ATLAS

ATLID Algorithms and Level 2 System Aspects

Algorithm Theoretical Basis Document (ATBD) for Lidar Aerosol inversion

Code : EC-TN-KNMI-ATL-19
Issue : 2.2
Date : 27/06/11
Authors : S. Berthier
D.P. Donovan
G-J van Zadelhoff

This page intentionally left blank

Issue	Revision	Date	Reason for Change	Author
1.0	Baseline	25/02/11	Version delivered to ESA	S.B., D.D., GJvZ [KNMI]
2.0		27/04/11	Version for the ATLAS final meeting	S.B., D.D., GJvZ [KNMI]
2.1		26/05/11	ATLAS Final Version	S.B., D.D., GJvZ [KNMI]
2.2		27/06/11	Update after Final review	S.B., D.D., GJvZ [KNMI]

Table of Contents

1. Purpose and Scope

This document describes the algorithms used to determine the large horizontal scale aerosol extinction, backscatter and depolarization ratio L2a ATLID products.

The general flow chart given in the figure 1. Red boxes give the part of the architecture corresponding to the present module. This module take as an input the L1b datasets and the L2a Featuremask (A-FM). This module makes a call to the Aerosol classification module to retrieve the aerosol typing (A-TC). The output of this module is ingested by the L2a Extinction module (A-EBD).

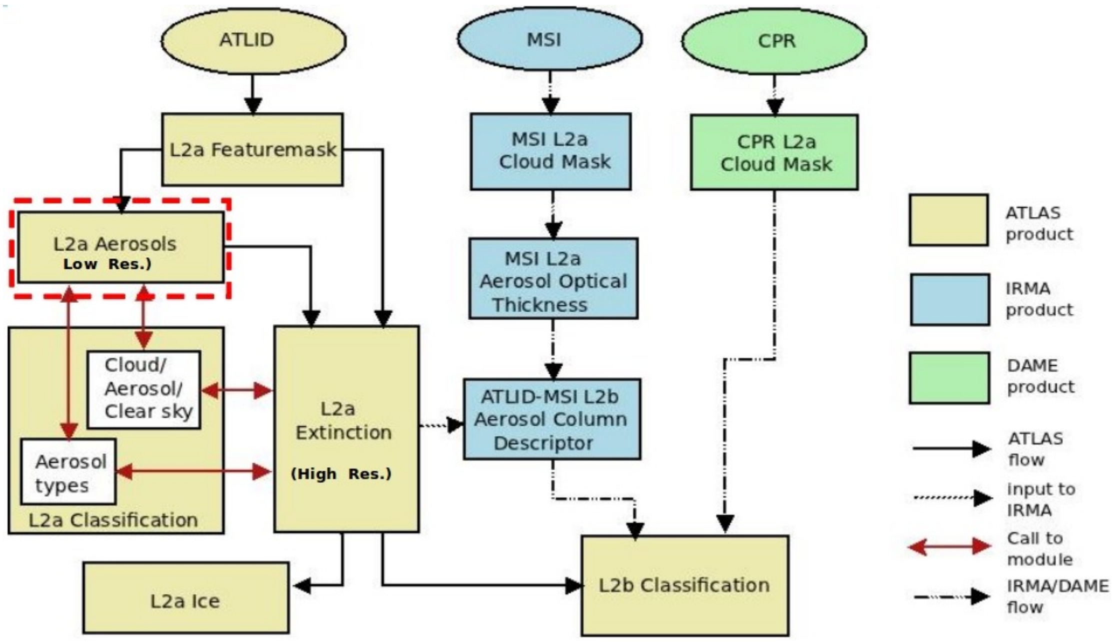

Figure 1: ATLAS General flow chart

2. Applicable and Reference Documents

2.1. Applicable documents

Reference	Code	Title	Issue	Date
MRD	ER-RS-ESA-SY-012	EarthCARE Mission Requirements Document	5	11/02/06

2.2. Reference & Related documents

Reference	Code	Title	Issue	Date
CASPER-PARD	CASPER-DMS-PARD-001	CASPER Products and Algorithms Requirement Document (PARD)	2.0	30-10-08
CASPER-FINAL	CASPER-DMS-FR-01	CASPER Final Report	1.1	30-01-09
ATL-PARD	EC-TN-KNMI-ATL-005	ATLAS Products and Algorithms Requirements Document (PARD)	1.1	10-03-10
EC-PTR	EC-ICD-ESA-SYS-0314	EarthCARE product table reference	1.3	06/15/10
A-FM	EC-TN-KNMI-ATL-010	ATLID L2a Feature mask	2.2	26-05-11
A-EBD	EC-TN-KNMI-ATL-021	ATLID L2a Extinction, Backscatter & Depolarization	1.2	26-05-11
A-TC	EC-TN-KNMI-ATL-022	ATLID L2a target classification	2.1	26-05-11
ATL-FINAL	EC-FR-KNMI-ATL-027	Final report	1.0	27-05-11

2.3. Reference

Keyword	Reference
[VPW05]	Mark A. Vaughan Kathleen A. Powell, David M. Winker, 2005, CALIOP Algorithm Theoretical Basis Document; Part 2: Feature Detection and Layer Properties Algorithms, PC-SCI-202 Part 2 [http://www-calipso.larc.nasa.gov/resources/pdfs/].
[Ackerman , 1998]	Ackermann, J., "The extinction-to-backscatter ratio of tropospheric aerosol: A numerical study", J. Atmos. Ocean. Tech., 15, 1046-1050, (1998).
[Evans, 1988]	Evans, B., "Sensitivity of the backscatter / extinction ratio to changes in aerosol properties: Implications for lidar", Appl. Opt., 27, 3299-3306, (1988).
[Klett, 1981]	Klett, J.D., "Stable analytical inversion solution for processing lidar returns", Appl. Opt., 20, 211-220, (1981).
[Müller, 2007]	Müller D., A. Ansmann, I. Mattis, M. Tesche, U. Wandinger, D. Althausen, and G. Pisani, "Aerosol-type--dependent lidar ratios observed with Raman lidar", J. Geophys. Res., 112, D16202, (2007)., doi:10.1029/2006JD008292
[Piironen, et al., 1994]	Piironen, P. and E. W. Eloranta, 1994: Demonstration of a high-spectral-resolution lidar based on a iodine absorption filter. Opts. Lett., 19, 234–236.
[Eloranta, et al., 2005]	Eloranta, E. W. High Spectral Resolution Lidar. in Lidar: Range-Resolved Optical Remote Sensing of the Atmosphere , Edited by C. Weitkamp, Springer-Verlag, New-York, 2005, 455p.

3. Scientific Background of the algorithm

3.1. Algorithm history

The present algorithm has been built upon a previous version of the lid_l2a ECSIM algorithm which included a method to retrieve aerosols extinction and backscatter. This approach is based on using a linear fitting procedure to determine the logarithmic derivative of the Lidar Rayleigh signal in order to determine the aerosol extinction.

3.2. Algorithm introduction

The objective of this algorithm is to retrieve the aerosols optical properties (the extinction coefficient $\boldsymbol{\alpha}$, the backscatter coefficient $\boldsymbol{\beta}$ and depolarization ratio $\boldsymbol{\delta}$) from the ATLID instrument on board the EarthCARE satellite. Unlike the 1-km scale

Cloud and Aerosol extinction, backscatter and depolarization ratio, described in A-EBD, the algorithm uses the lidar signals directly and does not rely on a priori assumptions on the backscatter-to-extinction ratio. The output of this algorithm will serve as an input to the 1-km scale algorithm.
Retrieving the value of the extinction coefficient is a relatively simple procedure for signals with a sufficient SNR (Signal-to-Noise ratio). This procedure involves the determination of the vertical differentiation of the logarithm of the range corrected Rayleigh signal. However, this direct procedure cannot be applied directly in the case of the EarthCARE signals. Due to the relatively low SNR associated with the EarthCARE Rayleigh channel, the direct differentiation will lead to an amplification of the extinction variance and to an unreliable retrieval of the extinction. It is therefore important to take into account the low expected SNR of the EarthCARE signals. It should be noted that, similar arguments apply to the derivation of the backscatter, however, the extinction retrieval is more sensitive to the SNR of the input lidar signals. Different methods will be used to deal with the noisy nature of the signals.

The SNR value is directly retrieved by processing, for a number of lidar shots, the ratio between the mean value of the signal, and the corresponding standard deviation.

To increase the SNR before any differentiation the following methods are used:

- **In the horizontal dimension**, the signal is averaged using a sliding window. The width of this sliding window is variable and is adjusted according to the actual signal to noise ratio of the signals and a configurable threshold SNR of the lidar signals needed to obtain the required accuracy of the extinction and backscatter products. The lower limit of the required SNR has been determined in a sensitivity study. The configuration and the accuracy of each of the methods are presented in Section 6.1.

- **In the vertical dimension**, the signal is smoothed using a linear fitting procedure, with a fixed window width, before the signal derivative is calculated. The configuration and the accuracy of this method is discussed in Section 6.1.

3.3. Physical/mathematical Background

Note : In the following of the documents, we use an aerosol scene to test the validity of our algorithm. This aerosol scene is a scene from the ICAROH project. It has been registered on the 4th of June 2006, above Morocco, with the input extinction shown in Figure 6. Dust constitute the main type in this scene.

3.3.1. General lidar equation

In this section a mathematical description of the lidar signal and the retrieval of the optical properties is given. We first discuss the case of a simple elastic backscatter signal. Under single scattering conditions, the power of atmospheric backscatter detected from line-of-sight range r can be written as:

$$P(r,\lambda)=\frac{C_{lid}}{r^2}\beta(r,\lambda)\exp[-2\int_0^r \alpha(R,\lambda)\,dR]. \quad (1)$$

Where C_{lid} is the lidar calibration constant, $\beta(r,\lambda)$ is the volume backscatter coefficient, and $\alpha(r,\lambda)$ is the volume extinction-coefficient. The integral of the extinction coefficient over range r is termed “optical thickness”, the exponential term is denoted “optical transmission”. For clarity of the presentation the spectral dependence of these quantities will not be noted below. Both the volume extinction and backscatter coefficients consist of molecular and particle scattering components :

$$\beta(r)=\beta_m(r)+\beta_a(r) \quad (2)$$

$$\alpha(r)=\alpha_m(r)+\alpha_a(r) \quad (3)$$

where the indexes *m* and *a* denote the molecular and particle scattering components, respectively. Thus, the basic lidar equation 1 includes four unknowns. The molecular scattering quantities $\alpha_m(r)$ and $\beta_m(r)$ can be calculated using Rayleigh theory when the atmospheric temperature and pressure along the measurement range are known. The resulting equation remains under-determined, such that the direct retrieval of $\alpha_a(r)$ and $\beta_a(r)$ is not possible using one measured signal only. Provided that there is a pre-defined relation between aerosol extinction and aerosol backscatter, the so-called lidar ratio :

$$S(r)=\frac{\alpha_a(r)}{\beta_a(r)} \quad (4)$$

Then Equation (1) can be solved using inversion techniques (i.e. [Klett, 1981]). The lidar ratio is an aerosol and cloud specific quantity, which depends on the cloud/aerosol size distribution, its complex index of refraction, and morphology ([Ackermann, 1998], [Evans, 1988]). Raman-lidar observations have shown lidar ratio variations from 23-29 sr in the case of marine aerosol, 35-59 sr in the case of mineral dust and 37-65 sr in the case of south/east Asian aerosol ([Müller et al., 2007]). Due to these variations, large errors of standard backscatter retrievals must be expected if the lidar ratio is not well known.

In contrast to the case of a simple single-wavelength elastic backscatter lidar,applying the HSRL method, aerosol extinction and backscatter coefficients can be directly retrieved without making assumptions on the lidar ratio [see Piironen et al, 1994, Eloranta et al., 2005]. A High spectral resolution lidar (HSRL) measures two atmospheric signals using the difference in molecular and particular backscatter behaviour: The received atmospheric backscatter is split into two paths using a spectrally filter. Figure 2 shows how the narrow bandwidth aerosol scattering peak is filtered by means of an Fabry-Perot Etalon (FPE) .

