Algorithms to retrieve aerosol optical properties using lidar measurements on board the EarthCARE satellite

Tomoaki Nishizawa[1], Rei Kudo[2], Eiji Oikawa[2], Akiko Higurashi[1], Yoshitaka Jin[1], Nobuo Sugimoto[1], Kaori Sato[3], Hajime Okamoto[3]

[1]National Institute for Environmental Studies, 16-2 Onogawa, Tsukuba, 305-8506, Japan
[2]Meteorological Research Institute, Japan Meteorological Agency, 1-1 Nagamine, Tsukuba, 305-0052, Japan
[3]Research institute for applied mechanics, Kyushu university, 6-1 Kasuga-koen, Fukuoka, 816-8580, Japan

Correspondence to: Tomoaki Nishizawa (nisizawa@nies.go.jp)

Abstract.

The Earth Cloud Aerosol and Radiation Explorer (EarthCARE) is a joint Japanese-European satellite observation mission for understanding the interaction between cloud, aerosol, and radiation processes and improving the accuracy of climate change predictions. The EarthCARE satellite was equipped with four sensors, a 355 nm high-spectral-resolution lidar with depolarization measurement capability (ATLID) as well as a cloud profiling radar, a multi-spectral imager, and a broadband radiometer, to observe the global distribution of clouds, aerosols, and radiation. In this study, we have developed algorithms to produce ATLID Level 2 aerosol products using ATLID Level 1 data. The algorithms estimated the following four products: (1) Layer identifiers such as aerosols, clouds, clear-skies, or surfaces were estimated by the combined use of vertically variable criteria and spatial continuity methods developed for the CALIOP (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation) analysis. (2) Aerosol optical properties such as extinction coefficient, backscatter coefficient, depolarization ratio, and lidar ratio at 355 nm were optimized to ATLID L1 data by the method of maximum likelihood. (3) Six aerosol types, namely smoke, pollution, marine, pristine, dusty-mixture, and dust were identified based on a two-dimensional diagram of the lidar ratio and depolarization ratio at 355 nm developed by cluster-analysis using the AERONET (AErosol RObotic NETwork) dataset with ground-based lidar data. (4) The planetary boundary layer height was determined using the improved wavelet covariance transform method for the ATLID analysis. The performance of various algorithms was evaluated using pseudo ATLID Level 1 data generated by Joint-Simulator (Joint Simulator for Satellite Sensors), which incorporates aerosol and cloud distributions simulated by numerical models. Results from applying the algorithms to the pseudo ATLID Level 1 data with realistic signal noise added for aerosol or cloud predominant cases revealed: (1) misidentification of aerosol and cloud layers was relatively low, approximately 10%; (2) the retrieval errors of aerosol optical properties were $0.08 \times 10^{-7} \pm 1.12 \times 10^{-7}$ m^{-1}sr^{-1} (2 ± 34% in relative error) for backscatter coefficient and 0.01 ± 0.07 (4 ± 27% in relative error) for depolarization ratio; (3) aerosol type classification was generally performed well. These results indicate that the algorithm's capability to provide valuable insights into the global distribution of aerosols and clouds, facilitating assessments of their climate impact through atmospheric radiation processes.

1 Introduction

Global measurements of the optical properties of atmospheric particles , encompassing aerosols and clouds, play a pivotal role in evaluating their climatic effects through atmospheric radiation processes. Zhang et al. (2022) analyzed the simulation results of the Coupled Model Intercomparison Project Phase 6 (CMIP6) global models and indicated that 20% of radiative effect uncertainty for anthropogenic aerosols was associated with aerosol vertical distribution under clear-sky condition. In addition, the presence of highly light-absorbing aerosols such as black carbon above or below the cloud layer can prodoundly the radiative transfer process of the upper atmosphere (Takemura et al., 2002; Oikawa et al. 2013, 2018). The Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation (CALIPSO) satellite, equipped with the two-wavelength (1064, 532nm) polarization Mie-scattering lidar CALIOP (Cloud-Aerosol LIdar with Orthogonal Polarization), has globally measured aerosols and cloud vertical distributions over the long term from 2006 to 2023 (Winker et al. 2010). This role will transition to ATLID (Atmospheric Lidar) (Heiliere et al. 2017; do Carmo et al. 2021), the 355nm high-spectral resolution lidar (HSRL) with polarization measurement function onboard the EarthCARE satellite (Earth Cloud Aerosol and Radiation Explorer). EarthCARE is a joint Japan-Europe satellite observation mission by JAXA (Japan Aerospace Exploration Agency), NICT (National Institute of Information and Communications Technology), and ESA (European Space Agency), which aims to conduct comprehensive global observations of clouds, aerosols, and atmospheric radiation (Wehr et al. 2023). The EarthCARE satellite features a 94 GHz cloud radar (CPR) with doppler measurement capability (Nakatsuka et al. 2023), a multiwavelength imager (MSI), comprising seven channels in the visible to infrared wavelength range (Wallace et al. 2016), and a broadband radiometer (BBR) designed for measuring shortwave and longwave radiation (Wallace et al. 2016). The scientific objectives emcompass (1) observing the global vertical distributions of natural and anthropogenic aerosols and their interaction with clouds; (2) observing global cloud distributions, cloud-precipitation interactions, and vertical motion characteristics within clouds; (3) evaluating the vertical profiles of radiative heating and cooling of the atmosphere (Illingworth et al. 2015; Okamoto et al. 2024a).

ATLID, a HSRL, can independently measure backscattered light from atmospheric particles (Mie signal) and backscattered light from atmospheric molecules (Rayleigh signal) separately, distinguishing it from CALIOP, a Mie-scattering lidar. This unique feature allows for the independent extraction of the vertical distributions of the extinction coefficient and backscattering coefficients of atmospheric particles by analyzing these two signals. Moreover, ATLID conducts polarization measurements, enabling the extraction of the co-polar component, parallel to the laser polarization (co-polar component), the cross-polar component (perpendicular to the laser polarization) of the Mie signal at a wavelength of 355 nm, and the Rayleigh signal. Signal calibrations are performed on these measured signals, resulting in calibrated attenuated backscatter coefficients.

Multichannel lidar observations using a combination of polarization measurements, Rayleigh/Mie separation measurements, and/or various wavelength measurements, including those of ATLID, CALIOP, and ground-based lidar measurements, facilitate simultaneous understanding of various optical and microphysical properties of atmospheric particles. This includes the identification of major atmospheric layers such as aerosol or cloud layers (Okamoto et al. 2008; Vaughan et al. 2009; Liu

et al. 2009; Hagihara et al. 2010) and the identification of aerosol and cloud types such as marine, dust, warm water, or 2-D ice (Omar et al. 2009; Yoshida et al., 2010; Illingworth et al. 2015; Okamoto et al. 2019). The size distribution and refractive index of total aerosols can be estimated using extinction coefficients and backscatter coefficients at two wavelengths (355 nm and 532 nm) observed by the Raman lidar and HSRL (e.g., Müller et al. 2014). The extinction coefficients of some major aerosol components (such as mineral dust and black carbon) have been estimated (Nishizawa et al. 2017; Kudo et al. 2023). Understanding of the climate impact of aerosols, necessitates not only comprehending the optical and microphysical properties of total aerosols but also those of individual aerosol components. For example, aerosols with strong light absorption properties (e.g., black carbon) have been reported to influence cloud formation and global atmospheric and water circulations in large fields (Menon et al. 2002; Koren et al. 2004). In addition, the top height of the planetary boundary layer (PBL) has been detected by various methods such as the gradient method (Lammert and Bösenberg, 2006), the wavelet covariance transform (WCT) method (Brooks, 2003), or other methods (e.g., the standard deviation method by Menut et al., 1999). Since aerosols emitted from the surface are trapped within the PBL because of the temperature inversion between the PBL and free troposphere, its top height, by which the transition point of the aerosol concentration gap is characterized, is an important parameter for understanding climate and air quality.

The EarthCARE mission generates Level 1 (L1) products, which are physical quantities derived after calibrating the measurements of each instrument, and Level 2 (L2) products, consisting of geophysical variables related to clouds, aerosols, etc, utilising L1 data either independenally or in combination (Eisinger et al. 2024). Each sensor development organization (JAXA and NICT for CPR, and ESA for others) generates the L1 product, and the L2 product is generated by individual agencies. For example, L2 products using ATLID, such as layer identifiers, particle (aerosol, cloud) type identifiers, and optical properties, such as extinction coefficient, backscatter coefficient, and depolarization ratio, and cloud top height, are estimated from the ESA (Irbah et al. 2023; Donovan et al. 2023, 2024; Zadelhoff et al. 2023; Wandinger et al. 2023). Synergy products are also generated by combining L1 data from the other instruments. JAXA also provides various ATLID standalone and synergistic products, while similar to ESA products, are produced using different independent retrieval methods developed independently.

To determine the globally vertical distribution of optical properties of aerosols and clouds needed to assess their climatic effects, we have developed atmospheric particle retrieval algorithms to generate JAXA L2 products using ATLID L1 data. In this study, we focus on the retrieval algorithms and products related to aerosols, whereas the description of cloud property estimation is planned for separate papers. Section 2 describes the products retrieved using the algorithms and their flows. Section 3 describes the algorithms developed to retrieve the individual products. Section 4 presents the products estimated using the simulated input data (ATLID L1 data) and the performance of the various algorithms. In Section 5, we summarize and discuss the prospects.

2 Algorithm flow and products

Figure 1 illustrates the products and flow of the algorithms. First, to improve signal quality, the algorithm reduces the signal noise using a discrete wavelet transform (DWT) (Fang and Huang, 2004). In this study, two types of Daubechies wavelets (support numbers 2 (D2) and 4 (D4)) are employed. Noise reduction is performed iteratively with D2, followed by D4, and then the noise reduction using D2 is applied again, followed by D4 and so on (D2=>D4=>D2=>D4=>D2...). A simulation analysis with added random noise indicated that the applying the DWT method could improve the signal-to-noise ratio (SNR) by a factor of two or more. Therefore, this method was adopted in the present study.

To further enhance the SNR, we integrate the data horizontally after applying the DWT method to the minimum horizontal resolution (~0.3 km) data from ATLID L1. This integration yields data at a 1 km horizontal resolution (ATLID L1(1) in Figure 1). Additionally, a moving average of this 1 km horizontal resolution data with a width of 10 km was produced (ATLID L1 (1*) in Figure 1). Consequently, the algorithm uses the three L1 data created (ATLID L1 (0.3), (1), and (1*) in Figure 1), with an altitude resolution of 0.1 km, akin to the original data, as input data when estimating each product. A quality check (QC) involving the SNR is performed on each of the generated data.

