Aeolus/ALADIN
Algorithm Theoretical Basis Document
Level 2A products

AEL-FM, AEL-PRO

D.P. Donovan, G-J. van Zadelhoff, P. Wang (KNMI)

Document Version V2.00, April 23, 2024

Applicable to Baseline-16 and 17

Change Log

Date	Author	Doc Version	SW Version	Description
16-09-2021	D.D.	0.1	1.xxxx	Initial version
30-03-2024	D.D	0.9	1.7	Consolidated version
04-04-2024	D.D.	1.0	1.7	Fixed typos in OE equations
04-15-2024	D.D.	1.01	1.7	Fixed typos
04-15-2024	D.D.	2.0	1.6-1.7	Retro added config and output info for baseline 16.

Abbreviations, symbols, and definitions

ACCD	Accumulation Charged Couple Device
ATB	Attenuated backscatter
MS	Multiple Scattering
MSP	Mie Spectrometer
RSP	Rayleigh Spectrometer
OE	Optimal Estimation

Table of Contents

1 Introduction

This Chapter describes two novel algorithms for cloud and aerosol property retrievals that have been implemented in the ALADIN operational processor. These two algorithms are known as AEL-FM and AEL-PRO. AEL-FM provides a high-resolution feature-mask based mainly on ALADIN crosstalk corrected particulate (Mie) profiles AEL-PRO, which uses AEL-FM as an input (as well as attenuated backscatter profiles), provides profiles of particulate backscatter and extinction coefficients and the lidar-ratio. Both these algorithms are based on algorithms previously developed for application to the cloud/aerosol lidar on-board EarthCARE (ATLID) [1].

Before AEL-FM and AEL-PRO could be successfully implemented, accurate pure-molecular and pure particulate profiles of the attenuated backscatter had to be retrieved. To this end, a procedure for producing crosstalk free attenuated backscatter profiles using the Mie Spectrometer (MSP) data alone was implemented. Part of this procedure involves deducing and applying an Effective Mie Spectral Response (EMSR) correction to the MSP measurements.

In this ATBD, AEL-FM and A-PRO are described, and representative results presented and discussed. First, though, the procedure used to generate the pure Rayleigh and Mie attenuated backscatters that are used as inputs to AEL-FM and AEL-PRO is described.

2 Aeolus measurement principle

Aeolus has two distinct detection units; the Mie spectrometer (MSP) and the dual-Fabry Pero based Rayleigh spectrometer (RSP) unit. The signals corresponding to each spectrometer are imaged onto an accumulation -CCD (ACCD) array and contain contributions from both Mie and Rayleigh scattering. For details, the reader is referred to e.g. [2] and [3].

The AEL-FM and AEL-PRO algorithms use "pure" crosstalk corrected profiles of molecular (Rayleigh) and cloud/aerosol (Mie) attenuated backscatter using data obtained from the Aeolus Mie Spectrometer (MSP) only. This contrasts with the "pure" attenuated backscatters generated by the procedure described in [3] using both the MSP and RSP signals.

In general, for a lidar system with crosstalk between two different detection channels we can write

$$B_{R,o} = K_{Ray}(C_1 B_R + C_2 B_M) \qquad\qquad 2.1$$

and

$$B_{M,o} = K_{Mie}(C_4 B_R + C_3 B_M), \qquad\qquad 2.2$$

where B_M and B_M are the intrinsic atmospheric values (e.g. true values of attenuated backscatter or related quantity), $B_{M,o}$ and $B_{M,o}$ are the corresponding detected (equivalent-)photocounts (K_{Ray} and K_{Ray} are defined accordingly). The solution to the above system can be written as:

$$B_R = \frac{-(C_3' B_{R,o} - C_2' B_{M,o})}{C_2' C_4' - C_1' C_3'}$$

2.3

$$B_M = \frac{C_4' B_{R,o} - C_1' B_{M,o}}{C_2' C_4' - C_1' C_3'}$$

2.4

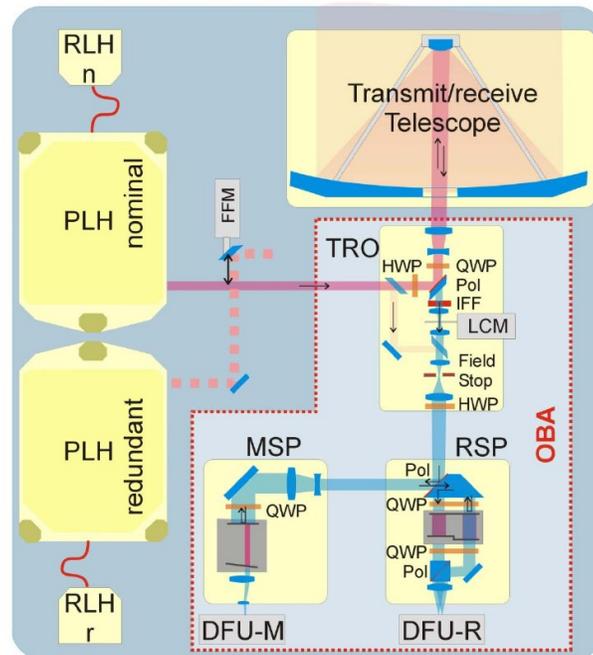where, here, $C_1' = K_{Ray} C_1$ etc.

Figure 1: Schematic of the optical layout of Aeolus (extracted from the L1b ATBD).

The default manner used to produce estimates of the profiles of pure Mie and Rayleigh backscatter signals is to use both the total signals from the MSP and RSP. Then, by using pre-computed crosstalk coefficients (C_1,C_2,C_3,C_4) (which can be a function of the Doppler shift) and MSP and RSP calibration coefficients (K_{Ray}, K_{Mie}) calibration coefficients, the observed total MSP and RSP signals are inverted to yield the pure Mie and Rayleigh signals. Using both the MSP and RSP signals is complicated by the fact that the RSP and MSP signals are generally not on the same vertical grid and the coefficients that define the system are a function of the Doppler shift and, esp., in the case of the C_4 term, not well constrained.

Moreover, the set of C terms is such that the error magnification (see Appendix A) associated with the default crosstalk correction procedure.
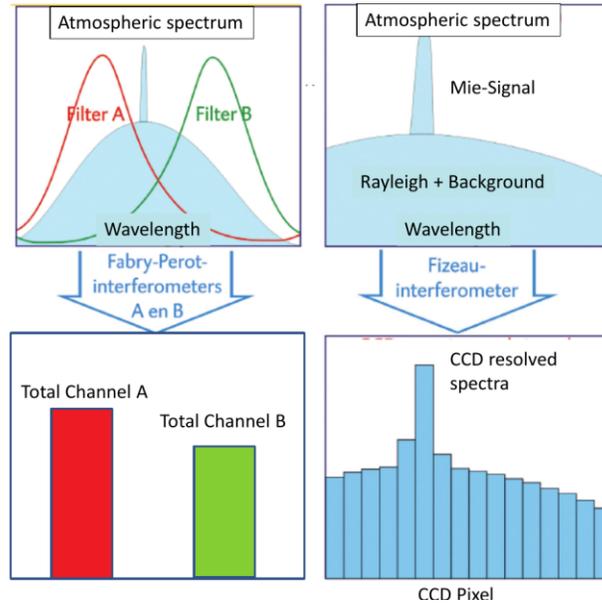

Figure 2: Schematic representation of the operation of the RSP and MSU units. Operationally, only the total signals (i.e., summed across the ACCD pixels) from the RSP A and B channels are available per range-gate. For the MSP unit, however, the 16 bin spectra per range-gate are available. Both the MSP and RSP ACCDs 2D arrays, however, the signals are summed by both rows and columns in the case of the RSP (yielding only the total return) and by row in the case of the MSP (yielding the return by column).

2.1   MSP-only based approach to ATB creation.

An alternative to estimating the pure Mie and Rayleigh attenuated backscatters using the combined MSP and RSP data can be realized by solely using the spectral data routinely delivered by the MSP. The idea is based upon using virtual Mie and Rayleigh channels formed by grouping appropriate ACCD pixels together. Referring to Figure 3, pixels 15 and 12-16 (the wing areas) are summed together to form the virtual Rayleigh channel and the central region pixels form the virtual Mie channel. Crosstalk between the virtual Mie and Rayleigh channels still exists, however, as described later, the associated set of crosstalk coefficients yields a favourable degree of error magnification. In addition, compared to using both the MSP and RSP data, only one calibration constant is present in the system (i.e. K_{ray} is not relevant).
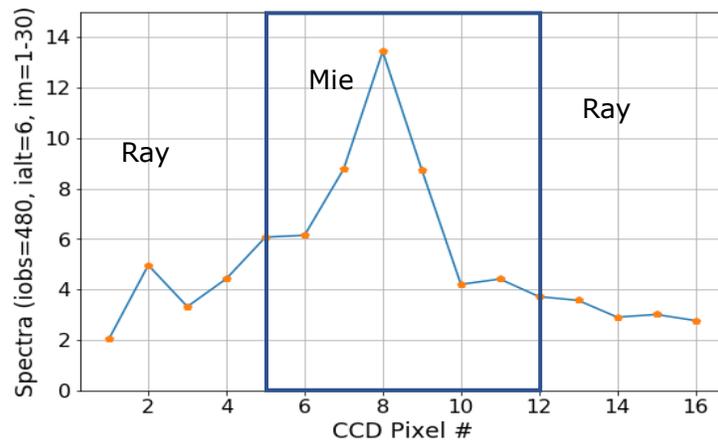
Figure 3: Sample MSP spectra averaged over one observation period (30 measurements profiles) for a single range-gate. The `Ray' and `Mie' labels denote the pixels assigned to the virtual Mie and Rayleigh channels.

In order to determine the necessary crosstalk Coefficients an optical model of the Fizeau-based MSP must be applied. The operation of the MSP is schematically depicted in Figure 4. Here the response to an input combined Mie and Rayleigh spectrum per ACCD pixel column is illustrated. An important point is that, due to the small FSR of the device, that several FSRs must be considered to accurately calculate the Rayleigh response. The set of crosstalk coefficients are found by calculating the per-pixel relative response to pure Rayleigh and Mie inputs and then respectively summing these over the appropriate pixels comprising the virtual channels. The resulting set of coefficients are functions of the optical/physical parameters of the FSP as well as the input spectra. The calculation of the crosstalk coefficients as well as the Fizeau model used is described in Appendix-B.
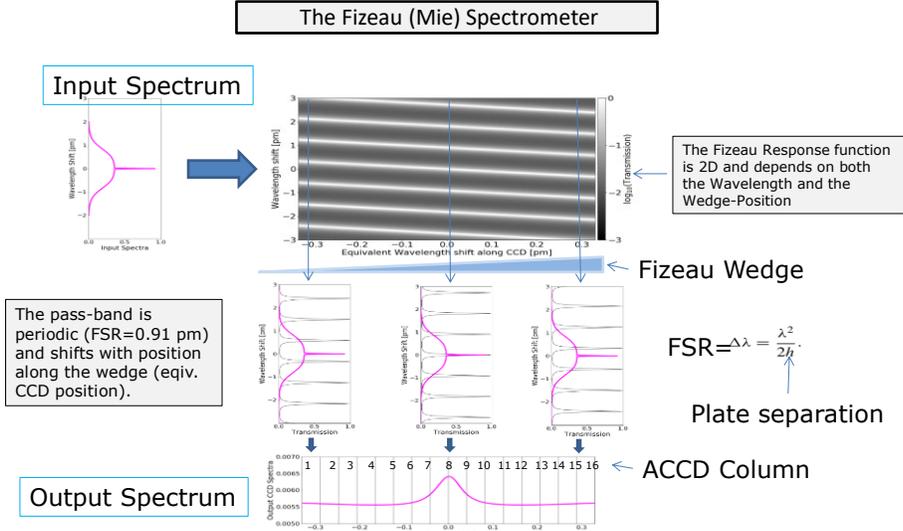
Figure 4: Schematic depiction of the optical transfer function of the MSP. The grey-scale image represents the spectral response for a single ACCD row (i.e., the image is 2-D, i.e. wavelength – wedge position space). The passband is a function of wedge position (which is here assumed to be the same for each ACCD row). The narrow Free-spectral-range (FSR) of the device is such that several FSRs much be taken into account to accurately calculate the response to an input Rayleigh spectrum signal.

Once the crosstalk coefficients are known, the summed signal levels (after background and dark-count subtraction) in the virtual Mie and Rayleigh channels can be inverted to yield the pure Rayleigh and Mie signals. The crosstalk coefficients correspond to a zero Doppler shift and a uniform intensity distribution across the MSP. The process of accounting for the non-uniform intensity distribution across the MSP is described in Appendix-C. Depending on the exact pixel boundaries chosen to define the Mie and Rayleigh virtual channels, the crosstalk coefficients are somewhat insensitive to expected Doppler shifts. Never-the-less, as a simple way to account for possible Doppler shifts, the centroid of the spectra is calculated and used to adjust the Mie and Rayleigh regions.

Putting the steps together one has (per time and range-gate) starting with the **mie_measurement_data** (which have already been corrected for dark counts) we have:

1. Screen for (and attempt to correct for) missed Hot-Pixels (See Appendix E).

2. Subtract the solar background from the spectra. For each measurement time we have:

$$N'_{i,j} = N_{i,j} - N_{i,25} \frac{\Delta t_{z,j}}{\Delta t_{Bak}},$$

2.5

where

- N^1 is the number of ACCD counts after Dark Count subtraction and background subtraction.

- $N_{i,j}$ is the number of ACCD counts after Dark Count subtraction. It is associated with the **mie_measurement_data** variable contained within the L1B data product.
- $3 \leq i \leq 18$ is the MSP spectral index.
- $1 \leq j \leq 24$ is the range index.
- N is the number of ACCD counts after Dark Count subtraction.
- $N_{i,25}$ is the number of ACCD counts accumulated in the background column.
- $\Delta t_{z,j}$ is the time interval associated with range gate j.
- Δt_{Bak} is the time interval associated with the background gate.

3. Apply the Effective MSP spectral response (See Appendix C):

$$N''_{i,j} = N^1{}_{i,j}\text{EMSR}_j \qquad\qquad 2.6$$

4. Estimate the errors associated with spectra. The approach used here is based on assuming that the errors in each pixel can be reasonably well-described by Poisson statistics. However, the DCO and the relationship between the reported **mie_measurement_data** values and the detected photo-counts must be "calibrated".

   a. DCO variance estimation: The operational code "knows" the exact value of the DCO offset applied to the **mie_measurement_data** spectral data. However, the prototype code does not have access to this information. It is known, however, that the DCO was calculated using either Pixels 1,2 and/or Pixels 19,20. Realizing that we do not strictly need the value of the DCO, but only its standard deviation (SD) instead we estimate the calculate the DCO SD by assuming it is constant per ACCD pixel per observation interval (but can vary with altitude. Accordingly, we estimate the DCO SD ($SD\_\{DCO\}$) using the SD values of the first and last two bins separately and choosing the higher of the two.

   b. Excess noise factor estimation:
      i. Using the same set of clear-air regions identified by the EMSR determination procedure (see Appendix-C) the variance of the selected data (per pixel) is calculated and the estimated contribution of the DCO is removed (using the result of Step a above).
      ii. An **excess_noise_factor** (F_N assumed valid for the whole orbit) is the calculated by assuming that the true variance is equal to the number of detected photocounts (i.e. Poisson statistics are valid). The per-pixel excessive noise factor is then determined by calculating the square-root of the observed variance divided by the expected variance. This is not an exact approach (since we do not know the DCO level) but is expected to be reasonably valid when the signals are not DCO dominated i.e.

$$F_{N,i} = \sqrt{\frac{\sigma^2_{i,Est}}{\sum N_i/n}}$$

<div align="right">2.7</div>

where n is the number of samples and the estimated observed per-spectral pixel variance of the observed ACCD counts (after DCO offset correction) is estimated by

$$\sigma^2_{i,est} = \sigma^2_{i,N} - \sigma^2_{DCO}$$

<div align="right">2.8</div>

If the statistics of N follow Poisson statistics, then $F_{i,N}$ will equal 1.

c. The per-pixel error estimates can now be found according to:

$$\sigma^2\left(N''_{i,j}\right) = \left(\frac{1}{EMSR_i}\right)^2 \left[\sum_{i_1+I_{\text{shift,j}}}^{i_2+I_{\text{shift,j}}} F_i^2 N_{i,j} + F_i^2 N_{i,25}\left(\frac{\Delta t_{z,j}}{\Delta t_{Bak}}\right)^2\right] + \sigma^2_{DCO}$$

<div align="right">2.9</div>

5. Calculate the spectrum centroid in order to calculate the shift in the spectrum from the central expected zero Doppler position.

$$I_{\text{center,j}} = \sum_3^{18}\left(i \times N''_{i,j}\right) / \sum_3^{18} N''_{i,j}$$

<div align="right">2.10</div>

6. If the spectrum shift is larger than one pixel, then shift the Mie and Rayleigh virtual channel boundaries as appropriate. This is calculated for each measurement interval.

$$I_{shift,\text{j}} = \left(\text{int}\left((I\}_{\text{center,j}} - 8.5\right) + 0.5\right)$$

<div align="right">2.11</div>

7. Sum the ACCD counts associated with the virtual Mie and virtual Rayleigh pixels. Accordingly, we have,

$$N_{\text{M,O,j}} = \sum_{i_1+I_{\text{shift,j}}}^{i_2+I_{\text{shift,j}}} N''_{i,j}$$

<div align="right">2.12</div>

and

$$N_{R,O,j} = \sum_{3}^{i_1+I_{\text{shift},j}-1} N_{i,j}'' + \sum_{i_2+I_{\text{shift},j+1}}^{18} N_{i,j}'' \qquad 2.13$$

where i_1 and i_2 are the spectral pixel limits assigned to the virtual Mie channel.

