Development of a cloud radiation database for EPS-SG ICI

Bengt Rydberg[2], Eleanor May[1], Hanna Hallborn[1], Patrick Eriksson[1], and Inderpreet Kaur[1]

[1] Department of Space, Earth and Environment, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
[2] Swedish Meteorological and Hydrological Institute, SMHI, SE-601 76 Norrköping, Sweden

February 2, 2023

Contents

Executive Summary

The EUMETSAT Polar System (EPS) Second Generation (EPS-SG) will provide continuity of observations of the current EPS, and will include the Ice Cloud Imager (ICI). ICI is a conically scanning radiometer that will measure at millimetre and sub-millimetre wavelengths. It will have 13 channels, operating within the range of 183 GHz to 664 GHz. The primary objective of ICI is the quantification of cloud ice. It is expected that the data will be used in support of numerical weather prediction and for the verification and parameterization of ice clouds within climate applications.

The main retrieval quantities to be produced at the EUMETSAT Central Facilities will be the integrated ice water path (IWP), the mean ice particle size by mass, and mean mass height. A retrieval algorithm based on Bayesian Monte Carlo integration (BMCI) has previously been developed. A crucial part of the retrieval algorithm is the cloud radiation retrieval database, consisting of pairs of atmospheric/surface states and corresponding simulated ICI observations. The aim of the study *Development of a cloud radiation database for EPS-SG ICI* is the generation of a state-of-the-art retrieval database ready for operational use upon the launch of EPS-SG.

The study was divided into three core tasks to be performed. This document is composed of technical reports written for each of the tasks. A brief description of the contents of each report are as follows:

- The Task 1 report is an exploration of the limitations of an existing preliminary database and a review of database generation methods. Generation of a completely new database was deemed necessary, and a path forward was defined regarding both generation and performance assessment. Several requirements were placed on the new database including the consideration of polarisation, three dimensional variability within cloud structure, and improved snow and sea ice surface emissivity models.

- The Task 2 report contains a description of the methods used to generate the new retrieval database. Focus is placed upon methods that differ from those used to generate a previous, preliminary database. In short, the report describes how:
  - A database covering $\sim 9.5 \times 10^6$ samples was generated, including descriptions of the input data, processes, and outputs of the scheme.
  - 3-D atmopsheric states, also referred to as scenes, were generated in order to capture the variability of cloud structure. Two-dimensional stretches of radar reflectivity from CloudSat were used as the primary source of information on the spatial structure of clouds, but were combined with MODIS multispectral data and ERA5 atmospheric/surface data to effectively widen the data and produce a three-dimensional scene.
  - A state of the art radiative transfer model was applied to simulate ICI. Variables taken into account include sensor characteristic data, atmospheric absorption, surface emissivity models, and hydrometeors.
  - Multiple particle models, consisting of a habit and a particle size distribution, are included within the database. The report describes how the models are selected for a simulation such that reality is statistically represented.

- The Task 3 report summarises the completed database, presenting a statistical overview of the simulations and cloud ice products. Validation of the database

is discussed, and it is shown that distributions of IWP samples in the database match distributions of IWP present in the DARDAR product. Also included in the report are validations performed for two test databases, consisting of simulated observations of The Global Precipitation Measurement (GPM) Microwave Imager (GMI) and The International Submillimetre Airborne Radiometer (ISMAR). This allowed for an assessment of the database generation techniques when confronted with real observations. Inversions of the real GMI observations showed realistic spatial distributions in agreement with DARDAR. Finally, an assessment of retrieval performance was carried out, applying BMCI alongside the new database. Retrievals are shown to satisfy the accuracy requirements stipulated in the study, and it is shown that retrieval accuracy is higher for lower latitudes.

In summary, the final cloud radiation database is in statistical agreement with data, performs well in retrieval tests, and demonstrates an improvement over the preliminary database. ICI will provide the first operational observations in the sub-millimetre region, and full operational coverage for 21 years is expected. During this time, retrievals of cloud ice products utilising this database will provide valuable data for weather and climate models.

Literature review to support the development of a cloud radiation database for EPS-SG ICI IWP retrieval

Inderpreet Kaur[1], Patrick Eriksson[1], and Bengt Rydberg[2]

[1] Department of Space, Earth and Environemnt, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
[2] Möller Data Workflow Systems AB (Molflow), Stampgatan 54E, SE-411 01 Gothenburg, Sweden

February 1, 2023

Contents

1 Introduction

1.1 Background

The Ice Cloud Imager (ICI) will be part of the European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT) Polar System's Second Generation (EPS-SG). ICI is a passive conically scanning radiometer that measures wavelengths from millimeter (mm) to sub-millimeter (sub-mm) range through 13 channels ranging from 183 GHz to 664 GHz. The main aim of ICI is to quantify cloud ice in support of climate monitoring, as well as to parametrize and validate ice clouds in weather and climate models.

A key aspect for the future ICI observations is global retrievals of ice water path (IWP). Previously, a retrieval algorithm based on a retrieval database and Bayesian Monte Carlo Integration (BMCI) was shown to be suitable for the task (Rydberg 2018). The retrieval database is a large set of synthetic scenes and the corresponding simulated observations. For the BMCI retrieval algorithm to be successful, the database should contain a large number of states such that each possible ICI measurement finds several matches. Further, these states should represent the atmospheric and surface conditions as realistically as possible. For instance, it is essential that vertical and horizontal cloud structures in the database follow the natural variability closely. In addition, pressures, temperatures, humidities, and surface emissivities must also be included in a realistic way.

1.2 Purpose of this document

The purpose of this document is to provide a review of the literature relevant to methods for the generation of cloud radiation databases for IWP retrieval. The information gathered herein shall be discussed with EUMETSAT to propose a way forward for the database generation.

This document is structured as follows. A description of the relevant IWP retrieval databases available in literature are discussed in Sect. 1.3 [R-2]. The limitations of the preliminary database are discussed in Sect. 1.4.1 [R-3]. Section 2 provides an assessment on overcoming the existing gaps and limitations of the preliminary database [R-3]. This section also covers the discussion on hydrometeor microphysical assumptions [R-6] and use of antenna patterns to avoid beam filling errors [R-8]. In Sect. 3, an assessment of other requirements necessary to generate the database are made [R-4].

1.3 Previous works

Several attempts at creating IWP retreival databases for sub-mm measurements can be found in the literature. Table 1 gives a detailed summary of such databases. Among the first attempts to retrieve IWP from sub-mm observations is by Rydberg et al. (2007). They presented an algorithm to create a database of synthetic one-dimensional (1D) atmospheric states. They merged radar data and statistics from *in situ* measurements to generate varying cloud structures. The algorithm also included oriented hydrometeors (as spheriods), but each channel was represented by a single frequency, and spectral differences between two side-bands were ignored. Additionally, the study was constrained to northern mid-latitudes due to unavailability of radar data in other regions. Later, Rydberg et al. (2009) extended the study by creating a retrieval database of three-dimensional (3D) atmospheric states using CloudSat measurements as input but with simplified cloud microphysics assumptions. They transformed two-dimensional (2D) Cloudsat profiles along with weather data from the European Centre for Medium-Range Weather Forecasts (ECMWF) to 3D through a stochastic iterative amplitude adjusted Fourier transform algorithm (Venema et al. 2006). In an another study, Evans et al. (2012) presented an algorithm to retrieve IWP using sub-mm measurements from flight campaigns. They created the retrieval database of 1D scenes based on CloudSAT and Calipso measurements and employed detailed microphysical assumptions. They utilized cloud *in situ* measurements of microphysical probability distributions to model realistic cloud profiles. However, the biggest constraint of their method is that the microphysical assumptions are either valid locally or seasonally. This hampers the extension of the database to represent global variability. Attempts at coupling a mesoscale numerical weather prediction (NWP) model and radiative transfer to generate the database also exist in literature. Wang et al. (2016) have attempted a statistical retrieval of cloud parameters for ICI. They simulated hydrometeor profiles with Weather Research and Forecasting (WRF) model to cover 12 different meteorological situations covering Europe. The hydrometeor microphysics was determined through WRF Single Moment - 6 class (WSM6) scheme (Hong and Lim 2006). A more recent study by Brath et al. (2018) also developed a retrieval database by using atmospheric profiles based on Icosahedral Nonhydrostatic (ICON, Zängl et al. 2015) model profiles, but with simplified microphysical assumptions. They assumed rain and cloud ice as Mie spheres and snow as aggregates. In all these studies mentioned above, the retrieval database reflects only a part of the global variability. These studies focus either on a particular season or certain latitudinal region. Among the retrieval databases which cover the global scenarios include the preliminary cloud radiation database (Rydberg 2018) and also Goddard Profiling algorithm (GPROF, Randel et al. 2020). These are described in the next sections.

Table 1: Brief description of the various existing IWP retrievals (using sub-mm) based on a retrieval database.

Retrieval database	**Input data**	**Microphysics**	**Scattering properties**	**Surface emissivity**	**Coverage**
Rydberg et al. (2009) (Odin SMR)	Cloudsat reflectivities + ECMWF	MH97	Mie Theory	Not relevant	Upper tropospheric cloud ice
Evans et al. (2012) (CoSSIR)	Cloudsat reflectivities + Calipso cloud fraction + ECMWF	PSD: multivariate gaussian distribution fitted with data from *in situ* cloud probes	DDA	Lambertian emissivity varied stochastically	Only particular cloud situations where the PSD is valid
Wang et al. (2016) (ICI)	WRF	WRF Single Moment - 6 class (WSM6) microphysical scheme	Ice: DDA aggregates, Other: Mie theory	Water: TESSEM2, Land: TELSEM2	12 diverse European mid-latitudinal situations
Brath et al. (2018) (ISMAR)	ICON	ICON one-moment microphysics scheme	Snow: DDA aggregates, LWC, ice, rain: Mie Theory	FASTEM	Regional: 50N – 75N, 30W – 5E
Rydberg (2018) (Preliminary database)	Cloudsat reflectivities + ECMWF	Combination of PSD and habits, F07, MH97, MGD + Six snow habits	ARTS database	Water: FASTEM, Land: Stochastic specular reflective surface	Global
Randel et al. (2020) (GPROF, only mm range)	DPR and radiometer data (observed 3D scenes)	?	Ice: DDA, Rain: Mie theory	Water: FASTEM, Land: empirical model based on soil moisture	Near global: 65S – 65N

1.3.1 Preliminary retrieval database

Previously, a preliminary cloud radiation retrieval database was developed by the Nowcasting Satellite Application Facility (Rydberg 2018). A method to generate 2D (vertically and horizontally) varying states consistent with stretches of radar reflectivities from the cloud profiling radar (CPR) of CloudSat was developed and used. Rain and ice water content fields were retrieved for a set of assumptions of particle size and shape distributions, and taking additional weather data (including liquid water content) from ERA-Interim into account. A land-sea mask was used to categorize surface type (land or water), and surface topography and wind speed data were taken into account. The Atmospheric Radiative Transfer Simulator (ARTS, Buehler et al. 2018) version 2.3, was applied to simulate ICI measurements in a slant down-looking pseudo two-dimensional (2D) geometry. Single scattering properties for the rain and ice hydrometeors categories were taken from Eriksson et al. (2018).

1.3.2 Other operational retrieval databases

The GPROF algorithm (Randel et al. 2020) has been providing instantaneous rainfall rates and vertical structures of precipitation and other related auxiliary parameters since 1996. GPROF makes use of BMCI and is the only existing operational inversion systems that can be compared with the ICI retrievals of concern here.

The operational version of GPROF uses 1D Eddington approximation to compute brightness temperatures (Tb). The simulations are based on retrievals performed in the central part of the swath where radar and radiometer data are at hand. The surface emissivity is calculated separately for ocean and land. For the former, fast microwave emissivity model (FASTEM, Liu et al. 2010) is used to calculate the ocean surface emissivities as a function of surface wind speed. While for the latter, the emissivities are defined by the soil moisture. When the surface is moderately dry or without any vegetation, the emissivity is assumed to be 0.9. However, for other regions, the emissivity for water covered surfaces is modified between 0 % to 20 % to represent various categories of soil moisture. The atmospheric absorption model is monoRTM (Moncet and Clough 1997). Scattering effects from cloud water are ignored. However, for rain drops, the scattering properties are calculated using Mie theory by assuming a spherical shape. Further, the ice particles are also assumed to be spheres at frequencies up to 89 GHz. For higher frequencies, an ensemble of non-spherical ice particles are considered and single scattering properties are calculated by using the discrete dipole method. Melting particles are not considered.

1.4 The way forward

The retrieval of IWP from satellite measurements is a non-unique and non-linear problem. The quality of the retrieval database is the most crucial element of the inversion algorithm. Often simplified assumptions made in forward modelling limit the a priori information for the BMCI retrieval. Thus, to develop a comprehensive cloud radiation database, it is important to identify gaps and limitations to find the best way forward.

1.4.1 Main limitations of the preliminary database

Even if the preliminary database represents the state-of-the-art, it clearly has several limitations. Some discussion of the matter, but not complete, is found in Eriksson et al.

(2020). Presently, the main limitations are considered to be:

- **Surface emissivity:** A complete surface emissivity model for land and snow is not available for sub-mm wavelengths. The preliminary database used an existing model (TELSEM) and simply assumed a constant emissivity over the frequency range covered by ICI. This is a simplification. In addition, the existing models are based on measurements at lower frequencies and there could be a systematic difference between the emissivity below and above 160 GHz.

- **Simplified particle model:** The clearest limitation is that the preliminary database did not consider particle orientation, and differences between vertically and horizontally polarised channels were strongly underestimated. Several ice particle habits were used but the selection needs to be revisited. The assumed particle size distribution (PSD) also needs to be revisited, specially if some random variation should be added.

- **Representation of footprint:** The preliminary database applied essentially a 1D antenna pattern, while the full 2D response must be represented as shown by Barlakas and Eriksson (2020). On the other hand, the independent beam approximation applied in the preliminary database is sufficient, true 3D radiative transfer is not required, also shown by Barlakas and Eriksson (2020).

1.4.2 Discussion

The main strength of the database applied in GPROF is that it is based on observed 3D scenes. This gives a good representation of both vertical and horizontal structures of rain hydrometeors. However, the radar onboard the Global Precipitation Measurement (GPM) core satellite lacks sensitivity to ice hydrometeors and we can not use these scenes for developing a retrieval database for ICI. To obtain global observation on ice hydrometeors, CloudSat is so far the only choice, the sensor used in the preliminary database. When it comes to the representation of ice hydrometeor scattering properties and the capability of the radiative transfer software applied, the quality is higher in the ICI preliminary database than in the GPROF simulations. That said, the GPROF simulations have other strengths, but they are less relevant for ICI. The above put together points towards developing the new retrieval database along the lines of the preliminary one, but addressing the limitations identified.

That is, several critical extensions, such as implementation of 3D atmospheric scenes, hydrometeor orientation, surface emissivity for complex surface types (snow/sea ice), representation of antenna patterns will be required. This will necessitate generation of a totally new database. Thus, the preliminary version will be completely replaced and a new version will be implemented in place. The new database will follow the same basic approach as used for the preliminary database, that is using CloudSat reflectivities as core data source. Further, in addition to the necessary modifications mentioned, several other smaller extensions will also be made.

2 Methods to address current gaps and limitations

2.1 Generation of 3D scenes

Studies (e.g. Bauer et al. 1998, Kummerow 1998, Battaglia et al. 2006) have analysed the impact of 3D-type effects in passive microwave radiative transfer for microwave frequencies (up to 183 GHz). Davis et al. (2007) were the first to extend the analysis towards sub-mm wavelengths. They followed a stochastic approach to generate 3D mid-latitude cirrus scenes using ground-based 2D radar observations as input. However, due to limited scope of their study, their results were inconclusive in certain cases. A more comprehensive study towards 3D effects at sub-mm wavelengths was made by Barlakas and Eriksson (2020). They used Iterative Amplitude Adjusted Fourier Transform (IAAFT) algorithm (Venema et al. 2006) to generate the 3D cloud fields from 2D Cloudsat radar observations. Another attempt at generating the 3D scenes has been made by Barker et al. (2011) using collocated A-train observations. In this study, we shall consider these two methods to generate the 3D scenes. These two are described below.

2.1.1 IAFFT

Barlakas and Eriksson (2020) generated 3D cloud fields from 2D Cloudsat fields using IAFFT algorithm (Venema et al. 2006). This algorithm is commonly used to generate surrogate data. To generate 3D cloud scenes, the algorithm inputs Cloudsat measurements in pressure and latitude grids and iterates the power spectra to adapt the amplitude distribution and Fourier coefficients to the 3D fields in pressure, latitude and longitude grid. ARTS 3D Monte Carlo (ARTS-MC) was used to perform the realistic full 3D pencil beam simulations for tropical and mid-latitude scenarios. The pencil beam calculations were integrated over the sensor field of view to get antenna weighted brightness temperatures. They made an analysis of the performance of 3D simulations against the simulations with independent beam approximations (IBA) and 1D simulations. In IBA, beam filling introduced a slight overestimation in the Tbs and the corresponding bias was between 0.1 and 1 K. On the other hand, 1D simulations had significant influence from beam filling effects which increased both with frequency and footprint size. For instance, at 15 km resolution, neglecting cloud heterogeneities led up to root mean square error of 4 K at 183 GHz, which increased to 13 K at 666 GHz.

The main limitations of IAFFT is that it cannot account for cloud classification information, if needed. Also, Rilemark and Svensson (2020) have shown that while filling in, discontinuities can arise between neighboring pixels.

2.1.2 The Barker 3D cloud-construction algorithm

Barker et al. (2011) present an algorithm that constructs 3D distributions of cloud from passive satellite imagery and collocated 2D nadir profiles of cloud properties inferred synergistically from lidar, cloud radar and imagery data. This section presents the most important aspects of the algorithm while more details are given in Appendix A. Effectively, the construction algorithm widens the active retrieved cross-section (RXS) of cloud properties, by using a match-and-substitute algorithm that fills in off-nadir pixels/columns with cloud property profiles from the swath center. For this purpose radiances from a passive multi spectral imager (MSI) for an off-nadir, recipient, pixel are compared with corresponding values for a range of pixels along the RXS; close matches are identified as potential donors with the closest to the recipient pixel being designated as the proxy to literally stand in for the recipient. This is repeated until all pixels in the required 3D domain are filled.

Barker et al. (2011) applied the algorithm on MSI data from Moderate Resolution Imaging Spectroradiometer (MODIS) onboard Aqua, and RXS data from the lidar Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) onboard Cloud-Aerosol Lidar and Infrared Pathfinder Satellite (CALIPSO) and the Cloud Profiling Radar onboard CloudSat. These sensors/platforms are all part of NASA's A-train constellation and have been operating in more than 10 years, and that implies that a collocated dataset covering all types of weather conditions is available.

The Barker et al. (2011) construction algorithm uses actual data along the across-track direction to widen the basic 2D cross-section into a 3D field, as opposed to the IAFFT algorithm by Venema et al. (2006) that only uses the properties of the basic 2D cross-section in order to generate a stochastic 3D field. An additional advantage of the Barker et al. (2011) construction algorithm is that it is basically designed to generate consistent 3D fields of many parameters, and this is less clear how to acheive using the algorithm by Venema et al. (2006).

The Barker et al. (2011) construction algorithm is based on the assumption that two closely spaced cloudy pixels having close to identical temperature and moisture profiles and only small differences in a number of spectral top of atmosphere radiances will have similar cloud property profiles. This is clearly an assumption that is not always valid. Another issue is that it can not be guaranteed that a good MSI match can be found at all if only pixels closely spaced in space and time are considered. Barker et al. (2011) points out that the algorithm is a 3D construction algorithm and not a 3D reconstruction algorithm, since the data used does not allow for a proper reconstruction. Anyhow, Barker et al. (2011) showed that the algorithm was able to reconstruct the cloud mask, with only small errors, of the RXS out to a distance of at least 20 km from the "local database". We therefore judge that the Barker et al. (2011) 3D construction algorithm, or a variant of it, is the most promising algorithm to deploy for the present study.

2.2 Surface emissivity

An important aspect for the retrieval database is modelling of the surface emissivities in a realistic manner. The impact of surface is mostly important for frequencies up to 325 GHz. For higher sub-mm frequencies, the weighting functions peak in the upper atmosphere, hence surface impact is mostly negligible. In this section, we describe the various studies available in literature that describe the surface emissivity for water, land and snow/sea ice surface types.

2.2.1 Water bodies

Over water, the emissivity is primarily related to the wind speed. In preparation for EPS-SG, a parametrisation of the surface emissivities for water bodies up to 700 GHz is available through Tool to Estimate Sea-Surface Emissivity from Microwaves to sub-Millimeter waves (TESSEM2, Prigent et al. 2017). Its performance has been tested by Prigent et al. (2017) using the International Submillimeter Airborne Radiometer (ISMAR) observations with some encouraging results. While some uncertainties can be expected when more observations will be available, we will employ TESSEM2 emissivity parametrisation to simulate surface effects over water bodies.

Table 2: Brief description of the various land (snow-free) emissivity estimates available in literature.

Study	Input data	Frequency	Comments
Hewison (2001)	airborne measurements (over Sweden)	24 to 157 GHz	Emissivity increases with frequency
Karbou et al. (2005)	Satellite based retrievals (AMSU)	24 to 150 GHz	150 GHz retrievals were noisy, Emissivity estimates at 89 GHz a good proxy for radiance simulations at both 150 and 183 GHz

Table 3: Brief description of various snow-cover emissivity estimates available for frequencies up to 183 GHz.

Study	Input data	Frequency	Comments
Hewison et al. (2002)	Airborne measurements over the Arctic	25 to 183 GHz	Emissivity at 183 GHz higher than 157 GHz
Harlow (2009)	Airborne measurements over Alaska	89 to 183 GHz	Emissivity increases with frequency except for fresh snow events
Munchak et al. (2020)	Satellite based retrievals (GMI)	10 to 166 GHz	OEM retrievals, 20 snow-cover classes, Emissivity increases with frequency for cold and dry snow, mountain snow, shallow/early snow, opposite trend
Camplani et al. (2021)	Satellite based retrievals (GMI, ATMS)	10 to 166 GHz	Emissivity increases with frequency for cold and dry snow

2.2.2 Land

Over land, snow-cover and sea ice surfaces, it is often a challenge to model surface emissivity due to strong temporal and spatial fluctuations. Analogous to TESSEM2, a new version of climatology based Tool to Estimate Land-Surface Emissivities at Microwave frequencies 2 (TELSEM2) has been developed by Wang et al. (2017) to provide surface emissivities over land, snow-cover and sea ice for frequencies up to 700 GHz. However, in TELSEM2, the emissivity estimates are empirical approximations. Thus, it is important to look for other emissivity estimates that exist in literature. This section describes the existing literature for surface emissivities over snow-free and snow-covered (both land and sea ice) surfaces.

Snow free surfaces A common approach for land emissivity estimates (snow-free) is to use retrieved values from surface-sensitive channels (below 89 GHz) as estimates for higher frequencies. A couple of studies (Table 2) have attempted to retrieve surface emissivities up to 157 GHz, but none have attempted to retrieve estimates at 183 GHz or higher.

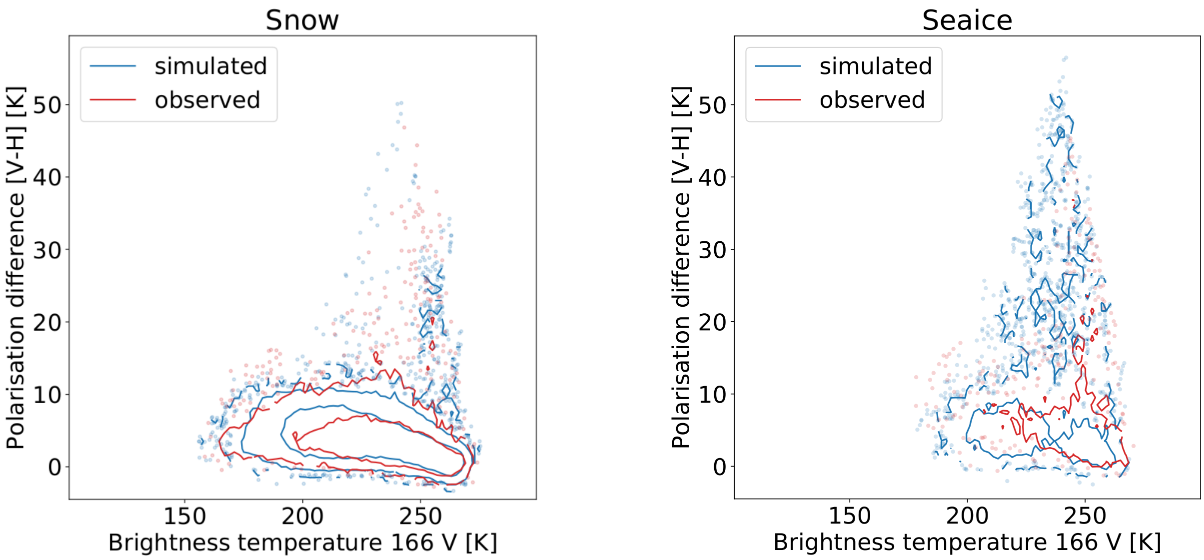

Figure 1: Polarisation difference contours (166V−166H) as a function of 166V GHz brightness temperatures (Tb) for observed GMI data (red) and simulated values (blue) over snow-covered land (left) and sea ice (right).

Hewison (2001) retrieved emissivities upto 157 GHz, using flight measurements over forest and agricultural lands in Sweden. They observed that emissivity increases with frequency. On the other hand, Karbou et al. (2005) retrieved land emissivities up to 157 GHz using satellite observations, but showed that emissivity estimates at 89 GHz are a good proxy for radiance simulations at both 150 and 183 GHz. Similarly, at the ECMWF, assimilation of all-sky radiances assumes that land emissivity is sufficiently invariant with frequency, and the emissivity estimates from surface-sensitive channels serve as a reasonable approximation for sounding channels (Baordo and Geer 2016). TELSEM2 also follows this trend and the emissivity values for snow free land are assumed to be constant for frequencies above 85 GHz. However, little is known about emissivity trend beyond 183 GHz in the sub-mm range. At higher frequencies (above 166 GHz) water vapour often obscures the surface signal. As a result, at high frequencies, impact of surface signal is low in the tropical belt and other areas with high water vapour content. However, in winter, and at higher latitudes, dry atmosphere with snow free surfaces requires modelling surface emissivities in a realistic way.

Snow-covered surfaces (land and sea ice) In TELSEM2, the emissivity values for snow-covered land are assumed to be constant for frequencies above 85 GHz. Further, the sea ice emissivities are parametrized upto 183 GHz using Special Sensor Microwave - Imager/Sounder (SSMIS) derived emissivities (Boukabara et al. 2011) but only for new ice and first year sea ice classes. Multi year ice emissivities are assumed constant for frequencies above 85 GHz. However, data from aircraft campaigns (e.g. Hewison et al. (2002)) have shown that emissivity at 183 GHz is consistently higher than at 157 GHz for both snow and sea ice surface types. In another study, Harlow and Essery (2012) studied the variability of snow emissivity over different snow-packs. They observed that the snow emissivities increase monotonically with increasing frequency for stratified snow. However, for fresh snow, the emissivity estimates follow a decreasing trend. Additional issue which is often raised while retrieving surface emissivity over snow-covered surfaces is assumption of specular or lambertian reflection. However, this is not important for ICI,

Table 4: Brief description of the different snow-cover emissivity estimates available in literature for frequencies above 183 GHz.

Study	Input data	Frequency	Comments
Tait et al. (1999)	Airborne measurements, (snow-covered regions northern US)	220 GHz	increasing Tbs with frequency over snow-covered surfaces - could be due to increasing emissivity
Haggerty and Curry (2001)	Ship/airborne measurements (snow-covered Arctic)	37 to 220 GHz	Emissivity over snow-cover increases beyond 150 GHz (Fig. 2)
Wang et al. (2016)	Airborne measurements (continental ice and sea ice Greenland)	up to 325 GHz	comparisons with TELSEM2 poor, emissivities increase between frequencies 150 and 183 GHz, but decrease at higher frequencies

as at conical scanning angles, the differences between specular and lambertian reflection vanish.

Based on the above studies based on measurement campaigns, we have developed a probabilistic snow-emissivity model that provides polarised emissivities for snow-covered surface types (both over land and sea ice) at frequencies between 150 and 190 GHz. This model gives a random estimate of the snow emissivity from a standard normal distribution depicting the valid range of emissivity values. The valid range was decided through emissivity values reported in literature (Hewison et al. 2002, Harlow 2009, Harlow and Essery 2012) and the model was fine-tuned by comparing observed and forward modelled radiances for snow-covered regions as observed by Global Precipitation Measurement (GPM) Microwave Imager (GMI). The basic idea is to find a set of emissivity estimates that simultaneously give radiance closest to the measurements. Figure 1 (left) shows a statistical comparison of the polarization differences (166V−166H) simulated and observed at 166 GHz for snow-covered land during January. The polarization differences above 20 K are most likely associated with mixed surface types e.g., snow/land boundaries, while the signals between 0 and 15 K are mostly pure surface contribution, but also some impact from hydrometeor scattering is expected. In spite of being random estimates, the range of emissivities for snow-covered regions obtained from our model is also comparable to emissivity estimates retrieved from passive microwave instruments. For instance, Munchak et al. (2020) retrieve emissivities for GMI using optimal estimation method. They define 20 snow-covered surface classes for which the emissivity estimates vary between 0.75 to 0.92 at 166 GHz. These values are quite comparable to snow emissivity estimates from our model. Similar values were also reported by Camplani et al. (2021) for snow-covered surfaces.

Over snow-covered sea ice regions, we obtained a similar performance with this empirical model. For sea ice, the structure of polarisation differences (PDs) is only slightly different than snow-cover (Fig. 1 (right)). Here, the flat arch region is a bit narrow and the concentration of observations along the arm is denser than snow-covered land. The PDs along the arch region originate from snow-covered sea ice, while the the higher PDs along the arm arise from sea ice and water mixture. Areas with lower fraction of sea ice

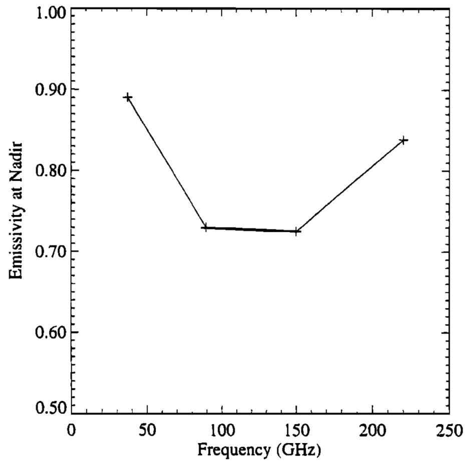

Figure 7. Nadir emissivity versus frequency on May 20, 1998.

Figure 2: Surface emissivity over snow covered regions as reported by Haggerty and Curry (2001).

shall have higher emissivity compared to areas with larger sea ice fraction. A detailed description of distribution of sea ice and water mixture inside the footprint is required to model these emissivities correctly.

An extension of such a model to frequencies upto 325 GHz will not be straightforward due to non-availability of satellite based observations. Few studies (Table 3) have attempted to estimate surface emissivity using data from aircraft campaigns, but it reflects only a small part of the global picture. For example, Wang et al. (2017) retrieved emissivities for sea ice and continental using ISMAR observations over Greenland. For the former, the emissivities were noisy, due to natural spatial variation in sea ice. However, for continental ice, the emissivity increases between frequencies 150 and 183 GHz, decreases at higher frequencies. The average emissivity at 243 GHz and 325 GHz is 0.85 and 0.87, respectively, and compared to 0.9 at 183 GHz. These estimates when compared with TELSEM2 (Wang et al. 2017) gave a very poor match. Further, in another study during a ship/flight campaign over the Arctic (May to July 1998, Haggerty and Curry 2001), measurements over snow-covered sea ice regions were made for frequencies between 37 to 220 GHz. They showed that the emission by snow-cover increases with frequency (Fig. 2). Another study by Tait et al. (1999) also analysed Tb from same instrument but flown over American subcontinent, and found that Tbs observed over snow-covered surfaces increased with frequency. This could be attributed to the increase in emissivities.

Based on the above trends reported in literature for emissivities over snow-covered regions, we will develop a probabilistic model to calculate the emissivity estimates. It should be noted that for the retrieval database the input data should match the simulations in an overall statistical sense. Thus, it is not totally necessary to know the emissivity of snow-covered surfaces at a given time and position. We can apply a random snow emissivity, as long as the applied distribution of emissivities follows reality.

2.3 Hydrometeor microphysics

2.3.1 Mixed phase clouds and particles

Implementing mixed phase particles (melting particles) is not possible due to unavailability of single scattering data for such particles at sub-mm frequencies. The preliminary

Table 5: Some commonly used particle size distributions (PSDs)

PSD	Comments
McFarquhar and Heymsfield (1997) (MH97)	Used in limb sounding community, only tropical settings
Field et al. (2007) (F07)	Tropical and mid-lat versions, Used at ECMWF
Delanoë et al. (2014) (D14)	Used in DARDAR v2.1 product, MGD
Exponential PSD	simple to use with dual or triple frequency radars, good for precipitation, but aircraft PSDs not exponential

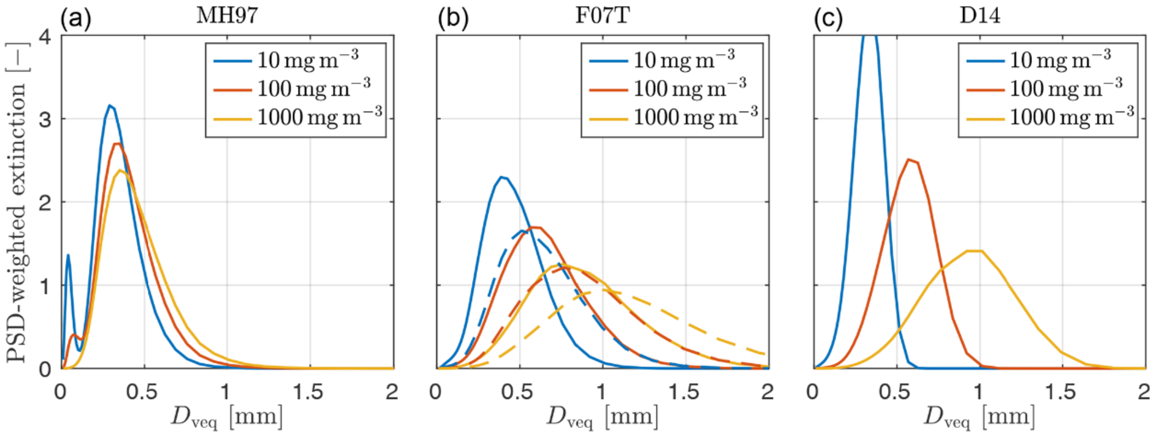

Figure 3: Normalised extinction as function of volume equivalent diameter ($D_{\rm veq}$) for three different particle size distributions. Figure from Ekelund et al. (2020).

database already includes the mixed phase clouds.

2.3.2 Particle size distributions

The particle size distribution (PSD) describes how the particle sizes are distributed within a volume element. Commonly used PSDs by the passive microwave community are described in Table 5, for example, the one by McFarquhar and Heymsfield (1997), usually referred as MH97. MH97 was also used by Rydberg et al. (2009) to retrieve IWPs using limb microwave measurements. It is mostly valid for anvil cirrus in thes tropics. Another parametrized PSD is by Field et al. (2007), also known as F07. It is a single-moment PSD parametrization based on data from multiple measurement campaigns. It has different settings for the tropics and mid-latitudes. The PSD used by DARDAR v2.1 product (Delanoë et al. 2014) denoted here as D14 is another parameterized PSD. It is a modified gamma distribution (MGD) fitted to insitu data. D14 is a 2-moment scheme, and requires the normalized number concentration (N_0^*) and mean volume-weighted diameter (D_m) as the two inputs. Either of N_0^* and D_m can be converted to ice water content (IWC). A 1-moment version of D14 was used by Ekelund et al. (2020). A comparison of the PSD-weighted extinction cross-section for these three PSDs is shown in Fig. 3. MH97 puts more weight on small particles, thus is less suitable for snow. On the other hand, both F07 and D14 put higher weight on larger particles with increasing value of IWC.

Out of F07 and D14, the latter emerges as a better choice for this study, due to

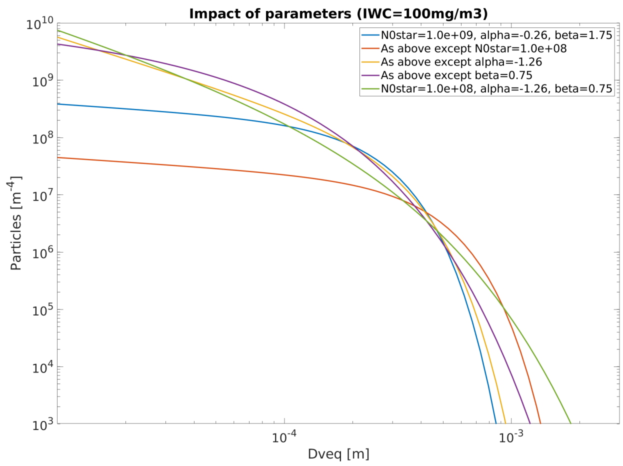

Figure 4: The impact of three parameters: N_0^*, α and β on the variability of D14.

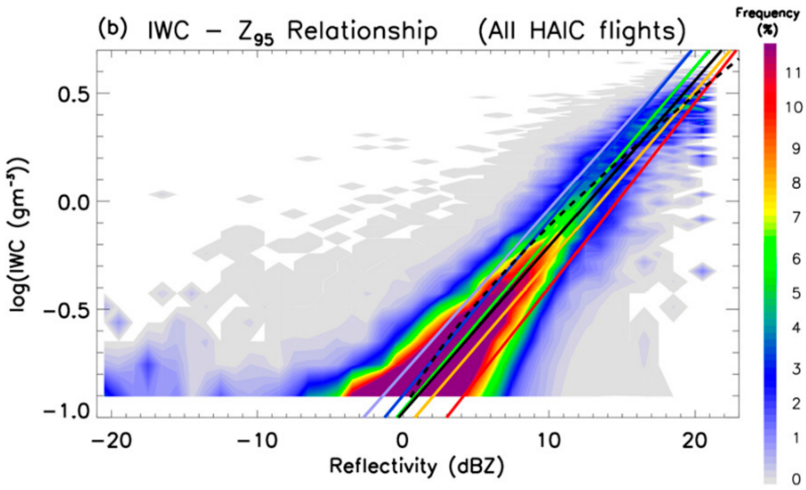

Figure 5: The distribution of joint IWC and dBZ measurements, as available from Protat et al. (2016).

various reasons. Firstly, DARDAR is the best satellite based reference available. Also, D14 is a modified gamma distribution (MGD) and can include continuous variability by randomising the background moments. However, some prior information should be at hand to assess the variability of background moments. D14 has four different moments. The first two moments (α and β) are related to μ and γ coefficients of the MGD, while the remaining two are related to N_0^* and D_m. The latter two moments can be set by selecting two out of the three variables: IWC, N_0^* and D_m. Figure 4 shows the impact of N_0^*, α and β on D14. At IWC of 100 mg m^{-3}, decreasing only N_0^* by 10% leads to +460% change in the rayleigh scattering effects. Similarly, when only α is decreased from -0.26 to -1.26, the corresponding change in scattering is only +14%. Changing only β from 1.75 to 0.75, increases the scattering by 57%. When these three changes are applied together, a 1000% increase in scattering can be obtained. However, the variability introduced by varying N_0^* and other moments should also follow the realistic dBZ-IWC variability. For example, Fig. 5 shows the joint frequency distribution of measured IWC and dBZ from a flight campaign (Figure from Protat et al. 2016).

