Additionally, HIRS might suffer from the difference in footprint location for HIRS and MHS, even if HIRS is not yet fully understood. One factor may be the difference in footprint location for HIRS and MHS, even if HIRS is not yet fully understood. One factor may be the

By eye, it is hard to see whether there is any improvement gained by adding them.

Figure 13 shows the fractional median error as a function of IWP for both cases. Here, the fractional median error is defined relative to CloudSat, so CloudSat is assumed to be true. The dashed line shows the error for the ANN where the input consists only of MHS channels, the dotted line shows the error for the ANN with an input consisting of MHS channels 3–5 and HIRS channels 8 and 11. For small values of IWP there is an improvement when adding the HIRS channels. However, the error is still large, since a median relative error of 2 means that the retrieved IWP is on average a factor 2 off. For larger values of IWP, the errors are roughly the same, as expected.

Why the retrieval does not strongly improve when adding HIRS is not yet fully understood. One factor may be the difference in footprint location for HIRS and MHS, even if only collocations with at least 5 CPR pixels in the HIRS pixel are considered. Additionally, HIRS might suffer from the

Fig. 10. Scatter plot showing the performance of the ANN using MHS channels 3 to 5. The retrieved IWP is plotted against the input IWP.

Fig. 11. ANN sensitivity to errors in the input brightness temperatures. Here, only the MHS channels are used as input to the ANN. See the text for an explanation and a discussion.

Fig. 12. Scatter plot to show the performance of the neural network, MHS 3–5, HIRS 8 and 11.

Fig. 13. Comparison of the median fractional error between independent and retrieved IWP, when only MHS channels are used or when both MHS and HIRS channels are used as input to the ANN. The median fractional error is defined as the median of all errors with a certain IWP, where the error is defined as $\frac{IWP_{NN} - IWP_{CPR}}{IWP_{CPR}}$. 
beam-filling problem: the sensor may be saturated if only a part of the pixel is cloud-covered, and be unable to tell the difference between a partly cloudy and a fully cloudy pixel. A further investigation is necessary and will be carried out.

5 Conclusions

The collocation-finding method described in this work finds many collocations between the NOAA-18 MHS and the CloudSat CPR. Those collocations are frequent and globally distributed. Other POES collocations with CloudSat are limited to the polar areas. Sampling effects due to different footprint sizes need to be taken into consideration.

There are numerous possible improvements to our procedure. The procedure to find the collocations can be refined by considering how the MHS footprint size depends on the scan angle. Even better, one can project the MHS sensor spatial response function onto the surface and calculate a weighted average of the collocated CPR pixels, similar to the procedure described by Nagle and Holz (2009).

In comparison with Nagle and Holz (2009), our algorithm is relatively simple. For example, it does not need satellite position data. It finds collocations even in the absence of simultaneous nadir observations.

Our method was designed for the case where one instrument is scanning and the other has a fixed viewing angle. It also works if both instruments are scanning, but in this case, it is slow and a different method is more suitable. If either satellite is geostationary or both instruments are on the same satellite, more optimised methods may be appropriate. The method does not depend on the nature of the sensor (active, passive) or the footprint size.

The collocations have various applications. They can be used to compare different IWP products. As an example, we have compared the NOAA NESDIS MSPPS MHS IWP product against the CloudSat CPR IWP product. IWP values from the CloudSat CPR were found to be significantly larger than those from the MSPPS. This may be partly attributed because thin clouds are transparent to radiation at MHS frequencies, but since the MSPPS underestimates IWP even for high values, there should be room for improvement.

As a second example, we have compared the IWP-BT relation for our collocations with the one for simulated radiances from synthetic atmospheric cases. The variability in the measured relation was found to be larger than the variability for the simulated relation.

The validation for simulated radiances was performed statistically. A stronger validation would be to simulate the radiances for the exact cases where a collocation exists.

As a final example, we have used the collocations to train an Artificial Neural Network to develop a new IWP product. We have shown that this method is promising. Finally, we have investigated the effect of adding HIRS channels 8 and 11 to such an ANN. Unexpectedly, this leads to only a small improvement in the retrieval quality.

The IWP retrieval using an Artificial Neural Network looks promising, but requires additional work. We can improve the retrieval in various ways. One can make a stronger restriction for homogeneous scenes by looking at MODIS or AVHRR pixels inside the MHS, although this is limited as infrared measurements do not detect the vertical extent of the cloud. Another alternative is to combine MHS with other HIRS channels than those explored so far, or to directly input a combination of MHS and AVHRR for the training. On the other hand, the ANN might be extended to work for more measurements. By having more input parameters or multiple neural networks, the retrieval could work globally.

One can extract additional information from other high-resolution data, such as from the Moderate Resolution Imaging Spectroradiometer (MODIS; King and Greenstone, 1999) or the Advanced Very High-Resolution Radiometer (AVHRR; Cracknell, 1997). to better characterise the collocations. Those can be used to make a stronger estimate as to how homogeneous the scene observed by MHS is.

All the applications can be expanded upon and many other applications can be developed.

These and other issues will be addressed in further research. In particular, future work will focus on developing a global IWP product from passive microwave and infrared sensors available on operational polar orbiting satellites.

The collocations are available for public use.

Acknowledgements. The bulk of the work was carried out as part of the Master’s Thesis by first author Gerrit Holl. Thanks to the Spacemaster education programme for making this possible. We thank the people, organisations and institutes that have helped us to obtain the satellite data. Thanks to Lisa Neclos for helping us with archived HIRS data, and to all involved with the NOAA CLASS archive for recent and current MHS, AMSU-B and HIRS measurements. Thanks to CloudSat for making available an accurate IWP product. We thank the UK MetOffice for providing the AAPP package and the ARTS radiative transfer community for its work on ARTS. We would also like to thank the National Graduate School in Space Technology at Luleå University of Technology. Thanks to the OpenStreetMap community for the freely usable map-data in Fig. 1. We would like to thank the AMT Editorial Board, associate editor Bernhard Mayer, and two anonymous reviewers, for their work in improving the article.

Edited by: B. Mayer
References


Rydberg, B., Eriksson, P., Buehler, S. A., and Murtagh, D. P.: Non-


Understanding inter-satellite biases of microwave humidity sounders using global simultaneous nadir overpasses

Understanding intersatellite biases of microwave humidity sounders using global simultaneous nadir overpasses

Viju O. John,1 Gerrit Holl,2 Stefan A. Buehler,2 Brett Candy,3 Roger W. Saunders,3 and David E. Parker1

Received 10 June 2011; revised 18 November 2011; accepted 23 November 2011; published 28 January 2012.

Simultaneous nadir overpasses (SNOs) of polar-orbiting satellites are most frequent in polar areas but can occur at any latitude when the equatorial crossing times of the satellites become close owing to orbital drift. We use global SNOs of polar orbiting satellites to evaluate the intercalibration of microwave humidity sounders from the more frequent high-latitude SNOs. We have found based on sensitivity analyses that optimal distance and time thresholds for defining collocations are pixel centers less than 5 km apart and time differences less than 300 s. These stringent collocation criteria reduce the impact of highly variable surface or atmospheric conditions on the estimated biases. Uncertainties in the estimated biases are dominated by the combined radiometric noise of the instrument pair. The effects of frequency changes between different versions of the humidity sounders depend on the amount of water vapor in the atmosphere. There are significant scene radiance and thus latitude dependencies in the estimated biases and this has to taken into account while intercalibrating microwave humidity sounders. Therefore the results obtained using polar SNOs will not be representative for moist regions, necessitating the use of global collocations for reliable intercalibration.

1. Introduction

Tropospheric humidity measurements from microwave humidity sounders such as Advanced Microwave Sounding Unit-B (AMSU-B) [Saunders et al., 1995] and Microwave Humidity Sounder (MHS) [Bonsignori, 2007] have been proven to have significant impact on the skill of numerical weather prediction [Andersson et al., 2007]. This is primarily due to their ability to measure humidity under all-sky conditions compared to clear-only sampling by infrared sounders [e.g., John et al., 2011]. Recently, some attempts have also been made to use microwave humidity sounding data for climate applications [e.g., Xavier et al., 2010; Eymard et al., 2010; Buehler et al., 2008]. However, although microwave temperature sounding data have been intercalibrated and extensively used for climate studies [Thorpe et al., 2010], this has not yet been done for the humidity sensors. The main reason for this is the short span of the data, primarily since late 1998; although Special Sensor Microwave Humidity Sounder (SSM/T-2) data began in 1994, these early measurements were not widely used except for research [e.g., Miao et al., 2001; Selbach et al., 2003; Sohn et al., 2003; Chung et al., 2011]. The error characteristics of SSM/T-2 radiances data are not fully understood, and careful validation is essential before they can be used to assess, in particular, long-term trends in upper tropospheric water vapor which is an important climate variable, yet poorly simulated by current climate models [e.g., John and Soden, 2007].

Unfortunately, there is a lack of stable and reliable ground-based or in situ reference measurements of atmospheric humidity to intercalibrate satellite instruments [Seidel et al., 2009]. Cao et al. [2004, 2005] have developed a method to find simultaneous nadir overpasses (SNOs) of polar orbiting satellite pairs and use them for intercalibration. There are regular near-polar SNOs and during an SNO, similar instruments on the different satellite platforms measure radiation emitted from the same area of Earth and/or its atmosphere at the same time. Therefore any difference in the radiance measured by the satellites can be used to intercalibrate the measurements. This is being developed in support of the Global Space-Based Inter-Calibration System (GSICS) initiative to provide climate quality satellite data sets [Goldberg et al., 2011].

SNO data have been proven useful for intercalibration of instruments such as HIRS and MSU/AMSU-A [e.g., Zou et al., 2006; Wang et al., 2007; Cao et al., 2005; Iacovazzi and Cao, 2007; Shi et al., 2008]. However, Iacovazzi and Cao [2008] showed that for those channels which are sensitive to the Earth’s surface, there are large uncertainties in...
the estimated intersatellite bias due to surface inhomogeneity which arises mainly from variable surface emissivity of SNO scenes at subpixel scales.

[5] The concerns expressed by Iacovazzi and Cao [2008] can be put in the context of microwave humidity sounders as follows. The peak emission for a sounding channel occurs at an atmospheric level for which the optical depth, integrated from the top of the atmosphere, becomes approximately one [e.g., Petty, 2006]. Therefore depending upon the amount of water vapor in the atmosphere, the peak emission levels of humidity sounding channels move up and down, in contrast to temperature sounding channels which use the absorption of well mixed gases such as oxygen or carbon dioxide. Thus the sounding height of a humidity channel is at its maximum in a wet tropical atmosphere and becomes lower as the satellite moves toward higher latitudes. Figure 1 shows how the total opacity, which is the vertically integrated absorption coefficient, varies as a function of the amount of water vapor in the atmosphere. For the dry atmospheres sampled by SNOs which normally occur between 70° and 80° latitudes [Cao et al., 2004], the opacity is of order one even for the channel closest to the 183.31 GHz water vapor line center. Analysis of ERA-Interim [Dee et al., 2011] four-times daily precipitable water vapor data for 2010-01 and 2010-07 showed that more than 50% of the values are below 3 mm at latitudes 70–80° except for the Arctic region in summer. This is consistent with the results of Melsheimer and Heygster [2008]. So for all channels on microwave humidity sounders, there is a significant contribution from the Antarctic surface and the Arctic surface in winter, and the radiation which reaches the satellite is then determined substantially by the surface skin temperature and the surface emissivity; the atmospheric contribution is relatively small. High-latitude surfaces are highly inhomogeneous, consisting of land, water, ice, or snow whose emissivities are significantly different [Weng et al., 2001; Weng and Yan, 2004]. Land surfaces have an emissivity close to 0.95 (note that surfaces with snow or sand have lower emissivity at these frequencies); ocean emissivity varies considerably depending on oceanic characteristics including surface roughness which is influenced by overlying atmospheric conditions; and snow and sea-ice emissivity also varies considerably [Mathew et al., 2008]. Therefore measures are necessary to reduce the noise related to surface inhomogeneity.

[6] Furthermore, near-polar SNOs only sample brightness temperatures which are not representative of lower latitudes. Owing to nonlinearity in the calibration, error in warm target measurements, and obstructed space view, intersatellite biases can vary with scene radiance [e.g., Zou et al., 2006]. Therefore there are several reasons why near-polar SNOs are inadequate for intercalibrating the microwave humidity sounders.

[7] Owing to atmospheric drag, the Earth’s nonsphericity, and gravitational pull from celestial bodies, the orbit of a polar orbiting satellite drifts and its local equator crossing time changes. When the equator crossing times of a pair of satellites become nearly the same, SNOs can occur at all latitudes for a short period, typically 1 or 2 months. We use these SNOs at all latitudes to estimate the adequacy of polar SNOs to intercalibrate microwave humidity sounders.

[8] Section 2 gives a short technical description of the humidity sounders and their channel characteristics. Section 3 revisits a recent comparison between simulated AMSU-B and MHS to show that global SNOs are necessary for reliable intercalibration. Section 4 describes the data and methods of analysis. Section 5 presents the results and section 6 provides a summary and conclusions.

2. Functional Description of Microwave Humidity Sounders

[9] The Advanced Microwave Sounding Unit-B (AMSU-B) and the Microwave Humidity Sounder (MHS) are five-channel microwave radiometers. They are designed to measure the radiation emitted from the Earth’s surface and atmosphere in order to estimate global fields of tropospheric humidity. The microwave absorption characteristics of the atmosphere are shown in Figure 1 and the instrument specifications are given in Table 1. AMSU-B is onboard NOAA-15 (N15), N16, and N17 and MHS is onboard N18, N19, and MetOpA (MA).

[10] Channels 1 and 2 at 89 GHz and 150 GHz (at 157 GHz for MHS), enable deeper penetration through the atmosphere to the Earth’s surface. Channels 3–5 are located in the strongly opaque water vapor absorption line at 183.31 GHz and provide information on the atmospheric humidity at different levels. The passbands of Channels 3, 4, and 5 are at 183.31 ± 1.00 GHz, 183.31 ± 3.00 GHz, and 183.31 ± 7.00 GHz (only at 183.31 + 7.00 GHz for MHS), respectively. The passbands of the channels are also shown in Figure 1. Note that the five channels on AMSU-B are formally numbered as channels 16–20 (channels 1–15 belong to AMSU-A which is a temperature sounding instrument), but in this article we call them channels 1–5 to be consistent with MHS channel numbers.
At each channel frequency, the antenna beamwidth is a constant 1.1 degrees (full width at half maximum). Ninety contiguous cells are sampled on the Earth’s surface, with each scan covering ±49.5 degrees on each side of the sub-satellite point. These scan patterns and geometric resolution translate to a 16.3 km diameter cell at nadir at a nominal altitude of ~833 km. Each channel is also sensitive to radiation of a particular polarization as defined in Table 1, the direction of which rotates with scan angle. The differences in the polarization for channels 3 and 4 on MHS compared to AMSU-B will only manifest themselves for very dry atmospheres (or high topography) where these channels become sensitive to surface radiation such as over Antarctica.

### Table 1. Channel Characteristics of the Instruments

<table>
<thead>
<tr>
<th>Channel</th>
<th>(f_c) (GHz)</th>
<th>(\Delta f) (GHz)</th>
<th>Passbands</th>
<th>NE(\Delta T) (K)</th>
<th>Beam Width (deg)</th>
<th>Polarization</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMSU-B</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>89.0</td>
<td>0.5</td>
<td>2</td>
<td>0.37</td>
<td>1.1</td>
<td>V</td>
</tr>
<tr>
<td>2</td>
<td>150.0</td>
<td>1.0</td>
<td>2</td>
<td>0.84</td>
<td>1.1</td>
<td>V</td>
</tr>
<tr>
<td>3</td>
<td>183.3±1.0</td>
<td>0.5</td>
<td>2</td>
<td>1.06</td>
<td>1.1</td>
<td>V</td>
</tr>
<tr>
<td>4</td>
<td>183.3±3.0</td>
<td>1.0</td>
<td>2</td>
<td>0.70</td>
<td>1.1</td>
<td>V</td>
</tr>
<tr>
<td>5</td>
<td>183.3±7.0</td>
<td>2.0</td>
<td>2</td>
<td>0.60</td>
<td>1.1</td>
<td>V</td>
</tr>
<tr>
<td>MHS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>89.0</td>
<td>0.5</td>
<td>2</td>
<td>0.22</td>
<td>1.1</td>
<td>V</td>
</tr>
<tr>
<td>2</td>
<td>157.0</td>
<td>1.0</td>
<td>2</td>
<td>0.34</td>
<td>1.1</td>
<td>V</td>
</tr>
<tr>
<td>3</td>
<td>183.3±1.0</td>
<td>0.5</td>
<td>2</td>
<td>0.51</td>
<td>1.1</td>
<td>H</td>
</tr>
<tr>
<td>4</td>
<td>183.3±3.0</td>
<td>1.0</td>
<td>2</td>
<td>0.40</td>
<td>1.1</td>
<td>H</td>
</tr>
<tr>
<td>5</td>
<td>190.3</td>
<td>2.0</td>
<td>1</td>
<td>0.46</td>
<td>1.1</td>
<td>V</td>
</tr>
</tbody>
</table>

*Here \(f_c\) is the central frequency of the channel (taken from Kleespies and Watts [2007]). \(\Delta f\) is the passband width, NE\(\Delta T\) is the noise equivalent temperature from the first flight models (NOAA KLM User’s guide [Goodrum et al., 2007]). Nominal polarizations are for nadir view only and rotate with view angle.

3. Comparison of Simulated AMSU-B and MHS Measurements

Kleespies and Watts [2007] compared the brightness temperatures simulated for MHS and AMSU-B using the 48 profiles of Strow et al. [2003]. Significant differences were found only for channels 2 and 5 and in both cases mean MHS brightness temperatures were colder than those of AMSU-B. We revisit the study to investigate the dependence of bias between the two instruments on surface and atmospheric conditions, enabling us to interpret the results of our SNO analysis for these channels.

Figure 2 shows simulated brightness temperature differences between MHS and AMSU-B using 5000 diverse profiles, sampled from ECMWF forecasts to span the natural variability of the real atmosphere [Chevallier et al., 2006]. The simulations used a line-by-line radiative transfer model [Buehler et al., 2005a] that was already used in a number of intercomparison studies [Buehler et al., 2004; John and Buehler, 2005; Moradi et al., 2010]. Surface emissivity at these frequencies varies considerably with surface type, with higher emissivity for land (~0.95) and lower emissivity for ocean (~0.6). Therefore in the simulations we used an emissivity of 0.95 for land profiles and 0.6 for ocean profiles. For profiles from a model grid box which has both land and sea we calculated the emissivity based on land cover linearly varying between 0.6 and 0.95. Results are shown only for channels 2 and 5 because for the other three channels the differences are negligibly small. The differences are shown as functions of brightness temperature (Figure 2, right) and precipitable water vapor (PWV; Figure 2, left) which is the vertically integrated water vapor density.

Figure 2 (top) shows the differences for channel 2 which is at 150 GHz on AMSU-B but at 157 GHz on MHS. Channel 2 is a sounding channel in a humid atmosphere but with a surface contribution which increases as atmospheric moisture decreases. With a very moist atmosphere, the surface has little effect and the brightness temperature of MHS is lower than that of AMSU-B because the atmosphere is slightly more opaque at 157 GHz than at 150 GHz (see Figure 1), raising the sounding altitude slightly. With a less moist atmosphere, the higher atmospheric opacity at 157 GHz than at 150 GHz makes the radiometrically cold surface have less influence on MHS than on AMSU-B, leading to higher brightness temperature for MHS. This is especially true for the ocean (blue symbols in Figure 2) because of its lower emissivity. When there is very little water vapor, the difference is close to zero because both instruments sample the surface which has a similar emissivity at 150 and 157 GHz. It is clear from Figure 2 that the differences can have a wide range from ~2 to 7 K depending on atmospheric and surface conditions (Kleespies and Watts [2007] reported ~1.54 ± 2.03 K bias) and thus it is not straightforward to combine AMSU-B and MHS channel 2 radiances by adding an offset to one of the measurements.

Figure 2 (bottom) shows the brightness temperature difference for channel 5. As for channel 2, there is little difference when the atmosphere is almost free of water vapor and it starts increasing with water vapor. When precipitable water is around 3 mm the trade off between surface and sounding channel effects come into play and the difference starts to decrease. When there is about 15 mm of precipitable water the channel becomes a sounding channel and insensitive to the surface. MHS channel 5 is measuring colder radiances compared to the AMSU-B one, by about ~0.6 K irrespective of surface type in an atmosphere with 20 mm or more precipitable water vapor.

Figure 2 (right) show simulated brightness temperature differences for channels 2 and 5 as a function of average scene brightness temperatures of AMSU-B and MHS. Transition from surface to sounding channel is clearly seen for sea points due to radiometrically colder surface which amplifies the water vapor signal from the atmosphere. Although it is possible to combine AMSU-B and MHS data...
for channel 5 [Kleespies and Watts, 2007] by adding a global offset to account for the frequency changes, the systematic major variations in bias between a dry polar atmosphere and a moist lower-latitude atmosphere lead us to conclude that biases from these channels estimated from polar SNOs cannot represent humid lower latitudes.

4. Collocation and Analysis Methods

[17] We used the collocation method based on Holl et al. [2010] (some details given in Appendix A). Sensitivity of distance and time thresholds for selecting collocations to the uncertainty in the estimated bias is shown in Figure 3. Consequently, to overcome spatial inhomogeneity we used only those pixel pairs whose centers are closer than 5 km, which is less than one third of the 16.3 km pixel diameter at nadir. We discard any measurements with time differences exceeding 300 s, to avoid changes in scene properties such as clouds. We used only four pixels each on either side of nadir, to avoid errors arising from limb effects and scan asymmetry. This also minimizes the impact of polarization differences. Collocations over both land and ocean are used throughout this study.

