Aeolus L2A Aerosol Optical Properties Product: Standard Correct Algorithm and Mie Correct Algorithm

Thomas Flament¹, Dimitri Trapon¹, Adrien Lacour¹, Alain Dabas¹, Frithjof Ehlers², and Dorit Huber³
¹CNRM, Université de Toulouse, Météo-France, CNRS, Toulouse, France
²ESA/ESTEC, Keplerlaan, Noordwijk, The Netherlands
³DoRIT, Munich, Germany

Correspondence: Thomas Flament (thomas.flament@meteo.fr)

Abstract. Aeolus carries ALADIN, the first High Spectral Resolution Lidar (HSRL) in space. Although ALADIN was optimized to measure winds, its two measurement channels can also be used to derive optical properties of atmospheric particles, including a direct retrieval of the lidar ratio.

This paper presents the two main algorithms of the optical properties product called Level 2A product, as they are implemented in version 3.12 of the processor, corresponding to the data labelled Baseline 12. The theoretical basis is the same as in Flamant et al. (2008). Here, we also show the in orbit performance of these algorithms. We also explain the adaptation of the calibration method, which is needed to cope with unforeseen variations of the instrument radiometric performance due to the in-orbit strain of the primary mirror under varying thermal conditions. Then we discuss the limitations of the algorithms and future improvements.

We demonstrate that the L2A product provides valuable information about airborne particles, in particular we demonstrate the capacity to retrieve a useful lidar ratio from Aeolus observations. This is illustrated on a case of Saharan dust emission, observed in June 2020.

1 Introduction

The Aeolus satellite from the European Space Agency was launched on 22 August 2018, after a long development phase. Carrying a single instrument, the Doppler lidar ALADIN (Atmospheric LAser Doppler INstrument), Aeolus is the first satellite that can measure vertical profiles of wind at the global scale, from the surface of the Earth up to the lower stratosphere (20km to 25km of altitude depending on the settings). The lidar operates in the UV at the wavelength $\lambda = 354.8\,\text{nm}$. This short wavelength was chosen in order to enhance the molecular backscatter and allow measurements at high altitudes where aerosols and clouds are scarce and cannot serve as wind tracers. ALADIN implements two detection channels, a narrow-band one for the analysis of the UV light back-scattered by the particles (hydro-meteors and aerosols) that display a narrow spectrum (the full-width at half-maximum is of the order of a few tens of MHz), and a wide-band channel for the light back-scattered by the air molecules with a large spectrum (the full-width at half-maximum is several GHz) (Dabas et al., 2008). The former is called the Mie channel, and the latter the Rayleigh channel. The lidar was designed and optimized for the measurement of the wind. With its two channels, it is also what is usually called a high-spectral resolution lidar (Shipley et al., 1983).
Both channels detect a mix of photons backscattered by particles and molecules, but their sensitivity to both contributions is different. In the Rayleigh channel, the detection efficiency of molecular photons is about two-times better than for particle-backscattered photons while the Mie channel does the inverse with a sensitivity 30% better for particulate photons. With a precise calibration of the instrument, the number of photons backscattered by both types of target can thus be separated. This allows the independent measurement of the backscatter and extinction coefficients of aerosols or clouds (Flamant et al., 2008), and thus provides a direct measurement of the extinction to backscatter ratio called the "lidar ratio". This ratio gives an additional piece of information on the type of particle (Ackermann, 1998; Noh et al., 2007; Yorks et al., 2011; Illingworth et al., 2015; Shen et al., 2021). CALIPSO (Winker et al., 2007) is an older lidar mission that has been in space since 2006 and has already produced an immense set of vertically resolved profiles of optical properties of aerosol and clouds. In the CALIPSO algorithms, the type of target is estimated using the available information, a default lidar ratio is specified for each type and extinction and backscatter coefficients are then inverted (Ali H. et al., 2009). In Aeolus, a specific algorithm has been developed to exploit the high spectral resolution capacity of ALADIN. It produces the Level-2A (L2A) product of the mission. It is implemented in the AEOLUS ground-segment and delivers data in near-real time. Aerosols impact the climate (Houghton et al., 2001), either directly by absorbing the down-welling visible light from the sun or up-welling infrared radiation emitted by the earth, or indirectly by serving as cloud-condensation nuclei, thus changing the concentration and size distribution of cloud drops, and hence their optical properties and life cycle (Haywood and Boucher, 2000; Yu et al., 2006; Oikawa et al., 2018). These impacts depend on the size, concentration, nature of the aerosols. In this context, the detection of aerosol plumes from space coupled with an information on their type is most welcome to better understand the interaction between aerosols and the earth climate. AEOLUS is a first step towards this goal before the launch of ESA EarthCare mission (Illingworth et al., 2015) that will also operate a high-spectral resolution lidar in the UV, this time specifically designed for the observation of aerosol and clouds. Studies have also been conducted on the assimilation of AEOLUS L2A data by numerical air-quality prediction models. They suggest they can improve air-quality forecasts.

