Authors' response to reviewer #2 of the manuscript "Exploring new EarthCARE observations for evaluating Greenland clouds in RACMO2.4" by Thirza N. Feenstra et al., Atmospheric Measurement Techniques (AMT): egusphere-2025-5623

We want to thank reviewer #2 for taking the time to review our manuscript. The provided comments and suggestions are addressed below and will definitely help improve the manuscript. Responses to the individual comments are shown in red, and changes we are planning to make in the manuscript are shown in blue.

General comments:

The manuscript "Exploring new EarthCARE observations for evaluating Greenland clouds in RACMO2.4" presents an evaluation of the regional climate model RACMO using observations from the active sensors onboard the recently launched EarthCARE satellite. Evaluating the macrophysical and microphysical characteristics of clouds in weather and climate models and constraining them using observations is one of the core objectives of the EarthCARE mission. In this sense, the study is timely and well aligned with the goals of the mission.

From a modeling perspective, this work demonstrates the potential value of EarthCARE observations for model evaluation and future model improvement. From an observational perspective, it provides a useful example of how model users can practically exploit EarthCARE measurements, which may also help inform product update planning within the satellite community.

The manuscript is generally well written and scientifically relevant. It also focuses on the question of how EarthCARE observations can be used to evaluate RACMO, which fits well within the scope of AMT. However, several important issues noted below need to be addressed before the manuscript can be considered for publication. Therefore, I recommend major revisions.

Specific comments

Major comments

#1. CPR reflectivity simulation

The authors use an ATLID simulator to compare ATLID backscatter with RACMO output, which is a reasonable choice. For the CPR reflectivity comparison, however, they rely on empirical relationships (Eqs. 1-7) rather than using a scattering-based radar forward model (e.g., PAMTRA; Mech et al., 2020), which may limit the robustness of the comparison.

These empirical relationships are statistical fits derived under specific conditions and do not represent variations in cloud microphysics, particularly changes in particle size distributions and densities. For instance, the Z-T-IWC relationship from Protat et al. (2007) used in Eq. (1) is known to exhibit regional variability at reflectivities above about -15 dBZ. Similarly, the Z-LWC relationship from Matrosov et al. (2004) used in Eq. (2) was derived primarily for non-precipitating marine stratiform liquid clouds. This relationship can be sensitive to CCN conditions, and its coefficients (i.e., 2.4^2 here) may therefore vary with region and season. Moreover, the Z~LWC^2 assumption is only valid for cloud droplets and breaks down once liquid water evolves into drizzle or rain, as scattering transitions away from the Rayleigh regime. Finally, the attenuation relationship applied to snow (Eq. 3) was derived under dry snow conditions.

As a result, part of the discrepancies between the observed and simulated reflectivity shown in Figs. 3 and 8 may result from errors in the reflectivity simulation itself, not only from deficiencies in the model. These errors are expected to increase with increasing reflectivity and may therefore have influenced the authors' conclusions.

Using a radar simulator (e.g., PAMTRA; Mech et al., 2020) together with an EarthCARE CPR instrument model (e.g., Orbital-Radar tool; Pfitzenmaier et al., 2025) would likely make this comparison more robust and reliable. This would be particularly great if the microphysical assumptions used in the simulator were aligned with those used in RACMO. If this is beyond the scope of the study, the authors should at least more clearly discuss the assumptions, limitations, and potential biases associated with the empirical relationships used, and carefully consider these aspects when interpreting the reflectivity comparisons shown in Figs. 3 and 8. Where possible, the authors could also consider whether alternative empirical relationships that are more appropriate for the cloud regimes considered here might be available.

For this part of the study, our initial plan was to use a scattering-based radar simulator, as you suggest here. However, this proved to be very complicated to combine with the relatively simple RACMO single-moment microphysics. In RACMO, number concentrations are not prognostic. Some (e.g., for liquid) are determined diagnostically in the radiation scheme, while others are not computed at all. Therefore, numerous assumptions had to be made regarding the construction of the local PSDs, introducing uncertainties of the same order of magnitude or larger than those associated with the use of Z(WC, T) parameterizations from literature. Therefore, we decided to use this simpler approach, as the aim of this radar simulation was more to determine whether the conclusions we draw based on the level 2 retrieved cloud properties are in line with what we see in the level 1 data, and to find limitations of the CPR (what it can observe and what not). Considering this, we would like to stick to the use of these empirical relationships, but we will address their uncertainty more, as these reflectivities are likely more uncertain than the backscatter we compute with the ATLID simulator.

We will rephrase and add to lines 211-213:

We simulate radar reflectivity using relationships between radar reflectivity and water content. Although using a scattering-based simulator might be more sophisticated, it would involve many assumptions regarding the particle size distributions, as these are not computed in RACMO's microphysical scheme. These assumptions could introduce large errors (Moradi et al., 2026). Therefore, we rely on empirical reflectivity relationships, although these are also associated with errors, as these relationships inhibit large regional variability and are often derived for specific cloud types (Matrosov et al., 2004; Protat et al., 2007). We correct the simulated reflectivity for attenuation from precipitation, liquid water, and atmospheric gases. We neglect attenuation from ice crystals, as this is small for W-band radars (Hogan and Illingworth, 1999).

We will add to line 483:

... to the analysis. However, it should be noted that the modeled radar reflectivities might come with relatively large errors, as the used relationships between water content and reflectivity are empirically based and are derived from observations in specific regions and of specific cloud types.

We included more on the uncertainty of the interpretation of Fig. 3e-f, which is explained in our answer to minor comment #14.

We will also update the discussion of Fig. 8e-f in lines 395-396:

Considering the radar reflectivity (Fig. 8e-f), RACMO simulates cloudy regions at roughly the same locations as the CPR observations, but the simulated reflectivity is too low, which might point to simulated ice and snow water contents that are underestimated.

#2. Lack of quantitative evaluation

Throughout the manuscript, the comparison between RACMO and EarthCARE observations is often described using qualitative terms like "underestimate" or "overestimate" without a clear indication of the magnitude of these biases. For example, it remains unclear whether the differences in backscatter, reflectivity, cloud top height, or ice water content correspond to systematic biases or regime-dependent behavior (e.g., stronger overestimation for higher water contents but reasonable agreement for weaker ones).

Including simple quantitative metrics (e.g., mean, median differences, relative biases, or percentile comparisons) would substantially strengthen the conclusions. In particular, such information would be very helpful when considered alongside forward model simulation errors and observational retrieval errors, as it would clarify whether the remaining discrepancies can reasonably be attributed to RACMO itself, or whether they are comparable in magnitude to forward model or retrieval errors.

Our analysis is indeed mainly qualitative. As this study only focuses on two cases, statistical metrics might not represent the model performance accurately, and, therefore, we initially refrained from adding statistical information (apart from the histograms in Fig. 5 and Fig. 10). We plan to do a larger scale evaluation of 1-2 year of EarthCARE overpasses in the future, which would include a more statistical analysis, as this would be able to represent a larger period and the whole domain. However, for this study, after this has been pointed out by both reviewers, we see that some statistical information can strengthen our conclusions. We will therefore add some statistics and will mention explicitly that these numbers only apply to the cases analyzed in this study. We will make the following changes to the text:

Line 329:

... western part of the GrIS. For this case, most ice clouds are detected (probability of detection of 0.61), and only a few ice clouds are modeled in the wrong location (false alarm rate of 0.17). Although the ...

Line 344:

... not captured by RACMO. Because these liquid and mixed-phase layers are relatively small, modeling them in exactly the right location is difficult, which is indicated by a low probability of detection of 0.11 and a high false alarm rate of 0.96.

Line 351:

... with a mid-range IWC. On average, over the entire vertical profile, the simulated IWC is underestimated with a bias of $-5.8 \cdot 10^{-5}$ kg m^{-3} (relative underestimation of 67%) and shows relatively weak correlation ($R^2 = 0.16$) with the observed IWC.

Line 359:

... for this overpass. In line with the modeled IWC, the modeled snowfall rates over the vertical profile are on average underestimated (bias of $-4.7 \cdot 10^{-5}$ kg m^{-2} s^{-1}, equaling a relative underestimation of 65%) but show a higher correlation ($R^2 = 0.39$) with the observations. The snowfall ...

Line 404:

...of the detected mixed-phase layers, indicated by a low probability of detection of 0.06 and a high false alarm rate of 0.96 for liquid water.

