https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

The Arctic Weather Satellite, introducing a new wavelength range for ice hydrometeor retrievals

Peter McEvoy[1], Eleanor May[1], and Patrick Eriksson[1]

[1]Department of Environmental and Energy Sciences, Chalmers University of Technology, Gothenburg, Sweden

Correspondence: Peter McEvoy (peter.mcevoy@chalmers.se)

Abstract. The first cloud property retrievals based on operational sub-millimetre measurements are presented, making use of the channels between 89 and 325 GHz of the Arctic Weather Satellite (AWS). The main quantities of the dataset are frozen water path (FWP) and the associated mass-weighted mean altitudes and particle sizes. In this first version, results are restricted to latitudes between 60° S and 60° N. The retrievals are based on detailed simulations of instrument observations. The actual inversion is made by a quantile regression neural network, and case-specific uncertainty estimates are provided.

Retrievals performed on simulations suggest that retrieved FWP values are essentially unbiased across a wide dynamic range, from $10\,\mathrm{kg\,m^{-2}}$ down to $40\,\mathrm{g\,m^{-2}}$. The associated mass-weighted mean altitude is also essentially unbiased for the entire relevant range of 2 km to 12 km. The particle size estimates, however, show a slight bias for sizes other than 400 µm. Comparisons with other datasets provide strong indications that these results also extend to retrievals from real observations; for example, local and zonal means match those of existing radar/lidar-based retrieval products.

The accuracy in FWP should be unprecedented among estimates based on passive satellite data, thanks to the new sensitivity afforded by sub-millimetre channels. The dataset complements cloud radar observations by providing a significantly broader spatial coverage. There is also potential to create a climate-relevant dataset, as the retrievals are directly applicable to the EPS-Sterna constellation, continuing the AWS observations up to 2045.

1 Introduction

For more than 30 years, frequencies above 300 GHz have been identified as being of high interest for improving measurements of ice hydrometeor properties, particularly their mass (Gasiewski, 1992; Evans and Stephens, 1995). During these three decades, insights have been gained from simulations (e.g. Evans et al., 2002; Miao et al., 2003; Ekelund et al., 2020) and air-borne instruments (Wang et al., 1998; Evans et al., 2005; Fox, 2020; Pfreundschuh et al., 2022). Observations from space have been made (Wu et al., 2008; Eriksson et al., 2014), but limited to limb sounding, a measurement geometry less suited for ice hydrometeor retrievals. This changed in 2024 with the launch of the Arctic Weather Satellite (AWS) (Eriksson et al., 2025). For the first time, regular downward-looking data with global coverage became available. Here, we present the first dataset of ice hydrometeor properties based on AWS.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

Satellite-based retrievals of cloud ice properties have been limited to measurements at each end of the electromagnetic spectrum. Optical and infrared passive sensors observe ice clouds at wavelengths up to about $\lambda = 15\,\mu\mathrm{m}$, and on the long wavelength side, operational microwave radiometers have been limited to frequencies below 200 GHz ($\lambda = 1.5\,\mathrm{mm}$). It is noteworthy that the frequency/wavelength ratio between these two limits is 100. This ratio is even higher for active measurements, about 3000. Cloud lidars and radars to date operate at 335 - 1064 µm and 94 GHz, respectively. Being restricted to relatively short and long wavelengths results in challenges for retrievals. Short-wave observations are characterised by the strong impact of the more numerous small particles, and the information provided by these measurements is often limited to the cloud top layer (e.g. Eliasson et al., 2011). On the other hand, microwave observations, both active and passive, can sound the interiors of cloud systems, but the measured signal is dominated by the locally largest particles.

These deviations in partial responses, in terms of particle size and cloud penetration, have led to poor agreement between satellite-based estimates of ice mass (Waliser et al., 2009). Retrievals combining satellite lidar and radar data are considered the most accurate, due to the high spatial resolution and sensitivity to both small and large ice hydrometeors. Still, the presently leading examples, DARDAR (Cazenave et al., 2019) and 2C-ICE (Deng et al., 2015), differ with 20 % even for global mean ice column mass (Eriksson et al., 2026). A main limitation of these datasets based on active observations is their limited spatial coverage (Eliasson et al., 2011), which underscores the need for ice mass retrievals based on passive observations.

Ice mass datasets based on passive data have, in general, been biased low compared to those based on active observations (Eliasson et al., 2013; Duncan and Eriksson, 2018). This can largely be understood as the fact that non-sensed particle sizes and altitudes are not included in the estimates. However, some newer passive retrievals compensate for this by applying machine learning, using e.g. 2C-ICE collocations as training data (Holl et al., 2014; Amell et al., 2024; Yang et al., 2026). Indeed, this approach can avoid biases with respect to the training data (Pfreundschuh et al., 2025). On the other hand, biases of the training data are inherited, and the retrievals do not provide any new physical insights. In addition, the retrieval precision is relatively poor due to the physical limitations of the passive observations used as input.

Sub-millimetre measurements have the potential to fill the sensitivity gap between the two conventional wavelength ranges (Wu et al., 2024). The sub-millimetre range is a continuation of the microwave domain, starting at 300 GHz ($\lambda = 1\,\mathrm{mm}$). The range's upper limit is often set to 3 THz ($\lambda = 0.1\,\mathrm{mm}$), but high absorption due to water vapour strongly limits the altitude coverage from about 1 THz ($\lambda = 0.33\,\mathrm{mm}$). Key features of sub-millimetre measurements as input for satellite-based ice mass retrievals are that they offer sensitivity over a broad range of ice mass (Evans et al., 1998; Jimenez et al., 2007) at the same time as providing altitude coverage (Evans et al., 2012; May and Eriksson, 2025).

To fully exploit the opportunities offered by the sub-millimetre range, satellite-based downward-looking observations are required. Several such satellite missions have been formulated (e.g. Miao et al., 2002; Buehler et al., 2012). Among the successful proposals, the most ambitious mission is the Ice Cloud Imager (ICI; Eriksson et al. 2020), with three such sensors to be launched as part of the second generation of the EUMETSAT Polar System (EPS-SG; Mattioli et al. 2019). The first launch involving ICI is planned for 2026. About one year later, the two small PolSIR satellites, each carrying a 325 and 680 GHz radiometer, are expected to be launched. The scientific objectives of PolSIR are related to diurnal variations of tropical ice clouds

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

(Bennartz and Wu, 2025). An even smaller satellite, the IceCube technology demonstration mission, performed measurements around 880 GHz for 15 months (2017-18; Gong et al. 2021).

AWS was not designed with ice mass retrievals in mind; its main justification is to provide microwave data for numerical weather prediction (NWP) in a more timely and cost-effective manner. Nevertheless, the design of the AWS radiometer came to include a receiver with four channels around 325 GHz, to improve the treatment of ice clouds (Kaur et al., 2021). Consequently, it became the first operational mission to provide sub-millimetre data (Eriksson et al., 2025). These observations will be continued by the EPS-Sterna constellation, consisting of copies of AWS (Rivoire et al., 2024). The first batch of satellites in this constellation is planned to be launched 2029, and there will be a continuous data record of both ICI and AWS measurements that extends to 2045.

Accurately representing atmospheric ice in atmospheric models remains a major challenge; they have long shown a large spread in their estimates of cloud ice mass (Waliser et al., 2009). Although improvements are seen in newer generations (CMIP6), the intermodel spread remains high, and models remain biased low relative to observations (Lauer et al., 2023). This is a problem with direct consequences for the estimation of longwave cloud radiative effects (Bjordal et al., 2020). Global storm-resolving models (GSRMs) represent a step forward. Operating at a finer spatial resolution, these models explicitly resolve clouds and convective dynamics. Importantly, GSRMs report all hydrometeor classes, allowing for the total column ice to be obtained. This is the only quantity that can reasonably be compared across models and with observations, since "cloud ice" is ambiguously defined. While GSRMs show improved agreement with observed FWP, differences remain, both between models and with observations (Eriksson et al., 2026). Further progress requires multiple, independent observational datasets to verify atmospheric models from global to local scales.

We present the “Chalmers hydrometeor inversion product from the Arctic Weather Satellite” (CHIP-AWS): a level 2 (L2) product of ice hydrometeor mass retrievals from AWS. This first version of the retrieval product is designed to be comparable to the operational product planned by EUMETSAT for ICI (Eriksson et al., 2020), but uses an updated machine-learning algorithm following May et al. (2024). The main retrieval is the frozen water path (FWP). The term FWP indicates that the aim is to estimate total ice mass, not just cloud ice. As a step towards providing profiles of frozen water content (May and Eriksson, 2025), FWP is accompanied by an estimate of the mass-weighted mean altitude. Size information is given as the full column mass-weighted mean diameter.

The AWS measurements are inverted into these quantities using machine learning (ML) in novel ways. A quantile regression neural network (QRNN; Pfreundschuh et al. 2018) is applied to provide case-specific retrieval uncertainties. Such a detailed error description is so far rare for machine learning-based datasets. Furthermore, to the best of our knowledge, operational machine learning models for cloud retrievals have exclusively been trained on existing retrievals. In contrast, our AWS retrievals are based on simulated observations. For reasons discussed below, this more complex approach could not be avoided. However, the approach should also be considered as a step forward, since it opens the way for a physical understanding of the retrievals.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

This first version of the dataset covers $\pm 60^\circ$ in latitude and off-nadir angles within 30°. These limitations are in place to prioritise good results for most areas before extending coverage. Retrievals are performed over all surfaces, though accuracy is lower over land, especially for higher latitudes and dry areas. The intention is to remove these constraints in future versions.

2 Source and reference data

This section introduces the instrument and input data used to produce the data product described in this article, as well as the reference data used for comparisons.

2.1 The Arctic Weather Satellite

AWS was launched into a sun-synchronous orbit in August 2024 and became fully operational at the beginning of 2025. The mean altitude of the orbit is about 590 km and the local time of the ascending node (LTAN) is 22:38. It carries, as its sole instrument, an across-track scanning radiometer. Here, a rotating parabolic mirror provides 145 so called field-of-views of Earth perpendicular to the satellite's flight direction. The radiometer has 19 channels over four frequency bands; AWS 11-18 (50 GHz), AWS21 (89 GHz), AWS 31-36 (183 GHz) and AWS 41-44 (325 GHz) corresponding to one feed-horn and receiver-chain for each band (Eriksson et al., 2025). The processed L1b files from EUMETSAT give measured radiances in terms of antenna temperatures T_a with unit kelvin (K) for each channel, field-of-view and scan.

In this work, channels AWS 21-44 are used, i.e. the 11 channels that include 89 GHz and above. The Earth-surface footprints for the 89 GHz, 183 GHz and 325 GHz receiver pairs have a full-width-at-half-maximum (FWHM) size of 20, 10, and 10 km, respectively, when the scanning mirror is pointing toward nadir (straight down).

In an effort to simplify the instrument, alignment optics were omitted and the four feed-horns are therefore slightly off the optical axis of the across-track scanning mirror. This results in different Earth-surface footprints for channels belonging to different receivers (see Fig. 5 of Eriksson et al., 2025 for more details). This has implications for using the measurements. To approximate a common propagation path between channel values, the values for AWS 21 and AWS 41-44 are re-interpolated onto the footprint locations of the grid defined by the 183 GHz channel footprints.

Mirror angles, the angles of rotation of the parabolic mirror from nadir, for the different measurements are not provided in the L1b files. Instead, zenith angles from Earth-surface footprint locations towards the satellite are available. The mirror angle can be approximated by geometrically determining the zenith angle in the satellite's reference frame at the expected altitude. These approximated mirror angles align with the simulated training data produced for this work (Sect. 4.1).

Systemic instrument biases in measured T_a values are corrected for. These bias corrections are informed by observation-minus-background (O-B) analysis of AWS T_a values performed by ECMWF (David Duncan, ECMWF, private communication). ECMWF provides a live plot of O-B values for AWS at https://charts.ecmwf.int/catalogue/packages/obstat/ (last access 2026-01-29). These corrections are mainly constant offsets, with two exceptions: the 183 GHz channels all have biases that change linearly across the scan's fovs and channel AWS35 (181.5 GHz) has an additional bias that varies with latitude. Our

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

corrections also take these variations into account. See the supplemental material for exactly how these bias corrections are applied.

For simulating the instrument, sensor characteristics are needed. In this work, measured instrument noise and measured spectral response functions are used as reported in Eriksson et al. (2025) (Table 4 and Sect. 4.1).

2.2 ERA5, static data and models

ERA5 hourly data on pressure levels and single level (Hersbach et al., 2023a, b) are used to define the atmospheric state in simulations (Sect. 4.1). The variables include single-level data such as skin temperature, 2-metre temperature, 10-metre U and V wind components, and surface pressure, as well as pressure-level data such as temperature, geopotential, specific humidity, ozone mass mixing ratio, and specific cloud liquid water content.

Furthermore, simulation surface properties are derived from a digital elevation map (GTOPO30; gto 2018) and a surface type classification (Copernicus land cover 100m; Buchhorn et al. 2020). The land cover classifications are further enhanced with dynamic effects by using ERA5 single-level variables: sea ice concentration and snow depth.

Additionally, various models and datasets are used to drive the simulations. For details and references, refer to May et al. (2024).

