ADM-Aeolus L2A Algorithm Theoretical Baseline Document

Particle spin-off products

Written by:

Pierre H. Flamant, Vincent Lever, Pauline Martinet, Thomas Flament, Juan Cuesta, Alain Dabas, Mathieu Olivier, Dorit Huber, Dimitri Trapon, Adrien Lacour

AE-TN-IPSL-GS-001

Version 5.7

Contract 18366/04/NL/MM CR 1

08 Jul. 2020

0.1 Document change log

Issue	Date	Modified pages	Observations	Name
V 4.0	11.01.08	--	Documentation & Format Scientific parts	Huber Flamant
V 4.1	14.04.08	--	New filling cases added (Sensitivity studies TBD). Correction of switched instrumental parameters. Correction of K_p and K_m into K_{mie} and K_{ray}	Olivier
V 4.2	17.07.08		New sensitivity study Corrections from PF by MO Add of parts 15 and 16 : Matchup with AUX_MET and AUX_CLM	Olivier
V 5.0	07.07.09		New version	Olivier Flamant
v.5.1	19 12 11		New version accounting for typo corrections, algorithm modifications, suggestions, also including new continuous mode operation. SCA error propagation.	Lever Flamant
v. 5.2	31.05.12	All but sections 7.4, 7.5 and 10.	Correction and reorganisation of previous version. Update of architecture section, ICA section, NWP match up algorithm.	Lever Flamant
v. 5.3	15.03.13		Sections 5 (architecture) and 9 (Scene Classification) updated. Section 6 (inversion algorithms) completed. Section 8 (Feature Finder) created.	Lever
v. 5.4	13.06.13		Various corrections.	Martinet
v. 5.5	16.01.2017	Feature finder section	Changed E_0 unit from J to mJ in accordance with actual code. Update of Feature Finder to new algorithm.	Flament
v. 5.6	08.07.2020		New section 6.4. Improvement of SCA products: estimation of calibration coefficients K_{ray} and K_{mie} and radiometric correction refinement on an empirical basis	Trapon
	01.02.2021		Section. 6 updated with the definition of the	Lacour

			heterogeneity_index	

0.2 Table of contents

1 INTRODUCTION

The Atmospheric Dynamics Mission (ADM-Aeolus) is the 2nd European Space Agency's Earth Explorer core mission. At the beginning of 2017, the satellite is undergoing testing, for a launch onega rocket around the end of the year ([RD 1]). Aeolus – inspired by the ancient Greek eponymous wind master – aims at measuring wind velocity profiles in the depth of Earth atmosphere as a first objective. In addition, it has the capability to measure clouds and aerosols optical properties as auxiliary mission objective. This second objective is addressed in the present Level 2 Algorithm Theoretical Basis Document (L2A ATBD).

The mission payload is made of ALADIN, a High Spectral Resolution (HSR) elastic backscatter Lidar operated at 355-nm laser wavelength with direct detection scheme. Wind profiles are derived from the laser light scattered by air molecules and particles moving with the wind. The characteristics i.e. abundance and size distribution of the two types of microscopic scatterers play a key role in the choice of operating lidar wavelength. The air molecules density is well distributed around the Earth and it varies slowly from place to place according to pressure. Its vertical distribution follow an (nearly) exponentially decrease with height above the surface. On the contrary, aerosol and cloud particle densities vary greatly in vertical, location and time. Aerosol particles from the main sources i.e. desert clay, marine environment, large cities pollution, volcanic and forest fire ashes, are lifted up by turbulence and convection and widely dispersed by atmospheric circulation. Sedimentation process occurs during the transport processes so the size distribution of flying aerosols particles varies from tenths to tens of micrometer at large distance from the sources. Depending on prevailing temperature condition, cloud particles are made of liquid droplets or ice crystals or a mix of the two. The size of cloud particles varies from tenths of microns to millimetres.

The lidar signal strength depends on molecules and particle number densities and sizes. The scattered spectra are different for molecules and particles and the two contributions can be put apart using a High Spectral Resolution multiple interferometers receiver. It results into two channels: i) a Rayleigh channel for molecular scattering and ii) a Mie channel for particle scattering. The two channels are used to deliver two independent wind velocity measurements, and to make unambiguous retrievals of particles backscatter and extinction coefficients [RD 8].

The present L2A ATBD presents the High Spectral Resolution Lidar technique and the relevant processing algorithms to derive clouds and aerosol layers optical characteristics. The L2A processor has already been presented briefly in [RD 3].

The L2A ATBD is organised as follow:

- Sections 2 and 3 present the applicable and reference documents, and the abbreviated terms and symbols, respectively.
- Section 4 presents the Aeolus mission and the lidar ALADIN. The Rayleigh and Mie channels are described and the scattered spectra from air molecules and particles are presented.
- Section 5 contains the overarching description of the L2A processor.

- Section 6 provides the core algorithms of the L2A processor, i.e. the Rayleigh channel standard correct algorithm (SCA) and the Mie channel algorithm (MCA), assuming a complete and homogeneous filling of the range bins. The SCA requires a crosstalk correction beforehand to enable the use of a local normalized integrated two-way transmission (NITWT). The MCA includes a pseudo-correction of the crosstalk and processing of Mie channel data only using an assumed particle extinction/backscatter ratio as input.
- Section 7 is dedicated to an iterative correct algorithm (ICA) which aims at correcting the hypothesis of an homogeneous filling of the bin used in the SCA. It processes crosstalk-corrected signals assuming various cases of partial filling of the range bins. A method is proposed to choose the closest case to reality. This algorithm can be used in the case where a clear layer exists under the aerosol layer.
- Section 8 details the rationale of the feature finder. For each height level, it locates the bins that contain a feature and packs them into one group. The SCA is applied on these groups to retrieve the Aeolus' aerosol and clouds spin-off products.
- Section 9 describes the principle of scene classification. Group-wise products are analysed together with NWP's to discriminate features between clouds and aerosols.

2 APPLICABLE & REFERENCE DOCUMENTS

2.1 Applicable documents

[AD 1] ESA/18366/04/NL/MM CR 1 ESA contract 'Aeolus Level 1B/2A Processor Refinement & Pre-Launch Validation'.

[AD 2] AE-RP-DLR-L1B-001 v4.0: ADM-Aeolus ATBD Level1B Products.

[AD 3] AE-IF-DLR-L2A-004 v2.1: ADM-Aeolus L2A Product - Input/Output Data Definition.

[AD 4] AE-IF-ECMWF-L2BP-001 v1.2: ADM-Aeolus Level2B/2C Processor Input/Output Data Definitions Interface Control Document.

[AD 5] AE-TN-ECMWF-L2BP-0023 version 2.1 ADM-Aeolus Level-2B Algorithm Theoretical Baseline Document.

2.2 Reference documents

[RD 1] http://www.esa.int/esaLP/LPadmaeolus.html

[RD 2] http://www.nasa.gov/mission_pages/calipso/main/index.html

[RD 3] Flamant P. H., J. Cuesta, M.-L. Denneulin, A. Dabas, D. Huber, *ADM-Aeolus retrieval algorithms for aerosol and cloud products*, Tellus, 60A, 273-286, 2008

[RD 4] Fiocco G., F. Congeduti, G. Benedetti-Michelangeli, *Measurement of Aerosol Motion and Wind Velocity in the Lower Troposphere by Doppler Optical Lidar*, J. Atmos. Sci., Vol 29, p. 906 – 910, 1972

[RD 5] Shipley S. T., D. H. Tracy, E. W. Eloranta, J. T. Trauger, J. T. Sroga, F. L. Roesler, and J. A. Weinman, *High spectral resolution lidar to measure optical scattering properties of atmospheric aerosols. 1: Theory and instrumentation*, Applied Optics, Vol. 22, 3716-3724, 1983

[RD 6] Sroga J. T., E. W. Eloranta, S. T. Shipley, F. L. Roesler, and P. J. Tryon, *High spectral resolution lidar to measure optical scattering properties of atmospheric aerosols. 2: Calibration and data analysis*, Applied Optics, Vol. 22, 3725-3732, 1983

[RD 7] ASTRIUM document, *ALADIN, FM instrument, Architecture, verification and performance*, ALD.NT.00039.T.ASTR.

[RD 8] ESA, 2008: *ADM-Aeolus Science Report*, SP-1311, 212p

[RD 9] Chanin M.L., A. Garnier, A. Hauchecorne, J. Porteneuve. *A Doppler lidar for measuring winds in the middle atmosphere*, Geophys. Res. Letters, 16, 1273-1276, 1989

[RD 10] Garnier A., M.L. Chanin, *Description of a Doppler Rayleigh Lidar for measuring winds in the middle atmosphere*, Appl. Phys., B55, 35-40, 1992.

[RD 11] Reitebuch O., U. Paffrath, I. Nikolau, AE.TN.DLR.A2D.TN51 *Technical Note TN 5.1* ADM-Aeolus Ground Campaigns Results.

[RD 12] Reitebuch O., U. Marksteiner, C. Lemmerz, AE.TN.DLR.A2D.TN52.240212 *Technical Note TN 5.2* ADM-Aeolus Airborne Campaigns Results.

[RD 13] Loth C., P. H. Flamant, A. Dabas, M.-L Denneulin, A. Dolfi-Bouteyre, A. Garnier, and D. Rees, *ILIAD, Impact of line shape on wind measurement and correction methods*, ESTEC contract 18334/04/NL/MM, Final Report, 2005.

[RD 14] G. Tenti, C. D. Boley, R. C. Desai, *On the Kinetic of Rayleigh-Brillouin Scattering from Molecular Gases*, Can. Journ. Phys. 52, pp. 285 - 290, 1974

[RD 15] Witschas B., M. O. Vieitez, E-J. van Duijn, O. Reitebuch, W. van de Water, W. Ubachs. *Spontaneous Rayleigh-Brillouin scattering of ultraviolet light in nitrogen, dry air and moist air*. Appl. Opt., 48, 4217-4227, 2010

[RD 16] Omar, A.H., et al., *The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm*, Journal of Atmospheric and Oceanic Technology 26:10, 1994-2014, (2009)

[RD 17] *LIVAS Technical Note 1*, ESTEC Contract No. 4000104106/11/NL/FF/fk

[RD 18] AE.TN.DLR.ACS-L1B.15.02.2006, *Iss. 1.3, Additional Computational Steps*

[RD 19] E Berthier S., P. Chazette, J. Pelon, B. Baum, *Comparison of cloud statistics from space borne lidar systems*. Atmos. Chem. Phys. Disc., 8, 6965-6977, 2008

[RD 20] Eguchi, N. and T. Yokota. *Investigation of clear-sky occurrence rate estimated from CALIOP and MODIS observations*, Geophysical Research Letters, 35, L23816, doi:10.1029/2008GL035897, 2008

[RD 21] AE-RP-DLR-L1B-001, *ATBD ADM-Aeolus Level 1B Products*, iss 3.0

[RD 22] SAGE II, http://badc.nerc.ac.uk/data/sage2

[RD 23] Nicolas F., L. R. Bissonnette, P. H. Flamant. *Lidar effective multiple-scattering coefficients in cirrus clouds*, Appl. Opt., 36, 3458-3468, 1997

[RD 24] ADM-IC-52-1666 issue 4.04 *Level 1B & E2S Input / Output Data Definition Document* (28/10/2011)

[RD 25] AE-TN-MFG-L2P-CAL-004_v20, *Generation of AUX_CAL,* Detailed Processing Model, Input/Output Data Definition

[RD 26] A. Dabas, M.-L. Denneulin, P. Flamant, C. Loth, A. Garnier and A. Dolfi-Bouteyre. *Correcting winds measured with a Rayleigh Doppler lidar from pressure and temperature effects,* Tellus 60A, 206-215, 2008

[RD 27] Collis R.T.H., P.B. Russell, *Lidar measurements of particles and gases by elastic backscattering and differential absorption*, in Laser Monitoring of the Atmosphere, Ed. E.D. Hinkley, Springer-Verlag, Berlin, pp. 380, 1976.

[RD 28] AE-IF-DLR-L2A-004, Aeolus Level 2A Processor Input/Output Definition

[RD 29] AE-IF-ECMWF-L2BP-001, ADM-Aeolus Level-2B / 2C Processor Input / Output Data Definitions Interface Control Document

[RD 30] D.Müller, A. Ansmann, I.Mattis, M.Tesche, U.Wandinger, D. Althausen and G.Pisani (2007).,*Aerosol-type-dependent lidar ratios observed with Raman lidar*, J.Geophys.Res.,112,D16202, doi:10.1029/2006JD008292

[RD 31] F.De Tomasi, A.M. Tafuro and M.R Perrone (2006), *Heigh and seasonal dependence of aerosol optical properties over southeast Italy*, J. Geophys.Res., 111, D1023, doi: 10.1029/2005JD006779

[RD 32] F.De Tomasi, A. Blanco and M.R Perrone (2003), *Raman lidar monitoring of extinction and backscattering of African dust layers and dust characterization*, Appl. Opt., 42, 1699-1709.

[RD 33] V. Amiridis, D.S Balis, S. Kazidis, A. Bais and E. Giannakaki (2005), Four-year aerosol observation with a Raman lidar at Thessaloniki, Greece, in the framework of the

European Aerosol Research Lidar Network (EARLINET), J. Geophys. Res., 110, D21203, doi:10.1029/2005JD006190.

[RD 34] P. Seifert, A. Ansmann, D. Müller, U. Wandinger, D. Althausen, A.J. Heymsfield, S.T. Massie and C. Schmitt (2007), *Cirrus optical properties observed with lidar radiosonde, and satellite over the tropical Indian Ocean during the aerosol-polluted northeast and clean maritime southwest monsoon*, J. Geophys. Res., 112, D17205, doi: 10.1029/2006JD008352.

[RD 35] D.N. Whiteman and B.Demoz (2004), *Subtropical cirrus cloud extinction to backscatter ratios measured by Raman Lidar during CAMEX-3*, Geophys. Res. Lett., 31, L12105, doi: 10.1029/2004GL020003.

[RD 36] E.J. O'Connor, A.J. Illingworth and R.J. Hogan (2004), *A technique for autocalibration of cloud lidar*, J. Atmos. Oceanic. Technol., 21, 777-786.

[RD 37] Groβ et al (2011), Characterization of Saharan dust, marine aerosols and mixtures of biomass-burning aerosols and dust by means of multi-wavelength depolarization and Raman lidar measurements during SAMUM 2, Tellus, 63B, 706-724.

[RD 38] M.Sicard, JL Guerrero-Rascado, F. Navas-Guzman et al (2012), *Monitoring of the Eyjafjallajökull volcanic aerosol plume over the Iberian Peninsula by means of four EARMINET lidar stations 2012*, Atmospheric chemistry and physics, 12.6, 3115-3130.

[RD 39] S. Groβ, V. Freudenthaler, M. Wiegner, J. Gasteiger, A. Geiβ, F. Schnell (2010), Dual-wavelength linear depolarization ratio of volcanic aerosols: Lidar measurements of the *Eyjafjallajökull plume over Maisach,* Atmospheric Environment, 48, 85-96.

[RD 40] H. Baars, A. Ansmann, D. Althausen, R. Engelmann, B. Heese, D. Müller, P. Artaxo, M. Paixao, T. Pauliquevis and R. Souza (2012), Aerosol profiling with lidar in the Amazon Basin during the wet and dry season, J. Geophys. Res., 117, D21201.

[RD 41] M. Vaughan, Winker,D.M. , Powell, K.A. CALIOP Algorithm Theoretical Baseline Document, Part 2: Feature Detection and Layer Properties Algorithms, PC-SCI-202 Part 2

[RD 42] Level 1b Processor Dtailed Processing Model, ADM-MA-52-1800

3 ABBREVIATIONS, SYMBOLS AND DEFINITIONS

3.1 Abbreviations

ABL	Atmospheric Boundary Layer
ACCD	Accumulation Charge Coupled Device
ADM-Aeolus	Atmospheric Dynamics Mission: Aeolus keeper of the winds
AERONET	AErosol RObotic NETwork
AHT	Accurate Housekeeping Telemetry

ALADIN	Atmospheric Laser Doppler Instrument
ATBD	Algorithm Theoretical Basis Document
BER	Backscatter-to-Extinction Ratio
BRC	Basic Repeat Cycle
CALIPSO	Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observation
CC	Credibility Criterion/Criteria
CCD	Charge Coupled Device
CIWC	Cloud Ice Water Content
CLWC	Cloud Liquid Water Content
DFP	Dual Fabry-Pérot
DFPI	Dual Fabry-Pérot Interferometer
DLR	Deutsches Luft- und Raumfahrtzentrum
E2S	End-to-end Simulator
EARLINET	European Aerosol Research Lidar NETwork
EM	Electromagnetic
ESA	European Space Agency
ESL	Elastic Scattering Lidar
ESTEC	European Space Technology and research Centre
FM	Flight Model
FOV	Field of View
FP	Fabry-Pérot
FSR	Free Spectral Range
FWHM	Full Width Half Maximum
GCM	General Circulation Models
HSR	High Spectral Resolution
HSRL	High Spectral Resolution Lidar
ICA	Iterative Correct Algorithm
IIR	Infrared Imager Radiometer
ILIAD	Impact of LIne shape on ADM-Aeolus Doppler estimates
IPSL	Institute Pierre Simon Laplace
IR	Infra Red
IRC	Instrument Response Calibration
ISR	Instrument Spectral Response
L1B	Level 1 B
L2A	Level 2 A
L2B	Level 2 B
L2C	Level 2 C
LiTE	Lidar in space Technology Experiment
LOD	Local Optical Depth per range bin
LOS	Line-of-sight

MCA	Mie Channel Algorithm
NA	Non Applicable
NITWT	Normalized Integrated Two-Way Transmission
OD	Optical Depth
ONERA	Office National d'Etudes et de Recherches Aérospatiales
RBS	Range Bin Setting
RH	Relative Humidity
RMA	Rosin Medium Aerosol
RUS	Rayleigh Useful Signal
SAGE	Stratospheric Aerosol and Gas Experiment II
SCA	Standard Correct Algorithm
SEBL	Standard Elastic Backscatter Lidar
SLOD	Slant Local Optical Depth per range bin
SNR	Signal to Noise Ratio
SR	Scattering Ratio
SSA	Single Scattering Assumption
TBD	To be done
TC	Thermal Control
USR	Useful Spectral Range
UV	Ultra Violet
WFC	Wide Field of view Camera

3.2 Symbols

Symbol	Definition	Unit
a_e	Effective albedo.	NA or %
B_i	Variable to lighten equations: $B_i = \Delta R_i \beta_{p,i}$	sr^{-1}
C_1, C_2, C_3, C_4	Calibration coefficients. C_1 corresponds to Rayleigh signal in the Fabry-Pérot, C_2 to Mie signal in Fabry-Pérot, C_3 to Mie signal in Fizeau and C_4 to Rayleigh signal in Fizeau. Here C_2,C_4 count for crosstalk effect between Rayleigh and Mie channels. They can change with Doppler shift.	NA
$C_{n,ray,i},\ C_{n,mie,i}$	n^{th} calibration coefficient resp. in the i^{th} bin of Rayleigh scale and the i^{th} bin of Mie scale.	NA
D_i	Variable to lighten equations: $D_i = (B_i + B_{i+1}) \times 2\alpha_{p,i+\frac{1}{2}} \Delta R_{i+\frac{1}{2}}$	sr^{-1}
D_{re}, D_c, D_{ta}	Diameter of receiver, coherence area on receiver, and laser footprint.	m
E_0	Laser pulse energy.	mJ
e_{xxx}	Relative error on the quantity xxx.	NA
f	Frequency	Hz
h_0	Satellite height	m
I	Light Intensity	$W.sr^{-1}$
(i, j)	Indices; often of height bin	NA
K_{ray}, K_{mie}	Overall calibration coefficients for the useful signal in the Rayleigh and Mie channels.	$m^2.s^{-1}$
k_{ray}, k_{mie}	Radiometric gain factor for the Mie and Rayleigh signals	pixel count/electron
k_m, k_p	Backscatter-to-extinction ratio for molecules, particles	sr^{-1}
l_{mask}	Length of the mask of the Gaussian blur, in the number horizontal bins.	NA
l_{min}	In the Feature Finder, minimal length of a group of bins.	NA
$L_{p,i}, \hat{L}_{p,i}$	Slant local optical depth for particles in range bin *i* and its	NA

Symbol	Definition	Unit
	estimate.	
M	Molar mass for air molecules	$g.mol^{-1}$
N	Number of measurements per observation	NA
N_c	Number of speckle cells	NA
N_p	Number of pulses accumulated to generate a profile	NA
n	Refractive index of aerosols	NA
$NITWT$	Normalised integrated two-way transmission	NA
P	Number of shots	NA
$p, p(z)$	Pressure	hPa
R_g	Ideal gas constant.	$J.K^{-1}.mol^{-1}$
R_i	Range from the satellite to the top gate of the bin *i*.	m
$R_{mean,i}$	Range to the middle of the bin i.	m
R_0	Satellite-to-the-surface range	m
$S_{ray,i}, S_{mie,i}$	Accumulated Lidar signal in a vertical range bin for Rayleigh and Mie channels respectively.	W
$T, T(z)$	Temperature.	K
T_A,T_B	Transmission characteristics of Fabry-Pérot A and B respectively	
$T_{m,sat,i}, T_{p,sat,i}$	Transmissions from satellite to the bottom of the bin i, for molecules and particles.	NA
xxx_{gr}	Average value of the variable xxx over the group gr.	
$xxx_{i+\frac{1}{2}}$	Averaged value of xxx between height bins i and $i+1$ and centred on z_{i+1}.	
$xxx_{p,i,k}$	Particle value of the variable xxx in the bin of the i^{th} height level and the k^{th} measurement.	
X, Y	Molecular and particle backscattered signals at telescope entry.	$m^{-3}.sr^{-1}$

Symbol	Definition	Unit
$X_i, X_{ray,i}, X_{mie,i}$	Accumulated molecular backscattered signal at telescope entry resp. in the bin i, in the ith bin of Rayleigh scale, in the ith bin of Mie scale.	$m^{-2}.sr^{-1}$
$Y_i, Y_{ray,i}, Y_{mie,i}$	Accumulated particle backscattered signal at telescope entry resp. in the bin i, in the ith bin of Rayleigh scale, in the ith bin of Mie scale.	$m^{-2}.sr^{-1}$
z, z_i	Altitudes.	m
δx	Error on the quantity x.	$[X]$
Δf	Frequency shift	MHz
$\Delta z_{\mathrm{p,i}}$	Geometrical thickness of a particle layer	m
$\alpha, \alpha_m, \alpha_p, \alpha_{m,i}, \alpha_{p,i}$	Extinction coefficient (depending on z), for molecules or particles, and for molecules or particles in the ith vertical range bin.	m^{-1}
$\beta, \beta_m, \beta_p, \beta_{m,i}, \beta_{p,i}$	Backscatter coefficient (depending on z), for molecules or particles, and for molecules or particles in ith vertical range bin.	$m^{-1}.sr^{-1}$
ϵ	Relative error of the NITWT.	NA
η	Attenuation factor of α in case of multiple scattering effect.	NA
λ	Wavelength	m
θ	Looking angle off nadir	°
ρ, ρ_{L1B}	Scattering ratio, L1B estimate of the scattering ratio.	
$\sigma_{ref}(\mu)$	Reference distribution of standard deviation on the smoothed β_p of a group of bins in function of its mean μ.	
σ_{xxx}	Standard deviation of the variable xxx.	$[X]$

3.3 Definitions and cross-references to other Aeolus product levels

3.3.1 Links between "accumulated Lidar signals" and Level 1B data

L1B input signals are expressed in engineering units. After proper calibration they are expressed as energy (J) or number of photons or photoelectrons. Level 1B data are expressed in number of photoelectrons in range bin after subtraction of mean background (solar for Rayleigh channel and solar + Rayleigh for Mie channel). According to the L1B ATBD ([AD 2]), Table 8.1, the L1A Mie atmospheric information are labelled $S_{mie,atm,i}$, whereas $S_{mie,i}$, is used in this ATBD, and the same for Rayleigh information i.e. $S_{ray,atm,i}$ and $S_{ray,i}$.

$$S_{ray,atm,i} \equiv S_{ray,i} = \int_{R_{ray,i-1}}^{R_{ray,i}} s_{ray}(r)dr \qquad \text{Eq. 3.1}$$

$$S_{mie,atm,i} \equiv S_{mie,i} = \int_{R_{mie,i-1}}^{R_{mie,i}} s_{mie}(r)dr \qquad \text{Eq. 3.2}$$

In practice, after correction for partial sampling by the dual Fabry-Pérot, the L1B processor provides

$$S_{ray,i} = Rayleigh\_Useful\_Signal\_A + Rayleigh\_Useful\_Signal\_B \qquad \text{Eq. 3.3}$$

$$S_{mie,i} = \left(\sum_{1}^{16} Mie\_Useful\_Signal\right) + Rayleigh\ Background \qquad \text{Eq. 3.4}$$

These signals include crosstalk contributions. As pointed out [RD 3] the Rayleigh channel only provides partial measurements of the total backscattered spectrum due to the dual Fabry-Pérot filter technique, so a correction is needed. Furthermore, the transmissions in the two FP filters are not equal due to the Doppler shift of the return signals, which is dependent on the local wind speed.

3.3.2 Cross references

Symbol	Algorithm Reference	Parameters	Document reference
C_1, C_4		Table 40	[AD 2]
C_2, C_3		Table 42	[AD 2]
K_{ray}		Table 37, K_ray	[AD 2]
K_{mie}		Table 41, K_mie	[AD 2]

$S_{m,i}\big\vert_{ct}$	section XX	$S_{m,i}\big\vert_{obs}$	
$S_{p,i}\big\vert_{ct}$		Table 6-50	[RD 24]

Table 3.1: Cross talk cross references

This section provides a cross-reference to the L1B input data and the L2A output data (TBD) and details some implementation aspects.

Symbol	**Algorithm Reference**	**Parameters**	**Document reference**
$X_{ray,i}\big\vert_{cal}$	section XX		
$X_{ray,i}\big\vert_{obs}$		Table 6-51: Useful_Signal_Channel_A + Useful_Signal_Channel_B	[RD 24]
$K_{R,i}$	section 6.1		
$T_{m,sat,i}$	section XX		
$\beta_{m,i}$	Equation (5.4a)		
$\alpha_{m,i}$	Equation (5.4b)		
$SatRange_i$		Table 5-9, linear interpolation to measurement scale	[RD 24]
$\frac{z_i - z_{i-1}}{\cos\theta_i}$	Calculated from $SatRange_i$		
tol		L2a processing parameters, tolerance limit to stop iterative inversion.	[AD 2], Table 55, named : Inversion_Tolerance_Value
$p(z)$		Table 82 values linearly interpolated	[AD 4]
$T(z)$		Table 82 values linearly interpolated	[AD 4]
E_0		Table 6-7 Avg_UV_Energy * 1e-3	[AD 2]

P		Table 6-3	[AD 2]

4 ADM-AEOLUS WIND LIDAR

The ADM-Aeolus mission is designed to fly ALADIN a High Spectral Resolution (HSR) Lidar operating at Ultra Violet laser wavelength 355 nm. The lidar combines a multiple interferometer receiver with a direct detection scheme. ALADIN has no depolarization capability according to a laser circular polarisation emission compliant with the multiple interferometer receiver design. ALADIN is the only instrument to be flown on ADM-Aeolus satellite.

The wind measurement principle is based on a pulsed laser beam illuminating moving atmospheric targets (molecules and particles) and subsequent determination of the Doppler frequency shift on the scattered light. The choice for a 355 nm laser emission is based on the lidar capability to measure wind profiles everywhere at all latitudes in Earth atmosphere using the light scattered by molecules. ALADIN will probe the atmosphere using one single line-of-sight (LOS) sets cross track at 37,5° from nadir. A 37,5° angle with respect to nadir provides with adequate horizontal wind projection on the Lidar Line-of-sight. A cross-track perspective with respect to the satellite direction of motion cancels the Doppler frequency shift effect associated to the platform velocity (about 7.7 km/sec while atmospheric winds can reach 100 m/sec). The multiple interferometers receiver splits the atmospheric signals into a Rayleigh and a Mie channels according to the difference in scattered spectra by molecules and particles. The Rayleigh and Mie channels provide two independent atmospheric wind measurements. In addition to wind information, ADM-Aeolus can provide information on clouds and aerosol layers. The Rayleigh channel can be used to derive the local optical depth (particle extinction coefficient) while the combination of the two channels enables to derive the particle backscatter-to-extinction ratio. It is important to recall that the lack of a depolarization capability in the ADM Aeolus mission will lead to a significant underestimation of the backscatter coefficient and an overestimation of the respective extinction-to-backscatter ratio in case of highly depolarizing cirrus clouds or desert dusts/ash containing aerosol layers. This is caused by the measurement of only the co-polar component the backscatter coefficient whereas the cross-polar component can be significantly higher. This is discussed in section 4.2.1.4.

The lidar receiver combines a dual Fabry-Pérot interferometer ([RD 9], [RD 10]) and a Fizeau interferometer ([RD 7]). The dual Fabry-Pérot analyzes the laser light scattered by molecules [RD 9] [RD 10], whereas the Fizeau spectrometer analyzes the light scattered by particles. The dual Fabry-Pérot samples the molecular spectrum into two band pass filters. Such spectral sampling needs to be taken into account to derive the backscatter coefficient. The FP spectral sampling varies slightly with wind velocity. Ideally the two channels are optically decoupled but in practice there exist an optical cross talk between the Rayleigh and Mie channels. This cross talk requires a careful calibration before launch. The ALADIN pulsed UV laser will be fired continuously at 50.5 Hz to result in a continuous mode operation. The atmosphere will be sampled in 24 range bins from about a 30-km altitude down to the surface. The vertical length of the range bins can vary between 0.25 km and 2 km. One observation is made of lidar signals accumulated over 12 sec or 90 km (see Figure 4-1). The horizontal measurement granularity within each observation is commendable. An observation is composed of N measurements of P pulses, with $2 \leq P \leq 600/N$, $N_{max} = 30$ and $(N \times P)_{max} = 600$.