8. The error corresponding to the summed counts in both virtual channels are calculated by quadratically summing the estimated spectrum pixel errors including the estimated DCO SD and the background signal level.

$$\sigma^2\left(N_{M,O,j}\right) = \sum_{i_1+I_{\text{shift},j}}^{i_2+I_{\text{shift},j}} \sigma^2\left(N_{i,j}''\right) \qquad 2.14$$

and

$$\sigma^2\left(N_{R,O,j}\right) = \sum_{3}^{i_1+I_{\text{shift},j}-1} \sigma^2\left(N_{i,j}''\right) + \sum_{i_2+I_{\text{shift},j+1}}^{18} \sigma^2\left(N_{i,j}''\right) \qquad 2.15$$

9. Using the appropriate set of crosstalk coefficients (See Appendix C), invert the observed signals from Step 8 to yield the crosstalk corrected pure Mie and Rayleigh signals. The errors in the crosstalk corrected signals are calculated using the errors calculated in Step 8 and applying standard (quadratic) error propagation. Accordingly, we have:

$$N_{R,j} = C_1^{inv} N_{R,O,i} + C_2^{inv} N_{M,O,i} \qquad 2.16$$

$$N_{M,j} = C_4^{inv} N_{R,O,i} + C_3^{inv} N_{M,O,i} \qquad 2.17$$

where

$$C_1^{inv} = \frac{-C_3}{C_2 C_4 - C_1 C_3} \qquad 2.18$$

$$C_2^{inv} = \frac{C_2}{C_2 C_4 - C_1 C_3} \qquad 2.19$$

$$C_3^{inv} = \frac{-C_1}{C_2 C_4 - C_1 C_3} \tag{2.20}$$

$$C_4^{inv} = \frac{C_4}{C_2 C_4 - C_1 C_3} \tag{2.21}$$

$$\sigma^2(N_{R,j}) = \left(C_1^{inv}\right)^2 \sigma^2(N_{R,O,i}) + \left(C_2^{inv}\right)^2 \sigma^2(N_{M,O,i}) \tag{2.22}$$

$$\sigma^2(N_{M,j}) = \left(C_4^{inv}\right)^2 \sigma^2(N_{R,O,i}) + \left(C_3^{inv}\right)^2 \sigma^2(N_{M,O,i}) \tag{2.23}$$

10. Divide the outputs by the range-bin widths and multiple by the square of the range to the lidar and the lidar calibration constant to generate profiles of the Mie and Rayleigh attenuated backscatters. The corresponding errors are scaled in a similar fashion.

    The calibration constant that is applied to both the Rayleigh and Mie ATB profiles (and their corresponding error profiles and error covariance matrix) is obtained via normalization to the return expected on the basis of the molecular density profile and the scattering ratio estimated by dividing the (un-normalized) Mie and Ray ATBs at altitudes above 20 km. When data is not available at these altitudes then interpolation in time is used (See Appendix-D).

    Accordingly, we have:

$$b_{R,j} = C_{lid} r_j^2 N_{R,j} \frac{1}{\Delta r_j}, \tag{2.24}$$

$$b_{M,j} = C_{lid} r_j^2 N_{M,j} \frac{1}{\Delta r_j}, \tag{2.25}$$

$$\{\sigma(b)_{R,j}\} = C_{lid} r_j^2 \{\sigma(N)_{R,j}\} \frac{1}{\Delta r_j}, \tag{2.26}$$

$$\{\sigma(b)_{M,j}\} = C_{lid} r_j^2 \{\sigma(N)_{M,j}\} \frac{1}{\Delta r_j}. \tag{2.27}$$

11. The error covariance matrices for each range-time bin for the Mie and Rayleigh ATBs are calculated using standard quadratic error propagation.

$$C_{i,j} = \sigma^2(b_{R,i}) \qquad : i = j \text{ and } i \le n \tag{2.28}$$

$$C_{i,j} = \sigma^2(b_{M,i}) \qquad : \quad i = j \text{ and } i > n \tag{2.29}$$

$$C_{i,j} = \frac{C_{lid}r_j^2}{\Delta r_j}\left[C_1^{inv}C_4^{inv}\sigma^2\left(N_{R,O,j}\right) + C_2^{inv}C_3^{inv}\sigma^2\left(N_{M,O,j}\right)\right] \qquad : \quad j = i + n \qquad\qquad 2.30$$

$$C_{i,j} = \frac{C_{lid}r_j^2}{\Delta r_j}\left[C_1^{inv}C_4^{inv}\sigma^2\left(N_{R,O,j}\right) + C_2^{inv}C_3^{inv}\sigma^2\left(N_{M,O,j}\right)\right] \qquad : \quad i = j + n \qquad\qquad 2.31$$

$$C_{i,j} = 10^{-10} \qquad : \quad \text{if } z_i \text{ or } z_j \ \le \text{Dem Height} \qquad\qquad 2.32$$

$$C_{i,j} = 0.0 \qquad : \quad \text{otherwise} \qquad\qquad 2.33$$

Where here i and j are both range indices and n is the number of range-gates.

Using the procedure outlined here, pure Mie signals with SNR ratios markedly superior to those produced using the default procedure are realized. This is somewhat unexpected since fewer photons are used in the process. However, the smaller number of photons is more than offset by the reduction in the associated crosstalk correction error magnification factor (See Figure 6). For the Rayleigh channel, the improved situation with respect to the crosstalk correction error magnification factor does not quite compensate for the increased noise due to the fact that the MSP total Rayleigh signal is about 4 times smaller than the corresponding RSP signal. This leads to about the same or somewhat less SNR in the crosstalk corrected Rayleigh channel. This, in turn, suggests that further work should be conducted to combine a variation of the procedure developed here with the RSP signals.
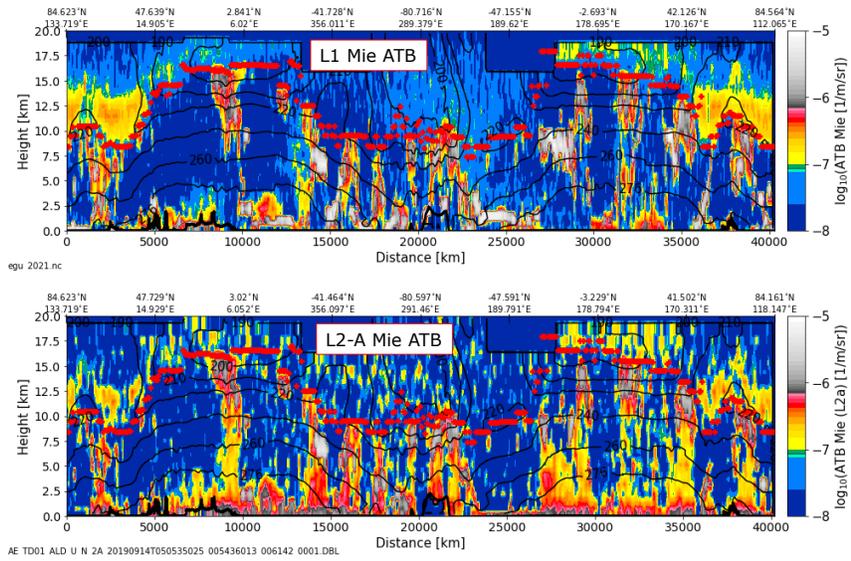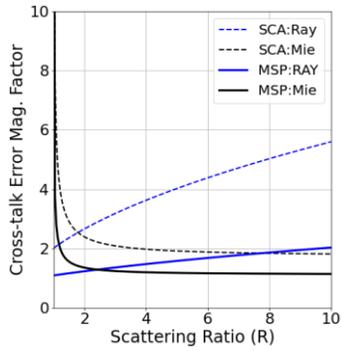
Figure 5: (Top) Mie ATBs calculated using the MSP only method described here. (Bottom) corresponding Mie ATBs estimated using the MSP+RSP method.

The L1 Fizeau set of coefficients yield much more favorable cross-talk error magnification factors

	L2a (SCA)	L1a (MSP-only)
C_1	0.9	0.5
C_2	0.5	0.1
C_3	1.3	1
C_4	1.0	0.3

However, this is balanced by the fact that the total useful Rayleigh signal is on the order of 3-4 times larger than the useful Mie signal.

So if one excludes the Rayleigh channels , that (depending on DCO and Bkgrnd levels) works out to something on the order of a factor of 2 reduction in SNR.

Combining this expected SNR reduction with the expected increase in SNR associated with the better Fns, one should expect about the same SNR in B_R and a definite improvement in the SNR in B_M when comparing the L1 and L2 estimates.

Figure 6: Crosstalk random error magnification factors (Defined in Appendix-A) for the default L2 approach and the MSP-only method (L1a).

3 AEL-FM algorithm

AEL-FM provides an index that corresponds to the probability of a target being present and identifies clear-air as well as attenuated regions. The product is reported at the native resolution of the input ATBs (about 3 km horizontal resolution). The procedure combines a number of techniques applied to the Mie ATBs including a hybrid-median edge preserving filter (in order to identify ``strong'' returns) and iterative adaptive Gaussian smoothing (which is targeted towards weaker extended features). The pure Rayleigh profiles are also used, mainly to determine attenuated regions. AEL-PRO is closely based on A-PRO. Both A-FM and AEL-FM are described in detail in [4]. The AEL-FM results are used to as input to the AEL-PRO algorithm. Sample AEL-FM results are shown later in Section 5

4 AEL-PRO algorithm

4.1 Overview and Theoretical Background

AEL-PRO is inspired by the corresponding A-PRO algorithm [5] developed for the EarthCARE lidar, ATLID. AEL-PRO is a multiscale, forward modelling optimal-estimation based approach to retrieving profiles of extinction and lidar-ratio. A major challenge to obtaining useful estimates of extinction and backscatter from Aeolus is the low (compared to terrestrial systems) SNR of the data. Along-track averaging can improve the SNR of the data, however, due to the highly non-linear nature of the lidar equation, indiscriminately averaging over both strong (e.g. cloud) and weak (e.g. aerosol) returns then performing an inversion usually leads to ambiguous (and perhaps physical meaningless) results. AEL-PRO avoids this issue by performing two passes. The first pass, applied to cloud screened data which can be averaged along-track, is aimed at retrieving the (homogeneous) aerosol and thin ice cloud. The 2nd pass, which uses information from the 1st pass, is applied at the highest available horizontal resolution is aimed at retrieving the extinction and backscatter associated with the strong (e.g. cloud) returns.

In the following sub-sections, the basic theoretical background components are outlined staring with the methods used to model the lidar returns. The optimal-estimation component of the procedure is then discussed. A more detailed implementation description of the entire AEL-PRO procedure is given in Section 4.3.

4.1.1 Single scattering lidar equations

After crosstalk correction and calibration, the Aeolus Rayleigh and Mie channels attenuated backscatter can be related to the atmospheric extinction and backscatters (neglecting multiple-scattering effects) as:

$$b_{R,\parallel}(z) = p_{R,\parallel}(z)r(z)^2 = \beta_{R,\parallel}(z) \exp\left[-2\int_{z_{sat}}^{z}\left(\alpha_M(z') + \alpha_R(z')\right)dr'\right] \qquad 4.1$$

and

$$b_{M,\parallel}(z) = p_{M,\parallel}(z)r(z)^2 = \beta_{M,\parallel}(z) \exp\left[-2\int_{z_{sat}}^{z}\left(\alpha_M(z') + \alpha_R(z')\right)dr'\right] \qquad 4.2$$

Where $b_{R,\parallel}$ is the co-polar Rayleigh attenuated backscatter, $b_{M,\parallel}$ is the co-polar Mie attenuated backscatter, $p_{R,\parallel}$ is the Rayleigh co-polar power, $p_{M,\parallel}$ is the Mie co-polar power. z is the atmospheric altitude and $r(z)$ is the range from the lidar. A_M is the aerosol and cloud extinction, α_R is the atmospheric Rayleigh extinction. $B_{M,\parallel}$ is the co-polar Mie backscatter, $\beta_{R,\parallel}$ is the co-polar Rayleigh backscatter.

Note that the above equations are appropriate for calibrated signals (i.e. explicit calibrations factors are all assumed to have a value of 1). Further note that Aeolus does not measure cross-polar power (thus the equations above only refer to the co-polar backscatters). This has the effect of reducing the received power and has implications when comparing e.g. the retrieved backscatter with other lidar systems.

4.1.2 Multiple Scattering

In general, for lidar cloud remote sensing multiple scattering effects must be taken into account. This is especially true for space-based observations. It is true even for the narrow instrument field-of-view (fov) associated with ALADIN. A comparison between attenuated backscatter signals corresponding to an idealized homogeneous ice cloud of optical thickness 1 and effective ice cloud particle size of 20 microns for Aeolus, EarthCARE and Calipso lidars is shown in Figure 7. The calculations were performed using the approach of Eloranta [6]. Here it can be seen that while the narrow fov of ALADIAN leads to the smallest relative contribution of MS to the total signal, the 2nd order signal is significant.

Idealized Ice cloud optical thickness of 1 and an effective particle size of 20 microns
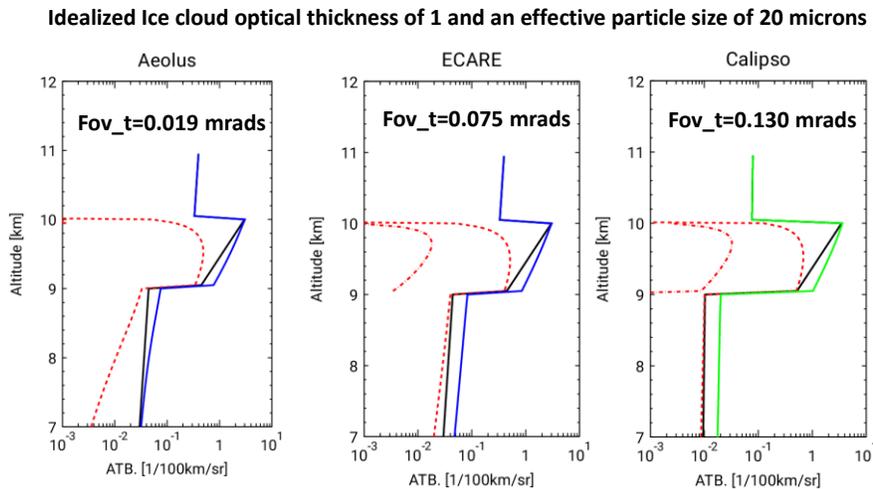

Figure 7: The above figure shows the single, 2nd order and 3rd order attenuated backscatter returns calculate using an semi-analytical model due to Eloranta.

In order to calculate lidar multiple-scattering, various approaches can be used. Monte-Carlo (MC) simulations are highly accurate. However, in general, they are too computationally expensive to use in an inversion procedure. Hogan [7] developed an analytical model, which is accurate and faster than the approach of Eloranta. The Platt approach **Error! Reference source not found.** is even faster, but less a

ccurate, especially for the signal below the cloud layer. Here we use a novel extension to Platt's approach which is faster than Hogan's method and sufficiently accurate for the purpose at hand.

In section 4.1.2.1 we describe the Platt's approach and our extension which forms the basis of the forward model used in AEL-PRO.

4.1.2.1   An Extension to Platt's approach (Platt+Tails)

Platt's approach

When the particles are large compared to the wavelength of the laser light so that half the scattered energy is scattered forward in a narrow diffraction lobe and largely stays within the lidar receiver file-of-view. This result was noted by Platt (1973) and forms the basis of a simple method for accounting for Multiple-Scattering (MS) effects.

Using Platt's approach, the multiple scattering factor $M_p(z)$ (the ratio between the lidar return and the single-scatter only return) is,

$$M_p(z) = \exp\left[2\int_{z_{Sat}}^{z}\left(1-\eta(z)\right)\alpha_M(z)dr'\right] \tag{4.3}$$

where η is the Platt coefficient which physically describes the fraction of scattered energy that remains within the lidar filed-of-view (and thus behaves like it has not been scattered).

Taking into account multiple scattering factor $M_p(z)$ in the lidar equation, we get

$$b_t(z) = \left(\beta_M(z) + \beta_R(z)\right)\exp\left[-2\int_{z_{sat}}^{z}\left(\eta(z')\alpha_M(z') + \alpha_R(z')\right)dr'\right] \tag{4.4}$$

Platt's approach performs well inside the cloud layer and is faster than the model of Hogan but it cannot capture the decaying structure below the cloud (structure of the 'tail'). Here bt refers to the total (cross+co-polar) signal, however, the form of the equations for the corresponding individual co- and cross- polar equations trivially follows.
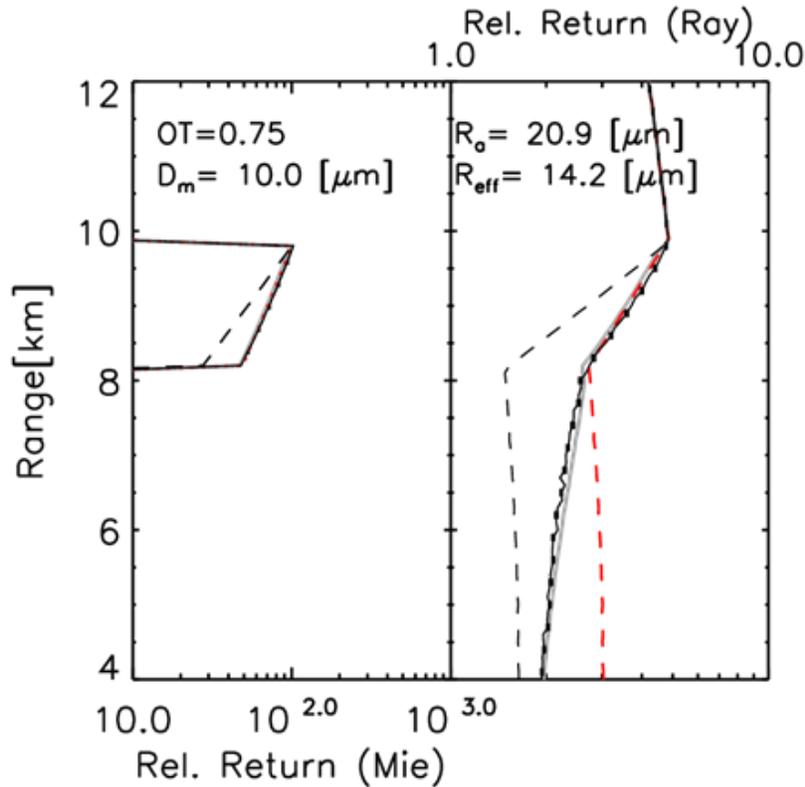
Figure 8: Mie co-polar and (Right) Rayleigh channel co-polar returns for an ice cloud of an optical thickness of 0.75 and an effective radius of 14.2 microns. Black solid: MC results. Dashed Black: Single scattering results. Solid Grey: Hogan's model results and Red-Dashed the results of Platt's method with a MS factor of 0.45. Note these results correspond to ATLID conditions (Altitude = 400km, laser field-of-view = and receiver field-of-view= ?.

