A unified synergistic retrieval of clouds, aerosols and precipitation from EarthCARE: the ACM-CAP product

Shannon L. Mason[1,2], Robin J. Hogan[1,3], Alessio Bozzo[1,4*], and Nicola L. Pounder[2,5*]

[1]European Centre for Medium-range Weather Forecasts (ECMWF), Reading, UK
[2]National Centre for Earth Observation (NCEO), University of Reading, Reading, UK
[3]Department of Meteorology, University of Reading, Reading, UK
[4]European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT), Darmstadt, Germany
[5]Assimila Ltd., Reading, UK
[*]*Denotes current affiliations*

Correspondence: Shannon Mason (shannon.mason@ecmwf.int)

Abstract. The ATLID-CPR-MSI retrieval of Clouds, Aerosols and Precipitation (ACM-CAP) product provides a synergistic "best-estimate" retrieval of the quantities and properties of all aerosols and hydrometeors detected by EarthCARE. While synergistic retrieval algorithms are now mature in many contexts, ACM-CAP is unique in providing a single unified retrieval product for all classes of hydrometeor—ice cloud and snow, drizzle and rain, and liquid clouds—and aerosol species, informed

5   by the synergistic target classification (AC-TC). The simultaneous retrieval of the entire atmosphere with a single optimal estimation retrieval, called Cloud, Aerosol and Precipitation from mulTiple Instruments using a VAriational TEchnique (CAP-TIVATE), allows for a robust accounting of observational and retrieval errors and the contributions of passive and integrated measurements in the context of layered and complex regimes, and for enforcing physical relationships between components (e.g. the conservation of precipitating mass flux through the melting layer).

10   We have demonstrated and evaluated the ACM-CAP product as applied to the three EarthCARE test scenes simulated from numerical weather model forecasts, using both case studies and statistical evaluation against the simulated measurements and the "true" quantities from the numerical model. We show that the retrievals are both strongly constrained by the observations from the active and passive instruments, and overall closely resemble the underlying model fields in terms of bulk quantities (e.g. cloud water content, precipitation mass flux, and aerosol extinction) and microphysical properties (e.g. cloud effective

15 radius, median volume diameter, and aerosols lidar ratio). The retrieval performs best where the active instruments have strong and unambiguous signal: in ice clouds and snow, which is observed by both ATLID and CPR, and in light to moderate rain, where CPR signal is strong. In precipitation, CPR's Doppler capability permits enhanced retrievals of snow particle density and raindrop size. In complex and layered scenes where ATLID is obscured, we have shown that making a simple assumption about the presence and vertical distribution of liquid cloud in rain and mixed-phase clouds allows improved assimilation of MSI solar

20 radiances. In combination with a constraint on CPR path-integrated attenuation from the ocean surface, this leads to improved retrievals of both liquid cloud and rain in mid-latitude stratiform precipitation. In the heaviest convective precipitation (i.e. greater than around 10 mm h^{-1}), both active instruments are strongly attenuated and dominated by multiple scattering; in these situations ACM-CAP provides a seamless retrieval of cloud and precipitation, but one which is subject to a high degree of

uncertainty. ACM-CAP's aerosol retrieval, constrained by ATLID and MSI solar radiances, is performed in hydrometeor-free parts of the atmosphere. The lidar backscatter is subject to high noise, while the solar radiances are expected to be dominated by uncertainties in surface properties especially over land. While the aerosol optical depth is well-constrained in the test scenes, there is a high degree of noise at the 1km resolution of the ACM-CAP product.

The use of numerical forecast models to simulate test scenes for testing and evaluation puts EarthCARE L2 processors at an unprecedented degree of readiness ahead of launch. While exposure to further simulated test scenes, campaign data, and ultimately in-flight EarthCARE measurements will motivate ongoing improvements to the representation of cloud and precipitation, the instrument forward-models, and their uncertainties, the present evaluation demonstrates that ACM-CAP will provide a novel unified and synergistic retrieval of clouds, aerosols and precipitation of high quality, including a robust accounting of the contributions of observations, and of measurement and retrieval errors.

1 Introduction

The scientific goals of the EarthCARE mission are to measure the global distribution of clouds, aerosols, and precipitation, to estimate their quantities and microphysical properties, and to quantify their radiative effects (Wehr et al., 2022). Within the ESA EarthCARE production model (Eisinger et al., 2022), the ACM-CAP product provides the "best estimate" retrieval of clouds, aerosols, and precipitation from the synergy of the atmospheric lidar (ATLID), cloud profiling radar (CPR), and multispectral imager (MSI). The retrieval framework underlying the ACM-CAP is the Cloud, Aerosol and Precipitation from mulTiple Instruments using a VAriational TEchnique (CAPTIVATE; Mason et al., 2017, 2018) algorithm, which is configurable for any combination of vertically-pointing radars, lidars, and radiometers. ACM-CAP exploits the complementary properties of EarthCARE's Doppler-capable CPR, high-spectral resolution ATLID, and solar and thermal infrared MSI channels to simultaneously retrieve all classes of hydrometeors and aerosols in each profile, and takes account of measurement errors and physical assumptions to report the uncertainties associated with all retrieved quantities for interpretation by users and downstream products. As is more fully described in Eisinger et al. (2022), ACM-CAP forms the basis for subsequent EarthCARE products quantifying cloud-aerosol-precipitation interactions with radiation: radiative transfer modelling for estimating broadband fluxes and heating rates (ACM-RT; Barker et al., 2022a), and the top-of-atmosphere radiative closure assessment (ACMB-DF; Barker et al., 2022b) when compared against EarthCARE's broadband radiometer (BBR).

Owing to the long-term success of the "Afternoon Train" (A-Train) of active and passive spaceborne remote sensors, algorithms exploiting the synergy of radars, lidars and radiometers to retrieve the properties of ice clouds and snow, rain, or liquid clouds are now mature. The active sensors in the A-Train facilitated an unprecedented survey of the atmosphere (Stephens et al., 2018), with the 94-GHz cloud profiling radar aboard *CloudSat* (Stephens et al., 2002) detecting ice clouds and snow, drizzle and light rain, and the 532-nm Cloud–Aerosol Lidar and Orthogonal Polarization (CALIOP) aboard Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO; Winker et al., 2003) sensitive to optically thin ice clouds, liquid clouds and aerosols. They were complemented by two radiometers aboard *Aqua*, Moderate Resolution Imaging Spectroradiometer (MODIS; Salomonson et al., 2002) providing solar and infrared radiances from clouds and aerosols across a wide swath, and

Clouds and the Earth's Radiant Energy System (CERES; Wielicki et al., 1998) broadband radiometer measuring radiative fluxes at top-of-atmosphere. One advantage of synergistic retrievals is in overcoming the limitation that single-instrument retrievals can be especially subject to uncertainties in complex or layered scenes. For example, MODIS cloud retrievals are affected by

60   the presence of drizzle (e.g. Zhang et al., 2011; Painemal and Zuidema, 2011) and in mixed-phase clouds (e.g. Khanal and Wang, 2018), and are subject to large biases in layered cloud scenes (e.g. Chang and Li, 2005; Naud et al., 2006). Similarly, *CloudSat* rain retrievals are subject to uncertainties due to liquid clouds, which contribute to radar attenuation (Leinonen et al., 2016; Matrosov, 2007; Matrosov et al., 2008), and *CloudSat* ice and snow retrievals are often blind to the presence of super-cooled liquid cloud (Battaglia and Delanoë, 2013; Battaglia and Panegrossi, 2020). In synergistic retrievals, complementary

65   measurements can be used to constrain multiple classes of hydrometeor simultaneously: for example, *CloudSat* and MODIS solar radiances are used to retrieve rain (Lebsock et al., 2011) and cloud water content (Austin et al., 2009; Leinonen et al., 2016). Synergistic retrievals can also be used to constrain additional properties within a class of hydrometeors: in ice clouds and snow, the complementary constraints of the radar reflectivity factor and lidar backscatter provides sufficient information to retrieve two parameters of the particle size distribution. DARDAR-CLOUD (Delanoë and Hogan, 2008, 2010) uses *CloudSat*,

70   CALIPSO and MODIS thermal infrared radiances to retrieve the profile of ice cloud and snow, with infrared radiances providing an integrated constraint on ice microphysical properties near cloud-top. Building upon the heritage of A-Train retrievals, and specifically on the optimal estimation framework used by DARDAR-MASK, ACM-CAP will take advantage of Earth-CARE's onboard synergy to assimilate all available ATLID, CPR and MSI measurements, and to retrieve all combinations of clouds, aerosols and precipitation simultaneously.

75   While the A-Train has yielded many single-instrument and synergistic retrievals, each product has been concerned with a subset of the full range of hydrometeors or aerosols in the atmosphere: therefore several data products must be combined in order to reconstruct the full distribution of clouds, aerosols and precipitation in the atmosphere and estimate their combined effects on the global radiation budget. The prominent effort to collate the A-Train retrievals and radiative transfer products based on composites of retrievals (Henderson et al., 2013) in the context of radiative flux measurements from CERES is the CALIPSO

80   *CloudSat* CERES and MODIS (CCCM; Kato et al., 2010, 2011) product. CCCM has been widely used to link profiles of clouds and aerosols to atmospheric heating rates and cloud radiative effects (e.g. Hill et al., 2018; Ham et al., 2017); however, a challenge when combining retrievals is that the different products are not necessarily based on inter-consistent physical assumptions, nor do they account for consistent contributions from each measurement. As a consequence the uncertainties in the retrieved quantities, and hence the derived radiative fields, are difficult to quantify (Kato et al., 2011). The ACM-CAP

85   product is novel in that all classes of hydrometeor and aerosols are retrieved simultaneously. This maximises the exploitation of EarthCARE instrument synergy, and allows the application of physical relationships between different parts of the retrieval. For example, retrieving snow and rain simultaneously means that a physical consistency condition can be applied ensuring that precipitation mass flux is conserved across the melting layer, as has been used in radar precipitation retrievals (Haynes et al., 2009; Mason et al., 2017; Mróz et al., 2021). In complex and layered scenes some integrated or passive measurements cannot

90   be adequately interpreted by species-specific retrievals. A unified approach ensures that the contributions of such constraints— for example, radar attenuation due to rain and liquid cloud, or solar radiance measurements with contributions from multiple

cloud or aerosol layers—are applied consistently. This allows for high-quality retrievals in the profiles where species-specific algorithms are most likely to report biased retrievals, large uncertainties, or to skip profiles entirely.

Using a single retrieval framework to simultaneously represent all hydrometeors and aerosols and assimilate all available measurements has the advantage of facilitating a detailed and consistent accounting of all measurement errors, uncertainties related to physical assumptions, and uncertainty estimates for all retrieved quantities. Retrieval uncertainties can be easily interpreted by users of the product, and included in downstream radiative transfer (ACM-RT) products.

In addition to its A-Train like measurements, EarthCARE's active instruments will have novel capabilities that will enhance the potential for cloud, aerosol and precipitation retrievals with ACM-CAP. The CAPTIVATE algorithm has already been used to demonstrate Doppler radars to retrieve information about the structure and density of snowflakes (Mason et al., 2018, 2019), and about the rain drop size distribution (Mason et al., 2017). The high-spectral resolution ATLID allows for the retrieval of ice extinction-to-backscatter ratio, and combined aerosol depolarization ratio and extinction-to-backscatter are used for advanced aerosol typing with Hybrid End-To-End Aerosol Classification (HETEAC; Wandinger et al., 2022).

In this study we describe the ACM-CAP processor and evaluate its performance over three synthetic EarthCARE scenes produced from a numerical weather model. Section 2 provides an outline of the ACM-CAP processor within in the ESA EarthCARE production model (Eisinger et al., 2022), and a detailed description of the retrieval framework, its representation of ice cloud and snow, liquid cloud, rain and aerosols, and its instrument forward models. In Section 4 we showcase and evaluate ACM-CAP using case studies in selected cloud, precipitation and aerosol regimes, and present a statistical evaluation of key retrieved quantities across the three test scenes. Finally in Section 5 we summarize the outlook for ACM-CAP and EarthCARE science.

2 ACM-CAP product

The CAPTIVATE retrieval scheme employs a classical variational (or "optimal estimation") approach (Rodgers, 2000), but is unique in that almost all aspects of the retrieval are configurable at run-time, including the observations to assimilate, the representation of the atmospheric constituents to retrieve (ice clouds and snow, liquid clouds, rain and aerosols), the state variables used to describe each constituent, and the additional constraints to apply. This means that the same algorithm can be applied to ground-based (Mason et al., 2018, 2019), airborne (Mason et al., 2017) and spaceborne platforms (Mason et al. A-Train paper). In Section 2.1 we provide an overview of how CAPTIVATE is configured for its application as the ACM-CAP processor, then we describe the representation of atmospheric constituents (Section 2.2) and the forward models for the EarthCARE instruments (Section 2.3).

2.1 Algorithm overview

Each EarthCARE orbit is divided into eight granules of length ~ 5000 km; the ACM-CAP processor runs one granule at a time, reading in six L1 and L2 data products and outputting one ACM-CAP data product for each. The ACM-CAP processor and its inputs and outputs is illustrated in Figure 1.
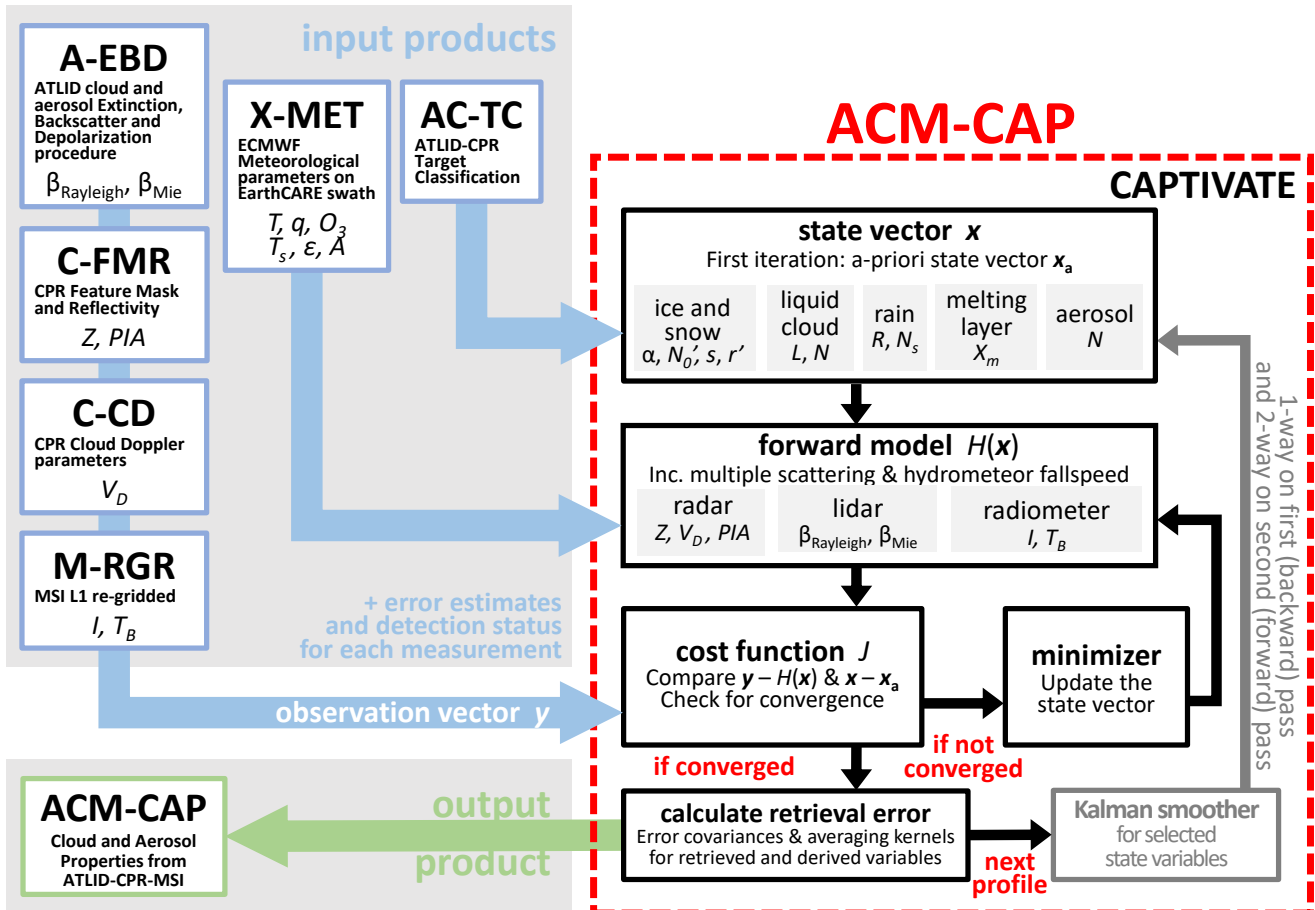
Figure 1. Flow chart showing the ACM-CAP processor and its input and output data products. The ACM-CAP processor uses the CAPTI-VATE retrieval framework, configured for EarthCARE's ATLID, CPR and MSI instruments.

Each profile in the granule is processed in turn. The synergistic target classification (AC-TC) product is used to define which
125 constituents will be retrieved in each grid volume. Each retrieved constituent is described by a number of state variables, the selection of which is described in Section 2.2. We write the vector of state variables describing a profile of constituent j as $\mathbf{x}_j$, and if we are retrieving the properties of n different constituents then the vectors are concatenated to obtain the full state vector $\mathbf{x}$.

The forward model $H(\mathbf{x})$ is used to simulate the observations made by each instrument based on the state. In Section 2.3
130 we describe how the state variables are used to forward model the measurements of active and passive instruments at a range of wavelengths. The forward model requires additional information about the atmosphere provided by the X-MET product. Atmospheric profiles of temperature, humidity and trace gas concentrations and surface properties are extracted from the European Centre for Medium-range Weather Forecasts (ECMWF) analysis along each granule. These data are needed to

estimate atmospheric transmission and scattering effects affecting active instruments and background radiances for passive

135 measurements. Each instrument makes a certain number of usable measurements in a profile, and we write the vector of usable measurements by instrument i as $\mathbf{y}_i$. If we are assimilating measurements by m different instruments the vectors are concatenated to obtain the full observation vector $\mathbf{y}$. These measurements are obtained from four input data files:

- A-EBD contains the post-processed ATLID backscatter measurements on the joint standard grid (JSG).

- C-FMR contains the CPR radar reflectivity and path-integrated attenuation (PIA) measurements

140 - C-CD contains the Doppler velocity measurements.

- M-RGR contains the re-gridded MSI reflectances and radiances.

Both CPR and MSI input data are provided on their own instrument grids, and are interpolated onto the JSG by ACM-CAP. To inform the interpretation and assimilation of each measurement, additional variables describing measurement uncertainties, and quality and detection statuses are also read from each data product.