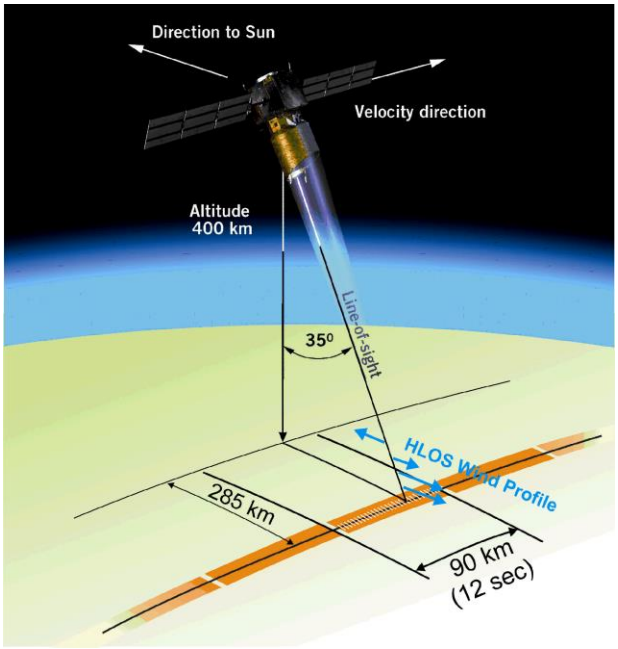

Figure 4-1: Artist's view of Aeolus and horizontal sampling.

4.1 Mission Concept

ADM-Aeolus is a dusk-dawn sun-synchronous orbit with a 7-day repeat cycle (i.e. 109 orbits) as shown in Figure 4-2. ALADIN is facing the night sector to keep the interfering background light as weak as possible to maximize signal-to-noise ratio. The main orbit parameters are:

- Inclination: 97°
- Mean altitude: 320 km
- Orbit duration: 90' 48"
- Velocity on orbit: 7.71 km/s
- Ground track velocity: 7.34 km/s
- Orbits per day: 15.86

The breakthrough of ADM-Aeolus with respect to current and past missions is to provide wind velocity profiles at medium scales everywhere and especially in the tropics. ADM-Aeolus, to be launched in 2017, will follow NASA ICESat/GLAS mission and NASA/CNES CALIPSO mission (launched in 2006 and still in operation 2014 – see [RD 2]). CALIPSO is integrated in the Afternoon Train for synergetic data processing with other space borne instruments (see Figure 4-3). ADM-Aeolus is to be followed by ESA EARTH-Care mission in late 2018. It is worth noticing that both ESA Aeolus and Earth-CARE missions will carry a High-Spectral-Resolution Lidar for unambiguous retrieval of particle optical properties while ICESat/GLAS and CALIPSO carry elastic backscatter lidar that do not separate molecules and particles contributions. Accordingly, ICESat/GLAS and CALIPSO require extra information on particle backscatter-to-extinction ratio (BER) to process the lidar signal. It results in an

inherent ambiguity of clouds and aerosols optical properties. On the contrary, ALADIN and EarthCARE will provide information on BER, a key microphysics variable for particle categorisation

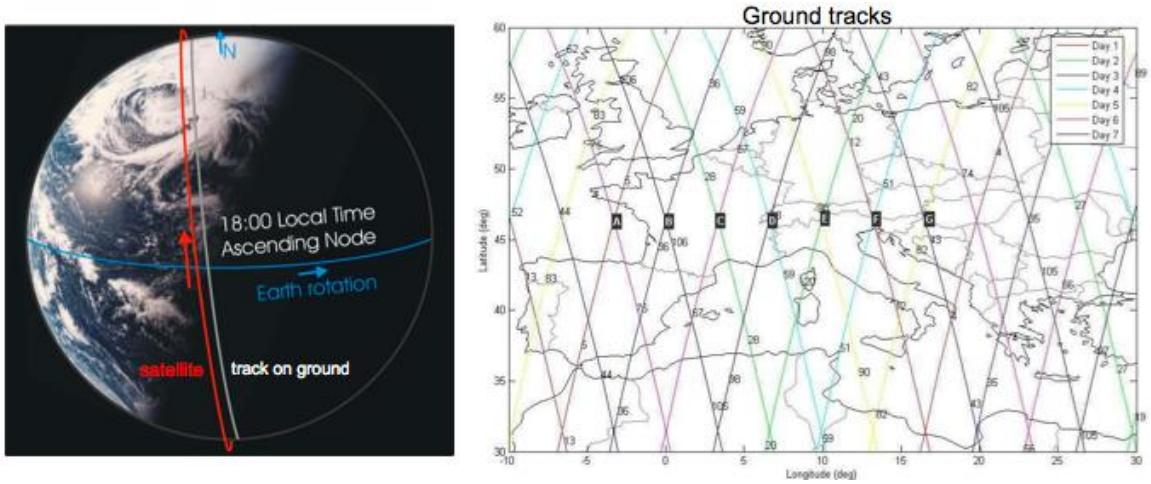

Figure 4-2: Example of ADM-Aeolus orbit over Western Europe.

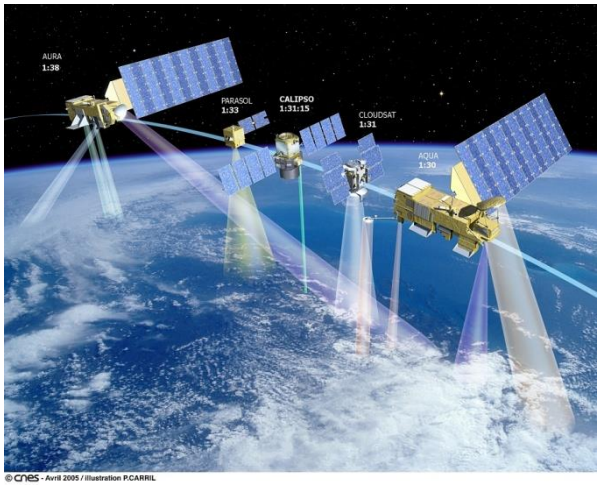

Figure 4-3: Artist's concept of the Afternoon-Train that includes CALIPSO with CALIOP Lidar, IIR and WAC, CLOUDSAT radar operating at 94 GHz, polarization radiometer POLDER, and AQUA and TERRA multiple radiometers platforms. Where IIR: infrared imager, WAC: wide angle camera. The time separation between the satellites is indicated.

The instrument variables for ALADIN on ADM-Aeolus and CALIOP on CALIPSO are listed in Table 4.1 for comparison.

The High Spectral Resolution Lidar concept was developed in the 60's (see [RD 4], [RD 5], [RD 6]) to derive unambiguous particle optical properties i.e. extinction and backscatter coefficients, with no assumption on particle backscatter-to-extinction ratio.

The ALADIN and CALIOP capabilities for particle optical properties are compared in Table 4.2.

Mission	**ADM-Aeolus**	**CALIPSO**
Lidar concept	**ALADIN** One single wavelength laser. High Spectral Resolution receiver separates the laser light scattered by molecules and particles into two signals in two channels	**CALIOP** 2 wavelengths laser. The receiver collects the laser light scattered by molecules and particles as one at each wavelength Polarization diversity provides with an additional channel
Nd-YAG laser Operating wavelength(s)	355 nm	532 nm 1064 nm
Transmitted energy per pulse	80 mJ	110 mJ at both wavelengths
Laser polarization	Circular *It precludes polarization diversity*	Linear It enables polarization diversity
Pulse duration	26 ns	20 ns
Pulse repetition frequency	50.5 Hz Continuous mode operation	20 Hz Continuous mode operation
Receiver telescope diameter	1.5 m	1 m
Receiver field-of-view (full angle)	18 µrad x 76 the telescope diffraction limit	130 µrad x 240 the telescope diffraction limit
Receiver	High Spectral Resolution *a*) Rayleigh channel: Dual Fabry-Pérot interferometer for light scattered by air molecules *b*) Mie channel: Fizeau interferometer for light scattered by particles	Total signal receiver (molecules and particles) at *a*) 532 nm // polarized *b*) 532 nm ⊥ polarized *c*) 1064 nm
Receiver spectral bandwidth (measured)	*a*) 0.63 pm for Rayleigh channel *b*) 0.067 pm for Mie channel	*a*) and *b*) 35 pm (at 532 nm) *c*) Standard interference filter (at 1064 nm)
Vertical resolution (range bin)	250, 500, 1000, 2000 m	30, 60, 180 m
Horizontal resolution (along satellite track)	3 to 7.5 km (accumulation of $P = 20\ to\ 50$ shots)	330 m (shot-to-shot)
Pointing of line-of-sight	37.5° off-nadir cross track	0 to 3° off nadir (to cancel spurious reflections)
Orbit height	408 km (498 km lidar range)	705 km
Lidar footprint at surface Footprint spacing	9 m 135 m	70 m 330 m

Table 4.1: Comparison of space borne lidar ALADIN and CALIOP.

Mission	Spatial sampling	Particle layer detection	Optical properties	Scene classification
ADM-Aeolus (HSRL)	Limited vertical resolution in range bins equal to 0.25, 0.50, 1 and 2 km.	Good. The Mie channel performs well at moderate SNR > 10.	Good. Based on HSRL capability to derive particle local optical depth per range bin LOD_p and the co-polar extinction-to-backscatter ratio *EBR* using the Rayleigh and Mie channels	Limited. Only two pieces of information provided by the Lidar (LOD_p and co-polar *BER*). No complementary instruments
CALIPSO (SEBL)	Vertical sampling at high resolution that provides flexibility. An accumulation is required to improve SNR.	Good (SNR > 10).	Limited. A colour ratio using 2 wavelengths and depolarization ratio are provided. But *a priori* knowledge of *EBR* is required to compute LOD_p.or backscatter or extinction coefficient	Good. Several pieces of information are provided by CALIOP, IIR and WFC, and other components of the A-Train (see Fig. 4.3)

Table 4.2: Comparison of retrieval capabilities for particle characteristics by ADM-Aeolus and CALIPSO.

4.2 Targets and signal reception

4.2.1 Atmospheric scattering

In this section we review the atmospheric scattering properties at 355 nm that are relevant to ALADIN Lidar measurements and those that raise an issue on the wind velocity measurements. The Lidar signal strength relies on air molecule backscatter coefficient for the Rayleigh channel and particle backscatter coefficient for the Mie channel. The main issue regarding the wind velocity measurement is due to the Rayleigh-Brillouin line shape that deviates significantly from a Gaussian shape in the lower atmosphere. Multiple scattering could impact the measurements of Doppler frequency shift and LOD, and so extinction coefficients, in case of Mie scatterers with large size, with respect to the 355-nm wavelength, e.g., dust, ice particles.

4.2.1.1 Scattered spectra by molecules and particles

Molecules and particles collide a huge number of times per second at pressure and temperature conditions prevailing in the lower atmosphere so they have the same mean kinetic energy. Accordingly particles are much slower due to the huge difference in mass between molecules (N_2 and O_2) and

particles. The thermal velocity distributions result in broad scattered spectra for molecules and narrow spectra for particles. The thermal velocities for molecules or particles are equally partitioned as positive and negative about the mean translational velocity v of the scattering medium.

In the middle atmosphere, above about 15 km of altitude, low pressure conditions prevail. It results in a Gaussian line shape for the molecular spectrum S_m:

$$S_m : \Delta f \longmapsto \frac{1}{\sigma_{\Delta f}\sqrt{2\pi}} e^{\frac{-(\Delta f)^2}{2\sigma_{\Delta f}{}^2}} \qquad \text{Eq. 4.1}$$

Where $\Delta f = f - fo - \Delta f_D$ with f_0 the laser emitted frequency and $\Delta f_D = -2\, v_r/\lambda_0$ the Doppler frequency shift associated to the radial wind velocity v_r. The standard deviations for frequency $\sigma_{\Delta f}$ and for radial velocity σ_{v_r}:

$$\sigma_{\Delta f} = \sqrt{\sigma_{las}^2 + 4\frac{\sigma_{v_r}^2}{\lambda_0^2}} \qquad \text{Eq. 4.2}$$

$$\text{with} \quad \sigma_{v_r}^2 = \frac{R_g T}{M} \qquad \text{Eq. 4.3}$$

Where $\lambda_0 = 355 \times 10^{-9}$ m, $R_g = 8.314\, J.mol^{-1}.K^{-1}$, T is the atmospheric temperature in Kelvin, M is the molar mass for (hypothetical) air molecules ($M = 2.9 \times 10^{-2}\, kg.mol^{-1}$) and $\sigma_{las} \approx 20 \times 10^6\ Hz$ is the spectral width of the emitted laser pulse (full width half maximum $FWHM \approx 50 \times 10^6\ Hz$).

The (FWHM) in Hz equal to:

$$\Delta f = \frac{4}{\lambda_0}\sqrt{\frac{2R_g\, T \ln 2}{M}} \qquad \text{Eq. 4.4}$$

At $355\, nm$, $\Delta f(MHz) = 224.8 \times \sqrt{T(K)}$. For atmospheric temperatures ranging between $210\, K$ and $320\, K$, the Gaussian linewidth varies between $3.25\, GHz$ and $4.0\, GHz$ in frequencies, and between $288\, m.s^{-1}$ and $357\, m.s^{-1}$ in thermal velocities. Thermal velocity definitely dominates the line width. Additional spectral broadening due to wind turbulence is usually small (a few $m.s^{-1}$).

The basic rationale to measure wind velocity using scattered spectra by molecules relies on a Gaussian line shape. Actually, a Gaussian function is fully determined by two parameters: the mean value ($f_0 + \Delta f_D$ here), and the standard deviation ($\sigma_{\Delta f}$ here). Therefore, wind estimation requires two independent lidar measurements to determine those parameters, especially Δf_D which contains wind information. In practice, two spectral intensities are measured using two sampling filters set at known frequencies. The two analyzing filters are provided by a Dual Fabry-Pérot interferometer. The two filter frequencies are set symmetrically about the laser emission frequency (see also Fig.4.11 for further explanation).

On the contrary, in the lower atmosphere below 15 km the spectra scattered by molecules have a Rayleigh-Brillouin line shape that is more pronounced as the pressure increases near the surface (see Fig. 4.4). Details on the inelastic Brillouin scattering can be found in [RD 13], [RD 14] and [RD 15]. A

Rayleigh-Brillouin line shape breaks the requirement on two independent measurements only to determine the line centre frequency. At UV wavelength and atmospheric pressure the Rayleigh-Brillouin line shape is not too different from a Gaussian line shape but it needs to be taken into account for accurate unbiased wind velocity measurements. In practice a Gaussian line shape is assumed first and then a correction is applied ([RD 26]). The Rayleigh-Brillouin spectrum displayed on Figure 4-4 is calculated using the S6 Tenti's Model proposed for N_2 molecules ([RD 14]). Recently, the S6 Tenti's Model has been verified experimentally for air at relevant pressure and temperature condition prevailing in the lower atmosphere [RD 15]. The results show that the S6 Tenti's Model is accurate and can be used for ADM-Aeolus. Notice that the total scattered optical power is the same for Rayleigh-Brillouin and Gaussian spectra. The Rayleigh-Brillouin line shape impacts the wind velocity measurements but not the molecular backscatter coefficient.

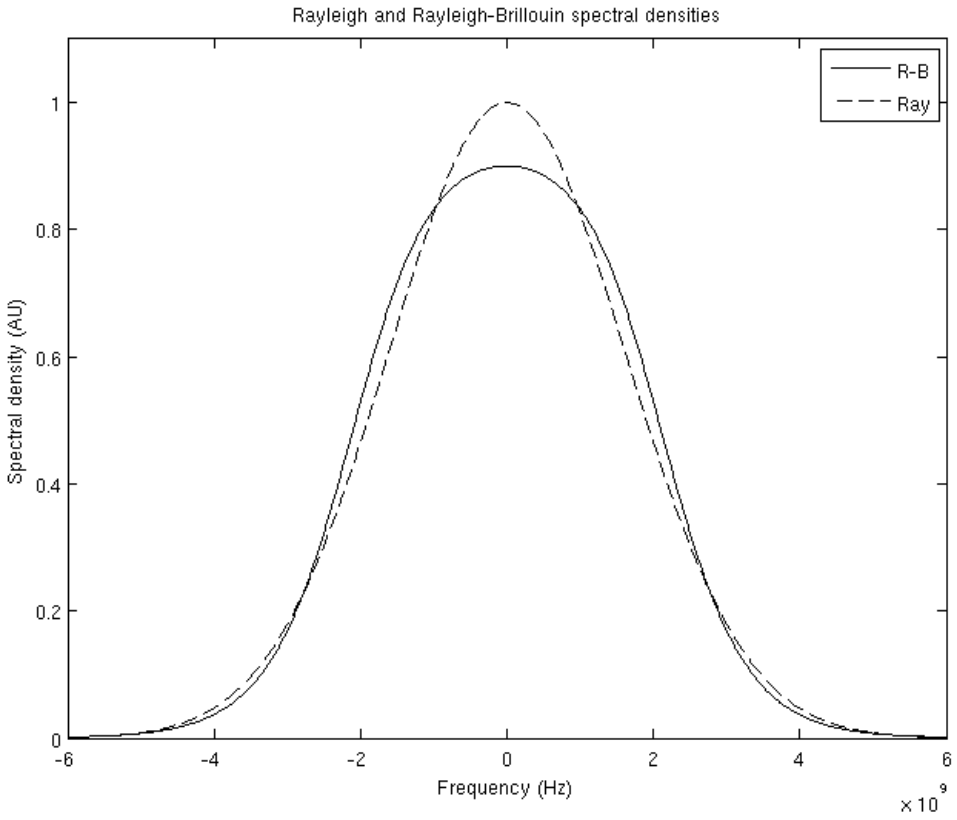

Figure 4-4: Spectral densities (arbitrary units) at $355\ nm$, $1013\ hPa$ and $288\ K$ for air molecules. Rayleigh-Brillouin (solid line) and Rayleigh (Gaussian) line shape (dashed line) are displayed.

In the lower atmosphere, the scattered spectrum due to particle Brownian motion is narrow (thermal velocity $\leq 1\ m.s^{-1}$). In practice the line width is limited by convolutions with i) the laser spectrum ($0.02\ pm$ and $25\ ns$ FWHM), ii) the Fizeau interferometer response function, iii) the wind turbulence in the scattering medium (a $\pm 1\ m.s^{-1}$ wind turbulence contributes for $5.5\ MHz$) Assuming that all these contributions have Gaussian line shapes, the overall line shape sums up the 3 contributions. For

ALADIN, the backscattered spectrum is ultimately limited by the Fizeau spectral bandwidth of $0.067\ pm$ or $159\ MHz$.

The overall scattered spectrum by molecules and particles displays a wide hill-looking spread of the initial laser pulse (about $3.82\ GHz$ FWHM) due to molecules and a sharp peak (about $50\ MHz$ FWHM, i.e. the laser's one) due to particles as shown on Figure 4-5.

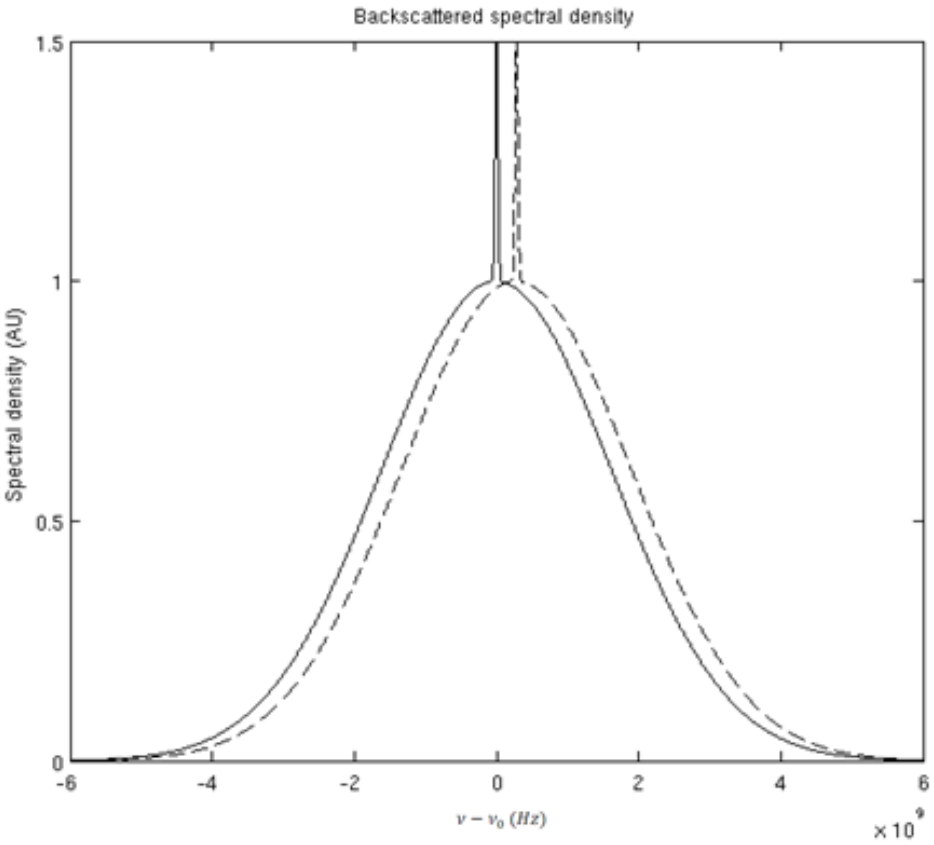

Figure 4-5: Example of backscattered spectrum by air molecules and particles at $\lambda = 355\ n$, $T = 300\ K$ and low pressure. The full line corresponds to zero wind ($v_r = 0$) and the dashed line to a wind radial velocity $v_r = 50\ m.s^{-1}$ (or $282\ MHz$ Doppler shift), respectively. The two spectra have Gaussian line shapes. The FWHM of the probing laser pulse is $0.02\ pm$; the mean masses are $m_m = 4.8 \times 10^{-26}\ kg$ for air molecules and $m_p = 4.2 \times 10^{-15}\ kg$ for particles (see [RD 26]).

4.2.1.2 Molecule and particle scattering coefficients

Molecular scattering is extensively used in ADM-Aeolus L2A algorithms so we provide here the equations to compute it using pressure and temperature information provided by NWP models. The molecular backscatter coefficient $\beta_m(m^{-1}.sr^{-1})$ and extinction coefficient $\alpha_m(m^{-1})$ for dry air at $355\ nm$ are computed using pressure $p(hPa)$ and temperature $T(K)$:

$$\beta_m(z) \cong 1.38 \left(\frac{550}{355}\right)^{4.09} \frac{p(z)}{1013}\frac{288}{T(z)}10^{-6} \qquad \text{Eq. 4.5}$$

$$\alpha_m(z) \cong 1.16 \left(\frac{550}{355}\right)^{4.09} \frac{p(z)}{1013}\frac{288}{T(z)} 10^{-5} \qquad \text{Eq. 4.6}$$

The coefficients 1.38 and 1.16 have been determined experimentally (see [RD 27]) and $\beta_m/\alpha_m = 3/8\pi$.

The atmospheric coefficients $\beta_m(z)$ and $\alpha_m(z)$ can be computed from surface pressure information assuming an exponential decay with height and temperature profile. Figure 4-6, left panel displays the molecular backscatter β_m and extinction α_m coefficients as a function of height (z) above surface (surface pressure equal to 1013 hPa, and atmospheric scale height equal to 8.5 km), whereas the right panel displays examples of particle extinction coefficients α_p for a cirrus cloud between 10 and 12 km and atmospheric boundary layer (ABL) from 0 to 2 km. The backscatter coefficient for cirrus and ABL can be calculated using appropriate extinction-to-backscatter values ([RD 35]).

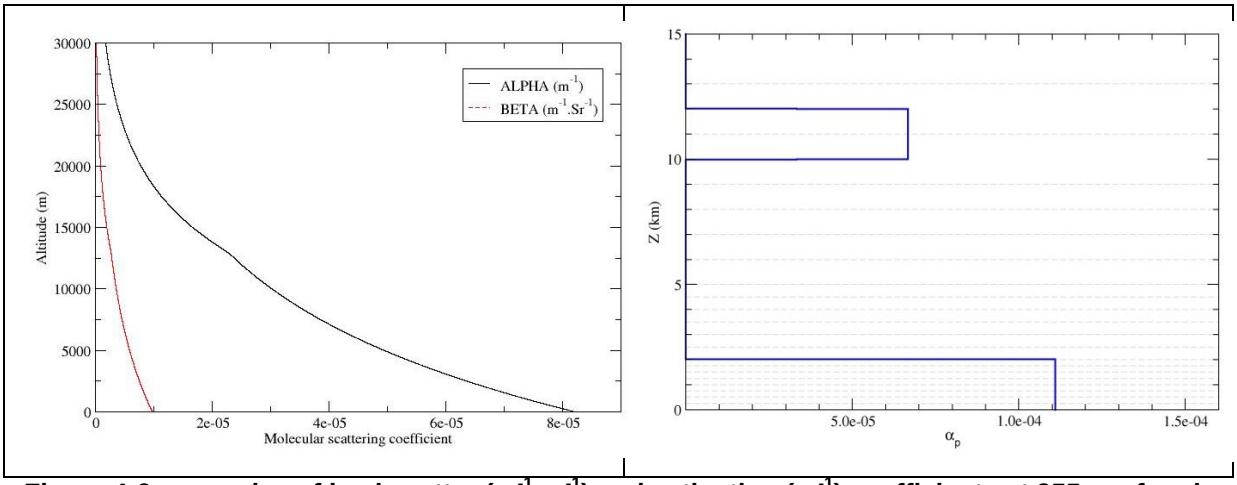

Figure 4-6: examples of backscatter (m^{-1} sr^{-1}) and extinction (m^{-1}) coefficients at 355 nm for air molecules (left), and extinction coefficient for particles in the atmospheric boundary layer from surface up to 2 km and in a cirrus cloud between 10 and 12 km (left). The cirrus optical depth is equal to 0.13, and the ABL optical depth equal to 0.22.

Figure 4-7 displays the two-way 355-nm molecular transmission (at 35°) from the satellite $T^2_{m,sat}(z)$ to the surface.

Figure 4-8 displays the apparent molecular backscatter coefficient at 355 nm not corrected for transmission $\beta_m(z)T^2_{m,sat}(z)$ at 35°.

$\beta_m(z)T^2_{m,sat}(z)$ Figure 4-9 displays the molecular backscatter at 355 nm $\beta_m(z)T^2_{m,sat}(z)T^2_{p,sat}(z)$ not corrected for molecular transmission at 35° in clear sky condition (blue) and cloudy condition (red) with a cirrus cloud between 10 and 12 km. The attenuation in the ABL has been disregarded for the sake of simplicity. The yellow lines outline the altitudes where the cirrus layer is present.

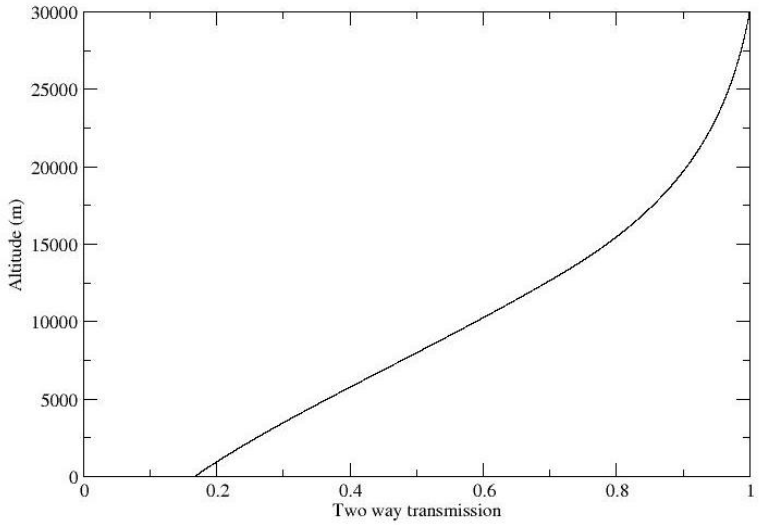

Figure 4-7: Two-way 355-nm molecular transmission (at 35°) from the satellite $T^2_{m,sat}(z)$ to the surface.

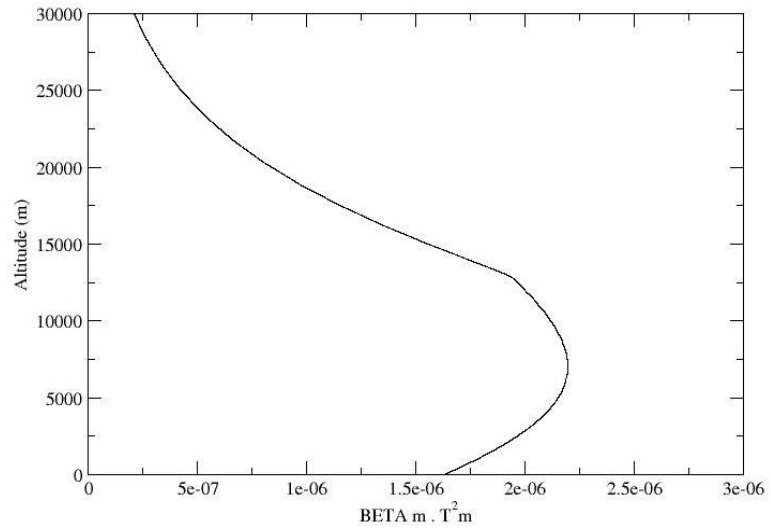

Figure 4-8: Apparent molecular backscatter coefficient at 355 nm not corrected for molecular two ways transmission.

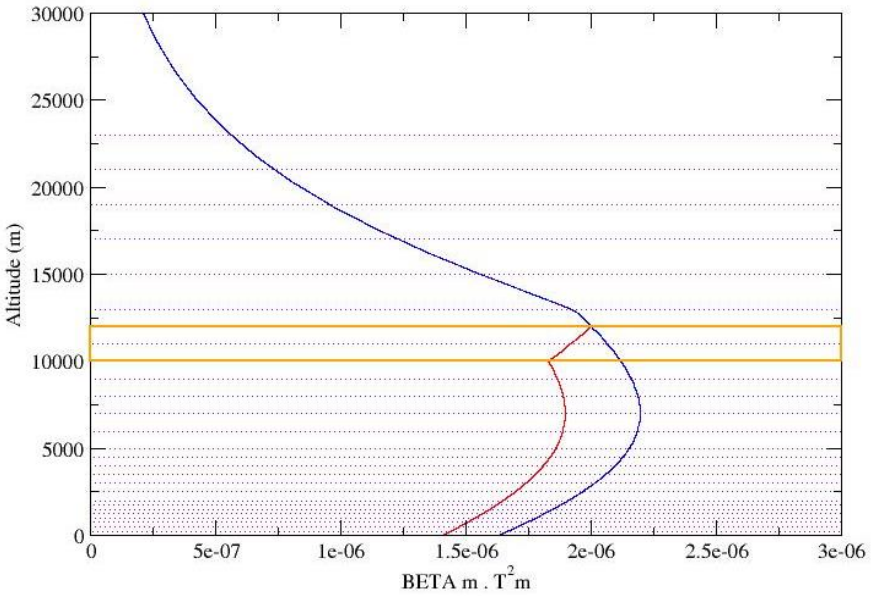

Figure 4-9: Apparent molecular backscatter coefficient at 355 nm not corrected for molecular 2 ways transmission effect in clear sky condition (in blue) and in presence of a cirrus cloud between 10 and 12 km (red). The yellow lines outline the altitudes where the cirrus cloud is present.