[18] Using those pixel pairs which satisfied these stringent criteria, we first calculated differences in brightness temperatures and then derived the mean difference or bias ($\Delta T_B$) and the standard deviation of the differences ($\sigma_{\Delta T_B}$). The standard deviation of collocated brightness temperatures (or in other words, SNO variability) has mainly two sources [Iacovazzi and Cao, 2008]: one is the combined radiometric noise ($NE\Delta T$, which is the smallest change in input brightness temperature that can be detected in the system output (i.e., calibrated brightness temperatures, including contributions from calibration noise)) of the two instruments and the other is the scene (spatial and temporal) inhomogeneity. In order to have robust statistics, we collected data for a month to calculate $\Delta T_B$ and $\sigma_{\Delta T_B}$. This is an advance over previous studies which consider individual SNO events which have fewer pixel pairs for computing statistics. We also calculate standard errors of mean values, namely $\sigma_{\Delta T_B}$ divided by the square root of the number of collocations.

[19] Clouds affect these channels [Sreerekha et al., 2008], but we have not screened for them. This is mainly because in polar conditions it is difficult to differentiate between clouds and the surface. Owing to our stringent spatiotemporal collocation criteria, we assume that measurements from both instruments are affected by clouds in a similar way.

5. Results

5.1. Selection of SNOs

[20] Figure 4 shows the equator crossing times of the ascending nodes [Ignatov et al., 2004] of NOAA and MetOp polar-orbiting satellites. The orbital parameters of these satellites are designed so that equator crossing time will drift away from local noon because if a satellite crosses the equator at noon, this can affect the functioning of both the satellite and the instruments on board owing to different solar illumination. The drift creates the possibility that

Figure 2. (left) Simulated brightness temperature differences between MHS and AMSU-B as a function of precipitable water vapor for (top) Channel 2 and (bottom) Channel 5 using a diverse atmospheric profile data set compiled from ECMWF forecasts [Chevallier et al., 2006]. Profiles are separated for ocean and land. Ocean emissivity is 0.6, land emissivity is 0.95, and emissivity of mixed grid point profiles varies linearly between 0.6 and 0.95. (right) Also shown is simulated brightness temperature differences as a function of mean brightness temperatures of AMSU-B and MHS. Note the simulations are only for nadir.
satellite pairs will have similar equator crossing times for short periods. During these time periods SNOs can occur globally. We have discovered that in recent years there have been SNOs at all latitudes and this is to our knowledge the first study to exploit this.

We have identified 4 months of data with sufficient number of collocations satisfying our stringent criteria (Δx less than 5 km and Δt less than 300 s) at all latitudes. They are 2008-08 for the N16-N15 pair, 2009-04 and 2009-05 for the MA-N17, and 2009-09 for the N19-N18. We assign newer platforms NOAA-16, MetOpA, and NOAA-19 as primary satellites and NOAA-15, NOAA-17, and NOAA-18 as secondary satellites. Bias is calculated as primary satellite minus secondary satellite.

We partitioned the collocations into 18 10° latitude bins. Figure 5 (top) shows the latitudinal distribution of the number of collocations. The number of collocations varies for each satellite pair, with the N19-N18 pair having the most, about 50,000–100,000, collocations in each latitude bin.

![Image showing sensitivity test to select distance and time threshold for collocations.](image)

**Figure 3.** Sensitivity test to select distance and time threshold for collocations. Standard deviation of brightness temperature difference in Kelvin for each grid box is shown. Distance grid is equally spaced with 1 km distance, but time grid has variable width. We have randomly selected 120 points from each grid box to calculate statistics. Grid boxes in white have too few collocations to make statistics. Collocations are taken from 70–80° latitudes in both hemispheres of MA–N17 collocations.
**Figure 4.** Equator crossing times of the ascending nodes of NOAA/MetOpA polar orbiting satellites for the ATOVS time period. Drifting of the orbits can be seen. MetOpA is maintained in a stable orbit.

**Figure 5.** Number of collocations in 10° latitude bins. Each collocation satisfies stringent spatiotemporal criteria ($\Delta x < 5$ km and $\Delta t < 300$ sec). Black circles show collocations of NOAA-15 and NOAA-16 during 2008-08, green and blue circles show collocations of MetOpA and NOAA-17 during 2009-04 and 2009-05, respectively, and red circles show the collocations of NOAA-19 and NOAA-18 during 2009-09. Note the logarithmic y-axis scale. Map plots show geographical distribution of SNOs.
bin. Most of the bins have 5000 or more collocations which are enough collocations to compute robust statistics.

5.2. Interpretation of SNOs

[23] Biases are expected to vary with scene-radiance (section 3), so estimates of biases derived from SNOs at all latitudes will be particularly valuable if they vary systematically with latitude. On the other hand, the estimates will be less useful if they are noisy. So before presenting our main results we consider these two aspects.

5.2.1. Meridional Distribution of Brightness Temperature

[24] In order to interpret the meridional distribution of biases in the measurements, we need to know the meridional distribution of the brightness temperatures. Figure 6 shows the mean and standard deviation of brightness temperatures for each latitude bin for each channel. A common feature is the very low brightness temperatures south of 70°S. Variability is greater over these southern polar regions because the heterogeneous surface conditions show through the very dry atmosphere. Channel 1 is a surface channel at all latitudes: as seen in Figure 1, the total opacity is less than one even for the very wet profile. Accordingly it also shows low brightness temperatures for the midlatitude southern hemisphere and for the Arctic, as does channel 2 for the same reason. This might be associated with less landmass in these latitudes and lower ocean emissivity. Because of its sensitivity to the surface, channel 1 also has high variability.

5.2.2. Uncertainties in SNO Method

[25] An important source of uncertainty for the SNO method is the radiometric noise of the instruments [Iacovazzi and Cao, 2008]. This is normally expressed as noise equivalent brightness temperature (NE\(D\)T): the first flight model values taken from Goodrum et al. [2007] for each channel are given in Table 1. NE\(D\)T is time varying, it generally increases as the instrument starts to degrade. It can also increase due to a change in the operating conditions of the satellite and due to radio frequency interference (RFI) from nearby transmitters or other instruments.

[26] Mean NE\(D\)T values for the analysis time period for all the channels are given in Table 2. NE\(D\)T were determined from the warm target views during the analysis time period. Note that for N19 channel 3 the noise was about 2.5 K for the first half of September 2009 but more than 7 K
for the second half of the month. Note the performance of this channel became better by the beginning of 2011.

Scene inhomogeneity is another source of uncertainty in the SNO method owing to spatial and temporal mismatches in collocated pixels. Figure 7 shows standard deviations of brightness temperature differences reflecting these uncertainties. Horizontal lines indicate the combined instrument noise based on values given in Table 2. Channel 1 shows standard deviations from 1 to 2 K which are higher than the combined instrument noise. For channel 2, N16–N15 (both having AMSU-B) and N19–N18 (both having MHS) show standard deviations from 1 to 1.5 K, approximately consistent with the specified NEΔT of these instruments. The MA–N17 pair (MA has MHS and N17 has AMSU-B) has higher standard deviation which can be explained by the differences between 150 and 157 GHz emissions for very different surface emissivities (land and sea) north of 40°S (see discussions in section 3).

Channel 3 shows different standard deviations for different satellite pairs. If NEΔT accords with prelaunch specifications (Table 1), we would expect the effective variability associated with NEΔT to be equivalent to 1.5 K for the AMSU-B–AMSU-B combination, 1.18 K for the AMSU-B–MHS combination (\(\sqrt{1.06^2 + 0.51^2} = 1.18\) K), and 0.7 K for the MHS–MHS combination. The MA–N17 pair comparing AMSU-B and MHS shows SNO variability close to instrument specification; even at high latitudes, where the surface is highly variable, standard deviations remain small. Thus there is very little contribution from scene inhomogeneity, given our stringent collocation criteria. The other two satellite pairs show significantly higher variability than expected from prelaunch instrument specifications. N19–N18 has the highest value of about 9 K, owing to known instrument problems causing exceptionally high noise in channel 3 on N19 at the time of the comparison. N16–N15 has about 4 K standard deviation which is also much higher than the specified noise of the instruments. Our analysis (Table 2) indicates that for both N16 and N15 the NEΔT have increased due to instrument problems.

<table>
<thead>
<tr>
<th>Channel</th>
<th>2008-08</th>
<th>2009-04/05</th>
<th>2009-09</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.33</td>
<td>0.40</td>
<td>0.22</td>
</tr>
<tr>
<td>2</td>
<td>0.71</td>
<td>0.49</td>
<td>0.37</td>
</tr>
<tr>
<td>3</td>
<td>2.32</td>
<td>1.00</td>
<td>0.52</td>
</tr>
<tr>
<td>4</td>
<td>1.38</td>
<td>0.72</td>
<td>0.41</td>
</tr>
<tr>
<td>5</td>
<td>1.10</td>
<td>0.82</td>
<td>0.36</td>
</tr>
</tbody>
</table>

*Units are in Kelvin.*

![Figure 7](image-url)
Channels 4 and 5 also show standard deviations consistent with NED<sub>T</sub> of the instruments, except for the N16–N15 pair which shows inflated standard deviations owing to instrument degradation (Table 2). The standard deviations are invariant with latitude for these channels as well which leads to the conclusion that there is little influence by scene inhomogeneity in the estimated bias, given our stringent collocation criteria.

5.3. Meridional Distribution of Bias

Figure 8 shows the bias and its standard error for all latitude bins and for all satellite pairs. Intersatellite bias is time varying, so the results shown here represent only the time period analysed.

5.3.1. Channel 1 (89 GHz)

Channel 1 shows very small intersatellite biases. N16 is about 0.15 K warmer than N15 for most latitude bins. MA measurements are also warmer than N17 measurements by about 0.2 K but with a few outliers. Though N19–N18 bias is small there is a latitude dependence, with negative bias for high latitudes and positive bias for low latitudes.

5.3.2. Channel 2 (150/157 GHz)

N16–N15 bias is very stable across all latitudes at about 0.5 K except for the two southernmost latitude bins. N19–N18 bias is clearly latitude-dependent, being as low as −0.3 K at high southern latitudes and 0.1 K at low latitudes, with a pattern similar to channel 1. The MA-N17 pair (AMSU-B and MHS combination) shows large biases, up to 4 K and high variability in bias with latitude, as expected from our analysis of simulated brightness temperatures for this channel in section 3. Biases are consistent for April and May 2009.

5.3.3. Channel 3 (183.3 GHz)

N16–N15 shows the largest biases, ranging from 1 to 2 K with a latitude dependence: 1.8 K bias near the South Pole which decreases to 1 K near the North Pole. N19–N18 has biases ranging from zero to 0.4 K with no obvious latitude dependence, but with large standard error. Note that N19 had exceptionally high noise for this channel. MA-N17 biases vary between −0.15 K to 0.4 K, being positive at high latitudes and near-zero or negative at low latitudes.

Figure 8. Black circles and vertical bars show bias (Δ<sub>Tb</sub>) for latitude bins and its standard error estimated using SNOs. First row: NOAA-19–NOAA-18 during 2009-09. Second and third row: MetOpA–NOAA-17 during 2009-05 and 2009-04, respectively. Fourth row: NOAA-16–NOAA-15 for 2008-08. Red circles show bias estimated from zonal mean brightness temperatures (see section 5.3.6 for details). Note that some of the red circles are out of the plot range.
5.3.4. Channel 4 (183.31 ± 3.00 GHz)
[34] N16-N15 has significant intersatellite bias which varies linearly from 4 K at the South Pole to near zero at the North Pole. N19–N18 biases are close to 0.2 K at low latitudes and the biases vary considerably with latitude. MA–N17 also show latitude dependence with high latitudes showing positive biases up to 0.5 K whereas at low latitudes the biases are very slightly negative.

5.3.5. Channel 5 (183.31 ± 7.00/+7.00 GHz)
[35] N16–N15 biases also vary linearly with latitude for channel 5, from ~3 K near the South Pole to ~5.5 K near the North Pole. N19–N18 biases are close to zero except for the two southernmost latitude bins where the bias is close to ~0.3 K. MA-N17 has the AMSU-B–MHS combination and this channel on MHS has only the upper sideband, so larger biases are expected. The biases show strong latitude dependence: positive at higher latitudes and near-zero or negative at lower latitudes.

5.4. Dependence of Bias on Scene Radiance
It is not certain whether it will work for other time periods if the method is found to be useful for the analyzed time period, in general except for channel 1 due to very small biases for estimated biases using this method. This method works well as in our SNO analysis. Red circles in Figure 8 show the estimated biases using this method. This method works well in general except for channel 1 due to very small biases for this channel. Thus it is confirmed that the latitude dependence of bias estimated using SNO method is correct and polar SNOs alone cannot be used to estimate intersatellite biases. Though the zonal average brightness temperature method is found to be useful for the analyzed time period, it is not certain whether it will work for other time periods due to differences in temporal sampling of the satellites and we are currently investigating this.

5.4.1. Channel 1
[39] N16–N15 bias tends to decrease with increasing scene radiance for N15 but not for N16. MA–N17 biases also tend to decrease with N17 T_B but increase with MA T_B. N19–N18 shows a strong increase in bias with both T_B.

5.4.2. Channel 2
[40] For both N16–N15 and N19–N18 biases generally increase with T_B of the primary satellites. MA-N17 has peak biases at a T_B of about 250 K for either satellite which is exactly what is expected based on our discussion in section 3 using simulated radiances (see Figure 2).

5.4.3. Channel 3
[41] For channel 3, N19–N18 bias varies with N19 T_B, from about ~20 K at 160 K to 45 K at 310 K. This clearly indicates the instrument problem for this channel. This apparent large bias can be explained by the large noise of N19. The range of brightness temperature is larger for N19 due to higher noise. This in turn will lead to a negative bias for colder N19 T_B bins and to a positive bias for warmer N19 T_B bins. N16–N15 bias also shows dependence on T_B of both satellites. The bias starts increasing with N16 T_B and then stays constant from 170 K to 230 K and then dips before increasing again. Note that T_B below 230 K are mostly from the two southernmost bins where the channel is a window channel. The bias varies from about ~2 K to 7 or 8 K with N16 T_B above 230 K. Bias seems to decrease with N15 T_B (3 K to ~1 K), so the overall bias is reduced due to contrasting bias dependence on N15 and N16 T_B. MA–N17 biases generally decrease with both T_Bs.

5.4.4. Channel 4
[42] Channel 4 bias patterns are similar to those of channel 3 for MA–N17. N19–N18 bias increases with N19 T_B but does not show any clear relationship with N18 T_B. N16–N15 bias stays constant with both T_Bs below 240 K and then starts to decrease up to 260 K. The bias then starts to increase with N16 T_B but continues to decrease with N15 T_B.

5.4.5. Channel 5
[43] Channel 5 on N19–N18 shows a strong increase in bias with both T_B, ~0.5 K at 160 K and increasing to ~0.1 K at 300 K. MA–N17 pair shows larger biases due to the frequency difference as discussed in section 3 with biases increasing with T_B and then starts decreasing when the channel becomes a sounding channel. N16–N15 biases show a very strong dependence on N15 T_B, ~6 K at 250 K and 0.5 K at 300 K.

5.4.6. Explaining Latitude Dependence of Bias
[44] The latitude dependence of intersatellite biases can be explained by their dependence on scene radiance. For example, the N19–N18 pair shows rather monotonically increasing bias with increasing T_B for all channels. This leads to similar meridional distribution of T_B and bias, except for channel 3 due to the large noise of N19. Another example is channels 3–5 of MA–N17 which show decreasing bias with increasing T_B, and thus shows opposite meridional patterns for bias and T_B.

5.5. Consistency Check on Estimated Bias
As an independent estimate to check the bias obtained from SNO method, biases were also calculated using zonal averaged brightness temperatures. This method was already used by Shi and Bates [2011] for infrared channels. We used only near-nadir brightness temperatures to avoid scan bias/ limb effect. This method works well when the sampling times of two satellites are similar which is the case of our analysis. If sampling times were different, differences would arise from the diurnal cycles of humidity and temperature [Zou et al., 2006]. Biases are calculated for 18 latitude bins as in our SNO analysis. Red circles in Figure 8 show the estimated biases using this method. This method works well in general except for channel 1 due to very small biases for this channel. Thus it is confirmed that the latitude dependence of bias estimated using SNO method is correct and polar SNOs alone cannot be used to estimate intersatellite biases.

5.6. Summary and Conclusions
Cao et al. [2004, 2005] have shown that the simultaneous nadir overpass (SNO) method is useful for inter-calibrating satellite instruments. Nevertheless, Iacovazzi and Cao [2008] expressed concerns over using SNOs for surface sensitive channels. Owing to their normal occurrence in the polar regions, SNOs have potential problems for their use in...
intercalibrating microwave humidity sounding channels which are surface sensitive under dry polar atmospheric conditions. But as a result of orbital drift, SNOs can occur globally for a short period of time for polar orbiting satellite pairs when their local equator crossing times become close. We used these global SNOs to evaluate intercalibration using only polar SNOs.

There are three satellite pairs with global SNOs for microwave humidity sounders. They are NOAA-16(N16)–NOAA15(N15) during 2008-08, MetOpA(MA)–NOAA-17(N17) during 2009-04 and 2009-05, and NOAA-19(N19)–NOAA-18(N18) during 2009-09. N15, N16, and N17 have the Advanced Microwave Sounding Unit-B (AMSU-B) and N18, N19, and MA have the Microwave Humidity Sounder (MHS). We have shown using simulated brightness temperatures that the differences for these channels between AMSU-B and MHS are dependent on the amount of water vapor in the atmosphere and on the scene radiance. The differences for channel 2 ranges between −2 and 7 K and for channel 5 from −1 to 3 K, but for other channels the differences are negligible.

The method used to obtain collocations is based on Holl et al. [2010]. We used only those collocations with spatial differences less than 5 km and temporal differences less than 300 s, based on sensitivity analyses, to avoid uncertainties due to scene inhomogeneities. We then partitioned the collocated measurements into 18 10° bins. All channels show a large range (100 K) in brightness temperature across the latitudes with coldest brightness temperature near the South Pole and the warmest in the tropics. The main source of uncertainty in the SNO method is the combined radiometric noise of the instrument pair. The standard deviations of brightness temperature differences are invariant with latitude indicating that scene inhomogeneities play only a minimal role, given our stringent collocation criteria. The sounding channels (channels 3–5) show different values of standard deviations across the satellite pairs which is consistent with their radiometric noise. For example, the largest standard deviation of about 9 K is shown by N19–N18 pair for channel 3 owing to the anomalously large noise of N19.

Channel 1 generally shows small intersatellite biases and less latitude dependence compared to other channels.

Channel 2 has higher bias (up to 3.5 K) for the AMSU-B–MHS combination which is consistent with the results based on simulated radiances shown in section 3. N19–N18 shows a strong latitude dependence for biases in channel 2. N16–N15 shows the largest biases for channels 3, 4, and 5 and also shows a strong latitude dependence. We have validated the biases estimated from global SNOs by biases estimated from zonal mean near-nadir brightness temperatures. We suggest that it is not appropriate to use SNOs over a restricted latitude range to intercalibrate humidity sounders. The reason for the latitude dependence of biases primarily originates from their dependence on scene radiance which themselves have a latitude dependence. It was shown that the dependence of biases on one satellite could be different from another. Channel 3 of N16–N15 shows this behavior with biases increasing with N16 brightness temperatures and decreasing with N15 brightness temperatures. We suggest that it is important to take into account the dependence of biases on scene radiance during the intercalibration procedure.

It has to be kept in mind that the present study explores the global SNOs which are available only for a short time during the life of satellites and thus cannot be used to estimate temporal evolution of bias [e.g., Zou et al., 2006]. Another method for intercalibration that is being developed is monitoring of satellite radiometer biases using NWP fields (R. W. Saunders et al., Monitoring satellite radiometer biases using NWP fields, manuscript in preparation, 2012) which allows global sampling during the entire life time of the satellites. Our plan is to combine the SNO method with the NWP method to intercalibrate microwave humidity sounders.

This work is being done as part of a project to homogenize radiances measured by microwave humidity sounders. The next step is to include SSM/T-2 data as well in our analyses. Intersatellite biases are generally estimated using only near-nadir radiances. To apply these bias estimates to measurements at other viewing angles requires that there are no scan-dependent biases, but in reality there are scan-dependent biases [e.g., Buehler et al., 2005b]. These asymmetries will also be estimated as part of this project.

Appendix A: Collocation Methodology

We used the collocation code from Holl et al. [2010]. This code is designed for any pair of satellite sensors and not specifically designed for the study performed here. However, the selection of data fulfilling spatial criteria was modified to improve performance and is different from the algorithm described by Holl et al. [2010]. The first steps are the same: orbits with temporal overlap are located, and then the segments within those pairs of orbits that have a time overlap (plus or minus the maximum time for a collocation) are located. This is described in detail by Holl et al. [2010]. In this study, all individual measurements are binned according to their latitude/longitude values, resulting in two “gridded” datasets for this orbital segment. For all grid cells where one sensor has measurements and the other sensor has measurements in a nearby grid cell, all time differences between measurements from the one and the other sensor are calculated. Here, “nearby” is a function of cell size and maximum collocation distance. The number of neighboring cells to explore for collocations is chosen such that no collocations can be missed. For example, if the maximum collocation distance is 15 km and cells are $1^\circ \times 1^\circ$, a cell at 85°N is only 9.7 km wide (note that most satellites do not reach so close to the pole). In order not to miss any collocations, this means that measurements from a cell centering at (0°, 85°N) are compared to all measurements in the 15 cells spanning from (2°W, 86°N) to (2°E, 84°N). However, measurements from a cell at (0°, 45°N), where the cell is 79 km wide, need to be compared only within the nine cells spanning from (1°W, 46°N) to (1°E, 44°N). If this temporal criterion is also met, the collocation is selected for further processing.

Acknowledgments. We thank three anonymous reviewers, Richard Allan, and John Eyre, Nigel Atkinson, Marc Shroder, and Joerg Shulz for valuable comments. Viju John and David Parker were supported by the Joint DECC/Defra Met Office Hadley Centre Climate Programme (GA01101) and Viju John was also supported by the UK JWCRP. This work contributes to COST Action ES0604-Water Vapor in the Climate System (WaVaCS). Thanks to Lisa Neclos of the NOAA CLASS for AMSU-B and MHS Level-1b data, EUMETSAT NWP-SAF for the AAPP software to process the data, and ARTS community for their radiative transfer model.

