AEOLUS LEVEL-2B ALGORITHM THEORETICAL BASIS DOCUMENT

(MATHEMATICAL DESCRIPTION OF THE AEOLUS LEVEL-2B PROCESSOR)

Authors

Michael RENNIE, David TAN, Erik ANDERSSON
ECMWF

Paul POLI, Alain DABAS
MÉTÉO-FRANCE

Jos De KLOE, Gert-Jan MARSEILLE, Ad STOFFELEN
KNMI

CHANGE LOG

Issue.	Date	New pages	Modified pages (after introducing new pages)	Observations	Name
0.1	25.11.2005	--	--	First draft	Poli, Tan
0.2	22.12.2005	29-31,32-33	16,21-22,23	Further details: centre-of-gravity height referenced to geoid, tripod obscuration correction in Mie processing, error quantifiers.	Tan
0.3	06.07.2006			Revisions to Mie Hlos error (Section 17).	Dabas
0.4	25.07.2006			L1B Screening, Retrieval of optical properties	Marseille
1.0	13.09.2006			Editorials	Tan
2.0	05.12.2006			Main title “Algorithm Theoretical Baseline Document”. Details of Sections 7-10.	Tan de Kloe Stoffelen
2.1	23.12.2007			Description of Matchup. Comments from ESA. Change bars with respect to Issue 2.0.	de Kloe Tan
2.2 draft	03.04.2008	16-18, 44-56		FP scene classification FP optical properties retrieval	Marseille
2.2	09.02.2009		Sections 11,12,13 and A.3. And Section 18.1.1.3. Section 6. Section 19.2.1	Editorials (change equation font size). And a Clarification. (ESA requests, AGS, 15-02-08) Update of Screening tests. Corrections in x/y/z to lat/lon.	Tan
2.3 draft A	08.04.2010	77-81	5, 43-55, 65-76	Added reference documents. Equation numbering, comments from ESA (mail 15/5/2008)	Marseille
2.3 non-draft	06.04.2011		Sections 12, 13 and 14	Updated description and formulae of Rayleigh and Mie processing. Changed equation numbering in section 18 and appendices C, D and E.	Rennie/Tan
2.4	26.11.2012			Many minor modifications. Added a section on the Grouping Algorithm.	Rennie
2.5	01.07.2014			New baseline Mie error quantifier	Rennie

				Review for L2BP v2.10 release	
2.6	17.03.2015			Update references	Rennie
2.7	11.11.2015			Update and cleaning of chapters 7 and 19, new section 7.1, remove Appendix C, update and clean new Appendix D, add references	Marseille
2.8	23.06.2016			A few editorial improvements and updated references	Rennie
2.9	04.10.2016			Added the Range Dependent Bias correction in Sections 13 and 0	Rennie
3.0	06.09.2017			Major update following response to ESA documentation review	Rennie
3.1	31.01.2018			Following ESA review recommendations: • Link inputs/outputs of algorithms to the source of the data in the IODDs • Add Mie core algorithm equations.	Rennie
3.2	14-2-2019	Appendix F		Derivation of the sensitivity of HLOS winds to the Mie Cross-talk (scattering ratio); Appendix numbering now in Table of Contents Cleaning of appendices C, D and E	Marseille
3.3	17-12-2019	Chapter 19		Update of Optical Properties Code algorithm	Marseille
3.31	29-1-2020	Chapters 13, 14		Add placeholders fordescription for M1 mirror temperature wind bias correction	J. de Kloe
3.40	16-7-2020			Clean up, plus remove Appendix C	Marseille

Table of Contents

1. Introduction 6
2. Documents 8
2.1. Applicable documents 8
2.2. Reference documents 8
2.3. Acronyms 9
2.4. Commonly used symbols for variables in this document 10
3. Overview of the Level-2B processing 11
3.1. File types used in relation to the Level-2B processing 11
3.2. Purpose of the Level-2B processing 12
4. High-level architecture of the Level-2B processing 14
5. Main Level-2B processing tasks 17
5.1. Determine reference geolocations for computing auxiliary meteorological data 17
5.2. Generate auxiliary meteorological data at reference geolocations 17
5.3. Pre-process auxiliary meteorological data 17
5.4. Aeolus L2B HLOS wind retrieval 17
6. Input product validation 18
6.1. Screening of Level-1B Data Input Product 18
6.2. Screening of Auxiliary Meteorological Data Input Product 21
6.3. Screening of Rayleigh-Brillouin Calibration Data Input Product 21
6.4. Screening of Calibration Coefficient Data Input Product 21
6.5. Screening of Climatological Data Input Product 22
7. The measurement Grouping Algorithm 23
8. Atmospheric scene classification 29
8.1. An introduction to the Optical Properties Code (OPC) retrieval 30
8.2. *Option 1:* Classification using scattering ratio 31
8.3. *Option 2:* Classification using particle feature detection 35
8.4. *Option 3:* Classification using particle extinction coefficient 36
9. Measurement-bin selection 39
10. Weighting of measurement-bins 41
11. Accumulation of measurement-bin information 42
12. Computation of (H)LOS wind observation geolocation and auxiliary parameters 43
12.1. Computation of the observation's centre-of-gravity geolocation 43
12.2. Observation start and stop geolocations 45
12.3. Computation of a representative L2B observation integration length 45
12.4. Computation of the reference pressure and temperature 46
13. Rayleigh channel (H)LOS wind processing — L2B Rayleigh wind observation retrieval 48
13.1. Computation of a weighted atmospheric Rayleigh response 50
13.2. Computation of a weighted scattering ratio 50
13.3. Computation of the atmospheric return LOS wind component (using the ILIAD scheme) 50
13.4. Computation internal reference and correction terms 51
13.4.1. Computation of a weighted internal reference LOS velocity component 51
13.4.2. Computation of a weighted spacecraft LOS velocity component 51
13.4.3. Ground return corrections 51
13.4.4. Rayleigh range dependent bias correction 52
13.4.5. Rayleigh M1 mirror temperature based wind bias correction 52
13.5. Computation of the Rayleigh channel (H)LOS wind component 52
14. Mie channel (H)LOS wind processing — L2B Mie wind observation retrieval 55
14.1. Computation of weighted Mie spectrometer readouts 57
14.2. Computation of the raw response (peak location) via the Mie core algorithm 57
14.3. Spectral nonlinearity correction 57
14.4. Computation of the atmospheric return LOS velocity component 57
14.5. Computation of post Mie core algorithm correction terms 58
14.5.1. Computation of the weighted internal reference LOS velocity component 58
14.5.2. Computation of a weighted spacecraft LOS velocity component 58
14.5.3. Ground return corrections 58
14.5.4. Mie range dependent bias correction 59
14.5.5. Mie M1 mirror temperature based wind bias correction 59

14.6. Computation of the post Mie core algorithm HLOS wind component 59
15. The ILIAD retrieval − for Rayleigh channel 62
15.1. Rayleigh channel atmospheric response inversion 62
15.2. Internal Reference response inversion 67
16. The Mie core algorithm 69
16.1. Detection Chain Offset subtraction 69
16.2. Telescope Obscuration correction 70
16.3. Background noise correction 70
16.4. Normalise the signal 70
16.5. Find the first guess at the peak location 70
16.6. Iterative optimisation search 70
16.7. Quality control flags 71
16.8. Renormalise peak height and offset 71
17. Rayleigh channel uncertainty estimation processing 72
17.1. Option 1: The Rayleigh HLOS wind error estimate (using output of the ILIAD algorithm) 72
17.2. Option 2: L2B Rayleigh error quantifier based on the L1B processor approach 76
18. Mie channel uncertainty estimation processing 78
18.1. Option 1: L2B Mie error quantifier based on L1B approach 78
18.2. Option 2: L2B Mie error quantifier using outputs of the Mie core algorithm (Tellus paper version) 79
18.3. Option 3: L2B Mie error quantifier based on least squares fit of an integrated Lorentzian and using outputs of the Mie core algorithm 81
19. Optical Properties Code: Retrieval of particulate optical properties from the Fabry-Pérot signal only 87
19.1. Estimation of the calibration factor 88
19.2. Particle layer detection 89
19.3. Particle layer optical properties 89
19.3.1. Simplified fast algorithm 92
19.4. Height dependent scattering ratio threshold 92
20. Matchup algorithm between AUX_MET profiles and L1B measurements 94
20.1. Dummy Matchup algorithm 95
20.2. Nearest Neighbour Matchup algorithm 95
20.2.1. Select the L1B measurement-bin geolocation 95
20.2.2. Apply time-window 95
20.2.3. Apply a distance range 96
20.2.4. Choose closest NWP profile 96
21. Potential future algorithm developments 97
Appendix A. Auxiliary meteorological data processing 98
Appendix B. Mie error quantifier derivation 101
Appendix C. Particle transmission computation from the Fabry-Pérot channel signal 104
Appendix D. Retrieval of atmosphere molecular backscatter and extinction (Rayleigh scattering) 111
Appendix E. Derivation of the sensitivity of HLOS wind to Mie contamination 113
Appendix F. Overview of AUX_PAR_2B settings and the control of the L2B processor algorithms 115

1. Introduction

This document is the Algorithm Theoretical Basis Document (ATBD) for the ADM-Aeolus Level-2B Processor. It has evolved over the years from Technical Note 2.3 (issue 1.0) of the "Development and Production of Aeolus Wind Data Products" contract study whose contents are fully described in [AD2]. It provides a mathematical description of the algorithms used for the Aeolus Level-2B (L2B) processing with information on where the input/output data are sourced/stored and details various processing options.

The nature of the L2B products and high-level architecture of the L2B processing is described in Mission Requirements Document [RD20]. File formats for L2B/C products and associated auxiliary data are described in [RD10].

The following diagram is provided to aid the explanation of the Level-2B processor algorithms.

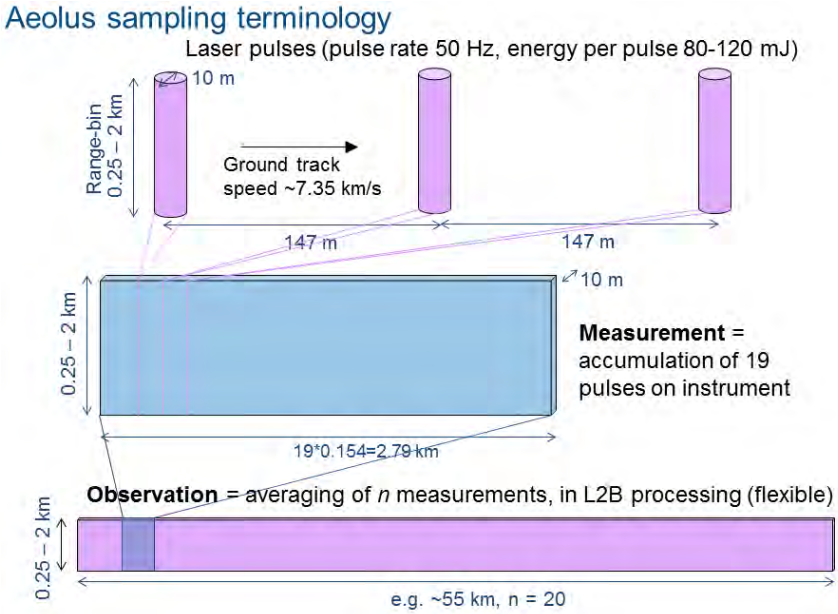


Figure 1. Diagram of Aeolus sampling terminology to aid in the description of the Level-2B processor

The Level-2B processor produces Line-of-sight (LOS) or Horizontal LOS (HLOS) wind components (a settings parameter) suitable for use in Numerical Weather Prediction (NWP) and atmospheric scientific research. The main inputs are Level-1B measurement data and calibration products. The Level-2B processor produces Aeolus wind products with a number of corrections beyond those provided in the Level-1B wind mode product:

- Measurements are gathered together into groups according to parameter settings to allow flexibility in the horizontal resolution and the noise characteristics of the L2B observations;
- Measurement-bins are classified into different meteorological categories (clear or cloudy are the current options) using atmospheric optical properties information (the default is to use L1B scattering ratio estimates); with the main aim of avoiding significant Mie signal (particulate) contamination of the Rayleigh (molecular) channel;
- Accumulation of the grouped and classified measurement-bin spectrometer counts prior to wind retrieval to obtain an appropriate resolution and noise level;
- (H)LOS wind retrievals are performed with the grouped and classified measurement spectrometer counts;
- Separate (H)LOS wind retrievals are performed for the Mie and Rayleigh channels resulting in (currently) four different types of (H)LOS winds: Rayleigh-clear, Rayleigh-cloudy, Mie-clear and Mie-cloudy. N.B. Mie-clear is produced because the classification process is imperfect due to noise in the optical properties information, meaning that cloudy measurements can be incorrectly assigned as clear;

- Rayleigh winds are corrected for the effects of temperature and pressure i.e. Rayleigh-Brillouin Correction using a priori temperature and pressure information (AUX_MET) in combination with an appropriate calibration look-up table (AUX_RBC). Without this correction 10 m/s HLOS wind biases can occur;
- Rayleigh winds are also corrected for Mie contamination (cross-talk) using optical properties information (the default is to use L1B scattering ratio estimates);
- Uncertainty estimates (wind component random error estimates) and quality flags for each wind observation;
- L2B wind observations are essentially independent – however profiles are also defined via lists of observation indices;
- The processing algorithm choices are controlled by an external settings file (AUX_PAR_2B)

2. Documents

2.1. Applicable documents

Ref	Document title	Document ref	Ver	Date
[AD1]	Statement of Work of the Development and Production of L2B/C Aeolus data.	AE-SW-ESA-GS-0117	1B	Sep 2004
[AD2]	Answer to RFQ/3-11094/04/NL/MM.			Jul 2004
[AD3]	ADM-Aeolus PDGS - L2/MetPF Interface Control Document	XADM-GSEG-EOPG-ID-04-0002	1.14	08/03/2017
[AD4]	Implementation of Level 2B/2C Processing Facility. Technical Requirements	XADM-GSEG-EOPG-RD-04-0003	2.0	Apr 2015
[AD5]	Earth Explorer Ground Segment File Format Standard	PE-TN-ESA-GS-001	1.4	Jun 2003

2.2. Reference documents

Ref	Document title	Document ref	Ver	Date
[RD1]	Aladin Instrument Operation Definition	AE-TN-ASF-AL-00044	3.0	Oct 2004
[RD2]	Level 1b Processor Detailed Processing Model	ADM-MA-52-1800	3/07	21-Jul-2017
[RD3]	Level1B Master Algorithm Document	AE-SW-ASU-GS-023	9.0	March 2017
[RD4]	Impact of line shape on wind measurements and correction methods	AE-SW-ESA-GS-016	1.0	Feb 2004
[RD5]	L1B & E2S Input/Output Data Definitions Interface Control Document	ADM-IC-52-1666	4/07	July 2017
[RD6]	Selection of L2B Parameters (Study TN2.1)	AE-TN-MFG-L2P-0021	1.2	Sep 2005
[RD7]	Aeolus Data Products Contents Guidelines	AE-TN-ESA-SY-007	1B	May 2004
[RD8]	Definition of Baseline Aeolus Level-2B Processing (Study TN2.2)	AE-TN-ECMWF-L2P-0022	1.2	Sep 2005
[RD9]	Level 1b Processor Detailed Processing Model	ADM-MA-52-1800	3/06	01 Jun 2016
[RD10]	ADM-Aeolus Level-2B Processor Input/Output Data Definitions Interface Control Document	AE_IF_ECMWF-L2BP-001	3.00	August 2017
[RD11]	Aeolus Product Modification for Implementation in L1bP CodeV2	AE.TN.DLR.APM-L1B.150206	1.3	Feb 2006
[RD12]	Calibration constants for the retrieval of optical properties of the atmosphere	AE-TN-MFG-L2A-CAL	1.0	Mar 2006
[RD13]	Raman Lidar Monitoring of Extinction and Backscattering of African Dust Layers and Dust Characterization	*Appl. Optics*, vol. 42, 9, pp 199-1709, 2003 by De Tomassi, F., Blanco, A., Perrone, M.R.,		2003
[RD14]	Simulation of wind profiles from a space-borne Doppler wind lidar	Q.J.R. Meteorol. Soc., 129, pp. 3079-3098, 2003, by Marseille, G.J. and Stoffelen, A		2003
[RD15]	Lidar sensing of aerosols and clouds in the troposphere and stratosphere	Proc. IEEE, 77, 433-448, 1989 by Reagan, J.A., McCormick, M.P., Spinhirne, J.D.,		1989
[RD16]	L2A PRODUCT ALGORITHM THEORETICAL BASIS DOCUMENT (ATBD)	AE-TN-IPSL-GS-001	5.5	Jan 2017
[RD17]	Performance of Aeolus in heterogeneous atmospheric conditions using high-resolution radiosonde data	Atmos. Meas. Tech., 7, pp. 2695-2717, doi:10.5194/amt-7-2695-2014	-	2014
[RD18]	The L2/Met PF: Aeolus Level-2B/C processing at ECMWF	AE-TN-ECMWF-L2BP-0091	2.0	April 2017
[RD19]	AEOLUS LEVEL-2B PROCESSOR SOFTWARE USER'S MANUAL	AE-MA-ECMWF-L2BP-001	1.80	August 2017
[RD20]	ADM-Aeolus Mission Requirements Document	AE-RP-ESA-SY-001	2.0	

2.3. Acronyms

ACCD	Accumulation Charge Coupled Device
ADS	Annotation Data Set
AMD	Auxiliary Meteorological Data
AOCS	Attitude and Orbit Control System
BRC	Basic Repeat Cycle
BUFR	Binary Universal Form for the Representation of meteorological data
DEM	Digital Elevation Model
E2S	End-to-End Simulator
ECMWF	European Centre for Medium-Range Weather Forecast
EE	Earth Explorer
EGM96	Earth Gravitational Model 1996
FSR	Free spectral range
FWHM	Full width half maximum
GADS	Global Annotation Data Sets
HLOS	Horizontal Line Of Sight
ILIAD	IMPACT of LINE SHAPE on AEOLUS-ADM DOPPLER ESTIMATES
LOS	Line of sight
L1B	Level-1B
L2B	Level-2B
L2C	Level-2C
L2/Met PF	Level-2 Meteorological Processing Facility (hosted by ECMWF)
MDS	Measurement Data Set
MPH	Main Product Header
MRC	Mie response calibration
N/A	Not applicable
NRT	Near Real Time
NWP	Numerical Weather Prediction
ODB	Observation DataBase (ECMWF software)
OPC	Optical Properties Code
PCD	Product Confidence Data
PDGS	Payload Data Ground Segment
PRF	Pulse Repetition Frequency
QRT	Quasi Real Time
RBC	Rayleigh Brillouin Correction
RDB	Range Dependent Bias
RRC	Rayleigh response calibration
SNR	Signal to Noise Ratio
SPH	Specific Product Header
SR	Scattering Ratio
TBD	To be determined
USR	Useful spectral range
WGS84	World Global System 84: Reference Ellipsoid for GPS data
WMO	World Meteorological Organization
WVM	Wind Velocity Measurement
ZWC	Zero Wind Correction

2.4. Commonly used symbols for variables in this document

Symbol	Meaning
i	Range-bin number
k	Measurement number
(i,k)	Index for measurement-bin
W	Weights for measurement-bins
j	Index for ACCD pixel
N	Number of measurement-bins used in an observation
λ_0	Laser wavelength
v	(H)LOS wind velocity component
V	LOS velocity component (e.g. for satellite-earth velocity correction)
α_p	Particle extinction coefficient
p	Atmospheric pressure
T	Atmospheric temperature
W	Scattering ratio
z	Altitude
β_p	Particle backscatter coefficient
β_m	Molecular backscatter coefficient
β_p	Particle backscatter coefficient
α_m	Molecular extinction coefficient
α_p	Particle extinction coefficient
τ_m	Molecular transmission
τ_p	Particle transmission

3. Overview of the Level-2B processing

Details on how to use the software can be obtained and installed are available in the SUM [RD19]. Here with provide a brief introduction to the L2B processor inputs/outputs for the benefit of this document.

3.1. File types used in relation to the Level-2B processing

A variety of files are involved in the input/output of the L2B processor. The following table lists files that are referred to within this document.

Table 1. List of file types relevant to the L2B processing that are referred to in the ATBD

File type	Short-hand name	Input/Output to L2B processor	Purpose
Aeolus Level-1B WVM Product	ALD_U_N_1B	I	Main input to L2B processor; e.g. spectrometer counts, geolocation info, MRC info, error estimates etc.
Predicted Ground Track File	GRND_TRACK	N/A	Contains one week of predicted orbit ground track geolocations (nadir and off-nadir). For use in L2/Met PF AUX_MET generation.
Aeolus Collocated Meteorological Product	AUX_MET_12	I	Necessary input to L2B processing. Provides temperature and pressure information needed in Rayleigh-Brillouin correction (and optionally used by the Optical Properties Code)
Rayleigh-Brillouin Correction Table	AUX_RBC_L2	I	Calibration look-up table input to L2B processor to enable correction of Rayleigh winds for temperature and pressure effects
Climatology Lookup Table	AUX_CLM_L2	I	Optical Properties Code retrievals in the future may use this file (provides climatology of lidar ratio)
L2 Calibration Coefficient Auxiliary File	AUX_CAL_L2	I	The Optical Properties Code can optionally use this calibration dataset
Level 2B Processing Parameters	AUX_PAR_2B	I	Contains settings and parameters for the L2B processing i.e. controls which options are used in the algorithms. A necessary input to the L2B processor. See [Appendix F. Overview of AUX_PAR_2B settings and the control of the L2B processor algorithms] for the options.
Aeolus Level-	ALD_U_N_2B	O	Output product of L2B

2B Product			processor. Meteorologically representative (H)LOS wind observations for use in data assimilation/scientific research.
Aeolus Level-2C Product	ALD_U_N_2C	N/A. This is produced by the L2.Met PF only	Aeolus assisted wind vector product (based on the L2B product plus information from the ECMWF analysis).

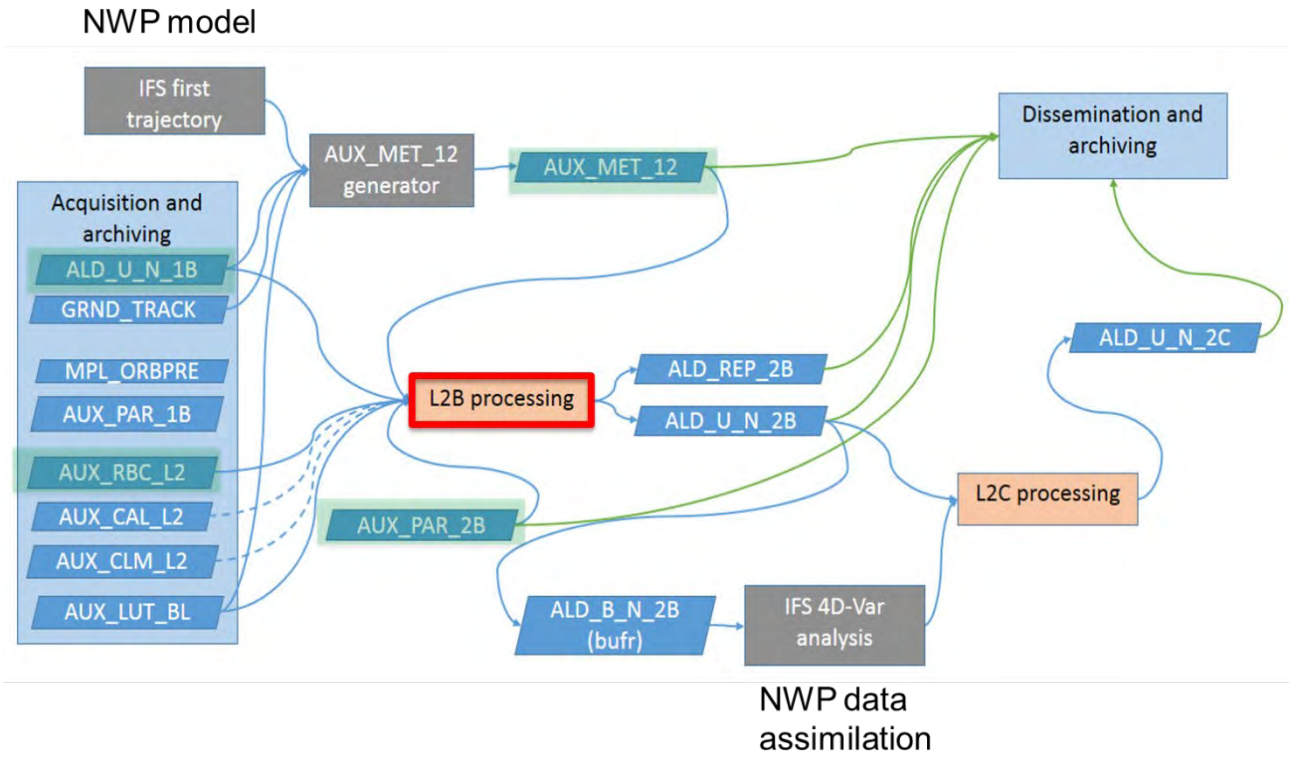


Figure 2. Diagram showing the inputs and outputs of the L2B processor as implemented in the L2/Met PF

3.2. Purpose of the Level-2B processing

The Level-2B processing of Level-1B Wind Velocity Measurement (WVM) data (referred to as ALD_U_N_1B in this document) is required for the generation of L2B data products (ALD_U_N_2B). L2B products are intermediate between L1B and L2C products.

The Level-2B processing HLOS wind retrieval involves taking L1B products (ALD_U_N_1B), auxiliary input parameters (AUX_PAR_2B), auxiliary calibration information (AUX_RBC_L2 and optionally AUX_CAL_L2 and AUX_CLM_L2) and auxiliary meteorological data (AUX_MET_12), and applying essential modifications, corrections and additions to generate a product of HLOS wind component observations which are suitable for data assimilation by the NWP systems and for more general scientific research. The L2B processor works on the principal of summing the signals (spectrometer counts) from Aeolus measurement-bins (effectively horizontal averaging) after classification into different atmospheric conditions and performing a wind retrieval on the integrated signals, rather than performing retrievals on individual measurements and then averaging the wind results.

A key input to the L2B processing HLOS wind retrieval is the auxiliary meteorological data file. Auxiliary Meteorological data (AUX_MET_12[1]) generation involves the generation/acquisition of auxiliary meteorological data as required as an input for the Level-2B Rayleigh HLOS wind retrieval. Information from the AUX_MET_12 file can also be applied in the optional L2BP Optical Properties Code retrieval algorithm and is also used as part of the Aeolus Calibration software and the L2A processing which is operated by the PDGS (Payload Data Ground Segments). Note that the Level-2B Mie wind retrieval does not require AUX_MET_12 data. Although AUX_MET_12 will have an indirect effect on the Mie results if the Optical Properties Code is applied in the scene classification algorithm. The auxiliary meteorological data take the form of *a priori* estimates of the atmospheric temperature and pressure (and some other variables not needed by the L2B processor e.g. wind, humidity, cloud ice/water content) geolocated to be close to the Aeolus L1B measurement data. 0 describes Auxiliary Meteorological Data generation as applicable in the operational mode (L2/Met PF).

During the Aeolus mission the Level-2B processing is envisaged to be operated in four environments, which are:

- **The operational environment:** Level-2B products to be produced in NRT by the L2/Met PF (part of the Aeolus Ground Segment) hosted by ECWMF
- **The operational re-processing environment:** reprocessed Level-2B products to be produced by the PDGS as necessary during the mission
- **The NWP centre environment**: NWP centres may wish to run their own instance of the L2B processor (and possibly generate their own AUX_MET_12) to produce products for their own data assimilation systems
- **The non-operational scientific research environment:** for more general scientific investigations e.g. CAL/VAL investigations

The operational environment will be the main source of Level-2B data (or L2B BUFR file format data derived from this) for NWP users. Operational re-processing will be done occasionally (as necessary) when the Ground Segments processors undergo improvements.

The L2B processing algorithms described in this document are applied in the same way in all environments.

Because Aeolus is an Earth Explorer demonstration mission the L2BP algorithms will probably undergo significant modifications during the lifetime of the mission. This version of the ATBD is in line with version 3.00 of the L2B processor (v3.00).

[1] The “12” in the file name refers to Level-1, Level-2 rather than the validity period of the file (which used to be 12 hours). Actually AUX_MET_12 files produced in the L2/Met PF will nominally have a validity period of 30 hours.

4. High-level architecture of the Level-2B processing

The high-level architecture of how to generate L2B products is shown in Figures 1 and 2. Figure 1 shows the main processing tasks for the generation AUX_MET_12 data and how that links to the HLOS wind retrieval, which is shown in Figure 2. The various steps of the L2B processor are indicated in Figure 2.

Different software in the operational L2/Met PF and possible NWP centre implementations is confined to differences in how the AUX_MET_12 is generated (Figure 1). The L2B processing (Figure 2) part will be the same wherever the processor is run (apart from different processing settings choices possible via the AUX_PAR_2B file).

Aspects of Figure 1 that are not expected to apply in non-operational NWP centre mode are a) the sending of data to the PDGS, and b) the use of predicted orbit tracks (GRND_TRACK files). NWP centres are free to modify the other aspects of auxiliary meteorological data processing to suit their own requirements (and also to modify/improve the L2B processing algorithms if they wish - however they should inform the L2B development team of any such improvements).

The dashed box in Figure 1 is not applicable to the PDGS re-processing mode. Instead, auxiliary meteorological data at reference geolocations (before interpolation to locations requested by the HLOS wind retrieval algorithm) are to be obtained from the archive within the PDGS. The re-processing mode will apply auxiliary pre-processing as described in Appendix A.3, consistent with the operational mode.

A high-level description of the main processing tasks is given in Section 5.

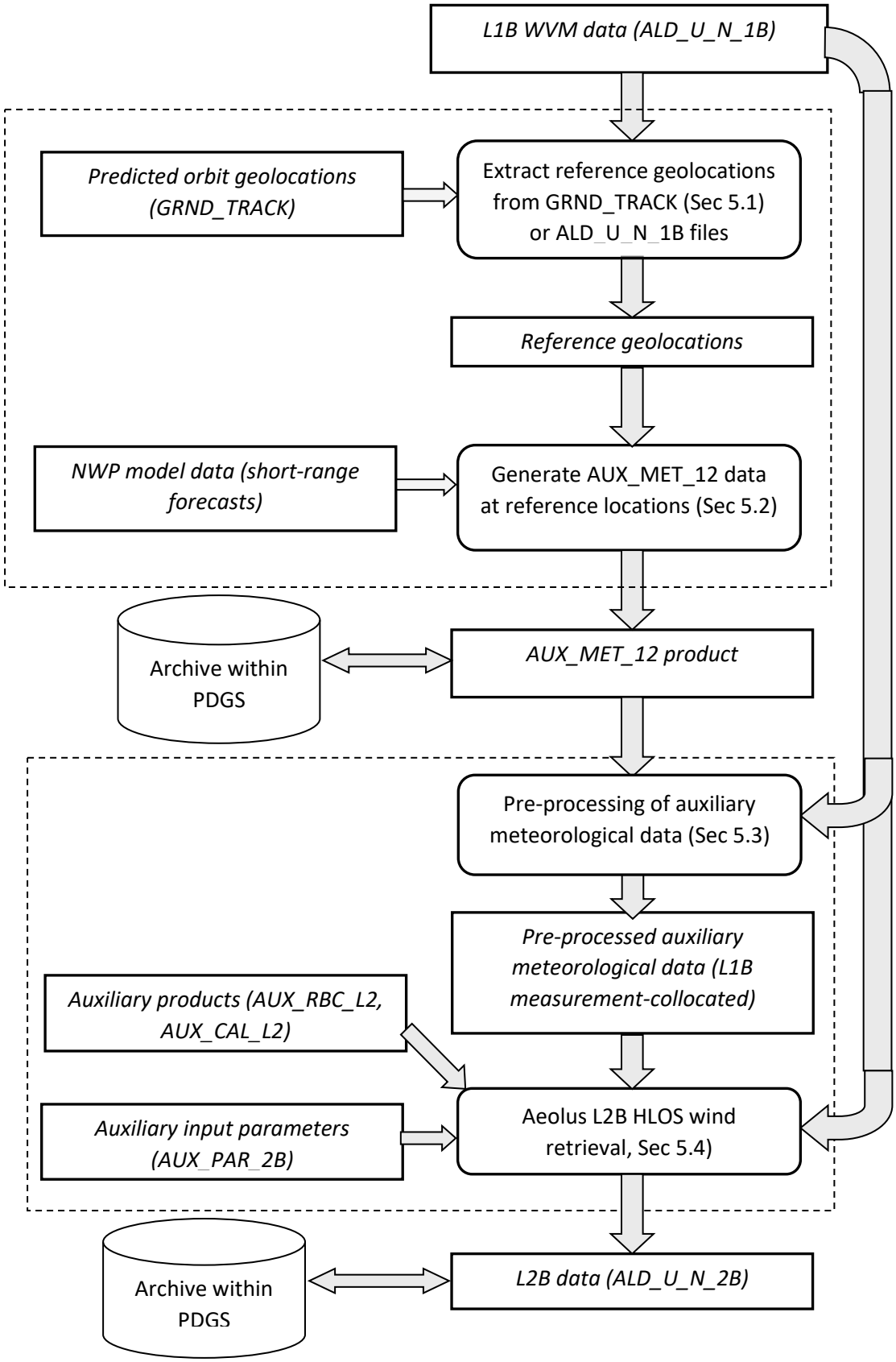


Figure 1: Data flow for Aeolus Level-2B processing. All the steps before HLOS wind retrieval comprise AUX_MET_12 data generation steps in operational mode. Re-processing mode does not need to regenerate AUX_MET_12 data. NWP centre mode may run their own AUX_MET_12 generation. Data flow for primary Level-2B processing (i.e. HLOS retrieval) is given in Figure 2.

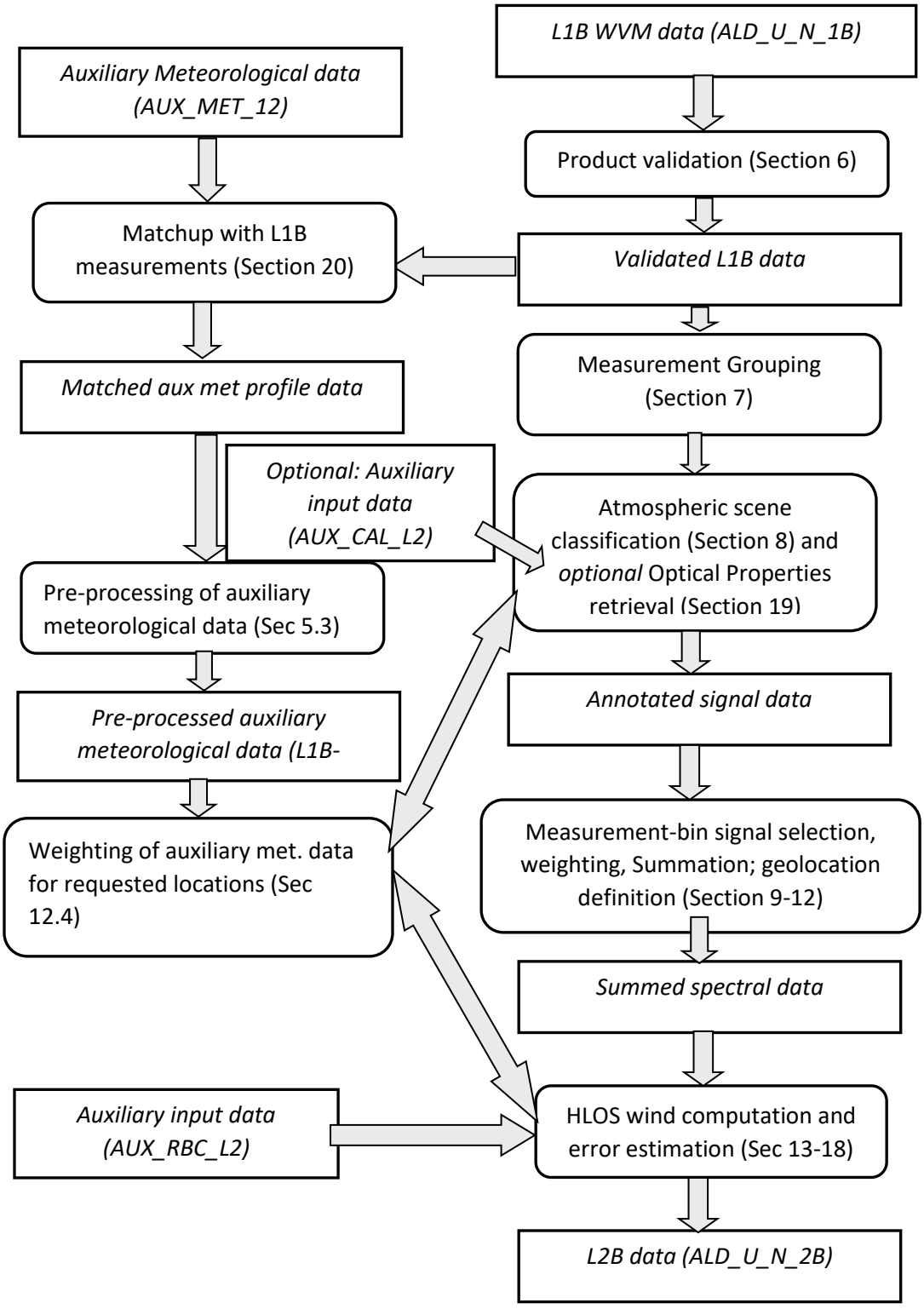


Figure 2: Data flow for Aeolus Level-2B HLOS wind retrieval.

5. Main Level-2B processing tasks

Mathematical descriptions of the main processing tasks are given in Sections 6 to 18. To show the correspondence of these algorithms with the high-level architecture provided in the previous Section and with the rest of the document we briefly discuss them in the task headings below.

AUX_MET_12 data is a compulsory input the L2B processing and hence it much be generated or obtained from the PDGS archive.

5.1. Determine reference geolocations for computing auxiliary meteorological data

This is part of auxiliary meteorological data processing and is described in the Appendix A.1.

Note: determining reference geolocations is a pre-processing step performed outside the L2BP processing code. The L2BP software package provides various tools which can help in implementing this step at NWP centres. For example, there are tools to read the L1B file and write out the geolocation information needed to help generate appropriate auxiliary meteorological data. There is also a tool for reading the L1B files (and if required the GRND_TRACK data) and extracting the geolocation information to the ECMWF Observation Database 2 (ODB-2) format — again this may be adapted by NWP centres for their own formats.

5.2. Generate auxiliary meteorological data at reference geolocations

This is part of auxiliary processing and is described in the Appendix A.2.

5.3. Pre-process auxiliary meteorological data

This is part of auxiliary processing and is described in the Appendix A.3.

5.4. Aeolus L2B HLOS wind retrieval

The steps of primary L2B processing are described in separate Sections:

- Input product validation, Section 6
- Measurement grouping, Section 7
- Atmospheric scene classification, Section 8
- Measurement-bin selection, Section 9
- Weighting of measurement-bins, Section 10
- Accumulation of measurement-bin data, Section 11
- Computation of L2B (H)HLOS wind observation geolocation and relevant parameters, Section 12
  - centre-of-gravity geolocation
  - Start/stop geolocations
  - representative integration length
  - Weighting of auxiliary meteorological data (reference pressure and temperature) for selected measurement-bins
- Computations of L2B variables: (H)LOS wind computation and uncertainty estimation, comprising
  - Rayleigh channel (H)LOS wind processing, Section 13 and 15
  - Mie channel (H)LOS wind processing, Section 14 and 16
  - Rayleigh channel (H)LOS wind uncertainty estimation processing, Section 17
  - Mie channel (H)LOS wind uncertainty estimation processing, Section 18
- Optical Properties Code retrieval using only the Rayleigh channel, Section 19
- AUX_MET data matchup algorithm, Section 20
- Potential future algorithm updates, Section 21

6. Input product validation

Any dynamic files that are read by the L2B processor undergo some basic screening checks to ensure the values are within expected ranges. The L2B processor produces warnings regarding unexpected values and potentially may not use the data in the HLOS wind retrieval depending on the context. Note that all input files are rigorously checked for the correct file format (according to their IODDs) and for any missing or infinite values.

