



# Looking up, looking down: zenith-to-nadir lidar comparisons for cloud scenes (IPRAL and ATLID/EarthCARE)

Artem G. Feofilov<sup>1</sup>, Arezki Haddouche<sup>1</sup>, Hélène Chepfer<sup>1</sup>, Sergey Khaykin<sup>2</sup>, Florian Lapouge<sup>3</sup>, Patricia Delville<sup>4</sup>, Marc-Antoine Drouin<sup>1</sup>, Martial Haeffelin<sup>4</sup>, Jean-Charles Dupont<sup>5</sup>, Cheikh Dione<sup>5</sup>, and Christophe Pietras<sup>6</sup>

<sup>1</sup> Laboratoire de Météorologie Dynamique/Institut Pierre Simon Laplace (IPSL), Sorbonne Université, CNRS, Institut Polytechnique de Paris, Paris, France

<sup>2</sup> LATMOS/IPSL, UVSQ Université Paris-Saclay, Sorbonne Université, CNRS, Guyancourt, France

<sup>3</sup> École Polytechnique de Paris, Institut Polytechnique de Paris, Paris, France

<sup>4</sup> IPSL, CNRS, École Polytechnique de Paris, Institut Polytechnique de Paris, Paris, France

<sup>5</sup> IPSL, École Polytechnique, UVSQ, Université Paris-Saclay, Palaiseau, France

<sup>6</sup> UVSQ, IPSL, LMD, Palaiseau, France

Correspondence to: Artem G. Feofilov ([afefilo@lmd.ipsl.fr](mailto:afefilo@lmd.ipsl.fr))

## Abstract

Intercomparisons between spaceborne lidar instruments and ground-based lidars are essential for calibration/validation of these instruments and for data quality assurance. However, zenith-to-nadir viewing geometry differences can create systematic discrepancies between space-borne and ground-based observations, even under perfect spatiotemporal collocation. This is particularly true for clouds: unlike aerosols, which are comparatively well-mixed and spatially homogeneous, clouds are structured and short-lived, so both the observation geometry and the collocation quality strongly affect whether a spaceborne and a ground-based profile can be meaningfully compared. Still, clouds are valuable targets because they push the signal into a different dynamic range, from strong backscatter to near-saturation, that clear-sky or aerosol scenes never exercise. The present paper addresses three questions: which parameters retrieved from lidar observations are better suited to zenith-to-nadir comparisons, which atmospheric scenes are favorable for such a comparison, and how the choice of collocation approach affects the comparison's reliability.

We address these questions with a simulation framework that generates 15 synthetic cloud scenes, propagates them through the lidar simulator of the Cloud Feedback Model Intercomparison Project (CFMIP) Observation Simulator Package, version 2 (COSP2), for both nadir and zenith geometries, and adds instrument-specific daytime/nighttime noise. We compare attenuated backscatter (Level 1) and retrieved particulate backscatter  $\beta_{part}(z)$  (Level 2) using two scattering-ratio (SR) definitions, and introduce a Comparability Score (CS) quantifying profile similarity via correlation and shape. Thin clouds yield high CS and comparable zenith/nadir measurements; thicker and multi-layer clouds do not, due to directional attenuation and retrieval instabilities. We also found that  $\beta_{part}(z)$  comparisons remain valid over a broader range of conditions than  $SR(z)$ .

In real collocations, spatiotemporal separation is rarely negligible, so we select the best-matched profile pair by maximizing cross-correlation. We estimate this method's false discovery rate (the fraction of selected pairs that do not sample the same cloud) at 10–20%, using both a theoretical null model and an empirical one built from mismatched-date collocations. Applied to real collocations at 355 nm between the Atmospheric Lidar (ATLID) on the Earth Cloud, Aerosol and Radiation Explorer (EarthCARE) and the IPSL Atmospheric Research Lidar (IPRAL) across four representative cases,  $\beta_{part}(z)$  again agrees better than native-geometry  $SR(z)$ , consistent with the simulation results.

We conclude that a spaceborne-to-ground-based cloud lidar comparison is meaningful, that is, not dominated by viewing geometry or by mismatched scenes, when three conditions are met: (1) candidate collocations are screened using the correlation coefficient and cloud vertical structure, rather than accepted by spatiotemporal proximity alone; (2) the collocated scene has a Comparability Score above 50; (3) the comparison variable,  $SR(z)$  or  $\beta_{part}(z)$ , is chosen according to the scene's optical depth: both variables are comparable below  $\tau = 0.7$  and only  $\beta_{part}(z)$  is comparable above this threshold. We propose this decision scheme, with its threshold values, as a starting point for other spaceborne-to-ground-based cloud lidar validations.



## 1. Introduction

In the field of remote sensing, active sounding by lidar (Light Detection And Ranging) instruments has been playing an important role since the 1960s (Collis and Russell, 1976; Measures, 1984; Weitkamp, 2005), due to its sensitivity and exceptional vertical resolution. Spaceborne lidar was first demonstrated in 1994 by the Lidar In-space Technology Experiment (LITE) (Winker, 1997), and matured over the following decades through the Geoscience Laser Altimeter System (GLAS) on the Ice, Cloud, and land Elevation Satellite (ICESat) (Abshire et al., 2005), the Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) on the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) (Winker et al., 2010), the Cloud-Aerosol Transport System (CATS) on the International Space Station (ISS) (Yorks et al., 2016), the Advanced Topographic Laser Altimeter System (ATLAS) on ICESat-2 (Markus et al., 2017), the Atmospheric Laser Doppler Instrument (ALADIN) on Aeolus (Stoffelen et al., 2005), the Aerosol and Carbon Detection Lidar (ACDL) on Daqi-1 (Zhou et al., 2022), and, most recently, the Atmospheric Lidar (ATLID) on the Earth Cloud, Aerosol and Radiation Explorer (EarthCARE) (Illingworth et al., 2015).

One of the challenges of a spaceborne lidar atmospheric mission is the ongoing need to calibrate and validate the lidar measurements and the retrieved geophysical properties. The principle of an atmospheric lidar measurement is sending a short package of sounding photons into the atmosphere, starting the internal clock, and collecting the backscattered light versus acquisition time. The registered signal contains information about the atmospheric scattering (molecules and particles), noise (parasitic solar photons, detector noise), and biases (background signal). It can also be saturated by a highly reflective target or completely zeroed due to a highly attenuating medium, and the conversion from the detector's millivolts or photocounts to physical units, like attenuated backscatter in  $[m^{-1}sr^{-1}]$ , is not trivial. Once this is done, it needs to be validated through a comparison with other calibrated and validated instruments, and a common method for such validation is nearly simultaneous common-volume measurements with ground-based lidars. This validation has two potential issues: (1) the spatiotemporal collocation is rarely perfect, and the collocation frequency is inversely proportional to its quality, so there is a common tradeoff between the number of collocations and the distance and temporal offset; (2) even in the case of a perfect collocation, the spaceborne lidar and the ground-based one do not observe the atmospheric column in the same way: even if they do it at the same wavelength with a similar instrument and a similar efficiency, they do not observe in the same direction, and the signal attenuation in zenith and nadir directions create differences in the compared values and/or make them completely incomparable. These calibration challenges have important implications for both scientific research and operational applications.

The ground-based lidars have been used as a source of "ground truth" for spaceborne validation. In the framework of the ATLID/EarthCARE calibration/validation campaign (<https://earth.esa.int/eogateway/missions/earthcare/data/calibration-validation> and references therein), dozens of international teams operate lidar systems at La Reunion and Cyprus for stratospheric aerosol validation (Khaykin et al., 2026), as well as the AD-Net (Asian Dust Network) stations, ACTRIS network sites at Mindelo, Julich, Ny Alesund, Lindenberg, and Munich, the RMR (Rayleigh-Mie-Raman) lidar at Andøya (Institut für Atmosphärenphysik, IAP), EMORAL, and IPRAL (Institut Pierre-Simon Laplace, France) as well as AERONET sun-photometer network (37 sites in South and Central America and Antarctica). These multi-platform comparisons enabled both case study validation and statistical assessments of ATLID performance across different atmospheric conditions and regions. For aerosols, the validation strategy relies on comparing spaceborne overpasses against a fixed spatiotemporal window of ground-based measurement, accumulated over one or more years into a statistical dataset (Mona et al., 2009; Pappalardo et al., 2010) or a more limited set of coincident cases (Mamouri et al., 2009), and, for ATLID, in stratospheric aerosol validation studies (Khaykin et al., 2026; Poncet et al., 2026). This works well because aerosol fields are spatially more homogeneous than clouds at the typical collocation distance ( $\sim 100$  km). However, even with aerosols the representativeness is not guaranteed: Gimmestad et al. (2017) reported that aerosol inhomogeneity affected intercomparison of ground-based lidar with CALIOP. To reduce this uncertainty, several studies have applied a trajectory analysis based on wind information to identify the true air parcel: Hybrid Single-Particle Lagrangian Integrated Trajectory (HYSPLIT) backward/forward trajectories have been used to match CALIOP overpasses to Aerosol Robotic Network (AERONET) sun photometers (Lopes et al., 2013) and to ground-based Raman lidars (Tesche et al., 2013), and a similar air-mass-matching philosophy is reflected in the recent validation protocol of the Committee on Earth Observation Satellites (CEOS) (Amiridis et al., 2025). Even this kind of wind-based tracking is harder to apply to clouds: contrary to aerosols, clouds are structured and short-lived, so proximity alone does not guarantee that the two instruments sample the same atmospheric structure. Both the observation geometry and the choice of collocation method strongly affect whether a spaceborne and a ground-based profile can be meaningfully compared, and this comparability is central to any calibration/validation effort.

Despite its importance, a systematic quantitative evaluation of how atmospheric conditions, noise characteristics, and collocation method affect zenith-to-nadir cloud comparisons remains lacking. Quantifying these dependencies is critical both for existing and upcoming missions and for optimizing ground-based network operations to support spaceborne lidar calibra-



tion. Knowing the conditions and variables suitable for a reliable calibration would be beneficial both from a scientific point of view, because it will yield a better quality lidar product, and from a practical point of view, because it will guide the design of validation protocols, optimize ground-based network operations, and provide practical criteria for assessing collocation quality in real-world scenarios.

The present paper addresses three questions: which parameters retrieved from lidar observations are better suited to zenith-to-nadir comparisons, which atmospheric scenes are favorable for such a comparison, and how the choice of collocation approach, spatiotemporal proximity versus correlation-based profile matching, affects the comparison's reliability. To address these challenges, we develop a simulation framework that systematically evaluates the influence of atmospheric conditions, noise, and collocation quality on zenith-to-nadir lidar comparisons. In Section 2, we introduce this framework and the formalism necessary to understand its elements. In Section 3, we present a theoretical study of Level 1 and Level 2 comparisons across all cloud scenes, and characterize the conditions under which zenith and nadir geometries remain comparable. In Section 4, we apply these findings to real ATLID/EarthCARE-IPRAL collocations, introducing a correlation-maximization method for profile matching and quantifying its false-discovery rate. Section 5 proposes practical recommendations for other spaceborne-to-ground-based validations, and we conclude in Section 6.

## 2. Method

To systematically evaluate zenith-to-nadir lidar comparisons across varying atmospheric cloud containing conditions, noise regimes, and collocation qualities, we developed a simulation framework based on the algorithm of Feofilov et al. (2023), enhanced for the purposes of the current study. The approach consists of three primary computational stages: (1) generation of synthetic cloud scenes with realistic atmospheric profiles, (2) noise-free lidar signal calculations using the Cloud Feedback Model Intercomparison Project Observation Simulator (COSP2), and (3) noise injection and signal degradation modeling for different observation geometries and wavelengths. The simulation workflow is illustrated in Fig. 1, which outlines the complete computational pipeline from atmospheric model data through noise-degraded lidar profiles to final comparability assessments. The pipeline consists of five stages and two processing routines described below.

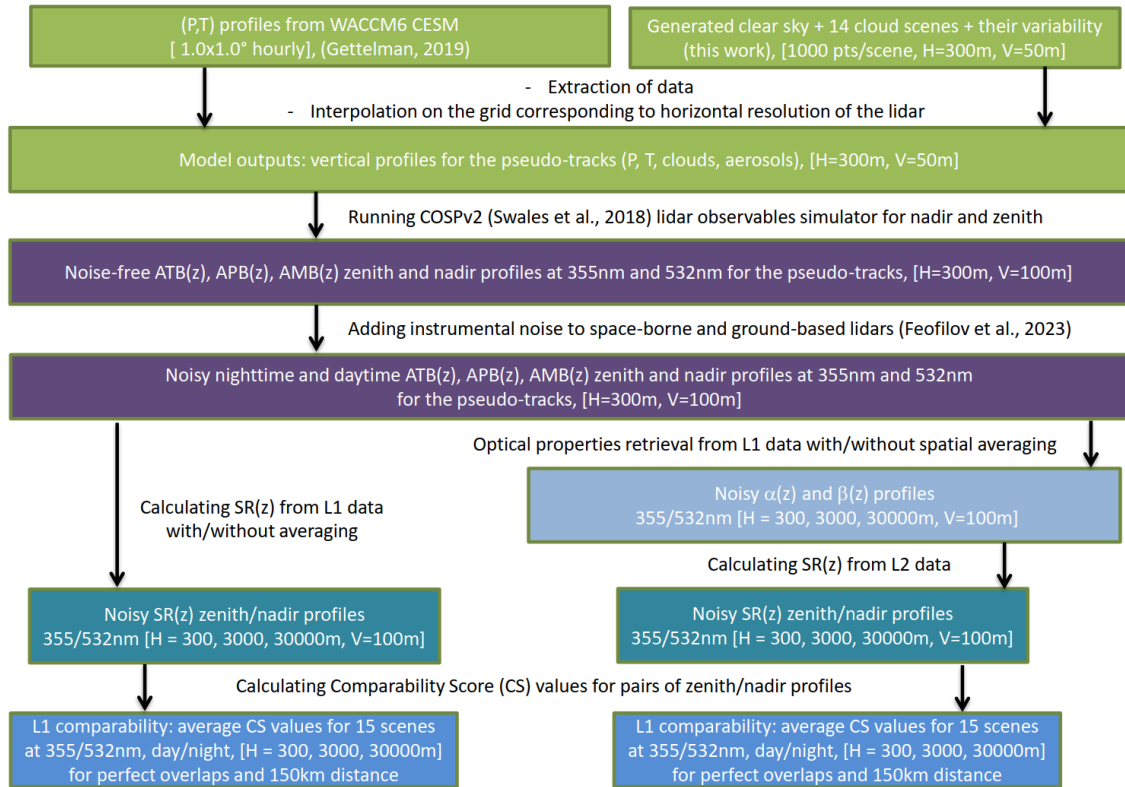


**Figure 1: Flowchart of the numerical experiment quantifying the comparability of ground-based and space-borne lidar observations across atmospheric scenes, noise regimes, and observation geometries. Green boxes list the input data. Violet boxes represent the noise-free and noisy L1 data. Light blue box stands for L2 optical properties data retrieved from L1. Water-green boxes show the SR(z) data. Dark blue boxes represent the Comparability Score data estimated from L1 and L2 profiles. Black text between boxes describes actions performed on each dataset. White text in square brackets in the boxes indicates horizontal (H) and vertical (V) resolutions of the datasets.**

## 2.1. Atmospheric profile and cloud scene generation

The workflow begins with generation of synthetic atmospheric profiles and cloud scenes. Since the atmospheric molecular scattering is significantly smaller than the particulate scattering, we opted to fix a pressure/temperature vertical profile and vary the properties and the positions of the clouds. For the atmospheric P et T profiles, we took the outputs of WACCM6 model (Gettelman et al., 2019) at [1°x1°] for the tropical conditions at noon, consistent with our previous simulations (Feofilov et al., 2023). The vertical resolution of this model is better than 1km up to 22km height. For the horizontal grid, we opted to use native resolution of CALIOP and ATLID lidars, which is about 300m, and we used 50m vertical grid. As in Feofilov et al. (2023), we created pseudo-orbits suitable for the processing by COSPv2.0 algorithm (Swales et al., 2018), which includes a lidar simulator (Chepfer et al., 2008; Reverdy et al., 2015; Guzman et al., 2017; Cesana et al., 2019) that mimics the measurements that would be obtained by a spaceborne lidar if it were overflying the atmosphere simulated by a climate model, or by a ground-based lidar if it were placed under the same atmosphere. To cover the variety of atmospheric scenes observed by cloud lidars, we chose 15 scenarios described in Table 1. We start from the most transparent case of a clear sky atmosphere at the top of the table and we gradually move from very thin cirrus clouds (Ci) through semi-transparent clouds to thick clouds. In addition, we consider multi-level cases when the upper cloud is one of single layer clouds listed in the Table 1 and the lower one is almost opaque that is the worst-case scenario for the zenith-to-nadir lidar comparisons.