2.3 CloudSat and DARDAR

CloudSat (Stephens et al., 2002) was operational between 2006 and 2023 and had the cloud profiling radar (CPR) as its main instrument. The satellite's track, measurement timestamps, and radar reflectivities are utilized to drive the simulations described in Sect. 4.1. The radar reflectivities are used to infer the frozen water content used to set the scene to be simulated. The complete catalogue of 2015 was used.

DARDAR is a data product that provides ice cloud microphysical properties (Cazenave et al., 2019). This product is derived by combining CPR measurements from CloudSat with lidar measurements from CALIPSO. DARDAR retrievals serve as a comparison dataset for the quantities presented in this work.

We explicitly note that, although CloudSat CPR reflectivities are used both as input for DARDAR and for the simulations in this article, the inversion from reflectivities to frozen water content in the simulations presented in this work is independent of the method employed by DARDAR. We argue that this independence makes DARDAR a valid product for comparison.

2.4 Chalmers Cloud Ice Climatology (CCIC)

CCIC provides global FWP (and other variables) estimates at high temporal resolution, spanning 60° S to 60° N every half hour (Amell et al., 2024). The method processes geostationary infrared images through a neural network model trained on 3.5 years of CloudSat-CALIPSO-derived 2C-ICE (Deng et al., 2015), another microphysical properties retrieval dataset. The network generates FWP values along with associated uncertainty estimates. CCIC uses a globally-merged IR dataset (Janowiak et al., 2001) that combines multiple IR images from various geostationary satellites into a single georing product.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

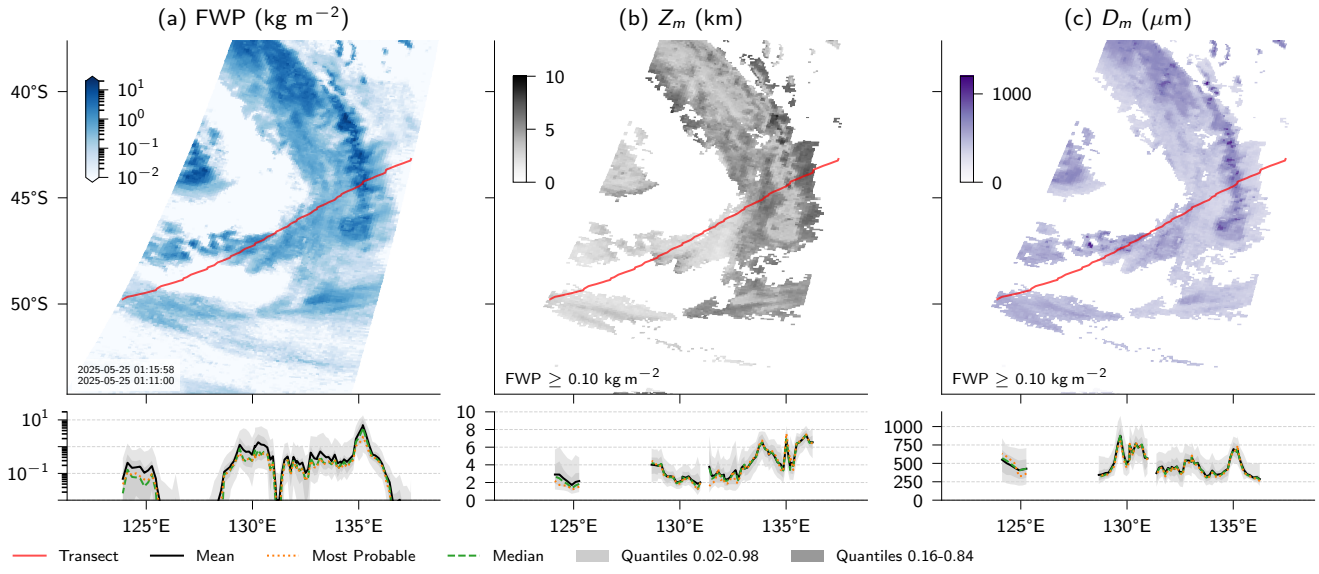

Figure 1. An example of CHIP-AWS data: a weather system in the Atlantic Ocean, south-east of South America. Top row shows the mean values derived from the retrieved CDF for three ice-mass-related variables: FWP, Z_m, and D_m. Bottom row presents dataset estimates sampled along the transect line and projected onto the longitude axis. These include the "mean" (the average of the retrieved CDF), the "most probable" value (corresponding to the maximum of the retrieved PDF), and the quantiles.

We include CCIC data as a comparison product for two reasons. First, it provides an independent passive FWP estimate. Second, it is the only FWP dataset with sufficient global coverage to provide co-locations with AWS.

3 The data product

3.1 Description

CHIP-AWS is a dataset of atmospheric ice-mass estimates that have been retrieved from calibrated antenna temperature (T_a) values from AWS's across-track scanning radiometer. AWS data from 2025 and onwards are continuously processed and available at http://clouds-and-precip.group/datasets/chip-aws. An archival copy of 2025 data can be accessed under data DOI https://doi.org/10.71870/v9dx-rm43 (McEvoy, May, and Eriksson, 2026). (Note that DOI is inactive during manuscript review, see the data availability section).

The retrievals are performed footprint by footprint and constitute a quasi-global ($\pm 60°$ latitude) level 2 processing of level 1b AWS data distributed by EUMETSAT.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

Variable	Dimensions	Description
Frozen water path ($\mathrm{kg\,m^{-2}}$)	(scan, fov)	Mass of ice hydrometeors in a vertical column through the atmosphere.
Z_m (m)	(scan, fov)	Mass-weighted mean altitude of ice hydrometeors inside the column.
D_m (m)	(scan, fov)	Mass-weighted mean volume-equivalent diameter of ice hydrometeors inside the column.

Auxiliary Variable	Dimensions	Description
Surface type fractions	(scan, fov, surface_type)	Fraction of footprint covered by: Ocean, Sea Ice, Land, Water, Snow and Glacier.
T_a (K)	(scan, fov, channel)	Reinterpolated and bias-corrected antenna temperatures used for the retrieval.
Remap distance (m)	(scan, fov)	Distance between AWS3X and the nearest AWS4X footprint in the reinterpolation window. This can be used to filter for cases that required little reinterpolation.
CCIC FWP ($\mathrm{kg\,m^{-2}}$)	(scan, fov)	Co-located FWP estimates from the Chalmers Cloud Ice Climatology (CCIC). An independent retrieval method based on geostationary IR.
Incidence angle (°)	(fov)	The median of "Satellite Zenith Angle" values reported by source L1b file.
Altitude (m)	scalar	Mean of "Satellite Altitude" values reported by source L1b file. Unreasonable values are removed before averaging.

Table 1. Estimated variables and auxiliary variables in CHIP-AWS. Estimates for the main variables are given as a “mean” (the mean of the retrieved CDF), ”most probable“ (value at peak PDF) and the 0.02, 0.16, 0.5, 0.84, and 0.98 quantiles of the retrieved CDF.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

Table 1 shows the variables included in this dataset. There are three main variables: frozen water path (FWP), Z_m, and D_m (see Sect. 3.2 for a definition of these quantities). These are given in the along-track "scan" dimension and the across-track "field-of-view" (fov) dimension with 88 elements corresponding to scan angles.

For each retrieval, a mean and a most probable estimate is provided, along with the 0.02, 0.16, 0.5, 0.84, and 0.98 quantiles to characterize the distribution of probable values. The underlying retrieval method generates an estimated cumulative distribution function (CDF) that characterizes the distribution of possible estimates (Sect. 4.2). This approach provides a representation of the uncertainty, addressing the ill-posed nature of the inversion problem, where multiple distinct scenarios can produce similar observed signals.

Auxiliary variables such as land cover type, antenna temperature, and interpolation distance are included to allow filtering data for specific applications.

3.2 Quantity definitions

Frozen water content (FWC) ($\mathrm{kg\,m^{-3}}$) can be expressed as

$$\mathrm{FWC} = \int_0^\infty m(d_\mathrm{veq}) n(d_\mathrm{veq}, z)\, \mathrm{d}d_\mathrm{veq}\,, \tag{1}$$

where n is the particle size distribution (PSD), m is the particle mass and d_veq is the volume-equivalent particle diameter ($d_\mathrm{veq} = \sqrt[3]{6m/\pi\rho}$), with m and ρ being the particle mass and the density of (solid) ice, respectively.

The main retrieval quantities are defined as follows. FWP is the vertical integral over FWC,

$$\mathrm{FWP} = \int_{z_0}^\infty \mathrm{FWC}\, \mathrm{d}z\,. \tag{2}$$

The mass-weighted mean altitude Z_m is defined as

$$\mathrm{Z_m} = \frac{\int_{z_0}^\infty z\mathrm{FWC}\, \mathrm{d}z}{\mathrm{FWP}}\,. \tag{3}$$

Finally, the mass-weighted mean particle diameter D_m is defined as

$$\mathrm{D_m} = \frac{\overbrace{\int_0^\infty d_\mathrm{veq}^4 \int_{z_0}^\infty n(d_\mathrm{veq}, z)\, \mathrm{d}z\, \mathrm{d}d_\mathrm{veq}}^{\text{4th moment}}}{\underbrace{\int_0^\infty d_\mathrm{veq}^3 \int_{z_0}^\infty n(d_\mathrm{veq}, z)\, \mathrm{d}z\, \mathrm{d}d_\mathrm{veq}}_{\text{3rd moment}}}\,. \tag{4}$$

This is consistent with Delanoë et al. (2014), Eq. (3), with the exception that their size parameter is defined with respect to the melted equivalent diameter.

For a more rigorous definition of these quantities that consider the spatial sensitivity of measurements, refer to Eq. (1–3) in Eriksson et al. (2020).

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

3.3 Spatio-temporal coverage and resolution

AWS measures continuously along its sun-synchronous orbit. In this dataset, measurements are provided in the instrument's native scan and field-of-view dimensions. However, this first version of the dataset is restricted to latitudes between 60° S and 60° N and the number of field-of-views are reduced from 145 to 88 elements to limit off-nadir incidence angles below about 37°. This results in a swath width of approximately 800 km. These restrictions are put in place to reduced the scope of this initial version. Future work aims to extend the retrievals with broader swaths and fully global coverage.

The combination of this swath width, the sun-synchronous orbit, and data processing results in median revisit times (per pass direction) of approximately one day globally. Variability increases towards the equator, with 75th percentile revisit times of five days at mid-latitudes and approximately seven days in the tropics. For further details on revisit time, see Appendix B.

This dataset includes estimates for all of 2025 and onwards as new retrievals are regularly processed. However, due to AWS commissioning tests and data processor updates, there are more frequent data gaps at the beginning of the year. From 2025-03-21 onwards, gaps are less frequent.

The instrument performs a full scan every 1.19 s, resulting in an along-track sampling distance of 8 km. The instrument's scanning motion leads to a varying sampling distance across-track, ranging from 8 to 12 km, for scan angles limited to 30° from nadir. The footprint size depends on the input channels. Out of the 11 input channels used to produce this dataset 10 have a nadir footprint FWHM size of 10 km, while only AWS21 has a footprint size 20 km. Due to having 10 out of 11 channels with the smaller footprint size, we argue that the effective resolution of the retrieval method is closer to 10 km at nadir.

The major d_a (across-track) and minor d_b (along-track) diameters of the ellipse approximating the footprint FWHM size depend on the Earth-surface incidence angle θ as follows:

$$d_b(\theta) \approx \left(\sqrt{(R+h)^2 - R^2\sin^2\theta} - R\cos\theta\right)\frac{d_n}{h}, \tag{5}$$

$$d_a(\theta) \approx \frac{d_b(\theta)}{\cos\theta}, \tag{6}$$

where R is Earth's radius, h is the satellite's altitude, and d_n is the footprint FWHM size at nadir. With these expressions, the footprint size can be calculated for each observation using incidence angles (satellite zenith angles) and satellite altitudes. The retrievals include averaged incidence angles for each field-of-view and averaged satellite altitudes, from each source L1b file. Unreasonable values from L1b are excluded from averaging. For the extreme scan-angles in this dataset 10 km nadir FWHM size is scaled to 14 km across-track and 11.7 km along-track.

Despite the fact that AWS observes the atmosphere from a slanted angle off-nadir, the dataset values are estimates of vertical columns from the Earth's surface. The model behind the retrieval procedure is trained to scale estimates to account for the extra integration path.

3.4 Example scene

We demonstrate the available variables and estimates using an example scene from the dataset. Figure 1 shows a weather system in the Atlantic, south-east of South America, with varying intensity of clouds across the scene.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

The retrievals from this scene reveal a wide range of values for all three variables: clear-sky conditions, areas with little ice, thick regions and intense centres. The mass-weighted mean properties Z_m and D_m likewise show considerable variation. Notably, while these properties exhibit trends that follow FWP (i.e. higher FWP is associated with larger particle diameters and height), multiple regions show variations in Z_m and D_m where the corresponding FWP is relatively constant.

For FWP (Fig. 1a) transect values close to $120°$ E, uncertainty is high (wide quantile span), while the estimate is relatively more certain for higher values near $135°$ E. This demonstrates how the dataset conveys uncertainty and that estimated values are more certain when there is enough ice to sense.

For mean mass-weighted altitude Z_m (Fig. 1b), note that values are masked in this figure when FWP values fall below the threshold of $0.1\,\mathrm{kg\,m^{-2}}$. Although the dataset still provides estimates for these cases, the high uncertainty renders these mass-weighted quantities uninformative. This also applies to D_m. In regions with higher FWP, we observe a low-altitude segment around $125°$ E and a high-altitude segment near $135°$ E.