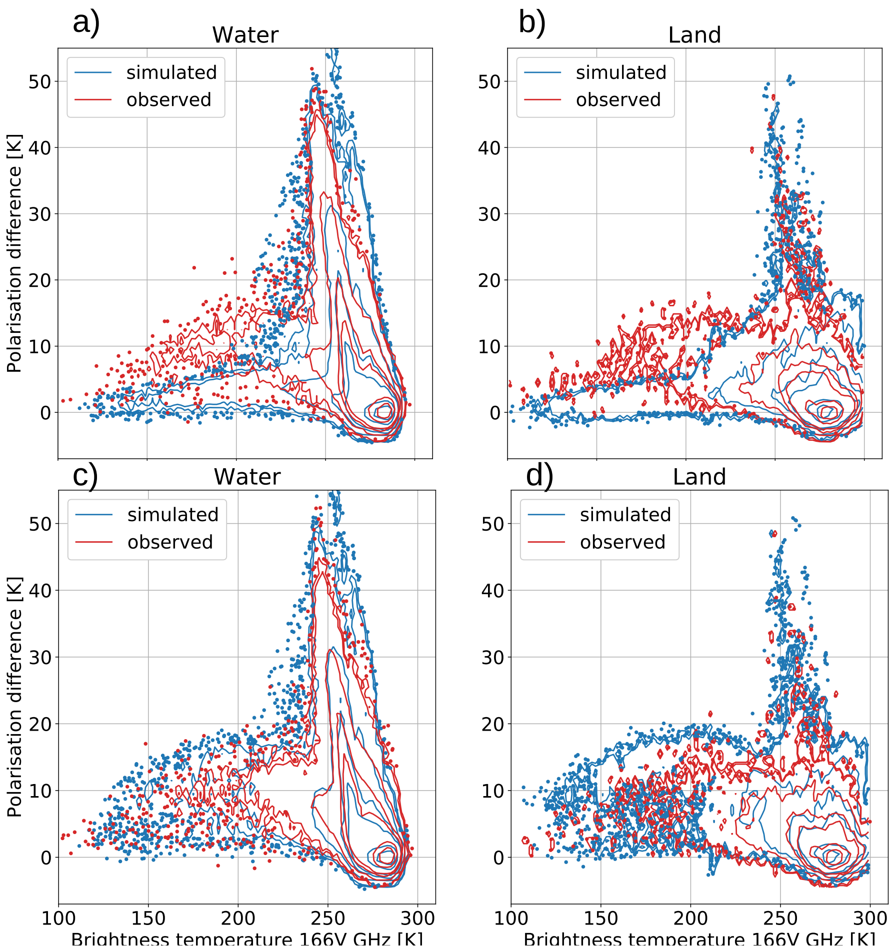

Figure 6: Polarisation difference contours (166V−166H) as a function of 166V GHz brightness temperatures (Tb) for observed GMI data (red) and simulated values (blue) for water and land. In (a) and (b) the simulations assume hydrometeors to have TRO, while in (c) and (d) the ARO effects are simulated using a scheme similar to Barlakas et al. (2021).

2.3.3 Oriented hydrometeors

Dual polarised measurements have consistently shown that ice hydrometeors are generally oriented in nature. However, in retrieval studies, ice hydrometeors have been repeatably being assumed with Total Random Orientation (TRO) to avoid heavy computations. With launch of ICI, the need to improve hydrometeor representation will increase manifolds.

So far, oriented particles have been ignored in microwave retrievals as such particles require polarised (vector) radiative transfer that is more computationally intensive than the standard unpolarised (scalar) calculations. In addition, there are few software packages available that can handle polarised microwave simulations. The lack of realistic scattering property data in the Azimuthal Random Orientation (ARO) category has been an even more significant limiting factor. Some such data were developed and explored by Adams and Bettenhausen (2012). The first publicly available data were provided by Brath et al. (2020), but just for two particle shapes (habits). This is in contrast with the main part of the accompanying scattering database that covers TRO with 36 shapes (Eriksson et al. 2018). The striking difference is explained by the fact that ARO requires much more resources, both in terms of production and storage of the data.

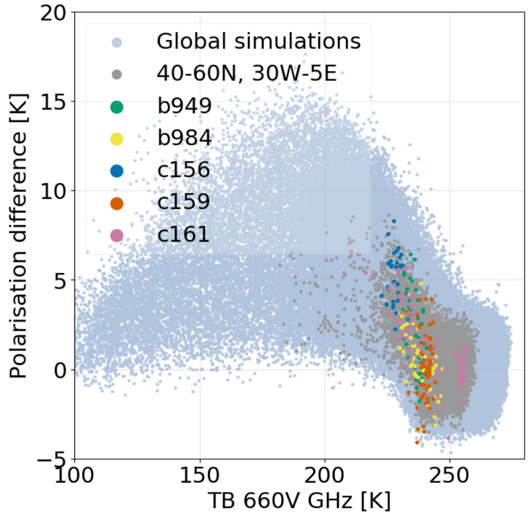

Figure 7: The scatter plot showing the PD vs Tb_V at 660 GHz. The light blue dots are for near global simulations; the gray dots denote simulations in 40°N - 60°N/30°W - 5°E. The five alphanumeric codes depict observations from five flight campaigns as described by Fox (2020).

While it is still too expensive to run a fully polarised radiative transfer, an alternate schemes to improve physical representation of hydrometeors have emerged. For example, Galligani et al. (2021) have attempted to parametrize the polarisation signals observed from GMI at 89 and 166 GHz, with good results. However, their method is non-physical and limited to only these two frequencies as it requires actual dual-polarisation observations for the frequency to be simulated. For instance, the GMI 183 GHz channels are V polarised and a correction for these channels can not be presented in lack of parallel H polarised channels. Further, an extension of their approach to sub-mm wavelengths is not viable today due to a lack of observational data. A more general method to approximate the polarisation signatures was introduced by Barlakas et al. (2021) for comically scanning radiometers. Their method mimics the effect of particles with ARO by scaling the optical properties of particles with TRO. The scheme was tested with RTTOV-SCATT and is available as an option in the latest RTTOV version.

The preferentially oriented hydrometeors exhibit strong dichroism effects (Davis et al. 2007), which introduces significant differences in Tbs measured at Vertical (V) and Horizontal (H) polarisations by microwave imagers. On the other hand, TRO hydrometeors have no dichroism effect, hence the polarisation signals are very small and arise only due to scattering effects (Emde et al. 2004). The overall level of extinction and the degree of polarization vary with the observation incidence angle at the Earth surface. However, for an incidence angle of about 55°, typically used by conically scanning radiometers, the overall amount of extinction is unaffected by hydrometeor orientation (Brath et al. 2020). Based on this, Barlakas et al. (2021) introduced a correction factor (α) to approximate the differences in vertical (V) and horizontal (H) polarisations, as measured by conical scanning radiometers. The correction factor increases and decreases the layer optical thickness (τ) in the H and V polarised channels, respectively. The correction mimics the differences between H and V polarisations caused by ARO, and the ratio of the modified layer optical thickness gives the polarisation ratio:

$$\rho = \frac{\tau_H}{\tau_V} = \frac{1+\alpha}{1-\alpha}\,. \tag{1}$$

The factor α ($>= 0$) weakens the extinction at V polarisation and strengthens it at H polarisation. For TRO, ρ is equal to one, and for a-ARO, $\rho > 1$. Since this scheme approximates the ARO behaviour without computing the fully polarised radiative transfer, henceforth, the oriented particles are referred as "approximated ARO" or "aARO". A modified version of the scheme has been implemented within ARTS and tested with the high frequency measurements of GMI. In the modified version, instead of applying a constant polarisation ratio to entire data, we scale the extinction for H and V polarisations by a variable polarization ratio. A random selection of ρ from a uniform distribution between (1, 1.4) is made. The variable selection allows to simulate a wide spectrum of polarisation signals as observed in reality. Figure 6 shows a comparison of observed and simulated PDs versus the Tbs at 166 GHz (PD-$\mathrm{Tb_V}$ distribution) with TRO and aARO assumption. The PD-$\mathrm{Tb_V}$ distribution is composed of two main parts. The long arm with very high PD reaching up to 50 K is due to the surface contamination, while the arch type structure for Tbs below 250 K is due to PDs arising from the ice hydrometeor scattering. With only TRO assumptions, the PDs from hydrometeors are quite flat and the maximum magnitude that can be simulated is around 5 K, which clearly falls short of the realistic distribution. On the other hand, with aARO, the wide spectrum of polarisation signals can be reproduced correctly in a statistical sense. In the results shown above, the performance of the scheme is tested with a simple particle model: large plate aggregate and the PSD by (Field et al. 2007).

We also investigated the performance of the scheme in sub-mm range using $\rho \in$ U(1, 1.4). Figure 7 shows the PD-$\mathrm{Tb_V}$ distribution at 660 GHz. The arch shape of PDs is preserved at 660 GHz and the maximum PD is around 15 K which occurs for Tb around 190 K. 660 GHz is not sensitive to surface (weighting function peaks in the upper troposphere), hence the high PDs associated with surface contamination are missing. Similar results were also observed by Gong and Wu (2017) from the Compact Scanning Submillimter-wave Imaging Radiometer (CoSSIR) 640 GHz measurements. Their observations were mostly over the Pacific Ocean near Central America, and the maximum amplitude of PDs was observed to be 10 K. With an identical range of ρ, the maximum observed PDs at 166 GHz are higher than at 660 GHz. This indicates that the upper limit of ρ might increase with increasing frequencies. This is not unexpected as at 660 GHz, the sensitivity to small oriented hydrometeors is stronger than at 166 GHz. Previously, Gong and Wu (2017) had also concluded a similar behaviour while studying the effect of aspect ratio on the bell/arch curve (ρ is an indirect representation of aspect ratio). While it is difficult to predict the range of ρ, which will reproduce the future observations from ICI, selecting $\rho > 1.4$ is not ruled out to maintain a buffer zone of variability. One can argue that with limited observations from flight campaigns it is not feasible to completely solve the entire variability of PDs.

2.4 Representation of footprint

For the retrieval database, the possible use of simulated antenna patterns (or simplified Gaussian) will be considered. Representation of antenna patterns will be done as a post-processing step, however it has to be decided which antenna pattern and incidence angle to assume. Due to variation in ICI incidence angles, there is still the open point on which angle should be considered to simulate footprints. The elevation angle offset is ± 0.9° between ICI channels, which is not negligible. It will require investigation if different angles should be considered. However a consistency between the option used to generate the database and the RTTOV call used in the operational processor will be probably maintained in order to avoid biases. Additionally, to represent the antenna footprint, a number of pencil beam calculations will be required. Ways to manage the additional computational load are discussed in Sect. 3.2.

3 Other considerations

3.1 Data

Different types of data on vertical and horizontal structures are required to generate the hydrometeor fields. In this section, a detailed description of the various static and non-static atmospheric/surface fields in clear-sky and all-sky conditions is presented.

3.1.1 Static clear-sky atmospheric data

Under clear-sky condition, gas absorption effects need to be considered. In addition to water vapour, oxygen and nitrogen absorption, ozone will also be included. A gas absorption model is being formulated for RTTOV and ICI inside another EUMETSAT study being led by UK Met Office. We will apply the same absorption model in the database generation. Contact with Stuart Fox will be made.

3.1.2 Non-static clear-sky atmospheric data

The vertical temperature profiles and gas profiles for water vapour, nitrogen, oxygen, and ozone will be taken from ERA5 reanalyses.

3.1.3 Static surface data

Surface elevation, land-sea mask and surface emissivity estimates together form the static surface data for the synthetic scenes. Surface elevation will be collected from SRTM30, which is a near-global digital elevation model (DEM). A module to extract elevation estimates is available as a part of typhon (Lemke et al. 2020). The land-sea mask shall be sourced from ERA5.

The various options available for surface emissivity are described in Sect. 2.2, but here we provide a brief outlook. Surface emissivities for water bodies will be taken from the TESSEM2. For snow-free land surface types, TELSEM2 will be considered but options to randomise the emissivities for higher frequencies will be included. For snow-covered surfaces, we shall extend the probabilistic model described in Sect. 2.2 to sub-mm range.

3.1.4 Non-static surface data

Non-static surface data required to generate synthetic scenes are snow depth, sea ice fractions, and skin temperatures. Snow depth and sea ice fraction are necessary to augment the surface classification, and calculate the associated emissivities. ERA5 reanalyses will be used for snow depth, sea ice, and skin temperature.

3.1.5 Static hydrometeor data

Absorption model for liquid water content The liquid water content (LWC) will be assumed to be totally absorbing and the liquid absorption model by Ellison (2007) is likely to be used as first choice. However, we might consider using the same absorption settings as used by RTTOV.

Single scattering data The ARTS single scattering database (SSD, Eriksson et al. 2018) that covers 34 particle habits and frequencies up to 886.4 GHz is most suitable for this study. The only constraint this database has that its scope is limited to TRO. However, as described in Sect. 2.3.3, the inclusion of hydrometeors with ARO will be done in an approximate manner, which rules out the requirement of a polarised SSD.

Additionally, for ice hydrometeors, multiple habits will be applied. The selection of habits can be latitude specific or random. However, some initial testing will be required to pick out the best combinations.

Particle size distribution As described in Sect. 2.3.2, D14 meets the requirements of the current study, but it would be important to match the measured IWC-dbZ variability for consistency.

3.1.6 Non-static hydrometeor data

The non-static hydrometeor data for generating 3D scenarios can be extracted either from satellite observation or model outputs. Collocated profiles of observations over footprints are usually not enough to generate a comprehensive database. On the contrary, global model outputs can overcome this constraint but they are often limited by their own inherent uncertainties. For example, numerical weather prediction (NWP) models cannot resolve cloud heterogeneities, typically the small scale cloud structures and are often associated discrepancies in IWP estimates.

For this study, Cloudsat radar reflectivities, together with multispectral imaging data from MODIS, will be used as an input to generate 3D synthetic scenes. The radar reflectivities will be converted to ice water content (IWC) fields conditioned by the microphysical assumptions defined for the forward modelling. Further, Cloudsat has operated in daytime mode since 2011 due to a battery anomaly. To avoid bias due to only daytime coverage, we shall utilize Cloudsat observations prior to 2011. Due to insensitivity of Cloudsat to liquid water in the atmosphere, ERA5 reanalyses for liquid water content (LWC) shall be used to complement the Cloudsat derived IWC. To avoid high LWC outside clouds and precipitation as detected by Cloudsat, some LWC filtering may be required.

3.2 Radiative transfer

The Atmospheric Radiative Transfer Simulator (ARTS, Buehler et al. 2018) will be used to simulate the radiative transfer. To represent the antenna footprint, a number of pencil beam calculations will be required. The input data has a resolution in the order of 1.5 km. If we want to sample an area of 30 km×30 km at this resolution, that would result in 400 pencil beams to simulate. There is a similar consideration for the inclusion of sideband passband responses. A single (monochromatic) calculation is not sufficient to correctly simulate a full passband. As shown in Eriksson et al. (2020), some of the ICI passbands are affected by ozone and the spectral representation becomes especially important for these bands. It is hard to set a general number on the required number of frequencies per passband. If we assume five frequencies per passband to keep noise low, it would give us 130 frequencies in total (5×2×13). The scattering solvers handle one position and one frequency at a time. This means that a naive setup would result in about 52000 (400*130) calls of the scattering solver for each database case to generate. This is a number more than two orders of magnitude higher than used for the preliminary database and is simply too high. We will need to investigate manners to perform the simulation using fewer pencil beam directions and monochromatic frequencies without affecting the calculation accuracy too much. Potential solutions include using fewer pencil beams outside the main antenna lobe and using fewer monochromatic frequencies for the all-sky calculation than the clear-sky one. We will also use the central computing facilities (at Chalmers) to manage the additional computational load.

3.3 A priori weights

The ICI retrieval database must in practice contain a finite number of cases/states, and these cases should cover all possible clear and cloudy weather conditions and should ideally be sampled from the *a priori* probability density function. In practise this means that the retrieval database will primarily contain clear sky cases where many of these are close to identical to each other, and these cases will have an insignificant contribution to the retrieval for observations over cloudy areas.

The ICI retrieval algorithm is designed to consider *a priori* weight (the term a in Eq. 8 of Eriksson et al. (2020)). The idea is to combine database cases that can be seen as duplicates into a single case in order to allow for using a retrieval database that is smaller in size but still cover all possible weather conditions and enables the same retrieval performance as the original database. Or stated differently, the use of *a priori* weight/"thinning" allows for using a more complete retrieval database given that the size of the retrieval database is limited by a computational cost reason.

No algorithm for thinning the retrieval database is readily available, but should be developed inside this study. It should be possible to at least thin cases matching clear-sky conditions. This can potentially be done by applying an algorithm that identifies cases where the simulated brightness temperatures of all ICI channels are close to identical, at the same time as the underlying state matches (e.g. surface type and temperate matches).

4 Assessments

4.1 Validation of the cloud radiation retrieval database

A true validation of the cloud radiation retrieval database is difficult or even impossible to accomplish in the absence of actual ICI data. However, the method to generate the retrieval database can be applied to existing instruments that are similar to ICI to allow for making some assessments. We judge that the most relevant instruments for this exercise are the following two:

- The Global Precipitation Measurement (GPM) Microwave Imager (GMI): is a multi-channel, conical scanning, microwave radiometer (`https://gpm.nasa.gov/missions/GPM/GMI`). GMI has thirteen microwave channels operating at frequencies from 10 GHz to 183 GHz. The satellite was launched at 2014-02-27 into a circular, non-sun-synchronous orbit at an altitude of 407 km, and with an inclination of 65° to the equator. The off-nadir-angle defining the cone swept out by the GMI is set at 48.565° which results in an observation incidence angle of 52.8°. Observations are acquired within an angle of ±70° in azimuth w.r.t. the for view which gives a swath of around 900 km at the Earth's surface. With a 1.2 m diameter antenna the footprint resolution is around 5 km for the highest frequency channels.

- The International Submillimetre Airborne Radiometer (ISMAR):
  is an airborne submillimetre radiometer (Fox et al. 2017). ISMAR has been developed as an airborne demonstrator for the ICI, and thus have channels that overlaps the mm and sub-mm channels of ICI.

The generation of a retrieval database for GMI allows for testing the simulations for various climate conditions, while a comparison with ISMAR will be limited to northern mid-latitudes but involve sub-mm radiances. A first assessment will be to verify that simulated radiances are consistent to actual radiances in a statistical sense, as done in Fig. 7 in Eriksson et al. (2020).

A second and higher level of assesment will be to perform retrievals and compare against existing cloud ice products and verify that the data is sufficiently consistent to each other. The DARDAR (raDAR/liDAR) project (`https://www.icare.univ-lille.fr/dardar/`) provides cloud properties derived from combining the CloudSat radar and the CALIPSO lidar measurments, and these cloud ice products are to date probably the most trustable data available. Retrievals from GMI observations will therefore be compared to DARDAR data.

4.2 Retrieval performance

The ICI retrieval performance can be estimated by picking out one part from the retrieval database into a reference dataset, and by performing retrieval simulation on this reference dataset using the remaining part of the retrieval database in the retrieval calculation. Using this approach will allow for a trustable retrieval performance characterisation only if the retrieval database states represent real condition and variability closely, as the reference dataset and retrieval database are not completely independent of each other. However, the approach described above is considered to be sufficiently accurate given that the retrieval database was validated according to the description in Sect. 4.1. The ICI retrieval performance can then, and will, be characterised for various climatological conditions, i.e. by picking out several reference datasets covering tropical, mid-latitude and arctic winter and summer conditions.

The reference datasets used for estimating ICI retrieval performance for various climatological conditions will also be used to demonstrate the increased retrieval performance by using the new cloud radiation retrieval database as compared to using the preliminary retrieval database developed by the NWCSAF (Rydberg 2018). By performing retrieval simulations on the reference datasets but using the preliminary retrieval database in the retrieval calculation we obtain data in such a way that it is straightforward to compare the associated retrieval performance of the two cloud radiation retrieval databases.

References

Adams, I. S. and Bettenhausen, M. H.: 2012, The scattering properties of horizontally aligned snow crystals and crystal approximations at millimeter wavelengths, *Radio Sci.* **47**(05), 1–15.

Baordo, F. and Geer, A.: 2016, Assimilation of SSMIS humidity-sounding channels in all-sky conditions over land using a dynamic emissivity retrieval, *Q. J. R. Meteorol. Soc.* **142**(700), 2854–2866.

Barker, H. W., Jerg, M. P., Wehr, T., Kato, S., Donovan, D. P. and Hogan, R. J.: 2011, A 3D cloud-construction algorithm for the EarthCARE satellite mission, *Q. J. R. Meteorol. Soc.* **137**(657), 1042–1058.

Barlakas, V. and Eriksson, P.: 2020, Three dimensional radiative effects in passive millimeter/sub-millimeter all-sky observations, *Radio Sci.* **12**(3), 531.

Barlakas, V., Geer, A. J. and Eriksson, P.: 2021, Introducing hydrometeor orientation into all-sky microwave and submillimeter assimilatifbrathon, *Atmos. Meas. Tech.* **14**(5), 3427–3447.

Battaglia, A., Simmer, C. and Czekala, H.: 2006, Three-dimensional effects in polarization signatures as observed from precipitating clouds by low frequency ground-based microwave radiometers, *Atmos. Chem. Phys.* **6**(12), 4383–4394.

Bauer, P., Schanz, L. and Roberti, L.: 1998, Correction of three-dimensional effects for passive microwave remote sensing of convective clouds, *J. Appl. Meteorol. Clim.* **37**(12), 1619–1632.

Boukabara, S.-A., Garrett, K., Chen, W., Iturbide-Sanchez, F., Grassotti, C., Kongoli, C., Chen, R., Liu, Q., Yan, B., Weng, F. et al.: 2011, MiRS: An all-weather 1DVAR satellite data assimilation and retrieval system, *IEEE T. Geosci. Remote* **49**(9), 3249–3272.

Brath, M., Ekelund, R., Eriksson, P., Lemke, O. and Buehler, S. A.: 2020, Microwave and submillimeter wave scattering of oriented ice particles, *Atmos. Meas. Tech.* **13**(5), 2309–2333.

Brath, M., Fox, S., Eriksson, P., Harlow, R. C., Burgdorf, M. and Buehler, S. A.: 2018, Retrieval of an ice water path over the ocean from ISMAR and MARSS millimeter and submillimeter brightness temperatures, *Atmos. Meas. Tech.* **11**(1), 611–632.

Buehler, S. A., Mendrok, J., Eriksson, P., Perrin, A., Larsson, R. and Lemke, O.: 2018, ARTS, the atmospheric radiative transfer simulator – version 2.2, the planetary toolbox edition, *Geosci. Model Dev.* **11**(4), 1537–1556.

Camplani, A., Casella, D., Sanò, P. and Panegrossi, G.: 2021, The passive microwave empirical cold surface classification algorithm (PESCA): Application to GMI and ATMS, *J. Hydrometeorol.* **22**(7), 1727–1744.

Davis, C., Evans, K., Buehler, S., Wu, D. and Pumphrey, H.: 2007, 3-d polarised simulations of space-borne passive mm/sub-mm midlatitude cirrus observations: a case study, *Atmos. Chem. Phys.* **7**(15), 4149–4158.

Delanoë, J., Heymsfield, A. J., Protat, A., Bansemer, A. and Hogan, R.: 2014, Normalized particle size distribution for remote sensing application, *J. Geophys. Res. Atmos.* **119**(7), 4204–4227.

Ekelund, R., Eriksson, P. and Pfreundschuh, S.: 2020, Using passive and active observations at microwave and sub-millimetre wavelengths to constrain ice particle models, *Atmos. Meas. Tech.* **13**(2), 501–520.

Ellison, W.: 2007, Permittivity of pure water, at standard atmospheric pressure, over the frequency range 0–25 THz and the temperature range 0–100 C, *J. Phys. Chem. Ref. Data* **36**(1), 1–18.

Emde, C., Buehler, S., Davis, C., Eriksson, P., Sreerekha, T. and Teichmann, C.: 2004, A polarized discrete ordinate scattering model for simulations of limb and nadir long-wave measurements in 1-D/3-D spherical atmospheres, *J. Geophys. Res. Atmos.* **109**(D24).

Eriksson, P., Ekelund, R., Mendrok, J., Brath, M., Lemke, O. and Buehler, S. A.: 2018, A general database of hydrometeor single scattering properties at microwave and sub-millimetre wavelengths, *Earth Syst. Sci. Data* **10**(3), 1301–1326.

Eriksson, P., Rydberg, B., Mattioli, V., Thoss, A., Accadia, C., Klein, U. and Buehler, S. A.: 2020, Towards an operational ice cloud imager (ICI) retrieval product, *Atmos. Meas. Tech.* **13**(1), 53–71.

Evans, K., Wang, J., O'C Starr, D., Heymsfield, G., Li, L., Tian, L., Lawson, R., Heymsfield, A. and Bansemer, A.: 2012, Ice hydrometeor profile retrieval algorithm for high-frequency microwave radiometers: application to the CoSSIR instrument during TC4, *Atmos. Meas. Tech.* **5**(9), 2277–2306.

Field, P. R., Heymsfield, A. J. and Bansemer, A.: 2007, Snow size distribution parameterization for midlatitude and tropical ice clouds, *J. Atmos. Sci.* **64**(12), 4346 – 4365.

Fox, S.: 2020, An evaluation of radiative transfer simulations of cloudy scenes from a numerical weather prediction model at sub-millimetre frequencies using airborne observations, *Radio Sci.* **12**(17), 2758.

Fox, S., Lee, C., Moyna, B., Philipp, M., Rule, I., Rogers, S., King, R., Oldfield, M., Rea, S., Henry, M., Wang, H. and Harlow, R. C.: 2017, ISMAR: an airborne submillimetre radiometer, *Atmos. Meas. Tech.* **10**(2), 477–490.

Galligani, V. S., Wang, D., Corrales, P. B. and Prigent, C.: 2021, A parameterization of the cloud scattering polarization signal derived from GPM observations for microwave fast radative transfer models, *IEEE T. Geosci. Remote* **59**(11), 8968–8977.

Gong, J. and Wu, D. L.: 2017, Microphysical properties of frozen particles inferred from global precipitation measurement (GPM) microwave imager (GMI) polarimetric measurements, *Atmos. Chem. Phys.* **17**(4), 2741–2757.

Haggerty, J. A. and Curry, J. A.: 2001, Variability of sea ice emissivity estimated from airborne passive microwave measurements during FIRE SHEBA, *J. Geophys. Res. Atmos.* **106**(D14), 15265–15277.

Harlow, R. C.: 2009, Millimeter microwave emissivities and effective temperatures of snow-covered surfaces: Evidence for lambertian surface scattering, *IEEE Geosci. Remote Sens.* **47**(7), 1957–1970.

Harlow, R. C. and Essery, R.: 2012, Tundra snow emissivities at MHS frequencies:MEMLS validation using airborne microwave data measured during CLPX-II, *IEEE Geosci. Remote Sens.* **50**(11), 4262–4278.

Hewison, T. J.: 2001, Airborne measurements of forest and agricultural land surface emissivity at millimeter wavelengths, *IEEE T. Geosci. Remote* **39**(2), 393–400.

Hewison, T., Selbach, N., Heygster, G., Taylor, J. and McGrath, A.: 2002, Airborne measurements of Arctic sea ice, glacier and snow emissivity at 24-183 GHz, *IEEE International Geoscience and Remote Sensing Symposium*, Vol. 5, pp. 2851–2855 vol.5.

Hong, S.-Y. and Lim, J.-O. J.: 2006, The WRF single-moment 6-class microphysics scheme (WSM6), *Asia-Pacific Journal of Atmospheric Sciences* **42**(2), 129–151.

Karbou, F., Prigent, C., Eymard, L. and Pardo, J. R.: 2005, Microwave land emissivity calculations using AMSU measurements, *IEEE T. Geosci. Remote* **43**(5), 948–959.

Kummerow, C.: 1998, Beamfilling errors in passive microwave rainfall retrievals, *J. Appl. Meteorol.* **37**(4), 356–370.

Lemke, O., Kluft, L., Mrziglod, J., Pfreundschuh, S., Holl, G., Larsson, R., Yamada, T., Mieslinger, T. and Doerr, J.: 2020, atmtools/typhon: Typhon release 0.8.0.
URL: *https://doi.org/10.5281/zenodo.3626449*

Liu, Q., Weng, F. and English, S. J.: 2010, An improved fast microwave water emissivity model, *IEEE T. Geosci. Remote* **49**(4), 1238–1250.

McFarquhar, G. M. and Heymsfield, A. J.: 1997, Parameterization of tropical cirrus ice crystal size distributions and implications for radiative transfer: Results from CEPEX, *J. Atmos. Sci.* **54**(17), 2187–2200.

Moncet, J. and Clough, S.: 1997, Accelerated monochromatic radiative transfer for scattering atmospheres: Application of a new model to spectral radiance observations, *J. Geophys. Res. Atmos.* **102**(D18), 21853–21866.

Munchak, S. J., Ringerud, S., Brucker, L., You, Y., de Gelis, I. and Prigent, C.: 2020, An active–passive microwave land surface database from GPM, *IEEE T. Geosci. Remote* **58**(9), 6224–6242.

Prigent, C., Aires, F., Wang, D., Fox, S. and Harlow, C.: 2017, Sea-surface emissivity parametrization from microwaves to millimetre waves, *Q. J. R. Meteorol. Soc.* **143**(702), 596–605.

Protat, A., Delanoë, J., Strapp, J., Fontaine, E., Leroy, D., Schwarzenboeck, A., Lilie, L., Davison, C., Dezitter, F., Grandin, A. et al.: 2016, The measured relationship between ice water content and cloud radar reflectivity in tropical convective clouds, *J. Appl. Meteorol. Clim.* **55**(8), 1707–1729.

Randel, D. L., Kummerow, C. D. and Ringerud, S.: 2020, *Satellite Precipitation Measurement*, Vol. 1, Springer, Gewerbestrasse 11, 6330 Cham, Switzerland, chapter The Goddard Profiling (GPROF) Precipitation Retrieval Algorithm, pp. 141–152.

Rilemark, R. and Svensson, C.: 2020, *Generation of atmospheric cloud fields using generative adversarial networks*, Master's thesis, Department of Space, Earth and Environment, Chalmers University OF Technology, Gothenburg, Sweden.

Rydberg, B.: 2018, EPS-SG ICI ice water path product: ATBD, *Technical Report Issue 2.1, Rev. 2*, EUMETSAT, NWCSAF.

Rydberg, B., Eriksson, P. and Buehler, S.: 2007, Prediction of cloud ice signatures in submillimetre emission spectra by means of ground-based radar and in situ microphysical data, *Q. J. R. Meteorol. Soc.* **133**(S2), 151–162.

Rydberg, B., Eriksson, P., Buehler, S. and Murtagh, D. P.: 2009, Non-gaussian bayesian retrieval of tropical upper tropospheric cloud ice and water vapour from Odin-SMR measurements, *Atmos. Meas. Tech.* **2**(2), 621–637.

Tait, A., Hall, D., Foster, J., Chang, A. and Klein, A.: 1999, Detection of snow cover using millimeter-wave imaging radiometer (MIR) data, *Rem. Sen. Env.* **68**(1), 53–60.

Venema, V., Ament, F. and Simmer, C.: 2006, A stochastic iterative amplitude adjusted fourier transform algorithm with improved accuracy, *Nonlinear Processes in Geophysics* **13**(3), 321–328.

Wang, D., Prigent, C., Aires, F. and Jimenez, C.: 2016, A statistical retrieval of cloud parameters for the millimeter wave ice cloud imager on board MetOp-SG, *IEEE Access* **5**, 4057–4076.

Wang, D., Prigent, C., Kilic, L., Fox, S., Harlow, C., Jimenez, C., Aires, F., Grassotti, C. and Karbou, F.: 2017, Surface emissivity at microwaves to millimeter waves over polar regions: Parameterization and evaluation with aircraft experiments, *J. Atmos. Oceanic Technol.* **34**(5), 1039 – 1059.

Zängl, G., Reinert, D., Rípodas, P. and Baldauf, M.: 2015, The ICON (ICOsahedral Non-hydrostatic) modelling framework of DWD and MPI-M: Description of the non-hydrostatic dynamical core, *Q. J. R. Meteorol. Soc.* **141**(687), 563–579.

Appendices

A The Barker 3D cloud-construction algorithm

Barker et al. (2011) presents an algorithm that constructs 3D distributions of cloud from passive satellite imagery and collocated 2D nadir profiles of cloud properties inferred synergistically from lidar, cloud radar and imagery data. Effectively, the construction algorithm widens the active retrieved cross-section (RXS) of cloud properties, by using a match-and-substitute algorithm that fills in off-nadir pixels/columns with cloud property profiles from the swath center. For this purpose radiances from a passive multi spectral imager (MSI) for an off-nadir, recipient, pixel are compared with corresponding values for a range of pixels along the RXS; close matches are identified as potential donors with the closest to the recipient pixel being designated as the proxy to literally stand in for the recipient. This is repeated until all pixels in the required 3D domain are filled.

Required data

The Barker 3D construction algorithm requires the following data:

- a series of profiles (i.e. the nadir RXS) of cloud properties retrieved from either active instruments alone or in synergy with MSI data
- MSI data at wavelengths typical of conventional imagers at resolutions ideally less than the coarsest active instrument and extending in the across-track direction on both sides of the RXS

Main data sources used by Barker et al. (2011):

- MSI data from MODIS onboard Aqua
- retrieved data from the active sensors CALIOP onboard Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) and the cloud profiling radar onboard CloudSat.

Even though Barker et al. (2011) used retrieved 2D nadir profiles of cloud properties, the algorithm can alternatively be applied to construct more basic 3D fields, like for example radar backscatter fields. This is relevant for the present study, as this allows for a full control of e.g. cloud microphysical assumptions, for further use of the the constructed 3D fields. Anyhow, Aqua, CALIPSO, and CloudSat are all part of NASA's A-train constellation, and hence collocated data from the instruments above are available.

The MODIS instrument has a viewing swath width of 2,330 km and views the entire surface of the Earth every one to two days. Its detectors measure 36 spectral bands between 0.405 and 14.385 μm, and it acquires data at three spatial resolutions; 250 m, 500 m, and 1000 m. The most relevant MODIS channels for the construction algorithm as identified by Barker et al. (2011) are the following:

- 0.62 - 0.67 [μm],
- 2.105 - 2.155 [μm],
- 8.4 - 8.7 [μm],

- and 11.77 - 12.27 [μm],

where it should be clear that data from a channel operating at visible wavelengths are only relevant for day time observations. CloudSat is an experimental satellite that uses radar to observe clouds and precipitation from space. The Cloud Profiling Radar (CPR) is a 94-GHz nadir-looking radar which measures the power backscattered by clouds as a function of distance from the radar, and the along-track sampling is 2 km (1.4 km x 1.7 km across and along track resolution), with a dynamic range of 70 dB, and 500 m vertical resolution.

CALIPSO combines an active lidar instrument with passive infrared and visible imagers to probe the vertical structure and properties of thin clouds and aerosols over the globe. CALIOP is a two-wavelength polarization-sensitive lidar that provides high-resolution vertical profiles of aerosols and clouds. The vertical and horizontal resolution is 30-60 m and 333 m, respectively.

Identification of donor columns

A main part of the actual 3D construction algorithm is the identification of a suitable donor column to use for an off-nadir pixel. Barker et al. (2011) came up with the idea to use a local database consisting of all data points along the subsatellite track and within a distance closer than 200 km from the current position, and to select one of these points as the donor. If we assume that a pixel along the RXS is located at position $(i, 0)$, the intention is to fill all pixels (i, j) with a donor column, and we begin by computing the following cost function:

$$F(i,j;m) = \sum_{k=1}^{K} w_k \left(\frac{r_k(i,i) - r_k(m,0)}{r_k(i,j)} \right)^2 : m \in [i - m_1, i + m_2] \tag{2}$$

where the summation is over the KMSI the channels, w_k is a weighting factor, $r_k(i, j)$ is the radiance of channel k at pixel position (i, j) and $r_k(m, 0)$ is the radiance of channel k at pixel position $(m, 0)$ and part of the local candidate database, and m_1 can be selected such that the database only includes data points from within a desired window. Barker et al. (2011) describes that just considering the radiance match can result in that one selects an inappropriate donor profile, and to limit this risk it is preferable to also consider the distance between between a potential donor at $(m, 0)$ and the recipient at (i, j), that can simply be calculated as:

$$D(i,j;m) = \Delta L \sqrt{(i-m)^2 + j^2}, \tag{3}$$

where ΔL is the imager resolution. Barker et al. (2011) recommends that one selects the donor with the closest distance to the recipient from the 3 % of the database states having the smallest radiance cost function value. It is also recommended that the selection of candidate donors take into account surface types such that the donor and recipient have the same surface type.

Testing the algorithm by reconstructing RXS

The 3-D construction algorithm can be tested in various ways. One straightforward test one can do is to try to reconstruct a piece of RXS but using a ”dead zone” around it, such that data from this zone is not allowed to be used for the reconstruction. This is also done in Barker et al. (2011), and it was found that cloud masks were retrieved with only small errors out to distances to at least 20 km from the local database.

Development of a cloud radiation database for EPS-SG ICI: Task 2 Report

Inderpreet Kaur[1], Patrick Eriksson[1], Bengt Rydberg[2], Eleanor May[1], and Hanna Hallborn[1]

[1] Department of Space, Earth and Environment, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
[2] Swedish Meteorological and Hydrological Institute, SMHI, SE-601 76 Norrköping, Sweden

February 8, 2023

Contents

1 Introduction

The EUMETSAT Polar System (EPS) Second Generation (EPS-SG) will provide continuity of observations of the current EPS in the timeframe from 2025 onward. The Ice Cloud Imager (ICI) [RD-1] will be one of the missions of the EPS-SG. ICI is a passive conical-scanning radiometer observing in the microwave to sub-millimetre wave range of the spectrum with 13 channels from 183 GHz to 664 GHz. Its primary objective is the quantification of cloud ice in support of climate monitoring, the validation of ice cloud models and the parameterization of ice clouds in weather and climate models. The core geophysical products of ICI will be data characterising the bulk mass of ice particles and their size. The main retrieval quantities to be produced at the EUMETSAT Central Facilities are the integrated ice water path (IWP), the mean ice particle size by mass and mean mass height.

1.1 Purpose of this document

A retrieval algorithm based on Bayesian Monte Carlo integration (BMCI) has been developed [RD-2, RD-3], providing the inversion as a description of the posterior probability distribution by utilising a cloud radiation retrieval database. The retrieval database is the key component of the ICI retrieval algorithm, and consists of pairs of atmospheric/surface states or scenes along with associated simulated observations.

This report builds upon a literature review report (Kaur, Eriksson and Rydberg 2022) and aims to detail the work performed within Task 2 of the project: *Development of a cloud radiation database for EPS-SG ICI*, in support of ICI Level2 processing at the EUMETSAT Central Facilities.

1.2 Structure of this document

Section 2 and 3 give a description of the main assumptions, data, algorithms, and models used for the generation of scenes and for the ICI simulation, respectively. The limitations and possible future developments of the retrieval database are discussed in Section 4.

1.3 Reference documents

[RD-1] Accadia C. et al. (2020) Microwave and Sub-mm Wave Sensors: A European Perspective. In: Levizzani V., Kidd C., Kirschbaum D., Kummerow C., Nakamura K., Turk F. (eds) Satellite Precipitation Measurement. Advances in Global Change Research, vol 67. Springer, Cham. https://doi.org/10.1007/978-3-030-24568-9_5 .