Table 1: Cloud scene configuration parameters. COT stands for Cloud Optical Thickness and p-t-p is point-to-point.



| N  | Scene name                             | Altitude range(s) (km)               | COT         | Particle type    | $R_{\text{eff}}$ ( $\mu\text{m}$ ) | Lidar ratio  | P-t-p variab (m) | Corr. length (m) |
|----|----------------------------------------|--------------------------------------|-------------|------------------|------------------------------------|--------------|------------------|------------------|
| 1  | Clear Sky                              | 0–30                                 | 0           | N/A              | N/A                                | N/A          | N/A              | N/A              |
| 2  | Very thin subvisible <i>Ci</i>         | 13.0–14.0                            | 0.05        | Ice              | 20                                 | 18.3         | 80               | 200              |
| 3  | Subvisible <i>Ci</i>                   | 11.5–12.5                            | 0.15        | Ice              | 24                                 | 19.2         | 90               | 300              |
| 4  | Thin <i>Ci</i>                         | 10.0–11.0                            | 0.3         | Ice              | 30                                 | 20.8         | 100              | 400              |
| 5  | Semi-transp. <i>Ci</i>                 | 12.0–13.0                            | 1.0         | Ice              | 37                                 | 23.3         | 200              | 600              |
| 6  | Opaque <i>Ci</i>                       | 15.0–16.0                            | 3.0         | Ice              | 45                                 | 26.3         | 350              | 800              |
| 7  | <i>Ci</i> above thick low              | 12.0–13.0 (upper)<br>2.0–2.5 (lower) | 1.0<br>10.0 | Ice<br>Liquid    | 37<br>37                           | 23.3<br>23.3 | 200<br>100       | 600<br>600       |
| 8  | Thin <i>Ac</i>                         | 4.5–5.0                              | 1.0         | Liquid           | 12                                 | 19.3         | 100              | 150              |
| 9  | Semi-transp. <i>As</i>                 | 5.5–7                                | 3.0         | Liquid           | 14                                 | 19.2         | 100              | 500              |
| 10 | Opaque <i>As</i>                       | 3.5–6.5                              | 10.0        | Liquid           | 16                                 | 18.9         | 150              | 600              |
| 11 | Semi-transp. <i>As</i> above thick low | 5.5–7.0 (upper)<br>2–2.5 (lower)     | 3.0<br>10.0 | Liquid<br>Liquid | 14<br>14                           | 19.2<br>19.2 | 100<br>100       | 500<br>500       |
| 12 | Thin <i>Sc</i>                         | 1.5–2.5                              | 3.0         | Liquid           | 9                                  | 19.3         | 100              | 150              |
| 13 | Semi-transp. <i>Sc</i>                 | 1.5–2.5                              | 5.0         | Liquid           | 12                                 | 19.3         | 120              | 200              |
| 14 | Opaque <i>Sc</i>                       | 1.8–2.5                              | 20.0        | Liquid           | 14                                 | 19.2         | 200              | 250              |
| 15 | Semi-transp. <i>Sc</i> above thick low | 1.5–2 (upper)<br>0.5–1 (lower)       | 5.0<br>10.0 | Liquid<br>Liquid | 12<br>12                           | 19.3<br>19.3 | 120<br>100       | 200<br>200       |

For each of the 14 cloud scenarios plus clear-sky conditions, we simulated 1000 vertical profiles with 300 m horizontal spacing and 50 m vertical resolution extending from the surface to 30 km altitude. To imitate the cloud height variability, needed for the estimates of the effects of non-perfect spatiotemporal overlap, we added a vertical variability modulated by a r.m.s. value given in the “point-to-point variability” column of Table 1. To ensure the continuity of cloud parameters and avoid uncorrelated random noise, we constrained the fluctuations using a “correlation length” parameter  $L_c$ , specific to each cloud type. The correlated random field was generated using a first-order autoregressive process described by:

$$Z(x_i) = Z(x_{i-1}) + \Delta Z_i; \quad \Delta Z_i = e^{-(x_i - x_{i-1})/L_c} \times \Delta Z_{i-1} + \delta_i, \quad i = 1, \dots, N_x; \quad (1)$$

where  $x$  denotes a horizontal position,  $Z(x_i)$  is a cloud height at the  $i^{\text{th}}$  point,  $\Delta Z_i$  is a point-to-point cloud height change at the  $i^{\text{th}}$  step, and  $\delta_i$  is a Gaussian random term with zero mean and standard deviation equal to the point-to-point variability value. The correlation length  $L_c$  represents the distance over which height fluctuations are scaled by a factor of  $e^{-1}$ . This is a simplistic model, which we built based on previous works like Atlas and Bretherton (2023) who studied the aircraft observations of gravity wave (GW) activity and turbulence in the tropical tropopause layer (TTL), Li et al. (2015) who used a combination of CloudSat and Calipso instruments to study the stratus cloud properties, Comstock et al. (2002) who used a combination of ground-based passive and active instruments to study the properties of tropical cirrus clouds, including their temporal variability, Hogan and Illingworth (2003) who used radar data to parameterize inhomogeneity of ice clouds, Wu and Liang (2005) who estimated cloud horizontal inhomogeneity from a cloud-resolving model simulation, Cesana et al. (2019) who created the CASCAD dataset for cumulus and stratocumulus clouds, based on CALIOP and CloudSat measurements, and Covert et al. (2022) who used a constrained model to simulate the vertical dependence of the horizontal variation of cloud water in stratocumulus clouds, and the outputs of an updated version of 3DCLOUD model (Szcza et al., 2014), 3DCLOUD\_V2 (Alkasem et al., 2017) used in Feofilov et al. (2023). These parameters are not related to observation geometry; we introduced them to get a rough estimate of what inconsistency one should expect for real-life overlaps described below.

The atmospheric scenes consisting of pressure (P), temperature (T), and variable-height cloud profiles are organized into a single pseudo-orbit that constitutes an input to COSP2 simulator. This process involves a recalculation from a bulk cloud optical thickness (COT or  $\tau$ ) into COSP2-specific parameters. In these calculations, we assumed a constant vertical profile of cloud properties within a given cloud layer at each horizontal step  $i$  (Eq. (1)); this is true for the majority of ice clouds (Feofilov et al., 2015). The remaining ice clouds, with “upper triangle” or “lower triangle” profiles (Feofilov et al., 2015), are rare and are not considered in the present manuscript for the sake of simplicity. As for the liquid clouds, a constant profile is a coarser approximation (Pawlowska et al., 2000; Wood, 2005); we opted to keep it to isolate the effect of nadir/zenith attenuation differences from cloud-internal microphysical variability. For each 50m sub-layer, we performed a series of calculations, inverse to those used in COSP2 package to obtain the same optical depths as in its *lidar\_optics()* routine:



$$MR = \frac{\Delta\tau \times \rho_p \times R_{eff}}{\eta \times \frac{3}{4} \times Q_{scat} \times \rho_{air} \times \Delta h}; \quad \rho_{air} = n_{air} \times m_p \times M_{air} \quad (2)$$

where  $MR$  is a unitless mass-mixing ratio,  $\Delta\tau$  is an optical depth per sublayer  $\Delta h$ ,  $\rho_p$  is a particle density (depends on particle type),  $R_{eff}$  is an effective particle's radius,  $\eta$  is a multiple scattering coefficient (0.75),  $Q_{scat}$  is scattering efficiency (2.0 both for ice and liquid),  $\rho_{air}$  is air density in the sublayer calculated from air molecule concentration  $n_{air}$ , proton mass  $m_p$ , and mean molecular mass of air  $M_{air}$  ( $\sim 28.8$  atomic mass units). Both in the COSP2 lidar simulator and in the calculations below, the backscatter and extinction coefficients at the corresponding wavelengths are calculated (Collis and Russell, 1976; Bucholtz, 1995):

$$\beta_{mol}(\lambda, z) = (d\sigma/d\Omega)_\lambda \times N(z); \quad \alpha_{mol}(\lambda, z) = \frac{4\pi}{1.5} \beta_{mol}(\lambda, z) \quad (3)$$

$$(d\sigma/d\Omega)_\lambda = \frac{\sigma(\lambda, z)}{4\pi} \times \frac{3}{4} (1 + \cos^2(\pi)) \quad (4)$$

$$\sigma(\lambda, z) = \frac{24\pi^3 (n_s^2(\lambda) - 1)^2 (6 + 3\rho(\lambda))}{\lambda^4 N_s^2 (n_s^2(\lambda) + 2)^2 (6 - 7\rho(\lambda))} \quad (5)$$

where  $(d\sigma/d\Omega)_\lambda$  is a differential cross section ( $\text{m}^2 \text{sr}^{-1}$ ),  $N(z)$  is a number density ( $\text{m}^{-3}$ ),  $\sigma(\lambda, z)$  is Rayleigh cross section ( $\text{m}^2$ ),  $n_s(\lambda)$  is the refractive index for standard air,  $\rho(\lambda)$  is the depolarization factor, and  $N_s$  is the number density of standard air ( $2.54743 \times 10^{25} \text{ m}^{-3}$ ). The resulting values of  $(d\sigma/d\Omega)_\lambda$  at 355 nm and 532nm are  $3.2897988 \times 10^{-31} \text{ m}^2 \text{sr}^{-1}$  and  $6.1668318 \times 10^{-32} \text{ m}^2 \text{sr}^{-1}$ , respectively.

In Fig. 2, we show the inputs generated with the help of the formalism outlined above. All the profiles are organized to one “pseudo-orbit” which is then transferred to the COSP2 package.

## 2.2. Noise-free lidar profiles simulated with COSP2

The interpolated atmospheric and cloud profiles shown in Fig. 2 were fed to the COSP2 lidar observables simulator (Swales et al., 2018), which computed noiseless attenuated backscatter (ATB), attenuated particulate backscatter (APB), and attenuated molecular backscatter (AMB) profiles at 355 nm and 532 nm for both nadir (space-borne) and zenith (ground-based) observation geometries using the CALIOP and ATLID simulators (Chepfer et al., 2008; Reverdy et al., 2015) included in this simulator. In zenith geometry, the simulator assumes that the ground based lidar operating at the corresponding wavelength (355 nm for ATLID and 532nm for CALIOP) observes the atmosphere overflying it, so the profile index can be considered as time index in this case. Output profiles maintained the original 300 m horizontal spacing and were saved at 100 m vertical resolution. It is important to note here that the simulators do not use any experimental setup parameters like telescope size, optical path throughput, or rejection filter bandwidth because they do not account for instrumental noise effects. To address the noise-related effects, their outputs were processed by a third part of the simulation chain (Fig. 1), which estimates noise and its propagation in the lidar system. For the sake of simplicity, we do not show noise-free lidar curtains at this step and pass directly to noise-perturbed profiles in the next section.

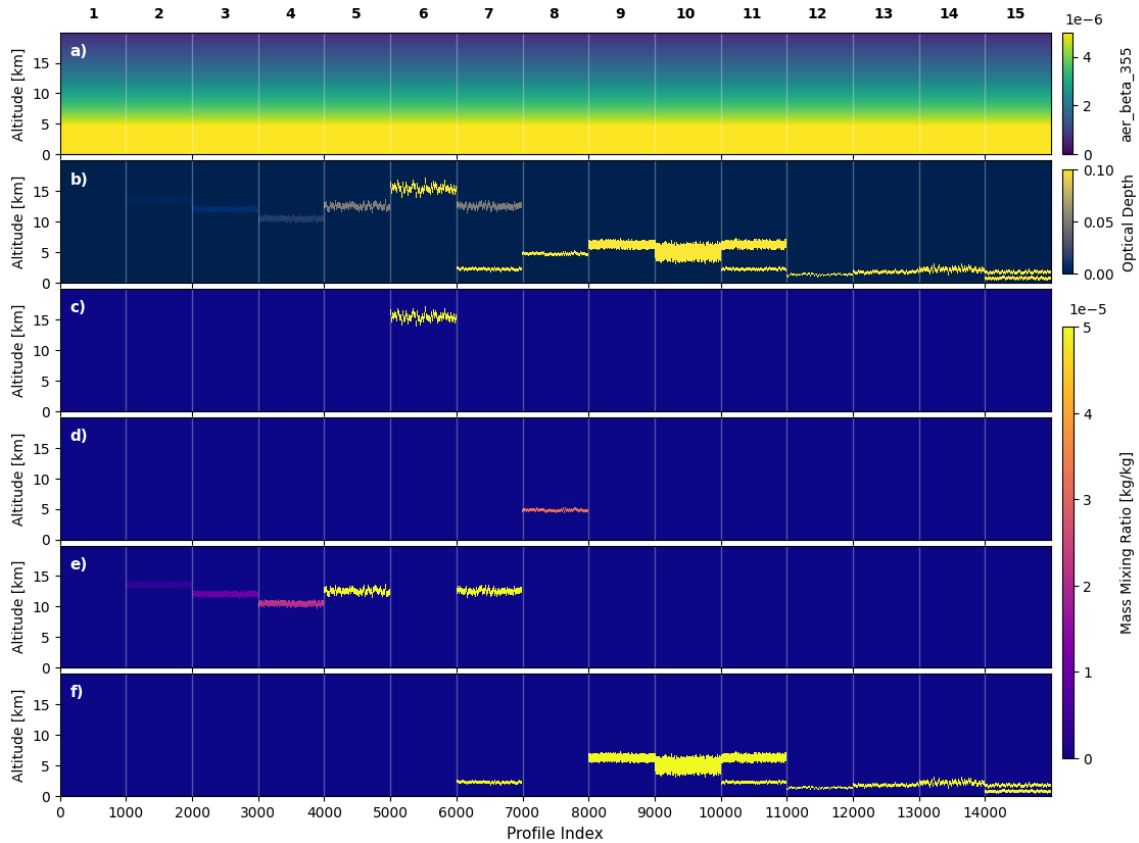


**Figure 2: Input data used in COSP2 calculations:** (a) molecular backscatter profile at 355 nm, aer\_beta\_355 in terms of COSP2 (aer\_beta\_532 is not shown for the sake of clarity); (b) optical depth per 50m layer both for convective clouds (dtau\_c) and large scale clouds (dtau\_s); (c, d, e, f): mass mixing ratios for convective core ice and liquid clouds and for large scale ice and liquid clouds (mr\_ccice, mr\_ccliq, mr\_lsice, and mr\_lsliq, respectively). The cloud scene numbers on top of the plot correspond to the first column of Table 1.

### 2.3. Adding instrumental noise

The detailed formalism for signal and noise calculations can be found in Feofilov et al. (2023). The parameters of CALIOP and ATLID lidars related to these calculations are listed in Table 1 of the same work. To exclude the instrument-based effects and address only the viewing geometry effects, we used the same characteristics for the pairs of ground-based and space-borne lidars, with only one exclusion: the field of view (FOV) of the lidar was changed from 65  $\mu$ rad for ATLID and 130  $\mu$ rad for CALIOP to 500  $\mu$ rad representing a typical value for ground-based atmospheric lidars (Whiteman, 2003; Weitkamp, 2005). We did not change the beam divergence, because wider FOV covers both narrow and wide beams that is necessary and sufficient for our noise estimates, whereas the rationale of changing FOV in these calculations was to correctly estimate the parasite solar photons. Below, we list only the main stages of instrumental noise estimation.

(a) *Photon emission calculation*: determining the number of emitted photons per laser pulse using instrumental parameters (pulse energy, wavelength, and optical throughput); (b) *Photon propagation*: calculating the number of photons reaching each detector (molecular and particulate channels) for each range gate, accounting for atmospheric backscatter, two-way transmission, and solid angle effects; (c) *HSRL (High Spectral Resolution Lidar) cross-talk correction* (for ATLID and its ground-based counterpart): apply the four cross-talk coefficients to account for non-ideal channel separation between molecular and particulate detection paths; (d) *Solar background estimation*: computing solar photon contributions to both detectors during daytime observations using radiative transfer calculations that account for surface albedo, viewing geometry, and atmospheric composition; (e) *Noise source inclusion*: adding dark current and readout noise components to the signal;



(f) *Variance propagation*: calculating signal and noise variances for each channel using error propagation formulas, including cross-channel covariance terms that arise from the HSRL cross-talk coupling.

Separate noise realizations were generated for daytime (Solar zenith angle of 45°) and nighttime conditions. The resulting noise-degraded profiles represent what actual lidar systems would observe under the specified atmospheric and instrumental conditions. Output files contain the noisy attenuated backscatter profiles (total, molecular, and clear sky, see the definitions in the next section) at 300 m horizontal spacing and 100 m vertical resolution for both observation geometries and both wavelengths (see the bottom part of Fig. 1).

## 2.4. Scattering ratio definitions

For a uniform comparison of zenith and nadir profiles, we need a backscatter-based quantity that would be independent of absolute signal strength and that would work both for L1 data (ATB/APB/AMB) and for L2 data (optical properties retrieved from L1). The scattering ratio (SR) serves this purpose, but one has to keep in mind that two definitions exist in the literature under this name. Both use total attenuated backscatter as the numerator, but they differ in the denominator: one employs a theoretical clear-sky molecular backscatter, the other a measured one. We adopt the former. It is applicable to both HSRL and non-HSRL systems, its denominator introduces no additional noise, and its physical meaning is unambiguous. The latter, while attractive for its matched attenuation path, is HSRL-only and it attenuates with the main signal in thick clouds, which increases the noise whereas the former gives a smooth attenuation down to zero that is convenient for the analysis of fully attenuated cases (e.g. Guzman et al., 2017).

Let us start with a lidar equation for total attenuated backscatter:

$$ATB(\lambda, z) = (\beta_{mol}(\lambda, z) + \beta_{part}(\lambda, z)) \times e^{-2 \int_{z_{lidar}}^z (\alpha_{mol}(\lambda, z') + \eta \alpha_{part}(\lambda, z')) dz'} \quad (6)$$

where  $\lambda$  is the wavelength, at which the lidar operates,  $ATB(\lambda, z)$  is a wavelength-dependent total attenuated backscatter for the layer at altitude  $z$ ,  $\beta_{mol}(\lambda, z)$  and  $\alpha_{mol}(\lambda, z)$  can be estimated from Eq. (3),  $\beta_{part}(\lambda, z)$  and  $\alpha_{part}(\lambda, z)$  are the particulate backscatter and extinction coefficients, which are the optical properties we seek to retrieve,  $z_{lidar}$  is an altitude of the satellite in the case of a space-borne lidar or a zero-altitude level in the case of a ground-based lidar, and  $\eta$  is a multiple scattering coefficient (e.g., Platt, 1973; Garnier et al., 2015; Donovan, 2016), the same as in Eq. (2). Similarly, the equation for the attenuated molecular backscatter that can be detected by a Rayleigh channel of an HSRL lidar is:

$$AMB(\lambda, z) = \beta_{mol}(\lambda, z) \times e^{-2 \int_{z_{lidar}}^z (\alpha_{mol}(\lambda, z') + \eta \alpha_{part}(\lambda, z')) dz'} \quad (7)$$

and the attenuated molecular backscatter for clear sky conditions is:

$$ATB_{mol}(\lambda, z) = \beta_{mol}(\lambda, z) \times e^{-2 \int_{z_{lidar}}^z \alpha_{mol}(\lambda, z') dz'} \quad (8)$$

Correspondingly, the first definition of scattering ratio is:

$$SR1(\lambda, z) = \frac{ATB(\lambda, z)}{ATB_{mol}(\lambda, z)} \quad (9)$$

and the second one is:

$$SR2(\lambda, z) = \frac{ATB(\lambda, z)}{AMB(\lambda, z)} \quad (10)$$

The interpretation of SR1 defined in accordance with Eq. (9) is straightforward. Since the  $ATB(\lambda, z)$  in Eq. (6) turns into  $ATB_{mol}(\lambda, z)$  in the absence of particles (Eq. (8)), the SR1 values for clear-sky conditions will be about 1. If the SR is larger than one, this indicates the presence of an aerosol or cloud layer. If SR1 is less than one, this indicates the attenuation of the backscattered radiance by a cloud or aerosol layer. Finally, if the SR1 is close to zero, this indicates a total attenuation of backscattered radiance (Guzman et al., 2017). For SR2 equal to 1, the interpretation is the same as for SR1. For the values larger than one, it also indicates the presence of aerosols, but the “attenuation” of SR2 within the cloud layer is not noticeable.



ble because both the numerator and denominator in Eq. (10) depend on attenuation by particles. However, in the case of a thick cloud the behavior of SR2 becomes unpredictable because both signals become strongly attenuated and one obtains the indeterminate form of zero over zero, which will be strongly affected by noise. In the past works (Chepfer et al., 2008, 2010, 2013; Feofilov et al., 2022, 2023, 2024), we advocated using SR1 rather than SR2 for spaceborne lidars. Leaping ahead, we can say that this metrics is not always sufficient for the comparison of the observations made in opposite viewing geometries, and other options should be considered.