A first description of the L2A processor can be found in Flamant et al. (2008). Since this publication, the processor has substantially evolved. Technical aspects of the real ALADIN instrument discovered after launch - the sensitivity of the Rayleigh receiver to the temperatures of the primary mirror of the telescope for instance - had to be taken into account. Additional parameters were also added, like the lidar ratio or the estimated level of error of the different products and new algorithms are being developed for the future. A detailed, up-to-date presentation of the L2A processor is available with the L2A Algorithm Technical Basis Document (ATBD) (Flamant et al., 2021) that can be found on-line and is updated every time a new version of the L2A processor is integrated in the operational processing chain. The purpose of this paper is to summarize in a less technical, more easily readable article, what parameters are available in the L2A product and how they are determined (section 2). This paper focuses on the Standard Correct Algorithm and Mie Correct algorithm, which are the historical L2A products developed at IPSL and MeteoFrance. We later discuss the known limitations of these algorithm, and what future improvements are planned (section 3). A section illustrating L2A products from the main algorithm on a real case follows (section 4). It is followed by a conclusion.
2 The aerosol optical properties product in context of the Aeolus mission

2.1 The ALADIN instrument, from the perspective of aerosol observation

A high-level description of the AEOLUS mission can be found in Stoffelen et al. (2005). Since the date of this publication, several parameters have changed. The satellite altitude has been lowered down to 320 km, the laser emission has gone from bursts of pulses of 12 seconds every 24 seconds to a continuous mode of operation with a pulse repetition frequency \( prf = 50 \text{ Hz} \), and the laser energy requirement has been reduced to 60 mJ (the current energy per pulse is around 70 mJ). The sounding geometry of AEOLUS is depicted in 1 and 2. The line-of-sight of the instrument is directed at 35 deg off-nadir. The signals detected from 20 consecutive pulses are accumulated directly on the detectors. The 20-pulse accumulated signals are transmitted to the ground. They are called measurements and set the finest resolution \( \approx 3 \text{ km} \), or granularity, of AEOLUS products. The sequence of 30 consecutive measurements form a "Basic Repeat Cycle" (BRC). The BRC notion is inherited from the previous, 24 s cycle of the burst-mode operation of the laser. In AEOLUS products, the measurements belonging to a BRC are stored in the same data set record and often processed altogether to form a single profile called observation that has a horizontal integration length of 87 km.

Signals detected in the Mie and Rayleigh channels are also integrated vertically in height-bins. The number of height-bins is fixed, equal to 24 on both channels. The altitude and thickness of these bins can be adjusted, separately for each channel. This capacity is used to refine the vertical resolution on particular features like low stratospheric winds in the tropical band for Quasi-Biennial Oscillation studies or extend the observed range higher up to study Polar Stratospheric Clouds, for instance. The bin thickness is always a multiple of 250 m. Fine bins are used preferably close to the ground where the wind is expected to vary more rapidly. The vertical resolution is relaxed higher up in order to reach the lower stratosphere. The settings of height bins usually ensure that there is one or several Mie bins in a Rayleigh bin, so that Mie and Rayleigh signals probe the same volumes. This correspondence is not possible for the top-most Rayleigh bin that must be above the top-most Mie bins. Aeolus also has a terrain-following system to avoid losing too many height-bins below the ground. This provides the best coverage given technical constraints, but also means the product is not provided on a regular vertical grid, which makes plotting and analysis slightly more complex.

The Rayleigh channel implements a dual Fabry-Perot interferometer. The Mie channel uses a Fizeau interferometer. The shape of the optical filters is drawn in Fig.3. The left column shows the spectral transmission of the two Fabry-Perot (top) and the Fizeau (bottom). In both panels, the maximum transmission is normalized to 1, but the overall transmission factor of the Fizeau is about \( \approx 4 \) times less than for the dual Fabry-Perot. The figure shows that the Mie peak in the spectrum is significantly filtered out by the dual Fabry-Perot as it stands half-way between the peak transmissions of Fabry-Perot A and B, where the sum of the two transmissions reaches a local minimum. In the Mie channel, the Mie peak is almost unfiltered because it stands where the transmission of the Fizeau is at its maximum value. Overall, the efficiency of the Rayleigh detection chain for the photons backscattered by particles is about 50% of what it is for molecular photons, while it is 130% through the Mie detection chain. This contrast is not very high but still allows the separation of particulate and molecular contributions to detected signals as explained below.
Figure 1. Sounding geometry of AEOLUS (copyright ESA)

Figure 2. Mie and Rayleigh height bins (copyright ESA)
Figure 3. Spectral transmission of the two Fabry-Perot (top-left) and Fizeau (bottom-left) interferometers, and filtering of a Mie-Rayleigh spectrum by the double Fabry-Perot (top-right) and Fizeau (bottom-right). The Rayleigh-Brillouin component of the spectrum is for a temperature $T = 300$ K and a pressure of $P = 1000$ hPa. The Mie contribution of the particulate backscatter is the narrow peak on top. The spectrum used here corresponds to a scattering ratio of 1.0025, i.e. it has a very small Mie peak, for illustration purposes.

2.2 L2A product overview

This is a presentation of the product structure and an overview of the algorithms. Most of the theoretical basis is consistent with Flamant et al. (2008).

The L2A product is subdivided in data sets, providing retrievals from several algorithms and ancillary information (quality indicators, geolocation, etc.). This is a quick introduction, before each algorithm and data set is described in more details in its own paragraph.