[RD-2] Rydberg, B.: EPS-SG ICI ice water path product: ATBD. SAF/NWC/LEO-EPSSG/ATBD/IWP-ICI Issue 2.1, Rev. 2, Tech. rep., EUMETSAT. NWCSAF, www.nwcsaf.org, -¿ documentation, code: ICI, 2018.

[RD-3] Eriksson, P., Rydberg, B., Mattioli, V., Thoss, A., Accadia, C., Klein, U., and Buehler, S. A.: Towards an operational Ice Cloud Imager (ICI) retrieval product, Atmos. Meas. Tech., 13, 53–71, https://doi.org/10.5194/amt-13-53-2020, 2020.

2 Generation of scenes

2.1 Main assumptions

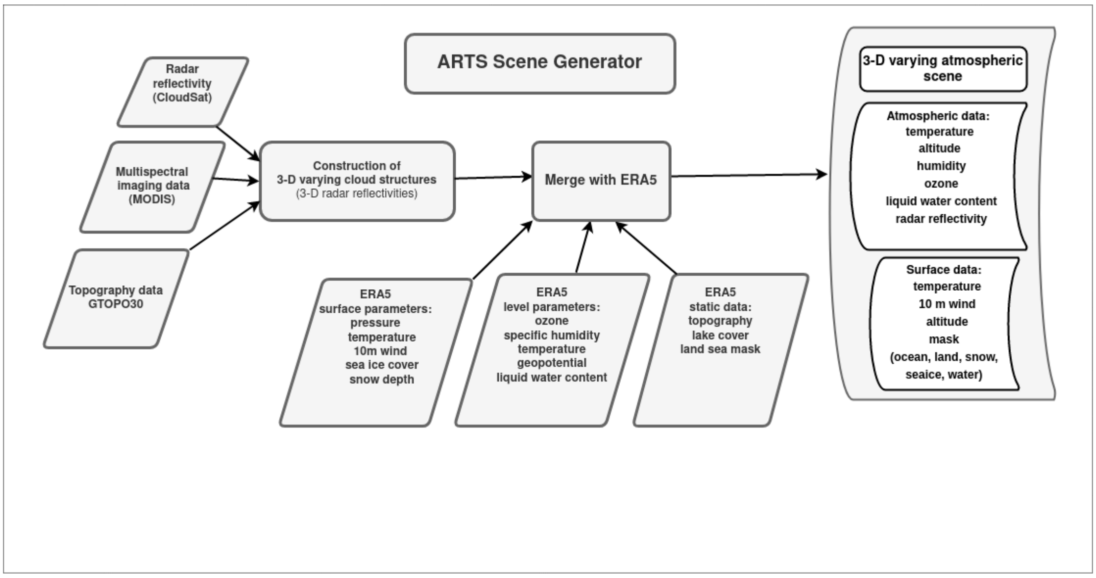

Figure 1: Flowchart summarising the input data, processes, and outputs of the scene generation scheme. The complete algorithm has been implemented within a Python toolbox denoted as the ARTS Scene Generator.

Atmospheric and surface states (or scenes) of the cloud radiation database must be realistic, include data for all parameters that have a significant impact on the ICI observation, and cover a wide range of conditions. It is assumed that the Cloud Profiling Radar (CPR), a 94 GHz near nadir-looking radar, onboard CloudSat, provides the best possible information on the vertical and horizontal cloud structures with a close to global coverage (82° S – 82° N). CPR data therefore serves as one of the main data source for the scene generation. However, the CPR only provides a two dimensional (2D) view of a scene, and the data must therefore be expanded in the across-track direction to allow for generating scenes that have a horizontal extension that is greater than the footprint size of ICI, of about 16 km. This is of importance in order to resolve 'beam filling issues' (Barlakas and Eriksson 2020) that otherwise ultimately can lead to a bias of the ICI level2 product. The generated scene must also include a description of atmospheric gases and temperature, and the surface type and state, and it is assumed that ERA5 is the best available source for this kind of data.

Figure 1 depicts the input data, processes, and output data of the scene generation scheme developed for the study. The scene generation scheme is responsible for generating 3D varying scenes, or input to the radiative transfer part of the ICI simulation that is handled by the Atmospheric Radiative Transfer Simulator (ARTS, Buehler et al. 2018). The scene generation is described in the following parts of this section. Note that the scene generation scheme is only responsible for generating 3D varying radar reflectivity fields and not underlying microphysical properties like ice and rain water content fields. ARTS provides functionality for retrieving ice and rain water content from radar reflectivity profiles, and this is described in more detail in Section 3. However, the main assumption/strategy is that ice and rain water content fields are constructed to be fully consistent to the generated 3D radar reflectivity fields, by making skilled assumptions on hydrometeor size distributions and scattering properties.

2.2 Construction of 3D varying cloud structure

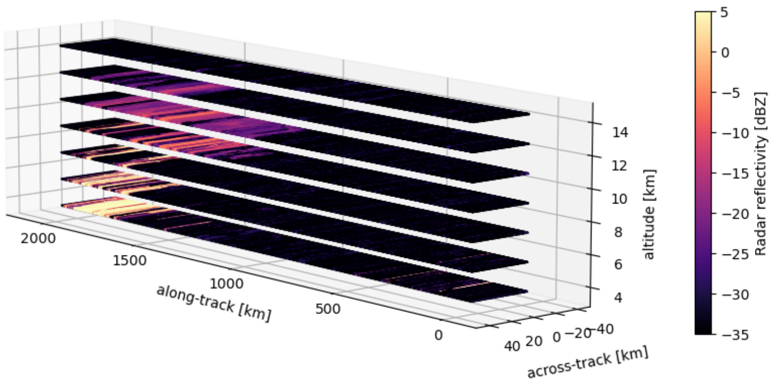

Figure 2: Example of a generated 3D varying radar reflectivity field.

Cloud structures having a three dimensional variability are generated following an algorithm originally developed by Barker et al. (2011). The details of the algorithm, and setup for the current study, is decribed in Appendix A. In short, the algorithm combines collocated 2D stretches of radar reflectivity from the CPR onboard CloudSat and multi-spectral imaging data from MODIS onboard Aqua, and effectively widens the radar data by using a match-and-substitute algorithm that fills in off-nadir pixels/columns with radar reflectivity profiles from the swath center. The algorithm applied for the current study is set up to generate 3D varying scenes with a horizontal extension of 2000 km x 50 km (in the along and across track direction of the CloudSat sub-satellite path), and one such example is shown in Figure 2. The distance on ground between two adjacent ICI scans is about 10 km, and the generated scene then covers and allows for simulating the center position of about 200 ICI scans, if we assume that the platform of ICI is flying within the CloudSat orbital plane (further described in Section 3).

CloudSat has been in operation since 2006, and was initially placed in a sun-synchronous orbit with a 13:31 h local time ascending node. Data covering the period 2009-01-01 to 2010-12-31, are used here to generate thousands of scenes like the one displayed in Figure 2. This period of time is chosen as the CPR was then operated during both day and night time part of the orbit, and the period covers both El Niño and La Niña phases.

2.3 Merging with background atmospheric/surface data

Data from ERA5 reanalysis, available at 0.25° resolution, are used to fully describe the atmospheric structure and surface condition of the 3D scene. The following ERA5 parameters are used: temperature, humidity, ozone, liquid water content, geopotential, skin temperature, 10 m wind, snow depth, and sea-ice concentration. Atmospheric data and 10 m wind speed and direction are interpolated onto the grid of the 3D scene. Snow depth and sea-ice concentration data are used to generate a surface type mask, but prior to this sea-ice concentration data are preprocessed, as described in Appendix C, to result in a sea ice mask (no sea-ice or sea-ice). Possible surface type values are: land, ocean, inland water, snow, and sea-ice. Surface type is classified as snow if the snow-depth (water equivalent) is more than 3 mm.

3 ICI simulation

3.1 Main assumptions

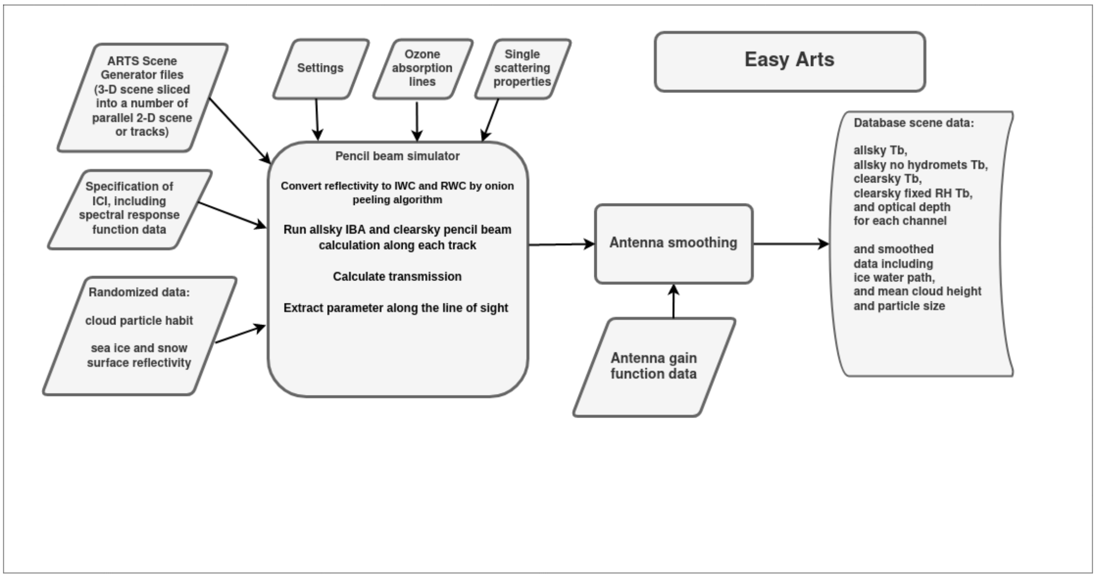

Figure 3: Flowchart summarising the input data, processes, and outputs of the scheme developed for the ICI simulation.

The Atmospheric Radiative Transfer Simulator (ARTS, Buehler et al. 2018) is used as the engine for the radiative transfer part of the ICI simulation, and Figure 3 depicts the input data, processes, and output data of the ICI simulation. In addition to this the simulation incorporates/takes into account

- state-of-the-art atmospheric absorption models,
- state-of-the-art surface emissivity models with an extension to cover variability over sea-ice and snow covered surfaces,
- multiple ice particle size distribution and habits (based on state-of-the-art single scattering property data), where choice of particle model and frequency of selection are chosen statistically, to represent reality,
- particle orientation is accounted for by mimicing the effect of azimuthally oriented particles, allowing for polarisation effects to be captured,
- antenna pattern and spectral response function data based on measured ICI data,

- scattering calculations using ARTS' interface to the DISORT scattering solver,
- simulation setup and antenna smoothing are applied to be consistent remapped data onto the field of view of ICI-01V.

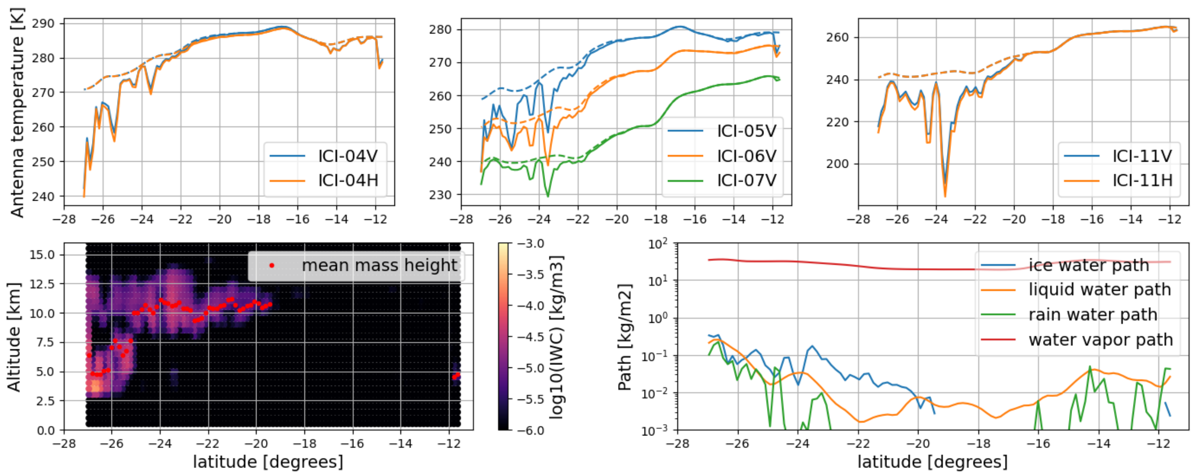

Figure 4: The upper panels show antenna temperature for seven of the ICI channels for a simulation based on the cloud structure displayed in Figure 2. The lower panel shows the underlying cloud, rain, and water vapour properties of the volume sampled by the simulated observation.

Thus, data are prepared and ARTS is setup to simulate ICI observation in a pseudo 3D mode, taking into account most or all relevant details, and Figure 4 displays a part of the final output of the ICI simulation for the scene displayed in Figure 2. The details of the simulation setup are described in the following parts of this section.

3.2 Surface emissivities

The Tool to Estimate Sea-Surface Emissivity from Microwaves to sub-Millimeter waves (TESSEM2, Prigent et al. 2017) and the Tool to Estimate Land-Surface Emissivities at Microwave frequencies (TELSEM2, Wang et al. 2017) are applied for the ICI simulation for ocean and land surfaces, respectively. For a surface type classified as snow or sea-ice covered an empirical and stochastic model described in Appendix C is applied.

3.3 Atmospheric absorption model

The absorption/emission of four gases are considered:

Nitrogen The continuum model of Liebe et al. (1993) is applied.

Oxygen The model of Rosenkranz (1993) is used, including both molecular transitions and a continuum term.

Water vapour Molecular transitions up to 1.65 THz are considered, of both the main and minor isotopologues. These data originate from the Atmospheric & Environmental Research group (`rtweb.aer.com`). Parameterisations for both "self" and "foreign" continuum are taken from the CKD MT v3.5 model (`github.com/AER-RC/MT_CKD`).

Ozone A selection of transitions from the "JPL line catalogue" (`spec.jpl.nasa.gov`) are considered.

The choices for water vapour and ozone are based on input and data provided by Stuart Fox (Met Office), based on experience from an ongoing EUMETSAT founded study. This study aims at updating and extending the microwave part of RTTOV.

3.4 Hydrometeor model

Particle Model	Habit	PSD	aARO Factor	p_i
AA1	Large plate aggregate	F07 - Tropics	1 - 1.6	0.3
AA2	Large column aggregate	F07 - Tropics	1 - 1.6	0.1
AA3	Large block aggregate	D14	1 - 1.6	0.13
IWC	Six bullet rosette	D14	1 - 1.6	0.2
Snow	Evans snow aggregate	F07 - Midlatitude	1.4 - 1.6	0.1
Graupel	Eight column aggregate	D14	1 - 1.2	0.17

Table 1: Particle models used within ICI simulations, given alongside habit and particle size distribution (PSD), where the source for the PSD is given in each case. The approximate azimuthally random orientation (aARO) scaling factor is randomly chosen within the given range. Also given is the occurrence fraction, or the probability of selection within a simulation, denoted as p_i. A further discussion of p_i is given in Appendix B

A set of six particle models representing frozen hydrometeors is used within the ICI simulations, where a model consists of a choice of particle size distribution (PSD), habit (shape), a factor for approximating particle orientation and an occurrence fraction. The six particle models are shown in Table 1. The selection among these six particle models, for use in a specific simulation, is made through a probability factor. The set of probability factors directly corresponds to the occurrence fraction, i.e. the fraction of simulations run with a given particle model, if the total number of simulations is high enough. Testing was performed to determine the best combination of probability factors, and discussed in Appendix B.

For rain a fixed particle model is used.

Particle size distribution

The PSD describes how the particle sizes are distributed within a volume element. For frozen hydrometeors three PSDs are applied. The first two are from Field et al. (2007), also known as F07, which has different settings for either the tropics or mid-latitudes. The third PSD used is based on Delanoë et al. (2014), here denoted as D14. A one-moment version of D14 is applied (with N_0^* parameterised as a function of temperature). Within the Literature Study, it was determined that F07 and D14 were the most suitable choices of PSD for snow, hence their inclusion. For rain a PSD described in Abel and Boutle (2012) is applied.

Habit/single scattering data

The single scattering data applied were taken from the ARTS single scattering database (Eriksson et al. 2018). Six of the database's "standard habits" were selected. The names of these habits are found in Table 1. The habits of the IWC, SWC and GWC particle models consist of particles with the same basic shape for all sizes. For the three AA (ARTS aggregate) particle models, the habit name refers to the particles covering sizes above about 150 m. For these later habits, single crystals are used as a complement to obtain full coverage of the relevant size span.

Approximation of particle orientation

Particle orientation is approximated following the scheme first developed by Barlakas et al. (2021) and later extended by Kaur, Eriksson, Barlakas, Pfreundschuh and Fox (2022) to operate in a random fashion. The basic idea of the scheme is simple, the extinction obtained assuming totally random orientation (TRO) is scaled differently for V and H polarisation to approximate azimuthally random orientation (aARO).

The main parameter of the scheme is the polarisation ratio, ρ:

$$\rho = \frac{\tau_H}{\tau_V} \tag{1}$$

where τ_H is the optical thickness after scaling (linearly proportional to the extinction) for polarisation H, and τ_H defined in the same way for polarisation V. The original TRO data matches to set $\rho = 1$, that gives $\tau_H = \tau_V$.

Observations of GMI at 166 GHz have been used to derive ρ. For assimilation a fixed value is preferred, and Barlakas et al. (2021) selected $\rho = 1.4$. The distribution of GMI 166 GHz data can be replicated even better by making a random selection of ρ, as shown by Kaur, Eriksson, Barlakas, Pfreundschuh and Fox (2022). A flat distribution between 1.0 and 1.5 was found to work well. The simulation performed here follows the later, random approach, but ρ is allowed to reach 1.6 as higher polarisation could potentially occur at the sub-millimetre wavelengths.

To what extent τ_H and τ_v deviates from TRO value, for a given ρ, is derived from:

$$\rho = \frac{\tau_H}{\tau_V} = \frac{1 + ra}{1 - a}. \tag{2}$$

This expression is applied with r as a fixed parameter, and the value of a is then implied by ρ. This is an extension of Barlakas et al. (2021), that devised the scheme by just comparing TRO and ARO extinctions and available data were fitted well without introducing r (then effectively applied $r = 1$). However, later full radiative transfer calculations revealed that, for final brightness temperatures, there is a considerably larger deviation from TRO for H and V (Barlakas et al. 2022). To mimic this behaviour, the parameter r was added, and the value $r = 5$ was eventually selected.

3.5 Radar onion peeling algorithm

An algorithm denoted as the radar onion peeling is used to convert radar reflectivity to cloud ice and rain water content. The algorithm first generates lookup-tables for ice and rain water content as function of radar reflectivity and temperature, taking into account the particle size distribution and single scattering data of the selected particle model. The ice water content and layer transmission are then retrieved sequentially from the top to bottom layer in an onion peeling approach. The reflectivity is assumed to origin from rain drop for temperatures above 273.15 K. To avoid surface clutter, reflectivies below 750 m over the ocean and below 1500 m above other surfaces, are replaced with the first value above the defined "clutter zone".

3.6 Frequency grid setup

Both the spectral response function and contamination by ozone absorption/emission varies within the passbands of the ICI channels (see Appendix D). The ARTS core radiative transfer simulation is made using a number of monochramatic frequencies and it is therefore of importance that the frequency grid applied is fine enough to capture these variations. Table 3 describes the estimated number of frequencies needed to obtain a clearsky simulation that has an error that is smaller than about 5 % of the $NE\Delta T$ value associated to the channel (Table 2). The number of frequencies varies between 3 and 25 for each sideband of the ICI channels, and these numbers are also used for the clear sky simulation, and integrated to obtain channel averaged values as described in Appendix D.

For the ICI database simulation we need to calculate the cloud signal (ΔTb) associated to each channel, or the difference between an allsky simulation and a reference clear sky simulation with a fixed relative humidity. To obtain ΔTb we run four type of simulations:

- Tb_{as}: allsky simulation using DISORT scattering solver of ARTS and only for the center frequency of the two sidebands
- $Tb_{as-no-hm}$: allsky simulation with no included hydrometeors using the DISORT scattering solver of ARTS and only for the center frequency of the two sidebands
- Tb_{cs}: clearsky simulation using the frequency grid setup in Table 3
- $Tb_{cs-fixed-rh}$: clearsky simulation with fixed relative humidity using the frequency grid setup in Table 3

and obtain the cloud signal as:

$$\Delta Tb = Tb_{cs} + Tb_{as} - Tb_{as-no-hm} - Tb_{cs-fixed-rh}. \tag{3}$$

The two allsky simulations are performed with the DISORT scattering solver of ARTS. They are more computational expensive than the ARTS clear sky simulation, and hence less frequencies are used. The simulation noise of the obtained ΔTb due to the frequency grid used is then mainly due to the scattering part of the simulation. The resulting simulation noise is assumed to be small compared to uncertainties related to the representativeness of the assumed microphysical and single scattering properties of the involved hydrometeors.

3.7 Geometry setup

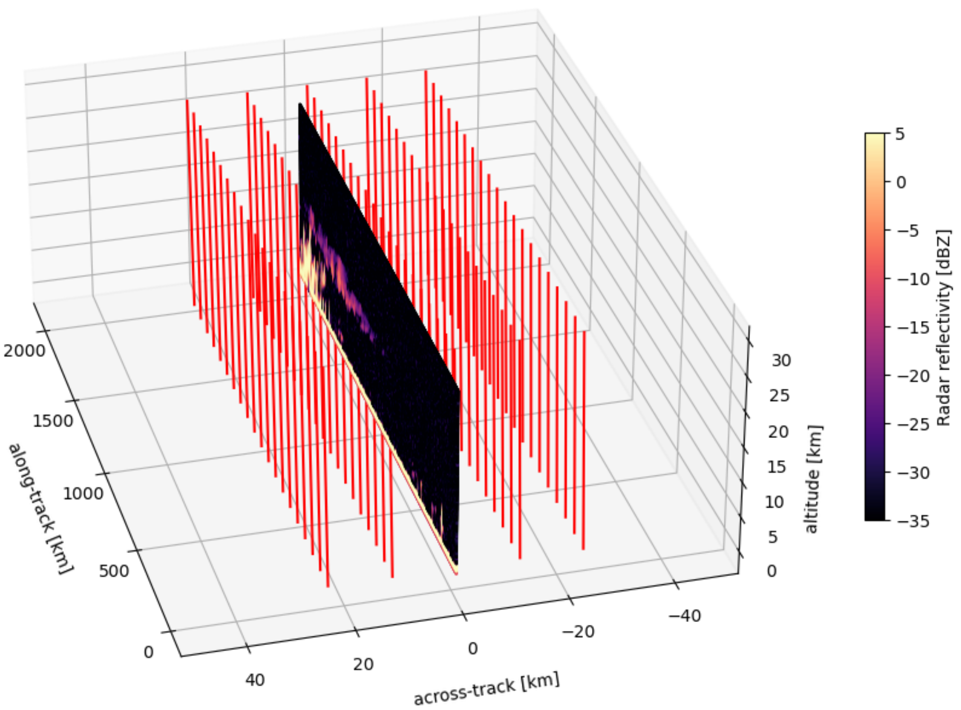

Figure 5: Schematic of the geometry setup of the radiative transfer simulation of a scene.

Figure 5 shows a schematic of the geometry setup of the core radiative transfer simulation covering a scene. The 3D scene is sliced into a number of parallel 2D slices, and a number of target positions are defined at the surface level along each 2D slice or along tracks. A number of pencil beam simulations are then performed for these target positions. The pencil beam simulation is performed using three different incidence angles taking into account that the viewing angle varies slightly among the ICI channels. The simulations assume a fixed Earth radius of 6371 km and a platform altitude of 832 km and the applied incidence angles are:

1. 53.84575349° for ICI-01V – ICI-03V and ICI-05V – ICI-10V
2. 51.80202686° for ICI-04V and ICI-11V
3. 51.70808917° for ICI-04H and ICI-11H

The setting of the number of 2D slices and distance between target positions to include in the simulation clearly impact the obtained accuracy of antenna weighted data, and more details around this are found in Section 3.8 and Appendix E.

3.8 Antenna smoothing

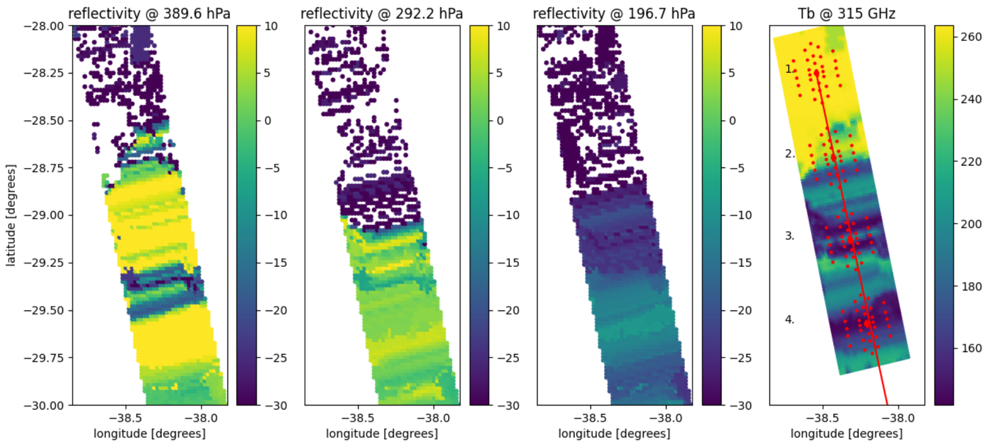

Figure 6: The three panels to the left shows radar reflectivity at three different pressure levels. The right panel shows simulated brightness temperatures at a frequency of 315 GHz. The simulation was done by dividing the scene into 4 smaller sub-scenes (denoted by numbers 1 to 4 in the right panel). The red markers around the center position of each smaller scene indicates the position of a subset of pencil beam simulations, used to generate the plots found in Figure 7. The red line from the center position of scene 1 points in the direction towards the sensor.

The previous section describes the geometry setup for the pencil beam simulation of a scene. The antenna smoothing calculation utilizes the pencil beam dataset, and is performed in order to be consistent to data remapped onto the view of ICI-01V. Let us then consider a target position along the center track of the scene and that the boresight of the antenna is directed towards this target point (e.g. consider one of the larger red markers of scene 4 in Figure 6) The antenna temperature can then be estimated in the following way:

1. Estimate the sensor position from the sensor azimuth and incidence angle (for that of ICI-01V) and platform height.

2. Estimate the relative zenith and azimuth angles (*dlos* in ARTS nomenclature) from the sensor position and for the surrounding samples of the pencil beam dataset.

3. Regrid brightness temperatures onto all positions of the ICI antenna pattern function, i.e. a function of relative zenith and azimuth angles. For the regridding we apply a linear interpolation within the convex hull of the dataset, and a nearest neighbour interpolation/extrapolation is applied outside the convex hull (see Figure 7 for example of regridded data)

4. Antenna temperature can then be obtained through integrating the gridded brightness temperature data over the antenna gain function (G) associated to ICI-01V, i.e:

$$T_a = \int_{\Omega} T_b(\Omega) G(\Omega) d\Omega, \tag{4}$$

   where Ω is the solid angle.

More details of the track setting and accuracies of antenna smoothed data are given in Appendix E.

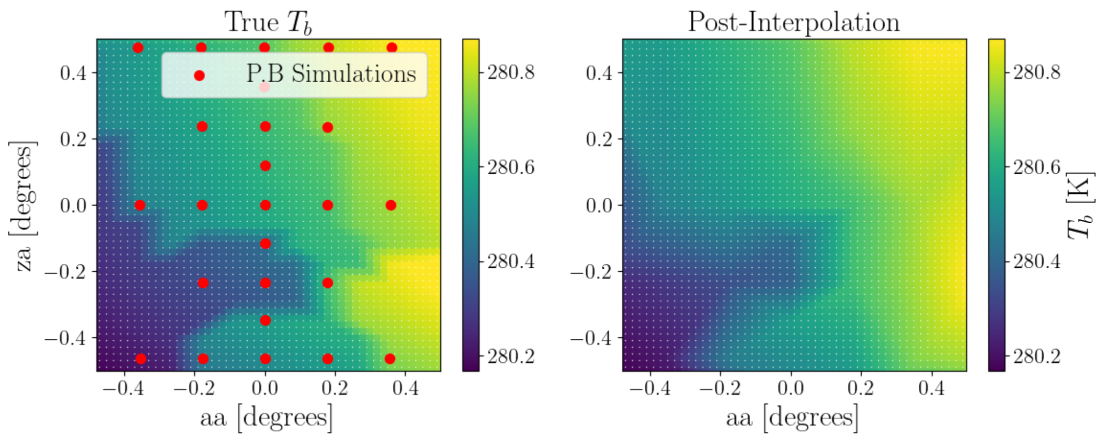

Figure 7: Both panels show brightness temperatures as function of relative zenith and azimuth angles. The red markers in the left panel are positions of pencil beam simulations, where the centre marker is the point at which the boresight of the antenna is pointed towards. A positive zenith and azimuth angle corresponds to a position below and to the right of the antenna boresight direction. The right panel shows the data re-gridded onto the positions of the ICI-01V antenna pattern grid through the use of interpolation, as described in step 3 of the antenna temperature procedure described above. The brightness temperatures in the left panel, denoted 'True', were obtained by running simulations on each grid point. They are included in the figure to act as a comparison to the interpolated temperatures.

4 Summary and outlook

A Python simulation environment (around the ARTS software) has been implemented to allow the simulation of ICI observations in great detail. To the best of our knowledge, on the overall level these should be the most realistic ICI simulations ever performed. Compared to the preliminary database, the level of detail of the simulations has been improved in several ways, including:

- The spectral response functions and interference of ozone are now fully included.
- The azimuthal dimension of the antenna pattern is now considered and the spatial variation of brightness temperatures is represented on a level corresponding to about 38 pencil beam calculations for each final antenna temperature.
- The set of particle models has been revised. In addition, the particle models now include an uneven occurrence fraction and an approximation of particle orientation.

For clarity it is stressed that the simulations mimic remapped ICI data (to the footprint of ICI-1V), and not original data.

Testing of the simulated data is part of Task 3 and details will be reported later, but the preliminary analysis indicates that both GMI (at 166 and 183 GHz) and ISMAR measurements can be statistically reproduced with the simulation toolbox developed.

The simulations that will form the retrieval database represent the state-of-the-art, but there, of course, exists room for improvement. The possible extensions can be divided into two categories. Useful development of technical nature includes:

- To generate more single scattering data matching the assumption of ARO (azimuthally random orientation). In a first step, this would allow to better test the scheme for approximating particle orientation (aARO), but would also open up for running the simulations using actual ARO data.
- Melting ice is currently not considered at all, due to lack of single scattering data for such particles. Models to generate melting particles are at hand, but the calculation of the single scattering properties appears to be highly computationally demanding (Kanngießer and Eriksson 2022).
- To make ARTS more computationally efficient, to allow for the generation of larger databases and a higher degree of flexibility in setting up the simulations. This is a general remark, but is particularly true for vector radiative transfer, see further below.
- When single scattering data of ARO type exist for at least three habits, abandoning the aARO scheme should be considered. In practice this means to perform vector radiative transfer, but inside ARTS such calculations have today two main "botlenecks":
  - The internal handling of ARO data is unefficient. Work to improve this handling is ongoing.
  - DISORT does not handle ARO and the solver of choice should instead be RT4. However, the later solver is less efficient and is difficult to use with parallelisation (as only at hand as Fortran code). The RT4 code should be fully integrated into ARTS.
- The scheme used here to extra- and interpolate pencil beam calculations to the antenna pattern should be integrated into ARTS. This would allow to more efficiently generate final antenna temperatures inside ARTS, and would make the simulation toolbox easier to use and extend.

All these improvements could be started today, but they also all represent a considerable undertaking and could not be performed inside this study.

A second category is to tune the settings of the simulations, to better represent reality. Further analysis of available ISMAR data should be considered, but that would still only cover northern mid-latitudes conditions. For this reason, it is here assumed that it is today very difficult to pinpoint the main weaknesses and a full revision is only possible after the launch of ICI. A rough plan for the work to be performed when ICI radiances are at hand is:

- Do there exist any situation where the observations consistently are outside the range of the database? Possible reasons and associated actions:
  - Deviations for "clear-sky" and channels with low surface sensitivity can largely be compensated with applying the bias terms of the retrieval algorithm, but should result in a revison of the gas absorption model.

  - Deviations for "clear-sky" and channels with significant surface contribution more likely depend on poor parameterisation of the surface emissivity. A first action is to screen out affected channels, and later trying to improve the emissivity model.
  - Deviations for data with clear signatures of hydrometeors can be of various complexity. For example, if higher differences between V and H channels (at 243 and 664 GHz) are seen in observations than simulations, the solution is simply to extend the database with cases having a higher polarisation ratio (ρ). If the deviations are more broadly observed among the channels, one or several additional habits should be the first option to consider.

- Reversely, is there any particle model in the database that have a very low match with observations (such as that the particle model is the best below a rate of e.g. 1%). If that is the case, it could be considered to remove, or replace, that particle model.

- The size of the database will be of the order 10^6. This equals roughly the number of ICI observations per day, and there will be data points outside of the database range just due to statistical reasons. That is, combinations of channel antenna temperatures occuring with a rate of e.g. 10^{-7} can not be expected to be covered well by the database. It should be investigated where these cases occur and how the retrieval algorithm handles these cases.

- When the quality of the database is understood on a general level, the next step should be to investigate it for particular conditions and to possibly start fine tuning the particle models. This work goes into uncharted territory and strategies for the analysis need to be formulated. But an example to clarify the type of studies in mind: Let's assume that the snow particle model is the one giving best match with observations in the anvil regions around tropical convection. A consequence will be that retrieved IWP will be unexpectedly high. As anvils are expected to contain relatively compact crystals and aggregates, this then indicates that none of the particle size distributions applied for the later habits is valid for anvils, and alternative size distributions should be tested. (Please note, just a hypothetical example).

- Looking further ahead, infrared radiances should be added to the database. That would allow for the use of geostationary infrared radiances to further constrain the particle models.

- Besides improving absorption, particle and surface emissivity models, it should be considered to simulate scenes of ICI antenna temperatures, at native observation geometry. That would allow the retrieval to make optional use of the overlap between footprints, but would require replacing the BMCI algorithm with machine learning (as well as a much faster ARTS).

References

Abel, S. J. and Boutle, I. A.: 2012, An improved representation of the raindrop size distribution for single-moment microphysics schemes, *QJRMS* **138**(669), 2151–2162.
URL: *https://rmets.onlinelibrary.wiley.com/doi/abs/10.1002/qj.1949*

Barker, H. W., Jerg, M. P., Wehr, T., Kato, S., Donovan, D. P. and Hogan, R. J.: 2011, A 3D cloud-construction algorithm for the EarthCARE satellite mission, *Q. J. R. Meteorol. Soc.* **137**(657), 1042–1058.

Barlakas, V. and Eriksson, P.: 2020, Three dimensional radiative effects in passive millimeter/sub-millimeter all-sky observations, *Radio Sci.* **12**(3), 531.

Barlakas, V., Geer, A. J. and Eriksson, P.: 2021, Introducing hydrometeor orientation into all-sky microwave and submillimeter assimilatifbrathon, *Atmos. Meas. Tech.* **14**(5), 3427–3447.

Barlakas, V., Geer, A. J. and Eriksson, P.: 2022, Cloud particle orientation and polarisation for cross-track microwave sensors in rttov, *EUMETSAT, NWCSAF* .

Buehler, S. A., Mendrok, J., Eriksson, P., Perrin, A., Larsson, R. and Lemke, O.: 2018, ARTS, the atmospheric radiative transfer simulator – version 2.2, the planetary toolbox edition, *Geosci. Model Dev.* **11**(4), 1537–1556.

Delanoë, J., Heymsfield, A. J., Protat, A., Bansemer, A. and Hogan, R.: 2014, Normalized particle size distribution for remote sensing application, *J. Geophys. Res. Atmos.* **119**(7), 4204–4227.

Ekelund, R., Eriksson, P. and Pfreundschuh, S.: 2020, Using passive and active observations at microwave and sub-millimetre wavelengths to constrain ice particle models, *Atmos. Meas. Tech.* **13**(2), 501–520.

Eriksson, P., Ekelund, R., Mendrok, J., Brath, M., Lemke, O. and Buehler, S. A.: 2018, A general database of hydrometeor single scattering properties at microwave and sub-millimetre wavelengths, *Earth Syst. Sci. Data* **10**(3), 1301–1326.

Field, P. R., Heymsfield, A. J. and Bansemer, A.: 2007, Snow size distribution parameterization for midlatitude and tropical ice clouds, *J. Atmos. Sci.* **64**(12), 4346 – 4365.

Fox, S., Lee, C., Moyna, B., Philipp, M., Rule, I., Rogers, S., King, R., Oldfield, M., Rea, S., Henry, M., Wang, H. and Harlow, R. C.: 2017, ISMAR: an airborne submillimetre radiometer, *Atmos. Meas. Tech.* **10**(2), 477–490.

Harlow, R. C.: 2009, Millimeter microwave emissivities and effective temperatures of snow-covered surfaces: Evidence for lambertian surface scattering, *IEEE Geosci. Remote Sens.* **47**(7), 1957–1970.

Harlow, R. C. and Essery, R.: 2012, Tundra snow emissivities at MHS frequencies:MEMLS validation using airborne microwave data measured during CLPX-II, *IEEE Geosci. Remote Sens.* **50**(11), 4262–4278.

Hewison, T., Selbach, N., Heygster, G., Taylor, J. and McGrath, A.: 2002, Airborne measurements of Arctic sea ice, glacier and snow emissivity at 24-183 GHz, *IEEE International Geoscience and Remote Sensing Symposium*, Vol. 5, pp. 2851–2855 vol.5.

Kanngießer, F. and Eriksson, P.: 2022, Cautious note on using the discrete dipole approximation for inhomogeneous, spherical scatterers with high optical contrast, *Optics Letters* **47**(16), 4203–4206.

Kaur, I., Eriksson, P., Barlakas, V., Pfreundschuh, S. and Fox, S.: 2022, Fast radiative transfer approximating ice hydrometeor orientation and its implication on IWP retrievals, *Remote sensing* **14**(7).

Kaur, I., Eriksson, P. and Rydberg, B.: 2022, Literature review to support the development of a cloud radiation database for eps-sg ici iwp retrieval, *Technical report*, Department of Space, Earth and Environemnt, Chalmers University of Technology.

Liebe, H. J., Hufford, G. A. and Cotton, M. G.: 1993, Propagation modeling of moist air and suspended water/ice particles at frequencies below 1000 GHz., *AGARD 52nd Specialists Meeting of the Electromagnetic Wave Propagation Panel, Palma de Mallorca, Spain.* ftp://ftp.its.bldrdoc.gov/pub/mpm93/.

Munchak, S. J., Ringerud, S., Brucker, L., You, Y., de Gelis, I. and Prigent, C.: 2020, An active–passive microwave land surface database from GPM, *IEEE T. Geosci. Remote* **58**(9), 6224–6242.

Prigent, C., Aires, F., Wang, D., Fox, S. and Harlow, C.: 2017, Sea-surface emissivity parametrization from microwaves to millimetre waves, *Q. J. R. Meteorol. Soc.* **143**(702), 596–605.