## 2.5. Optical property retrieval

To simulate the full L1→L2 retrieval chain and compare the nadir and zenith viewing geometries in the optical products domain, and retrieve the  $\beta_{part}(z)$  profiles used for the Level-2 comparisons in Sects. 3.2 and 4.4, we realized Klett's retrieval algorithm (Klett, 1981, 1985) in its reformulation by Ansmann and Müller (2005). More advanced multi-channel and iterative inversion schemes exist for L1→L2 retrieval (e.g. Donovan et al., 2024); we use Klett's single-channel inversion here for its closed-form solution and physical transparency, which keeps the propagation of viewing-geometry effects into the retrieved L2 quantities traceable.

$$\beta_{part}(R) = \frac{A}{B-C} - \beta_{mol}(R) \quad (11)$$

$$A = S(R) \times e^{-2 \int_{R_0}^R [LR_{part}(r) - LR_{mol}] \times \beta_{mol}(r) dr} \quad (12)$$

$$B = \frac{S(R_0)}{\beta_{part}(R_0) + \beta_{mol}(R_0)} \quad (13)$$

$$C = -2 \int_{R_0}^R LR_{part}(r) \times S(r) \times Tr(r, R_0) dr \quad (14)$$

$$Tr(r, R_0) = e^{-2 \int_{R_0}^r [LR_{part}(r') - LR_{mol}] \times \beta_{mol}(r') dr'} \quad (15)$$

where  $R$  is the distance from lidar,  $R_0$  is the distance to the reference layer,  $\beta_{part}$  is the aerosol backscatter coefficient,  $S(R)$  is a range corrected signal proportional to ATB,  $LR_{part}$  is the aerosol lidar ratio (extinction-to-backscatter ratio),  $LR_{mol}$  is the molecular lidar ratio. For the sake of simplicity, we dropped the wavelength index from the variables in Eqs. (11–15). As for the lidar ratios  $LR_{part}$ , in COSP2 it is estimated from particle type based on mass mixing ratios of water and ice and type of a cloud (convective or large scale), so in our numerical experiment we kept the consistency with COSP2 parameterization yielding  $LR_{part}$  values listed in the 7<sup>th</sup> column of Table 1, and we consider them constant within cloud layer.

It is known that Klett's method depends on the choice of a reference layer  $R_0$  and that the solution is stable only if  $R_0$  is chosen on the far end of the observation profile (Fernald, 1984). At the same time, it is clear that the method will not work if the reference layer is masked by an opaque cloud. If the reference layer is indeed obscured,  $B$  approaches zero according to Eq. (13), the first term in Eq. (11) becomes dominated by cloud extinction, the solution becomes numerically unstable, and the retrieval fails throughout the whole profile. In this case, it makes sense to take the farthest point of the profile that is not obscured by a thick cloud, as a reference one. We discuss this approach in more details in the companion technical note dedicated to real lidar's processing from the detector's level to L1 and L2 (Haddouche and Feofilov, 2026a).

## 2.6. Spatial and temporal averaging L1 and L2 data

As one can see from Table 1 and Fig. 2, our test cloud dataset includes scenes that are not favorable for observations due to low signal-to-noise ratio (SNR) like solar-illuminated scenes, very thin clouds, and clouds masked by other clouds. To improve the SNR and resulting cloud detection in these cases, both space-borne and ground-based lidars average the backscatter signal. For the former, averaging over flying distance leads to spatial resolution degradation (e.g. the optical properties retrieved from ALADIN/Aeolus have 3 to 90 km horizontal resolution, depending on the strength of backscatter signal, see Wang et al., 2024). For the latter, the integration periods over time typically range from minutes to tens of minutes depending on the lidar type and observation conditions. Longer averaging is generally required during daytime observations because of increased solar background noise, whereas nighttime measurements can achieve higher temporal resolution (Ansmann et al., 1992; Ansmann and Müller, 2005). Since the spatial atmospheric variability corresponding to long-term averaging is high, we opted to limit the averaging by 100 laser shots that corresponds to ~30km flight distance that is a



reasonable tradeoff between the SNR improvement and the degradation of the horizontal resolution. We note that this averaging is representative of the correlation lengths of the cirrus and mid-level scenarios in Table 1. It has not been validated against boundary-layer cumulus or stratocumulus, where CALIOP and ATLID retrievals are typically reported at native ~300 m resolution and where the constant-vertical-profile assumption of Sect. 2.1 is itself a coarser approximation (Sect. 2.1). The comparability results for the shallow-to-opaque stratocumulus scenes (12–15, Table 1) should therefore be read with this caveat; our aim was to keep the treatment uniform across scenes rather than tune it to each cloud regime. For the ground-based lidars, we used the same averaging, given that we use the same type of instruments with the same telescope size and optical throughput that are much closer to their targets, but during daytime the SNR is spoiled by a larger FOV leading to increased solar background signal. Further adjustment of averaging parameters is out of the scope of the present study, the main purpose of which is to show the differences in the information content of L1 and L2 profiles obtained in different viewing geometries. Both the L1 and L2 comparisons of observation geometries discussed below correspond to 100-element averaging when it comes to the quantitative characteristics of the agreement. To illustrate realistic noise, the lidar curtains in the next section are shown without averaging.

## 2.7. Zenith-to-nadir profile comparison: comparability score

To quantify the similarity between lidar profiles measured in zenith and nadir geometries, we define a metric based on similarity of height profiles, be it the scattering ratio from Eq. (9) (SR1) or Eq. (10) (SR2), or the retrieved  $\beta_{part}(\lambda, z)$ . For the former two cases, the input data for the numerator (and for Eq. (10), also the denominator) are sourced either directly from L1 or estimated from L2 data. The molecular backscatter and extinction coefficients,  $\beta_{mol}(\lambda, z)$  and  $\alpha_{mol}(\lambda, z)$ , are retrieved from L1 data (see Eqs. (2–5)). We term our agreement metric the "Comparability Score" (CS), which should satisfy the following requirements:

- It should be independent of absolute values or units of compared two input vectors (profiles)  $\vec{a}$  and  $\vec{b}$ .
- It should penalize the anti-correlation of profiles and favorize their correlation.
- It should favorize the close absolute values, given that the next condition is not fulfilled.
- It must approach zero if the profiles, while similar to each other, are close to a fixed reference value (1.0 for SR1(z), 0 for  $\beta_{part}(z)$  via a separate branch of Eq. (16) for the SR case); this prevents two clear-sky or aerosol-free profiles from scoring as highly comparable simply because both sit near their baseline, independent of whether either contains a real cloud or aerosol signal.
- It should decrease in proportion to the fraction of invalid or non-overlapping points, via the  $\frac{N_{joint}}{N_{tot}}$  factor in Eq. (17).

For the input vectors  $\vec{a}$  and  $\vec{b}$  (e.g. the SR1(z) calculated in zenith and the SR1(z) calculated in nadir geometry), comparability score CS is calculated as follows:

$$CS = 0 \text{ if } \vec{a} \approx 0 \text{ or } \vec{b} \approx 0 \quad (16)$$

$$CS = \frac{N_{joint}}{N_{tot}} \times \{(1 - w_{shape}) \times G_{corr} + w_{shape} \times G_{shape}\} \quad (17)$$

$$G_{corr} = \frac{\rho(\vec{a}, \vec{b}) + 1}{2}; \quad G_{shape} = e^{-2 \times \frac{|\vec{a} - \vec{b}|}{|\vec{a}| + |\vec{b}|}} \quad (18)$$

where  $\vec{a}$  and  $\vec{b}$  are the mean values over the valid elements of the input vectors,  $N_{joint}$  is the number of points whose value exceeds a physical validity threshold, set to zero for both SR(z) and for  $\beta_{part}(z)$  (Sect. 2.4),  $N_{tot}$  is a total number of points in the analyzed cloud layer (at the 100 m vertical resolution used for Table 1's altitude range),  $G_{corr}$  is a correlation term contribution to the score,  $\rho(\vec{a}, \vec{b})$  is Pearson's correlation coefficient,  $G_{shape}$  is a scaled difference term and  $w_{shape}$  is its contribution to the total score. The latter factor is chosen arbitrarily to make both terms in the brackets of Eq. (17) representative, and in the present work we used  $w_{shape} = 0.5$ .



### 3. Comparing zenith and nadir observations

#### 3.1. Comparison based on L1 data

In this section we quantify, rather than merely illustrate, how the viewing-geometry difference propagates from L1 to L2 quantities and into a numerical measure of comparability, laying the groundwork for the decision scheme in Sect. 5. In Figs. 3 and 4, we show the vertical profiles of  $AMB(z)$ ,  $APB(z)$ ,  $ATB(z)$ , as well as  $SR1(z)$  calculated in accordance with Eqs. (6–8) in nadir and zenith observation geometries, respectively. For the sake of simplicity, we show only the calculations at 355 nm. Qualitatively, the ones at 532nm look similar, and we discuss the agreement at this wavelength below. To illustrate the differences between solar illuminated scenes and nighttime ones on one plot, we split each scene in two halves: the left one shows the nighttime case, the right one is for day. Let us first qualitatively analyze the similarities and differences between the nadir and zenith geometries, between day and night, and between the scenes. For clear sky cases (column #1 of Figs. 3 and 4), the most remarkable feature is the difference between day and night for zenith geometry. This is not surprising given that the FOV of a ground-based instrument is large in our calculations. For nadir-looking analog of ATLID lidar, the difference is smaller that is in a good agreement both with our preliminary estimates (Feofilov et al., 2023) and with real ATLID data. In the discussion materials of Feofilov et al. (2023), we provided an explanation of this effect: it is a combination of relatively low Solar radiance flux at 355 nm, narrow FOV of this instrument, and narrow effective spectral width of its Mie channel. The difference in night/day noise remains valid for all scenes, but the absolute values differ. This is related to cloud reflection/scattering that depends on the scene and observation geometry. As stated before, we reduce the effects of noise by a factor of 10 through 100-point averaging in the comparisons below. A distinct feature of all scenes starting from scene 4 (thin *Ci*) is a “shadow” in attenuated backscatter panels of Figs. 3 and 4. In the cases of single-layer thick clouds (scenes 6, 10, and 14) and double-layer clouds (scenes 7, 11, and 15), the shadow is more pronounced, and the layers in it are barely accessible for the remote sensing. Finally, the biggest difference between Fig. 3 and Fig. 4 is the position of the area, which is partially or fully masked by a cloud: for nadir observations (Fig. 3), it is below the cloud layer whereas for zenith observations (Fig. 4) it is above the cloud layer. This has been already known since the beginning of atmospheric lidar observations, but this time we will address the issues linked with these differences.

To illustrate the comparability issues in L1, we use the  $SR1(z)$  profiles (Eq. (9)) built from the averaged  $ATB(z)$  and  $ATB_{mol}(z)$  profiles taken from the center of each scene's 1000-profile block, away from the transitions with adjacent scenes, both for nadir and lidar geometries (Fig. 5). Where it is appropriate, we also show the  $SR2(z)$  profiles (Eq. (10)) built from the averaged  $ATB(z)$  and  $AMB(z)$  profiles. We do not show them for the scenes, where the  $SR2(z)$  profiles are off the limits. Each time, we focus on the uppermost cloud layer to show the differences between single-layer scene and the same scene masked by an almost non-transparent cloud below. To illustrate the best possible cases of such comparisons and address only the observation geometry effects, we consider a perfect spatiotemporal overlap during nighttime.

As one can see in Fig. 5, the closest nadir-zenith curves are the  $SR2(z)$  for optically thin single-layer clouds. But, due to noisy values in the denominator of Eq. (10), this metric becomes unstable and does not work well for semi-transparent or thick clouds. From this point of view, both nadir- and zenith-geometry  $SR1(z)$  curves are more stable and they demonstrate an explainable physical behavior, but their comparability is worse than that of  $SR2(z)$  for thin cases: all the profiles for the transparent cases experience the same “opposite attenuation” problem, when the nadir and zenith profiles cross because the radiative transfer responsible for the formation of  $ATB(z)$  happens in opposite directions for these observation geometries. As for opaque cloud  $SR(z)$  profiles or the cloud  $SR(z)$  profiles over opaque clouds, they are not comparable at all, be it a  $SR1(z)$  or  $SR2(z)$ . This qualitative picture is not new to lidar remote sensing; what we add here is a quantitative measure of it, which makes it possible to decide whether the case is worth comparing with the collocated measurement, and if yes then what variables are comparable. The CS values we provide can be thresholded, compared across scenes, and ultimately built into the decision scheme of Sect. 5. We keep this framing in mind for the L2 comparisons below.

If we calculate the comparability score values for all the pairs of zenith and nadir profiles according to Eqs. (16–18) and the flowchart shown at the bottom of (Fig. 1), for the daytime and nighttime scenes, we will obtain the tables shown in Fig. 6. The rows of these tables correspond to cloud scenes whereas the columns correspond to combinations of wavelength (355 nm or 532nm), illumination conditions (day/night) and overlapping conditions: “close” representing nearly perfect spatiotemporal overlap, and “far” representing realistic separation typical of satellite-to-ground validation activities, yielding 8 columns for each table. As one can see, the values in the left-hand side table and middle table, which correspond to CS calculated from  $SR1$  (Eq. (9)) and  $SR2$  (Eq. (10)), respectively, follow the patterns visible in Fig. 5. Overall, the observed behavior fits well our general understanding of the physics of observations and the values chosen for the comparison. For the  $SR1$ : (1) the first three lines in  $SR1$  score demonstrate a good agreement between zenith and nadir, that is expected because  $SR1$  for the molecular case is equal to 1 and only the noise and spatial variability can affect the agreement, and the



attenuation in subvisible clouds is so weak that it cannot dramatically change the SR1 slope; (2) the cases of perfect overlap (“close” in the terms of Fig. 6) have expectably higher CS values than non-perfect ones (“far” in terms of Fig. 6); (3) CS values for the daytime scenes are lower than those for the nighttime. The SR2 scores (Fig. 6, middle panel) follow the same patterns as the SR1 scores, but there are several important differences: (1) the absolute CS values are in general higher than those based on SR1; (2) correspondingly, there are more scenes that one can compare using SR2; (3) daytime CS values at 532nm are strongly affected by AMB noise in the denominator of Eq. (10) for all the scenes, so the CS based on SR1 is preferable for subvisible and thin Ci at this wavelength during daytime.

### 3.2. Comparison based on L2 data

To quantify how geometry effects propagate through retrieval algorithms and into optical products and whether L1 or L2 is preferable for the comparison, we now examine the L2 data. In Fig. 7, we show the retrieved  $\beta_{part}(355nm, z)$  vertical profiles for nighttime scenes. As one can see, one of the obvious differences between SR1, SR2 and these profiles is the contrast between zenith and nadir profiles in the case of opaque cloud beneath the main scene. Since in these cases the reference layer for Klett’s method had to be below the opaque cloud for zenith and above it for nadir retrievals, the retrievals for the signals obtained in zenith observation geometry did not converge for the points above the cloud. This is an expected behavior for a signal strongly attenuated in opaque cloud. Some of the scenes with semi-transparent clouds yielded noisy retrievals. This might be improved by fine-tuning of Klett’s algorithm that is out of scope of the present work, because our task was to show the information content of the signal processed with the same algorithm in both observation geometries. As for the subvisible, thin, and certain semi-transparent cases, both curves look quite similar. The differences are due to noise and algorithmic stability of the retrieval. Switching to a global comparison (Fig. 6, right-hand side table), one notes a larger contrast for “good” and “bad” cases compared to SR1- and SR2-based CS values shown in the middle and left-hand side panels of this plot: for the optically thick clouds and scenes over them, the CS values are close to zero whereas for the optically thinner scenes the CS values are close to 80-90%.