The first data set contains the geolocation information. It is copied directly from the L1B. For the sake of traceability, some geolocation information specific to the L2A algorithms (from measurement groups for instance, see 2.2.4) are also written in other data sets.
The main optical properties data sets are derived by the Standard Correct Algorithm (SCA) and the accompanying Product Confidence Data (SCA_PCD). The SCA is described below in section 2.2.1.

The Mie Correct Algorithm (MCA) was designed to use the Mie channel only. It provides backscatter and extinction coefficients when Rayleigh signals are unavailable or when Rayleigh and Mie height-bins do not match. It does some sort of crosstalk correction based on the L1B-derived scattering ratio \( \rho = 1 + \beta_p/\beta_m \) (where \( \beta_p \) and \( \beta_m \) are the particulate and molecular backscatter), but it does not exploit the high-spectral capability of AEOLUS and uses a predefined lidar ratio for the inversion of Mie signals into extinction and backscatter profiles.

At last, two data sets aim at higher vertical or horizontal resolutions:

- the Iterative Correct Algorithm (ICA) tries to subdivide the height-bins vertically by testing assumptions on their partial filling by particles.
- the Group product accumulates signal over features with "high" Mie SNR (clouds for instance) within one BRC before applying the SCA. The idea is to remove portions of clear sky between features, in a simple attempt to retrieve more meaningful optical properties for the features.

### 2.2.1 Standard Correct Algorithm

The main product is the Standard Correct Algorithm (SCA). It uses both Mie and Rayleigh channels to derive crosstalk corrected signals, separating particulate and molecular contributions from the signal.

The ATBD (Flamant et al., 2021) provides a full development of the equations below. For the sake of concision, we only give the outline of the algorithm and important parts to understand the following paragraphs.

The Rayleigh and Mie signals measured by AEOLUS can be written

\[
S_{\text{ray},i} = K_{\text{ray}} N_p E_0 \ast (C_1,i X_i + C_2,i Y_i) \tag{1}
\]

\[
S_{\text{rie},i} = K_{\text{rie}} N_p E_0 \ast (C_4,i X_i + C_3,i Y_i) \tag{2}
\]

where \( i \) is the height-bin index (counted from top to bottom, \( i = 1 \) is the top-most bin), \( N_p \) is the number of accumulated pulses (\( N_p = 600 \) for an observation), \( E_0 \) is the pulse energy (monitored on-board and reported in the Level-1B product), \( C_1 \) to \( C_4 \), \( K_{\text{Ray}} \) and \( K_{\text{Mie}} \) are calibration constants, and

\[
X_i = \int_{z_{i+1}}^{z_i} \frac{1}{R^2(z)} \beta_m(z) T^2_{m}(z) T^2_{p}(z) dR(z) \tag{3}
\]

\[
Y_i = \int_{z_{i+1}}^{z_i} \frac{1}{R^2(z)} \beta_p(z) T^2_{m}(z) T^2_{p}(z) dR(z) \tag{4}
\]
are the molecular and particulate contributions to the signals (called the pure molecular and particulate signals), with $z_i$ the altitude of the top of height-bin $i$ and bottom of height-bin $i-1$, $R(z)$ the satellite-to-target range ($dR(z) = R'(z)dz$) and

$$T_m(z) = e^{\int z_{sat}^{z} \alpha_m(x)dx}$$  \(5\)

$$T_p(z) = e^{\int z_{sat}^{z} \alpha_p(x)dx}$$  \(6\)

are the one-way transmission through the atmosphere, with $\alpha_m$ and $\alpha_p$ the molecular and particulate extinction coefficients. The calibration coefficients $C_{1,i}$ and $C_{4,i}$ are functions of the pressure, temperature and line-of-sight wind velocity component inside height-bin $i$. They describe the changing portion of molecular spectrum that is transmitted through the Rayleigh and Mie receivers, respectively. $C_2$ and $C_3$ depend on the wind velocity in the satellite line of sight and give the transmission through each channel for narrow band spectra from particles. $C_1$ and $C_4$ coefficients are normalized to be 1 in standard conditions (1000 hPa, 300 K, 0 MHz Doppler shift). All $C_i$ coefficients are derived from an ad-hoc calibration procedure and tabulated for many sets of pressure, temperature and line-of-sight winds in the auxiliary file named AUX_CAL input to the L2A processor ((Dabas, 2017). The pressure $P$, temperature $T$ and line-of-sight wind velocity $v_{los}$ (from which the Doppler shift is computed) conditions inside the height-bins are provided by a specific product called the AUX_MET. This product is based on weather forecasts from the European Center for Medium Range Weather Forecasts (ECMWF). It contains predicted vertical profiles of the main atmospheric parameters along the predicted AEOLUS measurement track.

The determination of the radiometric calibration coefficients $K_{Ray}$ and $K_{Mie}$ is discussed below in section 2.2.2.