Line 409:

...western part of the GrIS. This results in a slightly lower probability of detection of 0.59 and a higher false alarm rate of 0.25 for ice clouds than for the March case. Looking at the ...

Line 418:

...than the March case. Although over the vertical profile, the bias of $-5.3 \cdot 10^{-5}$ kg m^{-3} is slightly lower for this case, the relative underestimation is larger (77%), but the correlation is slightly higher ($R^2 = 0.22$).

Line 420:

... for this overpass. This is reflected in the larger negative bias of $-5.1 \cdot 10^{-5}$ kg m^{-2} s^{-1} (relative underestimation of 72%) over the whole profile and slightly lower correlation ($R^2 = 0.37$) than for the March case. Even though ...

Rephrasing and adding to lines 494-496:

While these first case studies offer meaningful insights into cloud representation in RACMO, the small number of cases analyzed results in large uncertainty regarding the discrepancies between the EarthCARE observations and RACMO model results. The numbers presented in this study should thus be treated with caution, as they represent a small sample size. Therefore, a more comprehensive evaluation based on multiple months of EarthCARE observations will be necessary for a reliable evaluation and will guide model development.

#3. Curtain-based comparison

The manuscript focuses on detailed comparisons for two selected case studies. Using a limited number of cases is not a problem, and the authors provide useful information on the environmental context of each case. However, an important limitation of the current analysis is that the model-observation comparison is restricted to cross sections.

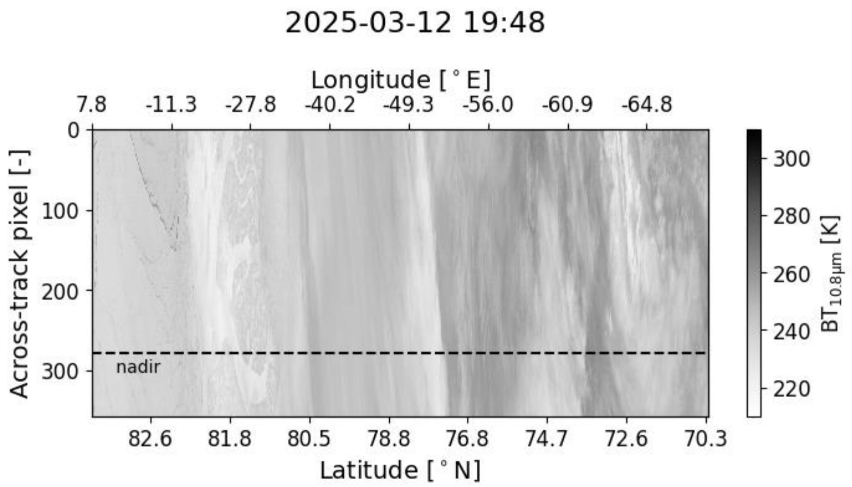
The RACMO may exhibit not only biases in cloud intensity, but also spatial displacement errors in the horizontal. For cloud systems with limited horizontal variability, this may not be a major issue. However, for more heterogeneous cloud fields, apparent underestimation or overestimation could partly reflect horizontal mismatches between the modeled and observed cloud fields. For example, in Section 3.2 (lines 304-314), the manuscript said that RACMO underestimates cloud top height and water content over the Baffin Bay region. But it is difficult to exclude the possibility that this discrepancy arises from horizontal differences in cloud location.

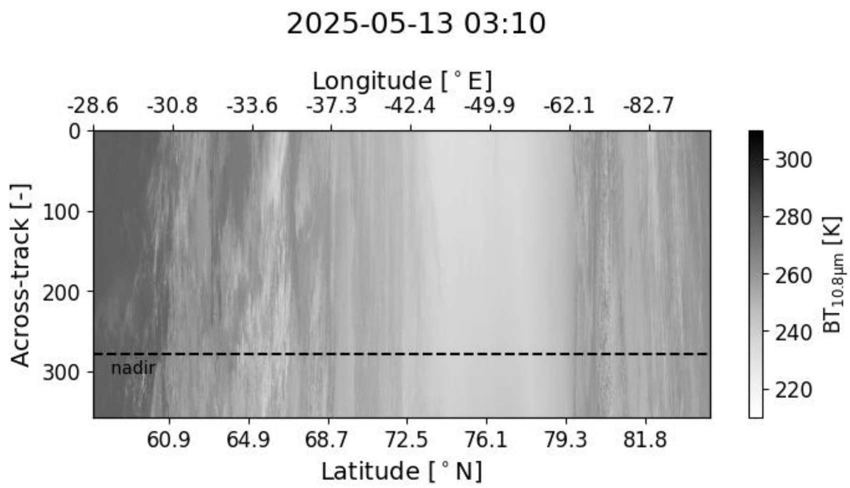
I therefore suggest that the authors more explicitly acknowledge this limitation in the manuscript. Alternatively, the authors could provide additional evidence that horizontal variability in cloud top height and water content (or simulated reflectivity) is limited for the selected cases, or include complementary analyses (e.g., CFAD-like comparisons) that better support the interpretation of systematic model biases.

Although it is indeed difficult for an RCM to model clouds exactly at the correct timing and location, the use of a relatively small domain with nudging at the upper boundary does allow for obtaining relatively good correspondence to the observed cloud state. Additionally, we do aim to get as close as possible to the correct timing and location when working towards an improved model version. Specifically for polar RCMs, it can be very important to match the timing and location as closely as possible. Specifically in areas where snow- or ice-melt occurs, capturing the location and timing of clouds is crucial, as they can strongly influence melt by altering the radiative fluxes.

However, when selecting cases, we aimed for cases where there was enough overlap in modeled and observed clouds. We indeed also found a few cases where the EarthCARE overpass was over or just next to a cloud edge, which could lead to large discrepancies with the RACMO output. Therefore, for this study, we have only considered cases where this was not the case and thus had lower across-track variability.

The low across-track variability for these cases can also be seen from the brightness temperature from the corresponding MSI observations (MSI-RGR-1C), which cover a wider (150 km) swath:





Considering this, we would like to refrain from including model data from neighboring times (as proposed by reviewer #1) and locations. We will add a sentence on the importance of co-location in general to the method in line 185:

... not to influence the analysis. As clouds strongly influence melt, it is crucial to model them in the correct location and at the correct time to capture melt patterns as accurately as possible. Therefore, using co-located profiles will yield the fairest comparison.

We will also add that, for this study, this was a constraint for the selection of cases to line 263:

... to compare the two. To achieve this, cases with low across-track spatial variability were chosen, since this prevented situations in which an EarthCARE overpass would be close to a cloud edge, where a small shift in timing or spatial patterns could result in large differences in the vertical profiles. Since models struggle ...

#4. Uncertainties of EarthCARE IWC products

In the manuscript, the EarthCARE IWC products (from ATL-ICE and CPR-CLD) are used as a reference when evaluating RACMO, but their uncertainties and possible biases are not discussed. One of the key messages of the paper is that RACMO underestimates IWC. However, the relatively high IWC observations (mainly from the CPR-CLD product) may themselves be biased high. Unfortunately, at this early stage of the mission, EarthCARE microphysical retrievals have not yet been fully validated.

Given this situation, it would be helpful if the authors treated the IWC comparison more cautiously and discussed the current level of confidence in these products. Also, pre-launch, forward model based studies (e.g., Mroz et al., 2023; Mason et al., 2023) provide useful guidance on the expected uncertainties and potential systematic biases and could be referenced in this context.

In addition, I understand that the authors used the most recent CPR-CLD baseline available at the time of their analysis (i.e., baseline BA). However, CPR-CLD is a rapidly evolving product, and noticeable changes in retrieved IWC have occurred between baseline BA and the more recent BB and BC versions (see the Product Disclaimer; https://earth.esa.int/eogateway/missions/earthcare/data). This raises the question of how sensitive the main conclusions of the manuscript are to the product version used. If possible, comparing two different CPR-CLD baselines for the selected cases would be informative. If substantial differences are found, updating the analysis to the latest available version would be recommended.