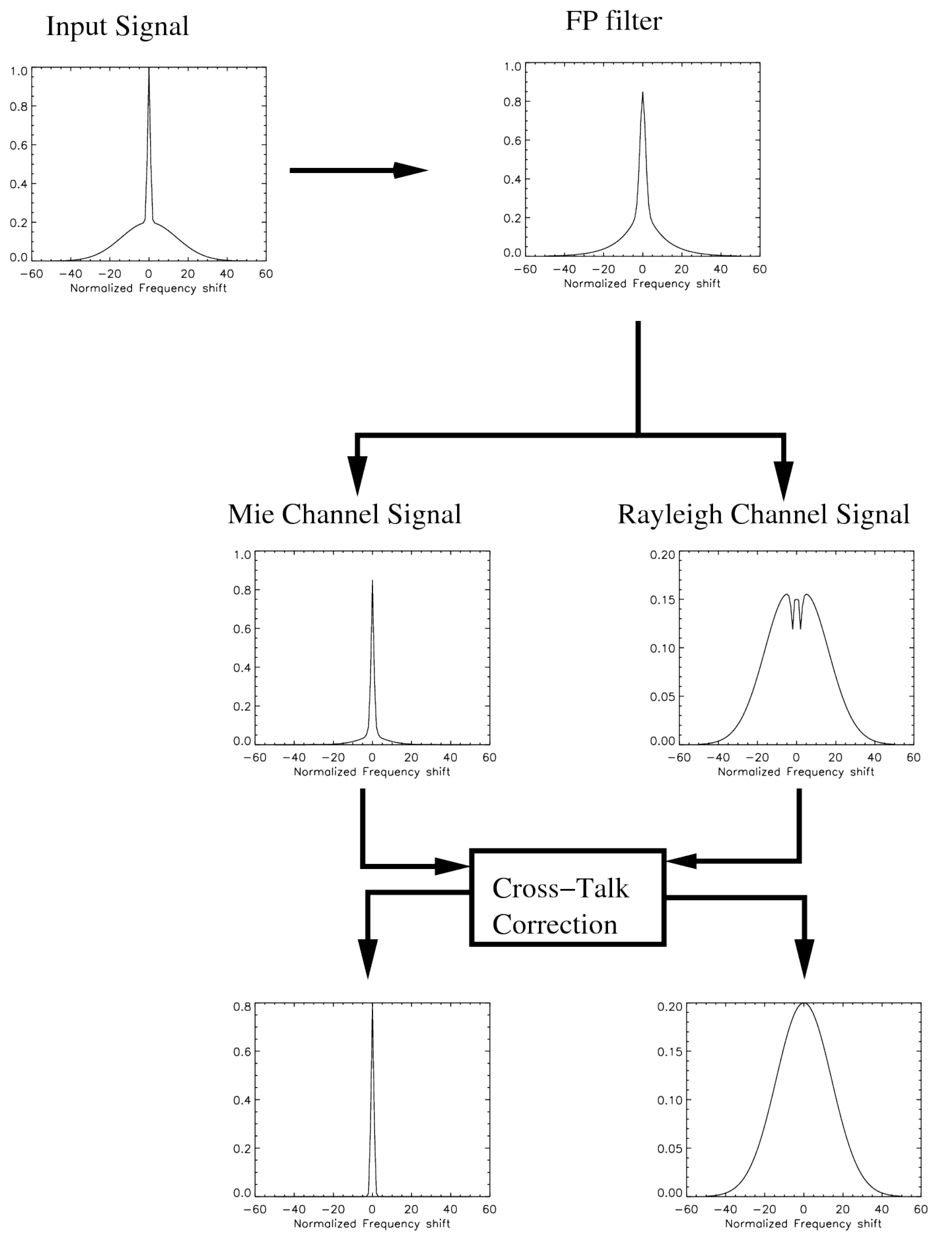

Figure 2: Principle of HSRL detection as employed by ATLID.

The FPE serves to (imperfectly) separate the thermally Doppler-broadened Rayleigh return from the much narrower Mie return. After a cross-talk correction is applied (part of the L1b processing) the Rayleigh and Mie backscatter signals are separately reported.

The lidar equations for the Mie and Rayleigh(Molecular) channels can be written as :

- For the Mie channel:

$$P_M(r)=\frac{C_M}{r^2}[\beta_a(r)]\exp[-2\int_0^r(\alpha_m(r)+\alpha_a(r))dR], \tag{5}$$

Where C_M is the lidar calibration constant for the Mie channel.

- For the Molecular channel:

$$P_R(r)=\frac{C_R}{r^2}[\beta_m(r)]\exp[-2\int_0^r(\alpha_m(r)+\alpha_a(r))dR], \tag{6}$$

Where C_R is the lidar calibration constant for the Rayleigh channel.

Since the Rayleigh backscatter is directly proportional to the (known) atmospheric density profile, equation 6 can be solved to retrieve aerosol extinction and subsequently equation 5 can be solved to calculate the aerosol backscatter.

3.3.2. Cloud Screening

Since the relationship between extinction and lidar signal is non-linear, "blindly averaging" the lidar signal over a significant horizontal domain will lead to biased results if large horizontal variations in the optical properties are present. Thus, in order to accurately retrieve aerosols properties, the algorithm must be able to deal with any potential cloud contamination inside the aerosols scenes. The clouds are detected and any potential aerosol regions below the detected clouds are not taken into account when smoothing the data. In this work clouds are identified by applying a threshold on the input feature mask [AFM-FM].

As a minimum SNR is required before any retrieval can be attempted, the algorithm must be able to resize the sliding window width that will be used in the next step (see next section 3.3.3.). The SNR ratio of the average data is calculated using the standard deviation of the cloud screened signal within the window.

The current algorithm has only been tested using cloud free aerosol fields. Future improvements of the algorithm are needed to test and validate the cloud screening algorithm. This will require a call to the (A-TC) module. The algorithm will then perform the retrievals of aerosols properties on aerosols type pixels.

3.3.3. Average in the horizontal dimension to retrieve required SNR

In order to perform the inversion to retrieve the aerosols properties, a minimum SNR is required. The signals are smoothed using a box-car window at every altitude level. The width of the box is increased until a minimum SNR threshold is reached for the entire profile.

The lidar signals can be binned and filtered to a horizontal resolution from a minimum of 10 km up to a maximum of 150 km (tbd) depending on the SNR within the profile, but will always be provided at a 1 km bin-size (via application of a sliding window).

It must be noted that in the case of the occurrence of cloud features inside the aerosols scene, the horizontal smoothing width will depend on the cloud screening itself [see Section 3.3.2.].

The dependence of the SNR on the bin size has been studied for a small number of scenes and an example is shown in Figure 3 [see ATL-FINAL document for more details], e.g. for a scene with a total aerosol optical thickness (AOT) of 0.05, 2km binning will result in SNR=2 and 10km binning a SNR=10.

It must be noted that increasing the size of the sliding window in the vertical dimension can compensate for reducing the width of the window in the horizontal dimension. As is shown in Section 6.1., a very high SNR in the Rayleigh signals is needed in order to retrieve an extinction with a SNR of 3. This requires a horizontal width of 20km and vertical smoothing window of 0.9km in the discussed example.

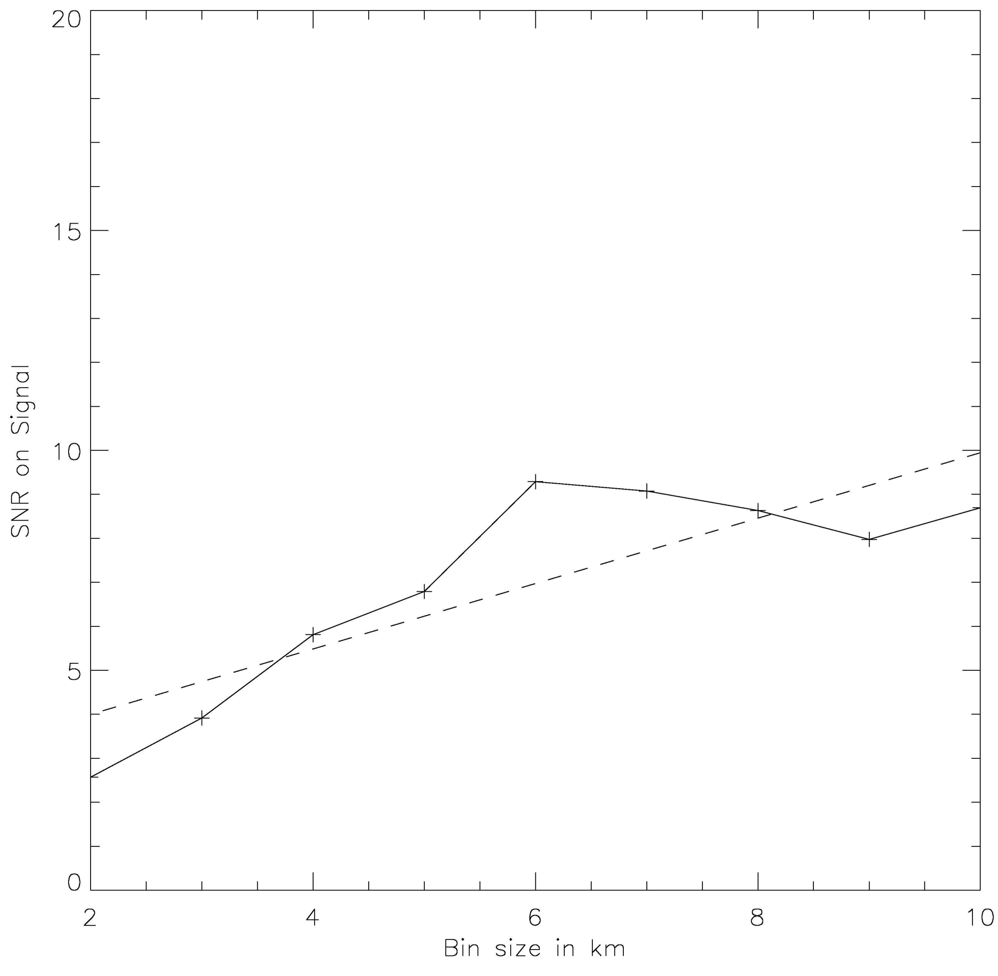

Figure 3: Relationship between the backscatter SNR and the bin size in km (AOT=0.05)

3.3.4. Extinction coefficient retrievals

3.3.4.1. Process the range corrected logarithm Rayleigh signal

The extinction may be estimated by evaluating the derivative of the logarithm of the range corrected Rayleigh signal, term that we call here $P(z)$: Starting from Equation 6 it can easily be shown that

$$\alpha(z)=-0.5\frac{d}{dz}\log\left(\frac{(r^2(z))P_R(z)}{\rho(z)}\right)=-0.5\frac{dB(z)}{dz} \tag{7}$$

where

$$B_l(z)=\log\left(\frac{(r^2(z))P_{Ray}(z)}{\rho(z)}\exp\left(2.\int_0^r \alpha_R(r)\,dr\right)\right) \tag{8}$$

β(z) is the range corrected Rayleigh backscatter signal corrected for Rayleigh attenuation.

3.3.4.2. Linear least Square fitting and derivative of the fitted signal

The main difficulty in deriving an accurate value for the extinction depends on having a ‘smooth’ Rayleigh signal. If the signal is too strongly influenced by noise, the errors become unacceptable large. Therefore, direct methods are normally not usable, except in the case of very high signal-to-noise ratio. So, the main issue in estimating the Rayleigh extinction lies in the ability to extract the ‘true’ derivative from the noisy signal. There are a number of possible solutions to this problem, each with their own pros and cons.

The resulting (smoothed) Rayleigh profile is used to calculate the extinction. This is similar to the Simple Direct Approach (as in the equation 7). However, a more sophisticated filtering strategy is employed to obtain the derivative of the range corrected logarithmic Rayleigh signal. The goal of the filtering is to retrieve a smooth signal to lower any noise in the process of extracting the extinction. In essence, the derivative of the range corrected logarithmic Rayleigh signal is estimated using a statistical fitting procedure, named the method of least square, applied to a sliding window (in altitude). The width of the sliding fitting window can vary for different approaches.