4.2.1.3 Multiple scattering effects

A single scattering assumption (SSA) is currently used in lidar technique to compute the instrumental budget. The instrumental budget that links the optical power collected by the telescope to the laser energy send out is known as the lidar equation. In the SSA, the laser photons are assumed to be scattered only once by the medium. The SSA is not valid for scenes containing large particle number densities or big particles. In lidar literature the denomination multiple scattering effects (MSE) is used indifferently for dense particle layers in which photons are scattered several times before they reach the receiver or forward diffraction effect by big particles. The practical effect of MSE is to reduce extinction losses by redirecting the scattered photons into the direction of propagation of the laser beam or back to the receiver. MSE associated to big particles do not impact Doppler frequency shift measurements.

Computation of MSE importance on lidar signal strength is complex and rarely practical in most experimental conditions. Pragmatically a multiplicative correcting factor $0 \leq \eta \leq 1$ is applied to reduce the extinction coefficient and so to increase the atmospheric transmission. MSE has been studied theoretically and numerically. It is shown that MSE combine in different ways the importance of the i) receiver telescope field-of-view (FOV) and/or laser beam divergence, ii) extinction coefficient (or optical depth) of the scattering layer and iii) lidar range [RD 23]. Forward diffraction is dominant for big particles, when the diffraction angle λ_0/d_p (FWHM) is of the order or smaller than the laser beam divergence, d_p is the particle equivalent diameter. For ADM-Aeolus with $\lambda_0 = 355\,nm$ and laser beam divergence $\cong 18\,\mu rad$, MSE are significant for very large particles with effective diameter $d_p \geq 1\,cm$. For such big particles $\eta \cong 0.5$ and the optical depth is reduced by a factor 2. For cirrus clouds with

smaller size particles, the particle number densities enter into consideration. The following examples are used to illustrate MSE for ADM-Aeolus using a lidar range $R \cong 500\ km$ (for a 400 km altitude and 37.5° perspective) and $\alpha_p \cong 0.5\ to\ 2\ km^{-1}$. Based on [RD 23], numerical applications show that SSA is valid i.e. $\eta \cong 1$ for i) particle diameters $d_p \leq 10\ \mu m$ in most practical conditions, ii) $d_p \cong 35\ \mu m$ with extinction coefficient $< 1\ km^{-1}$, iii) $d_p \cong 100\ to\ 350\ \mu m$, for optical depth $OD \leq 0.3$.

4.2.1.4 Polarisation

In ALADIN, the UV-laser pulse is linearly polarized but it is transmitted in the atmosphere with a circular polarization after passing through the quarter-wave plate set after the polarizer cube. If the atmospheric scattering processes do not change the state of polarization i.e. amplitude or phasing between the two EM radiation components, the backscattered light is circularly polarized in opposite direction of the incident light. Back to the receiver, and passing through the quarter wave plate then the polarization vector is linear but rotated by 90° with respect to the transmitted light polarization. Accordingly the scattered light is directed toward the DFP and Fizeau interferometers by the polarizer cube. Now, if the scattering processes modify the state of polarization of the two EM radiation components, after passing through the quarter wave plate the polarization is elliptical and only the fraction with cross linear polarization is directed to the DFP and Fizeau interferometers by the polarizer cube. The overall process is identical to what is happening with EM radiation with linear polarization scattered by ice crystals for example. The internal reflections rotate the plane of polarization.

The losses on lidar signals are the same for circular polarization as observed for linear polarization.It should be noted that only the co-polar component of the backscattered light is detected contrary to linearly polarized lidars that are able to retrieve both the co-polar and the cross-polar components of the backscattered signal. This is a drawback for ALADIN as for highly depolarizing targets (cirrus clouds, desert dusts or ash containing aerosols), the backscatter coefficient will systematically be underestimated by 50% to 75% in ice clouds and by up to 50% in dusts or ash aerosols. In fact, in such cases, the cross-polar component can be up to three times higher than the co-polar component and is not measured by the receiver. Consequently, the backscatter-to-extinction ratio will also be underestimated by the same factor by which the backscatter coefficient is underestimated.

4.2.2 High Spectral Resolution receiver

The ALADIN HSR receiver is displayed in Figure 4-10. More information can be found in [RD 7] and [RD 8].

The Rayleigh channel implements a dual Fabry-Pérot interferometer ([RD 9], [RD 10]), and the Mie channel implements a Fizeau interferometer. A Fizeau interferometer uses two plane reflecting surfaces with a slight deviation from exact parallelism. Fizeau interferometers are usually used in

wavemeters to measure wavelength to a guaranteed accuracy as high as ±0.0001 nm[1] in laboratory environment.

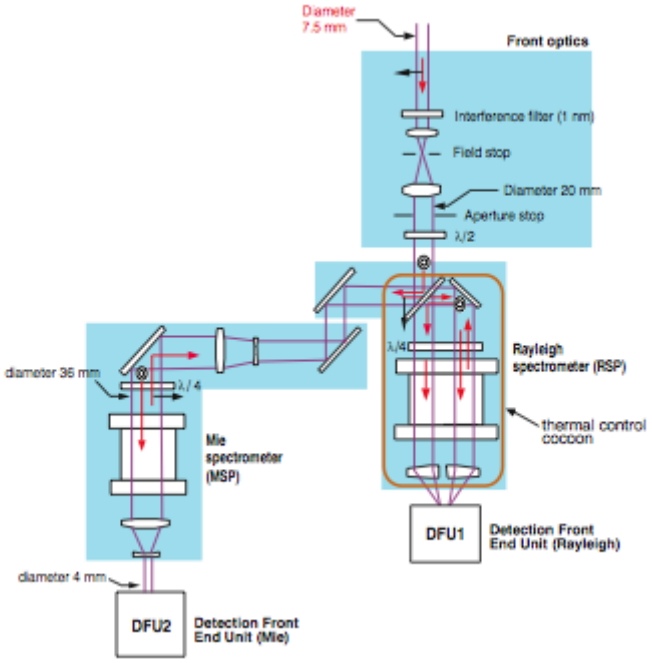

Figure 4-10: ALADIN receiver displaying the Rayleigh and Mie spectrometers based on a dual Fabry-Pérot and Fizeau interferometer, respectively.

In Figure 4-10 the incoming atmospheric signal is sent first into the Fizeau interferometer. The spectrum of the light sent onto the dual Fabry-Pérot interferometer is impacted by the reflection on the Fizeau interferometer. Then the light is sent into one part of the Dual Fabry-Pérot and then into the other part. Two receivers have the same field-of-view (full-angle) $FOV = 18\,\mu rad$. Considerations on FOV is important for 1) MSE (see 4.2) and 2) statistical signal fluctuations associated to speckle effects. The number of speckle cells (N_c) calculated onto the receiver pupil measure the coherence of the scattered light. N_c=1 for coherent light. N_c drives the lidar signal statistical fluctuations. N_c is calculated using $N_c \cong 1 + (D_{re}/D_c)^2$, with $D_c \cong \lambda\, R/D_{ta}$, and where D_{re}, $D_c$and $D_{ta} \cong 2(FOV)R$ are the diameter of the receiver, the speckle cell coherence diameter and the laser beam diameter, respectively. A numerical application with $D_{re} = 1.5\,m$ ([RD 8]) shows that $N_c \cong (150)^2$, indicating weak statistical fluctuations of the lidar signals.

[1] See for example i) Bristol Instrument (http://www.bristol-inst.com), ii) Burleigh, III) TOPTICA photonics, and relevant papers iv) Y.H. Meyer and M. N. Nenchev, "Tuning of dye lasers with a reflecting Fizeau wedge", Opt. Commun., 35, 119, 1980; C. Cahen et al., "Wavelength stabilization and control of the emission of pulsed dye lasers by means of multiple-beam Fizeau interferometer", Rev. Phys. Appl., 16, 353-358, 1981; B. Morris *et al.*, "Fizeau wavemeter for pulsed laser wavelength measurement", Appl. Opt. 23, 3862 (1984); D. F. Gray *et al.*, "Simple compact Fizeau wavemeter", Appl. Opt. 25 (8), 1339 (1986); C. Reiser and R. B. Lopert, "Laser wavemeter with solid Fizeau wedge interferometer", Appl. Opt. 27 (17), 3656 (1988); W Kedzierski *et al*, « A Fizeau wavemeter with single-mode optical fibre coupling", *J. Phys. E: Sci. Instrum.* 21, 796, 1988.

4.2.2.1 Dual Fabry-Pérot Spectrometer

The Rayleigh spectrometer is made of a dual Fabry-Pérot interferometer (DFPI). The dual FP is designed to have two band-passes centred symmetrically on each side of the scattered spectrum for zero wind velocity ([RD 8], [RD 9], [RD 10]). In Figure 4-11 the scattered spectrum is displayed with a Gaussian line shape for simplicity. The spectrum for zero wind is centred about the laser frequency f_0. The frequencies are normalized with respect to f_0. The spectral response of the two sub channels A (right) and B (left) are displayed in blue (FP-A) and green (FP-B) dashed lines, respectively. The blue and green shaded areas represent the number of photons N_A and N_B. The Rayleigh response is defined as $R_R = (N_A - N_B)/(N_A + N_B)$ where N_A and N_B are the number of photons collected in the two filters. In principle, for zero wind $N_A = N_B$ and $R_R = 1$. If the spectrum is positively Doppler-shifted by 282 MHz associated to a 50 m s^{-1} radial velocity, N_A increases while N_B decreases. The Doppler frequency shift can be measured using the Rayleigh response. The relationship is linear in the neighbourhood of $R_R = 1$ but deviate from linearity for large wind velocities. Also, because the incoming light is sent successively (and not simultaneously) on the two FP parts it results in $N_A \neq N_B$ for zero wind velocity.

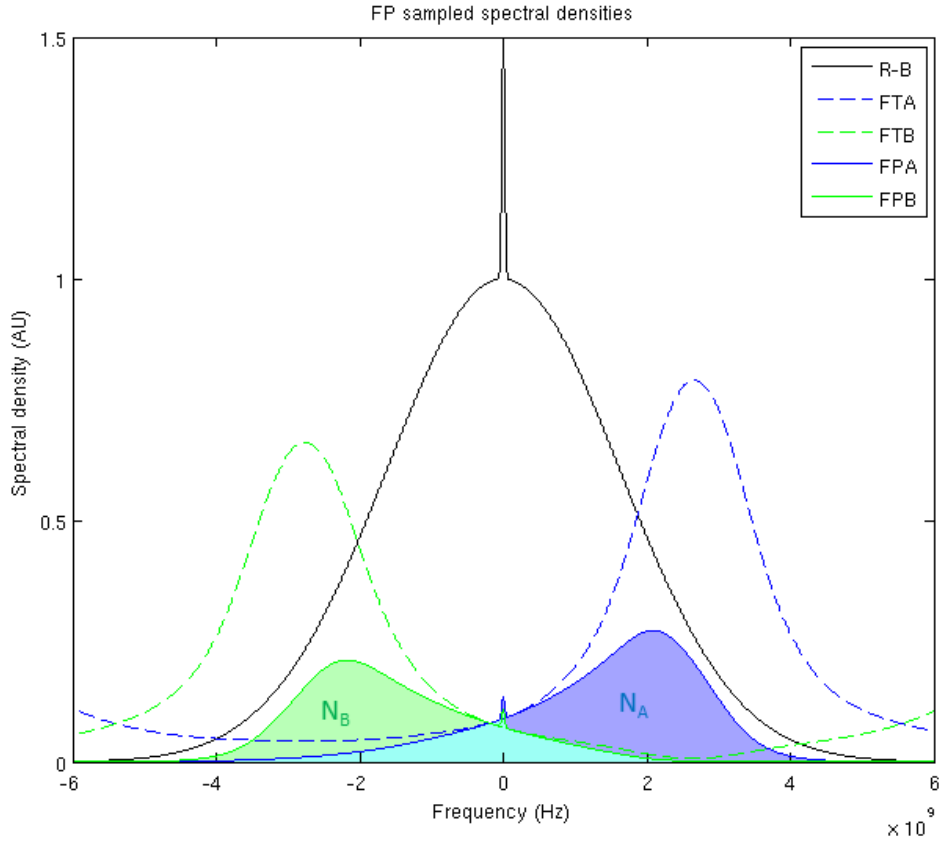

Figure 4-11: Total backscattered spectrum (black), dual Fabry-Pérot transfer functions FTA (dashed blue) and FTB (dashed green) and associated filtered spectral densities FPA (blue) and FPB (green). Molecular spectrum amplitude is unity and particle spectrum amplitude is 0.5 (scattering ratio of 1.5 to emphasize the phenomenon). Molecules are mean air molecules ($4.8 \times 10^{-26}\ kg$) and particles are water droplets of 1 μm radius ($4.2 \times 10^{-15}\ kg$). Transmission peaks are 0.79 (FTA) and 0.65 (FTB).

The two FP band-passes FP-A and FP-B have a FWHM = 0.7 pm (or 1.67 GHz) and are separated by 2.3 pm (or 5.47 GHz). As shown in Figure 4-11, the spectral sampling by the two Fabry-Pérot interferometers FP-A (blue) and FP-B (green) includes a small fraction of the Mie scattering peak. The two transfer functions FT-A and FT-B overlap, so the two measurements are not independent: it results in an inherent crosstalk between the two channels as discussed in section 4.2.2.4. The contributions of the residual Mie signal to FP-A and FP-B depend on frequency shift (wind velocity).

4.2.2.2 Fizeau spectrometer

The Mie spectrometer is a Fizeau interferometer. The free spectral range (FSR) of the Fizeau interferometer is equal to $0.90\,pm$ but only a fraction of it is imaged onto the detector so the useful spectral range is $USR = 0.63\,pm$ or $1500\,MHz$. The FWHM of the Fizeau interferometer transfer function is $0.067\,pm$ or about $159\,MHz$.

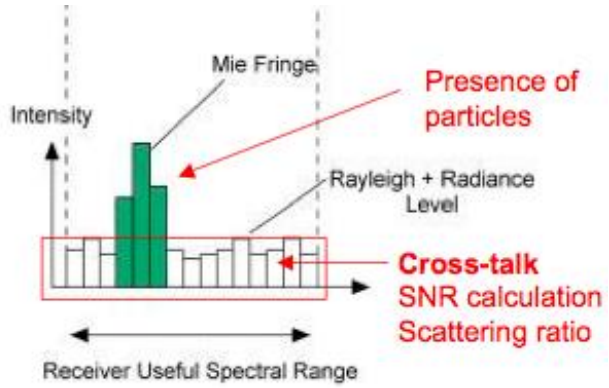

Figure 4-12: Fizeau interferometer spectrum.

As shown in Figure 4-12, the backscattered spectrum by particles is superimposed onto the molecular signal. There exists an optical crosstalk that needs to be calibrated.

4.2.2.3 CCD photo-detector

Accumulation of lidar signal on CCD photo-detector during 12 seconds results in a 2D pattern composed of 16 columns (for spectral resolution) and 25 rows (for vertical resolution). A separate and identical parallel register is added to store and accumulate the electric charges after detection. Each captured scene is shifted very quickly to the storage area. The expected quantum efficiency at 355 nm is 82%. This CCD architecture provides sensitivity and readout noise similar to those obtained by an image intensified camera. The readout noise is very low and allows quasi photon counting regime to be achieved. ACCD response non-linearity has been characterised and demonstrated as negligible [RD 12]. Due to memory effects, vertical cross-talk between the vertical bins will occur (see [RD 11]). This overlap has been assessed in the A2D campaign ($1\,\mu s$ for an integration time of $4.2\,\mu s$, see [RD 12] p153) but still needs to be characterised for ALADIN.

4.2.2.4 Rayleigh and Mie channels

For historical reasons in the mission and convenience in the nomenclature, both systems {spectrometer + detector} have been named accordingly to the scattering type they aim at measuring. On the one hand, the dual Fabry-Pérot, mainly sampling molecular contribution, has been named the *Rayleigh channel*; on the second hand, the Fizeau, focusing on the particulate scattering, has been named the *Mie channel*.

4.2.2.4.1 Rayleigh channel

The Rayleigh channel makes use of a dual Fabry-Pérot interferometer (DFPI). The atmospheric signal is transmitted successively through the first band-pass (A) of the Fabry-Pérot interferometer and then through the second band-pass (B). The two channels are physically separated after the Fabry-Pérot interferometer thanks to a Meslin lens made of two half lenses separated by a blind area. As the detector area is « separated » into half parts dedicated to A and B respectively (8 columns in each part) the pixels contributing to the « A » or « B » signals are well identified. The number of pixels covered by each spot may vary according to possible defocusing. The diameter of each spot for the best focus is 3.3 pixels. The lidar signals N_A and N_B transmitted through the band-passes A and B respectively are obtained after summation over a vertical range (accumulation time) chosen *a priori*.

The two channels A and B of the Dual Fabry-Pérot partially sample the full spectrum (see Figure 4-11). Figure 4-13 displays the ratio of the total backscattered spectrum sampled by the dual Fabry-Pérot to the molecular backscattered spectrum as a function of horizontal wind velocity. The ratio roughly varies from 0.325 to 0.333 when velocity varies from $-50\,m.s^{-1}$ to $50\,m.s^{-1}$. Therefore, the correcting factor has to be a function of Mie signal strength and wind horizontal velocity. Three values of the scattering ratio SR, 1, 1.02 and 1.04 are considered to highlight the impact of particle backscatter on the spectrum filtered by the DFPI. Larger variations are expected for larger scattering ratios. Using the extinction values in Figure 4-6, one can calculate that for cirrus cloud the scattering ratio varies between 1.60 (for BER = 0.02 sr^{-1}) and 5 (for BER = 0.05 sr^{-1}), while for ABL aerosols, the scattering ratio varies between 1.25 and 2.5 for BER equal to 0.025 and 0.04 sr^{-1} (continental and marine aerosols respectively).

4.2.2.4.2 Mie channel

The scattered light spectrum is displayed over the 16 columns of the CCD photo-detector (see Figure 4-12). It is assumed that the fringe is perfectly parallel to the 16 columns. The useful Fizeau spectral range is 1500 MHz; each column has an equivalent spectral width of 93.75 MHz or 17 m/s. The Mie spectrum, FWHM =159 MHz, covers 1.7 pixels. The raw data are 16 numerical values, one per column. The signal of each column is obtained for a given height bin, after accumulation over P shots of the photons detected in the 24 rows of that column. Mie SNR at BRC level (accumulation over 600 laser pulses) was found to vary from 15 to 40 for boundary layer aerosols, around 150 for cirrus clouds and 30 to 50 for desert dusts. These values were extracted from the L1B processor.

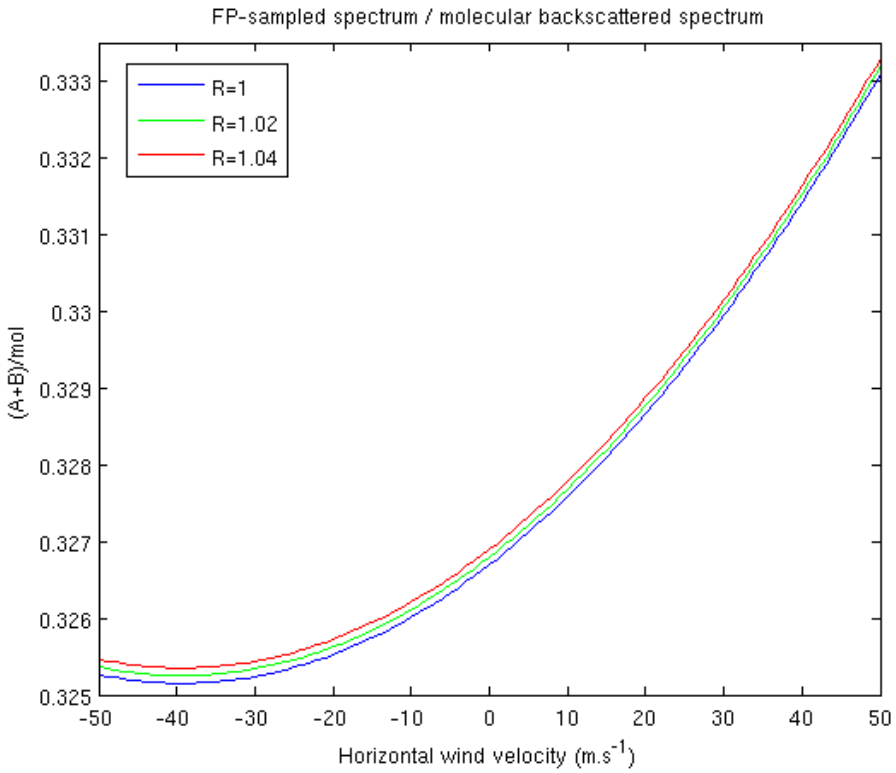

Figure 4-13: Ratio of molecular spectrum sampled by the dual Fabry-Pérot to the total backscattered spectrum as a function of frequency for three different scattering ratios R. Molecules and particles are the same as in Figure 4-11.

4.3 Optical property products from ADM-Aeolus

As described above, the Aeolus Mie and Rayleigh spectrometers enable the separation of the molecular and particle contributions to the total atmospheric backscatter signal, which makes it a so-called High Spectral Resolution Lidar (HSRL). HSRLs generally suffer from channel crosstalk, which must be quantified and corrected during the retrieval of optical properties products. The magnitude of the crosstalk between the Aeolus Mie and Rayleigh channels will be quantified during dedicated calibration procedures, which are further detailed in the following sections.

The molecular and particle backscatters are separated and crosstalk-corrected during a serie of signal inversions, as described below. Whereas the magnitude of the molecular scattering is of little scientific interest, the retrieval of aerosol backscatter and extinction is important input to air quality monitoring as well as climate and atmospheric process studies, including cloud formation.

This document describes the algorithms for the Aeolus operational retrieval of particle and molecular optical property products (the Aeolus L2a products). In the following, the molecular and particle scattering products are indexed $_m$ and $_p$ respectively:

- the particle extinction coefficient, α_p
- the particle backscattering coefficient, β_p

- the scattering ratio (SR) $\rho = 1 + \frac{\beta_p}{\beta_m} \geq 1$
- the particle backscatter to extinction coefficient (BER) $BER = \frac{\beta_p}{\alpha_p}$
- the Slant Local Optical Depth $L_p = \int_{R_1}^{R_2} \alpha_p(r)\, dr$, for a layer between ranges R_1 and R_2 from satellite.
- The local optical depth $\delta_p = \int_{z_2}^{z_1} \alpha_z(z) \frac{dz}{\cos\theta(z)}$, is the same as L_p but projected on the vertical axis. The angle $\theta(z)$ between the local vertical and the beam direction is depending on altitude because of the rotundity of the Earth, as shown on Figure 4-14

To retrieve these values, in addition to instrumental data, the processor has access to meteorological data (pressure, temperature), climatology data (BER) and calibration data to adjust instrumental constants.

As shown in Figure 4-14, the viewing angle is $\theta \approx 37.5°$, the satellite-to-the-surface range is $R_0 = h_0/\cos\theta$, the satellite height is h_0=320 km. In the lower atmosphere the looking angle $\theta(z)$ depends on altitude according to an Earth spherical shape.

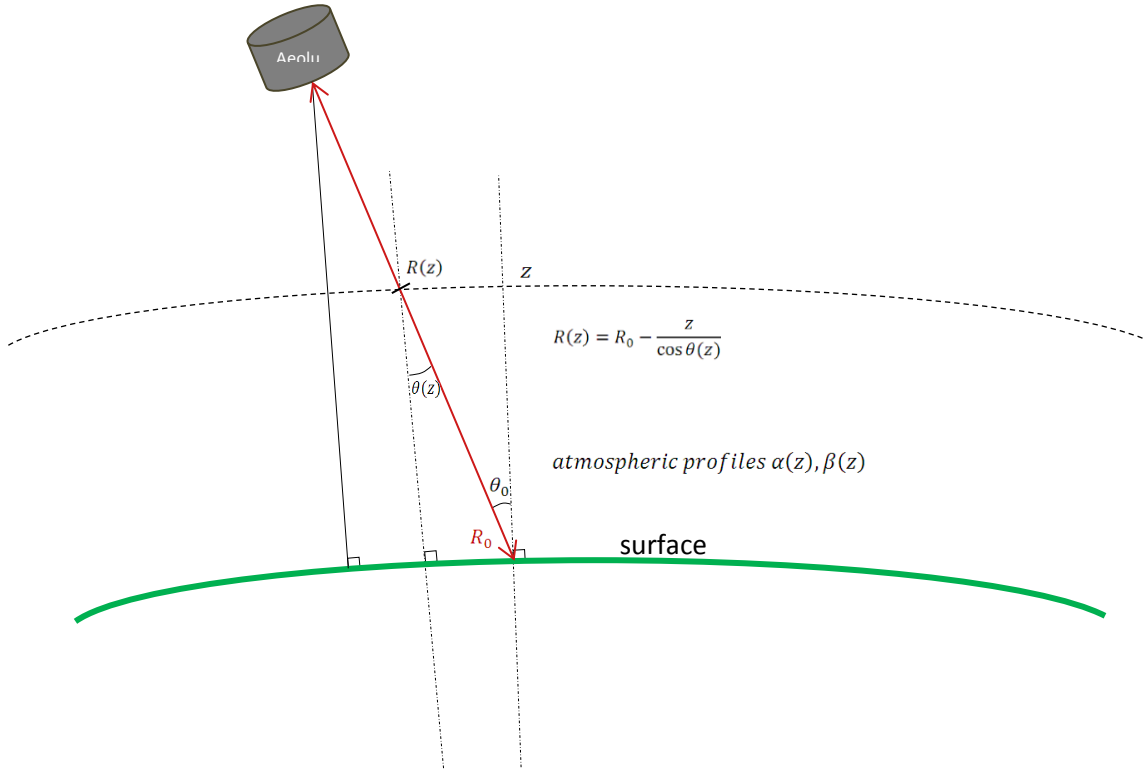

Figure 4-14: Schematic view of the geometry

4.4 ADM-Aeolus chain of processors

Instrumental data will be processed by the ground segment of the mission. There are three levels, 0, 1 and 2, divided in sublevels, A, B or C. Here is a description of their products:

- Level 0 product: unprocessed raw Annotated Instrument Source Packets data, time ordered, with header and annotation data included,
- Level 1A product: reconstructed measurement data and calibrated house-keeping information,
- Level 1B product: preliminary horizontal (HLOS) wind products (zero-wind corrected), processed calibration parameters, product confidence data and annotation data, this product is the main input of the Level 2A [RD 28].
- Level 2A product: Aeolus particle spin-off products, processed by the algorithms described in sections 6 and 7 [RD 29].
- Level 2B product: Aeolus consolidated wind products, taking into account corrections due to actual atmospheric pressure and temperature distributions, additional geophysical parameters and error quantifiers [RD 30].
- Level 2C product: Aeolus assisted two-component wind data, result of NWP assimilation processing [RD 30].

Data Product	Content
Level 0	Time-ordered source packet streams which are reorganized into different measurement data sets according to the system and instrument mode
Level 1A	Housekeeping source packet fully processed, AOCS source packets (geolocation) processed and assigned to measurement data, measurement data unprocessed.
Level 1B	Fully processed, calibrated and georeferenced measurement data including HLOS winds, viewing geometry, ground echo data and product confidence data (PCD).
Level 2A	Additional aerosol/cloud optical properties, as optical depth, extinction coefficient, backscatter coefficient and PCD.
Level 2B	L2B products represent “consolidated” HLOS wind data and include corrections using actual pressure and temperature information as obtained from numerical models from a NWP centre. Additional corrections are based on retrieved optical properties. Measurements are grouped after a scene classification.
Level 2C	L2C product contain two-component wind vector profiles on the location of the ADM-Aeolus ground track as obtained after the assimilation process of L2B products at a NWP centre. L2C products mainly contain information from the NWP model.

Table 4.1: Content of data products for ADM-Aeolus

5 AEOLUS L2A ARCHITECTURE

The level 2A processor of ADM-Aeolus will basically proceed as follows:

- Synthetic molecular signal computation
- Observation analysis:
    - SCA/ICA sequence:
        - Matching bin determination
        - Crosstalk correction,
        - SCA
        - ICA
    - MCA
- Group analysis
    - Feature finding
    - Matching bin determination
    - Crosstalk correction,
    - SCA
    - Scene classification

An overarching view of the level 2A processor is proposed in Figure 5-1.

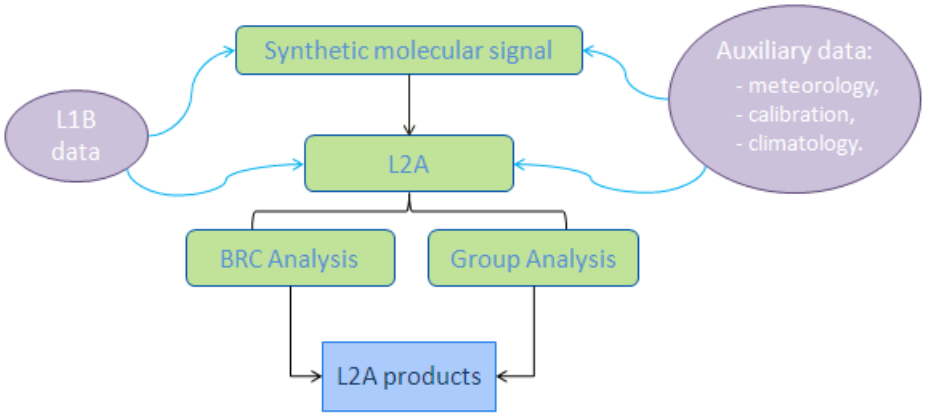

Figure 5-1: Simple flow chart of the L2A.

First of all, the synthetic molecular signal is computed from meteorological data and L1B geolocation data. Then, data are provided to the core of the L2A processor. Firstly, it runs an analysis at BRC level, accumulating 600 laser pulses. It enables to have sturdy, though widely averaged, products. Secondly, a group analysis is launched. Horizontal features are found and average products are

calculated over these features. By applying the L2A analysis at group level instead of BRC level, aerosol properties are retrieved with a higher resolution.

Detailed diagrams for both BRC and group analyses are proposed in Figure 5-2 and Figure 5-3 respectively.