Referring to Figure 8, it can be seen that Hogan's model and the MC results agree well. Platt's approach also works well, but only within the cloud. Below the cloud base, it cannot replicate the shape of the "tail" i.e. the gradual decay of the multiply scattered signal with range from cloud base.

Origin of the tails

The origin of the below-cloud tail can be understood as follows:

Within the cloud, the low mean-free-path of the photons ensures that the multiply-scattered light that contributes to the detected signal tends to be confined to within the field-of-view of the lidar. However, the angular variance of the lidar beam will be broadened as it propagates downwards through the cloud with more and more photons undergoing scattering events.

At cloud base the lidar beam emerges with an effective angular divergence which increases with the optical thickness of the cloud and decreases with the size of the cloud particles. This is due to that fact that the angular width of the cloud phase function forward lobe increases with decreasing particle size i.e.

$$\theta_{sc} \simeq \left(\frac{\lambda}{\pi Ra} \right) \qquad\qquad 4.5$$

Below cloud base the lidar beam will continue to propagate with a given divergence. However, the horizontal spread of the photons is no longer constrained by the presence of the cloud. As the beam continues to propagate downwards, depending on the lidar receiver footprint more and more of the multiple-scattered photons will travel outside of the receiver cone. Therefore, the signal below the clouds decays towards single scatter levels.

An Extension to Platt's approach

Our extension of the Platt's approach copes with the existence of MS tails. Based on the fact that the tails are fundamentally a consequence of the return signal decaying towards single-scatter values, we write the multiple scattering signal as,

$$b_M(z) = \beta_M(z)e^{-2\tau(z)}\left[\left(1 - f(z)\right) + f(z)e^{2\eta'\tau(z)} \right] \qquad\qquad 4.6$$

where

$$\eta' = 1 - \eta. \qquad\qquad 4.7$$

η is the (regular) Platt multiple-scattering extinction factor (here, for simplicity, assumed to be constant per layer) and $f(z)$ is the return signal multiple scattering fraction. The above equation is particular to the Mie channel, however, fundamentally similar equations can also be written for the Rayleigh channel. When $f(z) = 1$, Eq. 4.6 effectively reduces to Eq. 4.3 and when $f(z) = 0$, it reduces to the normal single-scattering lidar equation. When η' is zero then Eq. 4.6 also, reduces to the single-scatter case.
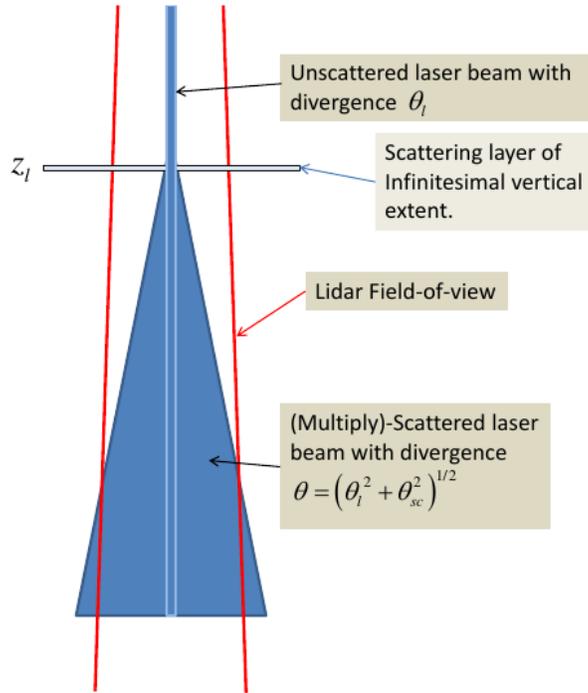
Figure 9: Schematic depiction of the angular broadening experienced by a lidar pulse as it interacts with a physically think scattering layer at altitude z_l.

Now we have to determine the profile of $f(z)$. We start by considering the case of a physically thin scattering layer. If we assume that the beam has a Gaussian profile and model the forward-scattering lobe of the effective layer phase function by a Gaussian, then the divergence of the forward scattered light will also be Gaussian with a divergence given by the convolution of the incoming beam divergence (θ_l) with the effective scattering forward-lobe width (θ_{sc}) so that the effective width of the multiply scattering radiation emerging from the layer bottom is given by

$$\theta_{eff} = \left(\theta_{sc}^2 + \theta_l^2\right)^{1/2} \qquad\qquad 4.8$$

where θ_{sc} is related to the effective area particle size

$$\theta_{sc} = \frac{\lambda}{\pi R_a} \qquad\qquad 4.9$$

The fraction of multiply scattered light that then remains within the lidar fov can easily be calculated and is given by

$$f(z, z_l) = 1.0 - \exp\left[-\left(\frac{\rho_t^2(z - z_{sat})^2}{\theta_{sc}^2(z - z_l)^2 + \rho_l^2(z - z_{sat})^2}\right)\right] \qquad\qquad 4.10$$

where ρ_t is the receiver telescope fov, ρ_l is the laser divergence, z_{sat} is the satellite altitude and z_l is the altitude of the scattering layer. This expression is only valid for a single thin scattering layer. Here we will use the information present in the signal itself to calculate the effective $f(z)$ under general conditions. Since the observed signal itself contains information on the location and relative strength of the scattering at each level, we postulate the form

$$f_e(z) = \frac{\int_{z_{sat}}^{z} f(z, z_l) b_t(z_l) dr_l}{\int_{z_{sat}}^{z} b_t(z_l) dr_l}$$

<div align="right">4.11</div>

where, b_t is the sum of Rayleigh and Mie channel attenuated backscatters. That is, we use the observed total backscatter itself as a weighting factor to determine the effective $f(z)$ profile. In the limit of a single thin scattering layer this expression yields the correct result.

As a further refinement, in order to explicitly account for the fact that the effective backscatter coefficient for particulate backscattering (which may be strongly peaked around 180 Deg) may, in general, be lower than the effective backscatter coefficient associated with molecular Rayleigh scattering (which is smooth in the backscattering direction), an additional factor is added which acts to adjust Eq. 4.11 for the Mie channel only

$$f_{M,e}(z) = f_{Msp}(z) f_e(z).$$

<div align="right">4.12</div>

Putting all the above elements together we have, specific for Aeolus calibrated crosstalk corrected attenuated backscatters

$$b_{M,\parallel}(z) = S_{\parallel}(z)\alpha_M(z)e^{-2\tau(z)}\left[\left(1 - f_{M,e}(z)\right) + f_{M,e}(z)e^{2\tau_\eta(z)}\right]$$

<div align="right">4.13</div>

and

$$b_{R,\parallel}(z) = \beta_{R,\parallel}(z)e^{-2\tau(z)}\left[\left(1 - f_e(z)\right) + f_e(z)e^{2\tau_\eta(z)}\right]$$

<div align="right">4.14</div>

where τ and τ_η are given by

$$\tau(z) = \int_{z_{sat}}^{z} \left(\alpha_R(z') + \alpha_M(z')\right) dr,$$

<div align="right">4.15</div>

and

$$\tau_\eta(z) = \int_{z_{sat}}^{z} \eta'(z')\alpha_M(z')dr,$$

<div align="right">4.16</div>

and the lidar-ratio is defined as

$$\alpha_M = S_\|(z)\beta_{M,\|}.$$

<div align="right">4.17</div>

In order to apply the preceding formalism both η and θ must be specified. In this work we fix η and model θ as a function of particle size. more detail can be found in Appendix B of [5].

4.2   Optimal Estimation procedure

In this work we employ the principle of Optimal Estimation [Rodgers 2000]. In general terms, we formulate a cost-function that characterizes the likelihood of the measurements being what they were given a particular parameter configuration of an appropriate forward model combined with our expectations concerning the forward-model parameters. The desired solution is then obtained by numerically minimizing this function.

Our cost function can be written as

$$\chi^2 = [\mathbf{y} - \mathbf{F(x)}]^T \, \boldsymbol{S}_e^{-1}[\mathbf{y} - \mathbf{F(x)}] + [\mathbf{x_r} - \mathbf{x_a}]^T \boldsymbol{S}_a^{-1}[\mathbf{x_r} - \mathbf{x_a}].$$

<div align="right">4.18</div>

$\mathbf{y}$ is the observation vector including the observed Rayleigh and Mie attenuated backscatters (note we have dropped the || subscripts) defined as

$$\boldsymbol{y} = \left(B_{R,1}, B_{R,2}, .. B_{R,n}, B_{M,1}, B_{M,2}, ... B_{M,n}\right)^T$$

<div align="right">4.19</div>

where n is the number of altitude bins and *the Rayleigh transmission corrected attenuated backscatters* (B) are defined as

$$B_i = b_i \exp\left(2\tau_{Ray}(z_i)\right)$$

<div align="right">4.20</div>

$\mathbf{x}$ is the logarithmic *state-vector*. Here defined such that

$$\mathbf{x} = \log_{10}[(\alpha_1, \alpha_2, \dots, \alpha_n, S_1, S_2, \dots, S_n, Ra_1, Ra_2, \dots, Ra_n, C_{lid})]^T$$

<div align="right">4.21</div>

$\mathbf{x}_a$ is the logarithmic a-priori state vector. Here defined as a vector consisting of the log base 10 values of the a-priori lidar-ratios, effective area particle sizes and the value of C_lid appropriate for calibrated attenuated backscatter signals (i.e. 1).

$$\mathbf{x_a} = \log_{10}\left[\left(S_{a,1}, S_{a,2}, \dots, S_{a,n}, Ra_{a,1}, Ra_{a,2}, \dots, Ra_{a,n}, 1\right)\right]^T.$$

<div align="right">4.22</div>

$\mathbf{x_r}$ is the reduced state-vector which is just the state vector without the extinction elements. This is consistent with the fact that no a-priori constraints are placed upon the log extinction values so that they are not present in the a-priori state-vector.

$\mathbf{S_e}$ is the observation error covariance matrix.

$\mathbf{S_a}$ is the a priori error covariance matrix appropriate to the logarithmic state-vector.

$\mathbf{F(x)}$ is the forward-model vector

4.2.1   Forward Model

The discrete forms of Eqs. 4.13 and 4.14 form the basis of the forward model. Accordingly, for $i \leq n$ we have

$$F_i = B_{R,i} = \frac{C_{lid}}{\Delta r_i} \exp(2\tau_{Ray,i}) \int_{r_{i,mid}-\Delta r/2}^{r_{i,mid}+\Delta r/2} \left(\beta_{R,\parallel}(z') e^{-2\tau(z')} \left[(1 - f_e(z')) \right. \right. \\ \left. \left. + f_e(z') e^{2\tau_\eta(z')} \right] \right) dr \tag{4.23}$$

where

$$\tau_{R,i} = \sum_{j=1}^{j=i-1} \alpha_{R,j} \Delta r_j, \tag{4.24}$$

and for $n < i \leq 2n$ we have

$$F_i = B_{M,i} = \frac{C_{lid}}{\Delta r_i} \exp(2\tau_{Ray,i}) \int_{r_{i,mid}-\Delta r/2}^{r_{i,mid}+\Delta r/2} \left(S_\parallel(z) \alpha_M(z) e^{-2\tau(z)} \left[\left(1 - f_{M,e}(z) \right) \right. \right. \\ \left. \left. + f_{M,e}(z) e^{2\tau_\eta(z)} \right] \right) dr'. \tag{4.25}$$

Assuming that for each range-bin, that the Mie and Rayleigh extinctions, lidar-ratio, and f terms can be treated as being constant, evaluating the integral in Eq. 4.23 then yields, for the Rayleigh signal,

$$i \leq n:\ F_i = B_{R,i} = C_{lid} \exp(-2\tau_{M,i}) \beta_{R,\parallel,i} \left(\Delta z_{c1,i} (1 - f_{e,i}) + \Delta z_{c2,i}\, f_{e,i} e^{2\tau_{\eta,i}} \right) \tag{4.26}$$

where

$$\tau_{M,i} = \sum_{j=1}^{j=i-1} \alpha_{M,j}\Delta r_j, \tag{4.27}$$

$$\tau_{\eta,i} = \sum_{j=1}^{j=i-1} \alpha_{M,j}\eta'_j\Delta r_j, \tag{4.28}$$

$$\Delta z_{c1,i} = \frac{1 - \exp\big(-2(\alpha_{M,i} + \alpha_{R,i})\Delta r_i\big)}{2(\alpha_{M,i} + \alpha_{R,i})\Delta r_i}, \tag{4.29}$$

and

$$\Delta z_{c2,i} = \frac{1 - \exp\left[-\left(2 - \big((1 - \eta'_i)\alpha_{M,i} - \alpha_{R,i}\big)\Delta r_i\right)\right]}{2\left((1 - \eta_i)\alpha_{M,i} - \alpha_{R,i}\right)\Delta r_i}. \tag{4.30}$$

For the Mie signal evaluating the integral in Eq. 4.24 gives

$$n < i \le 2n: F_i = B_{M,i}$$
$$= C_{lid}\exp(-2\tau_{M,i})\,\alpha_{M,i}S_i^{-1}\big(\Delta z_{c1,i}(1 - f_{M,e,i}) + \Delta z_{c2,i}\,f_{M,e,i}e^{2\tau_{\eta,i}}\big) \tag{4.31}$$

4.2.2 Observation error covariance matrix

The observational error matrix is calculated as described in Step 11 of Section 2.1 with the following refinement. Namely, that if a pixel is flagged as a surface return in the input AEL-FM data then the variance is set to a large value (10E-10 which is large with respect to the Mie and Ray channel calibrated attenuated backscatter but not so large that it creates numerical difficulties in the inversion of the matrix).

The (pseudo-)inverse of the observational error matrix (i.e. S_e^{-1}) is calculated using a SVD approach.

4.2.3 A priori error covariance matrix

The a-priori errors are assumed to be uncorrelated. The a priori values of the lidar-ratio and effective area particle size, the a priori errors depend the target classification.

The form of the a priori error matrix is that appropriate for log normally distributed errors [9] i.e.

$$Sa_{i,i} = \log_{10}\left(1 + \frac{\sigma_{x_i}^2}{x_{a_i}^2}\right) \tag{4.32}$$

where σ_{x_i} is the a-priori (linear)uncertainty assigned to the ith component of the state vector.

4.2.4   Gradient and Jacobian

In order to efficiently minimize the cost function, we must be able to compute its gradient with respect to the elements of the state vector. The gradient of the cost-function is related to the Jacobian of the forward model as:

$$\nabla \chi^2 = -2\mathbf{J}^T S_e^{-1}\big(\mathbf{y} - \mathbf{F}(\mathbf{x})\big) + 2S_a^{-1}(\mathbf{x_r} - \mathbf{x_a}), \qquad\qquad 4.33$$

where $\mathbf{J}$ is the forward model Jacobian with respect to the log state variables i.e.

$$J_{i,j} = \frac{\partial F_i(\mathbf{x})}{\partial log_{10}x_j} = \frac{\partial F_i(\mathbf{x})}{\partial x_j}\frac{\partial x_j}{\partial log_{10}x_j} = \frac{\partial F_i(\mathrm{x})}{\partial x_j}\ln(10)x_j. \qquad\qquad 4.34$$

The partial derivatives can be calculated using Eqs. 4 22 and 4.24 and are similar to those used in A-PRO (See Appendix C of [5]) except for the fact that the layers present in A-PRO correspond to single range gates in AEL-PRO (i.e. there are no multi-range-gate layers in AEL-PRO and each range-gate can be considered to be equivalent to an A-PRO layer).

4.3 Detailed Algorithm description

AEL-PRO is based mainly on the Extinction, Backscatter and Depolarization (EBD) component of the A-PRO ATLID processor [5]. However, since ALADIN does not measure the depolarization ratio of the backscatter light, the classification aspects of AEL-PRO are simplified with respect to A-PRO (which uses the depolarization ratio in its classification procedures). Like A-PRO, AEL-PRO uses a two-pass approach for processing both strong features (e.g. clouds) and weak features (e.g. aerosols). Unlike A-PRO (which uses a direct method to determine the weak feature extinction and backscatter fields) AEL-PRO uses the same optimal-estimation forward model approach applied to both passes but performed at different resolutions. Pass-1 of the algorithm is at a horizontal resolution of about 90 km and is applied to cloud-screened averaged ATBs while the Pass-2 is at the highest available resolution (about 3 km) and includes cloudy regions. This process is schematically depicted in Figure 10.