145 The optimal estimate is the state vector that minimizes the cost function

$$J = \frac{1}{2}\delta\mathbf{y}^T\mathbf{R}^{-1}\delta\mathbf{y} + \frac{1}{2}\delta\mathbf{x}^T\mathbf{B}^{-1}\delta\mathbf{x} + J_c(\mathbf{x}), \tag{1}$$

where, in the first term on the right-hand side, $\delta\mathbf{y} = \mathbf{y} - H(\mathbf{x})$, $H(\mathbf{x})$ is the forward model, while $\mathbf{R} = \mathbf{O} + \mathbf{M}$ is the error covariance matrix of $\delta\mathbf{y}$, and consists of the sum of the error covariance matrices of the observations $\mathbf{O}$ and the forward model $\mathbf{M}$. In the second term on the right-hand side $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_a$, where $\mathbf{x}_a$ is the prior estimate of the state, and $\mathbf{B}$ is the error

150 covariance matrix of these priors. The final term, $J_c(\mathbf{x})$, expresses other physical constraints on the state variables.

Two methods have been implemented for iteratively modifying the state to minimize the cost function. The first is the L-BFGS method (Liu and Nocedal, 1989), which requires the gradient of the cost function with respect to the state, $\partial J/\partial\mathbf{x}$ (a vector) to be computed. This is the approach used by most variational data assimilation systems where the state vector is very large. The second is the Levenberg-Marquardt method (Marquardt, 1963), which requires both $\partial J/\partial\mathbf{x}$ and the second

155 derivative of the cost function with respect to the state $\partial^2 J/\partial\mathbf{x}^2$ (a matrix known as the Hessian) to be computed. This curvature information leads to many fewer iterations being required, but each iteration is more computationally costly since the Hessian requires the full Jacobian matrix $\partial\mathbf{y}/\partial\mathbf{x}$ to be computed. In practice the Hessian and Jacobian matrices are computed very efficiently by coding CAPTIVATE in C++ making use of the combined array, automatic differentiation and optimization library *Adept* (Hogan, 2014, 2017). Both the Levenberg-Marquardt and L-BFGS implementations in *Adept* support bounding

160 values to be applied to any of the state variables. We presently use the L-BFGS method.

The method described so far allows all state variables to be modified in an attempt to minimize J. While we include in CAPTIVATE all the variables needed to describe each constituent, there are not always sufficient measurements to constrain their retrieval; i.e. there may be too many degrees of freedom. In these situations it is possible to designate a *model variable*, which is included in the state vector but not modified during the minimization. This reduces the degrees of freedom, while

165 allowing uncertainty in the model variable to be included in the cost function, and propagated to the retrieved and derived quantities.

A variational approach provides an elegant framework that takes rigorous account of uncertainties, but the fidelity of any retrieval is dependent on the appropriate choice of state variables and additional constraints, and the accuracy of the forward models. In Section 2.4 we describe the automatic computation of uncertainties and error covariance matrices for retrieved and 170 derived variables, error correlation scales, and additional metrics derived from the averaging kernel.

We have described the retrieval as carried out on each profile in turn; however, the retrieval of some state variables may be improved by representing a degree of coherence over larger spatial scales. For these state variables a Kalman smoother (Rodgers, 2000) can be applied, by which each retrieved profile is constrained on the first pass by the values retrieved in the previous profile, and on a second pass by the values retrieved in both directions. Kalman smoothing is especially beneficial for 175 retrieving state variables that are weakly constrained by noisy measurements, such as of aerosols from lidar backscatter.

2.2 Representation of atmospheric constituents

In this section we describe and justify how each of the atmospheric constituents is represented in ACM-CAP, although we stress that these representations are completely configurable and may be modified as needed. These are several overarching principles we maintain in selecting state variables:

180 1. Usually two variables are used to describe the size distribution, providing the degrees of freedom to allow total number density and mean size to vary. The shapes of the size distributions are configurable, but held fixed. The uncertainty associated with a fixed size distribution shape is secondary compared to those of number concentration and mean size (Delanoë et al., 2005), but does become relevant in, for example, triple-frequency radar retrievals (Mason et al., 2019).

2. Typically we retrieve one extensive variable E (e.g. water content or extinction coefficient) and one variable N that has 185 the properties of a number concentration. This means that only 1-D look-up tables are required, since all other extensive variables X can be written as $X/N = f(E/N)$, while all intensive variables can be written as $I = f(E/N)$ (see Delanoë et al., 2005, for further discussion).

3. Convergence is more rapid if the relationship between observations and the main state variables they are sensitive to is close to linear. Since the relationships between many variables are close to a power law (implemented as look-up tables 190 rather than an actual power law), they can be represented as close to linear if both $\mathbf{x}$ and $\mathbf{y}$ contain the natural logarithm of meteorological and observational quantities. This also ensures that retrieved quantities cannot go negative (Delanoë and Hogan, 2008).

4. Certain useful a-priori and physical constraints can be applied only if a constituent is described by a certain variable. For example, the constraint that the gradient of water content of liquid clouds with height should not exceed the adiabatic 195 rate can only be applied if liquid water content is a state variable.

Table 1 lists the state variables retrieved for each atmospheric constituent, along with their a-priori values and errors as presently configured for ACM-CAP. If a state variable is well constrained by an active instrument then independent values will be retrieved in each volume. However, frequently the observations will lack the information content to retrieve certain state variables at such high vertical resolution, so to ensure the retrieval is not ill-posed and converges quickly, the profile may 200 optically be described by fewer state variables, such as the coefficients of a set of cubic-spline basis functions (Hogan, 2007).

Table 1. List of state variables used to describe each of the constituents, with corresponding a-priori values and their uncertainties. Note that the natural logarithm of these values is usually used in the state vector, and the uncertainties expressed are the errors in the natural logarithm of the a-priori values. The physical constraints include the vertical representation and horizontal Kalman smoothing.

State variable	A-priori	RMS uncertainties	Spatial representation
Ice cloud and snow			
Geometric extinction coefficient $\ln \alpha_v$ (m^{-1})	$-6.9 + 0.0315T\,[^\circ\mathrm{C}]$	10.0	Cubic spline basis functions with $z_0 = 300$m
Primed number concentration $\ln N_0'$ (m$^{-3.4}$)	$22.46 - 0.08924T\,[^\circ\mathrm{C}]$ (Delanoë and Hogan, 2008)	1.0	Cubic spline basis functions with $z_0 = 500$m
Lidar ratio $\ln S$ (sr^{-1})	0.027	0.3	Cubic spline basis functions with $z_0 = 500$m
Density index r'	0.0	1.0	Cubic spline basis functions with $z_0 = 500$m
Liquid cloud			
Liquid water content $\ln L$ (kg m^{-3})	1×10^{-4}	2.0	*When detected by ATLID:* Direct
		1.0	*Collocated with rain:* One value per vertically-contiguous layer with linear gradient reducing to zero at the bottom of the layer *Collocated with ice:* Layerwise constant
Total number concentration $\ln N_T$ (m^{-3})	*Marine*: 74×10^6 *Continental*: 288×10^6 (Miles et al., 2000)	1.0	*When detected by ATLID:* Layerwise constant *Elsewhere:* model variable
Rain			
Rain rate $\ln R$ (kg m^{-2} s^{-1})	2.778×10^{-5}	4.0	Cubic spline basis function with $z_0 = 300$m
Number concentration scaling $\ln N_s$	1.0	0.5	Cubic spine basis function with $z_0 = 1$km; temporal smoothing error 0.15
Melting layer			
Thickness scaling factor $\ln X_m$	1.0	1.0	Model variable
Aerosols (dust, sea salt, continental pollution, smoke, dusty smoke, dusty mixture)			
Total number concentration $\ln N_T$ (m^{-3})	1×10^6	2.0	Direct; temporal smoothing error 0.1
Median volumetric diameter $\ln D_0$ (m)	*Dust:* 0.866×10^{-6} *Sea salt:* 0.789×10^{-6} *Cont. poll'n:* 0.545×10^{-6} *Smoke:* 0.241×10^{-6} *Dusty smoke:* 0.783×10^{-6} *Dusty mix.:* 0.802×10^{-6}	1.0	Layerwise constant

2.2.1 Ice clouds and snow

We follow Delanoë and Hogan (2008) and treat ice clouds and snow as a continuum described by extinction coefficient in the geometric-optics approximation, α_v, and a primed number concentration variable N_0' which is defined in terms of the normalized number concentration parameter N_0^* (Delanoë et al., 2005). The variable N_0' has the advantage that a reasonable a-priori estimate of it can be made from temperature alone (Delanoë and Hogan, 2008). This enables a seamless retrieval between regions where both radar and lidar detect the cloud, and regions where only one detects it. As ATLID has HSRL capability the independent information on backscatter and extinction allows vertical variations of the lidar extinction-to-backscatter ratio (S) to be retrieved. This quantity is represented by a cubic spline due to noise in the lidar measurements.

Doppler velocity can provide information on the riming of snowflakes, since rimed particles are denser and therefore fall faster than unrimed particles of a similar size. The retrieval of a density factor r to resolve variations in snow particle density due to riming in mixed-phase cloud layers was described in Mason et al. (2018). This single parameter is used to vary the prefactors and exponents of the mass- and area-size relations of ice particles, as well as assumptions about microphysical structure informing microwave scattering approximations. Snow with a density factor of $r = 0$ corresponds to unrimed aggregates with the mass-size relation given by Brown and Francis (1995) and the area-size relation of Francis et al. (1998), while precipitating ice with a density factor of $r = 1$ would correspond to spheres of solid ice. Intermediate values of r represent a continuum of snow particles from partially-rimed aggregates to lump graupel. While there are limited observational and theoretical constraints on how to best represent rimed snowflakes and the transition to graupel, CAPTIVATE retrievals of rimed snow from Hyytiälä, Finland assimilating dual-frequency Doppler radar measurements compared favourably in terms of snow rate and bulk density with in situ snow measurements at the surface (Mason et al., 2018).

In order that the parameter representing riming can be included in the minimisation without the possibility of reaching non-physical values, the retrieved state variable is a transformed density factor $r' = f(r)$, which also represents unrimed aggregates at $r' = 0$, but is physically meaningful at all values (Section 2.2.3 of Mason et al., 2018). This capability has been developed and evaluated using ground-based and dual-frequency Doppler radars. While Mason et al. (2018) demonstrated some skill in using 94 GHz Doppler radars to retrieve rimed snow in stratiform cloud scenes, the capacity to perform this kind of retrieval from EarthCARE is sensitive to the quality of Doppler velocity measurements. Corrections for radar mispointing and non-uniform beam filling errors, along-track integration and more sophisticated local smoothing techniques have been implemented to reduce Doppler velocity measurement noise and decompose an estimate of "sedimentation velocity" from vertical air motion (Kollias et al., 2022). The choice of which Doppler velocity variable to use in ACM-CAP—and a better characterisation of their associated uncertainties—will be informed by calibration and validation activities after launch. The synthetic test scenes used in this study do not include stratiform rimed snow in which to evaluate the contribution of Doppler velocity measurements to snow retrievals in more detail.

2.2.2 Liquid cloud

Liquid clouds present a significant challenge for spaceborne radar and lidar retrievals: while the radar signal is dominated by drizzle drops, the lidar signal is rapidly attenuated at the top of the layer, making the physical depth of a cloud layer difficult to establish. Irbah et al. (2022) have showed that for EarthCARE, around 20% of the volume of liquid cloud in the test scenes is directly detected by the synergy of the active instruments, representing around 10% of the liquid water content. Even when not directly detected by active instruments, integrated constraints on the liquid water path (LWP)—but not on the vertical distribution of liquid—may be obtained from the radar PIA (Lebsock et al., 2011), and on cloud optical depth from solar radiances (Leinonen et al., 2016).

Liquid water content (LWC) L is used as the main state variable, allowing for assumptions about the vertical distribution of cloud water even in cloud layers that are not directly observed by the active instruments (i.e. non-precipitating clouds not detected by CPR, or whenever ATLID is extinguished aloft). In ACM-CAP, liquid cloud is assumed to be collocated with precipitation in two situations:

– in rimed snow and convective cores, where the presence of supercooled liquid is very likely and will have a greater contribution to radar attenuation than ice alone;

– collocated with rain, where liquid cloud is not directly detected by ATLID, but cloud is very likely to be present and will contribute to radar attenuation.

Irbah et al. (2022) showed that this interpretation of the synergistic target classification resulted in the correct classification of around 60% of the liquid cloud by volume, representing almost 75% of liquid water content, across the three test scenes. The importance of these assumptions, and the capacity to constrain a retrieval of liquid cloud not directly detected by the active instruments, will be explored using case studies in Section 3.

The second variable retrieved is the total droplet number concentration, since a-priori estimates are available over land and sea (e.g. Miles et al., 2000). When ATLID detects a liquid cloud layer this variable is retrieved, assuming a constant value for each contiguous cloud layer; otherwise, the a-priori value is used.

2.2.3 Rain and drizzle

This constituent represents both "cold rain" originating from melting ice, and "warm rain" or drizzle from the collision and coalescence of cloud droplets within liquid clouds. The main variable retrieved is the rain rate, R. Since rain has a high fall speed, we can apply the physical constraint that R does not vary rapidly with height, achieved by adding to J_c a "flatness" term proportional to $(\partial \ln R/\partial z)^2$ using the approach of Twomey (1977). The result is that in moderate rainfall the retrieval can infer rain rates from the gradient of radar reflectivity factor with height (as proposed by Matrosov, 2007), while also being able to use the radar PIA derived from the surface reference technique when available (L'Ecuyer and Stephens, 2002).

The retrieval of warm and cold rain from airborne Doppler radars using CAPTIVATE was demonstrated by Mason et al. (2017). In that study, the second state variable for rain was the normalized number concentration parameter N_w as defined by

Testud et al. (2001) (see also Illingworth and Blackman, 2002). Informed by mean Doppler velocity and PIA, retrieved values
265 of N_w in that study varied over several orders of magnitude from near the Marshall and Palmer (1948) value of 8×10^6 m^{-4}
in cold rain, to much higher values in warm rain and drizzle. In order that the a-priori rain DSD realistically represents both
heavy rain and drizzle, in this study we implement the transition between a high concentrations of predominantly small drops
in drizzle and light rain, and fewer and larger raindrops in heavy rain, as described by Abel and Boutle (2012). This relation
constitutes our a-priori drop size distribution, for which the number concentration scaling parameter $N_s = 1$; the number
270 concentration scaling parameter is the retrieved as constrained by mean Doppler velocity and PIA measurements.

2.2.4 Melting layer

Retrieving an accurate physical description of the melting layer is very challenging because we have no direct measurements
of its properties, and current models for the scattering and attenuation behaviour of melting ice particles are very uncertain.
Since the radar is the only instrument affected by the melting, and there is no enhanced reflectivity 'bright band' at 94 GHz,
275 we treat the melting layer as a thin layer of radar attenuation that is applied across the infinitesimal layer between the lowest
volume in the profile classified as ice and snow, and the highest volume classified as rain—provided that the two are adjacent.
By default, we follow Matrosov (2008) and assume that the two-way attenuation of the melting layer A is proportional to the
rain rate R at the first volume just below the melting layer, such that

$$A = kX_m R \tag{2}$$

280 where at 94 GHz, $k = 2.2$ dB km^{-1}(mmh^{-1})$^{-1}$. This estimate has been supported using ground-based radars (Li and Moisseev,
2019). The physical depth and hence the total attenuation across the melting layer also depends on the local temperature profile;
variations in the strength of the melting layer attenuation can be represented by a thickness scaling factor X_m. Independent
information on melting-layer attenuation can sometimes be extracted from the combination of the radar PIA over the ocean
and the rain rate inferred from the reflectivity gradient; we therefore include the natural logarithm of X_m as either a retrieved
285 state variable or a model variable that resolves the effect of this uncertainty on the other retrieved variables and their errors.

To ensure physical consistency between retrieved constituents within the profile, a constraint can be included in J_c such
that the rain rate in the volume at the bottom of the melting layer is close to the mass flux of snow entering the melting layer.
This mass flux continuity constraint has been widely used in radar retrievals (Haynes et al., 2009; Mason et al., 2017); further
constraints on the continuity of snow and rain microphysical parameters across the melting layer have been demonstrated in
290 multiple-frequency radar retrievals (Mróz et al., 2021) but could prove beneficial even in this application.

2.2.5 Aerosols

The ACM-CAP treatment of aerosols takes as given the aerosol typing and properties of the HETEAC model (Wandinger
et al., 2016, 2022), which is incorporated in the ATLID aerosol classification (A-TC) and subsequently the synergistic target
classification (AC-TC). Thus, estimates of wavelength-dependent refractive index, size distribution and mean particle size are

295 predefined based on the aerosol classification, and the primary state variable retrieved is the total number concentration, which acts to scale all extensive variables such as aerosol extinction/optical thickness.

A major difficulty with using observations at 1 km along-track resolution is that at this scale the measurements are very noisy, especially when the signal is weak. The traditional approach is to average along-track before performing the retrieval, but this is not satisfactory if clouds are to be retrieved simultaneously at high spatial resolution. The retrieval of aerosols from 300 noisy lidar signal was the primary motivation for the implementation of the Kalman smoother, in which along-track smoothing is achieved by performing a first (backward) pass through the data during which the retrieval of a profile is constrained by the values retrieved in the previous profile, followed by a second forward pass in which the retrieval of a profile is constrained by the values in both directions.

Table 2. The scattering, fall-speed and size distribution assumptions made for each of the constituents retrieved in ACM-CAP. Certain categories are not applicable (N/A) if an instrument is not able to detect a constituent (e.g. radar and aerosol).

Constituent	Solar / infrared scattering model	Radar scattering model	Terminal fallspeed model	Size distribution shape
Ice cloud and snow	Baran (2004)	Aggregates: SSRGA (Hogan et al., 2017) Graupel: soft spheroids (Mason et al., 2018)	Heymsfield and Westbrook (2010)	Field et al. (2005)
Liquid cloud	Mie theory	Mie theory	Beard (1976)	Log-normal with shape factor 0.38
Rain	Mie theory	Mie theory	Beard (1976)	Gamma with shape factor $\mu = 5$
Melting ice	N/A	Matrosov (2008)	N/A	N/A
Aerosol	Mie theory	N/A	N/A	Lognormal with shape factor 2

2.3 Instrument forward models

305 The forward model $H(\mathbf{x})$ is a function that outputs the predicted observations $\mathbf{y}^f$ corresponding to a particular estimate of the state vector $\mathbf{x}$. Figure 2 shows the flow of information from $\mathbf{x}$ to $\mathbf{y}^f$. After outline the pre-calculated hydrometeor scattering and surface properties, the following sections describe the individual steps of forward modelling the instrument measurements from the state.