References


S. A. Buehler and G. Holl, Department of Space Science, Luleå University of Technology, SE-97187 Kiruna, Sweden.
B. Candy and R. W. Saunders, Met Office, Exeter EX1 3PB, UK.
V. O. John and D. E. Parker, Met Office Hadley Centre, Exeter EX1 3PB, UK.

(viju.john@metoffice.gov.uk)
Paper III

Optimised frequency grids for infrared radiative transfer simulations in cloudy conditions

Optimised frequency grids for infrared radiative transfer simulations in cloudy conditions

G. Holl, S.A. Buehler, J. Mendrok, A. Kottayil

Division of Space Technology, Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Rymdcampus 1, 98128 Kiruna, Sweden

A B S T R A C T

This paper shows that radiometer channel radiances for cloudy atmospheric conditions can be simulated with an optimised frequency grid derived under clear-sky conditions. A new clear-sky optimised grid is derived for AVHRR channel 5 ($12 \mu m, 833 cm^{-1}$). For HIRS channel 11 ($7.33 \mu m, 1364 cm^{-1}$) and AVHRR channel 5, radiative transfer simulations using an optimised frequency grid are compared with simulations using a reference grid, where the optimised grid has roughly 100–1000 times less frequencies than the full grid. The root mean square error between the optimised and the reference simulation is found to be less than 0.3 K for both comparisons, with the magnitude of the bias less than 0.03 K. The simulations have been carried out with the radiative transfer model Atmospheric Radiative Transfer Simulator (ARTS), version 2, using a backward Monte Carlo module for the treatment of clouds. With this module, the optimised simulations are more than 10 times faster than the reference simulations. Although the number of photons is the same, the smaller number of frequencies reduces the overhead for preparing the optical properties for each frequency. With deterministic scattering solvers, the relative decrease in runtime would be even more. The results allow for new radiative transfer applications, such as the development of new retrievals, because it becomes much quicker to carry out a large number of simulations. The conclusions are applicable to any downlooking infrared radiometer.

© 2012 Elsevier Ltd. All rights reserved.

1. Introduction

Radiative transfer modelling in cloudy atmospheres is an important tool for the development of satellite-based retrievals of cloud properties. Improved retrievals of properties such as Ice Water Path (IWP) are needed, because (ice) clouds are a major source of uncertainty in climate models (e.g. [30, Section 8.6.3.2]). For infrared radiation, a simple simulation of an instrument radiance is computationally expensive, because it requires thousands of monochromatic calculations. However, such a simulation is potentially more accurate than a simulation using band models, because band models include spectral averaging assumptions and parameterise optical properties (e.g. [25,31]). This reduces the flexibility of the model. Therefore, there is a need for a way to reduce calculation times, without compromising on accuracy or flexibility. This is particularly true for cloudy simulations, because a single monochromatic simulation is computationally much more expensive if cloud scattering is taken into consideration.

One way to reduce calculation times is to use an optimised frequency grid. By an optimised frequency grid, we mean a frequency grid that has significantly less (typically 100 to 1000 times less) frequencies than the full reference grid. The derivation of such an optimised
grid for infrared radiances is non-trivial, because it depends on atmospheric composition and viewing geometry, and because the optimisation is highly non-linear (i.e. there are many local minima that are not close to the global minimum). One method to derive an optimised grid is the correlated k method [18]. Another method is based on simulated annealing [6]. The simulated annealing approach is better suited to find a frequency grid that works well for a large variety of atmospheres. The grids used in this study are derived using simulated annealing, but the conclusions also hold for grids derived using other methods.

The primary aim of this study is to show that those clear-sky derived optimised frequency grids can be used for the simulation of cloudy radiances and that calculation times are significantly reduced. Although the spectral dependence of cloud optical properties is much smaller than the spectral dependence of gas absorption properties (as illustrated at the end of this paper), it is not self-evident that a clear-sky derived grid can be used for cloudy simulations. If the signal measured by a satelliteborne sensor is fully dominated by a cloud, the spectrum is flat and the exact choice of an optimised grid is not crucial. This is the situation for a high, thick cloud in a window channel. On the other hand, if the signal is a mixture between a cloudy and a clear-sky one, the situation may be more complicated. For example, if the cloud top is located as the same height as the peak of a broad weighting function for a particular channel, the atmospheric region from which the clear-sky signal originates is affected by the cloud. This means that the absorption due to different gases may be affected in different ways and that their relative contribution to the measured signal may change, which in turn has consequences for the most favourable weighting in an optimised frequency grid. Therefore, it is worth investigating whether the clear-sky derived grid is always a suitable choice for a simulation in the presence of clouds.

The secondary aims of the study are: (1) present a Advanced Very High Resolution Radiometer (AVHRR) setup for the radiative transfer model Atmospheric Radiative Transfer Simulator (ARTS), and (2) present the first cloudy infrared radiometer simulations using ARTS (infrared limb spectra were calculated with ARTS by Hoepfner and Emde [19]).

We focus on High-resolution Infrared Radiation Sounder (HIRS) and AVHRR, because those are sensors on operational satellites. Therefore, they allow for analysis of relatively long time series and they have a dense spatial and temporal coverage. However, the conclusions are valid of any downlooking infrared radiometer.

For cloudy simulations, we use a backward Monte Carlo model [12]. Nevertheless, the conclusions are applicable to any radiative transfer solver. The present study serves as a demonstration of the method.

More discussion on the validity of the conclusions follows later in the paper.

The structure of the paper is as follows. In Section 2, we describe the sensors, the radiative transfer model and the data used in the study. In Section 3, we describe the experiments we have carried out to test if the optimised grid derived using clear-sky simulations is valid for cloudy simulations as well. The results of those experiments are presented in Section 4 and discussed in Section 5. We finish in Section 6 by concluding remarks and a discussion of future work.

2. Methodology

2.1. Sensors

AVHRR is a downlooking imager in the visible/infrared that has been flown on polar-orbiting satellites from National Oceanic and Atmospheric Administration (NOAA) and MetOp since the late 1970s. It has a wide variety of atmospheric and non-atmospheric applications [11,24]. On the latest generation, AVHRR/3, channels 4 (10.8 μm, 926 cm⁻¹) and 5 (12 μm, 833 cm⁻¹) are window channels in the thermal infrared.

HIRS is a downlooking infrared sounder for temperature and humidity [33], carried on many operational satellites. It carries 20 channels ranging from the visible up to wavelengths of 15 μm (667 cm⁻¹).

In this study, we use a window channel and a sounding channel, in order to test our method under different conditions. For the window channel, we use AVHRR channel 5. For the sounding channel, we use HIRS channel 11, centred at 7.33 μm (1364 cm⁻¹) and with a Jacobian peaking in the mid-troposphere (varying in height depending on atmospheric humidity).

2.2. Radiative transfer simulations

2.2.1. ARTS

ARTS is a flexible and powerful model for the simulation of radiances at thermal infrared, sub-millimeter and microwave frequencies [3,16]. ARTS can perform radiative transfer simulations in one, two, or three dimensions and can consider full polarisation [13]. It performs monochromatic pencil-beam simulations and can convolve spectral irradiances with a sensor response function (SRF) to calculate instrument radiances. It can also consider the instrument antenna pattern. The implementation of ARTS infrared continua is described by Buehler et al. [4].

ARTS has been used for a wide variety of applications. For example, it has been used for the analysis of operational microwave satellite measurements [2,22,23,28,29], for sensor capability studies of future sub-millimeter radiometers [5,8,21,32], and for the modelling of infrared radiative transfer [4,19,26].

ARTS implements various optimisations to reduce calculation time, without significantly compromising accuracy. The sensor response is efficiently modelled by a matrix multiplication [15]. Atmospheric absorption can be treated by an absorption lookup table [7]. For radiance simulations of channels containing many spectral lines, such as channels on the infrared sensors HIRS and AVHRR, a large number of monochromatic simulations needs to be calculated in order to fully characterise the channel radiance. Buehler et al. [6] have developed a method to derive a small set of frequencies with associated weights that accurately represent the channel radiance. This allows the user to reduce the number of
monochromatic radiance calculations by a factor of roughly 100–1000. This in turn dramatically decreases the calculation time without introducing any significant errors in the calculated channel radiance. Thus, it becomes possible to quickly perform a large number of simulations, as required to get good statistics, to develop new retrievals, or to do retrievals. Buehler et al. [6] have derived optimised frequency grids for HIRS and under clear-sky conditions only. We show that the optimised grid produces results as accurate as the reference grid, even for cloudy simulations.

2.2.2. Cloudy radiative transfer

ARTS is able to perform radiance simulations in the presence of scattering particles, such as ice particles. Two alternative modules for calculating cloudy radiances are included.

One is the Discrete Ordinate Iterative (DOIT) module [14,34]. DOIT is a polarised Discrete Ordinate Method (DOM). DOIT has been developed for microwave and sub-millimeter radiances, where ice particle scattering phase functions are quite smooth. For infrared radiation, ice particle scattering phase functions have a strong forward peak (e.g. [9,35]). To properly characterise this with a DOM, a very fine angular grid is required, which makes the simulations very slow and therefore impractical. Although methods exist to alleviate this problem, none are currently implemented in ARTS-DOIT.

The other scattering module is a backward Monte Carlo (MC) method [12]. MC is often slower than DOM, but in the presence of a strong forward peak, MC was found to be faster for the problem at hand. For each monochromatic pencil-beam simulation, photons are generated, traced, and a radiance is calculated. This continues until either (1) a desired accuracy is reached, (2) a maximum number of photons is reached, or (3) a time limit is reached.

Although we use a MC-based method for cloudy simulations, the conclusions are valid for any solver. This is further discussed in Section 5.

2.3. Atmospheric data

For this study, we use two datasets containing atmospheric samples.

Firstly, we use atmospheric states from Garand et al. [17], which contain pressure, temperature, and concentrations for water vapour (H$_2$O), ozone (O$_3$), carbon dioxide (CO$_2$), nitrous oxide (N$_2$O), carbon monoxide (CO), and methane (CH$_4$). This dataset is clear-sky only.

Secondly, we use a dataset developed by Chevallier et al. [10]. This dataset is suitable for testing purposes, because it...
contains many different atmospheric states, including common and extreme cases. It consists of diverse atmospheric profiles, selected from reanalysis of the European Centre for Medium-range Weather Forecasting (ECMWF). We will refer to this as the Chevallier dataset. The dataset consists of several collections of 5000 profiles, each selected to maximise variability in a particular parameter: temperature, humidity, ozone, cloud condensate and precipitation. The datasets are sorted such that common cases, close to the mean atmospheric state, occur near the start of the dataset, whereas extreme cases occur near the end of the dataset. The total number of profiles is 25,000.

The Chevallier dataset includes profiles of cloud ice water and other hydrometeors. This is unusual, as most well-documented datasets of atmospheric profiles are clear-sky only. Regarding trace gases, it contains volume mixing ratios for water vapour and ozone. To accurately model infrared radiances, profiles for additional atmospheric gases are required, because they affect channel radiances (see Fig. 1). From the US standard atmosphere given by Anderson et al. [1], we have obtained profiles for carbon dioxide (CO₂), nitrous oxide (N₂O), carbon monoxide (CO), methane (CH₄), nitrogen (N₂), and oxygen (O₂), because those gases are present in the dataset from Garand et al. [17], but not in the Chevallier dataset. For those gases, we have used the same profile in every simulation. This does not affect the result, because those gases are either well-mixed or spectroscopically inactive in the region considered in this study.

3. Optimised infrared cloudy radiative transfer

3.1. Clear-sky derivation for AVHRR

We derive an optimised grid for AVHRR channel 5 (12.0 μm, 833 cm⁻¹) on NOAA-19 (there are small differences in the SRF between different copies of AVHRR). This channel is arbitrary chosen and serves only as an example; the method is applicable to any thermal infrared channel. We use the 42 profiles from Garand et al. [17] described earlier, and set volume mixing ratios for oxygen (O₂) and nitrogen (N₂) to 0.2095 and 0.7808, respectively. The derivation depends on atmospheric composition, so it is important that the atmospheric data represent considerable variability. The derivation is done for clear-sky conditions. No antenna pattern is applied, but all simulations are pencil-beam. All simulations are for a nadir-looking geometry. Surface emissivity is set to one. The reference setup consists of a spectrum ranging from 23.5681 THz (12.72 μm; 786 cm⁻¹) to 26.2981 THz.

Fig. 2. Comparison between simulations using the optimised and the reference grid. The top row shows the comparison for clear-sky simulations. The bottom row shows the comparison for cloudy simulations (note the difference in y-scale). For the cloudy simulations, for each profile, we took the mean of the 10 optimised runs and the mean of the 10 reference runs, and plotted the difference as a function of the mean reference radiance.
(11.40 μm; 877 cm⁻¹), with a constant grid spacing of 5 × 10⁴ Hz (0.016 cm⁻¹). In total, the reference grid is described by 5461 frequencies. Further decreasing the grid spacing does not affect the simulated channel radiance, so this is a sufficient number of frequencies. We derive an optimised frequency grid such that the relative error in the intensity (in W m⁻² Hz⁻¹ sr⁻¹) is at most 1 × 10⁻¹. This corresponds to a brightness temperature of approximately 0.07 K. To determine how many frequencies the optimised grid needs to contain, the procedure is repeated while increasing the number of frequencies, starting with five, until the condition stated above is met.

3.2. Cloudy test for AVHRR and HIRS

To investigate if the optimised frequency grid can be used for the simulation of cloudy radiances, we calculate radiances for a set of 50 cloudy atmospheres using both the reference grid and the optimised grid. The number 50 was chosen as a compromise between calculation speed and sufficient statistics, because the simulation for the reference grid takes a long time (see also Table 2). The atmospheres are from the Chevallier dataset [10] described earlier. We select each 100th profile (profile number 0, 100, ..., 4900) from the dataset that maximises variability in cloud condensate. Because of the way the dataset is sorted, choosing each 100th profile provides an appropriate slice-through of the total variability represented by the 5000 profiles (as shown by the full range of brightness temperatures in Fig. 2).

For this study, we include only cloud ice, i.e. no precipitation and no liquid water clouds. We assume the particle size distribution from McFarquhar and Heymsfield [27]. We arbitrarily choose a particle shape of solid hexagonal columns [35]. This does not affect the result, because the physical considerations are similar for other shapes (see Section 5).

As with the simulations used to derive the optimised frequency grid, all simulations are pencil-beam and nadir-looking and surface emissivity is set to one.

We perform simulations for AVHRR channel 5 and HIRS channel 11 using the reference grid and the optimised grid. The number of frequencies for each grid is given in Table 1. Cloudy simulations are carried out using a Monte Carlo solver. The optimised setup consists of 100 photons per frequency for HIRS and 1000 photons per frequency for AVHRR, corresponding to a total number of 1900 and 5000 photons for the channel radiance, respectively. For both sensors, the reference setup uses 10 photons per frequency corresponding to 46,110 photons for HIRS and 54,610 photons for AVHRR.

The error in the simulations consists of two components: (1) the grid error, or the error between the optimised and the full grid, and (2) the Monte Carlo error due to the stochastic nature of the simulation. It is important to distinguish between the two kinds of error, because type (1) is independent of radiative transfer solver, whereas type (2) is specific for the Monte Carlo method.

To investigate the MC error and runtimes for different numbers of photons, we repeat each simulation 10 times with identical input. By MC error, we mean the variability of the calculated brightness temperatures between subsequent runs with identical input. This variability arises from the random nature of MC simulations. This investigation will show if the numbers of photons used in the reference and the optimised runs are sufficient. A simulation with the optimised grid requires a larger number of photons per frequency than a simulation with the reference grid, in order for the total number of photons and therefore the MC error to be similar. However, to limit computation time, the number of photons should not be more than necessary. The number of photons for any simulation – with the reference grid or with the optimised grid – is a tradeoff between computation time and accuracy.

4. Results

We find that AVHRR channel 5 (henceforth: AVHRR-5) can be represented using a set of only five frequencies. In Fig. 1, the green bars show the frequencies and associated weights that accurately represent the channel radiance (relative error in radiance less than 1 × 10⁻³). The full list of frequencies and associated weights is also given in Table 1.

4.1. Accuracy of the optimised grid

For the cloudy simulations for both AVHRR-5 and HIRS channel 11 (henceforth: HIRS-11), we perform a number of tests to investigate the differences between the radiances calculated with the optimised grid and the radiances calculated with the reference grid. Some of those tests are presented here.

### Table 1

<table>
<thead>
<tr>
<th>Channel</th>
<th>AVHRR-5</th>
<th>HIRS-11</th>
</tr>
</thead>
<tbody>
<tr>
<td>reference</td>
<td>5461 Frequencies</td>
<td>4611 Frequencies</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>optimised</th>
<th>Freq. (cm⁻¹)</th>
<th>Weight</th>
<th>Freq. (cm⁻¹)</th>
<th>Weight</th>
</tr>
</thead>
<tbody>
<tr>
<td>792.40</td>
<td>0.408</td>
<td>1329.686</td>
<td>0.047</td>
<td></td>
</tr>
<tr>
<td>803.91</td>
<td>0.201</td>
<td>1329.919</td>
<td>0.015</td>
<td></td>
</tr>
<tr>
<td>822.42</td>
<td>0.489</td>
<td>1332.420</td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td>871.31</td>
<td>0.062</td>
<td>1334.921</td>
<td>0.049</td>
<td></td>
</tr>
<tr>
<td>874.29</td>
<td>0.207</td>
<td>1336.739</td>
<td>0.067</td>
<td></td>
</tr>
<tr>
<td>1342.925</td>
<td>0.031</td>
<td>1347.477</td>
<td>0.045</td>
<td></td>
</tr>
<tr>
<td>1348.411</td>
<td>0.051</td>
<td>1350.862</td>
<td>0.068</td>
<td></td>
</tr>
<tr>
<td>1351.946</td>
<td>0.031</td>
<td>1364.169</td>
<td>0.062</td>
<td></td>
</tr>
<tr>
<td>1364.202</td>
<td>0.009</td>
<td>1369.371</td>
<td>0.103</td>
<td></td>
</tr>
<tr>
<td>1370.605</td>
<td>0.188</td>
<td>1376.808</td>
<td>0.007</td>
<td></td>
</tr>
<tr>
<td>1385.496</td>
<td>0.082</td>
<td>1386.496</td>
<td>0.032</td>
<td></td>
</tr>
<tr>
<td>1397.735</td>
<td>0.003</td>
<td>1400.102</td>
<td>0.105</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 2 shows the differences between the optimised and the reference grid as a function of mean reference radiances. For the clear-sky simulations, AVHRR-5 has a small positive bias whereas HIRS-11 has a small negative bias. For the cloudy simulations, the situation is the opposite. For cloudy AVHRR-5, the bias between the optimised and the reference setup is $-0.032 \text{ K}$. The bias is similarly small for the runs with HIRS-11. In both cases, the cloudy bias is not a function of radiance.

Fig. 3 shows a histogram of the data presented in Fig. 2. The figure presents a histogram of the error of the radiance obtained with the optimised grid compared to the radiance obtained with the reference grid, and a histogram of the error due to the random nature associated solely with the Monte Carlo simulation method. Since both the reference-grid simulations and the optimised-grid simulations use the Monte Carlo method, both histograms include the Monte Carlo error. The two histograms look very similar and no outliers with a large error are present. This means that most of the variability in the error between radiances simulated with the optimised or the reference grid is due to the Monte Carlo error (maximum error around 1 K), and not due to errors introduced by the optimised frequency grid (root mean square error less than 0.03 K).

Fig. 3. Error statistics for the error due to the optimised frequency grid and the random nature of Monte Carlo simulations, for AVHRR-5 and High-resolution Infrared Radiation Sounder (HIRS)-11. The dashed line (- - -) shows the radiance obtained with the optimised grid minus the mean radiance obtained with the reference grid. This describes the sum of Monte Carlo error and grid error. The dotted line (.........) shows the radiance obtained with the optimised grid minus the mean radiance obtained with the optimised grid. This describes purely the Monte Carlo error.

Fig. 4. Variability of brightness temperatures for 10 × 50 cloudy simulations as described in the text for AVHRR-5. The top panel shows the range of brightness temperatures for the optimised run on the lower panel, the pluses show the standard deviations for 10 optimised runs, while the circles show the standard deviation for 10 reference runs. The legend shows the number of photons per frequency. In brackets are the total number of photons for the optimised grid and the reference grid, respectively.
4.2. Monte Carlo-specific runtime considerations

Monte Carlo simulations have a specific relation between number of frequencies, runtime, and accuracy. This is different for deterministic scattering solvers, where runtime per channel is proportional to the number of frequencies and the error does not depend on the number of simulations. As we use a MC solver in the present study (for technical reasons), we investigate how runtime and accuracy depend on the number of photons per frequency and per channel. In Section 5 we discuss the wider implications of the results obtained with the Monte Carlo solver.

Even with the optimised grid, Monte Carlo simulations are expensive. We investigate how many photons per frequency are needed to reduce the MC error to an acceptable level yet obtain a simulation runtime that allows for doing hundreds to thousands of simulations, as needed e.g. for developing or doing retrievals. We do so for both the optimised and the reference grid, because it is not a priori obvious that the optimised grid is faster. Since the optimised grid operates on a much smaller number of frequencies, the number of photons per frequency needs to be much higher in order to keep the MC error low, as the MC error is determined by the total number of photons for the channel radiance.

Figs. 4 and 5 show the MC error for AVHRR and HIRS for a varying number of photons per frequency, respectively. Unsurprisingly, the brightness temperatures converge and the variability goes down as the number of photons increases. The highest standard deviation for the AVHRR runs for the 50 selected atmospheres is 1.972 K for the runs with 500 photons and 0.666 K for the run with 5000 photons. A standard deviation of 2 K is still rather high, but a standard deviation of 0.7 K is acceptable for many applications. The standard deviations for the AVHRR reference setup with 5461 photons in total (1 per frequency) are similar to those for the optimised setup with 5000 photons. The HIRS runs show a similar pattern.

In practice, the MC error is not the only consideration for choosing the setup. For most applications, a large number of scenarios need to be simulated. Therefore, the runtime per simulation is of major practical importance.