Product validation comprises:

- Screening of Level-1B WVM Data Input Product (ALD_U_N_1B)
- Screening of Auxiliary Meteorological Data Input Product (AUX_MET_12)
- Screening of Rayleigh-Brillouin Calibration Data Input Product (AUX_RBC_L2)
- Screening of Calibration Coefficient Data Input Product (AUX_CAL_L2)

6.1. Screening of Level-1B Data Input Product

In the L1B-screening procedure L1B input variables of relevance to the L2B processing are checked for their validity i.e. whether the values meet our expectation. The L2B screening results indicate whether the L1B data have passed or failed the screening procedure. The results are stored in the L2B product (see [RD10]) as part of:

`L2B Meas Product Confidence Data (PCD) ADS`

- `L1B_Input_Screening`:
  - `L1B_Obs_Scr`
    - `Obs_Screening`
      For screening of L1B PCD data on the observation scale. N.B. observation scale L1B data is not used in the L2B HLOS wind retrieval; this screening has no effect on the HLOS wind retrieval.
  - `Mie_Meas`
    - `Meas_QC`
      For screening of L1B PCD Mie data on the measurement profile scale i.e. applicable to all range-bins in the measurement. N.B. the measurements are still used in the HLOS wind retrieval if problems are detected (warnings in the standard output will be given however)
    - `Bin_Screening`
      - `Bin_QC`
        For screening of L1B PCD Mie data on measurement range-bin scale. Measurement range-bins with problems as indicated by the `Bin_QC` are discarded from the HLOS wind retrieval (and warnings are given in the standard output)
  - `Rayleigh_Meas`
    - `Meas_QC`
      For screening of L1B PCD Rayleigh data on the measurement profile scale i.e. applicable to all range-bins in the measurement. N.B. the measurements are still used in the HLOS wind retrieval if problems are detected (warnings in the standard output will be given however).
    - `Bin_Screening`
      - `Bin_QC`
        For screening of L1B PCD Rayleigh data on measurement range-bin scale. Measurement range-bins with problems as indicated by the `Bin_QC` are discarded from the HLOS wind retrieval (and warnings are given in the standard output).

The values of the L2B screening output variables are determined by the values of the L1B variables and the threshold parameters defined in the AUX_PAR_2B. The AUX_PAR_2B threshold variables are obtained from the L2B Processing parameter GADS:

`Screening_Params`

- `L1B_Screening_Params`

- L1B_Obs_Screening_Params
- L1B_Mie_Meas_Screening_Params
- L1B_Rayleigh_Meas_Screening_Params

See [Appendix F. Overview of AUX_PAR_2B settings and the control of the L2B processor algorithms] for a list of the relevant AUX_PAR_2B parameters available.

In case of a missing variable or failure of the threshold test, the failure is reported and the L2B screening output variable is flagged. A threshold failure is reported when the L1B value is *larger than* a pre-defined upper threshold value or *smaller than* a pre-defined lower threshold value. The L1B input variables and corresponding threshold parameters are described in the following tables.

Table 2. Screening on the L1B observation scale. The L1B variables correspond to those in Table 2-20 of the L1B-IODD [RD11]. Thresholds given by a symbolic name refer to parameters in the AUX_PAR_2B processing parameters file. In case of a screening failure, the numerical value of the L2B variable in L1B_Obs_Screening (See table 17 of [RD10]) is set to the corresponding value in the first column of Table 18 of [RD10].

L1B variable	Lower threshold parameter	Upper threshold parameter
Laser related:		
L1B_Laser_Freq_Unlocked	0	L1B_Laser_Freq_Unlocked_Thresh
L1B_Ref_Pulses_Unlocked	0	L1B_Ref_Pulses_Unlocked_Thresh
L1B_Laser_Freq_Offset	0	L1B_Laser_Freq_Offset_Thresh
L1B_Laser_UV_Energy	0	L1B_Laser_UV_Energy_Thresh
L1B_ Laser_Freq_Offs_Stdev	0	L1B_Laser_Freq_Offs_StdevThresh
L1B_ Laser_UV_Energy_Stdev	0	L1B_Laser_UV_Energy_StdevThresh
L1B_Mie_Mean_Emit_Freq	L1B_Mie_Mean_Emit_Freq_Min	L1B_Mie_Mean_Emit_Freq_Max
L1B_ Mie _Emit_Freq_Stdev	0	L1B_Mie_Emit_Freq_Stdev_Thresh
L1B_Rayl_Mean_Emit_Freq	L1B_Rayleigh_Mean_Emit_Freq_Min	L1B_Rayleigh_Mean_Emit_Freq_Max
L1B_ Rayl _Emit_Freq_Stdev	0	L1B_RayleighEmitFreqStdevThresh
Satellite related		
L1B_Sat_Not_on_Target	0	L1B_Sat_Not_on_Target_Thresh
Measurement related		
L1B_Mie_Invalid_Meas	0	L1B_Mie_Corrupt_Thresh
L1B_Mie_Invalid_Ref_Pulses	0	L1B_Mie_Ref_PulsesCorruptThresh
L1B_Rayl_Invalid_Meas	0	L1B_Rayleigh_Corrupt
L1B_Rayl_Invalid_Ref_Pulses	0	L1B_Rayl_RefPulsesCorruptThresh
Combinations		
L1B_Mie_Invalid_Meas	0	L1B_Mie_Invalid_Meas_Thresh
L1B_Mie_ Invalid_Ref_Pulses	0	L1B_Mie_Invalid_RefPulsesThresh
L1B_Rayl_Invalid_Meas	0	L1B_Rayl_Invalid_Meas_Thresh
L1B_Rayl_ Invalid_Ref_Pulses	0	L1B_Rayl_InvalidRefPulsesThresh
Nr_Invalid_Mie_Peaks	0	N_range_gates

Table 3. Screening of L1B Mie data on measurement profile scale. The L1B variables correspond to those in Table 2-25 of the L1B-IODD [RD11]. Thresholds given by a symbolic name refer to parameters in the AUX_PAR_2B processing parameters file. In case of a screening failure, the numerical value of the L2B variable in L1B_Mie_Meas_QC (See table 19 of [RD10]) is set to the corresponding value in the first column of Table 20.

L1B variable	Lower threshold parameter	Upper threshold parameter
L1B_Mie_Meas_Invalid_Ref_Pulses	0	L1B_MieMeasInvalidRefPlssThresh
L1B_Avg_Laser_Freq_Offset	L1B_Avg_Laser_Freq_Offset_Min	L1B_Avg_Laser_Freq_Offset_Max
L1B_Avg_UV_Energy	L1B_Avg_UV_Energy_Min	L1B_Avg_UV_Energy_Max
L1B_Laser_Freq_Offset_Stdev	0	L1B_Laser_FreqOffsetStdevThresh
L1B_UV_Energy_Std_Dev	0	L1B_UV_Energy_Stdev_Thresh
L1B_Vel_of_Att_Uncertainty_Error	L1B_VelofAttUncertaintyErrorMin	L1B_VelofAttUncertaintyErrorMax
L1B_Mie_Mean_Emitted_Freq	L1B_Mie_Mean_Emitted_Freq_Min	L1B_Mie_Mean_Emitted_Freq_Max
L1B_Mie_Emitted_Freq_Stdev	This parameter is no longer screened (removed in L1B v1.3)	

L1B_Reference_Pulse_FWHM	New parameter added in L1B v1.3; no thresholds defined yet in the code.

Table 4. Screening of L1B Mie data on measurement-bin scale (see Chapter 7 for measurement-bin definition). The L1B variables correspond to those in Table 2-27 of the L1B-IODD [RD11]. Thresholds given by a symbolic name refer to parameters in the AUX_PAR_2B processing parameters file. In case of a screening failure the numerical value of the L2B variable in L1B_Mie_Meas_QC (See table 21 of [RD10]) is set to the corresponding value in the first column of Table 22. The logical variable has been set by L1BP and is either true (T) or false (F), but note that this screening test may be disabled until further notice.

L1B variable	Lower threshold parameter	Upper threshold parameter	Logical
L1B_Mie_Bin_Invalid	-	-	T/F
L1B_Scattering_Ratio	L1B_ScatteringRatio_Min	L1B_ScatteringRatio_Max	-
L1B_Mie_SNR	L1B_Mie_SNR_Thresh	No upper threshold	-

N.B. The L1B_Mie_Bin_Invalid results can be ignored by setting `Ignore_Mie_Meas_Invalid_Switch` to be `True` in the AUX_PAR_2B. See [Appendix E. Derivation of the sensitivity of HLOS wind to Mie contamination] for a list of the relevant AUX_PAR_2B parameters available.

Table 5. Screening of L1B Rayleigh data on measurement profile scale. The L1B variables correspond to those in Table 2-25 of the L1B-IODD [RD11]. Thresholds given by a symbolic name refer to parameters in the AUX_PAR_2B processing parameters file. In case of a screening failure, the numerical value of the L2B variable in L1B_Rayleigh_Meas_QC (See table 23 of [RD10]) is set to the corresponding value in the first column of Table 24.

L1B variable	Lower threshold parameter	Upper threshold parameter
L1B_Rayl_Meas_Invalid_Ref_Pulses	0	L1B_RayMeasInvalidRefPlssThresh
L1B_Avg_Laser_Freq_Offset	L1B_Avg_Laser_Freq_Offset_Min	L1B_Avg_Laser_Freq_Offset_Max
L1B_Avg_UV_Energy	L1B_Avg_UV_Energy_Min	L1B_Avg_UV_Energy_Max
L1B_Laser_Freq_Offset_Stdev	0	L1B_Laser_FreqOffsetStdevThresh
L1B_UV_Energy_Std_Dev	0	L1B_UV_Energy_Stdev_Thresh
L1B_Vel_of_Att_Uncertainty_Error	L1B_VelofAttUncertaintyErrorMin	L1B_VelofAttUncertaintyErrorMax
L1B_Rayleigh_Mean_Emitted_Freq	L1B_Rayl_Mean_Emitted_Freq_Min	L1B_Rayl_Mean_Emitted_Freq_Max
L1B_Rayleigh_Emitted_Freq_Stdev	0	L1B_Rayl_EmittedFreqStdevThresh

Table 6. Screening of L1B Rayleigh data on measurement-bin scale (see Chapter 7 for definition of measurement-bin). The L1B variables correspond to those in Table 2-27 of the L1B-IODD [RD11]. Thresholds given by a symbolic name refer to parameters in the AUX_PAR_2B processing parameters file. In case of a screening failure the numerical value of the L2B variable in L1B_Rayleigh_Bin_QC (See table 25 of [RD10]) is set to the corresponding value in the first column of Table 26. The logical variable has been set by L1BP and is either true (T) or false (F), but note that this screening test may be disabled until further notice.

L1B variable	Lower threshold parameter	Upper threshold parameter	Logical
L1B_Rayleigh_Bin_Invalid	-	-	T/F
L1B_Rayleigh_SNR_A	L1B_Rayleigh_SNR_Min	L1B_Rayleigh_SNR_Max	-
L1B_Rayleigh_SNR_B	L1B_Rayleigh_SNR_Min	L1B_Rayleigh_SNR_Max	-

N.B. The L1B_Rayleigh_Bin_Invalid results can be ignored by setting `Ignore_Rayleigh_Meas_Invalid_Switch` to be `True` in the AUX_PAR_2B. See [Appendix E. Derivation of the sensitivity of HLOS wind to Mie contamination] for a list of the relevant AUX_PAR_2B parameters available.

6.2. Screening of Auxiliary Meteorological Data Input Product

Screening is performed on a selection of the auxiliary meteorological data variables against the maximum and minimum allowed values. Only temperature, pressure and the associated geometric altitudes are used in the HLOS wind retrieval; hence only temperature and pressure are checked. The whole AMD vertical profile is flagged as invalid if one or more of the temperature or pressure values within the profile is outside the expected physical range and a warning message is given. Currently the AUX_MET_12 data profile can still be used in the HLOS wind retrieval, even if flagged invalid.

The variables checked from the AUX_MET_12 file format (see Table 81 of [RD10]) are:

`Meteorological MDS`

- `List_of_Profile_Data`
  - `AMD_pbase`
  - `AMD_T`

The threshold variables for the screening come from the L2B Processing parameter (AUX_PAR_2B) GADS:

`Screening_Params`

- `L2B_AMD_Screening_Params`

See [Appendix F. Overview of AUX_PAR_2B settings and the control of the L2B processor algorithms] for a list and for more details on the AUX_PAR_2B parameters.

The results of screening AMD variables are stored in the L2B product (see [RD10]) as part of the:

`L2B AMD Product Confidence Data (PCD) ADS`

- `L2B_AMD_Screening`
  - `L2B_AMD_Screening_QC:` set to `L2B_AMD_Unlikely_Profile` if any pressure or temperature values are outside the expected range
  - `L2B_AMD_Screening_QC_Flags:` Bit flag indicating the type of problem detected e.g. `AMD_T_out_of_bounds_nadir`

Table 7. Screening of AUX_MET_12 data. The AUX_MET variables can be found in Table 81 of [RD10]. The thresholds are part of the AUX_PAR_2B; see Table 136 of [RD10].

AUX_MET variable	Lower threshold parameter	Upper threshold parameter
AMD_pbase (Pressure at bottom of model layer)	L2B_AMD_p_min	L2B_AMD_p_max
AMD_T (Temperature at AMD_pnom)	L2B_AMD_T_min	L2B_AMD_T_max

6.3. Screening of Rayleigh-Brillouin Calibration Data Input Product

There are a number of checks on the AUX_RBC_L2 file content, for example:

- Checking the pressure and temperature grids are consistent with the values given in the SPH (maximum, maximum and number of elements) and that the spectrum array sizes match with the size of the temperature and pressure grids.
- Checking that the various frequency grids are consistent with the SPH (minimum, maximum and number of elements) and that the array is monotonically increasing or decreasing.

If any of the tests fail then the L2B processor exits with an error message explaining what the problem is.

6.4. Screening of Calibration Coefficient Data Input Product

This is not implemented yet. The AUX_CAL_L2 is optionally used by the Optional Properties Code (Section 19).

6.5. Screening of Climatological Data Input Product

This is not implemented yet. The AUX_CLM_L2 may be used by the Optional Properties Code (Section 19) in the future.

7. The measurement Grouping Algorithm

This section explains the Grouping Algorithm method. First, to help describe the Grouping Algorithm we need to provide some definitions related to the spatial sampling of Aeolus data. This is illustrated in Figure 3 which shows a representation of a two dimensional section of Aeolus data.

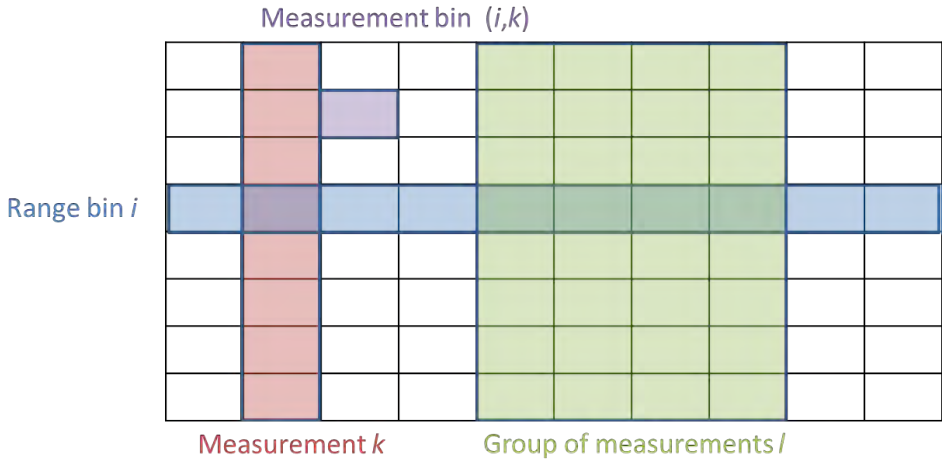


Figure 3. Spatial sampling of Aeolus data is defined in terms of measurements, groups of measurements, range bins and measurement-bins.

The smallest constituent, which we refer to as the measurement-bin, relates to the data of a particular range-bin of a particular measurement profile.

The (H)LOS wind retrieval (which is described in the following chapters) results in (H)LOS wind component “observations”. We refer to these observations also as wind results, to reflect the fact that they are independent wind observations, rather than profiles of wind observations i.e. have a separate geolocation for each wind result. The wind results are the output of a wind retrieval algorithm which is performed using the ALADIN instrument spectrometer data accumulated over a group of measurement-bins.

Prior to performing the wind retrievals, measurements are first assigned to groups; these groups are the output of the measurement Grouping Algorithm.

The Grouping Algorithm in effect provides flexibility in the horizontal resolution of L2B wind results and therefore also flexibility in the signal-to-noise ratio. Longer along-track accumulations may be required to produce more representative winds for NWP data assimilation or for research applications or to produce winds with a desired level of noise. The accumulation of measurement-bin data within the groups comes after the Grouping Algorithm and is explained in Section 11. Note that the grouping is followed by further partitioning of measurement-bins such as scene classification (as described in Section 8) and QC decisions (“measurement selection”); see the following chapters.

The Grouping Algorithm can produce groups of measurements irrespective of basic repeat cycles (BRCs) and hence the resultant L2B wind results may span across BRCs. This is possible due to the continuous nature of the Aeolus data along the orbit ground-track. Note that the L1B product wind observations are restricted to horizontal integrations within one BRC.

The measurement Grouping Algorithm is performed independently for each channel (Mie and Rayleigh; with independent settings). This is necessary because the two channels have unique properties e.g. potentially different vertical range-bin definitions and different horizontal averaging requirements to reach a certain precision. Typically Mie winds require fewer measurement-bins to achieve a given level of precision compared to the Rayleigh winds, with the expected levels of backscatter from e.g. clouds.

Inputs:

Symbol	Description	Input variable source
$\lambda_m(i,k)$	Measurement-bin latitude[2]	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/latitude_of_height_bin`
$\phi_m(i,k)$	Measurement-bin longitude	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/longitude_of_height_bin`
$z_m^u(i,k)$	Measurement-bin upper altitude[3]	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/altitude_of_height_bin`
$z_m^l(i,k)$	Measurement-bin lower altitude	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/altitude_of_height_bin`
z_{geoid}	Geoid offset (vertical distance) relative to WGS 84 ellipsoid valid for the BRC	ALD_U_N_1B: `Geoloc_ADS/Obs_Geoloc/Geoid_Separation`

Outputs:

Symbol	Description	ALD_U_N_2B file variable
N/A	Group structures defining which measurements belong to which group.	`Mie and Rayleigh Grouping ADS`, see Table 18 of the L2B/C IODD [RD10]

Algorithm:

Some features of the measurement Grouping Algorithm:

- A group of measurements can be formed from a minimum of one measurement up to all the measurements in the input L1B WVM file (although this latter choice is clearly not recommended!)
- The resultant groups are not required to be of a regular size in terms of the number of measurements they contain.
- The grouping algorithm starts from the first measurement in the L1B file — and groups are constructed, one by one, along the orbit.
- A minor consequence is that rather small groups may be formed at locations where grouping has to stop e.g. when the range bin definition has changed significantly, or at the end of the L1B file.
- The maximum horizontal extent (and therefore the number of measurements) of a given group is determined by parameters defined in the AUX_PAR_2B file and their interaction with the particular data in the L1B file.
- It is possible to select the group size to be equal to one BRC, which is referred to as the `Classic` setting. Given that nominally a BRC consists of 30 contiguous measurements, which corresponds to roughly 88 km horizontal distance of ground-track, then observations

[2] The latitude and longitudes reported in the L1B data are geodetic according to the WGS 84 ellipsoid.
[3] The altitude from L1B measurement-level data is the geometric height relative to the WGS 84 ellipsoid.

constructed from one BRC may actually be a sensible choice for e.g. global NWP model data assimilation. Mesoscale NWP model users may wish to have a better horizontal resolution (i.e. smaller group sizes); but with more noise.

The Grouping Algorithm method is selected via the AUX_PAR_2B settings file, see [Appendix F. Overview of AUX_PAR_2B settings and the control of the L2B processor algorithms], in particular:

- `WVM_Params`
  - `BRC_Grouping_Params`[4]
    - `Grouping_Method`

The chosen `Grouping_Method` determines the group size. The group size provides an upper limit on the horizontal extent of an observation. There are currently three Grouping methods available:

- `Classic`
  This mimics the L1B processor method, in that it creates groups which exactly match the BRCs. I.e. observations can only be constructed from the 30 (nominal) measurements in one BRC.
- `Advanced`
  This uses the parameters specified in the AUX_PAR_2B file, and constructs groups as large as possible given the constraints imposed from the parameters given the available set of measurements in a L1B product file.
- `Combine_BRCs`
  This uses a parameter specified in the AUX_PAR_2B (`WVM_Params/BRC_Grouping_Params/num_BRCs_to_merge`) file to construct groups from integer numbers of BRCs.

The "Advanced" method is controlled by the following parameters (separate parameters for Mie and Rayleigh channels):

- `WVM_Params`
  - `BRC_Grouping_Params`
    - `Grouping_Method`
      - `Max_Vertical_Rangebin_Misalignment_{Mie,Rayleigh}`

The maximum allowed vertical altitude difference between two {Mie, Rayleigh} measurement-bins with the same range-bin vertical index in a group. Moving along one measurement-bin at a time, if a measurement-bin's altitude differs by more than the threshold relative to the first measurement-bin of the group, then a new group will be started. Range bin height definitions can change due to the on-board Terrain Look-up table shifting (which shifts the whole profile of range-bins) and also due to changes in the definition of the vertical range bins which can occur up to seven times per orbit. This parameter is used to restrict the averaging of measurement-bins at different altitudes since this will reduce the vertical resolution of the observations.

- `Max_Horizontal_Accumulation_Length_{Mie, Rayleigh}`

The maximum horizontal distance between the first and last {Mie, Rayleigh} measurements in a group. If a measurement being considered for the group is a greater distance from the first measurement in the group than the threshold permits then a new group will be started. The distance is calculated using the measurement-bin latitude and longitude data and the great circle distance method (same formulae as in Section 20.2.3).

- `Max_Allowed_Gap_Between_{Mie,Rayleigh}_Measurements`

The maximum length of missing {Mie, Rayleigh} measurements before a group definition is closed and a new group is started. This ensures that large gaps between the data in a group can be eliminated e.g. do not want the situation with observations constructed from two widely spaced measurements, one at either end of the group. There should nominally not be many gaps in the data along the orbit. It is unclear how often missing measurements will occur; therefore this criterion may not have any influence in practice. The distance is calculated using

[4] "BRC" grouping parameters is a misnomer. In hindsight it should have been called "measurement" grouping parameters.

the measurement-bin latitude and longitude data and the great circle distance method (same formulae as in Section 20.2.3).

These AUX_PAR_2B parameters effectively control the maximum horizontal scale of the wind observations. Increasing the accumulation length allows more measurements into a group and hence should reduce random noise in the resultant winds, but at the cost of poorer horizontal resolution, which suppresses the detection of atmospheric variability. Note however that the final horizontal extent of a wind result also depends on the “measurement selection” algorithm and its settings (see Section 9).

Mie channel HLOS winds will most likely require smaller averaging lengths compared to the Rayleigh channel to achieve a given wind quality as has been demonstrated in end-to-end simulations.

Figure 4 illustrates the Grouping Algorithm via a so-called “measurement map”; this shows the location of measurement-bins in the vertical and horizontal (along-track) dimensions. Each small rectangular bin represents one measurement-bin (similar to Figure 3). The results of the grouping algorithm is shown by the groups which are indicated by green arrows (at the bottom), showing the horizontal extent of the groups. In this first cartoon example, the grouping algorithm distance criteria has been set such that groups are formed from five measurements (horizontally).

Measurement-bins within a group are later assigned to specific observations (see Section 9) — in Figure 4 the observation index to which a measurement-bin belongs to is also shown for clarity. This scenario could be for example the Rayleigh-clear observations obtained from a clear atmosphere, since all measurement-bins in a group on a given range bin are available for the HLOS wind retrieval. The light blue shading highlights for example the measurement-bins that go into observations numbers 1 and 54.

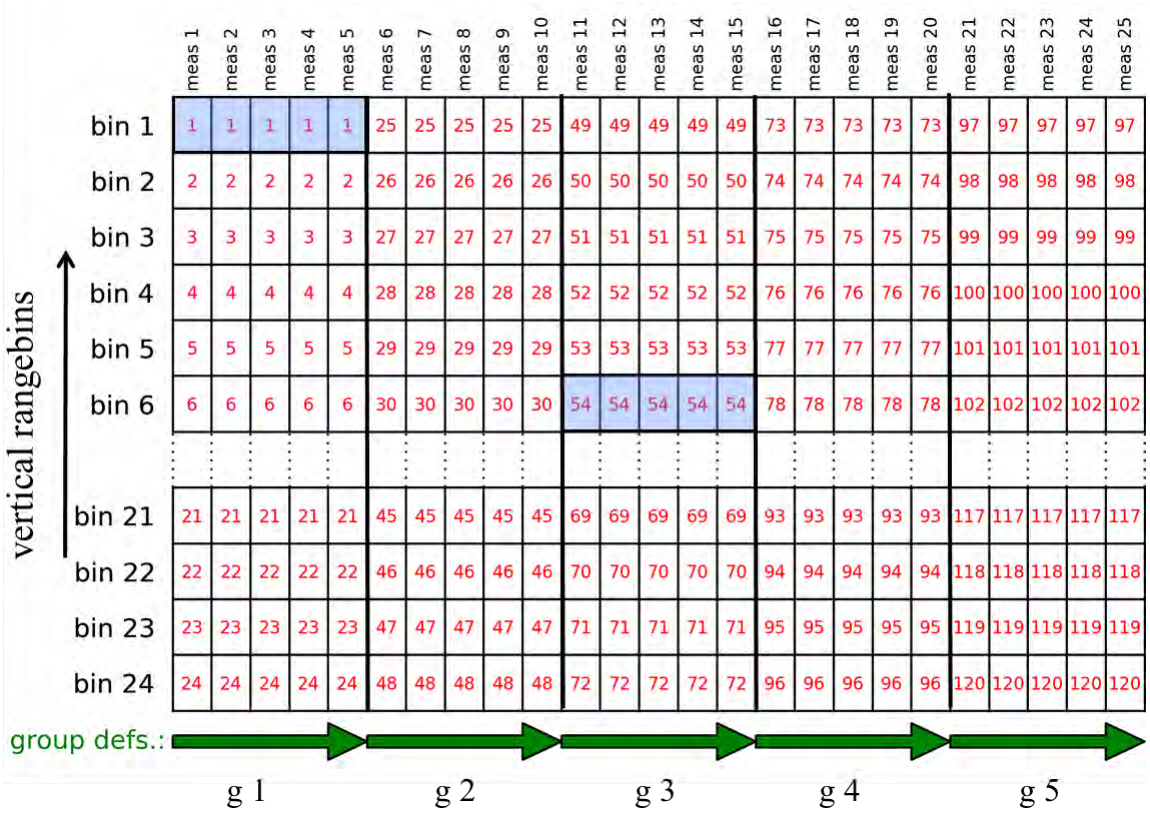


Figure 4. A “measurement map” to illustrate the how the measurement Grouping Algorithm works. The vertical axis represents the 24 vertical range bins of Aeolus and the horizontal axis represents either time or distance along the orbit track (or measurement number). The green arrows at the bottom indicate the number of measurements (and hence horizontal distance) over which groups are formed. The number within each measurement-bin shows to which wind observation it will be assigned to (in later processing decisions). The

`Max_Horizontal_Accumulation_Length_*` must have been chosen in this example to allow up to 5 measurements per group.

Figure 5 (in a similar way to Figure 4) illustrates how the grouping can vary with the "Advanced" grouping algorithm settings. Only one group is produced in this example which would happen with more relaxed distance criteria settings. Here all the measurement-bins on a given range bin will be assigned to one observation.

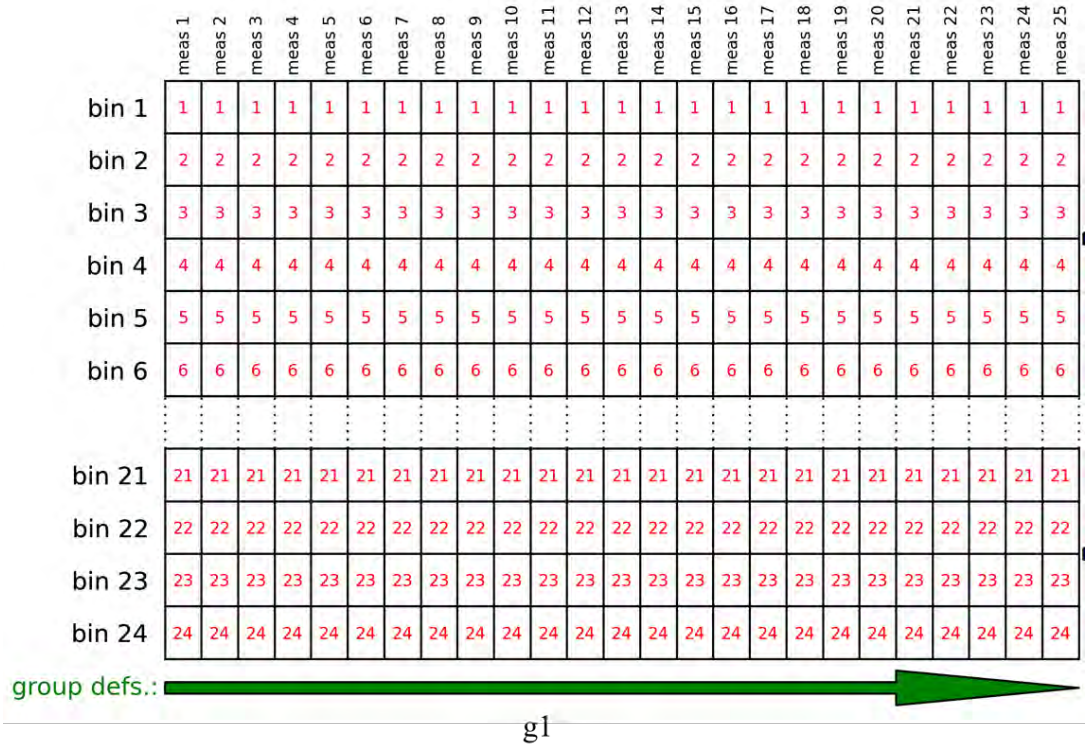


Figure 5. A "measurement map" illustrating a case where the Grouping Algorithm thresholds are not exceeded, so that one group has been formed from all the available measurements. The `Max_Horizontal_Accumulation_Length_*` must have been chosen in this example to exceed 25 measurements per group.

Figure 6 illustrates how new groups are started (and current groups are ended) if the vertical displacement of range bins heights is larger than a chosen threshold parameter. The observations (as a result of classification) could be for example Mie cloudy observations; hence not available on every range bin or perhaps the QC has rejected the other measurement-bins. In such a scenario the group sizes can be quite variable.

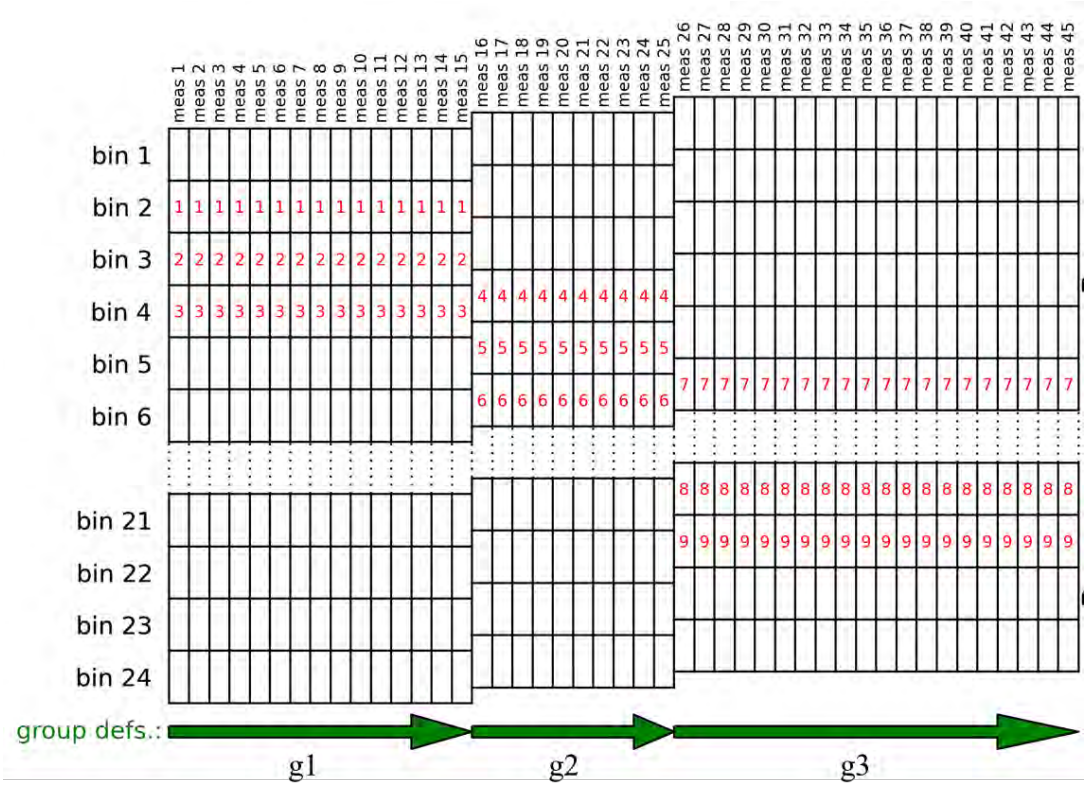


Figure 6. A "measurement map" illustrating the effect of changes of range-bin height definition on the grouping. Here only some of the measurement-bins are assigned to observations, which is possible with "measurement selection" (see section 9).

Variable group sizes may be a problem for some applications, since consistent quality and resolution may be preferred. Using the "Classic" method avoids this issue at the grouping stage, however the "measurement selection" may still lead to highly variable observation scales within one BRC of the "Classic" mode (particularly in inhomogeneous e.g. cloudy scenarios).

8. Atmospheric scene classification

One of the main features of the L2B processing is the accumulation of the signal from measurement-bins of the same atmospheric class. A wind retrieval is then performed on the accumulated data (Section 9). The atmospheric class for each measurement-bin (a measurement-bin is defined in Section 7; see Figure 3) is determined through a process referred to as "scene classification". By "atmospheric class" we are referring to classifying the atmosphere of the measurement-bin sensing volume according to its optical properties. In particular we want to know how much particulate backscatter relative to molecular backscatter is contributing to the signal received by ALADIN's channels.

Accumulating measurement-bins of the same atmospheric class is performed because it can improve the quality of the retrieved winds:

- Mie HLOS wind results require strong particulate backscatter (high Mie channel signal) to ensure a high signal to noise. Naturally they have lower signal to noise in an atmosphere with only low levels of particulates or no signal at all (clear air). Gross errors are more likely with very low SNR. Therefore it makes sense to only accumulate measurement-bins in which a reasonable level of particulate backscatter signal is expected.
- Rayleigh HLOS wind results are best in clear air conditions, because strong particulate backscatter contaminates the Rayleigh channel signal, and this requires an imperfect correction. The correction for Mie contamination (cross-talk) uses scattering ratio estimates, but this correction is imperfect, see Section 13). Therefore improved Rayleigh winds can be determined by avoiding particulate backscatter and only using measurement-bins not contaminated with particulate backscatter signal.

Currently the L2B processor produces four types of L2B Aeolus wind observations by classifying the measurement-bins into **Clear** or **Cloudy** conditions:

- Mie-cloudy:
  - Mie winds derived from measurement-bins classed as being cloudy i.e. have non-zero particle backscatter.
- Mie-clear:
  - Mie winds derived from measurement-bins classed as being clear i.e. having predominantly molecular backscatter. Of course Mie winds should not be possible in clear air, and hence Mie-clear wind results should only result if mistakes in the classification of measurement-bins have occurred (due to e.g. noisy scattering ratio estimates).
- Rayleigh-cloudy:
  - Rayleigh winds derived from measurement-bins classed as being cloudy i.e. have non-zero particle backscatter. For such cases the Rayleigh scattering signal is contaminated with Mie scattering signal. However, with the application of the Mie decontamination (cross-talk) correction (Section 13) these wind results are potentially useful (particularly for small scattering ratios in which the correction is more accurate).
- Rayleigh-clear:
  - Rayleigh winds derived from measurement-bins classed as being clear i.e. having predominantly molecular backscatter. The Mie decontamination (cross-talk) correction should only be a small correction in these cases.

Rayleigh-clear and Mie-cloudy wind results are expected to be of superior quality compared to the other two for the reasons mentioned above. Note that "cloudy" is a misnomer, in the sense that the particulate backscatter may also come from aerosols, not just cloud.

The atmospheric scene classification can be done by three alternative methods, each with specific algorithms:

1) **Classification by scattering ratio**[5]: Section 8.2

[5] The scattering ratio is defined as the ratio of total backscatter coefficient (molecular plus particulate) to molecular backscatter coefficient i.e. $\rho = \frac{\beta_m + \beta_p}{\beta_m}$. Values greater than 1 indicate the presence of particulate backscattering and hence some Mie signal contamination of Rayleigh signals.

2) **Classification by particle feature finding**[6]: Section 8.3
3) **Classification by particle extinction coefficient**[7]. Section 8.4

The algorithm used for scene classification is chosen via an AUX_PAR_2B setting, in particular: `Classification_Params/Classification_Type_{Mie,Rayleigh}`

- Option 1: `Class_Backscat_Ratio`
- Option 2: `Class_Layer_Detected`
- Option 3: `Class_Ext_Threshold`

8.1. An introduction to the Optical Properties Code (OPC) retrieval

For all three classification options listed above it is possible to obtain the relevant variables used for classification via the Optical Properties Code (OPC) algorithm, which is described in Section 19. The OPC uses only the Rayleigh interferometer (Fabry-Pérot) signal only (i.e. does not use the Mie channel) to retrieve particulate optical properties. provides information on the presence or not of particles (feature finding) and also can retrieve extinction and scattering ratio estimates that can be used in the scene classification algorithms. The main rationale behind this option is to have a plan-B in place in case of Mie channel failure, hence no scattering ratio input from the L1B processor needed for scene classification in the L2B processor

A Rayleigh measurement-bin alone is insufficient to detect particulate scattering, however the slant-path profile of Rayleigh measurement-bins can be used to retrieve particulate optical properties given a priori estimates of the atmosphere temperature and pressure at the measurement location. The AUX_MET_12 input file provides this atmospheric mass information from which estimates of molecular backscatter and extinction coefficients (in clear atmosphere) can be derived (see Appendix D. Retrieval of atmosphere molecular backscatter and extinction (Rayleigh scattering)). Deviations in the observed Rayleigh lidar signal from the forward modelled (lidar equation) clear air signal provides information on particulate layers.

The OPC derived optical properties are expected to be particularly beneficial where there are Rayleigh measurement-bins subject to particulate scattering, yet there are no Mie measurement-bins in close proximity (i.e. the Mie channel cannot be used to detect the magnitude of the particulate scattering). Rayleigh winds derived with an assumption of clear air when in fact particulate backscatter is present will be significantly biased (since a cross-talk correction requires accurate scattering ratio estimates). This may be a problem in atmospheric scenarios with e.g. Polar Stratospheric Clouds (PSC) or high cirrus clouds, because the nominal Mie range-bin definitions stop around the tropopause level altitude.

The OPC output depends on the provision of signal calibration constants. These are available in the auxiliary input data (AUX_CAL_L2 file) to the L2B processor or alternatively they can be calculated in the OPC itself. The source of calibration constants is controlled in the AUX_PAR_2B settings (see [Appendix F. Overview of AUX_PAR_2B settings and the control of the L2B processor algorithms]).

The expected signal and measured signal profiles are inputs for the OPC iterative scheme discussed in section 19 and [Appendix C. Particle transmission computation from the Fabry-Pérot channel signal]. The result is an estimate of particle pseudo transmission throughout the measurement profile. Deficiencies in the expected and measured FP signals results in an artificially oscillating structure in the retrieved particle transmission profile. Deficiencies in the measured profile are due to cross-talk and instrument noise. Deficiencies in the estimated profile is due to, e.g., imperfect pressure and temperature estimates from the AUX_MET input files and imperfect (estimation of) calibration parameters. The *k*-power method discussed in [Appendix C. Particle transmission computation from the Fabry-Pérot channel signal] damps the oscillations substantially.