The method of least squares assumes that the best-fit curve of a given type is the curve that has the minimal sum of the deviations squared (least square error) from a given

set of data. Suppose that the data points are $(x1, y1), (x2, y2), \ldots, (x_n, y_n)$ where x he independent variable and y is the dependent variable. The fitting curve $P(z)$ has the deviation(error) d for each data point, i.e., $d_1 = y_1 - f(x_1), d_2 = y_2 - f(x_2), \ldots, d_n = y_n - f(x_n)$. According to the method of least squares, the best fitting curve has the property that minimize the cost function :
.

$$\prod = d_1^2 + d_2^2 + \ldots + d_n^2 = \sum_1^n d_i^2 = \sum_{i=1}^n [y_i - f(x_i)]^2 \quad .$$

Polynomials are among the most common types of curves used in a regression. In the case of this algorithm, the order of the polynomial resulting in the best fitting results had to be determined. A sensitivity study was performed to compare the linear, quadratic and cubic least-squares line methods. For the lidar signals, created using the EarthCARE simulator, the linear regression retrieved the most accurate results (see Final Review Document), and therefore, this method is described within the remainder of this section.

The least-squares linear method uses a straight line $y = ax + b$ to approximate the given set of data $(x1, y1), (x2, y2), \ldots, (x_n, y_n)$, where $n \geqslant 2$.

The best fitting curve $f(x)$ has the least square error, i.e.,

$$\prod = \sum_{i=1}^n [y_i - f(x_i)]^2 = \sum_{i=1}^n [y_i - (a_0 + a_1 x_i)]^2 = min$$

with $a_{0,} a_1$ the unknown coefficients and x_i and y_i given. To obtain the least square error, the unknown coefficient $a_{0,} a_1$ must yield zero first derivatives.

$$\frac{\partial \prod}{\partial a_0} = 2 \sum_{i=1}^n [y_i - (a_0 + a_1 x_i)]^2 = 0 \qquad (9)$$

And :

$$\frac{\partial \prod}{\partial a_1} = 2 \sum_{i=1}^n x_i [y_i - (a_0 + a_1 x_i)]^2 = 0 \qquad (10)$$

The unknown coefficients a_0, a_1 can hence be obtained by solving the above linear equations.

In our practical case x_i is equivalent to $r(z_i)$, and yi is equivalent to $\ln\left(\frac{z_i^2 P_{Ray}(z)}{\rho(z)}\right)$

Yielding the least-squares estimates for this slope as :

$$a_1 = \frac{\left[N.\sum_{i=1}^{N} y_i x_i\right] - (\sum_{i=1}^{N} x_i)(\sum_{i=1}^{N} y_i)}{\left[N.\sum_{i=1}^{N} x_i^2\right] - (\sum_{i=1}^{N} x_i)^2} = \frac{Cov[x_i, y_i]}{\sigma_{x_i}} \quad (11)$$

The total extinction coefficient correspond simply to the $-\frac{a_1}{2}$, as developed inside Equation 7.

3.3.4.3. Error Assessment

As the calculation of the extinction is based upon parameters which each have their uncertainties, the propagation of these errors have to be assessed.

Most commonly the error on a quantity Δx, is given as the standard deviation, σ. Standard deviation is the positive square root of variance, σ^2.As the signals are smoothed in the vertical the variables are highly correlated, therefore the covariance must be taken in account. In this section we will describe how the error covariance matrix corresponding to the reported extinction product is derived.

<u>Calculation of the extinction error</u>

The extinction can be written as a function of the lidar log attenuated backscatter :

$$\alpha_k = \alpha(B_{k=1}, B_{k=2}, B_{k=3}, B_{k=4}, \ldots, B_{k=n}) \quad (12)$$

Where n is the size of the vertical profile and B_k is the Rayleigh lidar signal at the altitude level k.

Following standard practice, the covariance matrix for the extinction can be related to the covariance matrix of the Rayleigh Lidar signal (M_{ij}^B) as follows

$$M_{ij}^{\alpha} = A^T . M_{ij}^B . A \quad (13)$$

Where $[A_{ij}] = \frac{\partial \alpha_i}{\partial B_k}$

The signals at different altitudes are uncorrelated with each other so that $[M_{ii}^B] = \sigma_{B_{i,i}}^{\ 2}$ and $[M_{ij}^B] = 0$

However, due to the fact that a sliding window has been used, the extinction at one level **k** is dependent on the signals only within the window limits i.e. $\left\{B_{k-\frac{N-1}{2}}, \dots, B_k, \dots, B_{k+\frac{N-1}{2}}\right\}$ the errors between different extinctions are not all independent. In fact, for example in the case of N=3) the extinction error covariance matrix will have the form :

$$M_{ij}^{\alpha}=\begin{pmatrix} \sigma_{\alpha_1}^2 & cov_{\alpha_1,\alpha_2} & 0 & 0 & 0 & \dots & 0 \\ cov_{\alpha_2,\alpha_1} & \sigma_{\alpha_2}^2 & cov_{\alpha_2,\alpha_3} & 0 & 0 & \dots & 0 \\ 0 & cov_{\alpha_3,\alpha_2} & \sigma_{\alpha_3}^2 & cov_{\alpha_3,\alpha_4} & 0 & \dots & 0 \\ 0 & 0 & cov_{\alpha_4,\alpha_3} & \sigma_{\alpha_4}^2 & cov_{\alpha_4,\alpha_5} & \dots & 0 \\ 0 & 0 & 0 & cov_{\alpha_5,\alpha_4} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \sigma_{\alpha_{n-1}}^2 & cov_{\alpha_{n-1},\alpha_{n-1}} \\ 0 & 0 & 0 & 0 & \dots & cov_{\alpha_n,\alpha_{n-1}} & \sigma_{\alpha_n}^2 \end{pmatrix} \quad (14)$$

Partial derivatives :

To proceed further we must calculate the elements of A. For each altitude level k, partial differentiation of the extinction, by the elements B_i belonging to the sliding window must be determined.

The dependence of the extinction on P can be written as :

$$\alpha(k)=\frac{1}{2}\frac{\partial}{\partial z}(\log[\frac{B_{Ray}(k)}{\alpha_{Ray}(k)}\exp(\tau_{Ray}(k)).r^2(k)]) \quad (15)$$

so that

$$\frac{\partial\alpha_k}{\partial B_j}=\frac{1}{2}\frac{\partial}{\partial B_j}\frac{\partial}{\partial z}(\log[\frac{B_{Ray}(k)}{\alpha_{Ray}(k)}\exp(\tau_{Ray}(k)).r^2(k)]) \quad (16)$$

which can be written as :

$$\pi_i=\log[\frac{B_{Ray,i}}{\alpha_{Ray,i}}\exp(\tau_{Ray,i}).(r_i^2)] \quad (17)$$

The linear fitting is performed on the value $\boldsymbol{\pi_i}$ belonging to the sliding window. This fitting will thus lead to the linear expression $\boldsymbol{\pi} = \mathbf{a.z+b}$ whose differentiation by **z** will give the slope **a**.

The least-squares estimates for this slope is given by :

$$a=\frac{\left[N.\sum_{i=1}^{N} y_i r_i\right]-(\sum_{i=1}^{N} r_i)(\sum_{i=1}^{N} y_i)}{\left[N.\sum_{i=1}^{N} r_i^2\right]-(\sum_{i=1}^{N} r_i)^2}=\frac{Cov[r_i,y_i]}{\sigma_{r_i}} \tag{18}$$

Where N is the size of the sliding window. Hence it can be written :

$$\frac{\partial\alpha_k}{\partial B_j}=\frac{1}{2}\frac{\partial}{\partial B_j}\left(\frac{Cov[z_i,\pi_i]}{\sigma_{z_i}}\right) \tag{19}$$

The quotient can be derived. As the differentiation of a variance of z is null, it was found that :

$$\frac{\partial\alpha_k}{\partial B_j}=\frac{1}{2.\sigma_{z_i}}\frac{\partial Cov[z_i,\pi_i]}{\partial B_j} \tag{20}$$

As :

$$\frac{\partial\pi_j}{B_j}=\frac{1}{B_j} \tag{21}$$

Hence, equation 10 can be written again as :

$$\frac{\partial\alpha_k}{\partial B_j}=\frac{1}{2.\sigma_{z_i}}.\left(\frac{N.z_j-\sum_{i=1}^{N} z_i}{B_j}\right) \tag{22}$$

and

$$\boxed{[A_{i,j}]=\frac{\partial\alpha_k}{\partial B_j}=\frac{N(z_j-\bar{z}_i)}{2B_j\sigma_{z_i}}} \tag{23}$$

and finally we can then calculate the extinction error covariance matrix using Equation 9.

3.3.5. Backscatter coefficient retrievals

3.3.5.1. Retrieval algorithm

The HSRL measurements results in two signals : firstly the molecular scattering (Rayleigh scattering), and secondly the aerosol scattering (Mie scattering). From these measurements, the range corrected lidar return can be retrieved.

For the Rayleigh attenuation backscatter we have :

- $$B_{Ray}(r)=\beta_{Ray}(r).\exp(-2\tau(r))=\frac{r^2}{C_{Ray}}.P_R(r) \tag{24}$$

For the Mie attenuation backscatter we have :

- $$B_{Mie}(r)=\beta_{Mie}(r).\exp(-2\tau(r))=\frac{r^2}{C_{Mie}}.P_M(r) \tag{25}$$

From these quantities the unattenuated backscatter from the HSRL lidar signals can be retrieved using two different methods :

The Simple Direct Ratio Approach (Method 1) : The aerosol backscatter profile can be directly estimated from the HSRL Mie and Rayleigh signals. The ratio of the two signals can be directly related to the unattenuated backscatter :

$$\beta_{Mie}(r)=\frac{B_{Mie}}{B_{Ray}}.\beta_{Ray}(r) \tag{26}$$

This can be accomplished by taking the ratio of the calibrated Mie channel return and the Rayleigh return signals. This method is relatively simple to implement, and easy to use when characterizing the aerosol fields.

Extinction corrected backscatter approach (Method 2) : In the case of this method, the aerosol extinction as previously retrieved can be used (see the section 3.3.4.), and the backscatter equation can be written using equation (24)

$$\beta_{Mie}(r)=B_{Mie}.\exp\left(2\int_0^r \alpha_{Ray}(z').dr\right).\exp\left(2\int_0^r \alpha_{Mie}(z).dr'\right) \tag{27}$$

Where α_{Mie} is given by equation (7).

3.3.5.2. Error Assessment

Concerning the method 1 (Direct estimation of the backscatter), the error on the Mie backscatter can be directly written as the sum of other variances corresponding to each input parameters :

$$\left(\frac{\sigma_{\beta_{Mie}}}{\beta_{Mie}}\right)^2=\left(\frac{\sigma_{B_{Mie}}}{B_{Mie}}\right)^2+\left(\frac{\sigma_{B_{Ray}}}{B_{Ray}}\right)^2+\left(\frac{\sigma_{C_{Mie}}}{C_{Mie}}\right)^2+\left(\frac{\sigma_{C_{Ray}}}{C_{Ray}}\right)^2 \tag{28}$$

We can write the constant error term as :

$$\sigma_{C_{Ray},C_{Mie}}{}^{2}=\beta_{Mie}^{2}\left[\left(\frac{\sigma_{C_{Ray}}}{C_{Ray}}\right)^{2}+\left(\frac{\sigma_{C_{Mie}}}{C_{Mie}}\right)^{2}\right] \quad (29)$$

And hence :

$$\sigma_{\beta_{Mie}}{}^{2}=\beta_{Mie}^{2}\cdot\left[\left(\frac{\sigma_{B_{Mie}}}{B_{Mie}}\right)^{2}+\left(\frac{\sigma_{B_{Ray}}}{B_{Ray}}\right)^{2}\right]+\sigma_{C_{Ray},C_{Mie}}{}^{2} \quad (30)$$

As both signals are not smoothed in the vertical there is no need to calculate the covariance matrix for this product.

3.3.5.3. Retrieval algorithm for Low Resolution backscatter product

In order to allow the user to calculate the lidar ratio and the depolarization, it is needed to provide a low resolution backscatter product, whose resolution will fit the extinction coefficient one.