Finally, for mass-weighted particle mean volume-equivalent diameter D_m (Fig. 1c), we can identify regions with different particle sizes, yet have uniform FWP estimates (such as $50°$ S, $135°$ E).

In summary, the example scene illustrates dataset variables and the relationships between them. It is evident that both Z_m and D_m exhibit lower uncertainty only when FWP values are significant, emphasizing the importance of also considering FWP when interpreting these estimates.

4 Retrieval method

The retrieval method is based on the preparatory work for the Ice Cloud Imager (ICI) to be carried by Metop-SG-B described in May et al. (2024). In short, detailed radiative transfer simulations produce a database of simulated instrument antenna temperature (T_a) values and corresponding scene parameters (FWP, Z_m, D_m). The database is then used to train a neural network. Here, QRNNs give a model that produces estimated quantiles and thus CDFs of the physical parameters based on the T_a input. To get final estimates, real T_a observations from AWS are re-interpolated and bias-corrected (Sect. 2.1), before being fed into the trained QRNN model.

More detail on the simulation, training, and inference steps, as well as key differences from May et al. (2024), follows.

4.1 Simulation database

Radiative transfer simulations with ice-particles were performed with ARTS version 2.6 (Buehler et al., 2025). The ARTS database for single-scattering properties of ice hydrometeors (Eriksson et al., 2018) was used, which includes data for the sub-millimetre range.

For each simulation, an ice particle model, defined as a pairing of an ice particle habit and a particle size distribution (PSD), is chosen. Table 2 in May et al. (2024), shows the set of six particle models used for the simulations in that work. In this work, simulations use the same particle models, however, with particle model “AA3” replaced. “AA3” (habit: Large block aggregate; PSD: Delanoe et.al. 2014) is replaced with a new particle model consisting of habit “IconHail” from the same

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

Input	Purpose
CloudSat Radar Reflectivities	Set FWC. See Sect. 2.3.
ARTS single-scattering database	Simulate scattering with ice hydrometeors. See Sect. 2.2
ERA5 Variables	Set surface and atmospheric properties. See Sect. 2.2.
RTTOV v13	Gas absorption model.
TELSSEM / TESSEM	Surface emissivity model.
Landcover	Set surface emissivity based on surface type.
DEM	Set surface elevation.
Sensor Characteristics	Instrument modelling. AWS spectral response function, footprint sizes and noise. See Sect. 2.1.

Table 2. Simulation inputs.

scattering database and an exponential PSD (Marshall and Palmer, 1948); $n(d_{\text{veq}}) = N_0 \exp(-\Lambda d_{\text{veq}})$, with $N_0 = 2 \times 10^7$ and the slope-parameter Λ left as a free parameter that is set to achieve the simulation-scene's FWC value.

This change of particle model was motivated by early AWS observations revealing discrepancies in areas of deep convection. We leave this analysis and further investigations on what the AWS data reveal about particle microphysics to a forthcoming study, keeping this work focused on the data product itself.

The driving input to the simulations are radar backscatter data from CloudSat's cloud profiling radar. For an input CloudSat segment, a random particle model from the set of six is chosen, after which the frozen water content is estimated from the radar backscatter, populating the atmospheric state in the scene to be simulated. The particle models have different weighting in the random sampling (see May et al., 2024). The input CloudSat segment's lat-lon trajectory and timestamps are used to load complementary ERA5 data (see Sect. 2.2) setting humidity, surface skin temperature, and other properties. The scene is further refined by changing surface emissivity properties by looking up the trajectory in a land-cover dataset that has been augmented with dynamic ERA5 variables on sea-ice concentration and snow depth. All simulation inputs are listed in Table 2.

With the scene populated, radiative transfer simulations of AWS observations can begin. Here the AWS instrument is simulated observing along the same input CloudSat track segment. Even though the scene is 2D (altitude and along-track), a variation in the simulated across-track scan is achieved by the slant path effect from the varying scan angle. The instrument's frequency response is modelled using measured spectral response functions (SRFs) of the instrument channels (Eriksson et al., 2025).

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

A simulated segment results in multiple pencil beam brightness temperatures. These are seen as samples that are then integrated with weighting according to a Gaussian distribution to model the instrument's footprint, resulting in the simulated T_a.

The full database is generated by performing simulations on the entire CloudSat radar backscatter dataset from 2015 multiple times with randomised properties such as particle model, scan angle, and segment selection. This process yields a database of 6,086,928 simulated AWS T_a values for each of the 11 channels used as retrieval inputs, along with the corresponding scene parameters (FWP, Z_m, D_m).

4.2 Inversion with QRNN

A QRNN machine learning model performs the inversions. In this work, we adopt the remote sensing nomenclature, where the state vector $\boldsymbol{x}$ is retrieved from the observation vector $\boldsymbol{y}$. The network is trained to perform this inversion using a dataset of paired observations and states $\{\boldsymbol{y}_i, \boldsymbol{x}_i\}_i^n$.

The inverse problem considered here is non-linear and ill-posed. As a result, an observation vector does not map to a single state vector, but rather to a region in state space. It is therefore more appropriate to take a probabilistic approach to the inverse problem and retrieve a description of the probability density function (PDF) of the state. A QRNN achieves this by estimating quantiles of the CDF of the state vector $\boldsymbol{x}$. It does this through the use of a quantile loss function (Pfreundschuh et al., 2018), defined as

$$\mathcal{L}_\tau(x_\tau, x) = \begin{cases} \tau|x - x_\tau|, & x_\tau < x \\ (1-\tau)|x_\tau - x|, & \text{otherwise.} \end{cases} \tag{7}$$

where $\tau \in [0,1]$ denotes the quantile level. By minimising the expectation of $\mathcal{L}_\tau(x_\tau, x)$, the model can estimate the τth quantile x_τ, defined as

$$x_\tau = \inf\{x : F_{x|\boldsymbol{y}}(x) \geq \tau\}. \tag{8}$$

where $F_{x|\boldsymbol{y}}(x)$ represents the CDF. By training the retrieval model to estimate multiple quantiles, a discrete approximation of the CDF is obtained per observation.

The final model, through its training, will have learnt the distribution of the data used to train it, which serves as the prior in a Bayesian framework. When given measurements inputs, this prior is updated with the new information and produces the posterior distribution. Figure 2 illustrates this on a real scene: measurements from an intense part of the scene shift the quantile values so that high FWP become more probable, whilst a clear-sky measurement shifts quantile values to much lower FWP.

From the CDF, two single point estimate are obtained: either the mean (expectation value) μ and the most-probable value $\hat{x}$. The mean μ is given by integrating the quantile values over probability levels (integrating x over the CDF),

$$\mu = \mathbf{E}[X] = \int_{-\infty}^{\infty} x\, dF(x)\,. \tag{9}$$

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

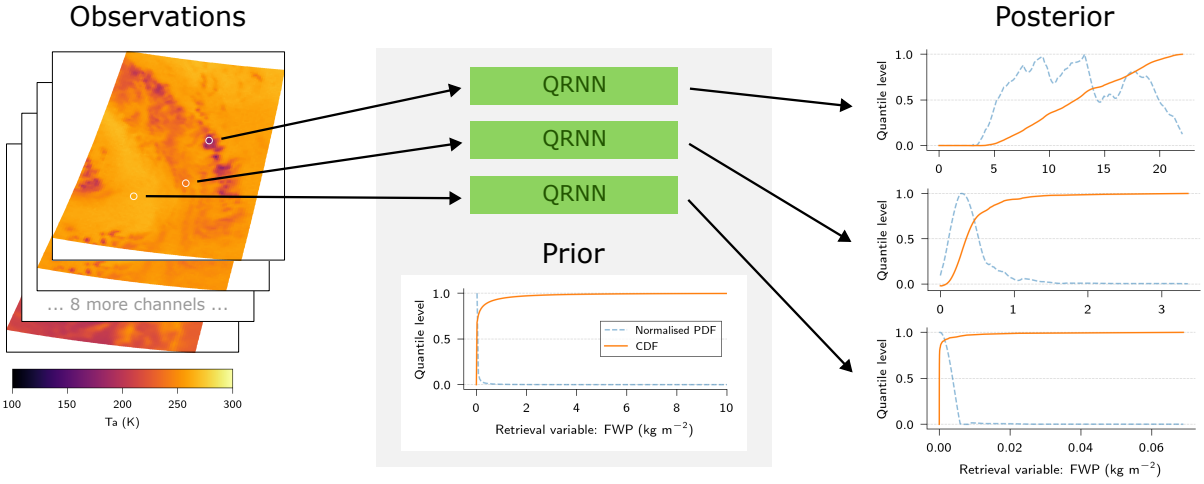

Figure 2. Illustration of QRNN updating quantiles for three example observations. Each observation comprises 11 T_a values. The model outputs quantile estimates interpretable as CDFs and PDFs, here PDFs are normalised so that maximum value equals 1. Both CDFs and PDFs are smoothed in this illustration.

The most-probable estimate $\hat{x}$ corresponds to the value at which the PDF is maximised,

$$\hat{x} = \arg\max_x \frac{d}{dx} F_{x|\boldsymbol{y}}(x)\,. \tag{10}$$

In practice, the integral for μ is evaluated using trapezoidal integration. For the most-probable estimate, noise in the predicted CDF can introduce significant artefacts, the CDF is therefore smoothed before numerical differentiation to more robustly identify the peak of the PDF.

Additionally, since the CDF represents the retrieval uncertainty arising from the ill-posed nature of the problem, i.e. the aleatoric uncertainty, the retrieved quantiles (or spread between them) constitute a observation-specific uncertainty measure. Accordingly, selected quantiles are provided alongside the single point retrieval estimates, as described in Sect. 3.1.

4.3 Training and inference

For CHIP-AWS, the QRNN input $\boldsymbol{y}$ is a vector consisting of a set of 11 antenna temperatures for channels in the 89, 183, and 325 GHz bands, and the mirror angle. The model output $\boldsymbol{x}$ corresponds to the scene parameters FWP, Z_m, and D_m. The list of inputs and outputs is presented in Table 3.

As described in Sect. 4.1, each simulation consists of a pairing of scene parameters $\boldsymbol{x}$ and corresponding simulated measurements $\boldsymbol{y}$, which together form the database $\{\boldsymbol{y}_i, \boldsymbol{x}_i\}_i^n$. The database is divided into 80 % training, 10 % validation and 10 % testing subsets. Subset division is performed in terms of CloudSat segments and not individual observations. The training and validation subsets are fed to the model during training, while the test subset is reserved for evaluation.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

Since FWP spans orders of magnitude, a log-linear transformation is applied to FWP prior to training: values less than $1\,\mathrm{kg\,m^{-2}}$ are transformed into logarithmic space. Cases of FWP less than $10^{-4}\,\mathrm{kg\,m^{-2}}$ are replaced with a random value sampled uniformly between 10^{-6} and $10^{-4}\,\mathrm{kg\,m^{-2}}$ and subsequently transformed into logarithmic space. As a consequence, the dataset has no estimates of zero FWP.

A single model is trained to output 99 uniformly space quantile levels ($\tau \in [0.01, 0.99]$) of the CDF for each retrieved variable. The network consists of 12 fully-connected layers that are shared across all output variables, followed by variable-specific two-layer multi-layer perceptron head that outputs the 99 quantile estimates.

For each individual AWS channel, a random noise sample is generated by sampling from a Gaussian distribution with zero mean and a standard deviation corresponding to NEΔT on ground (OG), which is given in Table 4 of Eriksson et al. (2025). New noise samples are generated each time a training sample is seen by the neural network, acting as a form of data augmentation.

Once trained, the model can perform retrievals on real observations. During inference, an AWS observation for a single scan and field-of-view is fed into the network. In the simulated training data, all input channels shared the same Earth-surface footprint, i.e. have a common propagation path. For consistency, the observation is re-interpolated onto the AWS3X footprint (see Sect. 2.1). Since the distance between the AWS3X and the nearest AWS4X footprint can be substantial, this distance is stored as an auxiliary variable (see Table 1). The observation T_a values are additionally bias corrected (Sect. 2.1).

Feeding the re-interpolated and bias-corrected T_a values into the trained QRNN returns an estimated CDF for each variable. From this CDF, he the mean estimate, the most probable estimate, and quantiles are derived, see Table 3. The produced dataset files includes five of the 99 quantiles: 0.02, 0.16, 0.5, 0.84, and 0.98.

5 Results

In this section, we first demonstrate that the underlying simulation database (DB) agrees with observations. We asses retrievals on a held-out test subset of simulations, demonstrating ideal synthetic retrieval performance. Finally, we show retrieval results from real AWS observations and compare against other datasets, both for individual cases and statistically.

5.1 Simulation database

Basing the retrieval training on simulations requires them to be accurate and cover the many atmospheric cases that can potentially be observed. To verify this, we look at the distribution of simulated T_a values and compare with observations from AWS. First, only cases over oceanic cases are considered, to focus on the accuracy of the radiative transfer simulations with the best performing surface model. Then other surface types are covered.

