Summary
17/07/2023 16:44:38

Differences exist between documents.

**New Document:**
Donovan_et_al_L1
59 pages (18.83 MB)
17/07/2023 16:43:53
Used to display results.

**Old Document:**
egusphere-2023-384
54 pages (32.40 MB)
17/07/2023 16:43:52

No pages were deleted

**How to read this report**

**Highlight** indicates a change.
~~Deleted~~ indicates deleted content.
▲ indicates pages were changed.
⟷ indicates pages were moved.

# The Generation of EarthCARE L1 Test Data sets Using Atmospheric Model Data Sets

David P. Donovan[1], Pavlos Kollias[2,3], Almudena Velázquez Blázquez[4], and Gerd-Jan van Zadelhoff[1]

[1]Royal Netherlands Meteorological Institute (KNMI), de Bilt, the Netherlands
[2]Department of Atmospheric and Oceanic Sciences, McGill University, Quebec, Canada
[3]Division of Atmospheric Science, Stony Brook University, Stony Brook, NY, USA
[4]Royal Meteorological Institute of Belgium (RIMB), Brussels, Belgium

**Correspondence:** D.P. Donovan (Donovan@knmi.nl)

**Abstract.** The Earth Clouds and Radiation Explorer mission (EarthCARE) is a multi-instrument cloud-aerosol-radiation process study oriented mission embarking a high spectral resolution lidar, a cloud profiling radar, a multi-spectral imager and a three-view broad-band radiometer. An important aspect of the EarthCARE mission is its focus on instrument synergy. Many L2 products are the result of L1 inputs from one or more instruments. Since no existing complete observational proxy-data

5    sets comprised of co-located and co-temporal "EarthCARE-like" data exists, it has been necessary to create synthetic data sets for the testing and development of various retrieval algorithms as well as the data processing chain. Given the synergistic nature of the processing chain, it is important that the test data are physically consistent across the various instruments. Within the EarthCARE project, a version of the EarthCARE simulator multi-instrument framework (ECSIM) has been used to create unified realistic test data frames. These simulations have been driven using high resolution atmospheric model data (described

10    in a companion paper). In this paper, the methods used to create the test data scenes are described. As well, the simulated L1 data corresponding to each scene is presented and discussed.

## 1   Introduction

The Earth Clouds, Aerosol and Radiation Explorer (EarthCARE; Illingworth et al., 2015) platform comprises a 94 GHz Doppler cloud profiling radar (CPR), 355 nm high spectral resolution atmospheric lidar (ATLID), a multispectral imager (MSI), and

15    a three–view broadband radiometer (BBR). The mission is complex in several aspects. In particular, the synergistic nature of many of the data products and the end-to-end nature of the processing chain (Eisinger et al., 2023; Barker et al., 2023) gives rise to the requirement to generate test data which is realistic and physically consistent from the point-of-view of all the instruments. For example, combined ATLID, CPR and MSI measurements will be used within the same synergistic algorithm in order to retrieve e.g. cloud particle size information. Thus, it is important, from the point of view of the various instruments,

20    that the simulations be conducted in a physically consistent and accurate manner. Crude instrument specific parameterizations (e.g. empirical radar reflectivity -vs- cloud ice water content ($Z_e$ -vs- $IWC$) relationships) should be avoided since they are not explicitly related to the basic physical atmospheric properties in a way that can be connected to the other instruments.

The requirement to generate physically consistent simulations with all of the virtual instruments "seeing" the same atmosphere in terms of both macro-and micro-physical considerations was a prime motivation in the original development of the EarthCARE simulator (ECSIM)[1] framework and its component models. ECSIM has been applied to a set of high resolution cloud resolving atmospheric model data in order to generate realistic L1 test data sets suitable for EarthCARE algorithm development an testing. Here L1 refers to calibrated and instrument corrected output generated by the instruments (e.g. lidar attenuated backscatter, radar reflectivity). L2 data refers to geophysical quantities retrieved using L1 data (e.g. optical extinction, target classification, water content).

The focus of this paper is on the general description of the L1 data sets and how they were generated rather than e.g. specific algorithm and instrument sensitivity study results. This paper thus serves as a central reference for the several other instrument/algorithm specific papers belonging to the EarthCARE Level 2 algorithms and data products special issue which rely on the ECSIM generated L1 test data sets. This paper first presents an overview of the ECSIM. Subsequently, the specific radiative transfer and instrument simulation methods used for each of the instruments is presented. The paper then concludes with a presentation of the simulated L1 data of the three main EarthCARE testing scenes. The input data used to build these testing scenes are described in a companion paper (Qu et al., 2022).

## 2 ECSIM overview

ECSIM is not a single model, but rather is a collection of tools, including radiative transfer models and instrument models, that cooperate to produce physically consistent simulations covering a range of diverse instruments.

Unlike many common radiative transfer models, ECSIM is, in principle, not tied to any particular size distribution representation. In particular, multi-wavelength, particle size resolved optical and physical characteristics (e.g. mass, maximum dimension, extinction coefficient, and phase-function) are stored in a database while the corresponding bin-resolved size distribution information is specified in separate size-distribution files. A separate master scene file stores the 3-D fields of temperature, pressure, velocity and the concentrations of relevant atmospheric gasses. This structure allows for the efficient generation of the optical properties necessary to drive diverse forward models. This process is schematically depicted in Fig.1. In the particular application described in this paper, namely the creation of test frames using Global Environmental Multi-scale (GEM) input fields as a base (Qu et al., 2022) and supplemented by aerosol data from Copernicus Atmosphere Monitoring Service (CAMS) reanalysis fields (Inness et al., 2019). The GEM and CAMS data were translated into ECSIM scene files using purpose built computer programs.

A key aspects of the ECSIM structure is that all the instrument simulations are driven by the same unambiguously defined atmosphere. All of the individual radiative transfer model read the same scene files and access the same scattering information data files. Whenever a model needs to calculate, for example, cloud extinction at a required wavelength, the appropriate binned

---

[1] It should be noted that The ECSIM framework was used as a basis for the development of a related framework called E3SIM. Within the context of this special issue the term E3SIM may be used in place of E3SIM. In this paper, the framework is referred to as ECSIM and the focus is not on the framework but rather the models and methods applied to the GEM/CAMS model inputs (Qu et al., 2022) in order to generate simulated EarthCARE L1 data are described.
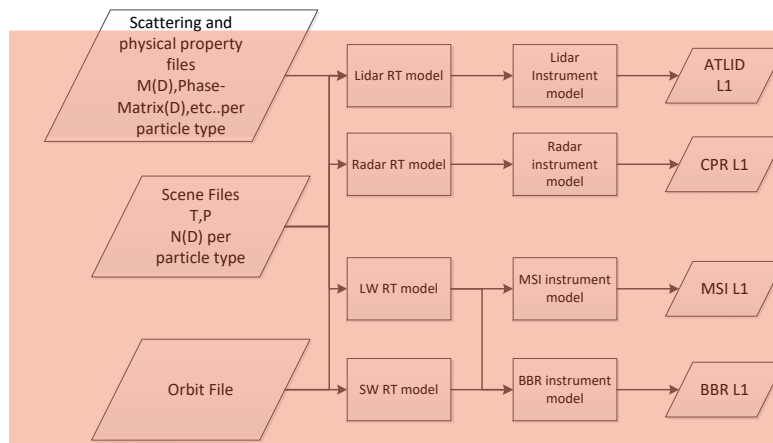
**Figure 1.** High level structure of ECSIM. All of the radiative transfer models read the same scene information. Another feature is that the radiative-transfer models are separate from the instrument simulation models.

size distribution (read from the scene files) is combined with the appropriate scattering properties (read from the scattering library). This structure facilitates the production of physically consistent simulations i.e. it helps insure that the radiative transfer models all "see the same atmospheric particles" and are "discouraged" from making their own ad-hoc assumptions regarding particle size distributions or optical/physical properties (e.g. different mass-vs-maximum dimension relationships for ice cloud particles).

## 2.1 Representation of the scene constituents

### 2.1.1 Hydrometeor microphysical properties

The Global Environmental Multi-scale (GEM) Numerical Weather Prediction (NWP) models runs for the EarthCARE algorithm development efforts (Qu et al., 2022) supports six different hydrometeor types: cloud droplets, raindrops, ice, snow, graupel and hail using a bin microphysical representation. The model bin warm microphysics scheme represents the cloud droplets particle size distribution in 50 bin sizes centered at values corresponding to cloud droplet diameters from 2 to 100 μm and the rain particle size distribution in 50 bin sizes centered at values corresponding to raindrop radius from 50 to 2500 μm. The Environment and Climate Change Canada (ECCC) cold microphysics scheme used by GEM supports four different types of particles: cloud ice, snow, graupel and hail. The particle size distribution of cloud ice is represented in 50 bin sizes centered at values corresponding to ice particle radius from 2 to 100 μm, every 2 μm. The particle size distributions of snow is represented by 100 bin sizes each from 25 to 2500 μm, every 10 μm. Table 1 summarizes the details of the bin representation

of the hydrometeor species in GEM. When translating the GEM cloud cloud information to ECSIM certain adjustments were
made to the cold cloud microphysics. As detailed in Qu et al. (2022) these adjustments improved the realism of the simulations.
Accordingly, the ECSIM ice and snow hydrometeors categories have particle mass (M) and cross-sectional area (Ac) according
to the Erfani and Mitchell (2016) relations for synoptic cirrus and temperatures between -40 and -20 C. The mass of the ice
and snow particles in grams is given by the expression $M(D) = \exp\left[a_0 + a_1 ln(D) + a_2 [ln(D)]^2\right]$ , where D is in cm and the
coefficients have values $a_0 = -6.72924$, $a_1 = 1.17421$ and $a_2 = -0.15980$.

| Hydrometeor type | $D_{min}$ (µm) | $D_{max}$ (µm) | Bin spacing (µm) | Number of bins | Density (gcm$^{-3}$) |
|---|---|---|---|---|---|
| Cloud | 2 | 100 | 2 | 50 | 1.0 |
| Rain | 100 | 5000 | 100 | 50 | 1.0 |
| Ice | 4 | 200 | 4 | 50 | (Erfani and Mitchell, 2016) |
| Snow | 50 | 5000 | 50 | 100 | (Erfani and Mitchell, 2016) |
| Graupel | 100 | 10000 | 100 | 100 | 0.4 |
| Hail | 100 | 10000 | 100 | 100 | 0.9 |

**Table 1.** GEM Cloud and precipitation microphysical parameters.

### 2.1.2 VIS-UV-IR hydrometer optical properties

In ECSIM, the optical scattering properties of cloud droplets and raindrops are calculated using Mie theory. The optical scattering and absorption properties of each bin are determined by averaging the scattering and absorption properties over 20 equally spaced sub-bins. Temperature dependent refractive indices were used (Hale and Querry, 1973) and calculations were done for temperatures of 240K and 300K and optical information at specific temperatures was found via interpolation.

For the GEM scenes, the UV-IR extinction, absorption, and phase functions for ice and snow were adapted from Baum et al. (2014), using the single-particle effective radius to interpolate in particle size. Here, the effective radius is defined as

$$R_{eff}(D) = \frac{3M(D)}{4\rho_i A_c(D)} \tag{1}$$

where $D$ is the particle maximum dimension, $M(D)$ is the particle mass, $A_c(D)$ is the cross-sectional area of the particle, and $\rho_i$ is the density of solid ice. The $M(D)$ and $Ac(D)$ relationships used are listed in Table 1 and are descibed in more detail in Qu et al. (2022).

For both cloud-ice and snow particle types, the Baum Aggregated Solid Columns properties were used. For cloud-ice, particle maximum dimensions between 4 and 200 microns were used in steps of 4 µm. For snow, sizes between 50 and 5000 µm in steps of 25 µm. The Aggregated solid-columns phase functions generally do not have strong halo features which is in accordance with observations (Baum et al., 2014). The aggregated solid-columns phase functions also lack a backscatter peak, however, this is likely not in accordance with observations. In particular, in Zhou and Yang (2015), the backscatter issue was investigated. They found that by directly solving Maxwell's equations numerically, a narrow backscattering peak

was present for a wide range of types of ice crystals (i.e. irregular smoothed, smoothed, roughened hexagons). The width of the backscatter peak is inversely proportional to the size parameter. They also showed that accounting for this peak produces more realistic values of the lidar-ratio and improves the agreement between lidar multiple-scattering coefficients derived using Calipso observations and theory (Zhou and Yang, 2015). According;y, for the lidar specific radiative transfer calculations (See Section 2.2) a backscatter-peak was added to the ice phase function following the procedure described in Zhou and Yang (2015).

For graupel and hail, for the UV-IR optical properties, Mie calculation were applied to equivalent effective radius spheres. This procedure produces reasonably accurate estimates of the extinction and absorption (Grenfell and Warren, 1999), however, the phase functions will not be accurate. This was not considered to be an important issue as, in the scenes considered here, occurrences of these scattering types are masked by significant amounts of cloud-ice and/or snow when viewed from above.

### 2.1.3 Microwave scattering and absorption

Mie scattering is used for the estimation of the backscattering and extinction coefficients for cloud droplets and raindrops. For ice and snow, the Self Similar Rayleigh-Gans Approximation (Hogan et al., 2017) is used for modelling of backscattering properties of ice and snow. The particles are assumed to be horizontally aligned with an aspect ratio of 0.6. The parameters for aggregates of bullet rosettes are as in (Hogan and Westbrook, 2014). Graupel and hail are modelled as the equal-mass spheres consisting of a blend of ice and air and using Mie theory. The dielectric properties of the mixture are calculated using Maxwell–Garnett approximation and assuming the ice inclusions in an air matrix. The gaseous attenuation is estimated using the Rosenkranz's method (Rosenkranz, 1998). The raindrops terminal velocity is estimated using (Brandes E. A. and J., 2002) and the ice and snow particles fall velocity is based on (Mitchell and Heymsfield, 2005).

### 2.1.4 Aerosol microphysical and optical properties

The aerosol fields were constructed using CAMS aerosol component fields mapped to the following "Hybrid End-To-End Aerosol Classification" (HETEAC) (Wandinger et al., 2022) base component types, more detail is given in Qu et al. (2022). We reiterate here that goal of this exercise is not to produce an "optimally accurate" mapping. An Ad-Hoc approximate approach in order to produce a range of "realistic enough" aerosol conditions with different optical characteristics was sufficient for the goals of the work presented here.

### 2.1.5 Optical Gaseous absorption

The ECSIM scenes contain the following (limited) number of gaseous species. $O_2$, $N_2$, $CO_2$, $H_2O$, $O_3$.

For the MSI and BBR simulations, the atmospheric transmission is calculated using a correlated-k method (Kato et al., 1999) for various narrow-bands. For each narrow-band, the aerosol and cloud optical properties are treated as being constant and represent a simple average across the band.

| | CAMS Field | | | | |
|---|---|---|---|---|---|
| Volume Fraction | DD1-DD3 | SS1-SS2 | SO4 | BCB | OMB |
| Coarse-Dust (Spheroids) | 100 | 9 | 5 | 3 | 3 |
| Coarse non-absorbing | 0 | 90 | -0 | 0 | 0 |
| Fine Weakly absorbing | 0 | 1 | 95 | 0 | 0 |
| Fine Strongly absorbing | 0 | 0 | 0 | 97 | 97 |

**Table 2.** Mapping between CAMS aerosol components and HETEAC aerosol component types.

### 2.1.6 Short-Wave Surface radiative properties

For the short-wave water surfaces, the azimuth dependent BRDF for the GEM derived scenes were calculated as a function of wind-speed based on look-up tables derived from the Matrix-Operator Model (MOMO) radiative transfer model (Hollstein and Fischer, 2012) (run in unpolarized mode). The MOMO ocean BRDFs (Fell and Fischer, 2001) themselves are based on a Cox-Munk approach.

For the non-snow/ice SW land surfaces, the Isometric–Volumetric–Geometric (Iso-Vol-Geo) Kernel-based approach is used (Schaaf et al., 2002). For the GEM scenes, Iso-Vol-Geo coefficients for non-ice/snow areas for a number of specific bands were supplied. The BRDFs at these bands then served as inputs to a Principle-Component based procedure for generating BRDFs for general SW frequencies (Vidot and Borbáas, 2014). For ice and snow surfaces, the surfaces were treated as being Lambertian with the albedo being assigned by underlying surface type (taken from the GEM model) following Moody et al. (2007).

### 2.1.7 Long-wave Surface radiative properties

For the Long-Wave radiative transfer calculations a much simpler approach was used. The surfaces were treated as being Lambertian and the emissivity was assumed to be a constant value of 0.93 over water and 0.98 over land.

### 2.2 Lidar Radiative Transfer Calculations

The lidar simulation is separated into two main steps. First a lidar specific radiative transfer model is used to calculate the time-spectral-polarization resolved lidar backscatter. Then a separate instrument model is applied to the output of the lidar radiative transfer model to simulate the effect of e.g. instrument noise and non-ideal spectral filtering. The lidar radiative transfer methods are described in this section and the lidar instrument model is then described in Section 2.3.

### 2.2.1 Basic Considerations

For a single wavelength non-polarized lidar, the single-scattered power incident upon the detection element for a completely elastic backscatter lidar (neglecting the background signal) for a an appropriate wavelength interval centered around $\lambda_l$ can be written as

$$P(z, \lambda_l) = C_{lid} z^{-2} \beta_\pi(z, \lambda_l) \exp\left[-2 \int_0^z \alpha(z', \lambda_l) dz'\right] \tag{2}$$

where $\lambda_l$ is the laser wavelength, $P(z, \lambda_l)$ is the power the lidar will receive from a target a distance $z = ct/2$ from the lidar (where $t$ is the elapsed time since the laser pulse was launched from the lidar), $\beta_\pi(z, \lambda_l)$ is the range and wavelength dependent backscatter coefficient and $\alpha(z', \lambda_l)$ is the extinction coefficient. Here both the backscatter and extinction coefficients represent appropriate averages over the wavelength interval of interest. $C_{lid}$ is a constant which takes into account factors related to the lidar's physical characteristics.

$$C_{lid} = T_r E_l A_o \tag{3}$$

where $T_r$ is the bulk effective receiver transmission, $E_l$ is the laser pulse energy and $A_o$ is the effective receiver area.

### 2.2.2 Multiple-scattering

Equation 2 is valid only when single-scattering is predominant. In cases where either a high proportion of scattered photons remain within the receiver Field of View (FOV) or the optical depth is such that photon mean-free-path is small compared to the instantaneous sampling volume then multiple scattering (MS) effects must be taken into account. In general, these conditions are often met for ground-based remote sensing of clouds and are always met for the case of space-based cloud remote sensing.

In order to account for multiple-scattering, the lidar radiative transfer simulations were conducted using the ECSIM-3D lidar Monte-Carlo radiative transfer code. The approach used is described within (Donovan et al., 2015) and additional relevant detail, not covered within Donovan et al. (2015) is given here. In particular, the approach used to determine the spectral properties of the return signal is discussed. Within the MS model, each photon packet is assigned both a mean relative spectral shift (initially zero) from its center wavelengths and a Gaussian spectral width. The Doppler shift of the packets is calculated according to the relative velocity of the scatterers encountered and the scattering geometry. The spectral width is initialised using the laser line-width and is hardly affected by particulate scattering but can be substantially widened if molecular scattering has occurred.

### 2.2.3 Doppler shift and Multiple Scattering

Even neglecting such processes involving the change of energy levels within the target scatterers (i.e. rotational–vibrational Raman scattering, fluorescence etc.) the scattered radiation will be Doppler shifted according to the relative motion of the scatterers with respect to the lidar. In general, a target moving with respect to a fixed lidar will 'see' a source whose frequency is given by (neglecting terms of order $(v/c)^2$)

$$f = f_o\left(1 + \frac{v}{c}\cos(\theta)\right) \tag{4}$$

where $v$ is the velocity of the scatterer, $c$ is the speed of light, $\theta$ is the angle between the scatterers velocity vector and the line-of-sight directed back toward the receiver, and $f_o$ is the frequency of the source light in the frame of the laser. When light

initially incident upon the moving scatterer, is scattered, it is equivalent to being instantaneously absorbed and re-emitted so that the scattered light will again be Doppler shifted in frequency depending on the motion of the scatterer with respect to the receiver. Accordingly, the frequency measured by an observer in the same frame as the lidar transmitter will be given by (again neglecting terms of order $(v/c)^2$)

$$f = f_o \left[ 1 +_- \frac{v}{c} (\cos(\theta) + \cos(\theta_{sc})) \right] \tag{5}$$

where $\theta_{sc}$ is the angle between the velocity vector and the scattered photon's trajectory. So that light scattered directly forward and measured by a stationary observer would have no Doppler shift while light scattered directly backwards to the lidar would be observed to have been shifted in wavelength by an amount twice that predicted by Eq. 4.
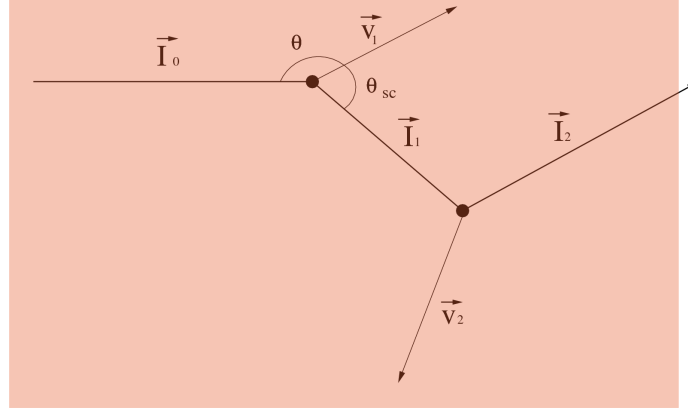
**Figure 2.** Geometry relevant for determining the observed Doppler shift after multiple scatters.