Table 2 shows the time the simulations took to run on an Intel(R) Xeon(R) X5482 Dual QuadCore (8 CPUs) with 3.20 GHz CPU and 16 GiB RAM. It shows that runtime increases with the number of photons, but not quite linearly in the studied range. For a small number of photons per frequency, runtime increases less than linearly with the number of photons, particularly when the number

<table>
<thead>
<tr>
<th>Run</th>
<th>Photons Per freq.</th>
<th>Time Total</th>
<th>Time User</th>
<th>Time cpu</th>
</tr>
</thead>
<tbody>
<tr>
<td>Optimised (5 freq.)</td>
<td>1 5</td>
<td>39 s 43 s</td>
<td>2 m 2 m</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 50</td>
<td>4 m 4 m</td>
<td>12 m 13 m</td>
<td></td>
</tr>
<tr>
<td></td>
<td>100 500</td>
<td>19 m 36 m</td>
<td>2 h 2 h</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1000 5000</td>
<td>3 h 3 h</td>
<td>15 h 15 h</td>
<td></td>
</tr>
<tr>
<td>Reference (5461 freq.)</td>
<td>1 5461</td>
<td>74 h 74 h</td>
<td>97 h 97 h</td>
<td></td>
</tr>
<tr>
<td></td>
<td>10 54,610</td>
<td>77 h 77 h</td>
<td>238 h 239 h</td>
<td></td>
</tr>
</tbody>
</table>

Table 2 Time to simulate 50 cloudy atmospheres for AVHRR channel 5 using ARTS-MC. “s” is seconds, “m” is minutes, and “h” is hours. User time refers to the time passed during the simulation. CPU time refers to the processor time. On a multi-core machine running parallelised code, CPU time may be considerably larger than user time.
of frequencies to simulate is large. For approximately the same total number of photons (5000), the optimised setup is still much faster than the reference setup. Therefore the optimised setup is a good choice for doing simulations.

5. Discussion

There are several reasons why the clear-sky derived optimised frequency grid can be used for cloudy simulations. Buehler et al. [6] require all weights to be (1) positive, (2) inside the channel, and (3) to add up to be exactly equal to one. These requirements make the result more consistent with our physical understanding of the problem, and if the spectrum is flat, a weighted mean gives the same result as a simple mean. Figs. 6 and 7 show calculated spectra for a window channel (AVHRR-5) and a sounding channel (HIRS-11), respectively. For both figures, panel (a) shows the spectrum for a thick cloud with a high cloud top height, whereas panel (b) shows the spectrum for a relatively low and thin ice cloud. Clear-sky spectral radiances vary strongly as a function of wavelength, because of the high number of relatively narrow spectral lines in the infrared. This effect is stronger for the sounding channel (Fig. 7) than for the window channel (Fig. 6). For a high and thick cloud, radiances emerge mostly from the cloud top. Since cloud optical properties are not strongly spectrally dependent, the spectrum for such a profile is quite flat. This can be seen in Figs. 6a and 7a. The spectrum for a profile with a low and thin cloud is much less flat. For the window channel (Fig. 6b), the spectrum is still relatively flat, although strong lines are apparent. The spectrum for the sounding channel with a low

![AVHRR-5 spectrum, clear-sky and cloudy](image)

**Fig. 6.** Infrared nadir-looking spectra for wavelengths in AVHRR channel five for atmospheric profiles 4300 (panel a) and 4400 (panel b) from the Chevallier dataset [10]. The IWC profiles are shown for reference. Channel response and weights are repeated from Fig. 1. The noise in the cloudy spectrum is due to the limited number of photons per frequency. Although 100 photons per frequency is more than enough to simulate the channel radiance using the full frequency grid (it would mean around 500,000 photons), there is still significant noise at each individual frequency. Due to constraints on computation power and time, a reference calculation with significantly more photons is not feasible: (a) Chevallier profile 4300 and (b) Chevallier profile 4400.
As discussed in the introduction, a more difficult situation may arise if the cloud top coincides with the peak of a broad weighting function. In this case, the atmospheric region from which the clear-sky signal originates changes, and therefore the relative contribution for different gases might, too. However, we find that the optimised grid works for all explored cases.

When considering clouds only, many combinations of frequencies and associated weights work, as long as the weights add up to one and the slight variation of optical properties with frequency is represented. Since gaseous absorption and emission still occur in a cloudy sky, also above the cloud, the clear-sky-derived optimised frequency grid is a good choice for cloudy simulations.

The time gain by using an optimised frequency grid with a Monte Carlo model is less than the time gain with a deterministic scattering solver. In fact, that there is a time difference at all between 5000 photons distributed over a large number of frequencies or the same number distributed over a small number of frequencies can be explained by the overhead associated with preparing the MC simulation for each frequency. Some tasks, such as extracting optical properties, need to be performed once for each frequency.

The time gain for deterministic scattering solvers is expected to be much larger. For such models, the runtime per channel is (close to) directly proportional to the total number of frequencies per channel. Therefore, an optimised frequency grid will give a reduction in runtime by a factor 100 to 1000.

6. Conclusions and outlook

In this study, we have shown that an optimised frequency grid derived for clear-sky conditions with the
method described by Buehler et al. [6] can be applied for cloudy simulations. For a newly derived optimised frequency grid for AVHRR channel 5 and a frequency grid derived by Buehler [6] for HIRS channel 11, we have investigated the differences between simulations using the optimised grid and the full grid, respectively. The optimised grid has 100–1000 times less frequencies than the full grid. We found the bias to be less than 0.03 K with no dependence on cloud properties.

For simulations with a deterministic scattering solver, it is evident that reducing the number of frequencies decreases the runtime. However, in the present study, we used the ARTS Monte Carlo model for the simulations. Therefore, we also studied how the choice of the number of photons affects the Monte Carlo errors and runtimes. For a similar number of photons per channel, the optimised grid has the same accuracy as the reference grid, but is approximately 10 times faster.

The results can be applied to any downloading infrared radiometer and are highly useful for further studies.

Like the optimised HIRS frequency grids developed by Buehler et al. [6], the new optimised frequency grid for AVHRR channel 5 is available along with the ARTS software distribution.

We plan to apply the results, for example, to a systematic study of the IWP signal in AVHRR thermal radiances. Such a study would compare statistics with those from a collocated dataset based on Holl et al. [20] (further developed by John et al. [23]) and those for microwave radiances, in particular Microwave Humidity Sounder (MHS) channels around the 183 GHz water vapour absorption line. This study might be carried out in the future.

Acknowledgments

The PhD student position of the main author is funded by the Swedish Vetenskapsrådet. We thank Ping Yang for providing the authors with updated single scattering data for the infrared. We thank the ARTS radiative transfer community for its work on ARTS. We would also like to thank the Swedish Microwave Humidity Sounder (MHS) channels around the 183 GHz water vapour absorption line. This study might be carried out in the future.

References


Paper IV

Systematic and random errors between collocated satellite ice water path observations

Systematic and random errors between collocated satellite ice water path observations

S. Eliasson,1 G. Holl,1 S. A. Buehler,1 T. Kuhn,1 M. Stengel,2 F. Iturbide-Sanchez,3 and M. Johnston4

Received 29 June 2012; revised 5 October 2012; accepted 15 November 2012.

[1] There remains large disagreement between ice-water path (IWP) in observational data sets, largely because the sensors observe different parts of the ice particle size distribution. A detailed comparison of retrieved IWP from satellite observations in the Tropics (±30° latitude) in 2007 was made using collocated measurements. The radio detection and ranging (lidar) (DARDAR) IWP data set, based on combined radar/lidar measurements, is used as a reference because it provides arguably the best estimate of the total column IWP. For each data set, usable IWP dynamic ranges are inferred from this comparison. IWP retrievals based on solar reflectance measurements, in the moderate resolution imaging spectroradiometer (MODIS), advanced very high resolution radiometer–based Climate Monitoring Satellite Applications Facility (CMSAF), and Pathfinder Atmospheres-Extended (PATMOS-x) datasets, were found to be correlated with DARDAR over a large IWP range (~20–7000 g m^-2). The random errors of the collocated data sets have a close to lognormal distribution, and the combined random error of MODIS and DARDAR is less than a factor of 2, which also sets the upper limit for MODIS alone. In the same way, the upper limit for the random error of all considered data sets is determined. Data sets based on passive microwave measurements, microwave surface and precipitation products system (MSPPS), microwave integrated retrieval system (MiRS), and collocated microwave only (CMO), are largely correlated with DARDAR for IWP values larger than approximately 700 g m^-2. The combined uncertainty between these data sets and DARDAR in this range is slightly less MODIS-DARDAR, but the systematic bias is nearly an order of magnitude.


1. Introduction

[2] Clouds have a dominant effect on the radiation entering and leaving the atmosphere [Hartmann et al., 1992]. Better understanding of the impact of ice clouds on the radiation budget and the hydrological cycle is paramount to improving climate models [e.g., Stephens et al., 1990]. Climate models are the most important tools for understanding long-term atmospheric processes and simulating climate scenarios. However, fundamental ice-cloud properties such as ice-water path (IWP) are difficult to measure accurately, making them poorly constrained. This leads to large differences of globally averaged IWP between models [Waliser et al., 2009]. Depending on their, e.g., microphysical properties, ice clouds may either cool or warm the atmosphere. The average radiative impact of all ice clouds is thought to be a net cooling effect, although semitransparent ice clouds, which may cover large areas, have a mostly warming effect on the atmosphere [Khvorostyanov and Sassen, 2002].

[3] In situ techniques used on aircraft and balloon campaigns provide the most detailed measurements of ice-water content (IWC), which, integrated by height, becomes IWP. However, the global coverage of such campaigns is sparse and limited. Ice-cloud retrievals based on satellite measurements are, in contrast, abundant. They provide macrophysical information on ice clouds such as their temporal variation and their spatial distribution, and are the most important source of validation of clouds in climate models. However, the uncertainties in satellite ice-cloud retrievals are still considerable, depending on the cloud characteristics and...
the instrument sensitivities [e.g., Zhao and Weng, 2002; Cooper et al., 2003; Austin et al., 2009]).

IWP (g m\(^{-2}\)) is one of the most important ice-cloud properties [Buehler et al., 2007, 2012, and references therein]. Previous studies have shown that the climate models show a very large spread in their reported cloud ice amounts [Waliser et al., 2009; John and Soden, 2006; Wyatt et al., 2006; Eliasson et al., 2011]. However, the models are in good agreement with observations in terms of one of the most fundamental quantities, top of atmosphere net radiative flux [Smith et al., 1994]. This implies that the models may be adjusting poorly constrained parameters, such as clouds, to achieve the correct top of atmosphere radiative flux, thus leading to unrealistic cloud characteristics in the models [Stephens et al., 2002; Li et al., 2012]

Atmospheric column integrated quantities, such as IWP, can be retrieved from radiances from passive down-looking sensors. Retrievals from passive sensors use two or more channels at microwave (MW), infra-red (IR), near-infra-red (NIR), or visible (VIS) wavelengths [e.g., Heymsfield et al., 2003]. Since June 2006, measurements from active instruments, a cloud profiling radar (CPR) on the CloudSat satellite, and a Cloud-Aerosol LIDAR and Infrared Pathfinder Satellite Observation (CALIOP) satellite, have greatly increased our knowledge of ice clouds [Heymsfield et al., 2008]. These instruments measure detailed information on the vertical structure of clouds and can detect multilayered clouds. This can, for example, provide valuable information on how the radiation is distributed in the atmospheric column, because it depends strongly on the vertical structure of clouds [L'Euzier et al., 2008; Mace and Benson, 2008].

In earlier studies [Waliser et al., 2009; Eliasson et al., 2011], a subset of climate models in the Fourth Assessment Report (AR4) of the Intergovernmental Panel on Climate Change were intercompared and compared with observational data sets. The models were in large disagreement on IWP (cloud ice). Not only were there large differences in magnitude, their spatial distribution were in poor agreement as well. The observational data sets were in good spatial agreement, but the differences in IWP magnitude was as large as the intramodel difference. This was also expected mainly because the observational data sets were based on measurements that (1) are made at different wavelengths, and thus sensitive to different parts of the particle size distribution (PSD) [Comstock et al., 2007; Waliser et al., 2009]; and (2) have different resolutions.

Retrievals from passive instruments (nadir viewing) are much more limited in terms of information on the vertical structure of clouds, and most have a much coarser measurement resolution (footprint). Despite their limitations, records from passive instruments cover 10 (moderate resolution imaging spectroradiometer [MODIS]) to 30 years (AVHRR), meaning that only these records can be used for actual climate studies. Data from CloudSat and CALIPSO are limited to approximately 6 years (at the time of writing) and are therefore more suitable for process studies. Furthermore, because passive instruments also have large swath widths, passive instruments provide a much better spatial coverage than active instruments, which measure only at nadir. These advantages motivate the inclusion of passive instrument retrievals in this study. However, passive instruments provide only IWP; therefore, vertical information from active instruments goes unused.

In practice, the longer the wavelength of the measurements, the deeper into a cloud one can measure, but this is at the cost of losing sensitivity to small particles, which are generally present at the top of the ice clouds [Wu et al., 2009]. Therefore, ice-cloud retrievals based on, e.g., passive microwave measurements can be made in deep clouds, but generally not in shallow clouds, which are mostly made up of small ice particles. In other words, microwave retrievals are only sensitive to the precipitation-sized ice particles (>~0.25 mm) in the IWP total column [Zhao and Weng, 2002]. Also, although CloudSat CPR measurements can be used to retrieve information deep into most clouds, it is insensitive to small particles. CALIOP, in contrast, is even sensitive to very small ice particles. Because the satellites are always in close proximity to one another, CALIOP can be complementary to CPR. CALIOP by itself is limited for retrieving IWP because the measurements are attenuated for moderately thick clouds [e.g., Delanoë and Hogan, 2010]. In summary, satellite retrievals based on a single measurement technique cannot measure the whole depth of the ice cloud [Wu et al., 2009], which would be desirable from a climate model perspective.

The models participating in the upcoming Fifth Assessment Report (AR5) have undergone some improvements in terms of cloud amount and distribution, but there still remains large differences in IWP [Jiang et al., 2012]. A similar study by Li et al. [2012] reports that climate models do not perform particularly well despite generally simulating cloud ice better in Coupled Model Intercomparison Project Phase 5 (CMIP5) than they did in CMIP3 (included in AR4). This is an added incentive to further improve the understanding of IWP.

This article can be seen as continuation of the work presented in Eliasson et al. [2011] and Waliser et al. [2009]. It builds on their findings but evaluates the discrepancies in the retrieval schemes rather than of averaged products. The latter, called level 3 monthly mean (L3), are representations that have already undergone a conversion from level 2 (L2) to L3, which, because different approaches and filters are applied, can contribute significantly to the differences between data sets. For instance, one important decision is the scan angle and solar zenith angle cutoffs one applies when creating the L3 data set, because there might be quality/bias dependencies of the IWP retrievals on these angles. The decision made in each of these approaches significantly alters the L3 product. Thus, this new study will once again survey satellite-observed IWP, but instead of comparing monthly mean values, we compared collocated measurements directly. By doing so, the same “cloud” is measured from different instruments, without the added complexity of different L2 to L3 conversions. This is not a validation study of any particular data set. The IWP retrievals assessed here, being from different sensors, are inherently different and difficult to compare in the rigorous manner a validation study requires [e.g., Stein et al., 2011; Zhang et al., 2009].

As reference for the satellite IWP data sets, the radar/ lidar (DARDAR) data set, which is based on CALIOP and CPR, was chosen [Delanoë and Hogan, 2010]. Data sets compared with this reference are from active sensors, two products from CloudSat’s L2 radio detection and ranging.
Table 1. List of Data Sets Used in This Study

<table>
<thead>
<tr>
<th>Short Name</th>
<th>Technique</th>
<th>Long Name</th>
<th>Satellite</th>
</tr>
</thead>
<tbody>
<tr>
<td>DARDAR</td>
<td>Active: combined radar and lidar</td>
<td>DARDAR-cloud</td>
<td>CloudSat and CALIPSO</td>
</tr>
<tr>
<td>IORO</td>
<td>Active: radar</td>
<td>Ice only radar only ice water path</td>
<td>CloudSat</td>
</tr>
<tr>
<td>RO</td>
<td>Active: radar</td>
<td>Radar only ice water path</td>
<td>CloudSat</td>
</tr>
<tr>
<td>MODIS</td>
<td>Passive: SRBS</td>
<td>MODIS</td>
<td>Aqua</td>
</tr>
<tr>
<td>CMSAF</td>
<td>Passive: SRBS</td>
<td>CMSAF</td>
<td>NOAA-18</td>
</tr>
<tr>
<td>PATMOS-x</td>
<td>Passive: SRBS</td>
<td>PATMOS-x</td>
<td>NOAA-18</td>
</tr>
<tr>
<td>MSPPS</td>
<td>Passive: Microwave</td>
<td>MSPPS</td>
<td>NOAA-18</td>
</tr>
<tr>
<td>MIRS</td>
<td>Passive: Microwave</td>
<td>MIRS</td>
<td>NOAA-18</td>
</tr>
<tr>
<td>CMO</td>
<td>Passive: Microwave</td>
<td>CMO</td>
<td>NOAA-18</td>
</tr>
</tbody>
</table>

*Data sets are further referred by the “short names” given here.

(radar) only (RO) data set (called 2B-CWC-RO), in this article called RO and ice only radar only (IORO) [Austin et al., 2009], and from passive sensors; three data sets retrieved from solar relectivities in the VIS and NIR spectral regions (MODIS [King et al., 2003], the advanced very high resolution radiometer [AVHRR] climate monitoring satellite applications facility [CMSAF] data set [Roebeling et al., 2006], and Pathfinder Atmospheres-Extended [PATMOS-x] [Walther and Heidinger, 2012]); and three data sets that are based on MW channels (microwave surface and precipitation products system [MSPPS] [Ferraro et al., 2005], microwave integrated retrieval system [MiRS] [Boukabara et al., 2011], and collocated microwave only [CMO] Holl et al. [2010]). An overview of all the data sets used in this study is given in Table 1.

Because of the combination of radar and lidar, DARDAR has a very large IWP value range that covers the range of all IWP data sets in this study. Except for DARDAR, it is assumed that each data set contributes information for certain IWP value ranges. These IWP retrievals from different techniques are compared “out of the box.” We do not attempt to make the best possible comparisons between data sets such as by limiting comparisons with optical depths that both data sets are sensitive to, as done in Stein et al. [2011] for DARDAR and MODIS. Instead, we acknowledge that all retrieval techniques have their merits and limitations, and compare them as they are, restricted only by the collocation rules described in section 3.

In this article, IWP measurements are compared at the smallest possible temporal and spatial scales by stringently collocating measurements using a flexible collocations toolbox based on work first described in Holl et al. [2010] and later expanded in John et al. [2012]. The main aim is to quantify which measurement techniques work for which IWP value ranges and to separate the systematic and random errors of the collocated data sets. The data sets are described in more detail in section 2. The collocation and comparison methodologies, including the systematic and random errors, are described in section 3. In section 4, first the uncertainty of DARDAR IWP is investigated before identifying the valid IWP ranges of the collocated data sets and finding their systematic and random errors. Section 5 gives a recollection and states our main conclusions.

2. Description of Data Sets

The data sets chosen for this study cover a wide range of measurement and retrieval techniques that are used to determine IWP. They are expected to report varying IWP magnitudes because they are sensitive to different ice particle sizes and shapes, and therefore are generally sensitive to different altitudes of the cloud. The data sets are based on instruments found on satellites flying in or near the afternoon train (A-train) [Stephens et al., 2002], so that a large number of collocated measurements can be found. For data sets that rely on visible reflectance measurements, only data from the daytime exist. Common to all data sets are the large uncertainties because of the assumptions on cloud microphysical properties that must be made to complete the cloud retrieval. Although these assumptions are best guesses mostly based on knowledge gained from in situ campaigns, many cloud properties, such as particle habit, are very variable [e.g., Heymsfield and McFarquhar, 2002]. Several studies [e.g., Zhang et al., 2009] have shown that different but widely accepted assumptions on the microphysical properties of ice particles lead to very large deviations in the retrieved IWP. In addition, the uncertainty because of the cloud particle phase, as the distinction between liquid and ice clouds, has a significant impact on the retrieved IWP as described later.

2.1. IWP From Active Instruments

Three data sets in this comparison originate from active measurements and are all based on CPR measurements from the CloudSat satellite.

2.1.1. DARDAR

The DARDAR data set is based on a combination of CPR, CALIOP, and MODIS measurements [Delanoë and Hogan, 2010]. By combining these techniques using the variational method described therein, DARDAR uses the very different particle size sensitivities of radar and lidar measurements. The retrieval is seamless and works as long as at least one of CPR or CALIOP detects a cloud. If the cloud is detected by both instruments, the cloud properties are retrieved using both measurements. Delanoë and Hogan [2010] showed that this combined retrieval approach is less sensitive to changes to assumed microphysical properties than retrievals based on CPR or CALIOP alone. The footprint size is the same as CPR because the CALIOP measurements are averaged horizontally in the CPR footprint. Further details on the retrieval technique are given in Delanoë and Hogan [2008], and a comparison of the DARDAR ice-cloud retrieval algorithms was performed in Stein et al. [2011]. The specific product used in this study is called DARDAR-cloud and is derived from CPR and CALIOP only.

DARDAR was chosen as the reference data set because it currently provides the best estimate of the total column IWP.
It is expected to cover a larger IWP range than any other data set assessed here. For details on the uncertainty associated with DARDAR IWP, see section 4.1.

2.1.2. CPR (RO and IORO)

[18] The CloudSat satellite carries a CPR that measures at 94 GHz [Stephens et al., 2002]. Measurements are made at a vertical resolution of 250 m, and the measurement footprint size is about 1.4 × 3.5 km, with the long axis along the satellite flight path. CloudSat is less sensitive to small ice particles than lidar, but it is not saturated unless the clouds are very thick and there is heavy precipitation.