[6] Particle feature finding is part of the Optical Properties Code (OPC).
[7] The extinction coefficient of the atmosphere characterizes how easily it can be penetrated by a beam of light. SI unit: m^{-1}

With the feature finding option, particle detection is performed by putting a threshold value on the retrieved particle pseudo transmission profile. This is done for all measurement-bins within a group. Next a 2-dimensional median filter is applied to filter outliers (this is optional).

Particle extinction is retrieved from the OPC measurement-bin transmission estimates. Then backscatter and subsequently scattering ratio are obtained from the retrieved particle extinction coefficient profile, assuming knowledge of the lidar ratio (extinction/backscatter). The lidar ratio information may be provided by the AUX_CLM_L2 files in future; however currently it uses hard-coded estimates. This will probably lead to biases in the estimated scattering ratios.

Once a particle layer is detected, additional properties are estimated as discussed in Section 19.3. These optical properties include particle backscatter, extinction, scattering ratio, lidar ratio (inverse of the backscatter-to-extinction ratio), the particle layer top and bottom location and the cross-talk correction factor. For isolated layers the transmission is obtained from the product of the retrieved particle transmission for bins capturing the particle layer.

The effects of noise can be mitigated with options to use a median filter (averaging along both the measurement and range-bin direction).

The Optical Properties Code iterative algorithm is controlled via settings in the AUX_PAR_2B file under `Optical_Properties_Params.`

The OPC settings include:

- `k_power`
- `FP_Xtalk_factor`
- `Median_Filter_Window_Width_x`
- `Median_Filter_Window_Width_y`
- `Calibration_from_AuxCal`
- `Cross_Talk_from_AuxCal`
- `Apply_Median_Filter`
- `Apply_2D_Feature_Finder`
- `FP_On_Upper_Bin_Std_Threshold`

See Table 102 of the IODD [RD10] for the meaning of these settings.

8.2. *Option 1:* Classification using scattering ratio

Inputs:

Symbol	Description	Input variable source
$\boldsymbol{A_m(i,k), B_m(i,k)}$	Rayleigh measurement-bin Useful Signal in Channels A and B	ALD_U_N_1B: `Useful_Signal_MDS/Measurement_Useful_Signal/ Rayleigh_Altitude_Bin_Useful_Signal_Info / Useful_Signal_Channel_{A,B}`
$\boldsymbol{\rho_m(i,k)}$	Mie channel measurement-bin scattering ratio estimates (either nominal or refined)	ALD_U_N_1B: `Product Confidence Data ADS/ List_of_Measurement_PCDs/List_of_Meas_Alt_Bin_PCDs/{Scattering_Ratio_Mie, Refined_Scattering_Ratio_Mie}`
$\boldsymbol{z_m^u(i,k)}$	Measurement-bin upper altitude[8]	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/altitude_of_height_bin`
$\boldsymbol{z_m^l(i,k)}$	Measurement-bin lower altitude	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/altitude_of_height_bin`
$\boldsymbol{z_{geoid}}$	Geoid offset	ALD_U_N_1B:

[8] The altitude from L1B measurement-level data is the geometric height relative to the WGS 84 ellipsoid.

	(vertical distance) relative to WGS 84 ellipsoid valid for the BRC	`Geoloc_ADS/Obs_Geoloc/Geoid_Separation`
N/A	Scattering ratio threshold values as a function of altitude	AUX_PAR_2B: `Classification_Params/List_of_{Mie,Rayleigh}_BackscatterRatio_Thresholds/{Mie,Rayleigh}_BackscatterRatio_Threshold/Threshold_Value/Altitude/`
N/A	Choice for source of scattering ratio estimates	AUX_PAR_2B: `Optical_Properties_Params/ScatRatio_Method`

Outputs:

Symbol	Description	ALD_U_N_2B file variable
N/A	Atmospheric class of each measurement-bin (for both Mie and Rayleigh); clear or cloud	`L2B Measurement Product Confidence Data (PCD) ADS/ L2B_{Mie,Rayleigh}_Classification_QC/L2B_{Mie,Rayleigh}_Meas_Bin_Classification/L2B_{Mie,Rayleigh}_Meas_Bin_Class Flags{1,2}` e.g. Table 40 of the IODD [RD10].

Algorithm:

This method of classification requires measurement-bin level scattering ratio estimates provided by either:

- L1B Mie measurement-bin level scattering ratio estimates. **This is the current recommended method for classification.**
- Optical Properties Code (OPC) retrievals of scattering ratio derived on Rayleigh measurement-bins. **This method of classification is still undergoing testing and is not yet recommended.**

The choice regarding which source of scattering ratio estimates to use is made via the AUX_PAR_2B file. Options include:

- `Scat_Ratio_from_L1B_Mie` − Use the L1B Mie measurement-bin nominal scattering ratio estimates.
- `Scat_Ratio_from_L1B_Mie_refined` − Use the L1B Mie measurement-bin refined scattering ratio estimates. **This is the recommended option; they are more accurate than the L1B nominal estimates**
- `Scat_Ratio_from_RaylOnly` − Use the scattering ratio estimates derived by the OPC iterative algorithm
- `Scat_Ratio_One_If_No_Mie`− Assign scattering ratio=1.0 (i.e. clear air) to Rayleigh measurement-bins if Mie measurement-bins are not in available (in close proximity; overlapping)

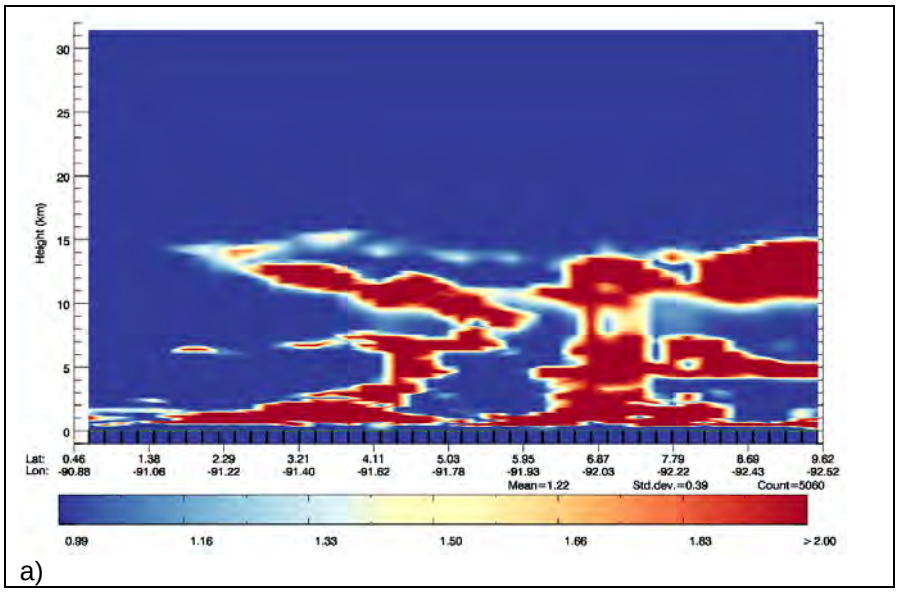
The scattering ratio threshold values (as a function of altitude) used in classification are defined in the AUX_PAR_2B file i.e. a list of threshold versus altitude.
An example of the AUX_PAR_2B file are provided in the software release e.g. see SUM [RD19].

This classification algorithm is based on detecting cloudy or clear scenes using threshold values of the scattering ratio (which can vary with altitude):

- **Cloud**: If the scattering ratio is greater than the threshold value then this indicates particulate scattering (Mie scattering) is present on top of the molecular scattering (Rayleigh scattering). N.B. Cloud is intended to encompass all particulate scattering, including hydrometeors and aerosols.

- **Clear:** If the scattering ratio is less than the threshold value then the measurement-bin is assumed to have only molecular scattering. N.B. Clear is intended to mean no particulate scattering, only molecular scattering.

Note this classification by threshold is prone to errors due to the noisiness of the estimates of measurement-bin scattering ratio. For example, Figure 7 illustrates how the L1B Mie measurement-bin refined scattering ratio estimates are rather noisy and may have a poor resolution compared to the atmosphere's true scattering ratio. Due to low transmission the scattering ratio estimates below optically thick clouds are very noisy if not just missing values (black areas). The classification based on a typical scattering ratio threshold of 1.25 (constant with altitude) is also shown.



a)

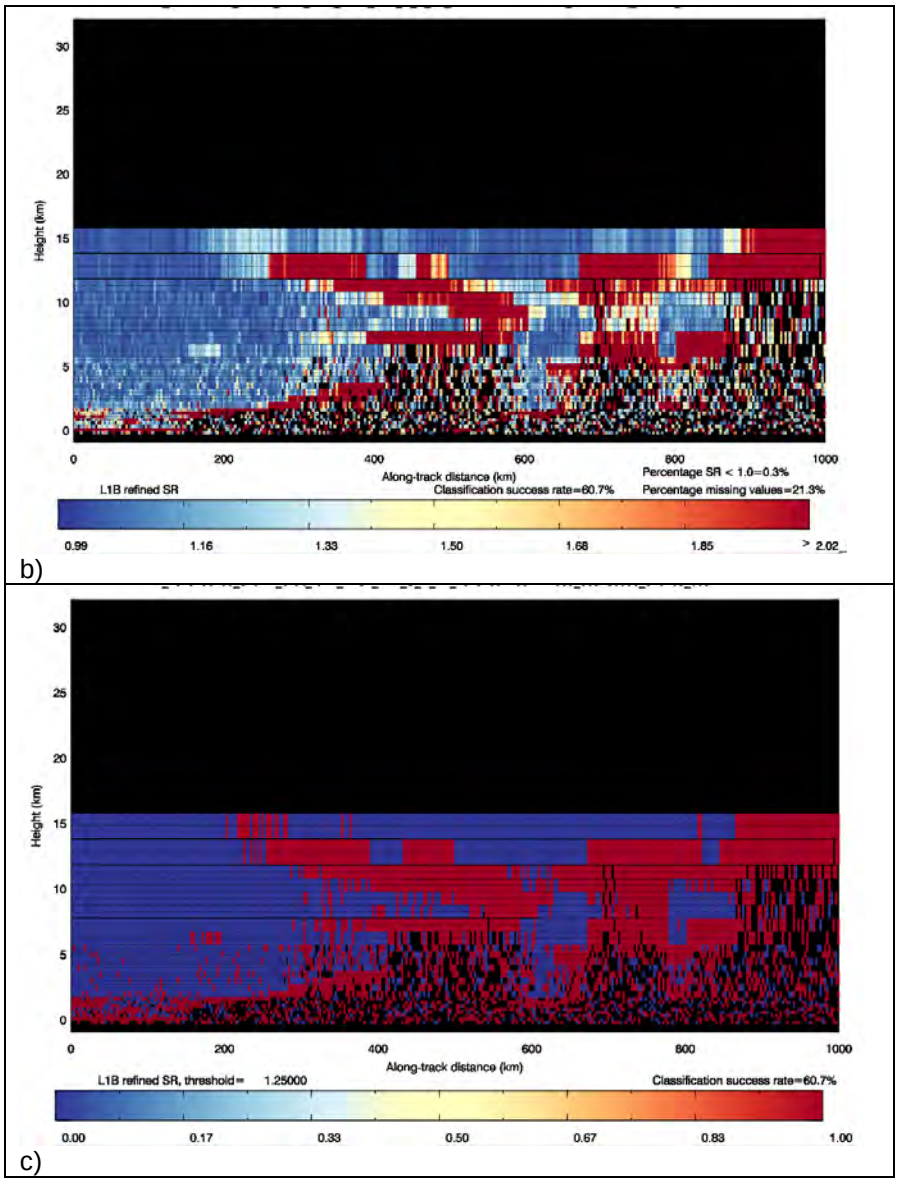


Figure 7. Realistic simulation of Aeolus for a cloudy atmosphere a) input "truth" scattering ratio b) example L1B Mie measurement-bin refined scattering ratio c) example of classification based on threshold scattering ratio of 1.25.

The L1B scattering ratio (SR) is only available on the Mie measurement-bins. Therefore, to allow classification of the Rayleigh measurement-bins, the Mie measurement-bin scattering ratio must be mapped onto the Rayleigh measurement-bins. Also, the OPC scattering ratio is only available on the Rayleigh measurement-bins. Therefore, to allow classification of the Mie measurement-bins, the Rayleigh measurement-bin scattering ratio must somehow be mapped onto the Mie measurement-bins.

Note that the Rayleigh measurement-bins are not compelled to coincide with the Mie ones; however it is typically expected that one Rayleigh measurement-bin can have an integer number of Mie measurement-bins within it. Therefore several situations can arise:

- Mie and Rayleigh measurement-bins match one-to-one: in which case just copy the Mie (Rayleigh) SR to the Rayleigh (Mie).
- Multiple Mie (Rayleigh) measurement-bins lie within one Rayleigh (Mie) measurement-bin: in which case average the Mie (Rayleigh) scattering ratios.
- Mie and Rayleigh are not properly aligned (average?)
- Mie (Rayleigh) but no Rayleigh (Mie) data: in which case there is nothing to do

There are options in the algorithm to set the scattering ratio to 1.0 (i.e. assume perfectly clear air) if there are no Mie measurement-bins available for assigning to the Rayleigh measurement-bins i.e. we do not extrapolate. This is typically done in the AUX_PAR_2B via setting the back-up method: `Optical_Properties_Params/ScatRatio_Method2` to be `Scat_Ratio_One_If_No_Mie`
There is also a minimum altitude for assuming scattering ratio is equal to one parameter:
`Optical_Properties_Params/Minimum_Altitude_for_Assuming_Rho_1`

Once the Mie and Rayleigh measurement-bins are assigned scattering ratios, then the classification can begin.

For classification of both Mie and Rayleigh measurement-bins: The middle altitude of the range gate, z_i, is determined from the average of the range gate top and bottom. Subsequently, a threshold scattering ratio value is determined from the threshold profile by linearly interpolating to the altitude z_i. If the scattering ratio exceeds the interpolated threshold, then the measurement-bin's atmospheric class is set to `Cloud`, else it is set to `Clear`.

8.3. *Option 2:* Classification using particle feature detection

This method of classification is still undergoing testing. Testing has shown that the method is mature enough to use in case of Mie channel failure, hence missing scattering ratio input from the L1B processor.

Inputs:

Symbol	Description	Input variable source
$\boldsymbol{A_m(i,k), B_m(i,k)}$	Rayleigh measurement-bin Useful Signal in Channels A and B	ALD_U_N_1B: `Useful_Signal_MDS/Measurement_Useful_Signal/ Rayleigh_Altitude_Bin_Useful_Signal_Info/Useful_Signal_Channel_{A,B}`
N/A	Particle feature detection mask (binary measurement-bin map)	Internal L2B processor output of the Optical Properties Code (OPC)
$\boldsymbol{z_m^u(i,k)}$	Measurement-bin upper altitude[9]	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/altitude_of_height_bin`
$\boldsymbol{z_m^l(i,k)}$	Measurement-bin lower altitude	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/altitude_of_height_bin`
$\boldsymbol{z_{geoid}}$	Geoid offset (vertical distance) relative to WGS 84 ellipsoid valid for the BRC	ALD_U_N_1B: `Geoloc_ADS/Obs_Geoloc/Geoid_Separation`

Outputs:

Symbol	Description	ALD_U_N_2B file variable
N/A	Atmospheric class of each measurement-bin (for both Mie and Rayleigh); clear or cloud	`L2B Measurement Product Confidence Data (PCD) ADS/ L2B_{Mie,Rayleigh}_Classification_QC/L2B_{Mie,Rayleigh}_Meas_Bin_Classification/L2B_{Mie,Rayleigh}_Meas_Bin_Class Flags{1,2}` e.g. Table 40 of the IODD [RD10].

[9] The altitude from L1B measurement-level data is the geometric height relative to the WGS 84 ellipsoid.

Algorithm:

There are algorithm choices regarding how the particle features are detected, which are controlled via the AUX_PAR_2B file. In particular under: `Optical_Properties_Params/`

- `Median_Filter_Window_Width_x`: controls how many measurements to apply a median filter to, to reduce noise
- `Median_Filter_Window_Width_y`: control how many range-bins to apply a median filter to, to reduce noise.
- `Apply_Median_Filter`: whether or not to apply a median filter, to reduce effects of noise
- `Apply_2D_Feature_Finder`: whether or not to allow feature finding in 2 dimensions.

The OPC requires molecular extinction estimates and the available methods are documented in 0.

The classification algorithm is based on detecting cloudy or clear scenes via threshold values of particle pseudo-transmission values (see section 19.2 for details of the algorithm). The thresholds are calculated internally in the algorithm. A typical threshold value is 0.1, which means that bins with transmission values > 1.1 and smaller than 0.9 are assigned: **Cloud** i.e. particulate scattering is present, otherwise **Clear**.

Note, however that the classification by threshold is prone to errors due to the noisiness of the estimates of measurement-bin pseudo-transmission. For this reason a two-dimensional (optional, see settings above) median filter, which extends over one observation group, is applied to the outcome of the above procedure, as such removing outliers and increase homogeneity of detected layers.

The classification measurement-bin map is provided on Rayleigh measurement-bins. Therefore, to allow classification of the Mie measurement-bins, the Rayleigh results must be mapped onto the Mie measurement-bins.

Note that the Rayleigh measurement-bins are not compelled to coincide with the Mie ones; however it is typically expected that one Rayleigh measurement-bin can have an integer number of Mie measurement-bins within it (or none at all if above a certain altitude). Therefore several situations can arise:

- Mie and Rayleigh measurement-bins match one-to-one: in which case just copy the Rayleigh classification result to the Mie.
- Multiple Mie measurement-bins lie within one Rayleigh measurement-bin: in which case copy the Rayleigh classification results to the Mie bins.
- Mie and Rayleigh are not properly aligned: TBD
- Mie but no Rayleigh data: this should not happen

8.4. *Option 3:* Classification using particle extinction coefficient

This method of classification is not recommended and no further development is done. The method presented in section 8.3 is preferred in case of Mie channel failure, hence missing scattering ratio input from the L1B processor.

Inputs:

Symbol	Description	Input variable source
$\boldsymbol{A_m(i,k), B_m(i,k)}$	Rayleigh measurement-bin Useful Signal in Channels A and B	ALD_U_N_1B: `Useful_Signal_MDS/Measurement_Useful_Signal/Rayleigh_Altitude_Bin_Useful_Signal_Info/Useful_Signal_Channel_{A,B}`
N/A	Retrievals of particle extinction coefficient derived on Rayleigh measurement-	Internal L2B processor output of the Optical Properties Code (OPC)

	bins.	
$z_m^u(i,k)$	Measurement-bin upper altitude[10]	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/altitude_of_height_bin`
$z_m^l(i,k)$	Measurement-bin lower altitude	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/altitude_of_height_bin`
z_{geoid}	Geoid offset (vertical distance) relative to WGS 84 ellipsoid valid for the BRC	ALD_U_N_1B: `Geoloc_ADS/Obs_Geoloc/Geoid_Separation`
N/A	Particle extinction threshold values as a function of altitude	AUX_PAR_2B: `Classification_Params/List_of_{Mie,Rayleigh}_Extinction_Thresholds/{Mie,Rayleigh}_Extinction_Threshold/Threshold_Value/Altitude/`
N/A	Choice for OPC method of particle extinction	AUX_PAR_2B: `Optical_Properties_Params/PartExt_Method/`

Outputs:

Symbol	Description	ALD_U_N_2B file variable
N/A	Atmospheric class of each measurement-bin (for both Mie and Rayleigh); clear or cloud	`L2B Measurement Product Confidence Data (PCD) ADS/ L2B_{Mie,Rayleigh}_Classification_QC/L2B_{Mie,Rayleigh}_Meas_Bin_Classification/L2B_{Mie,Rayleigh}_Meas_Bin_Class Flags{1,2}` e.g. Table 40 of the IODD [RD10].

Algorithm:

There are algorithm choices regarding the method of calculating both particle and molecular extinction in the OPC, which are controlled via the AUX_PAR_2B file.
Particle extinction method:
`Optical_Properties_Params/PartExt_Method/`
Options include (only one option):

- `PartExt_Method_Iterative`

Molecular extinction method: see [Appendix D. Retrieval of atmosphere molecular backscatter and extinction (Rayleigh scattering)] for details.
The particle extinction threshold values (as a function of altitude) for use in classification are defined in the AUX_PAR_2B file. An example of the AUX_PAR_2B file are provided in the software release e.g. see SUM [RD19].

The classification algorithm is based on detecting cloudy or clear scenes via threshold values of the particle extinction coefficient:

- **Cloud**: If the extinction coefficient is greater than the threshold value then this indicates particulate scattering (Mie) is present (intended to encompass all particulate scattering, including hydrometeors and aerosols).
- **Clear:** If the extinction coefficient is less than the threshold value then the measurement-bin is assumed to be particulate free.

[10] The altitude from L1B measurement-level data is the geometric height relative to the WGS 84 ellipsoid.

Note, however that the classification by threshold is prone to errors due to the noisiness of the estimates of measurement-bin particle extinction coefficient. For this reason a two-dimensional median filter, which extends over one BRC, is applied to the outcome of the above procedure, as such removing outliers and increase homogeneity of detected layers.

The OPC particle extinction coefficient is provided only on the Rayleigh measurement-bins. Therefore, to allow classification of the Mie measurement-bins, the Rayleigh measurement-bin particle extinction coefficient must be mapped onto the Mie measurement-bins.

Note that the Rayleigh measurement-bins are not compelled to coincide with the Mie ones; however it is typically expected that one Rayleigh measurement-bin can have an integer number of Mie measurement-bins within it (or none at all if above a certain altitude). Therefore several situations can arise:

- Mie and Rayleigh measurement-bins match one-to-one: in which case just copy the Rayleigh particle extinction coefficient to the Mie.
- Multiple Mie measurement-bins lie within one Rayleigh measurement-bin: in which case copy the Rayleigh particle extinction coefficient to the Mie bins.
- Mie and Rayleigh are not properly aligned (average?)
- Mie but no Rayleigh data: this should not happen

For classification of both Mie and Rayleigh measurement-bins: The middle altitude of the range gate, z_i, is determined from the average of the range gate top and bottom. Subsequently, a threshold particle extinction coefficient value is determined from the threshold profile by linearly interpolating to the altitude z_i. If the particle extinction coefficient exceeds the interpolated threshold, then the measurement-bin's atmospheric class is set to `Cloud`, else it is set to `Clear`.

9. Measurement-bin selection

This section relates to how the measurement-bins are chosen for use in the wind retrieval.

Inputs:

- The L2B processor grouping algorithm results for each channel (see outputs of Section 7)
- The L2B processor atmospheric class of each measurement-bin within the group (see outputs of Section 8)
- The L2B processor quality control decisions on the L1B measurement-bin data (see outputs of Section 6)

Outputs:

- Number of classification types detected for the group, per range bin
- Measurement map, relating each measurement-bin to the observation (wind result) index it is applied in. This is stored in the L2B product as: `Measurement Map` ADS, see Section 4.3.1 of the IODD [RD10]

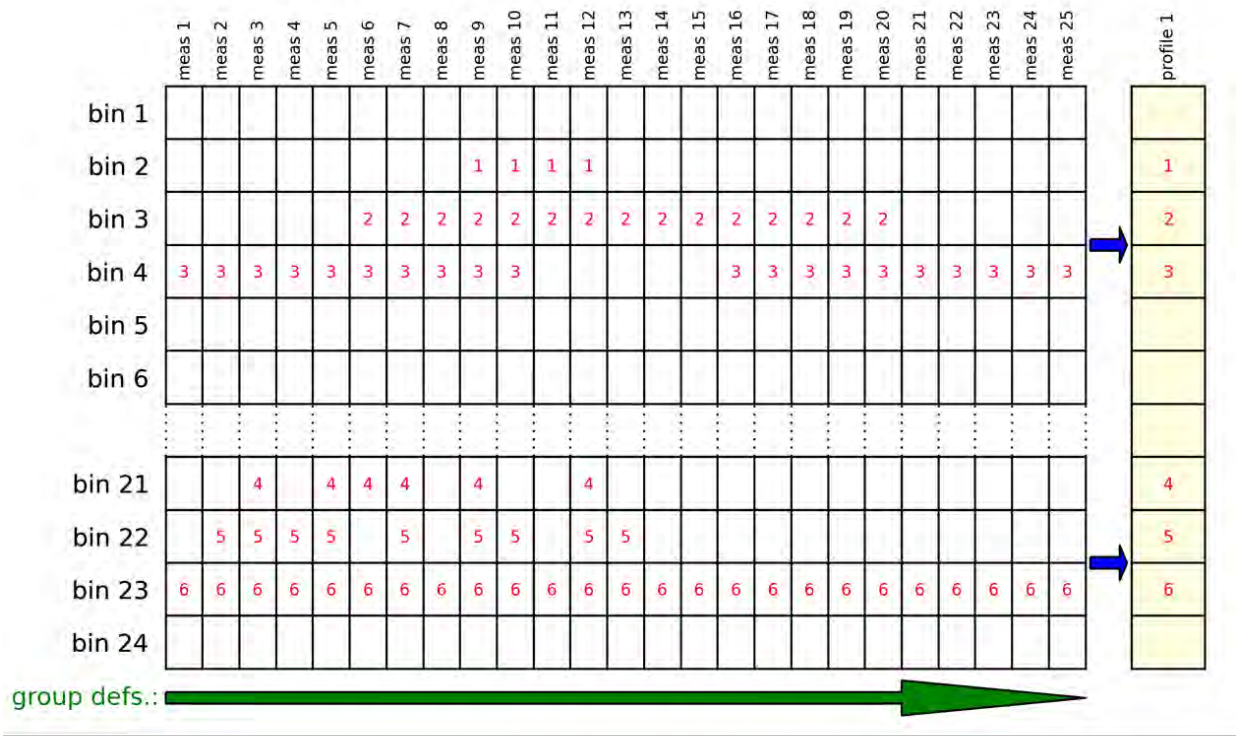
Algorithm:

For a given: channel; measurement group and range-bin, wind retrievals (see sections 13 and 14) are performed using the accumulated signal from the measurement-bins of a certain atmospheric classification. This wind retrieval from accumulated signal is restricted to using data of one vertical range bin at present, see Sections 10 and 11. Therefore there is potentially a L2B wind observation for each combination of:

- Rayleigh
  - Group ID
    - Range-bin number
      - Clear
      - Cloudy
- Mie
  - Group ID
    - Range-bin number
      - Clear
      - Cloudy

If all the measurement-bins in a group (for a given range-bin) are classified as clear, then a cloudy wind result will not be available and vice versa if all cloudy. Whether or not to use the spectrometer data of a given measurement-bin is also dependent on the quality control results of the L1B measurement-bin data (Section 6) i.e. a given measurement-bin can be rejected if deemed to have failed the QC.

For each measurement-bin, the observation in which it has been used is stored in a “measurement map” structure via an “observation index” (aka wind result ID). For example, Figure 9.1 shows how the “measurement map” could look based on a classification of cloudy measurements (e.g. Mie-cloudy). Notice how only one observation (of this type) is allowed per range bin and per group (here only one group is shown).



g1

Figure 9.1. Similar to Figure 8.4, but with measurement-bin selection such that not all measurement-bins are assigned to this type of observation e.g. could be a Mie-cloudy type of observation. The numbers indicate the "observation index" (aka "wind result ID").

Profiles of wind observations are defined in the L2B file format e.g. see `L2B Wind Profile MDS` in the IODD [RD10]. Profiles are defined as arrays of observation indices, as indicated on the right of Figure 9.1, that are sufficiently close in geolocation to be considered as a vertical profile. Of course the profiles are slant rather than vertical due to the off-nadir pointing of ALADIN in WVM mode. It may have benefits in NWP to associate wind results into a profile.

10. Weighting of measurement-bins

For each channel, measurement-bins are assigned a weight for its use in a particular observation index accumulation of signal (see Section 11). The summed up (or accumulated) ALADIN data e.g. spectrometer counts, is later passed to the wind retrieval algorithm to produce a wind result (e.g. see Section 13). For a given channel, each measurement-bin can contribute to only one wind observation i.e. to the clear or the cloudy observation. QC decisions can mean the measurement-bin is not used in any observation. This differs from the L1B processor observations for which all measurement-bins for one BRC are used to produce one wind observation i.e. no discrimination between clear/cloud. In other words the L2B processor has more flexible averaging over measurement-bins horizontally to produce wind observations with the resolution and noise levels preferred by the user and avoids mixing clear and cloudy measurement-bins.

Inputs:

- L2B processor mapping of each channel's measurement-bin data into the observations (L2B processor internal output of Section 9)

Outputs:

- For each L2B wind observation the weights of the measurement-bins in the accumulation:
  - Weights $W(i,k)$ for measurement *k* and range bin *i*
- Weights of the Rayleigh and Mie measurement-bins used in L2B wind observations are recorded in the L2B product in: `Measurement Map` ADS, see section 4.3.1 of the IODD [RD10]

Algorithm:
Currently the algorithm is very simple; each measurement-bin is assigned the same weight within its corresponding observation as follows:

$$W(i,k) = 1 \qquad \textbf{Eq. 10-1}$$

That is, in the current implementation of the L2BP, each measurement-bin of an observation is weighted equally with a factor of 1 i.e. just a straight-forward summation. Weights are set to zero if a measurement-bin is not part of the observation accumulation.

As mentioned previously, a given measurement-bin contributes only to a single observation (wind result) for a given channel. If a measurement-bin's class is different to the observation class being considered then its weight is set to zero; meaning for example that clear measurement-bins should not contribute to cloudy wind results.

For more efficient storage, the weights written to the L2B output file are:

$$W^*(i,k) = Int[1000 * W(i,k)] \qquad \textbf{Eq. 10-2}$$

11. Accumulation of measurement-bin information

Measurement-bin L1B data is accumulated in the horizontal dimension only (across measurements), not vertically (across range-bins), to produce observation level data, upon which the wind retrieval can be performed. Measurement-bin data can be noisy due to the small signals and is fairly small-scale (~3 km horizontal extent) making it difficult to use in scientific applications. Therefore accumulation of the signal across measurements is done in the L2B processor to improve the signal-to-noise ratio and produces more spatially representative wind observations for use in NWP data assimilation and atmospheric research. The resultant horizontal scale can be determined to a significant extent by the Grouping algorithm settings, as was described in Section 7. Note the accumulated data for wind retrievals consists of the L1B spectrometer data rather than the calculated response (or wind) of each measurement-bin; this process will be explained in Sections 13 and 14. This choice is made to reduce the chance of measurement-bin level gross errors contaminating the observation-level wind result.

Inputs:

- L2B processor internal: Measurement-bin weights for each observation of each channel (Section 10)

Outputs:

- Accumulated L1B spectrometer data for each channel and observation index (i.e. group, range-bin and class) from which a wind result can be retrieved - see sections 13.1 and 14.1
- Observation geolocation information (e.g. L2B wind result centre-of-gravity geolocation) - Section 12.1
- Representative horizontal averaging length of L2B wind estimate - Section 12.3. It may be useful for NWP users to know what the wind result represents in terms of horizontal resolution.

12. Computation of (H)LOS wind observation geolocation and auxiliary parameters

In the following chapters *N* refers to the number of measurements selected in the group that goes on to form the observation.

12.1. Computation of the observation's centre-of-gravity geolocation

Inputs:

Symbol	Description	Input variable source
$\lambda_m(i,k)$	Measurement-bin latitude[11]	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/latitude_of_height_bin`
$\phi_m(i,k)$	Measurement-bin longitude	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/longitude_of_height_bin`
$z_m^u(i,k)$	Measurement-bin upper altitude[12]	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/altitude_of_height_bin`
$z_m^l(i,k)$	Measurement-bin lower altitude	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/altitude_of_height_bin`
z_{geoid}	Geoid offset (vertical distance) relative to WGS 84 ellipsoid valid for the BRC	ALD_U_N_1B: `Geoloc_ADS/Obs_Geoloc/Geoid_Separation`
$t_m(k)$	Measurement centroid date-time	ALD_U_N_1B: `Geoloc_ADS/Meas_AOCS/Measurement_Centroid_Time`
α	Vertical centre-of-gravity height parameter	AUX_PAR_2B: `RBC_Algorithm_Params/Rayleigh_Height_Weight_Upper`
$W(i,k)$	Measurement-bin weights	Output of L2B processing described in Section 10

Outputs:

Symbol	Description	ALD_U_N_2B file variable
$\lambda_{cog}(i)$	Centre-of-gravity latitude	`{Mie,Rayleigh}_Geoloc_ADS/WindResult_Geolocation/Latitude_COG`
$\phi_{cog}(i)$	Centre-of-gravity longitude, reported in L2B product	`{Mie,Rayleigh}_Geoloc_ADS/WindResult_Geolocation/Longitude_COG`
$t_{cog}(i)$	Centre-of-gravity date-time	`{Mie,Rayleigh}_Geoloc_ADS/WindResult_Geolocation/DateTime_COG`
$z_{cog}^{u,l}(i)$	Centre-of-gravity upper/lower	`{Mie,Rayleigh}_Geoloc_ADS/WindResult_Geolocation/{Altitude_top,Altitude_Bottom}`

[11] The latitude and longitudes reported in the L1B data are geodetic according to the WGS 84 ellipsoid.
[12] The altitude from L1B measurement-level data is the geometric height relative to the WGS 84 ellipsoid.

	altitude (relative to the geoid)	
$z^{v}_{cog}(i,k)$	Vertical centre-of-gravity altitude[13] (relative to the geoid)	`{Mie,Rayleigh}_Geoloc_ADS/WindResult_Geolocation/Altitude_VCOG`
z_{geoid}	Height of geoid above WGS84 ellipsoid (repeat of L1B input value for nearest BRC)	`{Mie,Rayleigh}_Geoloc_ADS/WindResult_Geolocation/WGS84_to_Geoid_Altitude`

Algorithm:

The centre-of-gravity geolocation of the L2B observation is derived from the L1B measurement-bin geolocation data via a centre-of-gravity measurement index. This centre-of-gravity measurement index is a normalised weighted sum of the measurement index *k* of the observation group and is calculated as:

$$\bar{k}_i = int\left[\frac{\sum_{k=1}^{N} W(i,k)k}{\sum_{k=1}^{N} W(i,k)}\right] \quad \textbf{Eq. 12-1}$$

where *int*[] denotes the integer part of a real number.

Let $\lambda_m(i,k)$ denote the latitude at measurement-bin (*i,k*). Then the centre-of-gravity latitude for the observation is given by:

$$\lambda_{cog}(i) = \lambda_m(i,\bar{k}_i) \quad \textbf{Eq. 12-2}$$

Similarly, let $\phi_m(i,k)$ denote the longitude at measurement-bin (*i,k*). Then the centre-of-gravity longitude for the observation is given by:

$$\phi_{cog}(i) = \phi_m(i,\bar{k}_i) \quad \textbf{Eq. 12-3}$$

Similarly, let $t_m(k)$ denote the date-time for measurement-bin (*i,k*). Then the centre-of-gravity date-time for observation is given by:

$$t_{cog}(i) = t_m(\bar{k}_i) \quad \textbf{Eq. 12-4}$$

Similarly, let $z_m^{u,l}(i,k)$ denote the altitude at upper/lower edge of measurement-bin (*i,k*). Then the upper/lower edge of the centre-of-gravity altitude (relative to the geoid) for the observation is given by:

$$z_{cog}^{u,l}(i) = z_m^{u,l}(i,\bar{k}_i) - z_{geoid} \quad \textbf{Eq. 12-5}$$

Note that L2B centre-of-gravity altitudes are relative to the geoid height of the nearest BRC (L1B report geoid offset at the BRC level).

A more representative altitude for the observation can be found by weighting the upper and lower altitudes to give the vertical centre-of-gravity altitude for the observation:

$$z_{cog}^{v}(i) = z_{cog}^{l}(i) + \alpha(z_{cog}^{u}(i) - z_{cog}^{l}(i)) \quad \textbf{Eq. 12-6}$$

The resultant altitude is a fraction *α* up the range bin.

[13] The observation-level altitude is the geometric height with respect to EGM96 geoid.

For a fixed value of alpha for all the wind observations then set the AUX_PAR_2B parameters to be e.g.:

- For the Rayleigh channel:

```
<Rayleigh_Height_Assignment_Method>use_fixed_weight</Rayleigh_Height_Assign
ment_Method>
<Rayleigh_Height_Weight_Upper>0.49</Rayleigh_Height_Weight_Upper>
```

- For the Mie channel:

```
<Mie_Height_Assignment_Method>use_fixed_weight</Mie_Height_Assignment_Metho
d>
 <Mie_Height_Weight_Upper>0.5</Mie_Height_Weight_Upper>
```

For the Rayleigh wind observations, using an exponentially decaying weighting with altitude (with a scale height of 7 km), to represent the climatological molecular scattering profile (which is proportional to density), the weighting factor should be typically in the range 0.46 to 0.49 for Aeolus Rayleigh range-bins (assuming the extinction is negligible). For the Mie wind observations the scattering from particulates depends on the position of particles within the range bin which cannot be known a priori, therefore a 0.5 weighting is probably best.

An optional Rayleigh altitude assignment is available, which uses the AUX_MET_12 provided temperature and pressure versus altitude information to calculate a more appropriate vertical positioning for the assignment of the altitude; taking into account the vertical weighting function of attenuated backscatter through the vertical range-bin. Also, there is an option to use a climatological molecular backscatter profile. These options for the `Rayleigh_Height_Assignment_Method` are:

- `calculate_from_pressure_and_temperature`
- `calculate_from_climatological_beta_m`

There are further L2B wind observation geolocation values that are calculated in a similar fashion to that shown above using the centre-of-gravity measurement index and the vertical centre-of-gravity (if appropriate):

- Satellite to target range distance[14]
- Date-time
- Azimuth angle[15]
- Elevation[16] and incidence angle. Incidence angle equals 90 - elevation angle (in degrees).

12.2. Observation start and stop geolocations

The geolocations for the start and stop of an observation are simply taken from the L1B measurement-bin geolocations for the first and last index in which the measurement-bins have non-zero weight. The results are stored in the L2B product under:
`{Mie,Rayleigh}_Geoloc_ADS/WindResult_Geolocation/*_{Start,Stop}` e.g. `Latitude_Start`

12.3. Computation of a representative L2B observation integration length

Each wind observation includes a representative horizontal length-scale for the observation. This is obtained by calculating the great-circle distance between the latitude/longitude start and stop of the observation. The equations for the great circle distance were taken from: http://mathworld.wolfram.com/GreatCircle.html

The observation integration lengths are stored in the L2B product as: `{Rayleigh,Mie}_Wind_MDS/WindResult/Integration_Length`

[14] Distance along the LOS between the instrument and the lower edge of the height bin.

[15] Azimuth of target-to-satellite topocentric (horizontal reference frame) pointing vector of the height bin. It ranges from 0 deg to 360 deg and is measured from North over East.

[16] Elevation of target-to-satellite topocentric (horizontal reference frame) pointing vector of the height bin. It ranges from -90 deg to +90 deg and is measured positive from the North-East plane towards Zenith. Typical values for elevation angle in WVM are ~53 degrees.

12.4. Computation of the reference pressure and temperature

Pressure and temperature data are provided from NWP model forecasts by the AUX_MET_12 file. They are required for the Rayleigh-Brillouin correction of Rayleigh wind results; see Section 15. Values that are representative for the L2B wind observation scale need to be determined.