The occurrence of multiple scattering will have implications with respect to the spectral signature of the lidar return. By repeated application of Eq. 4 (see Fig. 2) it can be shown that the Doppler shifted frequency after $n$ scatters as measured by an observer in the same frame-of-reference as the lidar will be given by

$$f_n = f_o \left( 1 + \frac{1}{c} (\mathbf{v}_1 \cdot \mathbf{I}_0 + (\mathbf{v}_1 - \mathbf{v}_2) \cdot \mathbf{I}_1 + (\mathbf{v}_2 - \mathbf{v}_3) \cdot \mathbf{I}_2 + \cdots \mathbf{v}_n \cdot \mathbf{I}_n) \right) \tag{6}$$

where $\mathbf{v}_i$ is the velocity vector of the $i$th scatterer and $\mathbf{I}_i$ is the unit vector for the photon trajectory for the $i$th scatter. If the lidar is to observe the photon, to a good approximation, it must be scattered directly backwards towards the lidar receiver so that $\mathbf{I_n} = -\mathbf{I_o}$ and thus,

$$f_n = f_o \left( 1 +_- \frac{1}{c} (\mathbf{v}_1 \cdot \mathbf{I}_0 + (\mathbf{v}_1 - \mathbf{v}_2) \cdot \mathbf{I}_1 + (\mathbf{v}_2 - \mathbf{v}_3) \cdot \mathbf{I}_2 + \cdots - \mathbf{v}_n \cdot \mathbf{I}_o) \right) \tag{7}$$

### 2.2.4 Spectral Broadening and Multiple-scattering

In the MC model the spectral width of the packets is also tracked. The width is initialized using the laser spectral line width. The Gaussian width after subsequent scattering events is calculated by convolving the spectral width of the incoming packet with the spectral width induced by the scattering event (accounting the scattering geometry). For particulate scattering, the spectral width is assumed to be unaffected. However, for molecular scattering events thermal broadening is important. How this is modelled in ECSIM is described in the following subsection.
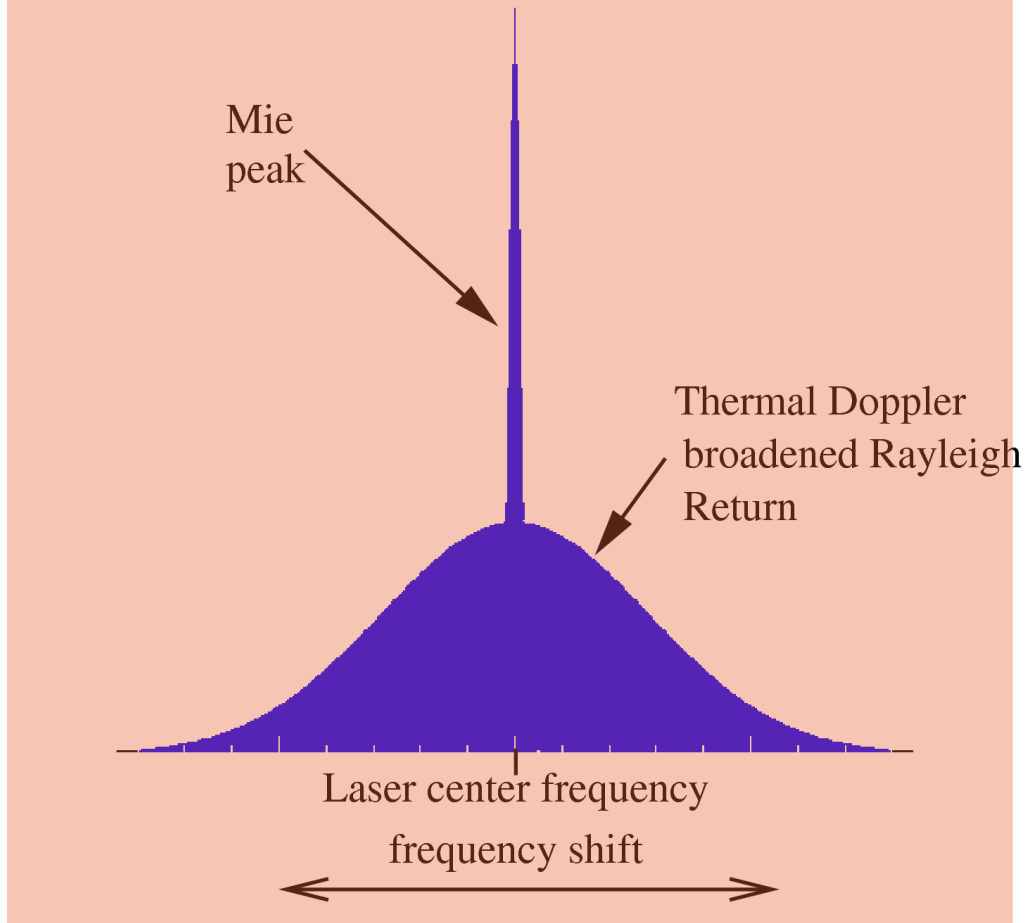
**Figure 3.** Idealized view of spectral signature of a lidar return.

A schematic representation of the spectral signature of a general lidar return signal is shown in Fig. 3. The spectral width of the particulate backscattering peak will be determined by the spectral width of the laser pulse itself along with any turbulence present in the sampling volume. The spectral width of the laser will be on the order of $10^{-7} f_o$, so that the laser line width will usually be the dominant factor. The molecular backscatter though, will be much spectrally broader than the particulate

 scattering return; this is due to the fact that the atmospheric molecules have a large thermal velocity. For low densities the Half-Maximum Half Width (HMHW) of the broadening produced by the thermal motions alone is given by

$$\Delta f = f_o \left( \frac{8kT \log(2)}{Mc^2} \right) \tag{8}$$

where, $k$ is Boltzmann's constant, $M$ is the average molecular mass, and $T$ is temperature. For typical values of atmospheric temperature the HMHW of the Doppler return will be on the order of $4 \times 10^{-6} f_o$ so that the Doppler broadening will be 10-20 greater than the laser line width. Eq. 8 applies at low gas densities.

In general, the central (non-Raman) Rayleigh line profile (Cabannes line) will consist of three components, a central peak together with two flanking "Brillouin-Mandel'shtam" peaks (Miles et al., 2001). In the low density or high temperature regimes the uncorrelated motion of the scatterers gives rise to a Gaussian velocity distribution centered around the mean velocity of the flow and Eq. 8 applies. As the pressure increases or the temperature decreases, density fluctuations on the order of the laser wavelength will appear. These density fluctuations travel at the speed of sound in the gas and will give rise to acoustic side bands.

The Rayleigh-Brillouin scattering line shape may be quantified in terms of the so-called $y$ parameter which is defined in terms of the ratio of the laser wavelength to the mean free-path. For the Earth's atmosphere

$$y = 0.230 \left[ \frac{T + 111}{T^2} \right] \left[ \frac{P\lambda}{\sin(\frac{\theta}{2})} \right] \tag{9}$$

where $T$ is the temperature in Kelvin, $P$ is the pressure in atmospheres, $\lambda$ is the laser wavelength in nm and $\theta$ is the scattering angle. Here $x$ is a normalized frequency parameter defined as

$$x = \frac{\sqrt{2}\lambda}{4 \sin\left(\frac{\theta}{2}\right)} \left( \frac{\nu}{\nu_o} \right) \tag{10}$$

where $\nu$ is the frequency shift from the line center and $\nu_o$ is the sound-speed

$$\nu_o = \sqrt{\frac{kT}{m}} \tag{11}$$

where $m$ is the molecular mass.

For low $y$ parameters, the lineshape has a simple Gaussian form. For larger values of $y$ (in the range that we must be concerned with) the line shape becomes complicated and somewhat costly to evaluate. Fortunately, for values of $y$ under about 10, the Rayleigh-Brillouin line shape may be accurately approximated by the sum of three independent Gaussian functions (Witschas, 2011).

For the single-scatter return the line shape is determined by the sum of three Gaussian profiles. These calculations are based on the results of pre-computed approximate Gaussian fits made to accurate calculations of the line profile (Pan et al., 2002) corresponding to various $y$ parameters (See Fig. 4). For the higher order scattering, at each molecular scattering event the probability of the scattered photon packet being shifted to one of the acoustic side bands is calculated. Whether or not the

center frequency of the photon packet is shifted is then determined stochasticly using the relative weights of the component Gaussians and the subsequent shift and width of the associated Gaussian profile is used to determine the line shape. This line shape and shift is then used in the subsequent for the next multiple-scattering scattering event.
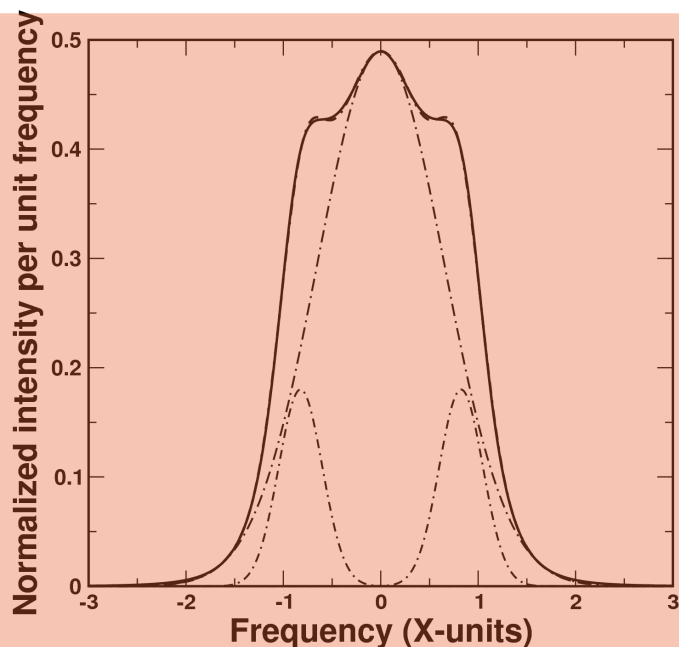
**Figure 4.** Exact Rayleigh-Brillouin line shape (Solid-Line) along with fitted sum of three Gaussian functions (Dashed–Dotted–line). The three component Gaussian functions are also shown. Here the $y$ parameter is equal to 1.0.

### 2.3 Lidar Instrument Model

ATLID transmitts a linearly polarized beam and possesses three receiver channels. A co-polar "Mie" channel which detects mainly particulate (i.e. clouds and aerosol) backscatter, a co-polar "Rayleigh" channel which detects mainly backscatter from molecules, and a cross-polar particulate+molecular channel. The Mie-Rayleigh separation is achieved using a Fabry-Perot Etalon. For details of ATLID's design see do Carmo et al. (2016) and do Carmo et al. (2021). Some of the important ATLID technical specifications are repeated in Table 3.

The ECSIM lidar instrument model ingests the output of the the lidar radiative transfer model and models the instrument response function including instrument noise. The model is capable of being configured to simulate various lidars, in this case ATLID. The instrument model ingests the time-spectral-polarization resolved output of the lidar radiative transfer model and models the three different optical ATLID detection channels. Explicit modelling of the Fabry-Perot Etalon HSRL spectrometer as well as various other optical elements in each detector path is performed. The effects of Poisson shot noise on the simulated observed signals is simulated. Instrument dark current and solar background effects are taken into account. Since ATLID's

| Parameter | Value |
|---|---|
| Telescope Diameter (m) | 0.62 |
| Wavelength (nm) | 355 |
| Receiver field-of-view (full-angle) (mads) | 66.5 |
| Laser Divergence (full-angle) (mrads) | 36 |
| Pulse Energy (mJ) | 35 |
| Range resolution (m) | 100m (0-20km)/500m(20km–40km) |
| PRF (Hz) | 51 |
| End-of-Life Co-polar-Mie transmission | 45 |
| End-of-Life Co-polar-Ray transmission | 43 |
| End-of-Life Cross-polar transmission | 43 |
| Quantum Efficiency | 79/75/79 |
| Molecular backscatter in Mie co-polar channel fraction | 25 % |
| Mie backscatter in Rayleigh co-polar channel fraction | 16 % |

**Table 3.** ATLID Technical Specifications

detector elements are Accumulation Charged Coupled Devices (ACCDs) they are subject to read-out noise. The effects of this read-out noise are also simulated.

### 2.3.1 Background Signal

The background signal refers to power registered by the lidar receiver that is due to the detection of photons from sources other than backscattered laser light. In the case here, the main source of background light will be scattered sunlight from the Earth's surface and atmosphere. As such, the background will depend on the solar angle, the surface type and the cloud cover. In this work, the lidar background values were based on an approximate look-up-table approach where the background irradiance at 355 nm was modelled using a radiative model for various cloud optical depths, surface albedos and solar zenith angle.

The background power incident upon the detector level is given by:

$$P_{back} = A_o \pi \rho_t \overline{I_b(\lambda) T_{rec}(\lambda)} \Delta\lambda \tag{12}$$

where $A_o$ is the effective receiver area, $\rho_t$ is the telescope $^1/_2$ angle field-of-view, and $\overline{I_b(\lambda) T_{rec}(\lambda)}$ is the average product of the receiver wavelength dependent transmission (including filtering) with the up-welling irradiance, and $\Delta\lambda$ is the wavelength interval that must be considered.

### 2.3.2 Receiver Optical Elements

The ATLID lidar receiver optical train is modelled as a sequence of elements (beam splitters, spectral filters, half-wave plates etc..) that operate on the spectral and polarization state of the lidar return. For simplicity, broad-band spectral filters are mod-

eled as having a rectangular passbands and are characterized by a single in-band transmission/reflection and a out-of-band reflection/transmission pair.

The Fabry-Perot (FP) etalon is the most critical optical element in ATLID's receiver chain. This element is used to separate the Mie signal from the Rayleigh signal.

The transmission function of an etalon without accounting for non-ideal effects such as the finite input beam collimation and surface roughness may be modeled (Saleh and Teich, 1991) as

$$T = \left(1 - \frac{A}{1-R}\right) \times \frac{1}{1 + \frac{4R}{(1-R)^2}\sin^2\frac{\delta}{2}} \tag{13}$$

where $R$ is the reflection coefficient of the etalon mirrors, $A$ is the relative absorption and scattering loss parameter and $\delta = 2\pi\frac{\lambda-\lambda_o}{\lambda_o}$ where $\lambda_o$ is the central wavelength. The corresponding etalon reflection function may be written as:

$$R = \frac{(1.0 - R - A)^2}{1.0 - 2.0 * R\cos(\delta) + R^2}. \tag{14}$$

The effects of non-parallel mirror flatness, diffraction and non-ideal beam collimation are taken into account by convolving the above absorbing etalon transmission and reflection functions by a top-hat function whose width in wavenumbers is given by:

$$\Delta\nu_{total} = \sqrt{\Delta\nu_{np}^2 + \Delta\nu_{ap}^2 + \Delta\nu_{dif}^2} \tag{15}$$

where $\Delta\nu_{np}$ accounts for the broadening due to non-parallel mirror alignment, $\Delta\nu_{ap}$ accounts for collimation or finite aperture effects

$$\Delta\nu_{ap} = 1.0 - \cos\left[\left(\frac{D_t}{D_{fp}}\right)^2 \frac{\rho_t}{2}\right] \tag{16}$$

where $D_t$ is the telescope diameter, $D_{fp}$ is the etalon diameter and $\rho_t$ is the telescope full-angle field-of-view in radians. $\Delta\nu_{dif}$ accounts for diffraction effects and is given by:

$$\Delta\nu_{dif} = 1.0 - \cos\left[\left(\frac{\lambda_o}{D_t}\right)\right] \tag{17}$$

An example etalon transmission and reflection profile as a function of $\nu - \nu_o$ is shown in Fig. 5. Here $T = 0.978$, $A = 0.0$, the Free-spectral-range (the distance between adjacent transmission maximums) is $0.5$ cm$^{-1}$, the flatness parameter is $\lambda_o/300$, the telescope field-of-view (FOV) is $5 \times 10^{-3}$mrad, the telescope diameter is 0.6 m and the etalon diameter is 0.05 m.

### 2.3.3 Noise Considerations

For a given time interval, the number of photons arriving at a given detector channel is given by

$$N_{det} = R_{det}\frac{\lambda_o}{hc}\left(P_{lid} + P_{back}\right)\Delta t \tag{18}$$
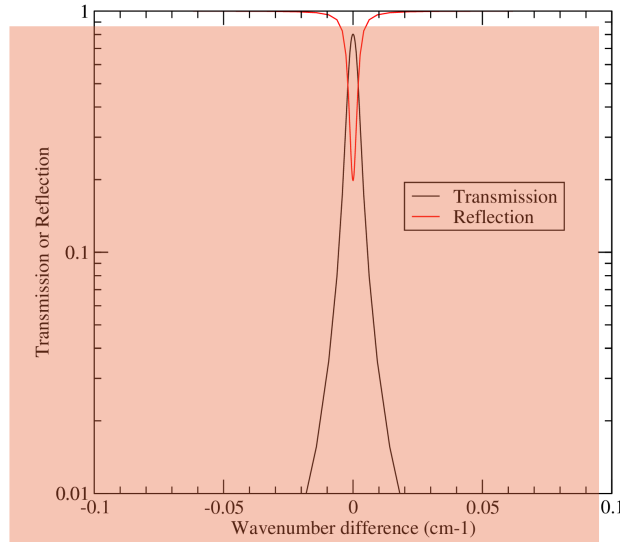
**Figure 5.** Etalon reflection and transmission showing non-ideal behavior.

where $\Delta t$ is the averaging time interval, $h$ is Planck's constant, $\lambda_o$ is an appropriate mean wavelength for the detector channel in question , $R_{det}$ is the detector quantum efficincy, $P_{lid}$ is the power received at the lidar detector and $P_{back}$ is the background power arising from the up-welling reflected solar radiance at 355 nm. According to Poisson statistics the standard deviation of $N_{\text{det}}$, shot-noise, is simply given by

$$\delta N_{det} = \sqrt{N_{det}} \tag{19}$$

In addition to the shot-noise, the effects of dark current noise and ACCD readout noise are also simulated. The average dark current is a characteristic of the detection electronic chain and its mean level is assumed to be constant in terms of effective noise photocounts per unit time. The dark-count rate is assumed to follow Poisson statistics. The read-out noise is a characteristic of CCD detectors is applied only at simulated read-out time. For ATLID, the envisioned default mode is for the ACCD to accumulate two shots before read-out. The mean read-out noise is assumed to be constant per read-out event and is modelled using Poisson statistics.

### 2.3.4 Spectral and Polarization Cross-talk

The characteristics of the HSRL filter ensure that a degree of spectral cross-talk between the Mie and Rayleigh channels exists (do Carmo et al., 2021). The effects of cross-talk are simulated within the lidar instrument model. The lidar instrument model

also applies a cross-talk correction, following the ATLID L0 to L1 processor step. This correction assumes perfect knowledge of the appropriate cross-talk coefficients. In flight, uncertainties in the coefficients will introduce additional errors, however, this is not envisioned to be a significant issue as the in-flight values of the coefficients are envisioned to be known within an accuracy of a few percent. This is to be achieved using a combination of a priori on ground characterization as well as in flight monitoring and adjustment e.g. using returns above 30 km (where only molecular scattering returns should be expected) to quantify the amount of Rayleigh-to-Mie cross-talk and using e.g. ground and suitably strong cloud returns (where to a good approximation only elastic particulate scattering returns exist) to quantify the Mie-to-Rayleigh cross talk. Due to the occurrence of variable but significant amounts of molecular Brillouin scattering in liquid water (Hostetler et al., 2018), water surfaces are not expected to be suitable cross-talk coefficient determination targets for ATLID.

For simplicity, the simulations presented in this paper did not consider polarization cross-talk. However, as is the case with lidars in general, due to e.g. non-ideal optical element characteristics as well as laser polarization purity and imperfect optical element orientation and alignments, a degree of polarization cross-talk will be present in the ATLID instrument (Bravo-Aranda et al., 2016; Donovan et al., 2015). The correction of spectral cross-talk effects will be conducted at the L1 stage by the L0 to L1 processor. The estimated errors will be reported in the L1 product.

The characterization of the ATLID polarization cross-talk has been the subject of extensive ground-testing and characterization e.g. do Carmo et al. (2016). In flight characterization and monitoring of the spectral cross-talk is planned to be accomplished using e.g. high-altitude returns where the Rayleigh depolarization signal is known and from solar background measurements over e.g. thick ice clouds at high sun conditions.

## 2.4 Radar Simulation

### 2.4.1 Radar Radiative Transfer Calculations

The radar scattering calculations described in Section 2.1.3 are used to estimate the radar backscatter cross section $\sigma_{b,j}(D)$ for each hydrometeor species $j$ and particle maximum diameter $D$ in units of $\mathrm{m}^2$. In addition, the extinction cross-section $\sigma_{e,j}(D)$ for each hydrometeor species $j$ and particle maximum diameter $D$ is estimated in units of $\mathrm{m}^2$. The GEM particle size distribution number concentration $n_j(D)$ is provided in units of $\mathrm{m}^{-3}$ (integrated across the bin diameter width) and the particle sedimentation velocity $V_j(D)$ is provided in units of $\mathrm{ms}^{-1}$.

Thus, the radar reflectivity factor $Z_e$ ($\mathrm{mm}^6\mathrm{m}^{-3}$), specific attenuation $A$ ($\mathrm{dBkm}^{-1}$) and reflectivity-weighted hydrometeor sedimentation velocity $V_{SED}$ ($\mathrm{ms}^{-1}$) are estimated for each hydrometeor species as follows:

$$Z_{e,j} = 10^{18} \frac{\lambda^4}{\pi^5 |K_w^2|} \sum_{i=1}^{Nbins} [\sigma_{b,j}(D)\, n_j(D)] \tag{20}$$

$$A_j = \frac{0.01}{\ln(10)} \sum_{i=1}^{Nbins} [\sigma_{e,j}(D)\, n_j(D)] \tag{21}$$

$$V_{SED,j} = \frac{\sum_{i=1}^{Nbins} [\sigma_{b,j}(D) \, n_j(D) \, V_j(D)]}{\sum_{i=1}^{Nbins} [\sigma_{b,j}(D) \, n_j(D)]} \tag{22}$$

where $\lambda$ is the wavelength in m, $j$ is the index for the hydrometeor species (cloud, rain, ice, snow, graupel and hail) and $|K_w^2|$ is the dielectric factor of water at 94-GHz. The aforementioned parameters are combined to produce their total value at each GEM grid point

$$Z_e = \sum_{j=1}^{6} Z_{e,j} \qquad A_h = \sum_{j=1}^{6} A_j \qquad V_{SED} = \frac{\sum_{j=1}^{6} [Z_{e,j} \, V_{SED,j}]}{Z_e} \tag{23}$$

In addition to hydrometeor attenuation $A_h$, gaseous attenuation $A_g$ can have a significant impact on the measured CPR reflectivity. For example, two-way gaseous attenuation from surface to upper troposphere of more than 5 dB is not unusual in tropics. (Josset et al., 2013) compared different models to estimate absorption at W-band by gases by taking advantage of the collocated CloudSat–Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) measurements. Their results indicate that the Rosenkranz (1998) model fits the observations best. The gaseous attenuation at each model grid point $A_g$ (dBkm$^{-1}$) is calculated using Rosenkranz (1998) and the water vapor mixing ratio profile from the EarthCARE auxiliary meteorological product (X-MET (Eisinger et al., 2023). Consistent with the nadir viewing of the EarthCARE CPR, the $A_h$ and $A_g$ estimates are computed assuming propagation along the GEM model vertical dimension.

The normalized (per unit of area) cross-section of the Earth's surface $\sigma_0$ [m$^{-1}$] is estimated. Over an ocean surface, the normalized cross section is estimated using the relationship from Li et al. (2005) as a function of the near-surface wind speed provided in the X-MET data product. CloudSat observations have shown that over the ocean surface $\sigma_0$ is known within 2 dB (Tanelli et al., 2008) and over land exhibits very large variability due to its dependency on vegetation, surface slope, soil moisture, snow cover and other factors (Haynes et al., 2009). At 94-GHz, the ocean surface $\sigma_0$ varies between 16 to 6 dB for near-surface wind speeds of 2 to 20 ms$^{-1}$ respectively (Tanelli et al., 2008). The corresponding radar reflectivity factor for the Earth's surface is estimated using the formula:

$$Z_{SFC} = 10^{18} \frac{\sigma_0 \lambda^4}{\pi^5 |K_w|^2 \Delta Z} \tag{24}$$

where $\Delta Z$ is the EarthCARE CPR range resolution (500 m). The $Z_e$ and $Z_{SFC}$ are intrinsic reflectivities (unattenuated). The measured (attenuated) CPR reflectivity $Z_{e,m}$ at a height $z$ in the GEM model can be defined in terms of intrinsic reflectivity (unattenuated) and the hydrometeors and gaseous specific attenuation using the formula:

$$Z_{e,m}(z) = Z_e(z) \exp^{-0.46 \sum_{z_{top}}^{z} (A_h(s) + A_g(s)) \Delta s} \tag{25}$$

where $z_{top}$ is the highest level in the GEM model (representing the top of the atmosphere). If the height in the model is set to zero, then this formula can be used to estimate the attenuated Earth's surface radar reflectivity.

### 2.4.2 Multiple scattering

Multiple scattering (MS) signatures have been observed in spaceborne radar observations Battaglia et al. (2010). The MS contribution to the observed spaceborne radar observations is more pronounced at higher radar frequencies and for the larger instantaneous field of view. Although the EarthCARE CPR has a smaller radar footprint compared to CloudSat, we do expect to see significant MS effects especially in deep convective cores (Battaglia and Tanelli, 2011). Here, the fast radar multiple scattering model for wide-angle scattering using the time-dependent two-stream approximation introduced by Hogan and Battaglia (2008) is used to estimate the CPR multiple scattering profiles. The MS model can only provide forward-simulated MS CPR radar reflectivity. The 'c' panels in Figs. 13, 21 and 28 show examples of CPR forward simulations using the MS model. We did not perform MS simulations of the CPR Doppler velocity since we plan to flag and remove the CPR profiles with significant MS contributions (Kollias et al., 2022b).