The first operation of the SCA consists in inverting equation (2)

$$X = \frac{K_{mie}C_3S_{ray} - K_{ray}C_2S_{mie}}{N_pE_0K_{ray}K_{mie}(C_1C_3 - C_2C_4)}$$ \(7\)

$$Y = -\frac{K_{mie}C_4S_{ray} - K_{ray}C_1S_{mie}}{N_pE_0K_{ray}K_{mie}(C_1C_3 - C_2C_4)}$$ \(8\)

This operation is called the cross-talk correction.

From these cross-talk correction signals, the derivation of the particulate backscatter is straightforward:

$$\beta_p = \frac{Y}{X} \beta_{m, sim}$$ \(9\)

where

$$\beta_{m, sim} = 1.38 \left( \frac{550 \mu m}{\lambda [nm]} \right)^{4.09} \frac{P [hPa]}{1013} \frac{288}{T[K]} 10^{-6} m^{-1} sr^{-1}$$ \(10\)

is the molecular backscatter coefficient expected from the pressure $P$ and temperature $T$ information from the AUX_MET file (Collis and Russell (1976)).

For the extinction, the derivation is done recursively from the top of the profile to the bottom. The assumption is made that the first bin contains no particles and hence that $\alpha_{p,1} = 0$. The L2A uses a so called "Normalized Integrated Two-Way
Transmission" (NITWT) defined as

\[ NITWT_i = \frac{X_i X_{1,\text{sim}}}{X_1 X_{i,\text{sim}}} \]  \hspace{1cm} (11)

where

\[ X_{i,\text{sim}} = \int_{z_i}^{z_{i+1}} R^{-2}(z) \beta_m(z) T_{m,\text{sim}}^2(z) dR(z) \]  \hspace{1cm} (12)

with

\[ T_m(z) = \exp \left( - \int_x^{z_{\text{sat}}} \alpha_{m,\text{sim}}(x) dR(x) \right) \]  \hspace{1cm} (13)

is the pure molecular signal expected from the \( P \) and \( T \) profile from the AUX_MET when there are no particles in the atmosphere (\( \alpha_{\text{sim}} = 8\pi \beta_{m,\text{sim}}/3 \) as expected from the Rayleigh theory for light scattering by molecules (Collis and Russell (1976)).

The extinction due to particles within bin \( i \) is then iteratively determined as:

\[ \alpha_{p,i} = \frac{1}{2 \Delta R_i} H^{-1} \left( \frac{1}{T_{p,1,i-1}^2} NITWT_i \right) \]  \hspace{1cm} (14)

where \( H(x) = \frac{1-e^{-x}}{x} \), \( \Delta R_i = R(z_{i+1}) - R(z_i) \) is the range thickness of bin \( i \), and \( T_{p,1,i-1}^2 \) is the two way transmission through particles between bin 1 and bin \( i - 1 \).

The normalization with the signal of the first bin is needed to compensate any extinction that could occur above the topmost bin. But this also makes the SCA extremely sensitive to noise in the first bin. This is discussed in 3.2.

Equations have been derived to estimate the impact of the detection noise on measured signals \( S_{\text{Ray}} \) and \( S_{\text{Mie}} \) on retrieved \( \beta_p \) and \( \alpha_p \) values. The derivation of these error estimates is fully explained in Flamant et al. (2021) but is too cumbersome to be reported here. It is based on the assumption that the uncertainty of \( S_{\text{Ray}} \) and \( S_{\text{Mie}} \) is purely due to the Poisson counting noise, and uses second-order developments. As a consequence, error estimates are valid as long as the level of noise is not too high, otherwise the approximation introduced by the second-order developments becomes too coarse. The errors estimates do not take into account the impact of atmospheric heterogeneity within the BRC that increases the random noise on the BRC-accumulation of observation level \( S_{\text{Ray}} \) and \( S_{\text{Mie}} \). It nevertheless remains that they are useful to identify the \( \beta_p \) and \( \alpha_p \) estimations that are reliable, and then give a good idea of their accuracy.

2.2.2 Radiometric calibration

Because of thermal constraint on the primary mirror (M1), radiometric sensitivity of the instrument varies with the orbit phase. Following the original plan (see Dabas (2017)), calibration measurements are acquired during a special mode, called Instrument Radiometric Calibration (IRC), during which the satellite is rotated to point at nadir. The objective is to use the molecular atmosphere as a well characterized target, with negligible Doppler shift on the vertical. This rotation of the satellite
causes a change in solar illumination, and results in a different thermal equilibrium of the spacecraft. Although the thermal control loops are performing well, gradients of a few tenths of a degree occur across the primary mirror. Pointing the satellite to nadir also changes these gradients, thus distorting very slightly the M1. This has an impact on the angle of incidence of the beam reaching the interferometers and disturbs the analysis of the backscattered UV light. This impacts both the winds (Weiler et al., 2021) and the aerosol optical properties retrieval.

In order to account for these changes, new calibration schemes have been designed. They are based on the evaluation of "clear sky" signals, i.e. signals from a pure molecular atmosphere and above any particulate feature. Such features would add an unpredictable amount of backscatter and extinction and spoil signal prediction. In their absence, the molecular atmosphere provides a predictable target and signals backscattered from these portions of the sky can be accurately compared to predicted molecular returns. The radiometric calibration coefficients are then the factor needed to scale the predicted signals to the observed ones.