Pass-1 and Pass-2

Perform column-by-column retrievals using Pass-I results

Separate "weak" and "strong" features

Average over "Weak-feature" areas and retrieve "background" aerosol/thin-cloud fields
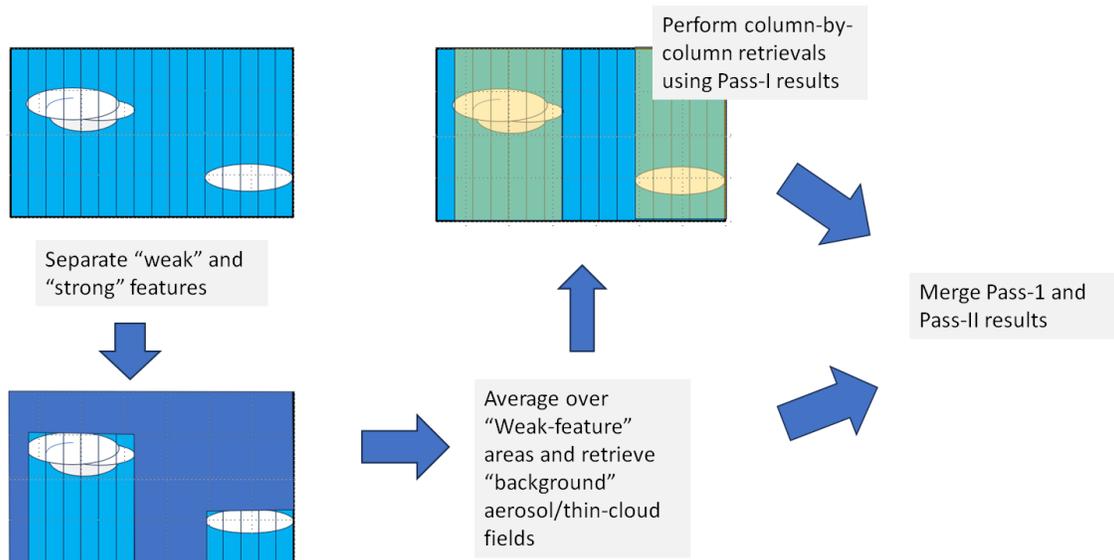
Merge Pass-1 and Pass-II results

Figure 10: Schematic depiction of the AEL-PRO Two-Pass approach and the merging of the observation resolution (aerosol) background state with the measurement level (cloud) information. l
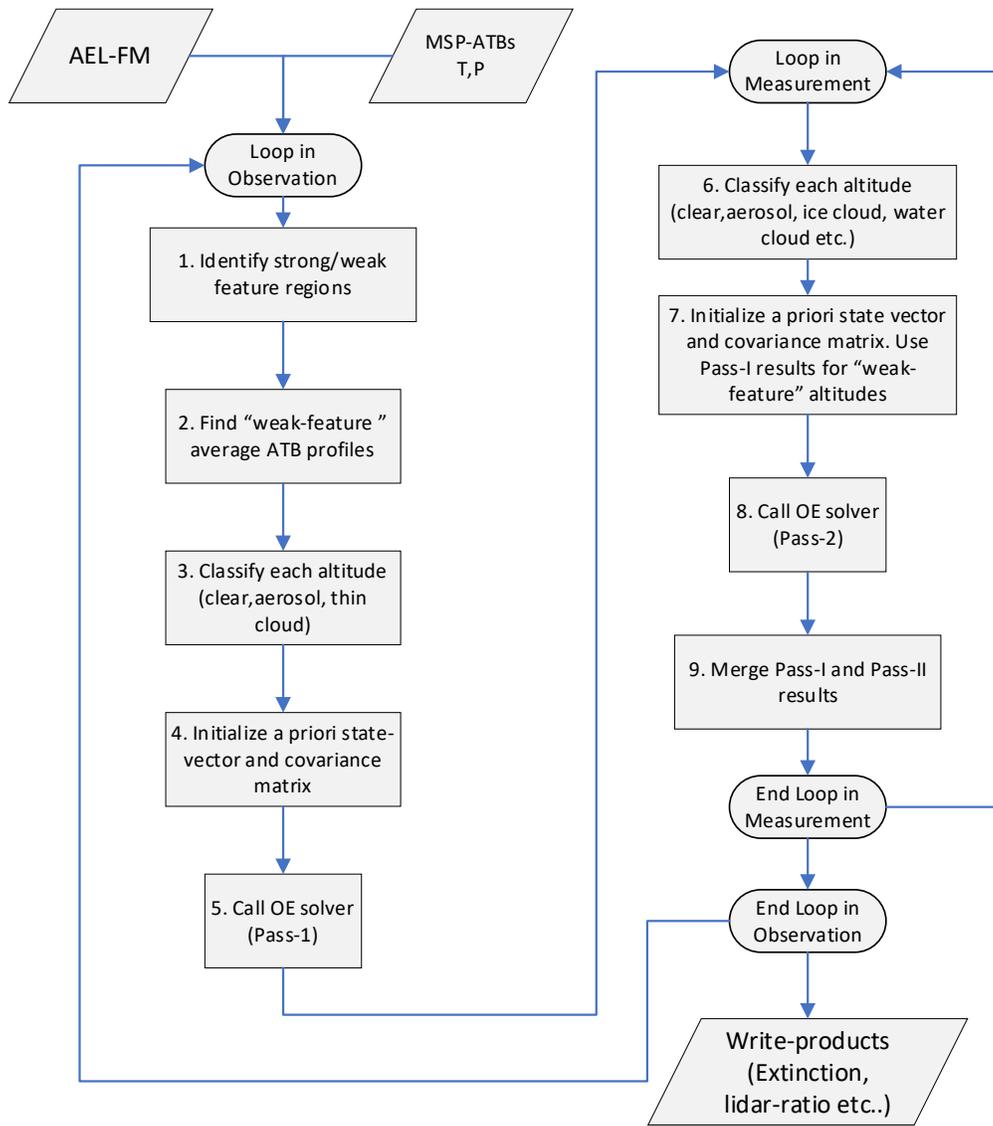
Figure 11: Flow chart of the main AEL-PRO steps.

A high-level flow chart of AEL-PRO is given in Figure 11.

Step 1: A threshold based on AEL-FM together with a threshold based on the lidar-scattering ratio (calculated at this stage directly by dividing the Mie ATB by the Ray ATB) is used to define a ``strong-feature" mask.

Step 2: For each measurement within the observation interval being treated, the highest altitude corresponding to the occurrence of a strong feature is found. Altitude bins above this level are used to determine the "weak-feature" average ATB profiles as well as their uncertainties.

Step 3: The profile is classified according to the following decision tree:
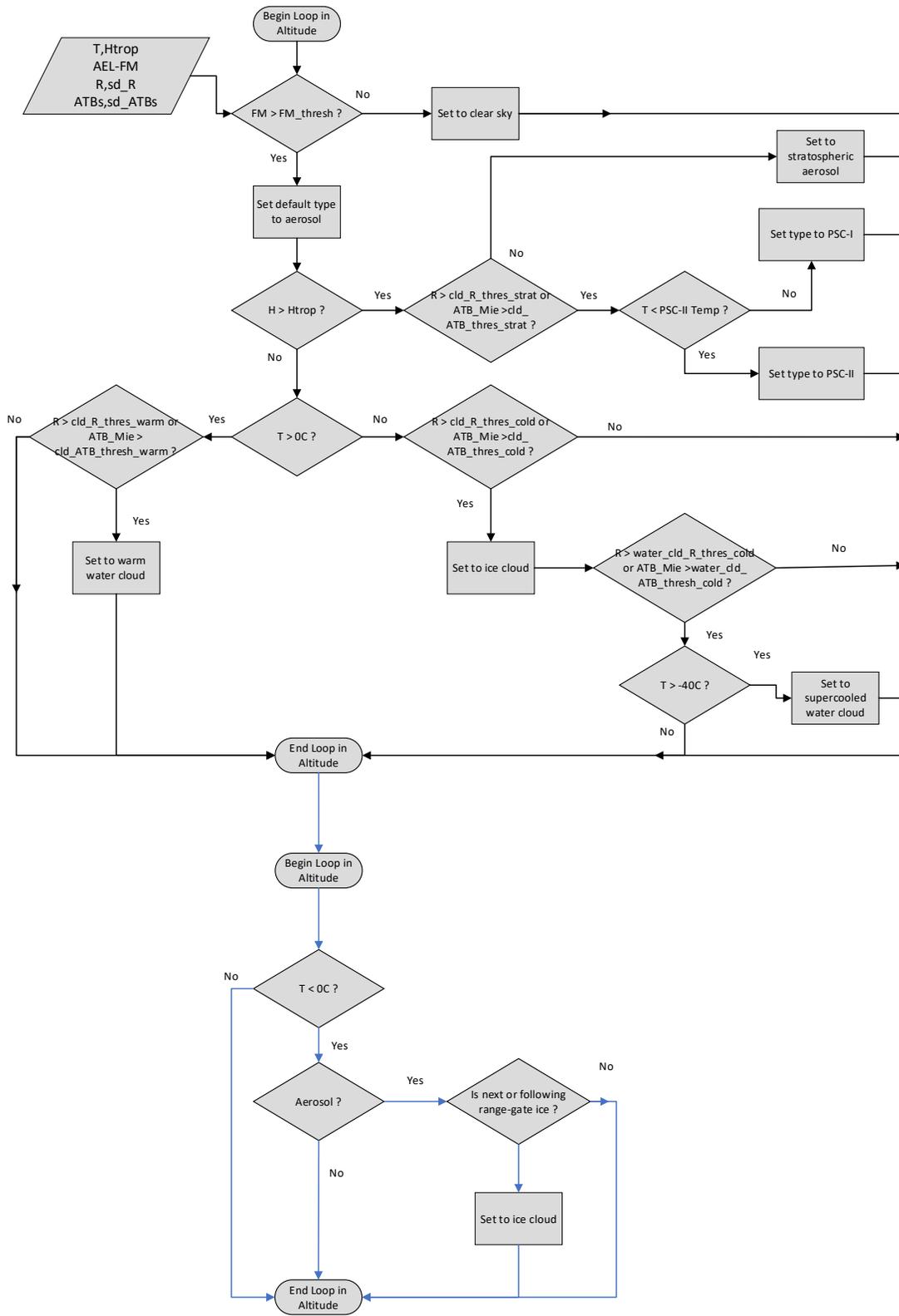

Figure 12: Schematic depiction of the classification procedure.

The inputs are the temperature and the Ray and Mie attenuated backscatters and their errors as well as the corresponding AEL-FM product. The comparisons with the thresholds are all conditioned with respect to the noise levels i.e. the threshold must be exceeded by the appropriate 1-sigma error estimate Here R is the scattering ratio which is computed using the ratio of the Mie and Rayleigh attenuated backscatters.

The tropopause height is computed from the temperature profile in a manner similar to that used in CALIPSO data processing (see page 52 https://gmao.gsfc.nasa.gov/GMAO_products/documents/GEOS-5.1.0_File_Specification.pdf) .

After the first loop in range-gate, a second loop is performed to prevent "aerosol-halos" around ice clouds. If a cold aerosol pixel is adjacent to an ice cloud then it is reclassified as being ice cloud.

Step 4: After the classification has been performed, the a priori values of the lidar-ratio and effective particle sizes and their associated errors are assigned based on the classification. The a priori values of the lidar scattering ratio are supplied as those values appropriate for non-polarized scattering. They are adjusted to take into account the fact that ALADIN uses circular depolarization and only detects the co-polar backscatter i.e.

$$S_{Circ} = S\left(1 + \frac{2\delta_{lin}}{1 - 2\delta_{lin}}\right) \qquad\qquad 4.35$$

where S is the unpolarized lidar-ratio and δ_{lin} is the linear depolarization ratio.

Step 5: The OE solver is now called. The OE cost function is minimized numerical using a using a version of the well-known BFGS Quasi-Newton numerical minimization procedure. The errors as well as the optimal parameters are also calculated in this step.

Step 6: Mirrors step 3 except at the measurement level.

Step 7: The a priori state-vector is initialized here. For weak-features, if they are deemed valid, the OE results from Step 5 are used. Otherwise, the priors are set using the Step 6 classification results.

Step 8: This step mirrors Step 5.

Step 9: The measurement level results are merged with the observation resolution background results from Step 5. This is accomplished by simply replacing the measurement level weak-feature results with the observational level background results.

4.3.1 Configuration parameters

The algorithm configuration parameters for Baseline 16 and 17 are listed below in the following table.

Parameter group	Name	Description	Units	Typical Value
atb_creation	x_talk_LUT	"Look-up table for C1,C2,C3,C4 coefficients	-	-
	ip_mie_1	Left pixel defining the Mie virtual channel on the ACCD (not counting the leading 2 pixels)	Pixel	6
	ip_mie_1	Left pixel defining the Mie virtual channel on the ACCD (not counting the leading 2 pixels)	Pixel	11
	cal_fit_order	Order of polynomial to use in the Rayleigh signal fitting procedure in the calibration routine	-	10
	cal_z_min	Minimum altitude to consider [m] for calibration procedure	m	14000
	cal_z_max	Maximum altitude to consider [m] for calibration procedure	m	30000
	cal_R_min	Minimum scattering ratio to consider for calibration procedure	-	0.8
	cal_R_max	Maximum scattering ratio to consider for calibration procedure	-	1.1
	cal_min_number_of_points	Minimum number of points required to determine the lidar Calibration constant	-	80
	cal_number_of_wrap_points	Number of wrap around points to use in the fitting procedure	-	50
	fraction_pixels_check	Fraction of minimum retrieved Clid profiles, below this fraction an additional check is made	-	0.8
	clid_bias_check	Value by which the Clid polynomial fit is allowed to swing with respect to median before a linear fit is used	-	5.0
	apply_dynamic_EMSR	If set to 1, then determine EMSR (TOBS replacement) using clear-sky values and apply that instead of the values in the L1b input	-	1
	z_trop_lower	Lower limit for tropopause search	m	5000
	z_trop_upper	Upper limit for tropopause search	m	18000
	dz_above_trop_EMSR	Seach window exclusion-zone above trop level for dynamic EMSR determination	m	5000
	dz_below_trop_EMSR	Seach window below trop level for dynamic EMSR determination	m	5000
	z_min_EMSR	Minimum altitude for dynamic EMSR	m	5000

		determination		
	R_thresh_EMSR	"Effective R threshold to be used to determine clear-air EMSR	-	1.2
	apply_excess_noise_factor	If set to 1, then apply the in-situ estimated excess noise factor in the noise calculations, if set to 0 then fix excess_noise_factor_per_pixel to 1	-	1
	detect_missed_hot_pixels	If set to 1, then try to detect and correct for missed Hot Pixels	-	0
	hp_half_window	1/2 window size to use in hot-pixel line detection	Pixel	5
	hp_line_test_thresh	Threshold to use for line detection in HP detection routine	LSB/pixel	4.0
multiple_ scattering	fov_t	Full-angle receiver effective field-of-view	mrads	0.022
	div_laser	Full-angle transmitter divergence	mrads	0.012
Masking	fm_fill_alt	Fill the FM input near the ground: e.g. assume there is always aerosol present	km	5
	fm_thresh	"If AEL-FM is greater than or equal to this threshold then a feature is assumed to be present	-	4
	strong_feature_beta_thresh	Mie ATB values above this level are assumed to be strong features	1/m/sr	1.0e-5
	strong_feature_R_thresh	R (scattering ratio) values above this level are assumed to be strong features	-	2
	strong_feature_FM_thresh	AEL-FM values at or above this level are assumed to be strong features	-	9.0
	R_fill_thresh	If R_profile-sd_R_profile gt R_fill_thresh then AEL-FM is adjusted to be max(5,AEL-FM)	-	1.02
	fm_fill_R_thresh	If R_profile lt 0 and FM lt fm_fill_R_thresh then we assume there is strong attenuation and assume a feature is present	-	7
Classification	sd_ATB_MIE_factor	In simple_class, if (ATB_MIE_profile-sd_ATB_MIE_profile)/sd_ATB_MIE_factor then there is at least aerosol there	-	2
	cld_R_thresh_warm	In simple_class R-sd_R > 2.0 then there may be a cloud of some sort (applicable for T gt 0 C)	-	2
	cld_R_thresh_cold	In simple_class R-sd_R > 2.0 then there may be a cloud of some sort (applicable for T lt 0 C)	-	1.05
	cld_R_thresh_strat	In simple_class R-sd_R > 2.0 then there may be a cloud of some sort (applicable for stratosphere)	-	1.02
	cld_ATB_thresh_warm	In simple_class ATB_MIE_profile gt	1/m/sr	5e-6

		cld_ATB_thresh then there may be a cloud of some sort (applicable for T gt 0 C)		
	cld_ATB_thresh_cold	In simple_class ATB_MIE_profile lt cld_ATB_thresh then there may be a cloud of some sort (applicable for T lt 0 C)	1/m/sr	5e-7
	cld_ATB_thresh_strat	In simple_class ATB_MIE_profile gt cld_ATB_thresh then there may be a cloud of some sort (applicable for stratosphere)	1/m/sr	2e-7
	temperature_strat_1	Stratospheric clouds with lower temperatures than this are assumed to be type 1 PSCs	K	
	temperature_strat_2	"Stratospheric clouds with lower temperatures than this are assumed to be type 2 PSCs	K	
	water_cld_R_thresh_cold	In simple class R-sd_R gt water_cld_R_thresh then there is a water cloud for sure (applicable for T lt 0 C ..supercooled clouds)	-	15
	water_cld_ATB_thresh_cold	In simple class ATB_MIE_profile gt water_cld_ATB_thresh then there is a water cloud (applicable for T lt 0 C ..supercooled clouds)	1e-6	1/m/sr
Priors Note: Depol. values are those appropriate for linear depolarization. The S values are those appropriate for unpolarised scattering (The adjustment to ALADIN's configuration is automatically performed within the algorithm)	dC_lid_adjust_factor	Ad-Hoc multiplicative adjustment of lidar calibration constant	-	1
	water_cld_depol	A priori value of water cld lin depol (due to MS)	-	0.1
	water_cld_S	A priori value of water S	sr	18
	water_cld_d_S	A priori value of water S relative uncertainty	-	0.2
	water_cld_Ra	A priori value of water effective area radius	microns	5
	water_cld_d_Ra	A priori value of water Ra relative uncertainity	-	1
	water_cld_eta_o	Fixed value of eta for water	-	0.375
	ice_cld_depol	A priori value of cld lin depol	-	0.3
	ice_cld_S	A priori value of ice S	sr	25.0
	ice_cld_d_S	A priori value of ice S relative uncertainty	-	0.4
	ice_cld_Ra	A priori value of ice effective area radius	microns	25.0
	ice_cld_d_Ra	A priori value of ice Ra relative uncertainty	-	2.0
	ice_cld_eta_o	Fixed value of eta for ice	-	0.375
	strat_1_cld_depol	A priori value of strat cld type-1 cld lin depol	-	0.1
	strat_1_cld_S	A priori value of strat cld type-1 S	sr	25.0
	strat_1_cld_d_S	A priori value of strat cld type-1 S relative	-	0.4