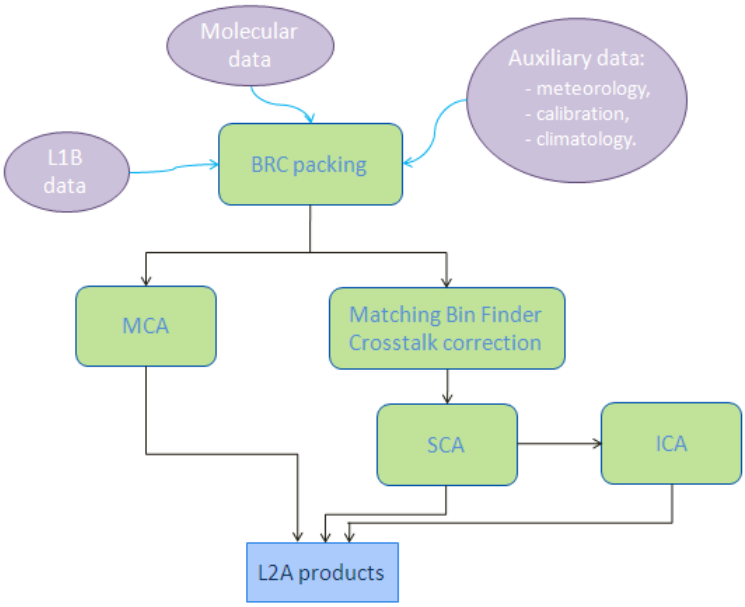

Figure 5-2: Flow chart of the BRC analysis.

The first step of the BRC analysis (Fig. 5.2) is data packing. All input data are reorganised at profile level. Afterwards, effective algorithms are applied. On the one hand, the Mie Channel Algorithm (MCA) will retrieve particle characteristics from Mie channel useful signals, the scattering ratio from the L1B, the calibration coefficients. This algorithm needs an a priori information on the BER taken from an appropriate climatology. In that sense, the MCA is similar to the retrieval algorithms used for standard elastic backscatter lidar like CALIPSO. These retrievals suffer from the uncertainty on the BER value that depends on many parameters such as the composition of particle mixture, particle size, shape, orientation and refractive index. The MCA algorithm is described in section 6.6 of the present document. No pre-processing is needed before it is applied. On the other hand, the SCA and the ICA need crosstalk-corrected data but provides unambiguous solutions by avoiding the use of a climatological value for the BER. Crosstalk correction requires matching bins between both Rayleigh and Mie scales. Hence the succession on the right branch of the flowchart of Figure 5-2. The SCA retrieves backscatter coefficient, extinction coefficient, local optical depth, scattering ratio and the backscatter-to-extinction ratio. Once the SCA has been run, the retrieved backscatter coefficient is provided to the ICA so that the particle-loaded bins can be located in the profile. The ICA is designed to sharpen inner-bin resolution and discriminates various filling cases for a loaded height level and selects the more relevant. Eventually, retrieved products from the different algorithms are packed together and saved.

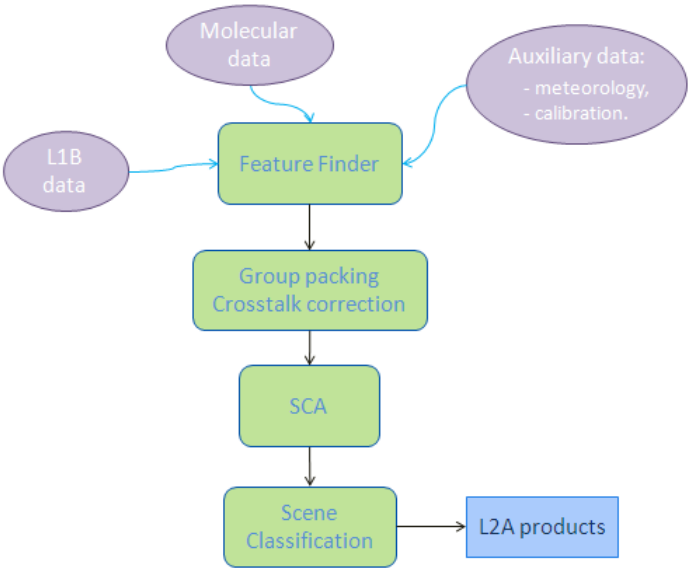

Figure 5-3: Flow chart of the group analysis.

The first step of the group analysis (Figure 5-3) is to locate the bins where the features are and to pack them into groups. Then, data are accumulated and averaged over these groups and the crosstalk between Rayleigh and Mie channels is corrected, paving the way to the SCA. Group-wise SCA products are provided to the scene classification algorithm. Together with NWP data, it discriminates features between clouds and aerosols.

6 CORE ALGORITHMS

6.1 Basic equations of the measurement principle

In this subsection, lidar equations are manipulated, under some hypotheses, to pave the way for the explanation of core algorithms.

6.1.1 Range resolved lidar equations

The range resolved atmospheric signals at telescope entry, $X(r)$ for molecules and $Y(r)$ for particles, expressed in $m^{-3}.sr^{-1}$, are the signals a perfect instrument would collect from the scattering of a pulse of a 1W power. As a matter of fact, a great use of these two quantities is made in the L2A. It enables to avoid handling instrumental calibration constants and, this way, to lighten equations. They involve four atmospheric variables α_p, β_p, α_m and β_m. They are all function of range r from the satellite:

$$X(r) = \frac{\beta_m(r)}{r^2}\exp\left(-2\int_0^r \left(\alpha_m(u) + \alpha_p(u)\right)du\right) \quad \text{Eq. 6.1}$$

$$Y(r) = \frac{\beta_p(r)}{r^2}\exp\left(-2\int_0^r \left(\alpha_m(u) + \alpha_p(u)\right)du\right) \quad \text{Eq. 6.2}$$

If both Rayleigh channel and Mie channel spectral samplings were perfect, the range resolved lidar signals $s_{ray}(r)$ and $s_{mie}(r)$ would be:

$$s_{ray}(r) = K_{ray}N_pE_0\frac{\beta_m(r)}{r^2}\exp\left(-2\int_0^r \left(\alpha_m(u) + \alpha_p(u)\right)du\right) \quad \text{Eq. 6.3}$$

$$s_{mie}(r) = K_{mie}N_pE_0\frac{\beta_p(r)}{r^2}\exp\left(-2\int_0^r \left(\alpha_m(u) + \alpha_p(u)\right)du\right) \quad \text{Eq. 6.4}$$

The signals $s_{ray}(r)$ and $s_{mie}(r)$ are expressed in engineering units. E_0 is the laser pulse energy, in Joules, and N_p is the number of pulses accumulated to generate a profile. After proper calibration they are expressed in optical power (W) or photons per sec by dividing the optical power by hf (h is the Planck's constant and f the optical frequency) or photoelectrons using the CCD quantum efficiency: $q.s_m(R)/hf$. Here, K_{ray}, and K_{mie} are the instrumental calibration constant parameters for the Rayleigh and Mie channels, respectively. Details about their computation can be found in [RD 25].

Yet, as shown in Figure 4-11 and Figure 4-12, dual Fabry-Pérot and Fizeau spectrometers do not respectively sample pure molecular and pure particle spectra and there is crosstalk between the Rayleigh and Mie channels. Accounting for this, the two range-resolved lidar equations for the Rayleigh and Mie signals are:

$$s_{ray}(r) = \frac{K_{ray}N_pE_0}{r^2}\left(C_1(P,T,f)\beta_m(r) + C_2(f)\beta_p(r)\right)\exp\left(-2\int_0^r \left(\alpha_m(u) + \alpha_p(u)\right)du\right) \quad \text{Eq. 6.5}$$

$$s_{mie}(r) = \frac{K_{mie} N_p E_0}{r^2} \left(C_4(P,T,f)\beta_m(r) + C_3(f)\beta_p(r)\right) \exp\left(-2\int_0^r \left(\alpha_m(u) + \alpha_p(u)\right) du\right) \qquad \text{Eq. 6.6}$$

Or, making use of X and Y to lighten expressions:

$$s_{ray}(r) = K_{ray} N_p E_0 (C_1(P,T,f)X(r) + C_2(f)Y(r)) \qquad \text{Eq. 6.7}$$

$$s_{mie}(r) = K_{mie} N_p E_0 (C_4(P,T,f)X(r) + C_3(f)Y(r)) \qquad \text{Eq. 6.8}$$

$C_1(P,T,f)$ and $C_4(P,T,f)$ are the fraction of molecular backscatter that is actually detected by respectively the Rayleigh and Mie channels, depending on the pressure P, the temperature T and the Doppler-shift frequency f. $C_2(f)$ and $C_3(f)$ are the fraction of particle backscatter that is actually detected by respectively the Rayleigh and Mie channels, depending on the sole Doppler-shift frequency because of the negligibility of Brownian motion on particle backscatter. The six instrumental coefficients, K_{ray}, K_{mie}, C_1, C_2, C_3 and C_4 are calibrated as described in [RD 25]. Note that the current version of the document does not take into account the Rayleigh-Brillouin effect. However, the atmosphere simulator now includes the Rayleigh-Brillouin scattering to simulate the lidar signals that would be received by ALADIN. Note that molecular backscatter and extinction coefficients are determined as a function of pressure (see section 4.2.1.2)

Here is a reminder of the computation of the instrumental coefficients, adapted from [RD 25].

The functions C_1 and C_4 are computed by convolving the transmission characteristics of the Fabry-Pérot, T_A and T_B[1], and the Fizeau, T_{Fiz}, by the expected shape of the molecular return.

$$C_1(P,T,f) = \frac{1}{K_1}\int \left(T_A(\mu) + T_B(\mu)\right) I_{mol}(P,T,\mu - f) d\mu$$

$$C_4(P,T,f) = \frac{1}{K_4}\int T_{Fiz}(\mu) I_{mol}(P,T,\mu - f) d\mu \qquad \text{[RD 25] (3)}$$

K_1 and K_4 are normalization constants, the normalization criteria are expressed by,

$$C_1(1000hPa, 300K, 0MHz) = C_4(1000hPa, 300K, 0MHz) = 1 \qquad \text{[RD 25] (2)}$$

so that:

$$K_1 = \int \left(T_A(\mu) + T_B(\mu)\right) I_{mol}(1000hPa, 300K, \mu) d\mu$$

$$K_4 = \int T_{Fiz}(\mu) I_{mol}(1000hPa, 300K, \mu) d\mu \qquad \text{[RD 25] (4)}$$

Constants C_2 and C_3 are directly proportional to $T_A + T_B$ and T_{Fiz} respectively:

[1]The Dual Fabry-Perrot receives light reflected from the Fizeau interferometer, TA and TB take the efficiency of the reflection onf the Fizeau interferomter into account.

$$C_2(\nu) = \frac{T_A(f) + T_B(f)}{K_1} \qquad \text{and} \qquad C_3(\nu) = \frac{T_{Fiz}(f)}{K_4}$$ [RD 25] (5)

In case the aerosol backscatter and extinction are negligible (SR almost 1), the term $Y(r)$ in Eq. 6.7 and Eq. 6.8 disappears and Kray and Kmie can be determined following:

$$K_{ray} = \frac{s_{ray}(r)}{N_p E_0 C_1\big(P(r), T(r), \nu(r)\big) X(r)} \quad \text{and} \quad K_{mie} = \frac{s_{mie}(r)}{N_p E_0 C_4\big(P(r), T(r), \nu(r)\big) X(r)}$$ Eq. 6.9

6.1.2 Range bin accumulated lidar signals

Due to the width of accumulation height bins (from $250\,m$ to $2000\,m$), atmosphere cannot be considered to be constant over these bins: sampling the whole bin does not boil down to sampling a point in the middle of the bin. Then, the classical average – for thin bins – of the equations Eq. 6.7 and Eq. 6.8 does not apply for ALADIN. The mechanism of signal accumulation has to be represented by the summation of these equations over the bins. Hence the accumulated lidar signals (capital letters) in range bins i of Mie and Rayleigh channels:

$$S_{ray,i} = \int_{R_{ray,i-1}}^{R_{ray,i}} s_{ray}(r)dr$$ Eq. 6.10

$$S_{mie,i} = \int_{R_{mie,i-1}}^{R_{mie,i}} s_{mie}(r)dr$$ Eq. 6.11

Pure molecular and particle accumulated signals at telescope entry for a bin i are written:

$$X_i = \int_{R_{i-1}}^{R_i} \frac{\beta_m(r)}{r^2} \exp\left(-2\int_0^r \big(\alpha_m(u) + \alpha_p(u)\big)\,du\right) dr$$ Eq. 6.12

$$Y_i = \int_{R_{i-1}}^{R_i} \frac{\beta_p(r)}{r^2} \exp\left(-2\int_0^r \big(\alpha_m(u) + \alpha_p(u)\big)\,du\right) dr$$ Eq. 6.13

6.1.3 Range bin accumulated lidar equations

6.1.3.1 Starting equations

Equations vary from one channel to another because bin scales differ. For instance, Rayleigh channel integration may start higher than for Mie channel, and Mie bins may be narrower than Rayleigh ones in the ABL. Instrumental data coming from/used by each channel is sampled (instrumental data) and/or interpolated (simulation and calibration data) along appropriate scale. From now, the indexation "*ray*" will refer to quantities interpolated in Rayleigh bins, and "*mie*" in Mie bins.

For instance, molecular and particle signals at telescope entry (Eq. 6.12 and Eq. 6.13), decomposed along Rayleigh scale, are written:

$$X_{ray,i} = \int_{R_{ray,i-1}}^{R_{ray,i}} \frac{\beta_m(r)}{r^2} \exp\left(-2\int_0^r \left(\alpha_m(u) + \alpha_p(u)\right) du\right) dr \qquad \text{Eq. 6.14}$$

$$Y_{ray,i} = \int_{R_{ray,i-1}}^{R_{ray,i}} \frac{\beta_p(r)}{r^2} \exp\left(-2\int_0^r \left(\alpha_m(u) + \alpha_p(u)\right) du\right) dr \qquad \text{Eq. 6.15}$$

And for Mie scale:

$$X_{mie,i} = \int_{R_{mie,i-1}}^{R_{mie,i}} \frac{\beta_m(r)}{r^2} \exp\left(-2\int_0^r \left(\alpha_m(u) + \alpha_p(u)\right) du\right) dr \qquad \text{Eq. 6.16}$$

$$Y_{mie,i} = \int_{R_{mie,i-1}}^{R_{mie,i}} \frac{\beta_p(r)}{r^2} \exp\left(-2\int_0^r \left(\alpha_m(u) + \alpha_p(u)\right) du\right) dr \qquad \text{Eq. 6.17}$$

L1B data, i.e. accumulated lidar signals, in the Rayleigh channel $S_{ray,i}$ and Mie channel $S_{mie,i}$ is not corrected from the inherent crosstalk between the two channels. The relevant lidar equations are then:

$$S_{ray,i} = K_{ray} N_p E_0 \left(C_{1,ray,i} X_{ray,i} + C_{2,ray,i} Y_{ray,i}\right) \qquad \text{Eq. 6.18}$$

$$S_{mie,i} = K_{mie} N_p E_0 \left(C_{4,mie,i} X_{mie,i} + C_{3,mie,i} Y_{mie,i}\right) \qquad \text{Eq. 6.19}$$

Where $C_{k,xxx,i}$ is the C_k coefficient averaged over the i[th] bin of the xxx channel. One can recognize a system of two equations with four unknown variables X_{ray}, Y_{ray}, X_{mie} and Y_{mie}. Provided there is a matching between some bins of both scales, for instance the Rayleigh bin i_{ray} and the Mie bin j_{mie}, this system locally boils down to two equations and two unknown variables and can be solved *in those bins*. Then:

$$S_{mie,j_{mie}} \equiv S_{mie,i_{ray}} \qquad \text{Eq. 6.20}$$

And the system

$$\begin{cases} S_{ray,i} = K_{ray} N_p E_0 \left(C_{1,ray,i} X_{ray,i} + C_{2,ray,i} Y_{ray,i}\right) \\ S_{mie,i} = K_{mie} N_p E_0 \left(C_{4,ray,i} X_{ray,i} + C_{3,ray,i} Y_{ray,i}\right) \end{cases} \qquad \text{Eq. 6.21}$$

Is inverted in:

$$\begin{cases} X_{ray,i} = \dfrac{K_{mie} C_{3,ray,i} S_{ray,i} - K_{ray} C_{2,ray,i} S_{mie,i}}{N_p E_0 K_{ray} K_{mie} \left(C_{1,ray,i} C_{3,ray,i} - C_{2,ray,i} C_{4,ray,i}\right)} \\ Y_{ray,i} = -\dfrac{K_{mie} C_{4,ray,i} S_{ray,i} - K_{ray} C_{1,ray,i} S_{mie,i}}{N_p E_0 K_{ray} K_{mie} \left(C_{1,ray,i} C_{3,ray,i} - C_{2,ray,i} C_{4,ray,i}\right)} \end{cases} \qquad \text{Eq. 6.22}$$

Or, if scale indices are temporarily left aside,

$$\begin{cases} X_i &= \dfrac{K_{mie}C_{3,i}S_{ray,i} - K_{ray}C_{2,i}S_{mie,i}}{N_p E_0 K_{ray} K_{mie}\left(C_{1,i}C_{3,i} - C_{2,i}C_{4,i}\right)} \\ Y_i &= -\dfrac{K_{mie}C_{4,i}S_{ray,i} - K_{ray}C_{1,i}S_{mie,i}}{N_p E_0 K_{ray} K_{mie}\left(C_{1,i}C_{3,i} - C_{2,i}C_{4,i}\right)} \end{cases} \quad \text{Eq. 6.23}$$

Finally, if there is bin matching, the even structure of an HSRL enables to separate molecular signal from particle signal and, this way, to start working on the retrieval of the coefficients α_p and β_p. This step is called crosstalk correction in the algorithms. If there is no crosstalk (ideal HSRL system), i.e. if $C_{2,i} = C_{4,i} = 0$, the inverted system boils down to

$$\begin{cases} X_i = \dfrac{S_{ray,i}}{N_p E_0 K_{ray} C_{1,i}} \\ Y_i = \dfrac{S_{mie,i}}{N_p E_0 K_{mie} C_{3,i}} \end{cases} \quad \text{Eq. 6.24}$$

6.1.3.2 Approximated equations

Before going forth, the averaging behaviour of the instrument has to be taken into account. The vertical resolution will be the size of a bin and will only permit direct retrieval of averaged quantities over the bin. From now on, quantities are assumed to be constant over bins. Then, for instance,

$$L_{p,i} = \int_{R_{i-1}}^{R_i} \alpha_p\,(r)dr \approx \alpha_{p,i}\Delta R_i \quad \text{Eq. 6.25}$$

And, for $R_{i-1} \le r \le R_i$, noting $L_{p,sat}$ the optical depth due to the presence of particles above the topmost bin:

$$L_{p,sat} = \int_0^{R_1} \alpha_p(r)dr \quad \text{Eq. 6.26}$$

$$\int_0^r \alpha_p(u)du = L_{p,sat} + \sum_{k=1}^{i-2} \int_{R_k}^{R_{k+1}} \alpha_p(u)du + \int_{R_{i-1}}^{r} \alpha_p(u)du \quad \text{Eq. 6.27}$$

$$\int_0^r \alpha_p(u)du \approx L_{p,sat} + \sum_{k=1}^{i-1} \alpha_{p,k}\Delta R_k + \alpha_{p,i}(r - R_{i-1}) \quad \text{Eq. 6.28}$$

The same stands for molecular quantities. It is then possible to define molecular and particle transmissions from the satellite (still assuming no significant attenuation over the first bin) to the bottom of the bin i, $T_{m,sat,i}$ and $T_{p,sat,i}$:

$$T_{m,sat,i} = \exp\left(-\int_0^{R_i} \alpha_m(u)du\right) = \exp\left(-\left(L_{m,sat} + \sum_{k=1}^{i} \alpha_{m,k}\Delta R_k\right)\right) \quad \text{Eq. 6.29}$$

$$T_{p,sat,i} = \exp\left(-\int_0^{R_i} \alpha_p(u)du\right) = \exp\left(-\left(L_{p,sat} + \sum_{k=1}^{i} \alpha_{p,k}\Delta R_k\right)\right) \quad \text{Eq. 6.30}$$

This way, pure molecular signal at telescope entry can be written slightly differently, taking constant terms out of the integral:

$$X_i = \int_{R_{i-1}}^{R_i} \frac{\beta_m(r)}{r^2} \exp\left(-2\int_0^r \left(\alpha_m(u) + \alpha_p(u)\right)du\right)dr \quad \text{Eq. 6.31}$$

$$X_i = T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 \int_{R_{i-1}}^{R_i} \frac{\beta_m(r)}{r^2} \exp\left(-2\int_{R_{i-1}}^r \left(\alpha_m(u) + \alpha_p(u)\right)du\right)dr \quad \text{Eq. 6.32}$$

Then, considering that molecular and squared range quantities are weakly varying over a range bin,

$$X_i = \frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 \beta_{m,i}}{R_{mean,i}^2} \int_{R_{i-1}}^{R_i} \exp\left(-2\int_{R_{i-1}}^r \alpha_m(u)du\right) \exp\left(-2\int_{R_{i-1}}^r \alpha_p(u)du\right)dr \quad \text{Eq. 6.33}$$

Inside the integral, variations of the molecular term $\int_{R_{i-1}}^r \alpha_m(u)du$ are limited and can be approximated by its mean value over the bin, i.e.:

$$\begin{aligned} \frac{1}{\Delta R_i}\int_{R_{i-1}}^{R_i}\int_{R_{i-1}}^{r} \alpha_m(u)du\,dr &\approx \frac{1}{\Delta R_i}\int_{R_{i-1}}^{R_i} \alpha_{m,i}(r - R_{i-1})dr \\ &\approx \frac{1}{\Delta R_i}\alpha_{m,i}\left[\frac{(r-R_{i-1})^2}{2}\right]_{R_{i-1}}^{R_i} \\ &\approx \alpha_{m,i}\frac{\Delta R_i}{2} \\ &\approx \frac{L_{m,i}}{2} \end{aligned} \quad \text{Eq. 6.34}$$

Hence,

$$\begin{aligned} X_i &= \frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 \beta_{m,i}}{R_{mean,i}^2} e^{-L_{m,i}} \int_{R_{i-1}}^{R_i} exp\left(-2\int_{R_{i-1}}^r \alpha_p(u)du\right)dr \\ &= \frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 \beta_{m,i}}{R_{mean,i}^2} e^{-L_{m,i}} \int_{R_{i-1}}^{R_i} exp\left(-2\alpha_{p,i}(r - R_{i-1})\right)dr \\ &= \frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 \beta_{m,i}}{R_{mean,i}^2} e^{-L_{m,i}} \left[\frac{e^{-2\alpha_{p,i}(r-R_{i-1})}}{-2\alpha_{p,i}}\right]_{R_{i-1}}^{R_i} \\ X_i &= \frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 \beta_{m,i}}{R_{mean,i}^2} e^{-L_{m,i}} \left(\frac{1 - e^{-2L_{p,i}}}{2\alpha_{p,i}}\right) \end{aligned} \quad \text{Eq. 6.35}$$

One can notice that a raw approximation would lead to $e^{-2L_{m,i}}$ instead of $e^{-L_{m,i}}$.

The same reasoning leads to a simplified expression of Y_i:

$$Y_i = \frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 \beta_{p,i}}{R_{mean,i}^2} e^{-L_{m,i}} \left(\frac{1 - e^{-2L_{p,i}}}{2\alpha_{p,i}} \right) \quad \text{Eq. 6.36}$$

In these approximated expressions, terms due to particles can be strictly isolated from molecular ones. The exploitation of this property is the base of the SCA.

6.2 The standard correct algorithm

We propose to retrieve the particle characteristics in a crosstalk-corrected range bin using a normalized integrated two-way transmission (NITWT) assuming a uniform particle layer filling of the entire range bin.

The NITWT concept is new. It is used to build a standard correct algorithm (SCA). The assumption of one single particle layer filling the entire range bin with a constant extinction coefficient is restrictive and will not be met in most practical situations. Partial fillings of range bin are addressed in section 7 with an iterative correct algorithm (ICA).

6.2.1 Normalised integrated two-way transmission

A quick handling of the molecular signal at telescope entry X_i (or $X_{i,obs}$ to insist on the fact it is observed signal) leads to a simplified expression. It has been noticed that molecular and particulate terms are well separated.

Let's now focus on synthetic molecular signal $X_{i,sim}$. It supposes an atmosphere only constituted of molecules and therefore no particulate term appears in the transmission to the satellite:

$$\begin{aligned} X_{i,sim} &= \int_{R_{i-1}}^{R_i} \frac{\beta_m(r)}{r^2} \exp\left(-2 \int_0^r \alpha_m(u) du \right) dr \\ &= \frac{T_{m,sat,i-1}^2 \beta_{m,i}}{R_{mean,i}^2} \int_{R_{i-1}}^{R_i} \exp\left(-2 \int_{R_{i-1}}^r \alpha_m(u) du \right) dr \end{aligned} \quad \text{Eq. 6.37}$$

To stay consistent with previous developments, the same approximation (slow and small variations of molecular characteristics) is made for the molecular transmission:

$$\begin{aligned} X_{i,sim} &\approx \frac{T_{m,sat,i-1}^2 \beta_{m,i}}{R_{mean,i}^2} \int_{R_{i-1}}^{R_i} e^{-L_{m,i}} dr \\ &\approx \frac{T_{m,sat,i-1}^2 \beta_{m,i}}{R_{mean,i}^2} \Delta R_i e^{-L_{m,i}} \end{aligned} \quad \text{Eq. 6.38}$$

The great similarity between the expressions of $X_{i,obs}$ and $X_{i,sim}$ suggests to make the ratio of both terms to remove some quantities:

$$\begin{aligned}
\frac{X_{i,obs}}{X_{i,sim}} &= \frac{\frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 \beta_{m,i}}{R_{mean,i}^2} e^{-L_{m,i}} \left(\frac{1-e^{-2L_{p,i}}}{2\alpha_{p,i}}\right)}{\frac{T_{m,sat,i-1}^2 \beta_{m,i}}{R_{mean,i}^2} \Delta R_i e^{-L_{m,i}}} \\
&= T_{p,sat,i-1}^2 \left(\frac{1-e^{-2L_{p,i}}}{2L_{p,i}}\right) \\
\frac{X_{i,obs}}{X_{i,sim}} &= T_{p,sat,i-1}^2 H(2L_{p,i})
\end{aligned} \qquad \text{Eq. 6.39}$$

With:

$$\begin{aligned}
H \;:\; & \mathbb{R} \to]0;+\infty[\\
& x \mapsto \frac{1-e^{-x}}{x} \\
& 0 \mapsto 1
\end{aligned} \qquad \text{Eq. 6.40}$$

Its graph is presented in Figure 6-1. This ratio still involves $L_{p,sat}$. This quantity is located above measured bins and remains inaccessible through provided lidar data. It can be removed by a normalisation by the value of this ratio in the first matching bin. For the sake of simplicity, let's suppose both Rayleigh and Mie scales perfectly match. The number of this bin is then 1. If it is not true, and this is always the case in practice, numbering suffers a translation which brings nothing but useless sophistication here. The resulting normalised ratio is called Normalised Integrated Two-Way Transmission ($NITWT$) and is written for the range bin $i \geq 2$:

$$\begin{aligned}
NITWT_i &= \frac{X_{i,obs}}{X_{1,obs}} \frac{X_{1,sim}}{X_{i,sim}} \\
&= \frac{\exp\left(-2\left(L_{p,sat} + \sum_{k=1}^{i-1} L_{p,k}\right)\right) H(2L_{p,i})}{\exp\left(-2L_{p,sat}\right) H(2L_{p,1})} \\
NITWT_i &= T_{p,1,i-1}^2 H(2L_{p,i})
\end{aligned} \qquad \text{Eq. 6.41}$$

$T_{p,i,j}$ is the particle transmission of bins i to j, $i \leq j$:

$$T_{p,i,j} = \exp\left(-\int_{R_i}^{R_{j+1}} \alpha_p(u)du\right) \qquad \text{Eq. 6.42}$$

$$T_{p,i,j} = \exp\left(-\sum_{k=i}^{j} L_{p,k}\right) \qquad \text{Eq. 6.43}$$

So eventually, the computation of the ratio of observed crosstalk-corrected molecular signal to simulated molecular signal leads to an expression made of sole accessible particle terms. This result is built assuming:

- small variations of molecular terms over the range bin,
- small variations of range squared terms,
- a homogeneous particle filling of the range bin,

- bin matching between Rayleigh and Mie scales,
- the accurate knowledge of molecular backscatter and extinction, i.e. of pressure and temperature. Air density at a given altitude is only weakly differing from auxiliary NWP forecast or analysis fields available to the Aeolus L2A algorithm.

6.2.2 Algorithm definition

6.2.2.1 Retrieval of extinction coefficient

The retrieval of a standard correct solution for $\alpha_{p,i}$ is based on the NITWT. Eq. 6.41 can be rewritten to emphasize the potential of the NITWT:

$$\frac{1}{T_{p,1,i-1}^2} NITWT_i = H\left(2L_{p,i}\right) \qquad \text{Eq. 6.44}$$

Or:

$$\alpha_{p,i} = \frac{1}{2\Delta R_i} H^{-1}\left(\frac{1}{T_{p,1,i-1}^2} NITWT_i\right) \qquad \text{Eq. 6.45}$$

With H^{-1} the inverse function of H on $]0;1[\cup]1;+\infty[$, extended by continuity on $]0;+\infty[$.

$$\begin{array}{rcl} H^{-1} & : &]0;+\infty[\to \mathbb{R} \\ & & x \mapsto H^{-1}(x) \\ & & 1 \mapsto 0 \end{array} \qquad \text{Eq. 6.46}$$

The right-hand term of Eq. 6.45 involves two kinds of data: ΔR_i and $NITWT_i$ that can be computed by sole data of the current bin, but $T_{p,1,i-1}$ requires information from all previous bins. A recursive retrieval is then proposed. In a first step, the system is initialised in the first matching bin and then a recurrence relation is proposed.

Initialisation:

Values must be set in the first matching bin. By construction, $NITWT_1 = 1$. Therefore, the information is lost in this bin. An hypothesis is required to initialize the system. The less hazardous and the easiest one to control (see 6.2.2.2) is $\alpha_{p,1} = 0$, i.e. the bin is clear. Eventually, the initialisation boils down to:

$$\begin{cases} NITWT_1 = 1 \\ \alpha_{p,1} = 0 \\ T_{p,1,1} = 1 \end{cases} \qquad \text{Eq. 6.47}$$

Recurrence relation

If we now suppose that the system is perfectly known in the bin -1, $i \geq 2$, let's see if it can be determined in the bin i.

$NITWT_i$ and ΔR_{i} are known from bin data and $T_{p,1,i-1}$ is given by the result of the previous bin. Then, $\alpha_{p,i}$ is computed from Eq. 6.45. In practice, after iterations on $L_{p,i}$ a satisfying value of the estimate $\hat{L}_{p,i}$ can be found. The criterion is that the error is less than an empirically chosen threshold th (10^{-5}):

$$\left| H\left(2\hat{L}_{p,i}\right) - \frac{1}{T_{p,1,i-1}^2} NITWT_i \right| < th \qquad \text{Eq. 6.48}$$

Afterwards, once $L_{p,i}$ is determined, it is floored to zero if it is negative. This point is discussed above in this section. Then, the transmission $T_{p,1,i}$ is computed:

$$T_{p,1,i} = T_{p,1,i-1} e^{-L_{p,i}} \qquad \text{Eq. 6.49}$$

The system is then determined for the rank *i* and the recurrence is demonstrated.