Inputs:

Symbol	Description	Input variable source
$p_{nwp}(i', k)$	Atmospheric pressure from NWP model output	AUX_MET_12: `Met_MDS_off_nadir/Met_Layer/pnom` N.B. the off-nadir rather than nadir AUX_MET data is chosen as this should be more closely matched in geolocation with the WVM L1B data.
$T_{nwp}(i', k)$	Atmospheric temperature from NWP model output	AUX_MET_12: `Met_MDS_off_nadir/Met_Layer/T`
$z_{nwp}(i', k)$	Geometric height (w.r.t. geoid) of the above NWP parameters	AUX_MET_12: `Met_MDS_off_nadir/Met_Layer/znom`
$z_m(i, k)$	Measurement-bin mid-level geometric height relative to WGS 84 ellipsoid	ALD_U_N_1B: Derived from : `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/altitude_of_height_bin`
z_{geoid}	Geoid offset (vertical distance) relative to WGS 84 ellipsoid valid for the BRC	ALD_U_N_1B: `Geoloc_ADS/Obs_Geoloc/Geoid_Separation`
$W(i, k)$	Measurement-bin weights	Output of L2B processing described in Section 10

Outputs:

Symbol	Description	ALD_U_N_2B file variable
$p_0(i)$	Reference atmospheric pressure at observation level	`{Rayleigh,Mie}_Wind_MDS/WindResult/Reference_pressure`
$T_0(i)$	Reference atmospheric temperature at observation level	`{Rayleigh,Mie}_Wind_MDS/WindResult/Reference_temperature`

Algorithm:

Computation of the reference pressure and temperature is done in two steps.

The first step is the vertical interpolation of the AUX_MET_12 profiles of pressure and temperature (on model levels) as a function of geometric height (with respect to geoid) to the observation geometric heights (with respect to geoid) producing intermediate values p_m and T_m. The second is a weighted average in the horizontal.

The intermediate pressures and temperatures are given by:

$$p_m(i,k) = p_{nwp}(i'(k),k) \quad \textbf{Eq. 12-7}$$

$$T_m(i,k) = T_{nwp}(i'(k),k) \quad \textbf{Eq. 12-8}$$

where $i'(k)$ is such that $z_m(i,k) - z_{geoid}$ (i.e. measurement-bin geometric height relative to geoid) lies between $z_{nwp}(i'(k))$ and $z_{nwp}(i'(k)-1)$ (in geometric height space). Note that this is effectively a nearest neighbour interpolation.

The reference pressure for the observation is given by a normalised weighting of the pressure values for the measurement-bins:

$$p_0(i) = \frac{\sum_{k=1}^{N} W(i,k) p_m(i,k)}{\sum_{k=1}^{N} W(i,k)} \quad \textbf{Eq. 12-9}$$

Similarly, the reference temperature is given by

$$T_0(i) = \frac{\sum_{k=1}^{N} W(i,k) T_m(i,k)}{\sum_{k=1}^{N} W(i,k)} \quad \textbf{Eq. 12-10}$$

13. Rayleigh channel (H)LOS wind processing — L2B Rayleigh wind observation retrieval

Inputs:

Symbol	Description	Input variable source
$\boldsymbol{A_m(i,k), B_m(i,k)}$	Measurement-bin Useful Signal in Channels A and B	ALD_U_N_1B: `Useful_Signal_MDS/Measurement_Useful_Signal/ Rayleigh_Altitude_Bin_Useful_Signal_Info / Useful_Signal_Channel_{A,B}`
$\boldsymbol{C_m(i,k), D_m(i,k)}$	Measurement-bin Internal Reference Signal in Channels A and B	ALD_U_N_1B: `Measurement_ADS/ List_of_Rayleigh_Reference_Pulse_{A,B}/ Rayleigh_Reference_Pulse`
$\boldsymbol{\rho_m(i,k)}$	Rayleigh measurement-bin scattering ratio. As derived from the L1B Mie measurement-bin scattering ratio values (either simple or refined estimates are available) or from the OPC derived scattering ratios.	L2B processor internal variable, see section 8
$\boldsymbol{\lambda_0}$	Laser wavelength	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/ Satellite_Characterisation_Data/ Laser_Wavelength`
$\boldsymbol{\varphi}$	Incidence angle of the line of sight (LOS) of the instrument at vertical centre-of-gravity of the range-bin	Derived from elevation angles of range-bins from ALD_U_N_1B, see section 12.1: `Geolocation ADS/ Observation_Rayleigh_Geolocation/List_of_Observation_Geolocation_of_Height_Bins/ Observation_Geolocation_of_Height_Bin/Topocentric_Elevation_of_ Height_Bin`
$\boldsymbol{V_{sat,m}(k)}$	Projection of the satellite apparent velocity onto the instrument LOS direction at measurement level	ALD_U_N_1B: `Geolocation_ADS/List_of_Measurement_Geolocations/AOCS_LOS_Velocity`
$\boldsymbol{C_R, W_{GR}, fit_{Ray}(lat), W_{HR}, C_M, W_{GMR}, fit_{Mie}(lat), W_{HMR}, C_{offset}}$	Ground speed derived correction parameters	ALD_U_N_1B: `Ground Wind Detection ADS/ Rayleigh_Ground_Correction_Velocity, Rayleigh_Ground_Correction_Weighting_Factor, HBE_Rayleigh_Ground_Correction_Velocity, Rayleigh_Harmonic_Correction_Factor, Mie_Ground_Correction_Velocity,`

		`Rayleigh_Correction_With_Mie_Ground_Echo_Weighting_Factor`, `HBE_Mie_Ground_Correction_Velocity`, `Rayleigh_Correction_With_Mie_Harmonic_Weighting_Factor`, `Mie_Rayleigh_Ground_Correction_Offset`
$\boldsymbol{T_0(i), p_0(i)}$	Atmospheric reference temperature and pressure	L2B processor output, as described in Section 12.4
$\boldsymbol{r(i)}$	Satellite to target range distance (metres) at vertical centre-of-gravity	Derived from ranges of range-bins from ALD_U_N_1B, see section 12.1: `Geolocation ADS/Observation_Rayleigh_Geolocation/List_of_Observation_Geolocation_of_Height_Bins/Observation_Geolocation_of_Height_Bin/Satellite_Range_of_Height_Bin`
$\boldsymbol{r_{ref}}$	Satellite to target reference range (metres) for off-nadir pointing	ALD_U_N_1B: `Calibration & Characterisation Data GADS/L1B_Characterisation_Data/Rdb_Characterisation_Data/Zero_Reference_Range_OffNadir`
$\boldsymbol{K_{Ray}}$	Coefficient for the gradient of frequency offset with respect to range (Hz/km) for off-nadir pointing for the Rayleigh channel	ALD_U_N_1B: `Calibration & Characterisation Data GADS/L1B_Characterisation_Data/Rdb_Characterisation_Data/Rayleigh_Slope_OffNadir`
$\boldsymbol{W(i,k)}$	Measurement-bin weights	Output of L2B processing described in Section 10

Outputs:

Symbol	Description	ALD_U_N_2B file variable
$\boldsymbol{v_{HLOS}(i)}$	L2B Rayleigh HLOS (horizontal[17] line of sight) wind component for the observation.	`Rayleigh_Wind_MDS/WindResult/Rayleigh_Wind_Velocity`
N/A	The classification type for the observation	`Rayleigh_Wind_MDS/WindResult/observation_type`
$\frac{\partial v_{HLOS}(i)}{\partial T}, \frac{\partial v_{HLOS}(i)}{\partial p}, \frac{\partial v_{HLOS}(i)}{\partial RR}, \frac{\partial v_{HLOS}(i)}{\partial \rho}$	Partial derivatives of HLOS wind with respect to various inputs to the retrieval; an output from ILIAD (Section 15).	`Rayleigh_Wind_MDS/WindResult/Rayleigh_Wind_to_{Temperature,Pressure,Backscatter_Ratio}`
$\boldsymbol{\rho_0(i)}$	The observation level scattering ratio used in the ILIAD scheme.	`Rayleigh_Wind_MDS/WindResult/Reference_Backscatter_Ratio`

[17] LOS wind component can also be produced, via a switch in the AUX_PAR_2B: `line_of_Sight_Wind_Flag`. Stored in same variable as HLOS wind.

The Rayleigh channel HLOS wind retrieval involves computation of:

- the weighted (observation level) Rayleigh response (Section 13.1),
- the ILIAD LOS wind component (Section 15.1),
- two post-ILIAD correction terms:
  - the weighted internal reference LOS velocity component (Section 13.4.1),
  - the weighted spacecraft LOS velocity component (Section 13.4.2),
  - the ground correction velocity component (Section 13.4.3),
  - the range dependent bias correction velocity component (Section 13.4.4),
- the post-ILIAD HLOS wind component (Section 13.5).

These are described separately below. The Rayleigh channel HLOS wind retrieval is done on the selected measurement-bins for each observation (see Section 9).

13.1. Computation of a weighted atmospheric Rayleigh response

The normalised weighted atmospheric Rayleigh response (effectively a mean average over the measurement-bins of the observation) is an input to ILIAD (the Rayleigh-Brillouin Correction code) and is computed as follows. The normalised weighting of useful signal in Rayleigh channels A and B are given by analogues of the quantities A2, B2 from L1B DPM document [RD2] eqn (103):

$$A_0(i) = \frac{\sum_{k=1}^{N} W(i,k) A_m(i,k)}{\sum_{k=1}^{N} W(i,k)} \quad \textbf{Eq. 13-1}$$

$$B_0(i) = \frac{\sum_{k=1}^{N} W(i,k) B_m(i,k)}{\sum_{k=1}^{N} W(i,k)} \quad \textbf{Eq. 13-2}$$

The weighted atmospheric Rayleigh response is given by the analogue of L1B DPM [RD2] *r1ro(i)*, eqn (105):

$$RR_{atm}(i) = \frac{A_0(i) - B_0(i)}{A_0(i) + B_0(i)} \quad \textbf{Eq. 13-3}$$

This response is used as input to ILIAD algorithm (Section 13.3).

13.2. Computation of a weighted scattering ratio

Following the previous normalised weighting expressions, an observation-level scattering ratio is constructed from measurement-bin estimates and is given by:

$$\rho_0(i) = \frac{\sum_{k=1}^{N} W(i,k) \rho_m(i,k)}{\sum_{k=1}^{N} W(i,k)} \quad \textbf{Eq. 13-4}$$

This is an input to the ILIAD algorithm (Section 13.3).

13.3. Computation of the atmospheric return LOS wind component (using the ILIAD scheme)

The ILIAD scheme is applied to give the atmospheric return LOS wind component:

$$V_{atm}(i) = ILIAD(RR_{atm}(i), \rho_0(i), T_0(i), p_0(i)) \quad \textbf{Eq. 13-5}$$

Further details on the ILIAD algorithm are given in Section 15.1. This is used as input to post-ILIAD corrections (Section 13.5).

Additional outputs are the LOS sensitivity with respect to temperature, pressure and scattering ratio: $\frac{\partial v_{LOS}}{\partial T}, \frac{\partial v_{LOS}}{\partial p}, \frac{\partial v_{LOS}}{\partial \rho}$.
They must be multiplied by $1/\sin(\varphi)$ to produce the L2B output parameters HLOS sensitivity with respect to temperature and pressure.

13.4. Computation internal reference and correction terms

Two steps are described in this section, in preparation for use in the post-ILIAD HLOS wind component calculation (Section 13.5).

13.4.1. Computation of a weighted internal reference LOS velocity component

Computation of a normalised weighted internal reference LOS velocity component $V_{INT}(i)$ is performed as follows. The normalised weighted summation of internal reference signal for Channels A and B are given by analogues of the quantities C2, D2 from L1B DPM [RD2] eqn (104):

$$C_0(i) = \frac{\sum_{k=1}^{N} W(i,k) C_m(i,k)}{\sum_{k=1}^{N} W(i,k)} \quad \textbf{Eq. 13-6}$$

$$D_0(i) = \frac{\sum_{k=1}^{N} W(i,k) D_m(i,k)}{\sum_{k=1}^{N} W(i,k)} \quad \textbf{Eq. 13-7}$$

The weighted Rayleigh internal reference response is given by an analogue of L1B DPM eqn (105):

$$RR_{int}(i) = \frac{C_0(i) - D_0(i)}{C_0(i) + D_0(i)} \quad \textbf{Eq. 13-8}$$

The internal reference LOS velocity component is calculated, using an analogue to the ILIAD scheme as:

$$V_{int}(i) = ILIAD(RR_{int}(i)) \quad \textbf{Eq. 13-9}$$

The required input for this method is provided by the AUX_RBC_L2 file. Further details of this are given in Section 15.2. The partial derivatives of LOS wind with respect to temperature, pressure, Rayleigh response and scattering ratio are also output from the ILIAD scheme.

This is used as an input to the post-ILIAD corrections (Section 13.5)

13.4.2. Computation of a weighted spacecraft LOS velocity component

The normalised weighted spacecraft LOS velocity component is given by:

$$V_{sat}(i) = \frac{\sum_{k=1}^{N} W(i,k) V_{sat,m}(k)}{\sum_{k=1}^{N} W(i,k)} \quad \textbf{Eq. 13-10}$$

This is used as an input to the post-ILIAD corrections (Section 13.5).

13.4.3. Ground return corrections

V_{GR} is the ground correction velocity component term, which accounts for various sources of bias such as errors in LOS pointing knowledge. This is optionally used in the final HLOS wind calculation. It is calculated as follows, with correction components taken directly from the L1B product file. All the

components are derived from ground return wind estimates (done by the L1B processor). The ground correction velocity component is:

$$V_{GR}(i) = W_{GR}C_R + W_{HR}fit_{Ray}(lat) + W_{GMR}C_M + W_{HMR}fit_{Mie}(lat) + C_{offset}$$ **Eq. 13-11**

W_{GR}, W_{GMR}, W_{HR} and W_{HMR} are weighting parameters (expected values 0 or 1 and sum 0 or 1) that are basically switches to control the use of the various correction terms. This makes the calculation very flexible for use. The $fit_{Mie,Ray}$(lat) are speeds which are the result of the Harmonic Bias estimation tool. C_M is the Mie ground correction speed and C_R is the Rayleigh ground correction speed (the ZWC results of the L1B product without harmonic function filtering). Note the total correction V_{GR}value that is provided in the L1B file can be used instead of re-calculating it from the coefficients in the L2Bp. The Mie channel correction factors are included as a possibility for the Rayleigh because the Rayleigh channel ground returns may not be accurate enough to use; whereas the Mie might be good enough. C_{offset} accounts for a possible constant offset between the Mie and Rayleigh ground correction speeds.
The weighting factors and whether or not to use the total L1B correction value are controlled by the various AUX_PAR_2B settings: `ZWC_Params`.

13.4.4. Rayleigh range dependent bias correction

V_{RDB} is the velocity component term which corrects for the so-called range dependent bias (RDB). This is calculated in the same way as for the L1B processor (see L1B DPM section 15.2.2) i.e. assuming the RDB varies linearly with satellite to target range (distance):

$$V_{RDB}^{Ray}(i) = \frac{\lambda_0}{2}K_{Ray}(r(i) - r_{ref})$$ **Eq. 13-12**

Use of this correction term is optional and is controlled via the AUX_PAR_2B switch: `RDB_Params/Do_Rayleigh_RDB_corr`

13.4.5. Rayleigh M1 mirror temperature based wind bias correction

This is a placeholder for the description of this algorithm. To be detailed later.

13.5. Computation of the Rayleigh channel (H)LOS wind component

The Rayleigh channel HLOS wind component observation is given by the analogue of L1B DPM [RD2] eqn (69):

$$v_{HLOS}(i) = \frac{1}{\sin(\varphi)}\left(V_{atm}(i) - V_{int}(i) - V_{sat}(i) - V_{GR}(i) - V_{RDB}^{Ray}(i)\right)$$ **Eq. 13-13**

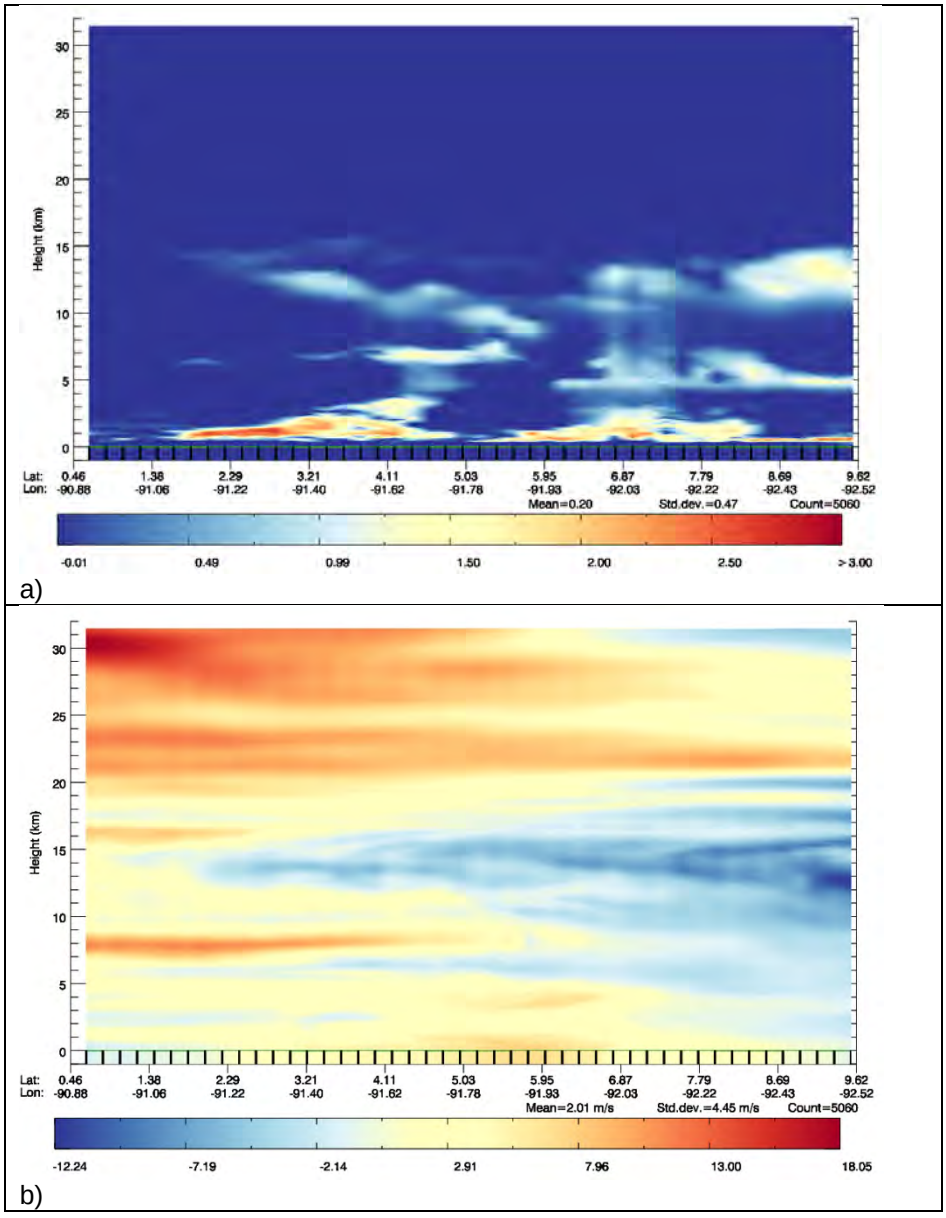
This is the main L2B result reported in L2B products: i.e. the HLOS wind component observation. The first two terms in the bracket account for the difference in frequency between emitted and backscattered light i.e. the Doppler shift. Other terms are basically corrections so that we are only left with a Doppler shift due to atmospheric wind. That is, residual Doppler shift due to the satellite to Earth velocity is subtracted (although this should be small if the satellite pointing is working well) and terms to account for various biases which can be measured using ground returns (which are a zero-wind reference) and the so-called range-dependent bias (an angular effect on the spectrometers).

If one wishes to produce line-of-sight rather than horizontal line-of-sight wind components then the $1/\sin(\varphi)$ factor is removed. This is controlled via the AUX_PAR_2B settings i.e.

- HLOS winds: `<Line_of_Sight_Wind_Flag>false</Line_of_Sight_Wind_Flag>`
- LOS winds: `<Line_of_Sight_Wind_Flag>true</Line_of_Sight_Wind_Flag>`

To give an impression of what the Aeolus Rayleigh L2B winds look like, Figure 8 shows an example from the same simulation scenario as used in Figure 7. Some key parameters used:

- Grouping algorithm was used with the advanced method with a maximum accumulation set to 100 km
- Classification was done with L1B refined scattering ratio with a constant threshold of 1.25.



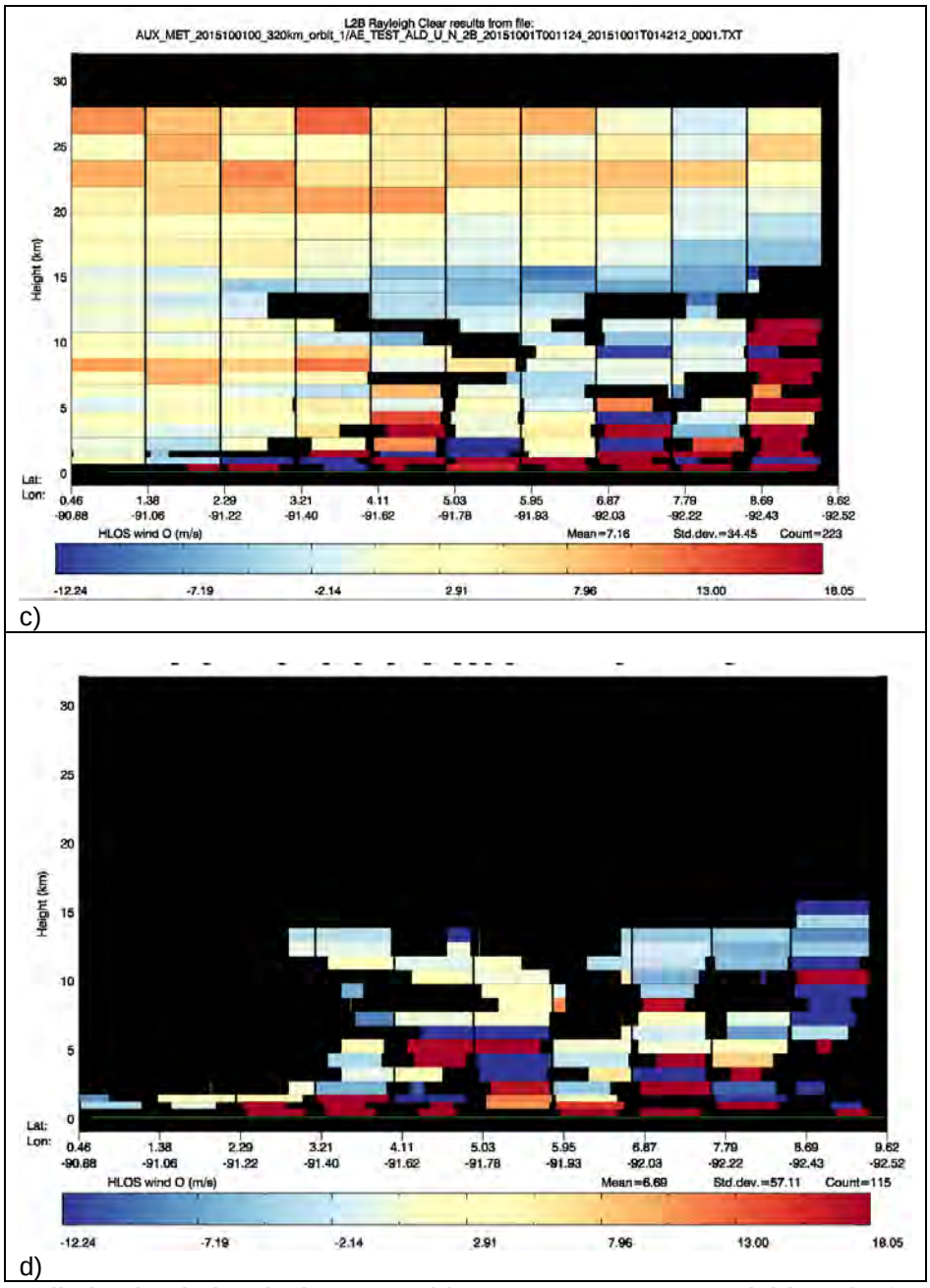


Figure 8. Realistic simulation (using E2S with TcO1279 ECMWF model input) of Aeolus for a cloudy atmosphere. Lidar cross-section plots of a) input "truth" logarithm of scattering ratio i.e. showing where cloud is b) input "truth" HLOS wind c) L2B Rayleigh-clear wind observations d) L2B Rayleigh-cloudy wind observations.

14. Mie channel (H)LOS wind processing — L2B Mie wind observation retrieval

Inputs:

Symbol	Description	Input variable source
$\boldsymbol{LID_{atm}(i,j,k)}$	Measurement spectrometer counts at measurement-bin level	ALD_U_N_1B: `Measurement_ADS/List_of_Mie_Measurement_Data/Mie_Measurement_Data`
$\boldsymbol{LID_{int}(j,k)}$	Internal Reference spectrometer counts at measurement level	ALD_U_N_1B: `Measurement_ADS/List_of_Mie_Reference_Pulse/Mie_Reference_Pulse`
$\boldsymbol{MRS_{atm}, \alpha, MRS_{int}, \delta}$	MRC linear fit coefficients i.e. gradient (MRS=Mie response slope) and intercepts for atmospheric and internal paths. One set of values per L1B file.	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/Mie_Response_Calibration_Data/Mie_Measurement_Response_Calibration/{Measurement_Mean_Sensitivity, Measurement_Zero_Frequency}` `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/Mie_Response_Calibration_Data/Mie_Reference_Pulse_Response_Calibration/{Reference_Pulse_Mean_Sensitivity,Reference_Pulse_Zero_Frequency}`
$\boldsymbol{EMR_{atm}, EMR_{int}}$	Spectral nonlinearity corrections for atmospheric and internal paths. One set of corrections as function of MRC frequency step per L1B file.	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/Mie_Response_Calibration_Data/ List_of_Mie_Frequency_Step_Results/ Mie_Frequency_Step_Result/ {Measurement_Error_Mie_Response, Reference_Pulse_Error_Mie_Response}`
$\boldsymbol{\lambda_0}$	Laser wavelength	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/ Satellite_Characterisation_Data/ Laser_Wavelength`
$\boldsymbol{\varphi}$	Incidence angle of the line of sight (LOS) of the instrument at vertical centre-of-gravity of the range-bin	Derived from elevation angles of range-bins from ALD_U_N_1B, see section 12.1: `Geolocation ADS/ Observation_Rayleigh_Geolocation/List_of_Observation_Geolocation_of_Height_Bins/ Observation_Geolocation_of_Height_Bin/Topocentric_Elevation_of_ Height_Bin`
$\boldsymbol{V_{sat,m}(k)}$	Projection of the satellite apparent velocity onto the instrument LOS direction at	ALD_U_N_1B: `Geolocation_ADS/List_of_Measurement_Geolocations/AOCS_LOS_Velocity`

	measurement level	
$C_M, W_{GM}, fit_{Mie}(lat), W_{HM}$	Ground speed derived correction parameters	ALD_U_N_1B: `Ground Wind Detection ADS/ Mie_Ground_Correction_Velocity, Mie_Ground_Correction_Weighting_Factor, HBE_Mie_Ground_Correction_Velocity, Mie_Harmonic_Correction_Factor`
N/A	Mie core algorithm processing parameters	AUX_PAR_2B: See [Appendix F. Overview of AUX_PAR_2B settings and the control of the L2B processor algorithms] on AUX_PAR_2B settings, in particular: `Mie_Algorithm_Params`
$r(i)$	Satellite to target range distance (metres) at vertical centre-of-gravity	Derived from ranges of range-bins from ALD_U_N_1B, see section 12.1: `Geolocation ADS/ Observation_Mie_Geolocation/List_of_Observation_Geolocation_of_Height_Bins/ Observation_Geolocation_of_Height_Bin/ Satellite_Range_of_Height_ Bin`
r_{ref}	Satellite to target reference range (metres) for off-nadir pointing	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/Rdb_Characterisation_Data/Zero_Reference_Range_OffNadir`
K_{Mie}	Coefficient for the gradient of frequency offset with respect to range (Hz/km) for off-nadir pointing for the Mie channel	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/Rdb_Characterisation_Data/Mie_Slope_OffNadir`
$W(i,k)$	Measurement-bin weights	Output of L2B processing described in Section 10

Outputs:

Symbol	Description	ALD_U_N_2B file variable
$v_{HLOS}(i)$	L2B Mie HLOS (horizontal[18] line of sight) wind component for the observation.	`Mie_Wind_MDS/WindResult/Mie_Wind_Velocity`
N/A	The classification type for the observation	`Mie_Wind_MDS/WindResult/observation_type`

Algorithm:

The Mie channel HLOS wind retrieval involves:

- Computation of weighted (observation level) Mie spectrometer readouts (Section 14.1)

[18] LOS wind component can also be produced, via a switch in the AUX_PAR_2B: `line_of_Sight_Wind_Flag.` Stored in same variable as HLOS wind.

  - the computation is performed for both the atmospheric signal and the internal reference signal
- Computation of the raw response via the Mie core algorithm (Section 14.2),
  - the computation is performed for both the atmospheric signal and the internal reference signal
- Spectral nonlinearity correction (Section 14.3)
  - the computation is performed for both the atmospheric signal and the internal reference signal
- Computation of the Mie core algorithm LOS wind velocity component (Section 14.4)
- Computation of two post-core correction terms
  - the weighted internal reference LOS velocity component (Section 14.5.1)
  - the weighted spacecraft LOS velocity component (Section 14.5.2)
  - the ground return correction (Section 14.5.3)
  - the range dependent bias correction (Section 14.5.4)
- Computation of the post-core HLOS wind component (Section 14.6).

These are described separately below. The Mie channel HLOS wind retrieval is done on the selected measurement-bins for each observation (see Section 9).

14.1. Computation of weighted Mie spectrometer readouts

The weighted Mie spectrometer readouts needed as input to the Mie core algorithm are computed as follows. The weighted summation spectrometer readouts (before tripod obscuration correction) for the atmospheric signal and for the internal reference are given by analogues of the quantities LID_0, INT_0 from L1B DPM [RD2] eqn (53) and (54):

$$LID_{atm}(i,j) = \sum_{k=1}^{N} W(i,k) LID_{atm}(i,j,k) \quad \textbf{Eq. 14-1}$$

$$LID_{int}(i,j) = \sum_{k=1}^{N} W(i,k) LID_{int}(j,k) \quad \textbf{Eq. 14-2}$$

14.2. Computation of the raw response (peak location) via the Mie core algorithm

The Mie core algorithm is applied to produce the raw response values (location of the Mie fringe peak expressed as pixel number) for the atmospheric and internal reference:

$$MR_{atm}^{raw}(i) = Mie\_Core[LID_{atm}(i,j)] \quad \textbf{Eq. 14-3}$$

$$MR_{int}^{raw}(i) = Mie\_Core[LID_{int}(i,j)] \quad \textbf{Eq. 14-4}$$

Further details on the Mie core algorithm are given in Chapter 16. It produces more output that just the peak location e.g. peak height, FWHM, offset and SNR.

14.3. Spectral nonlinearity correction

Spectral nonlinearity correction takes the same form as L1B DPM [RD2] eqn (57) and (58):

$$MR_{atm}^{lin}(i) = MR_{atm}^{raw}(i) - EMR_{atm}[MR_{atm}^{raw}(i)] \quad \textbf{Eq. 14-5}$$

$$MR_{int}^{lin}(i) = MR_{int}^{raw}(i) - EMR_{int}[MR_{int}^{raw}(i)] \quad \textbf{Eq. 14-6}$$

The raw responses (peak locations) are corrected to the values that would exist if the response curve was linear using a look-up table of non-linearity corrections as a function of raw response.

14.4. Computation of the atmospheric return LOS velocity component

Computation of a meteorologically-weighted LOS wind velocity component for the atmospheric return is performed as follows using the gradient and intercept of the atmospheric MRC:

$$V_{atm}(i) = \frac{-\lambda_0}{2}\left(\frac{MR_{atm}^{lin}(i)-\alpha}{MRS_{atm}}\right) \quad \textbf{Eq. 14-7}$$

This is used as input to post Mie core HLOS wind component computation (Section 14.6).

14.5. Computation of post Mie core algorithm correction terms

Two steps are described in this section, in preparation for use in the post Mie core algorithm HLOS wind component (Section 14.6).

14.5.1. Computation of the weighted internal reference LOS velocity component

Computation of a meteorologically-weighted internal reference LOS velocity component $V_{INT}(i)$ is given as follows by using the gradient and intercept of the internal reference MRC:

$$V_{int}(i) = \frac{-\lambda_0}{2}\left(\frac{MR_{int}^{lin}(i)-\delta}{MRS_{int}}\right) \quad \textbf{Eq. 14-8}$$

This is used as input to post Mie core algorithm HLOS wind component computation (Section 14.6).

14.5.2. Computation of a weighted spacecraft LOS velocity component

The normalised weighted spacecraft LOS velocity component is given by:

$$V_{sat}(i) = \frac{\sum_{k=1}^{N} W(i,k)V_{sat,m}(k)}{\sum_{k=1}^{N} W(i,k)} \quad \textbf{Eq. 14-9}$$

This is used as input to post Mie core algorithm HLOS wind component computation (Section 14.6).

14.5.3. Ground return corrections

The ground correction velocity component term, which accounts for various sources of bias such as errors in LOS pointing knowledge is calculated as follows. This is optionally used in the final (H)LOS wind calculation. It is calculated as follows, with correction components taken directly from the L1B product file. All the components are derived from ground return wind estimates (done by the L1B processor). The ground correction velocity component is:

$$V_{GM}(i) = W_{GM}C_M + W_{HM}fit_{Mie}(lat) \quad \textbf{Eq. 14-10}$$

W_{GM} and W_{HM} are weighting parameters (expected values 0 or 1 and sum 0 or 1) that are basically switches to control the use of the various correction terms. This makes the calculation very flexible for use. The fit_{Mie}(lat) functions are the result of the Harmonic Bias estimation tool. C_M is the Mie ground correction factor (the ZWC results of the L1B product without harmonic function filtering). Note the total correction V_{GM} value that is provided in the L1B file can be used instead of re-calculating it from the coefficients in the L2Bp.
The weighting factors and whether or not to use the total L1B correction value are controlled by the various AUX_PAR_2B settings: `ZWC_Params`.

14.5.4. Mie range dependent bias correction

V_{RDB} is velocity component term which corrects for the so-called range dependent bias (RDB). This is calculated in the same way as for the L1B processor (see L1B DPM section 15.2.2) i.e. assuming the RDB varies linearly with satellite to target range (distance):

$$V_{RDB}^{Mie}(i) = \frac{\lambda_0}{2} K_{Mie}(r(i) - r_{ref}) \quad \textbf{Eq. 14-11}$$

Use of this correction term is optional and is controlled via the AUX_PAR_2B switch: `RDB_Params/Do_Mie_RDB_corr`

This is used as input to post Mie core algorithm HLOS wind component computation (Section 14.6).

14.5.5. Mie M1 mirror temperature based wind bias correction

This is a placeholder for the description of this algorithm. To be detailed later.

14.6. Computation of the post Mie core algorithm HLOS wind component

The Mie channel HLOS wind component observation is given by the analogue of L1B DPM [RD2] eqn (62):

$$v_{HLOS}(i) = \frac{1}{\sin(\varphi)}\left(V_{atm}(i) - V_{int}(i) - V_{sat}(i) - V_{GM}(i) - V_{RDB}^{Mie}(i)\right) \quad \textbf{Eq. 14-12}$$

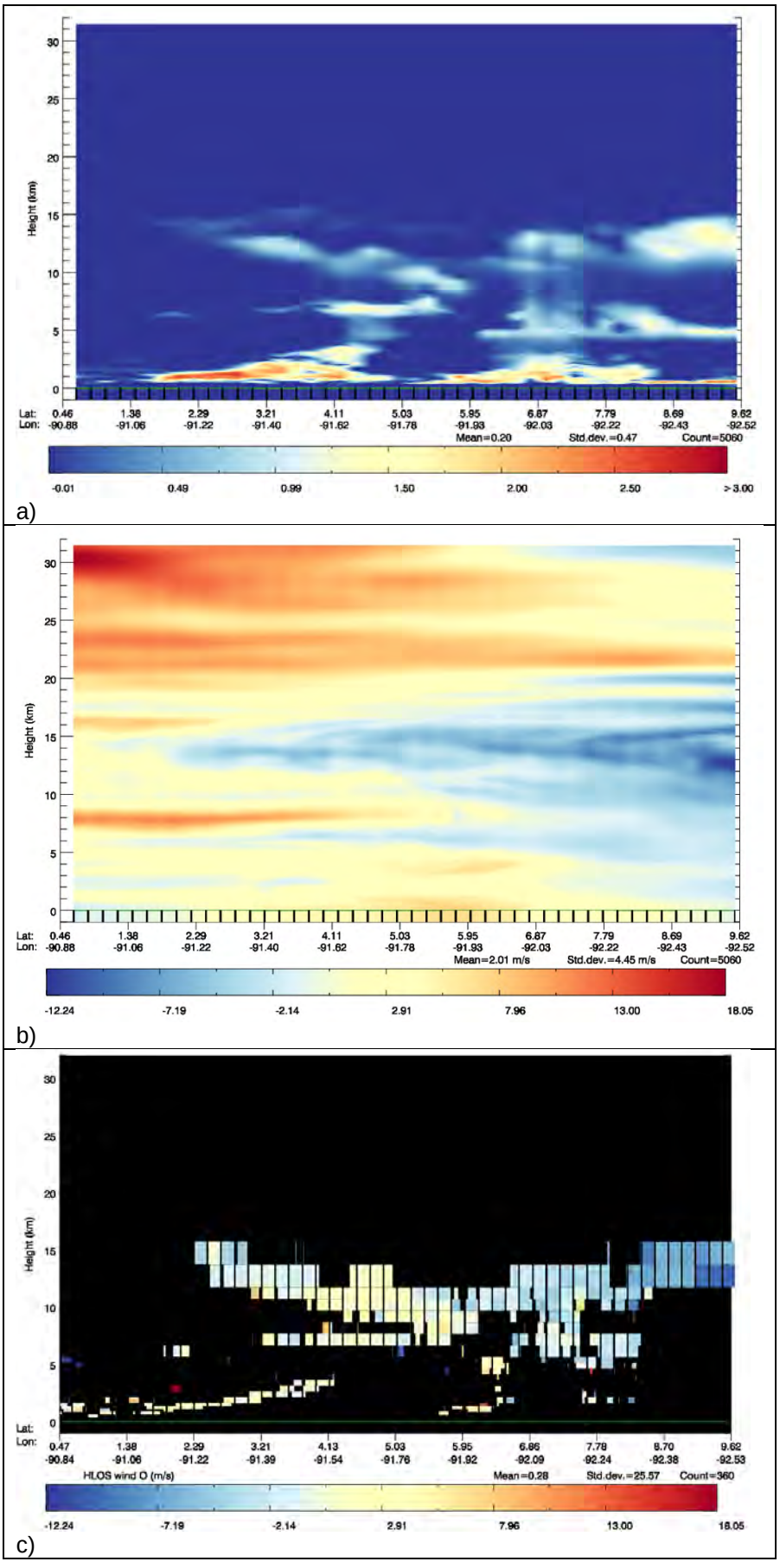
This is the main L2B Mie result reported in L2B products: i.e. the HLOS wind component observation. The first two terms in the bracket account for the difference in frequency between emitted and backscattered light i.e. the Doppler shift. Other terms are basically corrections so that we are only left with a Doppler shift due to atmospheric wind. That is, residual Doppler shift due to the satellite to Earth velocity is subtracted (although this should be small if the satellite pointing is working well) and terms to account for various biases which can be measured using ground returns (which are a zero-wind reference) and the so-called range-dependent bias (an angular effect on the spectrometers).

If one wishes to produce line-of-sight rather than horizontal line-of-sight wind components then the $1/\sin(\varphi)$ factor is removed. This is controlled via the AUX_PAR_2B settings i.e.

- HLOS winds: `<Line_of_Sight_Wind_Flag>false</Line_of_Sight_Wind_Flag>`
- LOS winds: `<Line_of_Sight_Wind_Flag>true</Line_of_Sight_Wind_Flag>`

To give an impression of what the Aeolus Mie L2B winds look like, Figure 9Figure 8 shows an example from the same simulation scenario as used in Figure 7 and Figure 8. Some key parameters used:

- Grouping algorithm was used with the advanced method with a maximum accumulation set to 20 km
- Classification was done with L1B refined scattering ratio with a constant threshold of 1.25. In this example the classification has done a fairly good job, hence the Mie-clear winds are mostly of poor quality.