In lines 471-483, we discuss the limitations of using single-instrument IWC retrievals. We do agree that it would be valuable to discuss the uncertainties coming with these retrievals in more detail here, but, as you mention, the validation is still ongoing work. We will include more on the uncertainty regarding IWC in line 490:

... clouds and precipitation (Mason et al., 2024). Although the presented IWC profiles in this study are based on the combined ATLID and CPR observations, their individual retrievals might be biased, especially since these are actively being developed. For example, Mason et al. (2024) showed that the CPR-CLD product might miss both the lower-end and higher-end IWC values. Therefore, including the observed ice water path from the MSI can provide an additional observation to reduce biases in the IWC profiles. Additionally, heating rates …

Regarding the chosen baselines, we took the most recent baselines in October 2025 for the individual products and cases. As reprocessing is only done for baselines [x]A, for most products, the most recent baseline was baseline BA. For CPR-NOM, these were baselines CA and CB. Now (January 2026), for the dates of these cases, baseline BA is still the most recent baseline for all products used except for CPR-NOM, which is now baseline DA. However, we will change from Level 1B to Level 2A for reflectivity and backscatter (minor comment #6), and therefore, for these dates, for all products, the latest baseline is BA.

Minor comments

#1. (lines 134-135)

The wording "profiles of clouds, aerosols, and radiation" is a bit misleading. While cloud and aerosol properties are treated as vertical profiles, radiation is not.

Thanks for indicating this. We will rephrase this:

EarthCARE carries four complementary instruments to obtain comprehensive profiles of clouds and aerosols, and top-of-atmosphere radiative fluxes

#2. (line 136)

The term "molecular" is used in connection with air density. While related, these are not strictly equivalent. Please clarify this description to avoid confusion.

We will clarify this by changing this part of line 136 to:

molecular (air molecules, i.e., nitrogen and oxygen)

#3. (lines 146-147)

The statement that "the CPR can fully penetrate through clouds" is somewhat too strong. While a 94 GHz w-band radar generally has much greater penetration capability than a lidar, significant attenuation can still occur in regions with heavy precipitation or high liquid water content, potentially leading to strong signal weakening or even signal loss. Please consider refining this statement.

We will rephrase lines 146-147:

The CPR is a 94 GHz W-band radar, which complements the ATLID observations by having a larger penetration capability, which can extend up to the surface. However, in thick liquid clouds or heavy precipitating systems, the CPR will suffer from attenuation below these layers.

#4. (lines 150-159)

Please add appropriate references for the instrument specifications mentioned here.

Compared to the CloudSat radar, the CPR has an increased sensitivity of about 5 dB (Wehr et al., 2023), allowing for the detection of smaller ice crystals and low-altitude clouds. With a footprint of 750 m, compared to CloudSat's 1.3–1.8 km (Stephens et al., 2008), EarthCARE's CPR has a significantly higher spatial resolution. The vertical sampling of both radars is around 500 m (Stephens et al., 2008; Wehr et al., 2023). However, because the CPR oversamples the radar echoes at 100 m, compared to 250 m for CloudSat, the vertical resolution of the retrieved cloud profiles is higher for the CPR (Wehr et al., 2023). Additionally, this allows the CPR to detect clouds closer to the surface, compared to CloudSat. The Multispectral Imager (MSI) provides observations in the four visible and near-infrared and three infrared channels over a 150 km wide swath for scene context and additional cloud and aerosol information (Wehr et al., 2023). The synergistic retrievals based on these three instruments will yield the most accurate 3D profiles of clouds and aerosols to date. From these, radiative fluxes can be modeled, which can be compared to the top-of-atmosphere fluxes measured by the Broadband Radiometer (BBR; Barker et al., 2025).

#5. (lines 160-161)

At the time the manuscript was written, synergy products (e.g., ACM-CAP) were not yet released. But these products became available as of 1 December. While it is not necessary to use them in this study, it may be worth briefly noting that these products have become available since 1 December (the same applies to lines 238-239 and 251-252).

We will update lines 160-163:

Hence, not all EarthCARE products were released when this study was done. At the time of writing, Level 1b (calibrated satellite measurements) and Level 2a (derived cloud and aerosol properties) single-instrument products and a few Level 2b combined instrument products are available. However, more multi-instrument products have become available in December 2025.

Lines 238-239:

During this study, a combined ATLID - CPR IWC product had not been released yet.

Lines 251-252:

Since the latter only became publicly available after this study was done, ...

We will also update this in the discussion in lines 484-486:

As this analysis is based on some of the first available EarthCARE observations, calibration and validation efforts are still ongoing. This not only implies that newer, more reliable baselines of the EarthCARE products used in this study will become available, but also that additional multi-instrument synergistic products have become available from the end of 2025.

And lines 491-492:

Therefore, in future work, these multi-instrument cloud and radiation products will be used to evaluate RACMO, ...

#6. (line 161)

Is there a specific reason why Level 1B data were used for lidar backscatter and radar reflectivity? Level 2A products provide things like corrected reflectivity, which might be more useful.

There is no specific reason for this, other than that we started working with the Level 1 data at the time it became available, and have not reconsidered this when the Level 2 data became available. We will therefore change from the ATL_NOM_1B to the ATL_EBD_2A product for Fig. 3a,c and Fig. 8a,c and from CPR_NOM_1B to CPR_FMR_2A for Fig. 3e and Fig. 8e. The new figures are shown below. Since in CPR_FMR_2A the areas that suffer from surface clutter are masked out, we show these areas in grey (as in the figures in our response to minor comment #24). We will also make a few additional changes to the text:

Lines 163-166:

Therefore, we use several ESA Level 2 ATLID and CPR products that have been available since March 2025. As we are primarily interested in cloud properties, this study focuses on the Level 2a ATL-EBD (lidar backscatter; Donovan et al., 2024), ATL-ICE (ice water content; Donovan et al., 2024), CPR-FMR (radar reflectivity; Kollias et al., 2023), and CPR-CLD (water content and precipitation rate; Mroz et al., 2023), and the Level 2b AC-TC (cloud, aerosol, and precipitation classification; Irbah et al., 2023) products.

Lines 195-197:

We compare both observed backscatter and reflectivity profiles and derived cloud properties. To compare RACMO model output with backscatter and reflectivity observations, we simulate lidar Mie and Rayleigh backscatter and radar reflectivity based on the RACMO output.

Lines 267-269:

For the chosen cases, we use EarthCARE data of baseline BA for all products.

Line 285:

3.2 Comparison of simulated and observed backscatter and reflectivity profiles

Lines 314-315:

The CPR observes very high reflectivity values just above the surface due to surface backscatter. Here, the observed reflectivity is not reliable and therefore masked out (grey, Fig. 3e).

Line 381:

4.2 Comparison of simulated and observed backscatter and reflectivity profiles

Lines 499-500:

Our evaluation includes a comparison of simulated backscatter and reflectivity profiles against ATLID and CPR observations, as well as an assessment of the modeled cloud and precipitation content and phase against the EarthCARE derived cloud properties for two case studies.

Lines 521-524 (data availability):

The products used are the ATL-EBD-2A product (baseline BA; European Space Agency, 2025a), the ATL-ICE-2A product (baseline BA; European Space Agency, 2025c), the CPR-FMR-2A product (baseline BA; European Space Agency, 2025d), the CPR-CLD-2A product (baseline BA; European Space Agency, 2025e) and the AC-TC-2B product (baseline BA; European Space Agency, 2025b).