Discussion: sole acceptance of $\alpha_{p,i} \geq 0$

Let's suppose that for some reason (see 6.2.3.1), $\frac{1}{T_{p,1,i-1}^2} NITWT_i$ is misestimated. To set ideas, let us suppose that it occurs in the first bin (supposed clear by the algorithm) and that the optical depth is $\hat{L}_{p,1} = \Delta L_{p,1} > 0$.

Then, for $i = 2$,

$$H\left(2\hat{L}_{p,2}\right) = \frac{1}{1} \frac{X_{2,obs}}{X_{2,sim}} < H\left(2L_{p,2}\right) = \frac{1}{e^{-2\Delta L_{p,1}}} \frac{X_{2,obs}}{X_{2,sim}} \qquad \text{Eq. 6.50}$$

A look at the graph of the function H (Figure 6-1) shows that it will result in an overestimation of the slant local optical depth: $\hat{L}_{p,2} > L_{p,2}$ and an underestimation of the transmission $\hat{T}_{p,1,2} < T_{p,1,2}$.

For $i = 3$, we would have:

$$H\left(2\hat{L}_{p,3}\right) = \frac{1}{\hat{T}_{p,1,2}^2} \frac{X_{3,obs}}{X_{3,sim}} > H\left(2L_{p,3}\right) = \frac{1}{T_{p,1,2}^2} \frac{X_{3,obs}}{X_{3,sim}} \qquad \text{Eq. 6.51}$$

Again, a look at the graph shows that there will be this time an underestimation of the SLOD. For instance, in case of a clear 3rd bin, $L_{p,3} = 0$, the estimate would be negative, *i.e.* the algorithm would find a fluorescence[1]!

This oscillation between underestimation and overestimation will propagate all along the calculation, as an inherent default of the recursive algorithm.

[1] The fluorescence spectrum is broad and shifted to longer wavelength. The contribution of the fluorescence light is negligible because a 1 nm bandwidth spectral filter centered at the laser emission is used at the receiver.

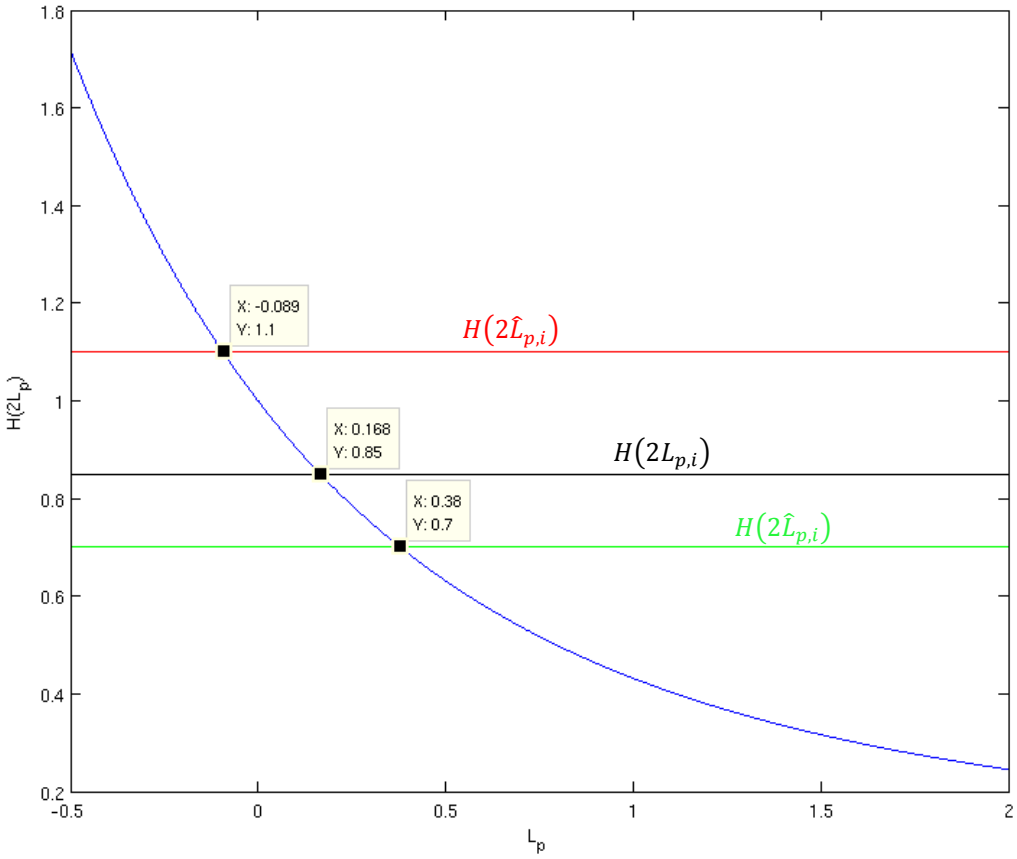

Figure 6-1: Function H (blue curve). If the error leads $H(2\hat{L}_{p,i})$ above unity (red line), the estimated optical depth is negative. On the contrary, if the transmission $T_{p,sat,i-1}$ is overestimated (*i.e.* $\hat{L}_{p,i-1}$has been underestimated), then $H(2\hat{L}_{p,i}) < H(2L_{p,i})$ and $L_{p,i}$ is overestimated.

Unfortunately, this phenomenon is hard to correct because observed oscillations could be due to irregular features of particles in the profile. Nonetheless, when the retrieved extinction is negative, i.e. when the algorithm suggests that aerosols emit light, this error[1] is adjusted by flooring extinction values to zero. This enables to partially limit the propagation of errors (calculated in section 6.2.3) in extinction retrieval.

Now, a mean to check for the presence of particles in this first matching bin is needed. The retrieval of the backscatter coefficient, in addition to its inherent value, plays very well this role.

[1] Fluorescence and phosphorescence are negligible for aerosols at considered wavelength.

6.2.2.2 Retrieval of backscatter coefficient

Coming back to the simplified expressions of the molecular and particulate signals at telescope entry (Eq. 6.35 and Eq. 6.36), one could notice their high similitude. Their ratio is written:

$$\frac{Y_i}{X_i} = \frac{\frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 \beta_{p,i}}{R_{mean,i}^2} e^{-L_{m,i}} \left(\frac{1-e^{-2L_{p,i}}}{2\alpha_{p,i}}\right)}{\frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 \beta_{m,i}}{R_{mean,i}^2} e^{-L_{m,i}} \left(\frac{1-e^{-2L_{p,i}}}{2\alpha_{p,i}}\right)} \qquad \text{Eq. 6.52}$$

$$\frac{Y_i}{X_i} = \frac{\beta_{p,i}}{\beta_{m,i}}$$

The ratio of crosstalk-corrected particulate signal to molecular signal simply leads to the ratio of the particle and molecular backscatter coefficients. The introduction of the synthetic molecular backscatter coefficient naturally gives an estimation of the particle backscatter coefficient:

$$\hat{\beta}_{p,i} = \frac{Y_i}{X_i} \times \beta_{m,i,sim} \qquad \text{Eq. 6.53}$$

This retrieval involves no recurrence and its accuracy is depending on the sole values of the range bin *i.* In addition, no normalisation is required and therefore, no bin data are spoiled by a hypothesis nor by a hypothesis enforcement. This property permits its use as a witness of the presence of particles in the first matching bin.

Besides, in the case of the backscatter coefficient, there is no need for a continue validity of bin data. If a bin is invalid, computation can go on if next one is valid. This is not the case for the retrieval of extinction, or at least not without arbitrarily considering it clear and thus adding oscillations and losing accuracy.

6.2.2.3 Example

The algorithm above was applied to a simple, horizontally homogeneous E2S scenario. The scenario is characterised by a standard atmosphere, a cloud between 5 and 7km, and the medium RMA aerosol profile (see Figure 6-2). The noise option was turned off.

The results of the SCA are shown in Figure 6-3. The backscatter and extinction profiles are both good. This is particularly true in the boundary layer. In the cloud, the retrieval is underestimating the backscatter, and on the contrary is overestimating the extinction in the fully loaded bin, but is underestimating it in both partially filled bins. However, a careful examination of X_i versus its expansion $T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 \beta_{m,i} R_{mean,i}^{-2} e^{-L_{m,i}} H(2L_{p,i}) \Delta R_i$ (not shown here) reveals that the reason for this does not reside in the algorithm itself but in the cross-talk correction. In the cloud, X_i is slightly overestimated. The reason is unknown at present. It could be the effect of small errors in the calibration constants, or in the E2S, in addition to the approximation of uniform values over a range bin.

Figure 6-4 shows the retrieval of the backscatter and the extinction coefficients for 100 observations simulated with the E2S from the same atmospheric scene as previously but this time the noise flag turned on. It can be seen the level of uncertainty is growing as the retrieval goes lower towards the ground. The reason for this is explained in the next section devoted to error propagation.

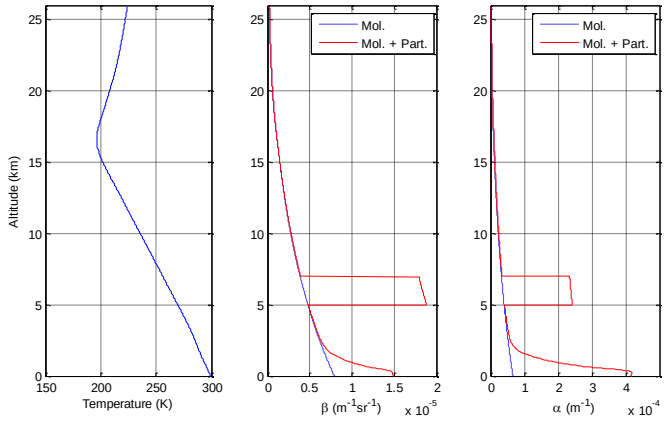

Figure 6-2: Temperature (left), backscatter (middle) and extinction (right) profiles of the atmosphere simulated for SCA tests.

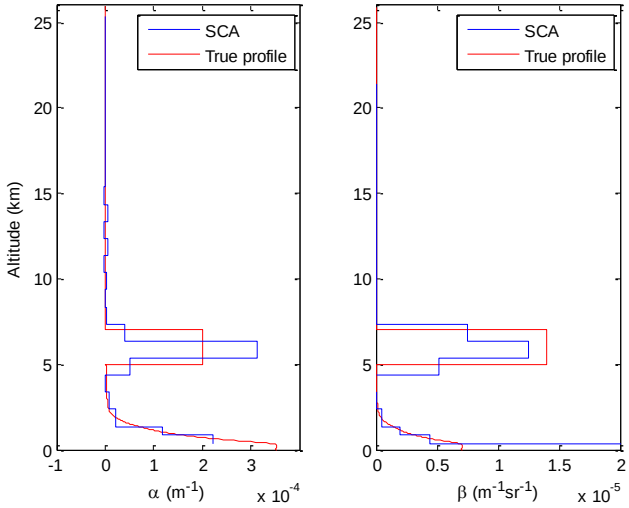

Figure 6-3: Extinction (left) and backscatter (right) profiles retrieved by the SCA on the scenario shown in Figure 6-2.

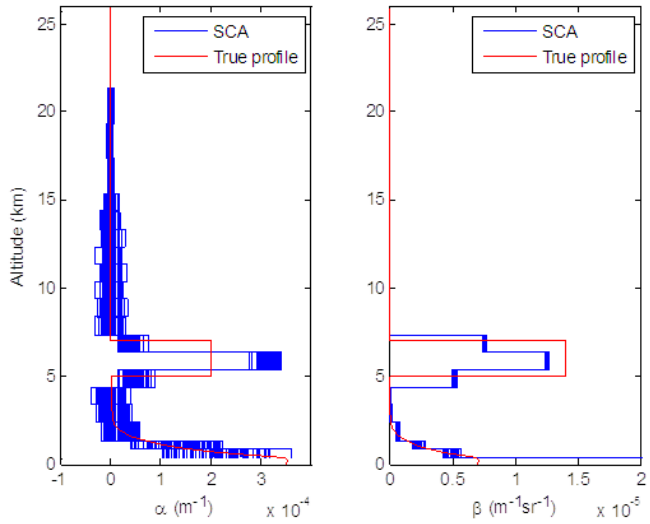

Figure 6-4: Same as Figure 6-3, but for 100 observations simulated with the E2S from the atmospheric scene in Figure 6-2 but with the noise flag turned on.

6.2.3 SCA approximations and error propagation

6.2.3.1 Error propagation

In practice, the useful signals registered on both Mie and Rayleigh channels contain uncertainties (due to the photocounting process for the major part). Let us denote the relative errors by $e_{ray,i} = \delta S_{ray,i}/S_{ray,i}$ and $e_{mie,i} = \delta S_{mie,i}/S_{mie,i}$. By definition, we have

$$\langle e_{ray,i}^2 \rangle = SNR_{ray,i}^{-2} \quad \text{Eq. 6.54}$$

$$\langle e_{mie,i}^2 \rangle = SNR_{mie,i}^{-2} \quad \text{Eq. 6.55}$$

Before going forth and for the sake of clarity, it is needed to rewrite the expressions of X_i and Y_i as a function of useful signals $S_{ray,i}$ and $S_{mie,i}$:

$$X_i = \chi_{3,i} S_{Ray,i} - \chi_{2,i} S_{mie,i} \quad \text{Eq. 6.56}$$

$$Y_i = -\chi_{4,i} S_{Ray,i} + \chi_{1,i} S_{mie,i} \quad \text{Eq. 6.57}$$

Where

$$\chi_{1,i} = \frac{1}{E_0 N_p K_{mie}} \frac{C_{1,i}}{C_{1,i}C_{3,i} - C_{2,i}C_{4,i}} \quad \text{Eq. 6.58}$$

$$\chi_{2,i} = \frac{1}{E_0 N_p K_{mie}} \frac{C_{2,i}}{C_{1,i}C_{3,i} - C_{2,i}C_{4,i}} \quad \text{Eq. 6.59}$$

$$\chi_{3,i} = \frac{1}{E_0 N_p K_{ray}} \frac{C_{3,i}}{C_{1,i}C_{3,i} - C_{2,i}C_{4,i}} \quad \text{Eq. 6.60}$$

$$\chi_{4,i} = \frac{1}{E_0 N_p K_{ray}} \frac{C_{4,i}}{C_{1,i}C_{3,i} - C_{2,i}C_{4,i}} \quad \text{Eq. 6.61}$$

Note that the calibration is supposed to be good enough to generate only negligible errors.

Then, the relative errors e_{X_i} and e_{Y_i} are:

$$e_{X_i} = \frac{\delta X_i}{X_i} = \frac{\chi_{3,i} e_{ray,i} S_{ray,i} - \chi_{2,i} e_{mie,i} S_{mie,i}}{\chi_{3,i} S_{ray,i} - \chi_{2,i} S_{mie,i}} \quad \text{Eq. 6.62}$$

$$e_{Y_i} = \frac{\delta Y_i}{Y_i} = \frac{-\chi_{4,i} e_{ray,i} S_{ray,i} + \chi_{1,i} e_{mie,i} S_{mie,i}}{-\chi_{4,i} S_{ray,i} + \chi_{1,i} S_{mie,i}} \quad \text{Eq. 6.63}$$

And the correlations:

$$\langle e_{X_i}^2 \rangle = \frac{\chi_{3,i}^2 SNR_{ray,i}^{-2} S_{ray,i}^2 + \chi_{2,i}^2 SNR_{mie,i}^{-2} S_{mie,i}^2}{{X_i}^2} \quad \text{Eq. 6.64}$$

$$\langle e_{Y_i}^2 \rangle = \frac{\chi_{4,i}^2 SNR_{ray}^{-2} S_{ray}^2 + \chi_1^2 SNR_{mie}^{-2} S_{mie,i}^2}{{Y_i}^2} \quad \text{Eq. 6.65}$$

$$\langle e_{X_i} e_{Y_i} \rangle = -\frac{\chi_{3,i}\chi_{4,i} SNR_{ray,i}^{-2} S_{ray,i}^2 + \chi_{2,i}\chi_{1,i} SNR_{mie,i}^{-2} S_{mie,i}^2}{X_i Y_i} \quad \text{Eq. 6.66}$$

6.2.3.2 Estimation of error in the Backscatter coefficient

Let us denote $\delta\beta_{p,i}$ the error on the backscatter coefficient retrieved by the SCA:

$$\begin{aligned} \hat{\beta}_{p,i} &= \beta_{p,i} + \delta\beta_{p,i} \quad \text{Eq. 6.67} \\ &= \frac{\hat{Y}_i}{\hat{X}_i}\beta_{m,i,sim} \\ &= \frac{Y_i + \delta Y_i}{X_i + \delta X_i}\beta_{m,i,sim} \\ &= \frac{Y_i}{X_i}\beta_{m,i,sim}\frac{1 + e_{Y_i}}{1 + e_{X_i}} \\ &= \beta_{p,i}\frac{1 + e_{Y_i}}{1 + e_{X_i}} \\ \hat{\beta}_{p,i} &\approx \beta_{p,i}\left(1 + e_{Y_i} - e_{X_i}\right) \quad \text{Eq. 6.68} \end{aligned}$$

Hence:

$$\delta\beta_{p,i} \approx \beta_{p,i}\left(e_{Y_i} - e_{X_i}\right) \quad \text{Eq. 6.69}$$

Its autocorrelation is:

$$\langle \left(\delta\beta_{p,i}\right)^2 \rangle = \beta_{p,i}^2 \left[\langle e_{X_i}^2 \rangle + \langle e_{Y_i}^2 \rangle - 2\langle e_{X_i} e_{Y_i} \rangle \right] \quad \text{Eq. 6.70}$$

which is reported in the L2a product according to [RD 28], chapter 3.5.3 (backscatter_variance expressed in $m^{-2}sr^{-2}$.).

6.2.3.3 Estimation of the error in the Extinction coefficient

Let ϵ_i be the relative error on the $NITWT_i'$:

$$NITWT_i'(1 + \epsilon_i) = \frac{\hat{X}_{i,obs}}{\hat{X}_{1,obs}}\frac{X_{1,sim}}{X_{i,sim}} \quad \text{Eq. 6.71}$$

Eq. 6.72

Hence:

$$\epsilon_i = \frac{1 + e_{X_i}}{1 + e_{X_1}} - 1 \approx e_{X_i} - e_{X_1} \quad \text{Eq. 6.73}$$

And its autocorrelation:

$$\langle \epsilon_i^2 \rangle \approx \langle e_{X_i}^2 \rangle + \langle e_{X_1}^2 \rangle \quad \text{Eq. 6.74}$$

Making the approximation in Eq. 6.44:

$$H(2x) \approx e^{-x} \quad \text{Eq. 6.75}$$

The $NITWT_i'$ becomes, without assumption on the filling of the first matching bin:

$$NITWT_i' = T_{p,sat,i-1}^2 \frac{e^{-L_{p,i}}}{e^{-L_{p,1}}} \quad \text{Eq. 6.76}$$

And:

$$\hat{L}_{p,i} = -\ln\left(\frac{NITWT_i'(1+\epsilon_i)}{\hat{T}_{p,sat,i-1}^2}\right) \quad \text{Eq. 6.77}$$

$$= -\ln\left(\frac{NITWT_i'}{T_{p,sat,i-1}^2}\right) - \ln(1+\epsilon_i) - 2\sum_{k=1}^{i-1}\left(\hat{L}_{p,k} - L_{p,k}\right) - L_{p,1}$$

$$\approx L_{p,i} - \ln(1+\epsilon_i) - 2\sum_{k=1}^{i-1}\left(\hat{L}_{p,k} - L_{p,k}\right) - L_{p,1} \quad \text{Eq. 6.78}$$

Hence the error $\delta L_{p,i}$ on the optical depth:

$$\begin{aligned}
\delta L_{p,1} &= -ln(1+\epsilon_1) - L_{p,1} \\
\delta L_{p,2} &= -ln(1+\epsilon_2) - 2\delta L_{p,1} - L_{p,1} \\
&= -ln(1+\epsilon_2) + 2\,ln(1+\epsilon_1) + L_{p,1} \\
&\vdots \\
\delta L_{p,i} &= -ln(1+\epsilon_i) + 2\sum_{k=1}^{i-1}(-1)^{i-k}\,ln(1+\epsilon_k) + (-1)^i L_{p,1}
\end{aligned} \quad \text{Eq. 6.79}$$

A first order expansion of the logarithm leads to:

$$\delta L_{p,i} \approx -\epsilon_i + 2\sum_{k=1}^{i-1}(-1)^{i-k}\epsilon_k + (-1)^i L_{p,1} \quad \text{Eq. 6.80}$$

$$\epsilon_i \approx e_{X_i} - e_{X_1} \Rightarrow \quad \delta L_{p,i} \approx -e_{X_i} + 2\sum_{k=2}^{i-1}(-1)^{i-k}e_{X_k} - (-1)^i e_{X_1} + (-1)^i L_{p,1} \quad \text{Eq. 6.81}$$

The terms $(-1)^i$ explain the oscillating behaviour of the SCA from a bin to another.

Then, taking into account that $\forall i \in [\![1;24]\!], \langle e_i \rangle = 0$ and $\forall (i,j) \in [\![1;24]\!]^2, i \neq j, \langle e_i e_j \rangle = 0$, its variance is:

$$\forall i \in [\![2;24]\!], \sigma^2_{L_{p,i}} = \langle \left(\delta L_{p,i}\right)^2 \rangle - \langle \delta L_{p,i} \rangle^2 \approx 4 \sum_{k=1}^{i} \langle e^2_{X_k} \rangle - 3\langle e^2_{X_i} \rangle - 3\langle e^2_{X_1} \rangle \qquad \text{Eq. 6.82}$$

which is reported in the L2a product according to [RD 28], chapter 3.5.3 (extinction_variance expressed in m^{-2}.). In addition to this variance, the bias between the estimation and the actual value is $(-1)^i L_{p,1}$. This result illustrates the oscillating behaviour amplified by the potential error due to the hypothesis on the first matching bin.

6.2.3.4 Scattering ratio

The scattering ratio ρ_i in the i^{th} bin is defined as follows:

$$\rho_i = 1 + \frac{\beta_{p,i}}{\beta_{m,i}}$$

Then the relative error $e_{\rho,i}$ is

$$e_{\rho,i} = \frac{e_{Y_i} - e_{X_i}}{1 + \frac{X_i}{Y_i}}$$

And the variance is

$$\langle e^2_{\rho,i} \rangle = \frac{\langle e^2_{X_i} \rangle + \langle e^2_{Y_i} \rangle - 2\langle e_{X_i} e_{Y_i} \rangle}{\left(1 + \frac{X_i}{Y_i}\right)}$$

6.2.3.5 Example

Previous error equations have been applied to the same E2S data as in section 6.2.2.3 (with the noise option turned on). The signal-to-noise ratios were taken equal to the square root of the useful signal:

$$SNR_{ray,i} = \frac{S_{ray,i}}{\sqrt{S_{ray,i}}} = \sqrt{S_{ray,i}} \qquad \text{Eq. 6.83}$$

$$SNR_{mie,i} = \frac{S_{mie,i}}{\sqrt{S_{mie,i}}} = \sqrt{S_{mie,i}} \qquad \text{Eq. 6.84}$$

This is an approximation as the useful signals are retrieved from the detector photocounts (subject to photocounting noise, following a Poisson distribution) by subtracting an estimate of a background noise level. In practice, one must thus expect that the SNR is somewhat higher than the square-root of the useful signal.

The error levels on the backscatter coefficients are displayed in Figure 6-5. The figure is made of two graphs. They are showing the same data but the topmost graph has a linear y-scale while the bottom one has a logarithmic y-scale. On both, the “real” backscatter profile (input to the E2S) is shown with a green curve. The blue stairs are giving the 100 profiles obtained by the SCA from the 100 BRCs of the scenario. The two red stairs are obtained by $\langle \beta_i \rangle \pm \langle \sigma_{\beta_i} \rangle$ where $\langle\ \ \rangle$ is the averaging operator, and σ_{β_i} is

the error level predicted by the equation. One can see that the predicted error level is in good agreement with the error level revealed by the 100 BRCs.

Figure 6-6 does the same for the extinction. There again, it appears that the equation produces a good estimate of the error level.

NB1: as regards extinction error estimation, the combination of the approximation along the various steps of the algorithm and of SCA consolidation (negative values forbidden) generates a constant divergence between estimated values and simulated values. After many simulations, the ratio has been assessed to 3. Therefore, the estimated standard deviation plotted on Figure 6-6 is the third of the result Eq. 6.82.

NB2: In this calculation, calibration coefficient accuracy is assumed perfect. The study of the propagation of errors coming from these values will slightly differ from the latter. Indeed, they are defined by integrals and the integrated functions T_A, T_B and T_{Fiz} (see [RD 25]) are assessed through the same procedure, hence the errors are not independent. This study will be carried later on.

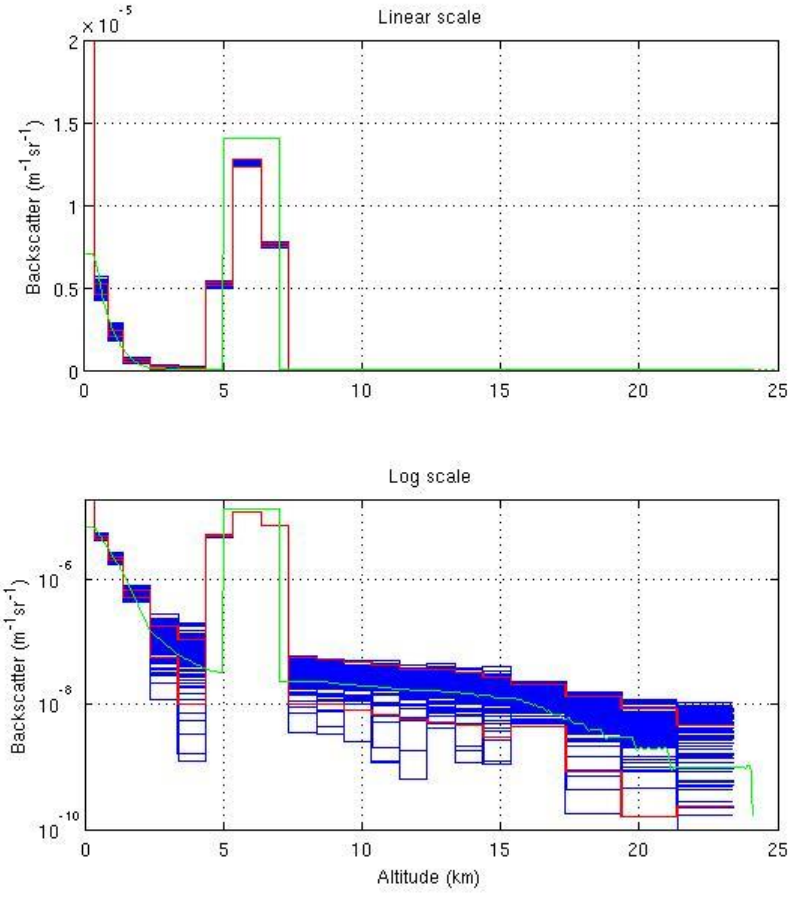

Figure 6-5: Backscatter coefficients retrieved by the SCA on a bunch of 100 BRCs simulated with the E2S on the basis of a single set of atmospheric profiles depicted in Figure 6-2. The bottom graph is identical to the top one except for the y-scale (log instead of linear). On both graphs, the green curve shows the E2S input backscatter profile. The blue stairs show the 100 backscatter profiles retrieved by the SCA on the 100 BRCs of the scenario. The two red stairs are given by the average of the SCA backscatter profile plus and minus the average error level predicted by the equations in section 6.2.3.1.

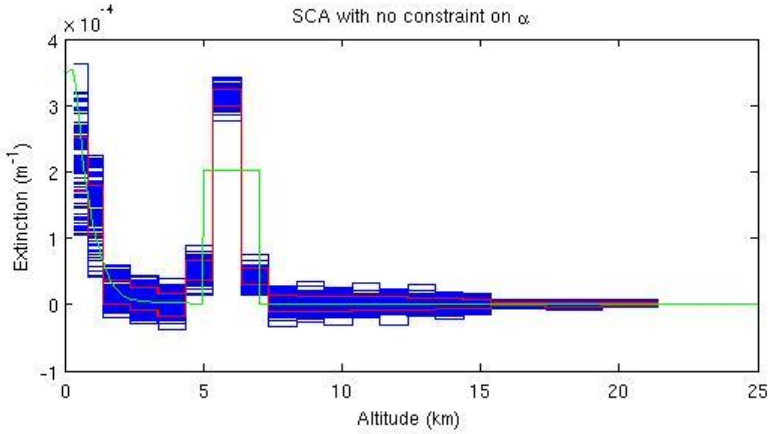

Figure 6-6: Same as the top graph in Figure 6-5 for the extinction profiles retrieved by the SCA.

6.2.3.6 Data Quality Flag

In order to provide users with information on the validity of the SCA retrievals, a data quality flag is provided for SCA products on Rayleigh bins and middle Rayleigh bins (see section 0). This quality check is also applied on group products. For each Rayleigh bin, this flag is made of 7 bits which give the validity of the following parameters in this order:

- Validity of the extinction coefficient retrieval;
- Validity of the backscatter coefficient retrieval;
- Mie Signal-to-Noise Ratio;
- Rayleigh Signal-to-Noise Ratio;
- The ratio of the expected error of the extinction retrieval (see section 6.2.3.3) to the value of the retrieved coefficient (relative error);
- The ratio of the expected error of the backscatter retrieval (see section 6.2.3.2) to the value of the retrieved coefficient (relative error);
- Attenuation of the signal;

For products in bins made from two halves of adjacent original Rayleigh bins (middle Rayleigh bins), this flag is made of 8 bits which give the validity of the following parameters in this order:

- Validity of the extinction coefficient retrieval;
- Validity of the backscatter coefficient retrieval;
- Validity of the BER retrieval;
- Mie Signal-to-Noise Ratio;
- Rayleigh Signal-to-Noise Ratio;
- The ratio of the expected error of the extinction retrieval (see section 6.2.3.3) to the value of the retrieved coefficient (relative error);
- The ratio of the expected error of the backscatter retrieval (see section 6.2.3.2) to the value of the retrieved coefficient (relative error);
- Attenuation of the signal.

The validity of the backscatter and extinction coefficient retrievals (for both Rayleigh bins and middle Rayleigh bins) depends on:

- The ratio of the expected error (see sections 6.2.3.2 and 6.2.3.3) to the value of the retrieved coefficient (relative error);
- Rayleigh SNR for the extinction coefficient and Mie SNR for the backscatter coefficient.