This average is directly based on the knowledge we gain during the retrieval of the extinction coefficient, with the linear fitting method. During these steps, and for each altitude range, the vertical size of the sliding window used is stored, and used to perform the average of the full resolution backscatter coefficient, as processed by the Simple Direct Ratio Approach (Method 1, see section 3.3.5.1.).

NB : During the average process, each value within the sliding window will have exactly the same weight (box-cart function). This is due to the fact that a linear fitting method is used during the assessment of the extinction. This process would be more complex in the case where higher order would be used inside the fitting function.

3.3.5.4. Error Assessment on Low Resolution backscatter product

The error made in the assessment of the low resolution backscatter product, at each altitude range k, is directly derived from the variance processed at full resolution in the Simple Direct Ratio Approach (see equation 8) :

$$\sigma_{\beta_{Mie,LR}}(k)=\beta_{Mie,LR}(k)\cdot\sqrt{\sum_{i=k-\frac{N}{2}}^{k+\frac{N}{2}}\left[\left(\frac{\sigma_{B_{Mie}}(i)}{B_{Mie}(i)}\right)^{2}+\left(\frac{\sigma_{B_{Ray}}(i)}{B_{Ray}(i)}\right)^{2}+\left(\frac{\sigma_{\delta_{C_M}}(i)}{C_M(i)}\right)^{2}\right]} \quad (31)$$

NB : As stated in the previous section, the retrieval of the backscatter is simplified due to the fact that a linear fitting method is used during the retrieval of the extinction. That give exactly the same weight among the value used to process the average of the backscatter coefficient.

The formalism of covariance matrix for the backscatter coefficient is similar to the extinction coefficient one (see equation 14).

To limit the product size, no covariance matrix for the low resolution backscatter profile is reported. Such information can be easily constructed by the end user since both the backscatter error at native resolution is reported along with the resolution of the low resolution backscatter product. However, this product can be provide if required.

3.3.6. Depolarization retrieval procedures

3.3.6.1. Retrieval algorithm

Discrimination between water and ice clouds can be achieved using the depolarization properties of the backscatter signal. The depolarization ratio (δ) correspond to the ratio of the Mie cross-polar and the Mie co-polar Backscatter. In the case of the HSRL lidar, separate depolarizations of the aerosol and the molecular backscattering can be measured.

This relation can be thus written as :

$$\delta = \frac{\beta_{Mie,cr}}{\beta_{Mie,co}} \tag{32}$$

<u>NB:</u> As, $\beta_{Ray,cr} \ll \beta_{Mie,cr}$ and because the Rayleigh part hasn't been separated yet inside the cross-polar channel, a good approximation of the particle depolarization is provided by the direct signal ratio :

$$\delta = \frac{\beta_{cr}}{\beta_{Mie,co}} \tag{33}$$

Figure 4 gives the histogram of the depolarization ratio as retrieved inside the aerosol layer on the same aerosols scene that has been used for Extinction and Backscatter retrievals. The mean depolarization ratio is equal to 0.25, with a standard deviation of 0.16.

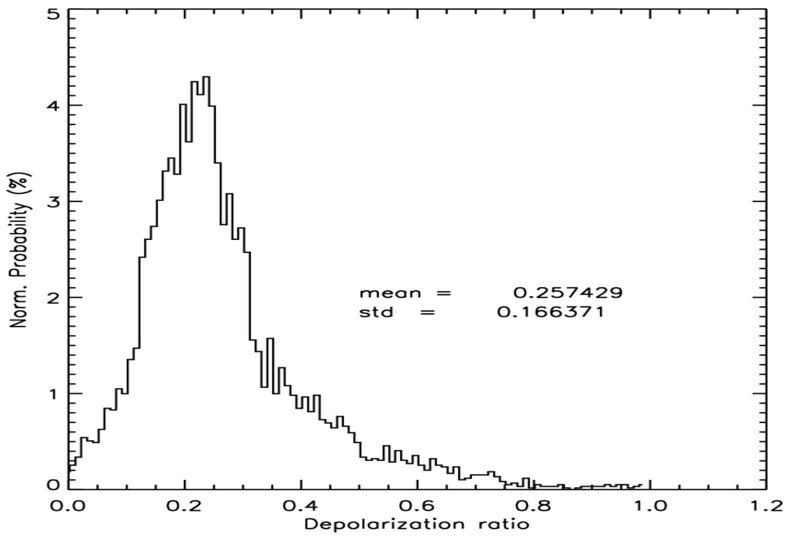

Figure 4: Histogram of the depolarization ratio inside the aerosol layer.

3.3.6.2. Depolarization ratio error assessment

The error on the depolarization ratio can be directly written as the sum of other variances corresponding to each input parameters :

$$\sigma_\delta{}^2 = \delta^2 . \left[\left(\frac{\sigma_{B_{Mie,cr}}}{B_{Mie,cr}} \right)^2 + \left(\frac{\sigma_{B_{Mie,co}}}{B_{Mie,co}} \right)^2 \right] + \sigma_{C_{CR},C_{Mie}}{}^2 \qquad (34)$$

With the constant error term :

$$\sigma_{C_{CR},C_{Mie}}{}^2 = \delta^2 . \left[\left(\frac{\sigma_{C_{cr}}}{C_{cr}} \right)^2 + \left(\frac{\sigma_{C_{Mie}}}{C_{Mie}} \right)^2 \right] \qquad (35)$$

4. Justification for the selection of the algorithm

This algorithm retrieves the extinction coefficient, backscatter coefficient, particle depolarization ratio and aerosol type probabilities.

The **extinction coefficient** is retrieved by a process based on finding the derivative of the range corrected logarithmic Rayleigh signal. Due to low expected SNR of that the Rayleigh signal, the method cannot be used to retrieve the optical properties using the signals directly. In other words, simple differencing techniques to determine the signal derivative usually lead to unsatisfactory results. The use of least square fitting methods to determine the vertical signal derivative profile is central to the extinction algorithm presented here.

Three different least-square fitting procedures have been studied: a **linear fitting**, a **quadratic fitting** and a **cubic fitting procedure**. A sensitivity study comparing the three procedures has been performed in order to decide which of these is the most suitable method. For each of the methods, the SNR of the retrieved extinction is given as a function of the sliding window size (in the horizontal and vertical dimension). As a conclusion of this study, it was found that the best and most robust results were obtained using the linear fitting method. This method is therefore adopted in the current algorithm development.

Considering the inversion of the **backscatter coefficient**, two methods have been tested : firstly a direct retrieval of the backscatter based on dividing the Mie and Rayleigh Signals, and secondly, one which calculates the unattenuated backscatter using the attenuated backscatter and the retrieved extinction coefficient (see section 5.5.4.1.). The two methods have been compared for a single scene, for which the details are presented in [ATL-FINAL] document. This sensitivity study favoured the first method, which has been implemented in the current processing chain. The backscatter is reported at two vertical resolutions. Namely, the backscatter is reported at native resolution and, for ease of comparison, at the same resolution as the corresponding extinction product.

The ratio of the Mie cross-polar and the Mie co-polar Backscatter coefficient is processed to retrieve the depolarization ratio parameter.

As the typing of the aerosol consist of a call to the L2a classification module (A-TC), the full description of the aerosol typing itself is not provided within this document. Only a short description is given in this document.

5. Mathematical algorithm Description

5.1. Input parameters

Input is read from the following products :

- The perpendicular and the parallel lidar signal, the satellite and lidar instrument configuration (Latitude, Longitude, time, height) is read from the L1b product.
- The A-FM product is read to get the Feature mask (separation of clouds, molecules and aerosols).

- The A-TC module is call by the Algorithm to get the target classification and the aerosol typing.

- The ancillary datasets from ECMWF.

Variable	Symbol	Description	Unit	Dim	Type	Sources	
Time	**UTC**	UTC time	s	Time	double	L1b	
Latitude	**LAT**	Latitude	deg	Time	real	L1b	
Longitude	**LON**	Longitude	deg	Time	real	L1b	
surface_altitude	$\mathbf{Z_{surf}}$	Height of surface above mean sea level	m	Time	real	L1b	
Height	**Z**	Height of each radar/lidar gate above mean sea level	m	Time	Real	L1b	
$\mathbf{C_{Ray}}$	$\mathbf{C_{Ray}}$	Lidar calibration constant for the Rayleigh Channel	-	Time	Real	L1b	
$\mathbf{C_{Mie}}$	$\mathbf{C_{Mie}}$	Lidar calibration constant for the Mie Channel	-	Time	Real	L1b	
Tot_perp	$\mathbf{B_{Tot,\perp}}$	Cross-Talk corrected background-subtracted Total Perpendicular return.	Photon Counts per Shot	Time, Height	Real	L1b	
Tot_perp_err	$\boldsymbol{\sigma_{B(Tot,\perp)}}$	Standard deviation Cross-Talk corrected background-subtracted Total Perpendicular return.	Photon Counts per Shot	Time, Height	Real	L1b	

Mie_para	$\mathbf{B_{Mie,//}}$	Cross-Talk corrected background-subtracted Mie parameter return.	Photon Counts per Shot	Time, Heigh t	Real	L1b	
Mie_para_err	$\boldsymbol{\sigma}_{\mathbf{B(Mie,//)}}$	Standard deviation of the Cross-Talk corrected background-subtracted Mie parallel return.	Photon Counts per Shot	time, height	Real	L1b	
Ray_para	$\mathbf{B_{Ray,//}}$	Cross-Talk corrected background-subtracted Rayleigh parallel return.	Photon Counts per Shot	time, height	Real	L1b	
Ray_para_err	$\boldsymbol{\sigma}_{\mathbf{B(Ray,//)}}$	Standard deviation of the Cross-Talk corrected background-subtracted Rayleigh parallel return.	Photon Counts per Shot	time, height	Real	L1b	

Table 1: Operational input parameters.

Variable	Symbol	Description	Unit	Dim	Type	Sources	Remarks
L2A-FeatureMask							
Time	$\mathbf{UTC_{FM}}$	UTC Time	S	Time	Real*8	A-FM	
Latitude	$\mathbf{LAT_{FM}}$	Latitude of the ATLID footprints	Deg.	Time	Real	A-FM	
Longitude	$\mathbf{LON_{FM}}$	Longitude of the ATLID footprints	Deg.	Time	Real	A-FM	
Height	$\mathbf{Z_{FM}}$	Height above mean sea level	m	Height	Real	A-FM	
Surface_Altitude	$\mathbf{Z_{surf,FM}}$	Height of surface above mean sea level	m	Time	Real	A-FM	
Mask_Pa	-	Cloud/ Aerosols Mask produced from A-FM product	None	Time, Height	Integer	A-FM	-2 = below ground surface -1 = totally extinguished data 0 = most likely molecular 1-5 = likely molecular but increasing chance of being a feature. 6-9 = likely feature, decreasing chance of being molecular 10 = most likely feature detection

Block_start_end	-	Boundaries of the data blocks for which the featuremask is derived.	None	nblocks	Integer	A-FM	
Ancillary datasets from ECMWF							
ECMWF_PRES	**P**	Pressure (ECMWF Field Code 54)	Pa	Height, Time	Real	ECMWF	
ECMWF_T	**T**	Temperature (ECMWF Field Code 130)	K	Height, Time	Real	ECMWF	
ECMWF_R		Relative humidity (ECMWF Field Code 157)	%	Height, Time	Real	ECMWF[-1000, 30000]	
ECMWF_Hgt	$\mathbf{Z_{ECMWF}}$	Height	m	Height, Time	Real	ECMWF	
ECMWF_WBT	T_{wb}	Wet-Bulb Temperature (ECMWF Bufr Field Code 012002)	K	Height, Time	Real	ECMWF	Along track
ECMWF_Trop	Z_{trop}	Tropopause	m	Time	Real	ECMWF	Along track
ECMWF_Surf_Alt	$\mathbf{Z_{Surf,ECMWF}}$	Surface Altitude (ECMWF Field Code 010001)	M	Time	Real	ECMWF	

Table 2: Operational input parameters.