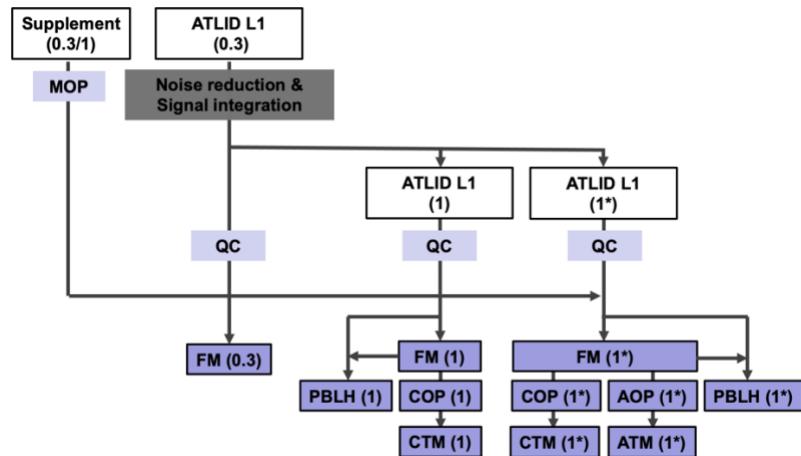



Figure 1. ATLID L2 products and the flow of algorithms.

Six different products (FM (Feature Mask), AOP (Aerosol Optical Properties), COP (Cloud Optical Properties), ATM (Aerosol Target Mask), CTM (Cloud Target Mask), PBLH (Planetary Boundary Layer Height) in Fig. 1) are created from the ATLID L1 data. First, the algorithm identifies which component of atmospheric molecules, aerosols, or cloud particles is predominantly present at each altitude, and outputs its identifier (FM). Additionally, the surface, subsurface, and fully attenuated layers are identified here. Secondly, the particle optical properties (POP), that is, the extinction coefficient (α_p), backscatter coefficient (β_p), depolarization ratio (δ_p), and lidar ratio (S_p) at 355nm, are retrieved from the ATLID L1 data. Whether the derived POP are of aerosol or cloud origin is determined according to the FM product (AOP or COP). Aerosol (ATM) and cloud (CTM) particle types are also identified at each altitude. The PBLH is then estimated. The product derived

from original 0.3 km data is only the FM because of the low SNR. To account for cloud heterogeneity and the low SNR of the aerosol layers, the cloud products (COP and CTM) and PBLH are derived from the 1 km horizontal resolution data. The aerosol and cloud products (AOP, ATM, COP, and CTM) and PBLH are created from the 1* km horizontal resolution data. In the following section, products and algorithms related to aerosols (i.e., FM, AOP, ATM, PBLH) are mainly described. To estimate these products, the optical properties of atmospheric molecules, such as extinction and backscatter coefficients (MOP in Figure 1) are required. Here, the EarthCARE auxiliary product, containing a subset of meteorological fields from an ECMWF forecast model (Eisinger et al. 2023) is used to compute the molecular optical propertes.

3 Algorithms

3.1 Layer identification (Feature mask)

The algorithm was developed based on retrieval methods developed using the CALIOP and ground-based lidar data (Okamoto et al. 2007 and 2008; Hagihara et al. 2010; Shimizu et al. 2004). The algorithm classifies the atmospheric layer into molecules (clear-sky), aerosol particles, or cloud particles, as well as the surface, subsurface, and layers, where the signal is fully attenuated under optically thick layers, such as clouds (Vaughan et al 2009). The co-polar ($\beta_{atn.M,co}$) and cross-polar ($\beta_{atn.M,cr}$) components of the Mie attenuated backscatter coefficient and Rayleigh attenuated backscatter coefficient ($\beta_{atn.R}$) are given as ATLID L1 products and are described by the following equations with atmospheric parameters:

$$\beta_{atn,M,co} = \beta_{p,co}(z) exp\left\{-2 \int_z^{Z_{ATLID}} \left(\alpha_p(z') + \alpha_m(z')\right) dz'\right\}, \qquad (1a)$$

$$\beta_{atn,M,cr} = \beta_{p,cr}(z) exp\left\{-2 \int_z^{Z_{ATLID}} \left(\alpha_p(z') + \alpha_m(z')\right) dz'\right\}, \qquad (1b)$$

$$\beta_{atn,R} = \beta_m(z) exp\left\{-2 \int_z^{Z_{ATLID}} \left(\alpha_p(z') + \alpha_m(z')\right) dz'\right\}. \qquad (1c)$$

$$\beta_{p,co}(z) = \beta_p(z) \frac{1}{1+\delta_p(z)}, \qquad (1d)$$

$$\beta_{p,cr}(z) = \beta_p(z) \frac{\delta_p(z)}{1+\delta_p(z)}, \qquad (1e)$$

$$\beta_p(z) = \frac{\alpha_p(z)}{S_p(z)}. \qquad (1f)$$

where, α_p, S_p, and δ_p are the extinction coefficient, lidar ratio, and linear depolarization ratio of particles (aerosols and clouds), respectively; $\beta_{p,co}$ and $\beta_{p,cr}$ are the co-polar and cross-polar components of the particular backscatter coefficient, respectively; α_m and β_m are the molecular extinction coefficient and the molecular backscatter coefficient, respectively; Z is the altitude; and Z_{ATLID} is the altitude of ATLID. The Mie-attenuated backscatter coefficient ($\beta_{atn.M}$) is defined as the sum of the co-polar and cross-polar components (i.e., $\beta_{atn.M} = \beta_{atn.M,co} + \beta_{atn.M,cr}$). We use $\beta_{atn.M}$ and $\beta_{atn.R}$ as the diagnostic parameters. The particle backscatter coefficient ($\beta_p = \beta_{p,co} + \beta_{p,cr}$) calculated from the following equation is also used as the diagnostic parameter :

$$\beta_P(z) = \beta_m(z) \beta_{atn,M}(z) / \beta_{atn,R}(z). \qquad (2)$$

First, a diagnosis is made on the 0.3 km horizontal resolution data in each layer sequentially from the upper layer to the lower layer according to the following criteria.

(a) If the SNRs of $\beta_{atn,M}$ and $\beta_{atn,R}$ are lower than the threshold (SNR$_{th}$) or data is missing, no diagnosis is made in that layer (classified as "Invalid").

(b) If the SNR of $\beta_{atn,R}$ is greater than the SNR$_{th}$ and the SNR of $\beta_{atn,M}$ is lower than SNR$_{th}$, the layer is classified as "Clear-sky."

(c) The SNR of $\beta_{atn,M}$ is greather than the SNR$_{th}$, the layer is classified as "Aerosol," "Cloud," or "Surface.

In this study, the value of SNR$_{th}$ is set to 3.

For the layer diagnosed as "Aerosol," "Cloud," or Surface," the following further diagnoses are made. For the surface detection, the two criteria are imposed: (1) the $\beta_{atn,M}$ is above the threshold value which is set based on actual ATLID observed data, and (2) the relevant layer is below +500 m with respect to the altitude of the Digital Elevation Model (DEM) surface altitude from the EarthCARE auxiliary data, to prevent misdetection due to clouds with large backscatter coefficients. If the above criteria are met, the layer is marked as "Surface," The layer below the surface is classified as "Sub-surface."

For the cloud detection, two procedures are performed based on the cloud-mask scheme developed for the CALIOP and shipborne lidar measurements (Hagihara et al. 2010 and Okamoto et al. 2008). First, the cloud layer is detected using the following vertical variable criteria.

$$\beta_{atn,M} > 0.5\beta_{c,th} \exp(-2\tau_m)\{1 - \tanh(z - z_c)\}, \quad (3a)$$

$$\beta_P > 0.5\beta_{c,th}\{1 - \tanh(z - z_c)\}. \quad (3b)$$

where, τ_m is the molecular optical thickness up to an altitude (z) from the altitude of the ATLID. z_c is given as 5 km, and $\beta_{c,th}=10^{-5.25}$ ($\sim5.6\times10^{-6}$) [m^{-1}sr^{-1}]. The validity of these threshold values was discussed in detail by Hagihara et al. (2010) and Okamoto et al. (2008). Statistical analyses of cloud and aerosol backscatter coefficient data over several years from long-term ground-based observations by HSRL (Jin et al. 2020) and Raman lidar (Nishizawa et al. 2017) also supported the value of $\beta_{c,th}$. If the SNR of $\beta_{atn,R}$ of the target layer is greater than SNR$_{th}$, the Equation 3(b) is used as the criteria; otherwise, Equation 3(a), is used. The advantage of using β_P as a diagnostic parameter is that it can eliminate attenuation due to clouds/aerosols from the ATLID to the target layer. However, the calculation of the β_P requires $\beta_{atn,R}$ in addition to $\beta_{atn,M}$, which has a disadvantage in that it cannot be diagnosed if the SNR of $\beta_{atn,R}$ is insufficient. Therefore, a hybrid method is used, where $\beta_{atn,M}$ is used if β_P cannot be used as a diagnostic parameter. After identifying the cloud layer using these critetia, a continuity test is performed to suppress misdetection due to signal noise (Hagihara et al. 2010). The continuity test is conducted on a 5-bin horizontal and 3-bin vertical window centered on the diagnostic layer, in accordance with the continuity test method used by Hagihara et al. (2010). If more than half (i.e. more than 8 bins) of the total 15 bins are clouds, the target layer is determined to be a "Cloud". However, if the number of cloud bins was less than half of the total, but at least one of the cloud bins was included, the target layer was designated as "Unknown."

For 0.3 km and 1 km resolution products, the "Clear-sky" or "Aerosol" layer are not diagnosed separately from the SNR's point of view, and together they are classified as "Clear-sky or aerosol" The "Fully attenuated" layer is diagnosed when no surface is detected; the layers below the lowest layer of "Clear-sky or aerosol" or "Cloud" are classified as "Fully attenuated."

Next, a diagnosis is made on the 1 km horizontal resolution data. Diagnostics similar to those performed on the 0.3 km horizontal data are performed; the layer types of "Cloud," "Clear-sky or aerosol," "Surface," " Sub-surface," "Fully attenuated," "Invalid," or "Unknown" are identified. The following method is used to identify the cloud layers. The 1 km horizontal resolution data is calculated by averaging several horizontal layers (~4 bins) of the 0.3 km horizontal resolution data. This improves the SNR, but it also causes dulling of the signal at the cloud edges, which leads to cloud layer misidentification. To suppress this, we use the FM product estimated using the 0.3 km horizontal resolution data (i.e., FM(0.3)); if more than half of the total number of the layers of the FM (0.3) product in the target layer for the 1 km horizontal resolution data is identified as "Cloud," the target layer for the 1 km horizontal resolution data is determined to be "Cloud" (Hagihara et al. 2010). However, if the number of cloud layers is less than half of the total, but at least one of the cloud layers is included, the target layer is designated as "Unknown." In addition, the possibility of identifying optically thin clouds or aerosols compared to the 0.3 km data, especially at high altitudes, also arises because of the improved SNR by the horizontal averaging. Equations 4a and 4b are then applied to the 1 km horizontal resolution data with the addition of a criterion for high altitude to Equations 3a and 3b.

$$\beta_{atn,M} > 0.5\beta_{c,th}\exp(-2\tau_m)\{1 - \tanh(z - z_c)\} + 0.5\beta_{c,th2}\exp(-2\tau_m)\{1 + \tanh(z - z_c)\}, \quad (4a)$$

$$\beta_P > 0.5\beta_{c,th}\{1 - \tanh(z - z_c)\} + 0.5\beta_{c,th2}\{1 + \tanh(z - z_c)\}. \quad (4b)$$

$\beta_{c,th2}$ is the threshold for identifying the layer that may be a cloud at high altitudes and is set based on actual ATLID data. If the criteria is satisfied, the layer is marked as "Unknown," and if not, it is marked as "Aerosol or clear-sky" if the threshold is unsatisfied.