		uncertainty		
	strat_1_cld_Ra	A priori value of strat cld type-1 effective area radius	microns	1.0
	strat_1_cld_d_Ra	A priori value of strat cld type-1 Ra relative uncertainty	-	2.0
	strat_1_cld_eta_o	Fixed value of eta for strat cld type-1	-	0.375
	strat_1_cld_depol	A priori value of strat cld type-1 cld lin depol	-	0.3
	strat_1_cld_S	A priori value of strat cld type-1 S	sr	25.0
	strat_1_cld_d_S	A priori value of strat cld type-1 S relative uncertainty	-	0.4
	strat_1_cld_Ra	A priori value of strat cld type-1 effective area radius	microns	10.0
	strat_1_cld_d_Ra	A priori value of strat cld type-1 Ra relative uncertainty	-	2.0
	strat_1_cld_eta_o	Fixed value of eta for strat cld type-1	-	0.375
	aerosol_depol	A priori value of aerosol lin depol	-	0.15
	aerosol_S	A priori value of aerosol S	sr	60.0
	aerosol_d_S	A priori value of aerosol S relative uncertainty	-	1.5
	aerosol_Ra	A priori value of aerosol effective area radius	microns	0.2
	aerosol_d_Ra	A priori value of aerosol Ra relative uncertainty	-	0.1
	aerosol_eta_o	Fixed value of eta for aerosol	-	0.375
	strat_aerosol_depol	A priori value of strat aerosol lin depol	-	0.01
	strat_aerosol_S	A priori value of strat aerosol S	sr	40.0
	strat_aerosol_d_S	A priori value of strat aerosol S relative uncertainty	-	1.0
	strat_aerosol_Ra	A priori value of strat aerosol effective area radius	microns	0.2
	strat_aerosol_d_Ra	A priori value of strat aerosol Ra relative uncertainty	-	0.1
	strat_aerosol_eta_o	Fixed value of eta for strat aerosol	-	0.375
dfpmin (numerical optimization solver)	maximum_step	Maximum allowed step size in line minimization search step: NOTE: VALUES ABOVE 0.5 OR SO CAUSE NUMERICAL PROBLEMS !	-	0.2
	gradient_tolerance	Check for convergence on gradient norm	-	1.0e-5
	relative_tolerance	Check for convergence on successive iterations	-	1.0e-5
	maximum_iterations	Maximum number of iterations in OE cost function minimization	-	1000
	high_chisq_qa_flag_threshold	If Chisq/dof resulting from the OE inversion is greater than this threhold then the corresponding QA bit is set	-	30.0

4.3.2   Output parameters.

Between baselines 16 and 17 a harmonization activity was conducted with respect to the product names and units. Also, more quality flags and information was included in the product. Here the output products are separately listed for both baseline 16 and 17.

4.3.2.1   Baseline 16
The AEL-PRO L2 product variables for Baseline 16 are outlined below.

Parameter	Description	Data Type	Dimension(s)
extinction [Mm-1]	Particle extinction	8 byte float	(measurement, mie_l1b_mid_bin)
error_extinction[Mm-1]	Error Estimate	8 byte float	(measurement, mie_l1b_mid_bin)
lr[sr]	Particle extinction-to-backscatter ratio	8 byte float	(measurement, mie_l1b_mid_bin)
error_lr [sr]		8 byte float	(measurement, mie_l1b_mid_bin)
particle_effective_area_radius [microns]		8 byte float	(measurement, mie_l1b_mid_bin)
error_particle_effective_area_radius [microns]		8 byte float	(measurement, mie_l1b_mid_bin)
tropopause_altitude [m]		8 byte float	measurement
starting_chisq_value_pass1		8 byte float	observation
ending_chisq_value_pass1		8 byte float	observation
starting_chisq_value_pass2		8 byte float	measurement
ending_chisq_value_pass2		8 byte float	measurement
number_of_iterationspass1		4 byte integer	measurement
number_of_iterations_pass2		4 byte integer	observation
classification		4 byte integer	(measurement, mie_l1b_mid_bin)
quality_index		4 byte integer (byte flag)	(measurement)

The AEL-PRO L2 product variables for Baseline 17 are outlined below.

Parameter	Description	Data Type	Dimension(s)
extinction [Mm-1]	Particle extinction	8 byte float	(measurement, mie_l1b_mid_bin)
extinction_variances[m-2]	Estimated error variance	8 byte float	(measurement, mie_l1b_mid_bin)
backscatter [Mm-1sr-1]	Particle Backscatter	8 byte float	(measurement, mie_l1b_mid_bin)
backscatter_variances[m-2sr-2]		8 byte float	(measurement, mie_l1b_mid_bin)
lr[sr]	Particle extinction-to-backscatter ratio	8 byte float	(measurement, mie_l1b_mid_bin)
lr_variances [sr2]		8 byte float	(measurement, mie_l1b_mid_bin)
ber[sr-1]	Particle backscatter-to-extinction ratio	8 byte float	(measurement, mie_l1b_mid_bin)
ber_variances [sr-2]		8 byte float	(measurement, mie_l1b_mid_bin)
sr[-]	Scattering Ratio	8 byte float	(measurement, mie_l1b_mid_bin)
sr_variances [-]		8 byte float	(measurement, mie_l1b_mid_bin)
particle_effective_area_radius [microns2]		8 byte float	(measurement, mie_l1b_mid_bin)
particle_effective_area_radius_variances [microns2]		8 byte float	(measurement, mie_l1b_mid_bin)
tropopause_altitude [m]		8 byte float	measurement
starting_chisq_value_pass1		8 byte float	observation
ending_chisq_value_pass1		8 byte float	observation
starting_chisq_value_pass2		8 byte float	measurement
ending_chisq_value_pass2		8 byte float	measurement
number_of_iterationspass1		4 byte integer	measurement
number_of_iterations_pass2		4 byte integer	observation

classification		4 byte integer	(measurement, mie_l1b_mid_bin)
quality index		4 byte integer (byte flag)	(measurement)

4.3.2.3   Classification Description

0	Clear sky tropospheric
100	Clear sky and used in Pass1 (tropospheric)
200	Clear sky and used in Pass1 (stratospheric)
1	Water cloud
101	Water cloud and used in Pass1 (tropospheric)
201	Water cloud and used in Pass1 (stratospheric)
2	Ice cloud tropospheric
102	Ice cloud tropospheric and used in Pass 1
3	Tropospheric aerosol
103	Tropospheric aerosol and used in Pass 1
13	Stratospheric Aerosol
213	Stratospheric Aerosol and used in Pass1
11	PSC-Type 1
211	PSC-Type 1 and used in Pass1
12	PSC-Type 2
212	PSC-Type 2 and used in Pass 1
1000 >= Index <2000	Attenuated and is tropospheric
2000 >= Index <10000	Attenuated and is stratospheric
Index > 10000	Surface

4.3.2.4 Quality_index definition

Bit	QA condition
0	No retrieval (e.g. no valid data)
1	High value of Chi_sq in optimal estimation minimization (obs level: Pass 1)
2	High value of Chi_sq in optimal estimation minimization (measurement level: Pass 2)
3	Maximum number of iterations reached (obs level: Pass 1)
4	Maximum number of iterations reached (measurement level: Pass 2)
5	(very-)Low SNR in retrieved extinction (ADDED in BASELINE 17)

5 Examples

Example Mie and Rayleigh attenuated backscatter fields for one orbit are shown in Figure 13. For context, both the MSP-only derived fields and the fields retrieved using both the MSP together with the RSP are shown. The MSP+RSP ATBs are produced as part of the SCA retrieval [4]. Here it can be seen that the MSP-only Mie ATBs appears to be more precise (i.e. less noisy) than their SCA counterparts while the MSP-only Rayleigh attenuated backscatter appears to share a similar level of precision with their SCA counterparts. This can be also seen in the example profiles shown in Figure 14. That the MSP-only crosstalk corrected ATBs have similar or even better levels of precision even though no photons collected by the RSP are used is a consequence of the fact that the set of cross-talk coefficients corresponding to the full MSP+RSP system is unfavourable when it comes to SNR inflation (see Appendix-A of [3]). The crosstalk and calibration procedures associated with the MSP-only approach are also simpler and easier to characterize than the more complicated (and more uncertain) full MSP+RSP system
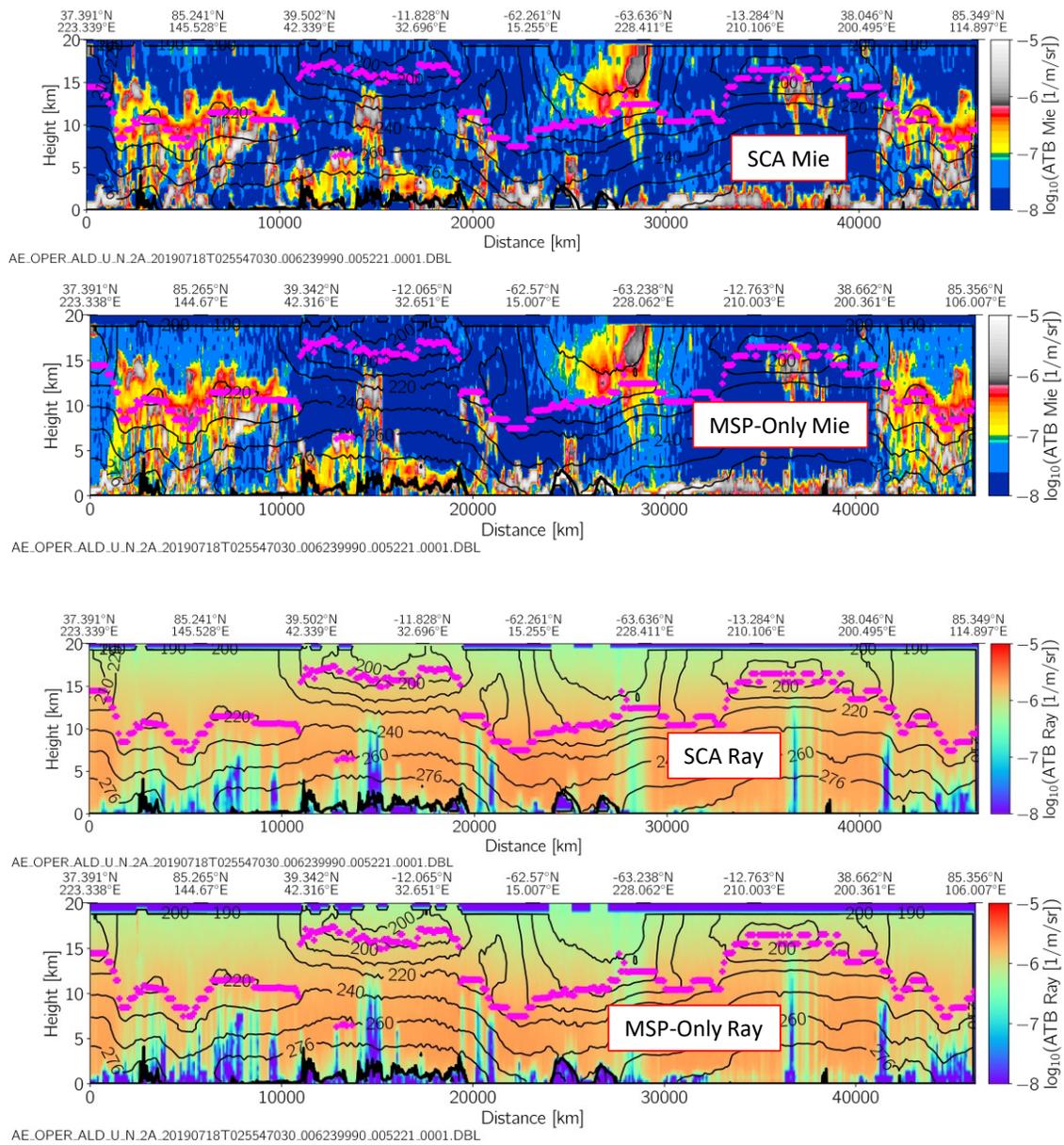
Figure 13: Attenuated backscatter fields for orbit number 5221 (20190718) produced using both the SCA method combining the RSP and MSP and the MSP only method described in this paper.
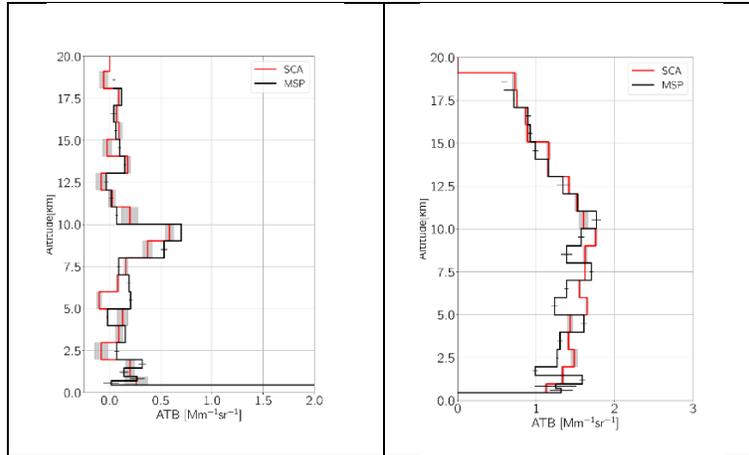
Figure 14: Profiles of Mie ATBs corresponding to averages from observations 51-54 (along-track from about 4521-4786 km) in Figure 13. The Grey shaded area represents the standard deviation of the SCA ATBs while the Black horizontal error bars represent the standard deviation of the MSP-only ATBs.

AEL-FM provides an index that corresponds to the probability of a target being present and identifies clear-air as well as attenuated regions. The product is reported at the native resolution of the input ATBs (about 3 km horizontal resolution). The procedure combines a number of techniques applied to the Mie ATBs including a hybrid-median edge preserving filter (in order to identify ``strong'' returns) and iterative adaptive Gaussian smoothing (which is targeted towards weaker extended features). The pure Rayleigh profiles are also used, mainly to determine attenuated regions. AEL-PRO is closely based on A-PRO. Both A-FM and AEL-FM are described in more detail in [5]. The AEL-FM results are used to as input to the AEL-PRO algorithm.

 An example of the AEL-FM "featuremask" product is shown in Figure 15. Here a number of interesting features are see to be present. High altitude clouds associated with tropopause heights above 16 km are seen to be present in the tropics. In the summer mid- and high-latitudes aerosol (forest file smoke) is present in the lower stratosphere and upper-troposphere. In the winter Antarctic stratosphere PSCs are seen to be present.
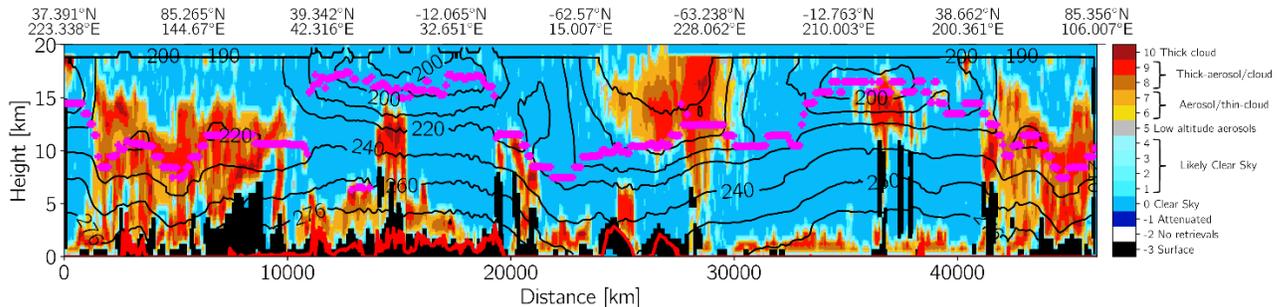
AE_OPER_ALD_U_N_2A_20190718T025547030_006239990_005221_0001.DBL

Figure 15: Example AEL-FM feature-mask index corresponding to the data shown in Figure 13. The Magenta symbols mark the Tropopause level while the lower Red line represents the surface elevation. The thin-Black contours show the atmospheric temperature. The temperatures are taken from the AUX-MET data. The DEM is taken from the L1B product.