[19] One of the two CloudSat products assessed in this study allows IWC and liquid-water content (LWC) to coexist between -20 °C and 0 °C (but only the IWC part is included in the integration), and in this study, this data set is called RO (field name: RO_ice_water_path) [Austin et al., 2009]. RO was also the reference data set in Eliasson et al. [2011].

[20] CloudSat also provides a combined radar and visible optical depth (RVOD) product with presumably better constrained retrievals. However, Protat et al. [2010] state that RO and RVOD are statistically virtually identical, and therefore only the RO data set is used here. The uncertainty of IWC retrievals using simulated CPR was determined to be about 40% [Austin et al., 2009]. If these were the only uncertainties in the retrieval, 40% would be the upper limit of the uncertainty for a column integrated RO IWP product. For RO data set, there is a substantial additional uncertainty because of the assumed cloud phase. A linear combination of IWC and LWC is used between the temperature range 0°C to −20°C, where the fractional ice phase increases from 0 to 1 (liquid-water path [LWP]: 1 to 0). Devasthale and Thomas [2012] showed that other realistic ice to liquid phase relationships in this temperature range lead to very different IWP retrievals.

[21] The other CloudSat product used in this study does not attempt to separate IWP and LWP, and instead reports the whole column in the above temperature range as IWP; we have called this product IORO (field name: IO_RO_ice_water_path). As mentioned earlier, the IORO data set does not suffer from cloud-phase uncertainties caused by the above rigid approach, but by assuming the whole layer contains only ice, the IWP is overestimated. The a priori temperature information that both data sets rely on is model auxiliary data from European Centre for Medium-range Weather Forecasting. Further uncertainties for both data sets are inherited from the a priori input used in the retrieval [Austin et al., 2009].

2.2. IWP From Solar Reflectances

[22] The data sets in this section are derived from solar reflectance and share the same retrieval technique. This method is called solar reflectance bispectral (SRBS) and is described in Nakajima and King [1990]. The method uses passive measurements of reflected solar radiation to retrieve visible cloud optical depth ($\tau_{\text{v}}$) and $\tau_c$ simultaneously, where $\tau_c$ is the mean effective radius. The solar reflectance of the nearly nonabsorbing wavelengths is used to retrieve $\tau_v$ and moderately absorbing solar reflectance measurements are used to retrieve $\tau_c$. This is done in conjunction with a lookup table of simulated reflectances.

[23] In SRBS retrievals, assumptions have to be made about the horizontal and vertical structure of the cloud, i.e., constant IWC and constant particle effective radius ($r_e$) throughout the cloud [Stein et al., 2011, and references therein]. These assumptions lead to uncertainties, because inhomogeneities in the vertical structure of clouds have a strong influence on the retrieved cloud properties [Zhang et al., 2010]. Also, the retrieved $r_e$ is biased toward the top of the thick clouds, because only the top four or five optical depths contribute to the reflectance in the moderately absorbing channel [McFarquhar and Heymsfield, 1998]. Because the smallest ice particles are generally at the top of the cloud, this may create systematic errors by underestimating $r_e$ for thick clouds.

[24] Nonetheless, from the retrieved $\tau_v$, and $\tau_c$, IWP can be derived using the following relationship:

$$\text{IWP} = \frac{4\pi r_e \rho_{\text{ice}}}{3Q_v}$$

where $\rho_{\text{ice}}$ is the density of ice, and $Q_v$ is the average extinction efficiency for ice at a wavelength of 0.66 μm ($Q_v \approx 2$) [e.g., Meyer et al., 2006]. We have used $\rho_{\text{ice}} = 930 \text{ kg m}^{-3}$, because this is the value used in the MODIS retrieval [King et al., 2006].

2.2.1. MODIS

[25] The MODIS cloud retrieval products are based on daytime measurements from 36 channels in the VIS, NIR, and IR spectral ranges from the MODIS instrument on board two polar-orbiting satellites, Terra and Aqua. However, only a subset of four channels is used to ultimately retrieve IWP, namely, either channel 1 (0.645 μm), 2 (0.858 μm), or 5 (1.24 μm), depending on the underlying surface, and channel 7 (2.13 μm) [King et al., 1997]. The lookup table used in the MODIS retrieval algorithm is based on the ice particle model by Baum et al. [2005] (Baum05).

[26] The MODIS L3 monthly mean IWP data set (called MYD08_L3) was assessed in Eliasson et al. [2011]. This article assesses the L2 data set (called MYD06_L2) collection 005 [King et al., 2003], which is from only the Aqua satellite. The IWP is extracted using the fields cloud_water_path, Cloud_Phase_Optical_Properties, and Quality_Assurance_1km. The resulting IWP data set is a 1 × 1 km pixel (footprint) product and, therefore, is of a similar size as the DARDAR retrieved footprint. MODIS data used in this study are described in documentation available through the MODIS Web site (http://modis-atmos.gsfc.nasa.gov/).

2.2.2. AVHRR PATMOS-x

[27] PATMOS-x IWP is derived from the PATMOS-Daytime Cloud Optical Microphysical Properties (DCOMP), based on passive measurements in the VIS and NIR spectral ranges from AVHRR global area coverage (GAC) data [Walther and Heidinger, 2012]. The lookup table is based on Baum05. GAC data are a reduced resolution data set based on AVHRR. Four adjacent footprints in every scan are averaged together; then the next three scan lines are skipped. The AVHRR instrument is installed on the polar-orbiting operational environmental satellites from National Oceanic and Atmospheric Administration (NOAA). In this study, the L2h cloud parameter product (version 5) is further sampled onto a 0.1° × 0.1° grid. The gridded data points are treated as “pseudo footprints” that are largest at the equator.
(roughly $10 \times 10$ km), shrinking laterally toward the poles. The product does not contain an IWP product, but contains $\tau_c$ and $r_c$. Therefore, IWP is extracted using equation 1 from these quantities from cloud types considered ice phase (called opaque ice, cirrus, overlapping, and overshooting in this data set, but slightly different terms are used in Walther and Heidinger [2012].

[28] PATMOS-x has only two solar reflectance channels for the $\tau_c$ and particle effective radius ($r_c$) retrievals: channels 1 (0.6 $\mu$m) and 3b (3.7 $\mu$m). The PATMOS-x L3 IWP data set (version 4) was assessed in Eliasson et al. [2011]. In this study, the L2b data set is based on measurements from NOAA-18, which flies close to the A-train; hence, the PATMOS-x measurements are often collocated with DARDAR. The PATMOS-x data used in this study are described and made available online at http://cimss.ssec.wisc.edu/patmosx/.

2.2.3. AVHRR CMSAF

[29] CMSAF IWP retrieval data are based on the cloud physical properties (CPP) algorithm developed at Koninklijk Nederlands Meteorologisch Instituut (KNMI) [Roebeling et al., 2006], which is used to retrieve cloud thermodynamic phase, cloud optical thickness, cloud particle effective radius, and LWP/IWP from AVHRR GAC data. Therefore, the CMSAF cloud products have a pseudo footprint of about $4 \times 1$ km. The CPP scheme, based on the SRBS method, uses lookup tables of water and ice-cloud reflectance simulated by the Doubling Adding KNMI radiative transfer model [Stamnes, 2001]. Although MODIS and PATMOS-x both use the Baum05 ice particle model, for CMSAF IWP, ice particles as given in Hess et al. [1998] were used. Cloud properties are retrieved by iteratively matching observed and simulated reflectance. The IWP is derived using equation 1 for all cloudy pixels, which were identified using the Polar Platform System cloud processing package developed by Swedish Meteorological and Hydrological Institute [Dybroe et al., 2005a, 2005b], and which were diagnosed to contain ice clouds (internally done in CPP). Also, notably, the AVHRR reflectances are recalibrated following Heidinger et al. [2010]. As PATMOS-x, and CMSAF IWP retrievals are based on the 3.7 $\mu$m channel of AVHRR, the availability of different NIR channels might significantly affect the IWP retrieval, because of different penetration depth of, e.g., 3.7 and 1.6 $\mu$m. This might explain some of the differences seen between MODIS and the AVHRR data sets later in this article.

2.3. IWP From Passive Microwave

[30] Retrieving IWP from passive microwave sensors is analogous to retrieving the IWP of the particles with a radius between approximately 250 and 1500 $\mu$m. These are precipitation-sized particles, which should be considered when comparing with other data sets that have sensitivity to small particles. In general, comparisons of microwave data sets where DARDAR IWP is relatively low will not be made, because such “clouds” are likely made up of small particles that are beyond the sensitivity of passive microwave measurements. The three data sets used in this study use measurements from the microwave humidity sounder (MHS) instrument, which has a footprint diameter of around 15 km at nadir on the ground.

2.3.1. MSPPS

[31] MSPPS provides a data set of IWP derived from the ratio of the amount of radiation at 89 and 150 GHz that is scattered out of the line of sight by large ice particles [Ferraro et al., 2005]. These two frequencies are measured by MHS channels 1 and 2. The particle-induced scattering intensifies with increasing frequency and is detected as a depression in brightness temperature [Vivekanandan et al., 1991]. Assumptions are made on the PSD and bulk volume density, and accurate knowledge on the surface temperature and its emissivity must be available [Zhao and Weng, 2002]. As described therein, errors in assumptions can lead to large errors in the IWP retrieval.

2.3.2. MiRS

[32] MiRS is a one-dimensional variational satellite data assimilation retrieval system that uses microwave observations for the retrievals of several atmospheric and surface geophysical parameters simultaneously, among them IWP [Boukabara et al., 2011]. Like MSPPS, MiRS is based on MHS sensor observations, but the MiRS retrievals are additionally constrained by measurements from the Advanced Microwave Sounding Unit A (AMSU-A). Because AMSU-A has a nadir footprint of 48 km, the MHS measurements made inside the larger footprint are averaged, so the footprint of MiRS is that of AMSU-A. MiRS is described in detail in Boukabara et al. [2011]. The MiRS version used in this study provides an IWP product called graupel water path, which indicates by name alone that only precipitating sized particles are retrieved. In MiRS, the IWP product is used as a predictor for the estimation of rainfall rate intensities in mm h$^{-1}$. As described in Iturbide-Sanchez et al. [2011], the validation/assessment of the MiRS rainfall rate is an indirect method to assess the quality of the retrieved MiRS hydrometeors, including the IWP.

2.3.3. CMO

[33] CMO is a data set based on a technique first introduced in Holl et al. [2010]. It is currently under development and will be described in detail in a paper that is in preparation. High-frequency measurements from MHS channels 3 to 5 located at 183(1) GHz, 183(3) GHz, and 190 GHz, which are traditionally used for water vapor retrievals, form the basis of this data set. This sets it apart from the other microwave-based data sets that use lower frequencies. Therefore, the CMO data set should be able to detect smaller ice particles and retrieve smaller IWP, a hypothesis that can be tested by including it in this study.

[34] Artificial neural networks are used for the retrieval of IWP rather than the depression of brightness temperature directly. The training database is obtained by collocating CloudSat with NOAA-18 MHS, using the collocation toolbox described in Holl et al. [2010] and later in section 3. Holl et al. [2010] showed that the MHS channels 3 to 5 are not sensitive to clouds with RO IWP less than 100 g m$^{-2}$, and the data set is planned to include radiances from IR measurements to potentially increase sensitivity to lower IWP values. However, this study uses CMO version 0.4, which does not yet include IR radiances.

3. Methodology

[35] To compare simultaneous measurements of the same cloud situation using several instruments, the measurements
need to be stringently collocated in time and space. To find such collocated measurements, we have used the collocation toolbox, a highly flexible toolkit that allows for easy collocation between different data sets, first presented in Holl et al. [2010]. The software, some precalculated collocated data sets, and detailed documentation can be accessed online at http://www.sat.ltu.se/docs/data/collocations. The toolbox has been used in several studies and continues to be developed further (V. O. John, G. Holl, N. Atkinson, and S. A. Buehler, Monitoring scan asymmetry of microwave humidity sounding channels using simultaneous all angle collocations [SAACs], Journal of Geophysical Research, in press 2013).

One very important consideration to take into account when collocating measurements that use different techniques is their different horizontal resolutions given by their “footprints.” Figure 1 illustrates footprint sizes for the instruments that the data sets in this study are based on. Technically, to collocate data sets, usually the one with the smallest footprints (secondary) is collocated to the other (primary). All the small-footprint measurements from the secondary data set that are close enough to the center of a larger footprint from the primary data set are found and are considered collocated. The user defines what “close enough” means in space and time in the presets (see later). Each collocated data set contains information such as distances, number of secondary measurements, and the statistics of each subset of small-footprint measurements such as standard deviation, mean, number of elements, or other user-defined functions such as “cloud amount.”

Comparing measurements with large footprints with measurements with small footprints introduces a sampling problem. For instance, a single CPR measurement (and DARDAR retrieval) covers around 0.65% of the area of an MHS footprint, and even the maximum possible number of DARDAR retrievals collocated with one MHS measurement cover less than 10% of its footprint [Holl et al., 2010]. In this article, large-footprint measurements are considered collocated with a set of small-footprint measurements only if the following two conditions are met, to mitigate this sampling problem: First, the number of smaller footprints inside the larger footprint must be at least 70% of the maximum possible collocations of small footprints to one large footprint within the set collocation criteria.

The temporal limit chosen is the time for advection to move by the radius of the footprint, so that at most half the footprint’s condition is changing. In this study, it is estimated that the air mass will be replaced at roughly a rate of 1 km min$^{-1}$ (16.7 m s$^{-1}$), which is a fairly conservative estimate for the Tropics [Koh et al., 2011]. Changes caused by other atmospheric processes such as convection, which, in most cases, will take much more than 10 minutes to change conditions completely, are also taken into account. Therefore, the largest permitted temporal difference between data sets is either the footprint radius [km] in minutes or 10 min, whichever value is smallest. The exact limitations used for each collocated data set are given in Table 2.
3.1. Systematic and Random Errors of Collocated IWP

The errors are split into systematic and random errors. What constitutes a random error or a systematic one depends on context. In this study, the systematic error is defined as the mean difference between the measurements (e.g., the bias, if one is taken as a reference) and the random error as the residual after subtracting the mean (e.g., the variance). Both may be a function of IWP or of other quantities. As an example of an expected systematic error, a product based exclusively on radar should retrieve systematically less IWP than a product based on combined radar and lidar, if all other factors (such as microphysical assumptions) are the same. Random errors between collocated data sets originate from a number of different sources. These include, but are not limited to, the representation error (poor collocation in space and time) and the random errors within each data set (such as noise). If the random errors of two data sets are Gaussian and uncorrelated, variances (square of the standard deviation) add up to the total variance for the collocated data set,

$$\sigma^2(Y - X) = \sigma^2(X) + \sigma^2(Y) + \sigma^2_n,$$

where $Y$ and $X$ are collocated measurements with uncorrelated random errors, and $\sigma_n$ is the representation error caused by imperfect collocation. In this study, $Y = \log_{10}(IWP_{\text{dataset}})$ and $X = \log_{10}(IWP_{\text{DARDAR}})$. Thus, the observed random error in the comparison puts an upper bound on the random error of the individual data sets. All active data sets and the CMO data set are based on the same CloudSat data, so that it would be wrong to assume the data sets have noncorrelated random errors. Therefore, approximating random errors for these collocated data sets (RO, IORO, CO) -DARDAR cannot be done in the above manner but can be considered applicable for the comparison of the other passive instruments with DARDAR. Further investigation on whether the IWP distributions are, in fact, Gaussian in logarithmic space (i.e., lognormal) is made in section 4.2.3.

4. Results and Discussion

4.1. DARDAR Uncertainties in IWP

DARDAR retrievals use the optimal estimation method; hence, the errors are retrieved for each IWC value [Delanoë and Hogan, 2008]. 1-sigma errors of lognormal quantities mean that the errors are also given in log space; therefore, the uncertainty is a fractional uncertainty (where, e.g., a 50% uncertainty for the value 100 g m$^{-2}$ means a value range between 100/1.5 and 100*1.5). However, it is not easy to assign an error to the column integrated IWP using only the fractional errors of each IWC, because the vertical autocorrelation of IWC errors is unknown. The error on the logarithm of IWP can be estimated using the provided fractional error (1-sigma random error in the natural logarithm of IWC called “ln_iwc_error”) and assuming that the IWC uncertainties are correlated throughout the column. The ±1-sigma values of IWC can be used,

$$\text{IWC}^+ = \text{IWC}_0 e^{\ln\text{iwc\_error}}$$

$$\text{IWC}^- = \text{IWC}_0 e^{-\ln\text{iwc\_error}},$$

where IWC$_0$ is the retrieved IWC, to get the corresponding column integrated quantities IWP$^+$ and IWP$^-$. From IWP$^-$, IWP$^+$, and IWP$_0$ (integrated from IWC$_0$), we can find the fractional errors, $ln(IWP/IWP_0)$ and $ln(IWP_0/IWP)$, and their average,

$$\sigma_{\text{IWP}} = \frac{ln(IWP) - ln(IWP_0)}{2}\sqrt{\text{IWP}/\text{IWP}_0}.$$

This error $\sigma_{\text{IWP}}$ is assigned to every IWP$_0$ value such that the ±1 – $\sigma_{\text{IWP}}$ limits (in log space) can be estimated as

$$\text{IWP}_0 \times e^{\pm\sigma_{\text{IWP}}}.$$
ranges from which statements about their systematic and random errors are made. The median and the 16th/84th percentiles are used to describe the distribution of IWP.

The median is used instead of log-mean because the IWP distributions are not quite lognormal, albeit nearly (see section 4.2.3). The choice of 16th/84th percentiles (“pseudo 1-sigma”) corresponds to ±1 standard deviation in log space if the IWP distribution would be lognormal. Furthermore, because the comparisons are made in log space, all measurements where the compared data sets report IWP = 0 are removed. Cases where DARDAR has cloud-free measurements have already been discarded to minimize the number of partly cloudy footprints (see section 3). The comparisons of collocated data sets reported in the following sections are grouped in active data sets, SRBS data sets, and MW data sets.

4.2.1. IWP Valid Ranges

By comparing collocated measurements directly against each other, one can approximately read out the actual overlap in IWP sensitivities for the different techniques against the reference data set, DARDAR. Figures 4, 6, and 8 (described in sections 4.3, 4.4, and 4.5) show the two-dimensional histogram distributions of IWP for each collocated data set. The median and spread (16th/84th percentiles) of the compared data sets in each DARDAR bin are shown in gray, and the median and spread of DARDAR IWP inside the bins of the compared data set are shown in black. To clarify, in each of the bins (along the x axis) there is a distribution of IWP values from the compared data set. For example, for all collocated measurements with DARDAR IWP in the bin between 50 and 70 g m⁻², the compared data set may have values ranging from 1 to 1000 g m⁻², but with most values centered in a fairly lognormal way around 40 g m⁻². This distribution can be viewed as a probability distribution function (PDF) of likely values that this observational data set may report when DARDAR values are between 50 and 70 g m⁻². By the same reasoning, the bins of the y axis also have a PDF of DARDAR IWP values.

[40] The color scale common to Figures 4, 6, and 8 shows the number of collocations per bin normalized by the largest bin value, and the total number of collocations is reported in each figure. For each group of comparisons (active, SRBS, and MW), the number of collocations that matches the data set with the fewest collocations are randomly selected to avoid data sets with fewer collocations appearing noisier.

[50] Because the data sets are based on instruments with different sensitivities, it only makes sense to discuss comparisons in IWP ranges where both the compared data set and DARDAR are sensitive to clouds. In IWP value ranges where either the compared data set or DARDAR are not sensitive to or cannot retrieve the cloud, the data sets are expected to be uncorrelated. Using Figure 6 (described in section 4.4), we show that the median IWP of the observed three data sets and the median of DARDAR values (shown as thick gray and black lines, respectively) appear to rapidly diverge for IWP values less than approximately 30 g m⁻² and diverge slightly again for values greater than approximately 2000 g m⁻². On one extreme, if the median lines are near perpendicular to one another, the measurements are uncorrelated within that IWP range, whereas if the median lines are parallel to each other, the measurements are correlated. For very low IWP values (<20 g m⁻²), all data set comparisons are uncorrelated. However, the median lines always converge at some point for increasing IWP values. Therefore, for each data set, the decision on the valid IWP ranges is based on where the median lines converge to being very close to each other and end at high IWP values if the median lines strongly diverge from each other again. The results of this IWP dynamic range analysis are summarized in Table 3.

4.2.2. Systematic and Random Errors

For easier comparison among the various data sets, the median and percentiles of the log ratio of IWP and DARDAR IWP are displayed together in Figures 5, 7, and 9 (described in sections 4.3, 4.4, and 4.5) as a function of DARDAR IWP. Thick lines show the median, and the same colored thin lines depict the random spread using the 16th/84th percentiles (see section 3.1 for description on random errors). Solid lines are drawn in the IWP ranges deemed valid to compare the two data sets (Table 3), and the lines become dashed outside this range. Horizontal thin gray lines showing factor of 2 steps in systematic error and a thick line showing equality are added for clarity.

[52] The 16th/84th percentiles shown in Figures 5, 7, and 9 represent the random error in each comparison. As mentioned earlier, the chosen percentiles match the standard deviation if the distribution is Gaussian.

4.2.3. Are IWP Distributions Lognormal?

In fact, the log-ratio distributions are often found to be close to Gaussian, as illustrated for the PATMOS-x versus DARDAR comparison in Figure 3. It shows the median and percentiles, as well as the mean and standard deviation in terms of log ratio between PATMOS-x and DARDAR. As Figure 3 shows, the log distribution of IWP is nearly Gaussian across the whole IWP range. This applies to most data sets in this study (data not shown). The fact that the log distributions are nearly Gaussian justifies using the
standard deviation (or percentiles) of the log ratio as a random error estimate (see section 3.1). Nonetheless, because the distributions are not exactly Gaussian, the median/percentile approach is chosen for better robustness. It means that the errors generally scale with the IWP value, and relative errors are more appropriate than absolute errors to characterize IWP.

4.3. Active Sensors

Using the median lines in the two-dimensional histogram of IWP in Figure 4, DARDAR is determined comparable with the active data sets in the IWP range of approximately 10 to 9000 g m\(^{-2}\), i.e., a large part of the range retrieved by DARDAR. However, the expected close level of agreement in terms of IWP of RO and IORO to DARDAR is due to large ice particles dominating the IWP (column integrated IWC). In terms of IWC, especially where there is low IWC, the CloudSat data sets diverge from DARDAR [Stein et al., 2011].