Finally, a diagnosis is made on the 1^* km horizontal resolution data. Diagnostics similar to those performed on the 1 km horizontal resolution data are performed; the layer types of "Cloud," "Aerosol," "Clear-sky," "Surface," "Sub-surface," "Fully attenuated," "Invalid," or "Unknown" are identified. Unlike the FM for the 1 km (and 0.3 km) horizontal resolution data, the layer diagnosed as "Aerosol or clear-sky" is classified into the "Aerosol" when the $\beta_{atn.M}$ is greater than SNR_{th} or "Clear-sky" when the $\beta_{atn.M}$ is lower than SNR_{th}.

3.2 Aerosol optical properties

As shown in Equations 1a-1c, the POP α_p, β_p, δ_p, and S_p can be directly derived using the L1 data of $\beta_{atn.M,co}$, $\beta_{atn.M,cr}$, and $\beta_{atn.R}$. The POP in the aerosol layer identified by the FM are classified as AOP. On the other hand, L1 data with a sufficient SN are needed to derive parameters with sufficient accuracy, especially for extinction coefficient retrieval (e.g., Liu et al. 1999]. The DWT method and moving average are applied to the L1 data, but the SNR is not sufficiently small. Therefore, we simultaneously estimated the vertical profiles of the POP from the L1 data by the method of maximum likelihood with a priori

smoothness constraints for the vertical profiles of the POP. The state vector $\boldsymbol{x}$, which comprises of $\alpha_\mathrm{p}(z_i)$, $\delta_\mathrm{p}(z_i)$, and $S_\mathrm{p}(z_i)$ at altitudes z_i, is optimized to the L1 data by minimizing the following cost function:

$$
f(\boldsymbol{x}) = \sum_i \frac{\left\{ ln\left(\beta_\mathrm{atn.M,co}^\mathrm{obs}(z_i) - \beta_\mathrm{atn.M,co}^\mathrm{min}\right) - ln\left(\beta_\mathrm{atn.M,co}^\mathrm{cal}(z_i) - \beta_\mathrm{atn.M,co}^\mathrm{min}\right) \right\}^2}{w_\mathrm{atn.M,co}^2(z_i)}
$$

$$
+ \sum_i \frac{\left\{ ln\left(\beta_\mathrm{atn.M,cr}^\mathrm{obs}(z_i) - \beta_\mathrm{atn.M,cr}^\mathrm{min}\right) - ln\left(\beta_\mathrm{atn.M,cr}^\mathrm{cal}(z_i) - \beta_\mathrm{atn.M,cr}^\mathrm{min}\right) \right\}^2}{w_\mathrm{atn.M,cr}^2(z_i)}
$$

$$
+ \sum_i \frac{\left\{ ln\left(\beta_\mathrm{atn.M,R}^\mathrm{obs}(z_i) - \beta_\mathrm{atn.M,R}^\mathrm{min}\right) - ln\left(\beta_\mathrm{atn.M,R}^\mathrm{cal}(z_i) - \beta_\mathrm{atn.M,R}^\mathrm{min}\right) \right\}^2}{w_\mathrm{atn.M,R}^2(z_i)}
$$

$$
+ \sum_i \frac{\left\{ ln\left(\alpha_\mathrm{p}(z_i)\right) - ln\left(-\alpha_\mathrm{p}(z_{i+1})\right) \right\}^2}{w_{\alpha_P}^2}
$$

$$
+ \sum_i \frac{\left\{ ln\left(S_\mathrm{p}(z_i)\right) - ln\left(-S_\mathrm{p}(z_{i+1})\right) \right\}^2}{w_{S_P}^2}
$$

$$
+ \sum_i \frac{\left\{ ln\left(\delta_\mathrm{p}(z_i)\right) - ln\left(-\delta_\mathrm{p}(z_{i+1})\right) \right\}^2}{w_{\delta_P}^2}, \tag{5}
$$

where z_i is i-th altitude, "obs" indicates the measurements, "cal" indicates the values calculated from $\boldsymbol{x}$ by Equations 1a-1f, where z_i is i-th altitude, "obs" indicates the measurements, "cal" indicates the values calculated from $\boldsymbol{x}$ by Equations 1a-1f, $w_\mathrm{atn.M,co/cr/R}$ is measurement uncertainties, $w_{\alpha_\mathrm{p}/\delta_\mathrm{p}/S_\mathrm{p}}$ is weight for determine the strength of the smoothness constraint, and "min" indicates the possible minimum value of the measurements. The measurements uncertaintis $w_\mathrm{atn.M,co/cr/R}$ are calculated from the signal noises such as shot noise, dark noise, and CCD read-out noise which are evaluated befor the lauch of the EarthCARE satellite. The values of $w_{\alpha_\mathrm{p}/\delta_\mathrm{p}/S_\mathrm{p}}$ are given by 1.0 in this study. A logarithmic transformation is applied to the measured and calculated values because the transformation reduces the differences of the dynamic ranges of each term in Eq. (5) and enables a fast and stable minimiation of $f(\boldsymbol{x})$ (Kudo et al., 2016). However, the L1 data may have negative values due to large SN, and the logarithmic transformation cannot be applied to negative values. Therefore, we subtracted the possible minimum values from the measured and calculated values. The fourth, fifth, and sixth terms in Equation 5 are the smoothness constraints for the vertical profiles of $\alpha_\mathrm{p}(z_i)$, $\delta_\mathrm{p}(z_i)$, and $S_\mathrm{p}(z_i)$. The smoothness of the vertical profiles is obtained by minimizing the differences of the values at two adjacent altitudes.

The minimization of $f(\boldsymbol{x})$ is conducted by the iteration of $ln(\boldsymbol{x}_{i+1}) = ln(\boldsymbol{x}_i) + \gamma \Delta \boldsymbol{x}_i$ in $ln(\boldsymbol{x})$ space. The vector $\Delta \boldsymbol{x}_i$ at the i-th step is determined by the Gauss-Newton method, and the scalar γ is determined by a line search method with Armijo rule (Nocedal and Wright, 2006). The convergence criterion for the iteration is that the difference between $f(\boldsymbol{x}_i)$ and $f(\boldsymbol{x}_{i+1})$ should be smaller than a given threshold.

In the actual analysis using the ATLID data, the optical properties of clouds and aerosols are simultaneously estimated using the developed POP retrieval algorithm. To estimate the optical properties of optically thick scatterers such as clouds, it is

essential to compute L1 data considering multiple scattering. The multiple scattering has been considered in CALIPSO and ground-based lidar analysis by introducing η-factor (e.g., Chen et al. 2002, Young 2013, Cairo et al. 2021), and thus the η-factor will be introduced into this algorithm to estimate the optical properties of clouds. Sato et al. (2018) developed a practical model to determine the time-dependent lidar attenuated backscatter coefficient, in which an analytical expression for the high-order phase function was implemented to reduce computational cost; furthermore, Sato et al. (2019) developed a vectorized physical model (VPM) which is a physical model extended with a polarization function, to analyze the observed depolarization ratio due to multiple scattering from water clouds. The introduction of these physical models is a promising as a more advanced and accurate approach in estimating COP, considering multiple scattering from clouds.

3.3 Aerosol type classification (Target mask)

The algorithm uses the derived FM and AOP products. The algorithm classifies aerosol type for each layer that the FM scheme identifies as "Aerosol." The algorithm uses differences in the light absorption and polarization properties of aerosol types; S_a strongly reflects the light absorption of aerosols, while δ_a strongly reflects the polarization of aerosols. Therefore, it is necessary to determine the aerosol types and their optical models. The optical properties and size distributions of aerosols are observed by the AErosol RObotic NETwork (AERONET) sun/sky photometer, and AERONET observation sites are located worldwide on all continents (e.g. Holben et al. 1998; Giles et al. 2019). In the retrieval of aerosol extinction in CALIPSO version 2 and 3 products, S_a of three aerosol types calculated by using size distributions and refractive indices of the clusters grouped by the cluster analysis of the AERONET dataset were used (Omar et al. 2005; Omar et al. 2009). Based on this assumption, we also perform a cluster analysis of the aerosol optical properties estimated from the AERONET data and determined the aerosol optical properties at 355 nm. In this study, aerosol particles are classified as six aerosol types ("Smoke," "Pollution," "Marine," "Pristine," "Dust," and "Dusty mixture") with "Unknown" in Figure 2.

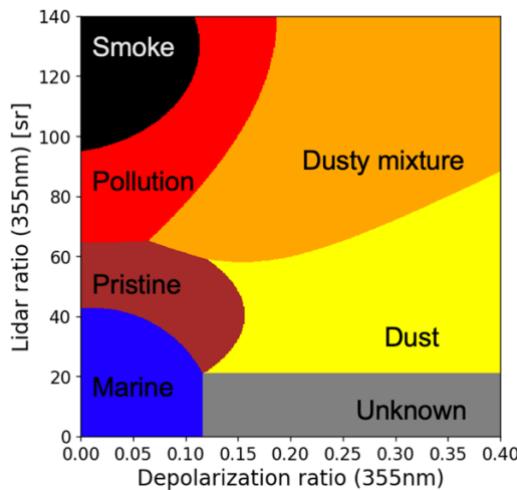



Figure 2. Lidar ratio – depolarization ratio diagram for classifying aerosol types.

The optical properties of the four aerosol types (smoke, pollution, marine, and pristine) are determined using cluster analysis of the AERONET level 2 product for years of 1992 to 2012. We perform a cluster analysis of the AERONET southern Africa sites for smoke type, that of Chinese sites for pollution type, and that of island sites for marine and pristine types. First, δ_a and S_a at 355, 532, and 1064 nm, which are usually observed by the HSRL and Raman lidar measurements, are calculated using the refractive indices at 440, 675, 870, and 1020 nm, size distributions of fine and coarse modes, and the sphericity of scattering light derived from the AERONET inversion algorithm (Dubovik et al., 2006) for each AERONET data sample. The refractive indices at 532 nm are interpolated and those at 355 and 1064 nm are extrapolated using those of AERONET product from 440 to 1020 nm. The non-spherical particle shape is assumed to be the AERONET spheroid model (Dubovik et al., 2006). Next, we adopt the fuzzy c-means method (Dunn, 1973; Bezdek, 1981) to conduct a cluster analysis. This method is based on minimizing the following objective function:

$$J = \sum_{i=1}^{N} \sum_{k=1}^{K} g_{ik}^2 \|x_i - c_k\|^2, \qquad (6)$$

where g_{ik} is the degree of membership of x_i to the k-th cluster, x_i is the observed data, c_k is the center of the k-th cluster. Partitioning is conducted through an iterative optimization of the J with updates of g_{ik} and c_k. Based on this cluster analysis, the center of cluster c_k is assumed to be the representative value of the selected cluster parameter. We select 12 parameters used in the cluster analysis to classify the aerosol properties. These parameters are δ_a and S_a at 355, 532, and 1064 nm and the imaginary part of the refractive index and fine mode fraction (FMF) to the total (fine+coarse) AOT at 440, 675, and 870 nm. Finally, we define the characteristic results of the cluster as the optical properties of the aerosol type.