The first version of this calibration algorithm used the comparison of signals averaged along a full orbit file. This yielded a single value of radiometric calibration for each channel and each orbit, which would work for a thermally stable mirror (see step 2 of Fig.5). Coefficients are evaluated independently from one orbit to the next.

Because the thermal constraints on the M1 vary along the orbit, the radiometric performance also vary. The second version of the algorithm uses a least square fit of the observed signals to the 12 temperatures $T_i$ read off the thermal sensors of the M1.

Radiometric coefficients are then written in the form:

$$
K_{Ray} = c_{Ray}^0 + c_{Ray}^1 \times T_1 + c_{Ray}^2 \times T_2 + \ldots + c_{Ray}^{12} \times T_{12} + \epsilon \\
K_{Mie} = c_{Mie}^0 + c_{Mie}^1 \times T_1 + c_{Mie}^2 \times T_2 + \ldots + c_{Mie}^{12} \times T_{12} + \epsilon
$$

This also allows us to interpolate the radiometric calibration in areas where the observed signals are not clear enough e.g. areas where clouds or particles reach too high in the atmosphere for the instrument to collect enough "clear sky" signals.

The successive improvements brought by these calibration techniques are shown in Fig.5. It can be seen that the distribution of the particulate signal in clear sky, which is supposed to be 0 in these conditions, is better centred around 0 after correction. The second version of the radiometric calibration produces a slightly narrower distribution, which means an improvement of the crosstalk correction.

This approach does not account for the residual contamination of supposedly "clear sky" and might overestimate the radiometric coefficients.

Attenuation by a particle layer that would be above the observable range is also not accounted for. The fit being made over a full observation file, which is at least half an orbit long and most of the time longer, this is enough to guarantee that a high reaching particulate feature in a given region would not bias the fit too much.
2.2.3 Mie Channel Algorithm

The Mie Channel Algorithm (MCA) uses observations from the Mie channel alone. It uses a pseudo cross-talk correction by using the L1B derived scattering ratio:

\[ Y_{mie,i} = \frac{S_{mie,i}}{K_{mie} N_p E_0 \left( C_{4,i} + C_{3,i} \right)} \]  

(16)

It then calls a recursive relation to derive the extinction in a given bin based on the extinction in bins above. It is a version of the Klett algorithm (Klett, 1981).

The initialisation is made in the topmost Mie bin, assuming there is no particulate attenuation between the satellite and the first bin \( L_{p,sat,1} = \int_0^{R_1} \alpha_p(r)dr = 0 \). The molecular attenuation, \( L_{m,sat,1} = \int_0^{R_1} \alpha_m(r)dr \) is know from the simulated
Figure 5. Figure showing the evolution of calibration schemes and the resulting improvement of error in signal prediction. The left panels show the crosstalk corrected "pure particulate signals" in particle-free bins, the histograms on the right show the distributions of these signals.

atmosphere. Then,

\[ L_{p,i} = -\frac{1}{2} \ln \left( 1 - \frac{2Y_{mic,i} R_{\text{mean}_i} e^{L_{m,i}}}{T_{m,sat,i}^2 T_{p,i}^2} \right) \]

where \( k_p = \beta_p/\alpha_p \) is the inverse of the prescribed lidar ratio and \( T_{m,sat,i} \) is the molecular transmission to the top of bin \( i \).

The MCA major drawback is the use of a prescribed lidar ratio. By fixing this ratio however, the MCA imposes that extinction and backscatter are co-located. This yield a more consistent picture of extinction that is useful to spot features and local variations of extinction. Fig. 6 shows an example of extinction retrieved by the MCA on the orbit discussed below in section 4.2. It also hints that the dust plume extinction (around BRC 90-110) is under-estimated due to the fixed lidar ratio (see Fig. 12 for comparison with the SCA lidar ratio).

2.2.4 Higher resolution : Group product and ICA

The Iterative Correct Algorithm (ICA) is an attempt to improve the vertical resolution. Indeed, if a particulate feature only occurs in the bottom part of the bin, the signal will be different than if the feature fills the bin completely. The ICA tests various filling possibilities and chooses the most likely one. The ICA is a subsequent step, coming after the retrieval of optical properties by the SCA. Due to actual radiometric performances that are lower than expected before launch, the ICA suffers from high noise levels, even more than the SCA. Since launch the focus has been on the improvement of radiometric calibration and no development time has been spent on this product, which is kept for historical reasons.
The group product uses the Mie channel SNR to identify features and accumulate signals on sub-BRC scale to try to separate features. The objective is to provide more meaningful optical properties than the ones from the SCA, which are averaged over different types of objects. The group product is an application of the SCA over subsets of horizontal bins.

3 Known limitations

3.1 Instrument limitations

ALADIN was designed primarily for wind determination. The fraction of light sent through the Fizeau interferometer of the Mie channel is smaller than for the Rayleigh channel. The Mie SNR is then lower than the Rayleigh SNR and limits the precision of signals calculated through the crosstalk correction.

Designed as a wind lidar, ALADIN does not have the ability to measure depolarization. The UV laser beam is linearly polarized, and analyzed only along the parallel direction. Any cross-polarized light is rejected. This means that, in order to compare Aeolus observations to other instruments, only the co-polar component must be considered.