When comparing active data sets, one needs to note that DARDAR is based on the same radar measurements as the CloudSat IWP products, so the random errors of these data sets are likely correlated; therefore, nothing can be said about the random errors of the individual data sets. However, because DARDAR uses lidar measurements in the retrieval, and the CloudSat data sets assume different particle phase assumptions, they are appreciably different from each other, as shown in Figure 5. This figure also shows that for thin clouds in the range of 30 to 40 g m\(^{-2}\), RO and IORO both report more than a factor of 2 less IWP than DARDAR. This demonstrates the strong impact of lidar measurements on DARDAR in this range. This was corroborated using the instrument flags of DARDAR. For instance, at 10 g m\(^{-2}\), on average, the fraction of IWC values in each vertical profile retrieved solely from lidar, radar, or from a mixture of both measurements is approximately 55%, 20%, and 25%, respectively.

The IORO and RO data sets converge with DARDAR for clouds around 500 g m\(^{-2}\) as the influence of lidar measurements becomes less dominating. The IORO data set reports more IWP than DARDAR between approximately 100 and 800 g m\(^{-2}\), which likely is due to an overestimation of IWP because all cloud water content where the temperature is colder than 0°C is assumed to be ice phase (see section 2). For values larger than 800 g m\(^{-2}\), RO and IORO report lower IWP than DARDAR, but RO decreases more than IORO.

DARDAR detects the cloud phase directly using the intensive backscatter by supercooled water compared with ice in the lidar signal. When the lidar signal is attenuated before the liquid phase level, the radar echo is assumed to be dominated by ice. Thus, only the IORO product was
considered in the Stein et al. [2011] study because it is a closer comparison with DARDAR.

### 4.4. Passive VIS/NIR Sensors

When comparing passive measurements of IWP with measurements from active instruments, one must bear in mind the limitations of IWP retrievals from passive sensors. In addition to the uncertainties introduced in assuming homogeneous clouds (see section 2.2), only measurements determined to be ice phase are used to retrieve ice-cloud properties, and the whole column is then assumed to be ice. These problems are known to cause large uncertainties in SRBS retrievals [e.g., Stein et al., 2011]. By comparison, as mentioned earlier, DARDAR profiles can contain both liquid and ice particles, and only the vertical bins that contain ice (IWC) are considered in the integration.

Three data sets using reflected solar measurements to derive IWP are shown in Figure 6. Using the median lines to find the IWP ranges where it may be valid to make comparisons with DARDAR, the SRBS data sets are determined to cover a large range of IWP values. In general, the retrievals at low IWP values (approximately \( \leq 30 \text{ g m}^{-2} \)) are not correlated at all between the compared data sets and DARDAR, meaning that either neither is sensitive to such thin clouds or only one of them is (probably DARDAR). There appears to be no clear upper limit to the SRBS data sets, although the data sets are slightly less correlated above approximately 2000 g m\(^{-2}\). At greater than these values, the systematic errors between the SRBS data sets and DARDAR rapidly increase. Because all three data sets are based on similar instruments, the retrievals also have approximately the same IWP ranges. However, there are still some notable differences between these data sets, which are best visualized in Figure 7.

Figure 7 shows the median and percentiles of the ratio between SRBS data sets and DARDAR IWP as a function of DARDAR IWP. Curiously, within the valid ranges, MODIS reports more IWP than DARDAR from approximately 20 to 100 g m\(^{-2}\), whereas CMSAF and PATMOS-x do not. In the vicinity of 20 g m\(^{-2}\), MODIS has a factor of 2 higher IWP values, and at the end of the range around 5500 g m\(^{-2}\),
MODIS reports IWP that is lower by around a factor of 4. PATMOS-x and CMSAF have a systematic negative bias for all DARDAR values, but from approximately 100 g m\(^{-2}\) the bias of PATMOS-x is nearly the same as the bias of MODIS, whereas CMSAF has nearly a factor of 2 more systematic bias than MODIS and PATMOS-x for the whole IWP range.

The combined random uncertainty (see section 3.1) of MODIS-DARDAR, depicted as thin red lines in Figure 7, is very constant and nearly parallel with the median throughout the valid range and ranges from a factor of 1.5 to 2. PATMOS-x-DARDAR has a somewhat larger but also quite constant random uncertainty (approximately a factor of 2–3) for the whole IWP range. CMSAF has a low random uncertainty below approximately 200 g m\(^{-2}\) (approximately a factor of 1.3), but then steadily increases to around a factor of 3 to 4 for the highest IWP values.

Taking MODIS-DARDAR as an example, Cooper et al. [2003, 2007] showed that the uncertainties in MODIS retrievals of \(r_c\), used to derive IWP, are of the order of 30 to 40%, and DARDAR’s random uncertainty was shown to have an overall random uncertainty of around 20 to 50% (see section 4.1). However, such general uncertainty estimates per data set are inadequate for IWP retrievals because they are clearly not valid for all IWP ranges. Using the approach described in this article (see section 3.1), the combined random error for the entire valid IWP range can be estimated as a function of IWP.

Assuming that the random errors between these two data sets are uncorrelated, the maximum random uncertainty of either data set is concluded to be smaller than a factor of 2 (100%). Because the combined error of MODIS-DARDAR is not too large, covering a wide range of IWP values, this demonstrates the likely strength of both data sets. This also helps to justify our choice of DARDAR as the reference data set, because if it was a bad reference, the combined uncertainties between DARDAR and any other data set would always be large. If both were assumed to have the same uncertainty, using equation 2, their uncertainty would be smaller than 1.63. An uncertainty of 1.63 on the ratio corresponds to a relative error of 63%, which is not too far from the earlier

![Figure 7](image-url)  
**Figure 7.** IWP from passive SRBS data sets. The median and “pseudo 1-sigma” of the ratios of MODIS versus DARDAR, PATMOS-x versus DARDAR, and CMSAF versus DARDAR are shown together. Thick lines indicate medians; think lines indicate “pseudo 1-sigma”; dashed lines indicate IWP ranges where the data sets are uncorrelated.

![Figure 8](image-url)  
**Figure 8.** Tropical IWP comparison 2007 for passive microwave techniques: two-dimensional histograms as in Figure 4. MSPPS IWP (top left), MiRS graupel water path (top right), and CMO IWP (bottom).
uncertainties from Cooper et al. [2003] and Austin et al. [2009] cited earlier for MODIS and DARDAR, respectively. Errors caused by collocations are neglected in this argument.

The differences between the SRBS data sets can be explained by mainly three factors. First, although PATMOS-x and CMSAF are based on the same measurements, the difference between them is systematic and large. This is due to the choice of ice particle model (see section 2.2.3). The largest uncertainties in these retrievals are a direct result of the assumed microphysics [Zhang et al., 2009]. Second, because MODIS IWP retrieval algorithm can use five channels compared with two for CMSAF and PATMOS-x, it is conceivable that it has better constrained retrievals. Third, the different sizes of the pseudo footprints of roughly 10 × 10 km, 4 × 1 km, and 1 × 1 km for PATMOS-x, CMSAF, and MODIS, respectively, may also play a role, and because larger footprints should have more collocation errors (see section 3), this may be the reason PATMOS-x has a larger random uncertainty than MODIS, although the systematic errors are the same.

4.5. Passive Microwave Measurements

Three data sets that retrieve IWP using passive microwave data are shown in Figure 8. Using the medians as before, comparisons of IWP can be made in the ranges from approximately 900 g m$^{-2}$ for MSPPS (left), from approximately 700 g m$^{-2}$ for MiRS (right), and from approximately 10 g m$^{-2}$ for CMO. There is no clear upper bound to valid IWP ranges for these data sets. It is well established that IWP retrievals using passive microwave are only sensitive to clouds that have precipitation-sized ice particles, and this is reflected in this comparison. As shown in the histogram of CMO versus DARDAR in Figure 8, CMO reports IWP even where DARDAR reports very low IWP values. This is purely an artifact of this current version of the data set because the retrieval will always retrieve IWP greater than 0.

As shown in both Figures 8 and 9, the systematic error between MSPPS versus DARDAR and the systematic error between MiRS versus DARDAR is very large even for clouds where both the DARDAR and these microwave data sets have sensitivity to clouds. The MSPPS and MiRS data sets are probably not suited for such a comparison, but it is worth noting that when clouds have an IWP larger than approximately 2000 g m$^{-2}$, they do appear to be fairly correlated with DARDAR, albeit offset to a factor of 5 to 6 lower IWP values. The correlation improves with increasing IWP because of the steady increase of precipitation. Once the IWP is larger than approximately 2000 g m$^{-2}$, the random error of the collocated microwave data sets is not particularly large. MSPPS has a random error of mostly less than a factor of 1.5, and MiRS has a larger random error. The increased random error of MiRS is likely due to the much larger footprint sizes of MiRS compared with MSPPS (see section 3.1).

The third microwave data set, CMO stands out because of its different retrieval approach and higher frequency microwave channels, and the valid IWP range defined here is very large (see Figures 8 and 9). Note that CMO is developed based on collocations with CPR; therefore, the random errors are not assumed to be uncorrelated. CMO reports higher values than DARDAR in the low IWP range. CMO is known not to be retrieving clouds here because it is insensitive to clouds in this range. The version compared in this study does not include any cloud detection and retrieves IWP in logarithmic space. Therefore, it retrieves a low value of IWP even when there is no ice at all. A good way of cutting off those values is still under development. Although the data set is still in a fairly early development stage, it already looks promising.

Table 3. Approximate Valid IWP Ranges: For Each Data Set Based on the IWP Bins or DARDAR, and the Range Selection Criteria Presented in Section 4.2.14

<table>
<thead>
<tr>
<th>Data Set</th>
<th>Minimum IWP [g m$^{-2}$]</th>
<th>Maximum IWP [g m$^{-2}$]</th>
</tr>
</thead>
<tbody>
<tr>
<td>RO</td>
<td>10</td>
<td>9000</td>
</tr>
<tr>
<td>IORO</td>
<td>10</td>
<td>9000</td>
</tr>
<tr>
<td>MODIS</td>
<td>40</td>
<td>--</td>
</tr>
<tr>
<td>CMSAF</td>
<td>40</td>
<td>--</td>
</tr>
<tr>
<td>PATMOS-x</td>
<td>40</td>
<td>--</td>
</tr>
<tr>
<td>MSPPS</td>
<td>900</td>
<td>8000</td>
</tr>
<tr>
<td>MiRS</td>
<td>700</td>
<td>--</td>
</tr>
<tr>
<td>CMO</td>
<td>10</td>
<td>--</td>
</tr>
</tbody>
</table>

*The missing upper bounds for CMO and MiRS indicate that a clear upper limit to their valid ranges was not found.

5. Conclusions

This article is a necessary continuation of the Eliasson et al. [2011] study where the satellite observations were compared in terms of monthly mean IWP. The results presented in this article help explain the differences between the L3 observational data sets found in Eliasson et al. [2011], which could only be inferred using the monthly mean data. However, comparing monthly mean products is problematic because different approaches are used to go from an L2 (“footprint” resolution) to L3 (gridded monthly mean) prod-

Figure 9. IWP from passive microwave data sets. The median and “pseudo 1-sigma” of the ratios of MSPPS versus DARDAR, MiRS versus DARDAR, and CMO versus DARDAR are shown together. Thick lines indicate medians; thin lines indicate “pseudo 1-sigma”; dashed lines indicate IWP ranges where the data sets are uncorrelated. Note that the y axis differs from Figures 5 and 7.
uct. IWP derived from different techniques were compared with DARDAR at a “footprint” resolution, using data from 2007 between ±30° latitude. In this study, DARDAR is assumed to provide the most accurate retrieval of IWP, because of using the combination of CALIPSO and CloudSat measurements. DARDAR IWP is further assumed to be valid over a wide range of IWP, thus being ideal for the investigations presented. Using the fractional errors of DARDAR IWC, we found that the 1-sigma fractional errors in column integrated IWP range between 20 and 50% and are generally decreasing with increasing IWP (see section 4.1).

The range of IWP where both DARDAR and the collocated data set are sensitive to clouds are first identified for each collocated data set and summarized in Table 3. The focus of this study is on the assessment of various available IWP data sets, i.e., investigating their valid IWP range using DARDAR as reference. In these IWP ranges, we have looked at the systematic bias of IWP and the combined random errors. These are then used to assess the quality of the data sets in Figures 5, 7, and 9.

The combined uncertainties in log space of the collocated data sets are reported in section 4. It is shown that data sets based on solar reflectances can be compared with DARDAR over a large IWP range starting from approximately 30 g m⁻². The combined uncertainty of DARDAR and MODIS is less than a factor of 2, and this sets the upper limit for the random error of either data set. CMSAF and PATMOS-x have larger combined uncertainties (see Figure 7).

All SRBS data sets’ systematic errors increase to a factor of 4 negative bias as their signals start to attenuate at high IWP values. CMSAF is biased low compared with MODIS and PATMOS-x because of different assumptions on the ice particle models in the retrievals (see section 2.2). Data sets based on passive microwave measurements are only comparable with DARDAR in ranges starting from approximately 700 g m⁻². In this range, their combined uncertainty is not larger than the SRBS data sets despite their larger footprints, but the systematic errors are very large. The bias of MSPPS/MIRS compared with DARDAR ranges from a factor of 5 to 7 and 8 to 10, respectively.

We have concluded that all assessed data sets provide valuable information on IWP. The results presented in Table 3 suggest that a synergistic passive retrieval using microwave and shortwave may be able to retrieve IWP for a larger range of values than retrievals based on either alone. Such a retrieval could be developed in a similar way to the CMO data set with a DARDAR reference, therefore truly exploiting the benefits of active and passive shortwave and microwave, and thus constrain IWP as much as possible from a spaceborne platform.

It is encouraging that MODIS, CMSAF, and PATMOS-x show reasonably good agreement over a large IWP range despite the remaining large systematic biases. This is especially encouraging because CMSAF and PATMOS-x are based on AVHRR measurements that span over 30 years. These data sets are suitable for long-term climate applications. Future work may also include comparing the data sets separated into, e.g., cloud types or specific atmospheric scenarios.

Acknowledgments. We thank the Swedish Science Council for cofunding the earlier phase and the Swedish National Space Board for cofunding the later phase of this study. We thank the ICARE Data and Services Center and L-AADS web for providing access to the data used in this study. We thank Patrick Eriksson from Chalmers University of Technology, Sweden, and Andy Heidinger from the Center for Satellite Applications and Research for numerous e-mail discussions related to this study. We thank the Swedish National Research School in Space Technology for offering courses that have helped J.E. in G.H. in their research.

References


Paper V

SPARE-ICE: synergistic IWP from passive operational sensors

Manuscript submitted 17 August 2013

SPARE-ICE: synergistic IWP from passive operational sensors

G. Holl¹, S. Eliasson¹,², J. Mendrok¹, S. A. Buehler¹

Corresponding author: G. Holl, Division of Space Technology; Department of Computer Science, Electrical, and Space Engineering; Luleå University of Technology; Kiruna; Sweden. gerrit.holl@ltu.se

¹Division of Space Technology,
Department of Computer Science, Electrical and Space Engineering, Luleå University of Technology, Kiruna, Sweden

²Atmospheric Remote Sensing Unit,
Research and Development Department,
Swedish Meteorological and Hydrological Institute (SMHI), Norrköping, Sweden
Abstract. This article presents SPARE-ICE, the Synergistic Passive Atmospheric Retrieval Experiment-ICE. SPARE-ICE is the first Ice Water Path (IWP) product combining solar, infrared, and microwave radiances. By using only passive operational sensors, the SPARE-ICE retrieval can be used to process data from at least the NOAA-15 – 19 and MetOp satellites, obtaining time series from 1998 onward. The retrieval is developed using collocations between passive operational sensors (solar, terrestrial infrared, microwave), the CloudSat radar and the CALIPSO lidar. The collocations form a retrieval database matching passive reflectances and radiances against the existing active combined radar-lidar product 2C-ICE. With this retrieval database, we train a pair of artificial neural networks to detect clouds and retrieve IWP. We show that a retrieval using solar, terrestrial infrared, and microwave performs better than a retrieval using only one or two of these techniques. The median fractional error between SPARE-ICE and 2C-ICE is around a factor 2, a figure similar to the random error between 2C-ICE Ice Water Content (IWC) and in-situ measurements. A comparison of SPARE-ICE with MODIS and MSPPS indicates that SPARE-ICE performs well even in difficult conditions. SPARE-ICE is available for public use.
1. Introduction

The systematic and global observation of cloud properties is essential for the understanding of the climate system [WMO, 2010]. A fundamental parameter in estimates of atmospheric ice is Ice Water Path (IWP) [g/m$^2$], defined as the vertical integral of Ice Water Content (IWC) [g/m$^3$] or the atmospheric column density of ice. Model estimates of IWP vary by more than an order of magnitude [Waliser et al., 2009], and show deviating spatial distributions [Eliasson et al., 2011]. One of the reasons for poor model performance is the lack of good constraints.

Only space-based remote sensing can provide global IWP measurements, but the remote sensing of atmospheric ice is severely underconstrained. Depending on the technique, the measured quantity (such as reflectance, radiance, or radar backscatter) depends in a complex manner on the vertical distribution of IWC, particle sizes, particle shapes, temperature, humidity, and other quantities. Individual measurements contain insufficient information to retrieve all those quantities directly and independently, and many assumptions need to be made [e.g. Stephens and Kummerow, 2007]. This leads to a considerable uncertainty in retrieved IWP. Moreover, the true uncertainty in IWP retrievals is very hard to quantify, not least because of the various difficulties of performing reliable in-situ measurements.

Current IWP products are based on reflected solar radiation [e.g. Rossow and Schiffer, 1991; King et al., 1997], passive microwave [e.g. Ferraro et al., 2005; Boukabara et al., 2011] or on active sensors such as radar and lidar [e.g. Stephens et al., 2002, 2008; Delanoë and Hogan, 2010; Deng et al., 2010]. Passive sub-millimetre measurements have been proposed
for their ability to sample the size distribution and their sensitivity to relevant particle
sizes, which allows for sensing the full ice column [e.g. Buehler et al., 2012], but no
down-looking instrument is currently in space. Active techniques have the capability to
determine the vertical structure of atmospheric ice, and IWP retrieved from active sensors
is likely more accurate than any IWP retrieval from passive measurements, although the
error remains very difficult to characterise in absolute terms. However, active sensors are
costly, require a lot of energy, and the technology is less mature than for many passive
technologies. They tend to have a small footprint and are (so far) exclusively carried on
scientific platforms, which have a limited lifetime. On the other hand, passive sensors
exist on both scientific and operational satellites. In particular, sensors on operational
satellites provide a much better spatial and temporal coverage than active sensors. A
nearly identical set of instruments is available the National Oceanic and Atmospheric
Administration (NOAA) 15 – 19 satellite series, as well as on MetOp-A, -B, and the
future MetOp-C satellite. Hence, any product based on those sensors can be readily
processed from 1998 until the present (and beyond). Each individual satellite provides
daily near-global coverage, Therefore, the potential data volume for such a product vastly
outstrips the data volume from active sensors, allowing entirely different applications.

Eliasson et al. [2013] use strict collocations between radar, lidar, solar, and microwave
sensors, in order to systematically compare spaceborne IWP retrievals on a footprint-
level basis. Thus, they quantify where various techniques share IWP sensitivity with
the active combined radar-lidar product raDAR/liDAR (DARDAR) [Delanoë and Hogan,
2010]. They confirm that different techniques have sensitivities for different IWP ranges.
A combination of different passive techniques may be capable of exploiting synergies, providing a more accurate IWP retrieval than by using any individual part of the spectrum. Synergies have been exploited in e.g. retrievals of rainfall [e.g. Kidd et al., 2003; Marzano et al., 2004; Rapp et al., 2009; Kidd and Levizzani, 2011] and liquid clouds [Taylor and English, 1995]. For retrievals of properties for atmospheric ice, synergies between radar and lidar have been used in at least two products [Delanoë and Hogan, 2010; Deng et al., 2010], but synergies between different passive techniques appear underutilised.

This article presents SPARE-ICE, the first IWP product exploiting the synergy between reflected solar, terrestrial infrared, and microwave. SPARE-ICE will be publicly available through the World Data Center for Remote Sensing of the ATmosphere (WDC-RSAT) under the Open Data Commons Attribution License (Note: data will be uploaded after manuscript submission and a persistent address, including DOI, will be added in the final manuscript).

The use of collocations between active, scientific instruments and passive, operational instruments to learn more about the latter has a limited history. Holl et al. [2010] introduce collocations between CloudSat Cloud Profiling Radar (CPR) and NOAA-18 Microwave Humidity Sounder (MHS) and briefly present three potential applications. Liu and Seo [2013] use radar reflectivities from CloudSat CPR as a proxy reference for the detection of snowfall over snow-covered surfaces. Other statistical retrievals do not use collocations, but retrieval simulations [e.g. Evans et al., 2012; Jiménez et al., 2007] or output from general circulation models (GCMs) [e.g. Surussavadee and Staelin, 2008].

The retrieval presented in this study builds on a collocation toolkit developed by Holl et al. [2010]. Whereas Eliasson et al. [2013] used collocations between IWP retrievals to...
characterise different sensitivities, we use collocations between the same sets of satellites
to develop a new retrieval algorithm. In the present study, we match reflectances from
Advanced Very High Resolution Radiometer (AVHRR) solar reflected channels, radiances
from AVHRR terrestrial infrared channels, and radiances from MHS humidity channels,
with IWP as reported by the official CloudSat product 2C-ICE [Deng et al., 2010] as
distributed by the CloudSat Data Processing Center (CDPC). Then, we use an Artificial
Neural Network (ANN) to train an IWP retrieval from passive measurements. Thus, we
reproduce actively retrieved IWP using only passive, operational sensors.

The primary aim of the study is to develop and analyse an improved IWP retrieval
from passive operational sensors. A secondary aim of the study is to quantify synergies
between different passive operational sensors.