Risse, N.: 2021, Microwave emissivity of sea ice from airborne observations, *Master thesis*, University of Bonn.

Rosenkranz, P. W.: 1993, Absorption of microwaves by atmospheric gases, *in* M. A. Janssen (ed.), *Atmospheric remote sensing by microwave radiometry*, John Wiley & Sons, Inc., pp. 37–90. ftp://mesa.mit.edu/phil/lbl_rt.

Wang, D., Prigent, C., Kilic, L., Fox, S., Harlow, C., Jimenez, C., Aires, F., Grassotti, C. and Karbou, F.: 2017, Surface emissivity at microwaves to millimeter waves over polar regions: Parameterization and evaluation with aircraft experiments, *J. Atmos. Oceanic Technol.* **34**(5), 1039 – 1059.

Appendices

A The Barker 3D cloud-construction algorithm

Barker et al. (2011) presents an algorithm that constructs 3D distributions of cloud from passive satellite imagery and collocated 2D nadir profiles of cloud properties inferred synergistically from lidar, cloud radar and imagery data. Effectively, the construction algorithm widens the active retrieved cross-section (RXS) of cloud properties, by using a match-and-substitute algorithm that fills in off-nadir pixels/columns with cloud property profiles from the swath center. For this purpose radiances from a passive multi spectral imager (MSI) for an off-nadir, recipient, pixel are compared with corresponding values for a range of pixels along the RXS; close matches are identified as potential donors with the closest to the recipient pixel being designated as the proxy to literally stand in for the recipient. This is repeated until all pixels in the required 3D domain are filled.

The Barker construction algorithm is based on the assumption that two closely spaced cloudy pixels have only small differences in a number of spectral top of atmosphere radiances and will have similar cloud property profiles. This is clearly an assumption that is not always valid, and it can not be guaranteed that a good MSI match can be found at all if only pixels closely spaced in space and time are considered. To take the latter into account the Barker methodology is here extended to include a fallback global database when no close matches are found in the local database.

Data

The modified version of the Barker 3D construction algorithm uses the following data:

- A series of radar reflectivity profiles from the cloud profiling radar onboard CloudSat.
- MSI data from MODIS onboard Aqua.
- Topography from the GTOPO30 dataset.

Aqua and CloudSat are both part of NASA's A-train constellation, and hence a large amount of collocated data from these instruments are available.

The MODIS instrument has a viewing swath width of 2,330 km and views the entire surface of the Earth every one to two days. Its detectors measure 36 spectral bands between 0.405 µm and 14.385 µm, and it acquires data at three spatial resolutions; 250 m, 500 m, and 1000 m. The most relevant MODIS channels for the construction algorithm as identified by Barker et al. (2011) are the following:

- 0.62 - 0.67 µm,
- 2.105 - 2.155 µm,
- 8.4 - 8.7 µm,
- and 11.77 - 12.27 µm,

It should be clear that data from a channel operating at visible wavelengths are only relevant for day time observations. The ICI retrieval database should cover both day time and night time observations and no channel at visible wavelengths is used.

CloudSat is an experimental satellite that uses radar to observe clouds and precipitation from space. The Cloud Profiling Radar (CPR) is a 94 GHz nadir-looking radar which measures the power backscattered by clouds as a function of distance from the radar, and the along-track sampling is 2 km (1.4 km × 1.7 km across and along track resolution), with a dynamic range of 70 dB, and 500 m vertical resolution.

Identification of donor columns

A central part of the 3-D construction algorithm is the identification of a suitable donor column to use for an off-nadir pixel. Barker et al. (2011) used a local database consisting of all data points along the subsatellite track and within a distance closer than 200 km from the current position, and selected one of these points as the donor. The choice of 200 km seems to have been chosen a bit arbitrarily, but it was found that the majority of samples will be selected from donor columns within a distance of 30 km, so the choice of the size of the local database is not critical. If we assume that a pixel along the RXS is located at position $(i, 0)$, the intention is to fill all pixels (i, j) with a donor column, and we begin by computing the cost function, defined as

$$F(i,j) = \sum_{k=1}^{K} \left(\frac{r_k(i,j) - r_k(m,0)}{r_k(i,j)} \right)^2, \tag{5}$$

$$: m \in [i - m_1, i + m_2] \tag{6}$$

where the summation is over the K MSI channels, $r_k(i, j)$ is the radiance of channel k at pixel position (i, j) and $r_k(m, 0)$ is the radiance of channel k at pixel position $(m, 0)$ and part of the local candidate database. Barker et al. (2011) describes that just considering the radiance match can result in the selection of an inappropriate donor profile, and to limit this risk it is preferable to also consider the distance between between a potential donor at $(m, 0)$ and the recipient at (i, j). Barker suggests that the profile with closest geographical distance shoud be selected from a set containing the 3 % best MSI matches, and that results in that the selection is made from around 10 samples.

The Barker algorithm is here extended in the following way:

- Disallow donor column with greater surface altitude than that of the recipient by using topography data
- Disallow donor column with poor MSI match (mean differences greater than 2 %)

  1. Use points within 200 km from recipient.
  2. If not any match above, use points within 2000 km from recipient.
  3. If not any match above, use a fallback global database holding data from hundreds of scenes.

An example of constructed 3D radar reflectivity structure is shown in Figures 8, 9, and 10, for data covering the hurricane Nicole in the Gulf of Mexico around 2016-10-12T17:50Z.

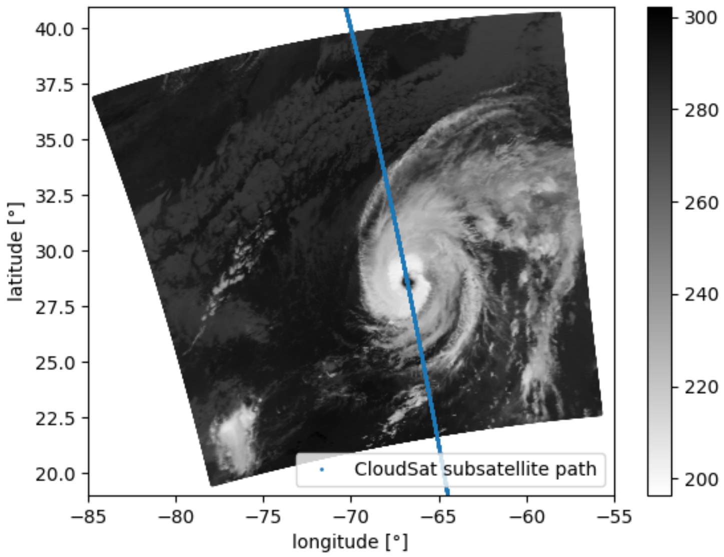

Figure 8: Modis band 29 data covering the hurricane Nicole in the Gulf of Mexico 2016-10-12T17:50Z, where the subsatellite path of CloudSat passes through the eye of the storm.

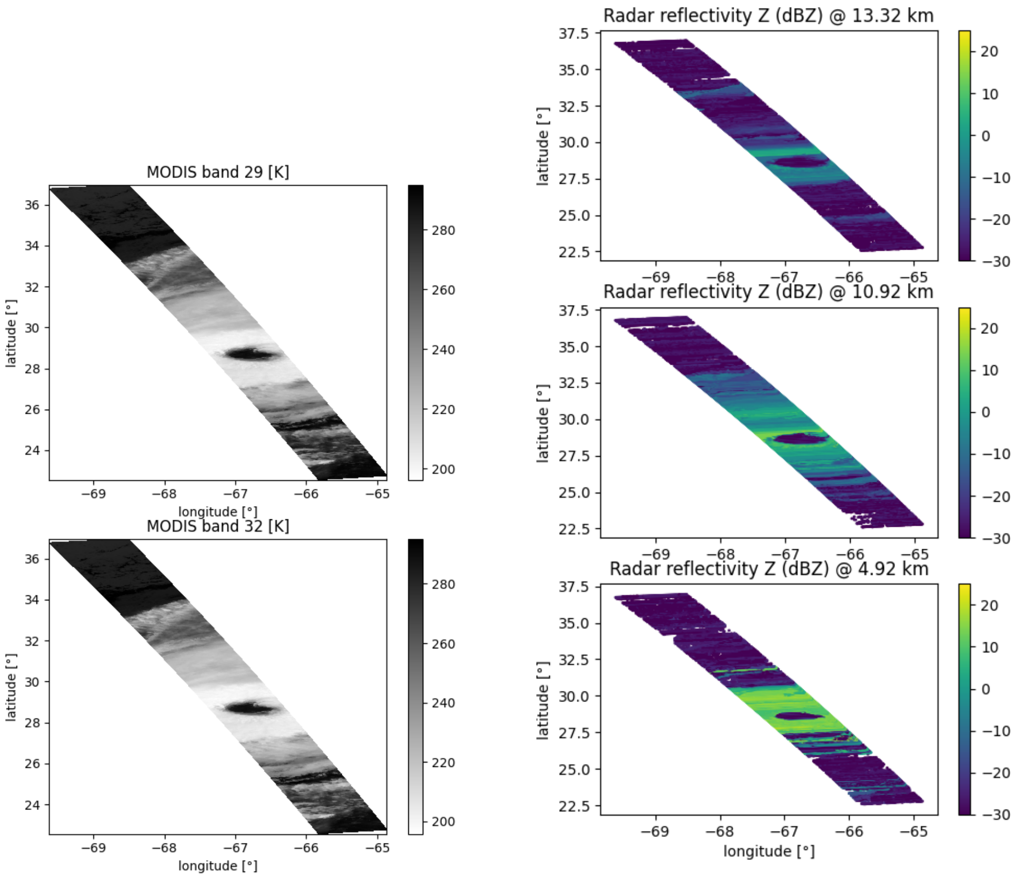

Figure 9: 'Barkerized' data for the scene shown in Figure 8. The left panels show Modis band 29 and 32 data extending ± 50 km in the across-track direction of CloudSat sub-satellite path. The right panels show constructed 2-D cross-sections of radar reflectivity at three different altitudes.

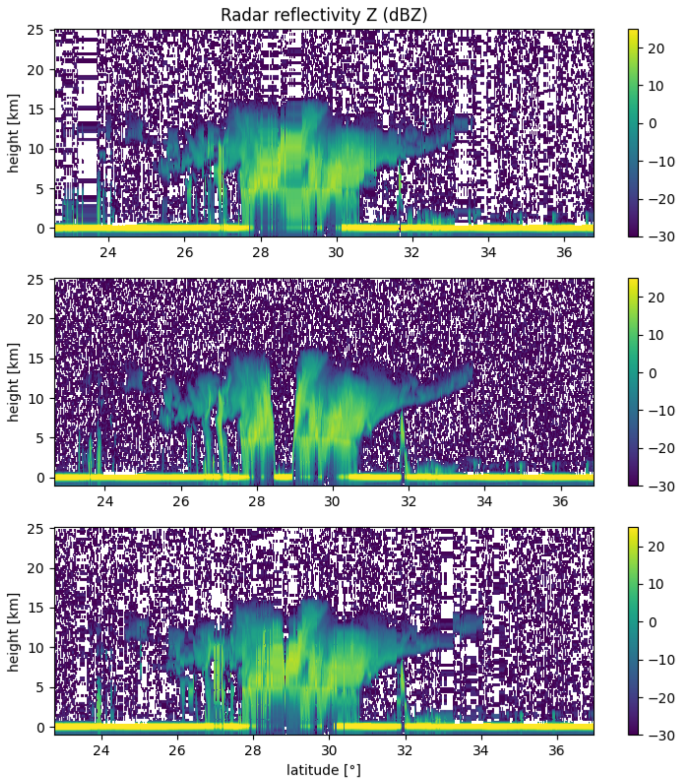

Figure 10: Continued from Figure 9, showing 'Barkerized' data for the scene in Figure 8. The figure shows constructed 2-D cross-sections of radar reflectivity along the westernmost, center, and easternmost path of the scene.

B Particle model occurence fraction

For each set of pencil beam simulations within a 3-D atmospheric scene, a random particle model is chosen. The particle model consists of a particle habit and a particle size distribution. A true random selection requires that each particle model has an equal probability of being selected. However, a random selection does not necessarily lead to a final dataset that statistically represents reality. The aim of this section is to determine an optimal set of probability factors for the selection of particle model. This set of probability factors directly corresponds to the occurrence fraction, i.e. the fraction of simulations run with a given particle model, if the total number of simulations is high enough.

To do so, simulations were run for a large set of atmospheric scenes, for four ICI channels at frequencies 183.31 GHz and 325.15 GHz at vertical polarisation and 243.2 GHz and 664 GHz at horizontal polarisation. The particle model was randomly chosen with equal probabilities for all models. Ice water content (IWC) values across the altitude grid were extracted following the onion peeling algorithm. Figure 11 shows the mean IWC values obtained across the range of altitudes, separated according to the particle model was used within the simulation. Also plotted are mean IWC values across the same altitude grid, extracted from the DARDAR-CLOUD product (Delanoë et al. 2014).

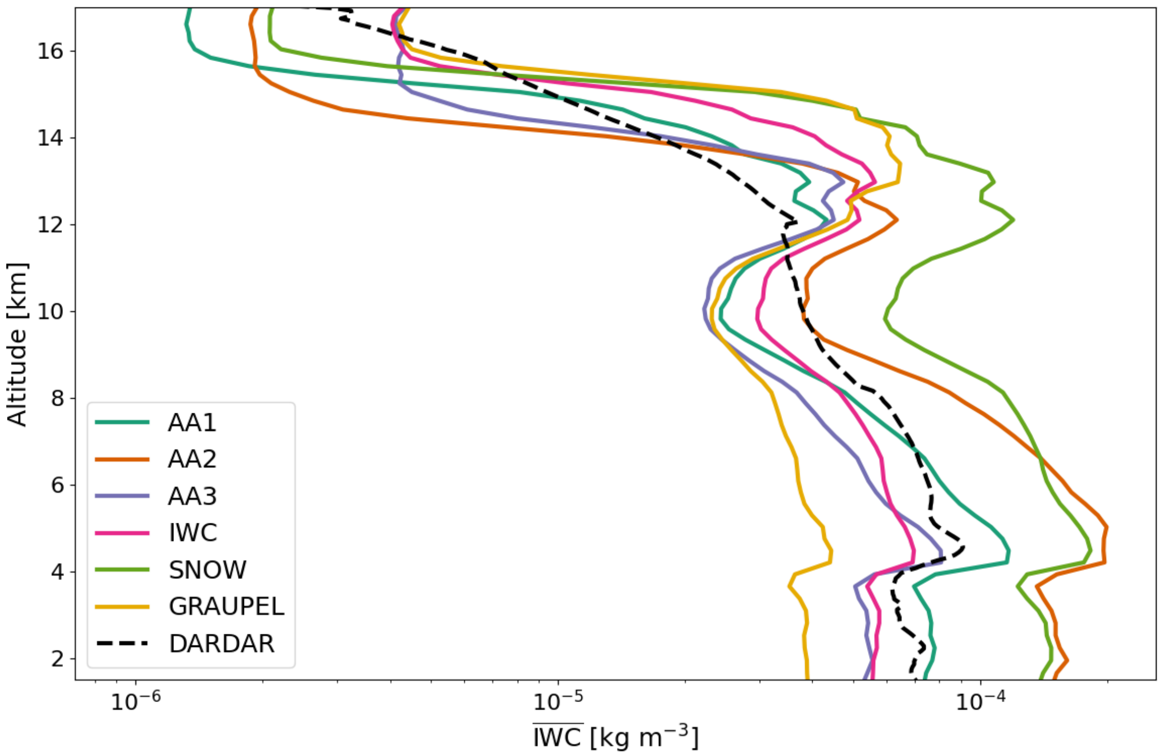

Figure 11: Mean IWC values obtained across the altitude grid. The mean was calculated over $\sim$ 1000 simulations *per particle model*. DARDAR IWC values shown are mean IWC values for global, year-round (2010) data.

Figure 11 shows only mean IWC values. However, the distribution of IWC is not Gaussian. Therefore, to determine a set of probability factors, the distribution of IWC for a given altitude was computed. An example of the IWC distribution at an altitude of

8 km is shown in Figure 12. A combination of the simulated IWC distributions, weighted with an optimal set of probability factors, should match the distribution obtained from the DARDAR product. However, it is noted that none of the simulated distributions achieve the higher IWC values seen in the DARDAR distribution. A perfect match will therefore not be achieved, although this can be partially attributed to the high uncertainties present in the DARDAR product.

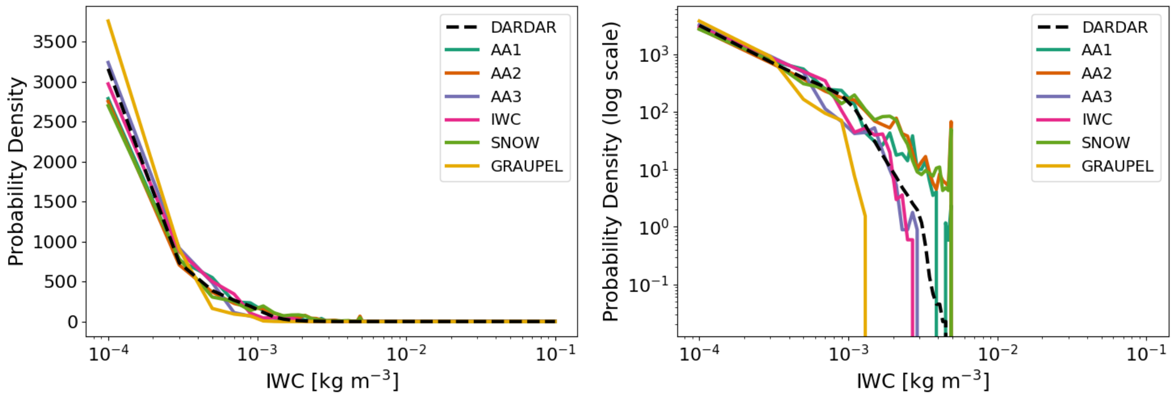

Figure 12: Distribution of IWC at an altitude of 8km. The plot on the right hand side shows the same distributions as the plot on the left, but with the probability density given on a log scale.

To compare the weighted distribution and the DARDAR distribution, the Kullback-Leibler divergence is implemented. The Kullback-Leibler divergence is defined as

$$D_{\mathrm{KL}}(Q||Q^{(0)}) = \sum_{x\in\mathcal{X}} P(x)\log\left(\frac{P(x)}{Q(x)}\right), \tag{7}$$

where P and Q are discrete probability distributions defined over $\mathcal{X}$. For our purposes, Q represents the distribution of IWC given by DARDAR data. P represents a distribution of IWC weighted by a set of probability factors. In a simplified sense, $D_{\mathrm{KL}}(P||Q)$ can be seen as a quantification of how much the distribution P differs from Q. Therefore, a weighted distribution generated with an optimal set of probability factors will minimise $D_{\mathrm{KL}}(P||Q)$.

In order to ensure that all six particle models are represented in the retrieval database, lower limits were placed on the probabilities to avoid any probability falling too low as a result of the numerical minimisation. The limits were: $p_{\mathrm{AA1}} > 0.3$, $p_{\mathrm{other}} > 0.1$. The choice of $p_{\mathrm{AA1}} > 0.3$ was motivated by the fact that this particle model has been tested extensively and has been shown to perform particularly well (Eriksson et al. 2018, Ekelund et al. 2020).

A set of $N \approx 3 * 10^4$ probability factor samples were generated according to these boundary conditions and the resulting weighted distributions are shown in Figure 13. Also shown is the distribution obtained using equal probability factors for all particle models, and the distribution receiving the lowest $D_{\mathrm{KL}}(P||Q)$. The set of probability factors is defined as as

$$p = \{p_{\mathrm{AA1}}, p_{\mathrm{AA2}}, p_{\mathrm{AA3}}, p_{\mathrm{IWC}}, p_{\mathrm{SWC}}, p_{\mathrm{GWC}}\}, \tag{8}$$

where we note that the set of probabilities, or occurrence fractions, p is not the same as the probability distribution P over IWC. The optimal set of probability factors was found to be

$$p = \{0.3, 0.1, 0.13, 0.2, 0.1, 0.17\}. \tag{9}$$

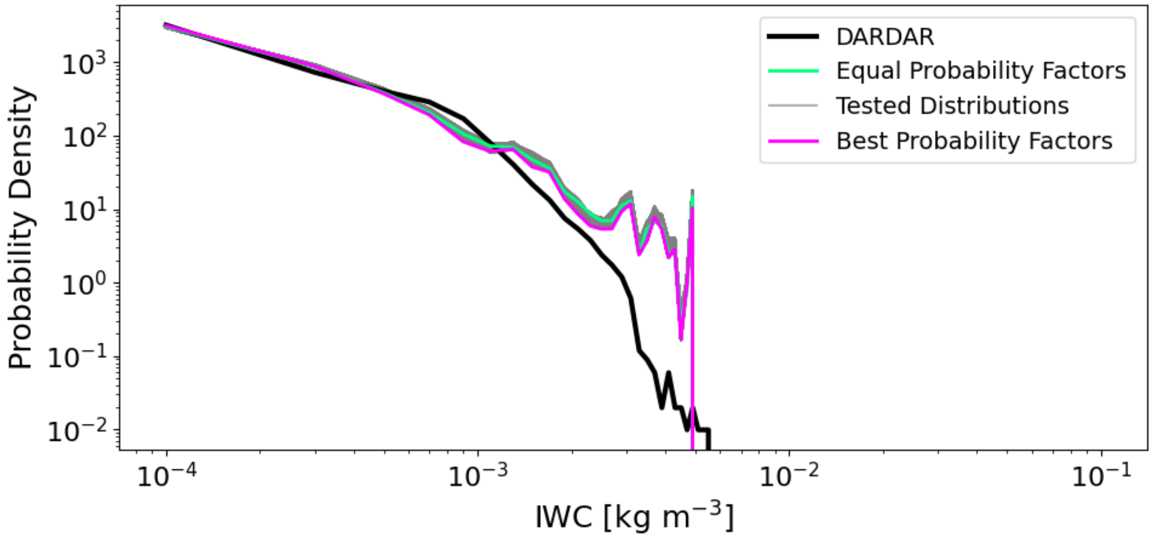

Figure 13: Distribution of IWC at an altitude of 8 km. $\sim 10^5$ distributions were tested. The distribution shown in green was obtained using all $p_i = 1/6$, which was the default setting before improvements were implemented. The 'best' distribution, shown in pink, fails to perfectly match the DARDAR distribution at high IWC values due to lower limits currently set on the probability factors. However, the DARDAR distribution also contains relatively high levels of uncertainty and must not be taken as the truth. It must also be noted that this discrepancy occurs at low probability density values.

C Generation of random sea ice distributions

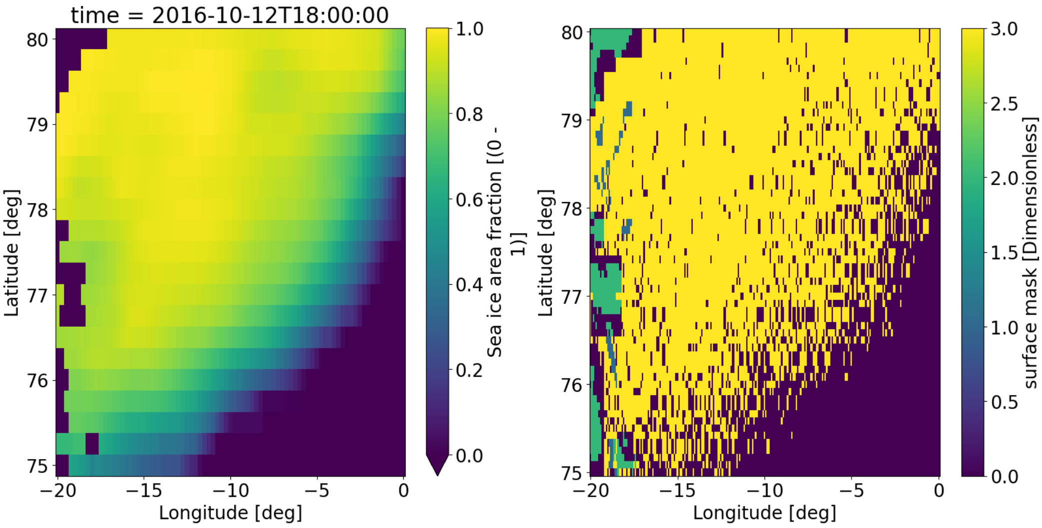

Figure 14: Sea ice concentrations up-scaled to 5 minute (∼8 km) resolution.

The sea ice distribution is taken from ERA5 reanalysis surface variable known as 'sea ice cover'. This parameter describes the fraction of a grid box which is covered by sea ice at a resolution of 25 km. The land water mask used in the database scenes is around 8 km resolution, which necessitates the mapping of sea ice cover from a lower to a higher resolution. We refer this mapping as up-scaling of sea-ice cover.

For mapping the sea ice distribution, a simple three-step algorithm is implemented. Firstly, the 25 km ERA5 sea ice cover is interpolated to a 10 km grid. This intermediate resolution can effectively represent sea ice concentration in steps of 0.17. In the next step, a random selection of the sea ice and open water is made according to the sea ice fraction. And finally, the sea ice and open water classification at 10 km resolution is interpolated to the final 8 km resolution in a nearest fashion. An example scene is shown in Figure 14.

Surface emissivities for snow and sea ice

In radiative transfer, it is crucial to model the surface contribution correctly. In preparation towards ICI, an updated version of The Tool to Estimate Land Surface Emissivity from Microwave to Submillimeter Waves (TELSEM), known as TELSEM2, was developed to also include parameterization of emissivities over continental snow and sea ice up to 700 GHz. However, for both surface types, the emissivities were fully parametrised only up to 85 GHz. For instance, emissivities for snow-covered land are assumed to be constant above 85 GHz. For sea ice, the emissivities are parametrised (up to 183 GHz) only for new ice or first year ice. For multi-year ice, constant emissivities above 85 GHz were assumed. However, in another study, Harlow and Essery (2012) studied the variability of snow emissivity over different snow-packs. They observed that the snow emissivities increase monotonically with increasing frequency for stratified snow. For fresh snow, the emissivity estimates follow a decreasing trend. Additionally, Wang et al. (2017) found that the comparisons of TELSEM2 with International SubMillimeter Airborne Radiometer (ISMAR, Fox et al. 2017) measurements at 325 GHz were inconclusive.

Previously, motivated by the above reason, we had developed a probabilistic snow-emissivity model that provided polarised emissivities for snow-covered surface types (both over land and sea ice) at frequencies between 150 and 190 GHz. This model gave a random estimate of the snow emissivity from a standard normal distribution depicting the valid range of emissivity values. The valid range was decided through emissivity values reported in literature (Hewison et al. 2002, Harlow 2009, Harlow and Essery 2012) and the model was fine-tuned by comparing observed and forward modelled radiances for snow-covered regions as observed by Global Precipitation Measurement (GPM) Microwave Imager (GMI). The basic idea is to find a set of emissivity estimates that simultaneously give radiance closest to the measurements. With this model, it was possible to provide a fair representation of polarisation differences observed over snow-covered surfaces.

However, an extension of the mentioned empirical model to frequencies up to 325 GHz is not straightforward due to non-availability of satellite based observations at frequencies above 183 GHz. In this study, we extend the existing emissivity model to cover frequencies from 10 to 1000 GHz using existing studies in literature. Two main studies which have been used as reference and the methodology are described in the sections below.

Data

Two main studies available in literature are used as a reference for the emissivity model: The study by Munchak et al. (2020) (referred as Munchak20) and the Master's thesis from Nils Risse (Risse 2021) (referred as Risse21). A brief summary of both studies is given below.

Munchak20

This study uses optimal estimation to retrieve the emissivity vector covering the frequencies between 10 to 166 GHz. They apply the algorithm to five years of global precipitation measurement (GPM) microwave imager (GMI) data over land, snow-covered land and sea ice regions. This retrieval dataset is used to estimate the mean and standard deviation of snow-cover and sea ice emissivities within a frequency range of 10 to 166 GHz.

Risse21

Only Risse (2021) have made detailed retrievals of sea ice emissivities in the sub-mm range by using measurements from an aircraft campaign. They retrieve emissivity for sea ice regions and snow-covered sea ice regions for 89, 183, 243, and 340 GHz by using passive microwave observations from Microwave Radar/radiometer for Arctic Clouds (MiRAC) for nadir viewing angle. They calculate the emissivities using radiative transfer model PAMTRA, by assuming both specular and lambertian reflection. The results from this study are only used to as a guidance to estimate the emissivities above 183 GHz.

Methodology

The range of the previous empirical model is extended to include all frequencies between 10 and 1000 GHz. The model generates emissivities at both V and H polarisations through a multivariate normal distribution. For an n dimensional random vector $\vec{Y}$, the multivariate normal distribution is defined as:

$$\vec{Y} = \mathcal{N}(\mu, \mathcal{K}_{\mathcal{YY}}), \tag{10}$$

where μ is the n-dimensional mean vector and K_{YY} is the covariance matrix of dimension $n \times n$. If σ_X is the standard deviation of $\vec{X}$, then

$$K_{YY} = \sigma_X * \sigma_X * \mathcal{C}_{YY}, \tag{11}$$

where $\mathcal{C}_{YY}$ is the corresponding correlation matrix. For generating a multivariate normal distribution between V and H frequencies, the correlation matrix $\mathcal{C}_{YY}$ is defined as:

$$\begin{vmatrix} \mathcal{C}_{\mathrm{V}} & \mathcal{C}_{\mathrm{VH}} \\ \mathcal{C}_{\mathrm{VH}} & \mathcal{C}_{\mathrm{H}} \end{vmatrix} \tag{12}$$

Here, $\mathcal{C}_{\mathrm{V}}$ and $\mathcal{C}_{\mathrm{H}}$ are the correlation matrices for V and H polarised channels, while $\mathcal{C}_{\mathrm{VH}}$ is the correlation matrix between V and H channels. $\mathcal{C}_{\mathrm{V}}$ is set up using the defined correlation lengths and an exponential correlation function. An identical correlation matrix is chosen for H polarised channels. The correlation matrix $\mathcal{C}_{\mathrm{VH}}$ is defined as $c * \mathcal{C}_{\mathrm{V}}$, where c is a constant which defines the correlation between V and H polarised channels. For this study, we have selected $c = 0.9$.

Results

Figure 15 shows the mean and standard deviation of the randomly generated emissivity estimates. The corresponding correlation matrix is also shown. The snow-cover emissivities at low microwave frequencies are higher than as compared to higher frequencies. The emissivities at higher frequencies are lower due to strong scattering in the snow, while at lower frequencies, the microwave penetration depths and insulating properties of the snow-packs contribute to higher emissivities. The trend in emissivities above 183 GHz follow the results reported by Risse (2021). They found that above 183 GHz the emissivities remain constant under specular reflection, while under lambertian reflection, a slight decrease with frequency is observed.

Similar statistics for sea ice emissivities are shown in Figure 16. For sea ice, the emissivities show an increasing trend over low microwave frequencies and up to 166 GHz. Risse (2021) have shown that the gradient between 89 and 183 GHz has a high variability

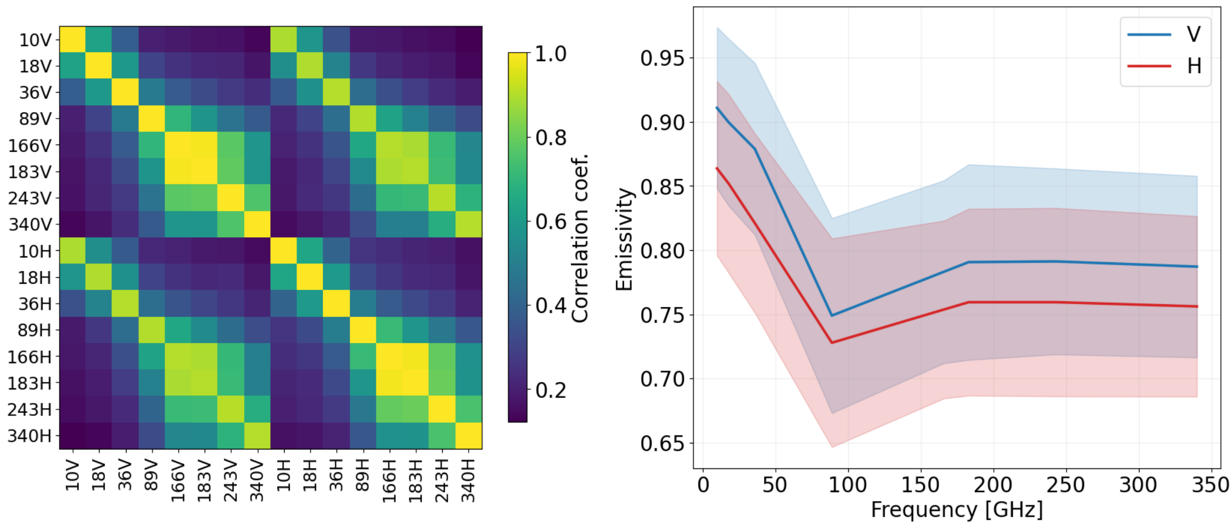

Figure 15: (left) The correlation matrix for snow-covered surface type, and (right) the mean emissivities and standard deviation.

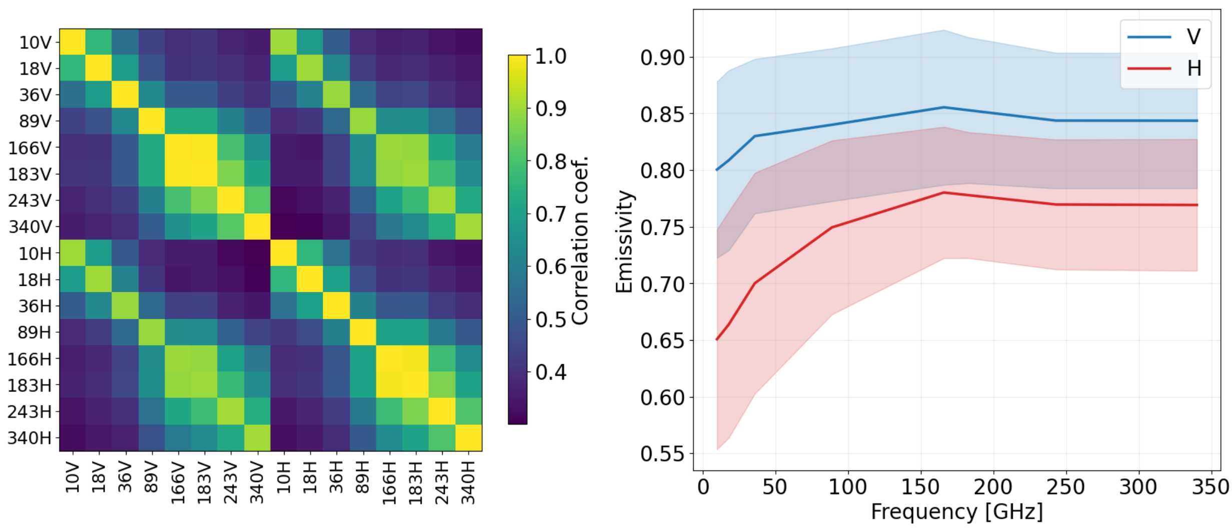

Figure 16: (left) The correlation matrix for sea ice surface type, and (right) the mean emissivities and standard deviation.

and is mostly positive. Above 183 GHz, the emissivity gradient is comparatively lower and slightly negative between 183 to 243 GHz. From 243 to 340 GHz, under lambertian reflection, slight negative gradients were reported and the vice-versa for specular reflection. Due to the contrasting behaviour for the two assumptions, we have assumed constant emissivities beyond 243 GHz.

D Spectral response function

The ICI radiometer consists of seven double sideband front ends, operating with local oscillator (LO) frequencies of 183.31, 243.20, 325.15, 448.0, and 664.0 GHz, and ICI will have 13 channels as described in Table 2. The spectral response function associated to each channel is shown in Figure 17. Simulated brightness temperature as function of frequency, covering the frequency range of the upper and lower sideband of each channel, is shown in Figure 18 for one scene of the ERESMAA dataset. The fine structure brightness temperature variation is due to ozone absorption/emission, and the strongest lines are found within the ICI-07V and both ICI-11 channels.

The aim of this section is to find out how important it is to include the spectral response function to simulate band average brightness temperature. The reference channel averaged spectral radiance, for channel i, is calculated as:

$$R_i = \frac{\int_{\gamma_0}^{\gamma_1} \frac{1}{\gamma^2} R(\gamma) H_i(\gamma) d\gamma}{\int_{\gamma_0}^{\gamma_1} \frac{1}{\gamma^2} H_i(\gamma) d\gamma}, \tag{13}$$

where γ is the wavenumber, $R(\gamma)$ is the spectral radiance, and $H_i(\gamma)$ is the channel i spectral response function displayed in Figure 17, and frequency grid used for the simulation includes all frequencies of the spectral response function data associated to the ICI channels (resolution of 10 MHz).

The brightness temperature is then obtained from

$$T_{bi} = \frac{B^{-1}(\gamma_{centre}, R_i) - b_i}{a_i} \tag{14}$$

where B^{-1} is the inverse planck function, γ_{centre} is the nominal centre wavenumber, and a_i and b_i are channel specific band corrections coefficients.

The channel averaged brightness temperature for a constant spectral response function, for channel i, is calculated as

$$T_{bic} = 0.5 \cdot \left(\frac{\int_{\gamma_0}^{\gamma_1} T_b(\gamma) d\gamma}{\int_{\gamma_0}^{\gamma_1} d\gamma} + \frac{\int_{\gamma_2}^{\gamma_3} T_b(\gamma) d\gamma}{\int_{\gamma_2}^{\gamma_3} d\gamma} \right), \tag{15}$$

where the integration limits are the outer limits of the lower and upper sideband for channel i according to Table 2. Histograms of the difference

$$\Delta T_{bi} = T_{bic} - T_{bi} \tag{16}$$

are displayed in Figure 19 for each ICI channel and for a simulation based on 1000 scenes from the ERESMAA dataset. The mean and standard deviation of this difference is displayed in the right column of Table 2. The greatest difference is found for ICI-09V (-0.446 ± 0.124 K). No particular strong ozone line is located within the frequency range assocated to ICI-09V, but the reason for the difference is probably explained by the fact that the slope of the spectral response function is greater for this channel than for all the other channels. The response is increasing for frequencies towards the outer part of the upper and lower sideband as measured from the LO frequency. Thus, applying a constant spectral response function will introduce a bias. Differences for other channels are significantly smaller than for ICI-09V. The second greatest difference is found for ICI-06V (-0.128 ± 0.056 K), that also has some features within the spectral response function that are greater than for other channels.