### 2.4.3 Radar Instrument Model

The EarthCARE CPR instrument model has two main modules. The first is the sampling geometry module that determines which part of the GEM model is sampled by the CPR at any given time step (or along track location) and accounts also for the along track displacement of the satellite sampling volume during signal integration. The second module is the CPR receiver module that introduces the instrument noise and estimates the CPR Doppler moments and their corresponding uncertainty (Kollias et al., 2007, 2022b). Some of the important EarthCARE CPR technical specifications are shown in Table 4.

| Parameter | Value |
| --- | --- |
| Antenna Diameter (m) | 2.5 |
| Frequency (GHz) | 94.05 |
| Noise ($mm^6 m^{-3}$) | 0.0071 |
| Altitude (km) | 400 |
| Pulse Length (m) | 500 |
| Range resolution (m) | 100 |
| Min Range (km) | -1 |
| Max Range (km) | 20 |
| PRF (Hz) | 6100 - 7500 |
| Speed ($ms^{-1}$) | 7600 |
| $\theta_{3dB}$ (deg) | 0.095 |
| IFOV$_{3dB}$ (m) | 700 |

**Table 4.** CPR Technical Specifications. The last two rows refer to the beam 3dB full width measured in degrees and the foot-print on the ground.

Following (Kollias et al., 2014, 2022b) the antenna weighting function $W_a(x,y)$ is shown in Fig. 6 where $x$ and $y$ are the along and cross track dimensions respectively. The $W_a(x,y)$ determines the contribution of GEM model grid point radar reflectivity $Z_e(x,y)$ to the total radar reflectivity observed by the CPR at a particular model grid. At a GEM model height $z$, the radar reflectivity contribution $Z_{GEM}(z)$ to the CPR is estimated using the equation:

$$Z_{GEM}(z) = \sum_{i=1}^{Xbins} \sum_{j=1}^{Ybins} [Z_e(i,j)\, W_a(i,j)] \tag{26}$$

where $Xbins$ and $Ybins$ are the number of GEM grid points that are illuminated by the CPR radiation. Next, the CPR range weighting function $W_r(x,y)$ that described the along range point target response of a radar is applied to estimate CPR measured radar reflectivity at a particular range $r$. The CPR range weighting function for the EarthCARE and CloudSat CPR's are shown in Fig. 6.
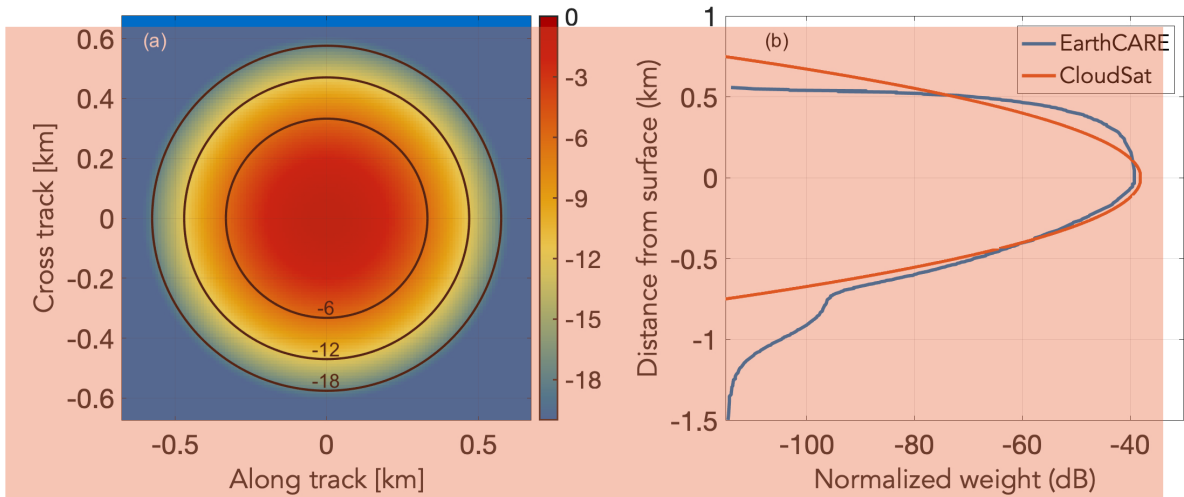
**Figure 6.** (a)the EarthCARE CPR antenna weighting function $W_a(x,y)$ distributions in the along and cross track direction. Units are dB below the peak value. (b) The range weighting function $W_r(x,y)$ for CloudSat and EarthCARE

The CPR range weighting function is the result of the transmitted waveform (the same for both radar) and the CPR receiver filter. The EarthCARE CPR receiver filter was designed to sharply cut off the $W_r(x,y)$ at 500 m above the Earth's surface to improve the detection of low-level clouds compare to CloudSat (Lamer et al., 2020). The CPR radar reflectivity at range $r$ is estimated using the equation:

$$Z_{CPR}(r) = \sum_{i=1}^{Gbins} [Z_{GEM}(z(i)-r)\, W_r(z(i)-r)] \tag{27}$$

where $Gbins$ is the number of GEM model levels within the limits of the CPR range weighting function centered at range $r$. The forward simulation of the EarthCARE CPR Doppler velocity and spectrum width requires to introduce the apparent Doppler velocity $V_{app}(x)$ introduced to each GEM model point within the CPR footprint will have due to the satellite motion $V_{sat}$. The $V_{app}(x)$ is estimated using the equation:

$$V_{app}(x) = -x\frac{V_{sat}}{H_{sat}} \tag{28}$$

where $x$ is the distance from nadir and $H_{sat}$ is the altitude of the satellite Fig. 7. The $V_{app}(x)$ is independent of the cross-track distance $y$, thus $V_{app}(x,y) = V_{app}(x)$. At a GEM model height $z$, the Doppler velocity contribution $V_{GEM}(z)$ to the CPR is estimated using the equation:

$$V_{GEM}(z) = \sum_{i=1}^{Xbins}\sum_{j=1}^{Ybins} [Z_{e,m}(i,j)\,W_a(i,j)\,(V_{app}(i,j) + W_{air}(i,j) + V_{SED}(i,j))] \tag{29}$$

where $W_{air}(i,j)$ is the vertical air motion (negative is up) and $V_{SED}(i,j)$ is the total sedimentation velocity at the (i,j) GEM grid point.

In the CPR receiver module, the product $S_v(i,j) = Z_{e,m}(i,j)W_a(i,j)$ and the sum of all the velocity components $V_{tol}(i,j) = V_{app}(i,j) + W_{air}(i,j) + V_{SED}(i,j)$ are used to construct the periodogram $S(V)$ $(mm^6 m^{-3}/ms^{-1})$ of the returned radar signal at each height $z$ of the GEM model following the methodology proposed first by (Zrnic, 1975). Since we are using radar reflectivity at power, the noise is also given in radar reflectivity units (see Table 3). The periodogram is a very useful tool for describing a time series data set. In a radar system, the PRF determines the temporal resolution of the time series data at a particular range gate. The periodogram is an estimate of the spectral density of a signal. The periodogram is interpolated at a spectral velocity resolution that is determined by the CPR sampling rate (PRF) which determines the higher sampled frequency. This is often called Nyquist frequency ($F_N = PRF/2$), which is half the sampling frequency of a discrete signal processing system. It is sometimes known as the folding frequency of a sampling system. Using the radar wavelength $\lambda$ (Table 3), we can convert the folding frequency to folding velocity or as often-called Nyquist velocity ($V_N = PRF\lambda/4$). If the Doppler velocity of a radar target exceeds the magnitude of the Nyquist velocity, folding occurs (aliasing).

Once the radar receiver noise is added in the frequency domain, the next step is to perform Inverse Fast Fourier Transform (IFFT) of the constructed Doppler spectrum in order to retrieve I (in-phase) and Q (quadrature-phase) voltage time series (Kollias et al., 2014). The in-phase channel includes the portion of the signal in phase with the reference sinusoid. The quadrature-phase channel includes the portion of the signal 90° out of phase with the reference sinusoid, is abbreviated Q The IFFT operator is applied to the amplitude spectrum. The temporal spacing of the I/Q voltages is $1/PRF$. As described in (Kollias et al., 2014), the I/Q time series at each GEM model height $z$ are convoluted along range with the square root of the range weighting function $W_r(x,y)$ and then are used as input to a pulse-pair Doppler moments estimator (Zrnic, 1977) to provide estimates of the CPR radar reflectivity, Doppler velocity and Doppler spectrum width.
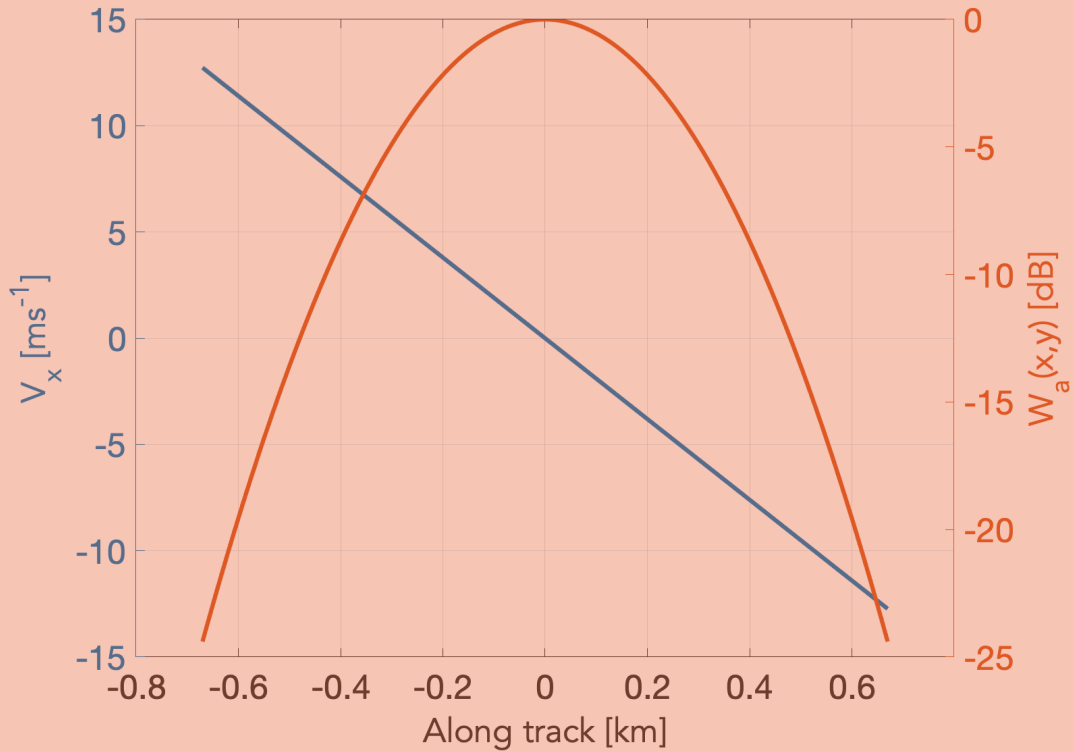
**Figure 7.** The EarthCARE CPR antenna weighting function $W_a(x, y)$ and apparent Doppler velocity $V_{app}(x)$ distribution as a function of distance in the along track direction

## 2.5 Multi-spectral imager (MSI) Simulations

### 2.5.1 Short-Wave Radiative transfer Calculations

The short-wave radiances for the first four EarthCARE Multi-Spectral-Imager (MSI) channels were performed by applying a one dimension radiative transfer model. ECSIM contains and option to perform 3D Monte-Carlo calculations, however, it was unfeasible to apply this option for the totality of the domain required. The 1D calculations used DISORT with 32 streams (Lin et al., 2015) driven by the atmospheric absorption, phase functions, and surface BRDFs extracted from the scene file. The band-limits and relevant atmospheric gasses for the MSI SW bands are listed in 5 and are based on an early instrument specification e.g. (Battrick., 2004). Here for simplicity, a top-hat response was assumed and it should be noted that the exact band widths and centers differ from that of the actual instrument. The MSI retrieval algorithms makes these same assumptions when applied to the simulated data, however, the detailed non-ideal spectral response of each channel will be fully accounted for when these algorithms are supplied with actual observations.

Even with a 1-D approach, the computation covering the 6000-by-150 km frame domain at a horizontal resolution of 0.25 km is computationally demanding. In order to speed-up the process, a sampling and interpolation process was implemented as follows,

1. The domain was divided into a number of sub-domains of n×m pixels. In this work n=12 and m=12.

2. For each pixel in the sub-domain the quantities $X_1(i,j) = A_s(i,j)\exp(-\tau(i,j))$ and $X_2(i,j) = \exp(-\tau(i,j))$ were calculated, where $\tau$ is the particulate optical depth and $A_s$ is the pixel black-sky albedo.

3. Within each subdomain, the pixels where DISORT would be applied to were selected accordingly,

   (a) The lower left pixels were selected by default

   (b) In the along-track direction each $i$th and $j$th pixels are selected by default. In this work, every 5th pixel was used.

   (c) The pixels along the spacecraft nadir track are selected by default.

   (d) The pixels with the maximum and minimum values of $X_1$ and $X_2$ respectively and selected.

4. For each of the selected pixels, DISORT was used to calculate the TOA radiances.

5. For each of the non-selected pixels (denoted by $i_n,j_n$) the radiance of the selected pixels $(i_s,j_s)$ are scanned through and the pixel that minimizes $abs(X_1(i_n,j_n) - (X_1(i_s,j_s))) + abs(X_2(i_n,j_n) - (X_2(i_s,j_s)))$ is used to fill in the TOA radiance.

6. The domain was shifted by half its width in the along-track direction and then shifted in the cross-track direction when the end of the frame was reached and the process repeated until the entire frame was covered. The overlap between domains was found to be useful for eliminating artifacts mimicking the sub-domain structure.

The procedure was able to speed-up the necessary calculation by an order of magnitude while retaining a suitable degree of accuracy. For example, for the "Halifax scene" (See Section 3), when smoothed to the MSI resolution of 1x1 km and referring to the MSI CH1 BRDF, 65% of the resulting errors were less than 0.1%, 87 % of the errors were within 1% and 99% of the errors were less than about 5%. As may be expected, the larger erros are associated with sharp cloud edges. The low errors are mainly due to the high degree of correlation in the clouds fields for distances less than 1.5 km or so. This echos the findings of Barker and Li (2019).

Of course, by neglecting 3-D radiative transfer effects the simulations are not as realistic as they might be. It is well known that 3D effects can have espically large effects on the measured radiances e.g. Barker and Liu (1995). However, it should be noted that apart from drastically reducing the computational requirements, there is one aditional technical benefit in performing the simulations using 1-D radiative transfer. That is the fact that the MSI retrieval algorithms (e.g. Hünerbein et al. (2023)), while still being state-of-the-art, are based on 1-D radiative transfer theory. Thus, the 1D simulations are better suited for the technical evaluation of these algorithms even though the 3-D simulations are less realistic. It should also be noted that the use

of the scenes for 3D RT studies is ongoing and 3D effects are being addressed in the EarthCARE radiative closure evaluation Cole et al. (2022).

After the MSI short-wave calculations were performed (at a resolution of 0.25 km), the radiances were binned to the MSI instrument resolution (1-km) and the effects of instrument random noise was simulated using a Gaussian pseudo-random number generator. Technical details of the MSI can be found in Chang (2019).

| Channel | $v_1$ cm$^{-1}$ | $v_2$ cm$^{-1}$ | Gasses |
|---------|------------|------------|--------|
| MSI_1 660 nm | 14948. | 15408. | $H_2O$, $O_3$ |
| MSI_2 865 nm | 11429. | 11696. | $H_2O$ |
| MSI_3 1.61 um | 6098. | 6329. | $CO_2$ |
| MSI_4 2.2 um | 4444. | 4651. | $H_2O$, $CH_4$ |

**Table 5.** Short-wave Multispectral Imager (MSI) bands. $v_1$ and $v_2$ are the effective band lower and upper wavenumber limits respectively.

### 2.5.2 Long-Wave Radiative transfer Calculations

Like the short-wave msi calculations, the TOA radiances for the long-wave MSI channels (see Table 6) were calculated using DISORT. Since the long-wave calculations are not as computationally demanding as the short-wave calculations, no sampling strategy was necessary and DISORT with 16 streams specified was applied to every pixel in the MSI frame domain.

As was the case with the short–wave channels, the idealized radiances at 0.25 km resolution were binned to the MSI instrument resolution (1-km) and the effects of instrument random noise was simulated using a Gaussian pseudo-random number generator.

| Channel | $v_1$ cm$^{-1}$ | $v_2$ cm$^{-1}$ | Gasses |
|---------|------------|------------|--------|
| MSI_5 8.85 $\mu$m | 1064. | 1205. | $H_2O$, $O_3$ |
| MSI_6 10.85 $\mu$m | 885. | 962. | $H_2O$, $O_3$ |
| MSI_7 11.85 $\mu$m | 813. | 877. | $H_2O$, $CO_2$ |

**Table 6.** Long-wave multispectral imager (MSI) Bands.

## 2.6 Broad-Band Radiometer (BBR) Simulations

The methods described for the MSI short and long–wave TOA radiances were applied to the BBR calculations for the spectral-band resolved BBR TOA radiances for each of the three BBR views as well as the TOA fluxes. The short-wave bands used are listed in Table 7 and long-wave bands are listed in Table 8.

| **Broad-Band Short Wave 0.2-4.0 $\mu$ m** | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| **Center $\lambda$** | Gasses | **Center $\lambda$** | Gasses | **Center $\lambda$** | Gasses | **Center $\lambda$** | Gasses | **Center $\lambda$** | Gasses |
| 0.215 | $O_3$ $O_2$ | 0.240 | $O_3$ $O_2$ | 0.264 | $O_3$ | 0.290 | $O_3$ | 0.315 | $O_3$ |
| 0.345 | $O_3$ | 0.380 | $O_3$ | 0.420 | $O_3$ | 0.460 | $O_3$ $H_2O$ | 0.500 | $O_3$ |
| 0.545 | $O_3$ $H_2O$ | 0.603 | $O_3$ $H_2O$ | 0.665 | $H_2O$ $O_3$ $O_2$ | 0.721, | $H_2O$ $O_3$ $O_2$ | 0.766 | $H_2O$ $O_3$ $O_2$ |
| 0.824 | $H_2O$ $O_3$ | 0.933 | $H_2O$ $O_3$ | 1.048 | $H_2O$ $O_3$ | 1.143 | $H_2O$ | 1.233 | $H_2O$ |
| 1.405 | $H_2O$ $CO_2$ | 1.589 | $H_2O$ $CO_2$ | 1.886 | $H_2O$ $CO_2$ | 2.254 | $H_2O$ | 2.644 | $H_2O$ $CO_2$ |
| 3.166 | $H_2O$ $O_3$ | 3.714 | $H_2O$ | 4.080 | $H_2O$ $CO_2$ | 4.239 | $H_2O$ $CO_2$ | 4.384 | $H_2O$ $CO_2$ |
| 4.5469 | $H_2O$ $CO_2$ | 4.7186 | $H_2O$ $CO_2$ | 4.9038 | $H_2O$ | 5.0252 | $H_2O$ | | |

**Table 7.** Center wavelengths used for broad-band short-wave calculations and relevant gasses.

The BBR ideal radiances at 250m resolution as calculated by the radiative transfer code were used to create simulated L1-b data, i.e, BBR filtered radiances: B-NOM, and BBR single pixel filtered radiances: B-SNG (see production model in Eisinger et al. (2023)). This process involved in a first step the convolution of the simulated radiances with the spectral responses of the BBR instrument to obtain Broadband SW, $L_{SW}$, and LW, $L_{LW}$, radiances at 250m. Secondly, as the BBR instrument will measure TW radiances, which are not simulated, the TW ($L_{TW}$) radiances at 250m resolutions are calculated as the weighted sum of the Long- and Short-wave chenels (i.e. $L_{TW} = L_{LW} + A \cdot L_{SW}$) (Velázquez-Blázquez et al., 2023).

In order to simulate realistic L1 B-SNG inputs, the Chopper Drum Mechanism (CDM) speed has to be taken into account. It has been configured at a 0.7 ratio of the original nominal speed, as recommended to maximise the lifetime of the mechanism, which defines a ground sampling distance for two consecutive SW or TW measurements of 1.1 km. The current BBR simulator software to produce the B-SNG input performs a bilinear interpolation of the 250m broadband radiances at the positions of the 30 detectors array and this for each of the BBR views (fore, nadir and aft). The resulting B-SNG SW and TW radiances for the Halifax scene (Section 3.1) are show in Figures 8 and 9. Finally, a domain integration or psf-weighting is done to pass from the single pixel radiances in B-SNG to the nominal BBR resolutions in B-NOM (standard, full and small), sampled every 1km.

## 3 The Tests Frames

In this section, simulated L1 data for the EarthCARE test scenes are presented and discussed. The Level-1 simulated data as well as various model truth fields for 3 GEM derived scenes are available in the available from https://doi.org/10.5281/zenodo.7117115.

### 3.1 Halifax

The high-latitude part of the Halifax scene features mixed-phase clouds at night-time, transitioning from deeper clouds with tops up to 6 km around $65°$N featuring supercooled liquid in convective cells, to mixed-phase clouds with tops around 3 km at temperatures as cold as $-30°$C, and finally to more broken shallow mixed-phase clouds toward $50°$N. Near the center of

| Long-Wave 4.0-400.0 $\mu$m | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Center $\lambda$ | Gasses | Center $\lambda$ | Gasses | Center $\lambda$ | Gasses | Center $\lambda$ | Gasses | Center $\lambda$ | Gasses |
| 4.07997 | $H_2O$ | 4.23729 | $H_2O$ | 4.38596 | $H_2O$ | 4.54545 | $H_2O$ | 4.71698 | $H_2O$ |
| 4.90196 | $H_2O$ | 5.02513 | $H_2O$ | 5.07614 | $H_2O$ | 5.12820 | $H_2O$ | 5.18135 | $H_2O$ |
| 5.23560 | $H_2O$ | 5.29101 | $H_2O$ | 5.34759 | $H_2O$ | 5.40541 | $H_2O$ | 5.46448 | $H_2O$ |
| 5.52486 | $H_2O$ | 5.58659 | $H_2O$ | 5.64972 | $H_2O$ | 5.71429 | $H_2O$ | 5.78035 | $H_2O$ |
| 5.84795 | $H_2O$ | 5.91716 | $H_2O$ | 5.98802 | $H_2O$ | 6.06061 | $H_2O$ | 6.13497 | $H_2O$ |
| 6.21118 | $H_2O$ | 6.28931 | $H_2O$ | 6.36943 | $H_2O$ | 6.45161 | $H_2O$ | 6.53595 | $H_2O$ |
| 6.62252 | $H_2O$ | 6.71141 | $H_2O$ | 6.80272 | $H_2O$ | 6.89655 | $H_2O$ | 6.99301 | $H_2O$ |
| 7.09220 | $H_2O$ | 7.19424 | $H_2O$ | 7.29927 | $H_2O$ | 7.40741 | $H_2O$ | 7.51880 | $H_2O$ |
| 7.63359 | $H_2O$ | 7.75194 | $H_2O$ | 7.87402 | $H_2O\ O_3$ | 8.00000 | $H_2O\ O_3$ | 8.13008 | $H_2O\ O_3$ |
| 8.26446 | $H_2O\ O_3$ | 8.40336 | $H_2O\ O_3$ | 8.54701 | $H_2O\ O_3$ | 8.69565 | $H_2O\ O_3$ | 8.84956 | $H_2O\ O_3$ |
| 9.00901 | $H_2O\ O_3$ | 9.17431 | $H_2O\ O_3$ | 9.34579 | $H_2O\ O_3$ | 9.52381 | $H_2O\ O_3$ | 9.70874 | $H_2O\ O_3$ |
| 9.90099 | $H_2O\ O_3$ | 10.1010 | $H_2O\ O_3$ | 10.3093 | $H_2O\ O_3$ | 10.5263 | $H_2O\ O_3$ | 10.7527 | $H_2O\ O_3$ |
| 10.9890 | $H_2O\ O_3$ | 11.2360 | $H_2O$ | 11.4943 | $H_2O$ | 11.7647 | $H_2O$ | 12.0482 | $H_2O\ CO_2$ |
| 12.3457 | $H_2O\ CO_2$ | 12.6582 | $H_2O\ CO_2$ | 12.9870 | $H_2O\ CO_2$ | 13.3333 | $H_2O\ CO_2$ | 13.6986 | $H_2O\ CO_2$ |
| 14.0845 | $H_2O\ CO_2$ | 14.4928 | $H_2O\ CO_2$ | 14.9254 | $H_2O\ CO_2$ | 15.3846 | $H_2O\ CO_2$ | 15.8730 | $H_2O\ CO_2$ |
| 16.3934 | $H_2O\ CO_2$ | 16.9492 | $H_2O\ CO_2$ | 17.5439 | $H_2O\ CO_2$ | 18.1818 | $H_2O\ CO_2$ | 18.8679 | $H_2O\ CO_2$ |
| 19.6078 | $H_2O\ CO_2$ | 20.4082 | $H_2O$ | 21.2766 | $H_2O$ | 22.2222 | $H_2O$ | 23.2558 | $H_2O$ |
| 24.3902 | $H_2O$ | 25.6410 | $H_2O$ | 27.0270 | $H_2O$ | 28.5714 | $H_2O$ | 30.3030 | $H_2O$ |
| 32.2581 | $H_2O$ | 34.4828 | $H_2O$ | 37.0370 | $H_2O$ | 40.0000 | $H_2O$ | 43.4783 | $H_2O$ |
| 47.6190 | $H_2O$ | 52.6316 | $H_2O$ | 58.8235 | $H_2O$ | 66.6667 | $H_2O$ | 76.9231 | $H_2O$ |
| 90.9091 | $H_2O$ | 111.111 | $H_2O$ | 142.857 | $H_2O$ | 200.000 | $H_2O$ | 333.333 | $H_2O$ |

**Table 8.** Center wavelengths used for broad-band long-wave calculations and relevant gasses.

the frame a storm system with supercooled layers, convective precipitation and ice clouds with tops up to altitudes of 13 km are present. South of about $45°N$ cloud-free and shallow low altitude water clouds are present. Also south of $45°N$ extended aerosol regions are present. From the ground to above 2 km a marine layer (mainly Sea-Salt) is present and above this a thinner continental pollution layer (mainly fin-mode non absorbing aerosol) is present. The differences in the lidar ratio associated with these two aerosol regions shown in the middle panel of Figure 10 is evident.