The following sections describe the work in detail. In section 2, we describe the instru-
ments used, the collocations, and the retrieval approach. Section 3 presents the results
for the investigation into synergies and shows SPARE-ICE retrievals for three case-studies
and a 2007 gridded mean. In section 4, we discuss the results. Finally in section 5, we
recommend tasks for further work, and conclude the article.

2. Methodology

In the following subsections, we will describe the SPARE-ICE retrieval algorithm in
detail. The overall approach is summarised in Figure 1. Broadly, we consider two phases:

1. The development of the retrieval system. This consists of:

   (i) Obtaining collocations as described by Holl et al. [2010],
(ii) selecting various combinations of input channels and auxiliary information, in
order to investigate synergies,

(iii) training of a pair of neural networks for each of those selections,

(iv) and finally, testing the performance of this pair of neural networks.

This step is repeated for many different combinations of input channels and auxiliary
information.

2. Processing measurements. This involves:

(i) Collecting and combining measurements for the needed sensors and channels,

(ii) selecting the input channels and auxiliary information,

(iii) applying the relevant set of neural networks to obtain IWP.

These steps are described in more detail in the following subsections.

2.1. Instruments and products

The instruments used in the study can be grouped in two categories: active and passive.
The active instruments are on scientific satellites inside the satellite constellation known
as the “A-Train”, whereas the passive instruments that were used are on the operational
satellite NOAA-18.

2.1.1. Radar and lidar

The CloudSat CPR is a 94 GHz radar launched in 2006 [Stephens et al., 2002, 2008].
It measures vertical profiles of backscattering, in particular from clouds and precipita-
tion. It observes only at (near) nadir with a footprint of 1.1 km. Operational products
include profiles and integrated quantities for ice microphysical properties derived from
backscattering and temperature information [Austin et al., 2009], in a product published
as 2B-CWC-RO by the CDPC, including IWC, IWP, and associated uncertainties. Retrieved profiles of IWC and Liquid Water Content (LWC), and therefore retrieved IWP, have a considerably error, reported by Austin et al. [2009] to be around 40%, but that is likely an underestimate. The error is large and hard to quantify, due to poorly known ice particle sizes and shapes, and due to a poor characterisation of supercooled liquid and melting particles.

The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) is a lidar, carried on the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO), measuring backscattering at 532 nm and 1064 nm from aerosols and cloud tops [Winker et al., 2003, 2009]. It measures only at nadir and has a footprint of 70 m.

CloudSat and CALIPSO fly close to each other in the A-Train satellite constellation. A combination of radar and lidar, or specifically between CPR and CALIOP, provides a more complete measurement of the ice column than any other presently existing space-borne measurement. Radar can penetrate thick systems of precipitating clouds, but is mainly sensitive to large particles and does not detect small ones. Lidar, on the other hand, is sensitive to small particles, but gets attenuated quickly. Therefore, radar and lidar are complementary. Two products exploiting this are publicly available at present. Delanoë and Hogan [2008] describe the radar-lidar product DARDAR, developed further in Delanoë and Hogan [2010]. A slightly newer product based on the same sensors is described by Deng et al. [2010] and published as an official CloudSat product by the name 2C-ICE. Both DARDAR and 2C-ICE retrieve IWP on the CPR footprint. For the present study, we choose to use 2C-ICE, but the principle can also be carried out with DARDAR.
or any other reference dataset. Deng et al. [2013] briefly reviews the differences between DARDAR and 2C-ICE.

Although DARDAR and 2C-ICE can, at least potentially, determine IWP better than any passive retrieval, their temporal and spatial coverage is inherently limited. Active systems require a lot more energy than passive ones, and both CPR and CALIOP observe only at an (almost) nadir geometry. The technology is less established and the lifetime for an active system may be less than for a passive one. Although their lifetimes have already exceeded design and expectations considerably, both CPR and CALIOP will come to an end. EarthCare [Bezy et al., 2005] is scheduled for launch in 2015 and will provide a continuation, but ultimately, active systems are not operational and hence, their continued existence is not guaranteed.

2.1.2. Passive systems

Considering the disadvantages of active systems, there is a need for continued development of IWP products based on passive sensors. In particular, a product based on passive operational sensors can be processed far back in time. In our approach, we use a combination of solar, terrestrial infrared, and microwave sensors, allowing for data to be processed back to 1998.

2.1.2.1. Reflected solar radiation and terrestrial infrared radiation

AVHRR [Cracknell, 1997] was first launched on-board NOAA-6 in 1979. Although the precise configuration has varied somewhat throughout the years, there may be no other meteorological satellite instrument with as many copies and as long a lifetime as AVHRR. Recent editions (from NOAA-15 onward) carry six channels, of which five are simultaneously operated. Three channels (1, 2, and 3A) measure reflected solar radiation. Two
channels (4 and 5) measure terrestrial infrared radiation (sometimes referred to as “thermal”). Channel 3B, covering 3.55–3.93 µm, measures a signal containing both reflected solar and emitted terrestrial radiation (Allen et al. [1990] present a technique to separate the two). Channel 3 can be switched between 3A and 3B. See Table 1 for the spectral characteristics of AVHRR data used in this study. Due to downlink limitations, global data for the NOAA satellites are only available at a limited resolution, through the Global Area Coverage (GAC) product. AVHRR GAC data only contains every third AVHRR scan, and averages four adjacent measurements [Robel et al., 2009], resulting in effective discontiguous footprints of 1 × 4 km², as illustrated in Figure 2. Both reflected solar and emitted terrestrial radiation contain information about clouds and other aspects of the Earth system. AVHRR has been used for a very wide variety of applications, including (but not limited to) the retrieval of cloud properties, such as in the Pathfinder ATMOSpheric (PATMOS)-X dataset, [Stowe et al., 2002; Jacobowitz et al., 2003; Heidinger and Pavolonis, 2009] available from the early 1980s to present.

The High-resolution Infrared Radiation Sounder (HIRS) also measures terrestrial radiation, and does so in more spectral bands than AVHRR. In Holl et al. [2010], we briefly explored retrieving IWP from a combination of HIRS and MHS. However, as shown by Figure 2, HIRS footprints are highly non-contiguous, and therefore HIRS is less suitable for usage in combination with other sensors and not used in SPARE-ICE.

2.1.2.2. Microwave

Advanced Microwave Sounding Unit (AMSU) and MHS are microwave sounders that have been carried on-board NOAA satellites since the launch of NOAA-15 in 1998. AMSU-B and MHS measure at frequencies from 89 to 190 GHz. At the highest frequencies,
around the 183 GHz water vapour absorption line, large ice particles can be detected from a spaceborne platform due to the scattering of radiation emitted by water vapour below the cloud base.

### 2.2. Collocations

In the context of this study, a collocation is an event where two spaceborne sensors observe the same place at almost the same time. Holl et al. [2010] have shown that, due to the proximity of NOAA-18 to the A-Train (the constellation including CloudSat, CALIPSO, and other satellites), collocations between sensors on NOAA-18 and CloudSat occur globally. The collocation algorithm is described in Holl et al. [2010] and updated in John et al. [2012].

In this study, we use collocations to obtain a database for training artificial neural networks. Figure 1 illustrates the role of collocations in the SPARE-ICE retrieval development, and Table 1 gives an overview of all measurements used.

For obtaining the collocations, we use the flexible and sophisticated toolkit initially developed by Holl et al. [2010], but since developed considerably further. The collocation algorithm is fully based on time and geolocation information contained in the data files. Others, such as Nagle and Holz [2009], use orbital parameters to calculate when collocations occur, before looking at data files.

We collocate NOAA-18 MHS with CloudSat CPR. To this, we add AVHRR measurements and auxiliary information. From CPR, we use the 2C-ICE product. Although this uses both CPR and CALIOP, it is retrieved on the CPR footprint, so we do not need to explicitly collocate with CALIOP.
As shown in Figure 2, the footprints for CPR and AVHRR are much smaller than the ones for MHS. Therefore, we “collapse” CPR and AVHRR onto the MHS, for all secondary footprints within 7.5 km and 10 minutes of the primary, as defined by the centre of the measurement. For each MHS footprint, we store the number of collocated CPR and AVHRR footprints, as well as the mean and standard deviation for each AVHRR channel measurement (reflection or radiance), as well as for IWP from both 2C-ICE and DARDAR. We also store the fraction of 2C-ICE or DARDAR footprints with IWP > 10 g/m². Even when the CloudSat ground track passes exactly through the centre of an MHS footprint, CPR covers less than 10% of MHS, and a random error from imprecise collocations is unavoidable. All subsequent processing is done on MHS footprints, so no oversampling occurs.

Although collocations occur globally, they occur only at particular ranges of MHS viewing angles, where the exact range is a function of latitude. Therefore, collocations do not fill the entire space of latitude and MHS viewing angle, which would be desirable for our neural network training approach described below. Additionally, collocations are not equally distributed over the globe, but occur more frequently near the poles, than around the equator (like any measurements from polar orbiting satellites). Due to the non-uniform distribution, collocations also contain correlations that do not occur in non-collocated data. Of course, even in non-collocated data, quantities like latitude and surface temperature may show correlations (of either sign) with IWP. In the present study, we use roughly $1.3 \times 10^6$ collapsed collocations with valid radiances occurring through 2007. Of those, approximately $5.8 \times 10^5$ have an MHS-averaged 2C-ICE IWP larger than zero. For details on the collocations, refer to Holl et al. [2010].
2.3. Retrieval development

The aim of the SPARE-ICE retrieval is to make a product that is globally available and at the full combined scan range of AVHRR and MHS. In developing SPARE-ICE, we also develop intermediate products using only a subset of channels and auxiliary information, as indicated in Figure 1. The process of training and testing the retrieval algorithm described below is repeated for each of these subsets.

As we find that reflected solar radiation improves the performance, as will be shown in section 3.1 and Figure 6, SPARE-ICE is limited to daytime only.

In the following pages, we describe our retrieval approach.

2.3.1. Neural networks

We use the collocations as a retrieval database for training Artificial Neural Networks (ANNs), as illustrated by Figure 1. An ANN is a network of interconnected processing units called nodes. We use Multi Layer Perceptrons (MLPs), feed-forward neural networks where the nodes are divided in multiple layers. MLPs are commonly used to statistically characterise the complex relation between a set of inputs and one or more quantities of interest, called targets. This makes MLPs popular for geophysical retrievals [e.g. Krasnopolsky, 2007]. We use an MLP with three layers. For the first layer, called the input layer, we use one node for each input quantity. See Table 1 for an overview of all input quantities used. The second layer is the so-called hidden layer, in which we use ten nodes. The third and final layer is the output layer, and contains a single node corresponding to our target, the logarithm of 2C-ICE IWP. All nodes in layer $n$ are connected to all nodes in layers $n+1$ via activation functions. The training algorithm iteratively assigns weights and biases to each activation function. This is done by minimising the
cost function, which is defined as the mean square error difference between the neural network output and the targets. The trained neural network forms a nonlinear regression, mapping the set of inputs to one or more target quantities. Such networks are prone to “overfitting”, where the network learns not only the statistical relation between the inputs and the target, but also between input noise and the target. To prevent overfitting, part of the data are set apart as validation data. If, for $N$ subsequent iterations, the fitting improves (i.e. the cost function becomes smaller) with the training data, but gets worse with the testing data, then the training is considered finished (this approach is called the “early stopping” criterion). In our implementation, we choose $N = 5$. Note that there are stochastic elements in the training (for example, the initial values of weights and biases) and that there is no guarantee that the fitted network is the best possible one. For the implementation, we use the Mathworks MATLAB® Neural Network toolbox V8.0.1 (R2013b).

For the training, we use all collocations where all inputs and the target have valid measurements, occurring during 2007 (see Table 1 for an overview of input measurements). We divide the collocations in three subsets: training, validation, and testing. The training data are used to minimise the cost function, and the validation data are used to prevent overfitting as described above. The testing data are not used in the training, and serve as independent data to characterise the performance of the regression. From the collocation database, we draw 200,000 random samples according to a uniform distribution. From those samples, we randomly assign $2/3$ to be used for training and $1/3$ to be used for validation. The remaining collocations are used for testing the neural network.
We split up the retrieval in two neural networks, as illustrated by Figure 1. The two networks operate independently, and the final IWP value is a result of combining the output from both nets.

1. One network classifies a scene as either cloudy or non-cloudy. During the training, we define a scene as cloudy if 2C-ICE IWP averaged over the covered part of the MHS footprint is larger than a threshold value $t$. From the sensitivity study by Eliasson et al. [2013] and confirmed later in this study (see Figure 7), $t = 10 \text{ g/m}^2$ is the lower sensitivity limit we can expect for SPARE-ICE, and is therefore a good choice as a threshold. During the retrieval, the classification network retrieves a value $p \in (0, 1)$, that we interpret as a cloud probability. For each value $p > c$, where $c$ is the cutoff value, we consider the scene as cloudy and proceed to retrieve IWP. For $p < c$, we consider a scene free of ice. For the classification network, we leave out microwave measurements, because microwave has no added value over solar and infrared measurements as far as cloud detection is concerned (see also Table 1).

2. The other network retrieves “raw” IWP. Because of the high dynamic range of IWP, and because we independently assess whether or not a scene is cloudy, we retrieve IWP in log space ($10 \log \text{IWP}$). Training and validation data are limited to cases with IWP $> 0$, but are otherwise a subset of the data used in the cloud classification network. The distribution of IWP is much closer to log-normal than to normal [Eliasson et al., 2013], which is beneficial to the neural network fitting algorithm.

When performing the actual retrieval, both the classification network and the network retrieving “raw” IWP are applied for all measurements. Then, we determine $IWP = 0$ where the classification network retrieves $p < c$, and $IWP = \text{“raw” IWP}$ otherwise.
2.3.2. Error analysis

Neural network retrievals do not provide a direct uncertainty estimate for an individual retrieval. Other methods do, but even for optimal estimation based retrievals, the reliability of the uncertainty estimate is only as good as the forward model and the error characterisation therein, and may very well be an underestimate of the true error, i.e. the error compared to the unknown truth. A reliable error estimate would require a ground truth, that is difficult to obtain for IWP. In our approach, we use the aforementioned testing data, that were not used in the training, to get an estimate of the error as a function of IWP.

2.3.2.1. Cloud classification error

For the cloud classification network, we explore the rate of false positives and the rate of false negatives. A false positive is an occasion where the neural net obtains a cloud \((p > c)\) while the independent reference states that there is not (according to the cloud definition described above). A false negative is an occasion where the neural net obtains no cloud \((p < c)\) while the reference states there is. A higher cutoff will decrease the rate of false positives, but increase the rate of false negatives. The final choice for \(c\) is somewhat subjective, depending on the balance desired for false positives and false negatives.

2.3.2.2. IWP retrieval error

We compare the IWP retrieved by the neural network against the reference IWP. For the IWP retrieval, we define the fractional error,

\[
FE = \exp \left\{ \ln \left( \frac{\text{IWP}_{\text{retrieved}}}{\text{IWP}_{\text{reference}}} \right) \right\} - 1.
\]
For example, if the reference IWP is 200 g/m$^2$, then a retrieved value of 50 g/m$^2$ would have a fractional error of 3 (300%), just like a retrieved value of 800 g/m$^2$ would. This should be kept in mind when comparing errors to other sources (such as Deng et al. [2013]), that are not always explicit in their definitions of the error, but that may consider a 50 g/m$^2$ retrieval for a 200 g/m$^2$ reference to be a 75% error rather than 300% one.

Because of the large errors in IWP and because many retrievals for IWP or IWC retrieve in log-space, we think our definition of the fractional error is more appropriate. A fractional error in linear space corresponds to a “classical” error in logarithmic space, and using a more classical error definition, one could ensure errors are at most 100% by retrieving 0 g/m$^2$.  

We divide the testing data (i.e. collocations not used in the training) in bins according to the reference IWP, and calculate the median fractional error for each bin. The bins are logarithmically spaced between $1 \times 10^{-1}$ and $1 \times 10^4$ g/m$^2$. Note that when shown as such, a low median fractional error for a certain IWP does not necessarily mean the retrieval is useful. In the hypothetical situation where a retrieval always results in a constant 10 g/m$^2$, then the median fractional error (or, indeed any other sensibly defined error) as a function of the reference IWP would be very small close to 10 g/m$^2$ (even a stationary clock states the correct time once or twice per day). Therefore, we also generate scatter plots for each retrieval, in order to identify the range where the retrieval has sensitivity to IWP.

### 2.3.3. Analysing multi-spectral synergies

We systematically explore synergies between three techniques: solar reflected, terrestrial infrared, and terrestrial microwave radiation. For each of the techniques, as well as for any
combination of two or three techniques, we train a retrieval (consisting of two networks, as described above). We also explore how adding other data affects the retrieval, in particular the viewing geometry (zenith and azimuth angle), the position of the Sun (zenith and azimuth angle), surface temperature (obtained from European Centre for Medium-range Weather Forecasting (ECMWF) re-analysis data through the CDPC product ECMWF-AUX), and surface elevation from Amante and Eakins [2009]. See Table 1 for a complete overview of all data considered.

Based on this exploration, we make an initial estimate of what information to use for the SPARE-ICE retrieval. The error analysis is not the only consideration for deciding what input to use for SPARE-ICE. For example, adding ECMWF re-analysis surface temperatures has the significant disadvantage of introducing a dependence on a model. However, to make a cost-benefit analysis, we still quantify how much surface temperature would improve a retrieval.

2.3.4. Characterising SPARE-ICE

After choosing the input data based on the error analysis and other considerations, we study the resulting SPARE-ICE IWP in more detail.

For a selection of atmospheric scenes, we compare SPARE-ICE IWP against IWP from two other products. This comparison is the first step away from the collocation world. Therefore, this comparison can indicate whether the retrieval can be extrapolated to cases where no collocations exist, such as towards the edge of the scan, or in geographic regions where retrievals may be difficult (such as mountain areas) that are too small to show up when doing global statistics. We choose MODerate resolution Imaging Spectroradiometer (MODIS) and Microwave Surface and Precipitation Products System (MSPPS) so that we...
have one product based on solar reflectances, and one on microwave radiances. MODIS retrieves IWP at a footprint of 1 km, i.e. much higher than SPARE-ICE, which retrieves at the MHS footprint (16 km at nadir). Note that MSPPS uses 89 GHz and 150 GHz, which is different from the channels centred around 183 GHz used by SPARE-ICE.

3. Results

The results are presented in two parts. First, we explore synergies and choose the inputs to SPARE-ICE. Then, we analyse the resulting product for individual scenes and gridded means.

3.1. Exploring synergies and choosing the inputs to SPARE-ICE

We will first show the performance for single-technique retrievals, and then for retrievals using a combination of two or three techniques. All results are for global retrievals for all angles for which we have collocations. For these results, we look only at the network retrieving IWP, i.e. the part of the retrieval assuming the presence of (cloud) ice. We will explore the performance of the cloud filter separately.

3.1.1. Single-technique retrievals

Figure 3 shows the performance of a retrieval using only measurements from AVHRR channels 1 (0.58–0.68 \(\mu\)m) and 2 (0.725–1 \(\mu\)m), using reflected solar radiation. As indicated before, for all retrievals, we observe a local minimum in the fractional error somewhere between 1 and 10 g/m\(^2\). This minimum does not indicate that the retrieval is truly performing well for these IWP values. Rather, this local minimum indicates that the neural network retrieves these values when it lacks information. If the neural network lacks information, it tends to the mean state of similar measurements for which it lacks...
information. In this case, that is the geometric mean (because retrievals are in log space) for clouds too thin to be detected; hence the values between 1 and 10 g/m$^2$. Therefore, one should be cautious in interpreting errors for small values of IWP based on figures like this one, and also look at scatter plots such as shown later in Figure 7. For a global retrieval using only AVHRR channels 1 and 2 and no additional information, the performance for all but the very thickest clouds is very poor. Adding latitude information improves this a little bit, and adding angular information improves this more, although errors remain high at over 600% compared to 2C-ICE. The improvement is not surprising, because solar angles are essential to interpret information from solar reflectances. Some of the improvement may be to correlations, either between quantities overall, or specifically in the collocations. Despite the improvements by adding auxiliary information, variations in surface brightness, as well as other factors, mean that solar reflected information from AVHRR channels 1 and 2 alone is not enough to obtain a quantitative estimate of IWP.

Figure 4 shows the performance of a retrieval using only the two AVHRR terrestrial infrared channels. Without surface temperature information, a global retrieval based on only the two terrestrial radiation AVHRR channels performs very poorly (it reaches 1500%). Adding latitude information improves results, but adding surface temperature information (from ECMWF re-analysis) improves them more. Adding latitude and surface temperature information does not help beyond adding only surface temperature information. The addition of latitude information might help mainly due to its correlation with surface temperature. In the atmospheric window region, terrestrial infrared measurements essentially give information on the target temperature. A large difference between target brightness temperature and surface temperature indicates the presence of a cloud. The
error remains around 200%, because this temperature difference alone is still not sufficient
to quantify the amount of atmospheric ice.

Figure 5 shows the performance of a retrieval using only the three MHS water vapour
channels around 183 GHz. This retrieval is similar to the one presented in Holl et al. [2010],
although Holl et al. [2010] did not attempt to do global retrievals. Without additional
information, a global microwave-only retrieval performs even poorer than a global solar-
only or a global terrestrial infrared-only retrieval. We see a significant improvement
by adding latitude or angular information, and even more improvement when we add
both, with the median fractional error going down to 100% for very thick clouds. A
retrieval at these frequencies is based on scattered radiation, so ice is detected through a
brightness temperature depression for a down-looking sensor. At high latitudes, where the
atmosphere is very dry compared to the tropics, these water vapour sounding channels
become surface channels, in particular when the surface elevation is high (such as on
Antarctica). The low surface emissivity (and low surface temperature) near 183 GHz then
results in a low brightness temperature temperature that, without additional information,
cannot be distinguished from a brightness temperature depression due to a thick ice
column, at least not from a single channel. Similarly, at off-nadir scan angles, the sounding
altitude is increased due to the longer path length through the atmosphere, results in a
lower radiance, again a low brightness temperature indistinguishable from scattering due
to atmospheric ice. Therefore, and as Figure 5 shows, information on latitude and viewing
geometry is valuable for any global microwave-based retrieval. Even more valuable should
be a combination of surface temperature and emissivity, but emissivities were not explored
in the present study.
3.1.2. Multi-technique retrievals

Figure 6 compares the performance single-technique and multi-technique retrievals, again according to the median fractional error as defined in equation 1. For this comparison, for each single-technique retrieval, we chose the best-performing (lowest error) combination of measurements and auxiliary information. The multi-technique retrievals are then simply combinations of the different single-technique ones. For IWP values between approximately 10 and 1000 g/m$^2$, out of the single-technique retrievals, solar performs worst, followed by microwave, and then terrestrial infrared. The combination of solar and terrestrial infrared performs better than either of those alone for any IWP between 10 and 4000 g/m$^2$. For the combination of solar and microwave, we also see a strong synergy between 30 and 2000 g/m$^2$. Out of the dual-technique retrievals, terrestrial infrared and microwave performs the best, and considerably better than either technique alone throughout the entire range of IWP values. The triple-technique retrieval, combining all three techniques, outperforms any dual-technique retrieval. However, a day-night retrieval, using terrestrial infrared and microwave, could do almost as good as a daytime-only retrieval.