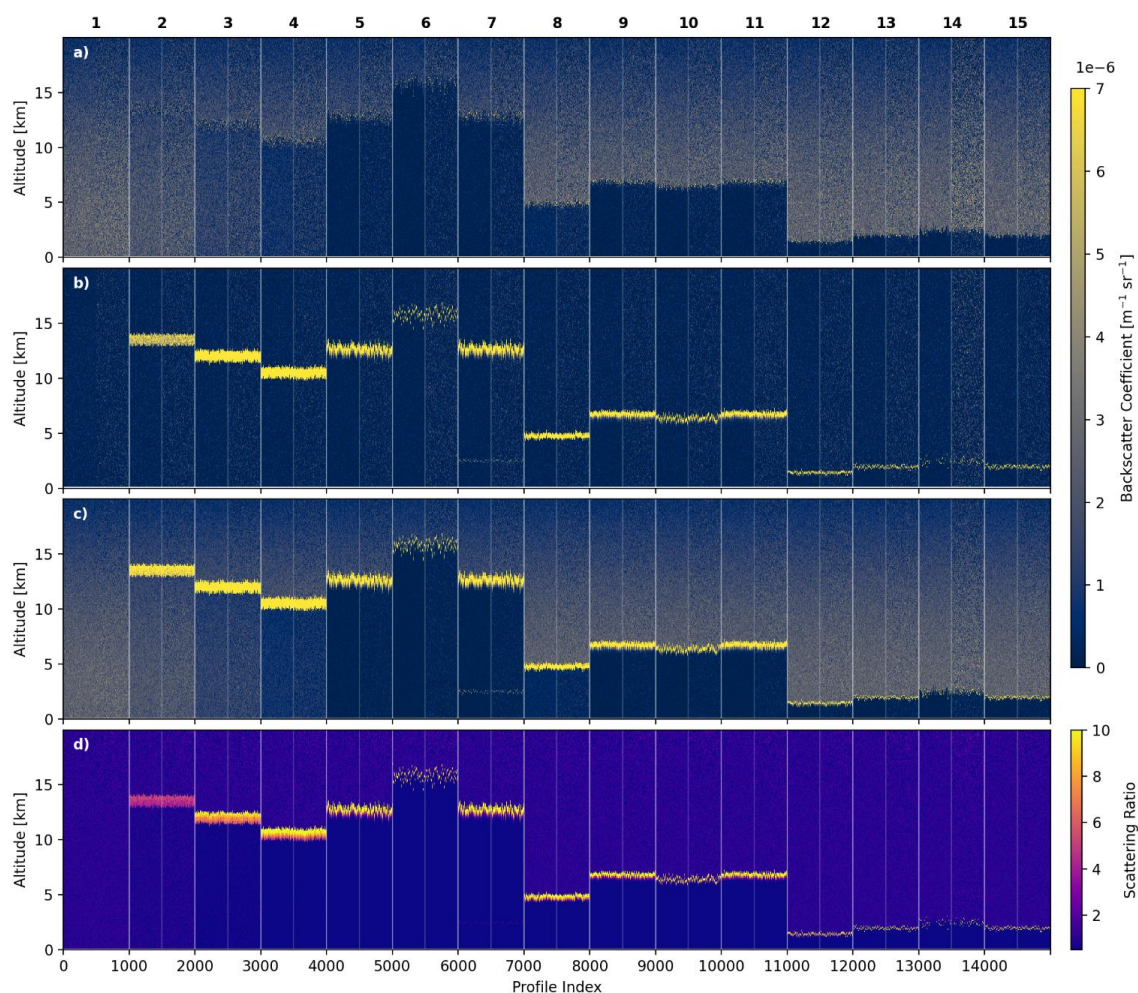


Figure 3: Noisy profiles of (a) attenuated molecular backscatter (AMB), (b) attenuated particulate backscatter (APB), (c) total attenuated backscatter (ATB), and (d) scattering ratio (SR, see Section 2.4). Nadir observation geometry, modeling for 355 nm. The cloud scene numbers on top of the plot correspond to the first column of Table 1. For each scene, the first half is shown for nighttime conditions, and the second half – for daytime conditions (SZA=45°).

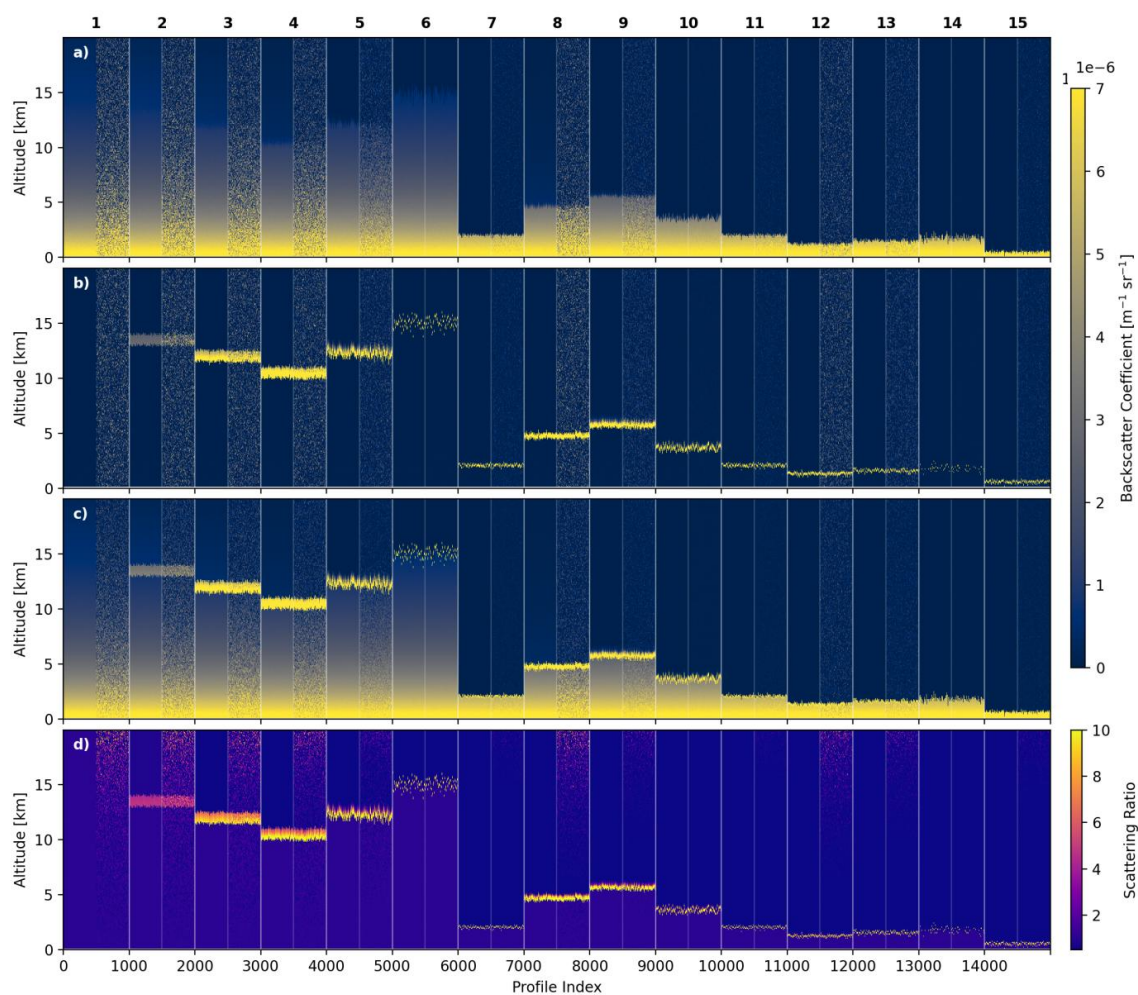


Figure 4: Same as Fig. 3, but for zenith-viewing geometry.

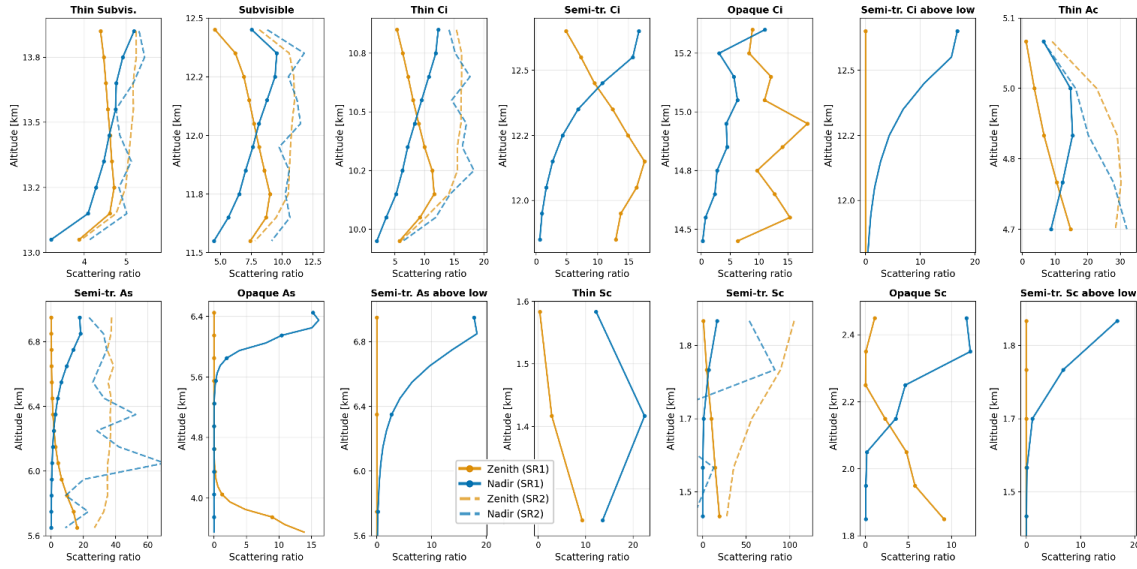


**Figure 5: Comparison of zenith and nadir viewing geometries using SR1 (Eq. (9)) and SR2 (Eq. (10)) definitions. The absence of SR2 curves for some scenes means that the SR2 values in these cases are beyond the axis limits (unphysical values, Sect. 2.4).**

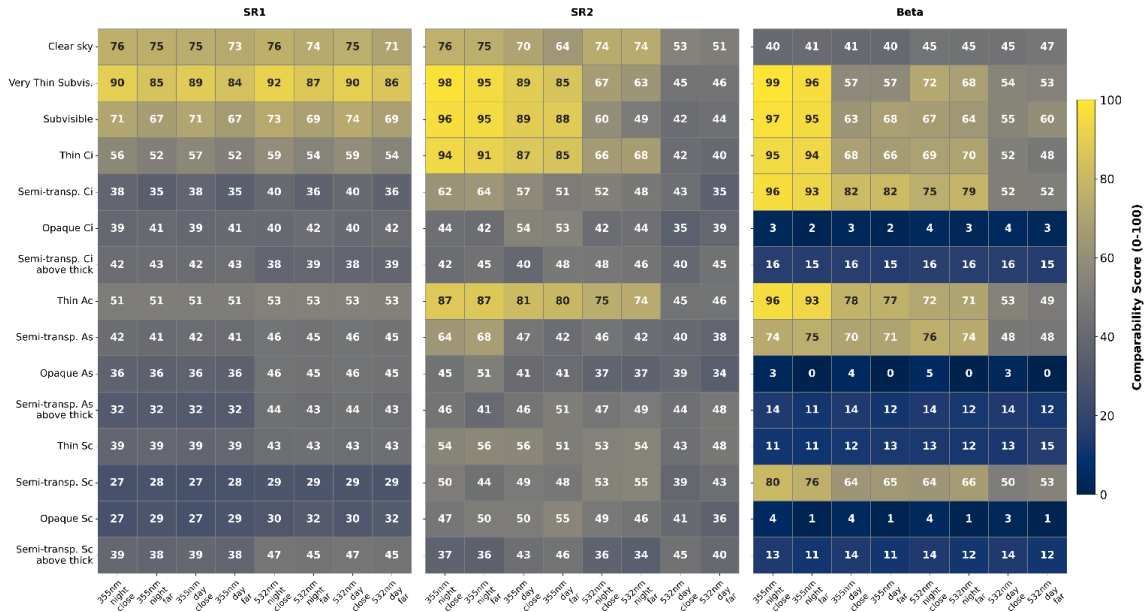


**Figure 6: Comparability score values calculated for all scenes. Left: CS values estimated for SR1 (Eq. (9)), non-HSRL lidar; middle: CS values estimated for SR2 (Eq. (10)), HSRL lidar; right: CS values estimated for the  $\beta_{part}(z)$  retrieved in nadir and zenith geometries.**

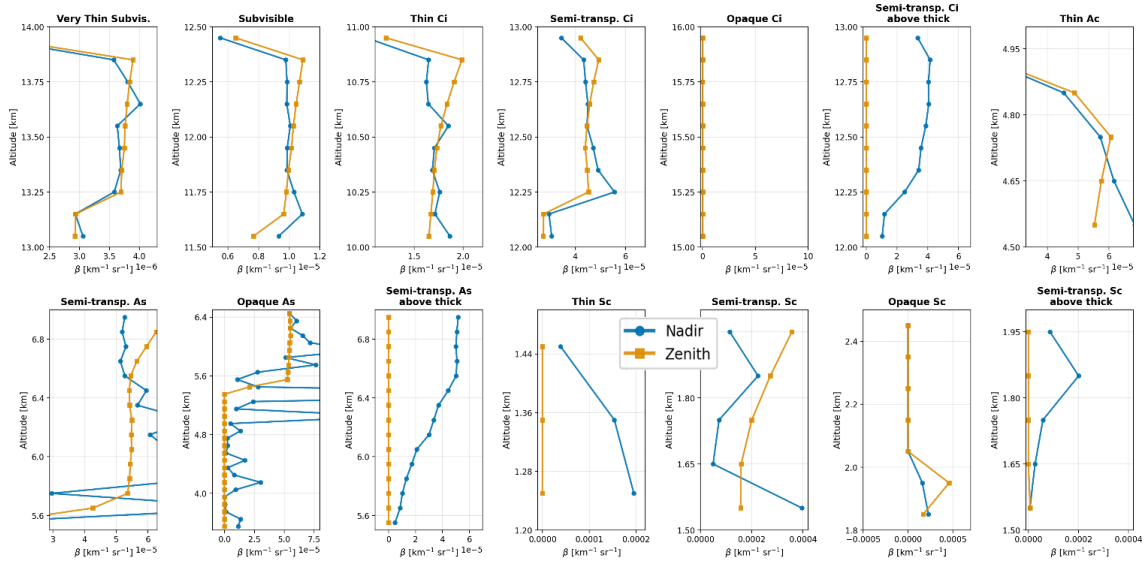


**Figure 7: Comparison of zenith and nadir viewing geometries using the  $\beta_{\text{part}}(355\text{nm}, z)$  profiles retrieved in zenith and nadir geometries. The absence of the  $\beta_{\text{part}}(355\text{nm}, z)$  nadir curve for opaque Ci scene means a failed retrieval in this case.**

#### 4. ATLID to IPRAL comparisons

To validate the theoretical conclusions from Section 3, we examine actual ATLID/IPRAL collocations from 2024–2026. The specifications of ATLID instrument used in calculations below can be found in Table 1 of Feofilov et al. (2023). The IPRAL setup is described in the next section.

##### 4.1. IPRAL 355 nm lidar and its data processing

IPRAL is a high-performance, multi-wavelength lidar deployed at the Site Instrumental de Recherche par Télédétection Atmosphérique (SIRTA) observatory (Haeffelin et al., 2005) in Palaiseau, France (48.71°N, 2.21°E). The system is capable of measuring multiple elastic and inelastic (Raman) channels (Alraddawi et al., 2025) across six wavelengths (355, 387, 408, 532, 607, and 1064 nm). However, for the validation of ATLID, we use exclusively the 355 nm elastic channel in order to match the operating wavelength of the spaceborne instrument. At 355 nm, IPRAL measures both the co-polarized (parallel) and cross-polarized (perpendicular) components of the backscattered signal using analog and photon-counting detection modes for each channel.

The IPRAL emitter utilizes a high-power Nd:YAG laser that emits laser pulses at 355, 532, and 1064 nm with an energy of approximately 120 mJ per pulse at 355 nm. The beams are directed into the atmosphere through a series of mirrors and an Achromatic Beam Expander. For the reception, the IPRAL system uses two telescopes to collect the backscattered signal: a Far-Range Telescope with a 60 cm diameter, and a Near-Range Telescope with a 20 cm diameter. The collected light then passes through a series of filters to separate each wavelength, followed by neutral density filters that control the signal intensity. Finally, the signals are detected by Photomultiplier Tubes (PMT), which output two simultaneous data streams: an analog signal and a photon-counting signal; the exceptions are the 1064 nm channel (avalanche photodiode, analog only) and the 607 nm channel (photon-counting only), neither of which is used in this work.

The optical configuration yields a total of 6 distinct raw signals at 355 nm, corresponding to the combinations of the two telescope ranges (Far/Near), the polarization states (parallel/perpendicular), and the two acquisition modes (analog/photon-counting). The processing chain from L0 to calibrated Level 1 and Level 2 products used in this work is described in Had-douche and Feofilov (2026a), and here we provide only the key elements of this chain.

First, the raw 355 nm signals are background-subtracted and range-corrected to obtain range-corrected signal  $\text{RCS}(z)$  for each of the four far-range channels (parallel/perpendicular  $\times$  analog/photon-counting). Calibration into physical units proceeds by fitting each channel, over clear-sky range gates only, against the theoretical molecular backscatter  $\text{ATB}_{\text{mol}}(z)$  computed from the co-located Trappes radiosonde profiles in accordance with Eq. (8). Next, for each channel we solve for a multiplicative coefficient  $k$  (lidar constant) and a residual background offset  $\Delta$  using a two-step procedure. First,  $N$  random



pairs of altitude points ( $\sim 100$ ) are drawn from two clear-sky windows, and  $k_i$  and  $\Delta_i$  are computed analytically for each pair  $i = 1, \dots, N$ ; the sample means give an initial estimate ( $k_0, \Delta_0$ ). This estimate then serves as the starting point for a Nelder–Mead minimization, which refines  $k$  and  $\Delta$  to minimize the mean absolute percentage error between  $k \times \{RCS(z) - \Delta\}$  and  $ATB_{mol}(z)$ .

A critical step at this stage is the correction of polarization cross-talk in the polarizing beam splitter: because the true perpendicular signal is only 0.4% of the parallel one in clear sky (e.g. She and Baumgartner, 2022), even a percent-level optical leakage from the parallel into the perpendicular channel is enough to dominate the latter. We solve the  $2 \times 2$  cross-talk system using two independent reference depolarization values: the molecular depolarization coefficient, and a mean depolarization cirrus value obtained from one year of CALIOP observations of the Palaiseau region, equal to  $0.3 \pm 0.05$ . This approach yields the calibration coefficients without any physical manipulation of the instrument, e.g. rotation of the half-wave plate as required by the classical  $\pm 45^\circ$  method (Freudenthaler et al., 2009; Freudenthaler, 2016).

The calibrated analog and photon-counting profiles are then merged over the range where the two profiles match, yielding a single calibrated profile for each polarization channel. The merged parallel and perpendicular profiles are finally summed to give the total attenuated backscatter  $ATB(z)$ , which constitutes the IPRAL Level 1 product used below.

Finally, Level 2 particle backscatter  $\beta_{part}$  is derived through the Klett–Fernald–Sasano backward inversion (Ansmann and Müller, 2005). Profiles are first averaged to improve the ATB signal-to-noise ratio (5 profiles for nighttime cases and 10 for daytime cases). The inversion is then applied starting from a clear-sky reference altitude, assuming an aerosol lidar ratio appropriate to the observed scene, selected in the range 30–50 depending on the case. The two polarization channels are inverted separately and summed; the limits of this approximation, the full derivation, the calibration method, and the numerical implementation are given in the companion technical note (Haddouche and Feofilov, 2026a).