The simulated ATLID Mie, Rayleigh and Cross-polar attenuated backscattters are shown in Figure 11. It can be seen that, as expected, the lidar penetration into the clouds is limited, especially in the central part of the frame. However, most of the ice clouds are well captured. The aerosol fields in the southern segment of the frame are also well captured. Quantitative measures of the ability to detect features in the ATLID data are presented in van Zadelhoff et al. (2023), which describes the
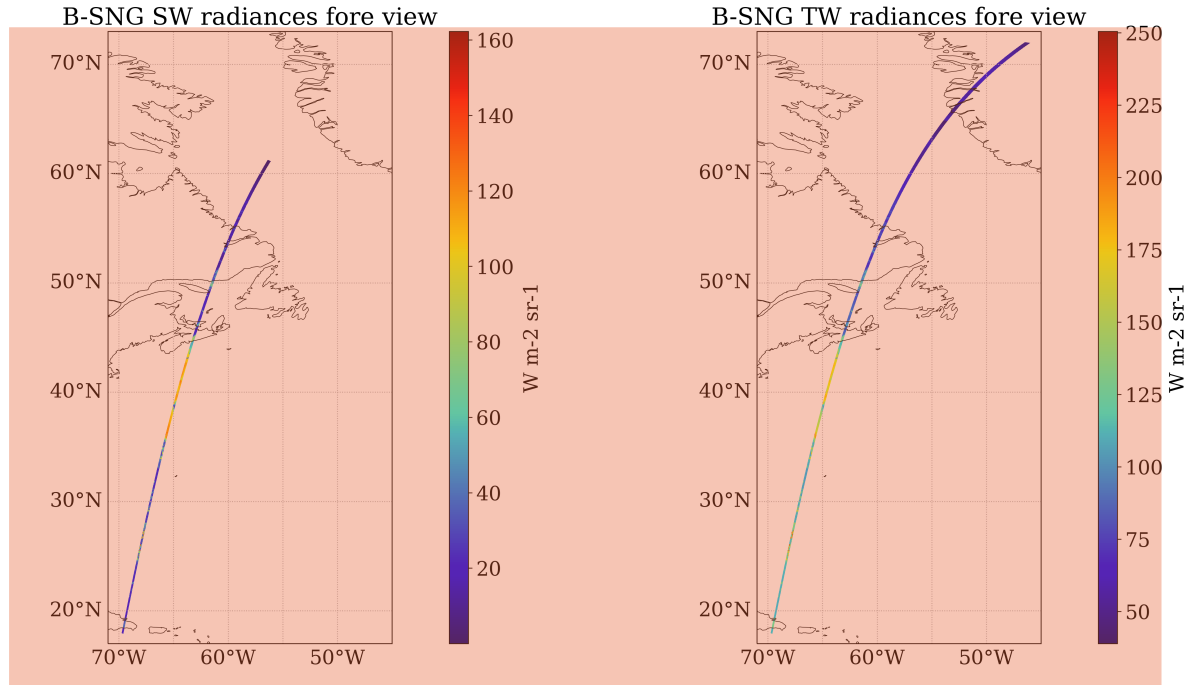
**Figure 8.** Simulated B-SNG TOA broadband radiances for the fore view of the BBR SW and TW channels for the Halifax scene. Similar plots are obtained for the nadir and aft views.

ATLID featuremask product (A-FM). The accuracy and sensitivity of quantitative cloud and aerosol extinction and backscatter retrievals applied to the ATLID data is discussed in e.g. Donovan et al. (2023) and Mason et al. (2023).

The nadir fields of particle mass, effective radius, and simulated observed Radar reflectivity for the full Halifax scene are shown in Fig. 12. The 'striped' area in the radar reflectivity present in the upper-left of the lowers panel is due to a simulated change in maximum height covered by the radar which is associated with a latitude dependent change in the operating PRF. Radar reflectivity is a strong function of the particle effective radius, hence, in general, the larger effective radii regions of the scene are well sampled, while areas containing relativity small water cloud droplets are not.

The GEM model and CPR simulated signals for a selected region of the Halifax scene are shown in Fig. 13. Due to the large horizontal extend of the scene, we focus on the $[30° - 48°N]$ simulated region that covers the frontal and convective systems. The top two panels shows the unattenuated 94-GHz radar reflectivity and the reflectivity weighted hydrometeor sedimentation velocity at the GEM grid resolution. The two systems are characterized by high cloud tops ( 12km), periods with thick high level clouds and extensive periods with light and moderate precipitation. The lower two panels show the raw
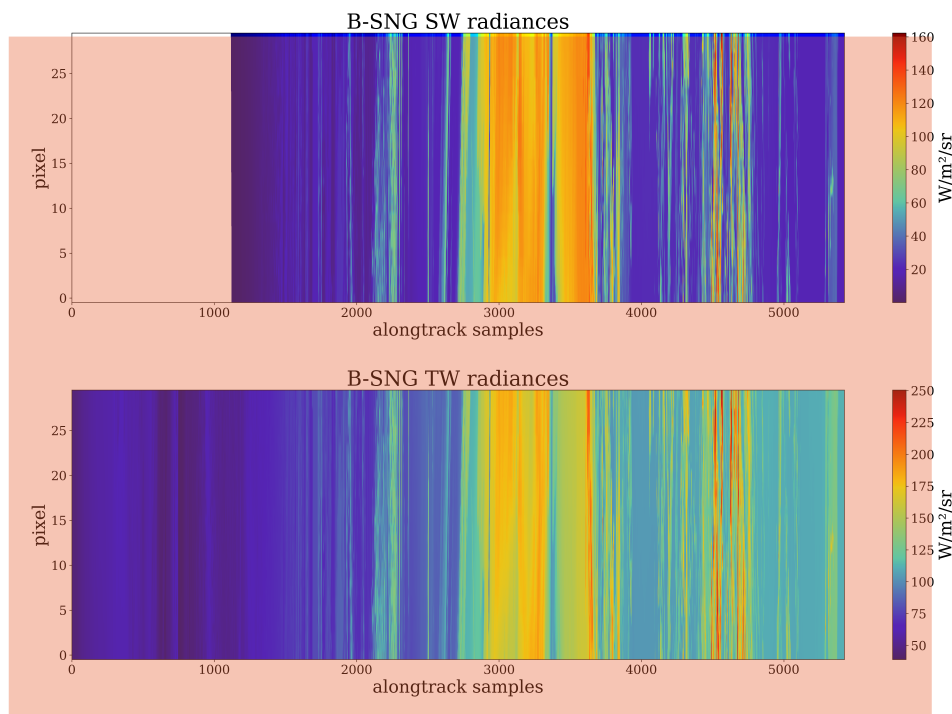
**Figure 9.** Simulated B-SNG TOA broadband radiances for the for view of the BBR SW and TW channels for the Halifax scene.

(uncorrected) CPR radar reflectivity and mean Doppler velocity measurements. The EarthCARE CPR has sufficient sensitivity to detect the hydrometeor layers except weak reflectivity echoes near the highest cloud tops. The strong 94-GHz attenuation by hydrometeors results to missed detections near the surface. This can be clearly seen by the depression of the surface echo radar reflectivity at 3700 km and the complete loss of the surface echo around 4100 km. The fainted CPR echoes that fill the surface echo gap around 4100 km are due to multiple scattering (Battaglia et al., 2010). As expected, the CPR raw Doppler velocity field (500 m along track integration) is noisy (Kollias et al., 2014, 2022a). The post-processing algorithm described in (Kollias et al., 2022b) are expected to significantly improve the quality of the CPR Doppler velocity measurements. Despite their noisiness, the CPR Doppler velocities reproduce the main features of the GEM model Doppler velocities, namely, the transition from solid to liquid hydrometeors and the low sedimentation Doppler velocities in the upper cloud levels. More detail concerning the performance of the CPR and the CPR cloud algorithms are provided in Mroz et al. (2023) and Kollias et al. (2002).

The simulated short-wave TOA radiances for the MSI short-wave channels are shown in Figure 14. Here, north of about 55°N, the sun is below the horizon so that the radiances are zero. Below about 42°N the swath is above ocean with the exception of passing over the Dominican Republic near the southern frame border. The NDVI (defined as the difference in
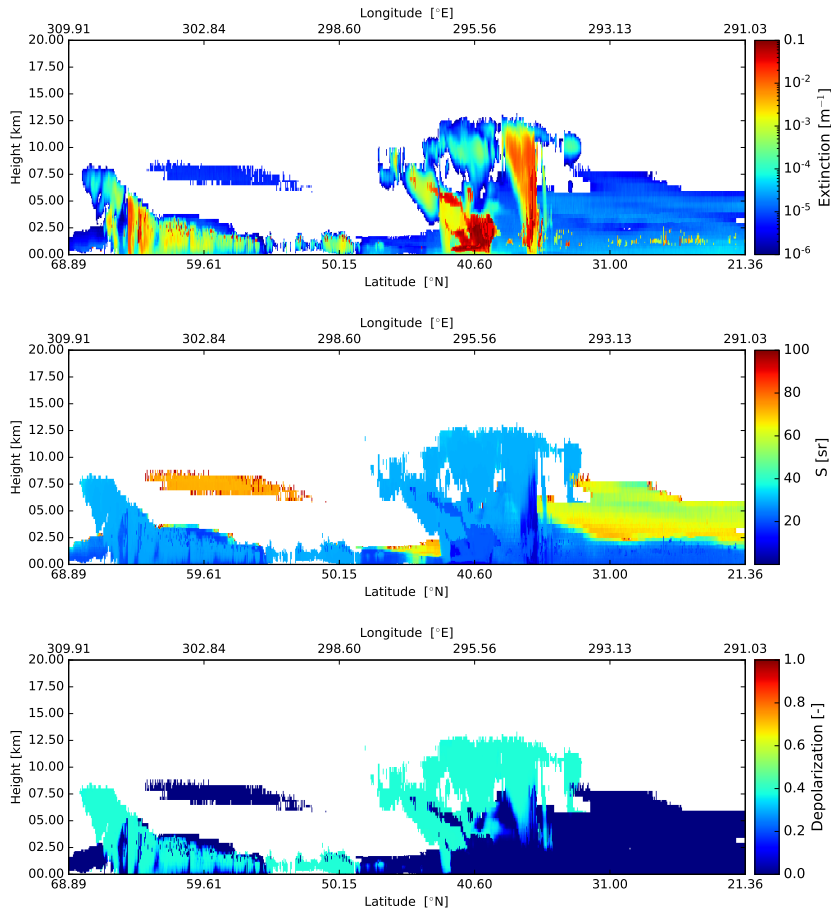
**Figure 10.** Cross-sections of the extinction at 355 nm, lidar extinction-to-backscatter ratio, and linear depolarization ratio, following the simulated EarthCARE orbit for the Halifax scene

reflectances divided by their sum for the two indicated channels) fields corresponding to the radiances shown in Fig. 14 are shown in Fig. 15.

The simulated long-wave TOA brightness temperatures for the MSI long-wave channels are shown in Fig. 16. Here it can be cold (but low altitude) clouds tops north of $55°$ N are visible, as are the cold (but high) cloud top temperatures between $35°$N and $45°$N. The warm land and ocean temperatures are also apparent in the mainly cloud-free region south of about $35°$ N. A close-up of the channel-1 radiances for the Halifax scene between $32.5°$N and $37.5°$N can be seen within Fig. 17