Not going through this extra step before comparing would make it seem that the total backscatter of highly depolarizing targets such as ice crystals or dust is largely underestimated by Aeolus.
Because the extinction is not affected, the corresponding Aeolus lidar ratio is going to be larger than the total lidar ratio.

### 3.2 High noise and extinction retrieval

Extinction can be calculated in a simple way from the molecular backscatter, or more precisely, from its derivative. Because we measure extinction by differentiating two bins, the retrieval is very sensitive to noise. We will see that the original algorithm works well in conditions with enough SNR, but it faces its limits in the "real world", where signals have shown to be weaker than expected before launch.

In order to retrieve the extinction, we compare the observed molecular signal to the molecular signal simulated from the atmospheric conditions (pressure and temperature, from auxiliary meteorological data, extracted from ECMWF NWP forecasts). The ATBD describes in detail how we can access the extinction within bin \( i \) from the available observations.

There are two important aspects to this algorithm: signal in one bin depends on the extinction within all overlying bins and bin \( i \) itself. Particle extinction is derived iteratively by propagating downwards the total extinction and comparison to simulated signal. The estimation of extinction is initialized at the topmost bin. By normalizing all molecular signals to the signal in the first bin, we can solve the problem of extinction above the sensing range, i.e. between the satellite and the topmost bin.

But this solution comes with a price: the molecular density, and hence the molecular backscatter, is also the lowest at the top of the sensed atmosphere, usually 20 to 25 km. The contribution of noise is then the strongest in the first bin. The retrieval of extinction by differentiating between the first two bins can yield unrealistically high values. This, in turn, means that the simulated signal is expected to be very low in the next bin. If the extinction was overestimated in the second bin, the signal in the third bin is measured to be larger than expected. This would lead to a negative extinction. If the first extinction is largely overestimated, it can take several bins before the measured signals reach the low expected signal level. See Fig. 7. The extinction is derived directly from the function on the right panel, extinction is positive only if this function is below 1. In the given example, we see that because of the overestimation in the first bin, the SCA extinction is only yields positive extinction around 2-3 km, whereas the MCA and the SCA backscatter coefficients suggest presence of Aerosols already above.

In the SCA, all bins that would have had negative extinction are reset to 0. This choice of thresholding has been largely discussed. It is justified by the error propagation in the iterative process. The large error in the retrieval of the first extinction would propagate downwards and create the stripes visible in Fig.8. Resetting negative extinction to 0 allows to reset also this propagation of error, but this comes at the price of bias, i.e., that the SCA extinction variable "reacts" delayed along the vertical and so lacks sensitivity (see paragraph above).

The propagation of errors through the algorithm, from the signals to the optical properties also shows the limitations of the SCA extinction retrieval. This is detailed in the ATBD (see eq. 6.82 in Flamant et al. (2021)) and can be summed up by the following formula:

\[
\sigma^2_{L_{p,i}} = \langle (\delta L_{p,i})^2 \rangle - \langle \delta L_{p,i} \rangle^2 \approx 4 \sum_{k=1}^{i} \langle e^2_{X_k} \rangle - 3 \langle e^2_{X_i} \rangle - 3 \langle e^2_{X_1} \rangle
\] (18)
**Figure 7.** Figure illustrating how extinction overestimation in a given bin leads to errors in bins below. The left panel shows observed signals (yellow), simulated signals (blue) and simulated signals with extinction from particles down to the previous bin and normalized by the first bin (dashed red). The extinctions retrieved by several means are shown in the middle panel (SCA in blue, derived from SCA backscatter with an arbitrary lidar ratio of 30 sr in red, from the MCA in red and the mid-bin SCA in purple). The value from which the extinction is derived is shown on the right panel.

**Figure 8.** Figure showing the extinction retrieved with the SCA (left) and with a modification that cancels the reset of negative values to 0 in the iterative retrieval process.
Figure 9. Particulate Backscatter input for the E2S. The white vertical lines separate the BRC.

Hence, it is not recommended to use the direct computation of extinction on the normal vertical scale. This drawback can be compensated by averaging two bins. This is done in the "mid-bin" part of the SCA data set. For users interested in extinction properties, it is highly recommended to use the "mid-bin" retrieval of the SCA.

The middle-bin algorithm allows negative values and averages the extinction values over two consecutive bins. The loss of vertical resolution is compensated by a substantial gain in errors (eq.8.86 of Flamant et al. (2021)):

$$\sigma^2 = \frac{1}{4}\langle e^2_{X_{i+1}} \rangle + \frac{9}{4}\langle e^2_{X_i} \rangle$$

(19)

In order to compute the lidar ratio by using this "mid-bin" extinction, the "mid-bin" backscatter is also derived by averaging the backscatter in two successive bins.

4 Example cases

4.1 Simulated example

The AEOLUS End-To-End Simulator has been developed before the satellite launch to help the development of the ground segment processors (Reitebuch et al., 2018). It simulates the ALADIN instrument in its nominal conditions and generates the downlinked data (Annotated Instrument Source Packet: AISP) used by the processor chain. The simulator also takes into account several noise sources including Poisson counting noise at the detector level. The following example is a scenario which is representative of tropical conditions with high convective clouds. The scene is constituted of 240 profiles which match the structure of the AEOLUS observation with 8 BRCs of 30 measurements each. This scenario presented in Fig. 9, presents heterogeneous scenes with optically thick features that presents a challenge for the L2A processor to retrieve the optical properties.