We estimate random uncertainty empirically rather than propagating it analytically through the full chain; the merge of the analog and photon-counting signals mixes two different noise regimes, so the resulting uncertainty no longer admits a simple closed form. For each profile, we subtract a running mean from the range-corrected signal to isolate the high-frequency residual, then scale its RMS with the smoothed signal to obtain a noise estimate that grows with signal strength, as expected for a shot-noise-limited detector; we use this approach for  $ATB(z)$ ,  $SR1(z)$ , and  $\beta_{part}(z)$ . Within the cloud layer, we deliberately hold the error at its value at the  $SR1(z)$  or  $\beta_{part}(z)$  peak rather than letting it evolve further: for the zenith-viewing (ground-based) profile, we carry the peak error upward above the peak, whereas for the nadir-viewing (spaceborne) profile, we carry it downward below the peak.

## 4.2. Profile selection: correlation as an objective criterion

For the typical distances involved in spaceborne-to-ground-based lidar collocations, an inhomogeneous scene can make direct intercomparison difficult: minimizing spatiotemporal distance alone does not guarantee that ATLID and IPRAL sample the same atmospheric structure: under a heterogeneous cloud, the closest overpass in space and time can still pair two genuinely different scenes. To address this, among the candidate ATLID profiles and IPRAL averaging windows around a given overpass, we select the pair that maximizes the Pearson correlation coefficient  $\rho$  between the two vertical profiles, following the approach used for spaceborne–ground collocations of SABER and a mesospheric lidar (Feofilov et al., 2012). This raises a natural question: how large a correlation can arise simply by chance, between two profiles that do not in fact sample the same air mass?

We address this with a decomposition of the compared profiles into background, cloud, and noise contributions (Appendix A). For a semi-transparent scene sampled at ATLID’s vertical resolution, the argument gives a one-sided probability of about 12% that a correlation as high as our selection threshold (correlation coefficient  $\rho > 0.9$ ) arises between two profiles that do not represent the same cloud, under conservative assumptions for the background correlation and the cloud’s effective degrees of freedom (DOF). In other words, roughly one pair in nine could still produce a spuriously high correlation by chance; in the other eight cases the matched pair genuinely represents the same cloud, displaced horizontally between the ATLID and IPRAL measurements. The estimate indicates that the correlation-maximization criterion is not a mere curve-fit: a high correlation between two genuinely different air masses remains possible, but is statistically disfavored.

To illustrate the validity of maximizing the correlation coefficient search approach, we performed a test using real data from IPRAL and ATLID. For “null-distributions”, we took the mismatched-date collocations ( $D0 \pm 3$ ) in the same geographic area ( $\Delta lat < 2^\circ$ ;  $\Delta lon < 10^\circ$ ) whereas the true collocations ( $D0$ ) were our standard collocations explained in details in the next section. For the reference IPRAL profile, we identified the cloud layer using the following set of rules: (a) the cloud layer(s) should lay between 3000m and 18000m to exclude any overlapping effects, low-level thick clouds, and noise effects; (b) the layer is defined from the 100pt smoothed  $SR1(z)$  profile calculated for IPRAL (Eq. (9)); (c) for this profile, the position of its maximal value  $SR1_{max}$  is searched as well as the two positions at half of the maximum; (d) the upper and the lower limits are calculated as doubled values of the height difference between the maximum and each of these half-maximum positions, added to (upper limit) or subtracted from (lower limit) the position of the  $SR1(z)$  maximum;  $SR_{edge}$  is the  $SR1(z)$



value at these limits; (e) the prominence value  $Pr$  calculated as  $Pr = (SR1_{max} - SR1_{edge})/SR1_{max}$  should be greater than 0.3.

The latter criterion is necessary: for  $Pr < 0.3$  an increasing fraction of reference windows are weakly structured or flat, so that the correlation maximum search might lead to false matches. We tested two null distributions: for the first one, we performed the same correlation maximum search as for the true collocations. For the second one, we just picked up a random profile from the collocated subset. These two null distributions serve different purposes. The search-and-maximize null reproduces the exact procedure used to select D0 matches, and therefore gives the relevant false-discovery rate for the method as actually applied. The single-draw random null distribution omits this search step. Comparing it to the first one will show how the search itself, rather than mere spatiotemporal proximity, inflates the apparent correlation between unrelated profiles.

Table 2: Empirical false discovery rate (FDR) as a function of the prominence threshold  $Pr$  and correlation coefficient threshold  $\rho_{threshold}$ . The hit rate is the fraction of cases in which the best-matched pair has  $\rho \geq \rho_{threshold}$ , among the cases with a reference-window prominence greater than or equal to corresponding  $Pr$  value. D0 stands for the true collocations hit rate at the day of the collocation, Null stands for the hit rate for the unrelated days, Enrichment column represents the ratio between D0 hit rate and null hit rate, FDR is the ratio of the Null to the D0 hit rate (equal to 1/Enrichment). Sample sizes ranged from  $N(D0)=23$  to 90 across the tested  $Pr$  values; the single-draw random-null hit rate remained at or below  $\sim 1\%$  throughout.

| Screening thresholds |                    | Match-rate statistics |          |            |         |
|----------------------|--------------------|-----------------------|----------|------------|---------|
| $Pr$                 | $\rho_{threshold}$ | D0 [%]                | Null [%] | Enrich [x] | FDR [%] |
| 0.3                  | 0.7                | 84.1                  | 29.4     | 2.86       | 35.0    |
| 0.3                  | 0.8                | 52.4                  | 14.2     | 3.68       | 27.2    |
| 0.3                  | 0.9                | 7.9                   | 1.3      | 6.13       | 16.3    |
| 0.4                  | 0.7                | 85.4                  | 17.0     | 5.03       | 19.9    |
| 0.4                  | 0.8                | 56.1                  | 10.6     | 5.32       | 18.8    |
| 0.4                  | 0.9                | 12.2                  | 1.8      | 6.65       | 15.0    |
| 0.5                  | 0.7                | 84.6                  | 17.6     | 4.80       | 20.8    |
| 0.5                  | 0.8                | 59.0                  | 11.0     | 5.38       | 18.6    |
| 0.5                  | 0.9                | 12.8                  | 1.9      | 6.73       | 14.9    |
| 0.6                  | 0.7                | 84.4                  | 16.1     | 5.24       | 19.1    |
| 0.6                  | 0.8                | 59.4                  | 9.4      | 6.29       | 15.9    |
| 0.6                  | 0.9                | 12.5                  | 2.2      | 5.63       | 17.8    |
| 0.7                  | 0.7                | 87.0                  | 17.1     | 5.07       | 19.7    |
| 0.7                  | 0.8                | 65.2                  | 11.4     | 5.71       | 17.5    |
| 0.7                  | 0.9                | 17.4                  | 2.9      | 6.09       | 16.4    |

Repeating the comparison for other prominence thresholds yields consistent results (Table 2): D0 exceeds  $\rho=0.8$  in 52–65% of cases, against 9–14% for the search-and-maximize null and  $\leq 1\%$  for the single-draw random null. This near-absence of high correlation under random pairing suggests that the search-and-maximize null is the appropriate, conservative baseline. The search step can inflate apparent correlation between unrelated profiles through ordinary multiple-comparisons effects, which is exactly why FDR is benchmarked against it rather than against single-draw random pairing. The D0/search-and-maximize gap itself holds across the full correlation distribution (Kolmogorov–Smirnov test,  $p < 10^{-15}$  at every prominence value tested), though a stricter correlation threshold is needed to compensate when the reference window is weakly structured. Analyzing Table 2, we conclude that the tradeoff between sample size and comparison quality lies near a prominence value of 0.3–0.4 and  $\rho_{threshold} \approx 0.8$ . For these values, the enrichment is about 5 and the false discovery rate is about 20%. This value is somewhat higher than the theoretical estimate made in the first half of this section; one plausible explanation is an overestimation of the effective DOF in Eq. (A5), since not all the cloud peaks are characterized by a well-developed internal structure. Still, the empirical null-hypothesis test made here indicates that the maximal correlation coefficient search method is valid in approximately 80% of cases and if one wants to increase this ratio, one should search for the structures with high DOF values (e.g. two peaks, mixture of slopes, and so on).

We also explored improving the method by identifying the common air mass independently, tracking cloud structures between the ATLID overpass and SIRTa with 15 min geostationary imagery from the Spinning Enhanced Visible and Infra-Red Imager (SEVIRI) via block matching, following the wind-advection logic of aerosol trajectory-matching studies (Lopes et al., 2013; Tesche et al., 2013). The method returns a time shift for the IPRAL averaging window and the coordinates of the matching track point. SEVIRI senses only cloud-top motion, which limited the results, so we do not pursue the approach



here. It remains a promising complement to the correlation-maximization search above and we leave its evaluation to future work.

### 4.3. EarthCARE/ATLID and SIRT/ATLID collocation and scene screening

First, we identify ATLID overpasses within a 300 km diameter search region centered on SIRT/ATLID (following the collocation approach explored in Haddouche, 2025); given EarthCARE's sun-synchronous orbit, this typically yields 10–13 overpasses per month, occurring at one of two near-fixed local times:  $\approx 14:00$  UTC and  $\approx 01:00$  UTC. Out of these overpasses, only a fraction has coincident ground-based data, because IPRAL does not operate continuously; it is shut down during precipitation, maintenance, and whenever aircraft overfly SIRT/ATLID. Overall, for the period considered, the acquisition rate is about 54%. For a genuinely close overpass, spatial mismatch and cloud advection become negligible; an example of this regime is the case reported in Baars et al. (2026) where the horizontal distance between the two beams was  $\approx 0.26$  km. However, our collocations mostly fall within 10–100 km overlap window, so the correlation-based window selection described in Section 4.2 is required to identify the comparable pair of profiles.

In the next step, we exclude scenes that are optically opaque. This restriction follows directly from the simulation results of Section 3: an opaque layer extinguishes the signal before it reaches the ground (or the satellite, depending on viewing geometry), so neither  $SR(z)$  nor  $\beta_{part}(z)$  can be used for the comparison. Retaining only optically thin to semi-transparent scenes therefore follows directly from the comparability criteria established earlier in the paper. The next filtering step is related to the shape of the ground-based lidar  $SR(z)$  and its agreement with the ATLID's  $SR(z)$  profile: at this step, we use the ATLID's  $\beta_{part}(z)$  and  $\alpha_{part}(z)$  profiles and feed them to forward simulator (Eq. (6)) to obtain a zenith-geometry  $SR(z)$ , which is directly comparable with ground-based lidar. This is an important conversion, and without it, the “native” geometry  $SR(z)$  might skew the shape (see Fig. 5) and the selection will prioritize a wrong shape. In accordance with previously found thresholds (Sect. 4.2), we do not consider the cases, for which the best pair of profiles yielded  $\rho < \rho_{threshold}$ , and we also skip the cases with prominence less than 0.3. In addition, we filter out the cases with low DOF values because these cases are prone to false mismatches (e.g. just a single-point peak or a peak without any features). Overall, these filtering steps leave us just a few profiles to be compared, but we are armed with a theoretical and experimental evidence that the cases passing these strict selection criteria correspond to an apple-to-apple comparison. We stress here that the selection is not based on the absolute values of the profiles (cherry-picking). Up to this moment, we operated only with the correlation coefficient and comparability score criteria. In the next section, we show several representative collocation cases, each of which illustrates a specific atmospheric scene and its projection to lidar variables.

### 4.4. Case studies: agreement and disagreement between IPRAL and ATLID L2

Table 3 lists the parameters of the four scenes and the corresponding variables; for each case it gives the identified cloud type, its vertical extent, and the correlation, prominence and CS metrics used to characterize the level of agreement between IPRAL and ATLID L2. The selected cases span the range of behaviors observed in the broader collocation dataset. The first case (2025-02-16) is a semi-transparent cirrus with close agreement between the instruments, which makes it our reference case. The second (2025-03-04) is a semi-transparent altostratus with a systematic amplitude offset. The third (2025-06-11) is a three-layer scene with partial, layer-dependent agreement. The fourth (2025-07-04) is a subvisible cirrus, for which a low prominence score understates a genuinely well-matched retrieved backscatter profile. We discuss each case individually below. For each case, Appendix B additionally shows the spatially/temporally averaged pairs and pairs closest-in-space/time and the, to make explicit what is the added value of the correlation-maximization step described in Sect. 4.2 versus these simpler alternatives.

Figures 8–11 compare the scattering ratio computed independently in each instrument's native viewing geometry (panel a) against the retrieved particle backscatter  $\beta_{part}(z)$  (panel b). Contrary to the selection step above, we use native geometry to illustrate our findings summarized in Fig. 6. We show a limitation of the SR1-based comparison (Eq. (9)) that the  $\beta_{part}(z)$  comparison largely avoids: SR1 depends on the attenuation along the optical path, which differs between zenith and nadir viewing geometries, and can therefore appear degraded even if the profile is normalized by its own molecular reference (Eq. (8)). This problem might be overcome if the  $AMB(\lambda, z)$  profile is measured (Eq. (7)), but it is not available for IPRAL because it is not an HSRL instrument. In contrast, the  $\beta_{part}(z)$  profile already accounts for layer-by-layer attenuation, so the  $\beta_{part}(z)$  comparisons are expected to give closer results.



Table 3: Summary of ATLID/IPRAL collocation cases discussed in Section 4.4. Cloud top and base altitudes are taken from the IPRAL peak-detection window;  $\tau_{col}$  is the integrated optical depth in the column;  $Pr$  and  $\rho$  are the prominence and Pearson correlation coefficient of the selected best-match pair;  $CS_{SR1}$  and  $CS_{\beta}$  are read from the left and right tables of Fig.6 for the simulated scene closest to the observed one;  $\Delta dist$  and  $\Delta time$  correspond to the best pair of measurements. Ci, As and Sc denote cirrus, altostratus and stratocumulus; Semi-tr. and Subvis. denote semi-transparent and subvisible.

| Date       | Day/<br>night | Cloud type   | Cloud<br>base [m] | Cloud<br>top [m] | $\tau_{col}$ | $Pr$ | $\rho$ | $CS_{SR1}$<br>[%] | $CS_{\beta}$<br>[%] | $\Delta dist$<br>[km] | $\Delta time$<br>[h] |
|------------|---------------|--------------|-------------------|------------------|--------------|------|--------|-------------------|---------------------|-----------------------|----------------------|
| 2025-02-16 | Day           | Semi-tr. Ci  | 8500              | 11500            | 0.91         | 0.92 | 0.95   | 35                | 82                  | 134                   | 2.2                  |
| 2025-03-04 | Day           | Semi-tr. As  | 5000              | 7600             | 1.25         | 0.96 | 0.96   | 41                | 71                  | 171                   | 0.6                  |
| 2025-06-11 | Night         | Thin (Ci+Sc) | 9500              | 14000            | 0.59         | 0.79 | 0.97   | 52                | 94                  | 241                   | 0.8                  |
| 2025-07-04 | Night         | Subvis. Ci   | 9000              | 12000            | 0.21         | 0.33 | 0.93   | 67                | 95                  | 52                    | 0.5                  |

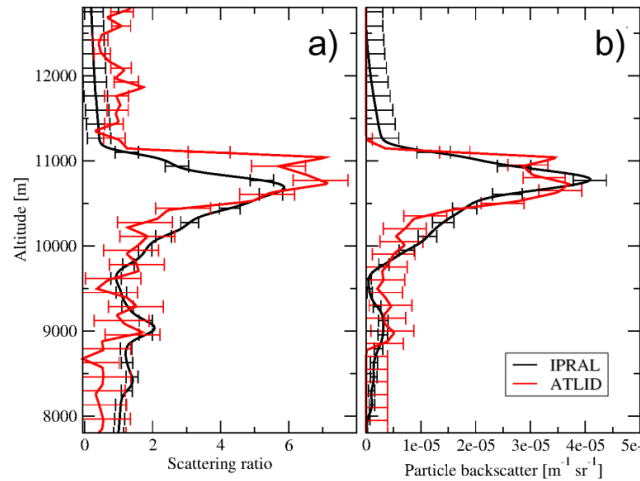


Figure 8: ATLID/IPRAL collocation for 2025-02-16, (the best ATLID profile is 134 km from SIRTa): (a) Scattering ratio in native (zenith) geometry for IPRAL (black) and native (nadir) geometry for ATLID (red); (b) particle backscatter  $\beta_{part}(z)$  for the same profiles. Error bars show the empirically estimated  $1\sigma$  random uncertainty (Sect. 4.1).

We start the discussion of the  $SR1(z)$  and  $\beta_{part}(z)$  profile agreement and disagreement across the spaceborne and ground-based lidars with the case shown in Fig. 8 (daytime ATLID overpass on 2025-02-16). We observe a semi-transparent cirrus layer in the absence of clouds below and above it. The total column opacity  $\tau_{col} \approx 0.9$ , and both instruments can observe the whole atmospheric column without attenuation of the backscattered signal. Both the  $SR1(z)$  and  $\beta_{part}(z)$  panels show that the cloud layer is detected between  $\sim 9.5$  km and  $\sim 11.3$  km by both lidars. But the agreement differs between the two variables. For  $SR1(z)$ , the peak estimated from the space-borne observation appears shifted upwards, whereas the lower shoulder estimated from the zenith observation starts earlier. This is expected and is explained by attenuation effects. On the contrary,  $\beta_{part}(z)$  profiles better agree in the position of the peak and in the shape of its shoulder, in full agreement with the estimates made above. With high CS score values (82), prominence value of 0.92, and reasonably extended structure, we would call this comparison a “recommended reference case” for similar collocated studies. The two  $\beta_{part}(z)$  curves agree within their  $1\sigma$  uncertainties over the full peak and both shoulders, which is consistent with the high CS score and a moderate spatiotemporal overlap of this case.