530
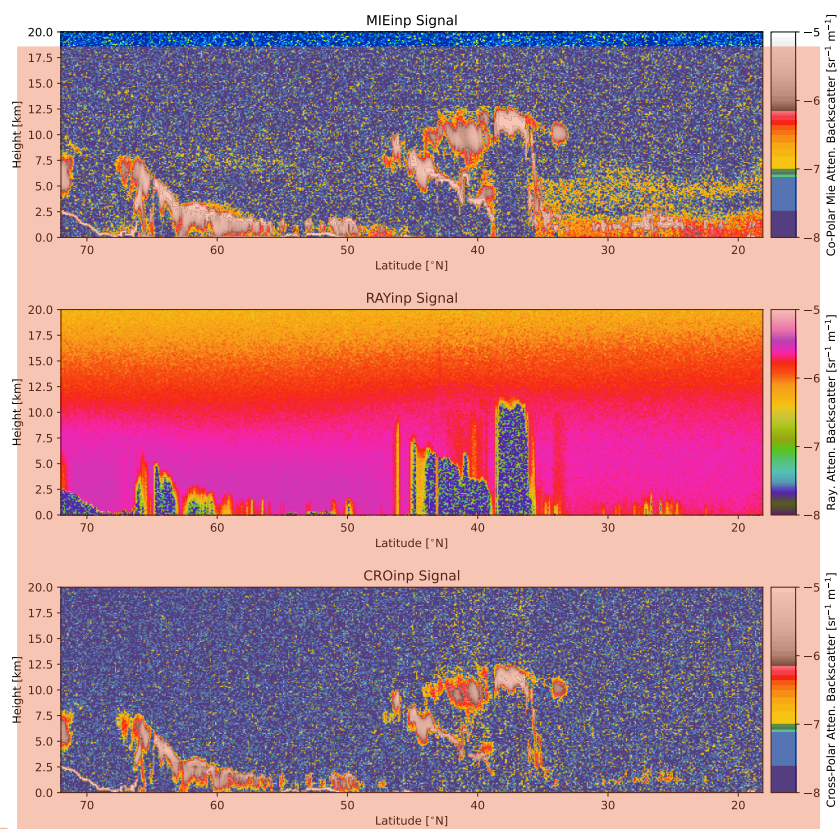
**Figure 11.** Simulated Mie channel, Rayleigh channel, and Cross-Polar ATLID attenuated backscatters for the Halifax scene.
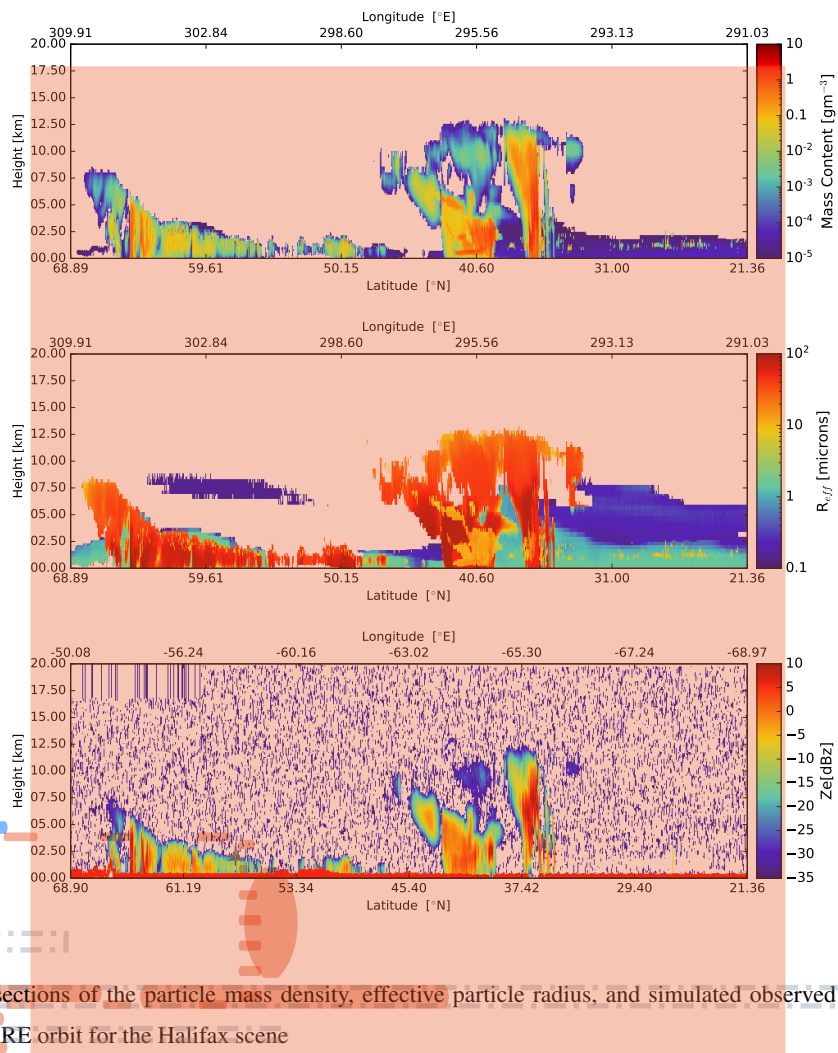
**Figure 12.** Cross-sections of the particle mass density, effective particle radius, and simulated observed Radar reflectivity, following the simulated EarthCARE orbit for the Halifax scene
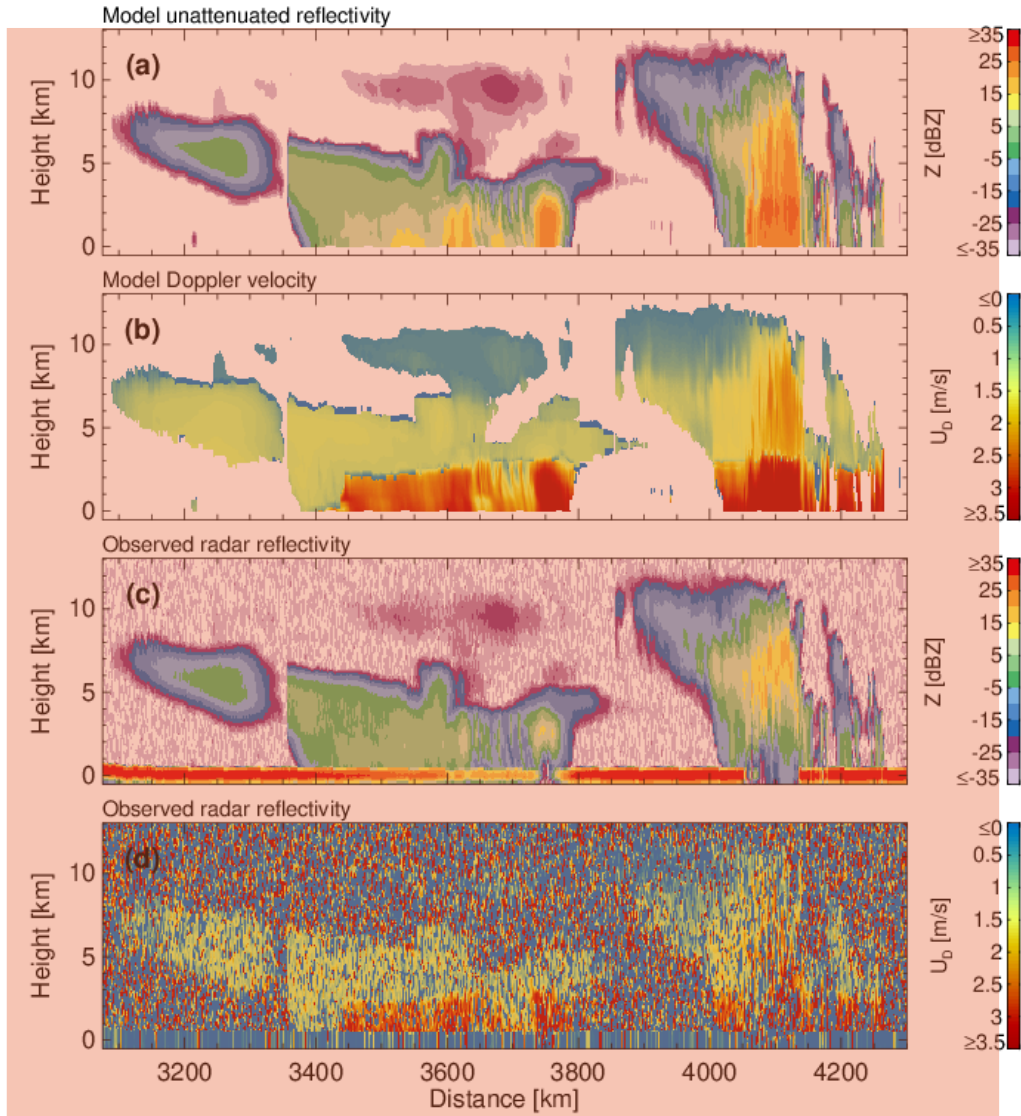
**Figure 13.** (a) GEM unattenuated 94-GHz radar reflectivity factor at the model resolution, (b) GEM sedimentation Doppler velocity at the model resolution, (c) CPR raw radar reflectivity factor at the CPR resolution with surface echo and (d) CPR raw Doppler velocity measurements with no correction applied at the CPR resolution with surface echo
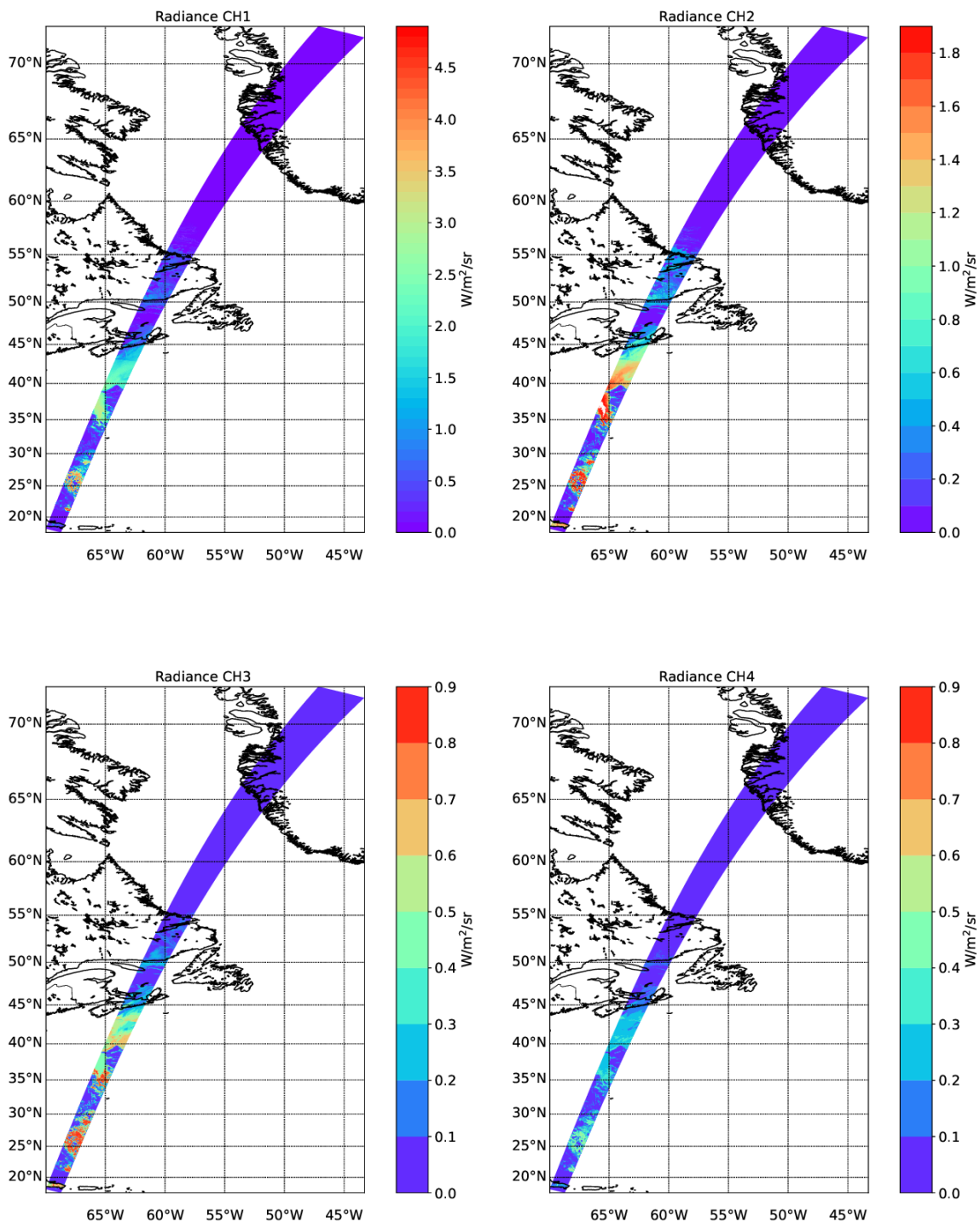
**Figure 14.** Simulated TOA radiances for the 4 SW EarthCARE MSI channels for the Halifax scene.
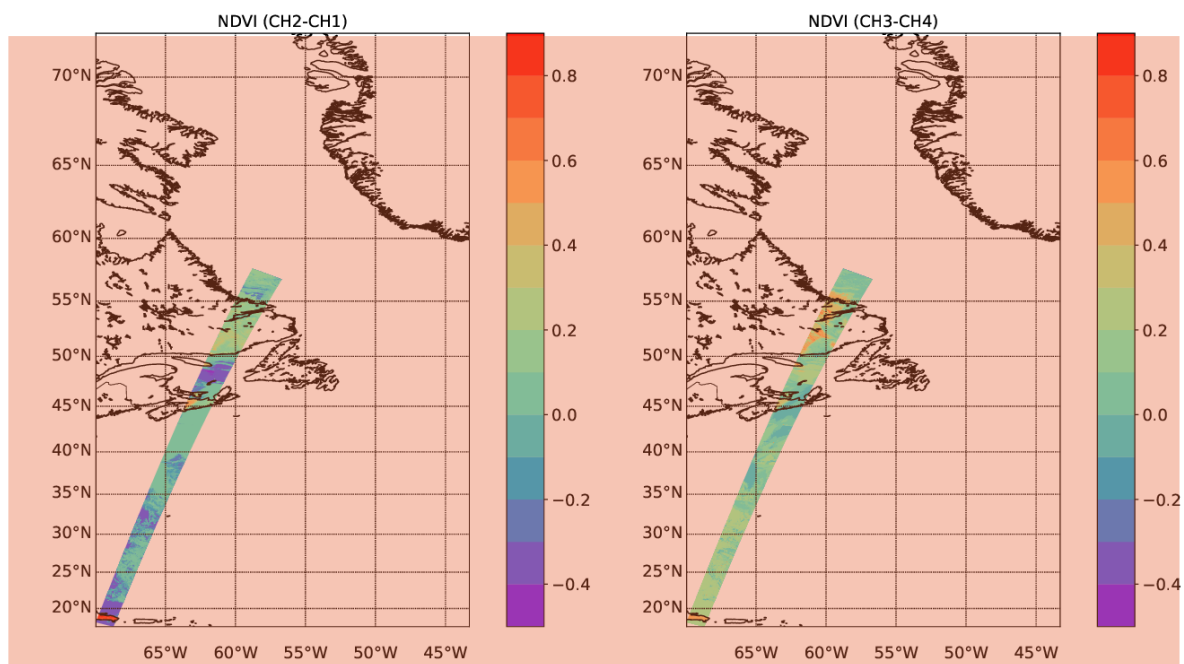
**Figure 15.** NDVI values corresponding to the simulated radiances shown in Fig. 14
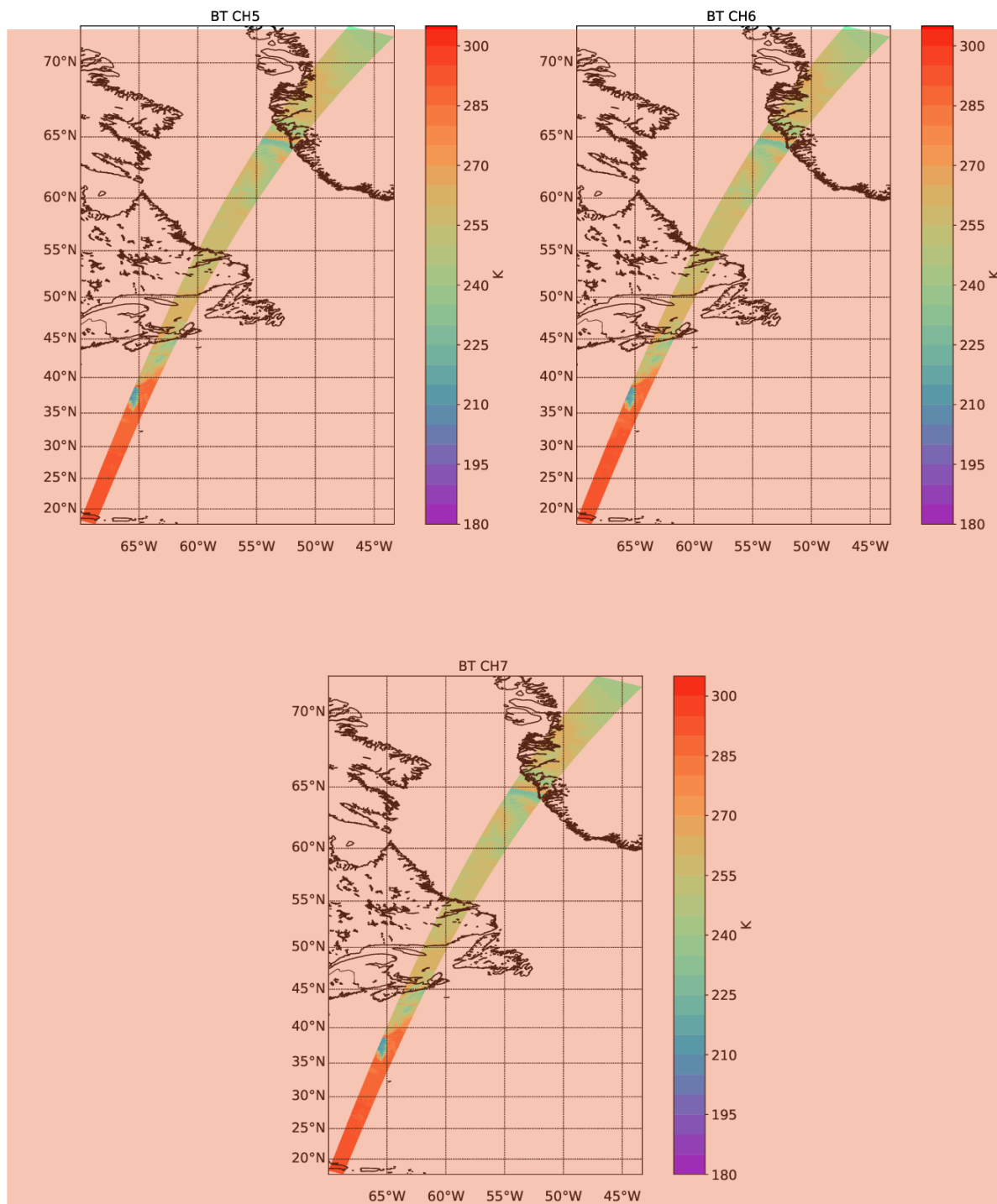
**Figure 16.** Simulated TOA brightness-temperatures for the three LW EarthCARE MSI channels for the Halifax scene.
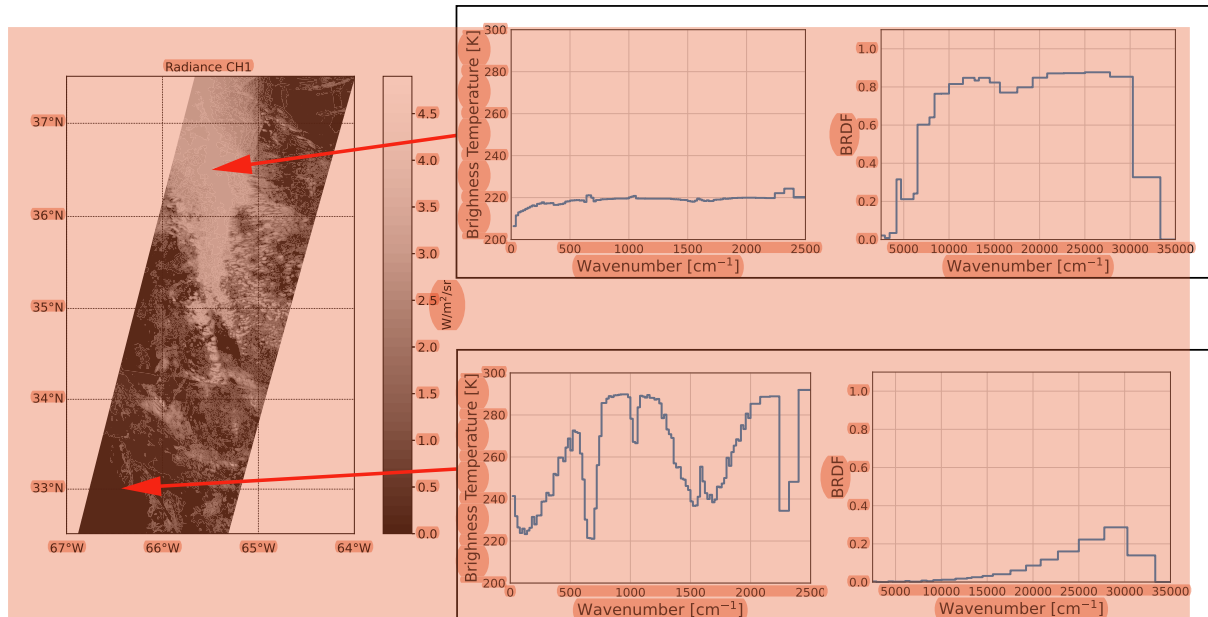
**Figure 17.** Left Panel: TOA channel 1 radiances for the Halifax scene between 32.5°N and 37.5°N. Right-Top: LW brightness temperatures and SW BRDFs as a function of wavenumber corresponding to a high altitude thick cloud. Right-Bottom: LW brightness temperatures and SW BRDFs for each BBR channel corresponding to cloud-free conditions over ocean.

## 3.2 Baja

The Baja scene is the 2nd GEM derived scene. This scene has a lot of topographical variation compared to the Halifax scene
535 and contains large regions of thinly distributed aerosols. In addition, high level ice clouds are present South of 35°N. Near the center of the scene, above the Rocky mountains, extended regions of optically thick ice and water clouds are present.

The extinction, lidar-radio and depolarization model-truth fields for the nadir track are shown in Fig. 18. The lidar observed attenuated backscatter fields are shown in Fig. 19 while the particle mass-density, effective radius, equivalent radar reflectivity are shown in Fig. 20.

540 The GEM model and CPR simulated signals for selected region of the Baja scene are shown in 21. Due to the large horizontal extend of the scene, we focus on the $[45°; 55°N]$ simulated region that covers the northern part of the GEM simulation over the Rockies which includes ice clouds. The top two panels shows the unattenuated 94-GHz radar reflectivity and the reflectivity weighted hydrometeor sedimentation velocity at the GEM grid resolution. The ice clouds are characterized by low radar reflectivity and Doppler velocities that increase towards the hydrometeor layer base. Due to their low radar reflectivity, a
545 significant fraction of the EarthCARE CPR echoes are close to its detection limit (-35 dBZ). The low radar reflectivities contribute to the noisiness of the CPR raw Doppler velocities.

The simulated short-wave TOA radiances for the MSI short-wave channels are shown in Fig. 22. Here north of about 55°N snow and ice surfaces are present leading to high channel 1 and 2 radiances even when clear-sky is present. The snow-ice surfaces also stand out in the CH3-CH4 NDVI images (Fig. 23).