The cluster, which is fine-mode dominated and light-absorbing aerosol, derived from the cluster analysis of the observations of AERONET sites located from 25°S to 35°S and 0°E to 40°E, which cover the source regions of African biomass burning aerosols, is assumed as the smoke type. The cluster, which is fine-mode dominated and light-absorbing aerosol, derived from the cluster analysis of the observations of AERONET sites located from 20°N to 40°N and 100°E to 125°E, which cover the source regions of Asian air pollution, is assumed as the pollution type. The marine and pristine types are derived from the cluster analysis of the observations of AERONET island sites far from aerosol source regions, where the marine type has the largest particle size with single scattering albedo (SSA) of 0.98 and the pristine type has a smaller particle size than the marine type, with the SSA of 0.98.

The difference in δ_a and S_a of non-spherical dust particles between observations and theoretical calculations remain large (Tesche et al., 2019), so that δ_a and S_a of the dust type at 355 nm are referred to as the averaged values of the Raman lidar observations in Morocco (Freudenthaler et al., 2009; Tesche et al., 2009), Germany (Wiegner et al., 2011), and Tajikistan (Hofer et al., 2017). δ_a of 0.24 ± 0.02 and S_a of 47 ± 8 sr for transported dust (Haarig et al., 2022) are also within the range of the dust type of the ATM. The dusty mixture type is defined as a mixture of dust and smoke in ratios of 0.65 and 0.35, respectively. Aerosol particles with high depolarization ratios and low lidar ratios are rarely observed; therefore, the unknown

aerosol type is defined as aerosols with $\delta_a > 0.12$ and $S_a < 21$ sr, which are determined by the values of the intersection of the border line between dust and pristine types and that between marine and pristine types, as shown in Figure 2.

The ATM consists of the major tropospheric aerosol types and the optical properties of the ATM are used in the radiative transfer calculation to estimate aerosol radiative effect (Yamauchi et al., 2024). For example, volcanic ash and stratospheric aerosols, which are not included in the ATM, are misclassified as one of the aerosol types of the ATM. This misclassification is one of the uncertainties of the estimated aerosol radiative effect. However, this uncertainty is smaller than the other uncertainties in the estimation of aerosol radiative effect, because the proportions of the volcanic ash and stratospheric aerosols are considerably small in the total aerosol amounts.

In the actual analysis using ATLID data, cloud types are classified together with aerosol types. Multiple cloud particle types such as warm water, supercooled water, two-dimensional ice, and three-dimensional ice, can be identified by applying a two-dimensional diagrammatic method of signal attenuation (or extinction coefficient) and depolarization ratio developed for CALIOP cloud type classification (Yoshida et al. 2010; Okamoto et al. 2024b).

3.4 Planetary boundary layer height

We adopted the WCT method because the WCT method is less affected by noise (Qu et al., 2017), and is promising for the spaceborne lidar (Kim et al., 2021). For the PBLH detection, lidar signals at wavelength of 532 or 1064 nm are mostly used; however, the ATLID wavelength is 355 nm. Due to relatively larger Rayleigh scattering at 355 nm (approximately five times larger than that at 532 nm), the proportion of Mie scattering by aerosols in the PBL in lidar backscatter becomes relatively smaller. In other words, the difference in 355 nm backscatter signals for aerosol-rich PBL and aerosol-poor free troposphere (FT) can be smaller. The larger Rayleigh scattering at 355nm also produces greater signal attenuation, resulting in the lower signal difference between the PBL and the FT. It should be noted that the signal attenuation due to Rayleigh scattering near the top of the PBL is generally larger for spaceborne lidar observations than for ground-based lidar observations. Thus, the detection of PBLH by spaceborne lidar observation at 355nm is more challenging than in the past, even when the WCT method is applied (Kim et al., 2021). In this study, instead of using attenuated backscatter signals as the input of the WCT method, the backscattering ratio (BR), which can be calculated as the ratio of Mie-attenuated backscatter to Rayleigh-attenuated backscatter from the ATLID L1 data (i.e., $BR = \beta_{atn.M} / \beta_{atn.R}$), is used to remove the effect of signal attenuation. Also, Rayleigh scattering component is eliminated by calculating BR minus 1 (i.e., $BR' = BR - 1$) to highlight aerosol scattering intensity in the PBL.

The BR' and FM are used as input data. Here, the target altitude (height above the surface) is set in between z_{min} and z_{max}, which are 0.1 and 5 km, respectively. If the target point is classified as a cloud in the FM at the target altitude, it is excluded from the PBLH detection. Then WCT is the calculated using the following equation:

$$WCT(a,b) = \frac{1}{a} \int_{z_{min}}^{z_{max}} BR'^{(z)} h\left(\frac{z-b}{a}\right) dz. \qquad (7)$$

where a and b are the dilation and the centered location of the Haar function h, which is defined as:

$$h\left(\frac{z-b}{a}\right) = \begin{cases} +1 : b - \frac{a}{2} \leq z < b, \\ -1 : b \leq z \leq b + \frac{a}{2}, \\ \quad 0 : \text{elswere.} \end{cases} \qquad (8)$$

In Equation 7, BR′ is normalized to 1.0 for the height from the surface to 1 km to reduce the dependency of aerosol concentrations. When the WCT exceeds a threshold value, the first WCT peak from z_{min} is determined as the PBLH. The threshold and dilation width are determined from a simulated BR′ profile based on ground-based HSRL data (Jin et al., 2020) with random noise according to the errors in the ATLID attenuated backscatter reported by do Carmo et al., (2021), and are set to 0.2 and 1.0 km, respectively.

4 Results and Discussion

To demonstrate the products and the performance of the algorithms, the ATLID L1 data (i.e., Equations 1a-1c) simulated using the Joint Simulator for Satellite Sensors (Joint-Simulator) (Hashino et al. 2013, Satoh et al. 2016, Row et al. 2023) were used. Cloud (and precipitation) distributions simulated using the Nonhydorstatic Icosahedral Atmospheric Model (NICAM; Satoh et al., 2014) and aerosol distributions simulated using NICAM Spectral Radiation Transport Model for Aerosol Species (NICAM-SPRINTARS; Takemura et al., 2000) were used as the input data. Signal noises such as shot noise, dark noise, and CCD read-out noise expected from the ATLID system were evaluated and added as gaussian random noise.

The ATLID L1 data for the cloud-predominant case simulated along the satellite path using Joint-Simulator are shown in Figure 3. The cloud and aerosol fields used for the signal simulation are also shown in Figure 4. Clouds are present from the surface to the altitude of 15 km. The SNR of the Mie co-polar- and cross-polar attenuated backscatter coefficients for the cloud layers are generally greater than 5 for 0.3 km horizontal resolution data; the SNR for 1 km horizontal resolution data is approximately twice than that for 0.3 km horizontal resolution data. The FM algorithm (Section 3.1) was applied to the simulated L1 data (Figure 5). To assess the effect of signal noise on the estimates, the algorithm was applied to L1 data with or without signal noise, and the agreement between them was evaluated. The results for the 0.3 km and 1 km horizontal resolution data showed generally good agreement for each layer type, such as "cloud" and "clear-sky or aerosol." For the 0.3 km horizontal resolution data, the total number of layers for the cloud type was 226,000, of which the number of misidentification was 24,000, corresponding to a 11% relative error. For the 1 km horizontal resolution data, the misidentification of the cloud layers was 9%. For the clear-sky or aerosol type, the misidentification reached to 41% for the 0.3 km horizontal resolution data; however, it improved to 5% for the 1 km horizontal resolution data. Many misidentification were noted at the edges of the cloud layers, highlighting the need for improvement is this aspect in future studies.

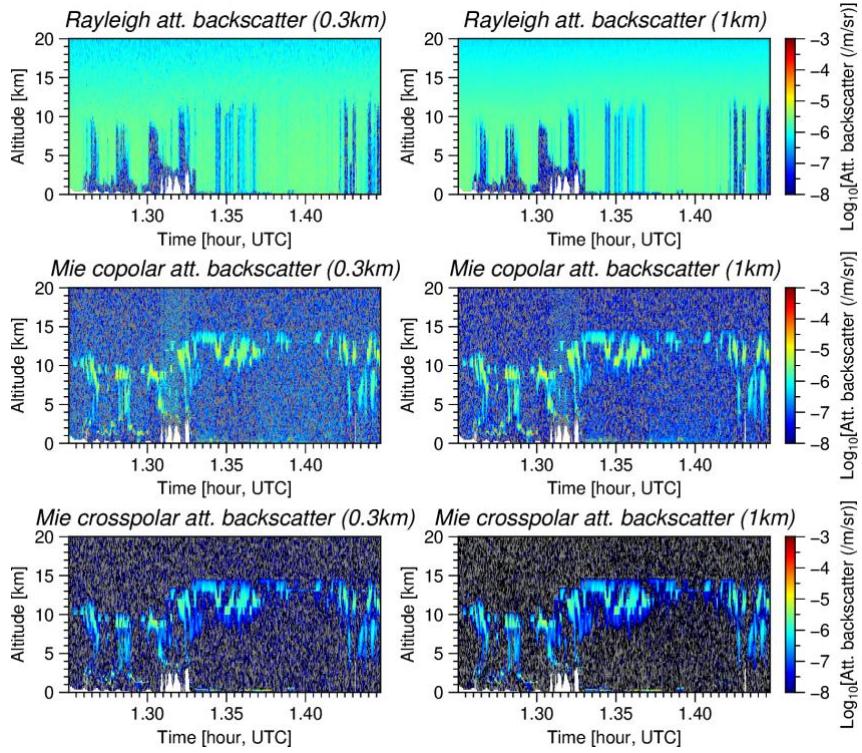
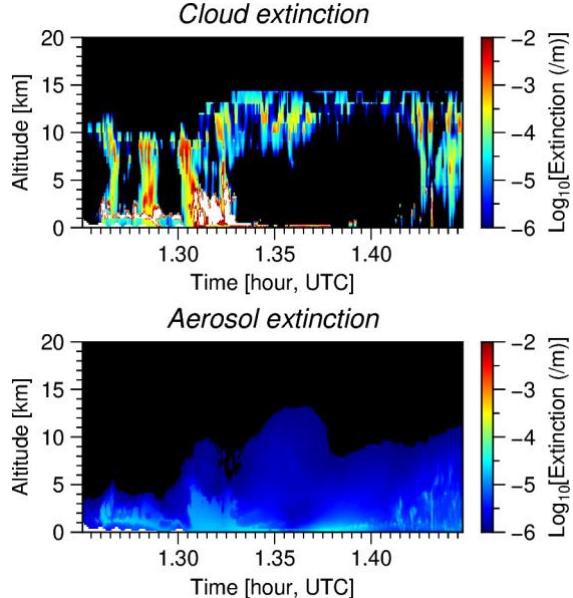
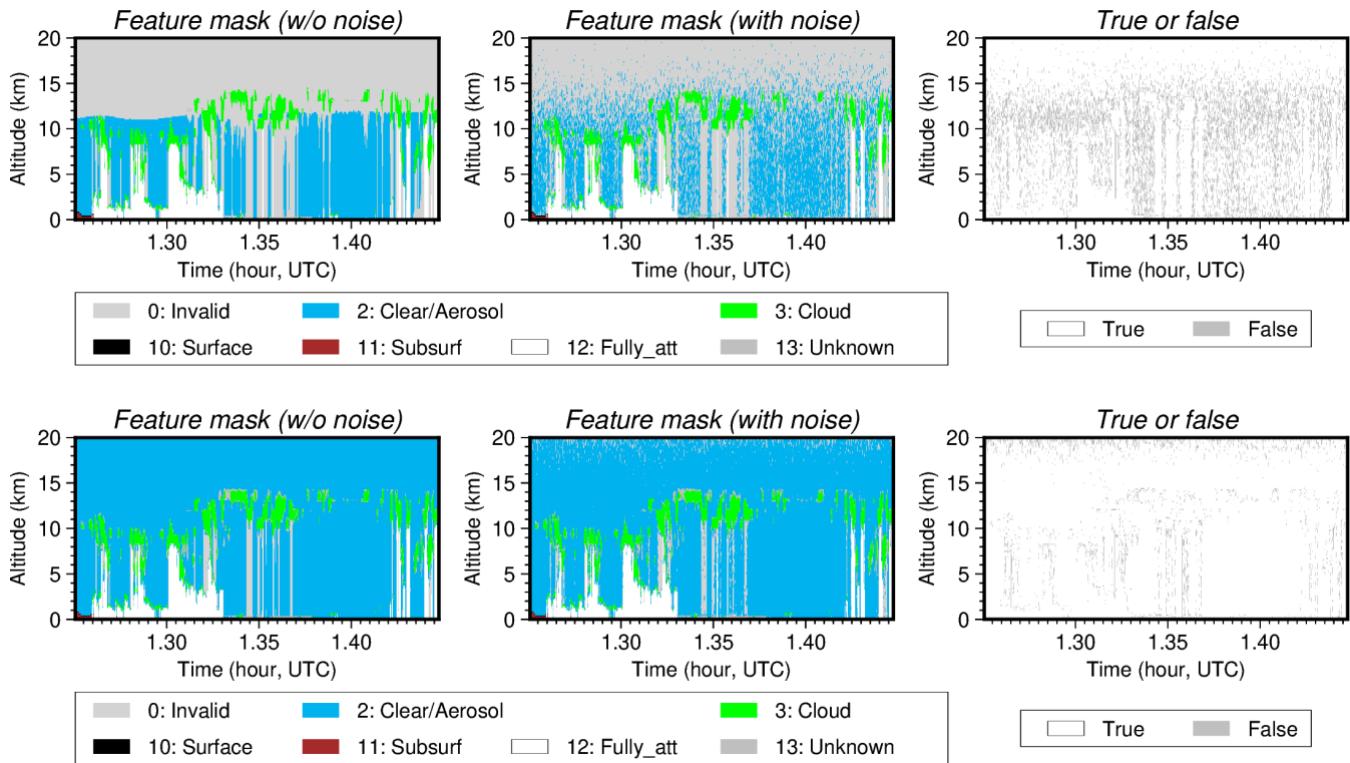