Height (km)
Lat: 0.46 1.38 2.29 3.21 4.11 5.03 5.95 6.87 7.79 8.69 9.62
Lon: -90.88 -91.06 -91.22 -91.40 -91.62 -91.78 -91.93 -92.03 -92.22 -92.43 -92.52
Mean=0.20 Std.dev.=0.47 Count=5060
-0.01 0.49 0.99 1.50 2.00 2.50 >3.00
a)
Height (km)
Lat: 0.46 1.38 2.29 3.21 4.11 5.03 5.95 6.87 7.79 8.69 9.62
Lon: -90.88 -91.06 -91.22 -91.40 -91.62 -91.78 -91.93 -92.03 -92.22 -92.43 -92.52
Mean=2.01 m/s Std.dev.=4.45 m/s Count=5060
-12.24 -7.19 -2.14 2.91 7.96 13.00 18.05
b)
Height (km)
Lat: 0.47 1.38 2.29 3.21 4.13 5.03 5.95 6.86 7.78 8.70 9.62
Lon: -90.84 -91.06 -91.22 -91.39 -91.54 -91.76 -91.92 -92.09 -92.24 -92.38 -92.53
HLOS wind O (m/s) Mean=0.28 Std.dev.=25.57 Count=360
-12.24 -7.19 -2.14 2.91 7.96 13.00 18.05
c)

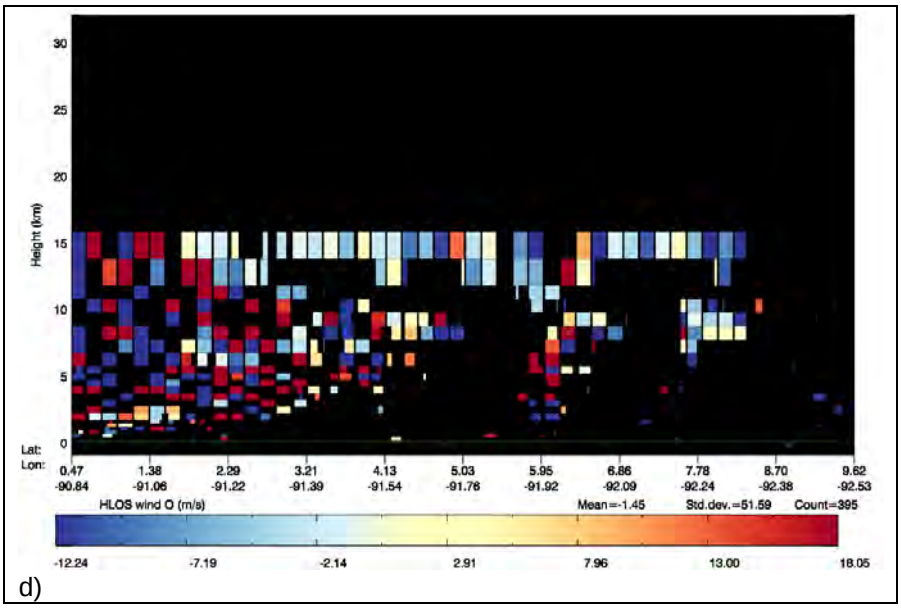


Figure 9. Realistic simulation (using E2S with TcO1279 ECMWF model input) of Aeolus for a cloudy atmosphere. Lidar cross-section plots of a) input "truth" logarithm of scattering ratio i.e. showing where cloud is b) input "truth" HLOS wind c) L2B Mie-cloudy wind observations d) L2B Mie-clear wind observations.

15. The ILIAD retrieval − for Rayleigh channel

ILIAD stands for IMPACT of LINE SHAPE on AEOLUS-ADM DOPPLER ESTIMATES. This Section concerns the Rayleigh-Brillouin correction of the winds for the effects of temperature and pressure. Section 15.1 describes the method of inverting the Rayleigh response to a Doppler frequency shift, given the *a priori* temperature and pressure information (from the AUX_MET_12 file) and the scattering ratio estimate (from either the L1B SR estimates or derived from the OPC). The ILIAD method is referred to in called via the algorithms of Section 13. Section 15.2 describes the inversion for the internal reference LOS velocity; this is referred to in section 13.4.

15.1. Rayleigh channel atmospheric response inversion

Inputs:

Symbol	Description	Input variable source
freq	One dimensional array of size n_{freq} between [-*FSR*,+*FSR*] e.g. =[-10.95 GHz, +10.95 GHz][19] with a step *df* = 25 MHz.	AUX_RBC_L2: `Rayleigh_Brillouin_ADS/F_FP`
T_A, T_B	One-dimensional arrays of transmission curves for the Fabry-Pérot interferometer channels A and B, defined for frequencies *freq*	AUX_RBC_L2: `Rayleigh_Brillouin_ADS/{TA_FP,TB_FP}`
P_{grid}	One-dimensional array of pressures. Of size n_P between e.g. [P_{min} = 10 hPa, P_{max} = 1100 hPa] with a step *dP* = 50 hPa. [20]	AUX_RBC_L2: `Rayleigh_Brillouin_ADS/P_grid`
T_{grid}	One-dimensional array of temperatures. Of size n_T between e.g. [T_{min} = 170 K, T_{max} = 330 K] with a step *dT* = 1 K	AUX_RBC_L2: `Rayleigh_Brillouin_ADS/T_grid`
RR0	One-dimensional array of Rayleigh responses of size *nRR*, covering e.g. between [-0.5, +0.5] with a step *dRR*=0.01	AUX_RBC_L2: `Rayleigh_Brillouin_ADS/RR`
fr	One dimensional array of size n_{fr} between [-*FSR-USR/2*, *+FSR+USR/2*] e.g. = [-11.70 GHz, +11.70 GHz] with a step *df*	AUX_RBC_L2: `Rayleigh_Brillouin_ADS/F_Gridtmp`
S	Three-dimensional array of pre-calculated Rayleigh-Brillouin spectrum estimates of size n_P . n_T . n_{fr} defined for: • Pressures P_{grid} • Temperatures T_{grid}	AUX_RBC_L2: `Rayleigh_Brillouin_ADS/Spec_Grid_PTF/ Spec_Grid_TF/Spec_Grid_F`

[19] FSR (free spectral range) values can vary, values here are for example.
[20] The pressure, temperature and frequency grids can vary with the AUX_RBC_L2 file.

	• Frequencies *fr*	
F	Three-dimensional array of frequencies of size n_P . n_T . n_{RR}, defined for • Pressures P_{grid}, • Temperatures T_{grid}, • Rayleigh responses *RR0*	AUX_RBC_L2: `Rayleigh_Brillouin_ADS/Fcalib_PTR/Fcalib_TR/Fcalib_R`
fd	One dimensional array of size *nfd* covering the range of potential Doppler shifts due to wind[21], i.e. detector frequencies between [-*USR*/2, +*USR*/2] e.g. = [-0.75 GHz, +0.75 GHz], with a step *df*	AUX_RBC_L2: `Rayleigh_Brillouin_ADS/Fd`
RRmes	One dimensional array (size *n*) of weighted atmospheric Rayleigh response(s) i.e. measurements.	Calculated in L2B processing; see Section 13.1
ρ	One dimensional array (size *n*) of weighted scattering ratio estimate(s) derived from measurements	Calculated in L2B processing; see Section 13.2
Pmod	One dimensional array (size *n*) of atmospheric pressure estimate(s) at observation geolocation	Calculated in L2B processing; see Section 12.4
Tmod	One dimensional array (size *n*) of atmospheric temperature estimate(s) at observation geolocation	Calculated in L2B processing; see Section 12.4

Outputs:

Symbol	Description	ALD_U_N_2B file variable
$\boldsymbol{V_{ILIAD}}$	Atmospheric return LOS wind component (corrected for temperature, pressure and scattering ratio effects)	N/A since only a temporary variable in L2B processing which is used in Section 13.3
$\frac{\partial V_{ILIAD}(i)}{\partial T}, \frac{\partial V_{ILIAD}(i)}{\partial p}, \frac{\partial V_{ILIAD}(i)}{\partial RR}, \frac{\partial V_{ILIAD}(i)}{\partial \rho}$	Partial derivative of atmospheric LOS wind with respect to temperature, pressure, Rayleigh response and scattering ratio	N/A since only a temporary variable in L2B processing which is used in Section 13 i.e. possibly converted to HLOS wind sensitivities.

Algorithm:

In what follows below we define an index *i* and loop from 1 to *n*.

<u>Correction of the Rayleigh signals for temperature and pressure effects:</u>

[21] 1 MHz corresponds to a LOS wind change of 0.18 m/s for Aeolus's wavelength.

The algorithm involves finding the frequency in the look-up table (F) grid that is closest to *RRmes*, *Tmod* and *Pmod* and then doing a linear interpolation to the actual values of *RRmes*, *Tmod* and *Pmod*.

The algorithm searches the arrays:

- P_{grid} for the closest point to $P_{mod}(i)$; iP is found ($n_P \geq iP \geq 1$)
- P_{grid} for the closest point immediately above $P_{mod}(i)$; iP^+ is found ($n_P \geq iP^+ \geq 1$)
- P_{grid} for the closest point immediately below $P_{mod}(i)$; iP^- is found ($n_P \geq iP^- \geq 1$)

- T_{grid} for the closest point to $T_{mod}(i)$; iT is found ($n_T \geq iT \geq 1$)
- T_{grid} for the closest point immediately above $T_{mod}(i)$; iT^+ is found ($n_T \geq iT^+ \geq 1$)
- T_{grid} for the closest point immediately below $T_{mod}(i)$; iT^- is found ($n_T \geq iT^- \geq 1$)

- $RR0$ for the closest point to $RR_{mes}(i)$; iR is found ($n_{RR} \geq iR \geq 1$)
- $RR0$ for the closest point immediately above $RR_{mes}(i)$; iR^+ is found ($n_{RR} \geq iR^+ \geq 1$)
- $RR0$ for the closest point immediately below $RR_{mes}(i)$; iR^- is found ($n_{RR} \geq iR^- \geq 1$)

The algorithm then numerically estimates the following partial derivatives:

- $$\alpha_P(i) = \frac{\partial f}{\partial P} = \frac{F(iP^+, iT, iR) - F(iP^-, iT, iR)}{P_{grid}(iP^+) - P_{grid}(iP^-)}$$ **Eq. 15-1**

- $$\alpha_T(i) = \frac{\partial f}{\partial T} = \frac{F(iP, iT^+, iR) - F(iP, iT^-, iR)}{T_{grid}(iT^+) - T_{grid}(iT^-)}$$ **Eq. 15-2**

- $$\alpha_{RR}(i) = \frac{\partial f}{\partial RR} = \frac{F(iP, iT, iR^+) - F(iP, iT, iR^-)}{RR0(iR^+) - RR0(iR^-)}$$ **Eq. 15-3**

The atmospheric temperature- and pressure-corrected frequency offset for the atmospheric response RR_{mes}(i) is then calculated as (tangent-linear approximation, between grid-points):

$$f_{\mathrm{mod}} = F(iP, iT, iR) + \alpha_T(i) \cdot \left(T_{\mathrm{mod}}(i) - T_{grid}(iT)\right) + \alpha_P(i) \cdot \left(P_{\mathrm{mod}}(i) - P_{grid}(iP)\right) + \alpha_{RR}(i) \cdot \left(RR_{mes}(i) - RR0(iR)\right)$$ **Eq. 15-4**

<u>Correction of the Mie cross-talk:</u>

It is important to correct the Rayleigh signal for any contamination by particulate (Mie) backscatter signal, as this can affect the measured Rayleigh response significantly. This correction is not done using a look-up table, but does still require some inputs from the AUX_RBC_L2 file.
We form a matrix **S1** of Rayleigh-Brillouin normalised spectrum estimates (for our given *a priori* temperature and pressure values) over the FSR of possible frequencies *freq* (n_{freq} of them) and over the range of possible Doppler shifts on top of this, *fd* (n_{fd} of them):

$$\mathbf{S1} = \begin{pmatrix} S(iP, iT, 1) & \cdots & S(iP, iT, n_{freq}) \\ \vdots & \ddots & \vdots \\ S(iP, iT, n_{fd} + 1) & \cdots & S(iP, iT, n_{freq} + n_{fd} - 1) \end{pmatrix}$$ **Eq. 15-5**

where the three-dimensional array *S* was calculated beforehand (stored in AUX_RBC_L2 file) using a parameterised version of the Tenti S6 model. We then do matrix multiplication of **S1** by the calibration

transmission curves (this is effectively a numerical integration of Rayleigh-Brillouin spectra with the calibration response curves) to get the counts.[22]

$$\underline{N}_A^{(1)} = df \cdot \left(\underline{\underline{S1}} \cdot \underline{T}_A\right)$$
$$\underline{N}_B^{(1)} = df \cdot \left(\underline{\underline{S1}} \cdot \underline{T}_B\right) \qquad \textbf{Eq. 15-6}$$

Which are vectors of size n_{fd}. We then calculate the Rayleigh response without yet accounting for the Mie signal in the Rayleigh i.e. as if the scattering ratio is 1

$$\underline{R}^{(1)} = \frac{\underline{N}_A^{(1)} - \underline{N}_B^{(1)}}{\underline{N}_A^{(1)} + \underline{N}_B^{(1)}} \text{ (vector of size } n_{fd}) \qquad \textbf{Eq. 15-7}$$

We then interpolate the range of frequencies *-fd* from the calculated responses $R^{(1)}$ to the measured response *RRmes*(i) value using spline interpolation to obtain a frequency value *fR1* to which the measured *RRmes(i)* value applies. Finally, we calculate the expected calibration response values without Mie contribution $T1_A$ and $T1_B$ valid at frequency *fR1* by interpolating the arrays T_A and T_B (respectively) from the frequency responses *freq* to *fR1*.

In order to calculate the frequency responses with a Mie contribution (cross-talk), we first form the following matrix, which is the spectra of the Rayleigh and the Mie signal added together (here (ρ_0-1) is the ratio of particulate backscatter to molecular backscatter i.e. $\rho_0$is the "measured" scattering ratio):

$$\mathbf{S2} = \mathbf{S1} + (\rho_0 - 1)\begin{pmatrix} g_{Mie}(1) & \cdots & g_{Mie}(n_{freq}) \\ \vdots & \ddots & \vdots \\ g_{Mie}(n_{fd}+1) & \cdots & g_{Mie}(n_{freq}+n_{fd}-1) \end{pmatrix}$$

Eq. 15-8

where we have prior to this calculated the following parameters from the laser wavelength λ_0:

$$FWHM = \frac{c \cdot 0.02 \cdot 10^{-12}}{\lambda_0^2} \text{ (c is the speed of light), } \sigma_{Mie} = \frac{FWHM}{2 \cdot \sqrt{\log(2)}} \qquad \textbf{Eq. 15-9}$$

This FWHM value appears to be hardcoded, does it depend on the Mie characterisation settings, therefore should it be flexible? The FWHM value in E2S is 50 MHz, whereas this calculation gives 47.57 MHz. Trying 50 MHz makes very little difference: of order 1.0e-5 m/s in HLOS wind.

The Mie contribution to the spectrum is assumed to be a normalised Gaussian function with variance $\sigma_{Mie}^2/2$ i.e.

$$g_{Mie} = \frac{1}{\sigma_{Mie} \cdot \sqrt{\pi}} \cdot \exp\left(-\frac{fr^2}{\sigma_{Mie}^2}\right) \qquad \textbf{Eq. 15-10}$$

and is a vector of size *nfr* defined for frequencies *fr*.

We apply the similar steps as previously, but to create the Rayleigh response that would be obtained with Mie contamination, namely:

$$\underline{N}_A^{(2)} = df \cdot \left(\underline{\underline{S2}} \cdot \underline{T}_A\right)$$
$$\underline{N}_B^{(2)} = df \cdot \left(\underline{\underline{S2}} \cdot \underline{T}_B\right) \qquad \textbf{Eq. 15-11}$$

[22] Vectors have been indicated by a single underline, and matrices by a double underline.

$$\underline{R}^{(2)} = \frac{\underline{N}_A^{(2)} - \underline{N}_B^{(2)}}{\underline{N}_A^{(2)} + \underline{N}_B^{(2)}} \text{ (vector of size } n_{fd}\text{)} \qquad \textbf{Eq. 15-12}$$

We then interpolate the range of frequencies *–fd* from the calculated responses $R^{(2)}$ to the measured response *RRmes*(i) using spline interpolation and obtain a frequency value *fR2*. Finally, we calculate the expected counts with Mie contribution $N2_A$ and $N2_B$ given the response *fR2* by interpolating the arrays $N_A^{(2)}$ and $N_B^{(2)}$ (respectively*)* from the detector frequencies *fd* to frequency *fR2*.

We then estimate the partial derivative of the frequency offset with respect to the scattering ratio (at scattering ratio $\rho = \rho_0$ using:

$$\left.\frac{\partial f}{\partial \rho}\right|_{\rho=\rho_0} = \frac{\left.\frac{\partial R}{\partial \rho}\right|_{\rho=\rho_0}}{\left.\frac{\partial R}{\partial f}\right|_{\rho=\rho_0}} \qquad \textbf{Eq. 15-13}$$

where $\left.\frac{\partial R}{\partial \rho}\right|_{\rho=\rho} \approx \frac{T1_A - T1_B - RR_{mes}(i)\cdot\left(T1_A + T1_B\right)}{N2_A + N2_B}$

Eq. 15-14

see [Appendix E. Derivation of the sensitivity of HLOS wind to Mie contamination] and

$$\frac{\partial R}{\partial f} = \frac{Rf2^+ - Rf2^-}{fR2^+ - fR2} \qquad \textbf{Eq. 15-15}$$

where $Rf2^+$ ($Rf2^-$) is obtained by interpolating $R^{(2)}$ from *fd* to *fR2*+1 (*fR2*, respectively).

In addition to the earlier correction for pressure and the temperature effects, the atmospheric frequency offset is now further corrected for the Mie contamination to produce a frequency value that only depends on the atmospheric wind. The dependence of the frequency offset upon scattering ratio (ρ) can be written as a Taylor expansion around $\rho = \rho_0$ as:

$$f(\rho) = f(\rho_0) + \frac{\partial f(\rho_0)}{\partial \rho}(\rho - \rho_0) + \frac{1}{2}\frac{\partial^2 f(\rho_0)}{\partial \rho^2}(\rho - \rho_0)^2 + \cdots \qquad \textbf{Eq. 15-16}$$

To determine the frequency offset without Mie contamination i.e. at $\rho = 1$ the above equation can be determined at f(1). Hence the corrected frequency offset is (to first order only):

$$f_{cor} = f_{\text{mod}} + \left(1 - \rho_0\right)\left.\frac{\partial f}{\partial \rho}\right|_{\rho=\rho_0} \qquad \textbf{Eq. 15-17}$$

A first order correction appears to be sufficient for Rayleigh-clear winds with $\rho < 1.8$
Hence the main outputs of the ILIAD scheme are produced: the LOS wind velocity component is computed from the Doppler frequency shift as:

$$V_{ILIAD}(i) = -f_{cor} \cdot \frac{\lambda_0}{2} \qquad \textbf{Eq. 15-18}$$

Also, the partial derivatives of frequency with respect to temperature, pressure, scattering ratio and Rayleigh response are multiplied by the same frequency to LOS wind factor to obtain the partial derivatives of LOS wind with respect to temperature, pressure, scattering ratio and Rayleigh response.

15.2. Internal Reference response inversion

Inputs:

Symbol	Description	Input variable source
RR0	One-dimensional array of Rayleigh responses of size *nRR*, covering e.g. between [-0.5, +0.5] with a step *dRR*=0.01	AUX_RBC_L2: `Rayleigh_Brillouin_ADS/RR`
F_{ref}	One-dimensional array of frequencies (of size *nRR)* defined for Rayleigh responses *RR0* for the internal reference	AUX_RBC_L2: `Rayleigh_Brillouin_ADS/Fint_R`
RR_{ref}	One dimensional array (size *n*) of weighted internal reference Rayleigh response(s)	Calculated in L2B processing; see Section 13.4.113.1

Outputs:

Symbol	Description	ALD_U_N_2B file variable
V_{INT}	Internal reference LOS wind component	N/A since only a temporary variable in L2B processing which is used in Section 13.4.1
$\frac{\partial V_{INT}(i)}{\partial RR_{ref}}$	Partial derivative of internal reference LOS wind with respect to internal reference Rayleigh response	N/A since only a temporary variable in L2B processing which is used in Section 13 i.e. possibly converted to HLOS wind sensitivities.

Algorithm:

In what follows below we define an index *i* and loop from 1 to *n*.

The algorithm searches the array *RR0*:

- for the closest point to $RR_{ref}(i)$; iR is found ($n_{RR} \geq iR \geq 1$)
- for the closest point immediately above $RR_{ref}(i)$; iR^+ is found ($n_{RR} \geq iR^+ \geq 1$)
- for the closest point immediately below $RR_{ref}(i)$; iR^- is found ($n_{RR} \geq iR^- \geq 1$)

The algorithm then calculates the following partial derivative:

$$\alpha_{ref}(i) = \frac{\partial f}{\partial RR_{ref}} = \frac{F_{ref}(iR^+) - F_{ref}(iR^-)}{RR0(iR^+) - RR0(iR^-)} \qquad \textbf{Eq. 15-19}$$

The frequency for the internal reference channel response RR_{ref}(i) is then calculated as (tangent-linear approximation between the grid points):

$$f_{ref} = F_{ref}(iR) + \alpha_{ref}(i) \cdot \left(RR_{ref}(i) - RR0(iR)\right) \qquad \textbf{Eq. 15-20}$$

The internal reference channel LOS velocity component is computed from the Doppler frequency shift as:

$$V_{INT}(i) = -f_{ref} \cdot \frac{\lambda_0}{2} \qquad \textbf{Eq. 15-21}$$

16. The Mie core algorithm

The L2B Mie core algorithm is a Fortran code implementation of the L1B Mie core algorithm Option-2 ("Downhill-Simplex"). Each set of ACCD data passed to the Mie core algorithm are fitted to a Lorentzian curve in order to derive the best estimate for the Mie fringe location, fringe width, offset and peak height.

Inputs:

Symbol	Description	Input variable source
$LID_{atm}(i,j)$	Atmospheric spectrometer counts at observation level	ALD_U_N_1B: `Measurement_ADS/List_of_Mie_Measurement_Data/Mie_Measurement_Data`
$LID_{int}(j)$	Internal Reference spectrometer counts at observation level	ALD_U_N_1B: `Measurement_ADS/List_of_Mie_Reference_Pulse/Mie_Reference_Pulse`
$Start\_FWHM$, res_err_thresh, $max\_it\_lor\_fit$, $FWHM\_upper\_thresh$ $FWHM\_lower\_thresh$ ph_upper_thresh, ph_lower_thresh, x_thresh, $nonlin\_opt\_thresh$, $max\_it\_nonlin\_opt$, NS	Mie core algorithm processing parameters	AUX_PAR_2B: `WVM_Params/Common_Processing_Params/Mie_Core_Algorithm_Params{,Reference_Pulse}/{Start_FWHM, Residual_Error_Threshold, Max_Iterations_Lorentz_Fit, FWHM_Upper_Threshold, FWHM_Lower_Threshold, Peak_Height_Upper_Threshold, Peak_Height_Lower_Threshold, Peak_Location_Threshold, Nonlinear_Optimization_Threshold, Max_Iterations_Nonlinear_Optimization, Num_Spectral_Sub_Samples}`

Outputs:

Symbol	Description	ALD_U_N_2B file variable
$x, peak\_height, offset, FWHM$	Solution Lorentzian fit parameters: peak location, peak height, offset and full-width half-maximum	`Mie_Wind_PCD_ADS%Mie_Wind_QC/{{IntRef}Fitting_Amplitude, {IntRef}Fitting_Offset,{IntRef}Fitting_FWHM}` N.B. the peak location is not stored in the L2B product. Of course it is passed internally in the L2B processor as an output of the Mie core algorithm to the Mie wind processing of Section 14.2

In the following derivations the index *j* refers to the ACCD pixel number.

16.1. Detection Chain Offset subtraction

The Detection Chain Offset (DCO) correction is applied to both atmospheric and internal reference data. It uses a weight factor, *w*, to calculate the offset value using both pixels 19 and 20 of the ACCD readout:

$$LID(j) = LID(j) - LID_{offset}$$
$$LID_{offset} = wLID(20) + (1-w)LID(19) \quad \textbf{Eq. 16-1}$$

where $0 \leq w \leq 1$.
This step is more flexible than the L1B processor calculation.

16.2. Telescope Obscuration correction

Telescope Obscuration correction needs to be applied for only the atmospheric data (i.e. not for the internal reference data):

$$LID_{atm}(j) = LID_{atm}(j)/TOBS(j) \quad \textbf{Eq. 16-2}$$

16.3. Background noise correction

The removal of the background noise level is performed by subtracting the lowest signal level found in the useful pixel range (from *j*=3 to 18):

$$LID(j) = LID(j) - \min(LID(3{:}18)) \quad \textbf{Eq. 16-3}$$

The minimum value is stored as LID_{min}

16.4. Normalise the signal

The signal is then normalised by dividing by the largest signal level found in the useful pixel range (from *j*=3 to 18):

$$LID(j) = LID/\max(LID(3{:}18)) \quad \textbf{Eq. 16-4}$$

The largest value and its index are stored as: LID_{max} and j_{max}

16.5. Find the first guess at the peak location

The first guess at the peak location (x_1) is calculated from a weighted mean of the three positions near the maximum:

$$x_1 = \frac{(j_{max}-1)LID(j_{max}-1)+j_{max}LID(j_{max})+(j_{max}+1)LID(j_{max}+1)}{LID(j_{max}-1)+LID(j_{max})+LID(j_{max}+1)} \quad \textbf{Eq. 16-5}$$

If j_{max} happens to be at position 3 resp. 18 use $LID(18)$ resp. $LID(3)$ for $j_{max}-1$ resp. $j_{max}+1$

16.6. Iterative optimisation search

Within this iterative approach no starting values for $peak\_height$ and $offset$ are needed, as they are calculated with a linear system of equations during the process. The starting value for the full-width-half-maximum $FWHM$ should depend on pre-determined instrument characteristics. Starting with initial estimates $x_i = x_1$ and $FHWM_i = Start\_FWHM$

Iterate with counter *i* the following calculations:

- $res\_error_i = \sum_{j=3}^{18}\left[\frac{1}{NS}\sum_{k=1}^{NS}\frac{peak\_height_i FWHM_i^2}{4(x_i - j - \frac{k}{NS}+0.5+1/(2NS))^2+FWHM_i^2} + offset_i - LID(j)\right]^2$
  Eq. 16-6
  Where *NS* = number of sub samples over a pixel. Stop the iteration if the absolute difference between the residual error expressed in the above equation and the residual error of the previous iteration is below $\boldsymbol{res\_err\_thresh}$ or *i* > $\boldsymbol{max\_it\_lor\_fit}$
- For each pixel *j* calculate:

$$f_j = \frac{1}{NS}\sum_{k=1}^{NS} \frac{FWHM_i^2}{4(x_i - j - \frac{k}{NS} + 0.5 + 1/(2NS))^2 + FWHM_i^2}$$ **Eq. 16-7**

- Solve the linear system of equations for the iterative parameters $peak\_height_i$ and $offset_i$

$$\begin{pmatrix} \sum_{j=3}^{18} f_j^2 & \sum_{j=3}^{18} f_j \\ \sum_{j=3}^{18} f_j & 16 \end{pmatrix} \begin{pmatrix} peak\_height_i \\ offset_i \end{pmatrix} = \begin{pmatrix} \sum_{j=3}^{18} LID(j) f_j \\ \sum_{j=3}^{18} LID(j) \end{pmatrix}$$ **Eq. 16-8**

- Apply a Downhill Simplex nonlinear optimization procedure [Press et al., Numerical Recipes in C++, Cambridge University Press 2003] to find the best approximations for peak location x_{i+1} and $FHWM_{i+1}$. The target function in this process is given by the same formula as in Eq.16-6.
  - The starting points for the Downhill Simplex are given by:

$$P1 = (x_i, FWHM_i); P2 = (1 + x_i, FWHM_i); P3 = (x_i, 1 + FWHM_i)$$ **Eq. 16-9**

  - The termination parameters for the Downhill simplex algorithm are when the thresholds are exceeded: $\boldsymbol{nonlin\_opt\_thresh}$ and $\boldsymbol{max\_it\_nonlin\_opt}$

16.7. Quality control flags

A check that reliable solution values have been found is performed. Reliable if:

- $ph\_lower\_thresh < peak\_height < ph\_upper\_thresh$
- $FWHM\_lower\_thresh < FWHM < FWHM\_upper\_thresh$
- $|x - j_{max}| < x\_thresh$

16.8. Renormalise peak height and offset

The peak height and the offset are returned to the original scale. Also for the background noise estimate is added back to the offset.

$$peakheight_{ACCD} = LID_{max} * peakheight_{relative}$$ **Eq. 16-10**

$$offset = offset * LID_{max} + LID_{min}$$ **Eq. 16-11**

17. Rayleigh channel uncertainty estimation processing

The recommended L2B Rayleigh HLOS wind error quantifier algorithm is Option 1 given in Section 17.1. An alternative, Option 2 is based on the approach taken for L1B HLOS wind error quantifiers, is retained as a back-up option (Section 17.2). ***Option 2 is not currently recommended due to using hardcoded parameters.***

Which algorithm to use is controlled via the AUX_PAR_2B file; in particular: `Error_Quantifier_Params/ErrorQuantMethod_Rayleigh`

- Option 1: `ErrorQuantMethod_Ray_iliad_sens`
- Option 2: `ErrorQuantMethod_Ray_1Bweighted`

17.1. Option 1: The Rayleigh HLOS wind error estimate (using output of the ILIAD algorithm)

This is the default Rayleigh HLOS wind error estimate algorithm.

Inputs:

Symbol	Description	Input variable source
$A_m(i,k), B_m(i,k)$	Measurement-bin Useful Signal in Channels A and B	ALD_U_N_1B: `Useful_Signal_MDS/Measurement_Useful_Signal/ Rayleigh_Altitude_Bin_Useful_Signal_Info / Useful_Signal_Channel_{A,B}`
$SNR_{A,m}(i,k)$, $SNR_{B,m}(i,k)$	Measurement-bin signal-to-noise (SNR) in Channels A and B for atmospheric signal	ALD_U_N_1B: `Product Confidence Data ADS/List_of_Measurement_PCDs/ List_of_Meas_Alt_Bin_PCDs/ Rayleigh_Signal_to_Noise_Ratio_Channel_{A,B}`
$C_m(i,k), D_m(i,k)$	Measurement-bin Internal Reference Signal in Channels A and B	ALD_U_N_1B: `Measurement_ADS/ List_of_Rayleigh_Reference_Pulse_{A,B}/ Rayleigh_Reference_Pulse`
$SNR_{C,m}(i,k)$, $SNR_{D,m}(i,k)$	Measurement-bin signal-to-noise (SNR) in Channels A and B for internal reference signal	ALD_U_N_1B: `Product Confidence Data ADS/List_of_Measurement_PCDs/Rayleigh_Ref_Pulse_Signal_to_Noise_Ratio_Channel_{A,B}`
ρ_0	The observation level scattering ratio used in the ILIAD scheme.	L2B processor internal variable, see section 13.2
λ_0	Laser wavelength	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/ Satellite_Characterisation_Data/ Laser_Wavelength`
φ	Incidence angle of the	Derived from elevation angles of range-bins from ALD_U_N_1B, see section 12.1:

	line of sight (LOS) of the instrument at vertical centre-of-gravity of the range-bin	`Geolocation ADS/ Observation_Rayleigh_Geolocation/List_of _Observation_Geolocation_of_Height_Bins/ Observation_Geolocation_of_Height_Bin/To pocentric_Elevation_of_ Height_Bin`
$\frac{\partial V_{ILIAD}(i)}{\partial T}, \frac{\partial V_{ILIAD}(i)}{\partial p}, \frac{\partial V_{ILIAD}(i)}{\partial RR}, \frac{\partial V_{ILIAD}(i)}{\partial \rho}$	Partial derivative of atmospheric LOS wind with respect to temperature, pressure, Rayleigh response and scattering ratio	L2B processor internal variable, see section 15
$\frac{\partial V_{INT}(i)}{\partial RR_{ref}}$	Partial derivative of internal reference LOS wind with respect to internal reference Rayleigh response	L2B processor internal variable, see section 15
σ_T	Standard error estimate for the reference temperature used in the ILIAD scheme	Potentially provided by AUX_MET_12 data. Currently not used.
σ_p	Standard error estimate for the reference pressure used in the ILIAD scheme	Potentially provided by AUX_MET_12 data. Currently not used.
σ_ρ	Standard error estimate for scattering ratio used in the ILIAD scheme	Source TBD. Currently not used.
$W(i,k)$	Measurement-bin weights for the observation	Output of L2B processing described in Section 10

Outputs:

Symbol	Description	ALD_U_N_2B file variable
$\sigma_{HLOS,tot}$	L2B Rayleigh HLOS wind error quantifier (standard error estimate i.e. 1-sigma) for each observation	`Rayleigh Wind Product Confidence DataADS/Rayleigh_Wind_Product_PC D/Rayleigh_Wind_QC/Hlos_error_es timate`

Algorithm

The standard error of the Rayleigh HLOS wind can be estimated from: the SNR of the useful Rayleigh responses *A* and *B* for both the atmospheric data and internal reference; the consequent error in the atmospheric and internal references responses; the ILIAD scheme that computes the LOS wind from *R,* temperature, pressure and scattering ratio; and the projection of the LOS wind onto the horizontal for the HLOS wind. That is from all the important terms that go into the HLOS wind formula of Section 13.5.

From a sensitivity analysis it follows that the error in Rayleigh response *R=(A-B)/(A+B)* is

$$\Delta R = \frac{\partial R}{\partial A}\Delta A + \frac{\partial R}{\partial B}\Delta B = \frac{2B}{(A+B)^2}\Delta A + \frac{-2A}{(A+B)^2}\Delta B \qquad \textbf{Eq. 17-1}$$

where A_k and B_k are the useful signals on measurement level of channels A and B, and expected standard deviations for ΔA and ΔB, σ_A and σ_B respectively, follow from the signal-to-noise ratio, *SNR*, respectively $\sigma_{A,k} = A_k / SNR(A_k)$ and $\sigma_{B,k} = B_k / SNR(B_k)$, with k the measurement index. This is done for both atmospheric signals and internal reference signals. The L1B provided SNR values account for Poisson noise both of the signal and the UV background (that is subtracted to get the signal for the atmospheric signals only).

Note that A_k and B_k may be negative due to the background subtraction (small risk and resulting (high) winds doubtful). As such, it would have been better to request the estimated error standard deviations from L1B at measurement level, rather than the SNR, for the above equation. Now the L2BP needs to check A_k, B_k, SNR_k < epsilon as a precaution. This is implemented in the screening function.

To obtain L2B observation-level useful signal *A* and *B*, A_k and respectively B_k are accumulated over *N* measurement-bins in a weighted sum, where each of the measurement-bins has a fractional weight *W(i,k)* (Section 12.1). σ_A and σ_B, on L2B observation level, are thus as well obtained from a weighted sum, respectively

$$\sigma_A^2 = \sum_{k=1}^{N} W(i,k)^2 \sigma_{A,k}^2, \quad \sigma_B^2 = \sum_{k=1}^{N} W(i,k)^2 \sigma_{B,k}^2 \qquad \textbf{Eq. 17-2}$$

This computation logically fits in with the other weighting computations (i.e., for *A* and *B*) in the code, and not at HLOS wind error computation level.

For independent errors in *A* and *B* the estimated error in the response *R* , $\sigma_R = \sqrt{<\Delta R\ \Delta R>}$ with <> the operator for the expected covariance, is obtained by

$$\sigma_R = \frac{2}{(A+B)^2}\sqrt{B^2\sigma_A^2 + A^2\sigma_B^2} \qquad \textbf{Eq. 17-3}$$

This is the first step in the HLOS wind error estimate computation.

For the atmospheric signal, the ILIAD scheme computes in the inversion step the line of sight wind, LOS^0, from response *R* , temperature *T* and pressure *p* . Using the sensitivity $\partial LOS^0 / \partial\rho$, for a zero scattering ratio, $\rho = 0$. For the given scattering ratio, $\rho \neq 0$, a linear correction is applied

$$LOS = LOS^0 + \Delta\rho\frac{\partial LOS^0}{\partial\rho} \qquad \textbf{Eq. 17-4}$$

As such that the sensitivity is:

$$\Delta LOS = \frac{\partial LOS^0}{\partial R}\Delta R + \frac{\partial LOS^0}{\partial T}\Delta T + \frac{\partial LOS^0}{\partial p}\Delta p + \Delta\rho\frac{\partial LOS^0}{\partial\rho} \qquad \textbf{Eq. 17-5}$$

where $\partial LOS^0 / \partial T$ and $\partial LOS^0 / \partial p$ are the local sensitivities of the LOS^0 to *T* and *p* respectively. For uncorrelated errors in the scattering ratio, temperature, pressure and the response, the resulting *LOS* error, $\sigma_{LOS} = \sqrt{<\Delta LOS\ \Delta LOS>}$ with <> the operator for the expected covariance, is estimated as

$$\sigma_{LOS} = \sqrt{\left(\frac{\partial LOS^0}{\partial R}\sigma_R\right)^2 + \left(\frac{\partial LOS^0}{\partial T}\sigma_T\right)^2 + \left(\frac{\partial LOS^0}{\partial p}\sigma_p\right)^2 + \left(\frac{\partial LOS^0}{\partial \rho}\sigma_\rho\right)^2} \qquad \textbf{Eq. 17-6}$$

$\partial LOS^0 / \partial R$ is computed in ILIAD and provided to this calculation. $\partial LOS^0 / \partial \rho$ and other sensitivities are available from ILIAD. The terms with the standard deviation of error for temperature and pressure, respectively σ_T and σ_p, are expected to be small, but this needs further testing. A climatological estimate of AUX_MET_12 errors might give σ_T = 1 K and σ_p = 1 hPa, but if the temperature and pressure sensitivities are incorporated in the data assimilation system, these values would have to be reset at that point. As such, it may be relevant to report the individual contributions as well in L2B. Of course the temperature, pressure and scattering ratio sensitivities are irrelevant for the internal reference error propagation and can be neglected from the above equation in that calculation.

In the error budget analysis part of the sensitivity study, KNMI plans to analyse the errors in the diverse backscatter ratio estimates as a first test for the scene classification. For the baseline KNMI suggests to take $\Delta\rho = 0$.