To examine the sensitivity of the L2A product to the noise, 20 independent E2S simulations are run from the same input scene. Figure 10 presents how the backscatter and extinction coefficient derived from the SCA mid-bin algorithm compare with the E2S inputs. Most of the time, the backscatter and extinction coefficients are correctly derived. Good results are obtained.
for the extinction retrievals, i.e. errors lie within the range of atmospheric heterogeneity, and the backscatter coefficients are also mostly correct, although with a slight low bias. The average of the 20 simulation overlap the expected values with a low dispersion meaning that one realization should be enough to characterize the atmospheric optical properties.

The largest discrepancies are seen below 4 km. In the low levels, the signal can be attenuated by optically thick features above. In addition, the bins below 2 km are only 250m thick: shorter accumulation time means that signals are lower and that the noise can become dominant. In practice the vertical resolution of the bins is seldom below 500 m. One may notice that some values of the retrieved coefficients are out of the graphics, they correspond to negatives values. Where noise dominates, the accumulated signal oscillates around zero and the average of the 20 realizations can be slightly negative.

In the BRC 4, the only one in near clear sky conditions, the averaged profile renders the backscatter and the extinction profiles correctly almost all the way down to the last bins. Nevertheless, a small systematic bias is visible when the bin height is 250 m, because the noise is large enough to have large negative values which are cut by the algorithm and not included in the average.

The BRCs 6 to 8 contain the optically thicker clouds and present the biggest heterogeneities of the scenario. They are also the BRCs with the less accurate SCA results; the profiles are consistently underestimated between 6 and 10 km, compared to the input profiles averaged over the BRC. The estimated errors are also too low and do not cover the expected values. This is because the variability in these BRCs is mostly due to the heterogeneity within the horizontal accumulation length, which is not taken into account in the error propagation.

In this example, the backscatter errors are underestimated by a factor 3 and the extinction errors are reasonably estimated. The bias in the errors can be traced back to the variability of the useful signal in the Mie and Rayleigh channels. In this scenario, the Poisson noise represents only a third of the simulated noise in the Mie channel but is the main contributor for the Rayleigh channel simulated noise. The bias is then propagated up to the calculation of the backscatter and extinction variances. To summarize, the accuracy of the error estimation depends on the significance of the Poisson noise contribution to the signal variability. Factors, like the heterogeneity, lead to an underestimation of the errors by the SCA algorithm.

This simulated example proves that, with the correct calibration coefficients, the SCA algorithm is able to render the clouds and aerosols optical properties of a complex scene and provide reasonable error margin as long as the Poisson noise is the main source of noise.

4.2 Saharan dust across the Atlantic ocean in June 2020

On June 2020 numerous instruments such as Copernicus Sentinel-5p TROPOMI (Zweers, 2018), NASA Suomi NPP VIIRS (Jackson et al., 2013) or NASA CALIPSO CALIOP (NASA, 2021) observed a massive dust plume ejected from Saharan Desert to the Atlantic ocean. The brown plume was highly visible from 13 to 20 June with true reflectance images. Aerosol Index (AI) product also showed the presence of UV-absorbing particles (i.e. Suomi NPP OMPS aerosol index up to 3 for thickest part of Saharan dust plume). Therefore this scene was selected in the frame of the evaluation of L2A product quality with external/validation data.
Figure 10. Backscatter (Top) and Extinction profiles (Bottom) comparison between E2S simulations processed up to the SCA mid-bin product and the E2S inputs. The black line is the simulation input parameters averaged within each BRC. Thin black line represents the associated standard deviation. Red profiles are the mean coefficients retrieved by the SCA algorithm from 20 realizations. Lighter shading is the associated standard deviation i.e. the true variability. Darker shading delimits the mean error estimated by the SCA error propagation algorithm.

On 19 June 2020 Aeolus crossed the North Atlantic Ocean between 08:00 and 08:30 UTC and the ALADIN instrument ground track (i.e. Aeolus displayed ground track is the intersection of ALADIN laser slant line of sight pointing 35° offset from nadir with the ground) intersected the western portion of the Saharan Air Layer (SAL) for approximately 20 BRCs. On the same day the dust layer can be seen with Aerosol index from 388 and 354 nm produced by two Copernicus Sentinel-5p TROPOMI overpasses from 14:50 to 16:32 UTC and from 13:09 to 14:50 UTC (noa, downloaded on 23 Feb. 2021). TROPOMI UV aerosol index product for the core of the plume (i.e. UV AI ≥ 2) is shown in Fig.11 overlaid with Aeolus and Calipso ground tracks.

The SCA lidar ratio has been processed from the mid-bin product and quality flags have been applied. The validity of the backscatter coefficient retrievals is determined by a threshold on the Mie SNR, while the validity of the extinction coefficient...
Figure 11. Figure showing Africa’s west coast and North east Atlantic Ocean from ArcGIS World Imagery from Online basemap (Esri et al., 2021) overlaid with Aeolus and Calipso ground tracks and BRC number between 8:08 to 08:16 UTC and Sentinel-5P UV Aerosol index from 388 and 354 nm spectral bands.