5.2. Configuration parameters

Variable	Symbol	Description	Unit	Dim	Type	Sources
Nmin	N_{min}	Minimum number of pixels to be applied to the sliding window in the fitting method (impair value: 3,5,7,...)	None	1	Integer	Algorithm
Nmax	N_{max}	Maximum number of pixels to be applied to the sliding window in the fitting method (impair values : 3,5,7,...)	None	1	Integer	Algorithm
SNRmin	SNR_{min}	Minimum SNR on the Signal to be taken in account before any signal processing	None	1	Real	Algorithm
Lim_Beta_Surface	$Beta_{min}$	Threshold on Backscatter value to serve as fast ground detection.	m^{-1}	1	Real	Algorithm

Table 3: Configuration parameters.

5.3. Output parameters

5.3.1. Operational output parameters

Variable	Symbol	Description	Units	Dim	Type	Destination
		1D Coordinate variables				
Time	**UTC**	UTC time	s	Time	double	A-EBD, A-TC
Latitude	**LAT**	Latitude	Deg.	Time	real	A-EBD, A-TC
Longitude	**LON**	Longitude	Deg.	Time	real	A-EBD, A-TC
Height	**Z**	Height above mean sea level	m	Height	real	A-EBD, A-TC
		Geographic information				
surface_altitude	Z_{surf}	Height of surface above mean sea level	m	Time	real	A-EBD, A-TC
		Aerosols optical properties parameter				
ext	α	Extinction coefficient	m^{-1}	Time, Height	real	A-EBD, A-TC
D_ext	σ_α	1-sigma-estimated error	m^{-1}	Time, Height	real	A-EBD, A-TC
beta	β	Backscatter coefficient	$m^{-1}.sr^{-1}$	Time, Height	real	A-EBD, A-TC
D_beta	σ_β	1-sigma-estimated error	$m^{-1}.sr^{-1}$	Time, Height	real	A-EBD, A-TC
beta_lr	β_{LR}	Backscatter coefficient at low res. (equivalent to extinction res.)	$m^{-1}.sr^{-1}$	Time, Height	real	A-EBD, A-TC
D_beta_lr	$\sigma_{\beta_{LR}}$	1-sigma-estimated error at low resolution (equivalent to extinction res.)	$m^{-1}.sr^{-1}$	Time, Height	real	A-EBD, A-TC
depol	δ	Depolarization ratio	None	Time, Height	real	A-EBD, A-TC
D_depol	σ_δ	1-sigma-estimated error	None	Time, Height	real	A-EBD, A-TC
Vert. Res.	-	Width of window on vertical dimension used to determine extinction and low resolution backscatter	m	Time, Height	real	A-EBD, A-TC
Hor. Res.	-	Width of window on horizontal dimension used to determine extinction and low resolution backscatter	m	Time, Height	real	A-EBD, A-TC
Hor_wind_start_T	-	UTC time of starting point for horizontal averaging window	s	Time	double	A-EBD, A-TC
Hor_wind_stop_T	-	UTC time of stopping point for horizontal averaging window	s	Time	double	A-EBD, A-TC

Table 4: Operational output parameters.

Variable	Symbol	Description	Units	Dim	Type	Destination	Remarks
Quality control variables							
Status	-	Retrieval status flag. 0➔ success, 1➔ no retrieval attempted (i.e. no aerosol present),2➔retrieval failed, 3➔No data	-	Time	integer	A-EBD, A-TC	
Error covariance matrix							
Ret_ERR_COV_MAT_EXT	**A**	Extinction error covariance matrix	-	Time, (height,8), (height,8)	Real	A-EBD, A-TC	
Aerosol Typing							
Aerosol_type_names	-	Aerosol Type Names	-	Ntypes	Char	A-EBD, A-TC	
Aerosol_type	-	Aerosol Type	-	Time, height	Byte	A-EBD, A-TC	
Aerosol_prob_direct	-	Aerosol-probabilities using signals only	-	Ntypes,Time,height	Real	A-EBD, A-TC	
Aerosol_prob	-	Aerosol-probabilities using signals and additional data	-	Ntypes,Time,height	Real	A-EBD, A-TC	

Table 5: Operational output parameters.

Due to the fact that the sliding widow will vary in size on the horizontal dimension, the horizontal resolution will also vary. The real vertical resolution can be deduced from the reported error covariance matrices.

In all the case, the product will be reported at the horizontal sampling of 1km (Product resolution of 10 to 100km, re-gridded to 1km on the horizontal) and 100 meter in the vertical.

5.4. Algorithm flow charts

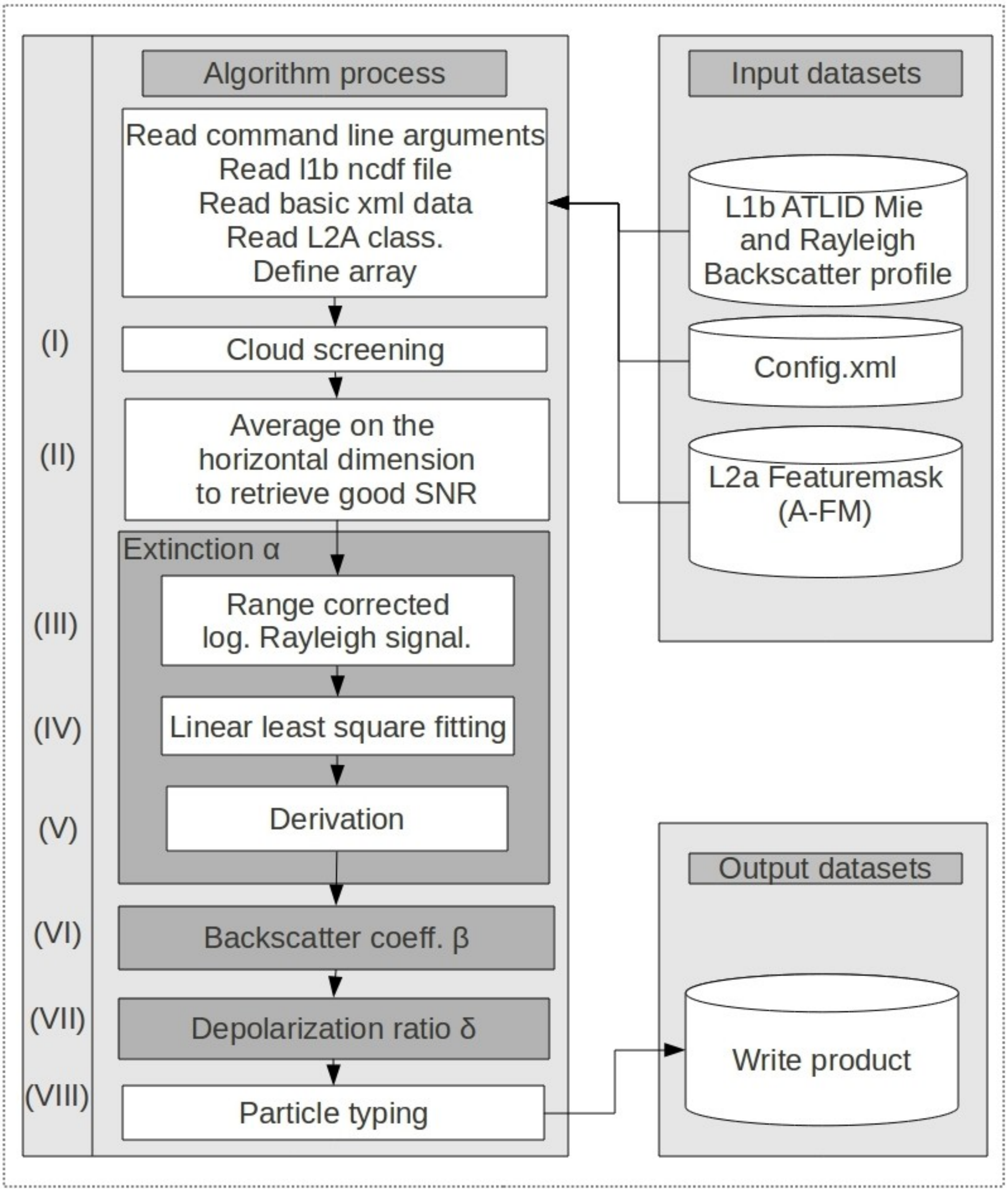

Figure 5: ATLID Inversion flow diagram.

The inversion algorithm can be defined by the succession of the following procedures. Each of these procedures are represented within the flowchart presented in Figure 5 by the block labelled with Roman number.

#	Task	Ref. Section
I.	**Cloud Screening**	**section 5.5.1.**
II.	**Average on the horizontal dim. to retrieve good SNR**	**section 5.5.2.**
III.	**α retrieval :** **Range corrected log. Rayleigh signal.**	**section 5.5.3.1.**
IV. And V.	**α retrieval :** **Linear least Square fitting and Differentiation.**	**section 5.5.3.2.**
VI.	**Backscatter coefficient retrieval.**	**section 5.5.4.**
VII.	**Depolarization ratio retrieval.**	**section 5.5.5.**
VIII.	**The Particle Typing.**	**Call to L2b Classification**

5.5. *Algorithm definition*

5.5.1. Cloud Screening

The clouds are detected and any potential aerosol regions below the detected clouds are not taken into account when smoothing the data.

Clouds are identified by applying a threshold (Th_{cld}) to the input target mask TM(t,z). Using this mask, an averaging mask is constructed such that :

- AM(t,z) = 0 where z < Zcld(t)
- AM(t,z) = 1 otherwise

Z_{cld}(t) is defined as the highest altitude bin where TM(t,z) is superior or equal to Th_{cld}.

The current algorithm has only been tested using cloud free aerosol fields. Future improvements of the algorithm are needed to test and validate the cloud screening algorithm.

5.5.2. Average in the horizontal dimension to retrieve required SNR

Signal-to-noise ratio is defined as the ratio between a signal and the background. It correspond to the the ratio of mean to standard deviation of a signal or measurement

In order to perform the inversion to retrieve the aerosols properties, a minimum SNR is required. The signals are smoothed using a box-car window at every altitude level.

The width of the box is increased until a minimum SNR [**SNRmin]** threshold is reached for the entire profile.
Here the following 3 steps are carried out :

(1) The window width is set to its default value.
(2) The average cloud scattering quantities are calculated. For example :

$$\bar{P}_M(t,z)=\frac{\left[\sum_{t=t_0-\frac{widht}{2}}^{t_0+\frac{widht}{2}} AM(t,z).P_M(t,z)\right]}{N_{samp}(z)}$$

Where :

$$N_{samp}(z)=\left[\sum_{t=t_0-\frac{widht}{2}}^{t_0+\frac{widht}{2}} AM(t,z)\right]$$

And :

$$\sigma^2_{\bar{P}_M}(t,z)=\frac{1}{N_{samp}(z)}\left[\sum_{t=t_0-\frac{widht}{2}}^{t_0+\frac{widht}{2}}\left[P_M(t,z)-(\bar{P})_M\right]^2\right]$$

similar expression hold for the calculation of $\bar{P}_R$ **,** σ_{P_R} **,** $\bar{P}_{CR}$ **,** $\sigma_{P_{CR}}$ **.**

SNR is processed as the ratio :

$$SNR_{\bar{P}_M}(z)=\frac{\bar{P}_M}{\sigma_{\bar{P}_M}}$$

(3) If $Min\left[SNR_{\bar{P}_M}(z)\right]$ **or** $Min\left[SNR_{\bar{P}_R}(z)\right]$ **> SNR$_{min}$, the widht of the window is increased and control is passed back to step (2) until the SNR threshold is reached or the maximum allowed window width is reached.**

The lidar signals can be binned and filtered to a horizontal resolution from a minimum of 10 km up to a maximum of 150 km (tbd) depending on the SNR within the profile, but will always be provided at a 1 km bin-size (via application of a sliding window).

5.5.3. Extinction coefficient retrievals

5.5.3.1. Process the range corrected logarithm Rayleigh signal

Here the attenuated backscatter profiles corresponding to the average cloud screened signals are calculated along with theirs respective errors :

$$\bar{B}_M = \frac{1}{C_M} . r^2(z) \bar{P}_M$$

and

$$\sigma_{B_M} = \frac{1}{C_M} . r^2(z) \sigma_{P_M}$$

similar expression hold for the calculation of $\bar{B}_R$, σ_{B_R} , $\bar{B}_{CR}$, $\sigma_{B_{CR}}$.