550 The simulated long-wave TOA brightness temperatures for the MSI long-wave channels are shown in Fig. 24. Here there are cold (but low altitude) clouds tops and surfaces north of 55° N visible, as are the cold (but high) cloud top temperatures around 45°N. The warm land and ocean temperatures are also apparent in the mainly cloud-free region south of about 40° N.
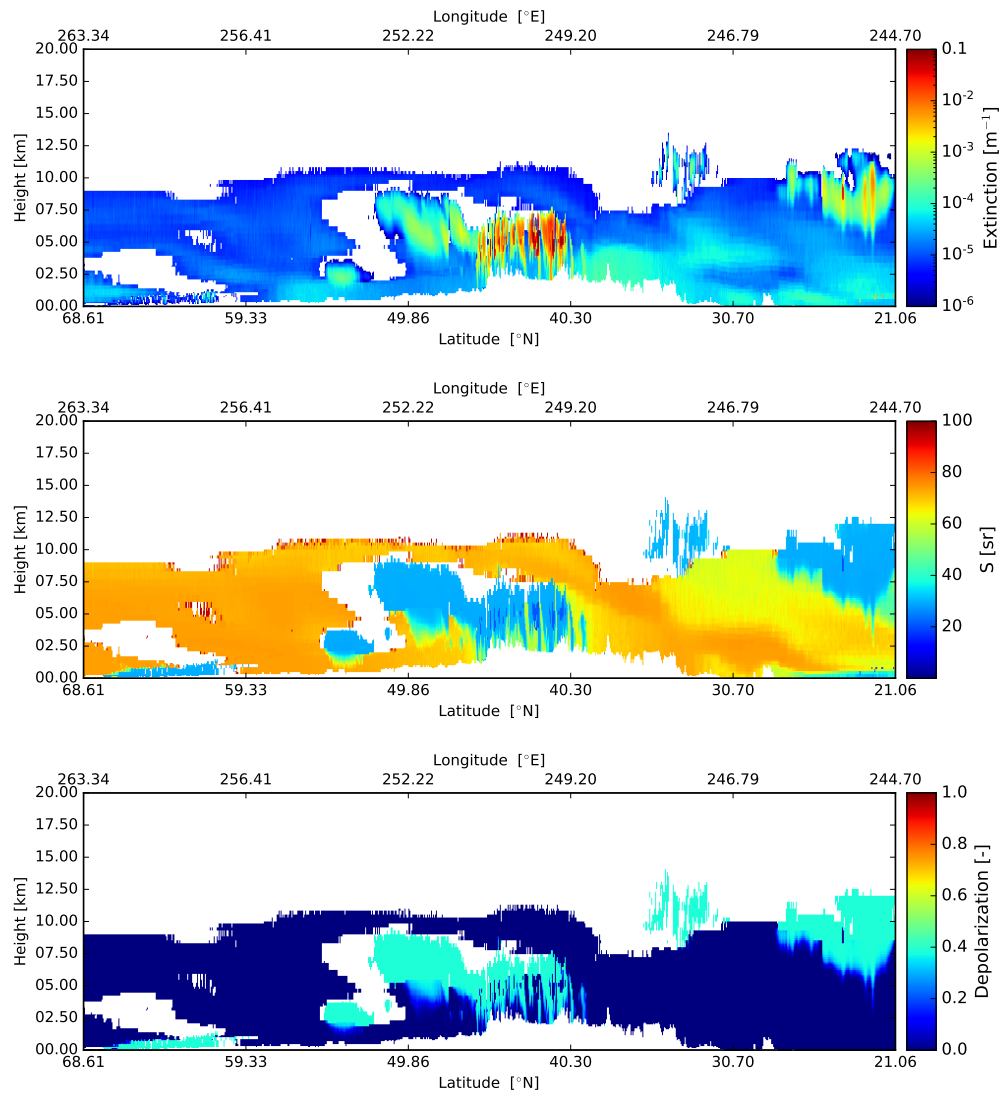
**Figure 18.** Cross-sections of the extinction at 355 nm, lidar extinction-to-backscatter ratio, and linear depolarization ratio, following the simulated EarthCARE orbit for the Baja scene
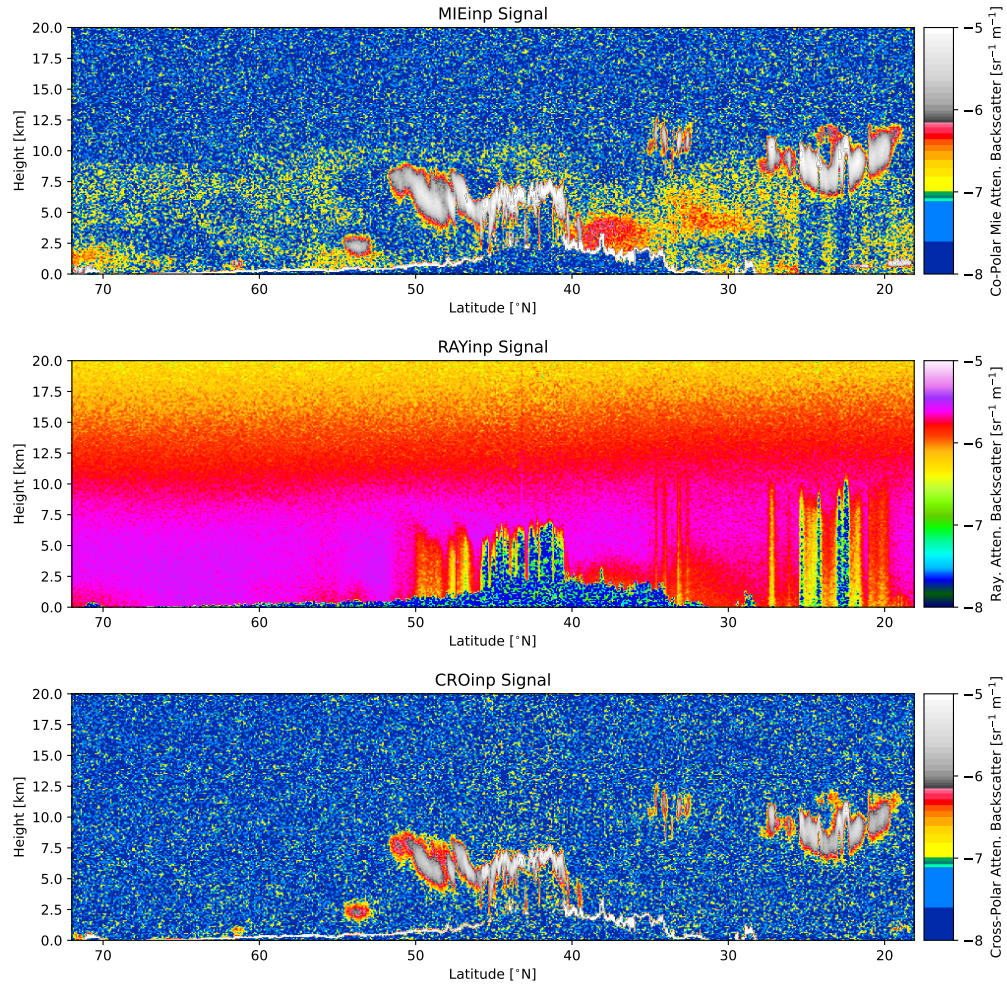
**Figure 19.** Simulated Mie channel, Rayleigh channel, and Cross-Polar ATLID attenuated backscatters for the Baja scene.
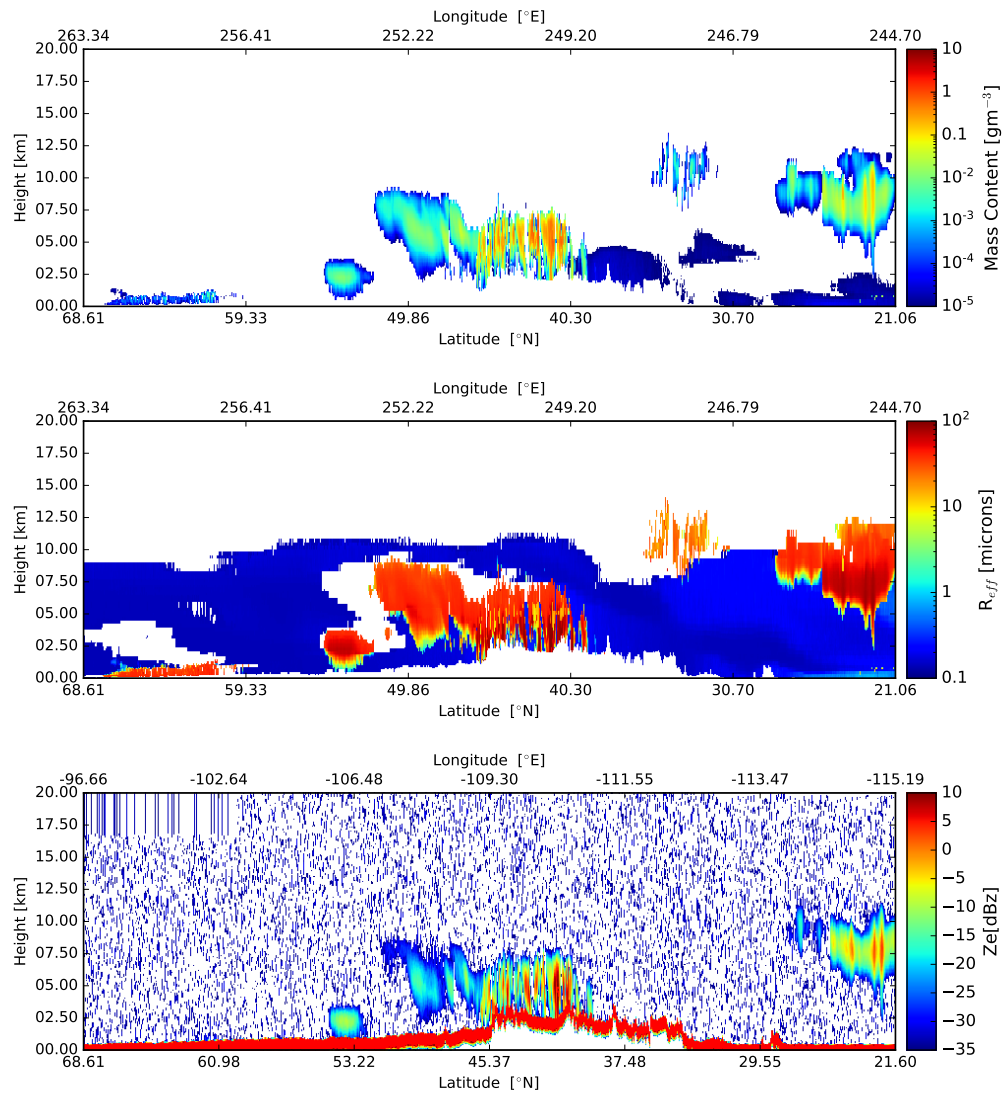
**Figure 20.** Cross-sections of the particle mass density, and effective particle radius, and observed Radar reflectivity, following the simulated EarthCARE orbit for the Baja scene.
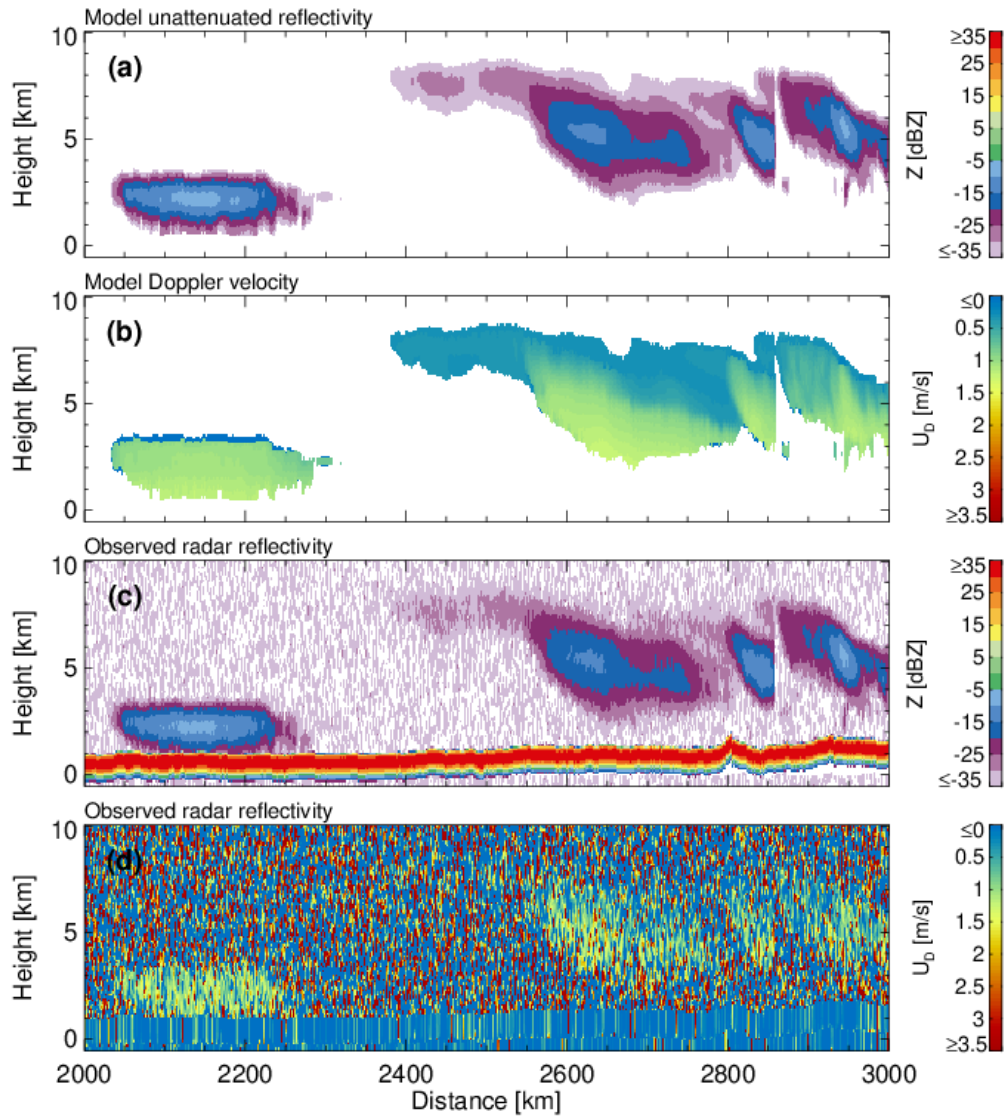
**Figure 21.** (a) GEM unattenuated 94-GHz radar reflectivity factor at the model resolution, (b) GEM sedimentation Doppler velocity at the model resolution, (c) CPR raw radar reflectivity factor at the CPR resolution with surface echo and (d) CPR raw Doppler velocity measurements with no correction applied at the CPR resolution with surface echo
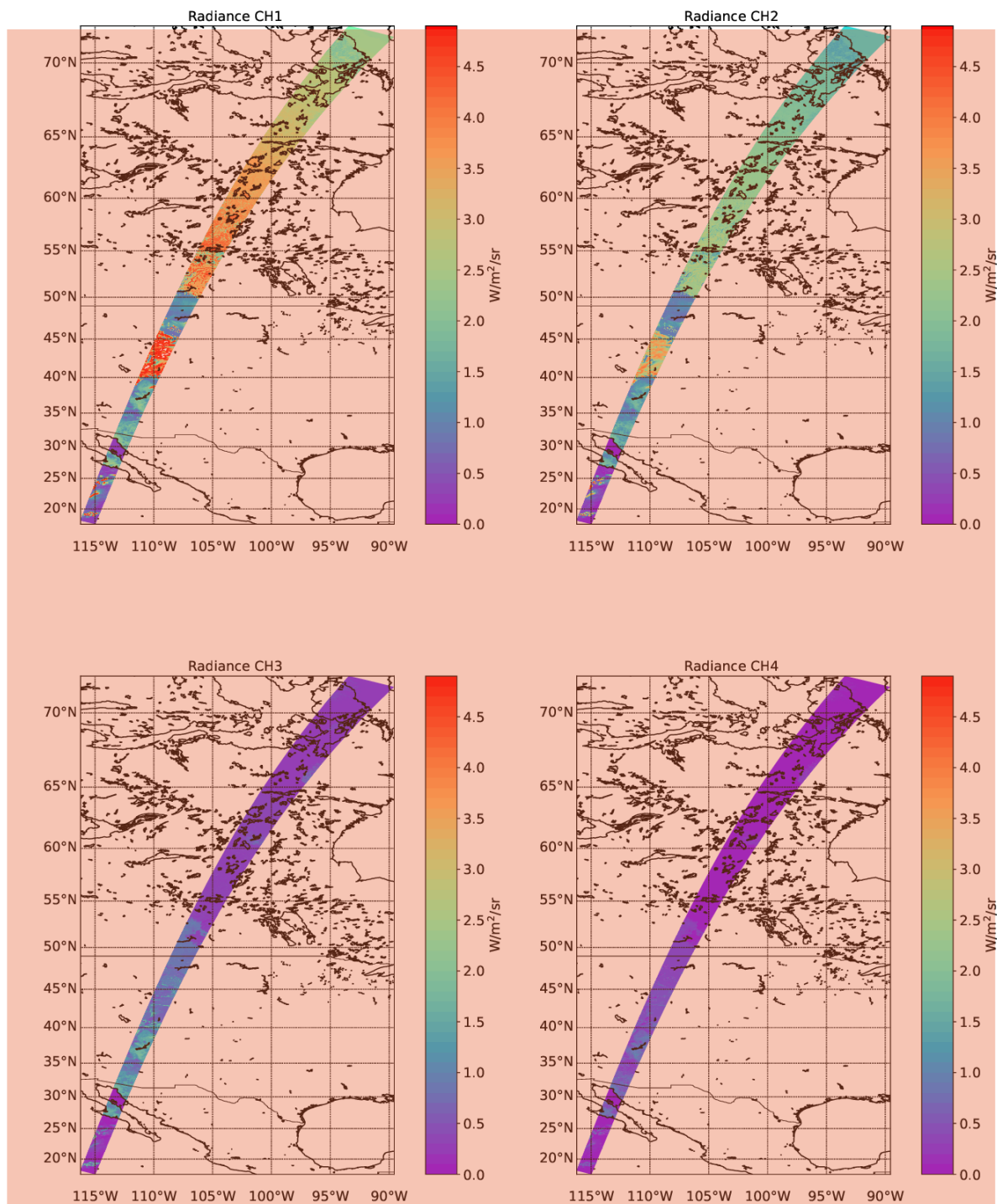
**Figure 22.** Simulated TOA radiances for the 4 SW EarthCARE MSI channels for the Baja scene.
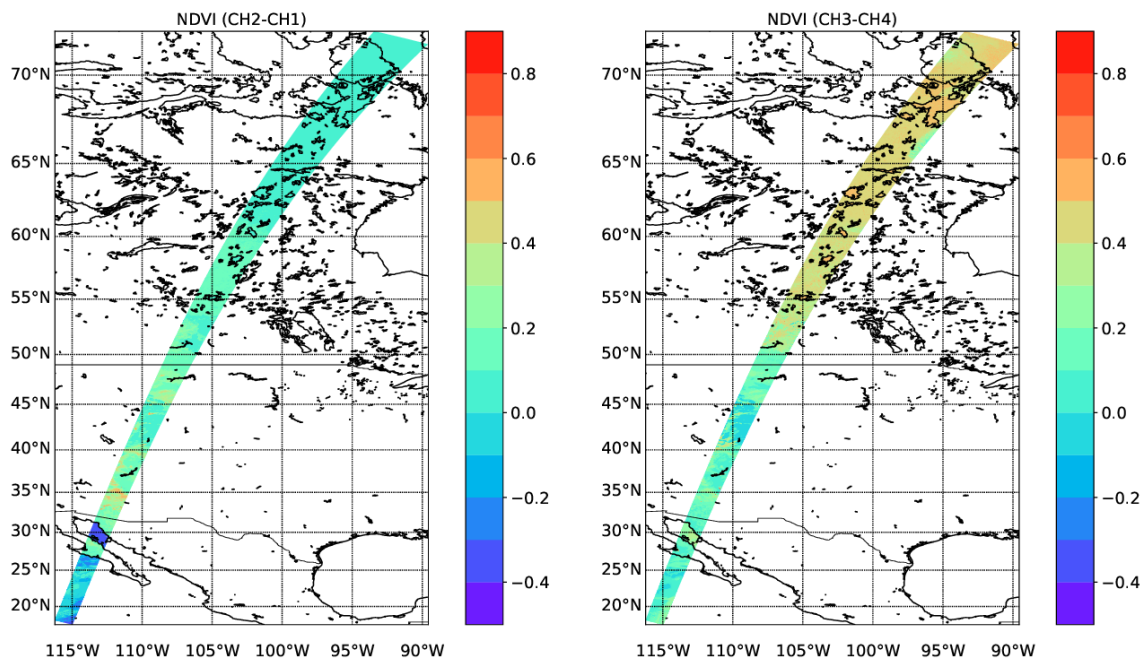
**Figure 23.** NDVI values corresponding to the simulated radiances shown in Fig. 22
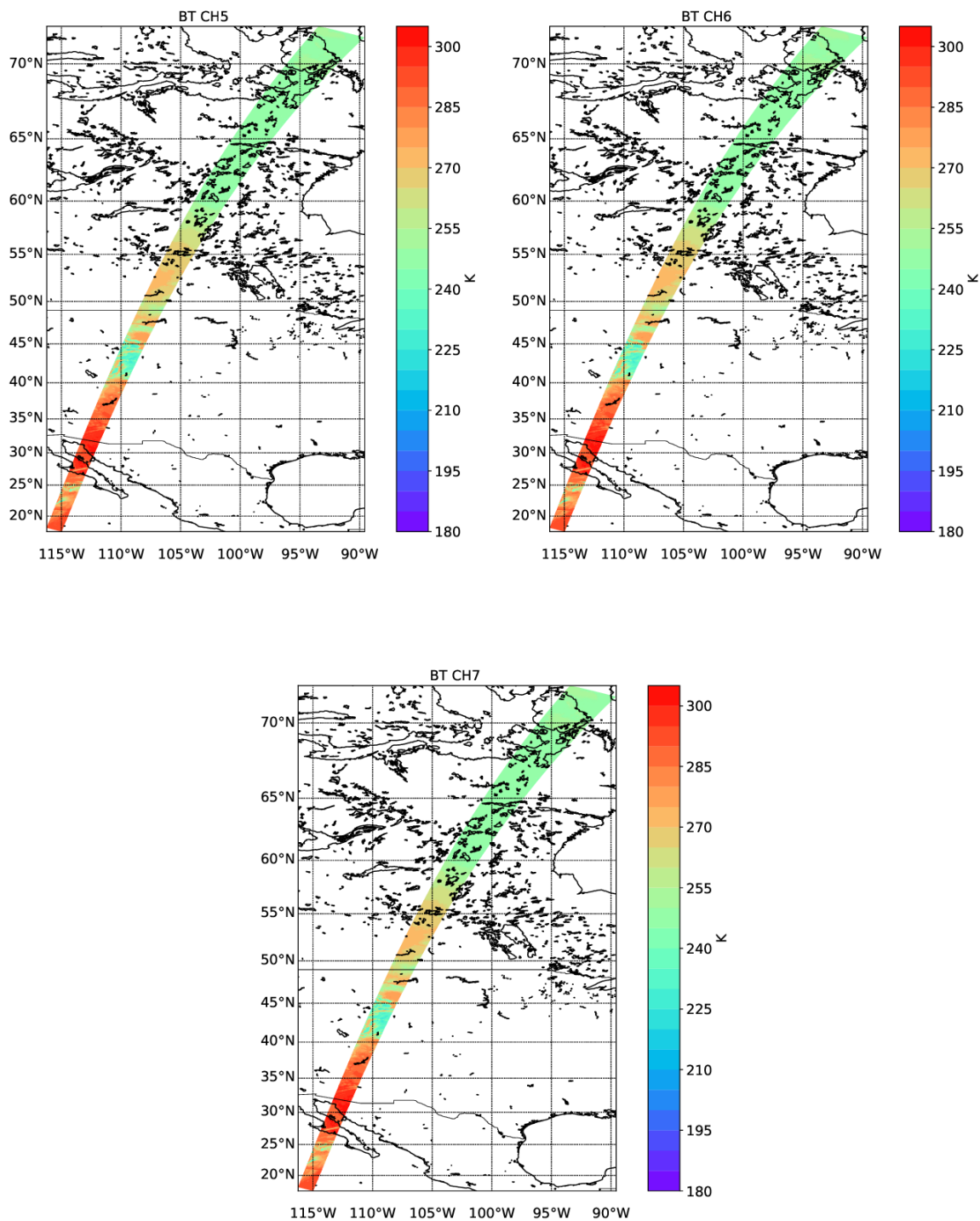
**Figure 24.** Simulated TOA brightness-temperatures for the 3 LW EarthCARE MSI channels for the Baja scene.

## 3.3 Hawaii

The 3rd GEM based scene is the Hawaii scene, situated almost completely over ocean, where the nadir track is completely over ocean while a few of the smaller Hawaiian islands are within the MSI track. This scene exhibits areas of clear-sky, upper level cirrus and a tropical convective system near the center of the scene. The extinction, lidar-ratio and depolarization model-truth fields for the nadir track are shown in Fig. 25. The lidar observed attenuated backscatter fields are shown in Fig. 26 while the particle mass-density, effective radius, equivalent radar reflectivity are shown in Fig. 27.
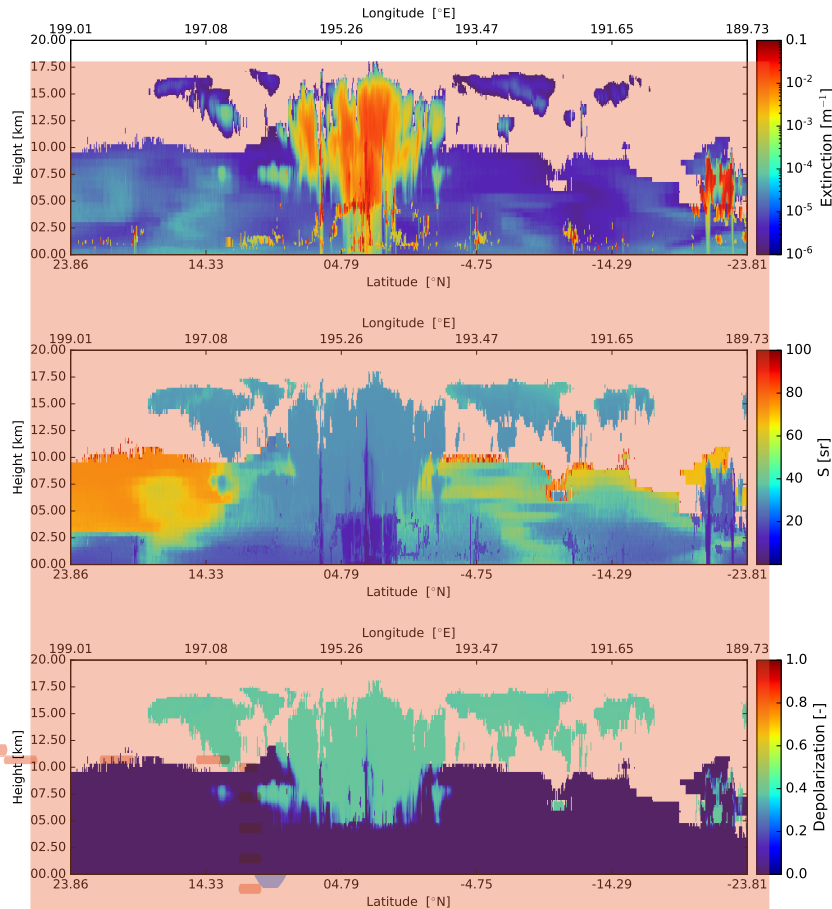
**Figure 25.** Cross-sections of the extinction at 355 nm, lidar extinction-to-backscatter ratio, and linear depolarization ratio, following the simulated EarthCARE orbit for the Hawaii scene

The GEM model and CPR simulated signals for the Hawaii scene are shown in Fig. 28. Due to the large horizontal extend of the scene, we focus on the center of the scene where a large tropical convective/stratiform precipitating system was simulated. The top two panels shows the unattenuated 94-GHz radar reflectivity and the reflectivity weighted hydrometeor sedimentation
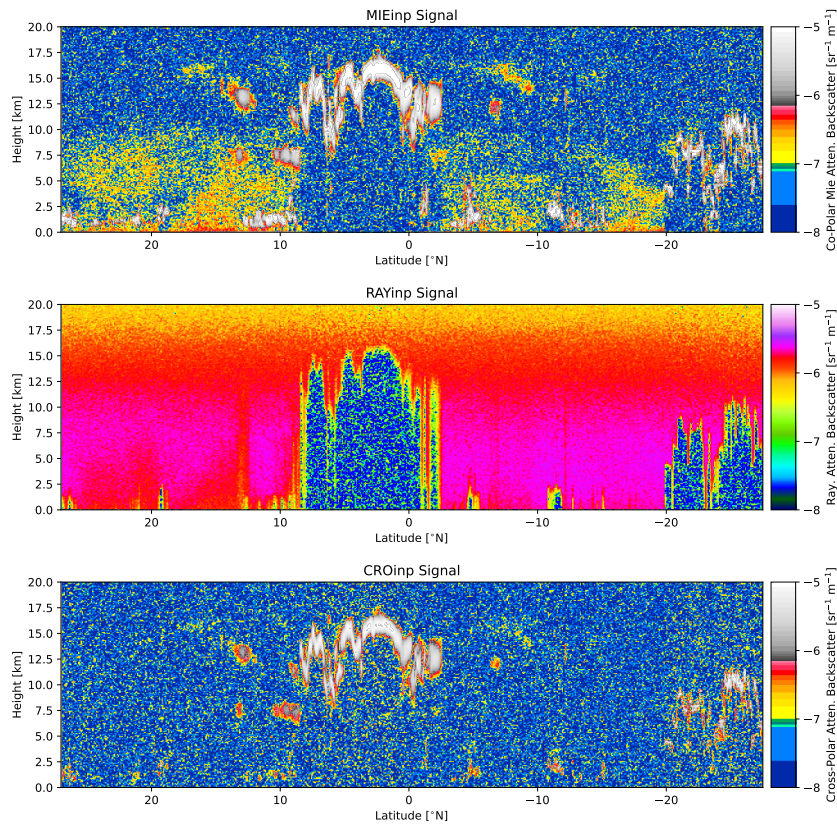
**Figure 26.** Simulated Mie channel, Rayleigh channel, and Cross-Polar ATLID attenuated backscatters for the Hawaii scene.

velocity at the GEM grid resolution. The widespread stratiform precipitation is extensive covering over 500 km and the precipitation system reaches tops of 16km. Embedded convective cells are simulated at 2375 and 2750 km along track distance. The lower two panels show the raw (uncorrected) CPR radar reflectivity and mean Doppler velocity measurements. The Earth-

565    CARE CPR has sufficient sensitivity to detect the hydrometeor layers except for weak reflectivity echoes near the highest cloud tops. The strong 94-GHz attenuation by hydrometeors results to significant attenuation in the rain layer and complete loss of the surface echo return in the two embedded convective cells. Significant multiple scattering is simulated at 2750 km, as seen by the fainted CPR echoes that fill the surface echo gap around 4100 km due to multiple scattering (Battaglia et al., 2010). As expected, the CPR raw Doppler velocity field (500 m along track integration) is noisy (Kollias et al., 2014, 2022a). Despite

570    their noisiness, the CPR Doppler velocities reproduce the main features of the GEM model Doppler velocities, namely, the transition from solid to liquid hydrometeors and the low sedimentation Doppler velocities in the upper cloud levels.
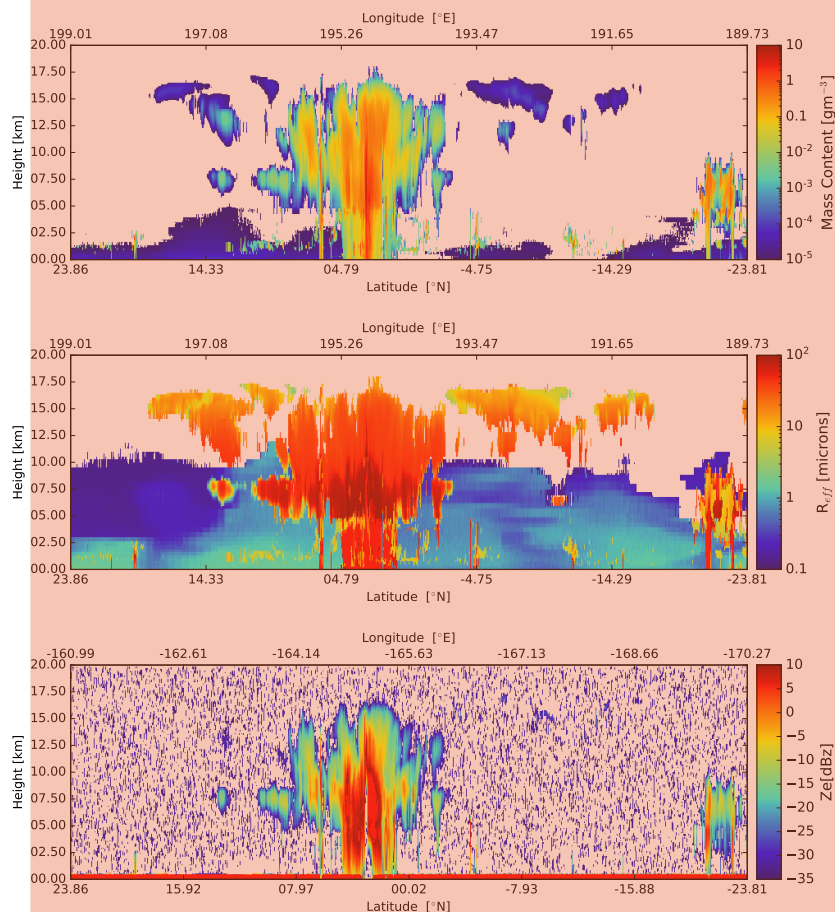
**Figure 27.** Cross-sections of the particle mass density, and effective particle radius, and observed Radar reflectivity, following the simulated EarthCARE orbit for the Hawaii scene

The simulated short-wave TOA radiances for the MSI short-wave channels are shown in Fig. 29. Here the cloud features are all clearly visible against the low albedo ocean. The NDVI fields are shown in Fig. 30. The high-values in the right-panel correspond to one of the few land-areas (the island of Niihau) within the MSI swath.