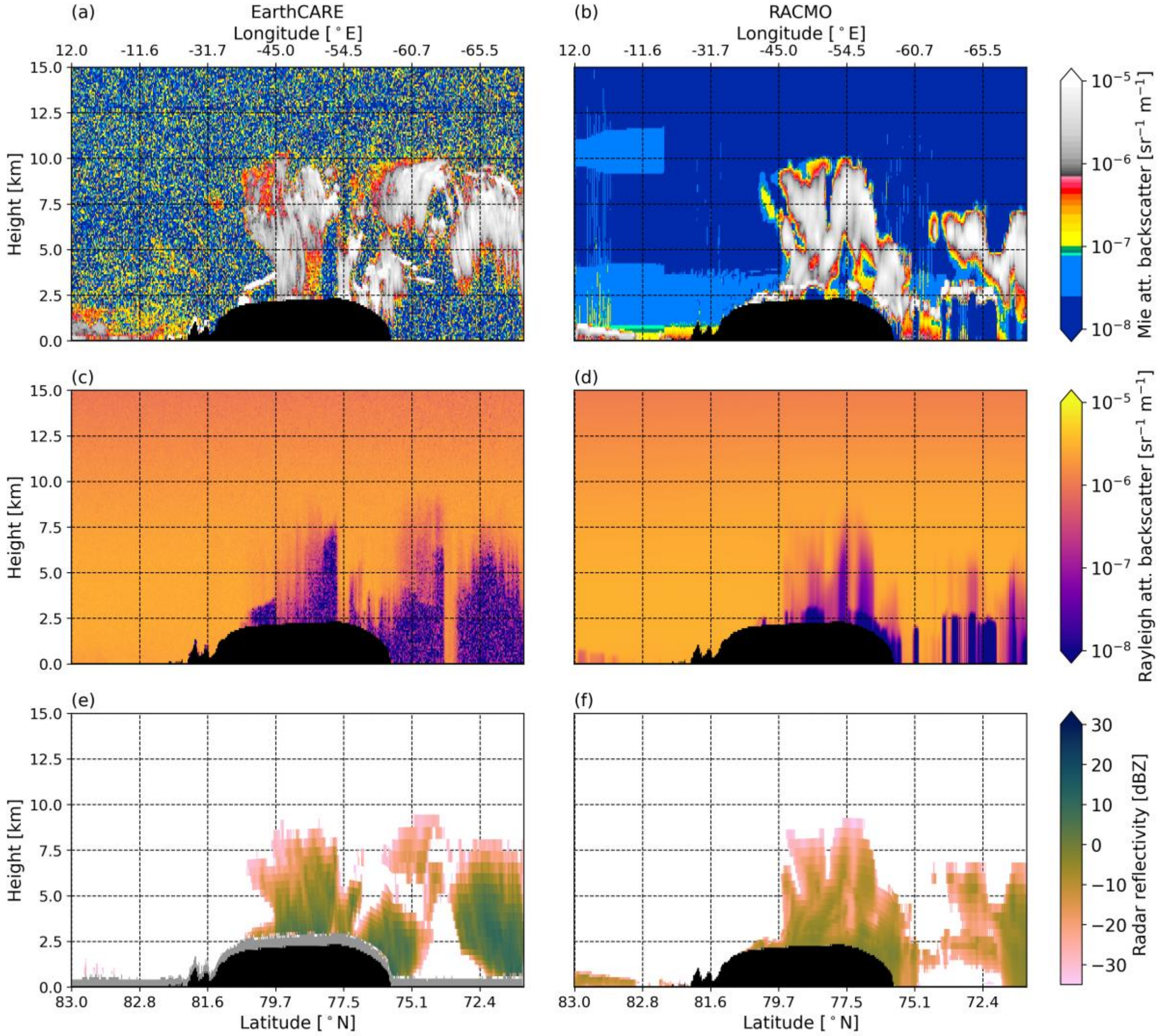


Figure 3. Profiles of the March 12th, 2025, 19:48 UTC (a,c,e) observed (EarthCARE, (a,c) ATL-EBD, baseline BA and (e) CPR-FMR baseline BA) and (b,d,f) modeled (RACMO) (a-b) Mie total (co- and cross-polar) attenuated backscatter [sr−1m −1], (c-d) Rayleigh attenuated backscatter [sr−1m −1] and (e-f) radar reflectivity [dBZ]. Note that the x-axis follows the time coordinates. Hence, the latitude and longitude coordinates do not vary monotonically. Also note that in (a-d), the vertical resolution is 100 m, while in (e-f) the vertical coordinates follow the RACMO hybrid-sigma levels. Black areas correspond to the topography. In (e), the surface clutter is shown in grey.

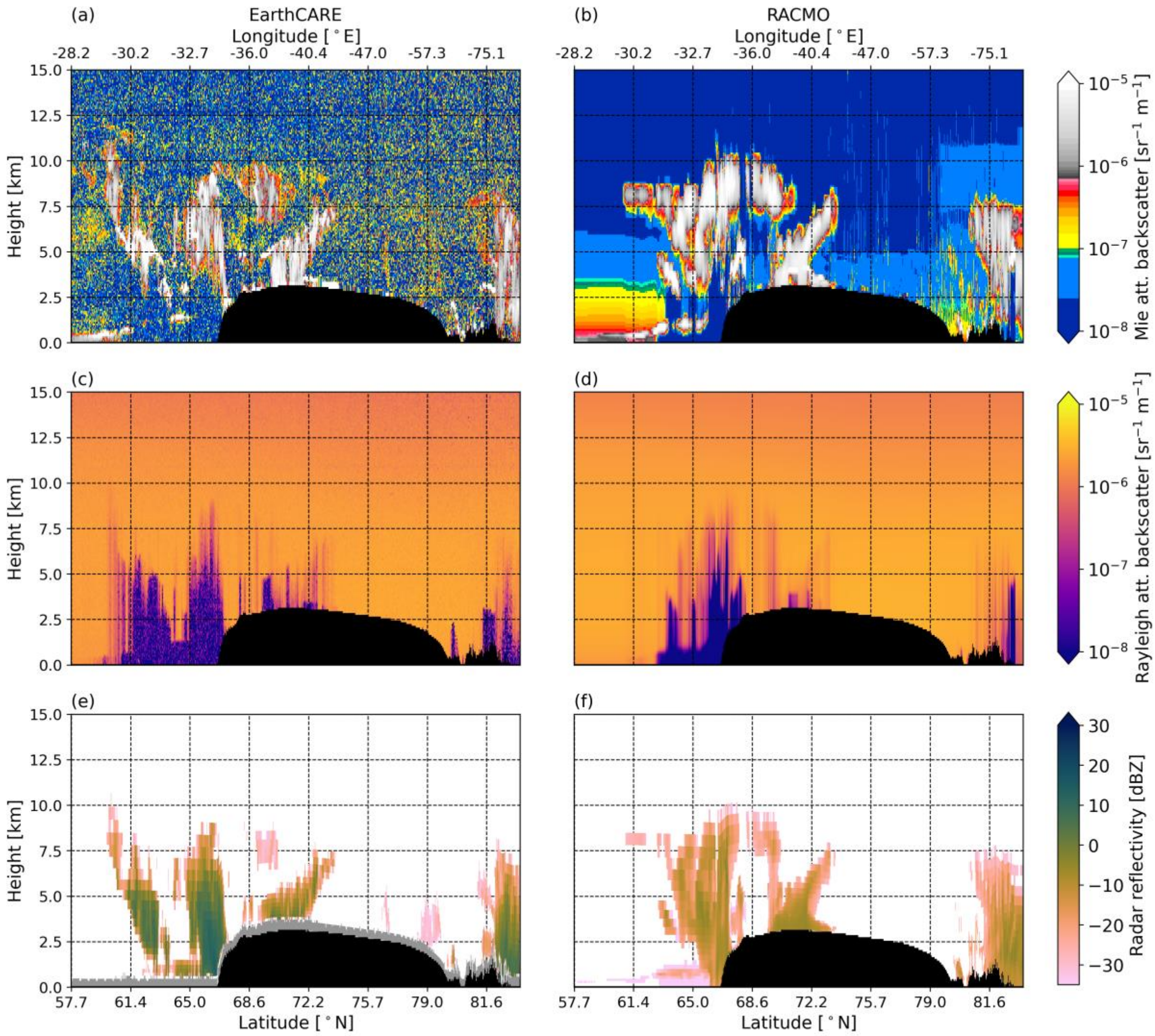


Figure 8. Profiles of the May 13th, 2025, 03:10 UTC (a,c,e) observed (EarthCARE, (a,c) ATL-EBD, baseline BA and (e) CPR-FMR, baseline BA) and (b,d,f) modeled (RACMO) (a-b) Mie total (co- and cross-polar) attenuated backscatter [sr−1m −1], (c-d) Rayleigh attenuated backscatter [sr−1m −1] and (e-f) radar reflectivity [dBZ]. Note that the x-axis follows the time coordinates. Hence, the latitude and longitude coordinates do not vary monotonically. Also note that in (a-d), the vertical resolution is 100 m, while in (e-f) the vertical coordinates follow the RACMO hybrid-sigma levels. Black areas correspond to the topography. In (e), the surface clutter is shown in grey.