The *LOS* wind error estimate is optionally projected onto the horizontal i.e. the *HLOS* wind error by the given projection angle, where the associated projection error is neglected (order 0.01 m/s if properly specified):

$$\sigma_{HLOS} = \frac{1}{\sin(\varphi)}\sigma_{LOS} \qquad \textbf{Eq. 17-7}$$

Very similar calculations are performed for the atmospheric Rayleigh response (to produce an atmospheric error estimate: $\sigma_{HLOS,atm}$) as for the internal reference Rayleigh response signal (to produce an internal reference error estimate: $\sigma_{HLOS,int}$). The total HLOS wind error estimate is then calculated as the quadratic sum of the atmospheric and internal reference standard deviations (N.B. the internal reference error was found to be non-negligible):

$$\sigma_{HLOS,tot} = \sqrt{\sigma_{HLOS,atm}^2 + \sigma_{HLOS,int}^2} \qquad \textbf{Eq. 17-8}$$

The above derivation considers sources of error from V_{atm} and V_{int} of the HLOS wind derivation in Section 13.5, but does not consider the errors coming from V_{sat}, V_{GR}, V_{RDB} , because it is unclear at this stage how significant they will be compared to the other terms.
An example of the Rayleigh-clear HLOS wind error estimates are given in Figure 10 below:

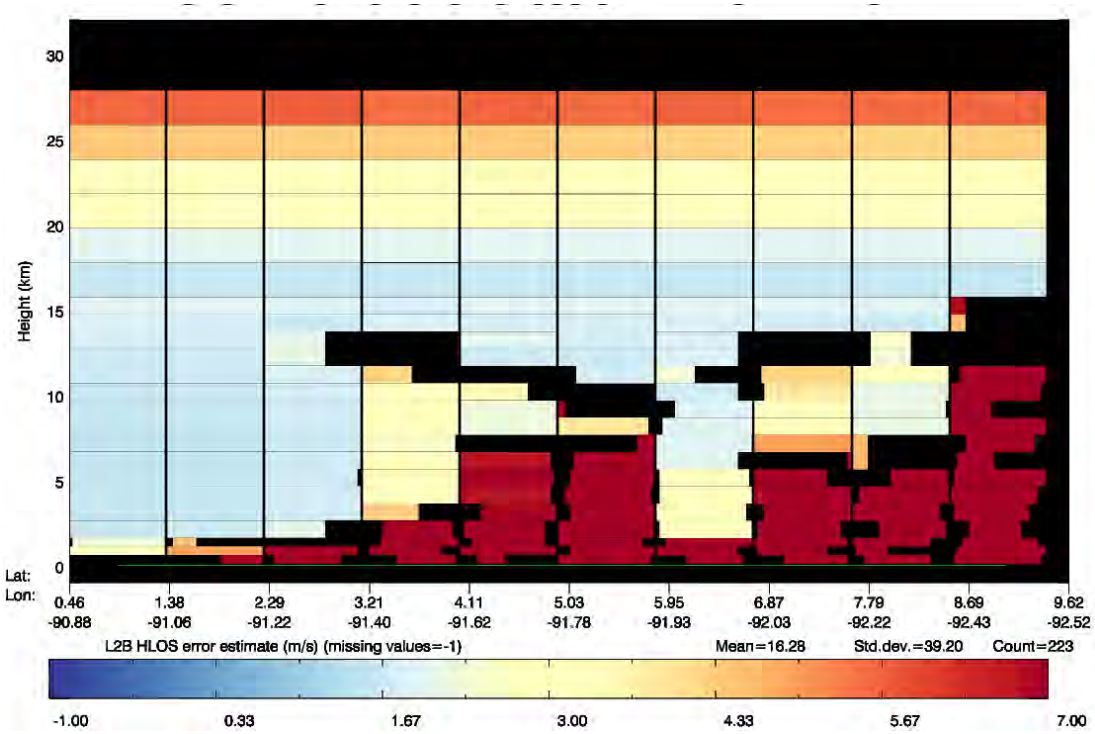


Figure 10. Example of L2B Rayleigh-clear HLOS wind standard error estimates (m/s), for the same simulation scenario as used in Figure 8. Errors estimates are much larger below clouds due to the low transmission of the laser light.

17.2. Option 2: L2B Rayleigh error quantifier based on the L1B processor approach

This method is kept as an option until both algorithms can be verified after launch. Currently it is not recommended due to the use of some hardcoded parameters in the code.

Inputs:

Symbol	Description	Input variable source
$\boldsymbol{A_m(i,k), B_m(i,k)}$	Measurement-bin Useful Signal in Channels A and B	ALD_U_N_1B: `Useful_Signal_MDS/Measurement_Useful_Signal/ Rayleigh_Altitude_Bin_Useful_Signal_Info / Useful_Signal_Channel_{A,B}`
$\boldsymbol{RRSR}$	Rayleigh response sensitivity factor	N.B. The current code does not actually read in this parameter from the L1B file - it is hardcoded to a value for now.
$\boldsymbol{Ka_2, Ka_3, Kb_2, Kb_3}$	Error quantifier parameters	N.B. The current code does not actually read in these parameters from the L1B file - it is hardcoded to a value for now.
$\boldsymbol{\lambda_0}$	Laser wavelength	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/ Satellite_Characterisation_Data/ Laser_Wavelength`
$\boldsymbol{\varphi}$	Incidence angle of the line of sight (LOS) of the instrument at	Derived from elevation angles of range-bins from ALD_U_N_1B, see section 12.1: `Geolocation ADS/ Observation_Rayleigh_Geolocation/List_of _Observation_Geolocation_of_Height_Bins/`

	vertical centre-of-gravity of the range-bin	`Observation_Geolocation_of_Height_Bin/Topocentric_Elevation_of_Height_Bin`
$\boldsymbol{W(i,k)}$	Measurement-bin weights for the observation	Output of L2B processing described in Section 10

Outputs:

Symbol	Description	ALD_U_N_2B file variable
$\boldsymbol{EQR(i)}$	L2B Rayleigh HLOS wind error quantifier (standard error estimate i.e. 1-sigma) for each observation	`Rayleigh Wind Product Confidence DataADS/Rayleigh_Wind_Product_PCD/Rayleigh_Wind_QC/Hlos_error_estimate`

Algorithm

The option based on the L1B approach consists of adapting the Error Quantifier (HLOS wind error estimate) approach detailed in section 13.1.3 of L1B DPM (version 1.2 dated 29 September 2005). This approach consists in relating the error estimate to the Rayleigh useful signal M(i) by using the generic function below:

$$EQRa(i) = \frac{\left(A(i) + Ka_2\right)^{Ka_3}}{A(i)}$$

$$EQRb(i) = \frac{\left(B(i) + Kb_2\right)^{Kb_3}}{B(i)}$$

$$EQR(i) = \frac{\lambda_0}{2\sin(\varphi)} \frac{1}{RRSR} \sqrt{EQRa(i)^2 + EQRb(i)^2} \qquad \textbf{Eq. 17-9}$$

Here, the index designates the height bin and Ka_2, Ka_3, Kb_2, and Kb_3, are coefficients determined empirically - stored in the Auxiliary Satellite Characterization Auxiliary and repeated in L1B products.

The use of this equation for the L2B processor is possible, taking into account the definition of meteorologically-weighted useful signal for channels A and B (Section 12.1). Accordingly, equation (77) of the DPM is modified to take into account the variable weight of measurements in the profile

$$A(i) = \frac{\sum_k W(i,k) A_m(i,k)}{\sum_k W(i,k)} \qquad \textbf{Eq. 17-10}$$

$$B(i) = \frac{\sum_k W(i,k) B_m(i,k)}{\sum_k W(i,k)} \qquad \textbf{Eq. 17-11}$$

The problem is that the parameters Ka_2, Ka_3, Kb_2, and Kb_3, are not known, they shall be established during a pre-launch, on-ground characterization of the instrument. Default values can be found in the Satellite Characterization files delivered with the L1Bp, but may be meaningless.

18. Mie channel uncertainty estimation processing

The recommended L2B Mie HLOS wind error quantifier algorithm is Option 3 as described in Section 18.3. Alternative algorithms are available and are retained as back-ups. Option 1 is based on the approach taken for the L1B HLOS wind error quantifier (Section 18.1). Option 2 (Section 18.2) is similar option 3, but with a few omissions.

Which algorithm to use is controlled via the AUX_PAR_2B file; in particular: `Error_Quantifier_Params/ErrorQuantMethod_Mie`

- Option 1: `ErrorQuantMethod_Mie_1Bweighted`
- Option 2: `ErrorQuantMethod_Mie_core_sens`
- Option 3: `ErrorQuantMethod_Mie_core_sens2`

18.1. Option 1: L2B Mie error quantifier based on L1B approach

Inputs:

Symbol	Description	Input variable source
$LID_{atm}(i,j,k)$	Measurement spectrometer counts at measurement-bin level	ALD_U_N_1B: `Measurement_ADS/List_of_Mie_Measurement_Data/Mie_Measurement_Data`
K_1, K_2, K_3	Error quantifier parameters	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/Satellite_Characterisation_Data/Error_Quantifiers/ Mie_Error_Quantifier_K{1,2,3}`
λ_0	Laser wavelength	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/ Satellite_Characterisation_Data/ Laser_Wavelength`
φ	Incidence angle of the line of sight (LOS) of the instrument at vertical centre-of-gravity of the range-bin	Derived from elevation angles of range-bins from ALD_U_N_1B, see section 12.1: `Geolocation ADS/ Observation_Rayleigh_Geolocation/List_of_Observation_Geolocation_of_Height_Bins/ Observation_Geolocation_of_Height_Bin/Topocentric_Elevation_of_ Height_Bin`
$W(i,k)$	Measurement-bin weights for the observation	Output of L2B processing described in Section 10

Outputs:

Symbol	Description	ALD_U_N_2B file variable
$EQM(i)$	L2B Mie HLOS wind error quantifier (standard error estimate i.e. 1-sigma) for each observation	`Mie Wind Product Confidence DataADS/Mie_Wind_Product_PCD/Mieleigh_Wind_QC/Hlos_error_estimate`

Algorithm

The option based on the L1B approach consists of adapting the Error Quantifier approach detailed in section 13.1.2 of L1B DPM (version 1.2 dated 29 September 2005). This approach consists in relating the error estimate to the Mie useful signal M(i) by using the generic function below:

$$EQM(i) = \frac{\lambda_0}{2\sin(\varphi)} K_1 \frac{(M(i) + K_2)^{K_3}}{M(i)} \qquad \textbf{Eq. 18-1}$$

Here, the index designates the height bin and K_1, K_2, and K_3 are coefficients determined empirically - stored in the Auxiliary Satellite Characterization Auxiliary and repeated in L1B products.

The use of this equation for the L2B processor is possible. The algorithm would be start with the removal of the offset read-out offset

$$\mathrm{LID}(\mathrm{i,j,k}) = \mathrm{LID}(\mathrm{i,j,k}) - \frac{1}{2}\sum_{\mathrm{j=19}}^{20} \mathrm{LID}(\mathrm{i,j,k}) \qquad \textbf{Eq. 18-2}$$

Then the background is removed

$$\mathrm{LID}(\mathrm{i,j,k}) = \left[\mathrm{LID}(\mathrm{i,j,k}) - \frac{\mathrm{t_i}}{\mathrm{t_i}} \mathrm{LID}(25,\mathrm{j,k}) \right] \qquad \textbf{Eq. 18-3}$$

Equation (72) of the DPM is modified to take into account the variable weight of measurements in the profile

$$LID(i,j) = \frac{\sum_k W(i,k) LID(i,j,k)}{\sum_k W(i,k)} \qquad \textbf{Eq. 18-4}$$

The weight-averaged spectra are then integrated to get the useful signal

$$\mathrm{M(i)} = \sum_{\mathrm{j=3}}^{18} \mathrm{LID(i,j)} \qquad \textbf{Eq. 18-5}$$

The problem is that the parameters K_1, K_2 and K_3 are not known, they shall be established during a pre-launch, on-ground characterization of the instrument. Default values can be found in the Satellite Characterization files delivered with the L1Bp, but may be meaningless.

18.2. Option 2: L2B Mie error quantifier using outputs of the Mie core algorithm (Tellus paper version)

N.B: the definition in this section is provisional and subject to future improvement.
In particular, additional terms are needed for contributions from internal reference and from post-core correction terms (see Section 14.5).

Inputs:

Symbol	Description	Input variable source
$\boldsymbol{LID_{atm}(i,j,k)}$	Measurement spectrometer counts at measurement-bin level	ALD_U_N_1B: `Measurement_ADS/List_of_Mie_Measurement_Data/Mie_Measurement_Data`
$\boldsymbol{T_{obs}(j)}$	Tripod obscuration factors	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/ Satellite_Characterisation_Data/ List_of_Tripod_Obscuration_Corrections/Tripod_Obscuration_Correction`
$\boldsymbol{peak\ height, FWHM, \hat{f}_{Mi}}$	Outputs of the	Internal L2B processor temporary outputs: See section

	Mie core algorithm i.e. parameters of the Lorentzian fit function: peak height, full-width half maximum, and position of peak in pixel space	14.2

$\boldsymbol{MRS_{atm}, MRS_{int}}$	MRC linear fit coefficients i.e. gradient (MRS=Mie response slope). One set of values per L1B file.	ALD_U_N_1B:
		`Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/Mie_Response_Calibration_Data/Mie_Measurement_Response_Calibration/Measurement_Mean_Sensitivity`
		
		`Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/Mie_Response_Calibration_Data/Mie_Reference_Pulse_Response_Calibration/Reference_Pulse_Mean_Sensitivity`

$\boldsymbol{\lambda_0}$	Laser wavelength	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/ Satellite_Characterisation_Data/ Laser_Wavelength`
$\boldsymbol{\varphi}$	Incidence angle of the line of sight (LOS) of the instrument at vertical centre-of-gravity of the range-bin	Derived from elevation angles of range-bins from ALD_U_N_1B, see section 12.1: `Geolocation ADS/ Observation_Rayleigh_Geolocation/List_of_Observation_Geolocation_of_Height_Bins/ Observation_Geolocation_of_Height_Bin/Topocentric_Elevation_of_ Height_Bin`
$\boldsymbol{W(i,k)}$	Measurement-bin weights for the observation	Output of L2B processing described in Section 10

Outputs:

Symbol	Description	ALD_U_N_2B file variable
$\boldsymbol{EQM(i)}$	L2B Mie HLOS wind error quantifier (standard error estimate i.e. 1-sigma) for each observation	`Mie Wind Product Confidence DataADS/Mie_Wind_Product_PCD/Mieleigh_Wind_QC/Hlos_error_estimate`

Algorithm

The baseline for this option of L2B Mie error quantifier is the error standard deviation given by the equation:

$$EQM(i) = \frac{\lambda_0}{2\sin\varphi}\left[\sum_{j=3}^{18}\alpha(i,j)^2\right]^{-1}\sqrt{\sum_{j=3}^{18}\sigma_{i,j}^2\alpha(i,j)^2} \qquad \textbf{Eq. 18-6}$$

where

$$\sigma_{i,j}^2 = \sum_{k=1} W^2(i,k)LID(i,j,k) \qquad \textbf{Eq. 18-7}$$

is an estimate for the variance (consistent with Poisson noise) of the weighted sum of photo-counts $\mathrm{LID(i,j,k)}$ (that is, after removal of the video offset), $W(i,k)$ is the weight associated to bin i, measurement k for the profile under consideration, and

$$\alpha(\mathrm{i,j}) = \mathrm{T_{obs}(j)}\ \mathrm{peak\_height}\left[\frac{1}{1+\frac{4\left(\mathrm{f_j^+ - \hat{f}_{mie}(i)}\right)}{\mathrm{FWHM}^2}} - \frac{1}{1+\frac{4\left(\mathrm{f_j^- - \hat{f}_{mie}(i)}\right)}{\mathrm{FWHM}^2}}\right] \qquad \textbf{Eq. 18-8}$$

Here, $\mathrm{peak\_height}$, FWHM and $\mathrm{\hat{f}_{mie}}$ (in frequency units; after conversion from pixel count units using the MRC gradient: *MRS*) are the outputs of the Mie core algorithm (peak height and FWHM of the fitted Lorentzian spectrum, and Mie frequency estimate), $\mathrm{T_{obs}(j)}$ is the tripod obscuration factor for CCD bin j, and $\mathrm{f_j^{+,-}}$ are the upper and lower frequencies of CCD bin j.

A derivation for the above equations is given in [Appendix B. Mie error quantifier derivation].

18.3. Option 3: L2B Mie error quantifier based on least squares fit of an integrated Lorentzian and using outputs of the Mie core algorithm

Following testing performed pre-launch, **this is the recommended** L2B Mie error quantifier.

Inputs:

Symbol	Description	Input variable source
$\boldsymbol{LID_{atm}(i,j,k)}$	Measurement spectrometer counts at measurement-bin level	ALD_U_N_1B: `Measurement_ADS/List_of_Mie_Measurement_Data/Mie_Measurement_Data`
$\boldsymbol{LID_{int}(j,k)}$	Internal Reference spectrometer counts at measurement level	ALD_U_N_1B: `Measurement_ADS/List_of_Mie_Reference_Pulse/Mie_Reference_Pulse`
$\boldsymbol{T_{obs}(j)}$	Tripod obscuration factors	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/ Satellite_Characterisation_Data/ List_of_Tripod_Obscuration_Corrections/Tripod_Obscuration_Correction`
$\boldsymbol{a, b, f_w, f_0}$	Outputs of the Mie core algorithm i.e. parameters of the Lorentzian fit	Internal L2B processor temporary outputs: See section 14.2. This output will be produced for both atmospheric signal and internal reference signal.

	function: • peak height (spectrometer counts) • offset (spectrometer counts) • full-width half maximum (pixels) • position of peak (pixels)	

MRS_{atm}, MRS_{int}	MRC linear fit coefficients i.e. gradient (MRS=Mie response slope). One set of values per L1B file.	ALD_U_N_1B:
		`Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/Mie_Response_Calibration_Data/Mie_Measurement_Response_Calibration/Measurement_Mean_Sensitivity`
		
		`Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/Mie_Response_Calibration_Data/Mie_Reference_Pulse_Response_Calibration/Reference_Pulse_Mean_Sensitivity`

λ_0	Laser wavelength	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/ Satellite_Characterisation_Data/ Laser_Wavelength`
φ	Incidence angle of the line of sight (LOS) of the instrument at vertical centre-of-gravity of the range-bin	Derived from elevation angles of range-bins from ALD_U_N_1B, see section 12.1: `Geolocation ADS/ Observation_Rayleigh_Geolocation/List_of_Observation_Geolocation_of_Height_Bins/ Observation_Geolocation_of_Height_Bin/Topocentric_Elevation_of_ Height_Bin`
r_g	Radiometric gain for the Mie channel	ALD_U_N_1B: `Calibration & Characterisation Data GADS/ L1B_Characterisation_Data/ Satellite_Characterisation_Data/ Radiometric_Gain_Mie`
$W(i,k)$	Measurement-bin weights for the observation	Output of L2B processing described in Section 10

Outputs:

Symbol	Description	ALD_U_N_2B file variable
$\sigma_{HLOS,tot}$	L2B Mie HLOS wind error quantifier (standard error estimate i.e. 1-sigma) for each observation	`Mie Wind Product Confidence DataADS/Mie_Wind_Product_PCD/Mieleigh_Wind_QC/Hlos_error_estimate`

Algorithm

This L2B Mie error quantifier derivation is similar to the method in Section 18.2 in that they are both derived using the output of the Mie core algorithm and using the assumption that the Mie core algorithm is effectively performing a least squares fit of a model to the useful ACCD count data. However this method is a more advanced in that it additionally considers:

- The fitting of the offset part of the Mie core fit
- Models the integration of a Lorentzian across the pixel width, which is effectively what the pixel readout gives.
- Takes account of the Mie radiometric gain factor
- Considers the errors from both the atmospheric and internal reference signals

Here we show how an estimate of the error on the Mie HLOS winds can be approximated from the solution error covariance of the weighted 'linear' least squares fit algorithm. The model is an integrated Lorentzian which is non-linear in terms of the state vector $\underline{x}$ that describes it, but an error estimate can be made by linearising the model about a given $\underline{x}$ (the model state as provided by the Mie core algorithm solution is chosen).

The integrated Lorentzian depends on four parameters: f_0 the position of the peak (in pixel space), f_w the FWHM of the Lorentzian (in pixel space), a the height of the peak (in spectrometer counts) and b the offset (in spectrometer counts). Hence the state vector can be written:

$$\underline{x} = \begin{bmatrix} f_0 \\ f_w \\ a \\ b \end{bmatrix} \qquad \textbf{Eq. 18-9}$$

In the following equations a double underline refers to a matrix and a single underline to a vector.

Some definitions:

- $\underline{x}$ is the state vector (variables) of the model (in our case an integrated Lorentzian)
- $\underline{y}$ is the vector of observations to be fitted by the model (in our case useful spectrometer counts). As obtained from a weighted summation of *LID* as in the previous section.
- $\underline{\underline{W}}$ is a weight matrix used in the least squares fit
- $\underline{\underline{O}}$ is the observation error covariance matrix (i.e. errors on the spectrometer counts)
- $\underline{\underline{H}}$ is the Jacobian matrix (i.e. $\partial \mathrm{H}(\underline{x})/\partial \underline{x}$) of the integrated Lorentzian forward model.

Weighted linear least squares minimises a cost function $J(\underline{x})$ to find the solution state vector $\underline{x}_s$ (see http://en.wikipedia.org/wiki/Linear_least_squares_%28mathematics%29)

$$\underline{x}_s = \arg\min_x J(\underline{x}) \qquad \textbf{Eq. 18-10}$$

where:

$$J(\underline{x}) = \left\| \underline{\underline{W}}^{1/2} \left(\underline{y} - \underline{\underline{H}}\,\underline{x} \right) \right\|^2 = \left(\underline{y} - \underline{\underline{H}}\,\underline{x} \right)^T \underline{\underline{W}} \left(\underline{y} - \underline{\underline{H}}\,\underline{x} \right) = \underline{y}^T \underline{\underline{W}}\,\underline{y} + \underline{x}^T \underline{\underline{H}}^T \underline{\underline{W}}\,\underline{\underline{H}}\,\underline{x} - 2\underline{x}^T \underline{\underline{H}}^T \underline{\underline{W}}\,\underline{y}$$

Eq. 18-11

Here the weighting matrix is not yet defined. The solution is found when the gradient is zero:

$$\frac{\partial J(\underline{x})}{\partial \underline{x}} = 2\underline{\underline{H}}^T \underline{\underline{W}}\,\underline{\underline{H}}\,\underline{x}_s - 2\underline{\underline{H}}^T \underline{\underline{W}}\,\underline{y} = 0 \qquad \textbf{Eq. 18-12}$$

rearranging for the solution:

$$\underline{x}_s = \left(\underline{\underline{H}}^T \underline{\underline{W}}\,\underline{\underline{H}}\right)^{-1} \underline{\underline{H}}^T \underline{\underline{W}}\,\underline{y} \qquad \textbf{Eq. 18-13}$$

The solution error covariance matrix can be calculated as follows (with $\underline{x}_t$ and $\underline{y}_t$ the true values):

$$\underline{\underline{X}}_s = E\left((\underline{x}_s - \underline{x}_t)(\underline{x}_s - \underline{x}_t)^T\right) = E\left(\left(\underline{\underline{H}}^T \underline{\underline{W}}\,\underline{\underline{H}}\right)^{-1} \underline{\underline{H}}^T \underline{\underline{W}}\left(\underline{y} - \underline{y}_t\right)\left(\left(\underline{\underline{H}}^T \underline{\underline{W}}\,\underline{\underline{H}}\right)^{-1} \underline{\underline{H}}^T \underline{\underline{W}}\left(\underline{y} - \underline{y}_t\right)\right)^T\right) \qquad \textbf{Eq. 18-14}$$

where E() is the expectation operator.
Since the observation error covariance is:

$$\underline{\underline{O}} = E\left((\underline{y} - \underline{y}_t)(\underline{y} - \underline{y}_t)^T\right) \qquad \textbf{Eq. 18-15}$$

the solution error covariance matrix is:

$$\underline{\underline{X}}_s = \left(\underline{\underline{H}}^T \underline{\underline{W}}\,\underline{\underline{H}}\right)^{-1} \underline{\underline{H}}^T \underline{\underline{W}}\,\underline{\underline{O}}\,\underline{\underline{W}}^T \underline{\underline{H}}\left(\underline{\underline{H}}^T \underline{\underline{W}}^T \underline{\underline{H}}\right)^{-1} = \begin{bmatrix} \sigma_{f_0}^2 & \sigma_{f_0}\sigma_{f_w} & \sigma_{f_0}\sigma_a & \sigma_{f_0}\sigma_b \\ \sigma_{f_0}\sigma_{f_w} & \sigma_{f_w}^2 & \sigma_{fw}\sigma_a & \sigma_{fw}\sigma_b \\ \sigma_{f_0}\sigma_a & \sigma_{fw}\sigma_a & \sigma_a^2 & \sigma_a\sigma_b \\ \sigma_{f_0}\sigma_b & \sigma_{fw}\sigma_b & \sigma_a\sigma_b & \sigma_b^2 \end{bmatrix} \qquad \textbf{Eq. 18-16}$$

If the weight matrix is set to be the identity matrix i.e. no weighting applied in least squares, then the solution error covariance matrix is:

$$\underline{\underline{X}}_s = \left(\underline{\underline{H}}^T \underline{\underline{H}}\right)^{-1} \underline{\underline{H}}^T \underline{\underline{O}}\,\underline{\underline{H}}\left(\underline{\underline{H}}^T \underline{\underline{H}}\right)^{-1} \qquad \textbf{Eq. 18-17}$$

This is the appropriate formula for the current L2Bp Mie core algorithm, which does not employ weighting in the least squares fit.

If the weight matrix is set to be the inverse observation error covariance matrix (which is appropriate for a maximum likelihood estimate assuming Gaussian error statistics for the observations), then the solution error covariance matrix simplifies to:

$$\underline{\underline{X}}_s = \left(\underline{\underline{H}}^T \underline{\underline{O}}^{-1} \underline{\underline{H}}\right)^{-1} \qquad \textbf{Eq. 18-18}$$

Using the observation error covariance matrix for the weighting may be employed in future Mie core algorithms. Both formulae are available as options in the L2Bp, with the default being the non-weighted.

The solution peak position error standard deviation (in pixels) is selected from an element of the solution covariance matrix i.e. $\sqrt{\underline{\underline{X}}_s}(1,1)$. This error can be propagated from pixel space to HLOS wind space as follows:

$$\sigma_{HLOS} = \left(\frac{\partial HLOS}{\partial LOS}\right)\left(\frac{\partial LOS}{\partial f}\right)\left(\frac{\partial f}{\partial f_0}\right)\sigma_{f_0} = \frac{\lambda_0}{2MRS_{A,I}\sin\varphi}\sigma_{f_0} \qquad \textbf{Eq. 18-19}$$

$MRS_{A,I}$ is the gradient of the Mie response slope (either atmospheric A, or internal reference I). Φ is the elevation angle and λ_0 is the laser wavelength, as in the previous section.

Since the solution error covariance matrix (both formulae) depends on the Jacobian of the forward model and the observation error covariance matrix we now describe how these components are

calculated. The forward model is an integrated Lorentzian over the pixel range i.e. for a given pixel *j* the spectrometer counts are simulated as:

$$H_j(\underline{x}) = \tau_j \int_{f_{j-0.5}}^{f_{j+0.5}} \left(\frac{a f_w^{\,2}}{4\left(f_0 - f\right)^2 + f_w^{\,2}} + b \right) df \qquad \textbf{Eq. 18-20}$$

where *t*=tripod obscuration coefficient for pixel *j*. The integral can be solved analytically to give:

$$H_j(\underline{x}) = \tau_j \left(b\left(f_{j+0.5} - f_{j-0.5}\right) - \frac{a f_w}{2} \left(\tan^{-1}\left(\frac{2}{f_w}\left[f_0 - f_{j+0.5}\right] \right) - \tan^{-1}\left(\frac{2}{f_w}\left[f_0 - f_{j-0.5}\right] \right) \right) \right)$$

Eq. 18-21

The Jacobian matrix i.e. partial derivatives of the forward model with respect to different variables of *x* (i.e. a 4 by 16 matrix) can be constructed from the following equations:

$$\frac{\partial H_j(\underline{x})}{\partial f_0} = \tau_j a f_w^{\,2} \left(\frac{1}{4\left(f_0 - f_{j-0.5}\right)^2 + f_w^{\,2}} - \frac{1}{4\left(f_0 - f_{j+0.5}\right)^2 + f_w^{\,2}} \right) \qquad \textbf{Eq. 18-22}$$

$$\frac{\partial H_j(\underline{x})}{\partial f_w} = \tau_j a \left(f_w \left[\frac{f_0 - f_{j+0.5}}{4\left(f_0 - f_{j+0.5}\right)^2 + f_w^{\,2}} - \frac{f_0 - f_{j-0.5}}{4\left(f_0 - f_{j-0.5}\right)^2 + f_w^{\,2}} \right] - \frac{1}{2} \left[\tan^{-1}\left(\frac{2}{f_w}\left[f_0 - f_{j+0.5}\right] \right) - \tan^{-1}\left(\frac{2}{f_w}\left[f_0 - f_{j-0.5}\right] \right) \right] \right)$$

Eq. 18-23

$$\frac{\partial H_j(\underline{x})}{\partial a} = \frac{\tau_j f_w}{2} \left(\tan^{-1}\left(\frac{2}{f_w}\left[f_0 - f_{j-0.5}\right] \right) - \tan^{-1}\left(\frac{2}{f_w}\left[f_0 - f_{j+0.5}\right] \right) \right) \qquad \textbf{Eq. 18-24}$$

$$\frac{\partial H_j(\underline{x})}{\partial b} = \tau_j \left(f_{j+0.5} - f_{j-0.5} \right) \qquad \textbf{Eq. 18-25}$$

When calculating the solution error covariance matrix, the observation error covariance matrix is assumed to be diagonal (and therefore so is its inverse) since the errors on the spectrometer counts for each pixel are assumed to be independent (if we have a need to include correlations then this can be done). The diagonal elements of the observation error covariance matrix are the error variances of the useful spectrometer counts for each pixel. The error variances are assumed to be consistent with Poisson statistics for the electron counts (i.e. the standard error for Poisson counting is the square root of the number of counts) and also taking account of the Mie radiometric gain in the conversion to the stored LSB. The error propagation from electron counts to useful spectrometer count (expressed in LSB), is as follows (showing the error variance on a given pixel for the spectrometer count):

$$LSB = r_g e_c \qquad \textbf{Eq. 18-26}$$

$$\sigma_{LSB}^2 = \left| \frac{\partial LSB}{\partial e_c} \right|^2 \sigma_{e_c}^2 = r_g{}^2 \sigma_{e_c}^2 = r_g{}^2 e_c = r_g LSB \qquad \textbf{Eq. 18-27}$$

Both weighted and non-weighted solution error covariance matrices require the inversion of a four-by-four matrix, either $\underline{\underline{H}}^T \underline{\underline{O}}^{-1} \underline{\underline{H}}$ or $\underline{\underline{H}}^T \underline{\underline{H}}$. If the matrix inversion is not well conditioned (using the L2Bp function `finvert_simple` which employs Gaussian elimination), then the error estimate is set to a missing value.

These calculations are performed for the atmospheric Mie signal (to produce an atmospheric error estimate: $\sigma_{HLOS,atm}$) and for the internal reference Mie signal (to produce an internal reference

error estimate: $\sigma_{HLOS,int}$). The total HLOS wind error estimate is then calculated as the quadratic sum of the atmospheric and internal reference standard deviations (N.B. the internal reference error was found to be non-negligible):

$$\sigma_{HLOS,tot} = \sqrt{\sigma_{HLOS,atm}^2 + \sigma_{HLOS,int}^2} \qquad \textbf{Eq. 18-28}$$

The above derivation considers sources of error from V_{atm} and V_{int} of the HLOS wind derivation in Section 14.6, but does not consider the errors coming from V_{sat}, V_{GM}, V_{RDB}, because it is unclear at this stage how significant they will be compared to the other terms.

An example of the Mie-cloudy HLOS wind error estimates are given in Figure 11 below:

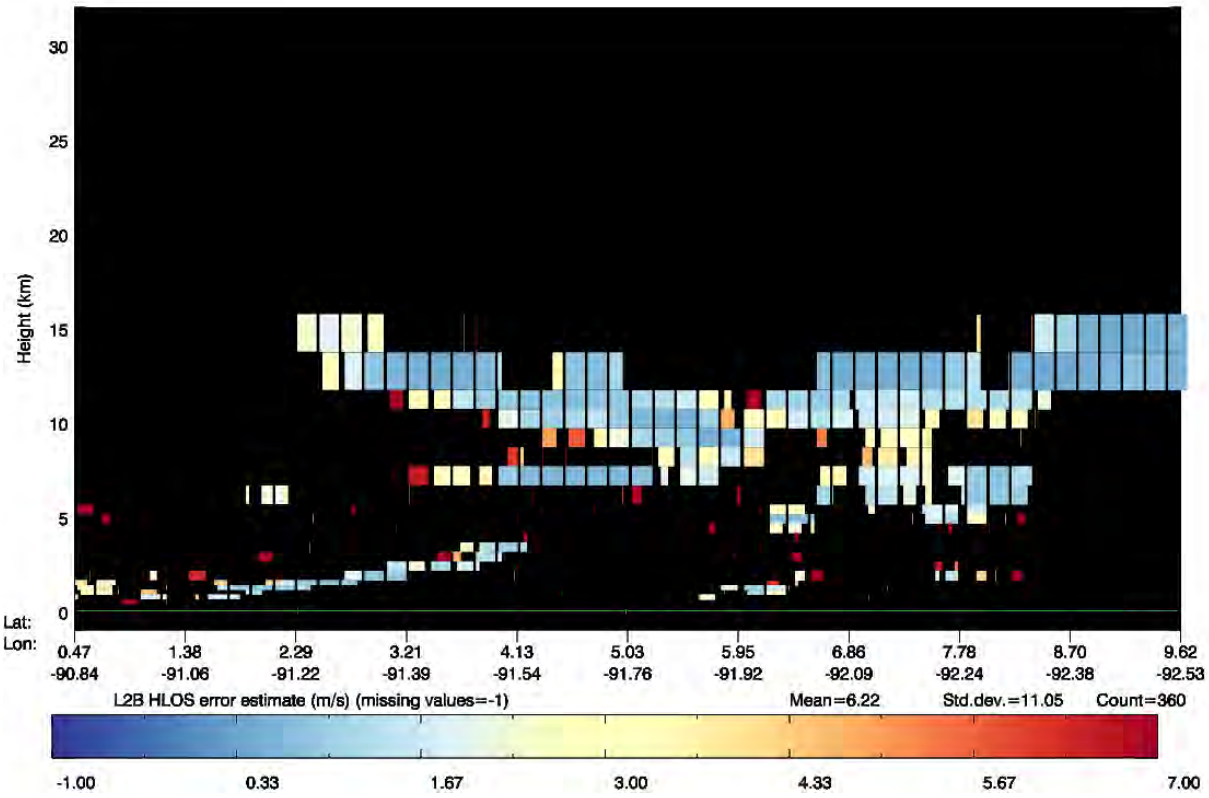


Figure 11. Example of L2B Mie-cloudy wind standard error estimates (m/s), for the same simulation scenario as in Figure 9.

19. Optical Properties Code: Retrieval of particulate optical properties from the Fabry-Pérot signal only

The optical properties code (OPC) is part of the L2B processor (L2Bp). Its main objective is to detect cloud and aerosol layers along the Aeolus track from the Fabry-Pérot (FP) i.e. Rayleigh channel signal only. The result is a binary map, used in the L2Bp classification procedure as input for the accumulation of Mie and Rayleigh signals from measurement to observation level. As such, the OPC is operated on measurement profile level. In addition to the binary map, the OPC provides for each measurement profile estimates of (i) aerosol backscatter and extinction, (ii) the scattering ratio, (iii) the positioning of particle layers inside the Aeolus bins, (iv) the cross-talk correction factor and in cases (v) the lidar ratio (inverse of the backscatter-to-extinction ratio). This section discusses the algorithm of the OPC implemented in the L2Bp.

Motivation for using the FP signal only in the OPC is the additional option of estimating the scattering ratio, and hence to enable classification, at altitudes of missing Mie range bins, e.g., in the UTLS where Polar Stratospheric Clouds may be present in some seasons. Also, in case of degradation of Mie channel performance and hence degradation of the L1B scattering ratio product, the OPC is a good alternative for a well-functioning FP channel. As such the OPC contributes to a mission risk reduction in case the Mie channel and/or the L1Bp detection procedure performs less than anticipated. Since classification is a crucial procedure (see section 8) in the L2Bp, the quality of produced winds strongly relies on the quality of the classification procedure.

It is noted that the scattering ratio is an important parameter in the ILIAD scheme to correct derived Rayleigh winds for Mie contamination in the Rayleigh channel signal, see Sections 15 and 17.

Output parameters include a binary map and for each measurement bin and various estimates of the optical properties e.g. extinction, backscatter, scattering ratio. The binary map may be used for scene classification as discussed in section 8.

From the L2A ATBD [RD16], page 37, or otherwise [ILIAD, RD12], one can write for the total Fabry-Pérot signal in bin *i*, (lidar equation) i.e., the total number of photon counts on Rayleigh channels A and B:

$$\mathrm{FP}(i) = K_{\mathrm{ray}} C(\nu(i), P(i), T(i)) \beta_m(i) \frac{\Delta z(i)}{R^2(i)} T_m^2(i) T_p^2(i) \tau_m(i) \tau_p(i) + \varepsilon(i)$$

$$C(\nu(i), P(i), T(i)) = C_1(\nu(i), P(i), T(i)) + C_2(\nu(i))\left(\rho(i) - 1\right)$$

Eq. 19-1

$$T_x(i) = \prod_{j=0}^{i-1} \tau_x(j), \qquad x = \{m, p\}$$

Eq. 19-2

with K_{ray} the calibration factor, $\beta_m(i)$ and $\beta_p(i)$ are molecular and particle (aerosol or cloud) backscatter coefficients ($m^{-1}sr^{-1}$) in bin number *i* respectively, $T_m(i\text{-}1)$ and $T_p(i\text{-}1)$ the respective total one-way molecular and particle transmission from bins starting at the topmost bin (numbered i=1) to bin number i-1, $\tau_m(i)$ and $\tau_p(i)$ is the respective one-way molecular and particle transmission in bin number *i* and $\varepsilon(i)$ is the total noise in bin number *i*, $\Delta z(i)$ is the distance of the laser beam through bin number *i*, i.e., the bin size divided by the cosine of the incidence angle and $R(i)$ is the range from the satellite to the bin centre. $C_1(\nu,P,T)$ is the fraction of molecular backscatter that is actually detected by the Rayleigh channel, depending on the pressure P, the temperature T and the Doppler-shift frequency ν. $C_2(\nu)$ is the fraction of particle backscatter that is actually detected by the Rayleigh channel, also denoted the cross-talk factor, depending on the sole Doppler-shift frequency because of the small width of Mie scattering spectral peak and ρ is the scattering ratio defined as the total backscatter divided by molecular backscatter: $\rho(\mathrm{i})=[\beta_m(i)+\beta_p(i)]/\beta_m(i)$.

Additional Note. The calibration parameter $C_1(\nu,P,T)$ is determined from interpolating a Fabry-Pérot (FP) lookup table. The lookup table has been generated using ILIAD output files. More specifically for a range of values for temperature, pressure and (Doppler shifted) frequency, the ILIAD software computes the Fabry-Pérot signal from the molecular (Rayleigh-Brillouin) spectrum and FP channel transfer functions. At the moment no pressure dependence is available in the lookup table. Currently $C_2(\nu)$ is set to zero, pending further developments in the optical calibration, see [RD12]. This means that the cross-talk factor is not available from AUX_CAL files. For this reason it is proposed to estimate it as part of the OPC, as discussed below.

For simplicity we introduce

$$K(i) = K_{\text{ray}} C_1(\nu(i), P(i), T(i)) \frac{\Delta z(i)}{R^2(i)}; \quad X = \frac{C_2(\nu(i))}{C_1(\nu(i), P(i), T(i))}$$

Eq. 19-3

hence, for now ignoring cross-talk dependence to ν, *P and T*. **Eq. 19-1** is then rewritten to

$$\text{FP}(i) = K(i)\left[\beta_m(i) + X\beta_p(i)\right] T_m^2(i-1) T_p^2(i-1) \tau_m(i) \tau_p(i) + \varepsilon(i), \quad i = 1,..,24$$

Eq. 19-4

Molecular backscatter and transmission are easily estimated from pressure and temperature, available from the AUX_MET data, see [Appendix D. Retrieval of atmosphere molecular backscatter and extinction (Rayleigh scattering)]. From these, estimates of the FP signal are obtained, e.g., assuming an atmosphere with molecules only, denoted $\hat{FP}_{\text{mol}}$, or assuming molecules only in bin *i*, but with possible particles in layers aloft, denoted $\hat{FP}$:

$$\widehat{FP}_{\text{mol}}(i) = K(i)\beta_m(i) T_m^2(i-1)\tau_m(i)$$

Eq. 19-5

$$\begin{aligned} \hat{FP}(i) &= K(i)\beta_m(i) T_m^2(i-1) T_p^2(i-1)\tau_m(i) \\ &= \hat{FP}_{\text{mol}}(i) T_p^2(i-1) \end{aligned}$$

Eq. 19-6

Clearly $\hat{FP}_{\text{mol}}$ only requires the AUX_MET data for calculation, for known $K(i)$, and $\hat{FP}$ requires an estimate of the transmission of particles in layers aloft in addition. $K(i)$ is discussed in the next section.