575     The simulated long-wave TOA brightness temperatures for the MSI long-wave channels are shown in Fig. 31. Here the very cold cloud tops near the center of the scene are especially prominent.

### 3.4 Halifax Aerosol Scene

As a last test scene based on GEM and CAMS inputs and used in the processor development within this special issue, a mainly aerosol scene was built from a sub-section of the Halifax scene. For this "Halifax-aerosol" scene the region south of 36°N was
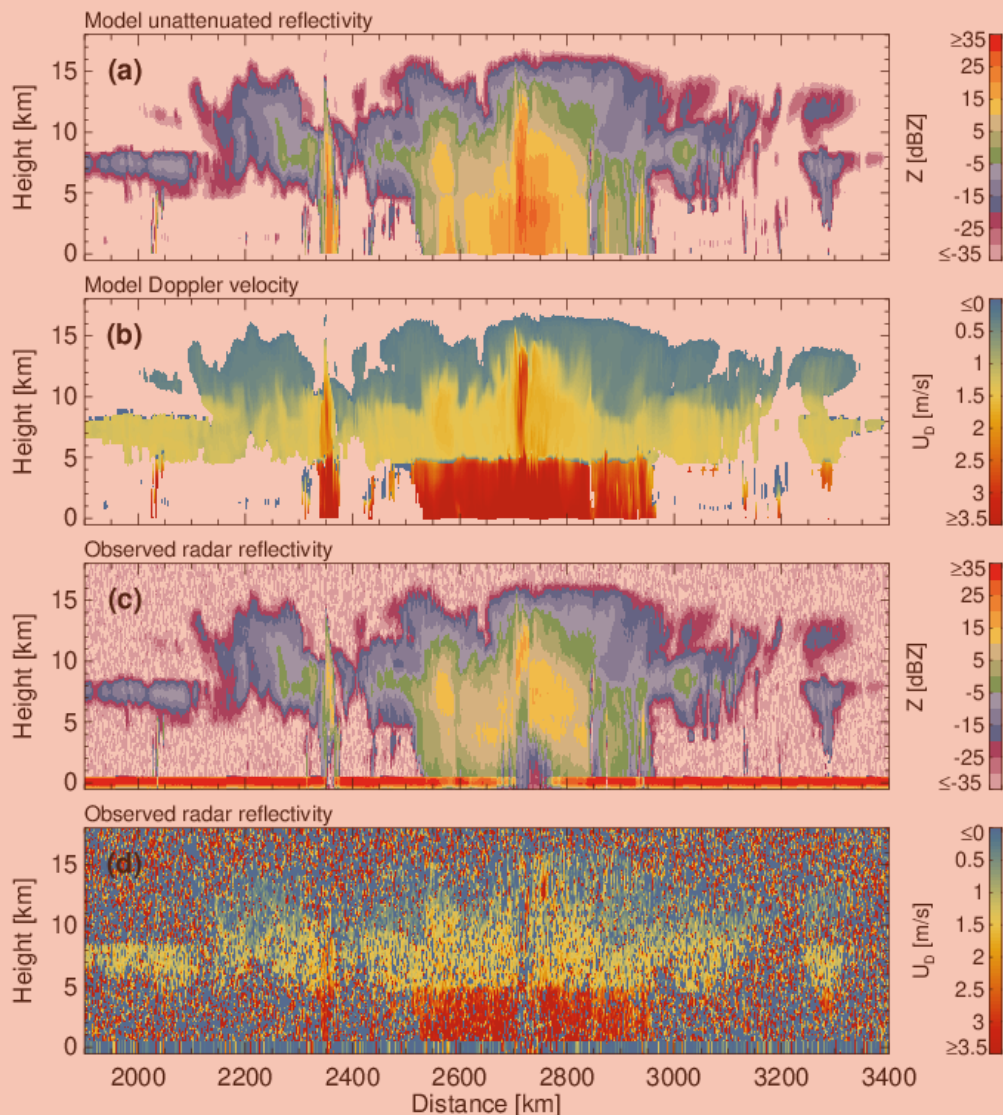
**Figure 28.** (a) GEM unattenuated 94-GHz radar reflectivity factor at the model resolution, (b) GEM sedimentation Doppler velocity at the model resolution, (c) CPR raw radar reflectivity factor at the CPR resolution with surface echo and (d) CPR raw Doppler velocity measurements with no correction applied at the CPR resolution with surface echo.

580  used, however, all liquid clouds and aerosol types were removed except for the coarse-mode non-absorbing aerosol (Sea-Salt). The remaining aerosol particle density was scaled by a factor of 2 to increase the total optical thickness and detectability by the MSI instrument. The resulting scene contains ice cloud north of about 33 °N and an aerosol rich boundary-layer marine aerosol layer limited to an altitude of about 2.5km with a very tenuous layer above. The lidar observed attenuated backscatter fields are shown in Fig. 32, while the MSI images are shown in Figs. 33–35.
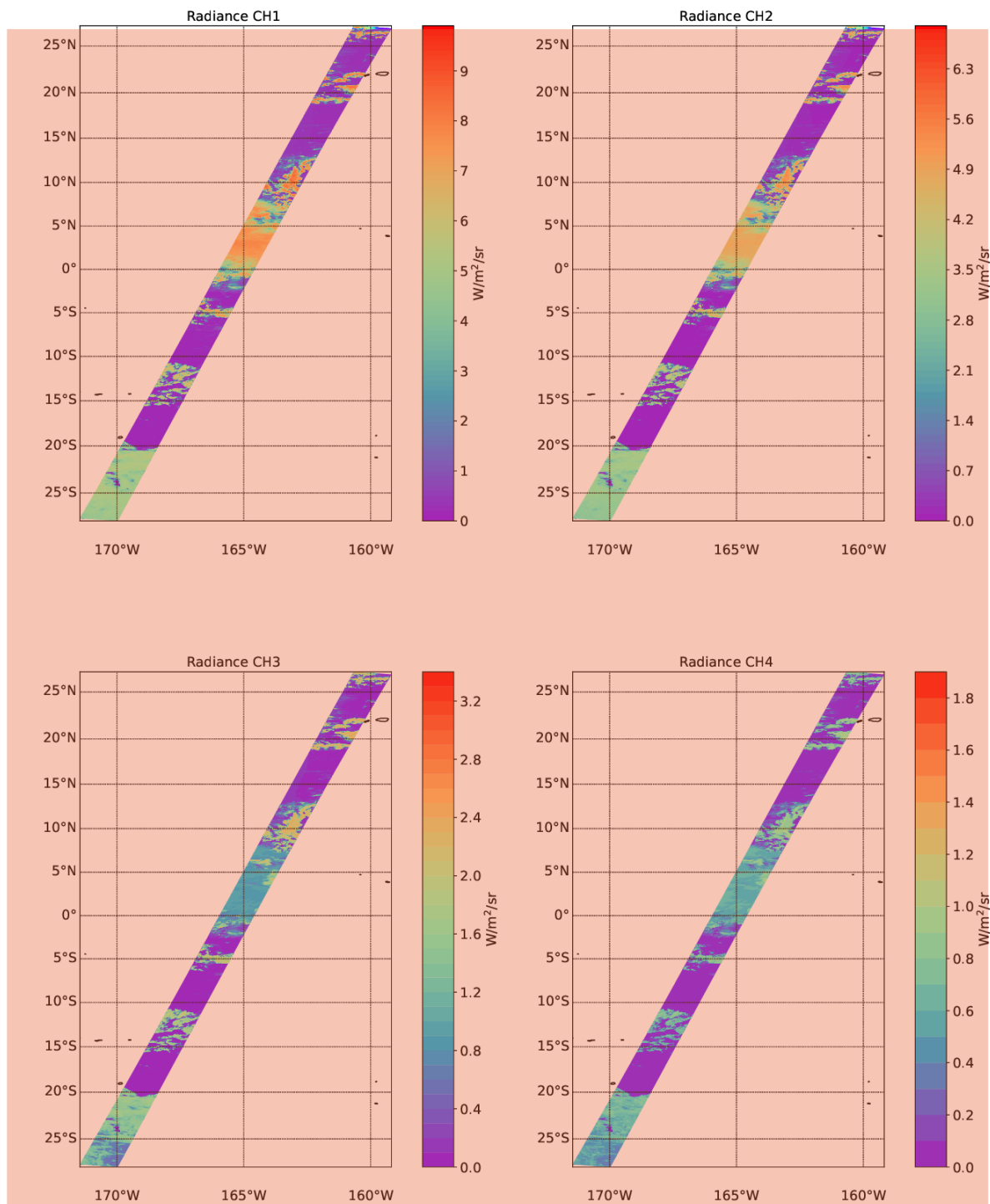
**Figure 29.** Simulated TOA radiances for the 4 SW EarthCARE MSI channels for the Hawaii scene.
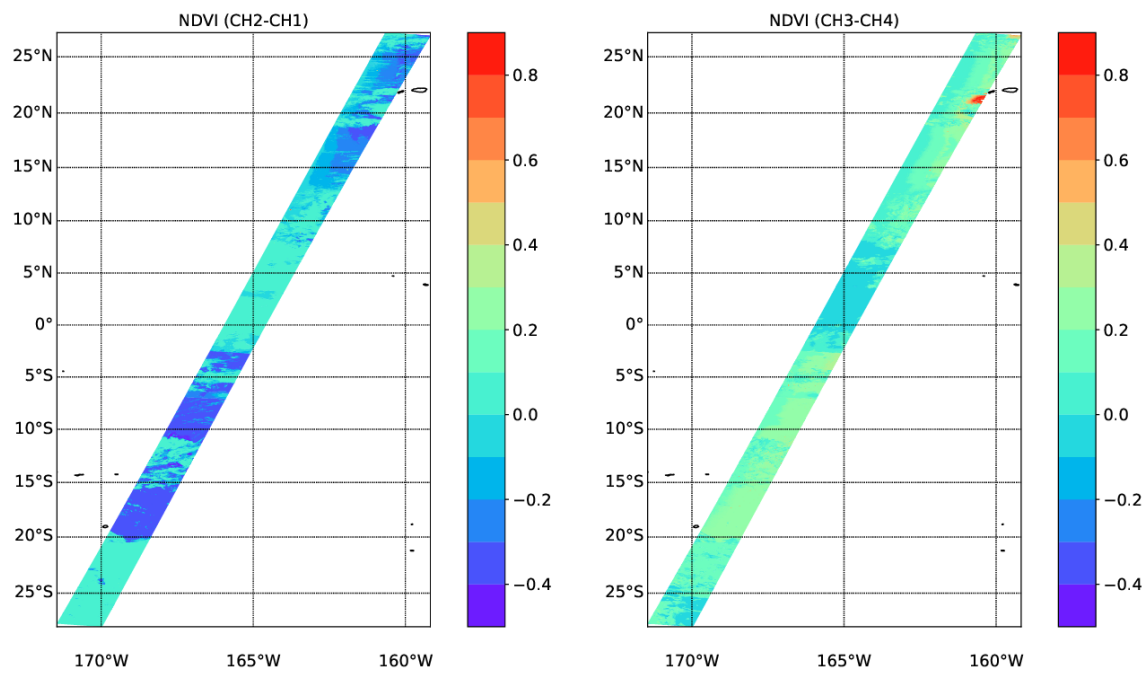
**Figure 30.** NDVI values corresponding to the simulated radiances shown in Fig. 29
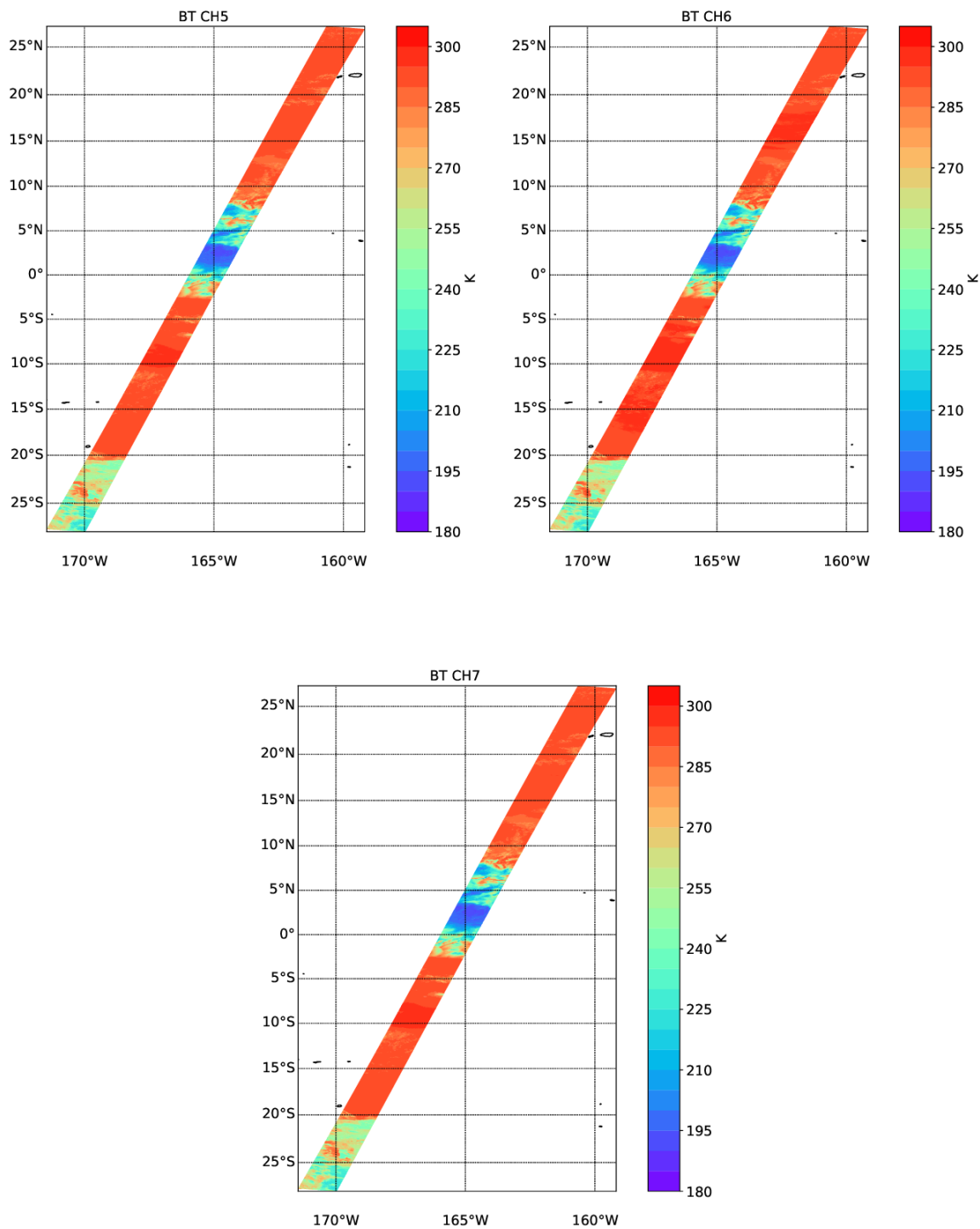
**Figure 31.** Simulated TOA brightness-temperatures for the 3 LW EarthCARE MSI channels for the Hawaii scene.
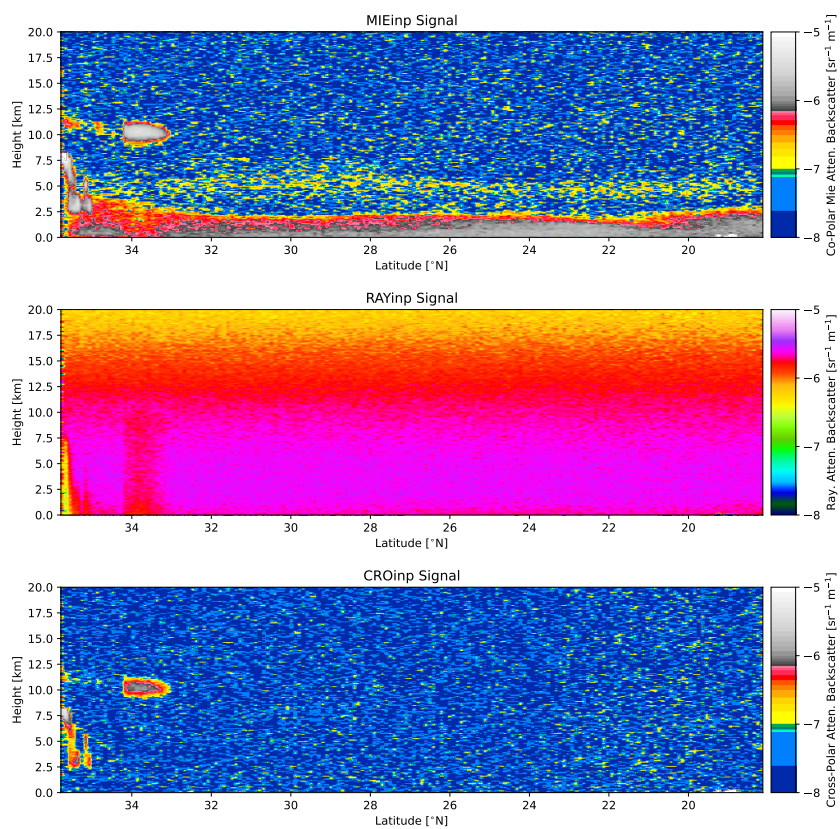
**Figure 32.** Simulated Mie channel, Rayleigh channel, and Cross-Polar ATLID attenuated backscatters for the Halifax-Aerosol scene.
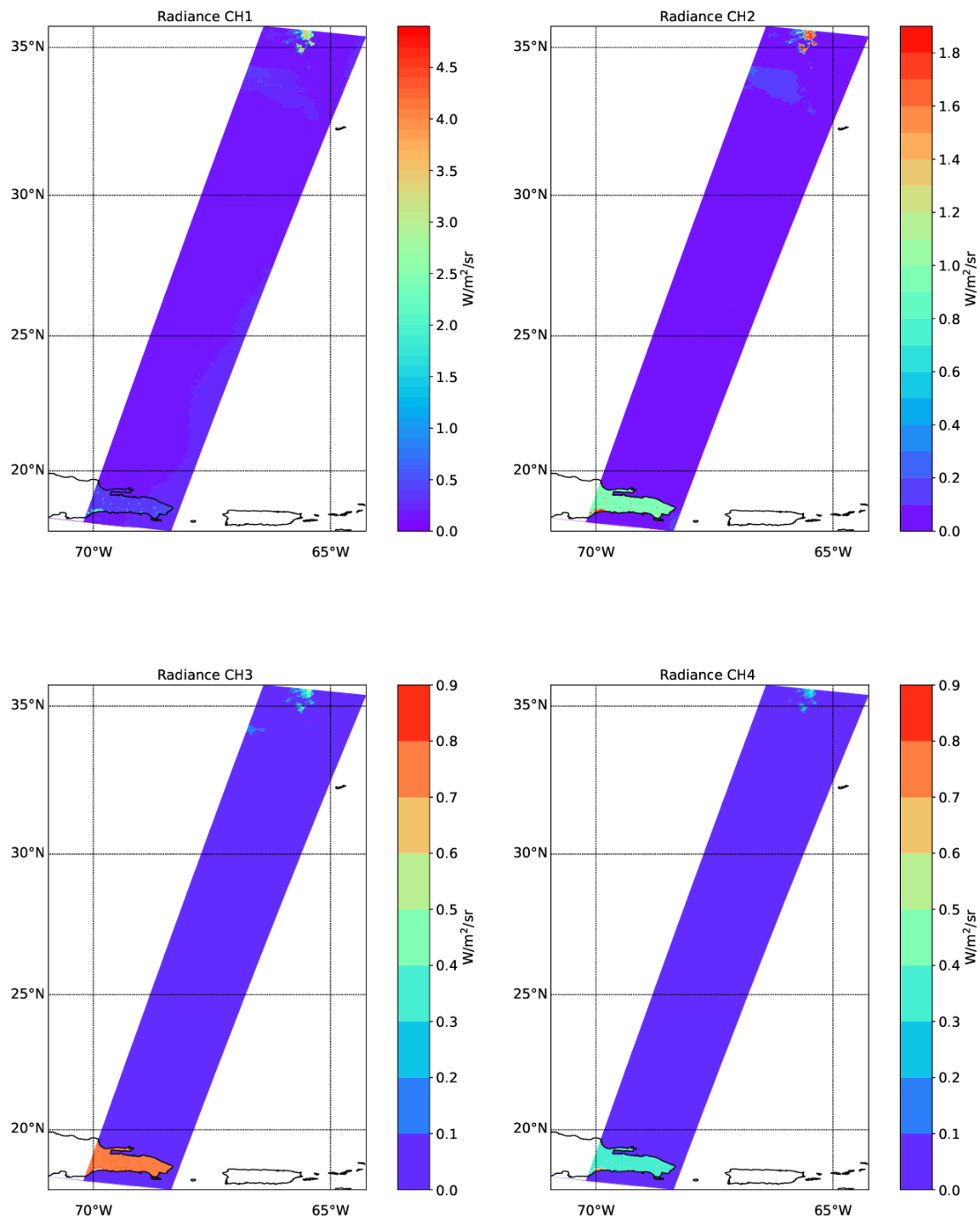
**Figure 33.** Simulated TOA radiances for the 4 SW EarthCARE MSI channels for the Halifax-Aerosol scene.
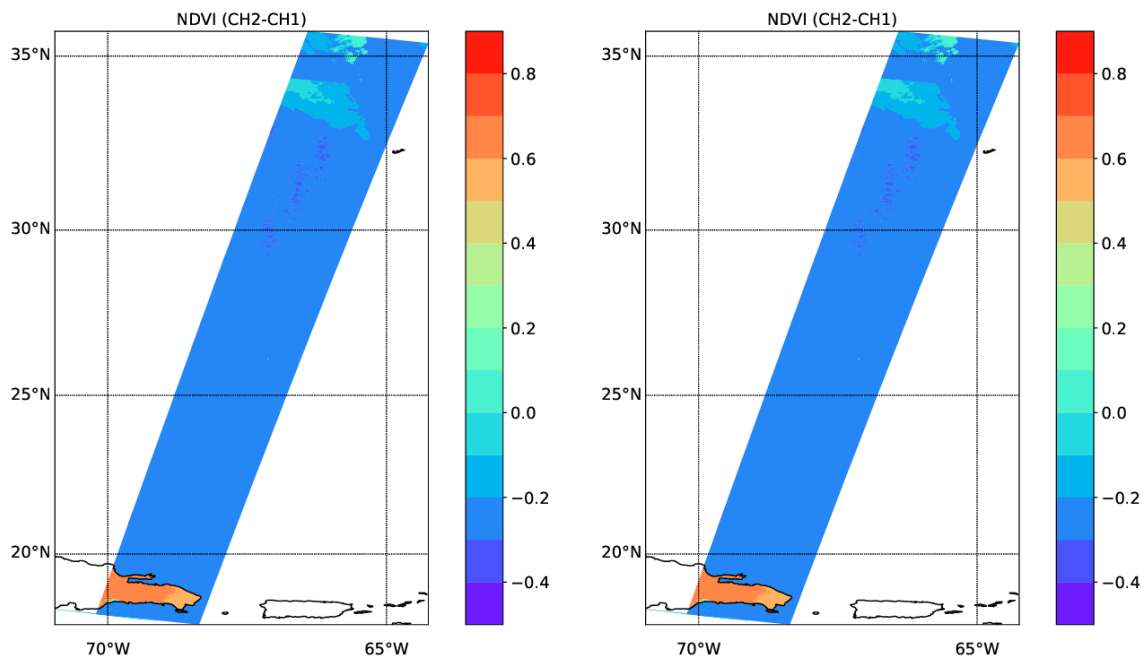
**Figure 34.** NDVI values corresponding to the simulated radiances shown in Fig. 33
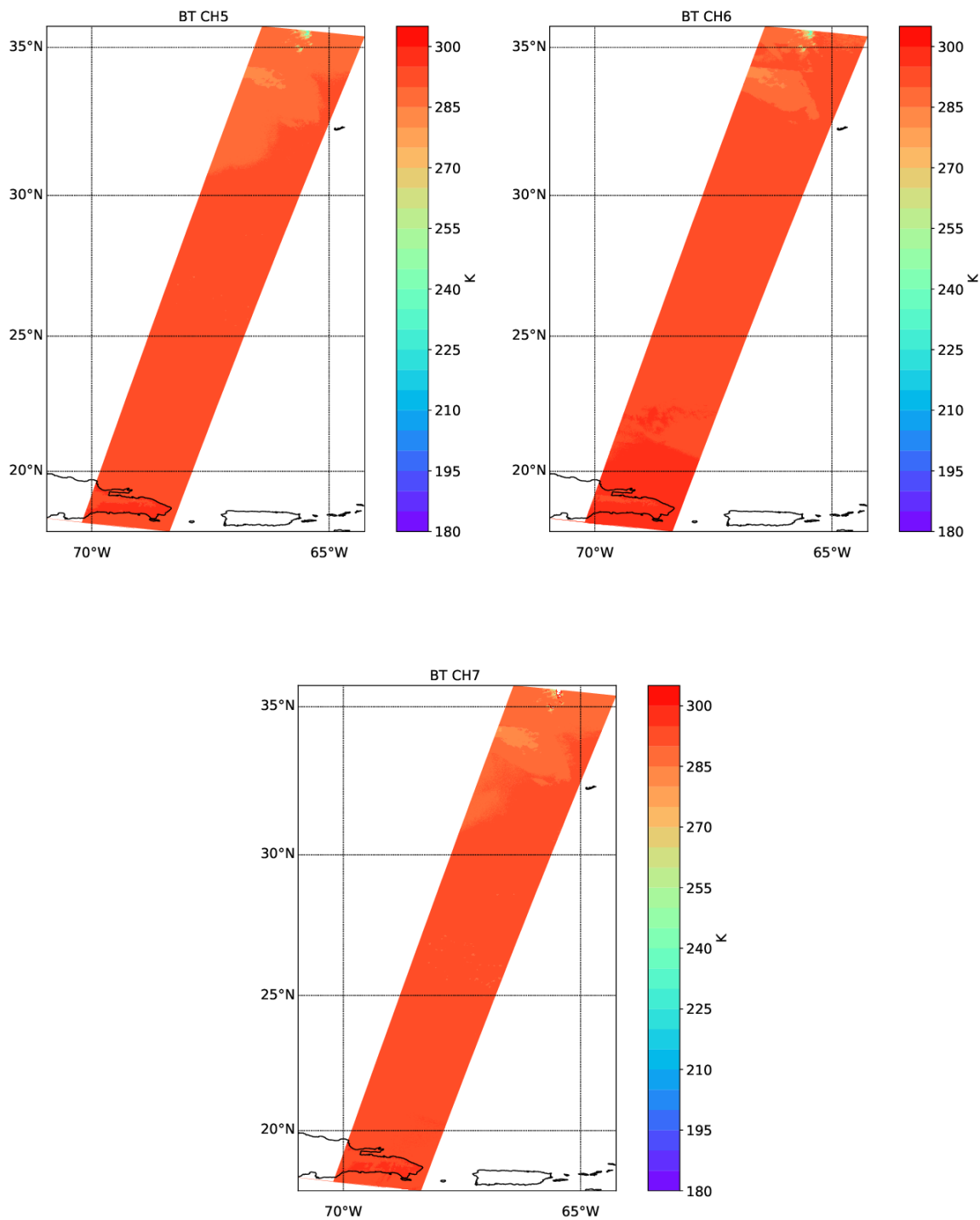
**Figure 35.** Simulated TOA brightness-temperatures for the 3 LW EarthCARE MSI channels for the Hawaii scene.