The signal is considered as valid in a specific bin if it has not been too much attenuated along the path. The attenuation is inferred from the cumulated local optical depth retrieved from the extinction coefficient.

The BER is considered valid if the retrieved value is between lower and upper limits defined in the AUX_PAR_2A file. For all parameters used in the data quality flag, some thresholds have been defined to consider a product as valid or not. These thresholds are defined in the AUX_PAR_2A file and have to be refined after launch.

6.3 Improvement of SCA products: the BER

6.3.1 Rationale

Particles will eventually be classified with respect to their BER. The retrieved backscatter coefficient is rather reliable, but because of oscillating error propagation through the SCA, it is not the case for the extinction coefficient. This leads to an unreliable BER.

The stabilisation of the 2-bin periodic oscillation of α_p, or L_p, can be carried out through an averaging over 2 sequent bins. It will decrease the resolution but will significantly increase the precision.

For instance, let us consider two adjacent bins i and $i+1$. The SCA-computed SLODs are affected by the respective errors $\delta L_{p,i}$ and $\delta L_{p,i+1}$. If the SCA SLODs L_p in the two adjacent bins are averaged, so are their errors:

$$\begin{aligned}\frac{1}{2}\left(\delta L_{p,i}+\delta L_{p,i+1}\right) &= \frac{1}{2}\left(-e_{X_i}+2\sum_{k=2}^{i-1}(-1)^{i-k}e_{X_k}-(-1)^i e_{X_1}+(-1)^i L_{p,1}\right)\\ &\quad+\frac{1}{2}\left(-e_{X_{i+1}}+2\sum_{k=2}^{i}(-1)^{i+1-k}e_{X_k}-(-1)^{i+1}e_{X_1}+(-1)^{i+1}L_{p,1}\right)\\ \frac{1}{2}\left(\delta L_{p,i}+\delta L_{p,i+1}\right) &= -\frac{1}{2}\left(e_{X_{i+1}}+3e_{X_i}\right)\end{aligned} \qquad \text{Eq. 6.85}$$

One can see from this equation that oscillating terms compensate to the benefit of a reduction of the error level. One can also note that the inherent bias due to the assumption of a clear first matching bin is also suppressed. The corresponding variance σ^2 is:

$$\sigma^2=\frac{1}{4}\langle e_{X_{i+1}}^2\rangle+\frac{9}{4}\langle e_{X_i}^2\rangle \qquad \text{Eq. 6.86}$$

To be compared to:

$$\sigma^2_{L_{p,i}} = 4\sum_{k=1}^{i}\langle e^2_{X_k}\rangle - 3\langle e^2_{X_i}\rangle - 3\langle e^2_{X_1}\rangle \qquad \text{(Eq. 6.82)}$$

One can see the drastic decrease of error.

6.3.2 Applied to ALADIN: towards a stabilised BER

In practice, the resolution is loosed because the averages are strictly valid over two sequent bins, but each averaged value will have a sphere of predominance smaller than the two bins, as shown in Figure 6-7. ALADIN's height bins have variable width and the profiles are limited. It will not affect the rationale – the SLOD of two adjacent bins are added – but the new bin corresponding to the sphere of predominance may vary in shape. Out of the borders, it is composed of two halves of the original lidar bins and those halves may have different size, leading to an intermediate width between the original ones. On profile borders, the new bin is composed of one half and for the full border bin: the resolution of average values is not as much artificially increased on borders than elsewhere.

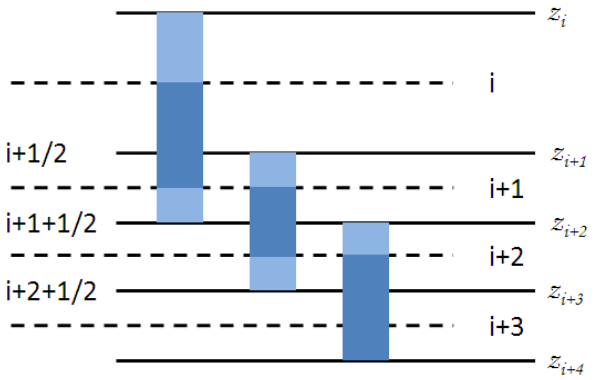

Figure 6-7: Height bins and averaged values. 3 cases are illustrated here: width variation in (i,i+1), constant width in (i+1,i+2) and end of profile in (i+2,i+3). The strict area of validity of the averages is in light blue and the sphere of predominance is in darker blue. Integers are the indices of the original lidar bins, half integers are the indices of the bins of predominance of averaged values.

The following equations formalise the approach for an average between bins i and $i+1$, with respective widths ΔR_i and ΔR_{i+1}. Intermediate bin related values are indexed $_{i+\frac{1}{2}}$. Then, the BER $k_{p,i+\frac{1}{2}}$ in $i+\frac{1}{2}$ is:

$$k_{p,i+\frac{1}{2}} = \frac{\beta_{p,i+\frac{1}{2}}}{\alpha_{p,i+\frac{1}{2}}} \qquad \text{Eq. 6.87}$$

The relative error $e_{k_{p,i+\frac{1}{2}}}$ is:

$$e_{k_{p,i+\frac{1}{2}}} = \frac{\delta k_{p,i+\frac{1}{2}}}{k_{p,i+\frac{1}{2}}} = e_{\beta_{p,i+\frac{1}{2}}} - e_{\alpha_{p,i+\frac{1}{2}}} \qquad \text{Eq. 6.88}$$

And the variance of the relative error is:

$$\sigma^2_{k_{p,i+\frac{1}{2}}} = \langle e^2_{\alpha_{p,i+\frac{1}{2}}} \rangle + \langle e^2_{\beta_{p,i+\frac{1}{2}}} \rangle - 2\langle e_{\alpha_{p,i+\frac{1}{2}}} e_{\beta_{p,i+\frac{1}{2}}} \rangle \quad \text{Eq. 6.89}$$

The following equations detail the steps to get the three components of Eq. 6.89. They are expressed in Eq. 6.95, Eq. 6.98, Eq. 6.100 respectively.

$$L_{p,i+\frac{1}{2}} = \frac{1}{2}\left(L_{p,i} + L_{p,i+1}\right) \quad \text{Eq. 6.90}$$

Then, the error $\delta L_{p,i+\frac{1}{2}}$ is:

$$\delta L_{p,i+\frac{1}{2}} = \frac{1}{2}\left(\delta L_{p,i} + \delta L_{p,i+1}\right) \quad \text{Eq. 6.91}$$

$$\delta\alpha_{p,i+\frac{1}{2}} = \frac{\delta L_{p,i+\frac{1}{2}}}{\Delta R_{i+\frac{1}{2}}} \quad \text{Eq. 6.92}$$

With $\Delta R_{i+\frac{1}{2}} = \frac{1}{2}(\Delta R_i + \Delta R_{i+1})$. Then the relative error on $\alpha_{p,i+\frac{1}{2}}$ is:

$$e_{\alpha_{p,i+\frac{1}{2}}} = \frac{\delta\alpha_{p,i+\frac{1}{2}}}{\alpha_{p,i+\frac{1}{2}}} = \frac{\delta L_{p,i} + \delta L_{p,i+1}}{2L_{p,i+\frac{1}{2}}} \quad \text{Eq. 6.93}$$

The injection of the value of $\delta L_{p,i}$ from Eq. 6.85 leads to:

$$e_{\alpha_{p,i+\frac{1}{2}}} = -\frac{1}{2L_{p,i+\frac{1}{2}}}\left(e_{X_{i+1}} + 3e_{X_i}\right) \quad \text{Eq. 6.94}$$

The variance of the relative error is then:

$$\langle e^2_{\alpha_{p,i+\frac{1}{2}}} \rangle = \frac{1}{\left(2L_{p,i+\frac{1}{2}}\right)^2}\left(\langle e^2_{X_{i+1}} \rangle + 9\langle e^2_{X_i} \rangle\right) \quad \text{Eq. 6.95}$$

Likewise, for $\beta_{p,i+\frac{1}{2}}$

$$\begin{aligned} \delta\beta_{p,i+\frac{1}{2}} &= \frac{\Delta R_i \times \delta\beta_{p,i} + \Delta R_{i+1} \times \delta\beta_{p,i+1}}{\Delta R_i + \Delta R_{i+1}} \\ &= \frac{\Delta R_i \beta_{p,i}\left(e_{X_i} - e_{Y_i}\right) + \Delta R_{i+1}\beta_{p,i+1}\left(e_{X_{i+1}} - e_{Y_{i+1}}\right)}{\Delta R_i + \Delta R_{i+1}} \end{aligned} \quad \text{Eq. 6.96}$$

$$e_{\beta_{p,i+\frac{1}{2}}} = \frac{\delta\beta_{p,i+\frac{1}{2}}}{\beta_{p,i+\frac{1}{2}}} = \frac{\Delta R_i\beta_{p,i}\left(e_{X_i} - e_{Y_i}\right) + \Delta R_{i+1}\beta_{p,i+1}\left(e_{X_{i+1}} - e_{Y_{i+1}}\right)}{\Delta R_i\beta_{p,i} + \Delta R_{i+1}\beta_{p,i+1}} \quad \text{Eq. 6.97}$$

With $B_i = \Delta R_i\beta_{p,i}$, the variance of the relative error on $\beta_{p,i+\frac{1}{2}}$ is expressed as follows:

$$\langle e^2_{\beta_{p,i+\frac{1}{2}}}\rangle = \frac{(\Delta R_i)^2\left\langle\left(\delta\beta_{p,i}\right)^2\right\rangle + (\Delta R_{i+1})^2\left\langle\left(\delta\beta_{p,i+1}\right)^2\right\rangle}{(B_i + B_{i+1})^2} \quad \text{Eq. 6.98}$$

The cross-correlation $\langle e_{\alpha_{p,i+\frac{1}{2}}} e_{\beta_{p,i+\frac{1}{2}}}\rangle$ is a bit heavier to get. Let's set $D_i = (B_i + B_{i+1}) \times 2L_{p,i+\frac{1}{2}}$. The expressions of Eq. 6.94 and Eq. 6.97 are multiplied together and the time average is taken. After having removed the terms that will lead to a null correlation, we get:

$$\langle e_{\alpha_{p,i+\frac{1}{2}}} e_{\beta_{p,i+\frac{1}{2}}}\rangle = -\frac{1}{D_i}\langle\left(B_i\left(e_{X_i} - e_{Y_i}\right) + B_{i+1}\left(e_{X_{i+1}} - e_{Y_{i+1}}\right)\right)\left(e_{X_{i+1}} + 3e_{X_i}\right)\rangle \quad \text{Eq. 6.99}$$

We get to:

$$\langle e_{\alpha_{p,i+\frac{1}{2}}} e_{\beta_{p,i+\frac{1}{2}}}\rangle = -\frac{1}{D_i}\left(B_{i+1}\left(\langle e^2_{X_{i+1}}\rangle - 3\langle e_{X_{i+1}} e_{Y_{i+1}}\rangle\right) + B_i\left(\langle e^2_{X_i}\rangle - 3\langle e_{X_i} e_{Y_i}\rangle\right)\right) \quad \text{Eq. 6.100}$$

6.4 Improvement of SCA products: Heterogeneity indexes

Similar SCA retrievals at the BRC levels may hide two different realities. One bin may be homogeneously populated by the same particles while the other may be only partially populated , e.g. a mix of clear sky in the first half of the BRC, and a cloud in the second half. The interpretation of the retrieved optical properties requires some information about the homogeneity of the measurement accumulated within the BRC. To provide more information on the sub-BRC level, we propose to implement a heterogeneity index that characterizes the variability within one BRC.

This heterogeneity index is implemented for both the Rayleigh and Mie channel. It is defined as the standard deviation of the useful signal with respect to the Poisson noise in order to get 1 if only the Poisson noise contribute to the variability of the scene:

$$heterogeneity\_index_{ray} = \frac{1}{\sqrt{k_{ray}}} \times \frac{\sigma\left(S_{ray}\right)}{\sqrt{\langle S_{ray}\rangle}} \quad \text{Eq. 6.101}$$

$$heterogeneity\_index_{mie} = \frac{1}{\sqrt{k_{mie}}} \times \frac{\sigma(S_{mie})}{\sqrt{\langle S_{mie}\rangle}} \quad \text{Eq. 6.102}$$

while, $\sigma(\ldots)$ and $\langle \ldots \rangle$ stand for the horizontal standard deviation and the mean over 30 measurements included in one observation.

In practice the value of these heterogeneous indexes are rarely equal to 1. Figure 6-8 and Figure 6-9, indicates that homogeneous bins yield a *heterogeneity_index* ranging between 3 and 5 respectively for the Rayleigh and Mie channel.

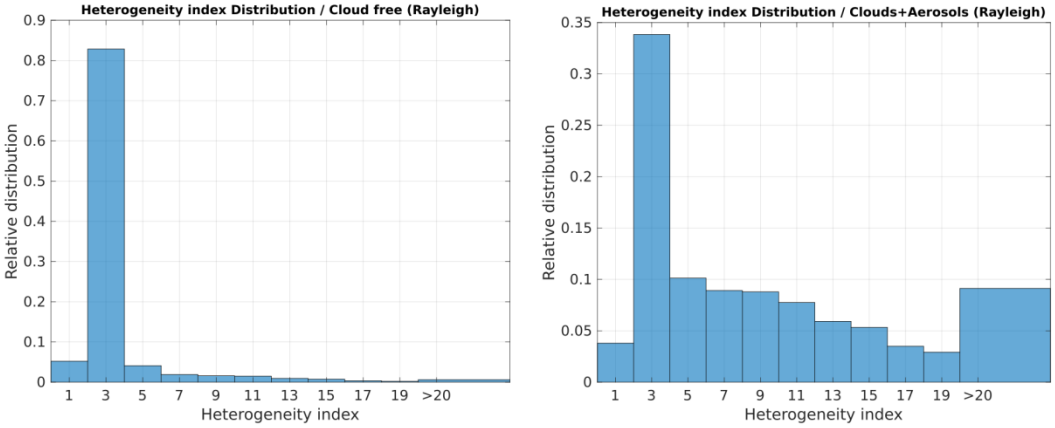

Figure 6-8: Distribution of the heterogeneity index computed on one orbit file and separated by cloud free (left) and cloudy (right) areas. Bins with SNR_{ray} <50 are rejected.

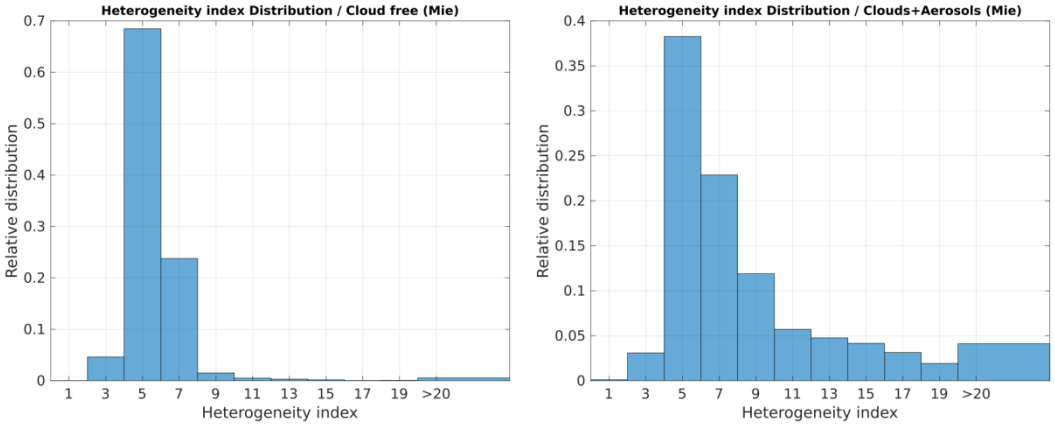

Figure 6-9: Same as Figure 6-8. Bins with SNR_{mie}<35 are rejected.

Figure 6-10 shows an example of the heterogeneity index applied on the Rayleigh channel for a simple scene. This example suggests that in cloudy areas, the bins can be considered as relatively homogeneously populated while the heterogeneously

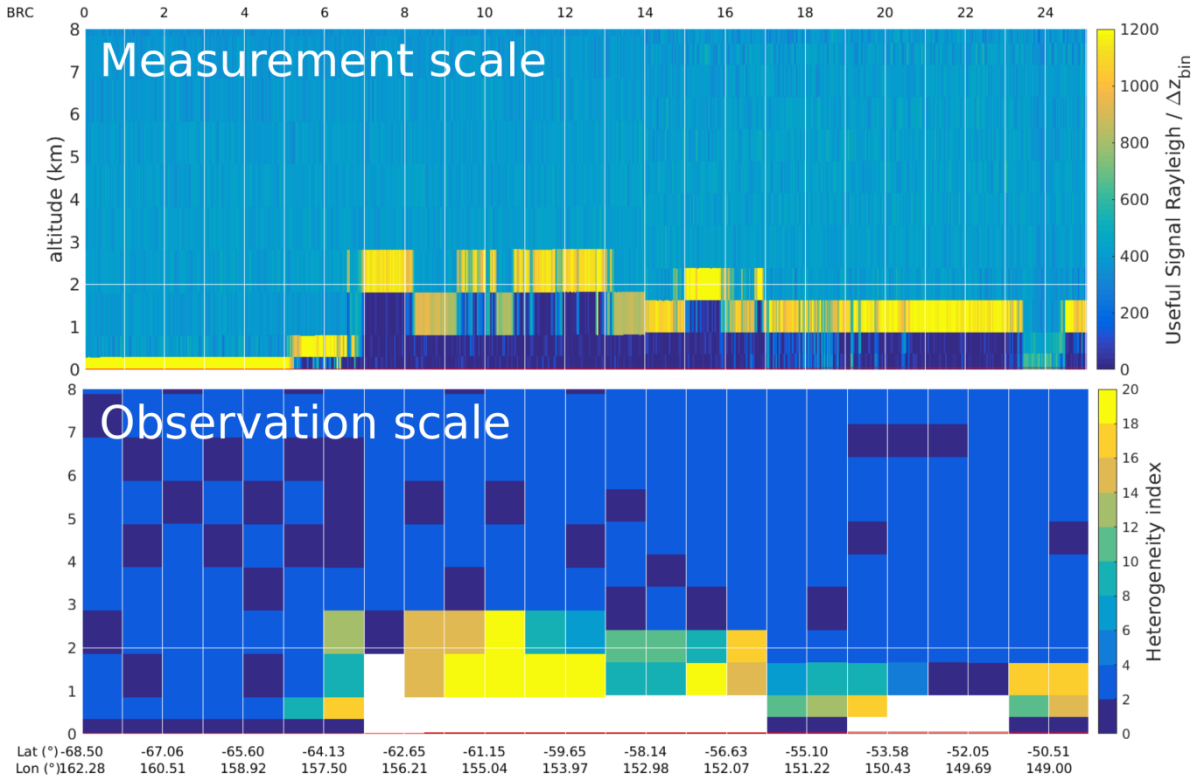

Figure 6-10: Comparison between the Rayleigh useful signal (top) and the associated heterogeneity index (bottom) on a scene.

6.5 *Improvement of SCA products: estimation of calibration coefficients K_{ray} and K_{mie} and radiometric correction refinement on an empirical basis*

6.5.1 Initial scheme: estimation of K_{ray} and K_{mie} from IRC mode

Initially the calibration coefficients for the useful signal in the Rayleigh and Mie channels, respectively K_{ray} and K_{mie} were estimated from auxiliary calibration data produced by the Instrument Response Calibration (IRC) mode for which the laser is pointing at NADIR (i.e. no Doppler shift, assuming the vertical wind averaged over the BRC is close to 0). This NADIR pointing mode is preferably conducted over land with high UV albedo and low cloud coverage (e.g. Antarctica). Geometric configuration corresponding to this NADIR pointing is showed with looking angle illustration in Figure 6-11.

From previous looking angle representation showed in Figure 4-14, the viewing angle with IRC mode and NADIR looking is $\theta_{(z)} = \theta_{(0)} \approx 0°$. Therefore the satellite-to-the-surface range is $R_0 = h_0/\cos(0) = h_0$ the satellite height is h_0=320 km. In the lower atmosphere the looking angle $\theta(z)$ is therefore in theory equal to zero degrees.

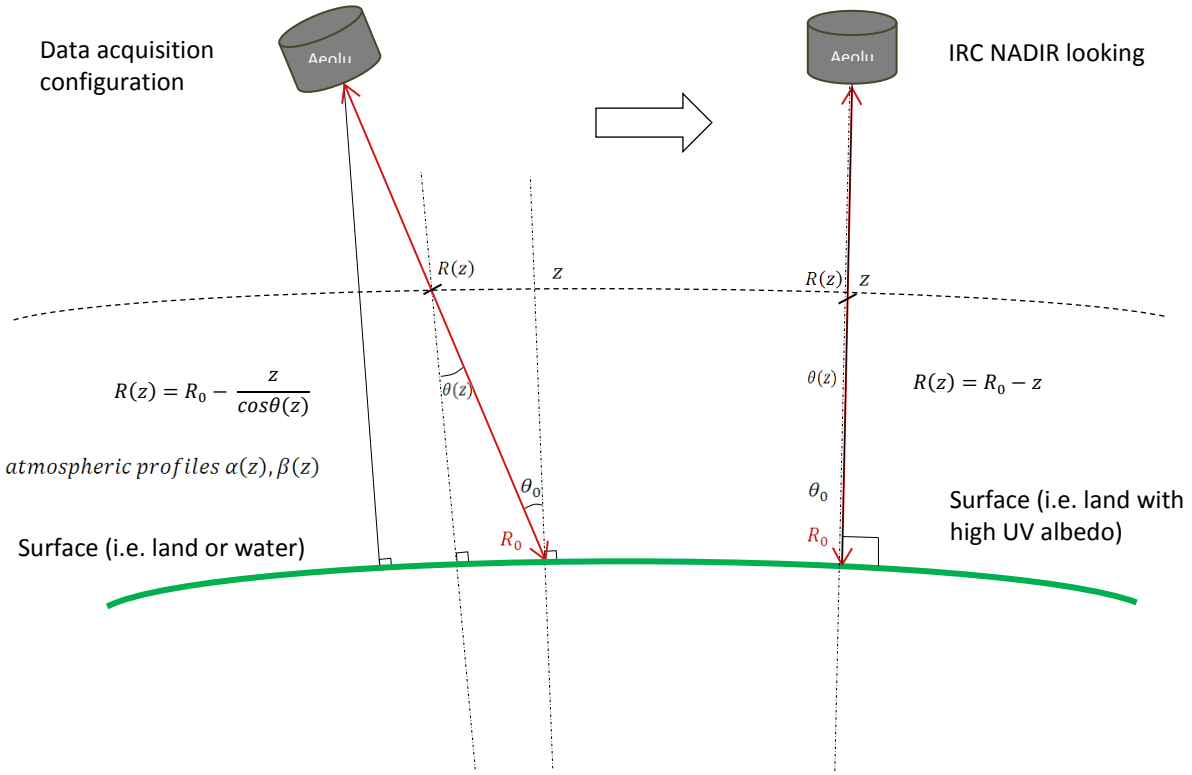

Figure 6-11: Schematic comparison of the viewing geometry of the laser in nominal Aeolus acquisition and in Instrument Response Calibration (IRC) mode (i.e. NADIR looking over land)

The coefficients K_{ray} and K_{mie} were considered as constant and values were estimated from IRC and taken to process the crosstalk correction (e.g. Eq. 6.5 and Eq. 6.6).

Unique default values for both K_{ray} and K_{mie} were therefore applied to all AEOLUS data until the next IRC (in principle once a week and the derivation of new K_{ray} and K_{mie} values). The comparison of the Mie and Rayleigh signals recorded during Wind Vector Mode operations (standard mode of operation of Aeolus with a 35 degrees off-nadir looking angle) in particle-free regions of the atmosphere (i.e. that is no particles in the region and above) with signal predictions from the known molecular backscatter and extinction showed, however, that the application of this calibration coefficients were overestimating the real signals in both Mie and Rayleigh channels. It thus appeared that a correction factor was needed.

This signal overestimation can be clearly observed when processing predicted signals. The relative errors (i.e.$(observed - predicted)\ /predicted)$) for Rayleigh and Mie useful signals have been calculated for selected bins in mid-altitudes (i.e. between 6km and 16km) in particle-free regions of the atmosphere. Example of relative errors on molecular signal (i.e. Rayleigh Useful Signal (RUS)) is showed in Figure 6-12. It can be seen that the distribution of relative errors appears biased with a largely negative median value, meaning that the calibration used for the signal prediction overestimated the molecular contribution to the signal (i.e.$(observed < predicted)$).

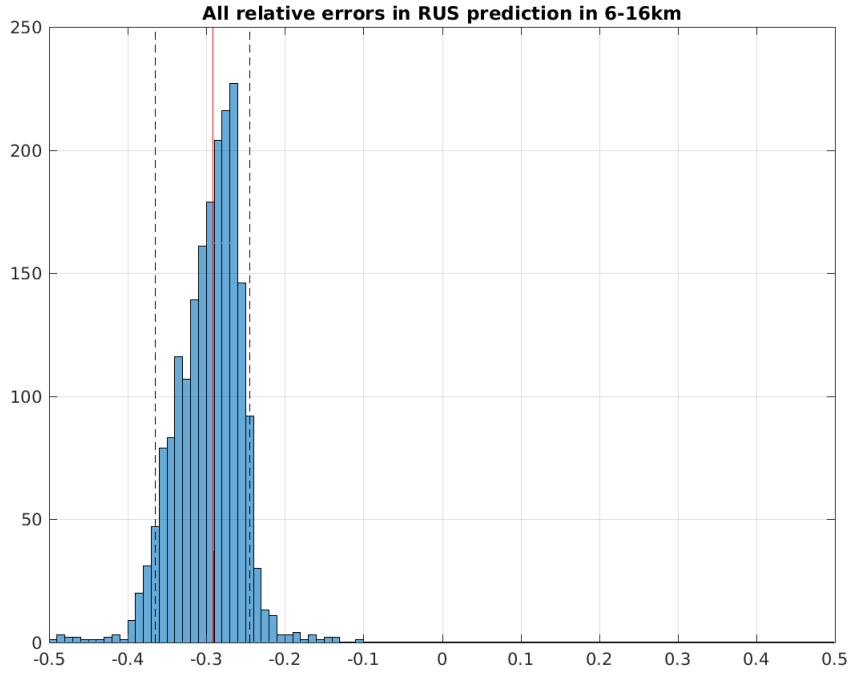

Figure 6-12: Relative errors ((observed-predicted) / predicted) for Rayleigh useful signal in mid-altitudes 6km to 16km in clear sky conditions

6.5.2 Orbit correction in particle-free condition

We observed that signal predicted using default K_{ray} and K_{mie} from the IRC mode are overestimated. A first tuning scheme of the radiometric correction has been implemented in L2A processor in order to reduce this signal overestimation. The idea is to perform signal prediction in particle-free conditions in

order to apply corrective factors for both K_{ray} and K_{mie}. We select particle-free bins in mid-altitudes 6 to 16km and estimate the relative errors for signal prediction of Rayleigh and Mie useful signal.

From Eq. 6.7 and Eq. 6.8 the accumulated lidar signals in vertical range bin for respectively Rayleigh and Mie channels S_{ray} and S_{mie} are given by:

$$s_{ray}(r) = K_{ray} N_p E_0 (C_1(P,T,f)X(r) + C_2(f)Y(r)) \quad \text{Eq. 6.103}$$

$$s_{mie}(r) = K_{mie} N_p E_0 (C_4(P,T,f)X(r) + C_3(f)Y(r)) \quad \text{Eq. 6.104}$$

In particle-free regions of the atmosphere β_{part} and ${T^2}_{part}$ are respectively equal to 0 and 1 and the range resolved atmospheric signals at telescope entry for particles $Y(r)$ can be considered as equal to zero. We would therefore have $C_2(f)Y(r) = C_3(f)Y(r) = 0$ which gives S_{ray} and S_{mie} proportional to well-known values C_1, C_4, N_p, E_0 and it comes:

$$s_{ray}(r) = K_{ray} N_p E_0 (C_1(P,T,f)X(r)) \quad \text{Eq. 6.105}$$

$$s_{mie}(r) = K_{mie} N_p E_0 (C_4(P,T,f)X(r)) \quad \text{Eq. 6.106}$$

Let's now focus on the synthetic molecular signal $X_{i,sim}$ (indices "$_{sim}$" to insist on the fact that it is simulated signal) introduced in section 6.2.1. It supposes an atmosphere only constituted of molecules and therefore corresponds to particle-free regions of the atmosphere with no particulate term in the equation of the transmission to the satellite (e.g. Eq. 6.37).

Let us denote for bin i the simulated molecular contribution to the signal return in the Fabry-Pérot interferometer for Rayleigh altitudes setting (i.e. Rayleigh Range Bin Setting (RBS)) by $X_{i,sim,ray}$.

Let us denote for bin i the simulated molecular contribution to the signal in the Fizeau interferometer entry for Mie altitudes setting (i.e. Mie Range Bin Setting (RBS)) by $X_{i,sim,mie}$.

Let us denote for bin i the relative errors for Rayleigh signal prediction by $e_{err,ray\ pred,i}$ and the relative errors for Mie signal prediction by $e_{err,mie\ pred,i}$. When applying the relative errors' formula $((observed - predicted)\ /predicted)$, the $X_{i,sim,ray}$ and $X_{i,sim,mie}$ correspond to the predicted terms and we would have:

$$e_{err,ray\ pred,i} = \frac{(S_{ray,i} - X_{i,sim,ray})}{X_{i,sim,ray}} \quad \text{Eq. 6.107}$$

$$e_{err,mie\ pred,i} = \frac{(S_{mie,i} - X_{i,sim,mie})}{X_{i,sim,mie}} \quad \text{Eq. 6.108}$$

Deviation from zero of the relative error's distribution (i.e. the median value) is taken as corrective factor for each K_{ray} and K_{mie} correction:

$$K_{ray} = K_{ray} \times (1 + median(e_{err,ray\ pred,i}) \quad \text{Eq. 6.109}$$

$$K_{mie} = K_{mie} \times (1 + median(\ e_{err,mie\ pred,i}) \quad \text{Eq. 6.110}$$

Unique values of K_{ray} and K_{mie} are therefore provided per orbit and computed in SCA processor.

Default K_{ray} and K_{mie} are showed in Figure 6-13 and Figure 6-14 with black dotted line: initial estimation from IRC mode significantly overestimate the mean oscillations of both K_{ray} and K_{mie} (i.e. respectively light blue and light orange curves) along the orbit. Corrected K_{ray} and K_{mie} by signal prediction performed in particle-free regions of the atmosphere in mid-altitudes are showed in same figures with black continuous line. The consistency of this first correction scheme based on signal prediction for Rayleigh channel signals can be visualized: corrected values given as adjusted constants are closer to the mean oscillations of both K_{ray} and K_{mie} along the orbit.