5.5.3.2. Linear least Square fitting and derivative of the fitted signal

Here the extinction coefficient profile and the associated error covariance matrix is found. This is accomplished using equation (7) and (18) for the extinction calculation, and the equation (13) and (23) for the error calculation.

5.5.4. Backscatter coefficient retrievals

5.5.4.1. Retrieval algorithm

The Simple Direct Ratio Approach (Method 1 discussed in Section 3.3.5.) : The ratio of the two signals can be directly related to the unattenuated backscatter :

$$\beta_{Mie}(r) = \frac{B_{Mie}}{B_{Ray}} . \beta_{Ray}(r) \tag{36}$$

Concerning the error on the Mie backscatter can be directly written as the sum of other variances corresponding to each input parameters, hence :

$$\sigma_{\beta_{Mie}} = \beta_{Mie} . \sqrt{\left(\frac{\sigma_{B_{Mie}}}{B_{Mie}}\right)^2 + \left(\frac{\sigma_{B_{Ray}}}{B_{Ray}}\right)^2 + \left(\frac{\sigma_{C_{Mie}}}{C_{Mie}}\right)^2 + \left(\frac{\sigma_{C_{Ray}}}{C_{Ray}}\right)^2} \tag{37}$$

We can write the constant term as :

$$\sigma_{C_{Ray},C_{Mie}}{}^2 = \beta_{Mie}^2 \left[\left(\frac{\sigma_{C_{Ray}}}{C_{Ray}}\right)^2 + \left(\frac{\sigma_{C_{Mie}}}{C_{Mie}}\right)^2\right] \tag{38}$$

And hence :

$$\sigma_{\beta_{Mie}}{}^2 = \beta_{Mie}^2 . \left[\left(\frac{\sigma_{B_{Mie}}}{B_{Mie}}\right)^2 + \left(\frac{\sigma_{B_{Ray}}}{B_{Ray}}\right)^2\right] + \sigma_{C_{Ray},C_{Mie}}{}^2 \tag{39}$$

5.5.4.2. Retrieval algorithm for Low Resolution backscatter product and error

This average is directly base on the knowledge gained during the extinction coefficient retrieval. During these steps, and at each altitude range, the vertical size of the sliding window used is stored , and used to perform the average of the full resolution backscatter coefficient, as processed by the Simple Direct Ratio Approach (Method 1, see section 5.5.4.1.).

The error made in the assessment of the low resolution backscatter product, at each altitude range k, is directly derived from the variance processed at full resolution in the Simple Direct Ratio Approach (see equation 37) :

$$\sigma_{\beta_{Mie,LR}}(k)=\beta_{Mie,LR}(k)\,.\sqrt{\sum_{i=k-\frac{N}{2}}^{k+\frac{N}{2}}\left[\left(\frac{\sigma_{B_{Mie}}(i)}{B_{Mie}(i)}\right)^2+\left(\frac{\sigma_{B_{Ray}}(i)}{B_{Ray}(i)}\right)^2+\left(\frac{\sigma_{\delta_{C_M}}(i)}{C_M(i)}\right)^2\right]} \tag{40}$$

5.5.5. Depolarization retrieval procedures

5.5.5.1. Retrieval algorithm

The depolarization ratio (denoted δ) correspond to the ratio of the Mie cross-polar and the Mie co-polar Backscatter. This relation can be thus written as :

$$\delta=\frac{\beta_{cr}}{\beta_{Mie,co}} \tag{41}$$

The error on the depolarization ratio can be directly written as the sum of other variances corresponding to each input parameters :

$$\sigma_\delta=\delta\,.\sqrt{\left(\frac{\sigma_{B_{Mie,cr}}}{B_{Mie,cr}}\right)^2+\left(\frac{\sigma_{B_{Mie,co}}}{B_{Mie,co}}\right)^2+\left(\frac{\sigma_{C_{cr}}}{C_{cr}}\right)^2+\left(\frac{\sigma_{C_{Mie}}}{C_{Mie}}\right)^2} \tag{42}$$

Writing the constant error term as :

$$\sigma_{C_{CR},C_{Mie}}{}^2=\delta^2.\left[\left(\frac{\sigma_{C_{cr}}}{C_{cr}}\right)^2+\left(\frac{\sigma_{C_{Mie}}}{C_{Mie}}\right)^2\right] \tag{43}$$

We can write :

$$\sigma_\delta{}^2 = \delta^2 . \left[\left(\frac{\sigma_{B_{Mie,cr}}}{B_{Mie,cr}} \right)^2 + \left(\frac{\sigma_{B_{Mie,co}}}{B_{Mie,co}} \right)^2 \right] + \sigma_{C_{CR}, C_{Mie}}{}^2 \quad (44)$$

5.5.6. Particle typing

After the determination of the direct retrieval products the aerosol-typing can be performed. This will be calculated using the module developed in the (A-TC) classification algorithm. A direct call to this algorithm will result in :

- the probabilities for all possible aerosol types based on the signals only
- the probabilities for all possible aerosol types based on signals and additional information
- a single type based on the highest probability

The description of the typing itself is not provided within this document. The details and assumptions are given in the (A-TC) document. The actual link between the two algorithms can only be made in a future ATLAS project when for both algorithms a working version exists at the same time.

6. Algorithm performance, sensitivity studies, limitations

6.1. Sensitivity to signal SNR, AOT on simulated datasets.

This part of the study is dedicated to the assessment of the sensitivity of the fitting method in order to retrieve the optical properties of the aerosol layers. The accuracy of each of the fitting methods can be directly linked to the Signal-to-Noise-Ratio (SNR) of the retrieved extinction.

This SNR will be mainly dependent of two parameters :

- **Sh** : The width of the sliding window that is applied in the horizontal dimension (in KM) .
- **Sv** : The width of the sliding windows that is applied in the vertical dimension to fit the signal (in pixel, Sz).

The testing and validation of the algorithm and or parts of the algorithm was performed using a specific transect from the SAMUM-1 campaign. The high resolution lidar data was observed by the DLR Falcon on the 4th of April 2006. This orbital transect concerns an area located in Morocco during a Saharan dust event. The retrieved extinction for this scene is shown in Figures 6.

The transect has been converted to an ECSIM scene using the measured lidar data and in-situ measured aerosol data. The constructed scene was used to simulated the equivalent EarthCARE lidar signals. Even though the scene shows a single dust layer up to 5 km it has a large vertical variation due to an orographical induced wave pattern at approximately 3 km. This makes the scene an interesting case with a single particle type with hardly any variation in the particle size distribution but with a large variety in the total particle number count. Any vertical (and horizontal) variation is therefore only due to the local variability in extinction. The lidar configuration used for the calculation of the EarthCARE ATLID signals can be found in Table 6.

The SNR for each of the [Sh,Sv] couple can be retrieved (see Figure 7). As expected, the SNR increases when these two parameters are increased.
A SNR equal to 3 can been seen as a reasonable retrieval accuracy in order to enable a good assessment of the aerosols optical properties. This correspond to the yellow color inside the Figure 7. The couple (Sh, Sv) should be chosen on this line. Larger horizontal width is preferable if this lowers the vertical width.

For the following of this study, we choose a parameters couple (Sh, Sv), whose values are given in the following table 7.

NB : As a limitation of this study, these values have been processed for only one aerosol scene, and therefore any scene dependence cannot be excluded. However, very few variations on these are expected in the case of aerosols only scenes. Variations on these values are expected when aerosols scenes start to be contaminated by cloud.

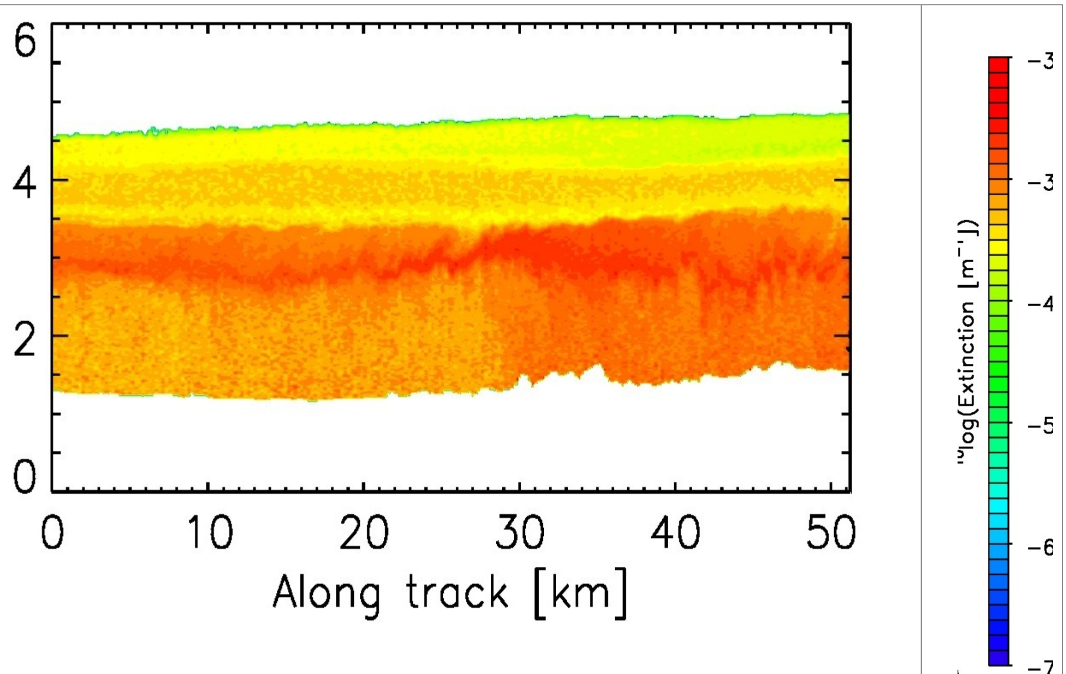

Figure 6: Input extinction.

parameter name	description	value	units
Width	Max radius from beam centre	0.25	km
Laser_PRF	Laser Pulse Rep Freq	74	Hz
nshot_sum	Number of shots to be summed	1	-
lid_wave	Lidar Wavelength	355	nm
dz_ins	Desired resolution of output	0.1	km
rho_t	Telescope full-angle FOV	0.03	mrad
rho_l	Laser full angle FOV 1/e width	0.0047	mrad
Ao	Laser effective receiver area	0.28	m^2
ND_fac	Neutral Density factor	1	-
El	Laser Pulse energy"	0.02	J
laser_lw	Laser Line width	30	MHz
hor_res	Horizontal resolution	0.25	km
vert_res	Vertical resolution	0.1	km
Detector_type	Type of detector	ACD	-
Quant_eff	Quantum Efficiency	0.74	-
Dark_counts	Photon Equivalent dark current	0.35	Hz

Table 6: Lidar Instrument configuration.