19.1. Estimation of the calibration factor

Calculation of **Eq. 19-5** and **Eq. 19-8** requires the calibration factor $K(i)$ which is defined in **Eq. 19-3**. The bin size and range to the satellite are known. For simplicity we assume that the calibration factor C_1 does not depend strongly on ν, *P* and *T* and is assumed constant. The remaining constant $K=K_{\text{ray}}C_1$ is then estimated at observation level obtained from

$$K = \frac{\frac{1}{M}\sum_{m=1}^{M} FP(m,1)}{\beta_m(M/2,1) T_m^2(M/2,1) \tau_m(M/2,1) \Delta z(1)/R^2(1)}$$

Eq. 19-7

where we introduced index *m* that denotes measurement number. So $FP(m,i)$ here denote the FP signal on bin number *i* for measurement *m*. The denominator is the uncalibrated molecules-only signal of the measurement at the observation centre, i.e., **Eq. 19-5** without the calibration factor $K_{\text{ray}}C_1$. *M* is the total number of measurements used in one observation, with a typical value of 30, but which can

be changed during the mission (and with the grouping algorithm settings). The calibration factor is thus calculated from the FP signal in the topmost bins of all measurement capturing one observation, assuming that this bin does not contain particles. This latter assumption should be checked by monitoring both the mean signal of the FP signal in the uppermost bin of the *M* measurements, i.e., the numerator of Eq. 19-7, and the standard deviation of the FP signal in the uppermost bin of the *M* measurements. A jump in the value of both the mean and standard deviation of adjacent observations might indicate the presence of particles. Also, a jump in the standard deviation only may indicate a change in the noise characteristics of the FP channel.

19.2. Particle layer detection

Initially, naively ignoring cross-talk and measurement noise in the FP signal in **Eq. 19-4** and taking the ratio of **Eq. 19-4** and **Eq. 19-6** yields an estimate of the particle transmission in bin *i*:

$$\hat{\boldsymbol{\tau}}_{\boldsymbol{p}}(\boldsymbol{i}) = \left[\boldsymbol{FP}(\boldsymbol{i})/\widehat{FP}(\boldsymbol{i})\right]^{\boldsymbol{k}}$$

Eq. 19-8

for *k* equal to 1. For Aeolus bins free of particles $\hat{\tau}_p$ should equal 1. If the Aeolus bin does contain particles, such as in case of a cloud or aerosol layers, then $\hat{\tau}_p$ deviates from 1, generally with values exceeding 1 (which is unphysical) in the uppermost bins (due to cross-talk, that is not part of the expected signal in the denominator of **Eq. 19-8**) and values smaller than 1 in the lowermost bins (due to extinction inside the particle layer and an transmission overestimation (values exceeding 1) in the uppermost bins).

A scheme based on **Eq. 19-8** has been implemented in the OPC. Details of this scheme are found in [Appendix C. Particle transmission computation from the Fabry-Pérot channel signal]. When using *k*=1 and applied to a typical FP measurement profile yields a strongly oscillating solution for the particle transmission profile as explained in [Appendix C. Particle transmission computation from the Fabry-Pérot channel signal]. Reducing the value of *k* shows (i) a damping of the oscillation and (ii) convergence of particle transmission to 1 below the cloud layer. For values closer to 0 the damping is stronger, however this is at the expense of getting transmission values substantially smaller than 1 for Aeolus bins below particle layers, i.e., potentially smearing the cloud to Aeolus bins below the actual cloud location.

Application of the *k*-power in **Eq. 19-8** thus strongly reduces oscillations in the particle transmission retrieval. These oscillations are the result of both cross-talk in the FP signal and measurement noise. The estimated particle transmission from **Eq. 19-8** can be used for the detection of particle layers, by putting thresholds. Clearly, the selection of threshold values is closely related to the selected value for *k*, which in turn is related to the noise level in the FP signal, i.e., *k* values closer to zero are needed to reduce noise contaminating the particle transmission estimate. However, this should be traded off against smearing the cloud location to lower bins.

A nice property of the algorithm is that the ***total transmission of an isolated particle layer*** is simply obtained from multiplying the transmission values obtained from **Eq. 19-8** for bins *i* that are assigned to the particle layer. The result is independent of the value of *k* (smaller than 1) used in **Eq. 19-8**, which is explained by the convergence of the estimated particle transmission to 1 below the isolated particle layer, for all values of *k* smaller than 1.

19.3. Particle layer optical properties

Taking the ratio of the measured Rayleigh channel signal, **Eq. 19-4**, and the expected signal in clean air, **Eq. 19-5**, inside the particle layer yields the normalized Rayleigh channel signal:

$$S_{mc}(i) = \frac{FP(i)}{\hat{FP}(i)} = SR_X(i)T_p^2(i-1)\tau_p(i), \quad i = 1, \ldots\ldots, N$$

Eq. 19-9

showing that the known normalized Rayleigh channel signal can be written as the product of the cross-talk included scattering ratio, SR_X, in the current bin, total particle transmission in overlaying bins and particle transmission in the current bin. T and τ both refer to one-way transmission through the Aeolus bin. The particle layer covers N Aeolus bins. The cross-talk included scattering ratio is related to the well-known scattering ratio through

$$SR(i) = \frac{SR_X(i) - 1}{X} + 1$$

Eq. 19-10

To enable to solve the equations for scattering ratio and transmission inside the particle layer we first assume the Aeolus bins, covering the particle layer, completely filled with particles and homogeneity inside the particle layer, i.e., $\beta_p(i)=\beta_p$, hence $\tau_p(i)=\tau_p$. A further simplification could assume the molecular backscatter to be constant within the particle layer and all bins completely filled with particles. Then, both $SR(i)$ and $SR_X(i)$ are constant values, leaving inside the particle layer the following equation:

$$S_{mc}(i) = SR_X {\tau_p}^{2i-1}, \quad i = 1, \ldots\ldots, N$$

Eq. 19-11

Meaning we are left with two unknowns SR_X and τ_p for N equations. The left hand side of **Eq. 19-11** is known from the measured Rayleigh channel signal and its estimate in clean air from AUX_MET. The right hand side is the model function, in the remainder denoted $\hat{S}_{mc}(i, SR_X, \tau_p)$. Estimation of SR_X and τ_p is done through a minimization of the following cost function J:

$$J = \frac{1}{2}\sum_{i=1}^{N}\left[\hat{S}_{mc}(i, Y) - S_{mc}(i)\right]^2$$

Eq. 19-12

with Y the vector of parameters to be estimated. From the above $Y=(SR_X, \tau_p)$.

In case that the particle layer one-way transmission, τ_L, is known, e.g., from the ratio of the measured and expected signal in clean air, τ_p can be computed explicitly: $\tau_p = {\tau_L}^{1/N}$. The particle layer transmission cannot be obtained in case of an optically thick layer with no signal below the layer, or in case of the closely separated particle layer with too small clean atmosphere in between for a valid calculation or in case the particle layer is close to or at the surface leaving too small or no bins below the layer for a valid calculation.

Solving for SR_X from minimizing the cost function **Eq. 19-12**, next SR is obtained from SR_X using **Eq. 19-10** and given X (from the AUX_CAL calibration file). Known SR gives β_p from known β_m (through AUX_MET pressure and temperature). Particle extinction, α_p, is obtained from particle transmission and next the lidar ratio, γ, from the ratio of particle extinction and backscatter:

$$\alpha_p(i) = -\ln(\tau_p)/R(i), \qquad \gamma = \alpha_p(i)/\beta_p$$

Eq. 19-13

with $R(i)$ the range which the laser beam travels through bin i, which equals the Aeolus bin size divided by the cosine of the incidence angle.The above shows the general idea of estimating the optical properties for a detected particle layer. The following paragraphs will tackle some of the assumptions, making the calculations more realistic.

First, molecular backscatter within the particle layer is not constant, implying a non-constant SR_X within the particle layer. With some calculations one can show

$$SR_X(i) = 1 + \frac{\beta_m(1)}{\beta_m(i)}[SR_X(1) - 1]$$

Eq. 19-14

showing that each i, $SR_X(i)$ can be expressed as a function of $SR_X(1)$. Since $\beta_m(i)$ is known for all i, this leaves still only one unknown for SR_X in the parameter vector Y. This value is denoted SR_{X1}. Ones SR_{X1} (=$SR_X(1)$) is known than the corresponding values in the others bins follow from **Eq. 19-14** and the corresponding SR from **Eq. 19-10**.

Next, we tackle the assumption of extending the particle layer over complete Aeolus bins. In general the top and bottom bin will be only partially filled with particles, the intermediate bins completely. Hereto we introduce fractions f_1 and f_N, denoting the fraction of the Aeolus bin filled with particles. Clearly $f_2,\ldots,f_{N-1}$ all equal 1 when assuming a homogeneous particle layer. For the model function of bin 1 we then have:

$$\hat{S}_{mc}(1, SR_{X1}, \tau_p) = \frac{\beta_m(1) + f_1 X \beta_p(1)}{\beta_m(1)} \tau_p^{f_1} = \left[1 + f_1 X \frac{\beta_p(1)}{\beta_m(1)}\right] \tau_p^{f_1} = [1 + f_1(SR_{X1} - 1)]\tau_p^{f_1}$$

Eq. 19-15

where for the latter equality we used the relation $SR_X(i) = [\beta_m(i)+X\beta_p(i)]/\beta_m(i) = 1+X\beta_p(i)/\beta_m(i)$, hence $X\beta_p(i)/\beta_m(i) = SR_X(i)-1$. For i=2,N-1 we get

$$\hat{S}_{mc}(i, SR_{X1}, \tau_p) = SR_X(i)\tau_p^{2f_1+2i-3}, \quad i = 2, \ldots, N-1$$

Eq. 19-16

and for i=N:

$$\hat{S}_{mc}(N, SR_{X1}, \tau_p) = [1 + f_N(SR_X(N) - 1)]\tau_p^{2f_1+2(N-2)+f_N}$$

These additional two parameters extend the parameter vector to Y=(SR_{X1}, τ_p,f_1,f_N).

Finally, so far it has been assumed that all Aeolus bins capturing the particle layer are of equal size, which is not necessarily true. To account for this a reference bin size is introduced, Δz_{ref} and bins with different bin size referenced to this bin size through: $\Delta z(i) = \alpha_i \Delta z_{ref}$. All these values are known a priori. Adding this additional complexity to the model function yields:

$$\hat{S}_{mc}(1, Y) = [1 + f_1(SR_{X1} - 1)]\tau_p^{a_1 f_1}$$

$$\hat{S}_{mc}(i, Y) = SR_X(i)\tau_p^{2\alpha_1 f_1 + 2\sum_{j=2}^{i-1} a_j + a_i} \quad , i = 2, \ldots, N-1$$

$$\hat{S}_{mc}(N, Y) = [1 + f_N(SR_X(N) - 1)]\tau_p^{2\alpha_1 f_1 + 2\sum_{j=2}^{N-1} a_j + \alpha_N f_N}$$

Eq. 19-17

Substituting these in the cost function, together with **Eq. 19-14** and minimizing the cost function yields the best possible estimate given the measured Rayleigh signal in the particle layer. An additional constraint comes from the total one-way layer transmission, which should be met in case it is known:

$$\tau_L = \tau_p^{\alpha_1 f_1 + \alpha_N f_N + \sum_{j=2}^{N-1} a_j}$$

Eq. 19-18

Minimization of the cost function requires the gradient of the cost function to the parameters: $\partial J/\partial Y$. Starting from initial estimate Y(0), subsequent estimates are obtained through iteration: $Y(k+1) = Y(k) - \mu\,\partial J/\partial Y(k)$, with μ the step to be taken in the search direction (linear search) toward the cost function minimum. For the gradient of **Eq. 19-12** we can write

$$\frac{\partial J}{\partial Y} = \sum_{i=1}^{N} \left[\hat{S}_{mc}(i,Y) - S_{mc}(i)\right] \frac{\partial \hat{S}_{mc}(i,Y)}{\partial Y}$$

Eq. 19-19

The gradient at the end of Eq. 19-19 follows (almost) readily from **Eq. 19-17**. At convergence of the iteration process, the estimate of *Y* translates into estimates of the particle backscatter, extinction, scattering ratio and lidar ratio inside the particle layer, following the procedure described above **Eq. 19-13**.

19.3.1. Simplified fast algorithm

Rather than estimating the parameters through the minimization of a cost function as described in the previous section, a more simple and computationally less expensive method is to prescribe the layer location in the Aeolus bins, i.e., prescribe f_1 and f_N. Parameter τ_p then follows directly from Eq. 19-18 for given total layer transmission τ_L. Then one unknown, SR_{X1}, remains, which can be obtained explicitly from setting the gradient of the cost function to SR_{X1} equal to zero. After some simple mathematics we can write for SR_{X1}:

$$SR_{X1} = \frac{\sum_{i=1}^{N} S_{mc}(i)g(i) - \sum_{i=1}^{N} g(i)^2 \left[\frac{\beta_m(i)}{\beta_m(1)} - 1\right]}{\sum_{i=1}^{n} g(i)^2}$$

Eq. 19-20

With *g*(*i*) the gradient of the model function $\hat{S}_{mc}$ to SR_{X1} for Aeolus bin number *i*.

19.4. Height dependent scattering ratio threshold

Two options have currently been implemented for particle detection in the OPC: (i) by putting thresholds on the retrieved particle transmission and (ii) by putting a threshold on the scattering ratio. The latter is default in the current L2Bp and the threshold value taken constant throughout the atmosphere, but height dependent thresholding is an option the L2Bp. In fact it was found that the application of a constant threshold resulted in many false alarms in the uppermost FP bins for the OPC derived scattering ratio, due to measurement noise and relative low molecular backscatter at these altitudes. Instead a height dependent threshold is proposed below and was tested offline.

The threshold value is set to a reference threshold value, T_{ref}, e.g. T_{ref} =1.2 at a reference height, e.g., at 10 km. Molecular backscatter at reference height, β_m^{ref}, is easily obtained from the AUX_MET data, typically temperature and pressure, see [Appendix D. Retrieval of atmosphere molecular backscatter and extinction (Rayleigh scattering)]. Next, the corresponding reference particle backscatter, β_p^{ref}, yielding a scattering ratio of T_{ref} is obtained from $\beta_p^{ref} = \beta_m^{ref}(T_{ref}\text{-}1)$. The height-dependent scattering ratio threshold is defined as: $T_{sr}(z) = [\beta_p^{ref} + \beta_m(z)]/\beta_m(z)$. Because β_m is decreasing with altitude z, $T_{sr}(z)$ is increasing with altitude according to the right panel of Figure 19-1.

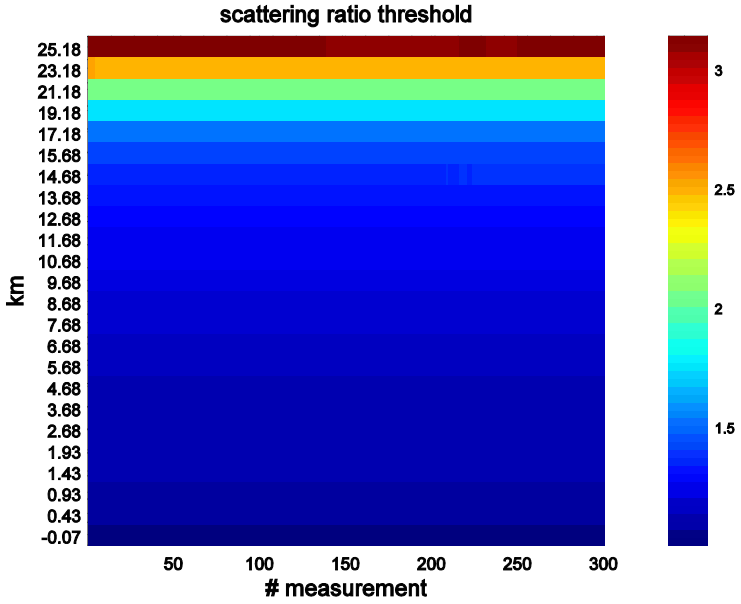


Figure 19-1. Height-dependent threshold for the scattering ratio, see the text for details.

20. Matchup algorithm between AUX_MET profiles and L1B measurements

Given the L1B measurement geolocations, and a set of auxiliary meteorological data available at reference geolocations, the purpose of the matchup algorithm is to obtain a selected subset of auxiliary meteorological data profiles that are the closest match in space and time to the L1B measurement locations.

The auxiliary meteorological data provided to the L2BP (via the AUX_MET_12 file) consists of typically 30 hours' worth of NWP model vertical profiles (and associated surface values). In general, the reference locations at which the meteorological profiles are given are close to the Aeolus measurement positions, but a one-to-one correspondence cannot be guaranteed. During the pre-processing step (see Appendix A.2) that generated the reference meteorological data, the locations may already have been interpolated to the exact BRC geolocations, but this is not a requirement. Even if an interpolation is done, only for the centre of the BRC will the profile match exactly (even then the perfect match is only for the near-surface range-bins). Due to the off-nadir LOS pointing angle of the laser (in wind mode) the vertical NWP model profiles will have a horizontal misplacement for range-bins at 30km altitude of around 23 km (however the ECMWF model effective resolution at the time of writing is ~40-80 km, so this will rarely matter).

Auxiliary meteorological data from the operational L2/Met PF is (at the time of writing) being provided along predicted orbit ground track geolocations every 3 seconds (every ~20 km).

Hence the L2BP needs an algorithm to choose which meteorological profile to use, and if required to interpolate between a set of nearby NWP profiles.

N.B. Other NWP centres may implement their own algorithm to enable for example interpolation between NWP profiles, in case use of high resolution data is desired. The L2BP provides two simple algorithms to be used here.

Inputs:

Symbol	Description	Input variable source
$\boldsymbol{\lambda_m(i,k)}$	Measurement-bin latitude	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/latitude_of_height_bin`
$\boldsymbol{\phi_m(i,k)}$	Measurement-bin longitude	ALD_U_N_1B: `Geoloc_ADS/Meas_Geoloc/{Mie,Rayleigh}_Rangebin_Geoloc/longitude_of_height_bin`
$\boldsymbol{t_m(k)}$	Measurement centroid date-time	ALD_U_N_1B: `Geoloc_ADS/Meas_AOCS/Measurement_Centroid_Time`
$\boldsymbol{D_{max}}$	The maximum allowed distance between the measurement and the NWP profile	AUX_PAR_2B: `AMD_Matchup_Params/Max_Allowed_Distance`
$\boldsymbol{\Delta T_{max}}$	The maximum allowed time difference between measurement and NWP profile	AUX_PAR_2B: `AMD_Matchup_Params/Max_Allowed_Time_Diff`
$\boldsymbol{\lambda_{nwp}(j)}$	AUX_MET profile latitude	AUX_MET_12: `Geoloc_ADS_off_nadir/Latitude`
$\boldsymbol{\phi_{nwp}(j)}$	AUX_MET profile longitude	AUX_MET_12: `Geoloc_ADS_off_nadir/Longitude`
$\boldsymbol{t_{nwp}(j)}$	AUX_MET profile date-time	AUX_MET_12: `Geoloc_ADS_off_nadir/DateTime`

Outputs:

Symbol	Description	ALD_U_N_2B file variable
N/A	The index of the selected NWP profile	`L2B Meas Product Confidence Data (PCD) ADS/L2B_AMD_Collocation`, see Table 26 of the IODD [RD10].

Currently two matching algorithms are available:

- Dummy Matchup
- Nearest Neighbour Matchup (**the recommended option**)

20.1. Dummy Matchup algorithm

This option is useful for development and testing purposes. The Dummy algorithm does not take any lat, lon or date, time information into account. It assumes that for every BRC an appropriate NWP profile is available, and matches them using the order in which they appear in the input files. The output index is simply the L1B BRC number. So BRC 1 in the L1B file is always combined with NWP profile 1 in the AMD file, BRC 2 in the L1B file is always combined with NWP profile 2 in the AMD file, etc.
The algorithm is selected by setting the AUX_PAR_2B `AMD_Matchup_Params/Matchup_Method` to: `Dummy`

20.2. Nearest Neighbour Matchup algorithm

The **recommended** matchup algorithm is the Nearest Neighbour Matchup algorithm.
The algorithm is selected by setting the AUX_PAR_2B `AMD_Matchup_Params/Matchup_Method` to: `Nearest_Neighbour`
This consists of the following steps for each measurement:

- Select a representative L1B Mie measurement-bin number latitude, longitude and centroid date-time
- Apply a time window on all available NWP profiles
- Apply a distance range check to all NWP profiles
- Loop through all NWP profiles that passed the above two checks and choose the profile closest to the L1B measurement-bin geolocation

No interpolation between NWP profiles is implemented at the moment.

The four steps are implemented as follows:

20.2.1. Select the L1B measurement-bin geolocation

The geolocation of the Mie channel measurement-bin half-way up the measurement profile is chosen to represent the measurement geolocation i.e. that of range-bin number 12 (of 24). This is a compromise that should be reasonably representative of the measurement in question.

20.2.2. Apply time-window

Inputs:

- Date and time of the L1B measurement-bin and of the current NWP profile

- Max_Allowed_Time_Diff $\Delta T_{\max}$ from AUX_PAR_2B file.

Outputs:

- decision: in time-window or not

Algorithm:

- Calculate the time difference in seconds between both date, time pairs

$$\Delta T = s_{day} * (d2 - d1) + (s2 - s1) \qquad \textbf{Eq. 20-1}$$

- If $\Delta T < \Delta T_{\max}$ then this NWP profile may be used

20.2.3. Apply a distance range

Inputs:

- lat, lon of the L1B measurement-bin and of the current NWP profile
- Max_Allowed_Distance $D_{\max}$ from AUX_PAR_2B

Outputs:

- profile within range or not

Algorithm:

- convert both lat, lon pairs to (x,y,z) unit vectors
-

$$x = \cos(\lambda)\cos(\delta)$$
$$y = \sin(\lambda)\cos(\delta) \qquad \textbf{Eq. 20-2}$$
$$z = \sin(\delta)$$

- calculate the angle between both vectors α in radians, assuming the earth is a perfect globe using the goniometric property that the inner product between two unit vectors equals the cosine of the angle between them

$$\alpha = \text{acos}(x1 * x2 + y1 * y2 + z1 * z2) \qquad \textbf{Eq. 20-3}$$

- convert the angle distance α to a surface distance D in km, using the earth's radius R in km: $D = R * \alpha$. We use the equatorial earth radius $R_{eq} = 6378.1\,\text{km}$
- if $D \le D_{\max}$ then this NWP profile may be used

20.2.4. Choose closest NWP profile

Inputs:

- lat, lon of the L1B measurement-bin and of the allowed set of NWP profiles that passed the distance range and time-window tests

Output:

- index of the selected NWP profile

Algorithm:

- Calculate the distance (as described above) between the L1B measurement-bin and each allowed NWP profile in a loop and select the NWP profile closest to the BRC centre position. Then report its index.

21. Potential future algorithm developments

- Input data screening
  - Better screening of AUX_RBC_L2 files:
    - This could be improved in future with the checking of expected ranges for the variables related to the Rayleigh response curves.
  - Add screening of the AUX_CAL_L2 and AUX_CLM_L2 files.
- Classification:
  - Could add in an extra observation type: “aerosol”, to cover conditions with low scattering ratio, but > 1.
- Optical Properties Code:
  - Use of AUX_CLM_L2 file in the Optical Properties Code. Currently, the lidar ratio is calculated explicitly by OPC, but in case valid a priori knowledge is available from AUX_CLM_L2 this could be used by the algorithm and improve the accuracy of the calculated remaining optical properties parameters.
- Measurement-bin selection:
  - It could be made more flexible in future to also allow vertical averaging over a number of range bins as well as measurements if this is found to be necessary (but this comes with a loss of vertical resolution).
- Selecting reference temperature and pressure data from RBC:
  - The intention is to replace nearest neighbour interpolation with a linear or higher order interpolation (more appropriate for more widely spaced NWP model levels at ~30 km).
- Observation level scattering ratio for use in the Rayleigh processing (Mie decontamination) could be derived via a new L2Bp scattering ratio algorithm using overlapping Mie-range bins (if available)
- Error quantifiers:
  - The error quantifier defined in Section 17.1 requires some further minor improvement. The details given in Section 17.1 are only for the error quantifier pertaining to the ILIAD HLOS wind estimate (Sections 13.1 and 13.3). Further contributions to the error are required for the post-ILIAD terms (Sections 13.4 and 13.5). The same applies to the post-Mie core algorithm terms of the Mie HLOS wind equation.

Appendix A. Auxiliary meteorological data processing

This appendix describes auxiliary meteorological data processing, i.e. provision of auxiliary meteorological data (AUX_MET_12 file), as performed at the operational Level-2B processing facility (i.e. L2/Met PF hosted by ECMWF). Other meteorological centres can adapt the processing methods described here to suit their own needs. For further guidance see [RD6].

A.1. Determine reference locations for computing auxiliary meteorological data

Inputs:

- L1B geolocation information (latitude, longitude, time) for each measurement and observation. As a minimum, the geolocation of the intersection of the DEM (digital elevation model) and the instrument's line-of-sight, applicable for the observation centroid (L1B tags "Geolocation_of_DEM_Intersection" and "Observation_Centroid_Time").
- (Operational L2/Met PF only). Predicted orbit derived ground-track locations. One GRND_TRACK file per 24 hours is planned to be provided to the L2/Met PF by ESA (see [AD3]), each covering a period of one repeat cycle (7 days) ahead. The most recent predictions are used.

Outputs:

- Reference locations (latitudes, longitudes and times) for computation of auxiliary meteorological data

Algorithm:

- For each valid L1B dataset record received in time for nominal NRT processing, a reference location (latitude, longitude, time) is selected, at which auxiliary meteorological data are to be computed/extracted. The reference location can be that of the DEM intersection available at the BRC centroid, as specified in L1B data, or even use the measurement-level geolocations for more highly sampled data. The meteorological data are to take the form of vertical profiles at each reference location, together with any model information (details of vertical and/or horizontal grid structure) that may be required to facilitate collocation to L1B measurement locations.
- (Operational L2/Met PF only). Reference locations are nominally selected from the predicted ground track file (GRND_TRACK) - hence no need to wait for L1B data. The GRND_TRACK file is provided by FOS to the L2/Met PF once per day with predicted ground track (nadir and off-nadir) every 3 seconds for 7 days ahead.
- (Operational L2/Met PF only). The locations determined by i) and ii) are merged. The merge sub-task is achieved with a L2B processing tool, which reads the available L1B files and the predicted orbit track data and merges the geolocation data such that predicted orbit geolocations are only used where L1B is absent (for whatever reason). The merged result is written to the ECMWF Observation DataBase 2 format (ODB-2), so it then is fed into the ECMWF data assimilation system for the model profile information to be extracted. This merging might be necessary if the accuracy of the predicted orbit ground track geolocations is not sufficient.

A.2. Generate auxiliary meteorological data at reference locations

Description:

i. At each reference location (input), compute the auxiliary meteorological data required for collocation to L1B measurement locations and for use in the subsequent HLOS retrieval (Table 4.1/4.2). The data to be computed also consists of meteorological parameters (related to humidity, cloud and winds) that are not required for L2B wind processing but which may be useful in L2A processing. In operational mode (L2/Met PF), the data will be vertical profiles (not along the slant path of the laser). Collocation to L1B measurement locations occurs in a separate task ("pre-processing", Item 5.3). The ECMWF model vertical profiles are provided with geometric altitudes above mean sea level which is approximately equal to the geoid (hence we say they are with respect to the geoid). Selection of the reference locations supplied as input is explained in Section 5.1. Consequently, such selection may be driven either by the geolocation information of L1B data, by predicted ground track information, or by the available meteorological model. Precise details will depend on the particular implementation of L2B processing at different sites,

Inputs:

- Reference locations (latitude, longitude, time) for computation of auxiliary meteorological data
- Meteorological fields

Outputs:

- Meteorological profile data at reference locations, i.e. prior to pre-processing (Table 4.1/4.2). The vertical coordinates of the meteorological input are retained.

Algorithm:

- The task is achieved through integration of the reference geolocations into dummy observations to be used in the ECMWF data assimilation system. A standard part of the system, known as screening job, involves 1) the computation of background meteorological fields, and 2) interpolation in space and time to the location of meteorological observations. The background fields are obtained from a short-term forecast initialized with the meteorological analysis from the previous assimilation window.
- By treating the L1B (and predicted orbit) geolocation data as meteorological observations, the meteorological profile data needed for Aeolus processing are obtained.

A.3. Pre-process auxiliary meteorological data

Preprocessing of auxiliary meteorological data involves storing auxiliary meteorological profile at the measurement scale of L1B data i.e. selecting the closest available AUX_MET data to the L1B measurement-bin geolocation. See section 20 on AUX_MET matchup algorithms.

Appendix B. Mie error quantifier derivation

This appendix clarifies the assumptions, and provides details of the mathematics, that lead to the expression proposed in Section 18.2 for predicting the level of error on Mie winds.

The Mie core algorithm estimates the frequency of the Mie return by minimizing the cost function:

$$J(f)=\sum_{j=3}^{18}\left[N(i,j)-\mu_j(f)\right]^2 \tag{B1}$$

where $N(i,j)$ is a weighted photo-count for i-th height-bin and j-th CCD bin and $\mu_j(f)$ is a prediction of the same number based on an ad-hoc model (at present, a Lorentzian on top a uniform level of background light).

The photo-counts $N(i,j)$ are computed from the measurement level CCD photo-counts $LID(i,j,k)$ (after the removal of the video offset):

$$N(i,j)=\sum_{k=1}^{N_{meas}} w_k LID(i,j,k) \tag{B2}$$

where N_{meas} is the total number of measurements and w_k are the weights allocated to measurement k.

In equation (B1), the minimization bears only on the frequency, while 4 parameters are optimized by the Mie core algorithm (the central frequency, the amplitude, the width and the uniform level of background light). Here, we simplify the problem by limiting the optimization to the single frequency parameter, hoping that the equation we derive will apply with no major deficiency to the more complex case.

Let us denote by f_0 the frequency that optimizes the cost function

$$\bar{J}(f)=\sum_{j=3}^{18}\left[\bar{N}(i,j)-\mu_j(f)\right]^2 \tag{B3}$$

where $\bar{N}(i,j)$ denotes the mathematical expectation of N_j. Then, let us approximate $J(f)$ around f_0 by a second order expansion:

$$J(f)\approx J(f_0)+\frac{\partial J}{\partial f}(f-f_0)+\frac{1}{2}\frac{\partial^2 J}{\partial f^2}(f_0)(f-f_0)^2 \tag{B4}$$

and assume that the frequency $\hat{f}$ that minimizes $J(f)$ is very close the frequency that minimizes (B3). This assumption can be written:

$$\hat{f}\approx f_0-\left[\frac{\partial^2 J}{\partial f^2}(f_0)\right]^{-1}\frac{\partial J}{\partial f}(f_0) \tag{B5}$$

Now, let us denote $J'(f)=J(f)-\bar{J}(f)$ and assume

$$\frac{\partial J'}{\partial f}(f_0)<<\frac{\partial \bar{J}}{\partial f}(f_0)\ \text{and}\ \frac{\partial^2 J'}{\partial f^2}(f_0)<<\frac{\partial^2 \bar{J}}{\partial f^2}(f_0) \tag{B6}$$

Equation (B4) can then be approximated by the first order expansion

$$\hat{f} - f_0 \approx -\left[\frac{\partial^2 \bar{J}}{\partial f^2}(f_0)\right]^{-1}\left[\frac{\partial \bar{J}}{\partial f}(f_0) + \frac{\partial J'}{\partial f}(f_0)\right] = -\left[\frac{\partial^2 \bar{J}}{\partial f^2}(f_0)\right]^{-1}\frac{\partial J'}{\partial f}(f_0) \tag{B7}$$

From which follows that

$$\left\langle \left(\hat{f} - f_0\right)^2 \right\rangle = \left[\frac{\partial \bar{J}}{\partial f}(f_0)\right]^{-2}\left\langle \frac{\partial J'}{\partial f}(f_0)\frac{\partial J'}{\partial f}(f_0)\right\rangle \tag{B8}$$

From (B1) and (B3), we can write

$$J'(f) = \sum_{j=3}^{18} N'(i,j)^2 - 2\sum_{j=3}^{18} N'(i,j)\left[\bar{N}(i,j) - \mu_j(f)\right] \tag{B9}$$

so

$$\frac{\partial J'}{\partial f}(f_0) = 2\sum_{j=3}^{18} N'(i,j)\alpha_j(f_0) \tag{B10}$$

where $\alpha_j(f_0) = \partial\mu_j(f_0)/\partial f$. Since $\langle N'(i,j)N'(i,k)\rangle = \sigma_{i,j}^2\,\delta(j-k)$ (the random fluctuations of the photo counts are independent), it follows that

$$\left\langle \frac{\partial J'}{\partial f}(f_0)\frac{\partial J'}{\partial f}(f_0)\right\rangle = 4\sum_{j=3}^{18}\sigma_{i,j}^2\alpha_j^2 \tag{B11}$$

From equation (B3)

$$\frac{\partial^2 \bar{J}}{\partial f^2}(f_0) = -2\sum_{j=3}^{18}\frac{\partial \alpha_j}{\partial f}(f_0)\left[\bar{N}(i,j) - \mu_j(f_0)\right] + 2\sum_{j=3}^{18}\alpha_j^2(f_0) \tag{B12}$$

The first term can be neglected if $\bar{N}(i,j) \approx \mu_j(f_0)$. This condition should be met as long as the model $\mu_j(f)$ is a good model for the photo-counts $\bar{N}(i,j)$, so we can make the approximation

$$\frac{\partial^2 \bar{J}}{\partial f^2}(f_0) = 2\sum_{j=3}^{18}\alpha_j^2(f_0) \tag{B13}$$

Now, combining (B11) and (B13) gives

$$\boxed{\left\langle \left(\hat{f} - f_0\right)^2 \right\rangle \approx \left[\sum_{j=3}^{18}\alpha_j^2(f_0)\right]^{-2}\sum_{j=3}^{18}\sigma_{i,j}^2\alpha_j^2(f_0)} \tag{B14}$$

This is the basis for the expression given in Section 18.2 which uses the photo-count model

$$\mu_j(f) = T_j^{obs}\int_{f_j^-}^{f_j^+}\left[\frac{2A}{\pi\Delta}\left[1 + \frac{4(x-f)^2}{\Delta^2}\right]^{-1} + B\right]dx \tag{B15}$$

where f_j^+ and f_j^+ are the upper and lower frequency bounds of CCD bin j, T_j^{obs} is the tripod obscuration factor for bin j, Δ is the FWHM of the Lorentzian spectrum assumed for the Mie return, A is its amplitude, and B is the uniform level of background light. Considering this model, it follows that

$$\alpha_j(f_0) = T_j^{obs}\frac{2A}{\pi\Delta}\left[\frac{1}{1+\frac{4\left(f_j^{+}-f_0\right)^2}{\Delta^2}} - \frac{1}{1+\frac{4\left(f_j^{-}-f_0\right)^2}{\Delta^2}}\right] \quad \text{(B16)}$$

In practice, A, Δ and f_0 are approximated with the parameters estimated by the Mie core algorithm, the link between both sets being

$$\begin{aligned} \text{peak\_height} &\leftrightarrow \frac{2A}{\pi\Delta} \\ \text{FWHM} &\leftrightarrow \Delta \\ \text{frequency estimate} &\leftrightarrow f_0 \end{aligned} \quad \text{(B17)}$$

It now remains to give an expression for $\sigma_{i,j}^2$. If we assume that the random fluctuations of $LID(i,j,k)$ follow independent, Poisson statistics, we have

$$\sigma_{i,j}^2 = \sum_{k=1}^{N_{meas}} w_k^2 \overline{LID}(i,j,k) \quad \text{(B18)}$$

which we can approximate by

$$\sigma_{i,j}^2 \approx \sum_{k=1}^{N_{meas}} w_k^2 LID(i,j,k) \quad \text{(B19)}$$

Appendix C. Particle transmission computation from the Fabry-Pérot channel signal

Particle transmission is computed from the measured and estimated Fabry-Pérot (FP) signal from Eq. 19-8. Here we describe the iterative scheme implemented in the optical properties code (OPC) to estimate the particle transmission profile, starting at the topmost bin. The used signals in the scheme have been defined in Eq. 19-2, Eq. 19-4, Eq. 19-5 and Eq. 19-6. For the measured signal in Eq. 19-4 we adopt a slightly different notation here that does not harm the calculations but will later on show an intriguing property of the implemented scheme. For the measured FP signal we write here:

$$FP(i) = \left(1-\delta(i)\right)FP_{\mathrm{mol}}(i)\overline{\mathrm{T}}_p^2(i-1)\bar{\tau}_p(i)$$

Eq. C-0-1

with $\delta(i)$ accounting for cross-talk and measurement noise in the FP measurement and the overbar denoting the (unknown) true particle transmission. The following iterative (i.e. layer-stepping) scheme has been implemented to retrieve particle transmission in bin *i* from the measured FP signal.

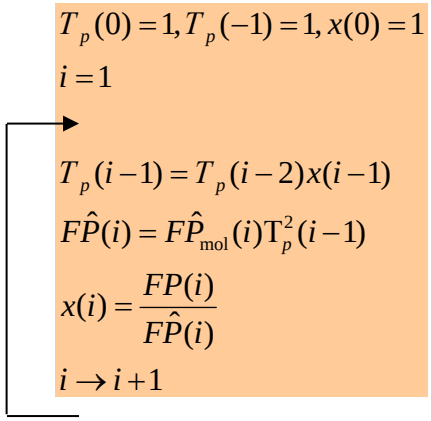


Figure C-1. Iterative scheme to retrieve particle transmission in bin *i*, from the measured Fabry-Pérot signal, *FP*(*i*).

In the scheme in Figure C-1, $F\hat{P}(i)$ denotes the expected FP signal in bin *i* when there are no particles in bin *i*. The idea is that the quotient of the measured and expected FP signal in bin *i* (parameter *x*) provides an estimate of the particle transmission in bin *i*.
To study the behaviour of the scheme above, a simplified atmosphere is assumed with a constant aerosol transmission profile: $\bar{\tau}_p(i) = \bar{\tau}_p$ and thus $\overline{T}_p(i) = \left(\bar{\tau}_p\right)^i$. From Eq. C-0-1 we may then write $FP(i) = \left(1-\delta(i)\right)FP_{\mathrm{mol}}(i)\left[\bar{\tau}_p\right]^{2i-1}$. Also the error δ is assumed constant throughout the profile for simplicity. The solution of the scheme in Figure C-1 with these assumptions is then easily obtained:

$$x(i) = \bar{\tau}_p(1-\delta)^{(-1)^{i+1}} \qquad i \geq 1$$

Eq. C-0-2

showing that the parameter *x* oscillates around the true particle transmission value. Subsequent estimate for aerosol extinction follows from

$$\begin{aligned}\alpha^{\mathrm{p}}(i) &= \frac{-\ln(x(i))}{\Delta z} \\ &= \frac{-\ln(\bar{\tau}^{\mathrm{p}})}{\Delta z} + \frac{(-1)^{i+1}\ln(1-\delta)}{\Delta z} \\ &\approx \frac{-\ln(\bar{\tau}^{\mathrm{p}})}{\Delta z} + \frac{(-1)^{i+1}\delta}{\Delta z} \\ &= \bar{\alpha}^{\mathrm{p}}(i) + \Delta\alpha^{\mathrm{p}}(i)\end{aligned}$$

Eq. C-0-3

showing that the retrieved aerosol extinction obviously also alternates around the true value. This behaviour is visualized by the red curves in Figure C-3. From Eq. C-0-3, the relative error in the retrieved aerosol extinction equals: $\left|\Delta\alpha^{\mathrm{p}}(i)\right| / \bar{\alpha}^{\mathrm{p}}(i) = \left|\delta\right| / \ln(\bar{\tau}^{\mathrm{p}})$.

A solution to damp the alternating behaviour of the particle optical properties solution is obtained by the alternative scheme proposed in Figure C-2 that only differs from Figure C-1 through the power *k* that is applied to the computation of *x*.

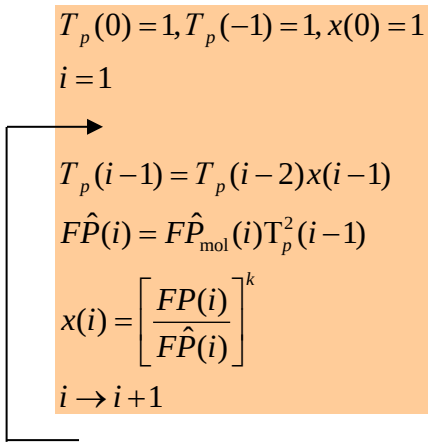


Figure C-2. Same as Figure C-1, but with a power *k* in the computation of parameter *x*.