For these comparisons, recall that the simulations have the atmospheric state driven by 2015 CloudSat radar reflectivities, whilst AWS observations are from 2025. Co-located simulations were not possible during this work due to CloudSat being decommissioned. Its successor, EarthCARE, was in commissioning phase at the time of writing and operates in an incompatible

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

Input variable	Description
AWS21 T_a	T_a at 89 GHz
AWS31 T_a	T_a at 165.5 GHz
AWS32 T_a	T_a at 176.311 GHz
AWS33 T_a	T_a at 178.811 GHz
AWS34 T_a	T_a at 180.311 GHz
AWS35 T_a	T_a at 181.510 GHz
AWS36 T_a	T_a at 182.311 GHz
AWS41 T_a	T_a at 325.15±1.2 GHz
AWS42 T_a	T_a at 325.15±2.4 GHz
AWS43 T_a	T_a at 325.15±4.1 GHz
AWS44 T_a	T_a at 325.15±6.6 GHz
Mirror angle	See Sect. 2.1.

Output variable	Description
FWP quantiles	Quantiles x_τ of the CDF of FWP, $\tau \in [0.01, 0.99]$.
Z_m quantiles	Quantiles x_τ of the CDF of Z_m, $\tau \in [0.01, 0.99]$.
D_m quantiles	Quantiles x_τ of the CDF of D_m, $\tau \in [0.01, 0.99]$.

Table 3. Inputs and outputs of the retrieval model. The T_a values listed as input variables are re-interpolated onto the AWS3X lat-lons and bias-corrected antenna temperatures, with the names corresponding to the channel ID.

Metric	FWP ($\mathrm{kg\,m^{-2}}$)	Z_m (km)	D_m (µm)
RMSE	0.64	1.55	122
Bias	-0.01	0.12	12
MFE	0.86	n/a	n/a
r	0.87	0.87	0.70

Table 4. Metrics of retrieval performance on the test set. All metrics are calculated only for cases with a true FWP value greater than $10^{-2}\,\mathrm{kg\,m^{-2}}$.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

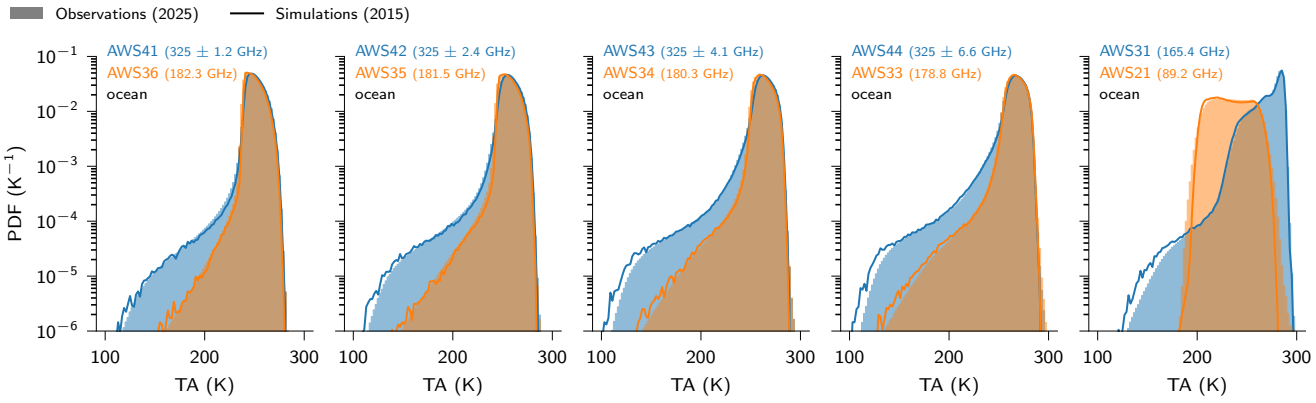

Figure 3. Distribution of simulated database T_a values (driven by CloudSat radar reflectivities from 2015) and distribution of actual observed T_a values from AWS 2025. Channels that have similar altitude of peak humidity sensitivity are paired. Both simulations and observations are filtered for ocean cases.

orbit. Despite this temporal mismatch, a sufficiently large sample size should ensure that the distribution of atmospheric states remains statistically representative, allowing for meaningful comparisons between the two datasets.

5.1.1 Simulations over ocean

Figure 3 shows the T_a distributions from the simulation database and AWS observations for each retrieval input channel over ocean. The simulated and observed T_a distributions agree remarkably well, both in overall shape and in absolute values, including at the tails of the distributions. Here, T_a values below $250\,\mathrm{K}$ for the sub-millimetre channels are associated with clouds.

A more detailed comparison between simulations and observations can be seen in the interplay between channels. Figure 4 illustrates the observed and simulated joint distribution of AWS33 ($178.811\,\mathrm{GHz}$) and AWS44 ($325.15 \pm 6.6\,\mathrm{GHz}$) T_a values. These channels are paired due to having similar clear-sky Jacobians for humidity (Eriksson et al., 2025, their Figure 12), i.e. can be seen to have a similar altitude sensitivity. The figure demonstrates that the simulated observations encompass the observed joint AWS33 and AWS44 T_a values with comparable density. This shows that the simulations not only replicate the distribution of individual channel values but also successfully simulate effects that affect multiple channels in a way that is consistent with real observations.

5.1.2 Simulations over other surfaces

Over land the performance is stilled deemed good, but misses some very intense convective cores. Scenes with cold surfaces and a dry atmosphere causes the most problems, often observed at higher latitudes and over sea-ice. Work is ongoing to address these limitations and handle such cases better.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

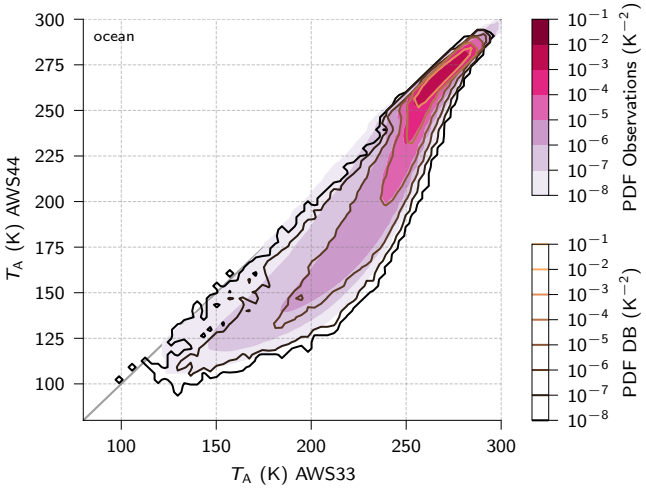

Figure 4. The 2D density plot of AWS33 and AWS44 T_a values from observed and simulated cases. The shaded areas represent the density of observed cases (from all of 2025), while the contour lines overlay the density of simulated cases. The data is filtered to include only include observations and simulations over ocean.

In the mean time, provided surface type fractions for each estimate are provided and can be used to filter out these cases. For more details on the performance over these surfaces, see Appendix A.

5.2 Retrieval performance on simulations

Retrievals in this section are performed on the held-out test-subset of the simulation database to evaluate model retrieval performance in isolation.

5.2.1 Occurrence fractions

Figure 5a-c shows the occurrence fractions of binned FWP, Z_m, and D_m values retrieved from simulated T_a values and are compared to the corresponding true database values (panels d-f are addressed in Sect. 5.3). Retrieved mean FWP values match the database distribution for values down to about $40\,\mathrm{g\,m^{-2}}$, after which retrievals overestimate the frequency of lower FWP values, indicating a loss of sensitivity in the observations.

As a consequence, Z_m and D_m values have been filtered to only include cases with significant FWP values, here the threshold is set at $0.1\,\mathrm{kg\,m^{-2}}$. For these mean estimates there is good agreement between the distribution of retrieved values and true values, although cases slightly under $5\,\mathrm{km}$ for Z_m and under $45\,\mu\mathrm{m}$ for D_m are overrepresented.

The model provides CDF estimates of the retrieved quantities, from which the mean estimated value Eq. (9) and the most probable value Eq. (10) are computed. For these statistical overviews, the most-probable estimate follows the true occurrence

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

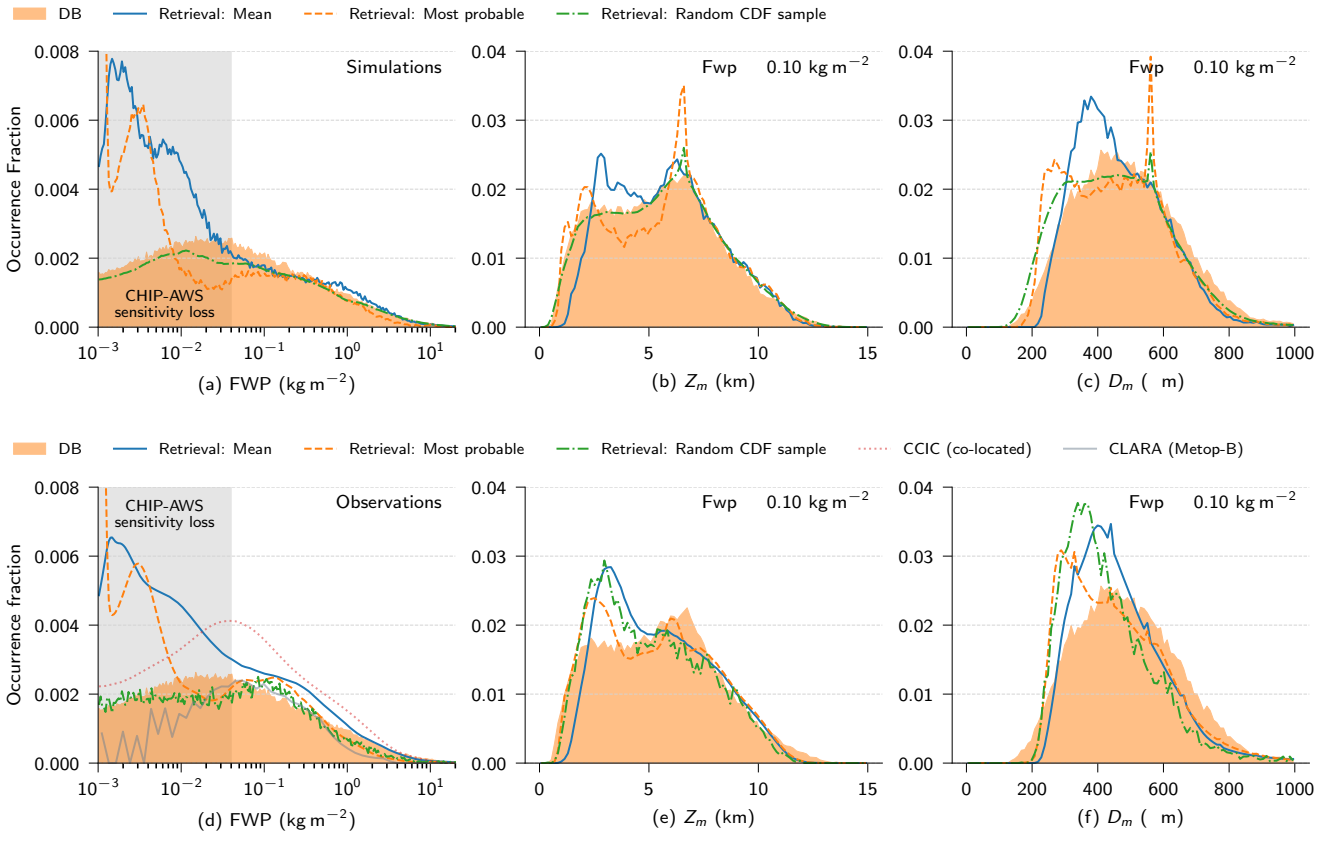

Figure 5. Occurrence fractions, i.e. the binned value counts divided by total sample count. Top row **(a-c)** shows retrievals (mean and most-probable) on simulated T_a values from the test-subset of the simulation database with the filled region (DB) being the "true" values. The "random sample" retrievals are obtained by taking a random sample according to the retrieved CDF. Bottom row **(d-f)** shows retrievals on actual observations. Co-located FWP values from CCIC and 2025 values from the CLARA record retrieved from Metop-B (ECT 09:31) are included for comparison. Same as in the top row, the filled region (DB) shows the distribution of "true" values from the test subset of the database and is included as a reference for the random CDF sampling.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

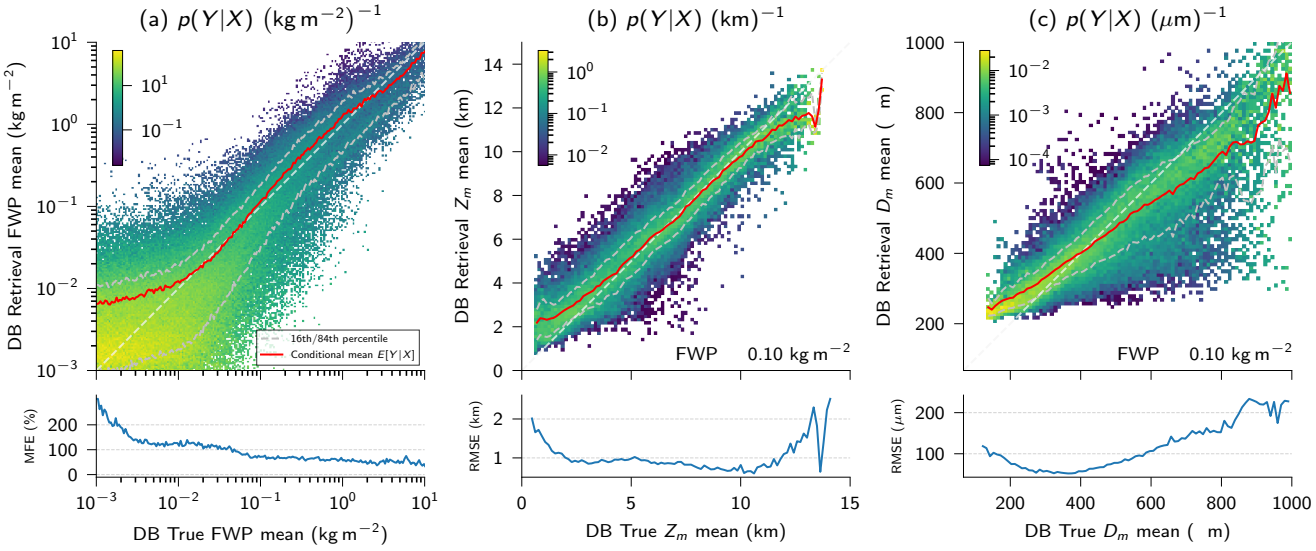

Figure 6. Conditional probability distributions of retrieved FWP, Z_m, and D_m from T_a values from the test-subset of the simulation database conditioned on the true value from the database. The bottom rows shows an error estimate for each respective quantity. For FWP, the error is expressed in median fractional error (MFE) (Holl et al., 2014, median of Eq. (1)), a metric more suitable to the exponential nature of the FWP values. For Z_m and D_m, the root mean square error (RMSE) of the retrievals are plotted for each "true" database value.

fraction distribution but also starts to deviate near the sensitivity limit. These estimates also exhibit a sharp peak of overrepresented values at 7 km for Z_m and 600 µm for D_m.