	Linear
Sv [pixel]	9
Sv [km]	0.9
Sh [km]	20

Table 7: Optimal Parameter couple Sv and Sh

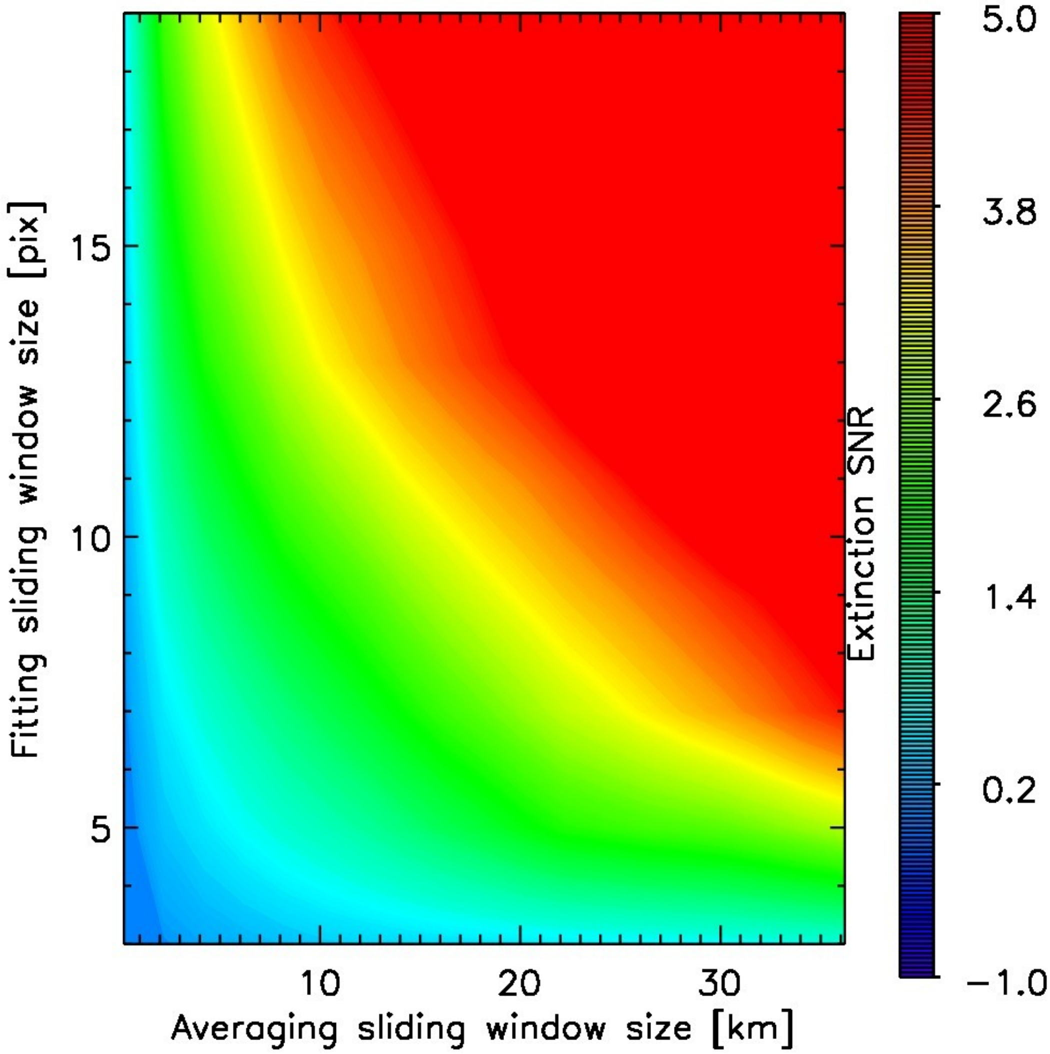

Figure 7: Extinction SNR retrieved after the application of the linear fitting method, in function of the size of the sliding windows use on the horizontal (horizontal axis) and the vertical dimension (vertical axis).

The Figure 8 depicts the results using the optimal set of Sv and Sh as was presented in Table 7. The upper panel shows the retrieved extinction, in the lower panels the area with an extinction SNR>3.

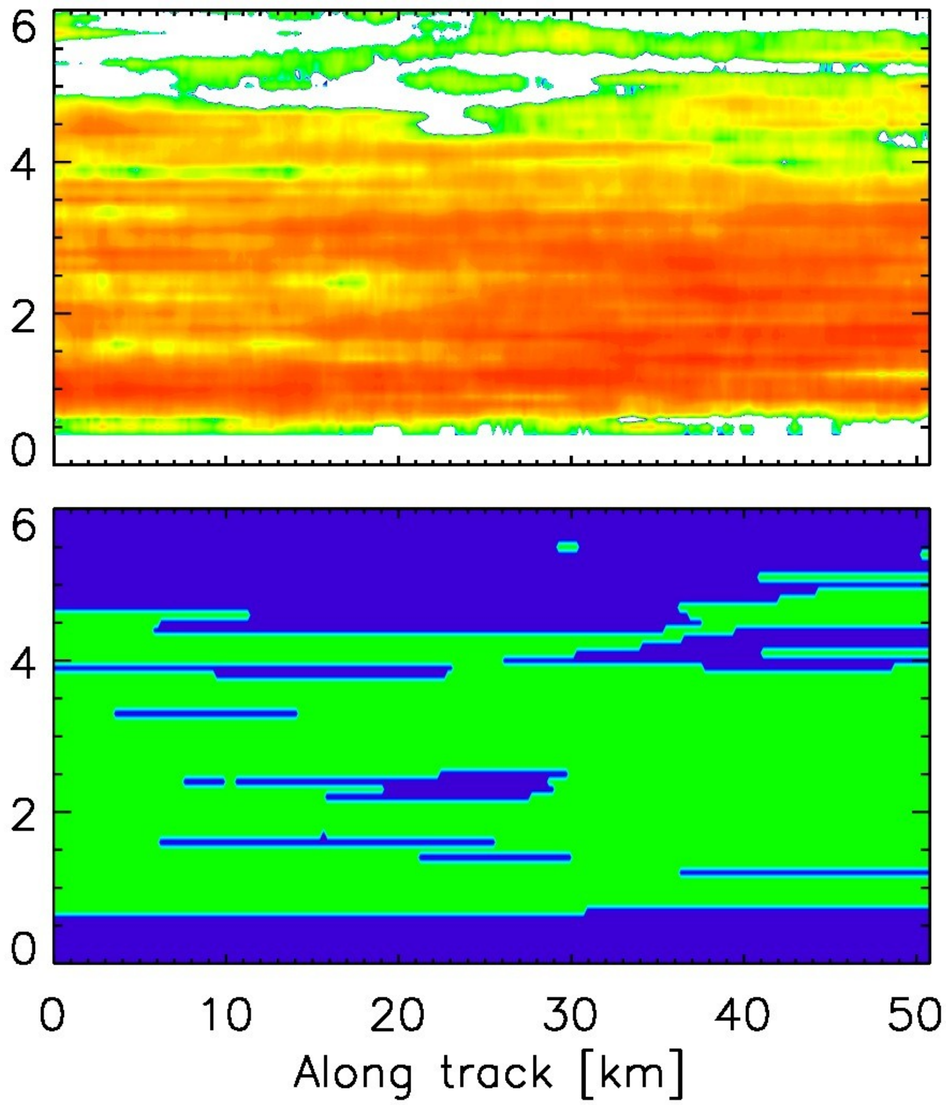

Figure 8: Linear fitting method, [Sv =9 pixels, Sh=20km] : Extinction coefficient (Upper panel), SNR>3 mask (Lower Panel). (NB: In the upper panel, extinction a retrieve also for SNR<3)

6.2. Comparisons in Backscatter product between method 1 and 2

Figure 9 shows the backscatter profile as retrieve by methods 1 & 2, as well as the compared to the 'true' input backscatter. The two method show a good correlation.

The ratio method (Method 1) is preferred due to the lack of extinction errors which can enter into the equation. Method 2 can be influenced by errors (or the resolution) of the retrieved extinction profile. By choosing Method 1, any additional contaminations are avoided in the retrieval of the backscatter coefficient. Sensitivity tests have been performed for both methods. Additional tests in the future are required to validate this conclusion.

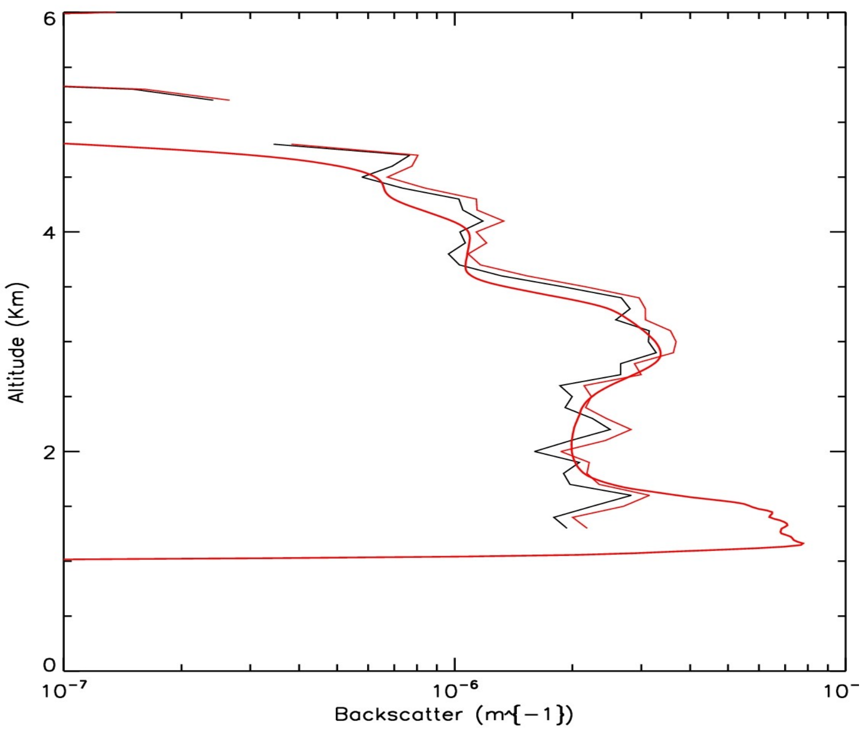

Figure 9: Backscatter coefficient : Thick line, in red : Input backscatter, black thin line : backscatter as retrieved by Method 1, red thin line : backscatter as retrieved by Method 2.

6.3. Comparison with DLR Falcon datasets

To validate the accuracy of the developed algorithm a comparison to the retrieved extinction from different algorithms is presented in this Section. Four different datasets are analysed (see Figure 10):

1. The extinction coefficient as retrieved by the DLR team using the Falcon HSRL data directly.
2. The extinction coefficient as retrieved with the ECSIM algorithm adopting the Falcon configuration.
3. The “true” extinction, after conversion of the Falcon DLR datasets to ECSIM format.
4. The extinction coefficient adopting the EarthCARE configuration retrieved using the current ECSIM algorithm.

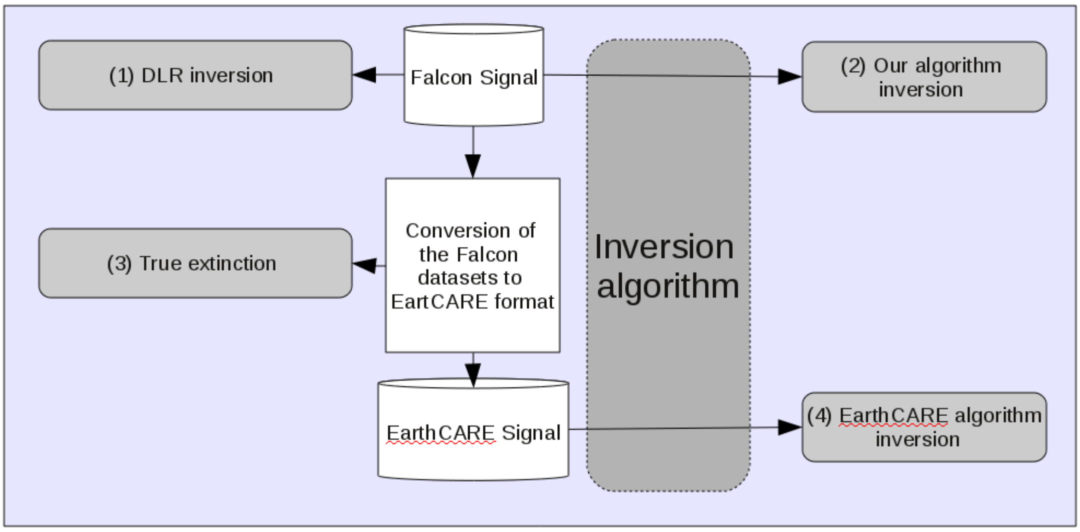

Figure 10: flowchart of the comparison study.

From the optimal Sh and Sv parameters presented in Table 4, adopting the linear fitting method, the extinction coefficient can be retrieved. The retrieval has been performed using the Falcon datasets (see Figure 11) and a similar retrieval is made for EarthCARE type signal (Figure 12).

Figures 6 and 13 show respectively the true extinction, and the HSRL extinction retrieved by the DLR team. Figures 14 gives the mean profile corresponding to each of the 2D extinction distributions.

First, it must be noticed the occurrence of surface contamination in the direct HSRL retrieval. The differences close to the surface are a direct result of the width of the vertical sliding window, affecting the inversion of the signals, and the width of the smoothing window in the horizontal.