Example extinction and lidar ratio retrievals corresponding to the same orbit as for Figure 13. Both results from the SCA mid-bin algorithm [3] and AEL-PRO results are shown in Figure 16. Here data where the estimated SNR is greater than 1 was aggregated to a resolution of 0.5 km (vertically) by 90 km (horizontally). There is a large degree of correspondence between the SCA and AEL-PRO results, however, the AEL-PRO results are more precise and sensitive, particularly with regards to the lidar-ratio retrievals. In particular, the SCA approach tends to only produce usable estimates of the lidar-ratio for extinction values above 0.05 km^{-1} while AEL-PRO supplied usable estimates of the lidar ratio for extinctions on the order of 0.002 km^{-1}. The more precise nature of the AEL-PRO results can again be seen in Figure 17. Here the difference in precision (noise) is evident between the SCA and AEL-PRO results. This difference in precision is due to the combined effects of both more precise attenuated backscatter profiles estimates as discussed earlier and the regularization (or stabilization) effect afforded by the optimal estimation approach used by AEL-PRO. It can also be seen in Figure 17 that the resolution of the AEL-PRO products is finer than the SCA products at lower altitudes. This is a direct consequence of the need to create a merged grid to combine the MSP and RSP signals used by the SCA process. The AEL-PRO approach uses the MSP vertical grid which tends to have a finer resolution than the RSP vertical grid.
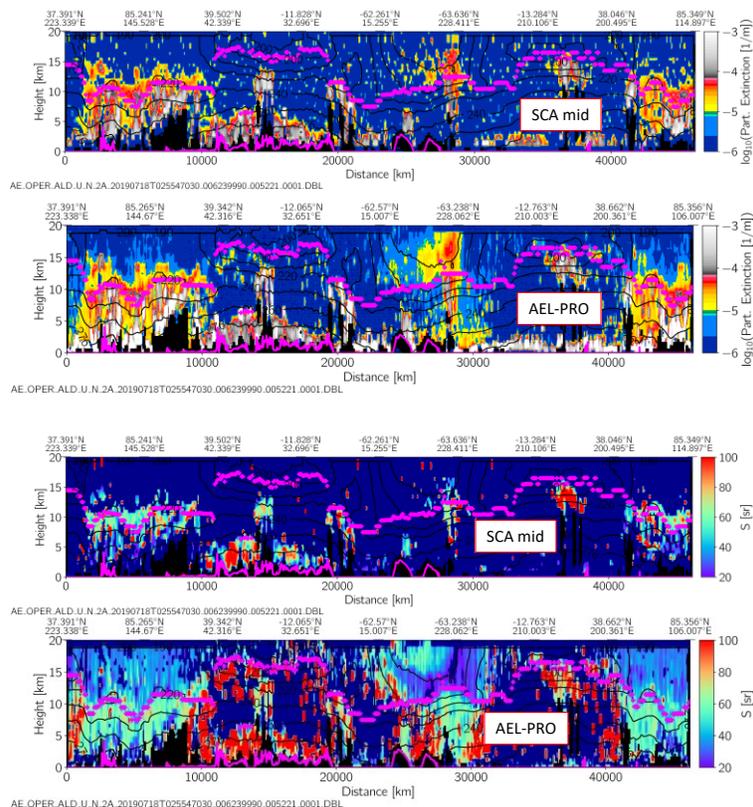

Figure 16: SCA and AEL-PRO (both baseline 2A16) retrievals o particulate extinction and lidar-ratio for the same orbit as previously presented. The black areas correspond to attenuated or below surface

altitudes. The lower Magenta line represents the surface elevation, the back contour lines the temperature, and the Magenta symbols the tropopause height.
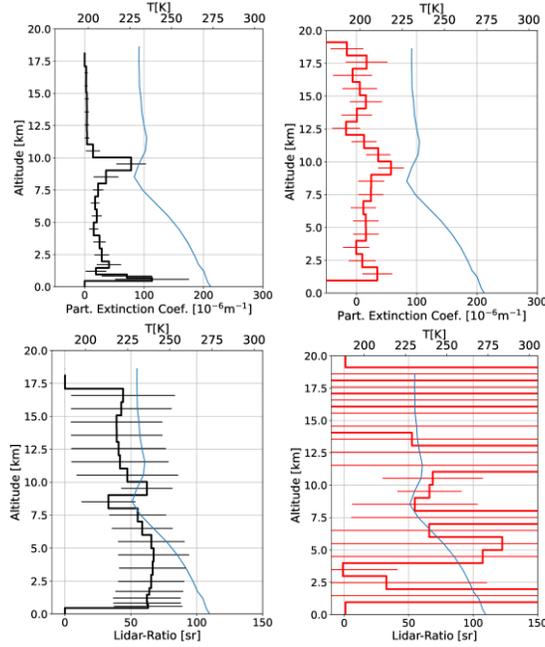

Figure 17: AEL-PRO (Black Left) and SCA Mid (Red Right) profiles of retrieved extinction and lidar-ratio for observation 51 (approx.. 76°N,201°E) for the same orbit as presented earlier. The Blue line is the AUX-MET temperature profile (upper x-axis scale).
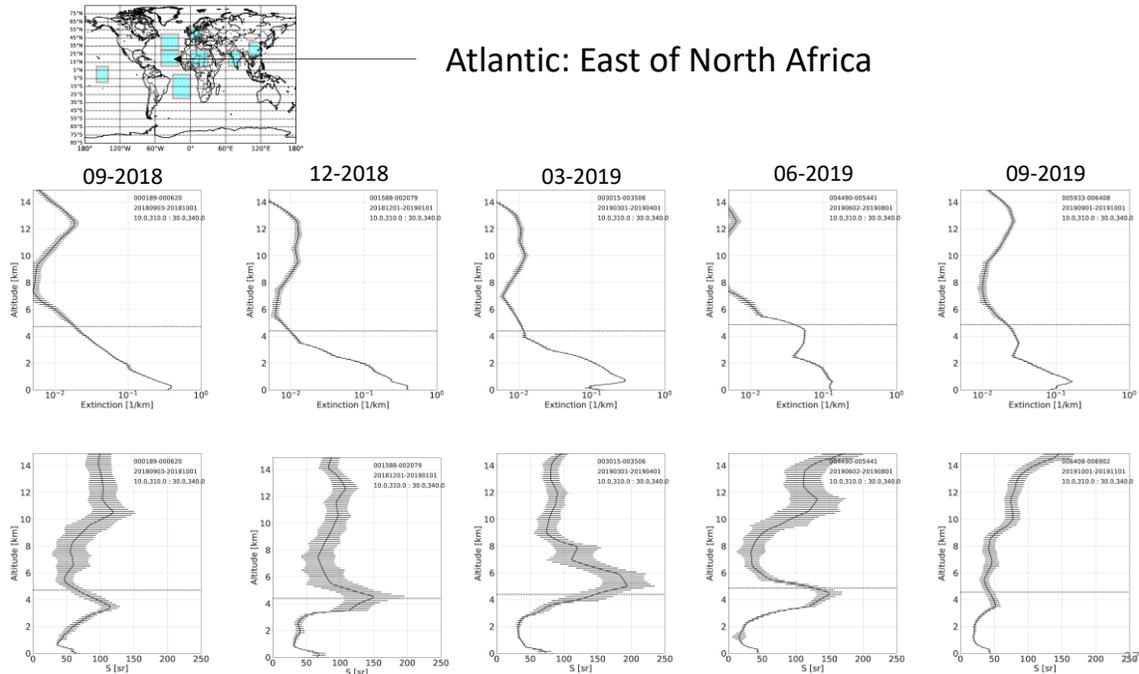

Figure 18: Average profiles of extinction and lidar ratio for the lat-lon area indicated in the figure for the labelled months. Here only clear-sky profiles have been used. The dotted horizontal line demarks the mean zero degrees Celsius level.

AEL-PRO data is of sufficient accuracy and precision to support longer-term studies. An example of the type of studies possible using AEL-PRO data is shown in in Figure 18. In Figure 18, a series of monthly averaged clear-sky profiles of aerosol extinction and lidar-ratio for an aerosol of the North-Atlantic Ocean east of Africa is shown. Aerosol scattering ratios above about 100 Sr are associated with desert-dust that is being transported from Africa. The yearly cycle of this transport is clearly visible in the time series.

6   Summary

The application of EarthCARE inspired approaches to ALADIN data has yielded benefits for both missions. Though A-FM and A-PRO had been developed using detailed realistic simulated ATLID data [6], the successful application of A-FM and A-PRO methods to real data has boosted confidence in the fundamental approaches as well as building practical experience that will prove useful when actual ATLID observations will be available.

From the ALADIN perspective, the work summarized in this chapter has yielded new insights into the impact of crosstalk on the precision and accuracy of the pure Mie and Rayleigh ATBs and has led to a successful alternative method for producing precise cross-talk corrected ATB profiles while correcting for the non-uniform spectral response of the MSP unit. Furthermore, AEL-FM and AEL-PRO have been successfully implemented into the operational Aeolus L2A processor and will yield a useful record of global aerosol extinction and lidar-ratio profiles. The examples shown here are not extensive, however, more evaluation work has been conducted and is been presented in "Evaluation of Aeolus feature mask and particle extinction profile products using CALIPSO data", Ping Wang, David Patrick Donovan, Gerd-Jan van Zadelhoff, Jos de Kloe, Dorit Huber, and Katja Reissig, submitted to AMTD.

Appendix A Crosstalk correction error magnification

Here we seek a relationship between the SNR of the crosstalk corrected signals and the SNR ratio of the signals that would be detected if no crosstalk was present in the system. We start with the basic equations,

$$B_{R,o} = K_{Ray}\left(C_1 B_R + C_2 B_M\right) \qquad (1.1)$$

and

$$B_{M,o} = K_{Mie}\left(C_4 B_R + C_3 B_M\right), \qquad (1.2)$$

where B_M and B_M are the intrinsic atmospheric values (e.g. true values), $B_{M,o}$ and $B_{M,o}$ are the corresponding detected (equivalent-)photocounts (K_{Ray} and K_{Ray} are defined accordingly). The solution to the above system can be written as:

$$B_R = \frac{-\left(C_3' B_{R,o} - C_2' B_{M,o}\right)}{C_2' C_4' - C_1' C_3'} \qquad (1.3)$$

$$B_M = \frac{C_4' B_{R,o} - C_1' B_{M,o}}{C_2' C_4' - C_1' C_3'} \quad \text{where } C'_1 = K_{Ray} C_1 \text{ etc.} \quad (1.4).$$

Applying standard error propagation to the above relationships leads to

$$SNR_{B_M} \equiv \frac{B_M}{\sigma_M} = \frac{\left|C_4' B_{R,o} - C_1' B_{M,o}\right|}{\left(\left(C_4' \sigma_{R,o}\right)^2 + \left(C_1' \sigma_{M,o}\right)^2\right)^{1/2}} \text{ (1.5) and } SNR_{B_R} \equiv \frac{B_R}{\sigma_R} = \frac{\left|C_2' B_{M,o} - C_3' B_{R,o}\right|}{\left(\left(C_2' \sigma_{M,o}\right)^2 + \left(C_3' \sigma_{R,o}\right)^2\right)^{1/2}} \text{ (1.6).}$$

Now, if one assumes that the noise is dominated by shot-noise (Poisson statistics) and the background and dark-count rates can be ignored, then one has $\sigma_{B_{o,M}}^2 = B_{o,M}$ (1.7) and $\sigma_{B_{o,R}}^2 = B_{o,R}$ (1.8). Substitution into (1.5) and (1.6) then yields

$$SNR_{B_M} = \frac{B_M}{\sigma_M} = \frac{\left|C_4' B_{R,o} - C_1' B_{M,o}\right|}{\left(C_4'^2 B_{R,o} + C_1'^2 B_{M,o}\right)^{1/2}} \text{ (1.9) and } SNR_{B_R} = \frac{B_R}{\sigma_R} = \frac{\left|C_2' B_{M,o} - C_3' B_{R,o}\right|}{\left(C_2'^2 B_{R,o} + C_3'^2 B_{R,o}\right)^{1/2}}. \text{ (1.10)}$$

Using the fact that $B_R(R-1) = B_M$ (1.11) together with (1.1) and (1.2) we can write

$$B_{o,R} = (C_1' + C_2'(R-1))B_R, \qquad (1.12)$$

$$B_{o,M} = (C_3' + C_4'(R-1)^{-1})B_M, \qquad (1.13)$$

$$B_{o,R} = (C_2' + C_1'(R-1)^{-1})B_M \qquad (1.14) \quad \text{and}$$

$$B_{o,M} = (C_4' + C_3'(R-1))B_R. \qquad (1.15)$$

Starting with Eq. (1.9) and substituting for $B_{o,M}$ using Eq. (1.13) and for $B_{o,M}$ using Eq. (1.14) respectively, then yields

$$SNR_{B_M} = \frac{\left| C_4'C_2' - C_1'C_3' \right|}{\left[C_4'^2 C_2' + C_1'^2 C_3' + (R-1)^{-1} \left(C_1'C_4'^2 + C_1'^2 C_4' \right) \right]^{1/2}} B_M^{1/2} \quad , \qquad (1.16)$$

While, in a similar fashion, using Eqns. (1.10), (1.12) and (1.15) yields

$$SNR_{B_R} = \frac{\left| C_4'C_2' - C_1'C_3' \right|}{\left[C_2'^2 C_4' + C_3'^2 C_1' + (R-1) \left(C_2'C_3'^2 + C_2'^2 C_3' \right) \right]^{1/2}} B_R^{1/2}. \qquad (1.17)$$

For a system without cross-talk present, $C_2' = C_4' = 0$ and we can write, using (1.16) and (1.17)

$$SNR_{B_R,ideal} = (C_1'B_R)^{1/2} \qquad (1.18)$$

and

$$SNR_{B_M,ideal} = (C_3'B_M)^{1/2} \qquad (1.19)$$

Then using Eqns. (1.16) -- (1.19) we have the following relationships between the SNR of the cross-talk corrected signals and the SNR ratio of the signals that would be detected if no cross-talk was present in the system.

$$F_{N_M}(R) \equiv \frac{SNR_{B_M,Ideal}}{SNR_{B_M}} = C_3'^{1/2} \frac{\left[C_4'^2 C_2' + C_1'^2 C_3' + (R-1)^{-1} \left(C_1'C_4'^2 + C_1'^2 C_4' \right) \right]^{1/2}}{\left| C_4'C_2' - C_1'C_3' \right|} \qquad (1.20)$$

and

$$F_{N_R}(R) \equiv \frac{SNR_{B_R,Ideal}}{SNR_{B_R}} = C_1'^{1/2} \frac{\left[C_2'^2 C_4' + C_3'^2 C_1' + (R-1) \left(C_2'C_3'^2 + C_2'^2 C_3' \right) \right]^{1/2}}{\left| C_4'C_2' - C_1'C_3' \right|} \quad . \qquad (1.21)$$

Appendix B   Calculation of the Fizeau Response

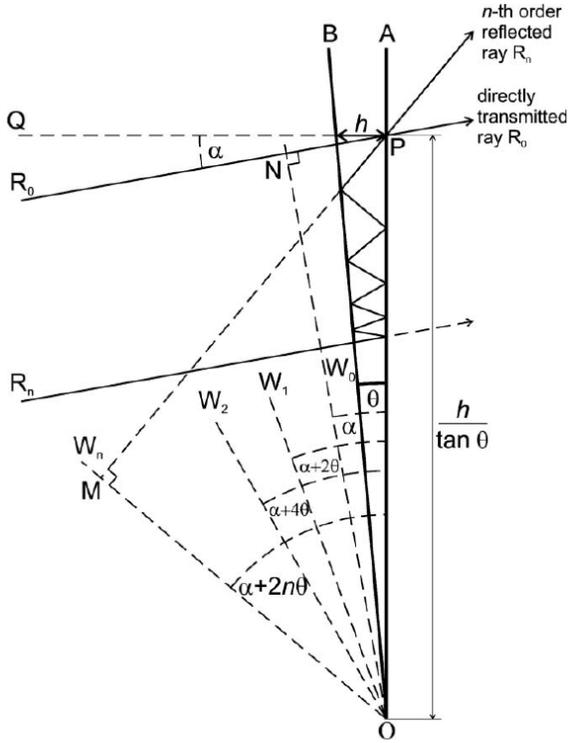The model of the Fizeau spectrometer transfer function is adapted from [10]

FIG. 2. Cross sectional diagram of a Fizeau interferometer. Parallel rays illuminate the interferometer plate OA at an angle α. Both the directly transmitted ray R_0 and a ray reflected n times (on each plate) R_n illuminate the point P of the interferometer plate OA. Wavefronts for various orders of reflection W_0, W_1, W_2, W_n are shown. The phase difference between rays is determined by the wavelength of the light and the path difference. The plate separation at P is h.
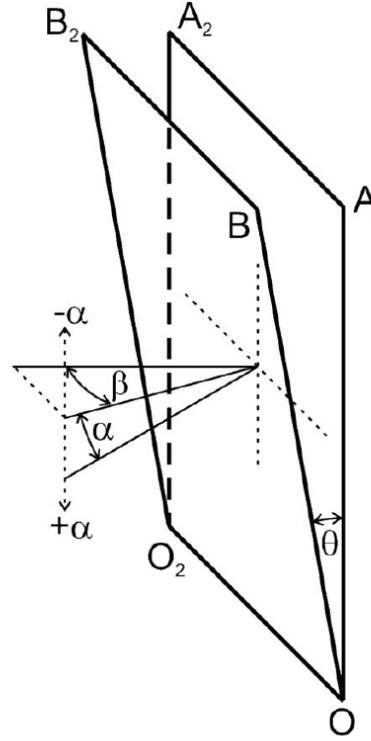
FIG. 3. Resolution of a ray of light into components in the wedge plane and in a plane perpendicular to the wedge plane. The angle α gives the inclination of the ray in the wedge plane relative to a normal to the second interferometer plate and the angle β the inclination of the ray in the plane perpendicular to the wedge plane relative to this normal. The phase shift due to the angle β is the same as that for $-\beta$. In contrast, the angle α *has a different effect* on the phase shift from the angle $-\alpha$ of the same absolute value.