Figure 3. Rayleigh-, Mie co-polar-, and Mie cross-polar attenuated backscatter coefficients at 355 nm for cloud predominant scenes simulated using Joint-Simulator. The left/right figure shows the 0.3/1 km horizontal resolution data.

355 Figure 4. Extinction coefficient at 355 nm of clouds and aerosols used for the simulations for cloud predominant case (Figure 3).

360 Figure 5. Feature mask products estimated for data with/without signal noise for cloud predominant scenes (Figure 3) and its true or false values. The upper/lower figures show the results for the 0.3/1 km horizontal resolution data.

The results of the aerosol layer identification are shown in Figure 7. Here, we used the the L1 data simulated for the aerosol-predominant case, where the dust layer is widely suspended over an altitude range of 3-20 km (Figure 6), and the FM algorithm
365 was applied to the L1 data with or without noise. The results for the data with signal noise are generally in good agreement with those for the data without signal noise. The total number of aerosol layers was approximately 280,000 and the number of misidentified layers was 30,000, resulting in 11% misidentification. As in the cloud-predominant case (Figure 5), much of the misidentification occurred at the edges of the aerosol layers, which is also an aspect that can be improved for aerosol identification.

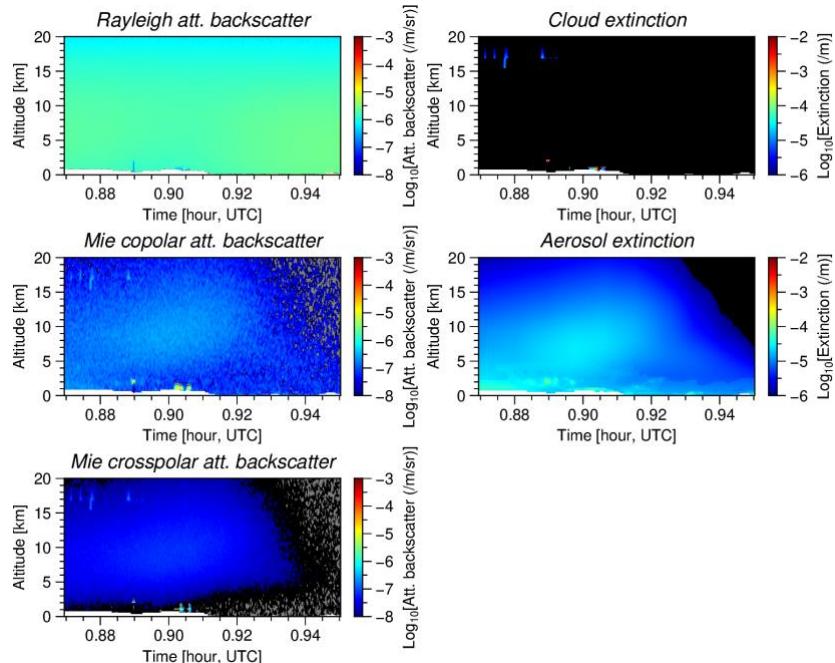



Figure 6. Rayleigh-, Mie co-polar-, and Mie cross-polar attenuated backscatter coefficients at 355 nm for aerosol predominant scenes simulated using Joint-Simulator and extinction coefficients at 355 nm of aerosols and clouds. The simulations shown used 1* km horizontal resolution data.

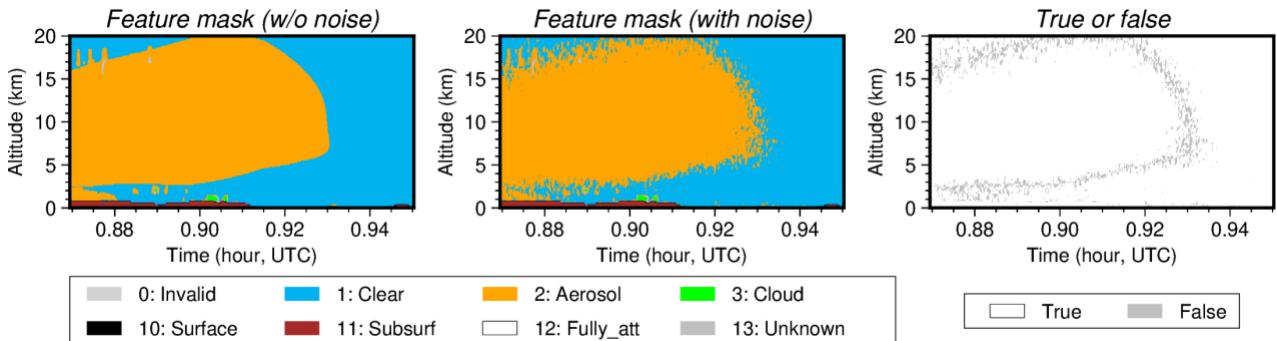



Figure 7. Feature mask products estimated for data with/without signal noise for aerosol predominant cases (Figure 6) and its true or false values.

The results of the AOP retrieval are shown in Figure 8. The AOP retrieval algorithm was applied to the aerosol-predominant
 case, as shown in Figure 6. The SNRs of the L1 data for the dust layer generally ranged from 5 to 20 for 1[*] km horizontal resolution data. The estimated backscatter coefficients and depolarization ratios generally agreed well with the actual values, which were the aerosol optical properties used in the L1 simulation. The mean of retrieved aerosol backscatter coefficient was

$3.26 \times 10^{-7} \, \text{m}^{-1}\text{sr}^{-1}$. The true value was $3.18 \times 10^{-7} \, \text{m}^{-1}\text{sr}^{-1}$ and the mean error (ME: retrieval − truth) was $0.08 \times 10^{-7} \, \text{m}^{-1}\text{sr}^{-1}$, corresponding to a relative error of 2%. The root mean square error (RMSE) was $1.12 \times 10^{-7} \, \text{m}^{-1}\text{sr}^{-1}$, corresponding to a relative

385 error of 34%. For the depolarization ratio, the means for the retrievals and true values were 0.27 and 0.26, respectively; ME = 0.01 (4%) and RMSE = 0.07 (27%). The figure also shows that the extinction coefficient and lidar ratio were more strongly affected by the signal noise than the backscatter coefficient and depolarization ratio. For the extinction coefficient, the means for the retrieval and true values were $1.32 \times 10^{-5} \, \text{m}^{-1}$ and $1.35 \times 10^{-5} \, \text{m}^{-1}$, respectively; ME = $-0.03 \, \text{m}^{-1}$ (2%) and RMSE = $1.05 \times 10^{-5} \, \text{m}^{-1}$ (78%). For the lidar ratio, the means of the retrieval and true values were 41 sr^{-1}, respectively; ME = 0 sr^{-1} (0%) and

390 RMSE = 25 sr^{-1} (61%). As with the backscatter coefficient and depolarization ratio, there was no significant bias error (ME) for the extinction coefficient and lidar ratio; however, the variation (RMSE) was relatively large owing to signal noise. Notably, the aerosol concentration used in this analysis was relatively low (i.e., $1.35 \times 10^{-5} \, \text{m}^{-1}$ on an average), and the RMSE value of the extinction coefficient is not large (i.e., $1.05 \times 10^{-5} \, \text{m}^{-1}$). For cases with higher aerosol concentrations, it can be expected that the relative error of the extinction coefficient and the RMSE of the lidar ratio decrease with an increase in the SNR of the