Additionally, retrieved CDFs can be randomly sampled, and these samples are expected to recreate the training data distribution (the prior). For FWP and Z_m the prior closely matches the underlying database, though with a slight underestimation of the number of samples for values below 1 $\mathrm{kg\,m^{-2}}$. For D_m, the prior is similarly close to the database distribution but shifted towards smaller values, with a small peak at 600 µm.

5.2.2 Conditional distributions

In Figure 6, conditional distributions of retrieved and true values from the test-subset of the database are shown together with their conditional mean. This is the estimated probability of a retrieved value given the true DB value. Here, retrieved and true FWP values follow each other closely along the 1:1 line from 10 $\mathrm{kg\,m^{-2}}$ down to values of 10 $\mathrm{g\,m^{-2}}$, after which the retrieved values tend to overestimate the mass of ice. This echoes the result for the occurrence fractions (Figure 5a), where a loss of sensitivity starts to be seen for values below 40 $\mathrm{g\,m^{-2}}$. The distributions for Z_m and D_m, after being filtered for cases with FWP $\geq 0.1\,\mathrm{kg\,m^{-2}}$, show agreement between retrievals and true values. However, for D_m, retrievals tend to overestimate particle sizes under 400 µm and underestimate the sizes above. These results give an indication of the retrieval bias compared to radar-derived estimates used for the simulation.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

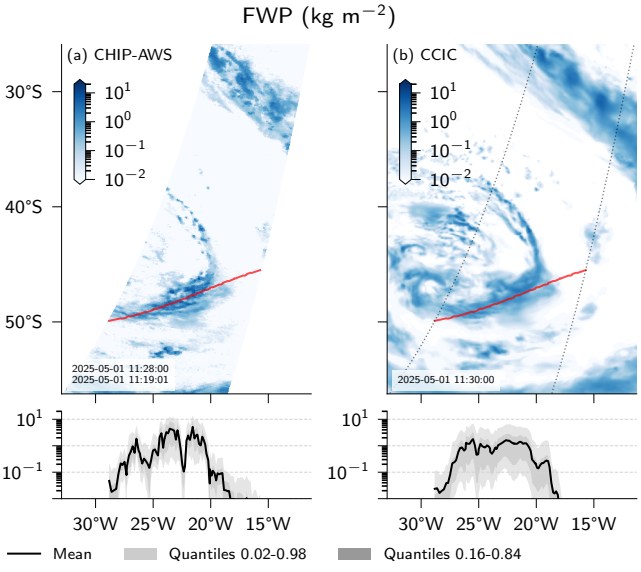

Figure 7. Comparison of real observations from the same scene between **(a)** CHIP-AWS and **(b)** CCIC. Top row shows the mean FWP estimate for each footprint in the scene. Bottom row shows the estimates and uncertainties sampled along the transect line and projected onto the longitude axis.

To get an indication of precision two metrics are used: the median fractional error (MFE) (May and Eriksson, 2025, Eq. (3)) for FWP and the root mean square error (RMSE) for Z_m and D_m. For FWP, the MFE is approximately 100% for lower values, 50% for values above $0.1\,\mathrm{kg\,m^{-2}}$, corresponding to median errors of $2\times$ and $1.5\times$ the true value, respectively. For Z_m, RMSE remains below 1 km across the 2–12 km range. For D_m, RMSE is below $200\,\mu\mathrm{m}$.

5.3 Retrieval performance on observations

5.3.1 Case comparison

Figure 7 presents a case-specific comparison between CHIP-AWS and CCIC, covering the same weather system. Both datasets capture the system and the main features in the overall shape and in the values sampled along the transect. Looking closer, the AWS-based retrievals exhibit a higher horizontal resolution compared to the CCIC retrievals, showing more detail and underlying structure, for example seen in the prominent large spiral arm at 30° S. Both datasets have per-observation uncertainties shown in the bottom row for the transect, in this case CHIP-AWS produces tighter uncertainty bounds.

Referring back to the conditional PDF of simulated retrievals (Fig. 6a), CHIP-AWS values are expected to exhibit little to no bias for all FWP values down to approximately $10\,\mathrm{g\,m^{-2}}$. However, as will be shown in the following section (Sect. 5.3.2), a bias-free estimate cannot be assumed for CCIC values away from $1\,\mathrm{kg\,m^{-2}}$.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

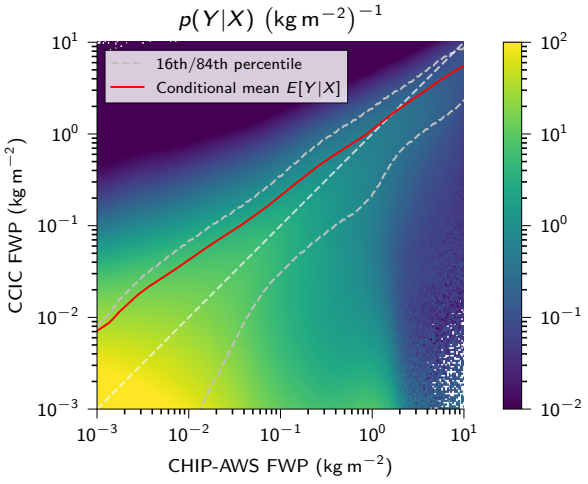

Figure 8. Conditional probability density function of co-located CCIC FWP mean estimates conditioned on CHIP-AWS FWP mean estimates. It is the probability of an observed CCIC FWP value given the corresponding CHIP-AWS estimate. The figure also includes the conditional mean $\mathbf{E}$[CCIC|CHIP-AWS].

5.3.2 Conditional PDF with CCIC

Figure 8 shows the probability of co-located CCIC FWP values conditioned on the corresponding CHIP-AWS estimates. The conditional mean shows no bias relative to CHIP-AWS for values close to $1\,\mathrm{kg\,m^{-2}}$, a positive and increasing bias for values below this, and a negative bias for values above.

This trend closely matches that in Amell et al. (2024) (their Fig. 6, equivalent of our Fig. 6), which shows a conditional PDF of CCIC retrievals given 2C-ICE values (CCIC's training data). There too, the conditional mean shows no bias at $1\,\mathrm{kg\,m^{-2}}$, but increasing bias for values below and negative bias for values above.

The fact that our comparison with CCIC (Fig. 8) recreates the conditional mean trend for CCIC and its training data suggests that CHIP-AWS is closer to 2C-ICE than CCIC. This is discussed further in Sect. 6.

5.3.3 Occurrence fractions

Figure 5d-f shows occurrence fraction plots of binned FWP, Z_m, and D_m values retrieved from actual AWS measurements for 2025. Overall, the retrieval mean and most-probable estimates from real measurements give an occurrence fraction that must be considered good for passive observations. The distribution follows the "true" distribution of radar-derived FWP but remains below the CCIC distribution. The other variables also follow the database distributions, though with a tendency to overestimate the frequency of lower values.

The retrieval mean and most-probable estimates behave similarly to the synthetic case (Fig. 5a-c). For FWP, however, retrieved values between 0.04 and $1\,\mathrm{kg\,m^{-2}}$ are more frequent than those retrieved from the simulated T_a values. This dis-

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

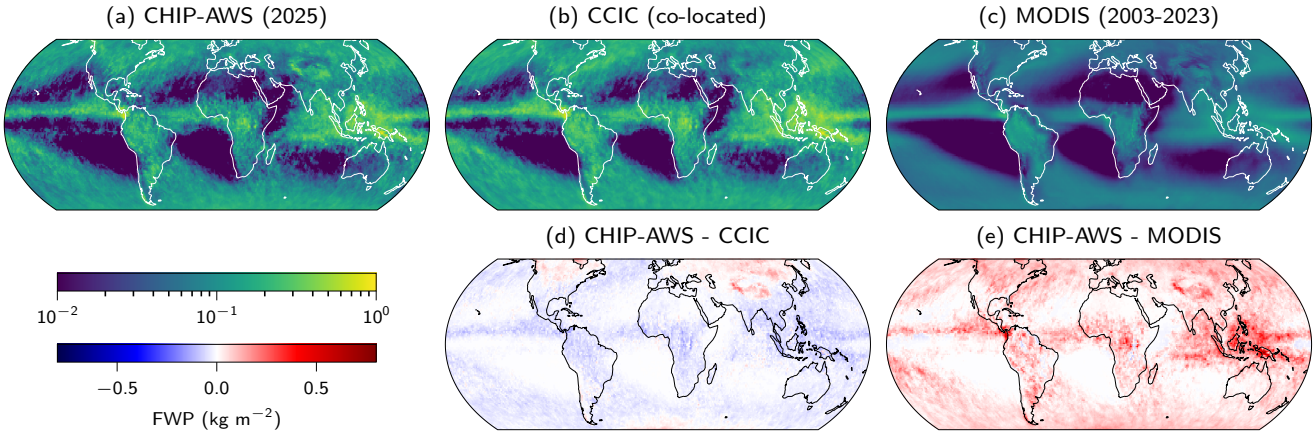

Figure 9. Spatial distribution of FWP estimates binned and averaged at $1^\circ \times 1^\circ$ resolution for 2025 from **(a)** CHIP-AWS and **(b)** CCIC. CCIC retrievals are collocated with CHIP-AWS footprints from all of 2025. A longer-timescale comparison is provided by **(c)** MODIS FWP averaged over two decades (2003–2023). Panels **(d)** and **(e)** show the differences obtained by subtracting CCIC and MODIS, respectively, from the 2025 CHIP-AWS estimates.

crepancy likely reflects differences between real observations and simulated training data, causing the retrieval method to overestimate the frequency of these when applied to real measurements.

In Figure 5d-f, the database serves as a rough reference rather than a direct benchmark. However, the random CDF sample estimates should recreate the database distribution. This holds for FWP down to the sensitivity limit. For Z_m and D_m, a peak at lower values deviates from the database distribution.

For comparison, estimates from two other passive FWP retrievals are included: co-located CCIC and 2025 CLARA estimates (Karlsson et al., 2023) from Metop-B that has an equatorial crossing time of 09:31, similar to AWS. However, CLARA's retrievals operate only during morning passes, whereas CHIP-AWS operates during both morning and evening passes. CHIP-AWS and CCIC show good agreement at the highest FWP values. Below that threshold, CHIP-AWS gives fewer cases than CCIC until it looses sensitivity at 40 $\mathrm{g\,m^{-2}}$. CLARA underestimates the number of high FWP values compared to the database, CCIC and CHIP-AWS.

5.3.4 Spatial distribution

Figure 9 shows the global spatial distribution of the mean FWP estimates from CHIP-AWS and CCIC (co-located) for the year 2025, averaged estimates from MODIS (2003–2023), as well as the difference between them. The CHIP-AWS estimates have no coastline artefacts that cannot be attributed to orographic features. This indicates that, even though simulations over land did not achieve as close of a match in distribution of simulated T_a as over the ocean (Sect. 5.2), the trained model still generalizes over the two cases and produces reasonable estimates, with one major exception: dry cases over snow or ice. Here, the retrieval method struggles.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

Looking at the difference between CHIP-AWS and CCIC (d), clear artefacts of overestimated values are apparent over Tibet in Asia, Hudson Bay and the Saint Elias Mountains in North America, as well as general snow-cover and sea ice towards Antarctica. This has been attributed to incorrect assumptions about surface reflectivities used in the simulations for the new sub-millimetre frequencies. The dataset provides surface type fractions for each observation, allowing users to filter out cases with snow or sea ice. In other areas, CHIP-AWS tends to estimate lower FWP values than CCIC. This is expected and echoes the result from the conditional PDF with CCIC (Sect. 5.3.2), where CCIC tends to have a higher bias than CHIP-AWS for most values. Comparing with MODIS (panel c and e), CHIP-AWS values are consistently higher.

This general bias, where CCIC tends to have higher and MODIS has lower values, also shows up when inspecting the zonal mean (following section).