The next case (Fig. 9) corresponds to a single mid-level layer, which we attribute to semi-transparent altostratus cloud with a total opacity of  $\approx 1.25$ . This layer is thicker than that in the previous example, and it clearly shows on the  $SR1(z)$  panel: the native SR peaks differ by more than a factor of two between instruments, and the peak positions are not the same. At the same time, the retrieved  $\beta_{part}(z)$  profiles agree much better: the absolute values differ by 15–20%, and the peak position differ by  $\approx 200$  m compared to  $\approx 320$  m in Fig. 9a. Above  $\approx 6.5$  km, the  $\beta_{part}(z)$  shapes start to diverge, with the ground-based profile falling off less rapidly than the spaceborne one, that might be explained both by a physical distribution of the cloud particles and by the retrieval issues in optically thickening medium. The case retains a high shape correlation ( $\rho = 0.96$ )



and the highest prominence in the set ( $Pr = 0.96$ ), but its CS value is lower than in the previous case (71 vs 82), and this is consistent with the observed behavior of the  $\beta_{part}(z)$  curves.

The two  $\beta_{part}(z)$  peaks in Fig. 9b do not agree within their  $1\sigma$  uncertainties, whereas the profiles converge below the peak. The  $\approx 200\text{m}$  offset noted above between the two peak altitudes is an indicator, rather than the cause, of this disagreement: it points to an imperfect air-mass match, with the cloud most probably evolving somewhat between the ATLID overpass and the IPRAL measurement, displacing the peak altitude by the time the two instruments sampled it. With respect to future validation studies, this kind of clouds is acceptable for L2 comparisons, but is completely unacceptable for L1 cases unless the  $AMB(\lambda, z)$  profile is available for both instruments and  $SR2(z)$  can be built (Eq. (10)).

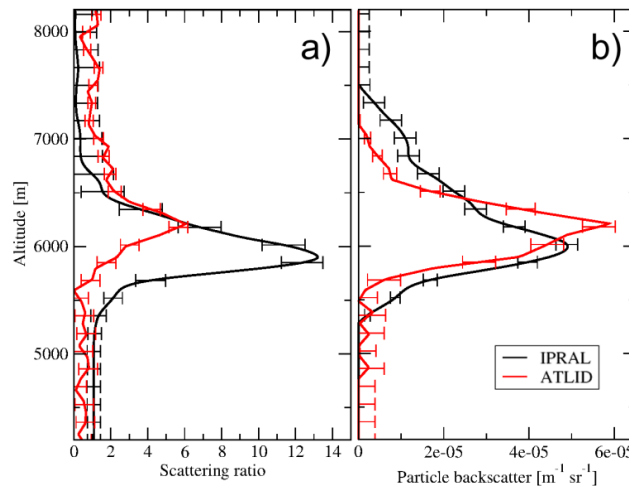


**Figure 9: Same as Fig. 8, but for 2025-03-04, (the best ATLID profile is 171 km from SIRTa)**

The third case shown in Fig. 10 is the most interesting one in the series because it covers several topics, from spatial variability to physical observability. It is a three-layer scene, which is composed of two thin cirrus clouds and one thin stratocumulus cloud with a total column opacity of about 0.59. As in two previous cases, this makes the scene observable for both lidars. The agreement between the profiles depend not only on the type of the profile (scattering ratio or backscatter), but also on the layer. In Fig. 10a, the  $(SR1-1)$  value for the uppermost peak at 13 km differs by  $\approx 40\%$ , whereas for the second peak at 10 km the differ is by a factor of 6, and the lowermost layer is invisible for ATLID in this representation. In Fig. 10b, the first peak differs by 20%, the second peak differs by the factor of 2, and the lowermost layer shows the comparable cloud characteristics. Considering the large distance between the best collocated profiles, we should note that most probably the explanation of the similarities and differences observed on this plot lies in different characteristics of the layers: the uppermost was the most stable both in space and time, and it did not change much while travelling from the ATLID's track to IPRAL's location whereas the second one partially dissipated. An indirect proof of the latter statement is the absence of an intermediate peak at 11.4 km for ATLID. Overall, this case shows a very close scene with three distinct features (and, therefore, a big DOF value), but it is not 100% the same. Within the two thin cirrus layers, the retrieved curves agree with each other within their  $1\sigma$  uncertainties; the greyed-out layer is excluded from this comparison for the reasons given above rather than for exceeding the uncertainty budget. With respect to future validations, the cases of this kind are recommended to be treated with care, and if the vertical structure of some features is not reproduced, then these features should be excluded from the comparison.

Finally, in Fig. 11 we show the case, that scored the worst on  $\rho$  and  $Pr$  diagnostics (0.93 and 0.33, respectively), but which is still interesting and promising for validations. A subvisible cirrus cloud with  $\tau_{col} \approx 0.2$  is the thinnest and closest-overpass case in the set. Native ATLID SR profile is dominated by spurious, noise-like fluctuations, especially below 11.5 km (Fig. 11a), and the peak is barely distinguishable from noise level. However, the ATLID's retrieval algorithms manage to suppress the noise effects and the retrieved cloud layer peaking at  $\approx 11.2$  km is reproduced with close agreement both in magnitude and in the altitude. It is the cleanest single-layer match of the set that is probably related to the spatiotemporal closeness of the compared profiles. The retrieved  $\beta_{part}(z)$  curves agree within their  $1\sigma$  uncertainties over the whole layer, consistent with this being the cleanest single-layer match of the set. Validation-wise, the cases of this type are good candi-



dates for the future comparisons. Across all four cases, the retrieved particle backscatter profiles comparisons show systematically closer ground-based/spaceborne lidar agreement than the native-geometry scattering ratio comparisons, regardless of cloud type, optical depth, or overpass distance. In all four cases, the difference between the  $SR1(z)$  and  $\beta_{part}(z)$  comparisons has the same origin.  $SR1(z)$  carries the attenuation history of each viewing geometry, which cannot be separated from the layer signal, whereas  $\beta_{part}(z)$  is a product of the retrieval algorithm, which removes everything not related to the properties of the local layer. In addition to this common origin, the four cases isolate distinct sources of divergence: differential attenuation along the zenith versus nadir optical path degrading  $SR1$  agreement (Fig. 8); the same attenuation effect aggravated by larger optical thickness leading to much stronger amplitude and shape difference (Fig. 9); layer-dependent agreement in a multi-layer scene, where large spatial separation between collocated profiles correlates with the evolution of certain layers (Fig. 10); and the instrument noise noticeable for a subvisible cloud and degrading  $SR1$  comparison (Fig. 11).

At the same time, the closest and averaged pairs for these same four cases (Appendix B) do not provide satisfactory agreement, except, perhaps, the closest case of the 2025-07-04 that is the smallest-separation case in the set. For 2025-03-04, the closest pair falls on an ATLID record with no corresponding IPRAL structure, and the averaged pair only partially recovers the shape. For 2025-02-16 and 2025-06-11, the averaging procedure mixes in structure unrelated to the matched layer, degrading the profile agreement below what the closest pair alone achieves. Summarizing, these cases support this paper's two central methodological arguments: (1) optically thinner cases are preferable for  $SR1(z)$  comparison and (2)  $\beta_{part}(z)$  comparison is a more robust diagnostic of space-borne-to-ground-based lidar agreement than  $SR1(z)$ . Another important observation is that the comparison results are in line with the CS scores obtained in the simulation chain. Across the four cases, the retrieved  $\beta_{part}(z)$  profiles agree within their propagated  $1\sigma$  uncertainties wherever the underlying layer is optically thin to semi-transparent and has not evolved between overpass and measurement; the one exception is Case 2's peak, where an  $\approx 200\text{m}$  altitude offset indicates an imperfect air-mass match rather than underestimated uncertainty or retrieval error. These findings allow us to formulate a set of recommendations for space-borne-to-ground-based lidar validations outlined in the next section.

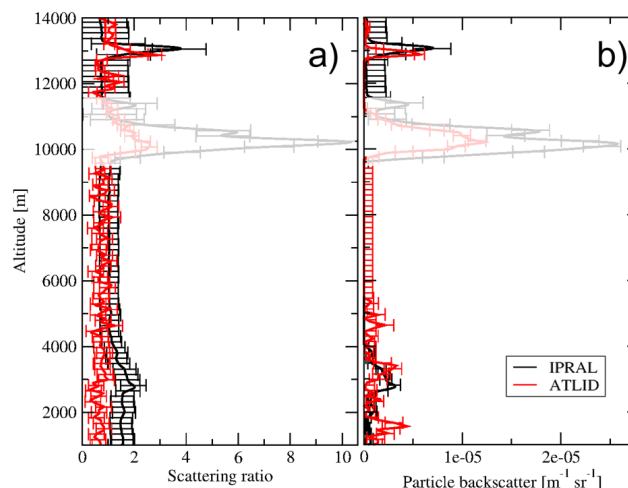


**Figure 10:** Same as Fig. 8, but for 2025-06-11, (the best ATLID profile is 241 km from SIRTa). Despite the qualitative similarity, the layer between 9500m and 11500m is greyed out as not comparable due to different dynamics for 3 layers.

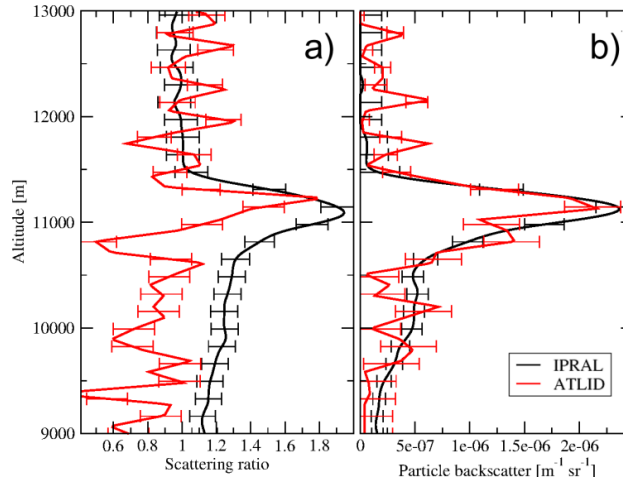


**Figure 11: Same as Fig. 8, but for 2025-07-04, (the best ATLID profile is 52 km from SIRTa)**

**Table 4: Optical depths  $\tau_{IPRAL}$  and  $\tau_{ATLID}$  of the cloud layers of the four cases of Table 3, with their uncertainties defined from the empirical noise estimate according to Sect. 4.1. Note the difference of column (Table 3) and layer values for 2025-07-04.**

| Date       | Layer       | Base-top [m] | $\tau_{IPRAL}$ | $\Delta\tau_{IPRAL}$ | $\tau_{ATLID}$ | $\Delta\tau_{ATLID}$ |
|------------|-------------|--------------|----------------|----------------------|----------------|----------------------|
| 2025-02-16 | Ci          | 8000–12400   | 9.3e-1         | 1.5e-2               | 8.7e-1         | 2.7e-1               |
| 2025-03-04 | As          | 5250–7550    | 1.3e0          | 1.2e-2               | 1.0e0          | 2.7e-1               |
| 2025-06-11 | Sc (low)    | 1000–4000    | 9.5e-2         | 2.5e-3               | 1.0e-1         | 7.1e-2               |
| 2025-06-11 | Ci (middle) | 9500–11400   | 5.8e-1         | 8.4e-3               | 2.5e-1         | 5.5e-2               |
| 2025-06-11 | Ci (upper)  | 12600–13500  | 6.1e-2         | 6.4e-3               | 3.5e-2         | 1.5e-2               |
| 2025-07-04 | Ci          | 9500–11600   | 4.9e-2         | 6.6e-4               | 4.3e-2         | 1.0e-2               |

In Table 4, we quantify the bulk agreement between cloud layers in terms of layer opacity obtained from  $\beta_{part}(z)$  of both lidars converted into an extinction coefficient with the same constant lidar ratio of 30 sr and integrated over the layer limits given in Table 4; the resulting optical depths are therefore a measure of the layer-integrated backscatter, and their ratio does not depend on the assumed lidar ratio. For the two semi-transparent single-layer cases, the ATLID optical depth is lower than the IPRAL one by 6% (2025-02-16) and 19% (2025-03-04); both differences are smaller than the  $1\sigma$  uncertainty of ATLID, which we consider an upper estimate (Sect. 4.1). The larger difference for 2025-03-04 is consistent with the 15–20% difference in the amplitude of the  $\beta_{part}(z)$  peaks discussed above. For the subvisible cirrus of 2025-07-04, the two values (0.049 and 0.043) are also within the uncertainty limits. In the three-layer scene of 2025-06-11, the optical depths of the low stratocumulus are similar (0.09 and 0.10) and those of the upper cirrus differ by  $1.5\sigma$  (0.06 and 0.04), whereas for the middle cirrus ATLID gives less than half of the IPRAL value (0.25 against 0.59). This is in line with our decision to exclude that layer from the comparison, and with its partial dissipation between the ATLID overpass and the IPRAL measurement. Overall, the agreement between the layer optical depths follows the same pattern as the  $\beta_{part}(z)$  profiles.



## 5. Recommendations for space-vs-ground lidar comparisons

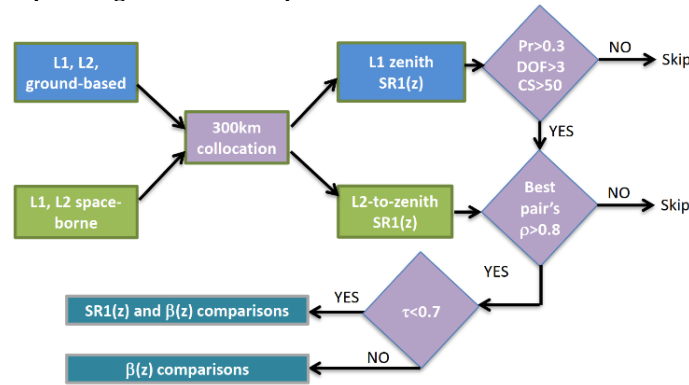


**Figure 12: Recommended scheme and thresholds for space-borne to ground-based comparison. L2-to-zenith means the conversion described in Sect. 4.3. The process is repeated for all pairs of ground-based and space-borne profiles passing the preliminary 300km collocation.**

The scheme in Fig. 12 formalizes the criteria established throughout this work: spatiotemporal collocation quality alone is a useful criterion, but it may lead to non-existent matches if it is too strict or to non-comparable profiles if it is too loose. On the other hand, observing the same cloud scene in the vicinity of the ground-based lidar can be ensured through a set of rules, which exclude the vertical profiles that do not belong to the same cloud. For optically thin and subvisible clouds,  $SR1(z)$  remains a valid, low-latency diagnostic; beyond a threshold optical depth, zenith-to-nadir attenuation differences render  $SR1(z)$  unreliable regardless of collocation quality, and  $\beta_{part}(z)$  must be used instead of  $SR1(z)$  for apples-to-apples comparison. We recommend this scheme as a practical starting point for spaceborne-to-ground-based lidar intercomparisons more broadly; while the specific thresholds defined here may require recalibration for other instrument pairs, the underlying logic of screening on correlation and profile properties first, then selecting the comparison variable by optical depth, should remain valid.

## 6. Conclusions

So, do we see the atmospheric column the same way from space and from the ground, and can we really compare the two observations apples to apples? Yes, but only under specific conditions. It is needless to say that both instruments should be altitude-calibrated and free of experimental artefacts. Still, even for a pair of instruments satisfying these requirements a simple collocation might not give the expected results. First of all, we should distinguish the perfect collocation when both instruments probe exactly the same atmospheric column at the same time from the more frequent scenario when the instruments measure close atmospheric scenes and the distribution of the scattering particle for the closest point is not the same. Second, the opposing viewing geometry combined with optical thickness of the atmospheric column imposes additional conditions: the atmospheric column should be transparent up to the compared cloud layer for ground-based lidar and down to this layer for space-borne lidar. Third, if the collocation is not perfect, we need to screen the profiles instead of matching by proximity.

From the simulation chain we developed and applied to the 15 synthetic cloud scenes, we found that comparability between zenith and nadir observations is primarily defined by the opposing geometries and to a secondary degree by the daytime or instrumental noise or sensitivity. The difference in viewing direction increases with optical depth: for thin and subvisible layers, two-way attenuation is negligible in either direction, and  $SR1(z)$ ,  $SR2(z)$ , and  $\beta_{part}(z)$  agree between geometries; for semi-transparent and thick layers, the attenuation asymmetry dominates and the  $SR1(z)$  profiles diverge even under perfect overlap. The  $SR2(z)$  and especially  $\beta_{part}(z)$  remain valid diagnostics under these conditions. For thicker clouds, even the  $\beta_{part}(z)$  fails to provide a good match. This is consistent with common logic:  $SR2(z)$  uses measured molecular backscatter and tracks the same attenuation path as the total signal, but is HSRL-only and becomes noise-dominated beyond that regime.  $SR1(z)$  doesn't suffer from highly attenuated value in the denominator, but the difference in geometries make it valid only below a threshold optical depth. At the same time, the retrieval procedures yielding  $\beta_{part}(z)$  remove the geometry-driven asymmetry by construction. The retrieval accounts for attenuation layer by layer, leaving only retrieval-specific limitations (reference altitude, obscured reference layer, assumed lidar ratio) as the residual source of disagreement.



The comparability score values we introduced in this work confirm this ordering:  $SR_1(z)$  is comparable only for the thinnest scenes,  $SR_2(z)$  extends further, and  $\beta_{part}(z)$  comparability stays high over the broadest domain except for daytime cloud observations at 532 nm.

To identify a common air mass under realistic spatiotemporal separation, we proposed a correlation-maximization search and quantified its false discovery rate against both a theoretical and an empirical null distribution. The empirical test, using real ATLID/IPRAL mismatched-date collocations, gives a false discovery rate of about 20 % at a correlation threshold of 0.8 and prominence above 0.4. Since this null reproduces the exact selection procedure, the rate includes the inflation of the apparent correlation by the search itself; the true collocations pass the same criteria about five times more often than mismatched-date ones.

The four ATLID/IPRAL case studies (Sect. 4.4) match the theoretical predictions:  $\beta_{part}(z)$  agreement is systematically closer than native-geometry  $SR_1(z)$  agreement, and the residual differences are primarily attributable to spatial decorrelation, cloud evolution, and noise rather than instrumental or calibration artefacts. This is itself a validation result, it indicates that IPRAL's L2 particle backscatter retrieval is consistent with ATLID's.