395 Mie-attenuated backscatter coefficient (e.g., Nishizawa et al. 2017).

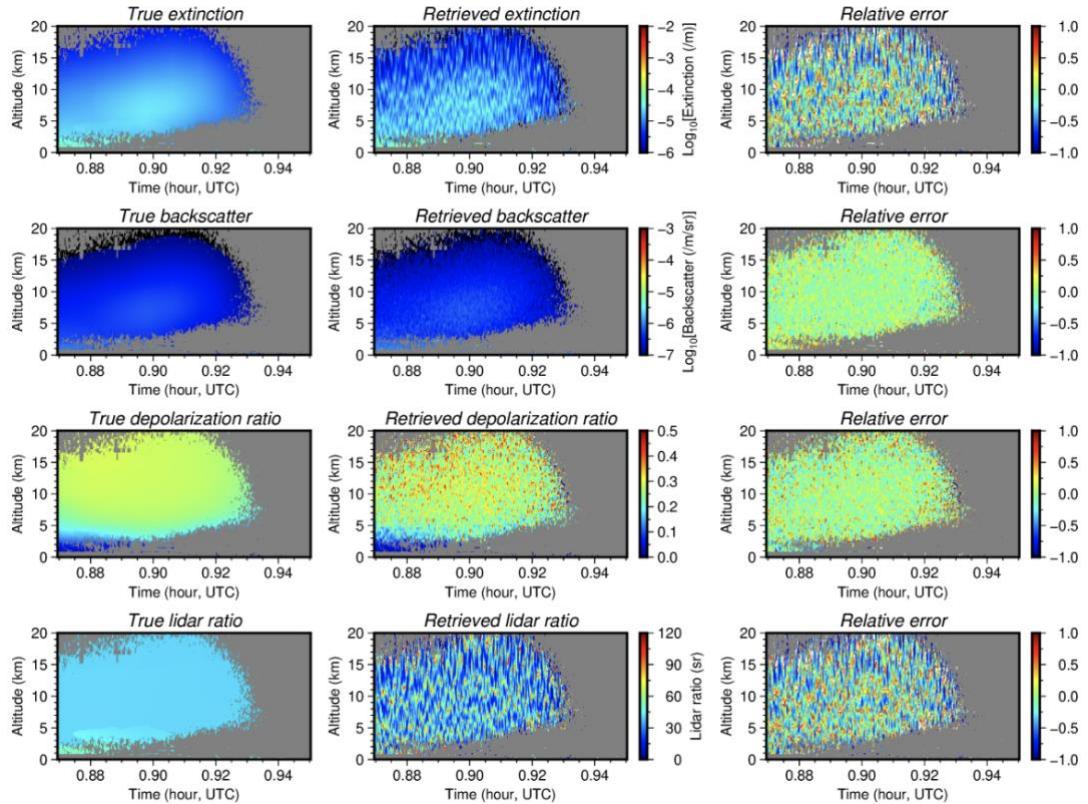



Figure 8. Extinction coefficient, backscatter coefficient, depolarization ratio, and lidar ratio of aerosols estimated for aerosol predominant cases (Figure 6) and its relative error.

The results of the aerosol-type classification are shown in Figure 9. The algorithm was applied to the L1 data with and without signal noise for the aerosol-predominant case, as shown in Figure 6. For the dust layer, there appeared some misclassification (e.g., Dusty-mixture and Unknown); however, there was a fair amount of agreement. The total number of layers classified as Dust type was approximately 225,000 and the number of misidentified layers was 84,000, resulting in a 37% misidentification. In addition, aerosol layers classified as Pristine type were also found below an altitude of 4 km; the total number of pristine layers was approximately 13,000 and the number of misidentified layers was 9,700, resulting in 74% misidentification. Since aerosol typing was classified using a two-dimensional diagram of the lidar and depolarization ratios (Figure 2), aerosol type misclassification is directly reflected by the errors in the lidar ratio and depolarization ratio. The error in the lidar ratio is considerably larger than that in the depolarization ratio (about 25 sr in RMSE). Therefore, the retrieved lidar ratios for the dust layer fall outside the range of the two-dimensional diagram for "dust", resulting in cases of misclassification as dusty-mixture or unknown. Similarly, pristine is misclassified as pollution or marine due to the errors in the lidar ratio. Thus, accurate estimation of the lidar ratio as well as depolarization ratio is required for aerosol type classification. For cases with higher aerosol concentrations, it can be expected that the misclassification of aerosol types decreases with an decrease of the retrieval error of the lidar ratio, as discussed in the AOP retrieval (Section 4).

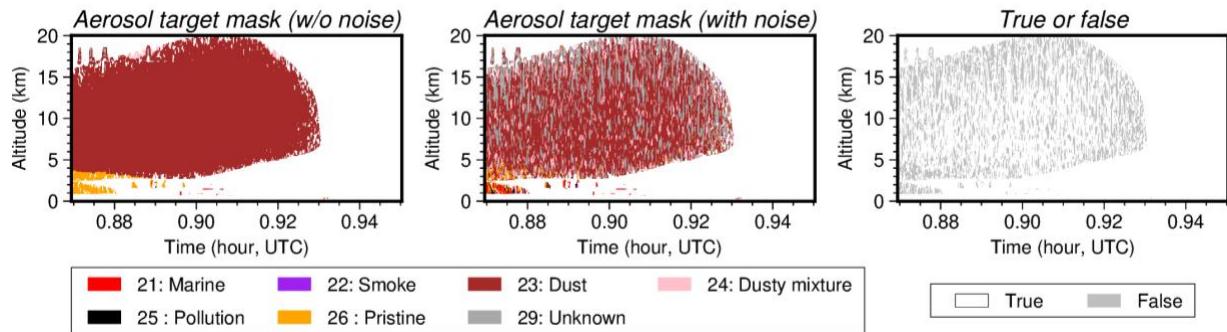



Figure 9. Aerosol target mask products estimated for data with/without signal noise for aerosol predominant scenes (Figure 6) and its true or false values.

Figure 10 shows an example of the PBLH detection for the L1 data simulated using Joint-Simulator, where the PBLH was approximately 2.1 km. The WCT was calculated using a Haar function with a dilation width of 1.0 km, and a WCT peak exceeding a threshold of 0.2, derived at 2.1 km. When signal noise was added to the L1 data (dotted line in Figure 10), another WCT peak at ~1 km appeared because of signal fluctuation; however, the peak was below the threshold. In this case, the estimated PBLH with noise agreed well with that without noise. In actual observations, the PBLH detection may be difficult for cases of noisy data, low aerosol concentrations, and the existence of residual layers at night because the gap of BR′ for the PBL and free troposphere may be small.

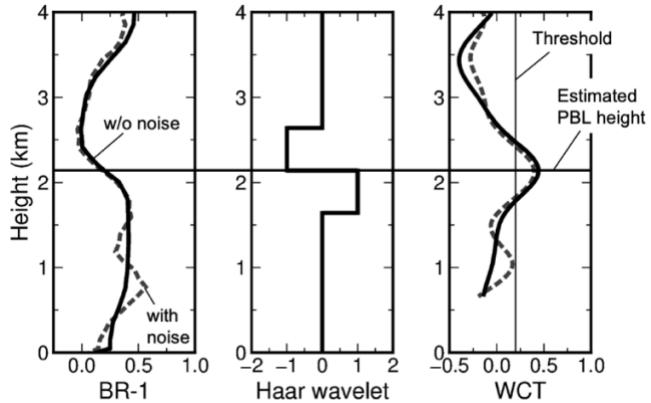



Figure 10. Vertical profiles of (left) backscattering ratio (BR) minus 1 for the simulated ATLID L1 data, (center) Haar function, and (right) wavelet covariance transform (WCT).

5 Conclusion

We developed algorithms to produce JAXA ATLID L2 products for aerosols and clouds using ATLID L1 data. The algorithms retrieve the following products: (1) feature mask; (2) particle optical properties such as extinction coefficient, backscatter coefficient, depolarization ratio, and lidar ratio at 355 nm; (3) target mask; and (4) planetary boundary layer height. The performance of the algorithms was demonstrated using pseudo-ATLID L1 data with realistic signal noise, which were calculated according to the ATLID specification for the aerosol or cloud dominated cases by Joint-Simulator.

The main findings of this simulation study are as follows: (1) the misidentification of the aerosol and cloud layers by the feature mask algorithm was relatively low (approximately 10%). (2) The retrieval errors of the AOP were $0.08 \times 10^{-7} \pm 1.12 \times 10^{-7}\,\mathrm{m^{-1}sr^{-1}}$ ($2 \pm 34\%$ relative error) for the backscatter coefficient and 0.01 ± 0.07 ($4 \pm 27\%$ relative error) for the depolarization ratio of aerosols; the relative errors of the extinction coefficient and lidar ratio were worse than those of the backscatter coefficient and depolarization ratio. (3) The aerosol-type classification generally performed well, with 37% misclassifications for the dust type, although there were more misclassifications for the pristine type. These misidentifications were mainly due to errors in the estimation of the lidar ratios. Accurate estimation of the lidar ratio as well as depolarization ratio is essential for aerosol type classification. (4) PBLH retrieval using the WCT method with ATLID L1 data was feasible. These results indicate that the algorithm's capability to provide valuable insights into the global distribution of aerosols and clouds, facilitating assessments of their climate impact through atmospheric radiation processes.

The validation of these aerosol and cloud products is essential, and JAXA and ESA will conduct validation using various platforms, such as airborne, shipborne, and ground-based measurements, and various observation instruments, including lidars. The validation studies include assuarance of the product accuracy and the improvement of the L2 algorithms mentioned above, including the verification and improvement of various assumptions (e.g., optical models of aerosols and clouds) and the criteria used in the algorithms.

To understand the climatic effects of aerosols, it is essential to understand the properties of individual aerosol components as well as the properties of total aerosols such as the JAXA ATLID L2 products described above. Therefore, we are currently developing an algorithm to estimate the extinction coefficients of dust, sea salt, carbonaceous (light-absorbing particles), and water-soluble aerosols at 355nm using the difference in the depolarization and light absorption properties of each aerosol component from the L2 products and ATLID L1 data. We are also currently developing an ATLID-MSI synergy algorithm to retrieve the vertical mean mode-radii of dust and fine-mode aerosols, and the extinction coefficients of the abovementioned four aerosol components. These algorithms are being developed based on the aerosol component retrieval algorithms that have been developed for the analysis of CALIOP and ground-based lidar data. (Kudo et al. 2023; Nishizawa et al. 2007, 2008, 2011, 2017). The aerosol component products derived by these algorithms using ATLID and MSI data are planned to be released as JAXA's L2 products and are expected to enhance the scientific value of the EarthCARE mission.

Author contributions.

TN developed the algorithm concept, a feature mask scheme, and managed the project. RK developed a scheme for retrieving the particle optical properties. EO developed a scheme to classify the aerosol types. YJ developed a scheme for estimating the planetary boundary layer height. AH, EO, YJ, and RK constructed code integrating the individual algorithms and performed an error analysis using the simulated data. NS, KS, and HO generated ideas that contributed to improvements in the retrieval schemes. TN prepared the paper with contributions from all the co-authors.

Competing interests.

The contact author has declared that none of the authors has any competing interests.

Special issue statement.

This article is part of the special issue "EarthCARE Level 2 algorithms and data products." However, this is not associated with conferences.

Acknowledgments.

The authors would like to thank D. Donovan and G. J. van Zadelhoff for their support and advice regarding ATLID signal-noise calculations. We thank the principal investigators and staff at the AERONET sites used in this study for maintaining the stations. We thank the members of the JAXA EarthCARE Science Team and the Joint-Simulator project. We would like to thank Editage (www.editage.jp) for English language editing.

Financial support.

480   This research was supported by the EarthCARE satellite study commissioned by the Japan Aerospace Exploration Agency (grant no. 23RT000223) and a Grant-in-Aid for Scientific Research (KAKENHI) (project numbers: JP18KK0289, JP17H06139, JPS17H04477, JP15H01728, JP15H02808, and JP25220101).