A key part of the forward model is the use of the state variables to calculate the profile of scattering properties at the 310 wavelengths of each instrument being used in the retrieval. Before the retrieval is run, offline calculations are performed to compute the scattering properties of individual hydrometeors, specifically the extinction, scattering and backscatter cross-sections. In the case of solar radiometers, we also compute and store the scattering phase function. The scattering models used for each constituent are listed in the second and third columns of Table 2. In order to forward-model the radar Doppler velocity we need a model for the terminal fall speeds of hydrometeors detectable to the radar, given in the fourth column of Table 2.
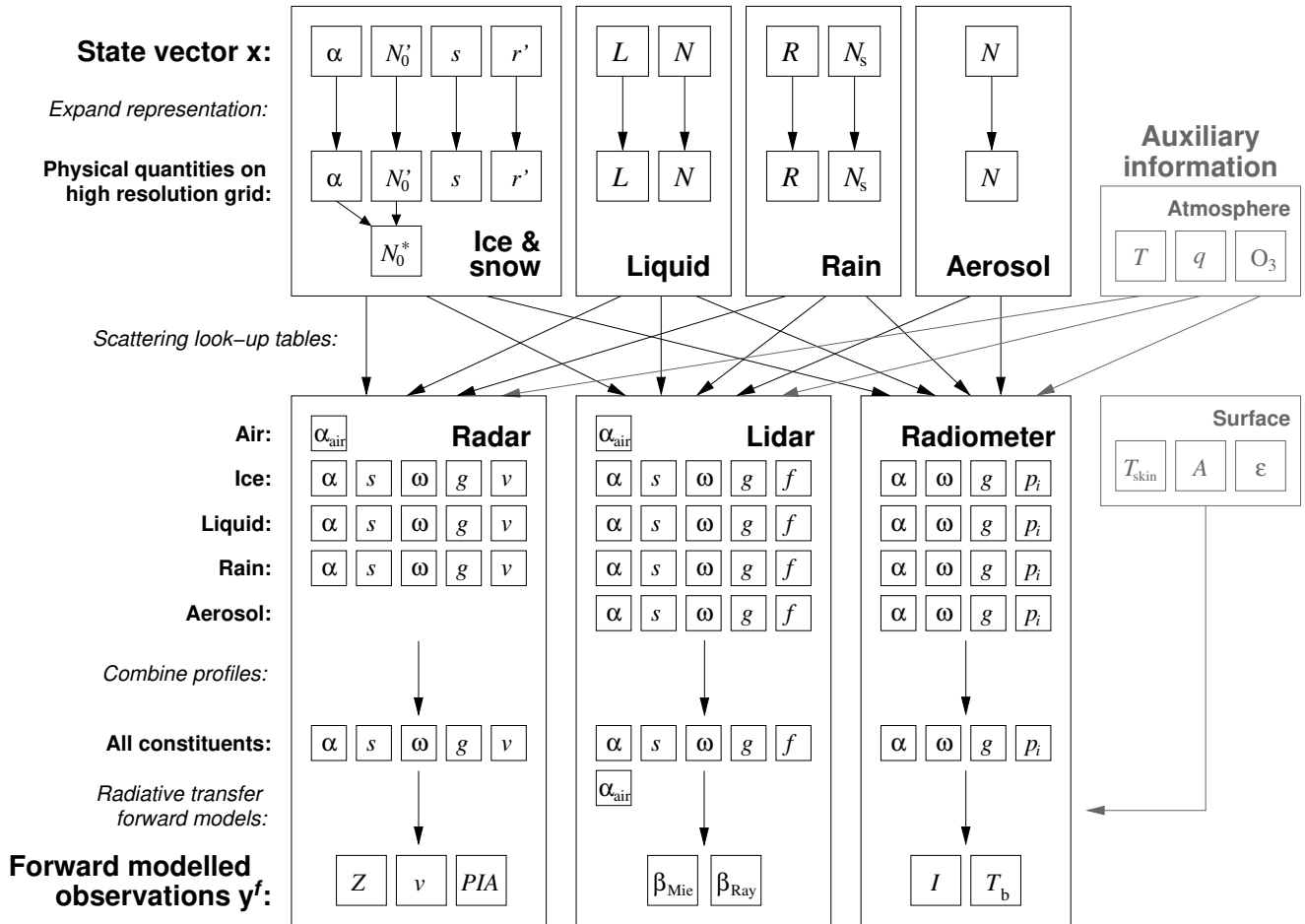
Figure 2. Flow chart depicting the flow of information through the forward model in ACM-CAP, translating the state variables $\mathbf{x}$ to the forward modelled observations $\mathbf{y}^f$. Auxiliary information about the state of the atmosphere and Earth's surface are shown in grey. The temperature and composition of the atmosphere is used to forward-model the profile of atmospheric scattering in each wavelength, while the surface skin temperature, albedo and emissivity is needed to simulate passive measurements.

Since liquid clouds, rain and aerosol can reasonably be treated as homogeneous spheres for the wavelengths under consideration, we may use Mie theory. The effect of representing large raindrops with a more realistic spheroidal geometry using the T-matrix scattering model had only minor effects on the retrieved rain in Mason et al. (2017), and is neglected here. The complex shapes of ice particles require more detailed careful consideration. For solar and infrared scattering from ice particles we use the Baran (2004) database, which takes account of surface roughness effects; however, the backscatter-to-extinction ratio S predicted by such a model is not regarded as accurate enough for the lidar forward model, so this variable is retrieved (see Section 2.2). For radar scattering by unrimed ice particles we use the Self-Similar Rayleigh-Gans (SSRGA) model of Hogan et al. (2017), appropriate for aggregates and other irregular particles. Following the evidence of Hogan et al. (2012) and

others, these particles are assumed to have aspect ratio of 0.6, to fall with a horizontal alignment, and to follow the mass-size and area-size relations of Brown and Francis (1995) and Francis et al. (1998), respectively. The mass and cross-sectional area

325 of snowflakes are both needed for the fall-speed model of Heymsfield and Westbrook (2010). Mason et al. (2018) described how when Doppler velocity is assimilated the density factor is used to transition from the unrimed aggregates above to heavily rimed graupel-like particles, represented as homogeneous spheroids for both radar scattering (Hogan et al., 2012) and mass- and area-size relations. All of these assumptions have uncertainties, which are represented approximately by adding a radar reflectivity forward model error to the appropriate diagonal elements of $\mathbf{M}$ (see eq. 1).

330 For the forward-modelling of passive solar and infrared radiances in clear-sky and optically thin profiles we require information about the surface, which is provided in the X-MET product generated from the ECMWF forecast model (Eisinger et al., 2022). For thermal infrared radiances the surface emissivities are taken as constant for wavelengths close to 10μm, with values of 0.96 over ocean, and 0.98 over the land (Fig. 3 of Feldman et al., 2014). The skin temperature and 2 m surface winds are from the same ECMWF forecast that provides the profile of atmospheric temperature and humidity.

335 In a synergistic retrieval the absence of a detection from one of the instruments can convey important information: if ATLID detects ice but CPR does not make a detection (classified "clear"), then there is a constraint that the forward-modelled CPR reflectivity should not exceed the instrument noise in that volume. This condition is applied for ice clouds to further constrain the properties of the lidar-only part of the profile.

In the following subsections we describe the steps shown in Fig. 2.

340 ### 2.3.1 Expanding vertical representation of variables

As indicated in the final column of Table 1, many state variables are not represented by separate values in every volume. Therefore, the first step in the forward model is to expand the representation of each state variable to compute its value in every volume. This process simply involves applying the operation $\mathbf{x}^{\text{full}} = \mathbf{W}\mathbf{x}$, where $\mathbf{x}$ contains the state variables for a particular quantity, $\mathbf{x}^{\text{full}}$ contains the corresponding values in each volume where that constituent is present, and $\mathbf{W}$ is a matrix describing

345 the representation. Hogan (2007) describe how $\mathbf{W}$ is formulated in the case of cubic splines.

After the state variables are computed in every volume, in the case of ice we then calculate the normalized number concentration parameter $N_0^* = N_0' \alpha_v^{0.6}$ (Delanoë and Hogan, 2008).

2.3.2 Scattering look-up tables

The next step is to compute the profile of scattering properties for each constituent (ice clouds and snow, liquid clouds, rain

350 and aerosol) at the wavelength of each instrument. All instruments require extinction coefficient α, single scattering albedo ω and asymmetry factor g. The active instruments also require extinction-to-backscatter ratio S. Furthermore, the Doppler radar requires reflectivity-weighted terminal fall speed v, and the lidar requires the fraction of the backscatter due to liquid droplets f in order to correctly describe small-angle multiple scattering (Hogan, 2008). Solar radiance modelling requires coefficients describing the full phase function p_i. These quantities are computed from the expanded state variables using look-up tables.

355 The scattering look-up tables are constructed when the algorithm is initialized.

2.3.3 Combining profiles

The profiles of scattering properties for each constituent, as well as the profile of scattering due to the atmosphere, are then combined into a single profile for the scattering at each wavelength. The extinction coefficients can be combined as a direct summation, while the other quantities must be combined as weighted sums. The extinction-to-backscatter ratio and single-scattering albedo are combined as weighted by the extinction coefficient; the combined asymmetry factors are weighted by the scattering coefficient (i.e. the extinction-to-backscatter ratio multiplied by the single-scattering albedo); and sum of the droplet fraction and mean Doppler velocity are weighted by the backscatter coefficient (i.e. the extinction coefficient multiplied by the extinction-to-backscatter ratio).

2.3.4 Radiative transfer

The final step in the forward model is to represent the propagation of radiation at all measured wavelengths through the combined profiles of scattering properties due to all hydrometeors and atmospheric gases. For inclusion in the forward model of the retrieval scheme, the radiative transfer model and its adjoint must be calculated accurately and efficiently. All of the radiative transfer methods are therefore written in C++ using the *Adept* automatic differentiation library (Hogan, 2017).

Multiple scattering is accurately treated within the forward model for all active measurements. Millimetre-wave radar is chiefly subject to multiple scattering in deep convective towers, while lidar multiple scattering can occur in all clouds. Wide-angle multiple scattering is modelled for both radar and lidar using the time-dependent two-stream method (TDTS; Hogan and Battaglia, 2008). Additional small-angle multiple scattering only affects lidar, and is represented using the photon variance–covariance method (PVC; Hogan, 2008).

The two-stream source-function (TSSF; Toon et al., 1989) approach is used for thermal infrared radiances, and has also been applied to model passive microwave radiances, although such measurements are not used in this study. For solar wavelengths the Forward-Lobe Two-Stream Radiance Model (FLOTSAM; Escribano et al., 2019) is used, wherein forward scattered solar radiation either remains close to the angle of incidence (the narrow forward-lobe), or is scattered into the wide forward-lobe at angles around $15°$; additional diffuse fluxes are calculated using the two-stream method, so FLOTSAM can be thought of as the equivalent of TSSF for solar wavelengths.

2.4 Calculation of retrieval errors

The state vector that minimises the cost function is called the "solution" of the optimal estimation retrieval. Once the cost function is minimized the errors in the retrieval can be estimated; however, we have often selected as state variables quantities that are not the most physically meaningful, e.g. the primed normalized number concentration parameter N_0' for ice. The scattering look-up tables are therefore used to convert the state variables into all the *derived variables* that might be of interest to users: as an example, to input to a radiative transfer code we may need to derive a vector $\mathbf{d}$ describing the profile of ice water content and effective radius, etc. To compute the retrieval RMS errors in $\mathbf{d}$, we first compute the error covariance matrix of $\mathbf{x}$ which is the inverse of the Hessian at the final iteration: $\mathbf{S_x} = \left(\partial^2 J/\partial \mathbf{x}^2\right)^{-1}$. The error covariance in $\mathbf{d}$ is given by

$\mathbf{S_d} = \mathbf{D S_x D}^T$, where $\mathbf{D} = \delta\mathbf{d}/\delta\mathbf{x}$ is a Jacobian matrix. The appendix of Delanoë and Hogan (2008) shows that $\mathbf{D}$ is very complex to implement manually; however, it is trivial to apply automatic differentiation to $\mathbf{D}$ in order to compute $\mathbf{d}$ and hence

390 $\mathbf{S_d}$. The square-root of the diagonal of $\mathbf{S_d}$ then provides the RMS error in $\mathbf{d}$, and error correlations between variables can also be computed.

In addition to the standard deviation error or RMSE for a particular quantity, the error covariance matrix yields the correlation between the errors of two variables at a particular gate, a value between -1 and 1. Secondly, the width of the diagonal band of the error covariance matrix around an element provides a measure of the vertical error correlation scale, given in metres.

395 Finally, the averaging kernel given by

$$\mathbf{A} = \mathbf{S_x H}^T \mathbf{R}^{-1} \mathbf{H}, \tag{3}$$

provides a measure of the information content of the retrieved state, such that an averaging kernel equal to the identity matrix would describe a retrieval in which all of the retrieved information comes from the observations. The effect of the priors, or of other physical constraints on the retrieval, are reflected by off-diagonal terms. The averaging kernel is used to derive the

400 averaging kernel sum, which reflects the contribution of the observations to the retrieved state, and the width of the diagonal, which indicates the smoothing of the retrieval compared to the true values (Pounder et al., 2012).

3   Case studies

Three simulated EarthCARE scenes have been produced by applying a state-of-the-art instrument simulator (Donovan et al., 2022) to a combination of high-resolution Global Environmental Multi-scale (GEM) numerical weather forecasts for clouds

405 and precipitation, merged with aerosols extracted from the Copernicus Atmospheric Monitoring Service (CAMS) (Qu, 2022). The test scenes have proved an invaluable tool for developing, testing, and evaluating EarthCARE retrieval algorithms and production model (Eisinger et al., 2022). Each scene corresponds to a granule, or a roughly 5000km or one eighth of an Earth-CARE orbit. The "Halifax" scene is a northern hemisphere midlatitude descending granule that passes over eastern Canada, the western Atlantic Ocean, and the Caribbean. The "Baja" scene is a northern-hemisphere midlatitude descending granule that

410 transects the North American continent and ends over the Baja peninsula. The "Hawaii" scene is a tropical descending granule over the central Pacific Ocean, beginning near Hawaii.

We have selected cloud, precipitation and aerosol regimes from within the test scenes as case studies for detailed evaluation. As these scenes have been generated from numerical models, we can access the model variables as "truth" for a more omniscient evaluation than is traditionally possible using in-situ measurements. This will help to demonstrate the performance

415 of ACM-CAP retrievals, as well as some of the challenges at the limits of the EarthCARE instruments; however, GEM is a numerical model that makes certain microphysical assumptions (e.g. the structure and density of snowflakes, and the drop size distribution of rain), which may not always be a good approximation to the real world, and which will differ from the prior assumptions and physical representations made in ACM-CAP. The details of some adjustments to the microphysical representation of ice, snow and supercooled liquid cloud output by the GEM model before input to the instrument simulators are given

420   in Section 7 of Qu (2022). Furthermore, as discussed in Qu (2022) and Donovan et al. (2022). the aerosol classes used in the CAMS model have been mapped to the HETEAC model for aerosol scattering properties before simulating the ATLID and MSI observations. Beyond ensuring that the scenes represent an adequate range of aerosol classes with distinct optical properties, it therefore cannot be ensured that the aerosol properties in the test scenes are realistic—especially in their interactions with cloud microphysical properties. These challenges will contribute to some of the differences between ACM-CAP and the
425   numerical model "truth" shown in the evaluation that follows.

3.1   Mid-latitude convective and stratiform precipitation

The first case features cold rain in convective and stratiform contexts from the Halifax scene. We show observed and forward-modelled EarthCARE measurements in Fig. 3, and retrieved and model quantities in Fig. 4. The first part of this case is dominated by light-to-moderate cold rain below stratiform mixed-phase cloud with tops around 5 km, beneath optically thin
430   ice clouds up to around 12 km. Heavier rain up to 10 mm h^{-1} is associated with an embedded convective cell around $39.5°$N, in which CPR is dominated by multiple scattering and attenuation. The second part of the scene features heavy precipitation up to 20 mm h^{-1} associated with deep convective clouds reaching around 13 km above sea level; physically and optically thick anvil cloud north of the deep convection overlays a shallow layer of liquid cloud at around 1 km.

The CPR and ATLID measurements are accurately forward-modelled across this scene (Fig. 3), indicating that the retrieval
435   is well-constrained by the available measurements—but not guaranteeing a unique solution in terms of retrieved quantities. In the retrieved IWC is systematically lower than the GEM model, especially in the optically thinnest cloud at 10 km above sea level (region A; Fig. 4c). While radar-lidar synergy is available in parts of this cloud the CPR signal is weak, so the retrieval is primarily constrained by ATLID. In the deepest ice clouds, IWC is also underestimated near cloud top (region C), around the level where lidar signal becomes fully attenuated; this Warm biases in infrared brightness temperatures (Fig. 3h) in these
440   regions and elsewhere with ice clouds (e.g. 44 to $45°$N, where the snow rate is also underestimated at 5 km, cf Fig. 4h), is likely related to these low IWC near the tops of these clouds lowering the effective radiative level of the clouds and increasing their infrared brightness temperature. This bias may be exacerbated by biases in atmospheric temperature used within the retrieval (X-MET data product, derived from ECMWF analysis), which can be 1 to 3 K warmer than that of the GEM model, especially in high clouds.

445   The retrieved snow rate at 5 km above sea level is a better match to the GEM model (Fig. 4i), but includes underestimates near the tops of stratiform cloud, such as at the poleward edge of region A. In region C deficits in retrieved IWC and snow rate are evident where CPR is extinguished: this illustrates the challenge of performing retrievals at the limits of the active sensors, and will be explored further in the tropical convection case (Section 3.2). While the retrieved snow is not sufficient to attenuate the radar, the retrieved rain rate in region C (Fig. 4j–l) is close to that in the GEM model, at least representing heavy enough
450   rain to saturate the forward-modelled PIA (Fig. 3e).