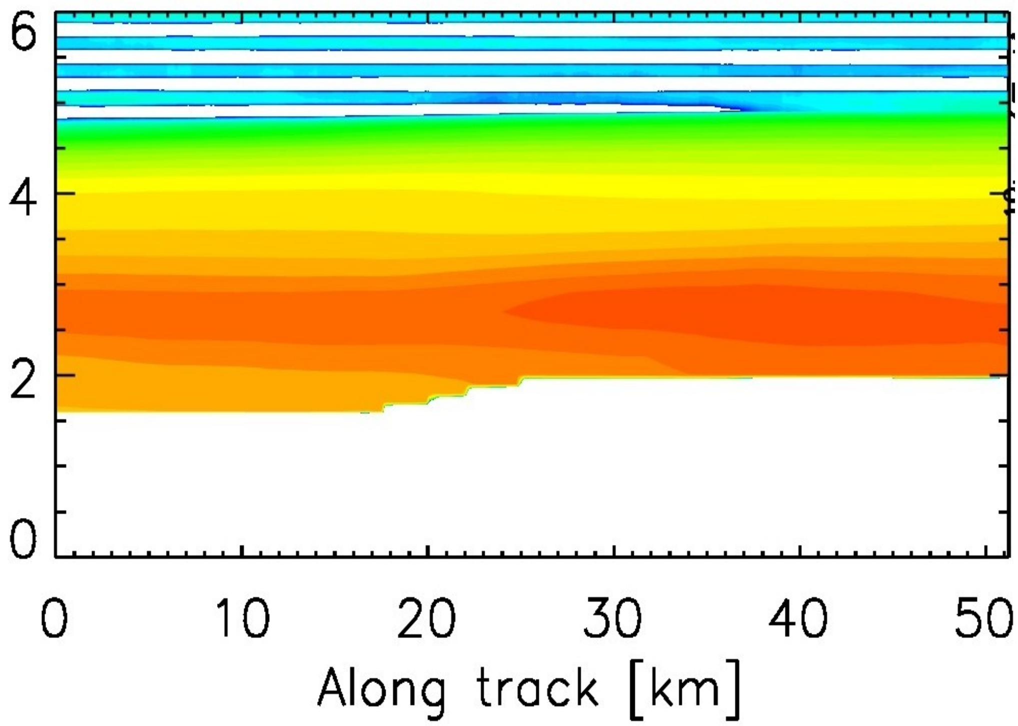

Figure 11: Extinction as derived from Falcon dataset. We use in this case the Linear fitting Method. The chosen parameters for the vertical and horizontal dimension of the sliding window are ones given inside the table 7, i.e. [Sv=9 pix, Sh=20].

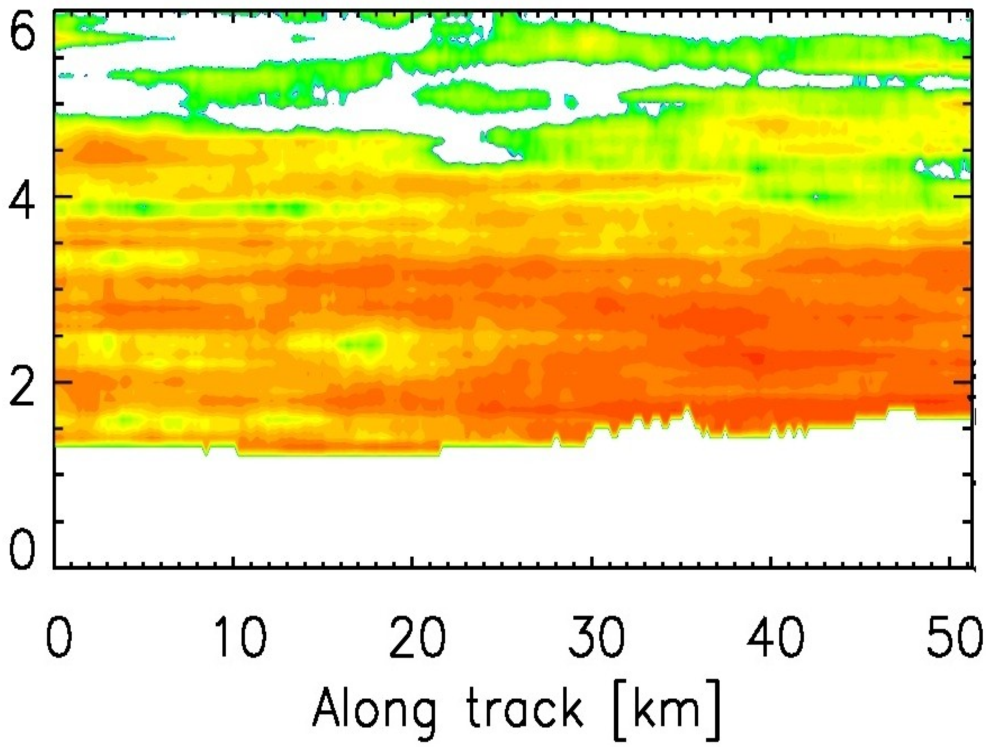

Figure 12: Extinction as derived from the EarthCARE signal. We use in this case the Linear fitting Method. The chosen parameters for the vertical and horizontal dimension of the sliding window are ones given inside the table 7, i.e. [Sv=9 pix, Sh=20].

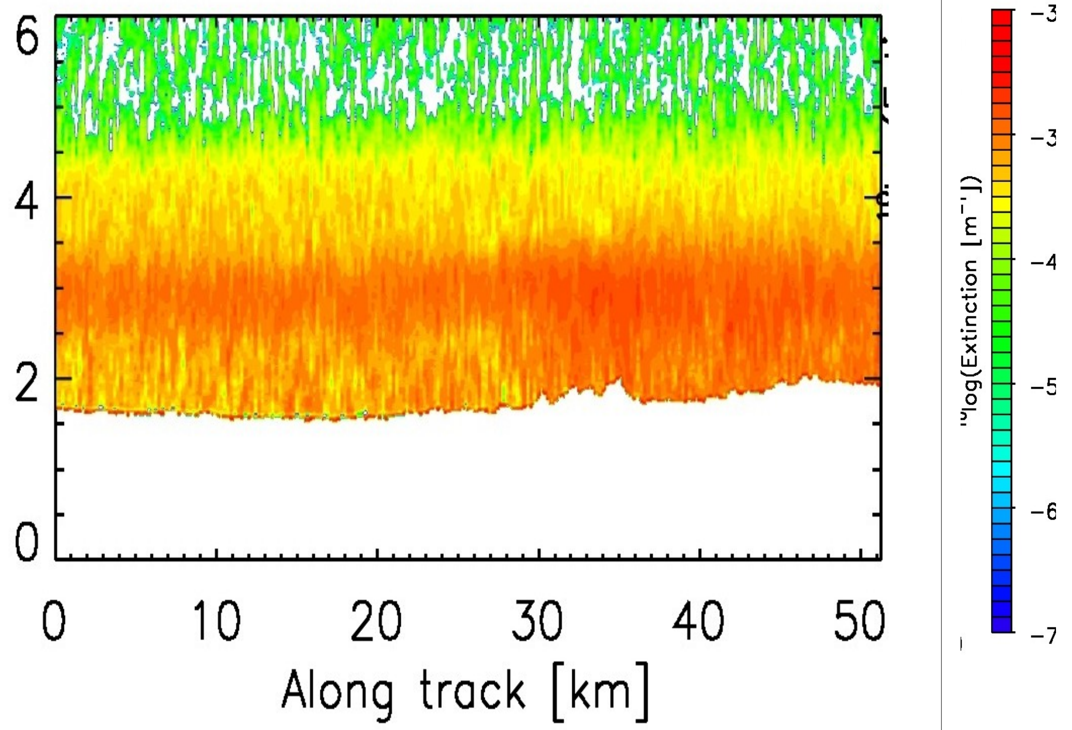

Figure 13: HSRL Extinction as derived by the DLR team.

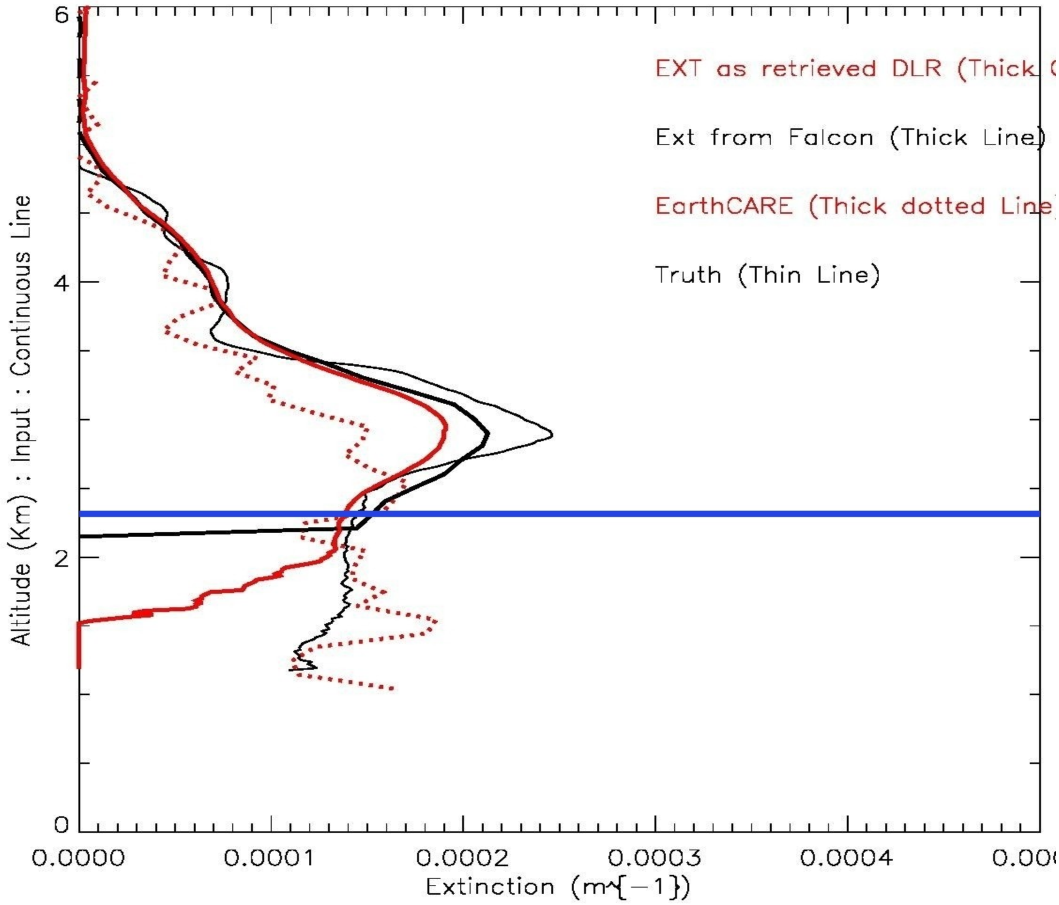

Figure 14: Extinction from all sources, averaged over all the scene. This plot give (1) The Extinction as retrieved by the DLR (Red thick continuous line), (2) the Extinction from the Falcon dataset (black thick continuous line), (3) the EarthCARE extinction (Red thick dotted line), and (3) the True extinction (black thin line). Plots under blue line couldn't be taken in account due to ground contribution.

In all the cases, the general geometry of the aerosol layer is visible.

Some horizontal “stripes”, due to the horizontal averaging, appear on the Falcon and the EarthCARE plot. The peak in Extinction is visible in the two retrievals based on the Falcon configuration, it is however hard to distinguish in the case of the EarthCARE retrieval.

7. Validation status

The algorithm presently exists in prototype form as a stand-alone IDL routine. It has not been integrated into the ECSIM environment yet. The validity of the algorithm applied to cloud-free actual HSRL lidar data and simulated ATLID signals corresponding to the same cloud-free scenes has been established. The proper functioning of the algorithm when clouds are present has not yet been fully established.

8. Future validation needs

The sensitivity study needs to be extended in the future to more and larger scenes (over 50km) containing similar realistic extinction profiles, multiple aerosol types and a combination of aerosols and clouds. These types of aerosols scenes have recently been produced. These last will help us to constraint the size of the sliding window, mainly in the case where the aerosol scene is cloud contaminated. In this last type of scenes, the algorithm will need to adapt the size of the sliding windows, to retrieve a reliable SNR.