6.5.3 Fit of K_{ray} and K_{mie} coefficients per observation by regression based on telescope temperatures

When looking at K_{ray} and K_{mie} behavior along the orbit significant deviations had been observed for both calibration coefficients (e.g. light blue curve for K_{ray} in Figure 6-13 and orange curve for K_{mie} in Figure 6-14). The need of a correction scheme per observation (i.e. per Basic repeat Cycle (BRC)) had been expressed.

It was then decided to take advantage of the telescope temperatures information provided by the Accurate Housekeeping Telemetry (AHT) and Thermal Control (TC) processes. The temperatures timelines given at observation scale for twelve sensors dispatched all over the mirror (i.e. outer and inner part) are therefore extracted from the Level 1B product. This second correction scheme is organized as follows: the K_{ray} and K_{mie} are firstly estimated per observation from signal prediction in particle-free regions of the atmosphere and then a multiple linear regression based on telescope temperatures is computed to fit the corrected K_{ray} and K_{mie} .

From equations Eq. 6.107 and Eq. 6.108 the relative errors for signal prediction are given per bin i:

$$K_{ray,i} = K_{ray} \times (1 + e_{err,ray\ pred,i}) \quad \text{Eq. 6.111}$$

$$K_{mie,i} = K_{mie} \times (1 + e_{err,mie\ pred,i}) \quad \text{Eq. 6.112}$$

Let us denote the calibration coefficients corrected for bin i in a same observation j by $K_{ray,i,j}$ and $K_{mie,i,j}$ .The median value of corrected K_{ray} and K_{mie} in selected bin (i.e. particle-free conditions) for observation j are taken and final output per observation $K_{ray,j}$ and $K_{mie,j}$ can be expressed as follows:

$$K_{ray,j} = median(K_{ray,i,j}) \quad \text{Eq. 6.113}$$

$$K_{mie,j} = median(K_{mie,i,j}) \quad \text{Eq. 6.114}$$

We thus obtain a couple ($K_{ray,j}$, $K_{mie,j}$) for each observation j per orbit. Then a multiple linear regression based on telescope temperatures given by twelve AHT and TC sensors is computed to fit these corrected values of both calibration coefficients.

Let us denote for the observation j the calibration coefficients by $K_{ray,fit,j}$ and $K_{mie,fit,j}$ and the telescope temperatures by $M1$. The multiple linear regression can be seen as an ordinary least square

fit in twelve dimensions and the fitted coefficients can be expressed as a function of $M1$ which depends on time t , longitude lon , and latitude lat:

$$K_{ray,fit,j}(t, lon, lat) = f(M1(t, lon, lat)) + \varepsilon \quad \text{Eq. 6.115}$$

$$K_{mie,fit,j}(t, lon, lat) = f(M1(t, lon, lat)) + \varepsilon \quad \text{Eq. 6.116}$$

The linear system of the regression model with 12 predictor variables $X_1, X_2, \ldots, X_{12}$ (i.e. the twelve temperatures timeline extracted from Level 1B product at observation scale) and fitted coefficients $K_{ray,fit,j}$ and $K_{mie,fit,j}$ as the response of the fitting model can be expressed as follows:

$$K_{ray,fit,j} = \beta_0 + \beta_1 X_1 + \ldots + \beta_{12} X_{12} + \varepsilon \quad \text{Eq. 6.117}$$

$$K_{mie,fit,j} = \beta_0 + \beta_1 X_1 + \ldots + \beta_{12} X_{12} + \varepsilon \quad \text{Eq. 6.118}$$

With detailed variables for Eq. 6.117:

$K_{ray,fit,j}$ = the input value of the dependent variable

$\beta_0 = K_{ray,j}$ = the intercept (i.e. value of $K_{ray,fit,j}$ when all other parameters are set to 0)

$\beta_1 X_1$ = the regression coefficient (β_1) of the first independent variable (X_1) (i.e. temperatures given by the first sensor)

$\beta_{12} X_{12}$ = the regression coefficient of the last independent variable (X_{12}) (i.e. temperatures given by the 12th sensor)

ε = least square fit residual (i.e. variations that the least square model cannot account for).

Fitted $K_{ray,fit,j}$ and $K_{mie,fit,j}$ per observation can be seen respectively in Figure 6-13 (e.g. dark blue curve) and in Figure 6-14 (e.g. red curve) and can be compared to the K_{ray} and K_{mie} estimated from the IRC mode (e.g. dark dotted lines) as K_{ray} and K_{mie} estimated from orbit correction with Eq. 6.109 and Eq. 6.110 (e.g. dark continuous line).

It can be observed that the multiple linear regression allows to reduce the amplitude of the observed K_{ray} and K_{mie} oscillations which are highly outlier contaminated (e.g. light blue curve in Figure 6-13 and orange curve in Figure 6-14). Moreover by the use of the regression model we can provide fitted $K_{ray,fit,j}$ and $K_{mie,fit,j}$ values per observation even in particle regions of the atmosphere (i.e. when particles feature or broken clouds occur in top-bin).

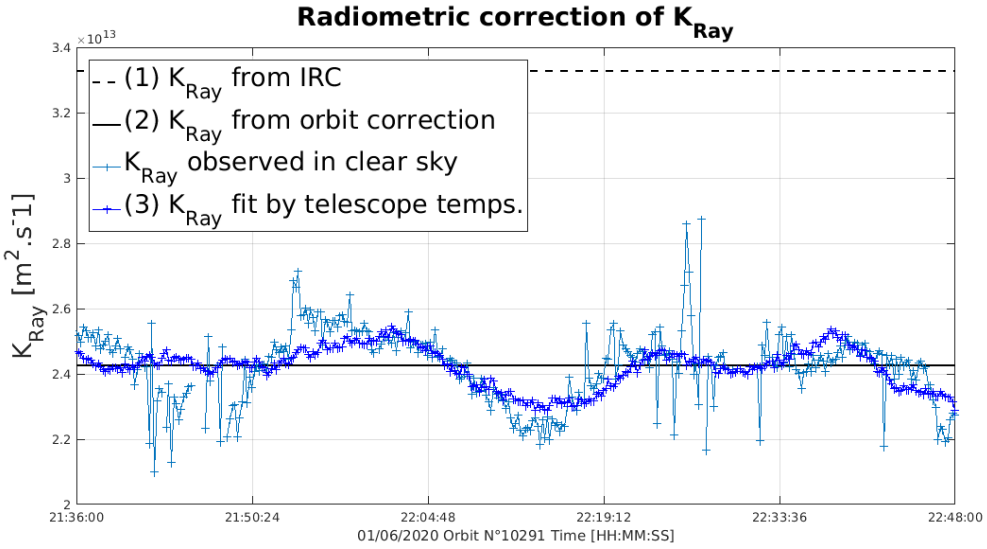

Figure 6-13: Radiometric correction of calibration coefficient Kray

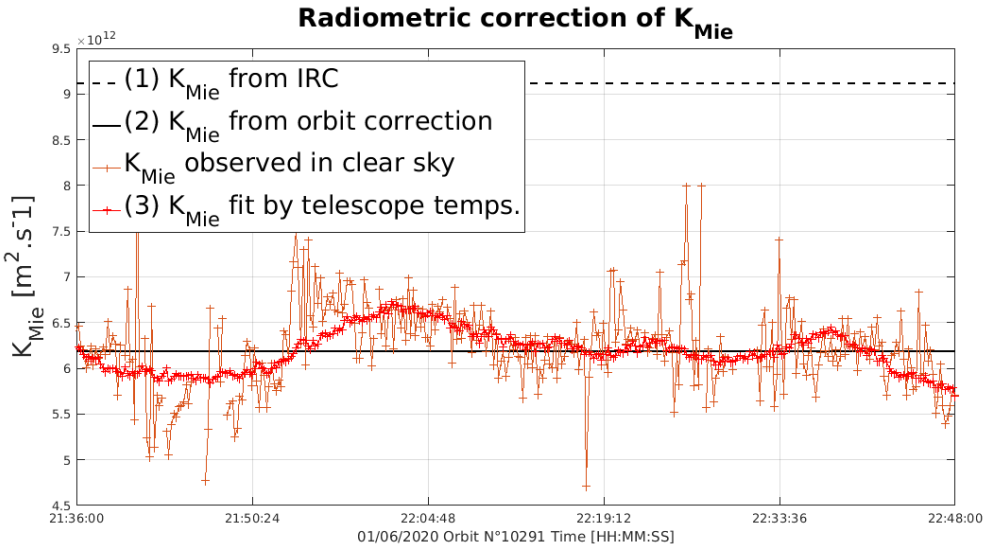

Figure 6-14: Radiometric correction of calibration coefficient Kmie

6.6 The Mie channel algorithm

The standard correct algorithm SCA applies on crosstalk-corrected data and matching bins of the Rayleigh and Mie channels to determine direct extinction and backscatter coefficients. Sometimes, there might not be valid L1B data to correct crosstalk, or there might be few bins matching. The Mie channel algorithm (MCA) aims at retrieving particle characteristics based on the sole Mie channel data, i.e., using ALADIN as backscatter lidar. It needs Mie channel data, synthetic data, calibration data, L1B scattering ratio and climatology data for the backscatter-to-extinction ratio.

Going back to the signal detected on the Mie channel (Eq. 6.19; *reminder*: the indices "$_{mie}$" mean quantities are interpolated along Mie scale), the extraction of $Y_{mie,i}$ leads to:

$$Y_{mie,i} = \frac{S_{mie,i}}{K_{mie} N_p E_0 \left(C_{4,mie,i} \frac{X_{mie,i}}{Y_{mie,i}} + C_{3,mie,i} \right)} \qquad \text{Eq. 6.119}$$

The ratio $\frac{X_{mie,i}}{Y_{mie,i}}$ is linked to the scattering ratio ρ, as it has been mentioned in 6.2.2.2:

$$\frac{X_{mie,i}}{Y_{mie,i}} = \frac{1}{\rho - 1} \qquad \text{Eq. 6.120}$$

So, depending on the accuracy of L1B scattering ratio ρ_{L1B}, a crosstalk pseudo-correction can be performed to get the particulate signal at telescope entry $Y_{mie,i}$:

$$Y_{mie,i} = \frac{S_{mie,i}}{K_{mie} N_p E_0 \left(\frac{C_{4,mie,i}}{\rho_{L1B} - 1} + C_{3,mie,i} \right)} \qquad \text{Eq. 6.121}$$

Then it is possible to work on the integral expression of $Y_{mie,i}$ detailed in Eq. 6.17. The same approximations as for the SCA are made: range squared terms and molecular terms vary slowly over a range bin and can be approximated by their mean values:

$$Y_{mie,i} = \frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2}{R_{mean,i}^2} e^{-L_{m,i}} \int_{R_{i-1}}^{R_i} \beta_p(r) \exp\left(-2 \int_{R_{i-1}}^{r} \alpha_p(u) du \right) dr \qquad \text{Eq. 6.122}$$

This time, the particle backscatter cannot be averaged and put out of the integral because there is no information about its shape (unlike β_m). Yet, using the particle BER $k_p = \beta_p / \alpha_p$, an analytic solution can be derived:

$$\begin{aligned} Y_{mie,i} &= \frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2}{R_{mean,i}^2} e^{-L_{m,i}} \int_{R_{i-1}}^{R_i} k_p \alpha_p(r) exp\left(-2 \int_{R_{i-1}}^{r} \alpha_p(u) du \right) dr \\ &= \frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2}{R_{mean,i}^2} k_p e^{-L_{m,i}} \int_{R_{i-1}}^{R_i} \alpha_p(r) exp\left(-2 \int_{R_{i-1}}^{r} \alpha_p(u) du \right) dr \end{aligned} \qquad \text{Eq. 6.123}$$

The integral involves the function $r \mapsto \alpha_p(r)$ and one of its primitives $r \mapsto \int_{R_{i-1}}^{r} \alpha_p(u) du$. Therefore, it can be written:

$$\int_{R_{i-1}}^{R_i} \alpha_p(r)\exp\left(-2\int_{R_{i-1}}^{r}\alpha_p(u)du\right)dr = \left[-\frac{1}{2}\exp\left(-2\int_{R_{i-1}}^{r}\alpha_p(u)du\right)\right]_{R_{i-1}}^{R_i} = \frac{1-\exp\left(-2\int_{R_{i-1}}^{R_i}\alpha_p(r)dr\right)}{2} \quad \text{Eq. 6.124}$$

Supposing again that the range bin is homogeneously filled with particles, it comes:

$$Y_{mie,i} = \frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2}{R_{mean,i}^2} k_p e^{-L_{m,i}} \frac{1-e^{-2L_{p,i}}}{2} \quad \text{Eq. 6.125}$$

Having a look at this expression, one can notice that there is a dependence on previous bin data. Again, a recursive algorithm will be used.

Initialisation:

The "bin" between the satellite and the top of Mie channel profile is supposed to be clear: $\int_0^{R_1}\alpha_p(r)dr = 0$.

Recurrence relation:

Once values are known for the range bin i-1, they are computed in bin i:

$$L_{p,i} = -\frac{1}{2}ln\left(1-\frac{2Y_{mie,i}R_{mean,i}^2 e^{L_{m,i}}}{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2 k_p}\right) \quad \text{Eq. 6.126}$$

Molecular quantities are computed from synthetic molecular data based on temperature and pressure profiles. The transmission from the satellite to the bin i is computed as in Eq. 6.59.

Eventually, even if it is not possible to perform real crosstalk correction, the HSRL improvement on result quality can still be exploited through ρ_{L1B}. A pseudo crosstalk correction is used to get closer to particle backscattered signal at telescope entry than with standard elastic backscatter lidars. The quality of this correction is depending on ρ_{L1B}'s, which is pretty satisfying. Yet the other dependency on the BER k_p hinders the algorithm from retrieving accurate values of optical depth. k_p is an *a priori* value and a too big difference to the real value could spoil the results. For instance, if it is too underestimated, then the argument of the logarithm will tend towards zero and the retrieved $L_{p,i}$ will tend towards infinity, or even get complex if the argument gets negative. On the contrary, if k_p is overestimated, SLOD variations will be flattened but in this case, results can still be used to attest the presence of particles and the relative variations of the local optical depth.

Finally, to be used as quantitative products, MCA outputs require a great attention to be paid to the backscatter-to-extinction ratio. If no accurate climatology data can be provided, then it is better to ensure overestimation of k_p so that the logarithm does not diverge. In this case, MCA outputs can only be used qualitatively. Error propagation has not been estimated yet. Note that only propagation may be predicted. The level of error is highly depending on the accuracy of the BER, i.e. on climatology, and local extraordinary events can significantly spoil this accuracy over a measurement.

NB: When the optical depth is large, e.g. for dense water clouds, it comes

$$Y_{mie,i} = \frac{T_{m,sat,i-1}^2 T_{p,sat,i-1}^2}{R_{mean,i}^2} a_e e^{-L_{m,i}} \qquad \text{Eq. 6.127}$$

It appears that the backscattered signal results of an effective albedo $a_e = k_p/2$. In presence of multiple scattering processes, the extinction coefficient is reduced by a factor $0 \leq \eta \leq 1$. Consequently the backscatter-to-extinction ratio is increased by the same factor so the effective albedo becomes $a_e = k_p/2\,\eta$.

7 ITERATIVE CORRECT ALGORITHM

As mentioned in 6.2.2.3, the SCA misestimates products in partially filled bins. The iterative correct algorithm (ICA) aims at making hypotheses on the filling of the range bins of a feature identified by the SCA, and at assessing which one gives the more relevant results in the first clear bin under the feature. Such a bin is therefore required and the ICA may be inapplicable in the ABL, for instance. Eventually, both the location of a particle layer and the slant local optical depth are retrieved. The algorithm uses the same inputs as the SCA. It needs crosstalk-corrected data and applies only in bins that match between both Mie and Rayleigh channels.

7.1 Partial filling equations

Considering a single layer located between altitudes R_a to R_b such $R_{i-1} \le R_a < R_b \le R_i$ with constant extinction coefficient $\alpha_{p,[a;b]}$, the integral that lead to equation Eq. 6.35 may be recalculated. Taking the layer into account and splitting the integral into three parts (over $[R_{i-1}; R_a]$, $[R_a; R_b]$ and $[R_b; R_i]$):

$$\begin{aligned}\int_{R_{i-1}}^{R_a} \exp\left(-2\int_{R_{i-1}}^{r} \alpha_p(u)du\right)dr &= \int_{R_{i-1}}^{R_a} \exp(0)\,dr \\ &= R_a - R_{i-1}\end{aligned} \qquad \text{Eq. 7.1}$$

$$\begin{aligned}\int_{R_a}^{R_b} \exp\left(-2\int_{R_{i-1}}^{r} \alpha_p(u)du\right)dr &= \frac{1-e^{-2L_{p,[a;b]}}}{2\alpha_{p,[a;b]}} \\ &= (R_b - R_a)\,H\left(2L_{p,[a;b]}\right)\end{aligned} \qquad \text{Eq. 7.2}$$

$$\begin{aligned}\int_{R_b}^{R_i} \exp\left(-2\int_{R_{i-1}}^{r} \alpha_p(u)du\right)dr &= \int_{R_{i-1}}^{R_a} \exp\left(-2\int_{R_a}^{R_b} \alpha_p(u)du\right)dr \\ &= e^{-2L_{p,[a;b]}}\int_{R_{i-1}}^{R_a} dr \\ &= (R_i - R_b)e^{-2L_{p,[a;b]}}\end{aligned} \qquad \text{Eq. 7.3}$$

With H the function defined in 6.2.1. Hence:

$$\int_{R_{i-1}}^{R_i} \exp\left(-2\int_{R_{i-1}}^{r} \alpha_p(u)du\right)dr = (R_a - R_{i-1}) + (R_b - R_a)\,H\left(2L_{p,[a;b]}\right) + (R_i - R_b)e^{-2L_{p,[a;b]}} \qquad \text{Eq. 7.4}$$

And:

$$\frac{1}{T_{p,sat,i-1}^2} NITWT_i = \frac{R_a - R_{i-1}}{R_i - R_{i-1}} + \frac{R_b - R_a}{R_i - R_{i-1}}\,H\left(2L_{p,[a;b]}\right) + \frac{R_i - R_b}{R_i - R_{i-1}} e^{-2L_{p,[a;b]}} \qquad \text{Eq. 7.5}$$

If $R_{i-1} = R_a$ and $R_b = R_i$, then Eq. 7.5 is the same as Eq. 6.44. Depending on the location of the particle layer in the range gate the solutions are not symmetric and quite different. Consequently, the retrievals of local optical depth are different as well. The same reasoning as above could be conducted for two or more layers.

Similarly to the SCA, $L_{p,[a;b]}$ is limited to positive values. Then, the quantity of Eq. 7.5 is less than or equal to unity. On the other hand, it is greater or equal to the constant term $(R_a - R_{i-1})/(R_i - R_{i-1})$. Hence the inequality Eq. 7.6 in a particle loaded range bin *i*:

$$\frac{R_a - R_{i-1}}{R_i - R_{i-1}} \leq \frac{1}{T_{p,sat,i-1}^2} NITWT_i \leq 1 \qquad \text{Eq. 7.6}$$

This criterion enables to reject some definitely unrealistic case assumptions and will be referred to as the credibility criterion (CC).

7.2 Proposed cases

The computing time rapidly increases over a wide feature because each case is checked in each bin and the algorithm is recursive, like the SCA. Considering N_{case} cases in a N_{bin} wide cloud, leads to consider ${N_{case}}^{N_{bin}}$ different arrays. Moreover, the higher is the segmentation of a bin, the more accurate must be input data so that the best case may be selected reliably enough. The algorithm is applied on noisy signals and the risk is to produce artificially highly resolved wrong data.

A good trade-off suggests ceiling to 7 cases. They are displayed in Figure 7-1.

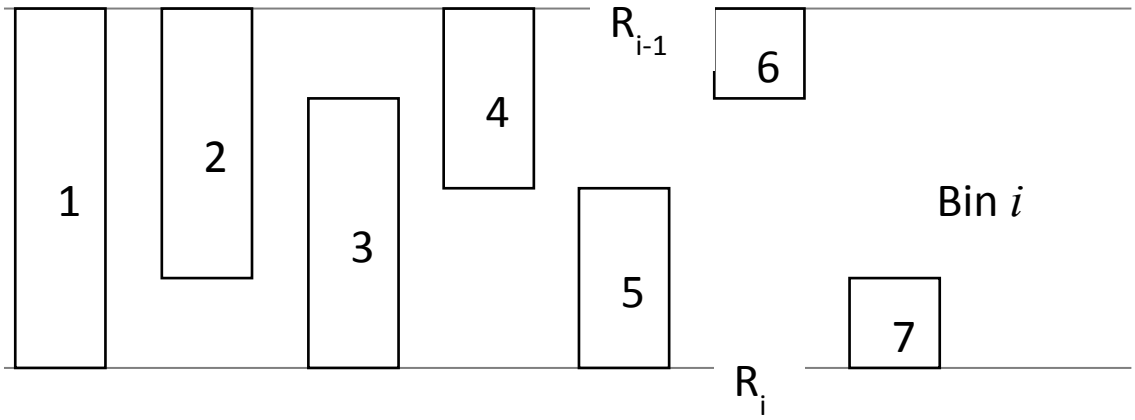

Figure 7-1: Illustration of the various filling cases analysed by the ICA. The fractions of the i^{th} bin are 1, $3/4$, $1/2$ and $1/4$ from the top (R_{i-1}) and from the base (R_i) of the bin.

7.2.1 Cases' detailed equations

This subsection details the expressions taken by the general equations – Eq. 7.5 to retrieve $L_{p,[a;b]}$, and the credibility criterion Eq. 7.6 – for the various selected cases.

Case 1. It is SCA full filling case.

Case 2.

$$\frac{1}{T_{p,sat,i-1}^2} NITWT_i = \frac{3}{4} H\left(2L_{p,[a;b]}\right) + \frac{1}{4} e^{-2L_{p,[a;b]}} \qquad \text{Eq. 7.7}$$

$$0 < \frac{1}{T_{p,sat,i-1}^2} NITWT_i \leq 1 \qquad \text{Eq. 7.8}$$

Case 3.

$$\frac{1}{T_{p,sat,i-1}^2} NITWT_i = \frac{1}{4} + \frac{3}{4} H\left(2L_{p,[a;b]}\right) \quad \text{Eq. 7.9}$$

$$\frac{1}{4} < \frac{1}{T_{p,sat,i-1}^2} NITWT_i \leq 1 \quad \text{Eq. 7.10}$$

Case 4.

$$\frac{1}{T_{p,sat,i-1}^2} NITWT_i = \frac{1}{2} H\left(2L_{p,[a;b]}\right) + \frac{1}{2} e^{-2L_{p,[a;b]}} \quad \text{Eq. 7.11}$$

$$0 < \frac{1}{T_{p,sat,i-1}^2} NITWT_i \leq 1 \quad \text{Eq. 7.12}$$

Case 5.

$$\frac{1}{T_{p,sat,i-1}^2} NITWT_i = \frac{1}{2} + \frac{1}{2} H\left(2L_{p,[a;b]}\right) \quad \text{Eq. 7.13}$$

$$\frac{1}{2} < \frac{1}{T_{p,sat,i-1}^2} NITWT_i \leq 1 \quad \text{Eq. 7.14}$$

Case 6.

$$\frac{1}{T_{p,sat,i-1}^2} NITWT_i = \frac{1}{4} H\left(2L_{p,[a;b]}\right) + \frac{3}{4} e^{-2L_{p,[a;b]}} \quad \text{Eq. 7.15}$$

$$0 < \frac{1}{T_{p,sat,i-1}^2} NITWT_i \leq 1 \quad \text{Eq. 7.16}$$

Case 7.

$$\frac{1}{T_{p,sat,i-1}^2} NITWT_i = \frac{3}{4} + \frac{1}{4} H\left(2L_{p,[a;b]}\right) \quad \text{Eq. 7.17}$$

$$\frac{3}{4} < \frac{1}{T_{p,sat,i-1}^2} NITWT_i \leq 1 \quad \text{Eq. 7.18}$$

7.2.2 Case-related solution comparison

Under the assumption of small optical depths, asymptotic developments enable to compare the expressions of L_p found in various cases. Results are presented through a comparison between the value found by the SCA, noted L_1, and by ICA for the case k, noted L_k.

Case k	2	3	4	5	6	7
Ratio L_k/L_1	$4/5$	$4/3$	$2/3$	2	$4/7$	4

Table 7.1: ratio of ICA optical depth over SCA optical depth for the considered cases.

So, for instance, if extinction coefficients are considered in the seventh case,

$$\begin{aligned} L_1 &= \frac{L_7}{4} \\ \alpha_1 \Delta R &= \frac{1}{4} \times \alpha_7 \frac{\Delta R}{4} \\ \alpha_1 &= \frac{\alpha_7}{16} \end{aligned} \qquad \text{Eq. 7.19}$$

The extinction coefficient would suffer an underestimation of a factor 16. This highlights the interest of the iterative correct algorithm.

7.3 Iterative correct algorithm for partial filling

The iterative correct algorithm boils down to applying the same steps as in the SCA but in a loop over all the possible cases for the considered features.

Firstly, β_p from the SCA is used to locate features that have a clear bin underneath. Their locations are stored.

Secondly, SCA values are taken until the first feature is reached.

Thirdly, the case matrix is built; each column stores a case combination for the whole feature. For each case array, the algorithm retrieves L_p bin per bin, recursively, as for the SCA but making use of matching forms of Eq. 7.5. The credibility criterion of Eq. 7.6 is used to reject candidates if they produce obviously wrong results in an intermediate bin. If the algorithm could run until the end of the feature, then the ratio $NITWT_i/T^2_{p,1,i-1}$ is computed in the bin below the feature, identified as clear. If it is close enough to unity, then this combination is kept. The best, in the sense of proximity to unity, valid combination is used to continue the computation.

Fourthly, the SCA is applied in clear bins between the feature and the next element. It is reinitialised in the first clear bin under the feature. This enables to get rid of errors accumulated over previous bins. If the next element is another feature, step 3 is repeated, if it is the ABL, the SCA is applied until the ground.

Note that 7 cases may be too many to reliably state in which bin fraction the layer is located. A reduction to 3 cases increases this reliability though results are quite feature-regularity-dependent. Note that, at the moment, only cases 1, 4 and 5 are tested in the delivered version of the ICA.

8 FEATURE FINDER

8.1 Principle

The feature finder is the first step in a second product chain that aims at providing products with a higher resolution than the default BRC-averaged product.

The idea is to identify blocks of measurements with similar properties in order to process them as a single homogeneous feature. This provides a resolution better than one BRC while using the best possible signal quality. Measurement level derivation of optical properties is not feasible because one measurement is, according to the current default settings, only the aggregation of 20 laser pulses and does not provide a signal of good enough quality for accurate determination of optical properties.

After being identified by the feature finder, “features”, are processed as “groups” of measurements.

The feature finder described below is tightly inspired from the EarthCARE mission. This aerosol lidar mission implements a feature finder that works in two successive steps. The feature finder presented below is inspired from the first step.

We aim to identify features as groups of measurements on a given range bin with a homogeneous content in particles. The algorithm is based on the detection of the presence of particles, estimated from the Mie channel SNR.

We assume a simple model for the Mie channel response: a given pixel is either “clear sky”, with no particles, or “particle-loaded”. The signal for both clear sky bins and particle-loaded bins are noisy and their signal levels follow a Gaussian distribution, with the Mie channel particle-loaded signal being generally higher than clear sky signal (see Figure 8-1). If we consider a given signal level as a lower threshold, the part of the clear sky signal distribution above the threshold (red area in Figure 8-1*)* is a false-alarm probability, i.e. the probability that signal in a clear sky pixel actually exceeds the threshold. The part of the particle-loaded distribution that falls below the threshold (dark green area in Figure 8-1*)* is the probability of missing detection.

The probability of detection on the Mie channel - the area of the bright green area in Figure 8-1 - is then:

$$P_{Mie} = 1 - \frac{1}{2} erfc\left(\frac{S - \delta S}{\sqrt{2}\,\delta S}\right) \qquad \textbf{Eq. 8.1}$$

where S is the expected useful signal level in the particle-loaded bin and δS the corresponding noise level, erfc is the complementary error function.

If we knew the behaviour of the instrument (signal levels for cloud and cloud free bins) and the corresponding noise intensity, we would be able to derive the distributions pictured in Figure 8-1 and from this to compute the values for the threshold corresponding to a specified good detection rate or false alarm rate.

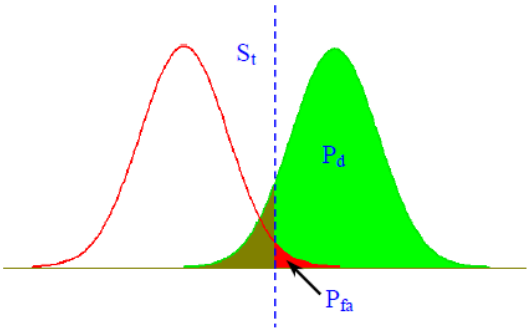

Figure 8-1: Probability of detection with the distribution for clear sky as a red line and the distribution for the particle-loaded signal as the green area. The red area right of the threshold is the false-alarm probability and the dark green area left of it is the probability of missed detection. From [RD 41]

If we do not have access to theoretical values for noise and noise-free signals corresponding to cloudy or clear conditions, we can take another point of view on the signal distributions. Below, we consider that S is the measured useful signal in a given bin and δS the noise in this bin. Noise takes a different value in each bin and we do not have access to the noise level of a given bin but we have access to both the signal in the said bin and its SNR.

$$SNR = \frac{S}{\delta S} \qquad \textbf{Eq. 8.2}$$

$$\text{or } \delta S = \frac{S}{SNR} \qquad \textbf{Eq. 8.3}$$

Then, we can compute the probability that the given signal level is reached because of the presence of particles, as opposed to only being reached because of noise. Equation Eq. 8.1 is seen as the probability that a given signal S is larger than noise.

Simplifying equation Eq. 8.1 with equation Eq. 8.3, we get:

$$P_{Mie} = 1 - \frac{1}{2} erfc\left(\frac{SNR_{Mie} - 1}{\sqrt{2}}\right) \qquad \textbf{Eq. 8.4}$$

And *erfc* being monotonic (see Figure 8-2), applying a threshold on P_{Mie} is the same as applying a threshold on SNR_{Mie}, i.e. any point on the curve to the right of the vertical red line is also above the horizontal red line.