#7. (lines 161-166)
EarthCARE Level-2 products are provided separately by ESA and JAXA. To avoid confusion, please clearly indicate whether the products used in this study are from ESA or JAXA.
We used the ESA level 2 products. We will changes lines 163-164:
Therefore, we use several ESA Level 2 ATLID and CPR products…
#8. (line 229)
For “attenuation for liquid water,” it might be helpful to clarify explicitly whether rain water is included or not here.
Rain is not included here. Attenuation from rain is computed using Eq. 4. We will make this distinction more clear by changing line 229:
Attenuation from liquid water clouds (excluding precipitation), …
#9. (lines 254-257)

As the authors mentioned earlier, lidar can only see the top in the presence of liquid water, and radar alone cannot directly detect supercooled liquid water. So, in mixed-phase clouds, EarthCARE can only identify the upper boundary of the supercooled liquid layer, while the phase below this layer remains uncertain. A brief reminder of this limitation here would be helpful.

We will add an additional sentence in line 254:

... the liquid water content. For the classification of liquid and mixed-phase clouds, it is important to consider that the ATLID can only detect the top of these clouds, and the CPR struggles to detect small liquid water droplets. Therefore, there is uncertainty regarding the thickness of these liquid layers and the cloud phase below the liquid and mixed-phase cloud tops. For a direct ...

#10. (line 259)

While it is correct that Level-2 data became publicly available in March 2025, observations are also provided for earlier periods. So, the data release date itself does not seem directly related to the choice of cases after March 2025.

At the time of writing, the reprocessing of the entire dataset had not been completed yet. Therefore, at that time, we could only use data from March 2025 onward. However, now that the reprocessing is completed, this is not a restriction anymore. We will add some clarification:

... data became available, and the reprocessing of the observations prior to this release date was not completed yet. The chosen cases ...

#11. (line 283)

The phrase “large snowfall amounts” is unclear. Do you mean snow water content or surface snowfall rate? Please clarify.

We mean snow water content (as in Fig. 2d). We will make this clearer:

RACMO simulates relatively high snow water content over northwest Greenland and Baffin Bay (Fig. 2d).

#12. (Figure 2)

If water content values below 10^{-7} kg m^{-3} were excluded, it would be clearer to remove the lower-end extension of the colorbar. In addition, panels (a) and (b-e) use different units (g m^{-2} vs kg m^{-3}). Using a consistent unit system (e.g., g m^{-2} and g m^{-3}, or kg m^{-2} and kg m^{-3}) would improve readability.

Thanks for pointing this out. We will change the unit in Fig. 2a and Fig. 7a to kg m^{-2} and we will remove the lower-end extension of the colorbars in Fig. 2, 5, 6, 7, 10, and 11 in the relevant panels.

#13. (Figure 2 caption)

In the phrase “shown in (a), for (b) cloud ice, (c) cloud liquid water, (d) cloud snow and (e) cloud rain,” please consider avoiding the terms “cloud snow” and “cloud rain.” Snow and rain are precipitation categories and can occur both within and below clouds.

This is indeed confusing. We will change this part of the caption of Fig. 2 and 7 to:

(b-e) Water content [kg m^{-3}] as simulated by RACMO, for the co-located satellite overpass shown in (a), for (b) cloud ice, (c) cloud liquid water, (d) snow, and (e) rain.

#14. (lines 313-314)

Please check whether “cloud water content” is the most appropriate term here. Snow and rain water content are also included and may occur below cloud base. In addition, the conclusion that RACMO underestimates water content is largely based on radar reflectivity, which depends not only on water content but also on particle density and size distribution. In mixed-phase clouds with supercooled liquid water, high reflectivity could also be associated with rimed ice particles.

Thanks for this clarification. We will change the term “cloud water content” in line 314 and the captions of Fig. 2 and Fig. 7 to:

Cloud and precipitation water content

We will also change lines 312-314 to make the uncertainty around the interpretation of the radar reflectivity clearer:

Looking again at the Baffin Bay area, RACMO underestimates the radar reflectivity (Fig. 3f). The missing reflectivity at high altitudes indicates that RACMO underestimates the cloud top height. The lower strength of the reflectivity signal might indicate that RACMO underestimates the cloud water content of the clouds in this region. This can, however, not be concluded from the radar reflectivity alone, as the radar reflectivity also depends on the number concentrations, particle size distributions, cloud phase, and presence of rimed particles. Additionally, relying on empirical relationships to simulate radar reflectivity also introduces uncertainties in the strength of the reflectivities, which might also explain part of the underestimation. The CPR shows ...

#15. (Figure 3)

Were reflectivity values below -35 dBZ masked out? If so, it would be clearer to remove the lower-end extension of the colorbar.

Yes, these were masked out. We will remove the lower-end extension of the colorbar in Fig. 3e-f and 8e-f.

#16. (line 320)

Is the term "snowfall" the most appropriate here? Snow may refer both to in-cloud snow and to precipitation below cloud base.

We will rephrase line 20:

... clouds and snow particles often coexist.

#17. (line 322)

It would be helpful to clarify what is meant by "snowfall" in this context. The AC-TC product distinguishes multiple ice categories (e.g., snow, rimed snow, heavy snow, snow + SLW...). Even when a radar gate is classified as snow within clouds, cloud ice may still be present but undetected. Below cloud base, however, snow is more likely to represent true snowfall (without cloud ice).

For snowfall, we consider all AC-TC classes that include snow, so the includes the mixed categories as well. The classes considered for snowfall are: snow, rimed snow, rimed snow and supercooled liquid, and snow and supercooled liquid. The snowfall classes that contain supercooled liquid water are added to the mixed-phase category as well. For ice, all classes that include snow are also considered (besides the ice cloud categories), as the presence of non-precipitating ice crystals can not be excluded.

We will make this clearer in the methodology and add the following to line 257:

... and a precipitation class. We also simplify the AC-TC classification by only considering the categories ice, mixed-phase and liquid cloud, and rainfall and snowfall. When an AC-TC category belongs to both a cloud and a precipitation class (e.g., snow and supercooled liquid), we count this towards both the corresponding cloud and the corresponding precipitation class. As no clear distinction between snow and ice cloud particles can be made, all classes that include snow are also counted towards the ice cloud category.

#18. (lines 322-324)

The description here does not seem fully correct. Because radar cannot reliably separate clouds from precipitation regions, AC-TC uses the term "snow" without explicitly distinguishing cloud and precipitation. This does not mean that precipitation regions are classified as clouds. Please consider rephrasing.

We will rephrase this to make the distinction clearer:

Therefore, in the simplified classification based on the AC-TC in Fig. 4, locations with snowfall are always co-occurring with ice cloud, as the presence of ice cloud crystals can not be excluded. Therefore, the regions below a cloud where snow particles are precipitating, which are found in RACMO (hatched regions with white background in Fig. 4b), will not occur in the EarthCARE classification, which occurs over Baffin Bay.

#19. (lines 334-337)

It may be useful to first clarify how EarthCARE distinguishes ice and snow, as this definition may differ from that used in the model. For example, optically thin ice detected by ATLID but not by CPR may correspond to very small particles, which would reasonably be classified as ice from