For those who plan similar spaceborne-to-ground-based validation campaigns, we propose a practical starting point (Fig. 12): screen candidate collocations on correlation and profile structure before looking at the absolute values, then select  $SR_1(z)$  or  $\beta_{part}(z)$  as the comparison variable according to the optical depth of the scene (use Fig. 6 as a rough estimate and do not use the scene if its CS is below 50). The specific thresholds given in Fig. 12 were calibrated for the ATLID/IPRAL pair at 355 nm and should be double-checked before applying to another pair of lidars, especially if the parameters of the lidars are different. But, the underlying logic should remain the same. Applying these conditions reduces the number of valid comparisons, but it makes the validation sound.

## Appendix A. Theoretical estimate of the false-correlation probability

We restrict the analysis to semi-transparent scenes, for which the compared  $SR(z)$  or  $\beta_{part}(z)$  profiles can be decomposed into three additive, mutually independent contributions ( $S(z)$  below generalizes the approach for L1 and L2 lidar data):

$$S(z) = S_{bck}(z) + S_{cld}(z) + S_{noise}(z) \quad (A1)$$

where  $S_{bck}(z)$  is a slow changing, well-mixed background contribution (molecular scattering plus a regional aerosol layer),  $S_{cld}(z)$  is the localized structure of interest, and  $S_{noise}(z)$  includes all noise-related effects including the instrumental noise and the retrieval uncertainty and represents the parts of the signals, which are completely uncorrelated between the two instruments. Under these assumptions, the covariance between two compared profiles  $S_1(z)$  and  $S_2(z)$  reduces to

$$cov\{S_1, S_2\} = \rho_{bck} \times \sigma_{bck}^2 + \rho_{cld} \times \sigma_{cld}^2 \quad (A2)$$

where  $\rho_{bck}$  and  $\rho_{cld}$  are the correlation coefficients of the background and cloud components between the two profiles, and  $\sigma_{bck}^2$  and  $\sigma_{cld}^2$  are their standard deviations (assumed equal in both profiles for simplicity). Normalizing by the total variance, and neglecting the noise term (assuming that it is a true noise that is not correlated by definition), gives the following estimate of the correlation coefficient of two collocated profiles:

$$\rho\{S_1, S_2\} = \delta \times \rho_{bck} + (1 - \delta) \times \rho_{cld}; \quad \delta = \frac{\sigma_{bck}^2}{\sigma_{bck}^2 + \sigma_{cld}^2} \quad (A3)$$

For two profiles that do not sample the same cloud structure, the cloud term is essentially decorrelated ( $\rho_{cld} \approx 0.0 \pm 0.2$ ), while the background term retains a moderate correlation because it reflects a broad, well-mixed regional field ( $\rho_{bck} \approx 0.5$ ). For a semi-transparent layer, where the cloud and background contributions have comparable amplitudes ( $\approx 0.5$ ), Eq. (A3) gives a "null" correlation that represents the correlation between two unrelated profiles:

$$\rho_{null} = 0.5 \times 0.5 + 0.5 \times (0 \pm 0.2) = 0.25 \pm 0.10 \quad (A4)$$

We now assess whether the probability of a "false correlation" that corresponds to the case when two cloud profiles are selected using the high correlation coefficient criterion alone (we take only the cases with  $\rho\{S_1(z), S_2(z)\} > 0.9$ ) whereas they do not represent the same cloud. For this experiment, we take a semi-transparent cloud occupying a 1500m thick layer that is sampled at ATLID's vertical step of 100m. Being considered independently, these 15 points would have given 15 "degrees of freedom" (DOF) for our comparison, but in reality, we should account for the fact that adjacent altitude bins are not independent: they co-vary over the correlation length intrinsic to the cloud structure. For a cloud considered in the example, the effective DOF is about 5 (1-2 peaks, 2-4 slopes). The approximate standard error of the sample correlation coefficient is then



$$SE(\rho) \approx \frac{1-\rho_{null}^2}{\sqrt{DOF-2}} \approx 0.54 \quad (A5)$$

and the standardized deviation of the observed correlation from the null value,

$$Z = \frac{\rho\{S_1(z), S_2(z)\} - \rho_{null}}{SE(\rho)} \approx 1.2 \quad (A6)$$

corresponds to a one-sided probability  $P(Z > 1.2) \approx 12\%$ . In other words, roughly one pair in nine could still produce a spuriously high correlation just by chance. In other eight cases, the matched pair of profiles represents the same cloud which moved horizontally between the measurements performed by ATLID and IPRAL. This estimate relies on simplified assumptions of DOF and  $\rho_{back}$ , but in fact we used rather conservative estimates, and for a scene without a pronounced background and/or with multi-peak cloud structure the one-sided probability P will be even lower.

## Appendix B. Closest and averaged collocation pairs

In this section, we show the supplementary figures to the correlation-maximized ("best") collocation pairs shown in Figs. 8–11 (Sect. 4.4). Figures B1–B4 contain both the spatially/temporally averaged pairs (panels (a) and (b)) and the closest-in-space/time pairs (panels (c) and (d)) for  $SR1(z)$  and  $\beta_{part}(z)$ . Each figure corresponds to one of the four cases discussed in Sect. 4.4, in the same order. These panels are provided to illustrate that the simpler pairing methods does not always provide comparable profiles.

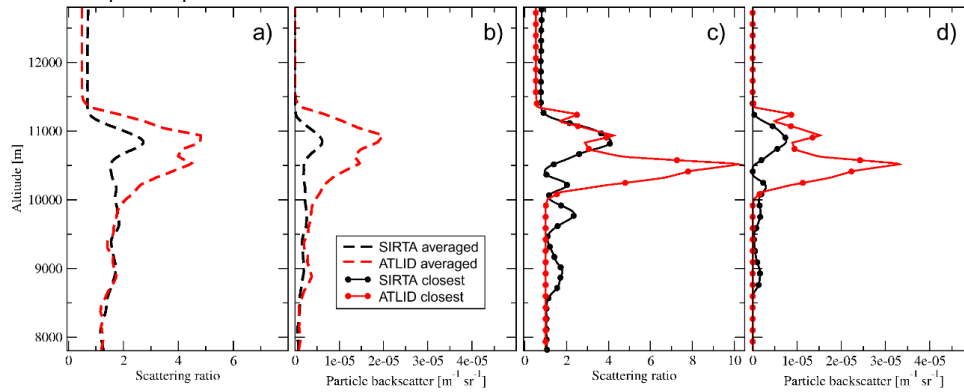


**Figure B1.** Same collocation date as Fig. 8 (2025-02-16), shown for the: (a)  $SR1(z)$  for the averaged pair; (b)  $\beta_{part}(z)$  for the averaged pair; (c)  $SR1(z)$  for the closest pair; (d)  $\beta_{part}(z)$  for the closest pair.

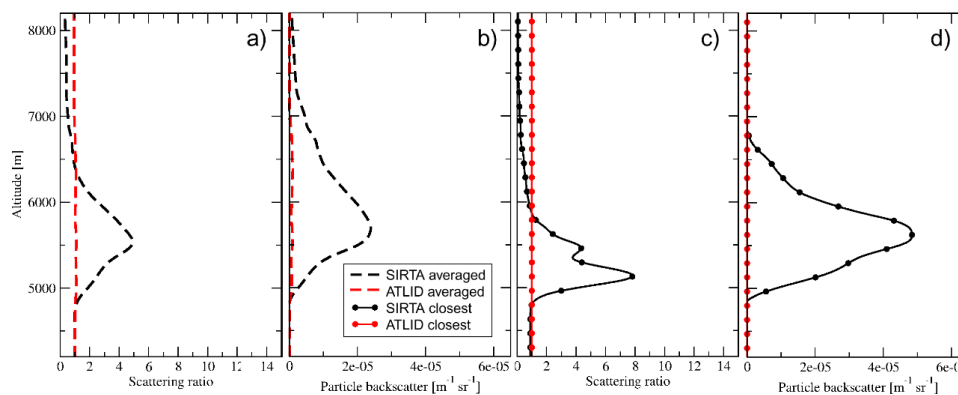
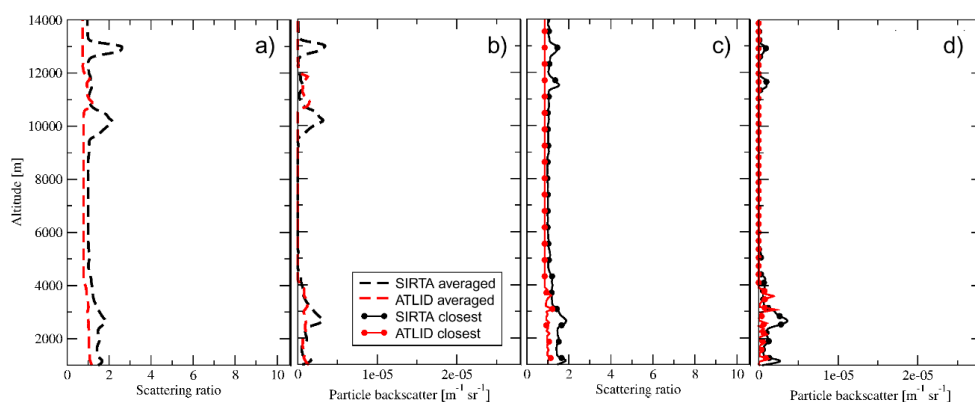
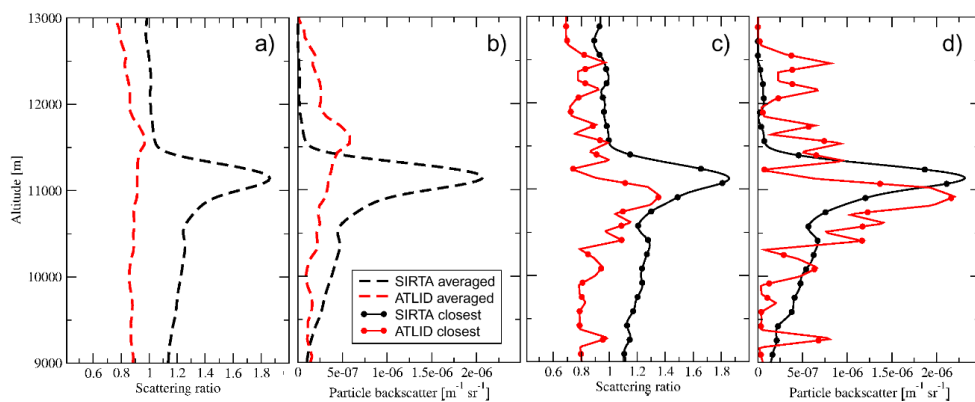


**Figure B2.** Same as Fig. B1, but for 2025-03-04 (compare with Fig. 9).



824  
825

826 **Figure B3. Same as Fig. B1, but for 2025-06-11 (compare with Fig. 10).**



827  
828

829 **Figure B4. Same as Fig. B1, but for 2025-07-04 (compare with Fig. 11).**



## 830 **Code and data availability**

831 The processing chain used to calibrate the IPRAL 355 nm signals and to retrieve the particle backscatter coefficient (version  
 832 1.0) is available at <https://doi.org/10.5281/zenodo.23387966> (Haddouche and Feofilov, 2026b) under the MIT license, and is  
 833 documented in the accompanying technical note (Haddouche and Feofilov, 2026a). The simulation framework is described  
 834 in Sects. 2 and 4.2 and in Feofilov et al. (2023); its code is not publicly released. The lidar simulations use COSP2 (Swales  
 835 et al., 2018), an open-source package available at <https://github.com/CFMIP/COSPv2.0> under a BSD license.

836 The ATLID Level 1 and Level 2 data (Baseline BA, NOM\_1B and EBD\_2A, respectively) were provided by ESA. The  
 837 IPRAL profiles, the collocated ATLID and IPRAL profiles, and the data underlying the figures are available from the corre-  
 838 sponding author upon request. The sources of the radiosonde and model inputs are given in Sects. 2.1 and 4.1.

## 839 **Author contributions**

840 AGF: conceptualization, methodology, software, formal analysis, visualization, writing (original draft). AH: data curation,  
 841 investigation, methodology (IPRAL calibration), validation, writing (original draft). HC: conceptualization, writing (review  
 842 and editing). SK: resources, writing (review and editing). FL, PD, MAD, MH, JCD, CD, CP: resources (operation and  
 843 maintenance of IPRAL), writing (review and editing).

## 844 **Competing interests**

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

## 846 **Acknowledgements**

847 We acknowledge ACTRIS CARS – Centre for Aerosol Remote Sensing for its support in setting up and optimizing the  
 848 IPRAL instrument. We thank ESA and the EarthCARE mission team for providing ATLID data. Claude (Anthropic) was  
 849 used for language editing, consistency checks, support in writing/debugging analysis/visualization code, and as a sounding  
 850 board for methodological choices. The authors reviewed all AI-assisted content and take full responsibility for the  
 851 manuscript.

## 852 **Financial support**

853 The apprenticeship of A. Haddouche was funded by the Académie Spatiale d'Île-de-France, a "Compétences et Métiers  
 854 d'Avenir" project of France 2030 managed by the French National Research Agency (ANR) under reference ANR-23-  
 855 CMAS-0041.

## 856 **References**

- 857 Abshire, J. B., Sun, X., Riris, H., Sirota, J. M., McGarry, J. F., Palm, S., Yi, D., and Liiva, P.: Geosci-  
 858 ence Laser Altimeter System (GLAS) on the ICESat Mission: On-orbit measurement performance,  
 859 Geophys. Res. Lett., 32, L21S02, <https://doi.org/10.1029/2005GL024028>, 2005.
- 860 Alkasem, A., Szczap, F., Cornet, C., Shcherbakov, V., Gour, Y., Jourdan, O., Labonnote, L. C., and  
 861 Mioche, G.: Effects of cirrus heterogeneity on lidar CALIOP/CALIPSO data, J. Quant. Spectrosc. Ra.,  
 862 202, 38–49, <https://doi.org/10.1016/j.jqsrt.2017.07.005>, 2017.
- 863 Alraddawi, D., Keckhut, P., Mandija, F., Sarkissian, A., Pietras, C., Dupont, J.-C., Farah, A.,  
 864 Hauchecorne, A., and Porteneuve, J.: Calibration of Upper Air Water Vapour Profiles Using the



- 865 IPRAL Raman Lidar and ERA5 Model Results and Comparison to GRUAN Radiosonde Observations,  
866 Atmosphere, 16, 351, <https://doi.org/10.3390/atmos16030351>, 2025.
- 867 Amiridis, V., Marinou, E., Hostetler, C., et al.: Best Practice Protocol for the Validation of Aerosol,  
868 Cloud, and Precipitation Profiles (ACPPV), Zenodo, <https://doi.org/10.5281/zenodo.15025627>, 2025.
- 869 Ansmann, A. and Müller, D.: Lidar and Atmospheric Aerosol Particles, in: Lidar: Range-Resolved Op-  
870 tical Remote Sensing of the Atmosphere, edited by: Weitkamp, C., vol. 102, Springer, New York, 105–  
871 142, 2005.
- 872 Ansmann, A., Wandinger, U., Riebesell, M., Weitkamp, C., and Michaelis, W.: Independent measure-  
873 ment of extinction and backscatter profiles in cirrus clouds by using a combined Raman elastic-  
874 backscatter lidar, Appl. Optics, 31, 7113–7131, <https://doi.org/10.1364/AO.31.007113>, 1992.
- 875 Atlas, R. and Bretherton, C. S.: Aircraft observations of gravity wave activity and turbulence in the  
876 tropical tropopause layer: prevalence, influence on cirrus clouds, and comparison with global storm-  
877 resolving models, Atmos. Chem. Phys., 23, 4009–4030, <https://doi.org/10.5194/acp-23-4009-2023>,  
878 2023.
- 879 Baars, H., Haarig, M., König, L., Floutsi, A. A., Basharova, E., Hofer, J., Gebauer, H., Engelmann, R.,  
880 Althausen, D., Skupin, A., Gast, B., Fritzsche, F., Ohneiser, K., Jimenez, C., Gaudek, T., Radenz, M.,  
881 Buholdt, H., Heese, B., Klamt, A., Seifert, P., Donovan, D. P., van Zadelhoff, G.-J., Abdullaev, S. F.,  
882 Khalifaeva, S. K., Nozirov, D. F., and Wandinger, U.: Validation of EarthCARE/ATLID aerosol profil-  
883 ing products with ground-based PollyNET lidars – case studies, Atmos. Meas. Tech., 19, 3831–3864,  
884 <https://doi.org/10.5194/amt-19-3831-2026>, 2026.
- 885 Bucholtz, A.: Rayleigh-scattering calculations for the terrestrial atmosphere, Appl. Optics, 34, 2765–  
886 2773, <https://doi.org/10.1364/AO.34.002765>, 1995.
- 887 Cesana, G., Del Genio, A. D., and Chepfer, H.: The Cumulus And Stratocumulus CloudSat-CALIPSO  
888 Dataset (CASCCAD), Earth Syst. Sci. Data, 11, 1745–1764, [https://doi.org/10.5194/essd-11-1745-](https://doi.org/10.5194/essd-11-1745-2019)  
889 [2019](https://doi.org/10.5194/essd-11-1745-2019), 2019.
- 890 Chepfer, H., Bony, S., Winker, D. M., Chiriaco, M., Dufresne, J.-L., and Sèze, G.: Use of CALIPSO  
891 lidar observations to evaluate the cloudiness simulated by a climate model, Geophys. Res. Lett., 35,  
892 L15704, <https://doi.org/10.1029/2008GL034207>, 2008.
- 893 Chepfer, H., Bony, S., Winker, D., Cesana, G., Dufresne, J.-L., Minnis, P., Stubenrauch, C. J., and  
894 Zeng, S.: The GCM-Oriented CALIPSO Cloud Product (CALIPSO-GOCCP), J. Geophys. Res., 115,  
895 D00H16, <https://doi.org/10.1029/2009JD012251>, 2010.
- 896 Chepfer, H., Cesana, G., Winker, D., Getzewich, B., Vaughan, M., and Liu, Z.: Comparison of two  
897 different cloud climatologies derived from CALIOP-attenuated backscattered measurements (Level 1):  
898 The CALIPSO-ST and the CALIPSO-GOCCP, J. Atmos. Ocean. Tech., 30, 725–744,  
899 <https://doi.org/10.1175/JTECH-D-12-00057.1>, 2013.
- 900 Collis, R. T. H. and Russell, P. B.: Lidar measurement of particles and gases by elastic backscattering  
901 and differential absorption, in: Laser Monitoring of the Atmosphere, edited by: Hinkley, E. D., Topics