5.3.5 Zonal means

Figure 10 shows the average FWP values binned by latitude for 2025, alongside the Z_m and D_m estimates. We compare these with historical data from the CloudSat- and CALIPSO-derived DARDAR dataset for 2007–2010, which include both morning and evening observations. After 2010, the cloud radar was deactivated for evening passes to address power issues. Co-located FWP values from CCIC are also included for comparison and can serve as a proxy for 2C-ICE estimates, another CloudSat- and CALIPSO-derived dataset.

Between 2C-ICE and DARDAR, two independent cloud property retrievals using the same instruments, there is a 20% difference in average values (Eriksson et al., 2026). Although cloud radar and lidar retrievals are expected to provide the best available benchmarks, considerable uncertainty remains about the true values. The CHIP-AWS FWP zonal mean closely follows the shape of the CCIC values, suggesting a near-constant ratio between them, though CCIC values are biased higher. The close agreement with CCIC is expected, since both datasets sample from the same atmospheric conditions due to being co-located. The CHIP-AWS dataset aligns more closely with DARDAR than CCIC, especially in 0° N – 20° N. Near the dataset's latitudinal bounds ($\pm 60^\circ$), CHIP-AWS FWP estimates increase above both CCIC and DARDAR, again attributable to poor retrieval performance for dry cases over ice and snow.

Other passive FWP estimates are also shown: MODIS for 2003–2023 and CLARA for 2025 (from Metop-B). Both fall substantially below DARDAR, CCIC and CHIP-AWS. That these passive retreivals fall below radar retrievals is consistent with previous results (Eliasson et al., 2011; Duncan and Eriksson, 2018).

For Z_m, our estimates match the DARDAR zonal mean, with a slight positive bias, here radar-lidar retrievals are expected to be a precise reference.

For D_m, again, our estimates appear to match DARDAR retrieved values, with some areas of deviation. However, here, radar-lidar retrievals have limitations with respect to particle sizes (see introduction). The combination of channels offered by AWS have the potential to constrain D_m better (Pfreundschuh et al., 2020), but this benefit is difficult to validate.

Figure 11 presents a Hovmöller diagram showing temporal variability in FWP and Z_m through zonal means averaged with a 3-week rolling window. The diagram reveals clear seasonal movement of the ITCZ. It also shows increased FWP poleward

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

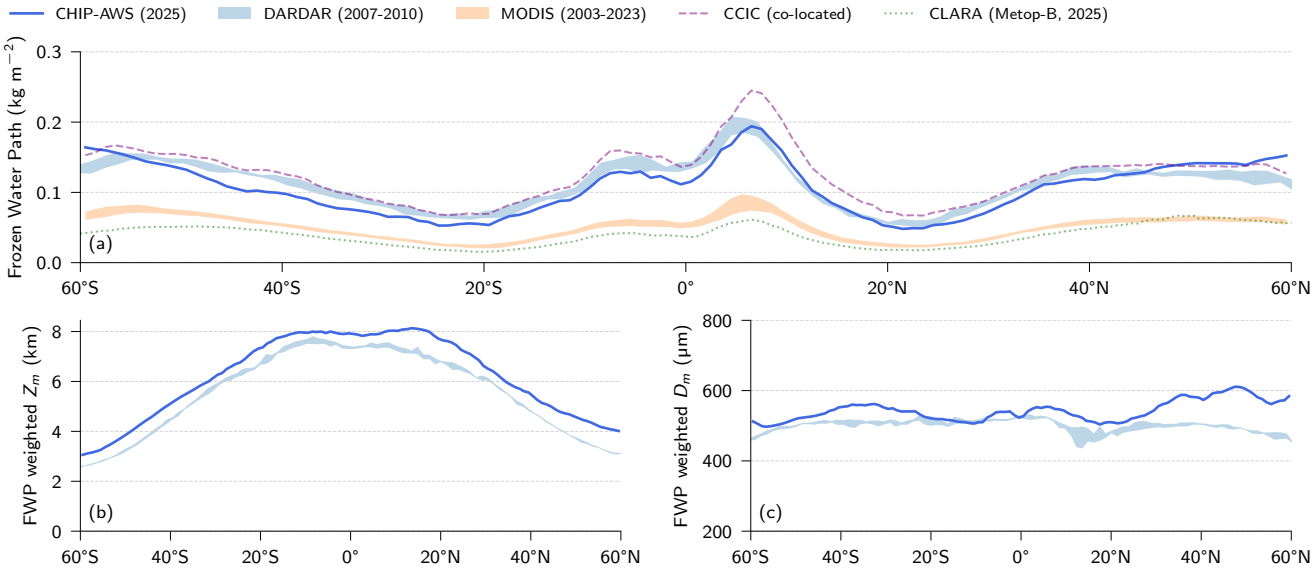

Figure 10. Zonal means of retrieved quantities from CHIP-AWS for 2025, binned by latitude. Z_m and D_m values are weighted by their corresponding FWP estimates, as these retrievals represent mass-weighted averages. For comparison, retrievals from CCIC, DARDAR, MODIS, and CLARA are also shown. CCIC has been co-located to the AWS footprints, enabling direct FWP comparison, while DARDAR allows historical comparison across all variables: FWP, Z_m, and D_m. MODIS and CLARA provide additional FWP comparisons from independent passive retrieval records. Metop-B retrievals from CLARA are selected for 2025 to match the current CHIP-AWS record, as its equatorial crossing time closely matches that of AWS. Note, however, that CLARA provides retrievals for the descending pass only.

of $\pm 40°$ N during local winter whilst Z_m decreases. This demonstrates that the dataset captures expected seasonal variability across the year. Notable, since the retrieval model receives no information about time or location.

5.3.6 Across-track bias

Figure 12 shows the deviation of the mean for each field-of-view (scan angle) from the overall mean of all retrieved values for 2025. This indicates a small varying bias across the scan for FWP and Z_m, albeit small; a change of $4\,\mathrm{g\,m^{-2}}$ and $75\,\mathrm{m}$, respectively, across the swath. A cyclic pattern is seen in the deviations. The remap distance subplot, hints that this is likely due to the re-interpolation process (Sect. 2.1). Still, the observed field-of-view dependent bias is negligible compared to the retrieval accuracy and can be ignored.

6 Discussions

Synthetic retrievals performed on simulated instrument measurements were used to characterise the retrieval performance. The conditional distributions of CHIP-AWS estimates and true database values (Fig. 6) shows negligible bias across three orders of

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

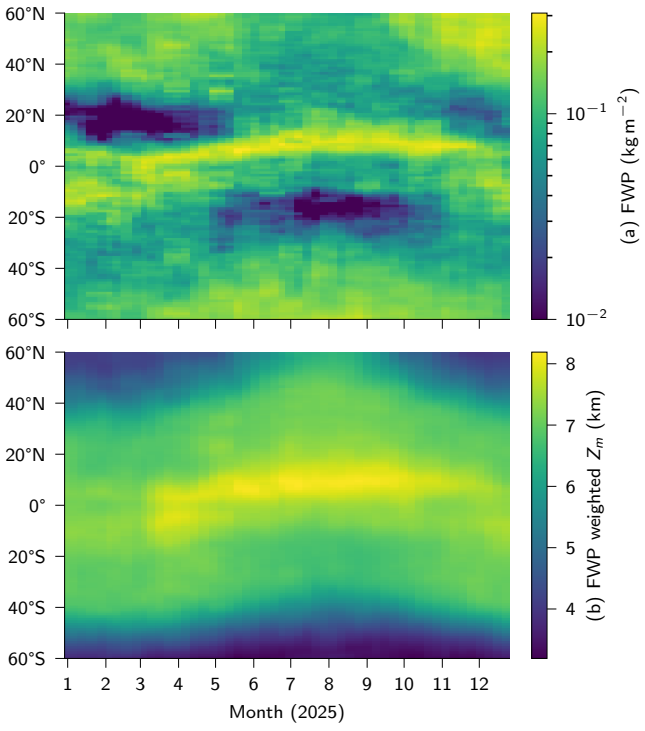

Figure 11. Hovmöller diagram (the zonal mean over time) of retrieved **(a)** mean FWP and **(b)** FWP weighted Z_m. Values are binned into latitudes for every week of the year and then averaged over time with a rolling 3 week window.

magnitude for FWP, negligible bias for almost the entire relevant range for Z_m, and D_m with a linearly varying bias with best performance at 400 µm. The comparison database values are radar retrievals. Precision metrics (MFE and RMSE) also show good results: 50 % MFE for FWP above 0.1 $\mathrm{kg\,m^{-2}}$, $\leq 1\,\mathrm{km}$ RMSE for Z_m, and $\leq 200\,\mu\mathrm{m}$ for D_m.

For comparison, SPARE-ICE, a passive FWP retrieval from combined microwave and infrared measurements, achieves an MFE for FWP of about 100% for values below 1 $\mathrm{kg\,m^{-2}}$, and 50% for values above (Holl et al., 2014, their Figure 7). Thus, CHIP-AWS achieves the same 50% MFE at a lower threshold of 0.1 $\mathrm{kg\,m^{-2}}$. However, SPARE-ICE maintains a 100 % MFE for values down to 1 $\mathrm{g\,m^{-2}}$ thanks to the IR information, whereas CHIP-AWS's MFE increases to about 250 % for the same values.

The strong agreement between simulated and observed T_a values (Fig. 3 and Fig. 4) make it likely that the characterisations made on synthetic data (Fig. 6) also hold for real observations, though compensating errors may exist in the simulated data. Ideally, we would validate against co-located radar retrievals. CloudSat was decommissioned in 2023, eliminating this option. EarthCARE (launched 2024) and the ACM-CAP (Mason et al., 2023) dataset could be used, but ACM-CAP remains under active development with no stable official retrievals available at the time of writing. Moreover, we have verified that EarthCARE

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

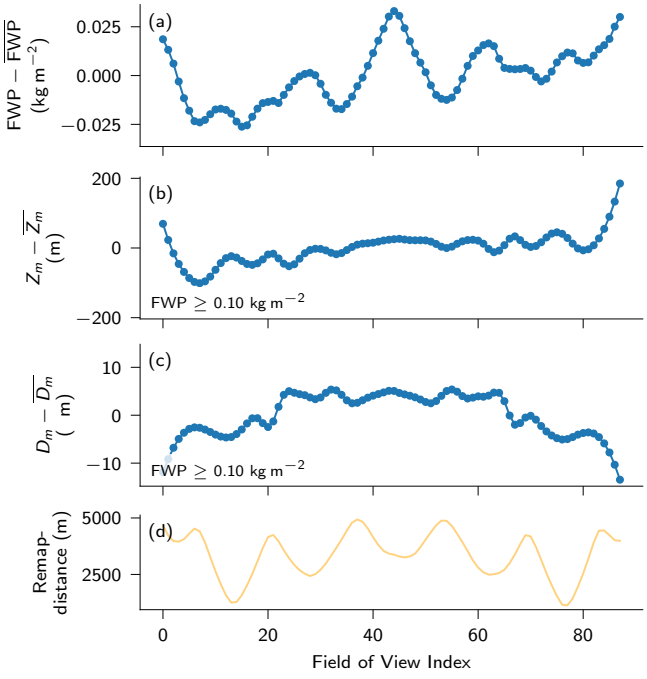

Figure 12. Per field-of-view deviation from the overall 2025 annual mean for **(a)** FWP, **(b)** Z_m, and **(c)** D_m, illustrating the magnitude of any field-of-view-dependent biases. Panel **(d)** shows the per field-of-view remap-distance between AWS channel footprints for the re-interpolation process.

has no co-locations with AWS outside of the poles. We therefore have to compare against historical datasets and other passive retrievals.

In a statistical comparison with zonal means (Fig. 10), CHIP-AWS achieves a very good match with historical DARDAR data. Recent machine-learning approaches trained on radar retrievals have achieved similar results (Yang et al., 2026; Amell et al., 2024), represented here by CCIC. Other passive retrievals, by contrast, show strong underestimation of mean FWP values, consistent with previous studies (Eliasson et al., 2011; Duncan and Eriksson, 2018).

CCIC also achieves a good result in the zonal mean and is currently the only dataset allowing substantial co-location with AWS. The conditional probability distribution (Fig. 8) reveals that CCIC exhibits varying bias relative to CHIP-AWS. Notably, the bias trend mirrors that between CCIC and its training data (2C-ICE) (Amell et al., 2024, their Figure 6). This suggests that CHIP-AWS is more consistent with 2C-ICE as an independent reference. This consistency suggests that the strong synthetic validation results likely extend to real observations to some degree.

Together, these case-specific and statistical results demonstrated good performance characteristics and provide the practically achievable validation that these characteristics extend to real observations. This datasets represents a passive retrieval method

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

of FWP, Z_m, and D_m with accuracy comparable to cloud profiling radars over a wide dynamic range, yet with a broad spatial coverage that only passive instruments can provide. This is a valuable complement to existing passive and active retrievals.

7 Conclusions

The first version of the "Chalmers Hydrometeor Inversion Product from the Arctic Weather Satellite" (CHIP-AWS) is introduced. This is a dataset providing estimates of frozen water path (FWP), mass-weighted mean altitude Z_m, and mass-weighted mean particle volume-equivalent diameter D_m retrieved from antenna temperature (T_a) values measured by AWS's across-track scanning radiometer, which features novel sub-millimetre channels. This marks the first regular dataset utilising improvements that sub-millimetre frequencies can offer to ice hydrometeor property measurements.