an EarthCARE perspective. It is not clear whether there is evidence that such cases should instead be interpreted as misclassified snow.
Lines 335-339 provide two possible reasons why there are more locations with co-occurring snow and ice in RACMO than are found in the observations. One is too rapid snowfall generation in RACMO (lines 338-339), the second is the fact that the lidar might be able to observe a thin snow layer, but cannot distinguish between snow and ice, as there is no measurement of sedimentation velocity. There is not sufficient evidence to say the latter is the case, but there is a possibility that it explains part of the differences found. But, as this would only be the case for very small snow water content, the first explanation is more likely, and we will therefore rephrase as:
Therefore, when the snow water content is too low to be observed by the radar, it might not be correctly classified in the ATLID-CPR classification, as ATLID cannot distinguish small precipitating snowflakes from in-cloud ice crystals. On the other hand, as this would only be the case for very small snow water contents, a more likely explanation would be that RACMO could generate snow too quickly when ice is present, which could also lead to ice clouds dissipating too quickly.
#20. (lines 337-338)
Radar reflectivity scales with the sixth power of particle size. If small cloud ice particles grow into larger snow particles, reflectivity does not necessarily have to remain very low (it depends on density and particle sizes though).
There might indeed be competing effects that change the radar reflectivity when snowfall would be generated at too low IWC. We propose to leave out the part of this sentence that links this to the radar reflectivity, as this is indeed too uncertain.
… RACMO could generate snow too quickly when ice is present, which could also lead to ice clouds dissipating too quickly.
#21. (line 341)
Please clarify what altitude range is meant by "mid-level altitudes."
We will add the altitude range:
… mid-level altitudes (3-6 km). …
#22. (line 342)
The statement that "RACMO produces mixed-phase layers that are too shallow" is not entirely clear, as the two appear rather similar. Please clarify this point.
This is indeed not entirely clear, and only true around 76 degree latitude, which we will add:
.. although, around 76°N, RACMO produces mixed-phase layers that are too shallow. …
#23. (lines 342-343)
Could the absence of detected supercooled liquid water at this altitude be due to overlying supercooled liquid layers or thick ice clouds?
Thanks for pointing this out; that is definitely a possibility. We will add this to the text:
Over Baffin Bay, RACMO simulates a mixed-phase cloud around 2 km altitude that is not observed by the satellite, likely because the lidar signal is fully attenuated here (Fig. 3c).
#24. (Figure 4)
In panel (a), the scattered "all ice" classification between 82-83 latitude is real? The AC-TC product may provide quality flags that could be used for quality control. In addition, regions labeled as "all liquid" are within the CPR surface clutter regions. These regions may be better described as areas where CPR observations are not reliable, rather than true "all liquid" clouds. I would recommend excluding surface clutter regions from the analysis.
We have checked this region with scattered "ice", which is also present in the ATLID IWC data (Fig. 5a) and this is not classified as bad quality data. These are regions where ATLID detects very low IWC, which is not picked up by the CPR. Regarding the surface clutter, these areas are indeed not reliable. Therefore, we will mask all surface clutter areas from Fig. 4, 5, 6, 9, 10, and 11 (shown below). For the histograms in Fig. 5c and Fig. 10c, we will also mask out the area that

suffers from surface clutter in the RACMO profiles, for fair comparison. For the classification, not all of the “all liquid” areas were found within the surface clutter region (Fig. 4 and Fig. 9).

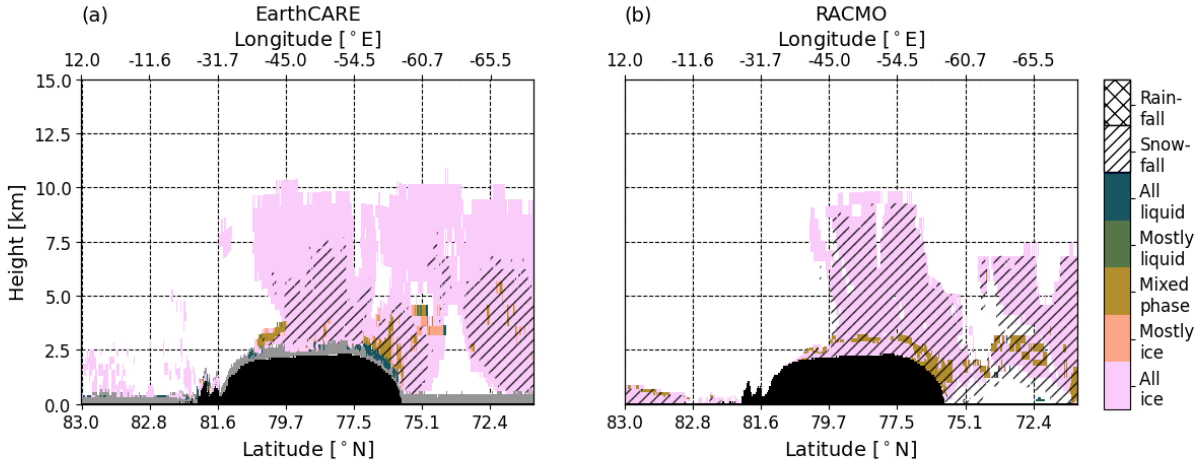


Figure 4. Cloud and precipitation classification for (a) EarthCARE (AC-TC, baseline BA) and (b) RACMO for March 12th, 2025, 19:48 UTC. Clouds are classified as ice (pink), liquid (blue), or mixed-phase (brown) clouds. Because the AC-TC is downsampled to the RACMO grid, some grid cells may fall between these categories and can then be classified as mostly ice (orange) or mostly liquid (green). The hatched areas indicate areas with snowfall or rainfall. Note that the x-axis follows the time coordinates. Hence, the latitude and longitude coordinates do not vary monotonically. Black areas correspond to the topography. In (a), gridcells classified as surface clutter are masked in grey.

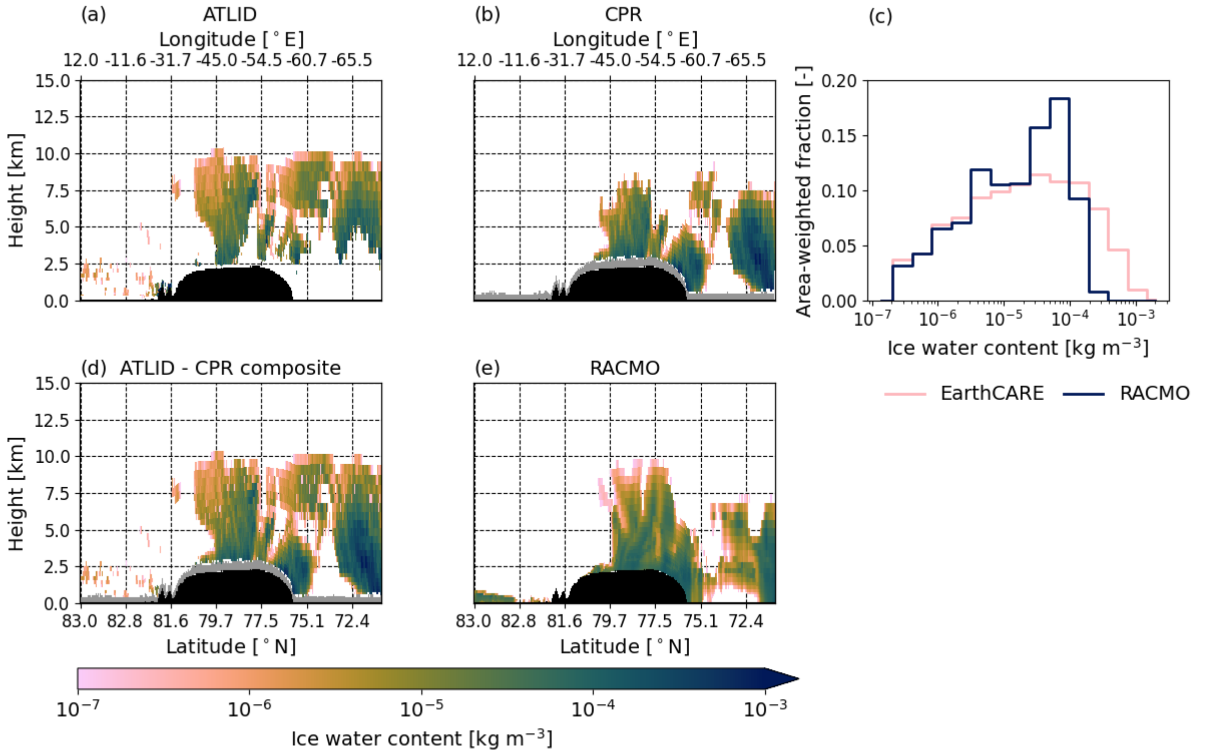


Figure 5. Ice water content [kg m−3] from (a) ATLID (ATL-ICE, baseline BA), (b) CPR (CPR-CLD, baseline BA), (d) ATLID-CPR composite and (e) RACMO (including snow water content) for March 12th, 2025, 19:48 UTC. (c) shows the gridcell area-weighted histogram of ice water content for EarthCARE (pink, ATLID-CPR composite) and RACMO (blue, including snow water content). When computing the histograms, areas that suffer from surface clutter are masked out in both the EarthCARE and RACMO profiles. Note that in (a,b,d,e) the x-axis follows the time coordinates. Hence, the latitude and longitude coordinates do not vary monotonically. In

(a,b,d,e), black areas correspond to the topography. In (b,d), gridcells classified as surface clutter are masked in grey.