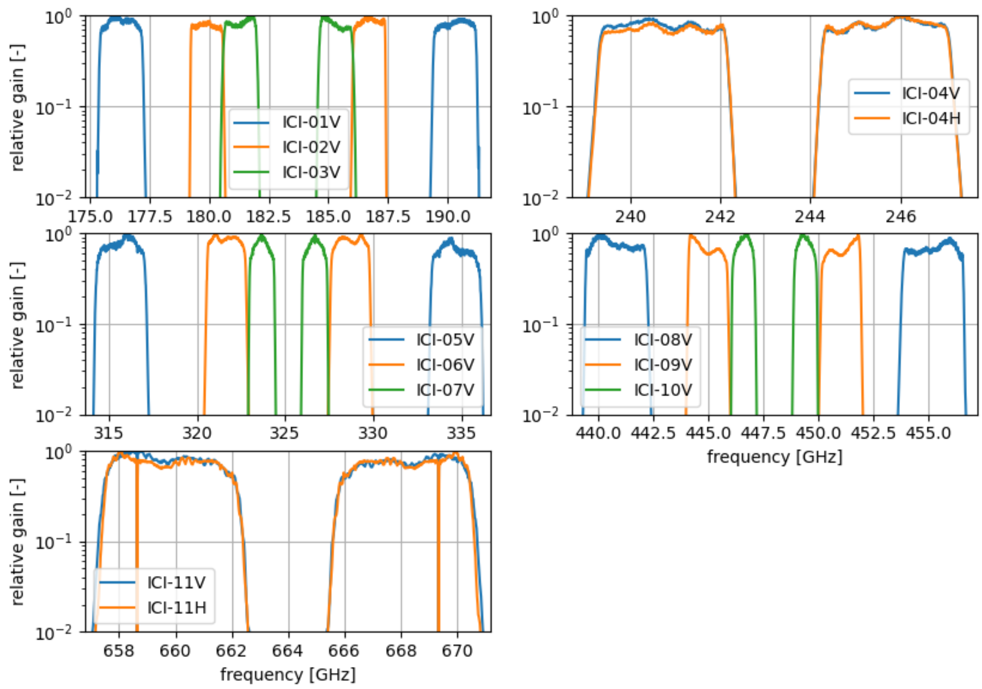

Figure 17: Spectral response function of the 13 ICI channels.

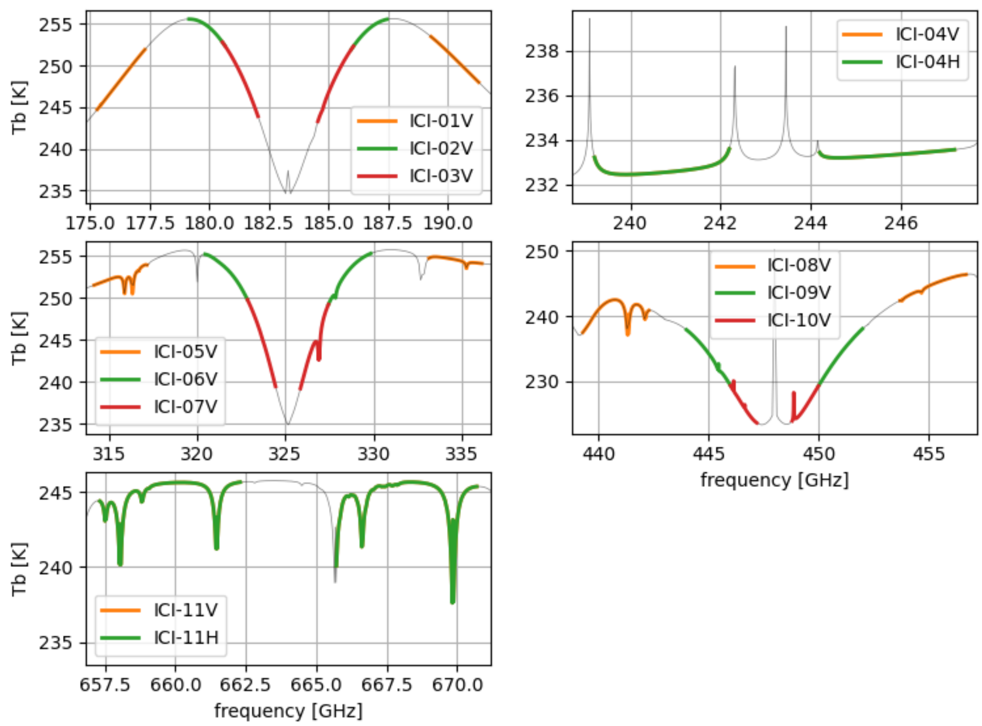

Figure 18: Simulated brightness temperature as function of frequency for one scene of the ERESMAA dataset.

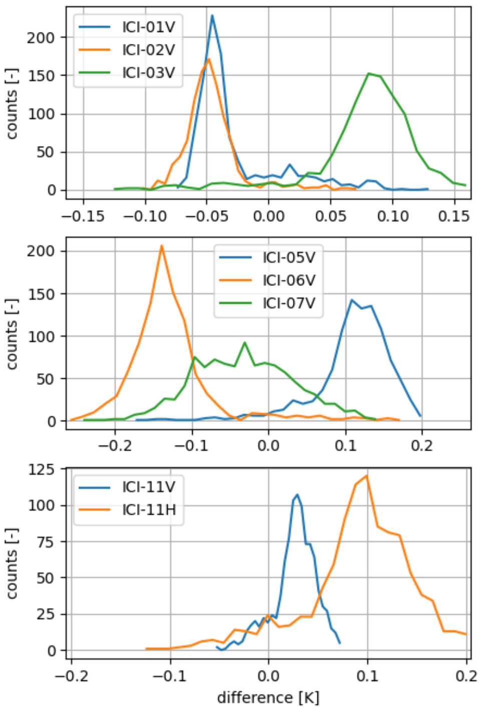

Figure 19: Histogram of brightness temperature differences due to neglecting the spectral response function for 1000 scenes of the ERESMAA dataset.

Table 2: ICI channel specification.

Channel	**Frequencies [GHz]**	**Bandwidth [GHz]**	**NEΔT [K]**	**Elevation offset [°]**	**Response impact [K]**
ICI-01V	183.31 ± 7.0	2.0	0.8	−0.780	−0.028 ± 0.035
ICI-02V	183.31 ± 3.4	1.5	0.8	−0.780	−0.046 ± 0.022
ICI-03V	183.31 ± 2.0	1.5	0.8	−0.780	0.079 ± 0.041
ICI-04V	243.2 ± 2.5	3.0	0.7	0.711	−0.020 ± 0.011
ICI-04H	243.2 ± 2.5	3.0	0.7	0.731	−0.026 ± 0.013
ICI-05V	325.15 ± 9.5	3.0	1.2	−0.822	0.109 ± 0.051
ICI-06V	325.15 ± 3.5	2.4	1.3	−0.822	−0.128 ± 0.056
ICI-07V	325.15 ± 1.5	1.6	1.5	−0.822	−0.032 ± 0.065
ICI-08V	448.0 ± 7.2	3.0	1.4	−0.822	−0.032 ± 0.024
ICI-09V	448.0 ± 3.0	2.0	1.6	−0.822	−0.446±0.124
ICI-10V	448.0 ± 1.4	1.2	2.0	−0.822	0.031±0.113
ICI-11V	664.0 ± 4.2	5.0	1.6	0.752	0.027±0.022
ICI-11H	664.0 ± 4.2	5.0	1.6	0.875	0.091 ± 0.054

Frequency grid setup for the ICI database simulation

The simulation in the previous section was performed using a fine resolution frequency grid. We here test the use of a much coarser frequency grid for the actual radiative transfer calculation, but then interpolate the obtained spectral radiance onto a high resolution (10 MHz) frequency grid. The interpolation is done individually for the upper and lower side band. Then we take into account the spectral response function, and obtain channel averaged spectral radiance as

$$R_i = \frac{\int_{\gamma_0}^{\gamma_1} \frac{1}{\gamma^2} R(\gamma) H_i(\gamma) d\gamma + \int_{\gamma_1}^{\gamma_2} \frac{1}{\gamma^2} R(\gamma) H_i(\gamma) d\gamma}{\int_{\gamma_0}^{\gamma_1} \frac{1}{\gamma^2} H_i(\gamma) d\gamma + \int_{\gamma_1}^{\gamma_2} \frac{1}{\gamma^2} H_i(\gamma) d\gamma}, \tag{17}$$

where the integration limits cover the lower and upper sideband, and then we apply Eq. 14 to obtain the brightness temperature. The result of this exercise is summarised in Table 3. Table 3 describes the number of frequencies per sideband needed to be included in the radiative transfer calculation to obtain errors that have a magnitude smaller than about 5 % of the NEΔT value of the channel as described in Table 2.

Table 3 includes results for using both linear and quadratic interpolation. The most demanding channels are the ICI-7 and the two ICI-11 channels, that are the channels that are most contaminated by ozone. Better results are obtained using quadratic interpolation for most channels.

Table 4 describes the frequency ranges of the lower and upper sideband for each channel according to specification and to the actual spectral response function data, here defined as the smallest and greatest frequency below (or above) the LO frequency where the relative spectral response function value is greather than 0.5. Using this definition gives that the width of the upper and lower sideband of the actual data is about 150 to 500 MHz smaller than that of the specification (500 MHz for the 664 GHz channel with 5 GHz bandwidth).

Table 3: Performance using a reduced resolution frequency grid.

Channel	**Nr. of frequencies per sideband**	**Linear interpolation, Difference [K]**	**Quadratic interpolation, Difference [K]**
ICI-01V	3	-0.030 ± 0.033	0.005 ± 0.007
ICI-02V	3	-0.052 ± 0.033	0.013 ± 0.011
ICI-03V	3	-0.154 ± 0.066	-0.011 ± 0.011
ICI-04V	10	-0.028 ± 0.019	-0.006 ± 0.003
ICI-04H	10	-0.035 ± 0.017	-0.012 ± 0.010
ICI-05V	5	0.004 ± 0.035	-0.040 ± 0.031
ICI-06V	4	-0.031 ± 0.023	0.033 ± 0.017
ICI-07V	15	-0.023 ± 0.007	-0.020 ± 0.007
ICI-08V	10	0.023 ± 0.023	0.036 ± 0.025
ICI-09V	10	0.012 ± 0.031	0.018 ± 0.032
ICI-10V	10	-0.023 ± 0.016	-0.017 ± 0.014
ICI-11V	25	-0.052 ± 0.082	-0.046 ± 0.078
ICI-11H	25	-0.047 ± 0.083	-0.043 ± 0.081

Table 4: Limits of upper and lower sideband, according to specification and actual data (relative gain > 0.5 is used for actual data).

Channel	**Lower sideband (spec / data) [GHz]**	**Lower width [GHz]**	**Upper sideband (spec / data) [GHz]**	**Upper width [GHz]**
ICI-01V	175.31 − 177.31 175.4306 - 177.1906	2 1.76	189.31 - 191.31 189.4306 - 191.1906	2 1.76
ICI-02V	179.16 - 180.66 179.2406 - 180.5706	1.5 1.33	185.96 - 187.46 186.0506 - 187.380666	1.5 1.33
ICI-03V	180.56 - 182.06 180.6206 - 181.9906	1.5 1.37	184.56 - 186.06 184.6406 - 186.0006	1.5 1.36
ICI-04V	239.2 - 242.2 239.3234 - 242.1334	3.0 2.81	244.2 - 247.2 244.2834 - 247.0834	3.0 2.8
ICI-04H	239.2 - 242.2 239.3534 - 242.1134	3.0 2.76	244.2 - 247.2 244.2934 - 247.0734	3.0 2.78
ICI-05V	314.15 - 317.15 314.3805 - 316.9905	3.0 2.61	333.15 - 336.15 333.3805 - 335.8205	3.0 2.44
ICI-06V	320.45 - 322.85 320.5805 - 322.7305	2.4 2.15	327.45 - 329.85 327.6205 - 329.7605	2.4 2.14
ICI-07V	322.85 - 324.45 323.0205 - 324.3305	1.6 1.31	325.85 - 327.45 326.0105 - 327.3205	1.6 1.31
ICI-08V	439.3 - 442.3 439.4522 - 442.1522	3.0 2.70	453.7 - 456.7 453.8622 - 456.5422	3.0 2.68
ICI-09V	444.0 - 446.0 444.1122 - 445.7822	2.0 1.67	450.0 - 452.0 450.2622 - 451.8922	2.0 1.63
ICI-10V	446.0 - 447.2 446.1322 - 447.0722	1.2 0.94	448.8 - 450.0 448.9322 - 449.8622	1.2 0.94
ICI-11V	657.3 - 662.3 657.5381 - 662.1181	5.0 4.58	665.7 - 670.7 665.9181 - 670.4181	5.0 4.50
ICI-11H	657.3 - 662.3 657.5581 - 662.0181	5.0 4.46	665.7 - 670.7 665.9781 - 670.4381	5.0 4.46

E Track Settings

Prior to running pencil beam simulations and performing antenna smoothing, track settings must be configured. The settings include a choice of track positions on which to perform pencil beam simulations, where the track is relative to the centre track. There must also be a choice of step size between each simulation along a track. In this section, we refer to a chosen track setting configuration as a *pattern*, and each pencil beam simulation as a *point* within the pattern.

The positioning of points can have a significant effect on the success of the interpolation of brightness temperatures and subsequent integration to obtain antenna temperature. The goal of the tests performed in this section is to evaluate both the *amount* of pencil beam simulations and the *configuration* of such simulations that are needed to simulate an ICI observation with an acceptable accuracy. The final choice of pattern must be chosen with care since there will be a trade-off between accuracy and computational cost.

Three categories of patterns were tested:

- Uniform - The centre track is always chosen. An even number of tracks on either side of the centre track are also chosen. The step size is constant for all tracks. Number of tracks and chosen step size were varied between patterns.
- Non-uniform - Patterns were chosen such that the centre track contains the highest density of simulations, i.e, the smallest step size. The further a track is from the centre, the larger the step size is chosen to be. Number of tracks and chosen step size were varied methodically between patterns.
- Pseudo-random - Three of the non-uniform patterns were selected, and the choice of step size in the pattern was randomised.

In total, 29 different patterns were tested. In order to compare patterns, antenna temperatures T_a were calculated for a set of chosen sensor positions within a scene. In the absence of observational data for T_a, simulations were also run for a full scene, where a full scene is defined as pencil beam simulations on every track, and with a step size of $0.02\,^\circ$. The resulting temperatures act as true observational data and are denoted by $T_{a,\text{true}}$. Then, prior to applying the antenna smoothing, the full set of simulations was filtered according to each of the patterns. An interpolation was performed between datapoints, providing predicted T_a values with which to compare to $T_{a,\text{true}}$. The reader can refer to Figure 7 to see a comparison of true and interpolated data.

The difference between predicted and true values is denoted by $T_{a,\text{pred}} - T_{a,\text{true}} = \Delta T_a$. The mean $\overline{\Delta T_a}$, standard deviation $\sigma_{\Delta T_a}$, and 99th percentile $P_{99}(\Delta T_a)$ were calculated for a dataset consisting of 100 atmospheric scenes, with three sensor positions per scene. Simulations were run for four ICI channels at frequencies 183.31 GHz and 325.15 GHz (vertical polarisation) and 243.2 GHz and 664 GHz (horizontal polarisation). Results are shown in Figure 20, in which the statistics are plotted as a function of the number of points in a pattern, given as a percentage of the number of points present in the full scene.

The patterns resulting in the lowest standard deviations are, as expected, those with the highest number of points. Due to the presence of a close to Gaussian antenna

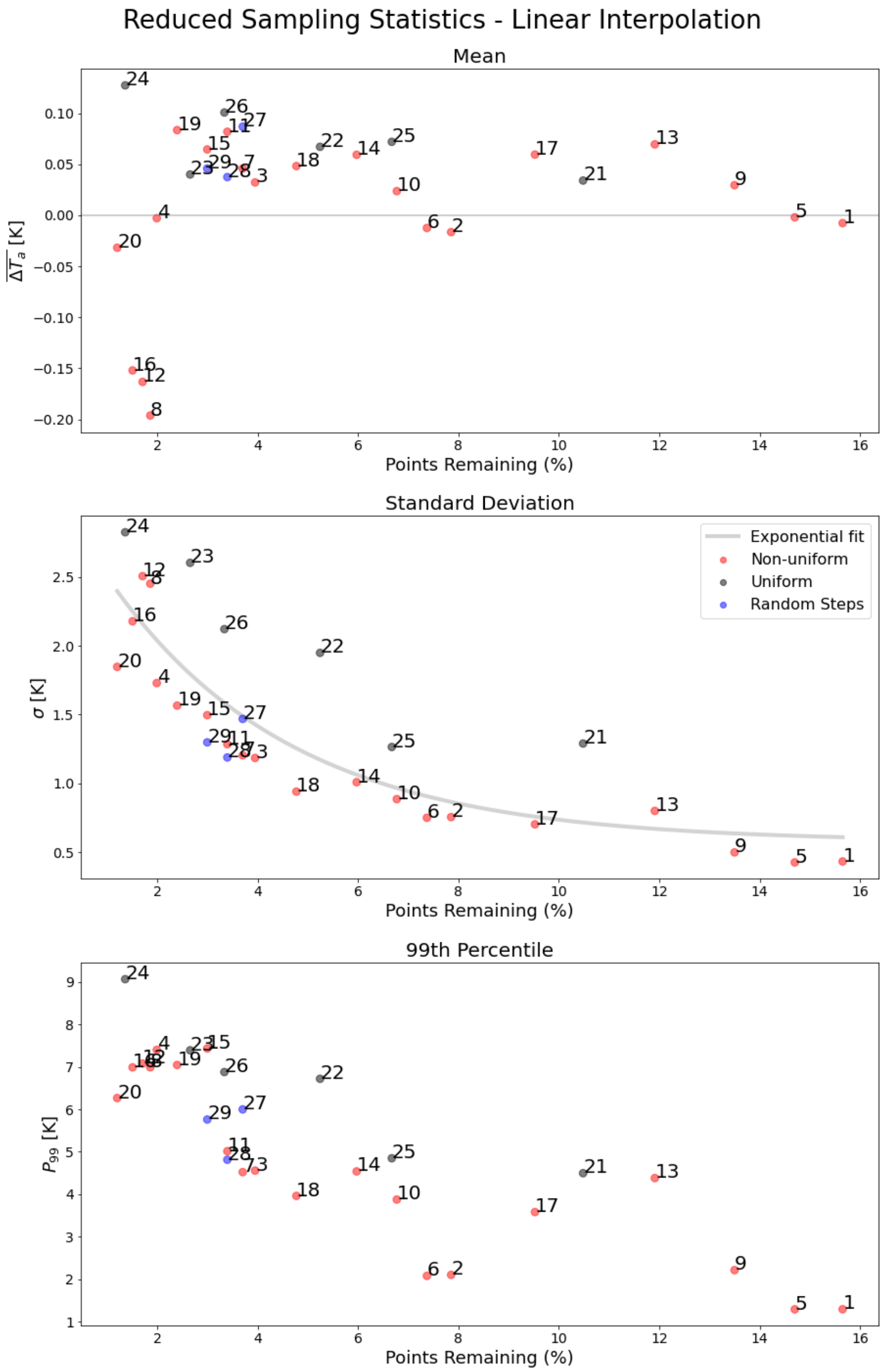

Figure 20: The mean, standard deviation, and 99th percentile of ΔT_a, defined as the difference between values of T_a computed using reduced track settings and full track settings. Each numbered point refers to a pattern, defined as a track setting configuration. Statistics were computed over 100 atmospheric scenes, 3 sensor positions, and 4 channels, producing 1200 samples in total *for each pattern*. The standard deviation appears to exponentially increase with a decrease in number of points in a pattern.

pattern, it was expected that the non-uniform patterns would perform better than uniform patterns. The results quantitatively confirm our expectations. There is no improvement seen in the case of the random patterns, but this is not surprising when the statistics were computed across a large set of samples.

It was found that the standard deviation decreases exponentially with the number of points. However, it must be stressed that the database should be sufficiently large such that it contains all relevant atmospheric states and all geographical regions and seasons. In turn, this will allow for a successful retrieval algorithm. This implies that achieving satisfactory computational times is an important consideration, so that enough simulations can be performed. Patterns 1 and 5 can be seen to give the lowest standard deviation and 99th percentile. However, the use of Pattern 2 or 6 would half the number of pencil beam simulations, and therefore thought to achieve a preferable balance between computational cost and accuracy. Patterns 1 and 2 are shown in Figure 21. There are other patterns that demonstrate similarly low standard deviations at a similar total number of points. However, patterns 2 and 6 perform notably well in the case of the 99th percentile, and as such may decrease the risk of outliers.

In order to validate the choice further, further simulations were run with the aim of comparing patterns 1, 2, 5, and 6 to the pattern implemented in the preliminary database, i.e. only points along the centre track. Results are shown in Figure 22, where the statistics are taken across $N = 100$ scenes, 3 sensor positions, and the same four ICI channels. Both patterns 2 and 6 perform better than just the central track in all cases. The differences between pattern 2 and pattern 6 are minimal. The same results were also seen upon investigation of separate channels.

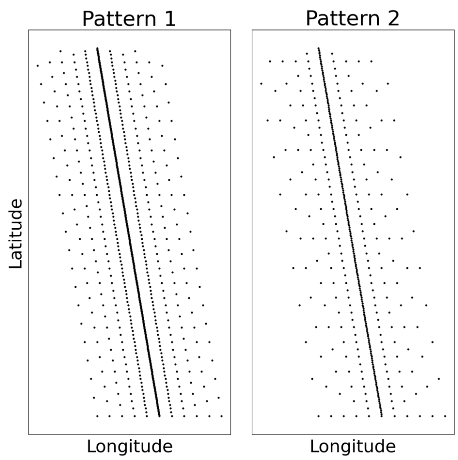

Figure 21: Patterns 1 and 2. Both patterns consist of the same tracks, but pattern 2 has double the amount of spacing between points as pattern 1. Patterns 5 and 6 differ from patterns 1 and 2, respectively, by the removal of the two outermost tracks.

Finally, ΔT_a was plotted as a function of $T_{a,\text{true}}$ in order to visualise how errors are distributed according to temperature. The result can be found in Figure 23. The higher the temperature, the more successful all patterns appear to be at predicting the true value. At low temperatures, the patterns appear to produce datapoints further from the true value than the 'centre' pattern does, explaining the higher 99th percentiles calculated in Figure 22. However, all patterns appear to perform better than the 'centre' pattern, i.e.

the case used in the preliminary database, at high temperatures, thus explaining the lower standard deviation seen in Figure 22. Patterns 2 and 6 differ slightly from one another at low temperature values, but appear to perform equally well at higher temperatures. However, pattern 2 displays a small improvement in standard deviation in comparison to pattern 6, and visually produces less extreme outliers in Figure 23. Therefore, pattern 2 is chosen to represent the track settings in the final database.

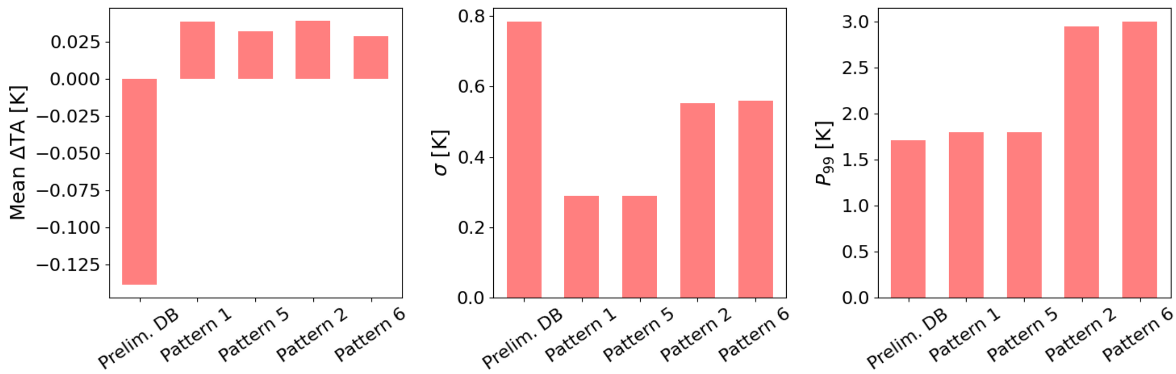

Figure 22: Comparison of mean $\overline{\Delta T_a}$, standard deviation σ, and 99th percentile P_{99}, of ΔT_a obtained through the use of patterns 1, 2, 5, and 6, and a pattern in which only the central track is used (denoted 'centre'). Statistics were computed over 100 atmospheric scenes, 3 sensor positions, and 4 channels, producing 1200 samples in total *for each pattern.*

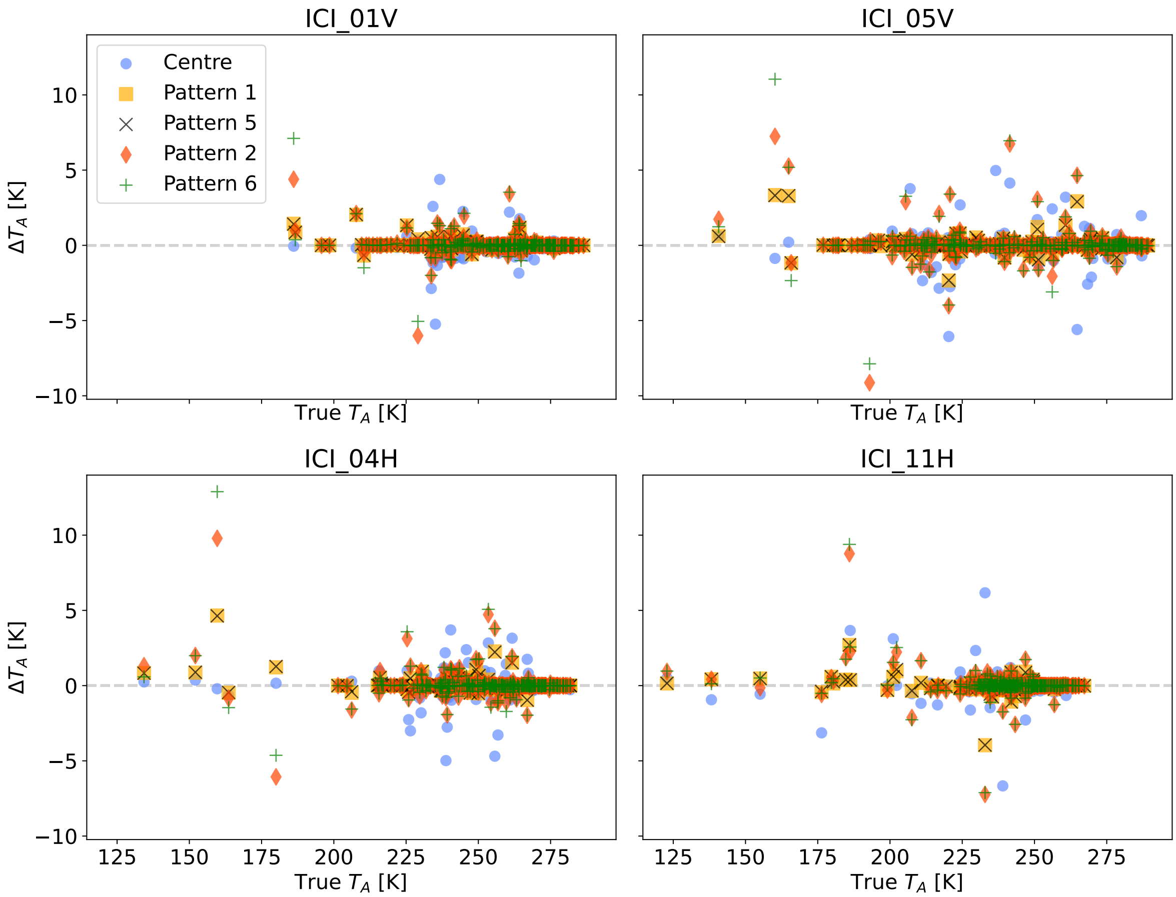

Figure 23: ΔT_a as a function of $T_{a,\mathrm{true}}$.

Development of a cloud radiation database for EPS-SG ICI: Task 3 report

Bengt Rydberg[2], Eleanor May[1], Hanna Hallborn[1], and Patrick Eriksson[1]

[1] Department of Space, Earth and Environment, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden
[2] Swedish Meteorological and Hydrological Institute, SMHI, SE-601 76 Norrköping, Sweden

February 3, 2023

Contents

1 Introduction

1.1 Purpose of this document

Previously, a retrieval algorithm based on Bayesian Monte Carlo integration (BMCI) was developed [RD-2, RD-3] with the aim of providing global retrievals of ice water path (IWP) from ICI observations. The retrieval database is an integral part of the algorithm, providing simulated ICI observations corresponding to pairs of atmospheric and surface states.

The purpose of this document is to present a validation of the retrieval database and subsequently assess the retrieval performance of the database. This work is part of Task 3 of the project: *Development of a cloud radiation database for EPS-SG ICI*. The report builds upon both a literature review report (Kaur, Eriksson and Rydberg 2022) and a report detailing the development of the retrieval database.

1.2 Structure of this document

Section 2 of this document gives an overview of the cloud radiation retrieval database, including distributions of the data contained in the database. In Section 3, a validation of the database is presented. Firstly, a statistical comparison to DARDAR is made. Then, a further comparison is made between simulated and actual measurements using two test databases simulating GMI and MARSS/ISMAR observations. In Section 4, the retrieval algorithm is tested using the new database.

1.3 Reference documents

[RD-1] Accadia C. et al. (2020) Microwave and Sub-mm Wave Sensors: A European Perspective. In: Levizzani V., Kidd C., Kirschbaum D., Kummerow C., Nakamura K., Turk F. (eds) Satellite Precipitation Measurement. Advances in Global Change Research, vol 67. Springer, Cham. `https://doi.org/10.1007/978-3-030-24568-9_5` .

[RD-2] Rydberg, B.: EPS-SG ICI ice water path product: ATBD. SAF/NWC/LEO-EPSSG/ATBD/IWP-ICI Issue 2.1, Rev. 2, Tech. rep., EUMETSAT. NWCSAF, www.nwcsaf.org, -¿ documentation, code: ICI, 2018.

[RD-3] Eriksson, P., Rydberg, B., Mattioli, V., Thoss, A., Accadia, C., Klein, U., and Buehler, S. A.: Towards an operational Ice Cloud Imager (ICI) retrieval product, Atmos. Meas. Tech., 13, 53–71, `https://doi.org/10.5194/amt-13-53-2020`, 2020.

[RD-4] EPS-SG End User Requirement Document (EURD) EUM/PEPS/REQ/09/01515 01/11/2019

2 Overview of database simulations

The final database contains $\sim 9.5 \times 10^6$ cases, where each individual case consists of a simulated ICI measurement and a corresponding set of atmospheric and surface states.

The primary quantities of interest are the integrated ice water path (IWP), the mean mass height Z_m, and the mean mass diameter D_m. The liquid water path (LWP), rain water path (RWP), and water vapour (WV) are also included in the database. Further details on the format and content of the cloud radiation retrieval database are detailed in the Data Format Specification Document (DFSD) [RD-2]. Distributions of the aforementioned quantities were calculated from all cases in the database, and are shown in Figures 1 and 2.

Distributions of brightness temperatures T_b present in the database are shown in Figure 5, split according to ICI channel. Both one- and two-dimensional probability density functions are shown. Although there are presently no ICI observations to compare with simulations, there is still value in performing a qualitative check of simulated measurements.

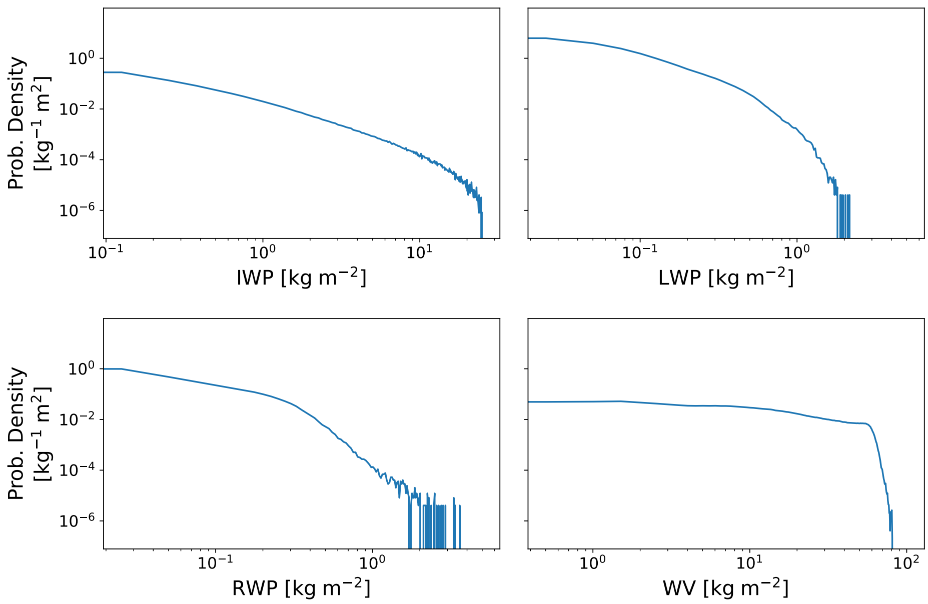

Figure 1: Probability density functions (PDF) of all ice water path (IWP), liquid water path (LWP), rain water path (RWP), and water vapour (WV) values present in the database.

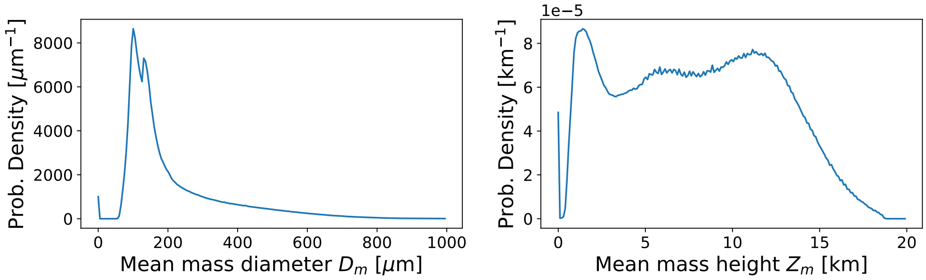

Figure 2: Probability density functions (PDF) of D_m and Z_m values present in the database. The peaks in probability density arising at $D_\mathrm{m} = 0$ μm and $Z_\mathrm{m} = 0$ km corresponds to cases in which $IWP = 0$ kg m^{-2}.

2.1 Developments from previous database

A major improvement to the new retrieval database is the inclusion of vertically and horizontally polarised channels 04V, 04H (243 GHz), 11V, and 11H (664 GHz). In contrast, the previous database [RD-3] contained only the intensity simulation for such channels, calculated as the mean of V and H. As shown in Figure 3, there exists a notable difference between simulated T_bs for 4V and 4H. A strong correlation between V and H remains, but slightly lower T_b values are now present at horizontal polarisation due to the introduction of azimuthally randomly oriented hydrometeors within the simulations. When comparing the one-dimensional distributions, the general shape of the distributions differs between the old and new databases. The largest difference occurs at medium to high T_b values, many of which generally correspond to clear-sky cases. The new database has been improved in regards to the handling of surfaces and the inclusion of more high latitude cases, both of which could contribute to the change in T_b distribution. However, the handling of surface emissivities may still lead to a failure to reproduce the very highest T_b values, although these would occur with a low probability density. A further investigation and discussion of this issue is made upon comparison of a test GMI database with GMI observations, and can be found in Section 3.2.3

In Figure 4, T_b from channels 04V and 04H from the new database are plotted against channel 04 from the old database. The two-dimensional distribution between 04V and 04H is also included to demonstrate the difference between the polarised channels, where the vertically polarised channel is shown to have slightly higher T_b than the horizontally polarised channel. Such a difference was not present in the old database due to the existence of only one non-polarised channel. A comparable plot is shown in Figure 3 for channels 11V and 11H from the new database, and 11 from the old database. In this case, distributions from the new database show a small shift towards colder temperatures for higher T_b values. Cases demonstrating high T_b generally correspond to clear-sky cases and so this shift could potentially be attributed to the full inclusion of the interference of ozone within the new database.

Improvements to the new database have also been made in regards to the surface emissivity model, the particle model (including some random variation), and the antenna footprint. Simulations for all channels benefit from these changes.

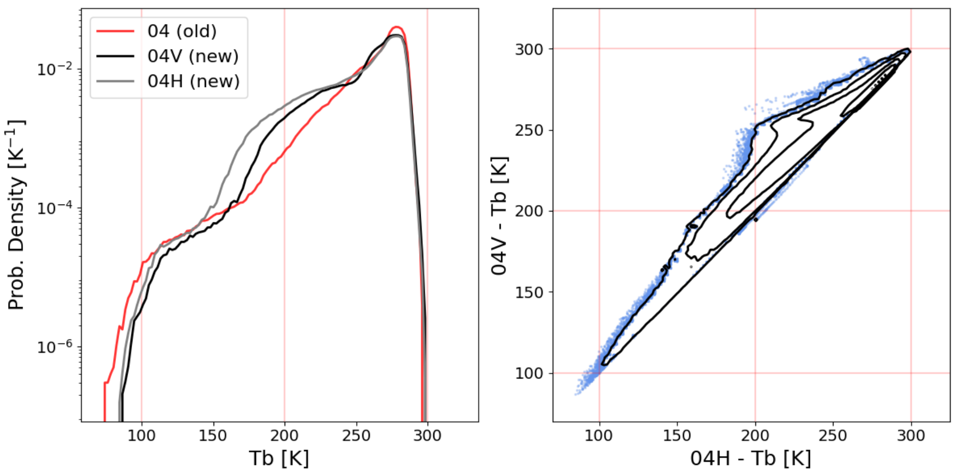

Figure 3: A comparison of the 243 GHz channel T_b distributions from the previous database and the new database (left). The old database contains only one channel at 243 GHz, denoted as 04. The new database contains two channels at 243 GHz, denoted as the vertically and horizontally polarised channels 04V and 04H, respectively. It is therefore possible to construct two-dimensional distributions in the case of the new database, illustrating that the difference between polarisations can now be captured using the new database. Samples in the two-dimensional distribution (right) that occur with a probability density less than 10^{-6} K^{-2} are represented by blue markers.

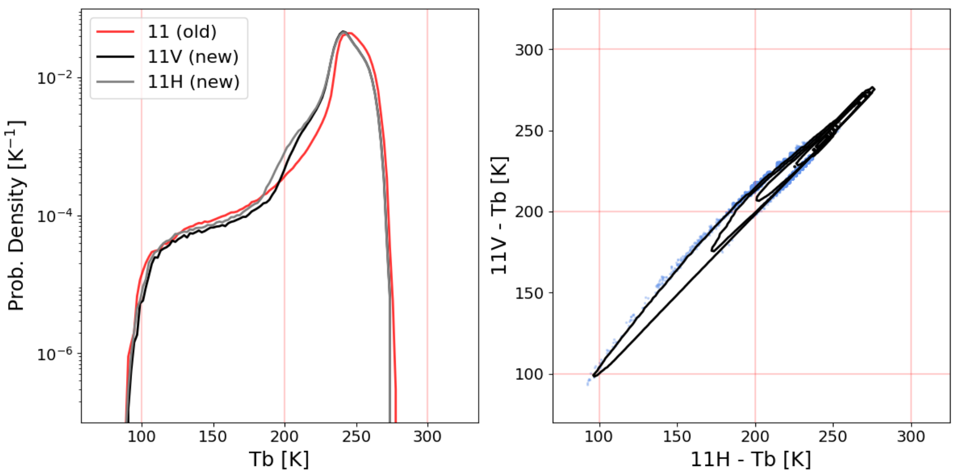

Figure 4: A comparison of the 664 GHz channel T_b distributions from the previous database and the new database (left). Only one distribution is available from the old database. This is denoted as 11 in the plot and denotes the intensity simulation computed as the mean of 11V and 11H. 11H and 11V are available from the new database and both are plotted in comparison to the old database. The two-dimensional distribution between 11V and 11H is also shown (right), where the blue markers indicate samples that occur with a probability density less than 10^{-6} K^{-2}.