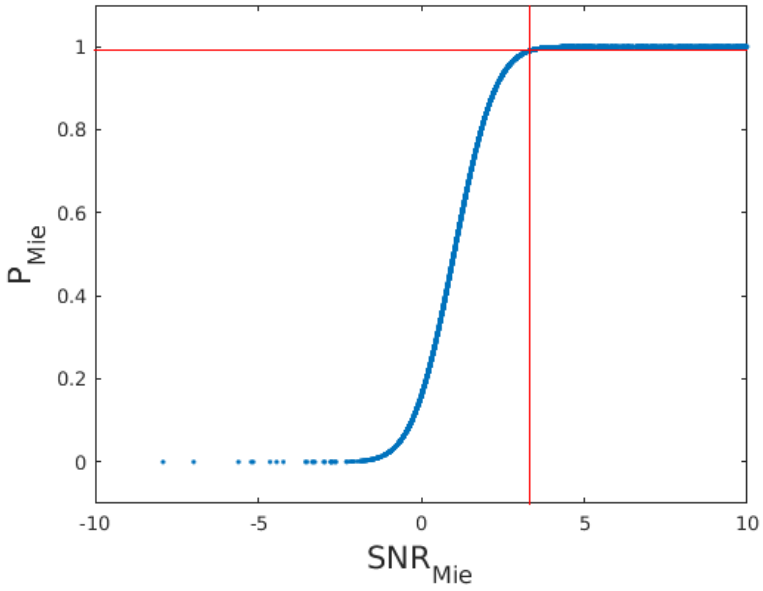

Figure 8-2: Probability that the signal level not due to noise as a function of SNR (Eq. 8.4).

If the rigorous framework from the CALIPSO algorithm [RD 41] provided the inspiration for this simple detection method, the EarthCARE algorithm suggested the use of a median filter. ADM-Aeolus vertical resolution being optimised for wind detection and already on the coarse side for cloud detection, we choose not to filter signals on a vertical scale and use only horizontal filtering.

The L2Ap feature finder currently needs 3 input parameters:

- The proposed default size for the median filter is 5 measurements wide but this can be modified as an input parameter to the L2Ap.
- The threshold on the Mie SNR was determined empirically from simulations on several scenarios representing various conditions, in order to get a reasonable ratio between false alarm and detection rates. The threshold is another input parameters to the L2Ap feature finder.
- At last, groups detected by this method are screened for minimum size. Any group smaller than a given number of measurements is rejected. The suggested minimum size is 5 measurements.

The algorithm is not able to detect light features (e.g. aerosols) under optically thick clouds.

Figure 8-3 gives an example of detection on a challenging scene derived from LiTE data over Myanmar, with thick scattered cloud and aerosols. For the test, it is considered that a pixel contains a

feature if the “true” backscatter in the simulation (specified in the simulator input atmosphere) for this pixel[1] is above 10^{-6} sr^{-1} m^{-1}.

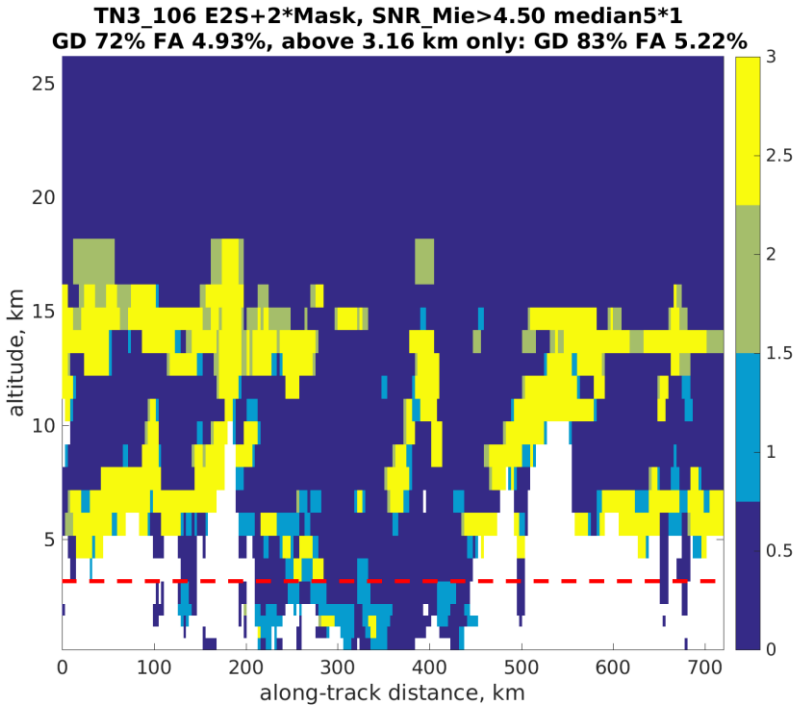

Figure 8-3: Feature detection on a simulated scene. The image shows good detections (yellow), false alarms (green), missed detection (light blue) and absence of feature (dark blue). The red dashed horizontal shows where the 3.16 limit is while the white pixels are areas where total two-way transmission is below 0.1. The detection score does not take this area into account.

This algorithm could be refined by considering an altitude-dependent threshold. For instance, the uppermost “false alarms” in Figure 8-3 are due to the fact that lighter clouds, with a backscatter below the 10^{-6} sr^{-1} m^{-1} threshold, are observed with a relatively higher SNR than the same feature would have lower in the atmosphere.

8.2 The SCA on groups

The SCA aims at characterising the group by its BER. It has been shown in 6.3 that it is centred on bin borders and has a predominance sphere of a half bin on each side. Subsequently, for a group at the height level i, base and top BERs are required and both $k_{p,i-\frac{1}{2}}$ and $k_{p,i+\frac{1}{2}}$ are calculated. Though it cannot be accessed directly, the feature’s microphysics is then bounded.

[1] One BRC is typically made of 24*30 pixels (number of range bins * number of measurements in a BRC). The input to the E2S contains many more vertical levels (typically ~300). For deriving the “true” backscatter coefficient at simulation input in the same grid as our products, we consider the average of E2S input values over the instrument bins.

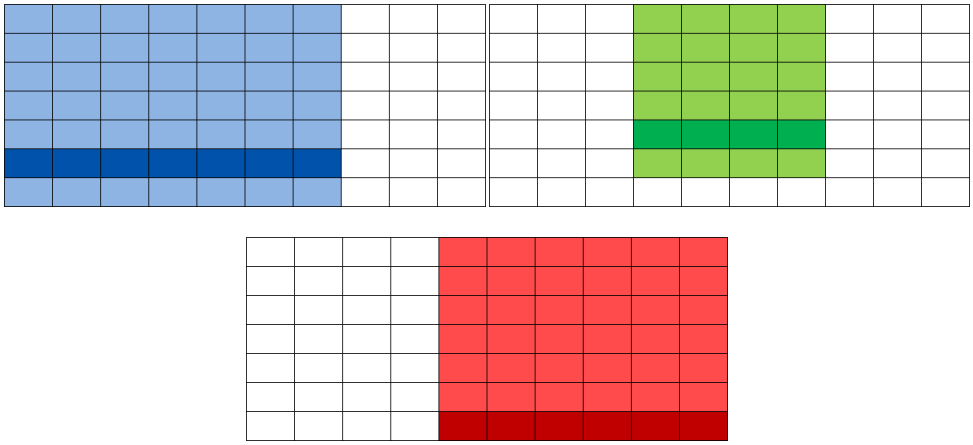

Figure 8-4: SCA applied on three groups to get BERs. Each one is processed independently from the others. Dark-coloured areas are the even group bins and light-coloured ones are required to retrieve group products. If a group is on the bottom of a profile, then only the upper border BER may be retrieved.

8.3 Discussion

8.3.1 General considerations

The immediate criticism to this algorithm is that it is based on SNR and not on k_p. This means that if a thick but weakly backscattering feature is close to a lighter but strongly backscattering one, they will not be discriminated and the resulting k_p will be an average over these features.

The main issue is to base a finding algorithm on k_p to aggregate signals before it is computed from these signals. A real BER-based algorithm could have relied for instance on a sampling of the height level, starting from seeds, each time calculating the BER and aiming at stabilising it around the value of the feature. A technique akin to simulate annealing could have been used. Unfortunately, the instrumental noise is too high at measurement level to expect a reliable convergence of such an algorithm, even on a homogeneous and rather thick and wide feature.

Information on isolated bins' particle characteristics are required to pack several of these bins into one feature. The issue is that bin-wise retrievals suffer from high noise levels. The Mie channel SNR ended up being the best compromise to easily select "particle-loaded" bins from reliable and accessible data.

The algorithm works well with the Mie channel SNR because of the way this Mie SNR is designed: it actually considers the crosstalk contribution from Rayleigh scattering as noise. The Mie SNR then actually indicates the "legibility" of particle signal from its background rather than the signal level in the so called "Mie useful signal". The Mie useful signal integrates all light received in the sensor behind the Fizeau interferometer, i.e. light scattered from both particles and molecules [RD 42].

Finally, it is to be noted that neither the depolarisation nor the multiple scattering effects are taken into account in the simulations that help test and develop the algorithm.

8.3.2 Error due to heterogeneity

The rationale of the algorithm could be questioned as follows: About packing on β_p and about SCA α_p retrieval, to what extent is it relevant to pack bins into horizontal groups with no attention paid to the upper features that may partially fill the column?

Indeed, for groups with heterogeneous upper profiles, the average retrieved products may be affected by an unequal weighting of the contribution of each bin by its own upper column.

More formally, with the index k to horizontally distinguish the bins of the group, the real expressions of the group signals $X_{i,gr}$ and $Y_{i,gr}$ are, assuming constant molecular values over the BRC for the height level i and constant laser energy per pulse:

$$X_{i,gr} = \frac{T_{m,sat,i-1}^2 e^{-L_{m,i}}}{R_{mean,i}^2}\beta_{m,i}\sum_{k\in gr}\left(\frac{1-e^{-2L_{p,i,k}}}{2\alpha_{p,i,k}}\exp\left(-2\sum_{j=1}^{i-1}L_{p,j,k}\right)\right)$$

$$Y_{i,gr} = \frac{T_{m,sat,i-1}^2 e^{-L_{m,i}}}{R_{mean,i}^2}\sum_{k\in gr}\left(\beta_{p,i,k}\frac{1-e^{-2L_{p,i,k}}}{2\alpha_{p,i,k}}\exp\left(-2\sum_{j=1}^{i-1}L_{p,j,k}\right)\right) \quad \text{Eq. 8.1}$$

Then, the expression of Eq. 6.52 becomes:

$$\frac{Y_{i,gr}}{X_{i,gr}}\beta_{m,i,sim} = \frac{\sum_{k\in gr}\left(\beta_{p,i,k}\frac{1-e^{-2L_{p,i,k}}}{2\alpha_{p,i,k}}\exp\left(-2\sum_{j=1}^{i-1}L_{p,j,k}\right)\right)}{\sum_{k\in gr}\left(\frac{1-e^{-2L_{p,i,k}}}{2\alpha_{p,i,k}}\exp\left(-2\sum_{j=1}^{i-1}L_{p,j,k}\right)\right)} \quad \text{Eq. 8.2}$$

It means that the β_p that is actually calculated for the group is the barycentre of the β_p of each bin weighted by the particle attenuation effect on the profile.

To get the particle coefficient, the group NITWT is calculated, with n_{gr} the number of bins in the group:

$$\frac{X_{i,gr,obs}}{X_{i,gr,sim}} = \frac{1}{n_g}\sum_{k\in gr}\left(\frac{1-e^{-2L_{p,i,k}}}{2L_{p,i,k}}\exp\left(-2\sum_{j=1}^{i-1}L_{p,j,k}\right)\right) \quad \text{Eq. 8.3}$$

Making use of Eq. 6.39, the approximation of averaged values for the group boils down to saying that:

$$\frac{1}{T_{p,sat,i-1,gr}^2}\frac{X_{i,gr,obs}}{X_{i,gr,sim}} = \frac{\frac{1}{n_g}\sum_{k\in gr}\left(\frac{1-e^{-2L_{p,i,k}}}{2L_{p,i,k}}\exp\left(-2\sum_{j=1}^{i-1}L_{p,j,k}\right)\right)}{T_{p,sat,i-1,gr}^2} = \frac{1-e^{-2L_{p,i,gr}}}{2L_{p,i,gr}} \quad \text{Eq. 8.4}$$

Let's set the ideas on a simple numerical example. If the group is 3-bin wide with bin SLOD of 0.25, 0.05 and 0.175 and upper profiles' cumulated SLOD of 0.05, 0 and 0.1 respectively, then the ratio of both right members (bottom/top) of Eq. 8.4 leads to 0.9918. If the upper scene is more heterogeneous, *e.g.* if the upper cumulated SLOD of 0.1 is set to 0.4, then the ratio gets 0.9388. Bringing even more

heterogeneity by one more bin with a SLOD of 0.075 and an upper SLOD of 0.15 leads to 0.9095, i.e. almost 10% of error on the estimation of $H(2L_{p,i,gr})$. The impact on the retrieved L_p can be estimated by a look at the Figure 8-5.

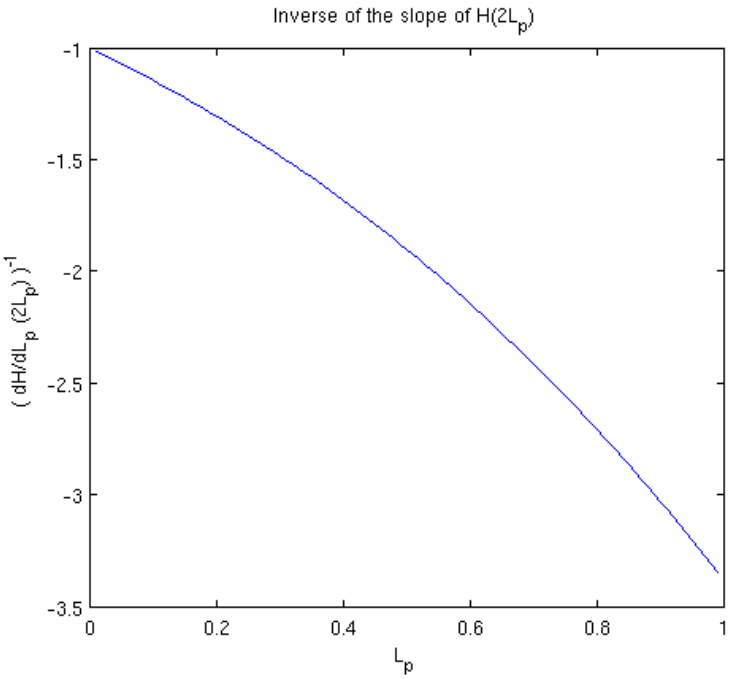

Figure 8-5: Inverse of the slope of $L_p \mapsto H(2L_p)$. For low L_p, the error on $H(2L_p)$ and on L_p are equivalent in magnitude, but beyond 0.5, the error on L_p is twice the one on $H(2L_p)$. It reaches 3 times at 0.9.

In addition, it is to be reminded that under a thick feature the SNR is low, and that sometimes the L1B calculates negative SNR and useful signals. In such cases, the bins and the bin under are removed and treated as if they were ground returns. This error will not appear for groups with a homogeneous upper profile (clear or large layers).

9 PRINCIPLE OF THE SCENE CLASSIFICATION

The Scene Classification Algorithm classifies the types of particle scenes occurring within each observation. The classification distinguishes backscatter from aerosols, water clouds and ice clouds. It is only applied on groups produced by the FF since they are the scene entities that are the most akin to be homogeneous.

ADM-Aeolus will only provide limited information of value for a comprehensive scene classification. This is a weak point of the L2A processor due to the lack of additional information provided by the lidar itself, *e.g.* depolarisation and multi-wavelength operation, and no additional payload instrument. The actual atmospheric scene complexity may be difficult to picture because of various low pass filtering effects associated to horizontal granularity and vertical sampling strategy.

The L2A scene classification will make use of the following variables:

- L2A products (BER and SR) help infer the presence of a cloud,
- NWP products (temperature T, relative humidity rh, cloud liquid and ice water content clwc & ciwc) add clues to discriminate between cloud and aerosol.

Main data, if available, are the BER and the SR. They are used as follows:

- If $k_p < 0.05\ sr^{-1}$: the BER is considered to be typical of a water cloud. The impact of the depolarisation on k_p is not accurately known, nor is its variability in function of crystal types. Subsequently, no relevant threshold value is proposed to discriminate between cirrus and aerosol directly on the BER.
- Lidar ratio of aerosols can vary from about 20 to about 80sr, which is marine aerosol (almost non absorbing) to dust (55sr) to smoke (efficiently absorbing).
- If $\rho > 2.5$, the SR is supposed to testify for a rather dense cloud. Some sensitivity tests have shown that the beam gets almost fully attenuated for cloud optical depth strictly larger than 5 (see).

Auxiliary meteorological data are used as follows:

- If $rh > 94\%$, then there is a high probability that a cloud be present. 94% is taken as 98%=in the cloud, minus 4% of error margin.
- clwc and ciwc are interpreted in function of the situation. Height bins may be wide and various cloud layers may be present at some altitudes. Cases span from 0 to 3:
    - 0: no cloud (clwc=0 and ciwc=0)
    - 1: only water cloud (clwc>0 and ciwc=0)
    - 2: mixed phase cloud (clwc>0 and ciwc>0)
    - 3: only ice cloud (clwc=0 and ciwc>0)
- As a last support when signals tell that there is a cloud-like feature but not the NWP, the temperature helps the user in the discrimination. Three cases are considered:
    - 1: only liquid water is possible ($T > 0°C$),

- 2: water may be mixed-phase ($-40°C < T \leq 0°C$),
- 3: only solid water is possible ($T \leq 40°C$).

These flags are provided to the users in the form of two flags: the Aladin cloud flag based on the BER, the SR and the NWP model relative humidity (figure 9.2) and a NWP cloud flag based on NWP temperature and cloud water content profiles (figure 9.3). These flags are reported in the L2a product according to [RD 28], chapter 3.5.13.

Thanks to the EARLINET project, our knowledge of the lidar ratio (extinction to backscatter ratio) at UV wavelengths has been improved in the last years. Typical lidar ratios for different aerosol and cloud types are reported in table 1.

Aerosol Type	**Lidar ratio (sr)**	**From**
Marine aerosol	20-25	RD33
Urban haze	58+/-12 (Central Europe) 52 +/-10 (North America)	RD31
Artic Haze	60+/-12	RD31
Desert dusts	55+/-6 (Sahara) 38+/-5 (Saudi Arabia)	RD31
Forest fire smoke	46+/-13	RD31
Biomass burning	78+/-5	RD38
Volcanic ash	39+/-10 50-60	RD39 RD40
Cirrus	33+/-9 (Northeast Indian monsoon) 29+/-11 (Southwest Indian monsoon) 19+/- 5.3 20.4+/-7.5	RD35 RD35 RD36 RD36
Water clouds	18.8+/-0.8	RD37

Table 1: Typical values of particle lidar ratios for different aerosol and cloud types.

At a later stage new studies based on particle layer consistency based on new results from CALIPSO and MODIS ([RD 11]) will be used.

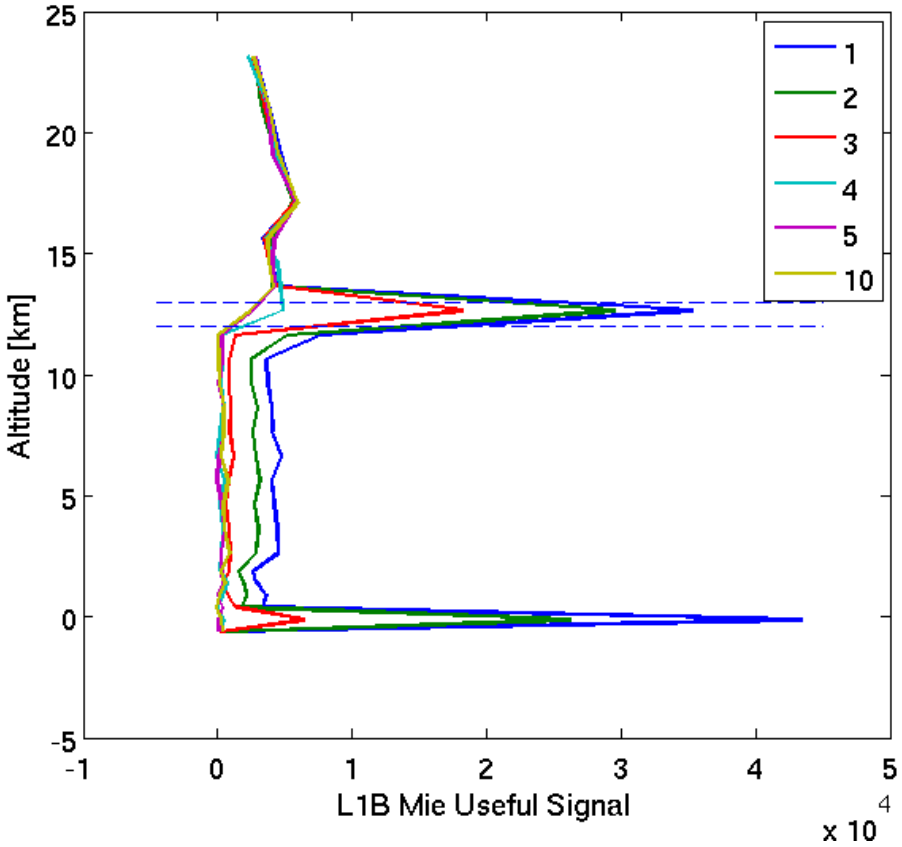

Figure 9.1: L1B Mie useful signal for a high cloud at 12 km altitude and a standard aerosol layer. The Mie useful signals simulated by the L1B processor are displayed for different cloud optical depths from 1 to 10.

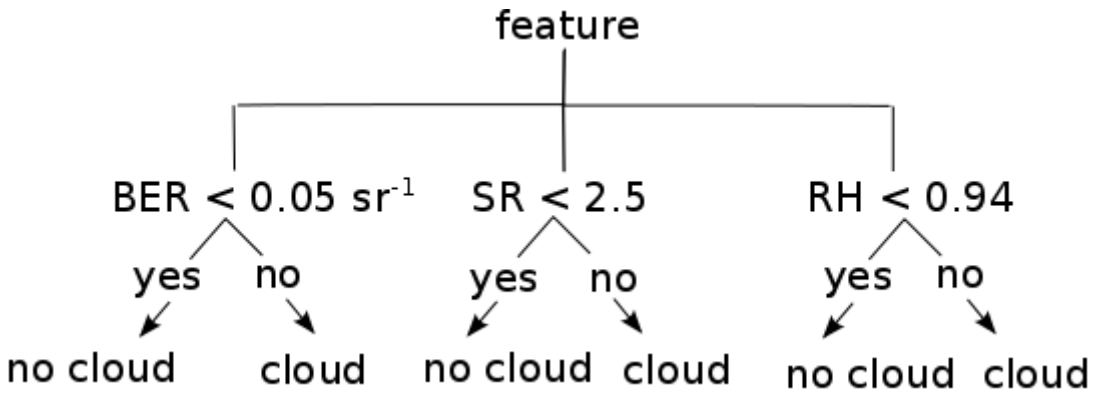

Figure 9.2: Flowcharts of the Aladin cloud flag

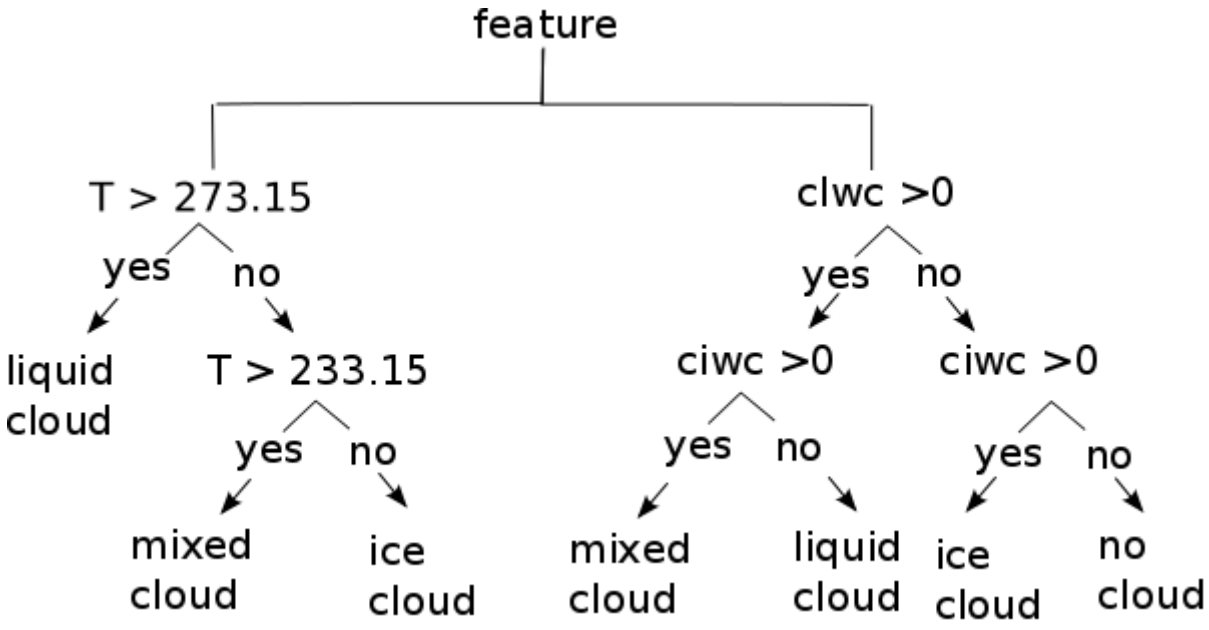

Figure 9.3: Flowcharts of the NWP cloud flag.

10 CONCLUSION

ADM-Aeolus is a wind lidar mission that carries the standalone high spectral resolution lidar ALADIN, continuously operating at one single wavelength (355 nm) with no polarization diversity capability, and no auxiliary payload like an imager. The wind lidar ALADIN is operated in direct detection for molecular and particle detection, the laser emission is circularly polarized, the High Spectral Resolution (HSR) receiver combines a dual (double edge) Fabry-Pérot interferometer and a Fizeau interferometer. According to the basic optical design there is a significant crosstalk between the two channels that calls for calibration. The dual Fabry-Pérot interferometer, called the Rayleigh channel, samples the molecular backscatter whereas the Fizeau interferometer, called the Mie channel, samples particle backscatter.

ADM-Aeolus can provide products on particle backscatter and extinction from the surface up to an altitude of 30 km, by range bin of 250 m, 500 m, 1000 m or 2000 m, depending on the height above the surface and the sampling strategy. The total number of vertical range bins is 24. The elementary horizontal sampling bin varies from a granularity of 2.9 km to 85.7 km, depending on the SNR. One observation is a 2D [i,j] matrix made of 24 rows (horizontal) and of a variable amount of columns (vertical), up to 30 measurements of 20 laser pulses each. The pieces of information to be retrieved are: i) the presence of particles in a range bin, and if so, ii) the optical depth, iii) backscatter coefficient, iv) the scattering ratio, v) the backscatter-to-extinction ratio.

Accumulation height bins can be wide. The validity of the range resolved lidar equations for Rayleigh and Mie channels written for average backscatter and extinction coefficients has been addressed in the present L2A ATBD and bin-accumulated lidar equations have to be applied to Aeolus data. In particular, the average of $\int_{R_{i-1}}^{r} \alpha_m(u)du$ over a bin improves accuracy.

Three different algorithms have been proposed to retrieve spin-off particle products from ALADIN signals at BRC level. Two of them, the Standard Correct Algorithm and the Iterative Correct Algorithm, start from crosstalk-corrected data and thus need signals from both Mie and Rayleigh channels, with a requirement on bin matching. The third one makes use of the sole Mie channel signal but performs a pseudo-crosstalk-correction based on level 1b scattering ratio. All of them need meteorological data: the SCA and the ICA to compute the Normalised Integrated Two-Way Transmission, the ratio of observed molecular signal to simulated molecular return, the foundation of these algorithms; the MCA to compute the molecular transmission. The latter also requires climatology data for an *a priori* value of the backscatter-to-extinction ratio.

The SCA enables to retrieve particle local optical depth and particle backscatter coefficient per range bin, assuming a homogeneous filling. It is able to correctly locate the particle layer while the range resolved equation assuming average layer values result in an underestimation of the $LOD_{p,i}$ by at least a factor 2 for the 1st range bin, and the occurrence of a virtual $LOD_{p,i}$ in the range bin immediately below the last range bin containing the actual particle layer. It results in a smearing effect that is significant for range bins on the order of 1 to 2 km. It may have incidental negative feedback on radiative transfer computation even if the rest of the shape is unchanged due to some self-

compensation between layers. The smearing effect is even more pronounced in the case of complex scenes with several particle layers distributed in the vertical and separated by virtually particle-free layers.

The ICA is very similar to the SCA but intends to detect partial filling by checking 7 (or 3) various cases and selecting the one that produces the closest results to observation. It also retrieves local optical depth and backscatter coefficient, taking the filling case into account. The occurrence of a clear range bin under the partially filled bin or array of bins is strictly required to operate the selection. A too large number of cases and a poor SNR in this bin shall complexifies the case selection: several cases may lead to close values and noise may lead the algorithm to select the wrong one.

The information on partial filling and layer position in the range bin may further be useful to the Aeolus wind profile processing in order to flag strong vertical wind shears. These cases are characterized by significant differences between the processed Rayleigh and the Mie channel winds.

The MCA retrieves particle local optical depth but strongly relies on climatology accuracy. The algorithm enables retrieving particle characteristics even if there is no bin matching or if the Rayleigh channel signal is unavailable. Besides, as long as BER input is not too overestimated (so that the logarithm is defined), the MCA gives qualitatively accurate results: layers are well located and their relative thickness is well assessed. Yet, only the exact local backscatter-to-extinction ratio allows an accurate LOD_p retrieval. In case of low $LOD_{p,i}$ the retrieval is not significantly biased by multiple scattering (assuming the multiple scattering coefficient is constant in a range bin). However for large particles and large LOD_p, the impact of multiple scattering effect is probably significant especially in the case of cirrus clouds and desert dusts. This is highly debated in the community ([RD 3], [RD 23]) but because the particle size is not known a priori, the retrieval problem is still complex when multiple scattering occurs.

A feature finding algorithm has been proposed to locate features in each height level, independently from the others. This choice has been made to bypass the sensitivity to noise excursion and the very low probability of finding enough measurements that are similar over the whole column to pack them into one profile. The issue of a heterogeneous upper column to a group has been addressed. An adapted SCA is applied on the groups of bins provided by the Feature Finder to extract feature-wise particle spin-off products.

A scene classification algorithm has been designed to discriminate features between clouds and aerosols in the groups that have been identified by the Feature Finder. It puts forward instrumental data, supported by NWP data, to infer a classification when it can be reliably achieved.