The forward-modelled ATLID Mie backscatter also broadly reproduces the measurements (Fig. 3g & f) in both optically thin (region A) and optically thick ice clouds (regions B & C), despite the underestimate of IWC. The rapid extinction of ATLID in
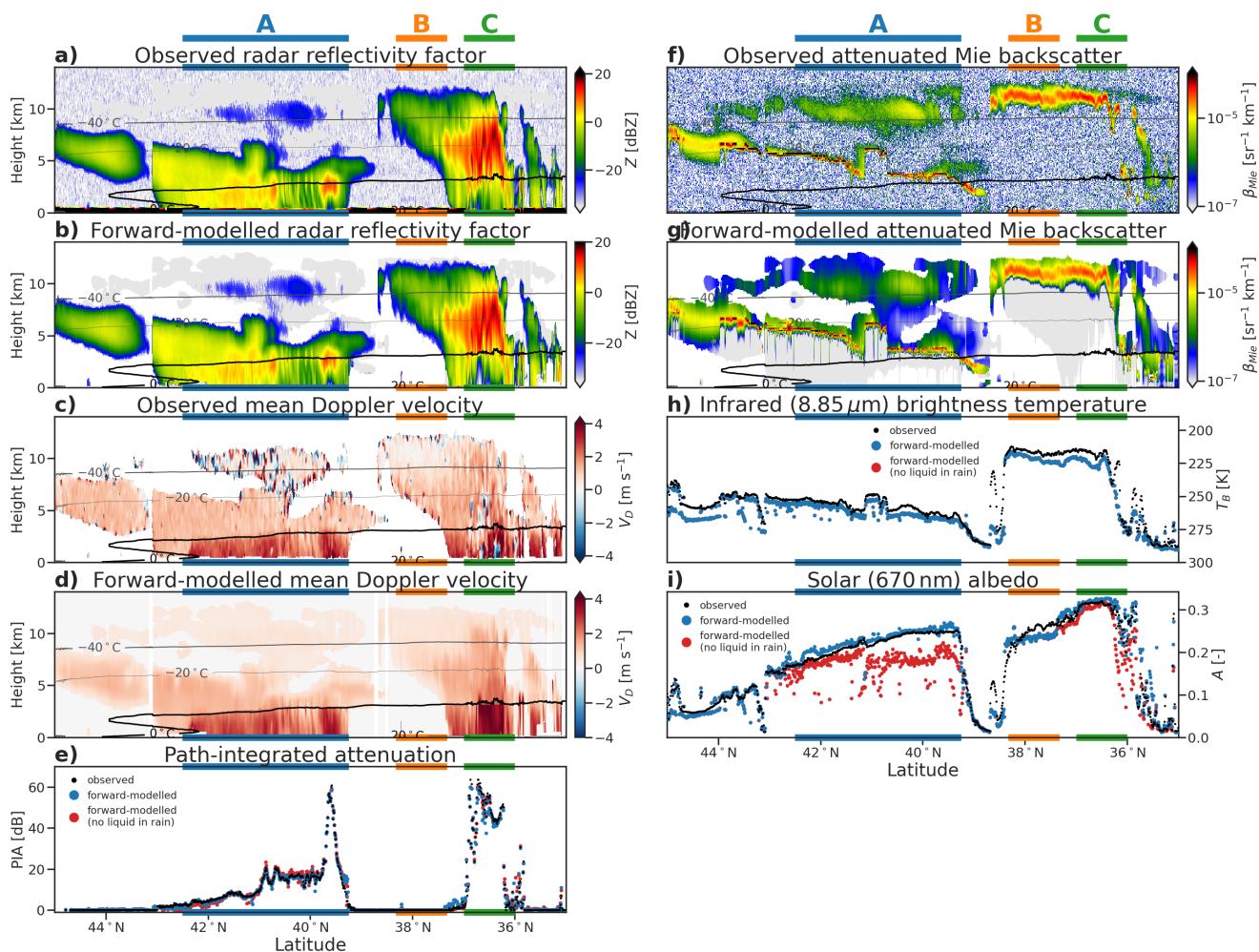
Figure 3. Simulated and forward-modelled CPR reflectivity factor (a & b), mean Doppler velocity (c & d) and PIA (e); ATLID attenuated Mie backscatter (f & g), MSI infrared brightness temperature (h) and solar albedo (i) for the midlatitude stratiform part of the Halifax scene. All profiling variables are overlaid with contours of atmospheric temperature from X-MET. Three areas of interest are highlighted.
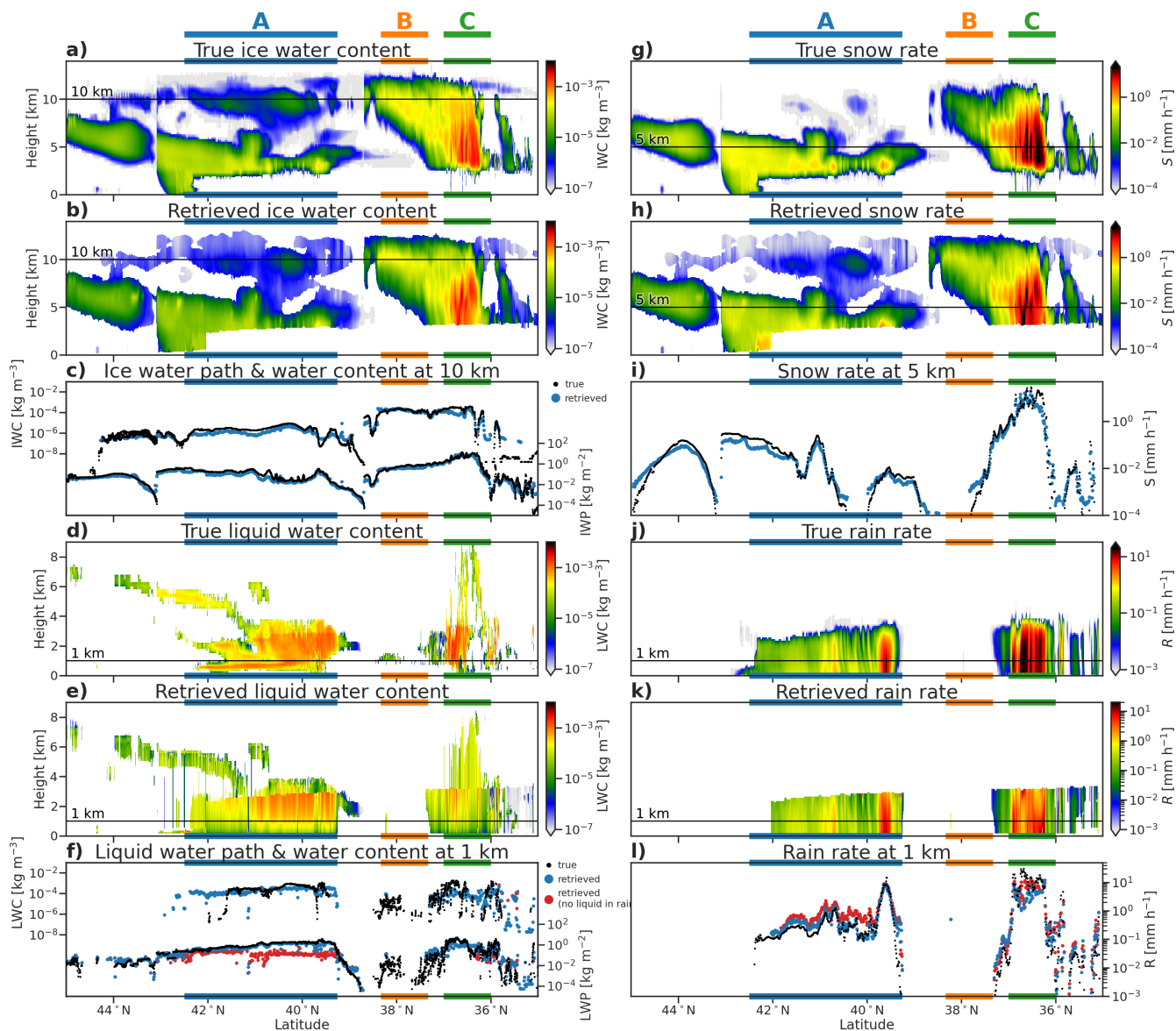
Figure 4. GEM model and ACM-CAP retrievals of IWC (a & b), LWC (d & e), snow rate (g & h), and rain rate (j & k), and comparisons of vertically-integrated and selected quantities for ice (c), liquid (f), snow (i) and rain (l), for the midlatitude stratiform part of the Halifax scene. Three areas of interest are highlighted.

the mixed-phase cloud-top layer of regime A is somewhat weaker than observed, corresponding to an underestimate in LWC in parts of these features (Fig. 4d & e).

455 The bulk of the retrieved liquid cloud, however, is collocated with rain in regions A and C (as described in Section 2.2.2) and its retrieval is not constrained by active measurements. Throughout this case, both the LWC at 1 km above sea level and the LWP (Fig. 4f) are remarkably close to the GEM model truth, even in convective precipitation (region C) where the spatial distribution of liquid water in the model is complex (Fig. 4d). As noted above, the deficit of CPR attenuation above the melting layer in region C is related to the deficit of IWC and snow rate: a remedy to this may be a more aggressive assumption to place

460 supercooled liquid throughout convective towers (up to and above 8 km above sea level in the GEM model). To illustrate the effect of retrieving liquid cloud in rain, an ACM-CAP retrieval in which liquid cloud is only retrieved where ATLID detects it is shown in red in Fig. 4f & l. The retrieved LWP throughout region A is underestimated by more than an order of magnitude, and the forward-modelled MSI shortwave channel (Fig. 3i) exhibits a 10% deficit in solar albedo. As both liquid cloud and rain contribute to the attenuation of CPR, which is strongly constrained within the retrieval by PIA, this deficit of liquid cloud

465 is also compensated by an overestimation of rain rate (Fig. 4l). In this stratiform cold rain regime solar radiance and radar PIA contribute to an accurate retrieval of LWP and rain rate.

3.2 Deep tropical convection

The equatorial part of the Hawaii scene is dominated by deep tropical cloud with tops around 18 km and a convective core with extreme precipitation beginning well above the melting level. This case provides an important check on the capacity

470 for synergistic retrievals in heavy precipitation (region A), where both ATLID and CPR are fully attenuated and passive and integrated measurements become saturated.

The retrieved IWC at 10 km above sea level (Fig. 6c) and rain rate at 1km above sea level (Fig. 6l) are remarkably close to the GEM model except in the convective core (region A). Here the retrieved rain rates are up to 10 mm h^{-1}, whereas the GEM model reaches values of 10 to 30 mm h^{-1}. As noted in the previous case, the greatest challenge is reproducing IWC and

475 snow rate within convective cores, where the radar reflectivity is affected by both attenuation and enhancement due to multiple scattering (Fig. 6c).

As in the mid-latitude stratiform precipitation, the presence and distribution of liquid cloud cannot be constrained by active instruments. The model truth includes a complex field of liquid cloud (Fig. 6d): scattered boundary layer clouds around 1 to 2 km and cloud layers close to the melting level, and in convective cores reaching from the surface to almost −40°C. To a

480 greater extent than in the mid-latitude case, where the top of a mixed-phase layer was detected by ATLID, the liquid clouds in this scene are almost completely obscured from the active instruments. Using the same approach to indirectly retrieve liquid cloud wherever rain is detected by CPR, the LWC at 1 km above sea level and LWP (Fig. 6f) are close to the GEM model in many parts of this scene (e.g. 5° to 3.5°N) but overestimated in others (2.5°N to 2°S) where the vertical distribution of liquid clouds is more limited. Comparing the forward-modelled solar albedo when liquid cloud is not retrieved in rain (Fig. 6f)

485 shows that this approach can greatly improve the representation of cloud in complex scenes. As in the previous scene, the most extreme mismatches to the MSI solar albedo are beneath non-precipitating ice clouds on either edge of the convective system,
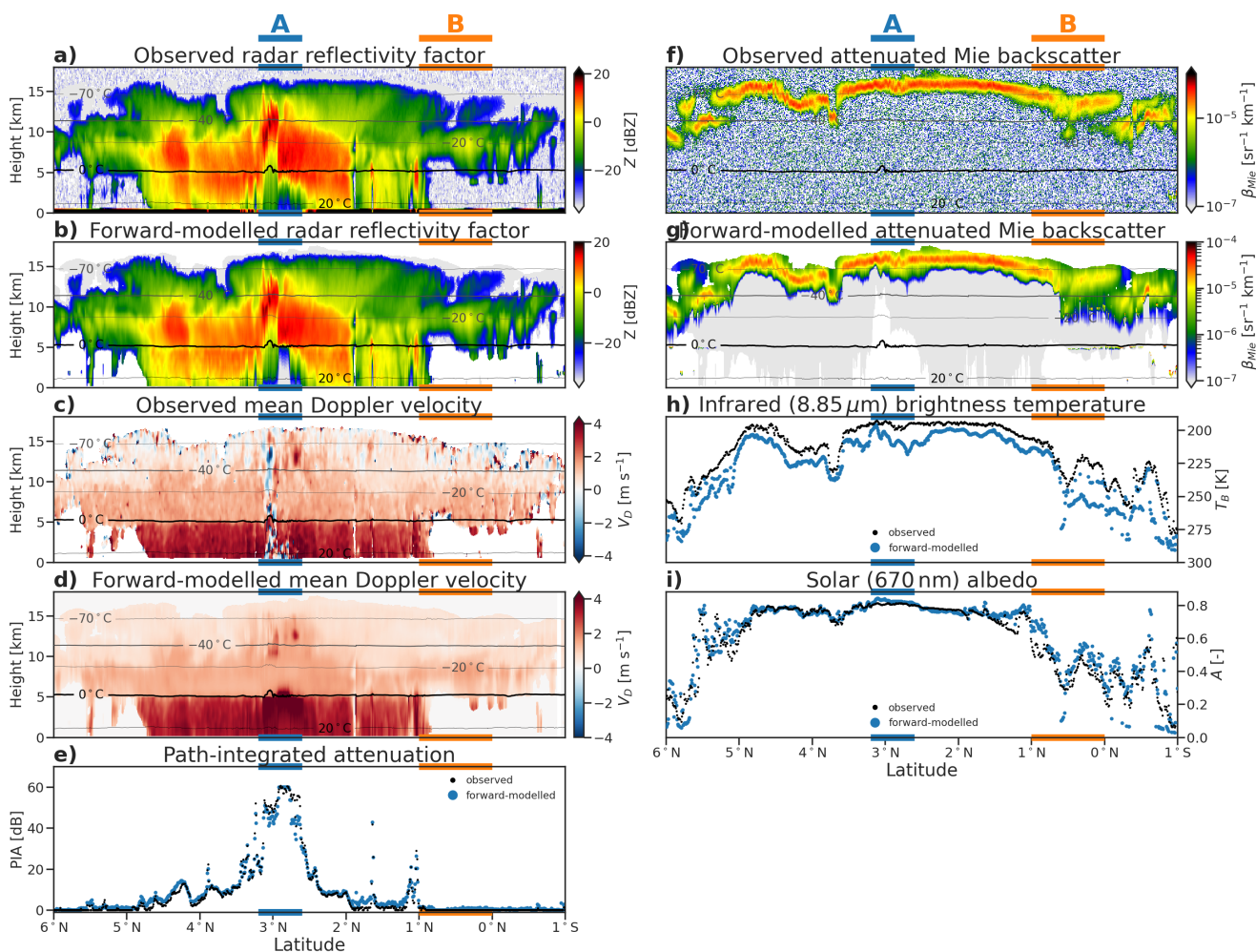
Figure 5. Simulated and forward-modelled CPR reflectivity factor (a & b), mean Doppler velocity (c & d) and PIA (e); ATLID attenuated Mie backscatter (f & g), MSI infrared brightness temperature (h) and solar albedo (i) for the deep convective part of the Hawaii scene. All profiling variables are overlaid with contours of atmospheric temperature from X-MET.
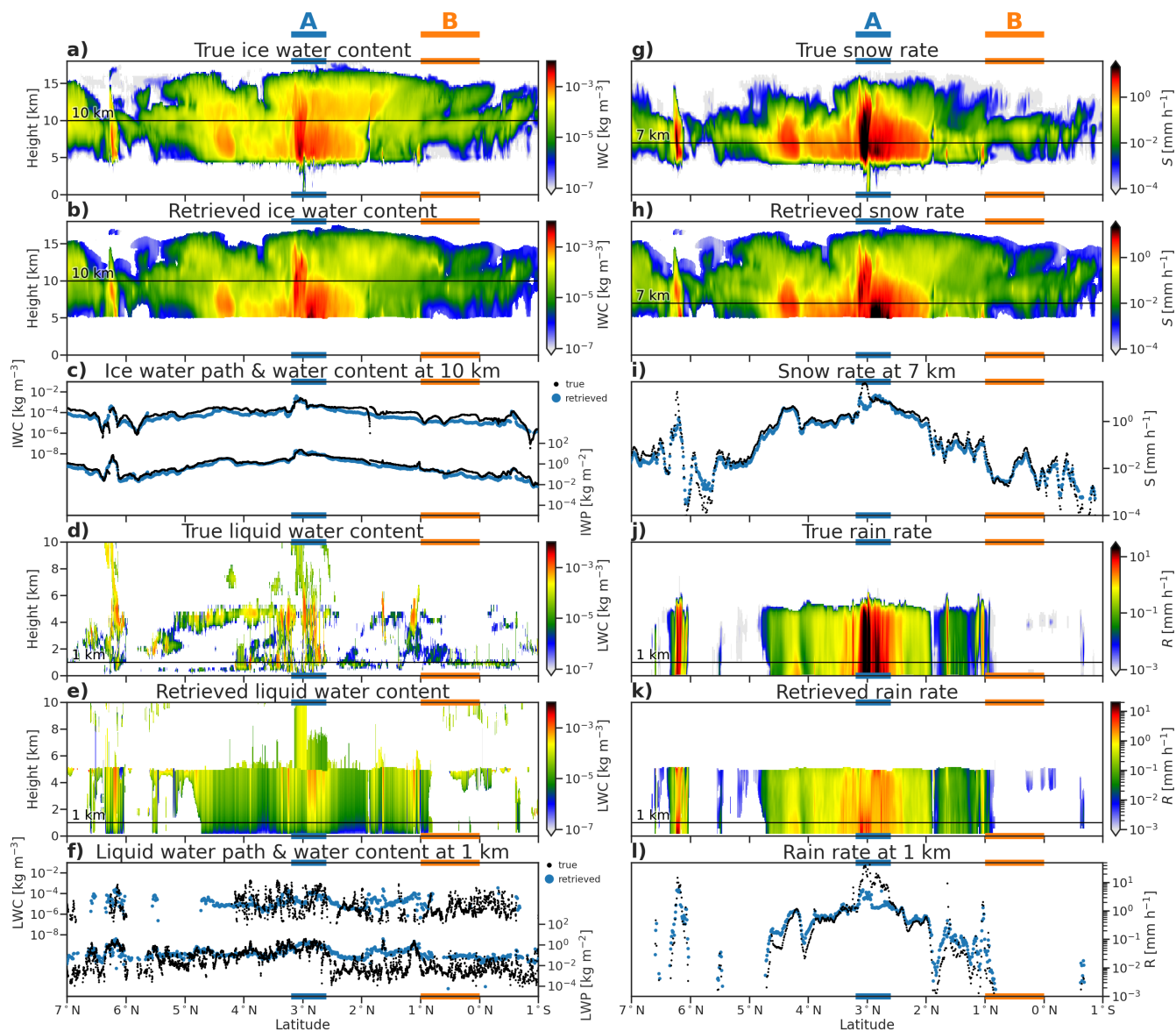
Figure 6. GEM model and ACM-CAP retrievals of IWC (a & b), LWC (d & e), snow rate (g & h), and rain rate (j & k), and comparisons of vertically-integrated and selected quantities for ice (c), liquid (f), snow (i) and rain (l), for the deep convective part of the Hawaii scene.

where neither liquid clouds nor rain are diagnosed in the target classification, but where the GEM model includes shallow layers of low non-precipitating cloud (region B)