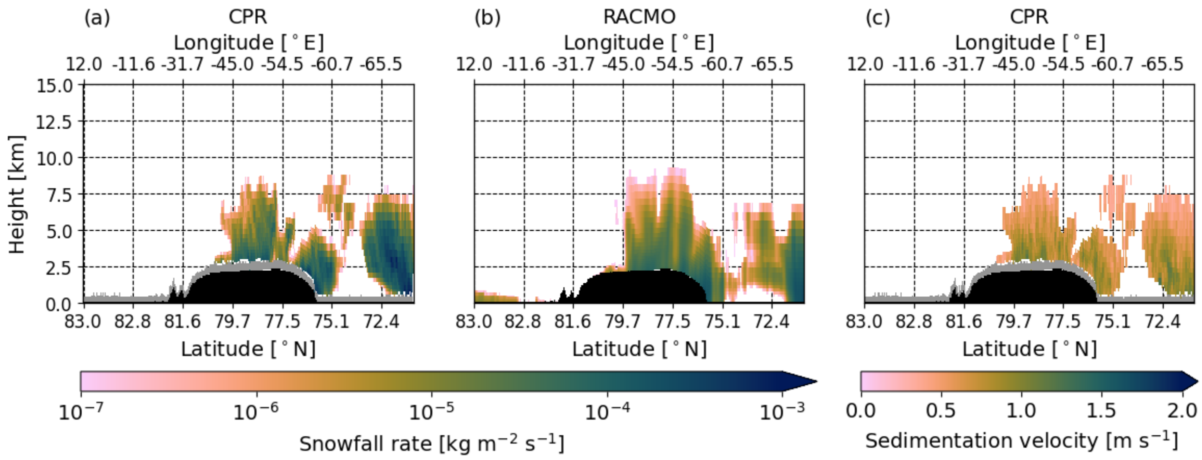


Figure 6. (a-b) Snowfall rate [kg m−2 s −1] from (a) CPR-CLD (baseline BA) and (b) RACMO for March 12th, 2025, 19:48 UTC. The RACMO snowfall rate is obtained by multiplying the snow water content by the sedimentation velocity. (c) Sedimentation velocity [m s−1] from CPR-CLD. Note that the x-axis follows the time coordinates. Hence, the latitude and longitude coordinates do not vary monotonically. Black areas correspond to the topography. In (a,c), gridcells classified as surface clutter are masked in grey.

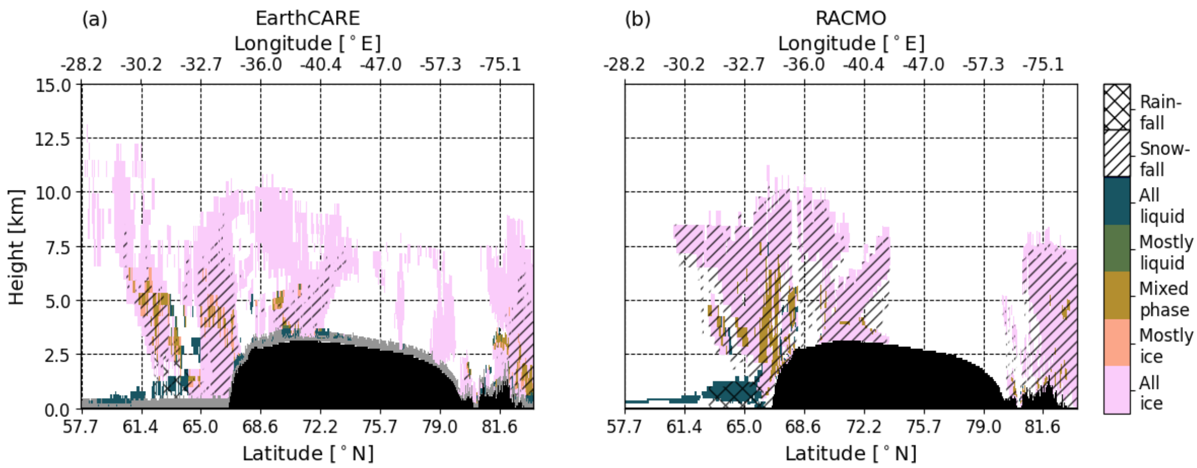


Figure 9. Cloud and precipitation classification for (a) EarthCARE (AC-TC, baseline BA) and (b) RACMO for May 13th, 2025, 03:10 UTC. Clouds are classified as ice (pink), liquid (blue), or mixed-phase (brown) clouds. Because the AC-TC is downsampled to the RACMO grid, some grid cells may fall between these categories and can then be classified as mostly ice (orange) or mostly liquid (green). The hatched areas indicate areas with snowfall or rainfall. Note that the x-axis follows the time coordinates. Hence, the latitude and longitude coordinates do not vary monotonically. Black areas correspond to the topography. In (a), gridcells classified as surface clutter are masked in grey.

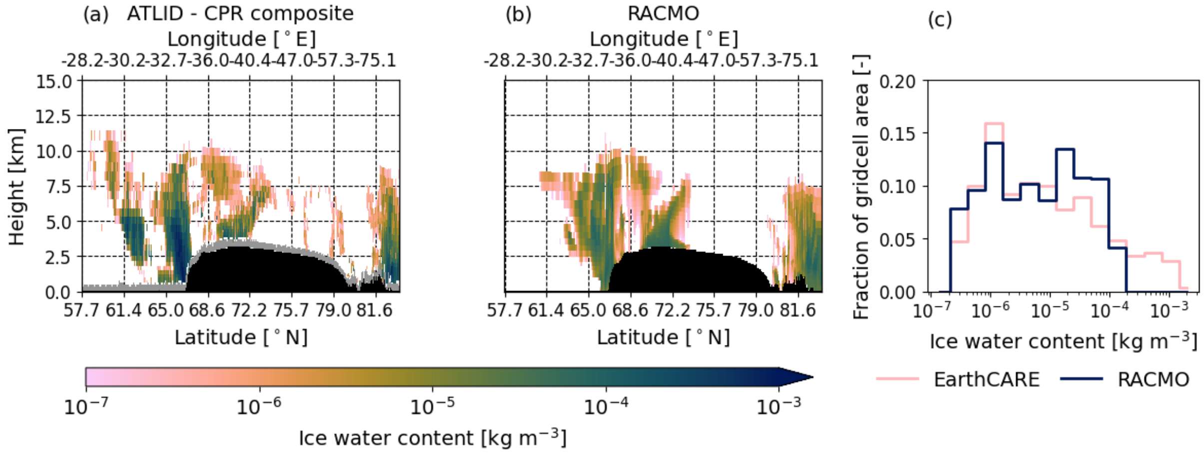


Figure 10. Ice water content [kg m−3] from (a) ATLID-CPR (from ATL-ICE and CPR-CLD, baseline BA) composite and (b) RACMO (including snow water content) for May 13th, 2025, 03:10 UTC. (c) shows the gridcell area-weighted histogram of ice water content for EarthCARE (pink, ATLID-CPR composite) and RACMO (blue, including snow water content). When computing the histograms, areas that suffer from surface clutter are masked out in both the EarthCARE and RACMO profiles. Note that in (a-b) the x-axis follows the time coordinates. Hence, the latitude and longitude coordinates do not vary monotonically. In (a-b), black areas correspond to the topography. In (a), gridcells classified as surface clutter are masked in grey.