Although SPARE-ICE uses solar, terrestrial infrared, and microwave, it is not identical to the solar+TIR+MW retrieval shown in the figure. Surface temperature information is omitted from SPARE-ICE, so that the retrieval can be based entirely on direct measurements combined with information that can be assumed constant through time (latitude and surface elevation). Furthermore, SPARE-ICE includes AVHRR channel 3B (3.55–3.93 µm), that is neither exclusively solar nor exclusively terrestrial infrared, and that is not included in any other retrievals shown in figures 3–6. Table 1 summarises the data.
used in SPARE-ICE. Figure 6 shows that for most values of IWP, the median fractional error in SPARE-ICE is around 100%. It is below 125% throughout the entire range of IWP, and briefly drops off to 75% for IWP between 1000–2000 g/m².

Figure 7 shows a scatter density plot for SPARE-ICE versus 2C-ICE. Note that, just like for the previous figures, this comparison does not yet take into account the cloud filter, but retrieves non-zero IWP everywhere. As expected, the correlation between 2C-ICE and SPARE-ICE is very good for IWP > 10 g/m². Our retrieval was trained against 2C-ICE IWP. Although the testing data were not used in the training, they come from the same dataset and therefore exhibit the same statistics as the training and validation data. Therefore, a good correlation between the retrieved and reference IWP is expected. For smaller values of IWP, we can see that SPARE-ICE IWP tends to be larger than 2C-ICE IWP, and levels off at a median of around 1–2 g/m² for 2C-ICE IWP < 1 g/m². SPARE-ICE is not sensitive to such small values of IWP. This justifies the choice of 10 g/m² as a cutoff in the cloud filter. This also illustrates the phenomenon visible in figures 3–6, where the median fractional error drops off as IWP approaches the sensitivity limit, which does not indicate a good retrieval, but rather a lack of information, as explained before.

Figure 7 also shows that the random error remains quite large; this is quantified by the aforementioned fractional error, shown in Figure 6, which has a median around 100% for SPARE-ICE, but with outliers up to two orders of magnitude away from the reference value.

### 3.1.3. Cloud filter performance

So far, all comparisons were in terms of IWP and in log-space, and therefore did not include cases where either retrieval is 0 g/m². Figure 8 shows the performance of the cloud filter.
filter. A scene is considered cloudy if the cloud probability $p$ is larger than the defined
cutoff. The sum of false positives and false negatives has a minimum for a cutoff of 0.46,
and the rate of false positives and false negatives are equal at a cutoff of approximately
0.51. The figure also shows reference and retrieved IWP values for the false positives and
false negatives. For a cutoff of 0.5, if the cloud filter falsely detects a cloud where there
really isn’t, SPARE-ICE retrieves a median IWP of around 25 g/m$^2$, while the reference
median IWP is close to 0 (recall that the definition of “cloudy” means the reference IWP
> 10 g/m$^2$). Similarly, when the cloud filter misses a cloud, SPARE-ICE retrieves a
“raw” value of less than 10 g/m$^2$, while the true IWP has a median of around 25 g/m$^2$.
Either way, the values of IWP for clouds where the cloud filter is wrong are quite small.
SPARE-ICE uses a cutoff of 0.5.

3.2. Characterising SPARE-ICE

We have processed measurements for NOAA-18 for 2007, 2008, and 2009, resulting in
three years of retrieved SPARE-ICE IWP. We show a selection of three case studies, all
from NOAA-18.

Although we have looked at all individual inputs, here we show only composites of three
AVHRR channels and three MHS channels. For AVHRR, we follow Dybbroe et al. [2005]
and combine channel 1, channel 2, and the inverse of channel 4 (so that clouds are white
in all channels) and put those in the red, green, and blue planes of a composite image.
Similarly, we make a composite image of MHS channels 3, 4, and 5 (where clouds are
black in all panels).

Figure 9 shows a fragment of a swath for a mid-latitude winter scenario in the North
Atlantic Ocean, off the coast of Newfoundland, Canada. The retrievals from MODIS
and SPARE-ICE show similar spatial structures, although there are some differences. The exact values of retrieved IWP differ considerably between MODIS and SPARE-ICE. For example, for a particular region (48–52 °W, 38–42 °N), the mean for MODIS IWP is 228 g/m², whereas the mean SPARE-ICE IWP is 116 g/m². Lacking an independent reference, there is no way of determining which one is more accurate.

The retrieval from MSPPS is clearly lacking in comparison to either MODIS or SPARE-ICE, detecting only a small fragment of the system (but note that MSPPS retrieves using 89 GHz and 150 GHz, and is known to retrieve mostly precipitating ice).

All features visible in SPARE-ICE are also visible to some degree in the AVHRR and MHS composites, or both, and in the individual channels (not shown). The cloud probability network mostly retrieves either probabilities close to 1 or close to 0. In the region where it is close to the cutoff, the corresponding SPARE-ICE retrieval is small, so the cutoff has no major impact on the overall IWP (see also Figure 8).

Another scene is shown in Figure 10. Here, the observation is of Tropical Cyclone Indlala in the south-west Indian Ocean, near Madagascar, on 14 March 2007. For this system, the microwave signature is significant enough for MSPPS to also obtain most of the spatial structure. This structure is also clearly visible in all input channels, although not nearly as cold in the 3.55–3.93 µm channel as in the longer wavelength AVHRR channels. This channel contains a mixed signal of solar reflected and terrestrial emitted, and for a very deep ice cloud, the solar reflected signal increases, while the terrestrial emitted signal decreases. Also another system on this map, at the coast of Africa, is seen in all input channels and recognised by MODIS, MSPPS, and SPARE-ICE.
For the final example, in Figure 11, we look at a system above Antarctica. Antarctica is a region where retrieving atmospheric ice is difficult with any passive technology. Using microwave humidity channels, it is difficult because the high surface elevation and the low temperature both result in a very low specific humidity throughout the atmospheric column. Combined with a low surface emissivity, this leads to a very low radiance. We lack the humidity background against which ice scattering is visible around 183 GHz.

The colour gradient of the MHS composite shows how the lack of humidity first affects the outermost channel (channel 5) first, and the innermost channel (channel 3) last. The MHS composite has regions that are black (all channels cold), and regions that are red (i.e. channel 3 still warm, but channels 4 and 5 cold). Brightness temperatures for channel 5 drop to 133 K. Solar channels are also hard to use, because the reflectance by ice is similar to the reflectance by ice clouds. Finally, the ice cloud signature from terrestrial infrared usually derives from the temperature difference compared to the surface; when the surface is very cold, this difference is no longer a clear indicator of clouds. Even so, SPARE-ICE seems to still perform quite well in this difficult region — at least the retrieved IWP is not clearly unphysical. From AVHRR channels 4 and 5, there is a region near the south pole with a radiance of around 240–250 K, and a region near the coast with a similar radiance.

At the coast, this region exists against a warmer background, and indeed, this is the area where a high cloud probability is obtained (larger than 0.7) and where SPARE-ICE retrieves non-zero IWP. MODIS and SPARE-ICE agree that there is non-zero IWP in this region, although SPARE-ICE sees a larger area with atmospheric ice than MODIS does. Even closer to the south-pole, SPARE-ICE observes clouds with low IWP. In the same region, MODIS retrieves some very high IWP values, up to 5.5 kg/m$^2$ as far south.
as 88°. This is not realistic, and likely due to measurement noise, since the signal-to-noise ratio is low at such low temperatures. For comparison, the highest retrieved IWP by SPARE-ICE south of 85° in this swath is 261 g/m². Although we don’t know whether or not this is correct, at least it is not obviously wrong. Above Antarctica, MSPPS sees virtually nothing (if these very high MODIS IWP retrievals were true, MSPPS should also detect IWP).

Figure 12 shows a gridded map for SPARE-ICE for 2007. For this map, all measurements between 1 January and 31 December were sorted into 1°x1° bins, and the mean IWP was calculated for each bucket. The IWP distribution appears in line with what one would expect, and there are no regions where SPARE-ICE is clearly wrong. For a comparison, we use the MODIS level-3 IWP for the same period, obtained from the MODIS Aqua (MYD08) dataset shown in Figure 13. Figure 14 shows the difference between SPARE-ICE and MODIS. We focus on MODIS, because Eliasson et al. [2013] show that MODIS compares favourably to DARDAR. Furthermore, Eliasson et al. [submitted 2013] show this is particularly true for ice-only clouds.

There are considerable differences between yearly mean SPARE-ICE and MODIS IWP. Over most of the planet, SPARE-ICE retrieves higher values of IWP than MODIS does. The difference is largest in tropical convective regions and other areas with very high IWP. MODIS IWP is obtained from a retrieval of optical depth and effective radius, which, by the nature of MODIS, are obtained for the upper layers of the cloud. The effective radius in the upper layers of the cloud is typically smaller than the effective radius for the entire column, because particles near the top are typically smaller than particles further down.
Therefore, MODIS underestimates IWP for these cases. On the other hand, SPARE-ICE includes microwave information, which helps to quantify cases with very thick clouds.

In polar areas, MODIS IWP is higher than SPARE-ICE IWP, something that can also be seen in the Antarctica observation in Figure 11. MODIS has difficulties at bright surfaces, such as those covered by snow or ice, and likely retrieves spuriously high IWP due to measurement noise. To the contrary, SPARE-ICE considers the polar areas to have very low IWP.

As mentioned before, lacking an independent measurement, we can not state with certainty which one is more correct (or less wrong).

4. Discussion

Compared to 2C-ICE, SPARE-ICE has a median fractional error of around 100% or a factor 2. This can be compared to the scatter found by Deng et al. [2013], who compared IWC retrievals between 2C-ICE and in-situ measurements, among other things. Deng et al. [2013] found that “[IWC is] strongly correlated with the in situ data (…) although the scatter is around a factor of 2”. Although Deng et al. [2013] do not quantify the random error more rigorously and use IWC rather than IWP, it does indicate that the error between SPARE-ICE and active retrievals may be of similar magnitude as the error between active retrievals and the truth. If both are random, uncorrelated, and 100%, that would put the true error for SPARE-ICE at $\sqrt{1+1} \times 100\% = 141\%$ according the law of error propagation.

From the case studies, it appears that SPARE-ICE does well in a wide variety of cases. Drawing information from solar, terrestrial infrared, and microwave, does not only lead to synergies as seen directly by the median fractional error (Figure 6), but also appears to
extend the IWP retrieval to climate regimes that normally pose large difficulties, such as over Antarctica (Figure 11). It is difficult to establish whether the system that SPARE-ICE concludes to be atmospheric ice, really is such, and not humidity or something else. However, unlike MODIS, SPARE-ICE appears not to obtain any IWP values that are clearly unphysical.

SPARE-ICE can be used for reasonably long time series, because the combination of sensors that it uses has been around since 1999. At the time of writing, AVHRR and MHS are carried on NOAA-15 through NOAA-19, MetOp-A, MetOp-B, and similar instruments are carried on other polar orbiting satellites. Therefore, SPARE-ICE has many potential applications for climatological studies.

A number of variations and improvements to SPARE-ICE can be made.

In the training, SPARE-ICE does not take into account the error in the reference dataset 2C-ICE. By adapting the neural network training algorithm, the calculation of the cost function could use a weighted error sum, giving higher weight to observations with a small error in 2C-ICE. This might improve the retrieval somewhat.

The development of a synergistic retrieval from a database need not be from collocations, but can also be done using simulations. By combining a sophisticated method of generating synthetic atmospheric profiles, such as presented by Evans et al. [2012], and a suitable radiative transfer model, one can also build a retrieval database. This could either use the same input channels, which would give a product using the same information as SPARE-ICE, but potentially independent from either SPARE-ICE or 2C-ICE (depending on how the retrieval database is created), or it could use additional inputs using channels that do not yet exist, such as those at sub-millimetre frequencies. In general, the addi-
tion of sub-millimetre frequencies should improve the SPARE-ICE retrieval. After 2022, European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) plans to fly the Ice Cloud Imager (ICI) on the Metop-SG-B1 satellite. ICI is very similar to instruments that have been proposed before, such as CloudIce \cite{Buehler2012}.

Another aspect that is worthwhile to study is to more formally investigate the information content for each of the channels used in SPARE-ICE, along the lines of Cooper et al. \cite{Cooper2006}.

The principle by which the SPARE-ICE retrieval works can be applied to other geophysical quantities. In general, it is very beneficial if scientific satellites fly in orbits near operational ones, because operational retrievals can learn from scientific ones, as SPARE-ICE illustrates.

5. Summary and conclusion

In this article, we have presented SPARE-ICE, a new IWP product based entirely on passive, operational sensors. By collocating NOAA-18 with the CloudSat 2C-ICE IWP product, we obtained a training database of AVHRR and MHS measurements on the one hand, and joint radar-lidar IWP on the other hand. With this database, we have trained a set of neural networks for the detection of atmospheric ice and the retrieval of IWP, using 2C-ICE IWP as a reference. By using five AVHRR channels, three MHS channels, and auxiliary information (solar and satellite angles, surface elevation), SPARE-ICE can retrieve IWP, while the median fractional error compared to 2C-ICE IWP is around 100 \% for IWP > 10 g/m$^2$.

Overall, SPARE-ICE should be a relevant new member to the family of IWP products. SPARE-ICE will be publicly available through the WDC-RSAT under the Open Data...
Acknowledgments. Thanks to the Swedish Vetenskapsrådet for financing the PhD position of the first author. Thanks to the CloudSat team for making available the CPR 2C-ICE product. We thank the UK MetOffice for providing the AAPP package for calibrating AMSU and MHS data from the NOAA CLASS archive, and thank Nigel Atkinson for practical help in reading data. We would also like to thank the National Graduate School in Space Technology at Luleå University of Technology for offering courses and workshops that helped the PhD student and first author in his research.

References


Amante, C., and B. Eakins (2009), ETOPO1 1 arc-minute global relief model: procedures, data sources and analysis, *Tech. rep.*, NOAA NESDIS NGDC.


Surussavadee, C., and D. H. Staelin (2008), Global millimeter-wave precipitation retrievals trained with a cloud-resolving numerical weather prediction model, Part I: Retrieval


WMO (2010), Implementation plan for the global observing system for climate in support of the UNFCCC, vol. GCOS-138 (WMO/TD No. 1532), WMO.
Figure 1. Diagram illustrating the approach for developing and retrieving SPARE-ICE and intermediate products. The horizontal line in the collocations box separates input measurements and auxiliary data from the target measurement. For a description of the input measurements and auxiliary data, including an explanation of the acronyms, refer to Table 1. All steps are described in detail in the text.
Table 1. Measurements considered in the development of SPARE-ICE. In the column “use in SPARE-ICE”, “both” means that the information is used both in the network to retrieve IWP, and in the network to detect a cloud, as described in the text. See also Figure 1.

<table>
<thead>
<tr>
<th>Sensor</th>
<th>Ch.</th>
<th>Spectral range</th>
<th>Technique</th>
<th>Use in SPARE-ICE</th>
</tr>
</thead>
<tbody>
<tr>
<td>AVHRR 1</td>
<td>1</td>
<td>0.58–0.68 μm</td>
<td>Solar</td>
<td>Both</td>
</tr>
<tr>
<td>AVHRR 2</td>
<td>2</td>
<td>0.725–1 μm</td>
<td>Solar</td>
<td>Both</td>
</tr>
<tr>
<td>AVHRR 3b</td>
<td>3b</td>
<td>3.55–3.93 μm</td>
<td>Mixed</td>
<td>Both</td>
</tr>
<tr>
<td>AVHRR 4</td>
<td>4</td>
<td>10.3–11.3 μm</td>
<td>Terrestrial IR</td>
<td>Both</td>
</tr>
<tr>
<td>AVHRR 5</td>
<td>5</td>
<td>11.5–12.5 μm</td>
<td>Terrestrial IR</td>
<td>Both</td>
</tr>
<tr>
<td>MHS 3</td>
<td>3</td>
<td>183 ± 1 GHz</td>
<td>Microwave</td>
<td>Raw IWP only</td>
</tr>
<tr>
<td>MHS 4</td>
<td>4</td>
<td>183 ± 3 GHz</td>
<td>Microwave</td>
<td>Raw IWP only</td>
</tr>
<tr>
<td>MHS 5</td>
<td>5</td>
<td>190 GHz</td>
<td>Microwave</td>
<td>Raw IWP only</td>
</tr>
</tbody>
</table>

Auxiliary information

| Solar Zenith Angle (SZA) | Both |
| Solar Azimuth Angle (SAA) | Both |
| Local Zenith Angle (LZA) | Both |
| Local Azimuth Angle (LAA) | Both |
| Surface temperature $T_{surf}$ | None |
| Surface elevation $z_{surf}$ | Both |
Figure 2. Illustration of various footprint sizes. Adapted from Holl et al. [2010]. Not shown is CALIPSO, which is smaller than CloudSat and falls mostly within the latter footprint.

Figure 3. Performance for a global retrieval based only on solar reflected measurements from AVHRR channels 1 and 2 (see Table 1), without or with additional angle or latitude information used in the retrieval. The fractional error is defined by equation 1.
Performance for a global retrieval based only on terrestrial emission measurements from AVHRR channels 4 and 5 (see Table 1), without or with additional surface temperature or latitude information used in the retrieval. TIR is short for Terrestrial Infra-Red radiation. The median fractional error for the retrieval using no additional information reaches more than 1500% for IWP between $10^3$ and $10^4$ g/m$^2$.
Figure 5. Performance for a global retrieval based only on MHS channels 3–5 near the 183 GHz water vapour absorption line (see Table 1), with or without additional latitude and viewing geometry information used in the retrieval. The viewing geometry information (“angles” in the legend) includes the zenith and azimuth angle for the satellite as seen from the surface. The median fractional error for the retrieval using no additional information reaches more than $10^4\%$ (i.e. a factor 100) for $\text{IWP} > 1 \times 10^3 \text{g/m}^2$. 
Figure 6. Comparison between the performance of the IWP retrieval networks using different combinations of input data. In this figure, solar refers to AVHRR channels 1 and 2, TIR refers to AVHRR channels 4 and 5, and MW refers to MHS channels 3–5 (see Table 1 for channel details). Here, all retrievals include latitude information, retrievals using solar or MW include information on solar and satellite angles, and retrievals including IR use surface temperature information. The difference between solar+TIR+MW and SPARE-ICE is that SPARE-ICE lacks surface temperature information, but includes AVHRR channel 3B.
Figure 7. Comparison between 2C-ICE and SPARE-ICE IWP. All testing data were binned first according to the reference 2C-ICE IWP, and then according to the retrieved SPARE-ICE IWP. The colour in each bin indicated the absolute number of retrievals with a particular combination of 2C-ICE / SPARE-ICE IWP. The black line represents the diagonal 1:1-line. The grey lines show the median for one product when the other product is kept fixed; for example, when 2C-ICE is 1 g/m², the median SPARE-ICE is 2 g/m². Any vertical or horizontal intersection is a histogram; for example, when 2C-ICE is 10 g/m², SPARE-ICE varies roughly between 1 and 100 g/m². Note that this figure considers “raw” IWP, meaning IWP before any cloud-filter is applied.
Figure 8. Performance of the SPARE-ICE cloud filter depending on the selected cutoff value. A false positive occurs when the cloud filter considers a scene to be cloudy (IWP > 10 g/m²) when it is not, and a false negative occurs when a scene is cloudy, but the filter concludes it is not. The solid lines show the rate of false positives, false negatives, and the total error rate, according to the axis line to the left. The dashed lines show the retrieved and reference IWP values for false positives (f.p.) and false negatives (f.n.), respectively, according to the axis line to the right.
Figure 9. Snapshot of a NOAA-18 swath at 1 January 2007, 16:40 UTC, of the coast of Canada. The top row shows colour composites for AVHRR and MHS, respectively, as described in the text. The top right panel shows the cloud probability, retrieved from the AVHRR channels. The bottom row shows retrieved IWP according to three different products: MSPPS, MODIS (Aqua), and, finally, SPARE-ICE. Inputs to the product not shown in the figure are satellite and solar angles, surface elevation, and the individual AVHRR and MHS channels.
Figure 10. Observation of 14 March 2007, 10:40 UTC, of Tropical Cyclone Indlala, making landfall at Madagascar in the south-west Indian Ocean. See Figure 9 for an explanation of the individual panels.
Figure 11. Example swath passing over Antarctica, 21 January 2007, 5:45 UTC. The south pole is visible as a cross near the bottom of the maps. Quantities shown are as in Figure 9.
Figure 12. SPARE-ICE IWP gridded mean for 2007, using NOAA-18. Map projection equal-area pseudo-cylindrical, according to Mollweide [1805].

Figure 13. MODIS Aqua IWP gridded mean for 2007. Projection as for Figure 12.
Figure 14. Difference between SPARE-ICE IWP and MODIS Aqua IWP, gridded mean for 2007. Dark tan and beige colours indicate areas where MODIS IWP is larger, whereas blue colours indicate areas where SPARE-ICE IWP is larger. Projection as for Figure 12.
Thank you for your attention.