Figure 19: Fizeau model and geometry taken from Novak O. et al [10].,

Referring to Figure 19, the intensity profile along the wedge is given by the following expression

$$I(\lambda, A_o, R, T, \theta, h, \alpha, \beta) = A_o^2 T^2 \left[\left(\sum_{n=0}^{\infty} R^n \cos(\delta_n) \right)^2 + \left(\sum_{n=0}^{\infty} R^n \sin(\delta_n) \right)^2 \right],$$ (1.22)

where the phase-difference for the nth reflection is given by

$$\delta_n = \frac{2\pi}{\lambda} \frac{h}{\tan(\theta)} \left[\sin(\alpha + 2n\theta) - \sin(\alpha) \right].$$ (1.23)

The plate separations (h) corresponding to the fringe maxima for each order (p) are given by

$$h_p = p\frac{\lambda}{2} \tag{1.24}$$

and the separation of the fringes along the plate is given by

$$d = \frac{\lambda^2}{2\tan(\theta)}. \tag{1.25}$$

Thus, using the above relationships the plate separation is determined by the free-spectral-range (FSR) according to

$$FSR = \frac{\lambda^2}{2h}. \tag{1.26}$$

The calculation of the response as a function of wavelength and position across the ACCD rows proceeds as follows (distances and wavelengths all in nm):

1. Find h at the at the centre of the model ACCD: $h = \lambda_c / FSR / 2$ (λ_c is the nominal laser centre wavelength)
2. Find the fringe order (p) associated with this location: $p = \text{int}(2h / \lambda_c)$
3. Find h_center so that the peak corresponding to order p is at the centre of the array: $h_c = p\lambda_c / 2$
4. Find the distance on the array corresponding to one FSR: $dd_{FSR} = \lambda_c / 2 / \tan(\theta)$
5. Create the wavelength array: $\lambda[:] = [\lambda_c - \Delta\lambda, ..., \lambda_c + \Delta\lambda]$ (linearly spaced with nlam components).
6. Create the h array: $h[:] = [h_c - ddp \times dd_{FSR} \times \tan(\theta), ..., +ddp \times dd_{FSR} \times \tan(\theta)]$ (linearly spaced with nd components). ddp is the fraction of a FSR covered by the ACCD.
7. Calculate the position on the ACCD corresponding to the wavelength of peak transmission: $\Delta\lambda_{ACCD}[:] = [-ddp \times FSR, ..., +ddp \times FSR]$ (linearly spaced with nd components).
8. Begin Loop in the elements of $\lambda[i_l]$:
    a. Begin loop in the elements of $h[j_h]$:
        i. Calculate the Transmission:
        $$T[i_l, j_h] = A_o^2 T^2 \sum_0^n \left[\left(R^n \cos(\delta[i_l, j_h, n]) \right)^2 + \left(R^n \sin(\delta[i_l, j_h, n]) \right)^2 \right] \text{ where}$$
        $$\delta[i_l, j_h, n] = 2\pi / \lambda[i_l] \times h[j_h] / \tan(\theta) \times [\sin(\alpha + 2n\theta) - \sin(\alpha)]$$

       b. End loop in the elements of $h[j_h]$:

9. End Loop in the elements of $\lambda[i_l]$:

10. In order to account for imperfections (surface roughness, imperfect collimation etc.) the response is convolved with a top-hat function (i.e. blurred) of a specified width (Blur).

The following parameters have been used in the calculation of the MSP response.

Parameter	Units	Value	Remarks
R	[-]	0.817	R and Blur were tuned to match the expected FWHM of 0.067 pm and a mean response to a Mie spectral signal of 0.135.
T	[-]	0.183	Assumed to be equal to (1-R) (i.e. no absorption losses)
FSR	nm	0.92e-3	
θ	Micro-rads	4.77	
Blur	pm	0.03	R and Blur were tuned to match the expected FWHM of 0.067 pm and a mean response to a Mie spectral signal of 0.135.
α	mrads	0.0	

Once the 2D (wavelength x h) transmission of the MSP has been calculated the appropriate crosstalk coefficients can be derived.

The calculation of the crosstalk coefficients proceeds as follows. For the C_2 and C_3 terms, first the per-column ACCD pixel response to the nominal laser spectra is calculated. Then, the response integrated over the appropriate ACCD columns (the virtual Mie and Rayleigh channels) is found. For the C_1 and C_4 terms a similar procedure is followed, however, since the calculation depends on the pressure and temperature dependent R-B line shape, the calculation includes a parameterized estimate of the R-B line-shape.

For the C_2 and C_3 terms we have:

1. For a given assumed laser line width (FWHM), the laser line shape is calculated assuming a Gaussian spectral profile i.e.

    a. $FWHM_{cm^{-1}} = FWHM_{MHz} \times 10^6 / (c*100)$

    b. $\sigma_{laser} = FWHM_{cm^{-1}} / 2.355$

    c. Create the wavenumber array (in inverse cm) corresponding to the wavelength array used in calculating the Fizeau response: $\delta v[:] = 10^7 \left[\dfrac{1}{\lambda[:]} - \dfrac{1}{\lambda_c} \right]$

2. Calculate the laser (Mie) spectral profile: $I_{laser}(\delta v[:]) = \dfrac{1}{\sqrt{2\pi}} \dfrac{1}{\sigma_{laser}} \exp\left[\dfrac{-\delta v[:]^2}{2\sigma_{laser}{}^2}\right]$

3. Calculate the Mie response as a function of wedge-position:

$$R_M[j_h] = \sum_{i_l} T[i_l, j_h] I_{laser}[\delta v[i_l]] \Delta v_{i_l}$$

4. Calculate the fraction of the Mie response for the virtual Mie channel: $C_3 = \sum_{jp_1}^{jp_2} R_M[j_h]/nd$.

   Here jp_1 and jp_2 are the indices of the array $\Delta\lambda_{ACCD}[:]$ corresponding to the lower and upper limits respectively of the virtual Mie region.

5. Calculate the fraction of the Mie response for the virtual Rayleigh channel:

$$C_2 = \sum_{0}^{jp_1-1} R_M[j_h]/nd + \sum_{jp_2+1}^{15} R_M[j_h]/nd .$$

6. Predict the Rayleigh-Brinillio line shape for the molecular backscatter ($I_{RB}[j_h]$). The method is based on a parameterizing the return as a sum of three Gaussians (a central peak together with and two symmetrically offset peaks). The FWHMs, offset and amplitude of the peaks depend on the atmospheric temperature and pressure (see Appendix F).

7. Calculate the R-B response as a function of wedge-position: $R_R[j_h] = \sum_{i_l} T[i_l, j_h] I_{RB}[\delta v[i_l]] \Delta v_{i_l}$

8. Calculate the fraction of the Rayleigh response for the virtual Mie channel: $C_4 = \sum_{jp_1}^{jp_2} R_R[j_h]/nd$

9. Calculate the fraction of the Mie response for the virtual Rayleigh channel:

$$C_1 = \sum_{0}^{jp_1-1} R_R[j_h]/nd + \sum_{jp_2+1}^{15} R_R[j_h]/nd .$$

Appendix C   Calculation of Effective Mie-spectrometer Spectral response (EMSR)

This Appendix briefly describes a method to use `clear-sky' returns to quantitatively determine the combined tripod/beam-shape/ACCD-efficiency response (or tripod factor for short) of the Aeolus Mie channel. Current methods of assessing the response involve e.g. operating the ACCD array in imaging mode and shifting the laser frequency out of the useful spectral range so that Mie backscatter is out-of-band (this is the so-called Mie Outside of Useful Spectral Range or MOUSR method). This method can only be performed at infrequent intervals and requires a special operation mode. A method using clear-air returns has the advantage that it can be performed during the course of normal observations.

Appendix C.1   The MSP

The Aeolus Mie Spectrometer (MSP) is an impressive instrument. Based upon a Fizeau spectrometer, the MSP delivers data with a spectral resolution of about 0.04 pm. The Fizeau element is comprised of a narrow wedge. Like a Fabry-Perot interferometer, the passband depends on the plate(wedge) thickness. Since the wedge thickness changes with position across the element the pass-band also changes. This is illustrated in Figure 20 for a `perfect' device i.e. the illumination across the spectrometer

In Figure 20, it can be seen that on the scale of the spectrometer useful spectral range (USR) that, in contrast to the Mie peak, the expected Rayleigh response is not exactly, but close to being flat. This can be seen in more detail in Figure 20 where the Rayleigh response corresponding to two sets of atmospheric conditions is shown. Some variation can be seen but the curves are flat to within a figure on the order of 1-1.5 %.
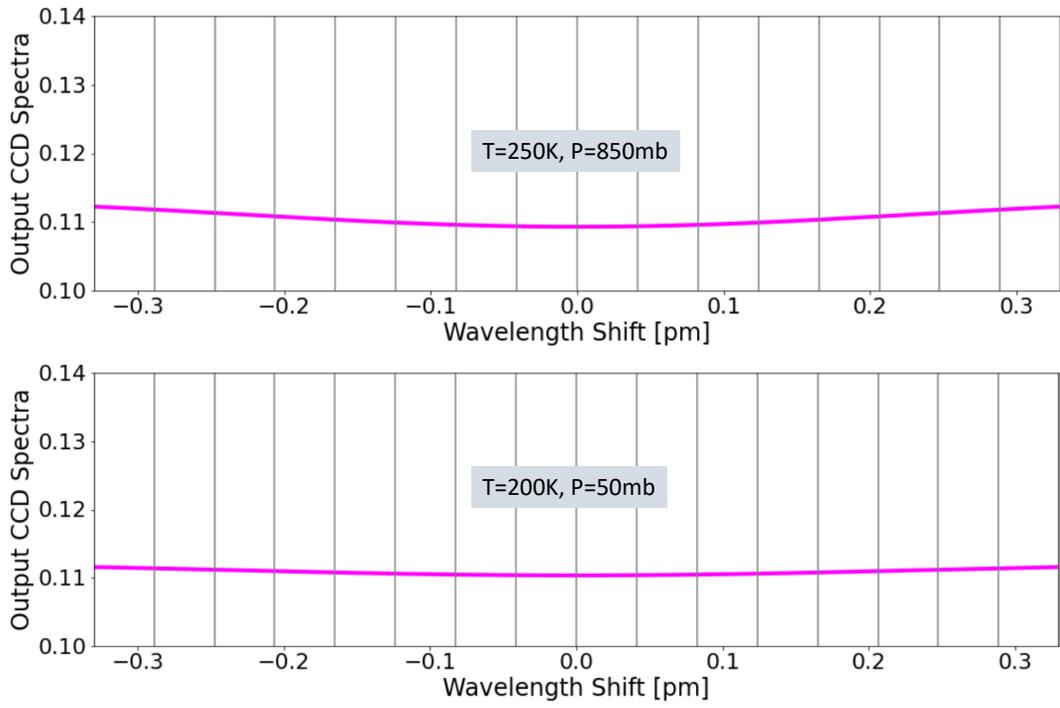
Figure 20: Response of an ideal MSP to a pure Rayleigh scattering input spectrum for two different atmospheric temperature/pressure conditions. The y-axis scale here corresponds to an input Rayleigh spectrum such that the integral over the input spectrum in frequency space is equal to 1.

The figures presented so far correspond to ideal conditions, that is, the ACCD response is uniform, and the MSP input aperture is uniformly illuminated. In practice, this is not the case, the response of the actual instrument is affected by intensity distribution of the light illuminating the MSP (See Figure 21). Thus, the spectral response will be affected by such factors as the effective beam-shape of the lidar backscatter signal, imperfections in the optical elements, and any obscuration in the optical path (e.g. the telescope tripod and secondary mirror assembly). This is clearly illustrated by the ACCD MOUSR image shown in Figure 22.
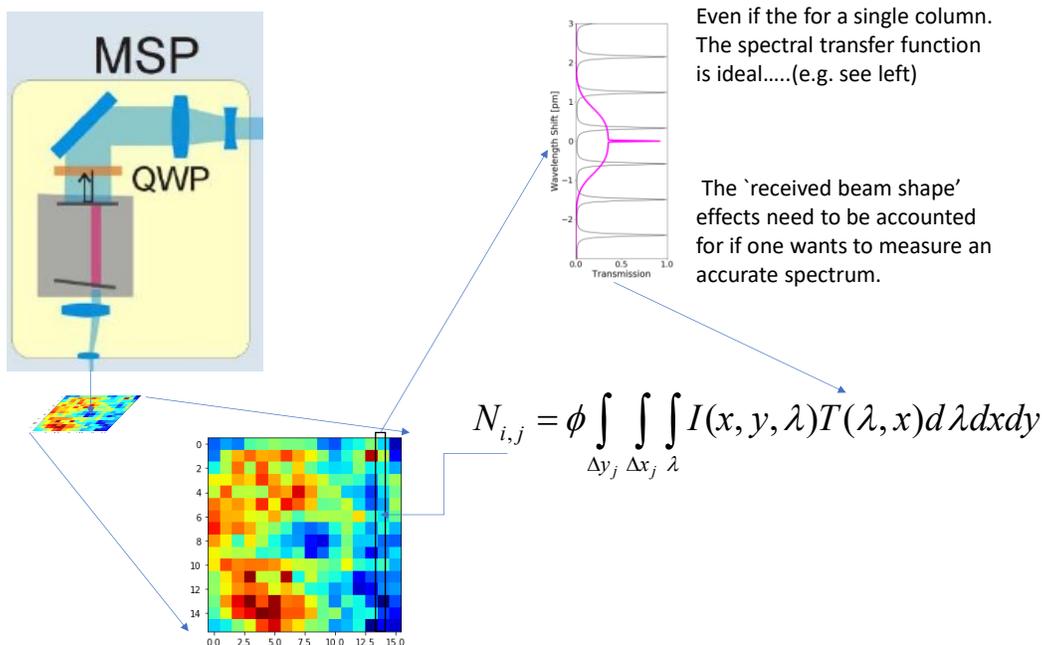
MSP

QWP

Even if the for a single column.
The spectral transfer function
is ideal.....(e.g. see left)

The `received beam shape'
effects need to be accounted
for if one wants to measure an
accurate spectrum.

$$N_{i,j} = \phi \int\limits_{\Delta y_j} \int\limits_{\Delta x_j} \int\limits_{\lambda} I(x, y, \lambda) T(\lambda, x) d\lambda dx dy$$

Figure 21: Illustration of the effect of non-uniform illumination of the MSP.
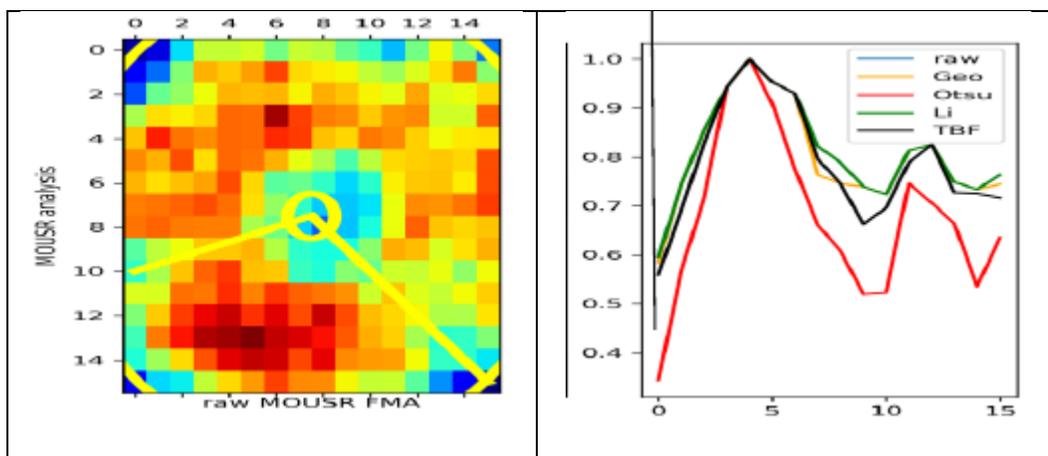
Figure 22: Left: Image of the 2D ACCD MSP acquired on 12/11/2018 obtained during a MOUSR (Mie Outside of Useful Spectral Range) measurement). Right. Corresponding normalized ACCD column totals using different data filtering strategies. Note: adapted from presentation by C.L. March 2021 L2 aerosol meeting.

Having seen that for an ideal system, the spectral response across the 16 signal ACCD columns is flat to a good approximation, estimating the tripod factor may be accomplished by averaging over observations where the signal is due mainly to Rayleigh scattering and then comparing the result to the expected

"flat" response. The task is then reduced mainly to formulating a robust means to select suitable areas that can be easily applied during the course of routine operations.

Appendix C.2   Algorithm

It is known that the 'cleanest' (referring to minimum background lidar scattering ratios) areas of the atmosphere in the altitude range routinely covered by Aeolus are

1. In the stratosphere at least a few km above the peak of the stratospheric aerosol layer (which usually occurs a few km above the tropopause.
2. In the upper tropopause where cirrus cloud formation and sedimentation actively scavenges aerosols out of the region.

Thus, we use the tropopause as a guide to selecting candidate regions. Further, we must also screen for the occurrence of high clouds and avoid low SNR observations. This does not need to be especially accurate but should be conservative with respect to excluding areas of elevated aerosols and especially clouds. To accomplish this, the **/measurement/mie_measurement_data[[454,30,25,20] "Spectrometer counts"** data from the L1B product are used. The Mie spectrometer data is averaged and the solar background removed using the last column of the ACCD (the DCO subtraction has already been applied to the "**mie_measurement_data**") on a per observation basis. The resulting spectra are segregated into two 'wing' regions and a central region (see Figure 23). The expected (Mie-free) signal in the central region can be predicted using the average wing signals. The ratio of the observed central region signal and the expected Mie-free signal then is used as a `quick-and-dirty' estimate of the lidar backscattering ratio (Note: this procedure is similar to the L1B (non-refined) scattering ratio).
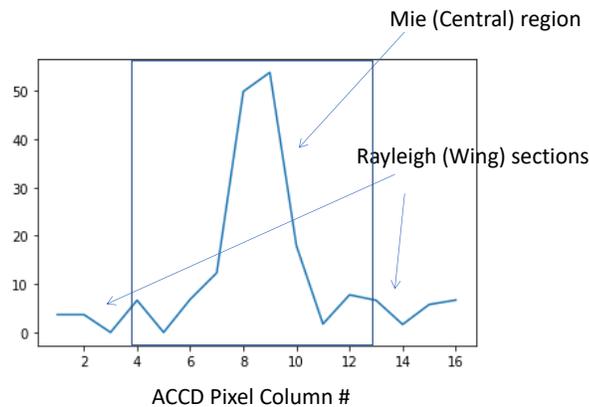

Figure 23: Sample ACCD spectrum in relative units.