3.3 High-latitude mixed-phase clouds

490   The high-latitude part of the Halifax scene features mixed-phase clouds at night, transitioning from deeper clouds with tops up to 6 km around 65°N with supercooled liquid in convective cells, to mixed-phase clouds with tops around 3 km at temperatures as cold as −30°C, and finally more broken shallow mixed-phase clouds toward 50°N.
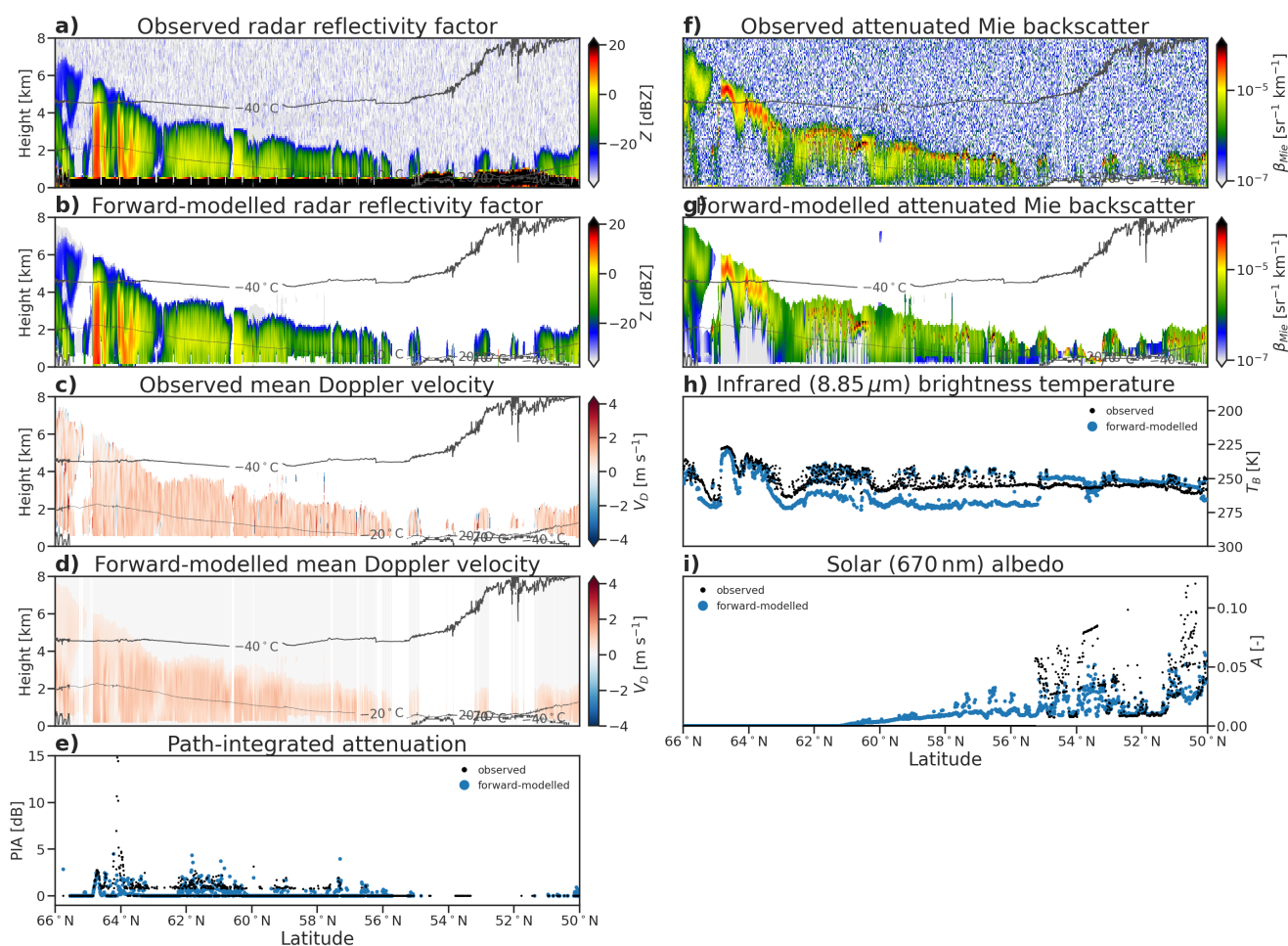
Figure 7. Simulated and forward-modelled CPR reflectivity factor (a & b), mean Doppler velocity (c & d) and PIA (e); ATLID attenuated Mie backscatter (f & g), MSI infrared brightness temperature (h) and solar albedo (i) for the high latitude mixed-phase part of the Halifax scene. All profiling variables are overlaid with contours of atmospheric temperature from X-MET.
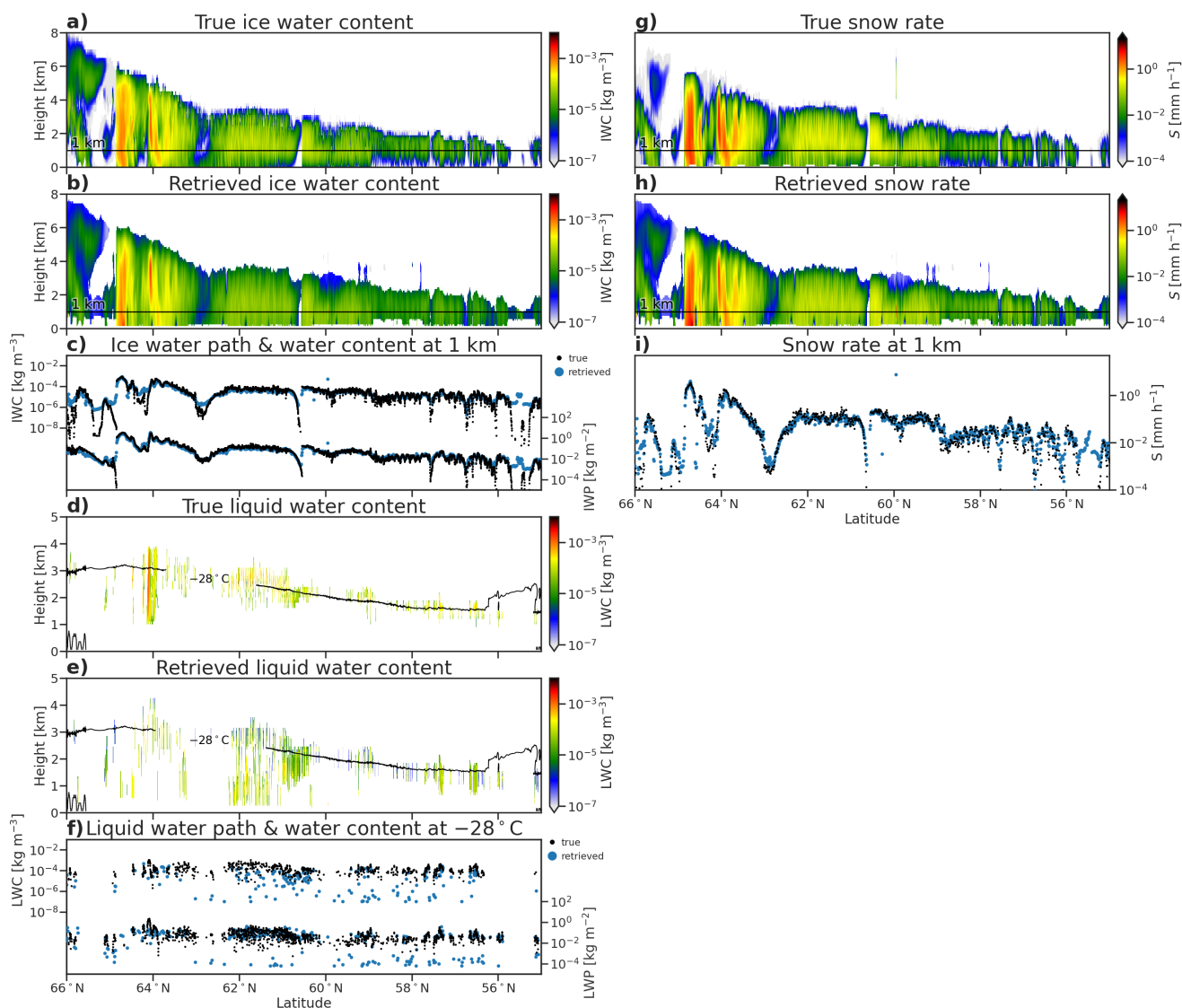
Figure 8. GEM model and ACM-CAP retrievals of IWC (a & b), LWC (d & e), snow rate (g & h), and rain rate (j & k), and comparisons of vertically-integrated and selected quantities for ice (c), liquid (f), snow (i) and rain (l), for the high latitude mixed-phase part of the Halifax scene.

Without solar radiances, the simultaneous retrieval of ice and supercooled liquid is constrained only by the active instruments. The retrieved IWC at 1 km above sea level (Fig. 8c) is close to the model truth throughout the scene, while the retrieved

495 LWC at the $-28°C$ isotherm (Fig. 8f) is under-estimated by three orders of magnitude, as is the LWP (Fig. 4f). In day-lit scenes solar radiances would provide a stronger integrated constraint on liquid water path; the PIA (Fig. 4c) provides little constraint in this part of the scene, where the CPR is only weakly attenuated by supercooled liquid clouds.

3.4 Maritime and continental aerosol layers

The subtropical part of the Halifax scene is dominated by two distinct overlapping layers of aerosols—sea salt from the ocean

500 surface up to 2 to 4km, with continental pollution aloft up to 6 to 8km—with shallow cumulus clouds embedded in the lowest 2km. Figure 9 shows the simulated and forward-modelled ATLID and MSI measurements through this scene, and Figure 10 the retrieved aerosol extinction, total aerosol optical thickness, and lidar (or extinction-to-backscatter) ratio.
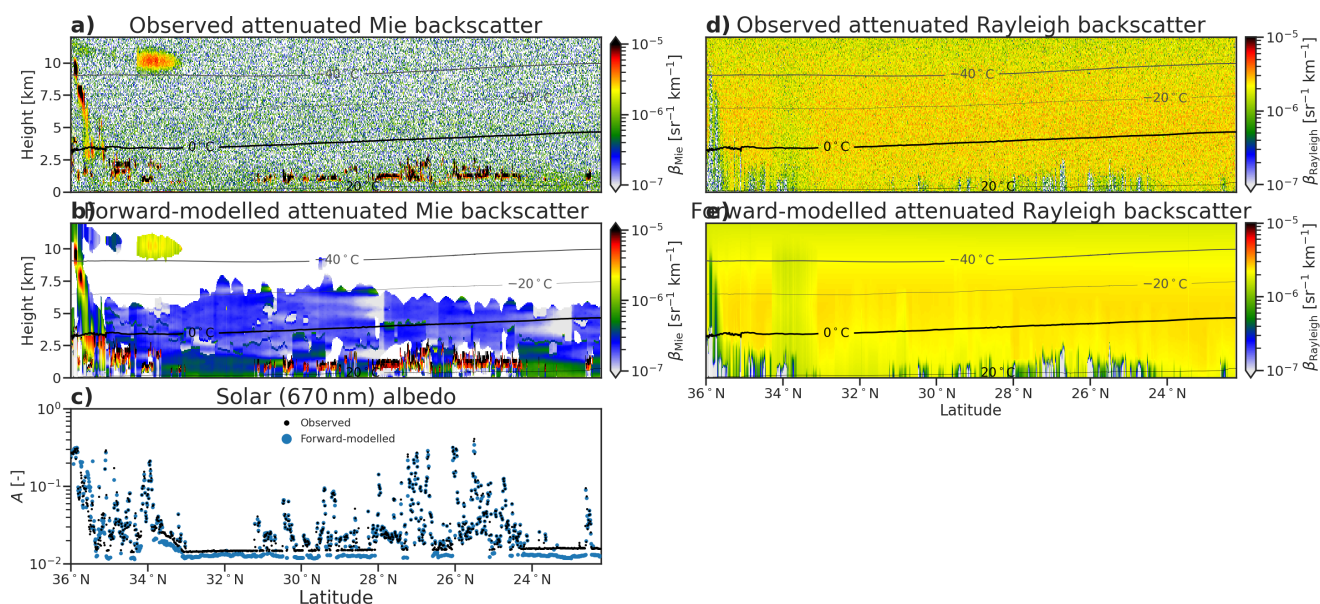
Figure 9. Simulated and forward-modelled ATLID attenuated Mie backscatter (a & b), MSI solar albedo (c) and ATLID attenuated Rayleigh backscatter (d &e e) for the subtropical part of the Halifax scene. All profiling variables are overlaid with contours of atmospheric temperature from X-MET.

The measured ATLID Mie backscatter (Fig. 9a) shows the high degree of measurement noise from which the signal of aerosol backscatter must be detected, in contrast to the clear signal from the high ice and liquid boundary layer clouds in the

505 same scene. The forward-modelled Mie backscatter (Fig. 9b) does not add noise, and shows that the signal is often less than 1×10^{-6} sr^{-1} km^{-1}.

The retrieved aerosol extinction (Fig. 10b) shows that the retrieval resolves some of the key vertical features within both the sea salt and continental pollution layers: in the sea salt the strongest extinction is within 1km of the surface on the equator-ward
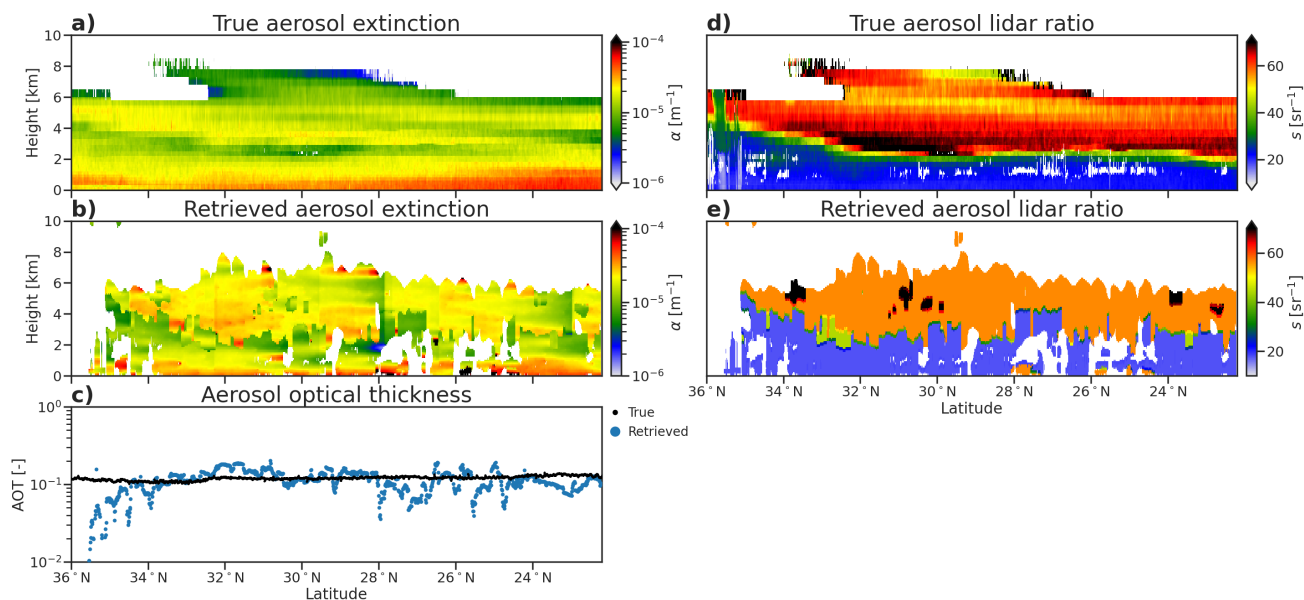
Figure 10. CAMS model and ACM-CAP retrievals of aerosol extinction (a & b), aerosol optical thickness (c), and aerosol lidar ratio (d & e) for the subtropical part of the Halifax scene. Note that in ACM-CAP the lidar ratio is not retrieved, but is a property of each HETEAC aerosol species, as mapped to the CAMS aerosol classes.

side of the scene, while embedded within the continental pollution layer are 2km-deep structures of stronger extinction. Many
510 horizontal features and discontinuities in the retrieved extinction are not found in the model variables, and reflect the challenges
of applying a Kalman smoother across large spatial scales when the target classification is interrupted by hydrometeors, mostly
liquid clouds in this case.

The forward-modelled lidar ratio in aerosols (Fig. 10e) shows that the model variables are only coarsely resolved. This
illustrates the approach taken within ACM-CAP to making the lidar ratio a fixed property of each HETEAC aerosol type. The
515 result is that some of the structure, especially within the continental pollution layer where S varies between 50 and 70 sr^{-1}
over around 3km of the layer, are not resolved. This likely contributes to the over-estimated aerosol extinction in the lowest
part of the continental pollution layer (around 4km above the surface between 33-28°N).

4 Statistical evaluation

In addition to case studies, we also evaluate the retrieval statistically in order to diagnose biases and sensitivities. Here we
520 combine all data from the three synthetic EarthCARE granules, to evaluate ACM-CAP retrievals of ice and snow, liquid
clouds, rain and aerosols against equivalent quantities from the GEM and CAMS models. We also statistically compare the
forward-modelled and observed measurements from ATLID, CPR and MSI.

4.1 Ice and snow

To evaluate the retrieved ice clouds and snow against the GEM model, and the forward-modelled EarthCARE measurements
525   against those simulated from the GEM model, we show probability density functions (PDFs; top panels) and joint histograms
(lower panels) of snow rate, IWC and effective radius (Figs. 11 & 12). The PDF of all "true" data (black PDFs) is compared
against those in volumes correctly classified by radar-lidar synergy (grey shading; these two will be identical for the measured
variables except due to measurement noise): the difference between these PDFs illustrates the portion of ice clouds that are
not recoverable from EarthCARE's instruments. The other PDFs show equivalent quantities retrieved or forward-modelled by
530   ACM-CAP everywhere (red lines) and specifically in volumes which do not contain ice or snow in the model (red shading). The
joint histograms show the degree to which the retrieved and forward-modelled quantities represent the true values, including
correlation coefficients (r) and root-mean-square (RMS) errors.
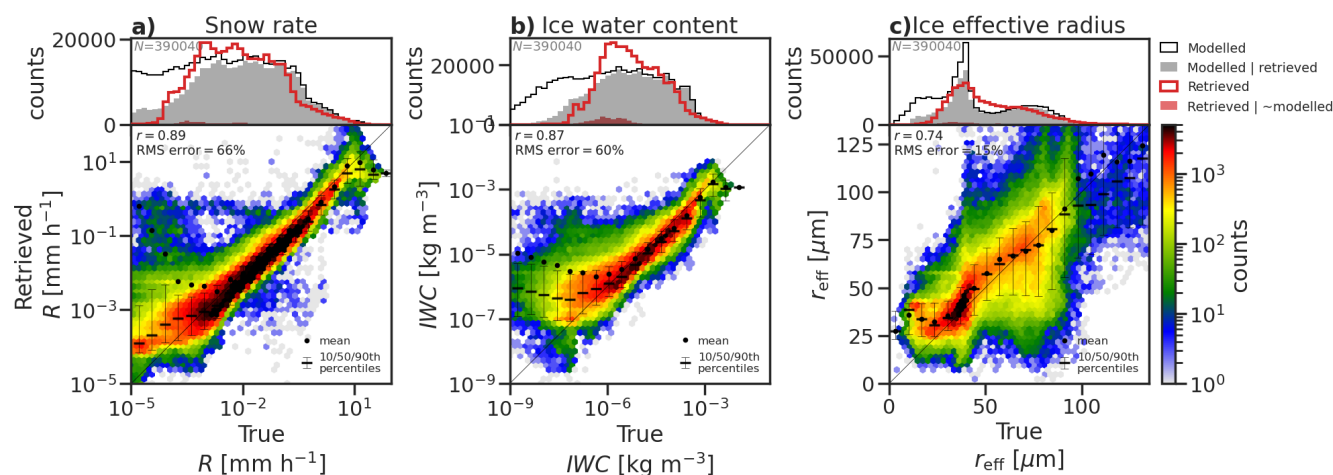
Figure 11. 1D (above) and joint (below) histograms comparing true (GEM model) and retrieved (a) snow rate, (b) IWC and (c) effective
radius for the three GEM scenes. The 1D histograms compare all GEM model data (black line) and GEM model data where ATLID or CPR
correctly detect ice cloud (grey shading) against retrieved ACM-CAP data (red line) and retrievals in pixels which do not contain ice cloud
(red shading).

The retrieved snow rates (Fig. 11a) and IWC (Fig. 11b) have similar high correlations ($r = 0.89$ & 0.87) and errors (66%
& 60%); this suggests that the variability in the fallspeeds of ice particles—as represented by the numerical model's four ice
535   species—does not dominate the uncertainty in the retrieval. The radar reflectivity is very faithfully reproduced (Fig. 12a), but
the mean Doppler velocity (Fig. 12b)—which performs well in stratiform snow where the measurement is dominated by the
terminal velocity of particles—has high degree of spread when the forward-modelled values are less than 1 m s^{-1}, such as in
ice clouds dominated by smaller particles.