Figure 12. SCA mid-bin valid Extinction to Backscatter ratio. The red box is the selection for the averaging of the lidar ratio of the dust plume is determined by a threshold on the Rayleigh SNR. The signal is considered as valid in a specific bin if these Rayleigh and Mie SNR are respectively superior to 90 and 40. This allows to reject the low signal bins for which background noise is large.

The L2A valid lidar ratio is shown in Fig. 12. Numerous valid bins within the dust plume (i.e. red box) show homogeneous high lidar ratios in mid-troposphere from altitudes $\approx 2.5$ km to $\approx 5.5$ km. The plume appears to be clearly visible in the data.
with reasonable contrast despite background noise (i.e. light green and yellow color code). The median lidar ratio within the red box is 130 sr. A ratio of the same order was computed for dust particles emitted from identical Saharan source on 30 June and observed by Aeolus above Cap-Verde (Ehlers et al., 2021). The ratio appears to be higher compared to other lidar measurements of such dust layer (Mona et al., 2012; Nisantzi et al., 2015) but this can be explained by Aeolus measurement principle: the backscatter coefficient is underestimated and therefore leads to lidar ratio overestimation by Aeolus in the presence of depolarizing particles because ALADIN only detects the backscatter co-polar component. A number of studies (Ansmann et al., 2003) have shown that light depolarization ratio of dust and marine particles mixture is significant. Moreover earlier on 19 June, CALIPSO flew over the North East Atlantic Ocean from 04:07 UTC and captured the massive Saharan Air layer as shown in Fig. 13. The depolarization ratio at 532 nm appeared to be large (i.e. up to 0.5) for the dust layer (the large feature on the right of the figure, between 0 and 6 km altitude) confirming the high concentration of depolarizing particles.

Despite being overestimated, some interesting observations can be made regarding the variability of the lidar ratio measured by Aeolus over the co-located layer. The horizontal distribution of the lidar ratio for the core plume appear to be uniform (i.e. constant lidar ratio of order \( \approx 80 \) to 120 sr along BRCs 95 to 110 as shown in Fig. 12). Same observation is valid for the vertical for which the ratio appears to be relatively constant.

Auxiliary meteorological information provided by ECMWF (Rennie et al., 2020) is useful to underline the cloud-free conditions. Mean relative humidity (Rh), cloud cover fraction (CC) and cloud ice/liquid water content at nominal pressure (Clwc, Ciwc) for the core of the plume are the following: Rh=42\%, CC=1.2\%, Ciwc=5.4e-6 kg.kg\(^{-1}\), Clwc=8.8e-8 kg.kg\(^{-1}\). The low relative humidity can be noticed and appears to be in line with previous studies on Saharan dust layer (Wong et al., 2009). A basic cloud classifier derived from cloud liquid and ice water content confirms the low cloud contamination in the area of the dust plume as shown in Fig. 14. The southern part of the plume, BRCs 110 to 120 (blue box in Fig. 12), is lightly cloud contaminated and has slightly lower median lidar ratio of 68 sr, as expected. Much lower lidar ratios are found where the cloud
mask reports clouds, e.g. median lidar ratio of 15 sr within the green box of Fig. 12. This simple cloud classification is also helpful in other cases to discriminate aerosols from clouds and will be implemented in a future release of the L2A.

5 Conclusions

The L2A product is still under development and new algorithms, better designed to cope with the high noise level of the instrument, are being developed.

Yet, we showed the ability of Aeolus to provide valuable information thanks to its HSRL design and despite the high noise and absence of depolarizing channel.

The lidar ratio derived directly from Aeolus observations can help discriminating between clouds and various types of aerosols. Depolarizing targets will appear to have a large lidar ratio; that is not directly comparable to other measurement because Aeolus misses the cross-polarized backscattered UV light. The comparison with other systems can still be made if the depolarization ratio is known.

Assimilation of Aeolus backscatter coefficient into atmospheric chemistry and transport models is being studied and first results are encouraging.

Data availability. Data is available to registered users from the Aeolus Data Dissemination Facility http://aeolus-ds.esa.int/oads/access/
Author contributions. The CNRM contributors are responsible for the code developments and maintenance. Frithjof Ehlers provided analyses of the L2A results and contributed to the improvement of algorithms. Dorit Huber is responsible for the coding of the operational processor deployed at ESA. The paper was written collegially by CNRM authors and reviewed by Dorit Huber and Frithjof Ehlers.

Competing interests. The authors declare no competing interests.

Acknowledgements. The L2A has been developed over more than 15 years by numerous contributor (Pierre Flamant, Marie-Laure Denneulin, Mathieu Olivier, Vincent Lever, Pauline Martinet, Hugo Stiegliz ...). Most of the development work for the L2A processor was carried out in the frame of contracts from the European Space Agency, but also benefited from the financial support of the French space agency CNES. We acknowledge the use of imagery from the NASA Worldview application (https://worldview.earthdata.nasa.gov/), part of the NASA Earth Observing System Data and Information System (EOSDIS).
References