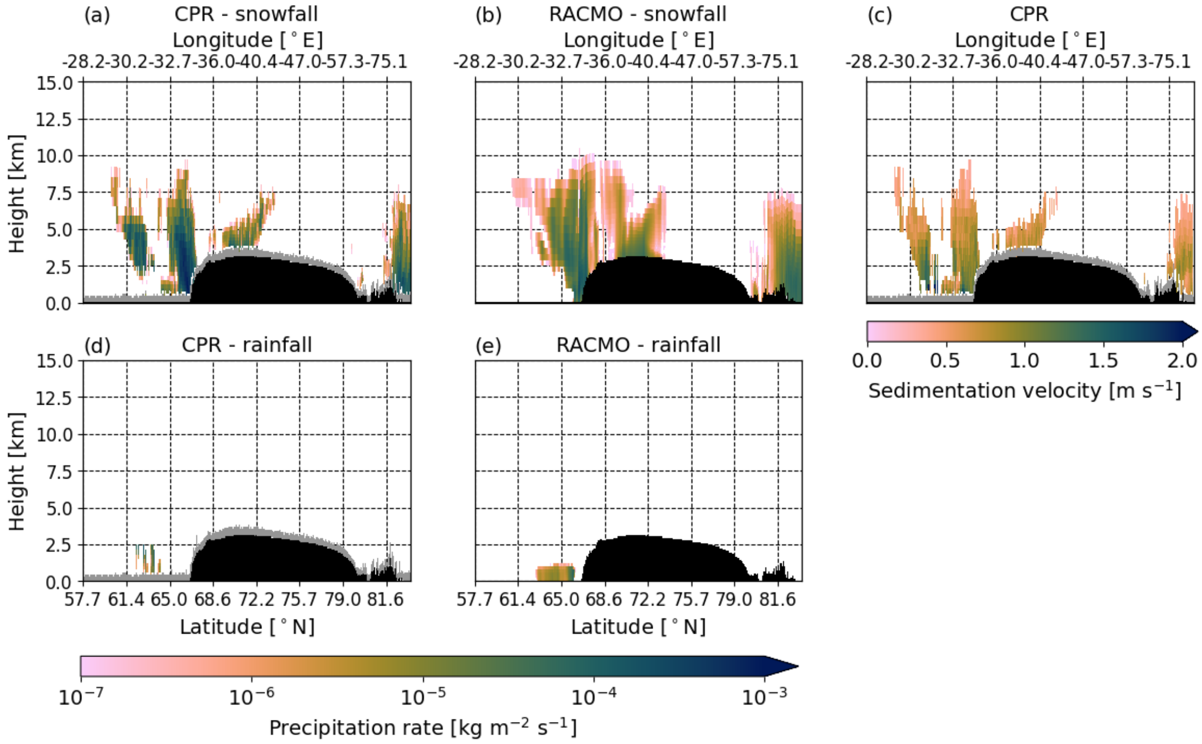


Figure 11. (a-b) Snowfall rate [kg m−2 s −1] from (a) CPR-CLD (baseline BA) and (b) RACMO for May 13th, 2025, 03:10 UTC. (c) Sedimentation velocity [m s−1] from CPR-CLD. (d-e) Rainfall rate [kg m−2 s −1] from (d) CPR-CLD and (e) RACMO. The snowfall and rainfall rates are obtained by multiplying the snow and rain water content by the sedimentation velocity. Note that the x-axis follows the time coordinates. Hence, the latitude and longitude coordinates do not vary monotonically. Black areas correspond to the topography. In (a,c,d), gridcells classified as surface clutter are masked in grey.

Regarding surface clutter, we will add to line 347:
... the lower and thicker clouds and precipitation. The few hundred meters closest to the surface are affected by surface clutter and are, therefore, not reliable and masked out from the observations. For a direct ...
#25. (lines 356-357)
When stating that RACMO overestimates snowfall rate along the western margin, please clarify the altitude level, as this is difficult to assess visually. In addition, snowfall estimates from EarthCARE are not reliable within the surface clutter region.
This is for altitudes below 2.5 km. We will add it to the text:
Contrastingly, the snowfall rate directly at the western margin (below 2.5 km altitude) is overestimated in RACMO.
As stated in the comment above, we will remove all surface clutter areas.
#26. (line 361)
Please specify in the main text which product was used to obtain the CPR sedimentation velocity.
This is CPR-CLD.
... observed by the radar (CPR-CLD; Fig. 6c).
#27. (Figure 6)
Please check whether sedimentation velocity is appropriately visualized. I can see sedimentation velocity in regions where CPR reflectivity is not shown in Fig. 3e, even though CPR Doppler measurements are only reliable for reflectivities above -15 dBZ. In addition, both snowfall and sedimentation velocity are not reliable within the surface clutter layer.
Thank you for indicating this. We masked out the invalid sedimentation velocities and the surface clutter regions, see the figures in the response to comment #24.

Technical corrections:
#1. (line 125)
In the phrase “the McICA method (McRad;)”, a reference is missing.
Thank you, the missing reference is Morcrette et al., 2008. We will add this reference in the revised manuscript.
#2. (line 226-227)
In “snow water content,” the word “snow” seems to be a typo. Please check.
Thank you for catching this. It should be rain instead of snow; we will correct this.
#3. (lines 233-234)
Units for LWP and WVP are missing.
Thank you for pointing this out. We will add the units [kg m^{-2}]. The unit mentioned here for the surface pressure is also incorrect. We will also change that to be hPa.
#4. (line 286)
In the phrase “in Fig. 2a-d,” panel (e) seems to be missing. Please check.
This should be Fig. 2b-e (which show the vertical profiles). We will update this.

References

Barker, Howard W., Jason N. S. Cole, Najda Villefranque, Zhipeng Qu, Almudena Velázquez Blázquez, Carlos Domenech, Shannon L. Mason, and Robin J. Hogan. 2025. 'Radiative Closure Assessment of Retrieved Cloud and Aerosol Properties for the EarthCARE Mission: The ACMB-DF Product'. *Atmospheric Measurement Techniques* 18(13):3095-3107. doi:10.5194/amt-18-3095-2025.

Kollias, P., Puidgomènech Treserras, B., Battaglia, A., Borque, P. C., and Tatarevic, A.: Processing reflectivity and Doppler velocity from EarthCARE's cloud-profiling radar: the C-FMR, C-CD and C-APC products, Atmos. Meas. Tech., 16, 1901–1914, https://doi.org/10.5194/amt-16-1901-2023, 2023.

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

Mech, M., Maahn, M., Kneifel, S., Ori, D., Orlandi, E., Kollias, P., Schemann, V., and Crewell, S.: PAMTRA 1.0: the Passive and Active Microwave radiative TRAnsfer tool for simulating radiometer and radar measurements of the cloudy atmosphere, Geosci. Model Dev., 13, 4229–4251, https://doi.org/10.5194/gmd-13-4229-2020, 2020.

Moradi, I., Kalluri, S., and Zhu, Y.: Forward modeling of spaceborne radar observations, Atmos. Meas. Tech., 19, 549–563, https://doi.org/10.5194/amt-19-549-2026, 2026.

Morcrette, J. J., H. W. Barker, J. N. S. Cole, M. J. Iacono, and R. Pincus. 2008. 'Impact of a New Radiation Package, McRad, in the ECMWF Integrated Forecasting System'. *Monthly Weather Review* 136(12):4773-98. doi:10.1175/2008MWR2363.1.

Mroz, K., Treserras, B. P., Battaglia, A., Kollias, P., Tatarevic, A., and Tridon, F.: Cloud and precipitation microphysical retrievals from the EarthCARE Cloud Profiling Radar: the C-CLD product, Atmos. Meas. Tech., 16, 2865–2888, https://doi.org/10.5194/amt-16-2865-2023, 2023.

Pfitzenmaier, L., Kollias, P., Risse, N., Schirmacher, I., Puigdomenech Treserras, B., and Lamer, K.: Orbital-Radar v1.0.0: a tool to transform suborbital radar observations to synthetic EarthCARE cloud radar data, Geosci. Model Dev., 18, 101–115, https://doi.org/10.5194/gmd-18-101-2025, 2025.

Stephens, Graeme L., Deborah G. Vane, Simone Tanelli, Eastwood Im, Stephen Durden, Mark Rokey, Don Reinke, Philip Partain, Gerald G. Mace, Richard Austin, Tristan L'Ecuyer, John Haynes, Matthew Lebsock, Kentaroh Suzuki, Duane Waliser, Dong Wu, Jen Kay, Andrew Gettelman, Zhien Wang, and Rojer Marchand. 2008. 'CloudSat Mission: Performance and Early Science after the First Year of Operation'. *Journal of Geophysical Research: Atmospheres* 113(D8). doi:10.1029/2008JD009982.

Wehr, Tobias, Takuji Kubota, Georgios Tzeremes, Kotska Wallace, Hirotaka Nakatsuka, Yuichi Ohno, Rob Koopman, Stephanie Rusli, Maki Kikuchi, Michael Eisinger, Toshiyuki Tanaka, Masatoshi Taga, Patrick Deghaye, Eichi Tomita, and Dirk Bernaerts. 2023. 'The EarthCARE Mission – Science and System Overview'. *Atmospheric Measurement Techniques* 16(15). doi:10.5194/amt-16-3581-2023.