The GEM model ice effective radius (r_{eff}) (Fig. 11c) has a bimodal distribution (black line) as an artefact of distinct ice and
540   snow habits in the model microphysics scheme. As ice and snow are represented as a continuum in ACM-CAP, the retrieved

high
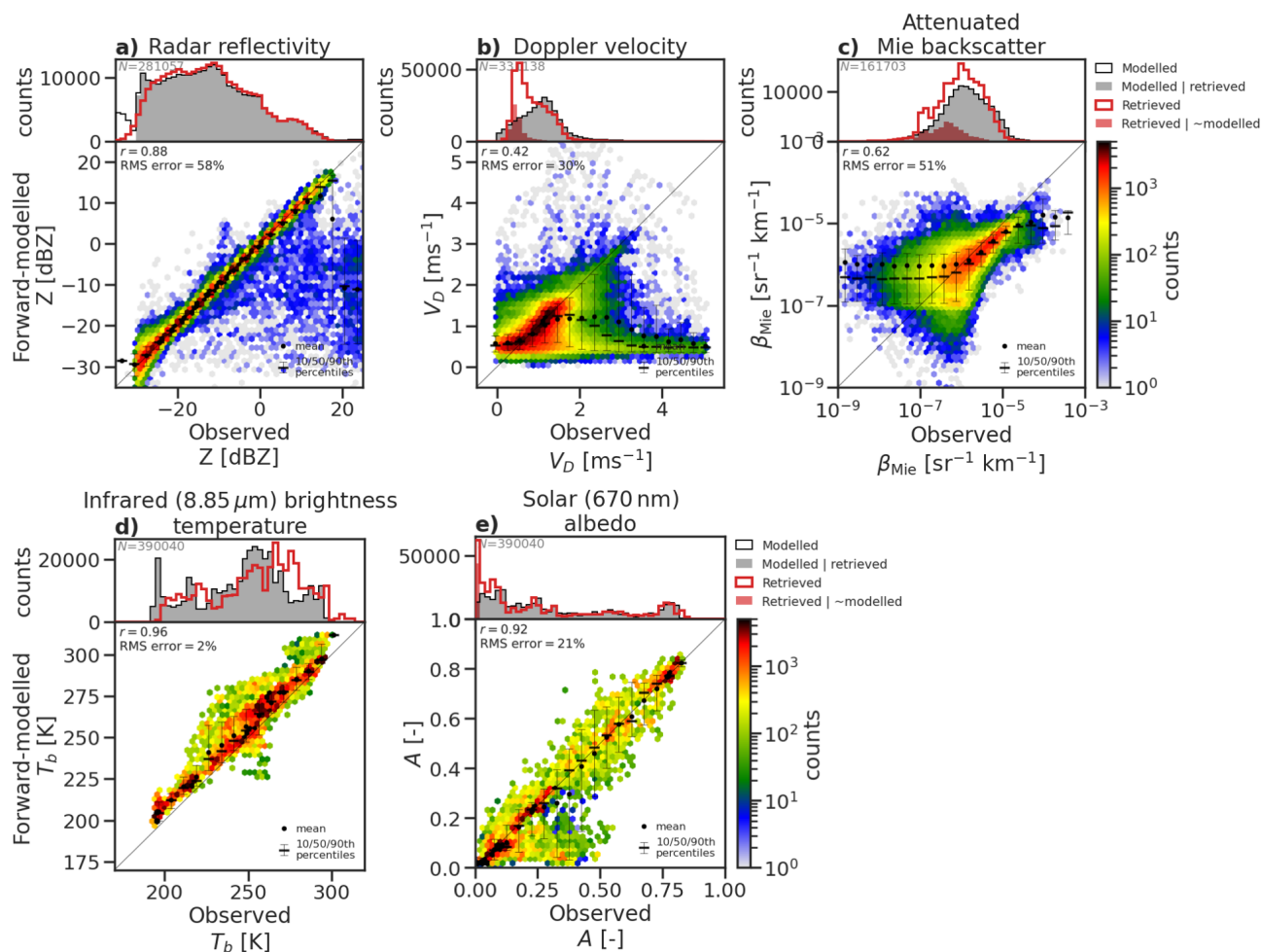
Figure 12. Histograms of observed and forward-modelled CPR radar reflectivity (a), mean Doppler velocity (b) and ATLID attenuated Mie backscatter (c) in pixels containing ice and snow across all three ECCC scenes. The 1D histograms compare all GEM model data (black line) and GEM model data where ATLID or CPR correctly detect ice cloud (grey shading) against retrieved ACM-CAP data (red line) and retrievals in pixels which do not contain ice cloud (red shading).

distribution of ice effective radius (red line) is not bimodal, but shows a tends to underestimate the frequency of the lowest and highest effective radii. The joint histogram (Fig. 11c) shows that the retrieved effective radius is well-correlated with the GEM model ($r = 0.74$) with low RMS error (15%); the greater apparent variability in this quantity is due to the linear, rather than logarithmic, scale. The thermal infrared and solar radiances (Fig. 12d & e) both exhibit strong correlations. The warm

545 bias in the infrared brightness temperatures includes a contribution from low IWC near cloud-top highlighted in parts of the case studies, but—since a similar bias is also evident for liquid clouds discussed next—we attribute this largely to systematic differences between the atmospheric temperatures in the GEM model and the ECMWF forecasts used to inform the retrieval. By contrast, the shortwave albedo has a higher degree of random error, but is not biased.

4.2 Liquid cloud
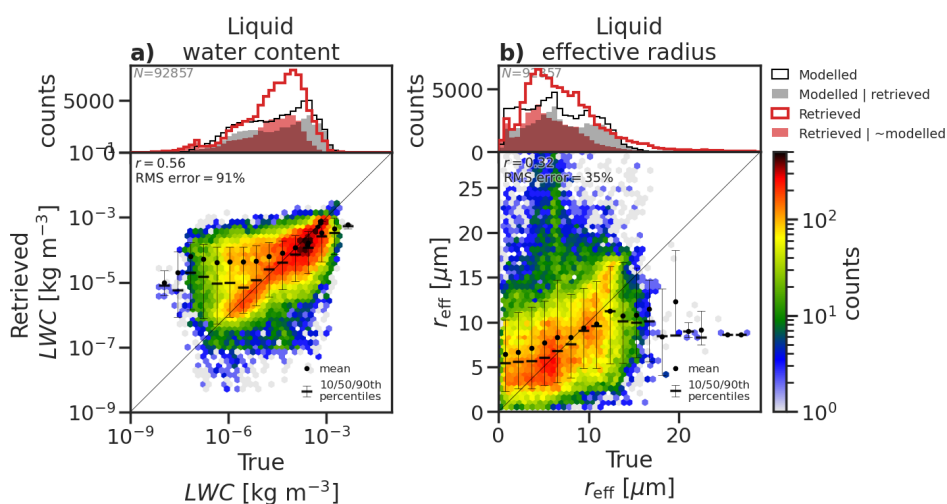
Figure 13. 1D (above) and joint (below) histograms comparing true (GEM model) and retrieved (ACM-CAP) liquid water content (left), extinction (middle) and effective radius (right) for the three ECCC scenes. The 1D histograms compare all GEM model data (black line) and GEM model data where ATLID or CPR correctly detect liquid cloud (grey shading) against retrieved ACM-CAP data (red line) and retrievals in pixels which do not contain liquid cloud (red shading).

550 In this section we evaluate the retrieved properties of liquid clouds (Fig. 13) and forward-modelled ATLID and MSI measurements (Fig. 14). As shown in the case studies, the retrieval of liquid cloud in rain results in an overall improvement in retrieved LWP, but the smooth spatial distribution of liquid cloud often differs from those in the GEM model (cf. red shaded PDF in Fig. 13). Overall the retrieval of liquid clouds are unbiased and moderately correlated with the GEM model variables, but with a high degree of random error.

555 An evaluation of the forward-modelled ATLID Mie backscatter and MSI shortwave albedo in liquid clouds helps evaluate the extent to which the available measurements are correctly interpreted in the retrieval. The attenuated Mie backscatter (Fig. 14a) reflects a moderately good fit to the ATLID measurements: the primary peak in backscatter from liquid clouds
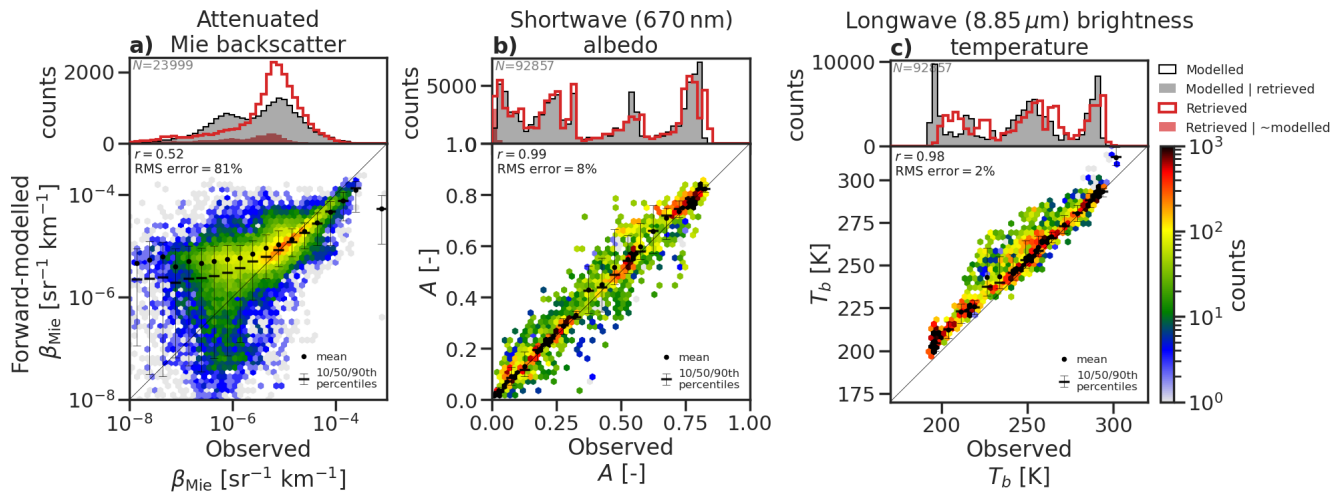
Figure 14. Histograms of observed and forward-modelled ATLID attenuated Mie backscatter (a) and MSI shortwave albedo (b) in pixels and profiles containing liquid cloud across all three ECCC scenes. The 1D histograms compare all GEM model data (black line) and GEM model data where ATLID or CPR correctly detect liquid cloud (grey shading) against retrieved ACM-CAP data (red line) and retrievals in pixels which do not contain liquid cloud (red shading).

around 1×10^{-5} sr^{-1}km^{-1} is well-represented, but the correlation rapidly deteriorates at lower values (i.e. where the signal is becoming extinguished). The fit to MSI shortwave albedo is extremely good ($r = 0.99$ with RMSE of 8%).

560   In common with the ice clouds, the MSI longwave (8.85 μm) channel has a slight warm bias for liquid clouds. A comparison of the temperature fields from the GEM model and X-MET data product derived from the ECMWF analysis revealed temperature differences of as much as several degrees, including positive biases near cloud-top in the three test scenes, which are likely to explain part of the observed bias in MSI longwave channels, and which contributes to the uncertainties in this evaluation.

While the benefit of retrieving liquid cloud in rain was clear from the case studies, it is important that the assumption of
565   liquid cloud in rain can be used at night without introducing a bias to the retrieval. The high-latitude mixed-phase case study (Fig. 8) showed that ACM-CAP may occasionally underestimate LWP at night, but it is not clear to what extent this would be improved by the availability of solar radiances. As a test, we ran all three scenes without assimilating solar radiances, and found that, while LWC exhibited more random error, the retrievals were not biased. This indicates that the priors and uncertainties used are broadly appropriate—at least across the cloud regimes sampled by the test scenes. A more robust check will be to
570   apply the same test using a large number of A-Train orbits.

4.3   Rain

The retrieved properties of rain (Fig. 15) and forward-modelled CPR measurements (Fig. 16) indicate that the rain rate (Fig. 15a) is strongly constrained by radar reflectivity and PIA (Fig. 16a & c) through light to moderate rain rates. A low bias is evident in the heaviest rain (10 mm h^{-1} and above), corresponding to an underestimate in the highest radar reflectivity
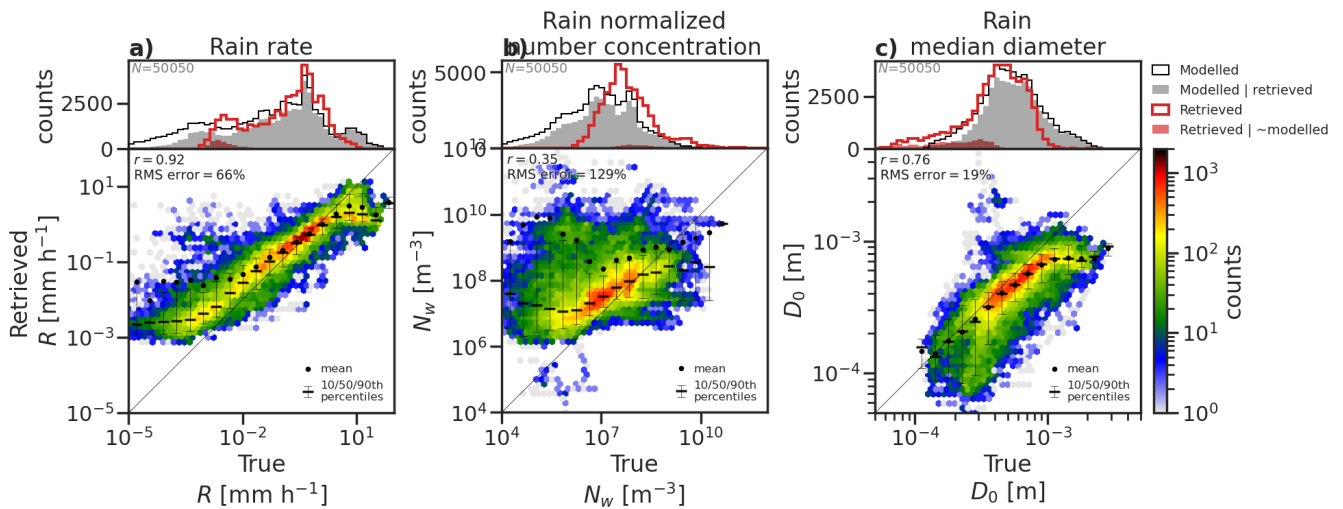
Figure 15. Histograms of GEM model quantities and ACM-CAP retrievals of (a) rain rate, (b) normalized number concentration, and (c) median diameter. 1D (above) and joint (below) histograms comparing true (GEM model) and retrieved (ACM-CAP) rain water content for the three ECCC scenes. The 1D histograms compare all GEM model data (black line) and GEM model data where ATLID or CPR correctly detect rain (grey shading) against retrieved ACM-CAP data (red line) and retrievals in pixels which do not contain rain (red shading).
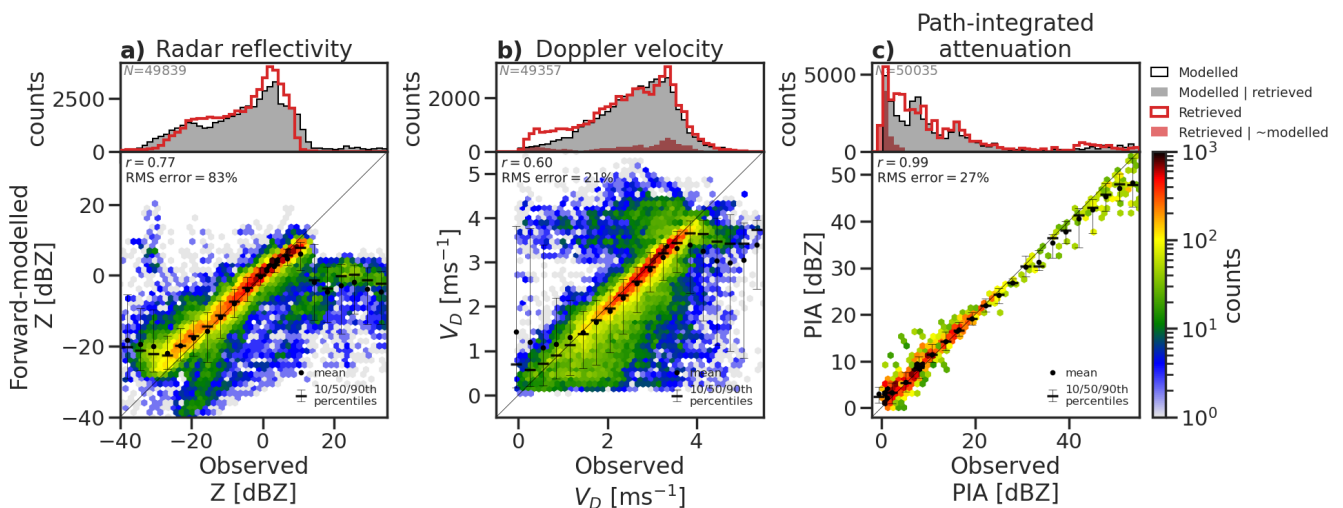
Figure 16. Histograms of observed and forward-modelled CPR radar reflectivity (a), mean Doppler velocity (b) and path-integrated attenuation (c) in pixels and profiles containing rain across all three ECCC scenes. The 1D histograms compare all GEM model data (black line) and GEM model data where ATLID or CPR correctly detect rain (grey shading) against retrieved ACM-CAP data (red line) and retrievals in pixels which do not contain rain (red shading).

575   factors in rain (above around 15 dBZ). The integrated constraint on attenuation due to both liquid cloud and rain has a correlation of $r = 0.99$ and an RMS error of 27%. In contrast, the mean Doppler velocity (Fig. 16b) and parameters of the rain drop size distribution (DSD; Fig. 15b & c) indicate some challenges in retrieving the microphysics of rain. While mean Doppler velocity has a moderate correlation with measurements and a weaker impact from measurement noise than was observed for ice cloud, a low bias is evident in the raindrop terminal velocity which is reflected in high biases in normalized number concentra-

580   tion and underestimates in median diameter, especially in heavy rain with relatively low concentrations of large raindrops (N_w less than the Marshall-Palmer value and D_m greater than 1 mm). This is likely related to the retrieval being over-constrained by priors when the measurements are near the limits of the CPR within heavy precipitation, and may suggest the need to set modified priors for rain within profiles identified as convective.

4.4   Aerosols

585   The retrieved properties of aerosols and forward-modelled ATLID and MSI measurements (Fig. 18) show that aerosols are the most challenging aspect of the retrieval, given the relatively weak signals and the related issues of characterising surface properties for the passive measurements. As discussed earlier, the aerosol quantities in the test scenes have been extracted from the CAMS model and mapped to the HETEAC species in preparation for inclusion in the simulated test scenes, and while care has been taken to represent a realistic range of aerosol properties, the realism of the model truth is not guaranteed.