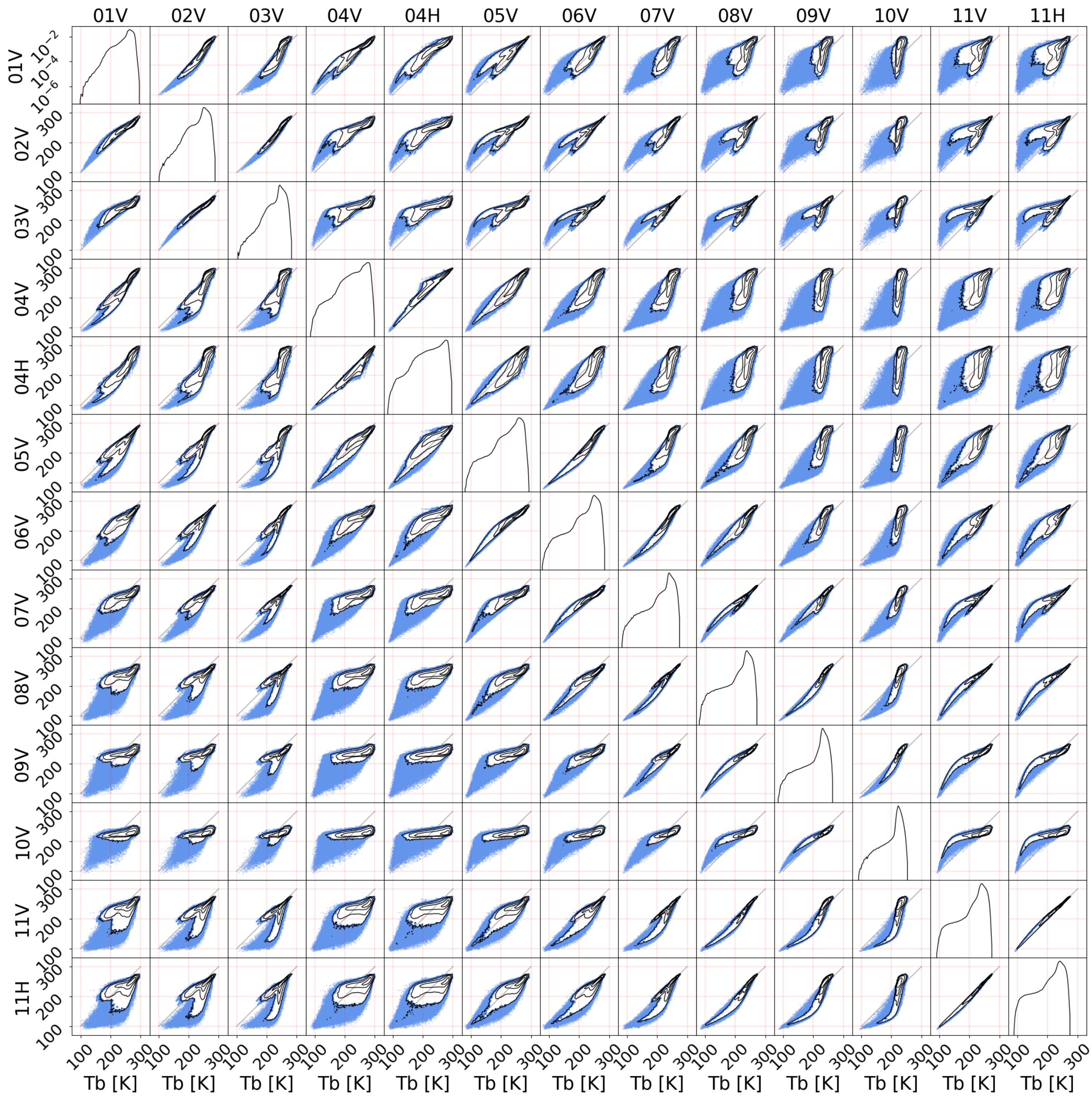

Figure 5: PDFs of simulated antenna-weighted brightness temperatures T_b in a version of the retrieval database consisting of $\sim 8.9 \times 10^6$ cases. On the diagonal, 1-dimensional PDFs are shown. Two-dimensional joint distributions are shown off-diagonal. Contour lines correspond to $[10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}]\ K^{-2}$, with $10^{-5}\ K^{-2}$ corresponding to the outermost contour line. Samples that occur with a probability density less than $10^{-6}\ K^{-2}$ are represented by blue markers.

3 Database validation

In the absence of ICI data, validation of the cloud radiation retrieval database is not immediately straightforward. However, there are several methods that allow for validation to a certain degree. Firstly, one can verify that the cloud ice products in the database are consistent to existing data in a statistical sense. For this purpose, data from the DARDAR product (Delanoë et al. 2014) are used.

A further assessment of the database can be made through an application of the database-generation method to existing instruments with sufficient similarity to ICI. For this purpose, retrieval databases are generated for the Global Precipitation Measurement (GPM) Microwave Imager (GMI) and for the International Submillimetre Airborne Radiometer (ISMAR) (Fox et al. 2017). It is then possible to verify that the simulated radiances from the test databases are sufficiently statistically similar to observed radiances.

3.1 Comparison to DARDAR

As discussed, one validation approach is the *statistical* comparison of cloud ice products in the database, such as ice water path (IWP), to pre-existing products. The DARDAR product (Delanoë et al. 2014) provides the most suitable satellite-based reference data available for this purpose. However, it must be noted that there are uncertainties present in the DARDAR product. The retrieval uncertainty in Cloudsat IWC is estimated at around ±40% (Heymsfield et al. 2008, Eliasson et al. 2011), and the same uncertainty is assumed for the DARDAR IWP in this analysis.

Comparisons of the probability density function (PDF) of IWP and the zonal means of IWP are shown in Figure 6. Also included are the same statistics computed on a preliminary version of the database [RD-3]. The current database produces a PDF which agrees more closely with DARDAR for lower IWP values. Although the previous database appears to agree better with DARDAR at higher IWP values, it must be noted that the probability density is given on a logarithmic scale and thus the difference between distributions at the higher end of the IWP scale is significantly smaller than differences at the lower end. Both the current and previous database contain more high IWP values than DARDAR, which is desirable when performing retrievals. In the case of the zonal mean, the new database is very consistent with DARDAR. It performs notably better than the old database, since there are a significant number of zonal means from the old database that fall outside of the (relatively large) DARDAR uncertainty range.

Global mean IWP maps for the new retrieval database, the old database, and the DARDAR product are shown in Figure 7. All three datasets agree in their representation of regions commonly associated with high IWP, such as the intertropical convergence zone (ITCZ). Similarly, all datasets show low average IWP in regions such as the Sahara desert and in regions generally associated with high amounts of statocumulus clouds. The new retrieval database shows a 'smoother' distribution of IWP than the old database, and agrees more closely with DARDAR in this sense.

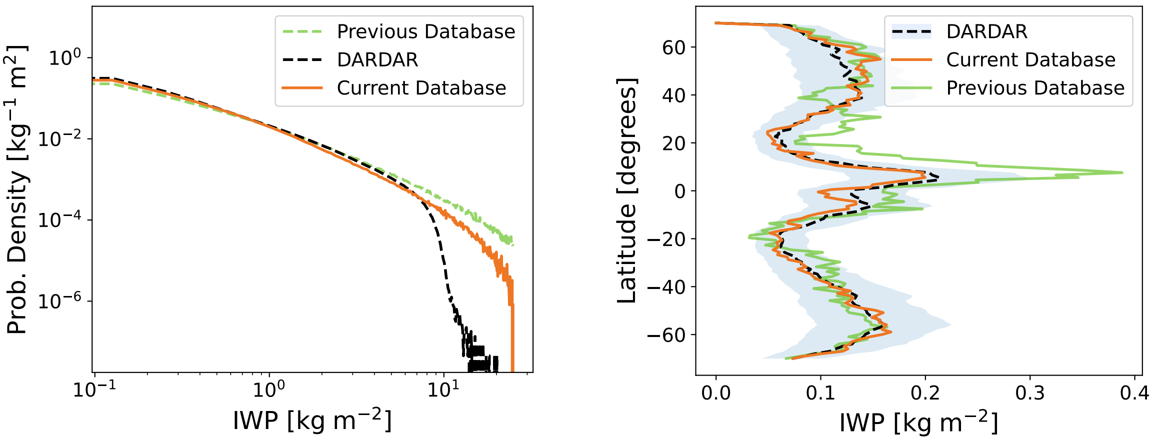

Figure 6: A comparison of (left) PDFs and (right) the zonal means of IWP values from the current cloud radiation retrieval database, the previous database, and IWP values from the DARDAR product. The shaded region shown for the zonal means (right) represents the ±40% uncertainty in DARDAR IWP. DARDAR data are taken for all months in 2010. Distributions were calculated across all latitudes and surface types available.

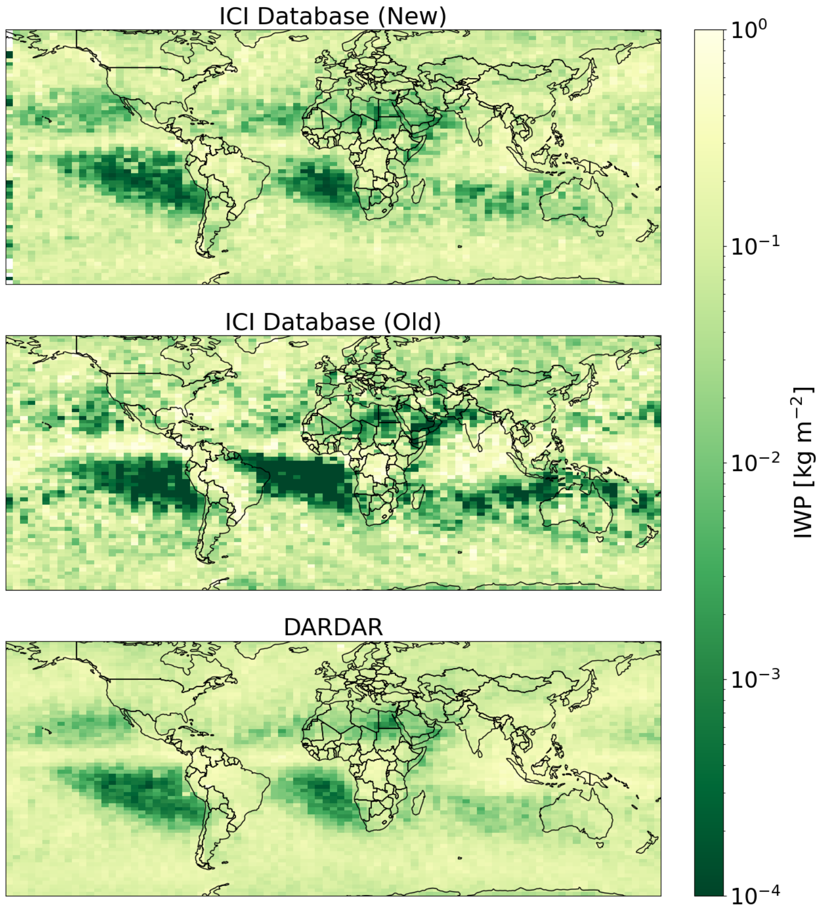

Figure 7: Global mean IWP from the new cloud radiation retrieval database (upper), the old database (middle), and the DARDAR product (lower). DARDAR data are taken for all months in 2010. Grid boxes span 2 degrees longitude by 2 degrees latitude.

3.2 GMI forward model simulations

A retrieval database with simulations of the high-frequency channels of GMI, listed in Table 1, was created for validation purposes. The GMI retrieval database consists of roughly 1.3 million cases covering the period January to April in 2009 and 2010.

First, simulated brightness temperatures and polarisation differences from the GMI database were compared to actual observations from GMI. These results are presented in Sec.3.2.1. Gaussian noise, with standard deviation according to NEΔT of Table 1, was added to the simulated brightness temperatures in the database to account for instrumental measurement errors when comparing to GMI observations. Then, retrievals based on the GMI database was compared to DARDAR products to further assess the database. These results are discussed in Sec.3.3

Channel	**Polarisation**	**NEΔT [K]**
166 GHz	V	0.70
166 GHz	H	0.65
183 ± 3GHz	V	0.56
183 ± 7 GHz	V	0.47

Table 1: The four high-frequency channels in GMI, simulated using the framework developed for generation of a cloud retrieval database for GMI. The channels are shown alongside their polarisation and the standard deviation of measurement noise. Noise values are taken from Kaur, Eriksson, Barlakas, Pfreundschuh and Fox (2022).

The simulated brightness temperatures (T_b) and polarisation differences (PD) in the GMI database were compared to actual observations from GMI during the same period of the year (January to April) but since GMI started to operate in 2014, data were taken from the year 2020 instead. The simulations do not correspond to individual observations but should rather agree with the observations in a statistical sense. The geographical sampling of observed GMI data is non-uniform with more data at the turning points of the satellite swath (between 60 and 70 degrees). To reduce the effect of this uneven sampling, the results in this section consider latitudes from -60 degrees to 60 degrees. All surface types are included in this section. To see the impact of different surface types, results where the data have been filtered on 'ocean', 'land', 'snow' and 'sea-ice' are presented in Appendix C where the surface classification for observations was taken from the GPROF retrieval product. Results from different latitude regions can also be seen there.

3.2.1 Brightness temperature distributions

Probability distributions for simulated and observed T_b are seen in Figure 8 for each of the high-frequency channels of GMI. For the 183 GHz channels, T_b distributions for corresponding ICI channels from the old (preliminary) database are also included. The fit to actual observations are generally quite good for clear-sky cases (the peak at high T_b values) although significantly higher T_b values, above 300 K for the 166V and 166H channels, are present in the observations. This difference has been identified with cases over land and is discussed in Sec. 3.2.3.

Figure 8 also shows that GMI simulations for cases with cloud impact agrees well with observations for 183±3V and quite well for 183±7V down to about 120 K. Cases below 120 K are very rare and are therefore not sufficiently well represented in the limited size

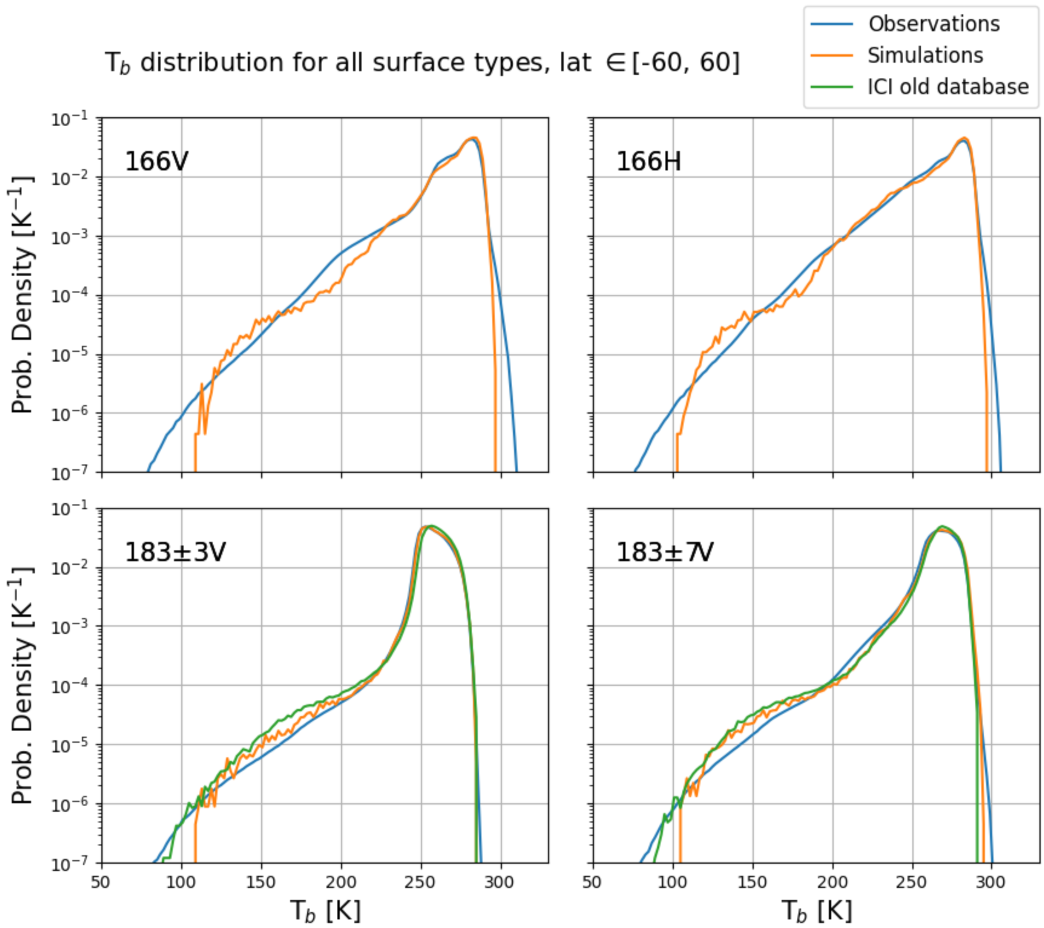

Figure 8: Brightness temperature (T_b) probability distributions for each of the high-frequency channels of GMI. Data from all surface types are used and from latitudes in the range -60 degrees to 60 degrees. Distributions for observations (blue) and simulations (orange) are seen for each channel and for the 183GHz channels, ICI T_b distributions for the old (preliminary) database are also included in green.

of the GMI database. This is in line with our expectations since it is not expected that the distributions should agree for probability densities below 10^{-6} as cases with such low occurrence frequency are not expected to appear in the simulated database. A larger GMI database would be required to fairly compare the occurrence of low T_b values.

For the 166 GHz channels in Figure 8, the distribution for GMI simulations deviates more from the observations. Especially for T_b values slightly lower than 200 K where there is a bump in the observations that is not captured by simulated T_b. The 166 GHz channels are more sensitive to surface radiation and this suggests that the observation data contain a larger fraction of snow and ice covered surfaces (cold surfaces). Looking at the latitude distribution of the observed data in Figure 28 in App. B, there are still considerably more data at latitudes around 50-60 degrees for the observations compared to in the GMI simulation database. The fit is improved when disregarding high latitudes as seen in Figure 29 in App. C.1 where only tropical latitudes are included. From the latitude distribution it is also seen that the simulation database contains relatively more cases at tropical latitudes where there are deep convection clouds. This could explain why there are more cases with really low T_b, around 150 K, in the simulations than in the observations as seen in Figure 8.

The joint probability distribution of simulated and observed T_b for the different channels is seen in Figure 9. The figure shows that the correlations between the simulated T_b for the different channels generally agrees well with the correlations between observed T_b. The simulated data are thus confirmed to cover the observed measurement space by GMI which is crucial for the retrieval algorithm since the database defines the prior knowledge of the atmospheric state.

Depending on surface type and hydrometeor impact there is a difference between measured T_b for the different polarisations of the 166 GHz channel as seen in the distribution between 166H and 166V in Figure 9. Surface impact results in relatively large variability of the distribution at higher T_b values seen as the arm towards lower T_b for 166H in the interval 200 K to 300 K for 166V. The variability that is seen for lower T_b (below 200 K identity point) in Figure 9 is due to hydrometeor impact. From the figure it is clear that the simulation database covers both these variabilities. The limited size of the simulation database is again the reason why rare cases at low T_b values are not captured by the distribution for simulations but as discussed above, it is not expected to do so.

3.2.2 Polarisation difference distributions

Polarisation difference (PD) distributions of database simulations and actual observations were also compared to see how well the simulation database captures variations in T_b for V and H polarisation. In this study, azimuthally random orientation (ARO) of hydrometeors is assumed to include this variation from cloud structures. The method for how ARO is approximated for this study is described in (Kaur, Eriksson, Barlakas, Pfreundschuh and Fox 2022). PD is defined as the difference between T_b for the 166V and 166H channels, i.e. as $\mathrm{PD} = \mathrm{T}_{b,166V} - \mathrm{T}_{b,166H}$.

In Figure 10, the joint probability distribution between PD and $T_{b,166V}$ is seen. The arc-shape seen between $T_{b,166V}$ of 100 K and about 230 K are as discussed above due to hydrometeor impact. Atmospheric states with aligned hydrometeors result in PD of about 20 K in this region as seen by the observed PD distribution. Low PD values in this region correspond to turbulent atmospheric states (such as deep convection) where the PD signal is reduced due to totally random orientation of the hydrometeors and multiple scattering. The distribution for the simulations in Figure 10 is seen to capture this arc-shape but is a bit broader towards larger PD. This is on purpose since the upper limit of the randomly chosen aARO factor (that scales the extinction of the two polarisations to approximate ARO of hydrometeors) have been intentionally increased compared to (Kaur, Eriksson, Barlakas, Pfreundschuh and Fox 2022). This is to ensure that the effect of ARO of hydrometeors on the other channels is also captured by the approximated ARO. The database can later be resampled to adjust this if needed. Negative PD values are likely due to thermal noise, where the lower number of negative values at low $T_{b,166V}$ correspond to the fact that there is very little data in this region. Negative PD values could also arise due to surface impact, but such a conclusion cannot be drawn from this figure alone. Horizontally aligned particles can also lead to negative PD values, but no evidence of this can be found in this figure.

The strong arm at $T_{b,166V}$ of about 250 K in Figure 10 is due to different radiative properties of different surface types. The largest PD values in the arm correspond to dry clear-sky cases where the PD signal is not attenuated by interaction with water vapor. For humid atmospheric conditions, the PD signal from the surface is attenuated corresponding to cases with low PD and high $T_{b,166V}$ in the lower right corner of the plot. It is seen that this arm is well represented in the distribution for the simulations in Figure 10 although

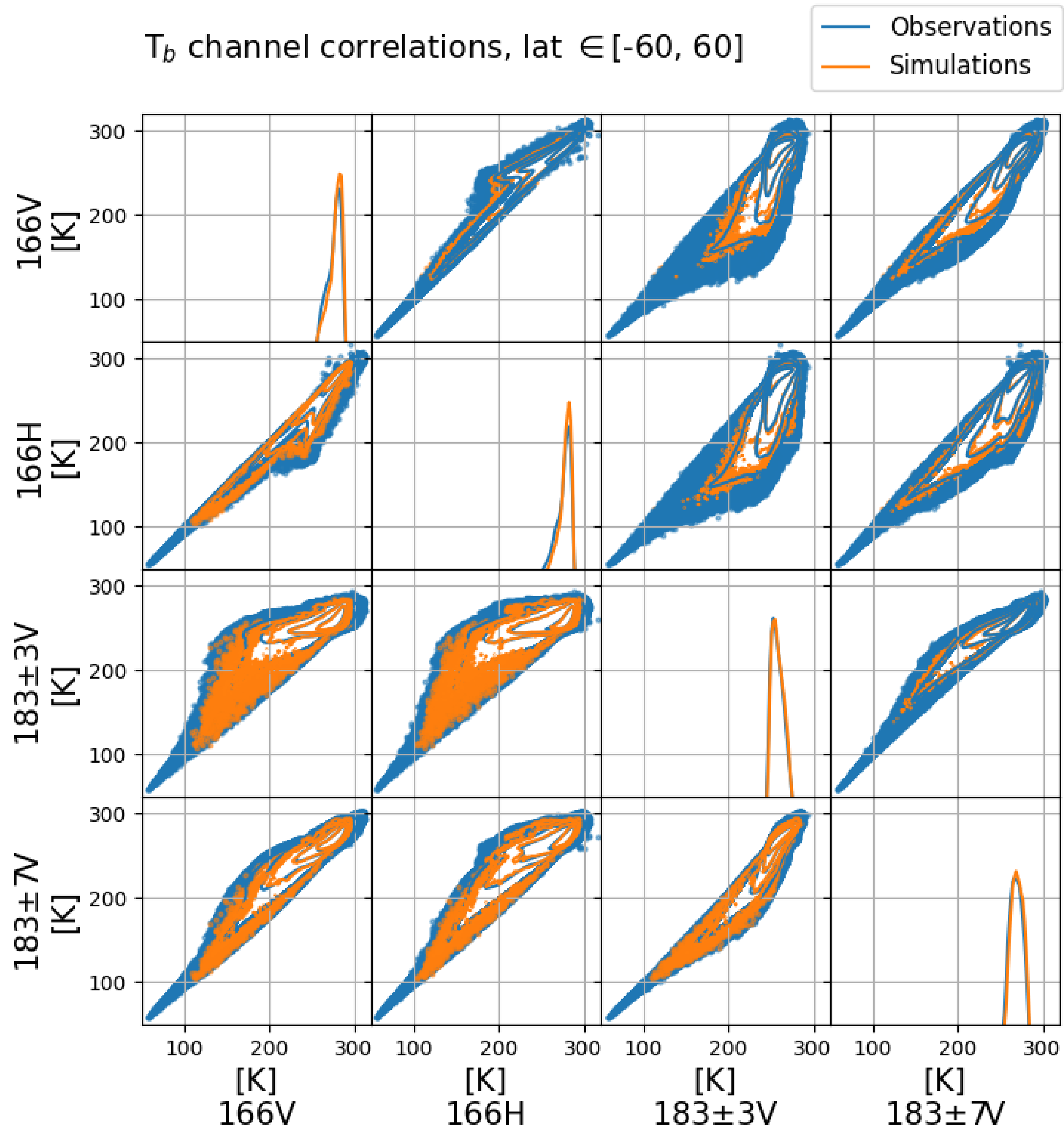

Figure 9: Brightness temperature (T_b) channel correlations for GMI observations (blue) and GMI simulations (orange). Joint probability distributions between each of the channels are seen where data from all surface types and from latitudes in the range -60 degrees to 60 degrees are used. Simulations are plotted on top of observations below the diagonal and above the diagonal, observations are plotted on top of simulations. The contour lines are at levels 10^{-6} K^{-2}, 10^{-5} K^{-2}, 10^{-4} K^{-2} and 10^{-3} K^{-2} and data points outside the 10^{-6} K^{-2} contour are plotted as dots. Marginal distributions are seen at the diagonal.

there seem to be a lack of cases for high $T_{b,166V}$ which again have been attributed to problem with surface modelling of land, see discussion in Sec. 3.2.3.

It is expected that the PD distributions for land should show smaller PD values for clear-sky cases than water surfaces. This is seen for the distribution for the simulations in Figure 36 and Figure 37 in Sec.C.2 where the PD distributions are plotted for ocean and land separately. The probability for large PD values is indeed smaller in the clear-sky arm for observations over land as well in these figures but there is a considerable amount of high PD values that are not present in the simulations in Figure 36. These are probably cases that have been misclassified by GPROF. For mid-latitudes (Figure 37) there are less really cold cases from deep convection clouds which is seen for both observations and simulations.

PD distributions where also plotted for snow covered land and sea-ice at northern latitudes, seen in Figure 38 but there are too few cases (snow about 42k and sea-ice about 5k) in the simulation database to properly compare the distributions to observations. Roughly the same set of PD values are seen for simulations and observations but there seem to be some problem for low $T_{b,166V}$ below 200 K for both snow and sea-ice.

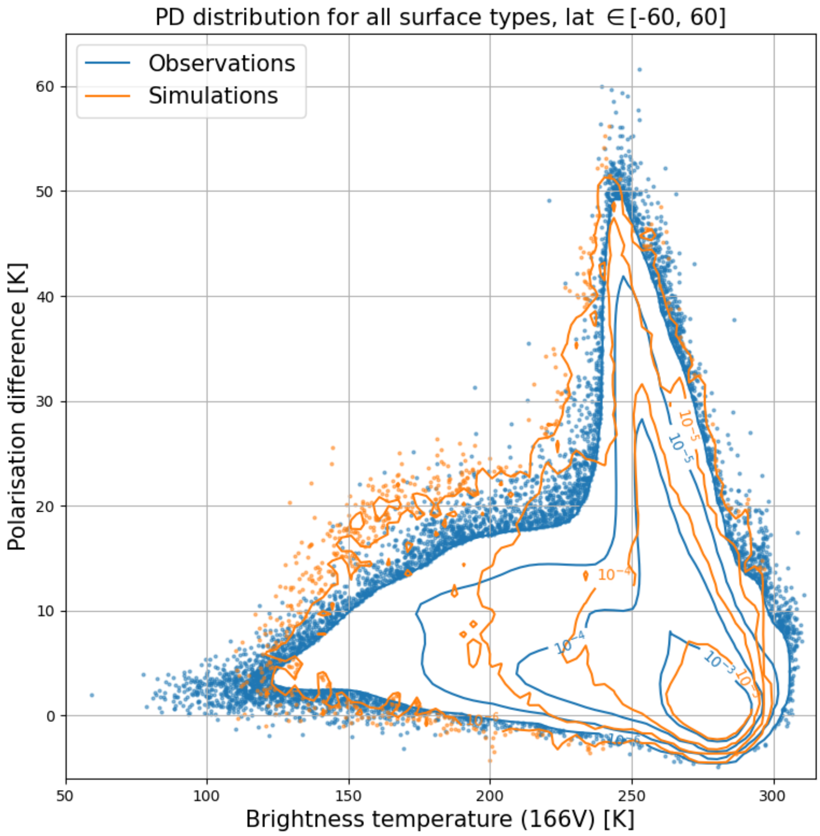

Figure 10: Joint probability distributions of polarisation difference (PD) and brightness temperature for the 166V channel ($T_{b,166V}$) are plotted for all surface types and latitudes in the range -60 degrees to 60 degrees. PD is defined as $T_{b,\mathrm{166V}} - T_{b,\mathrm{166H}}$ and details of the modelling of hydrometeors in the simulations can be seen in Sec. 3.4 in the Task2 report. Contour lines of the probability distributions for observations (blue) and simulations (orange) are seen at levels 10^{-6} K^{-2}, 10^{-5} K^{-2}, 10^{-4} K^{-2} and 10^{-3} K^{-2}. A random sample of 1% of the simulated data points with lower probability than 10^{-6} K^{-2} are plotted as blue dots. For the simulations all point with lower probability than 10^{-6} K^{-2} are plotted as orange dots.

3.2.3 Surface emissivity

In the distributions shown in Sec. 3.2.1 and Sec. 3.2.2 there have been significant differences between simulations and observations for really high T_b values. The problem is most significant in the 166 GHz channels where contributions from surface radiation have larger impact. Considering T_b distributions where only data from land have been included, as in Figure 30 and Figure 32 in App. C.1 where simulated values are less than 300 K while observed values clearly exceeds 300 K for 166H and 166V, this suggests that the difference could be due to problems in the modelling of surface radiative properties for land. The problem is significantly reduced when looking at T_b distributions where land is not included, e.g. Figure 31 in App. C.1.

The geographical locations for observations with $T_b > 300$ K for the 166V channel are seen in Figure 11. It is seen that most of these observations are in desert regions. Such high values are not obtained in the simulations. A first possible reason for this is too low emissivities for the regions of concern. Empirical values from TELSEM (Aires et al. 2011) are used to estimate the emissivities for land surfaces. These values are derived for 85 GHz and the validity for 166 GHz is not known. Another possible reason why the simulations do not include cases with $T_b > 300$ K is that the ERA5 skin temperatures are too low for arid and desert regions.

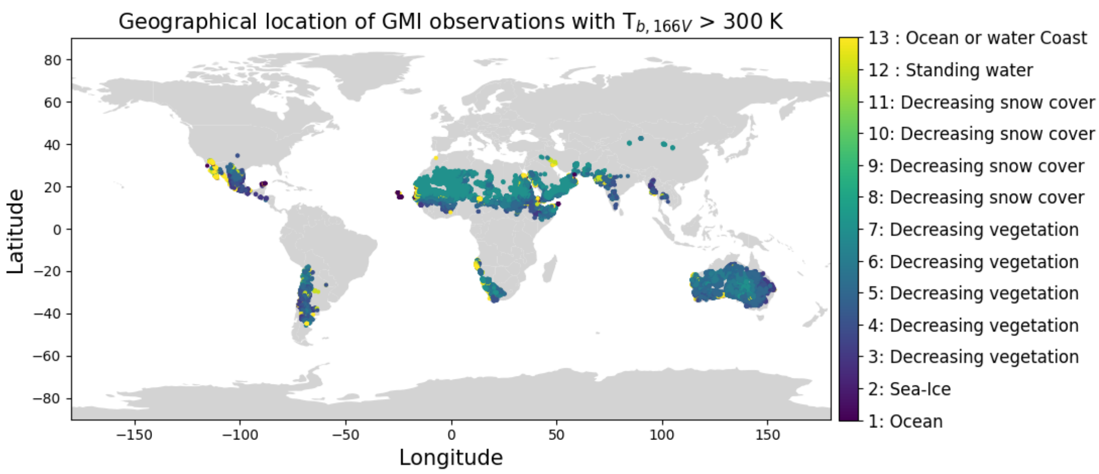

Figure 11: Geographical location of GMI observations with measured $T_b > 300$ K for the 166V channel. The GPROF surface classifications of the observations are indicated by the colorbar. Decreasing vegetation 3-7 are the classifications used for surface type 'land'.

3.2.4 Summary of forward model validation

Brightness temperature distributions for the simulation database agree very well with distributions for observations for the 183 GHz channels and quite well for the 166 GHz channels. Deviations were mainly attributed to problems in the modelling of surface radiative properties for land and differences in geographical sampling of observations and simulations. It was also noticed that a large database, with the number of samples approaching 10 million, is needed to represent rare cases with low T_b values.

The joint T_b distributions between the different channels of GMI in Figure 9 confirmed that the observed measurement space is covered by the simulation database. It was also shown that variability in T_b between the 166H and 166V channels due to both surface and cloud impact is captured by the simulations.

From the PD distributions in Figure 10 we conclude that the database captures the overall observed PD signature of GMI. Both clear-sky cases and cases with cloud impact agreed well with observations. A slightly broader distribution for simulations was seen for cloudy cases because the aARO factor has been intentionally increased. This is not a problem since the simulation database can be resampled to adjust this. However, PD distributions for snow and sea-ice were not completely satisfactory although more cases are needed to really compare the distributions to observations.

3.3 GMI retrievals

A test of the retrieval performance of the test GMI database is performed through the comparison of the test retrievals to DARDAR. For this purpose, quantile regression neural networks (QRNN) were used. Similarly to BMCI, QRNNs allow retrievals to be obtained alongside non-Gaussian retrieval errors. For a further discussion of QRNN and a comparison with BMCI, the reader can refer to Pfreundschuh et al. (2018).

3.3.1 Retrieval architecture

The input data for the QRNN consist of radiances for the four highest frequency GMI channels, as shown in Section 3.2. Surface type and surface temperature were included as additional inputs.

Simulations are noise-free. To imitate measurement uncertainty, noise was added to each measurement. The added noise was Gaussian and followed the channel-dependent standard deviations given in Table 1. To account for potential error when performing the surface classification, 2% of the surface type data were randomly selected and randomised. Both of these approaches were conducted with the aim of preventing overfitting on the dataset and therefore poor retrieval results. Therefore, the addition of randomly generated noise and the surface-randomisation was performed *for each batch and each epoch* during training of the QRNN.

The outputs of the QRNN are quantiles of IWP, which can be used to form a cumulative distribution function (CDF). To account for the fact that IWP values can differ by several orders of magnitude, a log-linear transformation was applied to all IWP values prior to training. An inverse transformation was applied to the retrieved values prior to plotting the results.

The database was split such that 80% of the datapoints were used for training, 10% for validation, and 10% for testing. This gave a training set containing 1.2×10^6 samples and validation and test sets both containing 1.5×10^5 samples. The test set was not seen by the neural net at any point during training, allowing for the performance of the retrievals to be fairly assessed. The training, validation, and test datasets were constructed through random selection of data and therefore follow the same underlying distribution.

3.3.2 Retrieval performance on simulated observations

To evaluate the performance of the retrievals, retrievals were made on the test set and compared statistically to DARDAR through the calculation of PDFs and zonal means. The results are shown in Figure 12. The deviation of retrieved values from reference values is shown in Figure 13. In order to compute a PDF of retrieved IWP, point estimates of IWP needed to be chosen from the quantiles outputted by the QRNN. Two PDFs were constructed. The first uses the mean as the point estimate of IWP, where the mean is computed as the first moment of the predicted CDF. The second PDF takes a random sample from the a posteriori distribution as the point estimate, where the posterior distribution is computed through an interpolation of the inverse CDF. This is henceforth referred to as 'posterior sampling'. However, it is stressed that this 'posterior sampling' is not equivalent to the method of repeated sampling of a posterior distribution to obtain an expectation value, since only a single sample is taken from each distribution in this case.

The test GMI database consisted of simulations of atmospheric scenes based on Cloudsat data from January to April. For a fair comparison, the DARDAR distributions were recomputed to be only over data from January to April.

The QRNN has not seen the test dataset during training. To be successful, the distribution of retrievals from the test dataset should follow the same distribution as IWP from the training dataset. A significant deviation from the training set would indicate issues with the retrievals, but no such deviation was present.

Sampling from the posterior distribution allows for the occasional sampling of higher IWP values that are not captured when taking the mean as a point estimate. Therefore, it is expected that the distribution given by the mean will not extend as high as the distribution given by posterior sampling. The posterior sampled distribution shown in Figure 12 is therefore a better representation of the capabilities of the database and retrievals. We note that it is also probable that the mean suffers as a result of quantile crossing, a phenomenon sometimes present in QRNNs where the estimated quantiles are not monotonic. For this study, any quantile crossing present was corrected with isotonic regression on the quantiles post-estimation.

The posterior sampled symmetric error distribution shows no clear bias towards over- or under-estimation. The larger error values likely correspond to IWP values comparable in magnitude. Since larger IWP values occur at very low probability densities, the lower accuracy of these retrievals is attributed to an under-representation of high IWP in this test GMI dataset.

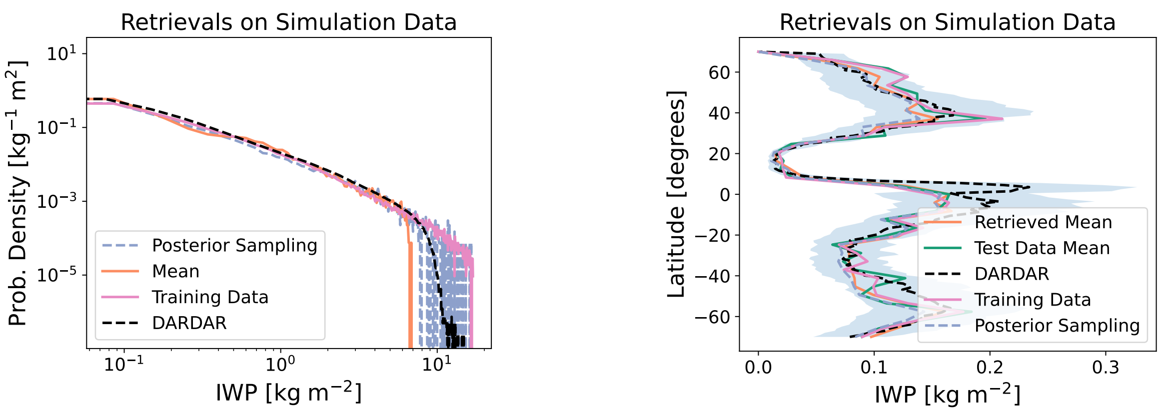

Figure 12: A comparison of (left) PDFs and (right) the zonal means of retrieved IWP using the test GMI database.

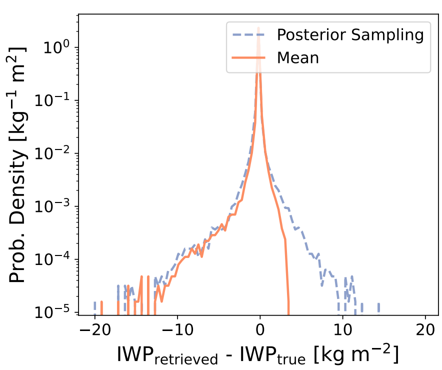

Figure 13: PDF of retrieval error. Distributions are shown using mean as the point estimate of IWP and using samples from the a posteriori distribution as the point estimate.