The data product begins in 2025 and is regularly expanded as more AWS data become available; an archival copy of 2025 retrievals is also published (see the data availability section).

The retrieval method employs QRNNs, which provide per-retrieval uncertainties in terms of quantiles. Training data were generated through rigorous radiative transfer simulations, yielding a physics-based database. This approach was necessitated by the absence of any prior Earth-facing sub-millimetre data with broad geographical coverage. Simulations also offer the advantage of making all physical assumptions explicit, enabling new insights into cloud microphysics. A key limitation of simulation-based training data is that erroneous assumptions can affect retrievals.

Simulation performance is best over oceans, where simulated T_a values agree remarkably well with observations. Over land, rare intense convective cores affect observations in a ways not captured by the simulations. Ice- and snow-covered surfaces in dry or high-altitude regions cause the most significant discrepancies. Users of the dataset should consider using the provided surface type fractions to filter retrievals according to their use-case.

CHIP-AWS achieves independent agreement with the zonal mean of radar-lidar based retrievals from DARDAR (Cazenave et al., 2019) for all three variables. This is, to the best of our knowledge, the first simulation-based passive retrieval method to do so (Duncan and Eriksson, 2018; Eriksson et al., 2026).

The dataset achieves negligible bias in retrieved FWP over a wide range, from $10\,\mathrm{kg\,m^{-2}}$ down to about $40\,\mathrm{g\,m^{-2}}$, which is unprecedented for passive retrievals. The estimates also show good precision: 50% MFE for values above $0.1\,\mathrm{kg\,m^{-2}}$ for FWP; RMSE below $1\,\mathrm{km}$ for Z_m; and RMSE below $200\,\mu\mathrm{m}$ for D_m.

This performance is achieved whilst providing a swath of $800\,\mathrm{km}$, substantially wider than the $1\,\mathrm{km}$ swath of cloud profiling radars. Compared to CCIC, CHIP-AWS measures ice mass properties more directly via the AWS instrument, which supports the results of more precise and unbiased local-scale estimates with higher horizontal resolution.

Ongoing work aims to improve the performance over ice and snow, extend the latitudinal coverage from $\pm 60°$ to global coverage, and expand the scan angle to increase the swath size. A future version of the dataset will also have aggregated level 2 files into a gridded level 3 dataset to facilitate analysis. Retrievals will also be extended to estimate vertical profiles of frozen water content (FWC) and D_m, following May and Eriksson (2025).

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

We conclude by noting that this work and future developments are directly applicable to EPS-Sterna, a confirmed constellation of AWS satellites with a 13-year mission duration beginning in 2029. This work is also relevant to the Ice Cloud Imager (ICI) onboard the Metop-SG-B satellites, which are sensitive to sub-mm frequencies up to 664 GHz. The first Metop-SG-B satellite is scheduled for launch in late 2026. We aim to develop consistent retrievals for both ICI and the EPS-Sterna radiometers.

These missions will support each other by increasing coverage and providing cross-reference measurements in multiple ways. First, the two instruments have different scanning geometries: AWS uses an across-track scanner while ICI is a conically scanning instrument, and these complementary views can help constrain particle model assumptions. Second, ICI measures polarisation in both H and V, yielding richer information on particles properties and orientation. Insights can be incorporated into the simulation assumptions, improving retrievals for both instrument types. Together, they will provide an extensive record of sub-millimetre data well suited for ice mass property retrievals extending to 2045.

Data availability. An archival copy of data for 2025 described in this work can be accessed from the Swedish National Data Service under data DOI https://doi.org/10.71870/v9dx-rm43 (McEvoy, May, and Eriksson, 2026). Manuscript reviewers can preview the dataset via https://researchdata.se/en/catalogue/dataset/2026-135/1?previewToken=859edaac-fe43-4f42-b1ad-f86a0d4dc4a2. Upon manuscript acceptance, the reserved DOI will be activated and dataset will be publicly published.

Data from 2025 and onwards are continuously processed and made available at http://clouds-and-precip.group/datasets/chip-aws. The processor version corresponding to this publication is `chip-aws-v1.0.0`.

Input AWS Level 1b data is distributed by EUMETSAT under collection id `EO:EUM:DAT:0905`.

Author contributions. All authors contributed to the development of the database generation and retrieval framework. PE and EM provided input on all results and contributed to revisions of the paper. PM led the writing of the paper and performed the visualisation and analysis of the results.

Competing interests. The authors have no competing interests to declare.

Acknowledgements. This work on AWS by PE, EM and PM is supported by the Swedish National Space Agency grants 2021-00077 and 2023-00139.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

Appendix A: Antenna temperature distributions over surface types

As mentioned in Sect. 5.1, over land, the distribution of simulated T_a values does not perfectly match the distribution of observed T_a. Figure A1 shows the T_a distributions of simulations and observations decomposed into the different surface types: all, ocean, land, and a combined category of snow, ice, and glacier.

Over land there is a discrepancy for the channels with altitude sensitivity closest to the surface, AWS 31 (165.4 GHz) and even more so for AWS 21 (89.2 GHz). For AWS 31, the simulated T_a values cover the range of all observed values, but underestimate the number of cases for values above 160 K. For AWS 21, simulated values only go down to 200 K, while observations show values down to 120 K. That the number of cases for each T_a value is not as frequent as observed is not ideal, but can be handled. Since there, at least there are simulations that produce these values, just not with the expected frequency. More problematic is when observed values ranges are not covered at all by simulations, as is the case for AWS21. The differences observed here have been attributed to unhandled surface effects, particularly for frozen (but non snow- or ice-covered) land and for certain cloud microphysics for very intense convection observed over land. This will impact retrieval performance for these cases.

For snow, ice, and glacier, there is a bigger problem. Here the distributions suggest shifted values, giving non-matching value ranges between simulations and observation for all channels except for AWS41 and AWS36. These simulated values all have a negative bias compared to observed values, causing values to be outside of the range of values seen in the observations. Since more channels have this discrepancy, the retrievals will be strongly impacted, especially when in dry atmospheric conditions or high altitude surfaces, where surface effects impact more. These artefacts are seen in the spatial distribution plot (Fig. 9, Sect. 5.3). This mismatch between observations and simulations is attributed to incorrect assumptions on surface reflectivities for snow and ice.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

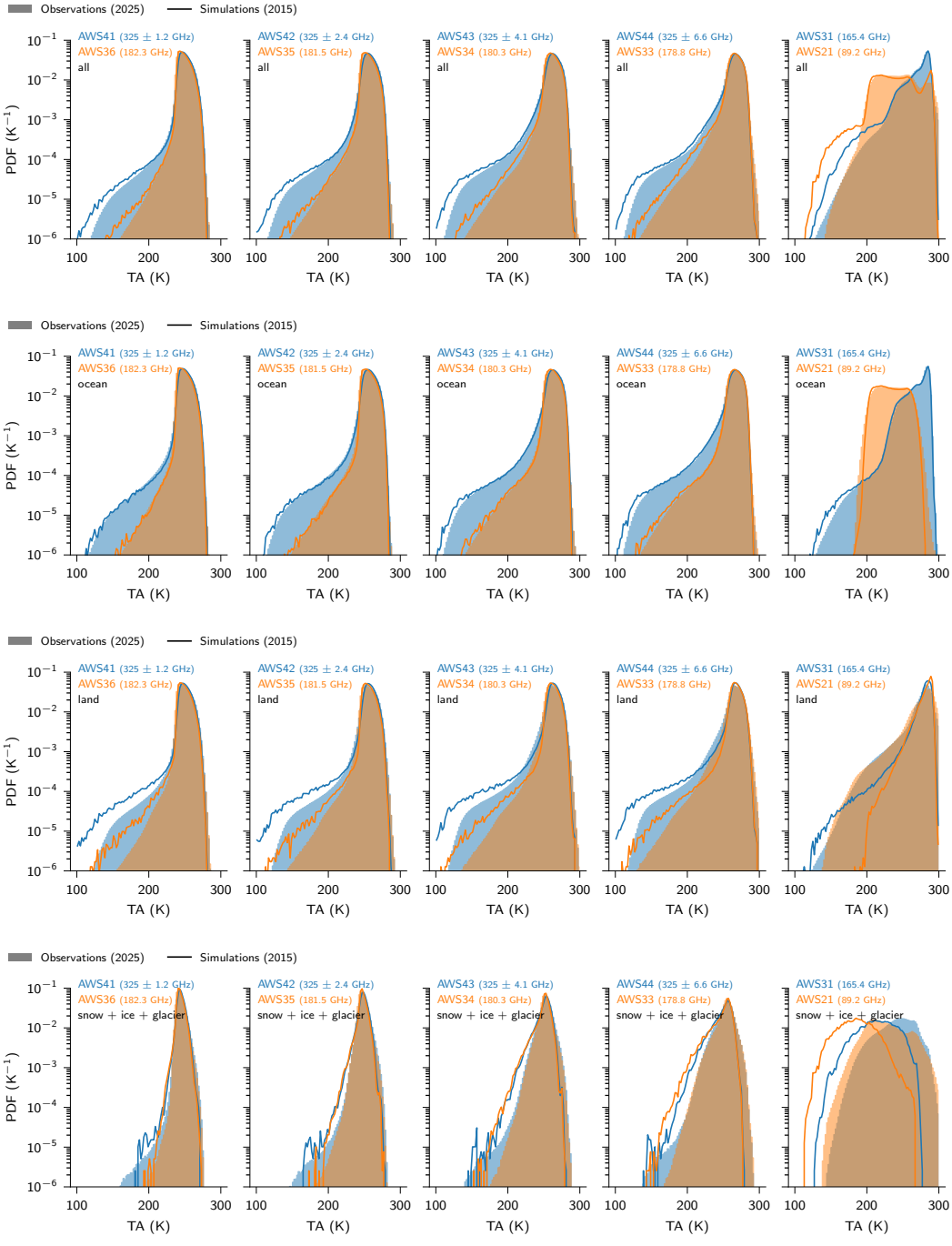

Figure A1. Distribution of simulated database T_a values (driven by CloudSat radar reflectivities from 2015) and distribution of actual observed T_a values from AWS 2025. Channels that have similar altitude of peak humidity sensitivity are paired together in the figure.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

Latitude zone	Median	Q75	Q95
30°–60° N	24 h	143 h	168 h
0°–30° N	24 h	168 h	189 h
0°–30° S	24 h	164 h	189 h
30°–60° S	24 h	120 h	164 h

Table B1. Zonal medians of dataset revisit time statistics (descending only) for April – August 2025.

Appendix B: Revisit-time analysis

The revisit time analysis uses level 2 data from April to August 2025, excluding scans with bad quality flags. Only descending passes are included.

A grid of equal-area latitude-longitude cells of 50 km is established. For each satellite pass, observations and their timestamps are binned into these cells. A pass contributes no data to a bin if fewer than three observations fall within it.

For each bin, only observations separated by more than 30 minutes are retained to avoid counting multiple measurements from the same pass. Basic statistics for the time differences are then calculated.

These statistics reflect both the satellite's orbital behaviour and data processing effects, including data gaps and quality assurance filtering. See Table B1 and the supplemental material for the analysis code.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

References

Global 30 Arc-Second Elevation (GTOPO30), https://doi.org/10.5066/F7DF6PQS, 2018.

Amell, A., Pfreundschuh, S., and Eriksson, P.: The Chalmers Cloud Ice Climatology: retrieval implementation and validation, Atmos. Meas. Tech., 17, 4337–4368, https://doi.org/10.5194/amt-17-4337-2024, 2024.

Bennartz, R. and Wu, D. L.: NASA's Polarized Submillimeter Ice-Cloud Radiometer (PolSIR): Observing the Diurnal Cycle of Ice Clouds in the Tropics and Sub-Tropics, in: 105th Annual AMS Meeting, vol. 105, p. 446951, 2025.

Bjordal, J., Storelvmo, T., Alterskjær, K., and Carlsen, T.: Equilibrium climate sensitivity above 5 °C plausible due to state-dependent cloud feedback, Nature Geoscience, 13, 718–721, https://doi.org/10.1038/s41561-020-00649-1, 2020.

Buchhorn, M., Smets, B., Bertels, L., Roo, B. D., Lesiv, M., Tsendbazar, N.-E., Herold, M., and Fritz, S.: Copernicus Global Land Service: Land Cover 100m: collection 3: epoch 2019: Globe, https://doi.org/10.5281/zenodo.3939050, 2020.

Buehler, S. A., Defer, E., Evans, F., Eliasson, S., Mendrok, J., Eriksson, P., Lee, C., Jiménez, C., Prigent, C., Crewell, S., Kasai, Y., Bennartz, R., and Gasiewski, A. J.: Observing ice clouds in the submillimeter spectral range: the CloudIce mission proposal for ESA's Earth Explorer 8, Atmos. Meas. Tech., 5, 1529–1549, https://doi.org/10.5194/amt-5-1529-2012, 2012.

Buehler, S. A., Larsson, R., Lemke, O., Pfreundschuh, S., Brath, M., Adams, I., Fox, S., Roemer, F. E., Czarnecki, P., and Eriksson, P.: The atmospheric radiative transfer simulator ARTS, version 2.6 — Deep python integration, J. Quant. Spectrosc. Radiat. Transfer, 341, 109 443, https://doi.org/10.1016/j.jqsrt.2025.109443, 2025.