Putting together the points discussed above as well as a few refinements, results in the procedure depicted in Figure 24. Here N_w refers to the number of pixels assigned to the `wings', Sum_c is the sum of the signal in the central region, Sum_w is the sum of the signals belonging to the wing regions, and Std is used to denote the standard deviation. R_thresh is the scattering ratio estimate threshold

(currently set to 1.2) and dZ is the 'above-tropopause blind-zone.', P_sum(ipix) is the pixel dependent sum of the signals across all the selected clear-sky observation-altitude spectra. A mask which tracks which observation-altitude points that were used to produce P_sum is also generated (see Figure 25). The tropopause level is diagnosed using the T and P information already gridded to the MSP heights contained in the L2A product (although the AUX-MET data file could also be used). The procedure to determine the tropopause level is similar to that used by CALIPSO.

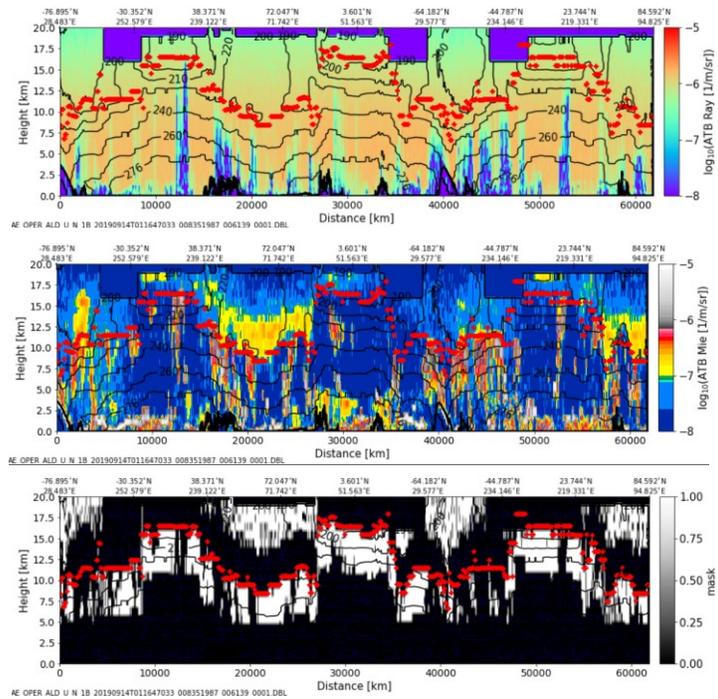*Figure 24: Flow chart of the tripod factor estimation procedure.*
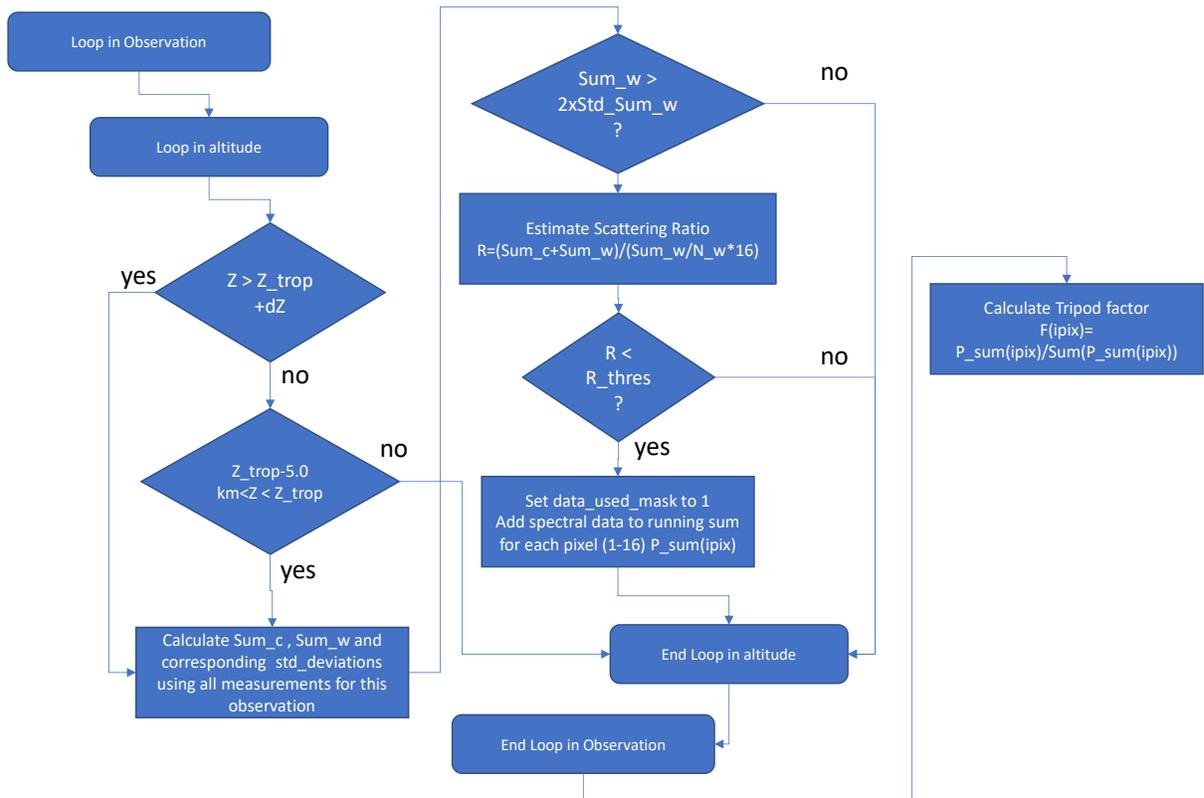
Figure 25: Example results of the masking procedure. The white areas in the bottom panel correspond to the areas used in determining the tripod factor. The Red points denote the tropopause height.
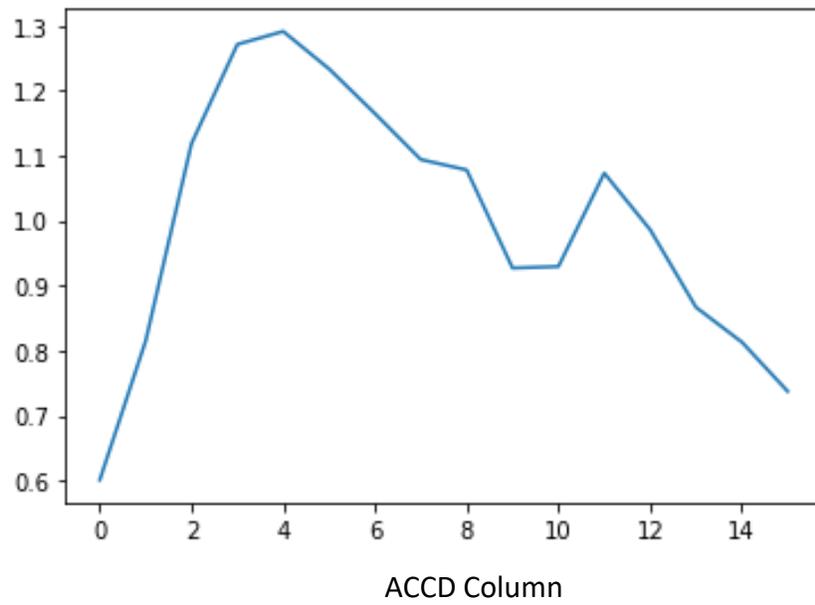
ACCD Column

Figure 26: Normalized tripod correction factor corresponding to the data shown in Figure 25.

Appendix D   ATB Calibration

Here the procedure used to (approximately) calibrate the ATBs is described. The method is based on normalizing the total observed scattering ratio to the result (neglecting aerosol attenuation) expected on the basis of the meteorological atmospheric density profile.

The main steps are as follows:

1. For each observation interval, the Rayleigh atmospheric optical depth array (τ_{Ray}) is calculated starting from the top of the atmosphere. The calculation uses the mod_od variable in the L2a product.

2. For each observation interval, points are selected based on a minimum height and an upper and lower limit to the observed scattering ratio $R = 1.0 + B_M / B_R$

3. For each observation group where more than the lidar calibration constant is estimated:

$$C_{lid}[i_{obs}] = \frac{R[i_{obs}]\beta_R[i_{obs},i_{alt}]\Delta r[i_{obs},i_{alt}]\exp[-2\tau_R[i_{obs},i_{alt}]]}{\left(B_M[i_{obs},i_{alt}] + B_R[i_{obs},i_{alt}]\right)\Delta r[i_{obs},i_{alt}]}$$

4. Outliers are filtered out be calculating the mean and the standard deviation of all the points calculated in the previous step. Points that deviate by more than 2 standard deviations are discarded.

5. An array containing all the remaining estimates is formed and padded using a wrap-around strategy (in order to reduce the impact of edge effects in the next step).

6. A high-degree polynomial fit (degree 10) of the selected observation as a function of observation number is conducted.

7. For each observation the polynomial fit is used to generate an estimation (along with predicted error) of the lidar effective lidar calibration constant.
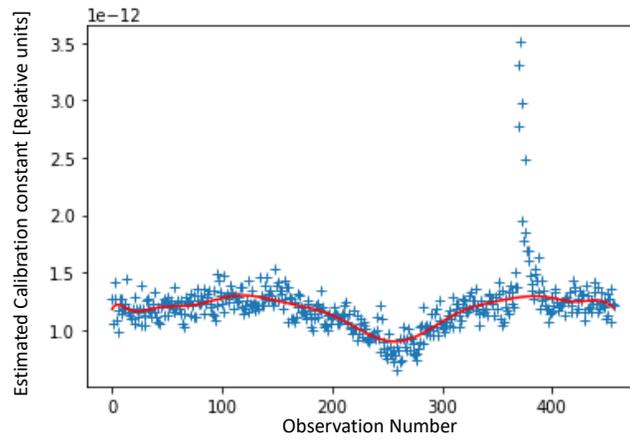
An example is shown in Figure 27

Figure 27: Example calibration constant as a function of observation number.

Appendix E   Missed Hot-Pixel detection

Here the procedure used to detect suspected missed hot-pixels and an approximate correction procedure is described. This procedure is necessary since hot pixels can sometimes appear or change levels between DUDE measurements. This can lead to corrupted EMSR determinations and other undesirable results.

The strategy is based upon tracking the minimum signal per-pixel/per-altitude per observation interval for the 16 signal pixels.

The main steps are:

1. Calculation of the minimum pixel levels (after background subtraction) per observation interval and height bin.
2. Applying a simple convolution kernel horizontal-line detection algorithm and applying a threshold to generate detection mask (1==> possible hot-pixel, 0==> ok).
3. Cleaning up the detection mask by:
    a. Allowing only one height per pixel to be `hot' (detection of multiple candidate lines for the same pixel for different heights is considered to be a sign of bad chosen thresholds and the whole height range is considered to be hot-pixel free for all pixels.)
    b. Convolving the mask for the remaining pixel rows where a line has been detected with a top-hat function and imposing a threshold of 1/2 the width of the mask. This joins segments with small gaps and filters out small, isolated detections.
4. For all pixels with a positive detection mask, a correction is applied. An additive adjustment is made so based on the difference between the hot-pixel minimum and the average minimum of the non-hot pixels for the same observation and height interval.

Appendix F   R-B Line shape parameterization

Here the procedure for estimating the R-B line shape is briefly described.

At low densities, the shape of the spectral shape of the Rayleigh backscatter is well-described by a single Gaussian profile whose FWHM is given by

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

Where, k is Boltzmann's constant, M is the average molecular mass, and T is temperature. However, in In general, the central (non-Raman) Rayleigh line profile (Cabannes line) will consist of three components, a central peak together with two flanking "Brillouin-Mandel'shtam" peaks [Miles, Lempert and Forkey (2001)]. In the low density or high temperature regime the uncorrelated motion of the scatterers gives rise to a Gaussian velocity distribution centred around the mean velocity of the flow and the above equation applied. As the pressure increases of the temperature decreases density fluctuations on the order of the laser wavelength may appear. These density fluctuations travel at the speed of sound in the gas and will give rise to acoustic side bands.
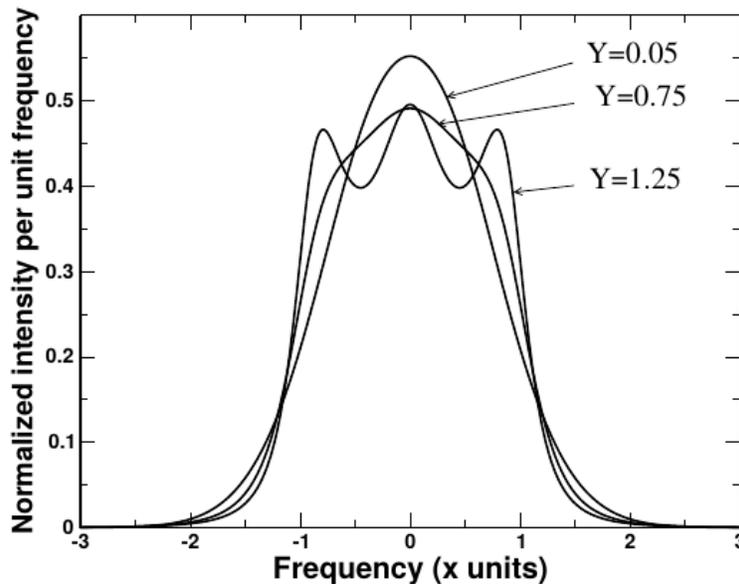

Figure 28:Rayleigh-Brillouin profiles for various y-parameters. x is a normalized frequency param-

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

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

where T is the temperature in Kelvin, P is the pressure in atmospheres, λ is the laser wavelength and θ is the scattering angle. Sample Rayleigh-Brillouin profiles for various y-parameters are shown in Figure 28 The profiles were calculated using a program due to [12]. Here x is a normalized frequency parameter defined as

$$x = \frac{\sqrt{2}\lambda}{4\sin(\theta/2)} \left(\frac{\mu}{v_o} \right).$$

Where μ is the frequency shift from the line centre and v_o is the sound speed ($v_o = \sqrt{(kT/m)}$)where m is the molecular mass.

For the range of y-values we are concerned with the RB line-shape may be accurately approximated using the sum of three Gaussians (see Figure 29). Accordingly, explicit runs using the approach of Pan were done for a range of temperatures and pressures. Three peak Gaussian fits were then performed on this data and the resulting fit coefficients were stored in a look-up-table (LUT) as a function of y-parameter. For general use, the Gaussian coefficients are deduced by linearly interpolation based on the LUT y-values and the y value corresponding to the desired set of T,P, wavelength and scattering angle.
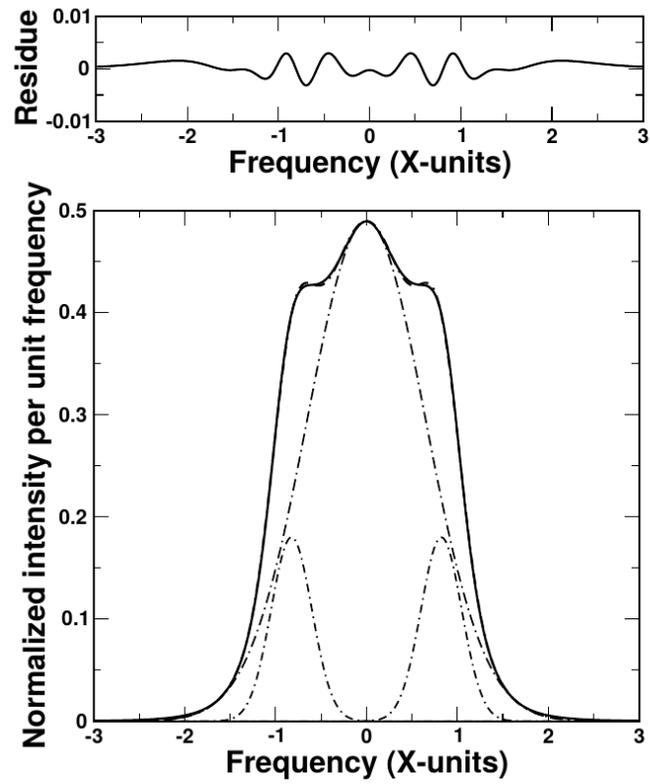
Figure 29: Figure 28: (Bottom Panel) Exact Rayleigh-Brillouin line shape (Solid-Line) along with fitted sum of three Gaussian functions (Dashed Dotted line). The three component Gaussian functions are also shown. Here the y parameter is equal to 1.0. The top panel shows the difference between the fitted function and the exact calculation.

7 References

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

[2] Reitebuch O., D. Huber, I. Nikolaus, ADM-Aeolus ATBD Level1B, AE-RP-DLR-L1B-001.

[3] Flament, T., Trapon, D., Lacour, A., Dabas, A., Ehlers, F., and Huber, D.: Aeolus L2A aerosol optical properties product: standard correct algorithm and Mie correct algorithm, Atmos. Meas. Tech., 14, 7851–7871, https://doi.org/10.5194/amt-14-7851-2021, 2021.

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

[5] Donovan, D. P., van Zadelhoff, G.-J., and Wang, P.: The EarthCARE lidar cloud and aerosol profile processor (A-PRO): the A-AER, A-EBD, A-TC and A-ICE products, EGUsphere [preprint], https://doi.org/10.5194/egusphere-2024-218, 2024.

[6] Eloranta E.W., 1998: Practical model for the calculation of multiply scattered lidar returns, Appl. Opt. 37, 2464-2472 (1998).

[7] Hogan, R., Fast, approximate calculation of multiply scattered lidar returns, Applied Optics, 45, 5984-5992 (2006).

[8] Platt, C. M. R., 1973: Lidar and radiometric observations of cirrus clouds, J. Atmos. Sci., 30, 1191–1204.

[9] Bocquet, M., Pires, C. A., and Wu, L. (2010). Beyond Gaussian Statistical Modeling in Geophysical Data Assimilation. Monthly Weather Review 138, 8, 2997-3023

[10] Novak O. et al., ``Fizeau interferometer system for fast high-resolution studies of spectral line shapes'' Rev. Sci. Ins., 82, 023105 (2011).

[11] Miles, R.B., W.R. Lempert and J.N. Forkey, Laser Rayleigh scattering, Meas. Sci. Technol., 12, R33–R51 (2001).

[12] Pan X., Coherent Rayleigh-Brillouin Scattering, Ph.D. thesis, Princeton University, 2003.