3.3.3 Retrieval performance on real observations

The QRNN was trained on simulations of GMI observations. To further test the retrieval performance, the same QRNN was used to retrieve IWP from *actual* GMI observations. For input to the QRNN, GMI brightness temperatures were paired with surface temperatures and surface type classifications taken from the Goddard Profiling Algorithm (GPROF). Due to differences in surface classifications used in this study and in GPROF, some datapoints were forced to be classed with an unknown surface type. This was expected to slightly decrease the performance of the retrievals.

As before, the PDF and zonal mean of retrievals of IWP are computed. Results are shown in Figure 14. Retrievals were performed on $\sim 10^6$ observations from GMI in 2019. Observations were randomly selected from the entire year and across all latitudes and longitudes with the aim of obtaining a comprehensive distribution of IWP values.

The distribution of IWP retrievals agrees well with DARDAR in a statistical sense. It must be noted that the database was not specifically designed to match with DARDAR and therefore this is a welcome result. The IWP PDF obtained through posterior sampling is slightly more skewed towards higher IWP values than DARDAR. Also shown is IWP given by GPROF for the same observations used for the retrievals. Although DARDAR cannot be claimed to be the 'truth', GPROF retrievals are based upon a database focused on hydrometeor precipitation and are thus generally regarded as less accurate than DARDAR. Therefore, the fact that our IWP retrievals from the same dataset agree more closely with DARDAR than GPROF does indicates a significant improvement in the realm of cloud ice retrievals.

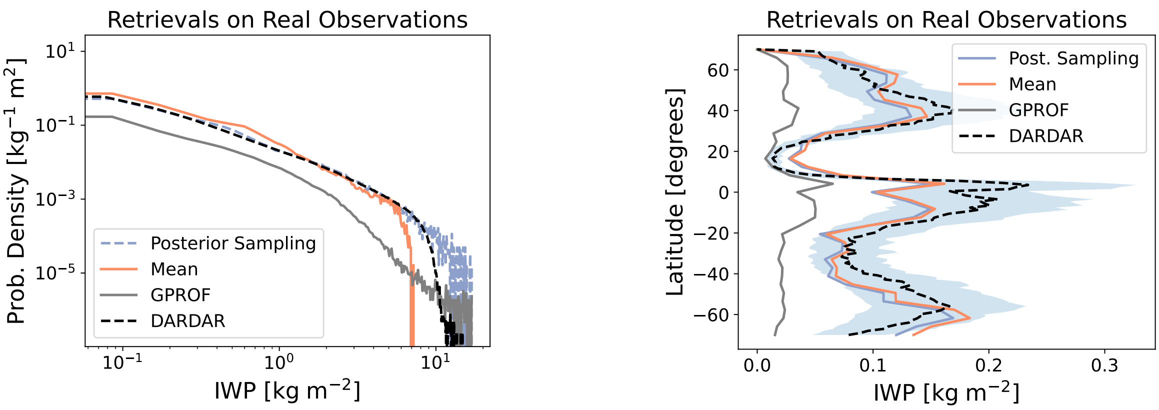

Figure 14: A comparison of (left) PDFs and (right) the zonal means of retrieved IWP from actual GMI observations.

3.4 ISMAR and MARSS

The Microwave Airborne Radiometer Scanning System, MARSS (McGrath and Hewison 2001), and the International Submillimetre Airborne Radiometer, ISMAR (Fox et al. 2017), is a microwave and submillimetre radiometer, respectively. ISMAR was developed as an airborne demonstrator for ICI, and has been flown, together with MARSS and other sensors, on the FAAM BAe-146 research aircraft. The first flight with ISMAR was succeeded in 2014 and since then a number of flights were realized, primarily covering the region around UK. Figure 15 displays tracks from 15 flights performed during 2015 to 2019 and where the altitude was above 9 km, and the observations from these parts of the flights are here compared to simulated observations.

Our simulations are not performed to match each flight individually. Instead, the MARSS/ISMAR simulation is performed for the scenes generated for the ICI database, but using a flight altitude of 9.4 km, an observation zenith angle of 140 °, and channel specification matching the MARSS and ISMAR ones. Figure 16 shows scatter plots of simulated and actual ISMAR and MARSS observations for data from ten channels. The simulated data include scenes covering the region of $45\,^{\circ}N$ – $75\,^{\circ}N$ and $40\,^{\circ}W$ – $5\,^{\circ}E$. The aim with this comparison is to verify that the simulated data cover the measurement space as observed by ISMAR and MARSS. That is, the aim is that each ISMAR/MARSS sample is found within the cluster of simulated samples, and in general Figure 16 confirms that this is the case.

It can further be noted that most of the data points are found along two "arms" for some of the scatter plots, and that both for the actual observation and the simulated data. The explanation to this is due to two different effects. First, for channels that have little or no sensitivity to the surface, like for example the 183.25±3.0 and 325.15±3.5 GHz channels, one arm is aligned close to the one-to-one line, while the other arm contain points where data from the 325.15±3.5 GHz channel is colder than that of the 183.25±3.0 channel. The points close to the one-to-one line are associated to clear sky conditions, while the points along the other arm are affected by cloud ice scattering. The scattering strength increases with frequency, and hence, the data from the 325.15±3.5 GHz channel is colder than that of the 183.25±3.0 channel. The clustering of data points along two arms are not that noticeable in the plots for channel pairs that are more close in frequency (e.g. the 325.15±1.5 GHz and 325.15±3.5 GHz channels).

A second effect that can give rise to clustering of data points along an arm is due to the surface contribution to the observation. This is most obvious in the plot showing data associated to the 243.20±2.5 GHz V and H channel, where some data from the V channel are up to 50 K warmer than that of the H channel in the simulated data, although not at all that high in the observed data. This high polarisation difference is associated to dry condition over an ocean surface, and this type of condition is over-represented in the simulated data compared to the observation dataset. This is not a big issue here, as the aim primarily is to verify that the simulation cover the observed measurement space.

The scatter plot for data associated to the 183.25±7.0 and 325.35±3.5 GHz channels looks distinctly different from the previously discussed plots, but also here the simulated data cover the observed data. It can be noted that for high and low temperatures data associated to the 183.25±7.0 GHz channel, are mainly warmer and colder, respectively, than that of the 325.35±3.5 GHz channel. This is explained by the fact that the atmospheric transmission is higher at 183.25±7.0 GHz than at 325.35±3.5 GHz, and hence, the influence of cloud ice can be greater at the lower frequency channel since the cloud ice may effectively not be seen at the higher frequency channel.

It can be noted that some ISMAR data from the 664 GHz channels are found a bit outside the simulated data. This is most likely explained by the fact that these channels sometimes are a bit more noisy than assumed in the simulation (noise in the form of NEΔT values according to Table 3 in Fox et al. (2017) was added to the simulation). Overall, the agreement between the observed and simulated data are judged to be satisfying and we conclude that the simulation setup is consistent to MARSS/ISMAR observation.

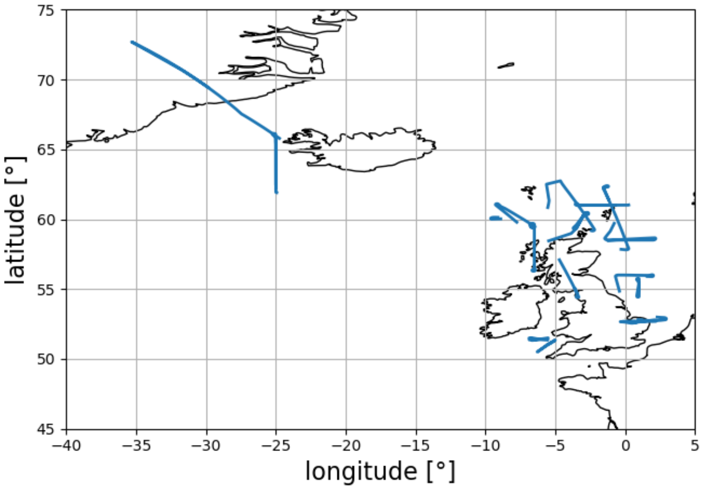

Figure 15: Flight tracks of the FAAM BAe-146 research aircraft where ISMAR and MARSS observations are available and flight altitude is above 9 km.

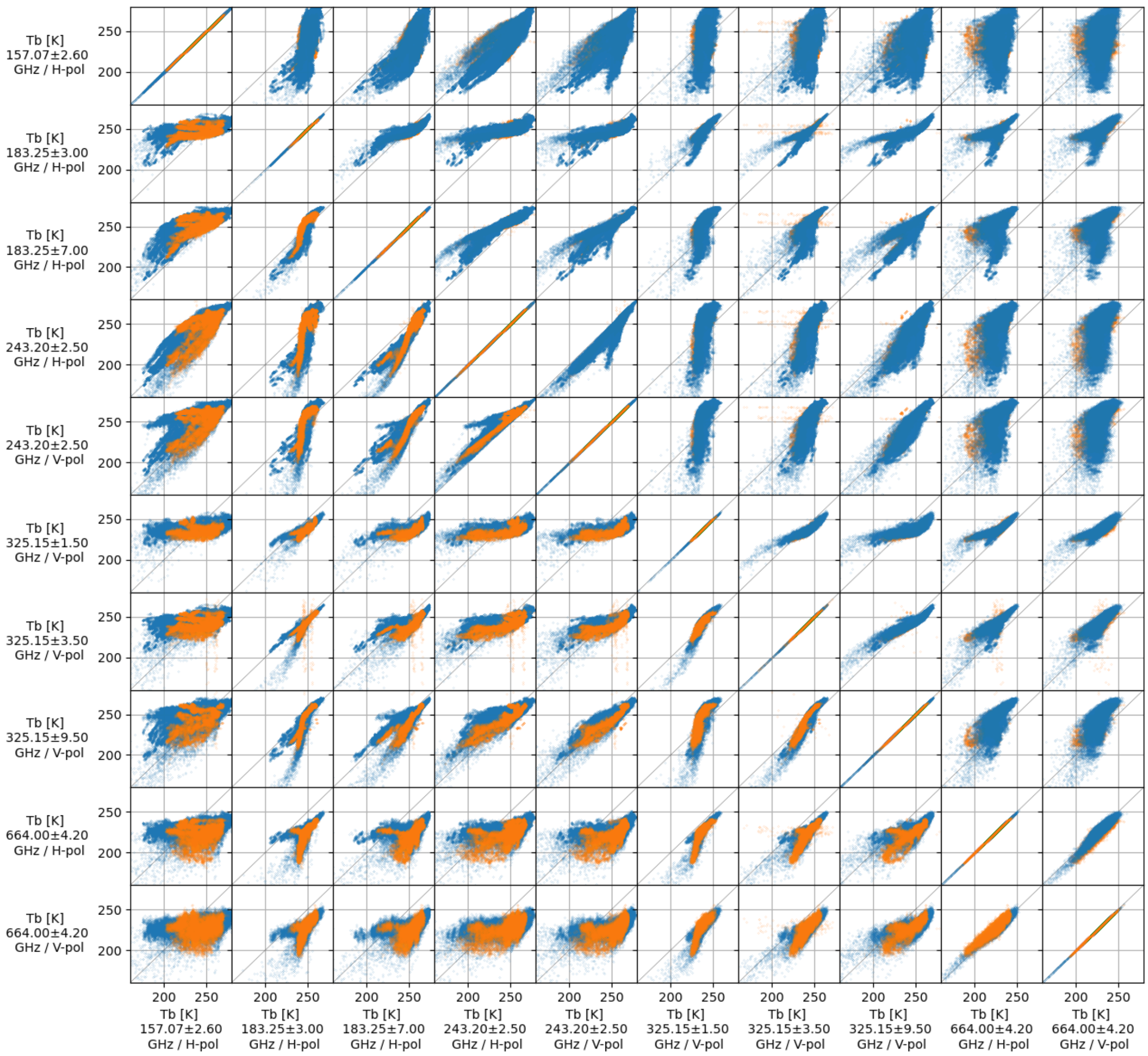

Figure 16: Scatter plots of simulated (blue dots) and actual ISMAR and MARSS observations (orange dots) for data from ten channels, and for a flight altitude around 9.4 km, and an observation zenith angle of 140 °. The simulation is not performed to match each actual observation individually. See text for more details.

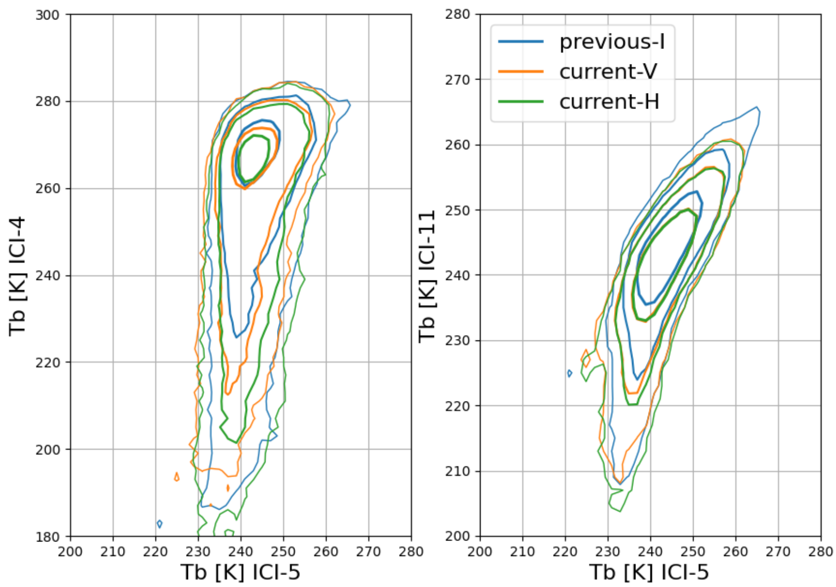

Figure 17: Contour plots comparing distribution of simulated observations associated to the previous and current database for the ICI-04, ICI-05, and ICI-11 channels. The current database contains data for both the V and H channels for ICI-04 and ICI-11, while the previous database only contains intensity simulation (mean of V and H). Three contour lines (0.0001 , 0.001, 0.01) are displayed.

Improvement with respect to the previous database

Data from the previous database can not directly be compared to the actual ISMAR/MARSS observation, as the allsky simulation is done for the ICI observation geometry. Anyhow, we here attempt to identify some improvements of the current database with respect to the previous database for some of the ICI channels and for data covering the region used in the previous section, where we found a good agreement betwen simulated and actual ISMAR observations for the simulation setup used for the current database.

Figure 17 shows a comparison of distribution of simulated observations associated to the previous and current database for the ICI-04 (243 GHz), ICI-05 (325±9.5 GHz), and ICI-11 (664 GHz) channels. The agreement of the Tb distributions associated to the current and previous database is relatively high on a general level, but there are a few differences. A main difference is that the current database contains data for both the V and H channels for ICI-04 and ICI-11, while the previous database only contains intensity simulation (mean of V and H). The outer contour lines do not deviate significantly for the current and previous database. The agreement is not perfect, but the contour line associated to the previous database is at least partly found between the H and V contour lines of the current database. The outer contour lines of the H and V data differ by up to $\sim$20 and 4 K for the ICI-04 and ICI-11 data. The polarisation difference is due to the impact of both the surface and hydrometeors impact on the "observation" for the ICI-4 channels, and primarily only due to hydrometeors for ICI-11 as atmospheric transmission is close to 0. Both the surface and hydrometeor contribution are modelled using state-of-the-art models /data as described in the Task2 report, and the inclusion of H and V data in the current database is a main improvement of the current database.

The 0.01 contour lines (the thickest lines) are shifted by a few K, indicating a bias between clearsky Tbs of the current and previous database. The simulation of the current database is more detailed w.r.t. channel response functions, and spectroscopy than the previous one, and hence considered to be a second main improvement of the current database.

4 ICI Retrieval Performance

The retrieval performance simulation is done by randomly picking out one million states from the extended current retrieval database with a true and known IWP above 0.0001 kg/m^2, adding noise to the simulated observation according to the $NE\Delta T$ specification of the ICI channels, and then applying the BMCI retrieval algorithm.

4.1 Comparison with previous database

The previous database lacks separated simulations for the dual polarised ICI-04 and ICI-11 channels, and the mean of the H and V data was then used for the retrieval using the previous database. Figure 18 shows the retrieved IWP as function of the true IWP, and Figure 19 shows the true and retrieved probability density function (PDF) of IWP, for mixed climatic conditions, and for the retrieval based on both the current and previous database. It is noted that the retrieval performance varies with climatic condition, but here we focus on the "average" performance, and more details is given in Section 4.2. The difference between the retrieved posterior median values and the true IWP is smaller than 50 % for an IWP greater than about $0.02\,kg/m^2$ for the retrieval based on the current database. The retrieval based on the current database provides a significant lower uncertainty estimate than the one based on the previous database above this value and up to about $1\,kg/m^2$. For example, the upper percentile is about 60 % lower for an IWP above $0.03\,kg/m^2$ and the lower percentile is more close to the true IWP, for the retrieval based on the current database compared to the one using the previous database.

The median retrieval is biased low compared to the true IWP at IWPs above about $3\,kg/m^2$ and for both of the two retrievals. The difference between the upper and lower percentiles is somewhat smaller for the retrieval based on the previous database. At high IWPs the retrieval can be significantly influenced by the *a priori* IWP distribution and the previous database contains more cases with high IWP than the current database (see Figure 6), that effectively results in a lower retrieval uncertainty at high IWP for the retrieval based on the previous database. The number of cases with high IWPs in the previous database is probably overestimated with respect the true distribution, and the retrieval works relatively well for cases with high IWPs, but the "cost" of this is likely also a higher retrieval uncertainty for cases with relatively low IWP values.

It is kind of expected that the retrieval based on the current database provides a better retrieval performance, since the test data are selected from this database, and the purpose of this test is mainly to confirm that this is actually the case. In addition to this, the current database is an improved version of the previous database and contains both more information (H and V data) and more realistic data such as the IWP apriori distribution that matches DARDAR data (Section 3.1). The latter was achieved by including a revised set of hydrometeor models compared to what was used for the previous database. The accuracy and uncertainty estimate depends to a large extent on the variability of the hydrometeor models included in the databases, and also on the ability of the retrieval to discriminate between the models. The previous database contains a somewhat higher hydrometeor model spread than the current database. In addition, the current database contains H and V information, and it is likely that the retrieval can use this information, particularly for cases with low and moderate high IWPs, to identify some particle model as likely or unlikely given the observation, and hence decrease the retrieval uncertainty compared to that of the previous database.

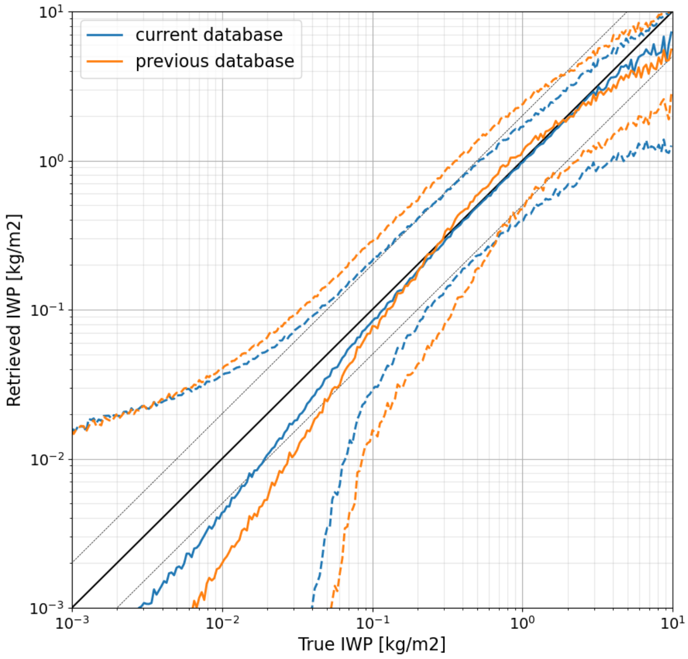

Figure 18: Estimated retrieval performance for IWP using the current and previous database and for mixed climatic conditions. The solid lines show the median of the retrieved posterior median values, the colored dashed lines show the median of the retrieved 5th and 95th percentile.

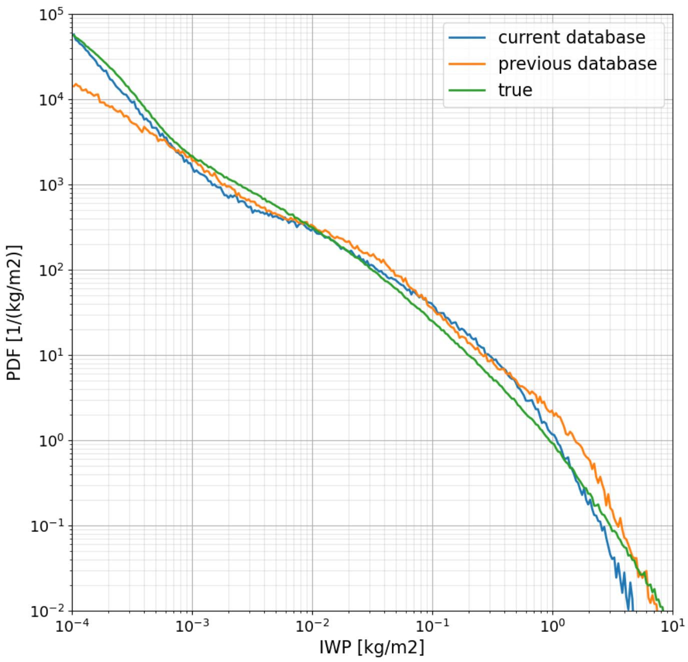

Figure 19: Retrieved (posterior median value) and true probability density function of IWP, for the data displayed in Figure 18.

4.2 Test retrievals with new database

We here consider statistics derived from the retrieved 5th, 50th, and 95th percentile of IWP, Dmean, and Zmean, and only considering the current database, and for five climatic regions, defined as:

- low-latitude: latitude within $30\,°S$ – $30\,°N$
- warm mid-latitude: latitude within $60\,°S$ – $30\,°S$ or $30\,°N$ – $60\,°N$ and surface temperature above $10\,°C$
- cold mid-latitude: latitude within $60\,°S$ – $30\,°S$ or $30\,°N$ – $60\,°N$ and surface temperature below $0\,°C$
- temperate mid-latitude: latitude within $60\,°S$ – $30\,°S$ or $30\,°N$ – $60\,°N$ and surface temperature between $0\,°C$ – $10\,°C$
- high-latitude: latitude greater than $60\,°S$ or $60\,°N$

4.2.1 IWP detection limit

Figure 20 shows estimated retrieval performance for IWP and five "climatic regions". The lower (5th percentile) and upper (95th percentile) dashed lines can be interpreted as the level where it is a 95 and 5 % probability, respectively, that the IWP is above this level. For low true IWP values (i.e. below 0.001 kg/m^2) the gradient of the log-log plot of these levels as function of true IWP is low, and the values reflect the IWP *a priori* distribution as the sensitivity is low. For high true IWP values (i.e. above 0.3 kg/m^2) the gradient of the percentiles curves is close to unity and this indicates a high sensitivity to IWP. The detection limit is clearly located within the range of 0.001 to 0.3 kg/m^2, and one way to define this limit/threshold is to identify where the gradient of the 5th percentile curve (lower curve) is at a maximum, and this varies with "climatic region" according to:

- low-latitude: $\sim$0.015 kg/m^2,
- warm mid-latitude: $\sim$0.02 kg/m^2,
- temperate mid-latitude: $\sim$0.045 kg/m^2,
- cold mid-latitude: $\sim$0.055 kg/m^2,
- high-latitude: $\sim$0.09 kg/m^2.

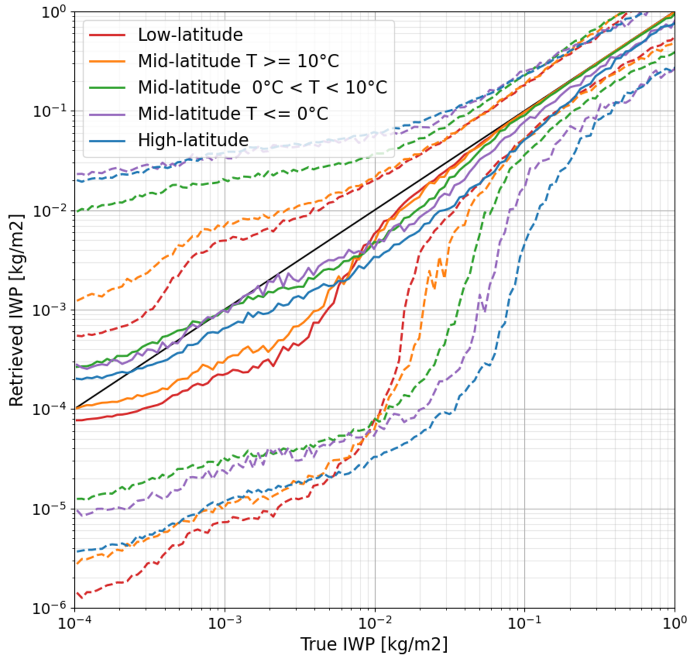

Figure 20: Estimated retrieval performance for IWP and five climatic regions. The solid lines show the median of the retrieved posterior median values, and the dashed lines show the median of the retrieved 5th and 95th percentile.

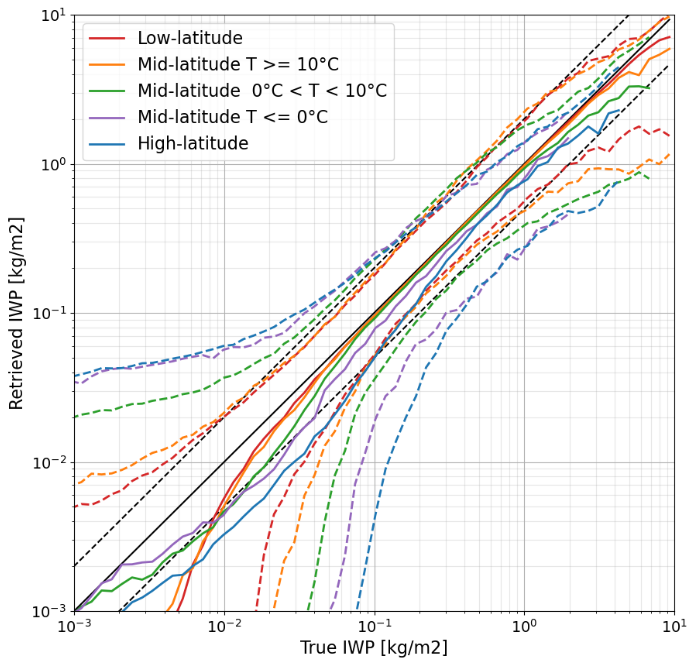

Figure 21: As Figure 20, but showing the data for a somewhat different IWP range.

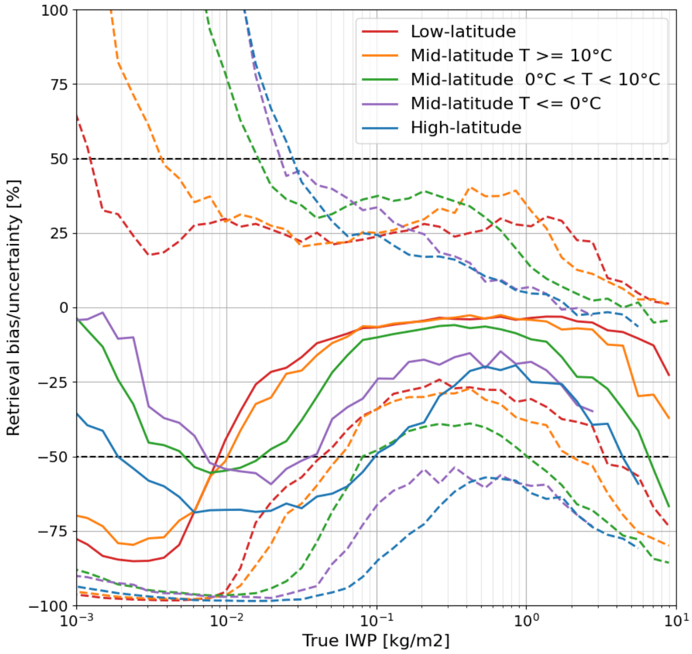

Figure 22: As Figure 21 but showing the performance in terms of a relative error and the dashed lines show data associated to the retrieved 16th and 84th percentile.

4.2.2 IWP retrieval performance

Figure 21 shows estimated retrieval performance for IWP and five "climatic regions". It can clearly be noted that the retrieval accuracy or performance vary with climatic region. Best performance is found for the tropics and warm mid-latitudes, and worst performance is found for the high-latitudes and cold mid-latitudes. The obtained performance is in line with results reported in Eriksson et al. (2020) for the tropics and mid-latitudes, but we here also include estimated performance at high-latitudes for the first time (as far as we know). The performance is summarised below:

- low-latitude and warm mid-latitude: High accuracy in retrieved median values for IWP above $\sim$0.02 kg/m^2. The median of the retrieved 5th and 95th percentiles is located within 0.5 and 2 times the retrieved or true IWP value, and hence the upper uncertainty is on average greater than the lower uncertainty level in absolute terms.

- temperate mid-latitude: High accuracy in retrieved median values for IWP within the range of $\sim$0.05 to 1 kg/m^2. The median of the retrieved 5th and 95th percentiles is here mainly located within 0.5 and 2 times the retrieved or true IWP value. For IWP values above 1 kg/m^2 the retrieved median values are on average a bit lower than the true IWP, i.e. the sensitivity drops off and the lower and upper uncertainty levels increases and decreases, respectively, and hence the retrieval becomes significantly influenced by the *a priori* distribution.

- cold mid-latitude and high-latitude: The accuracy in retrieved median values is within 50 % for IWP values above $\sim$0.06 to 0.1 kg/m^2 for the cold mid-latitudes and high-latitudes, respectively. The median of the retrieved median values is systematically a bit underestimated compared to the true IWP. The lower uncertainty level is significantly greater for these climatic regions compared to the others. The median of the retrieved 5th and 95th percentiles is located within about 0.3 and 2 times the retrieved or true IWP value for true IWP above $0.3\,kg/m^2$.

The end-user requirement for IWP is a 50% accuracy, according to Section 6.1 in RD-4 (EURD), and without any further description about how the accuracy is defined. The estimated retrieval bias is below 50% for an IWP above $0.01\,kg/m^2$ and $0.1\,kg/m^2$ for low and high latitudes, respectively. On the other hand, the retrieval uncertainty presented in Figure 21 is in general greater than 50% and for all conditions and for the full IWP range. However, it is reminded that the uncertainty is reported as the posterior 5th and 95th percentiles, and the distance between these roughly corresponds to a $\pm$ two standard deviations uncertainty estimate and for a normal distribution. Figure 22 shows the estimated retrieval performance as a relative measure, and for the retrieved 16th and 84th percentiles, corresponding to $\pm\,1$ standard deviations. The estimated retrieval performance is then close to be in compliance with the end-user requirement, except for high-latitudes and cold mid-latitudes conditions and high IWP cases in general, where the lower percentile tends to be found at a level that is more than 50% lower than the true IWP value.

The main reason for the poorer performance at high-latitudes and cold mid-latitudes is that the cloud ice mass is here located closer to the surface, resulting in that the cloud ice impact on the observation is relatively low and the cloud ice is potentially not even sensed by some of channels at high frequency (≥ 448 GHz) and the inner 183 and 325 GHz channels. The channels that actually have sensitivity to the complete cloud ice column will consequently also be sensitive to the surface, and the surface contribution to the observation varies with frequency and is not perfectly known, and effectively adds an extra retrieval uncertainty compared to retrieval at e.g. low latitudes. Figure 23 shows estimated retrieval performance for IWP and for cold surface temperature and for five different surface types. The best performance is clearly obtained over the ocean, where the lower percentile is significantly closer to the true IWP than for the other surface types. For example, an IWP of $0.5\,kg/m^2$ is clearly detected over all surface types, but the lower percentile is located around $0.25\,kg/m^2$ and $0.1\,kg/m^2$ for ocean and the snow/ice/mixed surface types, respectively.

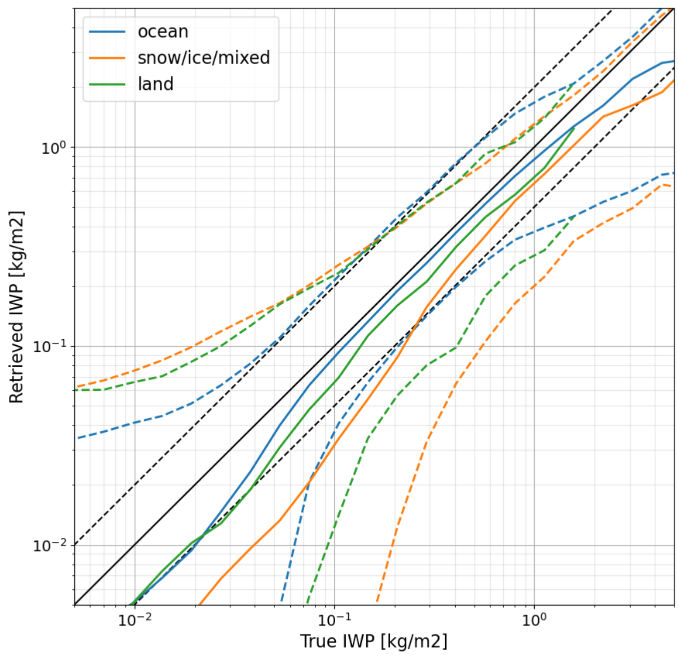

Figure 23: Estimated retrieval performance for IWP and for a surface temperature lower than 0 °C and for different surface types. The performance for snow ice, and mixed surface types were similar and combined to result in a less noisy plot. The solid lines show the median of the retrieved posterior median values, and the dashed lines show the median of the retrieved 5th and 95th percentile.

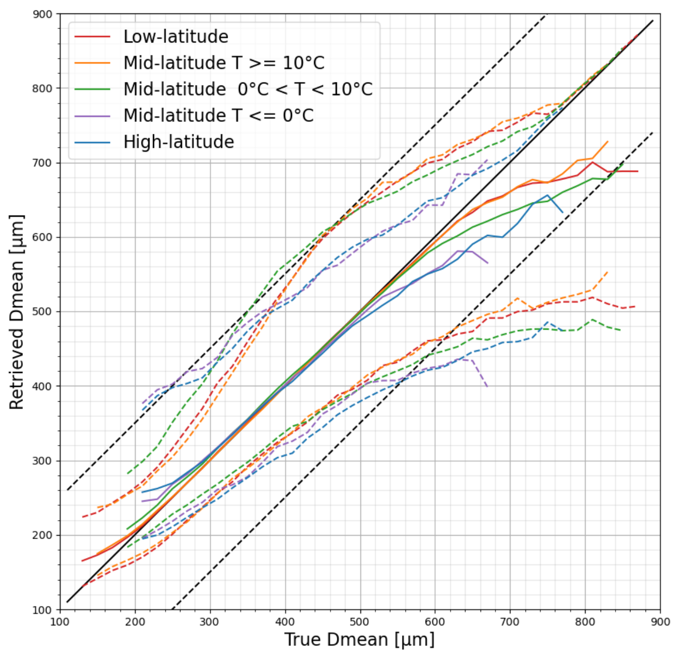

Figure 24: Estimated retrieval performance for Dmean and five climatic regions. The solid lines show the median of the retrieved posterior median values, and the dashed lines show the median of the retrieved 5th and 95th percentile. Only data with a true IWP above 0.02 kg/m^2 are used for low-latitude and warm mid-latitude, above 0.05 kg/m^2 for intermediate warm mid-latitude, above 0.06 kg/m^2 for cold mid-latitude, and above 0.1 kg/m^2 for high-latitude.

4.2.3 Dmean retrieval performance

Figure 24 shows estimated retrieval performance for Dmean and five "climatic regions". The accuracy of the retrieved median Dmean as function of the true Dmean is mainly better than 20 μm, and the upper and lower percentiles are found within 150 μm from the true Dmean, within the Dmean range of 100 – 550 μm, and for all "climatic regions". For a Dmean greater than 550 μm the median Dmean retrieval is biased low compared to the true Dmean.

For e.g. the low-latitudes, the difference between the upper and lower percentiles is significantly smaller for small Dmean than for large Dmean values (i.e 100 and 250 μm around Dmean values of 100 and 250 μm, respectively). The lower and upper percentiles are naturally bounded by the *a priori* distribution, for low and high Dmean values, respectively, but the upper percentile at low Dmean values is located closer to the true Dmean value than the lower percentile at high Dmean values. Low and high Dmean values are generally associated to cloud ice mass at high and low altitudes, respectively. ICI provide a more complete observation for cloud ice mass at high altitude than at low altitude as all channels can then be sensitive to the cloud ice mass. Thus, the ICI observation provides more information at the same time as the Dmean *a priori* distribution is more narrow for cloud ice mass at high altitudes, and consequently the Dmean retrieval uncertainty is lower at low Dmean than at high Dmean values.

However, it can be noted that the difference between the upper and lower percentiles around small and large Dmean values is similiar to each other (200–250 μm) for the cold mid-latitudes and high-latitudes. The Dmean retrieval performance is poorer for the cold mid-latitudes and high-latitudes than for the low-latitudes and for small Dmean values. At the higher latitudes the cloud ice mass is located at a lower altitude than at low latitudes, and the cloud ice mass is consequently not "sensed" by all channels of ICI, and the information in the observation is less complete. However, the Dmean retrieval uncertainty is lower for large Dmean values and for high latitudes compared to that of low latitudes. This can be explained by a more narrow Dmean *a priori* distribution at the higher latitudes, compared to that of lower latitudes.

4.2.4 Zmean retrieval performance

Figure 25 shows estimated retrieval performance for Zmean and five "climatic regions". The retrieval performance clearly varies with climatic region as the cloud ice mass appears at different altitudes. The accuracy of the retrieved median values are within 500 m within the altitude range of 0.5 –12 km for all regions. The lower and upper percentiles are naturally bounded by the *a priori* distribution, for low and high Zmean values, respectively. The retrieved 5th and 95th percentiles are found within ± 2 km from the true Zmean for most altitudes and regions, but with some variations and deviations.

- low-latitude and warm mid-latitude: The retrieved 5th and 95th percentiles are primarily found within ± 1 km from the true Zmean, except for the lowermost and uppermost 1.5 km where cloud ice mass occur, where the upper and lower level, respectively increases to about 2 km.
- cold mid-latitude, and high-latitude: The difference between the upper and lower percentile is about 2.5 km within the altitude range of 2 –6 km, and the slope of the percentile is here lower than the one-to-one line.

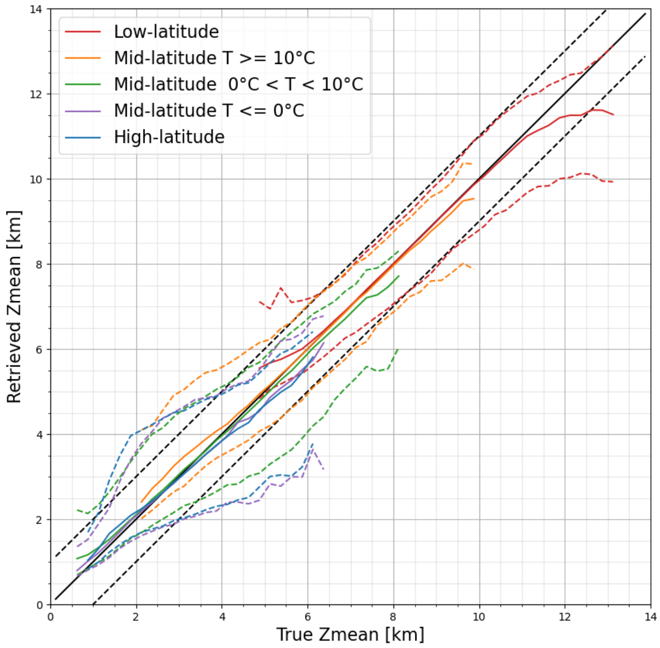

Figure 25: Estimated retrieval performance for Zmean and five climatic regions. The solid lines show the median of the retrieved posterior median values, and the dashed lines show the median of the retrieved 5th and 95th percentile. Only data with a true IWP above 0.02 kg/m^2 are used for low-latitude and warm mid-latitude, above 0.05 kg/m^2 for intermediate warm mid-latitude, above 0.06 kg/m^2 for cold mid-latitude, and above 0.1 kg/m^2 for high-latitude.