590   ACM-CAP relies on the ATLID classification of aerosols to determine the physical properties of each aerosol type including their scattering properties and size distribution, and retrieves its number concentration by which all other quantities such as extinction and mass content are determined. Hence the quantized distribution of forward-modelled aerosol lidar ratio, with values corresponding to each aerosol classification (Fig. 17b).

      The forward-modelled attenuated Mie and Rayleigh backscatter (Fig. 18a & b) have little relation to the measurements at

595   the scale of the JSG. This is consistent with the forward-modelled measurements in the aerosol case study (Fig. 9), where noise dominates the simulated ATLID measurements in the aerosol. This demonstrates the importance of the Kalman smoother for extracting information on larger spatial scales. Errors in the solar albedo appear to be dominated by the land surface (i.e. the Baja scene, crossing North America) with a higher degree of scatter at moderate to high albedos, but a relatively close correlation over the ocean (i.e. where $A < 0.2$).

600 ## 5   Discussion and conclusions

      The ACM-CAP product uses the synergy of all available measurements from EarthCARE's active and passive sensors to retrieve profiles of ice and snow, rain, liquid clouds and aerosols simultaneously. Such a unified retrieval product has never been produced from the synergy of spaceborne instruments, and has the advantage of facilitating retrievals even in mixed-phase, layered, and heavily precipitating scenes. This is a priority for the EarthCARE production model, in which the retrievals are

605   used to compute broadband heating rates and inform a top-of-atmosphere radiative closure assessment. In this study we have described the innovative CAPTIVATE optimal estimation retrieval framework in its configuration for the ACM-CAP processor,
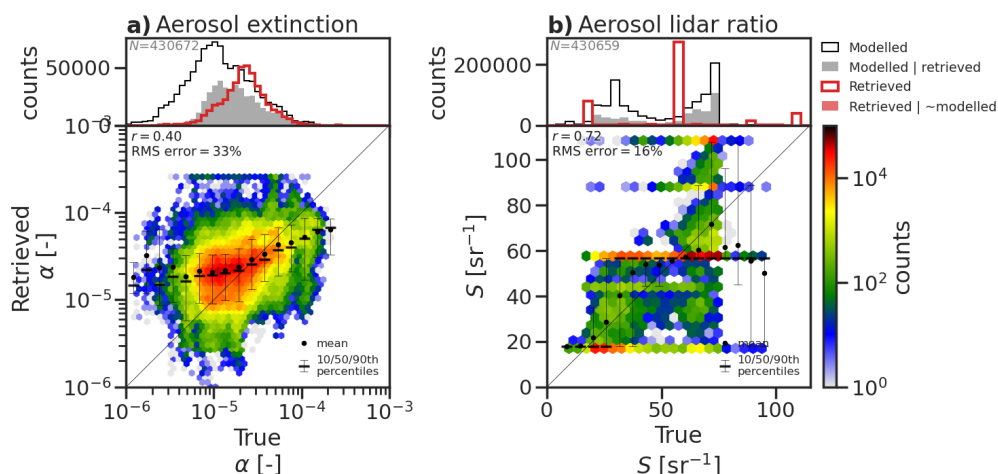
Figure 17. Histograms of GEM model quantities and ACM-CAP retrievals of (a) aerosol extinction and (b) aerosol lidar ratio. 1D (above) and joint (below) histograms comparing true (GEM model) and retrieved (ACM-CAP) aerosol quantities for the three ECCC scenes. The 1D histograms compare all GEM model data (black line) and GEM model data where ATLID or CPR correctly detect rain (grey shading) against retrieved ACM-CAP data (red line) and retrievals in pixels which do not contain rain (red shading).
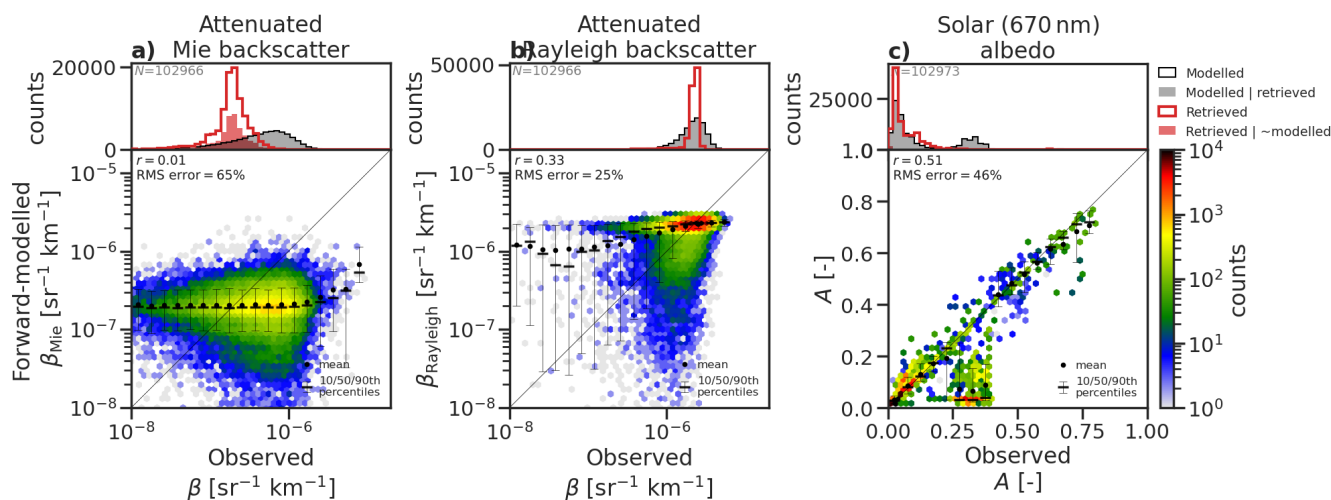
Figure 18. Histograms of observed and forward-modelled (a) ATLID attenuated Mie backscatter, (b) attenuated Rayleigh backscatter, and (c) MSI shortwave albedo in pixels and profiles containing aerosols across all three ECCC scenes. The 1D histograms compare all GEM model data (black line) and GEM model data where ATLID or CPR correctly detect aerosols (grey shading) against retrieved ACM-CAP data (red line) and retrievals in pixels which do not contain aerosols (red shading).

and presented a detailed evaluation of ACM-CAP's performance across three EarthCARE granules. Taking as "truth" the numerical model fields used to generate the test scenes, it was possible to evaluate the retrieval more thoroughly than will be possible using in-situ or remotely-sensed measurements to evaluate in-flight EarthCARE retrievals—with the caveat that the numerical weather model is not a perfect physical representation of aerosols, hydrometeors or the atmosphere, and some biases in the model may contribute to apparent errors and uncertainties presented here. Over the test scenes the retrieved IWC and snow rate were closely correlated with the model truth; however, the snow rate was subject to more than twice the random error as IWC, reflecting the high degree of natural variability in the structure and terminal fallspeed of snowflakes. The rain retrieval is constrained chiefly by CPR measurements, with important contributions from Doppler velocity and PIA. The retrieved rain rate was very highly correlated with model truth, with a similar random error to that of IWC. Liquid clouds and aerosols are more challenging to retrieve: both rely on detection by ATLID, which is obscured by cloud aloft and rapidly extinguished in liquid cloud layers. Retrieved LWC was moderately correlated with that in the model, with RMS error around 110%. The retrieved aerosol extinction is moderately correlated with the model, but has a lower RMS error of around 32%.

The quality of ACM-CAP retrievals is dependent on the accuracy of the target classification used to determine where different species are to be retrieved. A significant limitation of radar-lidar synergy, whether from CloudSat-CALIPSO or EarthCARE, is that of detecting liquid cloud which is embedded within ice clouds or rain, or beneath layered cloud scenes, after ATLID is extinguished. An evaluation of the target classifications in the same test scenes (Irbah et al., 2022) showed that liquid cloud is correctly identified by ATLID in around 22% of sampling volumes—less than 10% of LWC—primarily at the tops of mixed-phase clouds and shallow liquid clouds unobscured by optically thick clouds. We have demonstrated a simple assumption to compensate for this missing liquid cloud, which is to assume liquid cloud is present wherever rain is detected by CPR. While this does not resolve the small-scale structure of liquid clouds within rain, it is sufficient to correctly identify around 54% of volumes containing liquid cloud, or 65% of LWC (Irbah et al., 2022). We have showed that making the assumption of liquid cloud in rain within ACM-CAP greatly improves the assimilation of MSI solar radiance channels in stratiform rain, constraining the retrieval of a simple profile of LWC. An equally important benefit is that the contribution of this liquid cloud to radar attenuation helped to reduce a positive bias in retrieved rain rate in the stratiform precipitation regime. Finally, the fact that liquid cloud in the context of stratiform rain can be coarsely constrained by MSI solar radiance channels suggests its importance for the shortwave broadband fluxes at top-of-atmosphere. This contributes to ACM-CAP's superior shortwave TOA radiative closure when compared to the composite of single-instrument retrievals (Barker et al., 2022b).

We have described the configuration of the ACM-CAP retrieval prior to EarthCARE's launch, and evaluated its performance using test scenes generated from a numerical weather prediction model. The test scenes have proved invaluable in facilitating the maturity of the retrieval processor ahead of EarthCARE's launch; however, the three orbital granules represent around 5% of one day of in-flight EarthCARE data, and cannot provide a thorough coverage of all the regions and regimes that are of interest for EarthCARE science. Prior to launch, field campaign data, additional simulated scenes, and A-Train data will provide further potential to test and evaluate ACM-CAP, including against in-situ measurements of cloud, aerosol and precipitation properties. Ultimately the configuration of the ACM-CAP retrieval algorithm will be updated as necessary in response to the exposure

of the entire EarthCARE production model to in-flight measurements, in coordination with EarthCARE calibration/validation campaigns.

Data availability. The EarthCARE Level-2 demonstration products from simulated scenes, including the ACM-CAP product discussed in this paper, are available from https://doi.org/10.5281/zenodo.7117115

645 *Author contributions.* All authors contributed to the development of the ACM-CAP product and CAPTIVATE algorithm, and to the writing and correction of the manuscript. Shannon Mason led on the analysis and evaluation of the ACM-CAP product and the writing of the manuscript, while Robin Hogan provided the detailed description of the algorithm.

Competing interests. The authors have no competing interests.

Acknowledgements

650 This work has been funded by ESA grants 4000104528/11/NL/CT (VARSY), 4000112030/15/NL/CT (DORSY) and 4000134661/21/NL/AD (CARDINAL), and the National Centre for Earth Observation grant NE/R016518/1 (NCEO LTS-S).

We thank Tobias Wehr, Michael Eisinger and Anthony Illingworth for valuable discussions and their support for this work over many years.

References

655 S. J. Abel and I. A. Boutle. An improved representation of the raindrop size distribution for single-moment microphysics schemes. *Quarterly Journal of the Royal Meteorological Society*, 138(669):2151–2162, 10 2012. ISSN 00359009. https://doi.org/10.1002/qj.1949. URL http://onlinelibrary.wiley.com/doi/10.1002/qj.1949/fullhttp://doi.wiley.com/10.1002/qj.1949.

R. T. Austin, A. J. Heymsfield, and G. L. Stephens. Retrieval of ice cloud microphysical parameters using the CloudSat millimeter-wave radar and temperature. *Journal of Geophysical Research: Atmospheres*, 114(D8), 4 2009. ISSN 2156-2202.
660 https://doi.org/10.1029/2008JD010049. URL https://onlinelibrary.wiley.com/doi/full/10.1029/2008JD010049https://onlinelibrary.wiley.com/doi/abs/10.1029/2008JD010049https://agupubs.onlinelibrary.wiley.com/doi/10.1029/2008JD010049.

A. J. Baran. On the scattering and absorption properties of cirrus cloud. *Journal of Quantitative Spectroscopy and Radiative Transfer*, 89 (1-4):17–36, 11 2004. ISSN 00224073. https://doi.org/10.1016/j.jqsrt.2004.05.008.

H. Barker, Z. Qu, J. Cole, and OTHERS. Acm-rt paper tbd. *Atmospheric Measurement Techniques*, to be submitted, 2022a.

665 H. Barker, Z. Qu, J. Cole, and OTHERS. Acmb-df paper tbd. *Atmospheric Measurement Techniques*, to be submitted, 2022b.

A. Battaglia and J. Delanoë. Synergies and complementarities of CloudSat-CALIPSO snow observations. *Journal of Geophysical Research: Atmospheres*, 118(2):721–731, 1 2013. ISSN 2169897X. https://doi.org/10.1029/2012JD018092. URL http://doi.wiley.com/10.1029/2012JD018092.

A. Battaglia and G. Panegrossi. What can we learn from the CloudSat radiometric mode observations of snowfall over the ice-free ocean?
670 *Remote Sensing*, 12(20), 2020. https://doi.org/https://doi.org/10.3390/rs12203285. URL https://www.mdpi.com/851406.

K. V. Beard. Terminal Velocity and Shape of Cloud and Precipitation Drops Aloft. *Journal of the Atmospheric Sciences*, 33(5):851–864, 5 1976. ISSN 0022-4928. https://doi.org/10.1175/1520-0469(1976)033<0851:TVASOC>2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(1976)033%3C0851:TVASOC%3E2.0.CO;2http://journals.ametsoc.org/doi/abs/10.1175/1520-0469%281976%29033%3C0851%3ATVASOC%3E2.0.CO%3B2.

675 P. R. A. Brown and P. N. Francis. Improved Measurements of the Ice Water Content in Cirrus Using a Total-Water Probe. *Journal of Atmospheric and Oceanic Technology*, 12(2):410–414, 4 1995. ISSN 0739-0572.

F. L. Chang and Z. Li. A Near-Global Climatology of Single-Layer and Overlapped Clouds and Their Optical Properties Retrieved from Terra/MODIS Data Using a New Algorithm. *Journal of Climate*, 18(22):4752–4771, 11 2005. ISSN 0894-8755. https://doi.org/10.1175/JCLI3553.1. URL https://journals.ametsoc.org/view/journals/clim/18/22/jcli3553.1.xml.

680 J. Delanoë and R. J. Hogan. Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. *Journal of Geophysical Research: Atmospheres*, 115(4):D00H29, 7 2010. https://doi.org/10.1029/2009JD012346. URL http://www.met.rdg.ac.uk/~clouds/publications/delanoe_hogan_atrain.pdfhttp://www.met.rdg.ac.uk/~clouds/publications/delanoe_hogan_atrain.pdfhttp://www.met.rdg.ac.uk/~clouds/publications/delanoe_hogan_atrain.pdf.

J. Delanoë, A. Protat, J. Testud, D. Bouniol, A. J. Heymsfield, A. Bansemer, P. R. A. Brown, and R. M. Forbes. Statisti-
685 cal properties of the normalized ice particle size distribution. *Journal of Geophysical Research*, 110(D10):D10201, 5 2005. https://doi.org/10.1029/2004JD005405. URL http://doi.wiley.com/10.1029/2004JD005405.

J. M. E. Delanoë and R. J. Hogan. A variational scheme for retrieving ice cloud properties from combined radar, lidar, and infrared radiometer. *Journal of Geophysical Research*, 113(D7):D07204, 4 2008. ISSN 0148-0227. https://doi.org/10.1029/2007JD009000. URL http://onlinelibrary.wiley.com/doi/10.1029/2007JD009000/fullhttp://doi.wiley.com/10.1029/2007JD009000.

690   D. Donovan, P. Kollias, and G.-J. van Zadelhoff. Generating satellite test datasets - 2. the earthcare end-to-end simulator (e3sim). *Atmospheric Measurement Techniques*, to be submitted, 2022.

M. Eisinger, T. Wehr, T. Kubota, D. Bernaerts, and K. Wallace. The earthcare production model and auxiliary products. *Atmospheric Measurement Techniques*, to be submitted, 2022.

J. Escribano, A. Bozzo, P. Dubuisson, J. Flemming, R. J. Hogan, L. C.-Labonnote, and O. Boucher. A benchmark for testing the accuracy and
695   computational cost of shortwave top-of-atmosphere reflectance calculations in clear-sky aerosol-laden atmospheres. *Geoscientific Model Development*, 12(2):805–827, 2 2019. ISSN 1991-9603. https://doi.org/10.5194/gmd-12-805-2019. URL https://www.geosci-model-dev. net/12/805/2019/.

D. R. Feldman, W. D. Collins, R. Pincus, X. Huang, and X. Chen. Far-infrared surface emissivity and climate. *Proceedings of the National Academy of Sciences*, 111(46):16297–16302, 11 2014. ISSN 0027-8424. https://doi.org/10.1073/PNAS.1413640111. URL https://www.
700   pnas.org/content/111/46/16297https://www.pnas.org/content/111/46/16297.abstract.

P. R. Field, R. J. Hogan, P. R. A. Brown, A. J. Illingworth, T. W. Choularton, and R. J. Cotton. Parametrization of ice-particle size distributions for mid-latitude stratiform cloud. *Quarterly Journal of the Royal Meteorological Society*, 131(609):1997–2017, 7 2005. ISSN 00359009. https://doi.org/10.1256/qj.04.134. URL http://doi.wiley.com/10.1256/qj.04.134.

P. N. Francis, P. Hignett, and A. Macke. The retrieval of cirrus cloud properties from aircraft multi-spectral reflectance measure-
705   ments during EUCREX'93. *Quarterly Journal of the Royal Meteorological Society*, 124(548):1273–1291, 4 1998. ISSN 00359009. https://doi.org/10.1002/qj.49712454812. URL http://doi.wiley.com/10.1002/qj.49712454812.

S. H. Ham, S. Kato, F. G. Rose, D. Winker, T. L'Ecuyer, G. G. Mace, D. Painemal, S. Sun-Mack, Y. Chen, and W. F. Miller. Cloud occurrences and cloud radiative effects (CREs) from CERES-CALIPSO-CloudSat-MODIS (CCCM) and Cloud-Sat radar-lidar (RL) products. *Journal of Geophysical Research: Atmospheres*, 122(16):8852–8884, 8 2017. ISSN 21698996.
710   https://doi.org/10.1002/2017JD026725. URL https://agupubs.onlinelibrary.wiley.com/doi/full/10.1002/2017JD026725https://agupubs. onlinelibrary.wiley.com/doi/abs/10.1002/2017JD026725https://agupubs.onlinelibrary.wiley.com/doi/10.1002/2017JD026725.

J. M. Haynes, T. S. L'Ecuyer, G. L. Stephens, S. D. Miller, C. Mitrescu, N. B. Wood, and S. Tanelli. Rainfall retrieval over the ocean with spaceborne W-band radar. *Journal of Geophysical Research*, 114:D00A22, 2 2009. ISSN 0148-0227. https://doi.org/10.1029/2008JD009973. URL http://onlinelibrary.wiley.com/doi/10.1029/2008JD009973/fullhttp://doi.wiley.com/
715   10.1029/2008JD009973http://onlinelibrary.wiley.com/doi/10.1029/2008JD009973/pdf.

D. S. Henderson, T. L'Ecuyer, G. Stephens, P. Partain, and M. Sekiguchi. A Multisensor Perspective on the Radiative Impacts of Clouds and Aerosols. *Journal of Applied Meteorology and Climatology*, 52(4):853–871, 4 2013. ISSN 1558-8424. https://doi.org/10.1175/JAMC-D-12-025.1. URL http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-12-025.1.