References

485   Bezdek, J. C.: Pattern recognition with fuzzy objective function algorithms, Plenum Press, 1981.

Brooks, I. M.: Finding Boundary Layer Top: Application of a wavelet covariance transform to lidar backscatter profiles, J. Atmos. Oceanic Technol., 20, 1092-1105, 2003.

Cairo, F., Muro, M. D., Cnels, Marcel, Liberto, L. D., Bucci, S., Legras, B., Kottayil, A., Soccione, A., Chisu, S.: Lidar observations of cirrus clouds in Palau ((7°33′N, 134°48′E), Atmos. Chem. Phys., 21, 7947-7961,
490   https://doi.org/10.5194/acp-21-7947-2021, 2021.

do Carmo, J. P., de Villele, G., Wallace, K., Lefebvre, A., Chose, K., Kanitz, T., Chassat, F., Corselle, B., Belhadj, T., 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, 76, 2021.

Chen, W.-N., Chiang, C.-W., Nee, J.-B.: Lidar ratio and depolarization ratio for cirrus clouds, Appl. Opt., 41, 6470-6476,
495   2002.

Donovan, D. P., Kollias, P., Velázquez Blázquez, A., and van Zadelhoff, G.-J.: The generation of EarthCARE L1 test data sets using atmospheric model data sets, Atmos. Meas. Tech., 16, 5327-5356, https://doi.org/10.5194/amt-16-5327-2023, 2023.

Donovan, D. P., 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, Atmos. Meas. Tech., 17, 5301-5340, 2024.

500   Dubovik, O., Sinyuk, A., Lapyonok, T., Holben, B. N., Mishchenko, M., Yang, P., Eck, T. F., Volten, H., Muñoz, O., and Veihelmann, B.: Application of spheroid models to account for aerosol particle nonsphericity in remote sensing of desert dust, J. Geophys. Res. Atmos., 111, D11208, https://doi.org/10.1029/2005JD006619, 2006.

Dunn, J. C.: A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, J. of Cybernetics, 3, 3, 32-57, 1973.

505   Eisinger, M., Marnas, F., Wallace, K., Kubota, T., Tomiyama, N., Ohno, Y., Tanaka, T., Tomita, E., Wehr, T., and Bernaerts, D.: The EarthCARE mission: science data processing chain overview, Atmos. Meas. Tech., 17, 839-862, 2024.

Fang, H. T. and Huang D, S.: Noise reduction in lidar signal based on discrete wavelet transform, Opt.Commun., 233, 1, 67–76, 2004.

Freudenthaler, V., Esselborn, M., Wiegner, M., Heese, B., Tesche, M., Ansmann, A., Müller, D., Althausen, D., Wirth, M.,
510   and Fix, A.: Depolarization ratio profiling at several wavelengths in pure Saharan dust during SAMUM 2006, Tellus B, 61, 165–179, https://doi.org/10.1111/j.1600-0889.2008.00396.x, 2009.

Giles, D. M., Sinyuk, A., Sorokin, M. G., Schafer, J. S., Smirnov, A., Slutsker, I., Eck, T. F., Holben, B. N., Lewis, J. R., Campbell, J. R., Welton, E. J., Korkin, S. V., and Lyapustin, A. I.: Advancements in the Aerosol Robotic Network

(AERONET) Version 3 database – automated near-real-time quality control algorithm with improved cloud screening for

515   Sun photometer aerosol optical depth (AOD) measurements, Atmos. Meas. Tech., 12, 169–209, https://doi.org/10.5194/amt-12-169-2019, 2019.

Haarig, M., Ansmann, A., Engelmann, R., Baars, H., Toledano, C., Torres, B., Althausen, D., Radenz, M., and Wandinger, U.: First triple-wavelength lidar observations of depolarization and extinction-to-backscatter ratios of Saharan dust, Atmos. Chem. Phys., 22, 355–369, https://doi.org/10.5194/acp-22-355-2022, 2022.

520   Hagihara, Y., Okamoto, H., and Yoshida, R.: Development of a combined CloudSat-CALIPSO cloud mask to show global cloud distribution, J. Geophys. Res., 115, D00H33, https://doi.org/10.1029/2009JD012344, 2010.

Hashino, T., Satoh, M., Hagihara, Y., Kubota, T., Matsui, T., Nasuno, T., and Okamoto, H.: Evaluating cloud microphysics from NICAM against CloudSat and CALIPSO, J. Geophys. Res. Atmos., 118, 7273–7292, https://doi.org/10.1002/jgrd.50564, 2013.

525   Héliere, A., Gelsthorpe, R., Le Hors, L., and Toulemont, Y.: ATLID, the Atmospheric Lidar on board the EarthCARE Satellite, Proc. SPIE 10564, International Conference on Space Optics - ICSO 2012, 105642D, https://doi.org/10.1117/12.2309095, 2017.

Hofer, J., Althausen, D., Abdullaev, S. F., Makhmudov, A. N., Nazarov, B. I., Schettler, G., Engelmann, R., Baars, H., Fomba, K. W., Müller, K., Heinold, B., Kandler, K., and Ansmann, A.: Long-term profiling of mineral dust and pollution aerosol

530   with multiwavelength polarization Raman lidar at the Central Asian site of Dushanbe, Tajikistan: case studies, Atmos. Chem. Phys., 17, 14559–14577, https://doi.org/10.5194/acp-17-14559-2017, 2017.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y., Nakajima, T., Lavenue, F., Jankowiak, I., and Smirnov, A.: AERONET – A federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1–16, https://doi.org/10.1016/S0034-4257(98)00031-5, 1998.

535   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, B. Am. Meteorol. Soc., 96, 1311–1332,

540   https://doi.org/10.1175/BAMS-D-12-00227.1, 2015.

Irbah, A., Delanoë, J., van Zadelhoff, G.-J., Donovan, D. P., Kollias, P., Puigdomènech Treserras, B., Mason, S., Hogan, R. J., and Tatarevic, A.: The classification of atmospheric hydrometeors and aerosols from the EarthCARE radar and lidar: the A-TC, C-TC and AC-TC products, Atmospheric Measurement Techniques, 16, 2795–2820, https://doi.org/10.5194/amt-16-2795-2023, 2023.

545   Jin, Y., Nishizawa, T., Sugimoto, N., Ishii, S., Aoki, M., Sato, K., and Okamoto, H.: Development of a 355-nm high-spectral-resolution lidar using a scanning Michelson interferometer for aerosol profile measurement, Opt. Express, 28(16), 23209-23222, 2020.

Kim, M.-H., Yeo, H., Park, S., Park, D.-H., Omar, A., Nishizawa, T., Shimizu, A., and Kim, S.-W.: Assessing CALIOP-derived planetary boundary layer height using ground-based lidar, Remote Sens., 13, 1496, 2021.

550 Koren, I., Y. J. Kaufman, L. A. Remer, and J. V. Martins, Measurement of the effect of Amazon smoke on inhibition of cloud formation, Science, 303, 1342–1344, 2004.

Kudo, R., Nishizawa, T., and Aoyagi, T.: Vertical profiles of aerosol optical properties and the solar heating rate estimated by combining sky radiometer and lidar measurements, Atmos. Meas. Tech., 9, 3223–3243, https://doi.org/10.5194/amt-9-3223-2016, 2016.

555 Kudo, R., Higurashi, A., Oikawa, E., Fujikawa, M., Ishimoto, H., and Nishizawa, T.: Global 3-D distribution of aerosol composition by synergistic use of CALIOP and MODIS observations, Atmos. Meas. Tech., 16, 3835–3863, https://doi.org/10.5194/amt-16-3835-2023, 2023.

Lammert, A., and Bösenberg, J.: Determination of the convective boundary-layer height with laser remote sensing, Boundary-Layer Meteorol., 119, 159-170, 2006.

560 Liu, Z., Matsui, I., Sugimoto, N.: High-Spectral-Resolution Lidar using an iodine absorption filter for atmospheric measurements, Optical Engineering, 38, 1661-1670, 1999.

Liu, Z., Vaughan, M. A., Winker, D. M., Kittaka, C., Getzewich, B. J., Kuehn, R. E., Hostetler, C. A.: The CALIPSO lidar cloud and aerosol discrimination: Version 2 algorithm and initial assessment of performance. J. Atmos. Oceanic Tchn., 26, 1198–1213. https://doi.org/10.1175/2009JTECHA1229.1, 2009.

565 Menon, S., J. Hansen, L. Nazarenko, Y. Luo, Climate effects of black carbon aerosols in China and India, Science, 297, 2250–2253, 2002.

Menut, L., Flamant, C., Pelon, J., and Flamant P. H.: Urban boundary-layer height determination from lidar measurements over the Paris area, Appl. Opt., 38(6), 945-954, 1999.

Müller, D., Hostetler, C. A., Ferrare, R. A., Burton S. P., Chemyakin, E., Kolgotin, A., Hair, J. W., Cook, A. L., Harper, D.B.,
570 Rogers, R. R., Hare, R. W., Cleckner, C. S., Obland, M. D., Tomlinson, J., Berg, L. K., and Schmid, B.: Airborne Multiwavelength High Spectral Resolution Lidar (HSRL-2) observations during TCAP 2012: vertical profiles of optical and microphysical properties of a smoke/urban haze plume over the northeastern coast of the US, Atmos. Meas. Tech., 7. 3487-3496, 2014.

Nakatsuka, H., Ohno, Y., Koopman, R., Rusli, S., Kikuchi, M., Eisinger, M., Tanaka, T., Taga, M., Deghaye, P., Tomita, E.,
575 and Bernaerts, D.: The EarthCARE mission – science and system overview, Atmos. Meas. Tech., 16, 3581–3608, https://doi.org/10.5194/amt-16-3581-2023, 2023.

Nishizawa, T., Okamoto, H., Sugimoto, N., Matsui, I., Shimizu, A, Aoki, K.: An algorithm that retrieves aerosol properties from dual-wavelength polarized lidar measurements, J. Geophys. Res., 112, D06212, 2007.

Nishizawa, T., Sugimoto, N., Matsui, I., Shimizu, A., Tatarov, B., and Okamoto, H.: Algorithm to retrieve aerosol optical
580 properties from high-spectral-resolution lidar and polarization Mie-scattering lidar measurements, IEEE T. Geosci. Remote, 46, 4094–4103, 2008.

Nishizawa, T., Sugimoto, N., Matsui, I., Shimizu, A., and Okamoto, H.: Algorithms to retrieve optical properties of three component aerosols from two-wavelength backscatter and one-wavelength polarization lidar measurements considering nonsphericity of dust, J. Quant. Spectrosc. Radiat. Transfer, 112, 254–267, 2011.

585 Nishizawa, T., Sugimoto, N., Matsui, I., Shimizu, A., Hara, Y., Itsushi, U., Yasunaga, K., Kudo, R., Kim, S.-W.: Ground-based network observation using Mie–Raman lidars and multi-wavelength Raman lidars and algorithm to retrieve distributions of aerosol components, J. Quant. Spectrosc. Radiat. Transfer, 188, 79-93, https://doi.org/10.1016/j.jqsrt.2016.06.031., 2017,

Nocedal, J., and Wright, S. J.: Numerical optimization, 2nd edition, Springer Series in Operations Research and Financial

590 Engineering, 664 pp., Springer Science+Busiginess Media, LCC, New York, 2006.

Oikawa E., Nakajima T., Inoue T., Winker D.: A study of the shortwave direct aerosol forcing using ESSP/CALIPSO observation and GCM simulation, J. Geophys. Res. Atmos., 118, 3687-3708, diu:19.1992/jgrd.50227, 2013.