5 Summary

A retrieval database to be applied for the ICI level2 product processing at the EUMETSAT Central Facilities has been developed, validated against available observations, and compared to a preliminary database.

The database produced contains 9.5 million cases, where each individual case consists of a simulated ICI measurement and a corresponding set of atmospheric and surface states. The methodology applied for generating the database is inspired by the one used for developing the preliminary database [RD-2, RD-3], but resolves some of the main shortcomings of the previous database. The improvements include incorporation of full spectral response functions, interference of ozone, two-dimensional variation of brightness temperatures within the footprint, and orientation of ice hydrometeors (Kaur, Eriksson, Rydberg, May and Hallborn 2022).

The most important aspect of the database is that it shall statistically represent reality. The clear-sky atmospheric state and dynamic surface data are taken from ERA5, while the spatial distribution of clouds and precipitation are based on CloudSat reflectivities. The ERA5 and CloudSat inputs ensure that the database on an overall level follow reality, but do not fully constrain the radiative transfer needed to simulate ICI radiances. For example, several microphysical quantities need to be added, such as particle shape (habit) and sizes (i.e. particle size distribution. PSD). Already the preliminary database operated with multiple particle models, to represent the variability in ice hydrometeor microphysics. The particle models were for the new database version fully revised and an occurrence fraction has been introduced. These changes gave a better agreement in mean IWP to DARDAR retrievals, especially for tropical latitudes (Sec. 3.1). Accordingly, the IWP implied by the database constitutes now a better a priori distribution for the retrievals.

A completely new feature is to consider particle orientation, following Barlakas et al. (2021). In lack of constraint from CloudSat, the factor describing degree of orientation is selected randomly (Kaur, Eriksson, Barlakas, Pfreundschuh and Fox 2022). As a complement, dedicated calculations for V and H polarisation have been performed. The effect of including polarisation is obvious for the dual polarisation channels (4V/H and 11V/H), but it is stressed that this signifies an improvement also for the single polarisation channels. Also the surface generates polarisation and a rough polarised emissivity model for sea ice and snow was developed, in lack of any existing model to use.

For validation purposes, also ISMAR/MARSS and GMI observations were simulated. It was found that simulated data cover the measurement space as observed by ISMAR and MARSS, including data from the dual polarised 243 and 664 GHz channels, that were missing in the previous database. This gives confidence in that the ICI database has good quality at least for northern mid-latitude conditions.

GMI observations are limited to frequencies below 200 GHz, but have close to global coverage. The simulated GMI data match the real observations relatively closely for all four channels considered. The main exception is measurements over snow and sea ice. The surface classification involving sea ice and snow is uncertain and there are relatively few simulations including these surface types so the interpretation of the noted deviations is not straightforward but it could be the case that the emissivity models created should be revised. In any case, the simulated polarisation differences above ocean and land agree well with observations. As remarked above, the preliminary database did not represent polarisation differences at all. Another improvement compared to the preliminary database was noted for 183 ± 3 GHz. This a channel with very low surface influence and the improved fit with GMI for antenna temperatures around 150 K (Fig. 8) gives further confirmation on that the new set of particle model is more realistic.

Inversion of real GMI observations was performed as a full scale test of the database generation. These retrievals gave realistic spatial distributions and zonal mean IWP in good agreement with the DARDAR dataset.

Retrieval simulations for ICI show that the performance vary quite significantly with climatological condition. Best performance is obtained when the surface temperature is warm ($> 10°$C) or the cloud ice mass is found well above the surface. The IWP detection limit is then around 0.02 kg/m^2, and the retrieved 5th and 95th percentile is located about 50% and 100% below and above, respectively, the 50th percentile (median) or the true value. The accuracy in retrieved Dmean and Zmean values are high within the range of $100 - 550\,\mu m$ and $2 - 12$ km, respectively, and with uncertainties of about 150 μm and 1 km, respectively. These performance estimates are in line with result reported in Eriksson et al. (2020).

The performance for high-latitudes is estimated for the first time, and the performance is effectively decreasing with decreasing surface temperature or cloud ice mass height. For surface temperature well below 0°C the IWP detection limit is almost five times greater (0.1 kg/m^2) than that at warm temperatures, and the retrieved 5th percentile is further away from the median value than at warmer temperatures.

References

Aires, F., Prigent, C., Bernardo, F., Jiménez, C., Saunders, R. and Brunel, P.: 2011, A tool to estimate land-surface emissivities at microwave frequencies (telsem) for use in numerical weather prediction, *Q. J. R. Meteorol. Soc.* **137**(656), 690–699.

Barlakas, V., Geer, A. J. and Eriksson, P.: 2021, Introducing hydrometeor orientation into all-sky microwave and submillimeter assimilatifbrathon, *Atmos. Meas. Tech.* **14**(5), 3427–3447.

Delanoë, J., Heymsfield, A. J., Protat, A., Bansemer, A. and Hogan, R.: 2014, Normalized particle size distribution for remote sensing application, *J. Geophys. Res. Atmos.* **119**(7), 4204–4227.

Eliasson, S., Buehler, S. A., Milz, M., Eriksson, P. and John, V. O.: 2011, Assessing observed and modelled spatial distributions of ice water path using satellite data, *Atmos. Chem. Phys.* **11**(1), 375–391.

Eriksson, P., Rydberg, B., Mattioli, V., Thoss, A., Accadia, C., Klein, U. and Buehler, S. A.: 2020, Towards an operational ice cloud imager (ICI) retrieval product, *Atmos. Meas. Tech.* **13**(1), 53–71.

Fox, S., Lee, C., Moyna, B., Philipp, M., Rule, I., Rogers, S., King, R., Oldfield, M., Rea, S., Henry, M., Wang, H. and Harlow, R. C.: 2017, ISMAR: an airborne submillimetre radiometer, *Atmos. Meas. Tech.* **10**(2), 477–490.

Heymsfield, A., Protat, A., Austin, R., Bouniol, D., Hogan, R., Delanoë, J., Okamoto, H., Sato, K., Zadelhoff, G.-J., Donovan, D. and Wang, Z.: 2008, Testing iwc retrieval methods using radar and ancillary measurements with in situ data, *J. Appl. Meteorol. Climatol.* **47**, 135–163.

Kaur, I., Eriksson, P., Barlakas, V., Pfreundschuh, S. and Fox, S.: 2022, Fast radiative transfer approximating ice hydrometeor orientation and its implication on IWP retrievals, *Remote sensing* **14**(7).

Kaur, I., Eriksson, P. and Rydberg, B.: 2022, Literature review to support the development of a cloud radiation database for eps-sg ici iwp retrieval, *Technical report*, Department of Space, Earth and Environemnt, Chalmers University of Technology.

Kaur, I., Eriksson, P., Rydberg, B., May, E. and Hallborn, H.: 2022, Development of a cloud radiation database for eps-sg ici: Task 2 report, *Technical report*, Department of Space, Earth and Environment, Chalmers University of Technology.

McGrath, A. and Hewison, T.: 2001, Measuring the accuracy of MARSS - an airborne microwave radiometer, *J. Atmos. Oceanic Technol.* **18**(12), 2003–2012.

Pfreundschuh, S., Eriksson, P., Duncan, D., Rydberg, B., Håkansson, N. and Thoss, A.: 2018, A neural network approach to estimating a posteriori distributions of bayesian retrieval problems, *Atmospheric Measurement Techniques* **11**(8), 4627–4643.

Appendices

A Retrieval Database Size

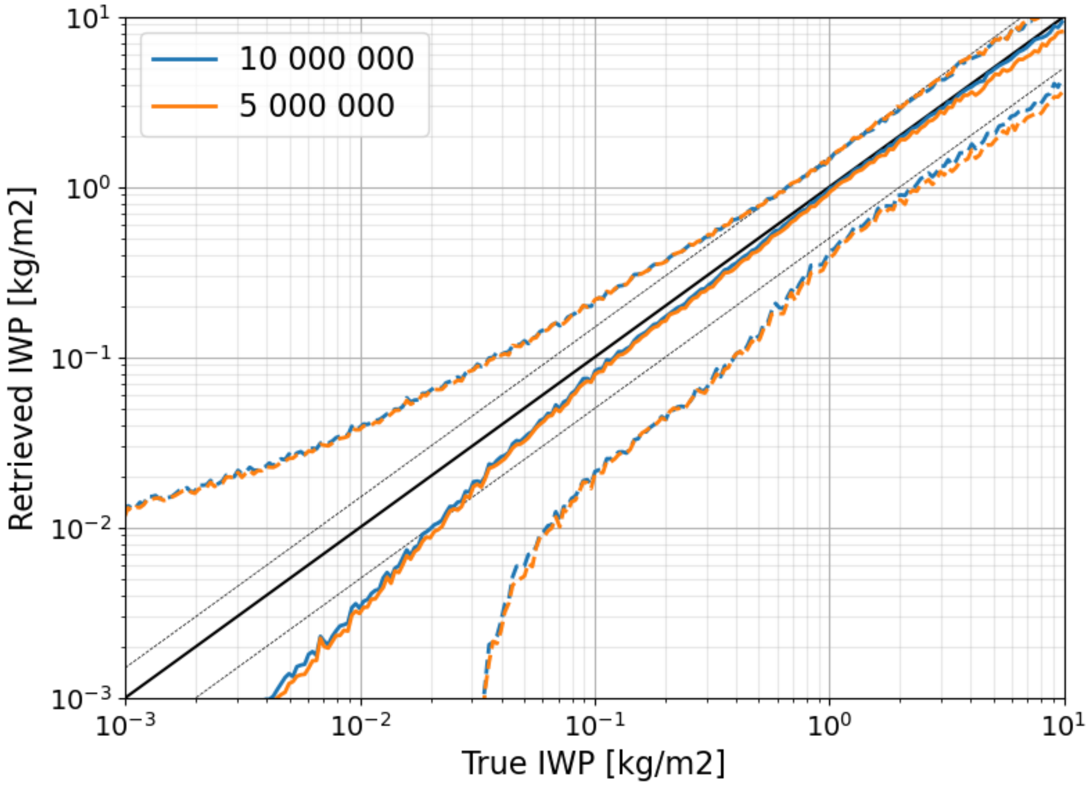

Figure 26: Estimated retrieval performance for IWP using a retrieval database consisting of around 10 and 5 million states. The solid lines show the median of the retrieved posterior median values, and the dashed lines show the median of the retrived 5th and 95th percentile.

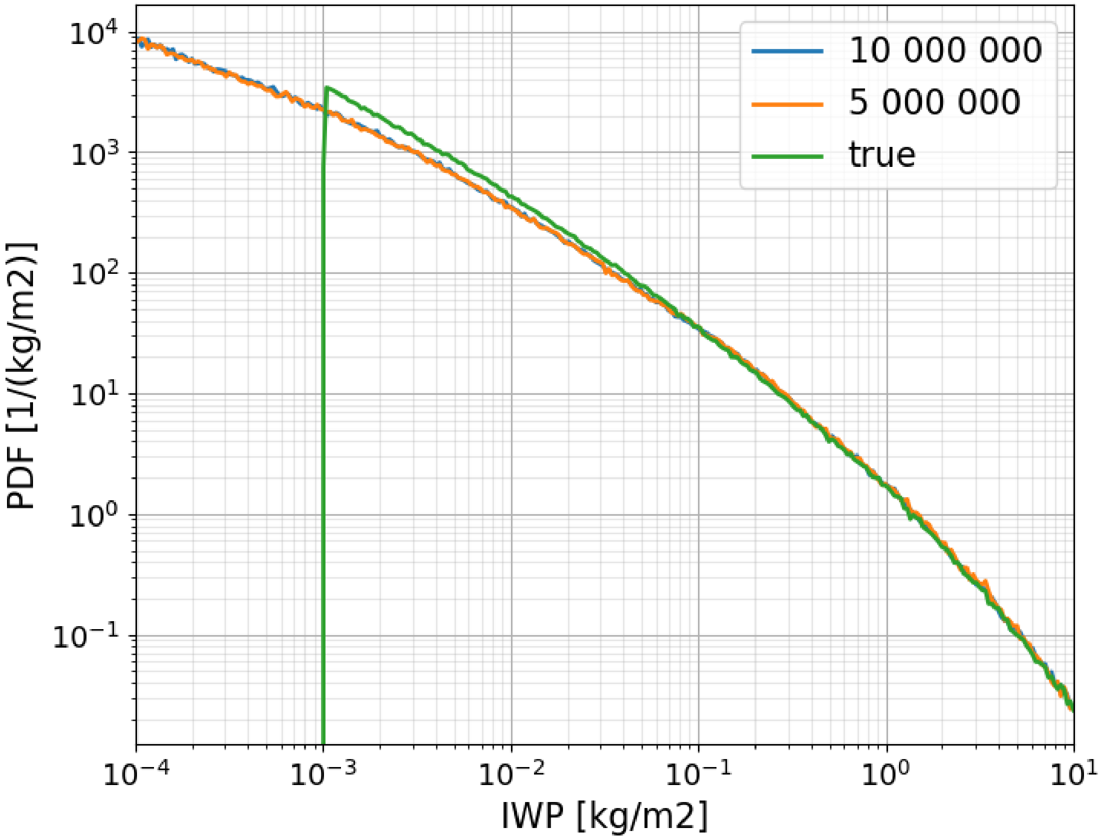

Figure 27: Retrieved (posterior median value) and true probability density function of IWP, for the data displayed in Figure 26.

The computational cost for generating a database of many millions of states is high, and here we make a test to check the difference in retrieval performance by using a retrieval database of 5 and 10 million states. For this exercise we use the preliminary ICI retrieval database, and pick out 500 000 states with an IWP greater than 0.001 kg/m^2 and run the BMCI retrieval algorithm for the two databases (the smaller database is a subset of the larger one).

Figure 26 shows the retrieved IWP as function of the true IWP, and Figure 27 shows the true and retrieved probability density function (PDF) of IWP. The retrieval performance is found to be very similar for the two databases. The retrieved PDFs are more or less located on top of each other (Figure 27), and for both cases the occurrence frequency of low IWP values are underestimated, as the sensitivity is low for low IWP values. This is also clearly seen in the retrieved vs. true IWP plot (Figure 26). The accuracy of the median retrieved IWP seems to be a bit better for IWP greater than $5 kg/m^2$ for the 10 million database, but uncertainties are of quite similar magnitude, so no big difference.

Thus, the result presented here indicates that the retrieval performance will, in practise, be close to identical for a database of 5 and 10 million states, and a database size of 5 million should therefore be sufficient to use.

B Latitude distribution of GMI simulation database and GMI observations

The latitude distributions of observed and simulated data are plotted as probability distributions in Figure 28.

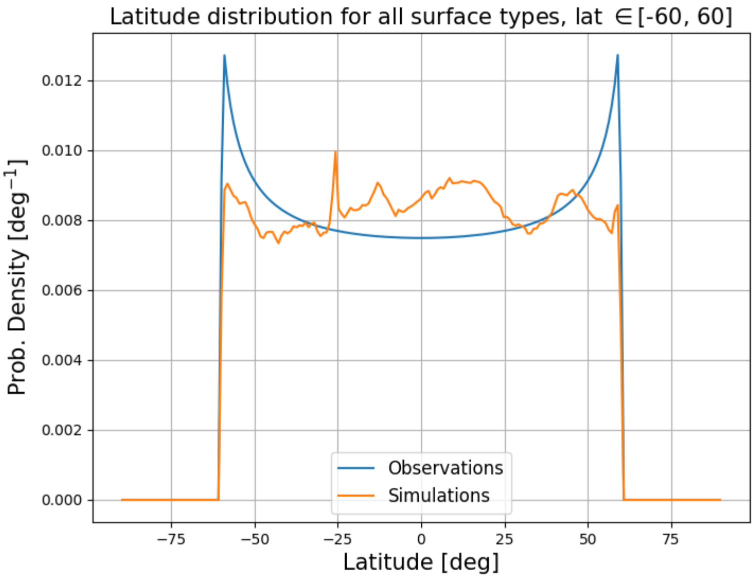

Figure 28: Latitude probability distributions for all surface types and latitudes in the range -60 degrees to 60 degrees.

C Division of GMI distributions according to latitude region and surface type

This section provides complementary figures for the assessment of the GMI validation database discussed in Sec. 3.2. Brightness temperature distributions for different surface types and different latitude regions are presented in Sec. C.1 and polarisation difference distributions for different surface types and different latitude regions are presented in Sec. C.2. The surface type classifications that were used are listed in Table 2.

Surface type	**GPROF classifications**	**Simulation clssifications**
land	Decreasing vegetation 3-7	LAND
ocean	Ocean	OCEAN
snow	Decreasing snow cover 8-11	SNOW, OCEAN_AND_SNOW, LAND_AND_SNOW
sea-ice	Sea-ice	SEAICE, OCEAN_ADN_SEAICE, LAND _AND_SEAICE, SEAICE_MIXED

Table 2: Surface classifications

C.1 Brightness temperature distributions for different surface types and latitude regions

Figure 29 shows the T_b distribution for all surface types at tropical latitudes where only data from latitudes -30 to 30 are included. The T_b distributions for land at tropical latitudes is seen in Figure 30 and for ocean at tropical latitudes is seen in Figure 31. Figure 32 and Figure 33 shows the T_b distributions for land and ocean respectively at mid-latitudes between -60 to -30 and 30 to 60 degrees. Figure 34 shows T_b distributions for snow covered land surfaces and Figure 35 shows T_b distributions for sea-ice. Both these figures are for northern latitudes between 50 and 70 degrees.

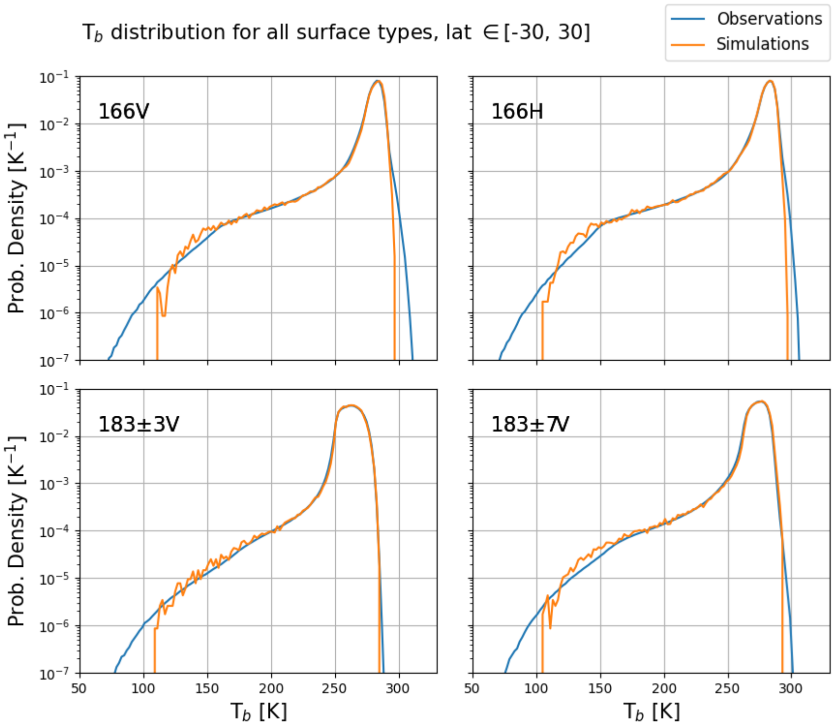

Figure 29: Brightness temperature (T_b) probability distributions for each of the high-frequency channels of GMI for all surface types at tropical latitudes, i.e. only data from latitudes -30 to 30 are included.

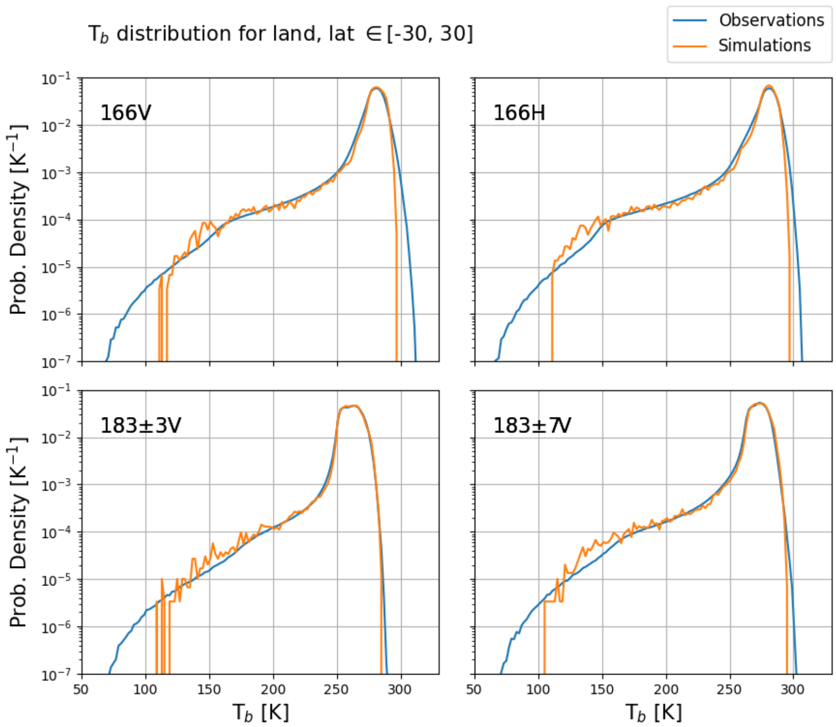

Figure 30: Brightness temperature (T_b) probability distributions for each of the high-frequency channels of GMI for land at tropical latitudes, i.e. only data from latitudes -30 to 30 are included.

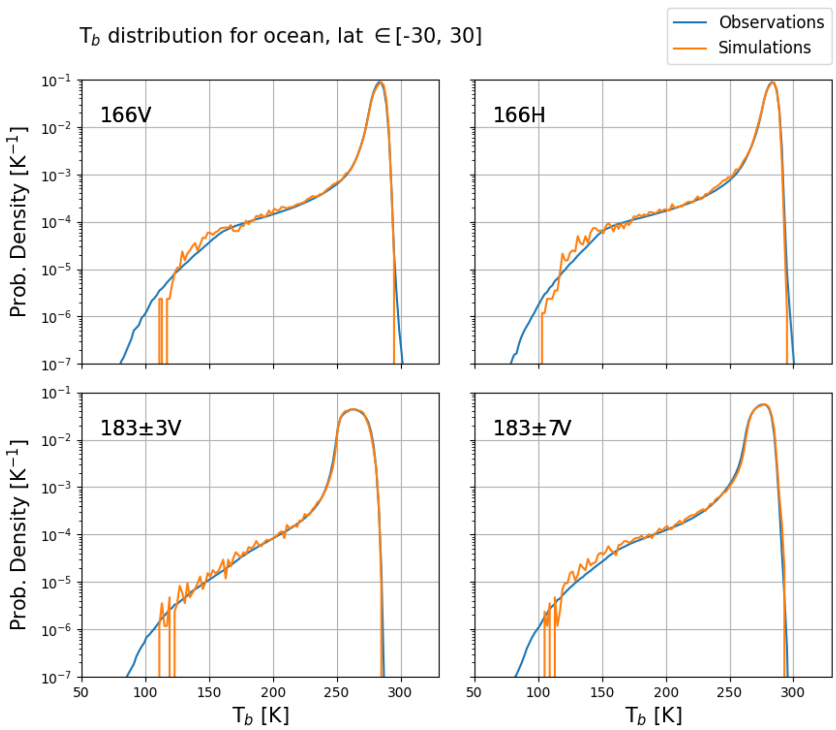

Figure 31: Brightness temperature (T_b) probability distributions for each of the high-frequency channels of GMI for ocean at tropical latitudes, i.e. only data from latitudes -30 to 30 are included.

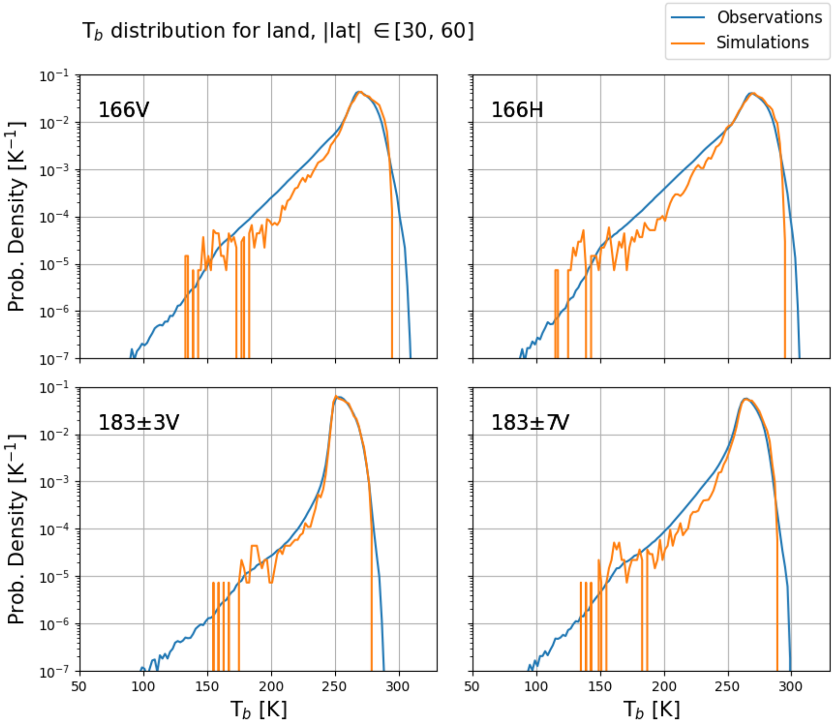

Figure 32: Brightness temperature (T_b) probability distributions for each of the high-frequency channels of GMI for land at mid-latitudes i.e. only data from latitudes -60 to -30 and 30 to 60 are included.

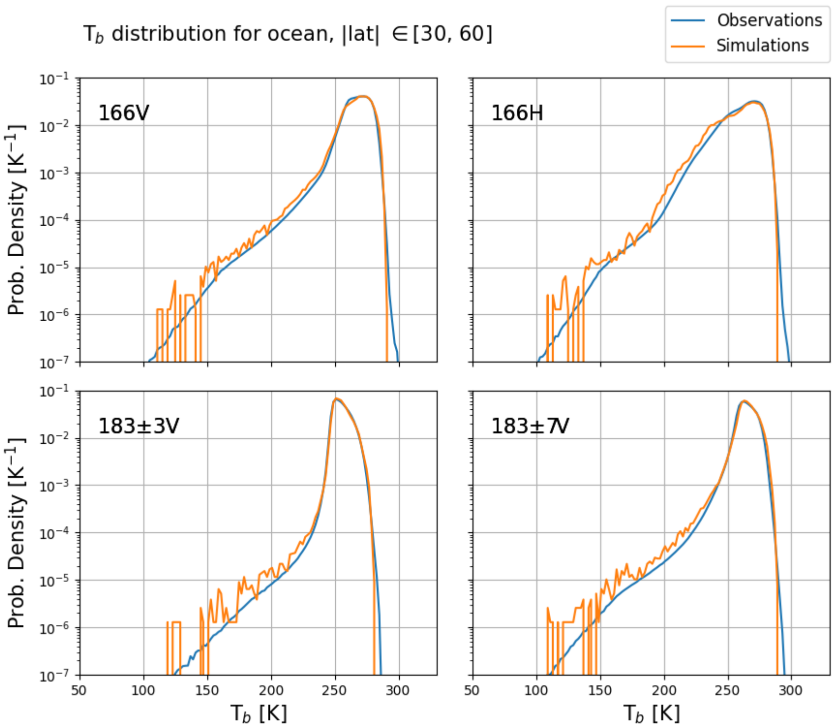

Figure 33: Brightness temperature (T_b) probability distributions for each of the high-frequency channels of GMI for ocean at mid-latitudes i.e. only data from latitudes -60 to -30 and 30 to 60 are included.

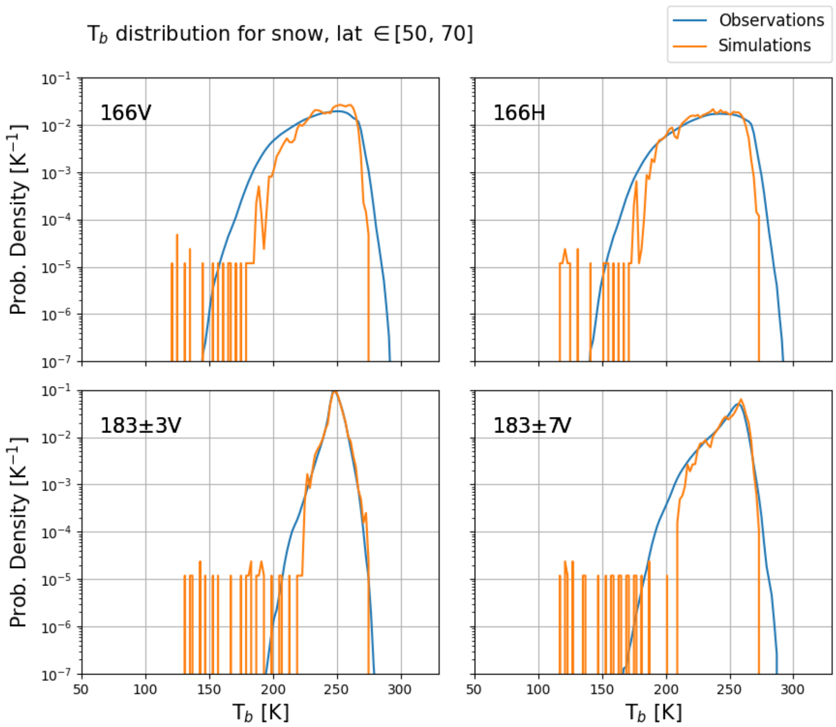

Figure 34: Brightness temperature (T_b) probability distributions for each of the high-frequency channels of GMI for snow at northern latitudes i.e. only data from latitudes 50 to 60 are included. There are only about 42 000 cases in the simulation dataset.

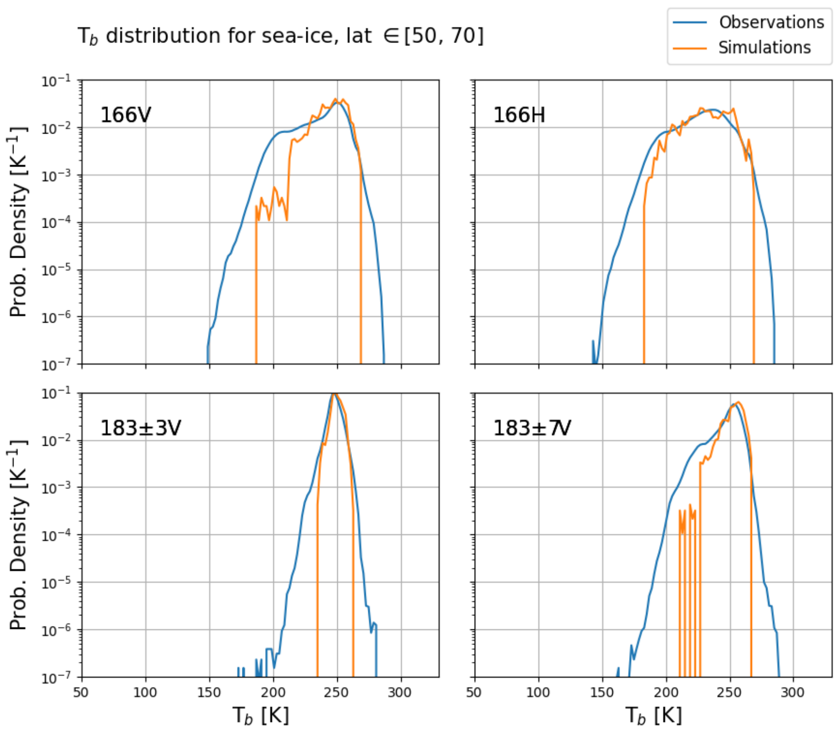

Figure 35: Brightness temperature (T_b) probability distributions for each of the high-frequency channels of GMI for sea-ice at northern latitudes i.e. only data from latitudes 50 to 60 are included. There are only about 5000 cases in the simulation dataset.

C.2 Polarisation difference distributions for different surface types and latitude regions.

PD distributions for land and ocean are seen in Figure 36 for tropical latitudes and in Figure 37 for mid-latitudes. Figure 38 shows PD distributions for snow covered land and sea-ice at northern latitudes.

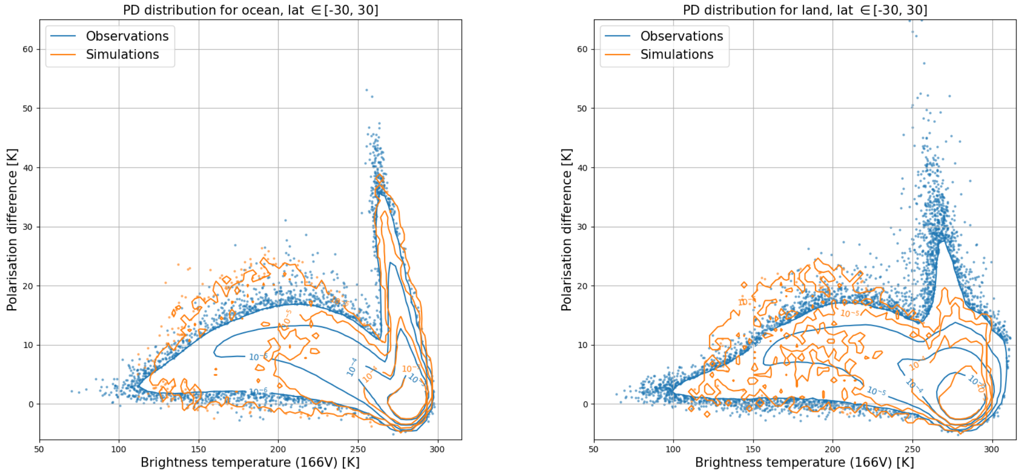

Figure 36: Joint probability distributions of polarisation difference (PD) and brightness temperature for the 166V channel ($T_{b,166V}$) for ocean (left) and land (right) at tropical latitudes in the range -30 degrees to 30 degrees. Contour lines of the probability distributions for observations (blue) and simulations (orange) are seen at levels 10^{-6} K^{-2}, 10^{-5} K^{-2}, 10^{-4} K^{-2} and 10^{-3} K^{-2}. A random sample of 1% of the simulated data points with lower probability than 10^{-6} K^{-2} are plotted as blue dots. For the simulations all points with lower probability than 10^{-6} K^{-2} are plotted as orange dots.

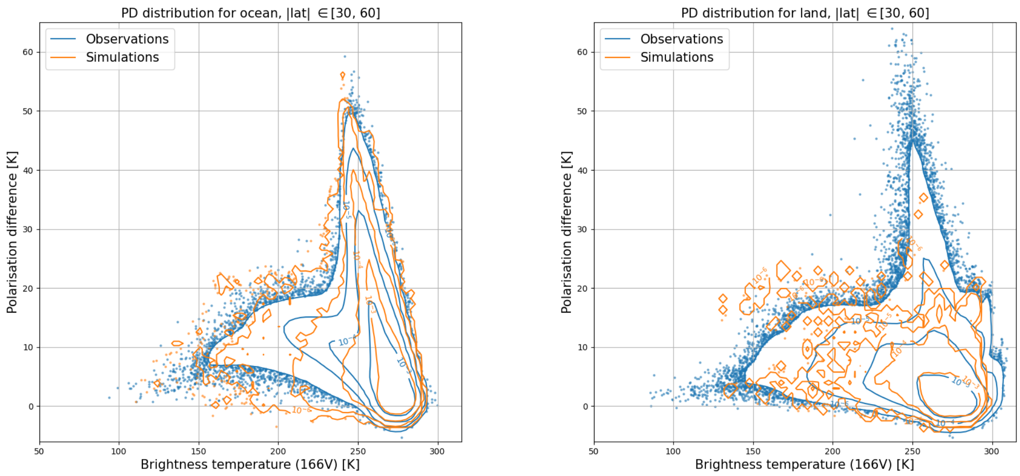

Figure 37: Joint probability distributions of polarisation difference (PD) and brightness temperature for the 166V channel ($T_{b,166V}$) for ocean (left) and land (right) at mid-latitudes in the range -60 to -30 and 30 to 60 degrees. Contour lines of the probability distributions for observations (blue) and simulations (orange) are seen at levels 10^{-6} K^{-2}, 10^{-5} K^{-2}, 10^{-4} K^{-2} and 10^{-3} K^{-2}. A random sample of 1% of the simulated data points with lower probability than 10^{-6} K^{-2} are plotted as blue dots. For the simulations all points with lower probability than 10^{-6} K^{-2} are plotted as orange dots.

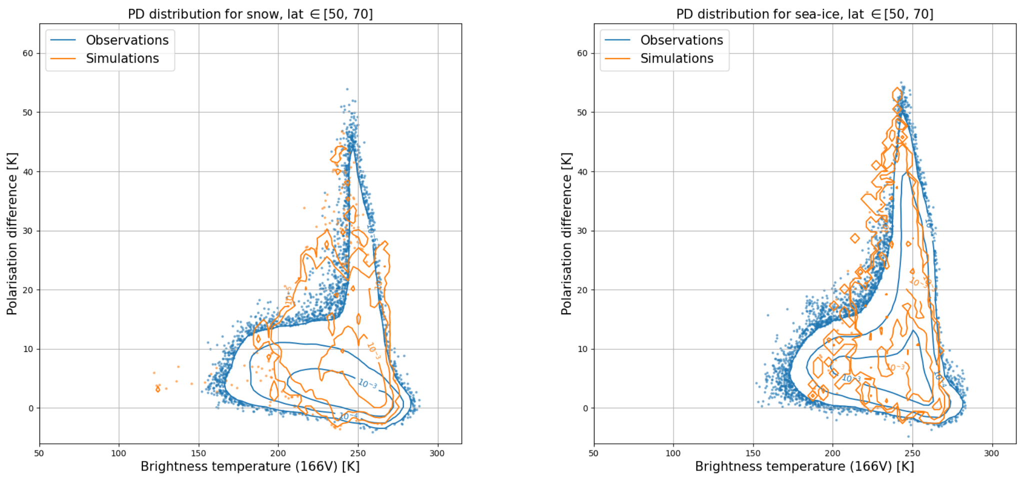

Figure 38: Joint probability distributions of polarisation difference (PD) and brightness temperature for the 166V channel ($T_{b,166V}$) for snow (left) and sea-ice (right) at northern-latitudes in the range 50 to 70 degrees. Contour lines of the probability distributions for observations (blue) and simulations (orange) are seen at levels 10^{-5} K^{-2}, 10^{-4} K^{-2} and 10^{-3} K^{-2}. A random sample of 1% of the simulated data points with lower probability than 10^{-5} K^{-2} are plotted as blue dots. For the simulations all points with lower probability than 10^{-5} K^{-2} are plotted as orange dots. There are only about 42 000 cases in the simulation dataset for snow and only about 5000 cases in the simulation dataset for sea-ice.