- 902 in Applied Physics, vol. 14, Springer, Berlin, Heidelberg, 71–151, [https://doi.org/10.1007/3-540-](https://doi.org/10.1007/3-540-07743-X_18)  
903 [07743-X\\_18](https://doi.org/10.1007/3-540-07743-X_18), 1976.
- 904 Comstock, J. M., Ackerman, T. P., and Mace, G. G.: Ground-based lidar and radar remote sensing of  
905 tropical cirrus clouds at Nauru Island: Cloud statistics and radiative impacts, J. Geophys. Res., 107,  
906 4714, <https://doi.org/10.1029/2002JD002203>, 2002.
- 907 Covert, J. A., Mechem, D. B., and Zhang, Z.: Subgrid-scale horizontal and vertical variation of cloud  
908 water in stratocumulus clouds: a case study based on LES and comparisons with in situ observations,  
909 Atmos. Chem. Phys., 22, 1159–1174, <https://doi.org/10.5194/acp-22-1159-2022>, 2022.
- 910 Donovan, D. P.: The Expected Impact of Multiple Scattering on ATLID Signals, EPJ Web Conf., 119,  
911 01006, <https://doi.org/10.1051/epjconf/201611901006>, 2016.
- 912 Donovan, D. P., van Zadelhoff, G.-J., and Wang, P.: The EarthCARE lidar cloud and aerosol profile  
913 processor (A-PRO): the A-AER, A-EBD, A-TC, and A-ICE products, Atmos. Meas. Tech., 17, 5301–  
914 5340, <https://doi.org/10.5194/amt-17-5301-2024>, 2024.
- 915 Fernald, F. G.: Analysis of atmospheric lidar observations: Some comments, Appl. Optics, 23, 652–  
916 653, <https://doi.org/10.1364/AO.23.000652>, 1984.
- 917 Feofilov, A. G., Kutepov, A. A., She, C.-Y., Smith, A. K., Pesnell, W. D., and Goldberg, R. A.:  
918 CO<sub>2</sub>(v<sub>2</sub>)-O quenching rate coefficient derived from coincidental SABER/TIMED and Fort Collins li-  
919 dar observations of the mesosphere and lower thermosphere, Atmos. Chem. Phys., 12, 9013–9023,  
920 <https://doi.org/10.5194/acp-12-9013-2012>, 2012.
- 921 Feofilov, A. G., Stubenrauch, C. J., and Delanoë, J.: Ice water content vertical profiles of high-level  
922 clouds: classification and impact on radiative fluxes, Atmos. Chem. Phys., 15, 12327–12344,  
923 <https://doi.org/10.5194/acp-15-12327-2015>, 2015.
- 924 Feofilov, A. G., Chepfer, H., Noël, V., Guzman, R., Gindre, C., Ma, P.-L., and Chiriaco, M.: Compari-  
925 son of scattering ratio profiles retrieved from ALADIN/Aeolus and CALIOP/CALIPSO observations  
926 and preliminary estimates of cloud fraction profiles, Atmos. Meas. Tech., 15, 1055–1074,  
927 <https://doi.org/10.5194/amt-15-1055-2022>, 2022.
- 928 Feofilov, A. G., Chepfer, H., Noël, V., and Szczap, F.: Incorporating EarthCARE observations into a  
929 multi-lidar cloud climate record: the ATLID (Atmospheric Lidar) cloud climate product, Atmos. Meas.  
930 Tech., 16, 3363–3390, <https://doi.org/10.5194/amt-16-3363-2023>, 2023.
- 931 Feofilov, A., Chepfer, H., Noël, V., and Hajiaghazadeh-Roodsari, M.: Towards Establishing a Long-  
932 Term Cloud Record from Space-Borne Lidar Observations, in: Space-based Lidar Remote Sensing  
933 Techniques and Emerging Technologies, LIDAR 2023, edited by: Singh, U. N., Tzeremes, G., Refaat,  
934 T. F., and Ribes Pleguezuelo, P., Springer Aerospace Technology, Springer, Cham,  
935 [https://doi.org/10.1007/978-3-031-53618-2\\_6](https://doi.org/10.1007/978-3-031-53618-2_6), 2024.
- 936 Freudenthaler, V.: About the effects of polarising optics on lidar signals and the  $\Delta 90^\circ$  calibration, At-  
937 mos. Meas. Tech., 9, 4181–4255, <https://doi.org/10.5194/amt-9-4181-2016>, 2016.



- 938 Freudenthaler, V., Esselborn, M., Wiegner, M., Heese, B., Tesche, M., Ansmann, A., Müller, D., Al-  
939 thausen, D., Wirth, M., Fix, A., Ehret, G., Knippertz, P., Toledano, C., Gasteiger, J., Garhammer, M.,  
940 and Seefeldner, M.: Depolarization ratio profiling at several wavelengths in pure Saharan dust during  
941 SAMUM 2006, *Tellus B*, 61, 165–179, <https://doi.org/10.1111/j.1600-0889.2008.00396.x>, 2009.
- 942 Garnier, A., Pelon, J., Vaughan, M. A., Winker, D. M., Trepte, C. R., and Dubuisson, P.: Lidar multiple  
943 scattering factors inferred from CALIPSO lidar and IIR retrievals of semi-transparent cirrus cloud opti-  
944 cal depths over oceans, *Atmos. Meas. Tech.*, 8, 2759–2774, <https://doi.org/10.5194/amt-8-2759-2015>,  
945 2015.
- 946 Gettelman, A., Mills, M. J., Kinnison, D. E., Garcia, R. R., Smith, A. K., Marsh, D. R., Tilmes, S., Vitt,  
947 F., Bardeen, C. G., McInerney, J., Liu, H.-L., Solomon, S. C., Polvani, L. M., Emmons, L. K., La-  
948 marque, J.-F., Richter, J. H., Glanville, A. S., Bacmeister, J. T., Phillips, A. S., Neale, R. B., Simpson,  
949 I. R., DuVivier, A. K., Hodzic, A., and Randel, W. J.: The Whole Atmosphere Community Climate  
950 Model Version 6 (WACCM6), *J. Geophys. Res.-Atmos.*, 124, 12380–12403,  
951 <https://doi.org/10.1029/2019JD030943>, 2019.
- 952 Gimmetstad, G., Forrister, H., Grigas, T., and O'Dowd, C.: Comparisons of aerosol backscatter using  
953 satellite and ground lidars: implications for calibrating and validating spaceborne lidar, *Sci. Rep.*, 7,  
954 42337, <https://doi.org/10.1038/srep42337>, 2017.
- 955 Guzman, R., Chepfer, H., Noel, V., Vaillant de Guélis, T., Kay, J. E., Raberanto, P., Cesana, G.,  
956 Vaughan, M. A., and Winker, D. M.: Direct atmosphere opacity observations from CALIPSO provide  
957 new constraints on cloud-radiation interactions, *J. Geophys. Res.-Atmos.*, 122, 1066–1085,  
958 <https://doi.org/10.1002/2016JD025946>, 2017.
- 959 Haddouche, A.: Colocalisation, comparaison et validation des données du lidar spatial AT-  
960 LID/EarthCARE avec des lidars au sol, M1 ISI apprenticeship report, Sorbonne Université / LMD,  
961 Palaiseau, <https://www.researchgate.net/publication/404065765>, 2025.
- 962 Haddouche, A. and Feofilov, A.: Calibration procedure and backscatter retrieval for IPRAL 355 nm  
963 far-range signals: technical note for the ATLID/EarthCARE zenith-to-nadir comparison, version 1.0,  
964 Laboratoire de Météorologie Dynamique (LMD/IPSL), École Polytechnique, Palaiseau, France,  
965 <https://doi.org/10.13140/RG.2.2.26904.15360>, 2026a.
- 966 Haddouche, A. and Feofilov, A.: IPRAL-Lidar-Processing: calibration and backscatter retrieval for  
967 IPRAL 355 nm far-range signals, version 1.0, Zenodo [code],  
968 <https://doi.org/10.5281/zenodo.23387966>, 2026b.
- 969 Haefelin, M., Barthès, L., Bock, O., Boitel, C., Bony, S., Bouniol, D., Chepfer, H., et al.: SIRTa, a  
970 ground-based atmospheric observatory for cloud and aerosol research, *Ann. Geophys.*,  
971 <https://doi.org/10.5194/angeo-23-253-2005>, 2005.
- 972 Hogan, R. J. and Illingworth, A. J.: Parameterizing ice cloud inhomogeneity and the overlap of inho-  
973 mogeneities using cloud radar data, *J. Atmos. Sci.*, 60, 756–767, [https://doi.org/10.1175/1520-0469\(2003\)060<0756:PICLAT>2.0.CO;2](https://doi.org/10.1175/1520-0469(2003)060<0756:PICLAT>2.0.CO;2), 2003.



- 975 Illingworth, A. J., Barker, H. W., Beljaars, A., Ceccaldi, M., Chepfer, H., et al.: The EarthCARE satel-  
976 lite: the next step forward in global measurements of clouds, aerosols, precipitation, and radiation, B.  
977 Am. Meteorol. Soc., 96, 1311–1332, <https://doi.org/10.1175/BAMS-D-12-00227.1>, 2015.
- 978 Khaykin, S., Sicard, M., Leblanc, T., Sakai, T., Balugin, N., Berthet, G., Chevrier, S., Chouza, F.,  
979 Feofilov, A., Gantois, D., Godin-Beekmann, S., Haddouche, A., Jin, Y., Morino, I., Kadygrov, N.,  
980 Lecas, T., Liley, B., Querel, R., Taha, G., and Yushkov, V.: Global transport of stratospheric aerosol  
981 produced by Ruang eruption from EarthCARE ATLID, limb-viewing satellites and ground-based lidar  
982 observations, Atmos. Chem. Phys., 26, 607–622, <https://doi.org/10.5194/acp-26-607-2026>, 2026.
- 983 Klett, J. D.: Stable analytical inversion solution for processing lidar returns, Appl. Optics, 20, 211–220,  
984 <https://doi.org/10.1364/AO.20.000211>, 1981.
- 985 Klett, J. D.: Lidar inversion with variable backscatter/extinction ratios, Appl. Optics, 24, 1638–1643,  
986 <https://doi.org/10.1364/AO.24.001638>, 1985.
- 987 Li, J., Huang, J., Stamnes, K., Wang, T., Lv, Q., and Jin, H.: A global survey of cloud overlap based on  
988 CALIPSO and CloudSat measurements, Atmos. Chem. Phys., 15, 519–536,  
989 <https://doi.org/10.5194/acp-15-519-2015>, 2015.
- 990 Lopes, F. J. S., Landulfo, E., and Vaughan, M. A.: Evaluating CALIPSO's 532 nm lidar ratio selection  
991 algorithm using AERONET sun photometers in Brazil, Atmos. Meas. Tech., 6, 3281–3299,  
992 <https://doi.org/10.5194/amt-6-3281-2013>, 2013.
- 993 Mamouri, R. E., Amiridis, V., Papayannis, A., Giannakaki, E., Tsaknakis, G., and Balis, D. S.: Valid-  
994 ation of CALIPSO space-borne-derived attenuated backscatter coefficient profiles using a ground-based  
995 lidar in Athens, Greece, Atmos. Meas. Tech., 2, 513–522, <https://doi.org/10.5194/amt-2-513-2009>,  
996 2009.
- 997 Markus, T., Neumann, T., Abdalati, W., et al.: The Ice, Cloud, and land Elevation Satellite-2 (ICESat-  
998 2): Science requirements, concept, and implementation, Remote Sens. Environ., 190, 260–273, 2017.
- 999 Measures, R. M.: Laser Remote Sensing: Fundamentals and Applications, John Wiley & Sons, New  
1000 York, 1984.
- 1001 Mona, L., Pappalardo, G., Amodeo, A., D'Amico, G., Madonna, F., Boselli, A., Giunta, A., Russo, F.,  
1002 and Cuomo, V.: One year of CNR-IMAA multi-wavelength Raman lidar measurements in coincidence  
1003 with CALIPSO overpasses: Level 1 products comparison, Atmos. Chem. Phys., 9, 7213–7228,  
1004 <https://doi.org/10.5194/acp-9-7213-2009>, 2009.
- 1005 Pappalardo, G., Wandinger, U., Mona, L., Hiebsch, A., Mattis, I., Amodeo, A., Ansmann, A., Seifert,  
1006 P., Linné, H., Apituley, A., et al.: EARLINET correlative measurements for CALIPSO: First intercom-  
1007 parison results, J. Geophys. Res.-Atmos., 115, D00H19, <https://doi.org/10.1029/2009JD012147>, 2010.
- 1008 Pawlowska, H., Brenguier, J.-L., and Burnet, F.: Microphysical properties of stratocumulus clouds,  
1009 Atmos. Res., 55, 15–33, [https://doi.org/10.1016/S0169-8095\(00\)00054-5](https://doi.org/10.1016/S0169-8095(00)00054-5), 2000.
- 1010 Platt, C. M. R.: Lidar and radiometric observations of cirrus clouds, J. Atmos. Sci., 30, 1191–1204,  
1011 1973.



- 1012 Poncet, E., Khaykin, S., Godin-Beekmann, S., and Leblanc, T.: Validation of ATLID stratospheric aer-  
1013 osol retrievals using SAGE III/ISS and ground-based lidar observations, EGUsphere [preprint],  
1014 <https://doi.org/10.5194/egusphere-2026-4358>, 2026.
- 1015 Reverdy, M., Chepfer, H., Donovan, D., Noël, V., Cesana, G., Hoareau, C., Chiriaco, M., and Bastin,  
1016 S.: An EarthCARE/ATLID simulator to evaluate cloud description in climate models, J. Geophys.  
1017 Res.-Atmos., 120, 11090–11113, <https://doi.org/10.1002/2015JD023919>, 2015.
- 1018 She, C.-Y. and Baumgartner, J. N.: Atmospheric Lidar Fundamentals: Weather, Climate and Environ-  
1019 mental Monitoring, Cambridge University Press, Cambridge, 2022.
- 1020 Stoffelen, A., et al.: The Atmospheric Dynamics Mission for global wind measurement, B. Am. Mete-  
1021 orol. Soc., 86, 73–87, 2005.
- 1022 Swales, D. J., Pincus, R., and Bodas-Salcedo, A.: The Cloud Feedback Model Intercomparison Project  
1023 Observational Simulator Package: Version 2, Geosci. Model Dev., 11, 77–81,  
1024 <https://doi.org/10.5194/gmd-11-77-2018>, 2018.
- 1025 Szczap, F., Gour, Y., Fauchez, T., Cornet, C., Faure, T., Jourdan, O., Penide, G., and Dubuisson, P.: A  
1026 flexible three-dimensional stratocumulus, cumulus and cirrus cloud generator (3DCLOUD) based on  
1027 drastically simplified atmospheric equations and Fourier transform framework, Geosci. Model Dev., 7,  
1028 1779–1801, <https://doi.org/10.5194/gmd-7-1779-2014>, 2014.
- 1029 Tesche, M., Wandinger, U., Ansmann, A., Althausen, D., Müller, D., and Omar, A. H.: Ground-based  
1030 validation of CALIPSO observations of dust and smoke in the Cape Verde region, J. Geophys. Res.-  
1031 Atmos., 118, 2889–2902, <https://doi.org/10.1002/jgrd.50248>, 2013.
- 1032 Wang, P., Donovan, D. P., van Zadelhoff, G.-J., de Kloe, J., Huber, D., and Reissig, K.: Evaluation of  
1033 Aeolus feature mask and particle extinction coefficient profile products using CALIPSO data, Atmos.  
1034 Meas. Tech., 17, 5935–5955, <https://doi.org/10.5194/amt-17-5935-2024>, 2024.
- 1035 Weitkamp, C. (Ed.): Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere, Springer Se-  
1036 ries in Optical Sciences, Vol. 102, Springer, New York, <https://doi.org/10.1007/b106786>, 2005.
- 1037 Whiteman, D. N.: Examination of the traditional Raman lidar technique. I. Evaluating the temperature-  
1038 dependent lidar equations, Appl. Optics, 42, 2571–2592, <https://doi.org/10.1364/AO.42.002571>, 2003.
- 1039 Winker, D. M.: LITE: results, performance characteristics, and data archive, Proc. SPIE 3218, Laser  
1040 Radar Ranging and Atmospheric Lidar Techniques, <https://doi.org/10.1117/12.295639>, 1997.
- 1041 Winker, D. M., Pelon, J., Coakley Jr., J. A., Ackerman, S. A., Charlson, R. J., Colarco, P. R., Flamant,  
1042 P., Fu, Q., Hoff, R. M., Kittaka, C., Kubar, T. L., Le Treut, H., McCormick, M. P., Mégie, G., Poole,  
1043 L., Powell, K., Trepte, C., Vaughan, M. A., and Wielicki, B. A.: The CALIPSO Mission: A Global 3D  
1044 View of Aerosols and Clouds, B. Am. Meteorol. Soc., 91, 1211–1229,  
1045 <https://doi.org/10.1175/2010BAMS3009.1>, 2010.
- 1046 Wood, R.: Drizzle in stratiform boundary layer clouds. Part I: Vertical and horizontal structure, J. At-  
1047 mos. Sci., 62, 3011–3033, <https://doi.org/10.1175/JAS3529.1>, 2005.



- 1048 Wu, X. and Liang, X.-Z.: Radiative effects of cloud horizontal inhomogeneity and vertical overlap  
1049 identified from a monthlong cloud-resolving model simulation, *J. Atmos. Sci.*, 62, 4105–4112,  
1050 <https://doi.org/10.1175/JAS3565.1>, 2005.
- 1051 Yorks, J. E., McGill, M. J., Hlavka, D. L., et al.: The Cloud-Aerosol Transport System (CATS): Data  
1052 products and performance overview, *J. Atmos. Ocean. Tech.*, 33, 2225–2241, 2016.
- 1053 Zhou, L., et al.: Dqi-1 satellite and its aerosol and CO<sub>2</sub> lidar measurements, *Adv. Atmos. Sci.*, 39,  
1054 1011–1024, 2022.