A. J. Heymsfield and C. D. Westbrook. Advances in the Estimation of Ice Particle Fall Speeds Using Laboratory and Field Measurements.
720   *Journal of the Atmospheric Sciences*, 67(8):2469–2482, 8 2010. https://doi.org/10.1175/2010JAS3379.1. URL http://journals.ametsoc. org/doi/abs/10.1175/2010JAS3379.1.

P. G. Hill, R. P. Allan, J. C. Chiu, A. Bodas-Salcedo, and P. Knippertz. Quantifying the Contribution of Different Cloud Types to the Radiation Budget in Southern West Africa. *Journal of Climate*, 31(13):5273–5291, 7 2018. https://doi.org/10.1175/JCLI-D-17-0586.1. URL http://journals.ametsoc.org/doi/10.1175/JCLI-D-17-0586.1.

725   R. J. Hogan. A Variational Scheme for Retrieving Rainfall Rate and Hail Reflectivity Fraction from Polarization Radar. *Journal of Applied Meteorology and Climatology*, 46(10):1544–1564, 10 2007. https://doi.org/10.1175/JAM2550.1. URL http://journals.ametsoc.org/doi/ abs/10.1175/JAM2550.1http://journals.ametsoc.org/doi/abs/10.1175/JAM2550.1.

R. J. Hogan. Fast Lidar and Radar Multiple-Scattering Models. Part I: Small-Angle Scattering Using the Photon Variance–Covariance Method. *Journal of the Atmospheric Sciences*, 65(12):3621–3635, 12 2008. ISSN 0022-4928. https://doi.org/10.1175/2008JAS2642.1. URL http://journals.ametsoc.org/doi/abs/10.1175/2008JAS2642.1.

R. J. Hogan. Fast Reverse-Mode Automatic Differentiation using Expression Templates in C++. *ACM Transactions on Mathematical Software*, 40(4):1–16, 7 2014. ISSN 00983500. https://doi.org/10.1145/2560359. URL http://dl.acm.org/citation.cfm?id=2560359http://dl.acm.org/citation.cfm?doid=2639949.2560359.

R. J. Hogan. Adept 2.0: a combined automatic differentiation and array library for C++, 10 2017. URL https://zenodo.org/record/1004730.

R. J. Hogan and A. Battaglia. Fast Lidar and Radar Multiple-Scattering Models. Part II: Wide-Angle Scattering Using the Time-Dependent Two-Stream Approximation. *Journal of the Atmospheric Sciences*, 65(12):3636–3651, 12 2008. ISSN 0022-4928. https://doi.org/10.1175/2008JAS2643.1. URL http://journals.ametsoc.org/doi/abs/10.1175/2008JAS2643.1.

R. J. Hogan, L. Tian, P. R. A. Brown, C. D. Westbrook, A. J. Heymsfield, and J. D. Eastment. Radar Scattering from Ice Aggregates Using the Horizontally Aligned Oblate Spheroid Approximation. *Journal of Applied Meteorology and Climatology*, 51(3):655–671, 3 2012. ISSN 1558-8424. https://doi.org/10.1175/JAMC-D-11-074.1. URL http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-11-074.1.

R. J. Hogan, R. Honeyager, J. Tyynelä, and S. Kneifel. Calculating the millimetre-wave scattering phase function of snowflakes using the self-similar Rayleigh-Gans Approximation. *Quarterly Journal of the Royal Meteorological Society*, 143(703):834–844, 1 2017. ISSN 00359009. https://doi.org/10.1002/qj.2968. URL http://doi.wiley.com/10.1002/qj.2968.

A. J. Illingworth and T. M. Blackman. The Need to Represent Raindrop Size Spectra as Normalized Gamma Distributions for the Interpretation of Polarization Radar Observations. *Journal of Applied Meteorology*, 41(3):286–297, 3 2002. ISSN 0894-8763. https://doi.org/10.1175/1520-0450(2002)041<0286:TNTRRS>2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0450(2002)041<0286:TNTRRS>2.0.CO;2.

A. Irbah, J. Delanoe, G.-J. van Zadelhoff, D. Donovan, P. Kollias, A. Tatarevic, B. P. Tresseras, S. Mason, and R. Hogan. The classification of atmospheric hydrometeors and aerosols from earthcare's radar-lidar synergy: the ac-tc product. *Atmospheric Measurement Techniques*, to be submitted, 2022.

S. Kato, S. Sun-Mack, W. F. Miller, F. G. Rose, Y. Chen, P. Minnis, and B. A. Wielicki. Relationships among cloud occurrence frequency, overlap, and effective thickness derived from CALIPSO and CloudSat merged cloud vertical profiles. *Journal of Geophysical Research*, 115(D4):D00H28, 7 2010. ISSN 0148-0227. https://doi.org/10.1029/2009JD012277. URL http://doi.wiley.com/10.1029/2009JD012277.

S. Kato, F. G. Rose, S. Sun-Mack, W. F. Miller, Y. Chen, D. A. Rutan, G. L. Stephens, N. G. Loeb, P. Minnis, B. A. Wielicki, D. M. Winker, T. P. Charlock, P. W. Stackhouse, K.-M. Xu, and W. D. Collins. Improvements of top-of-atmosphere and surface irradiance computations with CALIPSO-, CloudSat-, and MODIS-derived cloud and aerosol properties. *Journal of Geophysical Research*, 116(D19):D19209, 10 2011. ISSN 0148-0227. https://doi.org/10.1029/2011JD016050. URL http://onlinelibrary.wiley.com/doi/10.1029/2011JD016050/abstracthttp://onlinelibrary.wiley.com/doi/10.1029/2011JD016050/fullhttp://doi.wiley.com/10.1029/2011JD016050.

S. Khanal and Z. Wang. Uncertainties in MODIS-Based Cloud Liquid Water Path Retrievals at High Latitudes Due to Mixed-Phase Clouds and Cloud Top Height Inhomogeneity. *Journal of Geophysical Research: Atmospheres*, 123(19):154–11, 10 2018. ISSN 2169897X. https://doi.org/10.1029/2018JD028558. URL http://doi.wiley.com/10.1029/2018JD028558.

P. Kollias, B. P. Tresseras, A. Battaglia, and A. Tatarevic. The earthcare cloud profiling radar processor: C-cpr. *Atmospheric Measurement Techniques*, to be submitted, 2022.

M. D. Lebsock, T. S. L'Ecuyer, and G. L. Stephens. Detecting the Ratio of Rain and Cloud Water in Low-Latitude Shallow Marine Clouds.
765   *Journal of Applied Meteorology and Climatology*, 50(2):419–432, 2 2011. ISSN 1558-8424. https://doi.org/10.1175/2010JAMC2494.1.
URL http://journals.ametsoc.org/doi/abs/10.1175/2010JAMC2494.1.

T. S. L'Ecuyer and G. L. Stephens. An Estimation-Based Precipitation Retrieval Algorithm for Attenuating Radars. *Journal of Applied Meteorology*, 41(3):272–285, 3 2002. ISSN 0894-8763. https://doi.org/10.1175/1520-0450(2002)041<0272:AEBPRA>2.0.CO;2. URL   http://journals.ametsoc.org/doi/abs/10.1175/1520-0450(2002)041<0272:AEBPRA>2.0.CO;2http://journals.ametsoc.org/doi/abs/
770   10.1175/1520-0450%282002%29041%3C0272%3AAEBPRA%3E2.0.CO%3B2.

J. Leinonen, M. D. Lebsock, G. L. Stephens, K. Suzuki, J. Leinonen, M. D. Lebsock, G. L. Stephens, and K. Suzuki. Improved Retrieval of Cloud Liquid Water from CloudSat and MODIS. *Journal of Applied Meteorology and Climatology*, 55(8):1831–1844, 8 2016. ISSN 1558-8424. https://doi.org/10.1175/JAMC-D-16-0077.1. URL http://journals.ametsoc.org/doi/10.1175/JAMC-D-16-0077.1.

H. Li and D. Moisseev. Melting Layer Attenuation at Ka-and W-Bands as Derived From Multifrequency Radar Doppler Spectra Observations.
775   *Wiley Online Library*, 124(16):9520–9533, 8 2019. https://doi.org/10.1029/2019JD030316. URL https://doi.org/.

D. C. Liu and J. Nocedal. On the limited memory BFGS method for large scale optimization. *Mathematical Programming*, 45(1-3):503–528, 8 1989. ISSN 0025-5610. https://doi.org/10.1007/BF01589116. URL http://link.springer.com/article/10.1007/BF01589116http://link.springer.com/10.1007/BF01589116.

D. W. Marquardt. An Algorithm for Least-Squares Estimation of Nonlinear Parameters. *Journal of the Society for Industrial and Applied*
780   *Mathematics*, 11(2):431–441, 6 1963. ISSN 0368-4245. https://doi.org/10.1137/0111030. URL http://epubs.siam.org/doi/pdf/10.1137/0111030http://epubs.siam.org/doi/abs/10.1137/0111030.

J. S. J. Marshall and W. M. K. Palmer. The distribution of raindrops with size. *Journal of Meteorology*, 5(4):165–166, 8 1948. ISSN 0095-9634. https://doi.org/10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2.

S. Mason, R. Hogan, C. Westbrook, S. Kneifel, D. Moisseev, and L. Von Terzi. The importance of particle size distribution and internal
785   structure for triple-frequency radar retrievals of the morphology of snow. *Atmospheric Measurement Techniques*, 12(9), 2019. ISSN 18678548. https://doi.org/10.5194/amt-12-4993-2019.

S. L. Mason, J. C. Chiu, R. J. Hogan, and L. Tian. Improved rain rate and drop size retrievals from airborne Doppler radar. *Atmospheric Chemistry and Physics*, 17(18):11567–11589, 9 2017. ISSN 1680-7324. https://doi.org/10.5194/acp-17-11567-2017. URL https://www. atmos-chem-phys.net/17/11567/2017/https://doi.org/10.5194/acp-17-11567-2017.

790 S. L. Mason, C. J. Chiu, R. J. Hogan, D. Moisseev, and S. Kneifel. Retrievals of Riming and Snow Density From Vertically Pointing Doppler Radars. *Journal of Geophysical Research: Atmospheres*, 123(24):807–13, 12 2018. ISSN 2169-897X. https://doi.org/10.1029/2018JD028603. URL http://doi.wiley.com/10.1029/2018JD028603https://onlinelibrary.wiley.com/doi/abs/10. 1029/2018JD028603.

S. Matrosov. Assessment of Radar Signal Attenuation Caused by the Melting Hydrometeor Layer. *IEEE Transactions on Geoscience and*
795   *Remote Sensing*, 46(4):1039–1047, 4 2008. ISSN 0196-2892. https://doi.org/10.1109/TGRS.2008.915757. URL http://ieeexplore.ieee. org/xpls/abs_all.jsp?arnumber=4459186http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4459186.

S. Matrosov, A. Battaglia, and P. Rodriguez. Effects of multiple scattering on attenuation-based retrievals of stratiform rainfall from CloudSat. *Journal of Atmospheric and Oceanic Technology*, pages 2199–2208, 2008. https://doi.org/https://doi.org/10.1175/2008JTECHA1095.1. URL https://journals.ametsoc.org/view/journals/atot/25/12/2008jtecha1095_1.xml.

800 S. Y. Matrosov. Potential for attenuation-based estimations of rainfall rate from CloudSat. *Geophysical Research Letters*, 34(5):L05817, 3 2007. ISSN 00948276. https://doi.org/10.1029/2006GL029161. URL http://onlinelibrary.wiley.com/doi/10.1029/2006GL029161/fullhttp://doi.wiley.com/10.1029/2006GL029161.

N. L. Miles, J. Verlinde, and E. E. Clothiaux. Cloud Droplet Size Distributions in Low-Level Stratiform Clouds. *Journal of the Atmospheric Sciences*, 57(2):295–311, 1 2000. ISSN 0022-4928. https://doi.org/10.1175/1520-0469(2000)057<0295:CDSDIL>2.0.CO;2.
805 URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0469(2000)057%3C0295:CDSDIL%3E2.0.CO%3B2http://journals.ametsoc.org/doi/abs/10.1175/1520-0469%282000%29057%3C0295%3ACDSDIL%3E2.0.CO%3B2.

K. Mróz, A. Battaglia, S. Kneifel, L. von Terzi, M. Karrer, and D. Ori. Linking rain into ice microphysics across the melting layer in stratiform rain: a closure study. *Atmospheric Measurement Techniques*, 14(1):511–529, 2021. URL https://amt.copernicus.org/articles/14/511/2021/.

C. M. Naud, A. D. Del Genio, and M. Bauer. Observational constraints on the cloud thermodynamic phase in midlatitude storms. *Journal*
810 *of Climate*, 19(20):5273–5288, 2006. ISSN 08948755. https://doi.org/10.1175/JCLI3919.1. URL http://journals.ametsoc.org/doi/abs/10.1175/JCLI3919.1.

D. Painemal and P. Zuidema. Assessment of MODIS cloud effective radius and optical thickness retrievals over the Southeast Pacific with VOCALS-REx in situ measurements. *Journal of Geophysical Research Atmospheres*, 116(24), 2011. ISSN 01480227. https://doi.org/10.1029/2011JD016155.

815 N. L. Pounder, R. J. Hogan, T. Várnai, A. Battaglia, and R. F. Cahalan. A Variational Method to Retrieve the Extinction Profile in Liquid Clouds Using Multiple-Field-of-View Lidar. *Journal of Applied Meteorology and Climatology*, 51(2):350–365, 2 2012. ISSN 1558-8424. https://doi.org/10.1175/JAMC-D-10-05007.1. URL http://journals.ametsoc.org/doi/abs/10.1175/JAMC-D-10-05007.1.

Z. Qu. Generating earthcare test datasets - 1. high resolution simulations using the global environmental multiscale model. *Atmospheric Measurement Techniques*, to be submitted, 2022.

820 C. D. Rodgers. *Inverse methods for atmospheric sounding: theory and practice*. World Scientific, Singapore, 2000. ISBN 981-02-2740-X.

V. Salomonson, W. Barnes, J. Xiong, S. Kempler, and E. Masuoka. An overview of the Earth Observing System MODIS instrument and associated data systems performance. In *IEEE International Geoscience and Remote Sensing Symposium*, volume 2, pages 1174–1176. IEEE, 2002. ISBN 0-7803-7536-X. https://doi.org/10.1109/IGARSS.2002.1025812. URL http://ieeexplore.ieee.org/document/1025812/.

G. Stephens, D. Winker, J. Pelon, C. Trepte, D. Vane, C. Yuhas, T. L'Ecuyer, and M. Lebsock. CloudSat and CALIPSO within the A-Train:
825 Ten Years of Actively Observing the Earth System. *Bulletin of the American Meteorological Society*, 99(3):569–581, 3 2018. ISSN 0003-0007. https://doi.org/10.1175/BAMS-D-16-0324.1. URL http://journals.ametsoc.org/doi/10.1175/BAMS-D-16-0324.1.

G. L. Stephens, D. G. Vane, R. J. Boain, G. G. Mace, K. Sassen, Z. Wang, A. J. Illingworth, E. J. O'Connor, W. B. Rossow, S. L. Durden, S. D. Miller, R. T. Austin, A. Benedetti, C. Mitrescu, T. CloudSat Science Team, and others. THE CLOUDSAT MISSION AND THE A-TRAIN. *Bulletin of the American Meteorological Society*, 83(12):1771–1790, 12 2002. ISSN 0003-0007. https://doi.org/10.1175/BAMS-83-12-
830 1771. URL http://journals.ametsoc.org/doi/10.1175/BAMS-83-12-1771http://journals.ametsoc.org/doi/abs/10.1175/BAMS-83-12-1771.

J. Testud, S. Oury, R. A. Black, P. Amayenc, and X. Dou. The Concept of "Normalized" Distribution to Describe Raindrop Spectra: A Tool for Cloud Physics and Cloud Remote Sensing. *Journal of Applied Meteorology*, 40(6):1118–1140, 6 2001. ISSN 0894-8763. https://doi.org/10.1175/1520-0450(2001)040<1118:TCONDT>2.0.CO;2. URL http://journals.ametsoc.org/doi/abs/10.1175/1520-0450(2001)040<1118:TCONDT>2.0.CO;2.

835 O. B. Toon, C. P. McKay, T. P. Ackerman, and K. Santhanam. Rapid calculation of radiative heating rates and photodissociation rates in inhomogeneous multiple scattering atmospheres. *Journal of Geophysical Research*, 94(D13), 1989. ISSN 01480227. https://doi.org/10.1029/jd094id13p16287.

S. Twomey. Introduction to the Mathematics of Inversion in Remote Sensing and Indirect. *Elsevier Scientific Publishing*, 1977. URL https://scholar.google.co.uk/scholar?hl=en&q=twomey+1977+Introduction+to+the+Mathematics+of+Inversion+in+Remote+Sensing&btnG=&as_sdt=1%2C5&as_sdtp=.

U. Wandinger, H. Baars, R. Engelmann, A. Hünerbein, S. Horn, T. Kanitz, D. Donovan, G.-J. Van Zadelhoff, D. Daou, J. Fischer, J. Von Bismarck, F. Filipitsch, N. Docter, M. Eisinger, D. Lajas, and T. Wehr. HETEAC: THE AEROSOL CLASSIFICATION MODEL FOR EARTHCARE. *EPJ Web of Conferences*, 119:01004, 2016. https://doi.org/10.1051/01004. URL https://www.epj-conferences.org/articles/epjconf/abs/2016/14/epjconf_ilrc2016_01004/epjconf_ilrc2016_01004.html.

U. Wandinger, A. Floutsi, H. Baars, M. Haarig, A. Hünerbein, N. Docter, D. Donovan, and G.-J. van Zadelhoff. Hybrid end-to-end aerosol classification model for earthcare. *Atmospheric Measurement Techniques*, to be submitted, 2022.

T. Wehr, T. Kubota, R. Oki, D. Bernaerts, K. Wallace, R. Koopman, M. Eisinger, and P. Deghaye. The earthcare satellite and science data products. *Atmospheric Measurement Techniques*, to be submitted, 2022.

B. A. Wielicki, B. R. Barkstrom, B. A. Baum, T. P. Charlock, R. N. Green, D. P. Kratz, B. Lee Robert, P. Minnis, G. Louis Smith, T. Wong, D. F. Young, R. D. Cess, J. A. Coakley, D. A. Crommelynck, L. Donner, R. Kandel, M. D. King, A. J. Miller, V. Ramanathan, D. A. Randall, L. L. Stowe, and R. M. Welch. Clouds and the earth's radiant energy system (CERES): Algorithm overview. *IEEE Transactions on Geoscience and Remote Sensing*, 36(4):1127–1141, 1998. ISSN 01962892. https://doi.org/10.1109/36.701020.

D. M. Winker, J. R. Pelon, and M. P. McCormick. The CALIPSO mission: spaceborne lidar for observation of aerosols and clouds. volume 4893, page 1. International Society for Optics and Photonics, 3 2003. https://doi.org/10.1117/12.466539. URL http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.466539.

Z. Zhang, S. P. J. o. G. Research, and u. 2011. An assessment of differences between cloud effective particle radius retrievals for marine water clouds from three MODIS spectral bands. *Wiley Online Library*, 116(20):20215, 2011. https://doi.org/10.1029/2011JD016216. URL https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/2011JD016216.