The solution of the iterative scheme in Figure C-2 is

$$x(j) = \bar{\tau}_p^{\,\beta(j,k)} \left\{ \left[(1-\delta)\bar{\tau}_p \right]^k \right\}^{1-\beta(j,k)}$$

Eq. C-0-4

with $\beta(j,k)$ a polynomial of *k* as follows

$$\begin{aligned}\beta(1,k) &= 0 \\ \beta(2,k) &= 2k \\ \beta(3,k) &= 4k - 4k^2 \\ \beta(4,k) &= 6k - 12k^2 + 8k^3 \\ \beta(5,k) &= 8k - 24k^2 + 32k^3 - 16k^4\end{aligned}$$

etc.

A general expression for $\beta(j,k)$ is described later in the text. Here we constrain to the properties of $\beta(j,k)$

- for k=1, the value of $\beta(j,k)$ alternates between 0 and 2 and Eq. C-0-4 is in agreement with Eq. C-0-2.
- For k=1/2: $\beta(j,1/2)$=1 for j>0.
- $\lim_{j\to\infty}\beta(j,k)=1, \quad \text{for } 0<k<1$

The last property implies that for suitable values of k, the impact of the error δ in the estimate for particle transmission damps with increasing j, i.e., further down into the profile. This is visualized in Figure C-3 that shows aerosol extinction retrieval for different values of the power k.

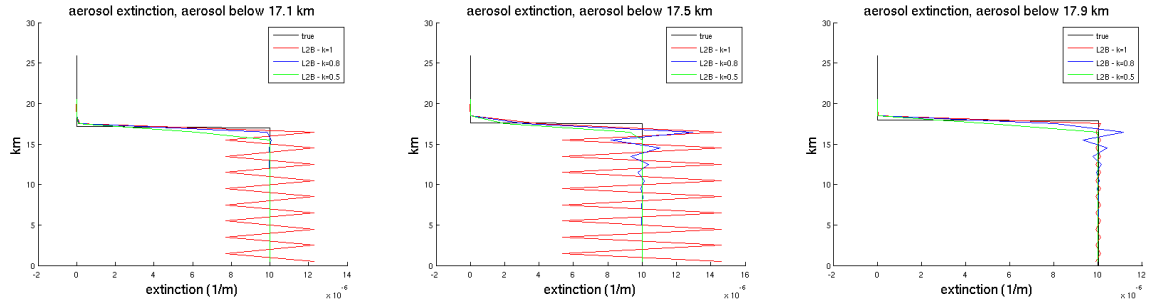


Figure C-3. L2B retrieval of aerosol extinction from the FP signal. No aerosol backscatter contamination is assumed. There is no aerosol attenuation at the top of the atmosphere. Aerosol extinction has a constant value of 10^{-5} 1/m (black lines) below 17.1 km (left panel) corresponding to a moderate value for δ, below 17.5 km (middle panel) corresponding to a large value for δ, and below 17.9 km (right panel) corresponding to a small value for δ. The red, blue and green lines show the L2B retrieval for k=1, 0.8 and 0.5 (see Figure C-2) respectively.

From Figure C-3 it follows that the aerosol optical properties retrieval depends on the position in the first range bin where substantial aerosol is found. This location determines the value of δ. In any case the solution for k=1 alternates around the true value, the amplitude depending on the location of aerosols within the first aerosol contaminated bin. For values of k smaller than one, the retrieved aerosol extinction converges to the true value at the expense of a bias in the first aerosol contaminated bin. This is explained later in the text.

Figure C-4 shows similar results for a linear increasing aerosol extinction.

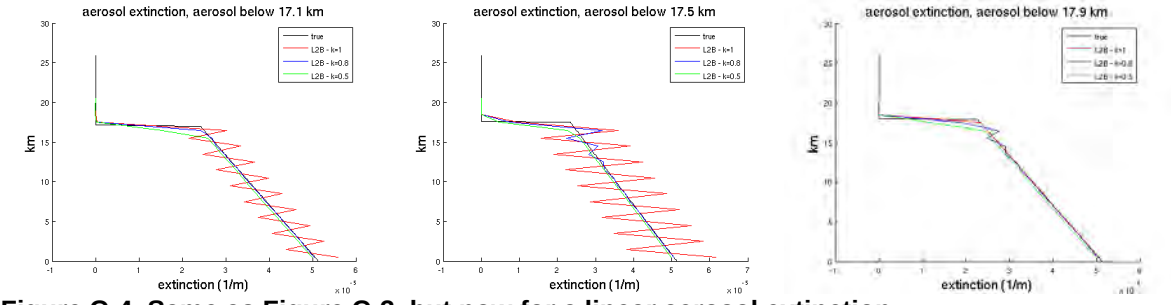


Figure C-4. Same as Figure C-3, but now for a linear aerosol extinction.

The behaviour of the retrieval algorithm is further explored in the next section for noisy FP signals.

D.1 Iterative scheme for non-constant particle transmission profile and noisy FP signals

In general, particle backscatter and transmission are a non-constant function of altitude. In addition, measured FP signals are contaminated with random noise. The solution of the scheme in Figure C-2 with non-constant $\delta(i)$ and non-constant (unknown) true particle transmission profile $\bar{\tau}^{\mathrm{p}}(j)$ is

$$x(i) = \gamma(i)\prod_{j=1}^{i}\left[\frac{\bar{\tau}_p(j)}{\gamma(j)}\right]^{\alpha(i-j+1,k)}$$
$$\gamma(i) = \left[(1-\delta(i))\bar{\tau}_p(i)\right]^k$$

Eq. C-0-5

with for α:

$$\alpha(1,k) = 0$$
$$\alpha(2,k) = 2k$$
$$\alpha(3,k) = 2k - 4k^2$$
$$\alpha(4,k) = 2k - 8k^2 + 8k^3$$
$$\alpha(5,k) = 2k - 12k^2 + 24k^3 - 16k^4$$
$$\text{etc.}$$

Parameter α has the following properties:

$$\sum_{j=1}^{i}\alpha(j,k) = \beta(j,k)$$

Eq. C-0-6

$$\alpha(j,k) = 2k\left(1-\beta(i-j,k)\right)$$

Eq. C-0-7

For the special case of a constant particle transmission profile, $\bar{\tau}_p(i) = \bar{\tau}_p$ and constant error $\delta(i)=\delta$, $\gamma(j) = \gamma = \left[(1-\delta)\bar{\tau}_p\right]^k$ and Eq. C-0-5 becomes $x(i) = \gamma\left(\bar{\tau}_p/\gamma\right)^{\alpha(1,k)+\alpha(2,k)+\cdots+\alpha(i,k)} = \gamma\left(\bar{\tau}_p/\gamma\right)^{\beta(i,k)}$, using Eq. C-0-6. Substituting γ into this equation yields Eq. C-0-4. Also from Eq. C-0-6 and Eq. C-0-7 a general expression for β is found:

$$\beta(1,k) = 0$$
$$\beta(i,k) = 2k\sum_{l=1}^{i-1}\left(1-\beta(i-l,k)\right), \qquad i>1$$

Eq. C-0-8

Substituting Eq. C-0-7 into Eq. C-0-5 yields

$$x(i) = \gamma(i)\prod_{j=1}^{i-1}\left[\frac{\bar{\tau}_p(j)}{\gamma(j)}\right]^{2k\left[1-\beta(i-j,k)\right]}$$
$$\gamma(i) = \left[(1-\varepsilon(i))\bar{\tau}_p(i)\right]^k$$

Eq. C-0-9

In other words, for *i* substantially larger than *j*, $\beta(i\text{-}j,k)$ is (close to) zero meaning that the impact of errors $\delta(j)$, e.g., measurement noise, in bin *j* (through $\gamma(j)$) does (almost) not contaminate $x(i)$, i.e., the particle transmission estimate in bin *i*, the amount of contamination depending on the value of *k* and the difference *i-j*.

Examples

1. *k*=1

For *k*=1, $\beta(j)$ alternates between 0 and 2. Then Eq. C-0-9 becomes

$$\begin{aligned} x(i) &= \gamma(i)\prod_{j=1}^{i-1}\left[\frac{\bar{\tau}_p(j)}{\gamma(j)}\right]^{2(-1)^{i-j+1}} \\ &= \gamma(i)\left[\frac{\bar{\tau}_p(i-1)}{\gamma(i-1)}\right]^{2}\left[\frac{\bar{\tau}_p(i-2)}{\gamma(i-2)}\right]^{-2}\cdots\left[\frac{\bar{\tau}_p(1)}{\gamma(1)}\right]^{2} \\ &= \bar{\tau}_p(i)\big(1-\delta(i)\big)\big[1-\delta(i-1)\big]^{-2}\big[1-\delta(i-2)\big]^{2}\cdots \end{aligned}$$

Eq. C-0-10

This yields for the expected value and standard deviation of $x(i)$:

$$\begin{aligned} \mathrm{E}\big[x(i)\big] &= \bar{\tau}_p(i) \\ \mathrm{std}\big[x(i)\big] &\approx \sqrt{\big[\sigma(i)\big]^2 + \sum_{j=1}^{i-1}\big[2\sigma(j)\big]^2} \\ &\approx \sigma\sqrt{1+4(i-1)} \quad \text{for constant } \sigma(i) \end{aligned}$$

Eq. C-0-11

assuming for the relative (noise/signal) measurement error δ, $\mathrm{E}[\delta(i)]=0$ and with $\sigma(i)$ denoting the standard deviation of $\delta(i)$. For the computation of the standard deviation in **Eq. C-0-11** a Taylor expansion for the negative powers has been applied and uncorrelated noise between subsequent bins was assumed.
Thus from Eq. C-0-10, Eq. C-0-11, the estimate $x(i)$ for particle transmission in bin *i* alternates around the true value. The error increases with increasing bin number, i.e., further down into the profile since the particle transmission estimate in the current bin is contaminated by measurement noise in the current and all bins above.

2. *k*=1/2

For *k*=1/2, $\beta(1)=0$, $\beta(j)=1$ j>1. Then Eq. C-0-9 becomes

$$
\begin{aligned}
x(i) &= \gamma(i)\left[\frac{\bar{\tau}_p(i-1)}{\gamma(i-1)}\right]^{2k[1-\beta(1,k)]} \\
&= \gamma(i)\frac{\bar{\tau}_p(i-1)}{\gamma(i-1)} \\
&= \left[(1-\delta(i))\bar{\tau}_p(i)\right]^{1/2}\frac{\bar{\tau}_p(i-1)}{\left[(1-\delta(i-1))\bar{\tau}_p(i-1)\right]^{1/2}} \\
&= \left[\frac{(1-\delta(i))}{(1-\delta(i-1))}\right]^{1/2}\left(\bar{\tau}_p(i)\bar{\tau}_p(i-1)\right)^{1/2}
\end{aligned}
$$

Eq. C-0-12

This yields for the expected value and standard deviation of $x(i)$:

$$
\begin{aligned}
\mathrm{E}[x(i)] &= \left[\bar{\tau}_p(i-1)\bar{\tau}_p(i)\right]^{1/2} \\
\mathrm{std}[x(i)] &= \sqrt{\left[\frac{1}{2}\sigma(i)\right]^2+\left[\frac{1}{2}\sigma(i-1)\right]^2} \\
&\approx \sigma/\sqrt{2}
\end{aligned}
$$

Eq. C-0-13

i.e. the estimate $x(i)$ for particle transmission in bin i has a bias $\bar{\tau}_p(i)\left(1-\sqrt{\bar{\tau}_p(i-1)/\bar{\tau}_p(i)}\right)$, which is small for a smooth particle attenuation profile but large for e.g. aerosol gradients. However in case of large aerosol gradients cross-talk on the FP signal through aerosol contamination will be detected by the algorithm and subsequent flagging or advanced processing is applied. Besides the bias note the small error due to measurement noise contamination, i.e., only noise in the current and previous bin contaminate the particle transmission estimate.

For values of k larger than 1/2 and smaller than 1, it can be shown that the bias reduces with increasing values of k but at the expense of an increased standard deviation of error in the particle transmission estimate.

Results

The left panel in Figure C-5 shows a typical noisy realization from a noise-free FP signal based on the linear aerosol extinction profile of Figure C-3. Gaussian distributed random noise with zero mean and standard deviation equal to the square root of the number of photon counts of the noise-free signal has been added to the noise-free profile. The right panel of Figure C-5 shows the corresponding retrieved aerosol extinction for different values of the power k in the iterative scheme.

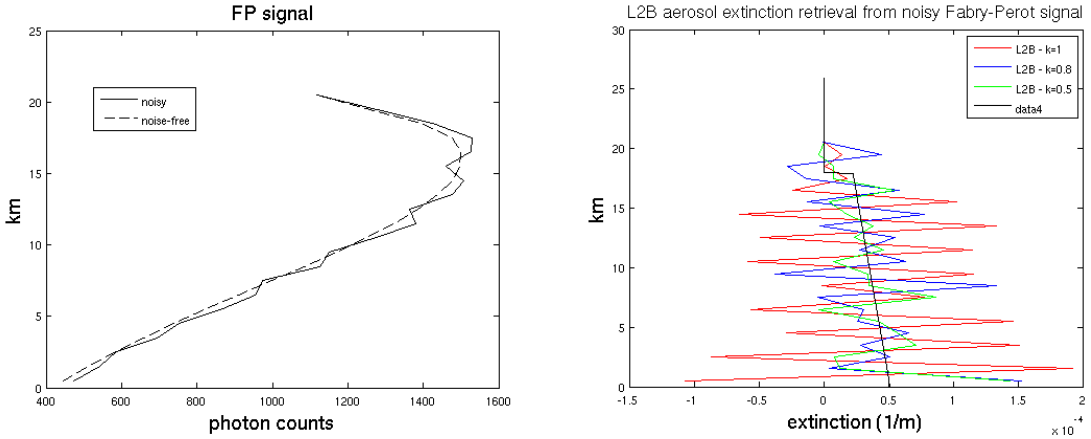


Figure C-5. Left panel, simulated Fabry-Perot signal with linearly increasing aerosol attenuation according to Figure C-4. The dashed line corresponds to the noise-free signal, the solid line to a typical noisy realization, assuming Gaussian distributed noise with the mean and standard deviation equal to the square of the number of photon counts. The right panel shows the corresponding aerosol extinction retrieval for *k*=1 (red), *k*=0.8 (blue) and *k*=0.5 (green). The black line denotes the true aerosol extinction.

Figure C-6 show the statistics of retrieved aerosol extinction obtained from 10000 retrievals each with another noise realization applied to the FP signal. Depending on the location of substantial aerosol attenuation within the first aerosol contaminated bin a large (left panel) or negligible (middle panel) bias is found for *k*=1 in the iterative scheme in Figure C-2. Using other values for *k* in the iterative scheme introduces a relative large negative bias in the first aerosol contaminated bin but smaller biases of less than 10% in the lower bins, this percentage quickly decreasing in subsequent bins, in agreement with the analysis above. Larger biases are found for smaller values of *k*. On the other hand the standard deviation of error decreases substantially with decreasing values for *k*. A reduction of almost a factor of 10 is achieved when using *k*=0.5 instead of *k*=1.

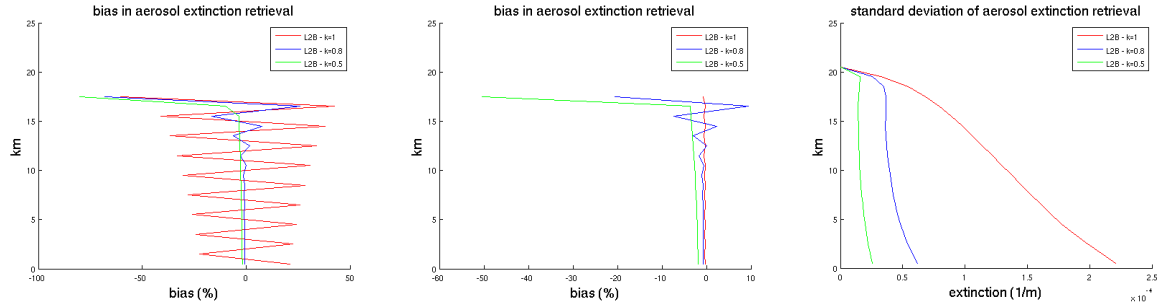


Figure C-6. Statistics of aerosol backscatter retrieval from noisy Fabry-Perot channel data. Bias for the case of aerosol attenuation below 17.5 km (left panel) and below 17.9 km (middle panel). The right panel show the standard deviation that is similar for both cases.

Appendix D. Retrieval of atmosphere molecular backscatter and extinction (Rayleigh scattering)

The optical properties code (OPC) as described in Section 19, requires as input estimates of molecular backscatter and extinction throughout the measurement profile. OPC has two options discussed in the following subsections to obtain these inputs. The choice is controlled via the AUX_PAR_2B file: `Optical_Properties_Params/MolExt_Method`

Option 1: `MolExt_Method_Parametrised`
Option 2: `MolExt_Method_Proper`

Option 1: Basic parameterization

A simple parameterization for molecular backscatter coefficient can be used in case information of atmospheric pressure and temperature is not available (however it should always be available from the compulsory AUX_MET_12 data). The following parameterization is obtained from literature, e.g., [RD17]

$$\beta_m(z) = 10^{-7}\left(\frac{\lambda_0}{\lambda}\right)^{4.09} e^{-z/z_0}$$

$$\alpha_m(z) = \frac{8\pi}{3}\beta_m(z)$$

Eq. D-0-1

with β_m molecular backscatter ($m^{-1}sr^{-1}$), α_m molecular extinction (m^{-1}), z altitude (m), z_0 the atmosphere density scale height with a value of 8000 m, λ the Aeolus laser beam wavelength (m) and λ_0 a reference wavelength of $1.06 \cdot 10^{-6}$ m. The relation between molecular backscatter and extinction follows from Rayleigh scattering theory.

*Option 2: AUX_MET data based parameterization (**the recommended option**)*

Additional prior knowledge of the atmosphere is available from auxiliary NWP data (AUX_MET_12) and is used for the retrieval of atmosphere optical properties, in particular profiles of pressure and temperature from NWP models. Note that a priori wind information is not used in the L2B processing. Molecular backscatter and extinction may be determined from Rayleigh scattering theory given profiles of atmospheric pressure and temperature between the earth's surface and satellite altitude as follows [RD15].

Estimates of the atmosphere temperature and pressure at measurement location are available from the AUX_MET_12 input files, i.e., from short-range forecasts of an NWP model. Molecular backscatter and extinction coefficients may be determined from Rayleigh theory given profiles of atmospheric pressure and temperature:

$$n(z) = p(z)/kT(z)$$

$$m_s = 1 + 10^{-8}\left[5791817/\left(238.0185 - \left(10^6\lambda\right)^{-2}\right) + 167909/\left(57.362 - \left(10^6\lambda\right)^{-2}\right)\right]$$

$$\sigma_{\mathrm{mol}}(z) = \left(24\pi^3/n_s^2\left(10^2\lambda\right)^4\right)\left[\left(m_s^2-1\right)^2/\left(m_s^2+2\right)^2\right]\left[(6+3\rho)/(6-7\rho)\right]$$

$$\alpha_m(z) = \sigma_{\mathrm{mol}}(z)n(z)$$

$$\beta_m(z) = \frac{3}{8\pi}\alpha_m(z)$$

Eq. D-0-2

The top equation is the ideal gas law. Furthermore:

k=1.3806503·10^{-23}	: [J/K]	; Boltzmann constant
n_s= 2.54743·10^{-19}	: [m^{-3}]	; molecular number density at standard p and T
p	: [Pa]	; pressure
T	: [K]	; temperature
n	: [m^{-3}]	; molecular number density
λ=355·10^{-9}	: [m]	; laser wavelength
m_s=1.00028569773896	: [-]	; refractive index of air for 355 nm laser wavelength
ρ=0.03178	: [-]	; yielding the depolarization constant
σ_{mol}	: [m^2]	; molecular scattering cross-section per molecule

Both options described in sections have been implemented in the L2BP OPC. Examples are found in [RD17].

Appendix E. Derivation of the sensitivity of HLOS wind to Mie contamination

Eq. 15-14 is needed for the calculation of the sensitivity of HLOS to the scattering ratio, $\frac{\partial f(\rho_0)}{\partial \rho}$, with *f* the Doppler shifted frequency and ρ the scattering ratio, see Eq. 15-13. The solution can be filled in the equation below, which described the dependency of the Doppler shifted frequency upon the scattering ratio as a Taylor expansion of the second order around ρ_0. Note that the second order contribution is ignored so far.

$$f(\rho = 1) = f(\rho_0) + \frac{\partial f(\rho_0)}{\partial \rho}(1 - \rho_0) + \frac{1}{2}\frac{\partial^2 f(\rho_0)}{\partial \rho^2}(1 - \rho_0)^2 + \cdots$$

The sum of the Rayleigh and Mie signal returning from the atmosphere can be written simplified as

$$I(f) = \beta_m Y I_m(f) + \beta_p Y I_p(f)$$

i.e. as the sum of the broadband molecular spectrum, I_m, and the small band Mie spectrum, I_p multiplied by the molecular backscatter, β_m, and particle backscatter, β_p, respectively. Y denotes all other parameters, determining the strength of the atmospheric signal such as laser energy, telescope diameter, atmospheric range, etc., all independent of the laser frequency *f*. Noting that β_p / β_m equals (ρ_0-1), the above equation can be written as

$$I(f) = \beta_m Y I_m(f) + \beta_p Y I_p(f)$$
$$= \left(I_m(f) + \frac{\beta_p}{\beta_m} I_p(f) \right) \beta_m Y = \left(I_m(f) + (\rho_0 - 1) I_p(f) \right) \beta_m Y$$

The atmospheric return signal measured on the Rayleigh detector is obtained by convolving the atmospheric return signal with the channel A and B transmission functions. For simplicity, we ignore the signal strength, $\beta_m Y$, which is frequency independent. Then the signal measured by Rayleigh channel A and B can be written as:

$$A, B = \int_{-\infty}^{\infty} T_{A,B}(f) I(f, f_d) df = \int_{-\infty}^{\infty} T_{A,B}(f)(I_m(f, f_d) + I_p(f, f_d)) df \approx$$
$$\int_{-\infty}^{\infty} T_{A,B}(f)(I_m(f, f_d) + (\rho - 1)\delta(f, f_d)) df = \int_{-\infty}^{\infty} T_{A,B}(f) I_m(f, f_d) df + (\rho - 1) T_{A,B} = A_m, B_m + A_p, B_p$$

with $T_{A,B}$ the transmission curves of channels A and B respectively, *I* the total spectrum of the measured atmospheric wind (in frequency space), which can be written as the sum of the broadband molecular spectrum, I_m, and the small band Mie spectrum, I_p. f_d Is the Doppler shifted frequency, i.e., the centre of both spectra. The small band Mie peak is assumed a delta function, δ.

Then we can write for the frequency response, *R*

$$R = \frac{A-B}{A+B} = \frac{A_m + (\rho-1)T_A - (B_m + (\rho-1)T_B)}{A_m + (\rho-1)T_A + B_m + (\rho-1)T_B}$$

and for the sensitivity of *R* relative to *f*, after some calculation:

$$\frac{\partial R}{\partial \rho} = \frac{2(T_A B_m - T_B A_m)}{(A_m + B_m + (\rho-1)(T_A + T_B))^2}$$

Noting from the above that $\mathrm{A}_m = \mathrm{A} - (\rho_0 - 1)T_A$ and $\mathrm{B}_m = \mathrm{B} - (\rho_0 - 1)T_B$, i.e., the molecular contribution to the signal is the total contribution minus the contribution from particles, and substituting this into the above equation gives

$$\frac{\partial R}{\partial \rho} = \frac{2(T_A B_m - T_B A_m)}{(A_m + B_m + (\rho - 1)(T_A + T_B))^2} = \frac{2T_A(B - (\rho_0 - 1)T_B) - 2T_B(A - (\rho_0 - 1)T_A)}{(A + B)^2}$$
$$= \frac{2BT_A - 2AT_B}{(A + B)^2} = \frac{(T_A - T_B) - R(T_A + T_B)}{A + B}$$

where the last equal sign follows from bringing $(A + B)$ out of brackets in the numerator. This equation equals Eq. 15-14.

Appendix F. Overview of AUX_PAR_2B settings and the control of the L2B processor algorithms

The AUX_PAR_2B parameters which are relevant for the L2B processing is shown in Table 8, along with brief comments or references related to the use of the parameter. [RD10] provides a complete list of the options.

Not used and obsolete

Not used but may be in the future

Table 8. List of AUX_PAR_2B parameters relevant to the L2B processing algorithms

Top level	Second level	Third level	Fourth level	Recommended option	Description/Status
WVM_ Params					
	Rangebin_Mismatch_Tolerance				Not used in L2Bp
	Line_of_Sight_Wind_Flag			False i.e. HLOS winds, common choice for NWP purposes	Determines whether the wind results are LOS or HLOS winds. See sections 13 and 14
	N_Obs_Mie_Max				Obsolete
	N_Obs_Rayleigh_Max				Obsolete
	BRC_Grouping_Params				See section 7
		Grouping_Method			
		Max_Vertical_Rangebin_Misalignment_Mie			
		Max_Vertical_Rangebin_Misalignment_Rayleigh			
		Max_Horizontal_Accumulation_Length_Mie			
		Max_Horizontal_Accumulation_Length_Rayleigh			
		Max_Allowed_Gap_Between_Mie_Measurements			
		Max_Allowed_Gap_Between_Rayleigh_Measurements			
		num_BRCs_to_merge			
	Classification_Params				
		Classification_Type_Mie		Class_Backscat_Ratio	See section 8.2. Or see section 8.3 if interested in using OPC feature finder
		Classification_Type_Rayleigh		Class_Backscat_Ratio	See section 8.2. Or see section 8.3 if interested in using OPC feature finder
		List_of_Mie_Backscatter Ratio_Thresholds			A flexible number of Threshold_Value

Top level	Second level	Third level	Fourth level	Recommended option	Description/Status
					and associated Altitude. Can have different thresholds for the Mie and Rayleigh.
		List_of_Rayleigh_BackscatterRatio_Thresholds			Similar to that of Mie above.
		List_of_Mie_Extinction_Thresholds			Similar to above, but for extinction values if using the OPC derived extinction.
		List_of_Rayleigh_Extinction_Thresholds			See above, but for Rayleigh.
	Optical_Properties_Params				
		ScatRatio_Method		Scat_Ratio_from_L1B_Mie_refined	Method for getting scattering ratio. See section 8.2
		ScatRatio_Method2		Scat_Ratio_One_If_No_Mie	Backup in case above gives no result. See section 8.2
		PartExt_Method		PartExt_Method_Iterative; if want extinction derived from OPC	Method to retrieve particle extinction; which can be used in classification. See section 8.3
		MolExt_Method		MolExt_Method_Proper; the method that uses the AUX_MET data	Method to get molecular extinction. See [Appendix D. Retrieval of atmosphere molecular backscatter and extinction (Rayleigh scattering)]
		Minimum_Altitude_for_Assuming_Rho_1			Lowest altitude above which it is OK to set the scattering ratio to 1.0 (i.e. clear air); if using Scat_Ratio_One_If_No_Mie
		k_power			OPC related setting. See section 19
		FP_Xtalk_factor			OPC related setting. See section 19
		Median_Filter_Window_Width_x			OPC related setting. See section 19
		Median_Filter_Window_Width_y			OPC related setting. See section 19
		Calibration_from_AuxCal			OPC related setting. See section 19
		Cross_Talk_from_AuxCal			OPC related setting

Top level	Second level	Third level	Fourth level	Recommended option	Description/Status
					See section 19
		Apply_Median_Filter			OPC related setting. See section 19
		Apply_2D_Feature_Finder			OPC related setting. See section 19
		FP_On_Upper_Bin_Std_Threshold			OPC related setting. See section 19
	Error_Quantifier_Params				See description of choices in sections 17 and 18
		ErrorQuantMethod_Mie			
		ErrorQuantMethod_Rayleigh			
	Common_Processing_Params				
		Mie_PCD_Params			Not used
			Alpha_Correction		
			Summation_Index		
		Mie_Core_Algorithm_Params{,Reference_Pulse}			Settings for the Mie core algorithm for atmospheric return and the reference pulse can be defined
			SNR_Threshold		
			Start_FWHM		
			Residual_Error_Threshold		
			Max_Iterations_Lorentz_Fit		
			FWHM_Upper_Threshold		
			FWHM_Lower_Threshold		
			Peak_Height_Upper_Threshold		
			Peak_Height_Lower_Threshold		
			Peak_Location_Threshold		
			Nonlinear_Optimization_Threshold		
			Max_Iterations_Nonlinear_Optimization		
			Num_Spectral_Sub_Samples		
		Corrupt_Data_Detection_Params			
			Max_Signal_Derivative		
			Pixel_Saturation_Threshold		
	Mie_Algorithm_Params				
		Copy_L1B_Mie_Core_Algorithm_Params		True	Switch to import L1B Mie Core algorithm Downhill simplex settings. It

Top level	Second level	Third level	Fourth level	Recommended option	Description/Status
					is usually recommended to set to true
		Copy_MieCoreAlgParams_to_IntRef		true	Switch to be able to use the atmospheric Mie core settings for the internal reference channel
		Mie_Height_Assignment_Method		use_fixed_weight	Currently only one option
		Mie_Height_Weight_Upper		0.5	Since the distribution of attenuated back-scatter within the bin is unknown, then best to assign half-way through bin.
		Skip_Mie_Non_Linearity_Correction		false	The non-linearity of the Mie response should be considered i.e. not skipped
		Extrapolate_Mie_Calibration_Data		false	Switch as to whether the non-linear response can be extrapolated
		Use_Ref_Pulse_Zero_Freq		true	Switch to use the zero frequency value of the internal calibration
		Use_Meas_Zero_Freq		true	Switch to use the zero frequency of the atmospheric calibration
		Offset_Subtraction_Col20_Weight		0.5	How much to weight column 20 versus column 19 in the offset subtraction
		List_of_SNR_Thresholds		TBD with real Aeolus data	List made up of SNR thresholds for given altitude ranges. For QC
		Flag_Clear_Mie_Results_Invalid		false	Option to flag all Mie-clear winds as invalid
	RBC_Algorithm_Params				
		Rayleigh_Height_Assignment_Method		use_fixed_weight	Choose which method for Rayleigh height assignment
		Rayleigh_Height_Weight_Upper		0.49	A good compromise value to account for typical attenuated backscatter profile in a bin.
		Do_Mie_Decontamination		true	Switch to control if the effect of Mie signal in the

Top level	Second level	Third level	Fourth level	Recommended option	Description/Status
					Rayleigh channel is corrected for
		Flag_Cloudy_Rayleigh_Results_Invalid		false	Option to flag all Rayleigh-cloudy winds as invalid
		LOS_min			This and the following 11 parameters are used to detect problems occurring in the RBC algorithm. Exact figures are to TBD with real Aeolus data.
		LOS_max			
		LOS_ref_min			
		LOS_ref_max			
		dLOSdR_min			
		dLOSdR_max			
		dLOSdT_min			
		dLOSdT_max			
		dLOSdP_min			
		dLOSdP_max			
		dLOSdrho_min			
		dLOSdrho_max			
	AMD_Matchup_Params				See section 20
		Matchup_Method		Nearest_Neighbour	Choose method to select AMD data to be used in the HLOS wind retrieval
		Max_Allowed_Time_Diff		TBD	Parameter used when searching for matching AMD profiles
		Max_Allowed_Distance		TBD	Parameter used when searching for matching AMD profiles
		Max_Analysis_Time_Diff		0 (does not check if zero)	Can restrict the HLOS retrieval from using AMD data forecasts older than a defined time period
	CLM_Matchup_Params				
		Matchup_Method			
	ZWC_Params				Zero Wind Correction
		ZWC_Scheme_Mie		ZWC_scheme_use_L2B_settings	Method for doing the ZWC; either copy the total L1B correction, or recalculate using L2B weightings (see below). See section 14.6.
		ZWC_Scheme_Rayleigh		ZWC_scheme_use_L	Method for doing the ZWC; either

Top level	Second level	Third level	Fourth level	Recommended option	Description/Status
				2B_settings	copy the total L1B correction, or recalculate using L2B weightings (see below). See section 13.5.
		Mie_Ground_Correction_Weighting		TBD	
		Mie_HBE_Ground_Correction_Weighting		TBD	
		Rayleigh_Ground_Correction_Weighting		TBD	
		Rayleigh_HBE_Ground_Correction_Weighting		TBD	
		Mie_Rayleigh_Ground_Correction_Weighting		TBD	
		Mie_Rayleigh_Ground_Correction_Offset		TBD	
	RDB_Params				Normally RDB correction should be applied
		Do_Mie_RDB_corr		true	Switch for RDB correction for the Mie channel
		Do_Rayleigh_RDB_corr		true	Switch for RDB correction for the Rayleigh channel
Screening_Params					
	L1B_Screening_Params				
		L1B_Geolocation_Screening_Params			
			Latitude_Min		
			Latitude_Max		
			Longitude_Min		
			Longitude_Max		
			Altitude_Min		
			Altitude_Max		
			Altitude_DEM_Min		
			Altitude_DEM_Max		
			Geoid_Separation_Min		
			Geoid_Separation_Max		
			Topocentric_Elevation_Min		
			Topocentric_Elevation_Max		
			Topocentric_Azimuth_Min		
			Topocentric_Azimuth_Max		
			AOCS_LOS_Velocity_Min		
			AOCS_LOS_Velocity_Max		

Top level	Second level	Third level	Fourth level	Recommended option	Description/Status
		L1B_Obs_Screening_Params			See section 6.1. Values to be determined with real Aeolus data.
			L1B_Laser_Freq_Unlocked_Threshold	TBD	
			L1B_Ref_Pulses_Unlocked_Threshold	TBD	
			L1B_Laser_Freq_Offset_Threshold	TBD	
			L1B_Laser_UV_Energy_Threshold	TBD	
			L1B_Laser_Freq_Offs_Stdev_Threshold	TBD	
			L1B_Laser_UV_Energy_Stdev_Threshold	TBD	
			L1B_Mie_Mean_Emit_Freq_Min	TBD	
			L1B_Mie_Mean_Emit_Freq_Max	TBD	
			L1B_Mie_Emit_Freq_Stdev_Threshold	TBD	
			L1B_Rayleigh_Mean_Emit_Freq_Min	TBD	
			L1B_Rayleigh_Mean_Emit_Freq_Max	TBD	
			L1B_Rayleigh_Emit_Freq_Stdev_Threshold	TBD	
			L1B_Sat_Not_on_Target_Threshold	TBD	
			L1B_Mie_Corrupt_Threshold	TBD	
			L1B_Rayleigh_Corrupt_Threshold	TBD	
			L1B_Mie_Ref_Pulses_Corrupt_Threshold	TBD	
			L1B_Rayl_Ref_Pulses_Corrupt_Threshold	TBD	
			L1B_Mie_Invalid_Meas_Threshold	TBD	
			L1B_Mie_Invalid_Ref_Pulses_Threshold	TBD	
			L1B_Rayl_Invalid_Meas_Threshold	TBD	
			L1B_Rayl_Invalid_Ref_Pulses_Threshold	TBD	
		L1B_Mie_Meas_Screening_Params			
			L1B_Mie_Meas_I	TBD	

Top level	Second level	Third level	Fourth level	Recommended option	Description/Status
			nvalid_Ref_Pulses_Threshold		
			L1B_Avg_Laser_Freq_Offset_Min	TBD	
			L1B_Avg_Laser_Freq_Offset_Max	TBD	
			L1B_Avg_UV_Energy_Min	TBD	
			L1B_Avg_UV_Energy_Max	TBD	
			L1B_Laser_Freq_Offset_Stdev_Threshold	TBD	
			L1B_UV_Energy_Stdev_Threshold	TBD	
			L1B_Vel_of_Att_Uncertainty_Error_Min	TBD	
			L1B_Vel_of_Att_Uncertainty_Error_Max	TBD	
			L1B_Mie_Mean_Emitted_Freq_Min	TBD	
			L1B_Mie_Mean_Emitted_Freq_Max	TBD	
			L1B_Mie_Emitted_Freq_Stdev_Threshold	TBD	
			L1B_Meas_Reference_Pulse_FWHM_Min	TBD	
			L1B_Meas_Reference_Pulse_FWHM_Max	TBD	
			L1B_Scattering_Ratio_Min	TBD	
			L1B_Scattering_Ratio_Max	TBD	
			L1B_Mie_SNR_Threshold	TBD	
			Mie_Ground_Bin_Thickness_Threshold		
			Max_Signal_Derivative		
			Pixel_Saturation_Threshold		
			Ignore_Mie_Meas_Invalid_Switch	true	An important switch. To avoid using the L1Bp definitions of an invalid measurement range-bin
		L1B_Rayleigh_Meas_Screening_Params			
			L1B_Rayleigh_Meas_Invalid_Ref_Pulses_Threshold	TBD	
			L1B_Avg_Laser_Freq_Offset_Min	TBD	

Top level	Second level	Third level	Fourth level	Recommended option	Description/Status
			L1B_Avg_Laser_Freq_Offset_Max	TBD	
			L1B_Avg_UV_Energy_Min	TBD	
			L1B_Avg_UV_Energy_Max	TBD	
			L1B_Laser_Freq_Offset_Stdev_Threshold	TBD	
			L1B_UV_Energy_Stdev_Threshold	TBD	
			L1B_Vel_of_Att_Uncertainty_Error_Min	TBD	
			L1B_Vel_of_Att_Uncertainty_Error_Max	TBD	
			L1B_Rayleigh_Mean_Emitted_Freq_Min	TBD	
			L1B_Rayleigh_Mean_Emitted_Freq_Max	TBD	
			L1B_Rayleigh_Emitted_Freq_Stdev_Threshold	TBD	
			L1B_Rayleigh_SNR_Min	TBD	
			L1B_Rayleigh_SNR_Max	TBD	
			Rayleigh_Ground_Bin_Thickness_Threshold		
			Pixel_Saturation_Threshold		
			Ignore_Rayleigh_Meas_Invalid_Switch	true	An important switch. To avoid using the L1Bp definitions of an invalid measurement range-bin
	L1B_Cal_Char_Data_Screening_Params				None of the parameters under this tag are used in the L2Bp. **N.B. have shifted one column to the left to fit it in.**
		Sat_Char_Data_Screening_Params			
			Tripod_Obscur_Corr_Min		
			Tripod_Obscur_Corr_Max		
		Mie_Resp_Calib_Data_Screening_Params			
			Meas_Resp_Min		
			Meas_Resp_Max		
			Meas_Err_Mie_Resp_Min		

Top level	Second level	Third level	Fourth level	Recommended option	Description/Status
			Meas_Err_Mie_Resp_Max		
			Ref_Pulse_Resp_Min		
			Ref_Pulse_Resp_Max		
			Ref_Pulse_Err_Mie_Resp_Min		
			Ref_Pulse_Err_Mie_Resp_Max		
			Ref_Pulse_Zero_Freq_Min		
			Ref_Pulse_Zero_Freq_Max		
			Ref_Pulse_Mean_Sensitivity_Min		
			Ref_Pulse_Mean_Sensitivity_Max		
			Meas_Zero_Freq_Min		
			Meas_Zero_Freq_Max		
			Meas_Mean_Sensitivity_Min		
			Meas_Mean_Sensitivity_Max		
		L1B_GWD_ADS_Screening_Params			
			Zero_Wind_Corr_Min		
			Zero_Wind_Corr_Max		
			Zero_Wind_Corr_Weight_Factor_Min		
			Zero_Wind_Corr_Weight_Factor_Max		
			Mie_Rayl_Correction_Offset_Min		
			Mie_Rayl_Correction_Offset_Max		
	L2B_AMD_Screening_Params				See section 6.2. Some reasonable values suggested below.
		L2B_AMD_p_min		1	Pa
		L2B_AMD_p_max		1100*1E2	Pa
		L2B_AMD_T_min		150	K
		L2B_AMD_T_max		400	K