Cazenave, Q., Ceccaldi, M., Delanoë, J., Pelon, J., Groß, S., and Heymsfield, A.: Evolution of DARDAR-CLOUD ice cloud retrievals: new parameters and impacts on the retrieved microphysical properties, Atmospheric Measurement Techniques, 12, 2819–2835, https://doi.org/10.5194/amt-12-2819-2019, 2019.

Delanoë, J. M. E., Heymsfield, A. J., Protat, A., Bansemer, A., and Hogan, R. J.: Normalized particle size distribution for remote sensing application, Journal of Geophysical Research: Atmospheres, 119, 4204–4227, https://doi.org/10.1002/2013JD020700, 2014.

Deng, M., Mace, G. G., Wang, Z., and Berry, E.: CloudSat 2C-ICE product update with a new Ze parameterization in lidar-only region, Journal of Geophysical Research: Atmospheres, 120, 12,198–12,208, https://doi.org/10.1002/2015JD023600, 2015.

Duncan, D. I. and Eriksson, P.: An update on global atmospheric ice estimates from satellite observations and reanalyses, Atmos. Chem. Phys., 18, 11 205–11 219, https://doi.org/10.5194/acp-18-11205-2018, 2018.

Ekelund, R., Eriksson, P., and Pfreundschuh, S.: Using passive and active observations at microwave and sub-millimetre wavelengths to constrain ice particle models, Atmos. Meas. Tech., 13, 501–520, https://doi.org/10.5194/amt-13-501-2020, 2020.

Eliasson, S., Buehler, S. A., Milz, M., Eriksson, P., and John, V. O.: Assessing observed and modelled spatial distributions of ice water path using satellite data, Atmos. Chem. Phys., 11, 375–391, https://doi.org/10.5194/acp-11-375-2011, 2011.

Eliasson, S., Holl, G., Buehler, S., Kuhn, T., Stengel, M., Iturbide-Sanchez, F., and Johnston, M.: Systematic and random errors between collocated satellite ice water path observations, J. Geophys. Res., 118, 2629–2642, 2013.

Eriksson, P., Rydberg, B., Sagawa, H., Johnston, M. S., and Kasai, Y.: Overview and sample applications of SMILES and Odin-SMR retrievals of upper tropospheric humidity and cloud ice mass, Atmos. Chem. Phys., 14, 12 613–12 629, https://doi.org/10.5194/acp-14-12613-2014, 2014.

Eriksson, P., Ekelund, R., Mendrok, J., Brath, M., Lemke, O., and Buehler, S. A.: A general database of hydrometeor single scattering properties at microwave and sub-millimetre wavelengths, Earth Syst. Sci. Data, 10, 1301–1326, https://doi.org/10.5194/essd-10-1301-2018, 2018.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

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.

Eriksson, P., Emrich, A., Kempe, K., Riesbeck, J., Aljarosha, A., Auriacombe, O., Kugelberg, J., Hekma, E., Albers, R., Murk, A., Møller Pedersen, S., John, L., Stake, J., McEvoy, P., Rydberg, B., Dybbroe, A., Thoss, A., Canestri, A., Accadia, C., Colucci, P., Gherardi, D., and Kangas, V.: The Arctic Weather Satellite radiometer, Atmos. Meas. Tech., 18, 4709–4729, https://doi.org/10.5194/amt-18-4709-2025, 2025.

Eriksson, P., Baró Pérez, A., Müller, N., Hallborn, H., May, E., Brath, M., Buehler, S. A., and Ickes, L.: Advancements and continued challenges in observations and global modelling of atmospheric ice mass, Atmos. Chem. Phys., 26, 2741–2768, https://doi.org/10.5194/acp-26-2741-2026, 2026.

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

Evans, K. F. and Stephens, G. L.: Microwave radiative transfer through clouds composed of realistically shaped ice crystals. Part II. Remote sensing of ice clouds, J. Atmos. Sci., 52, 2058–2072, 1995.

Evans, K. F., Walter, S. J., Heymsfield, A. J., and Deeter, M. N.: Modeling of submillimeter passive remote sensing of cirrus clouds, J. Appl. Meteorol., 37, 184–205, 1998.

Evans, K. F., Walter, S. J., Heymsfield, A. J., and McFarquhar, G. M.: Submillimeter-wave cloud ice radiometer: Simulations of retrieval algorithm performance, J. Geophys. Res., 107, AAC–2, 2002.

Evans, K. F., Wang, J. R., Racette, P. E., Heymsfield, G., and Li, L.: Ice cloud retrievals and analysis with the compact scanning submillimeter imaging radiometer and the cloud radar system during CRYSTAL FACE, J. Appl. Meteorol., 44, 839–859, 2005.

Fox, S.: 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, 2758, 2020.

Gasiewski, A.: Numerical sensitivity analysis of passive EHF and SMMW channels to tropospheric water vapor, clouds, and precipitation, IEEE T. Geosci. Remote, 30, 859–870, https://doi.org/10.1109/36.175320, 1992.

Gong, J., Wu, D. L., and Eriksson, P.: The first global 883 GHz cloud ice survey: IceCube Level 1 data calibration, processing and analysis, Earth Syst. Sci. Data, 13, 5369–5387, https://doi.org/10.5194/essd-13-5369-2021, 2021.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on pressure levels from 1940 to present, https://doi.org/10.24381/cds.bd0915c6, 2023a.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., and Thépaut, J.-N.: ERA5 hourly data on single levels from 1940 to present, https://doi.org/10.24381/cds.adbb2d47, 2023b.

Holl, G., Eliasson, S., Mendrok, J., and Buehler, S.: SPARE-ICE: Synergistic ice water path from passive operational sensors, J. Geophys. Res., 119, 1504–1523, 2014.

Janowiak, J. E., Joyce, R. J., and Yarosh, Y.: A Real-Time Global Half-Hourly Pixel-Resolution Infrared Dataset and Its Applications, Bulletin of the American Meteorological Society, 82, 205 – 218, https://doi.org/10.1175/1520-0477(2001)082<0205:ARTGHH>2.3.CO;2, 2001.

Jimenez, C., Buehler, S. A., Rydberg, B., Eriksson, P., and Evans, K. F.: Performance simulations for a submillimetre wave cloud ice satellite instrument, Q. J. R. Meteorol. Soc., 133, 129–149, https://doi.org/10.1002/qj.134, 2007.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

Karlsson, K.-G., Riihelä, A., Trentmann, J., Stengel, M., Solodovnik, I., Meirink, J. F., Devasthale, A., Jääskeläinen, E., Kallio-Myers, V., Eliasson, S., Benas, N., Johansson, E., Stein, D., Finkensieper, S., Håkansson, N., Akkermans, T., Clerbaux, N., Selbach, N., Marc, S., and Hollmann, R.: CLARA-A3: CM SAF cLoud, Albedo and surface RAdiation dataset from AVHRR data - Edition 3, https://doi.org/10.5676/EUM_SAF_CM/CLARA_AVHRR/V003, 2023.

Kaur, I., Eriksson, P., Pfreundschuh, S., and Duncan, D. I.: Can machine learning correct microwave humidity radiances for the influence of clouds?, Atmos. Meas. Tech., 14, 2957–2979, 2021.

Lauer, A., Bock, L., Hassler, B., Schröder, M., and Stengel, M.: Cloud Climatologies from Global Climate Models—A Comparison of CMIP5 and CMIP6 Models with Satellite Data, Journal of Climate, 36, 281 – 311, https://doi.org/10.1175/JCLI-D-22-0181.1, 2023.

Marshall, J. S. and Palmer, W. M. K.: The Distribution of Raindrops with Size, Journal of Atmospheric Sciences, 5, 165–166, 1948.

Mason, S. L., Hogan, R. J., Bozzo, A., and Pounder, N. L.: A unified synergistic retrieval of clouds, aerosols, and precipitation from Earth-CARE: the ACM-CAP product, Atmospheric Measurement Techniques, 16, 3459–3486, https://doi.org/10.5194/amt-16-3459-2023, 2023.

Mattioli, V., Accadia, C., Ackermann, J., Di Michele, S., Hans, I., Schlüssel, P., Colucci, P., and Canestri, A.: The EUMETSAT polar system-second generation (EPS-SG) passive microwave and sub-mm wave missions, in: 2019 PhotonIcs & Electromagnetics Research Symposium-Spring (PIERS-Spring), pp. 3926–3933, IEEE, 2019.

May, E. and Eriksson, P.: The Ice Cloud Imager: retrieval of frozen water mass profiles, Atmos. Meas. Tech., 18, 7243–7266, https://doi.org/10.5194/amt-18-7243-2025, 2025.

May, E., Rydberg, B., Kaur, I., Mattioli, V., H., H., and Eriksson, P.: The Ice Cloud Imager: retrieval of frozen water column properties, Atmos. Meas. Tech., 17, https://doi.org/10.5194/amt-17-5957-2024, 2024.

McEvoy, P., May, E., and Eriksson, P.: CHIP-AWS v1.0.0 2025 archive: Atmospheric ice mass properties, https://doi.org/10.71870/v9dx-rm43, 2026.

Miao, J., Rose, T., Kunzi, K., and Zimmermann, P.: A Future Millimeter/Sub–Millimeter Radiometer for Satellite Observation of Ice Clouds, Int. J. Inf. Millim. Waves, 23, 1159–1170, 2002.

Miao, J., Johnsen, K.-P., Buehler, S., and Kokhanovsky, A.: The potential of polarization measurements from space at mm and sub-mm wavelengths for determining cirrus cloud parameters, Atmos. Chem. Phys., 3, 39–48, 2003.

Pfreundschuh, S., Eriksson, P., Duncan, D., Rydberg, B., Håkansson, N., and Thoss, A.: A neural network approach to estimating a posteriori distributions of Bayesian retrieval problems, Atmos. Meas. Tech., 11, 4627–4643, https://doi.org/10.5194/amt-11-4627-2018, 2018.

Pfreundschuh, S., Eriksson, P., Buehler, S. A., Brath, M., Duncan, D., Larsson, R., and Ekelund, R.: Synergistic radar and radiometer retrievals of ice hydrometeors, Atmos. Meas. Tech., 13, 4219–4245, https://doi.org/10.5194/amt-13-4219-2020, 2020.

Pfreundschuh, S., Fox, S., Eriksson, P., Duncan, D., Buehler, S. A., Brath, M., Cotton, R., and Ewald, F.: Synergistic radar and sub-millimeter radiometer retrievals of ice hydrometeors in mid-latitude frontal cloud systems, Atmos. Meas. Tech., 15, 677–699, https://doi.org/10.5194/amt-15-677-2022, 2022.

Pfreundschuh, S., Kukulies, J., Amell, A., Hallborn, H., May, E., and Eriksson, P.: The Chalmers Cloud Ice Climatology: A novel robust climate record of frozen cloud hydrometeor concentrations, J. Geophys. Res., 130, e2024JD042 618, https://doi.org/https://doi.org/10.1029/2024JD042618, 2025.

Rivoire, L., Marty, R., Carrel-Billiard, T., Chambon, P., Fourrié, N., Audouin, O., Martet, M., Birman, C., Accadia, C., and Ackermann, J.: A global observing-system simulation experiment for the EPS–Sterna microwave constellation, Q. J. R. Meteorol. Soc., 150(762), 2991–3012, https://doi.org/doi.org/10.1002/qj.4747, 2024.

https://doi.org/10.5194/egusphere-2026-2456
Preprint. Discussion started: 11 June 2026

Stephens, G. L., Vane, D. G., Boain, R. J., Mace, G. G., Sassen, K., Wang, Z., Illingworth, A. J., O'connor, E. J., Rossow, W. B., Durden, S. L., et al.: The CloudSat mission and the A-Train: A new dimension of space-based observations of clouds and precipitation, Bull. Amer. Met. Soc., 83, 1771–1790, 2002.

Waliser, D. E., Li, J.-L. F., Woods, C. P., Austin, R. T., Bacmeister, J., Chern, J., Del Genio, A., Jiang, J. H., Kuang, Z., Meng, H., et al.: Cloud ice: A climate model challenge with signs and expectations of progress, J. Geophys. Res., 114, 2009.

Wang, J., Racette, P., Spinhirne, J., Evans, K., and Hart, W.: Observations of cirrus clouds with airborne MIR, CLS, and MAS during SUCCESS, Geophys. Res. Lett., 25, 1145–1148, 1998.

Wu, D. L., Jiang, J. H., Read, W. G., Austin, R. T., Davis, C. P., Lambert, A., Stephens, G. L., Vane, D. G., and Waters, J. W.: Validation of the Aura MLS cloud ice water content measurements, Journal of Geophysical Research: Atmospheres, 113, 2008.

Wu, D. L., Gong, J., Deal, W. R., Gaines, W., Cooke, C. M., De Amici, G., Pantina, P., Liu, Y., Yang, P., Eriksson, P., and Bennartz, R.: Remote sensing of ice cloud properties with millimeter and submillimeter-wave polarimetry, IEEE Journal of Microwaves, 4, 847–857, https://doi.org/10.1109/JMW.2024.3487758, 2024.

Yang, Y., Dou, T., Xu, G., Zhou, R., Li, B., Husi, L., Wang, W., and Xiao, C.: FYAI: a Fengyun satellite-based dataset for atmospheric ice water path, Earth Syst. Sci. Data, 18, 1287–1305, https://doi.org/10.5194/essd-18-1287-2026, 2026.