Oikawa, E., Nakajima, T., and Winker, D.: An evaluation of the shortwave direct aerosol radiatve forcing using CALIOP and MODIS observations, J. Geophys. Res., 123, 1211–1233, https://doi.org/10.1002/2017JD027247, 2018.

595 Okamoto, H., Nishizawa, T., Takemura, T., Kumagai, H., Kuroiwa, H., Sugimoto, N., Matsui, I., Shimizu, A., Emori, S., Kamei, A., and Nakajima, T.: Vertical cloud structure observed from shipborne radar and lidar: Midlatitude case study during the MR01/K02 cruise of the research vessel Mirai, J. Geophys. Res., 112, D08216, https://doi.org/10.1029/2006JD007628, 2007.

Okamoto, H., Nishizawa, T., Takemura, T., Sato, K., Kumagai, H., Ohno, Y., Sugimoto, N., Shimizu, A., Matsui, I., and

600 Nakajima, T.: Vertical cloud properties in the tropical western Pacific Ocean: Validation of the CCSR/NIES/FRCGC GCM by shipborne radar and lidar, J. Geophys. Res.-Atmos., 113, D24213, https://doi.org/10.1029/2008JD009812, 2008.

Okamoto, H., Sato, K., Borovoi, A., Ishimoto, H., Masuda, K., Konoshonkin, A., Kustova, N.:Interpretation of lidar ratio and depolarization ratio of ice clouds using spaceborne high-spectral-resolution polarization lidar, Optics Express, 27, 36587-36600, 2019.

605 Okamoto, H., Sato, K., Nishizawa, T., Jin, Y., Nakajima, T., Wang, M., Satoh, M., Suzuki, K., Roh, W., Yamauchi, A., Horie, H., Ohno, Y., Hagihara, Y., Ishimoto, H., Kudo, R., Kubota, T., and Tanaka, T.: JAXA Level2 algorithms for EarthCARE mission from single to four sensors: new perspective of cloud, aerosol, radiation and dynamics, Atmos. Meas. Tech. Discuss. [preprint], https://doi.org/10.5194/amt-2024-101, in review, 2024a.

Okamoto, H., Sato, K., Nishizawa, T., Jin, Y., Ogawa, S., Ishimoto, H., Hagihara, Y., Oikawa, E., Kikuchi, M., Satoh, M., and

610 Roh, W.: Cloud masks and cloud type classification using EarthCARE CPR and ATLID, Atmos. Meas. Tech. Discuss. [preprint], https://doi.org/10.5194/amt-2024-103, in review, 2024b.

Omar, A. H., J.-G. Won, D. M. Winker, S.-C. Yoon, O. Dubovik, and M. P. McCormick: Development of global aerosol models using cluster analysis of Aerosol Robotic Network (AERONET) measurements. J. Geophys. Res., 110, D10S14, doi:10.1029/2004JD004874, 2005.

615 Omar, A. H., Winker, D. M., Kittaka, C., Vaughan, M. A., Liu, Z., Hu, Y., Trepte, C. R., Rogers, R. R., Ferrare, R. A., Lee, K.-P., Kuehn, R. E., and Hostetler, C. A.: The CALIPSO automated aerosol classification and lidar ratio selection algorithm, J. Atmos. Ocean. Tech., 26, 1994–2014, https://doi.org/10.1175/2009JTECHA1231.1, 2009.

Qu, Y., Han, Y., Wu, Y., Gao, P., and Wang, T.: Study of PBLH and its correlation with particulate matter from one-year observation over Nanjing, Southeast China, Remote Sens., 9, 668, 2017.

620 Roh, W., Satoh, M., Hashino, T., Matsugishi, S., Nasuno, T., Kubota, T.: Introduction to EarthCARE syntheric data using a global storm-resolving simulation, Atmos. Meas. Tech., 16, 3331-3344, https://doi.org/10.5194/amt-2023-16-3331-2023, 2023.

Sato, K., Okamoto, H., and Ishimoto, H., Physical model for multiple scattered space- borne lidar returns from clouds, Opt. Exp., 26, 322629, doi.org/10.1364/OE.26.00A301, 2018.

625 Sato, K., Okamoto, H., and Ishimoto, H., Modeling the depolarization of scapce-borne lidar signals, 27, 345166, doi.org/10.1364/OE.27.00A117, 2019.

Satoh, M., Tomita, H., Yashiro, H., Miura, H., Kodama, C., Seiki, T., Noda, A. T., Yamada, Y., Goto, D., Sawada, M., Miyoshi, T., Niwa, Y., Hara, M., Ohno, T., Iga, S., Arakawa, T., Inoue, T., and Kubokawa, H.: The non-hydrostatic icosahedral atmospheric model: description and development, Prog. Earth Planet. Sci., 1, 18. doi:10.1186/s40645-014-0018-1, 2014.

630 Satoh, M., Roh, W., and Hashino, T.: Evaluations of clouds and precipitations in NICAM using the Joint Simulator for Satellite Sensors, CGER Supercomput. Monogr. Rep., 22, 110, https://doi.org/CGER-I127-2016, 2016.

Shimizu, A., Sugimoto, N., Matsui, I., Arao, K., Uno, I., Murayama T., Kagawa, N., Aoki, K., Uchiyama, A., and Yamazaki, A.: Continuous observations of Asian dust and other aerosols by polarization lidars in China and Japan during ACE-Asia. J. Geophys. Res., 109 , D19S17. doi:10.1029/2002JD003253, 2004.

635 Takemura, T., T. Nakajima, A. Higurashi, S. Ohta, and N. Sugimoto, Aerosol distributions and radiative forcing over the Asian-pacific region simulated by Spectral Radiation-Transport Model for Aerosol Species (SPRINTARS), J. Geophys. Res., 108(D23), 8659, doi:10.1029/2002JD003210, 2002.

Takemura, T., Okamoto, H., Maruyama, Y., Numaguti, A., Higurashi, A., and Nakajima, T.: Global three-dimensional simulation of aerosol optical thickness distribution of various origins, J. Geophys. Res. Atmos., 105(D14), 17853–17873, 640 2000.

Tesche, M., Ansmann, A., Müller, D., Althausen, D., Mattis, I., Heese, B., Freudenthaler, V.,Wiegner, M., Esselborn, M., Pisani, G., and Knippertz, P.: Vertical profiling of Saharan dust with Raman lidars and airborne HSRL in southern Morocco during SAMUM, Tellus B, 61, 144–164, https://doi.org/10.1111/j.1600-0889.2008.00390.x, 2009.

Tesche, M., Kolgotin, A., Haarig, M., Burton, S. P., Ferrare, R. A., Hostetler, C. A., and Müller, D.: 3+2 + X: what is the most 645 useful depolarization input for retrieving microphysical properties of non-spherical particles from lidar measurements using the spheroid model of Dubovik et al. (2006)?, Atmos. Meas. Tech., 12, 4421–4437, https://doi.org/10.5194/amt-12-4421-2019, 2019.

Vaughan, M. A., Powell, K. A., Winker, D. M., Hostetler, C. A., Kuehn, R. E., Hunt, W. H., Getzewich, B. J., Young, S. A., Liu, Z., and McGill, M. J.: Fully automated detection of cloud and aerosol layers in the CALIPSO lidar measurements, J. Atmos. Ocean. Tech., 26, 2034–2050, https://doi.org/10.1175/2009JTECHA1228.1, 2009.

Wallace, K., Hélière, A., Lefebvre, A., Eisinger, M., and Wehr, T.: Status of ESA's EarthCARE mission, passive instruments payload, in: Earth Observing Systems XXI, edited by: Butler, J. J., Xiong, X. J., and Gu, X., vol. 9972, p. 997214, International Society for Optics and Photonics, SPIE, https://doi.org/10.1117/12.2236498, 2016.

Wandinger, U., Haarig, M., Baars, H., Donovan, D., and van Zadelhoff, G.-J.: Cloud top heights and aerosol layer properties from EarthCARE lidar observations: the A-CTH and A-ALD products, Atmos. Meas. Tech., 16, 4031-4052, https://doi.org/10.5194/amt-16-4031-2023, 2023.

Wehr, T., Kubota, T., Tzeremes, G., Wallace, K., Nakatsuka, H., Ohno, Y., Koopman, R., Rusli, S., Kikuchi, M., Eisinger, M., Tanaka, T., Taga, M., Deghaye, P., Tomita, E., and Bernaerts, D.: The EarthCARE mission – science and system overview, Atmos. Meas. Tech., 16, 3581–3608, https://doi.org/10.5194/amt-16-3581-2023, 2023.

Wiegner, M., Groß, S., Freudenthaler, V., Schnell, F., and Gasteiger, J.: The May/June 2008 Saharan dust event over Munich: Intensive aerosol parameters from lidar measurements, J. Geophys. Res.-Atmos., 116, D23213, https://doi.org/10.1029/2011JD016619, 2011.

Winker, D. M., Pelon, J., Coakley, J. A. Jr., Ackerman, S. A., Charlson, R. J., Colarco, P. R., Wielicki, B. A.: The CALIPSO mission: A global 3D view of aerosols and clouds. Bulletin of the American Meteorological Society, 91, 1211–1230, https://doi.org/10.1175/2010BAMS3009.1, 2010.

Yamauchi, A., Suzuki, K., Oikawa, E., Sekiguchi, M., Nagao, T. M., and Ishida, H.: Description and validation of the Japanese algorithm for radiative flux and heating rate products with all four EarthCARE instruments: pre-launch test with A-Train, Atmos. Meas. Tech., 17, 6751–6767, https://doi.org/10.5194/amt-17-6751-2024, 2024.

Yoshida, R., Okamoto, H., Hagihara, Y., Ishimoto, H.: Global analysis of cloud phase and ice crystal orientation from CALIPSO data using attenuated backscattering and depolarization ratio, J. Geophys. Res., 115, D00H32, doi:10.1029/2010JD014032, 2010.

Young, S., A., Vaughan, M. A., Kuehn, R. E., Winker D. M.: The retrieval of profiles of particulate extinction from Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) Data: Uncertainty and Error Sensitivity Analyses, J. Atmos. Ocean. Tech., 30, 395-428, DOI: 10.1175/JTECH-D-12-00046.1, 2013.

van Zadelhoff, G.-J., Donovan, D. P., and Wang, P.: Detection of aerosol and cloud features for the EarthCARE atmospheric lidar (ATLID): the ATLID Feature Mask (A-FM) product, Atmos. Meas. Tech., 16, 3631–3651, https://doi.org/10.5194/amt-16-3631- 2023, 2023.

Zhang, L., Li, J., Jiang, Z., Dong, Y., Ying, T., and Zhang, Z.: Clear-sky direct aerosol radiative forcing uncertainty associated with aerosol vertical distribution based on CMIP6 models, Journal of Climate, 35, 3021-3035, 10.1175/JCLI-D-21-0480.1, 2022.