## 4 Conclusion

In this paper the main testing and development scenes used for EarthCARE L2 algorithm scientific and technical development and testing have been presented. The model data used as the basis for the scenes is discussed in a companion paper (Qu et al., 2022). The focus of this paper has been to present the modelled L1 signals for each of the EarthCARE instruments and to document the methods used in their creation. The L1 signals together with the corresponding model-truth filed, have in turn, been used to develop a suite of single- and multi-instrument retrieval processors. These processors are described by various companion papers. Further insight into the expected precision, accuracy and coverage of the EarthCARE data can be found in an algorithm inter-comparison paper Mason et al. (2023).

Numerous other studies, beyond what has been presented within this special issue, could be conducted using this simulated data set. Post-launch, the simulations presented here will continue to play at role in the evolution of the various processors. Post-launch activities including e.g. investigating the effects of in-orbit EarthCARE instrument degradation may be conducted using the present simulations as a base. In addition, the simulations could be expanded to include other past instruments e.g. Calipso or future proposed lidars or radars. Such activities may be useful in understanding the relationship between e.g. the existing Calipso aerosol record and expected EarthCARE results.

The detail, realism, physical consistency and scale of the tests scenes developed for EarthCARE algorithm development, implementation, and testing comprise a unique effort. The creation of detailed realistic test scenes has involved considerable work but should be judged as time well spent. Not only have they served to develop new scientific inversion approaches, but have also proved very useful in terms of technical development. It is true that the success of any inversion procedure must be evaluated using real data, however, the ability to compare against a "model truth" is invaluable when constructing new inversion algorithms, both in a scientific sense and in a technical (e.g. debugging) sense.

When actual EarthCARE data is available, there will, no doubt, be surprises to be dealt with. The extensive algorithm development process, aided by the simulations, will help ensure that these unexpected issues will be handled efficiently.

*Data availability.* The EarthCARE Level-2 nadir model-truth data and the simulated L1 products discussed in this paper, are available from https://doi.org/10.5281/zenodo.7117115.

*Acknowledgements.* This work has been funded by ESA grants 22638/09/NL/CT (ATLAS), ESA ITT 1-7879/14/NL/CT (APRIL) and 4000134661/21/NL/AD (CARDINAL). We thank Tobias Wehr, Michael Eisinger and Anthony Illingworth for valuable discussions and their support for this work over many years. A special acknowledgement is in order for Tobias Wehr, ESA's EarthCARE Mission Scientist who unexpectedly passed away recently. Tobias was eagerly looking forward to EarthCARE's launch, a mission to which he had dedicated a considerable span of his career. His support for the science community, collaborative approach and enthusiasm for the mission science will not be forgotten.

615  *Competing interests.* The authors declare that they have no conflict of interest.

# References

Barker, H. and Li, J.: Accelerating Radiative Transfer Calculations for High-Resolution Atmospheric Models, Quarterly Journal of the Royal Meteorological Society, 145, https://doi.org/10.1002/qj.3543, 2019.

Barker, H. W. and Liu, D.: Inferring Optical Depth of Broken Clouds from Landsat Data, Journal of Climate, 8, 2620 – 2630, https://doi.org/https://doi.org/10.1175/1520-0442(1995)008<2620:IODOBC>2.0.CO;2, 1995.

Barker, H. W., Cole, J. N. S., Qu, Z., Villefranque, N., and Shephard, M.: Radiative closure assessment of retrieved cloud and aerosol properties for the EarthCARE mission: the ACMB-DF product, Atmospheric Measurement Techniques, to be submitted, 2023.

Battaglia, A. and Tanelli, S.: DOMUS: DOppler MUltiple-Scattering Simulator, IEEE Transactions on Geoscience and Remote Sensing, 49, 442–450, https://doi.org/10.1109/TGRS.2010.2052818, 2011.

Battaglia, A., Tanelli, S., Kobayashi, S., Zrnic, D., Hogan, R. J., and Simmer, C.: Multiple-scattering in radar systems: A review, Journal of Quantitative Spectroscopy and Radiative Transfer, 111, 917–947, https://doi.org/10.1016/j.jqsrt.2009.11.024, 2010.

Battrick., B.: ESA SP-1279(1) – EarthCARE – Earth Clouds, Aerosols and Radiation Explorer, Tech. rep., 2004.

Baum, B. A., Yang, P., Heymsfield, A. J., Bansemer, A., Cole, B. H., Merrelli, A., Schmitt, C., and Wang, C.: Ice cloud single-scattering property models with the full phase matrix at wavelengths from 0.2 to 100$\mu$m, Journal of Quantitative Spectroscopy and Radiative Transfer, 146, 123–139, https://doi.org/https://doi.org/10.1016/j.jqsrt.2014.02.029, electromagnetic and Light Scattering by Nonspherical Particles XIV, 2014.

Brandes E. A., Zhang, G. and J., V.: Developing and bounding ice particle mass- and area-dimension expressions for use in atmospheric models and remote sensing, J. Appl. Meteor. Climatol., p. 674–685, https://doi.org/10.1175/1520-0450(2002)041, 2002.

Bravo-Aranda, J. A., Belegante, L., Freudenthaler, V., Alados-Arboledas, L., Nicolae, D., Granados-Muñoz, M. J., Guerrero-Rascado, J. L., Amodeo, A., D'Amico, G., Engelmann, R., Pappalardo, G., Kokkalis, P., Mamouri, R., Papayannis, A., Navas-Guzmán, F., Olmo, F. J., Wandinger, U., Amato, F., and Haeffelin, M.: Assessment of lidar depolarization uncertainty by means of a polarimetric lidar simulator, Atmospheric Measurement Techniques, 9, 4935–4953, https://doi.org/10.5194/amt-9-4935-2016, 2016.

Chang, M.: Overview of the EarthCARE multi-spectral imager and results from the development of the MSI engineering model, in: International Conference on Space Optics — ICSO 2012, edited by Cugny, B., Armandillo, E., and Karafolas, N., vol. 10564, p. 105643Z, International Society for Optics and Photonics, SPIE, https://doi.org/10.1117/12.2551545, 2019.

Cole, J. N. S., Barker, H. W., Qu, Z., Villefranque, N., and Shephard, M. W.: Broadband Radiative Quantities for the EarthCARE Mission: The ACM-COM and ACM-RT Products, Atmospheric Measurement Techniques Discussions, 2022, 1–37, https://doi.org/10.5194/amt-2022-304, 2022.

do Carmo, J. P., Hélière, A., Le Hors, L., Toulemont, Y., Lefebvre, A., and Lefebvre, A.: ATLID, ESA Atmospheric LIDAR Developement Status, EPJ Web of Conferences, 119, 04 003, https://doi.org/10.1051/epjconf/201611904003, 2016.

do Carmo, J. P., de Villele, G., Wallace, K., Lefebvre, A., Ghose, K., Kanitz, T., Chassat, F., Corselle, B., Belhadj, T., and Bravetti, P.: ATmospheric LIDar (ATLID): Pre-Launch Testing and Calibration of the European Space Agency Instrument That Will Measure Aerosols and Thin Clouds in the Atmosphere, Atmosphere, 12, https://doi.org/10.3390/atmos12010076, 2021.

Donovan, D., van Zadelhoff, G.-J., and Wang, P.: The EarthCARE lidar cloud and aerosol profile processor: the A-AER, A-EBD, A-TC and A-ICE products, Atmospheric Measurement Techniques, in preparation, 2023.

Donovan, D. P., Klein Baltink, H., Henzing, J. S., de Roode, S. R., and Siebesma, A. P.: A depolarisation lidar-based method for the determination of liquid-cloud microphysical properties, Atmospheric Measurement Techniques, 8, 237–266, https://doi.org/10.5194/amt-8-237-2015, 2015.

Eisinger, M., Wehr, T., Kubota, T., Bernaerts, D., and Wallace, K.: The EarthCARE Mission - Science Data Processing Chain Overview, Atmospheric Measurement Techniques, to be submitted, 2023.

Erfani, E. and Mitchell, D. L.: Developing and bounding ice particle mass- and area-dimension expressions for use in atmospheric models and remote sensing, Atmospheric Chemistry and Physics, p. 4379–4400, https://doi.org/10.5194/acp-16-4379-2016, 2016.

Fell, F. and Fischer, J.: Numerical simulation of the light field in the atmosphere–ocean system using the matrix-operator method, Journal of Quantitative Spectroscopy and Radiative Transfer, 69, 351–388, https://doi.org/https://doi.org/10.1016/S0022-4073(00)00089-3, 2001.

Grenfell, T. C. and Warren, S. G.: Representation of a nonspherical ice particle by a collection of independent spheres for scattering and absorption of radiation, Journal of Geophysical Research: Atmospheres, 104, 31 697–31 709, https://doi.org/https://doi.org/10.1029/1999JD900496, 1999.

Hale, G. M. and Querry, M. R.: Optical Constants of Water in the 200-nm to 200-$\mu$m Wavelength Region, Appl. Opt., 12, 555–563, https://doi.org/10.1364/AO.12.000555, 1973.

Haynes, J. M., L'Ecuyer, T. S., Stephens, G. L., and Jakob, C.: Cloud and precipitation regimes revealed by combined CloudSat and ISCCP observations, AGU Fall Meeting Abstracts, 44, 04, http://adsabs.harvard.edu/abs/2009AGUFM.A44B..04H, 2009.

Hogan, R. J. and Battaglia, A.: 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, 3636–3651, https://doi.org/10.1175/2008JAS2643.1, 2008.

Hogan, R. J. and Westbrook, C. D.: Equation for the Microwave Backscatter Cross Section of Aggregate Snowflakes Using the Self-Similar Rayleigh-Gans Approximation, Journal of the Atmospheric Sciences, 71, 3292–3301, https://doi.org/10.1175/JAS-D-13-0347.1, 2014.

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

Hollstein, A. and Fischer, J.: Radiative transfer solutions for coupled atmosphere ocean systems using the matrix operator technique, Journal of Quantitative Spectroscopy and Radiative Transfer, 113, 536–548, https://doi.org/https://doi.org/10.1016/j.jqsrt.2012.01.010, 2012.

Hostetler, C. A., Behrenfeld, M. J., Hu, Y., Hair, J. W., and Schulien, J. A.: Spaceborne Lidar in the Study of Marine Systems, Annual Review of Marine Science, 10, 121–147, https://doi.org/10.1146/annurev-marine-121916-063335, pMID: 28961071, 2018.

Hünerbein, A., Bley, S., Horn, S., Deneke, H., and Walther, A.: Cloud mask algorithm from the EarthCARE Multi-Spectral Imager: the M-CM products, Atmospheric Measurement Techniques, 16, 2821–2836, https://doi.org/10.5194/amt-16-2821-2023, 2023.

Illingworth, A. J., Barker, H. W., Beljaars, A., Ceccaldi, M., Chepfer, H., Clerbaux, N., Cole, J., Delanoë, J., Domenech, C., Donovan, D. P., Fukuda, S., Hirakata, M., Hogan, R. J., Huenerbein, A., Kollias, P., Kubota, T., Nakajima, T., Nakajima, T. Y., Nishizawa, T., Ohno, Y., Okamoto, H., Oki, R., Sato, K., Satoh, M., Shephard, M. W., Velázquez-Blázquez, A., Wandinger, U., Wehr, T., and van Zadelhoff, G.-J.: The EarthCARE Satellite: The Next Step Forward in Global Measurements of Clouds, Aerosols, Precipitation, and Radiation, Bulletin of the American Meteorological Society, 96, 1311–1332, https://doi.org/10.1175/BAMS-D-12-00227.1, 2015.

Inness, A., Ades, M., Agustí-Panareda, A., Barré, J., Benedictow, A., Blechschmidt, A.-M., Dominguez, J. J., Engelen, R., Eskes, H., Flemming, J., Huijnen, V., Jones, L., Kipling, Z., Massart, S., Parrington, M., Peuch, V.-H., Razinger, M., Remy, S., Schulz, M., and Suttie, M.: The CAMS reanalysis of atmospheric composition, Atmospheric Chemistry and Physics, 19, 3515–3556, https://doi.org/10.5194/acp-19-3515-2019, 2019.

Josset, D., Tanelli, S., Hu, Y., Pelon, J., and Zhai, P.: Analysis of Water Vapor Correction for CloudSat W-Band Radar, IEEE Transactions on Geoscience and Remote Sensing, 51, 3812–3825, https://doi.org/10.1109/TGRS.2012.2228659, 2013.

Kato, S., Ackerman, T. P., Mather, J. H., and Clothiaux, E. E.: The k-distribution method and correlated-k approximation for a shortwave radiative transfer model, Journal of Quantitative Spectroscopy and Radiative Transfer, 62, 109–121, https://doi.org/https://doi.org/10.1016/S0022-4073(98)00075-2, 1999.

Kollias, P., Albrecht, B. A., and Marks, F.: Why Mie?, Bulletin of the American Meteorological Society, 83, 1471–1483, https://doi.org/10.1175/BAMS-83-10-1471, 2002.

Kollias, P., Clothiaux, E. E., Miller, M. A., Albrecht, B. A., Stephens, G. L., and Ackerman, T. P.: Millimeter-Wavelength Radars: New Frontier in Atmospheric Cloud and Precipitation Research, Bull. Amer. Meteorol. Soc., 88, 1608–1624, https://doi.org/10.1175/BAMS-88-10-1608, 2007.

Kollias, P., Tanelli, S., Battaglia, A., and Tatarevic, A.: Evaluation of EarthCARE Cloud Profiling Radar Doppler Velocity Measurements in Particle Sedimentation Regimes, Journal of Atmospheric and Oceanic Technology, 31, 366–386, https://doi.org/10.1175/JTECH-D-11-00202.1, 2014.

Kollias, P., Battaglia, A., Lamer, K., Treserras, B. P., and Braun, S. A.: Mind the Gap - Part 3: Doppler Velocity Measurements From Space, Frontiers in Remote Sensing, 3, https://doi.org/10.3389/frsen.2022.860284, 2022a.

Kollias, P., Treserras, B. P., Battaglia, A., Borque, P. C., and Tatarevic, A.: Processing reflectivity and Doppler from Earth-CARE's cloud-profiling radar: the C-FMR, C-CD and C-APC products, Atmospheric Measurement Techniques, to be submitted, https://doi.org/10.5194/amt-16-1901-2023, 2022b.

Lamer, K., Kollias, P., Battaglia, A., and Preval, S.: Mind the gap – Part 1: Accurately locating warm marine boundary layer clouds and precipitation using spaceborne radars, Atmospheric Measurement Techniques, 13, 2363–2379, https://doi.org/10.5194/amt-13-2363-2020, 2020.

Li, L., Heymsfield, G. M., Tian, L., and Racette, P. E.: Measurements of Ocean Surface Backscattering Using an Airborne 94-GHz Cloud Radar—Implication for Calibration of Airborne and Spaceborne W-Band Radars, Journal of Atmospheric and Oceanic Technology, 22, 1033 – 1045, https://doi.org/https://doi.org/10.1175/JTECH1722.1, 2005.

Lin, Z., Stamnes, S. A., Jin, Z., Laszlo, I., Tsay, S.-C., Wiscombe, W. J., and Stamnes, K.: Improved discrete ordinate solutions in the presence of an anisotropically reflecting lower boundary: Upgrades of the DISORT computational tool, Journal of Quantitative Spectroscopy & Radiative Transfer, 157, 119–134, 2015.

Mason, S. L., Hogan, R. J., Donovan, D. P., van Zadelhoff, G.-J., Kollias, P., Treserras, B. P., Qu, Z., Cole, J. N. S., Hünerbein, A., and Docter, N.: An intercomparison of EarthCARE cloud, aerosol and precipitation retrieval products, Atmospheric Measurement Techniques, to be submitted, 2023.

Miles, R. B., Lempert, W. R., and Forkey, J. N.: Laser Rayleigh scattering, Measurement Science and Technology, 12, R33–R51, https://doi.org/10.1088/0957-0233/12/5/201, 2001.

Mitchell, D. L. and Heymsfield, A. J.: Refinements in the Treatment of Ice Particle Terminal Velocities, Highlighting Aggregates, Journal of the Atmospheric Sciences, 62, 1637–1644, https://doi.org/10.1175/JAS3413.1, 2005.

Moody, E. G., King, M. D., Schaaf, C. B., Hall, D. K., and Platnick, S.: Northern Hemisphere five-year average (2000–2004) spectral albedos of surfaces in the presence of snow: Statistics computed from Terra MODIS land products, Remote Sensing of Environment, 111, 337–345, https://doi.org/https://doi.org/10.1016/j.rse.2007.03.026, remote Sensing of the Cryosphere Special Issue, 2007.

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

Pan, X., Shneider, M. N., and Miles, R. B.: Coherent Rayleigh-Brillouin Scattering, Phys. Rev. Lett., 89, 183 001, https://doi.org/10.1103/PhysRevLett.89.183001, 2002.

Qu, Z., Donovan, D. P., Barker, H. W., Cole, J. N. S., Shephard, M. W., and Huijnen, V.: Numerical Model Generation of Test Frames for Prelaunch Studies of EarthCARE's Retrieval Algorithms and Data Management System, Atmospheric Measurement Techniques Discussions, 2022, 1–31, https://doi.org/10.5194/amt-2022-300, 2022.

Rosenkranz, P. W.: Water vapor microwave continuum absorption: A comparison of measurements and models, Radio Science, 33, 919–928, https://doi.org/https://doi.org/10.1029/98RS01182, 1998.

Saleh, B. E. A. and Teich, M. C.: Fundamentals of photonics; 1st ed., Wiley, New York, NY, https://cds.cern.ch/record/244535, 1991.

Schaaf, C. B., Gao, F., Strahler, A. H., Lucht, W., Li, X., Tsang, T., Strugnell, N. C., Zhang, X., Jin, Y., Muller, J.-P., Lewis, P., Barnsley, M., Hobson, P., Disney, M., Roberts, G., Dunderdale, M., Doll, C., d'Entremont, R. P., Hu, B., Liang, S., Privette, J. L., and Roy, D.: First operational BRDF, albedo nadir reflectance products from MODIS, Remote Sensing of Environment, 83, 135–148, https://doi.org/https://doi.org/10.1016/S0034-4257(02)00091-3, the Moderate Resolution Imaging Spectroradiometer (MODIS): a new generation of Land Surface Monitoring, 2002.

Tanelli, S., Durden, S. L., Im, E., Pak, K. S., Reinke, D. G., Partain, P., Haynes, J. M., and Marchand, R. T.: CloudSat's cloud profiling radar after two years in orbit: performance, calibration, and processing, IEEE Transactions on Geoscience and Remote Sensing, 46, 3560–3573, https://doi.org/10.1109/TGRS.2008.2002030, 2008.

van Zadelhoff, G.-J., Donovan, D. P., and Wang, P.: Detection of aerosol and cloud features for the EarthCARE lidar ATLID: the A-FM product, EGUsphere, 2023, 1–29, https://doi.org/10.5194/egusphere-2023-145, 2023.

Velázquez-Blázquez, A., Baudrez, E., Clerbaux, N., and Domenech, C.: Unfiltering of the EarthCARE Broadband Radiometer (BBR) observations: the BM-RAD product, Atmospheric Measurement Techniques, to be submitted, 2023.

Vidot, J. and Borbás, E.: Land surface VIS/NIR BRDF atlas for RTTOV-11: model and validation against SEVIRI land SAF albedo product, Quarterly Journal of the Royal Meteorological Society, 140, https://doi.org/10.1002/qj.2288, 2014.

Wandinger, U., Floutsi, A. A., Baars, H., Haarig, M., Ansmann, A., Hünerbein, A., Docter, N., Donovan, D., van Zadelhoff, G.-J., Mason, S., and Cole, J.: HETEAC – The Hybrid End-To-End Aerosol Classification model for EarthCARE, EGUsphere, 2022, 1–40, https://doi.org/10.5194/egusphere-2022-1241, 2022.

Witschas, B.: Analytical model for Rayleigh–Brillouin line shapes in air, Appl. Opt., 50, 267–270, https://doi.org/10.1364/AO.50.000267, 2011.

Zhou, C. and Yang, P.: Backscattering peak of ice cloud particles, Opt. Express, 23, 11 995–12 003, https://doi.org/10.1364/OE.23.011995, 2015.

Zrnic, D. S.: Simulation of weatherlike Doppler spectra and signal, J. Appl. Meteorol., 14, 619–620, 1975.

Zrnic, D. S.: Spectral moment estimates from correlated pulse pairs, IEEE Trans. Aerosp. Electron. Syst., AES-13, 344–354, 1977.