

ADM-Aeolus, VAMP

Vertical Aeolus Measurement Positioning
Technical note TN3

Name code: AE-TN-KNMI-VAMP-003

Authors: **Jos de Kloe**, Gert-Jan Marseille, Ad Stoffelen, KNMI



Koninklijk Nederlands
Meteorologisch Instituut
Ministerie van Verkeer en Waterstaat

issue number: 0.3
issue date: 21-Apr-2009

Change log

Version	Date	Comment
0.1	10-Apr-2009	A first draft.
0.2	17-Apr-2009	Adapted to internal KNMI comments. Added new KNMI logo.
0.3	21-Apr-2009	Included chapters on Task 4b and LIPAS written by G.J.Marseille

Contents

Contents	3
List of Tables	4
List of Figures	6
1 General introduction	7
2 Documents and acronyms	8
2.1 Applicable documents	8
2.2 Reference documents	8
2.3 Literature	8
2.4 Acronyms	9
2.5 Document preparation	9
3 Task 4a: Mie Core QC	10
3.1 Introduction	10
3.2 Setup of the tool	10
3.3 Simulation of transmission and spectrum	12
3.4 Mie Core Algorithm	14
3.5 Input data	16
3.6 Error multiplier	16
3.7 Horizontal gradients and turbulence	17
3.8 Results: LITE data; signal scaling	18
3.9 Results: CALIPSO data; data selection	21
3.10 Results: CALIPSO data; Mie Core outputs	21
3.10.1 Outer loop iterations	22
3.10.2 Residual error	23
3.10.3 FWHM	24
3.10.4 Altitude effects	28
3.11 Task 4a: QC conclusions and recommendations	28
4 Task 4a: Rayleigh QC	30
5 Task 4b: LIPAS upgrade	31
5.1 Background	31
5.2 From LIPAS to LIPAS-2009	31
5.2.1 Atmospheric return signal	32
5.2.2 Fizeau interferometer	33
5.2.3 Fabry-Perot transfer function	36
5.2.4 Results	40
6 Task 4b: Quantification of the L2 HLOS information content for Mie only, Rayleigh only and collocated Mie and Rayleigh sampling	47
6.1 Assessment of vertical sampling scenarios	47
7 Task 4c: Sampling scenarios	48
8 Conclusion	49

List of Tables

1	Parameter values used to simulate the Mie spectral peak, and the Fizeau transmission curves.	12
2	Mie Core algorithm settings.	15
3	Overview of effects of applying the different thresholds	27
4	Overview of effects of applying the different thresholds on different regions of the atmosphere.	29
5	Parameter values used to simulate the Fizeau transmission function ,see also Table 1.	35
6	Parameter values used to simulate the Fabry-Perot transfer function.	37
7	Parameter values used for the ADM performance simulation.	47

List of Figures

1	LOS Wind shear plotted as function of the wind speed error, for the CALIPSO data discussed in section 3.9. Scatterplot (left panel) and contour plot (right panel). The data has been selected on $SNR > 100$, which means this result includes mainly reflections on optically thick clouds. The error in wind speed is therefore mainly caused by height assignment errors.	10
2	Mie spectral shape for a wind speed of 50 m/s (blue), compared to the much broader Fizeau transmission curve for accd column 3 calculated in 3 different ways (red: using the Lorentzian function as is done in LIPAS; cyan: using the Airy function as in done by the E2S; gray: using the Airy function but without integration over the accd bin surface). The 2 vertical green lines denote the edges of the accd bin used for this example, and the vertical blue line denotes the frequency that corresponds to the Doppler shift of a LOS wind of 50 m/s. Finally, the small blue peak below the Mie spectral shape, is the amount of light entering accd bin 3.	13
3	Mie spectral shape for a wind speed of 50 m/s (blue), compared to the the Fizeau transmission curves for all 16 accd columns (cyan). The blue curves below the Mie peak denote the fraction of light detected by each of the accd bins. Clearly, due to the long tails of the transmission curves, all accd bins will receive some signal. The expected signal that will be measured by each accd channel is denoted by the red squares.	15
4	Input LOS winds derived from ECMWF u,v,w components, colocated at the locations of the CALIPSO data used for the majority of simulations, plotted against the wind speed error found by the simulations (for results with $SNR > 10$ and ValidityFlag set).	17
5	First simulation results using LITE data. The data shows nonsense retrievals for very low SNR and a clear overflow problem.	18
6	Simulation results using LITE data, after applying a SNR threshold of 10 (left panel), and after discarding data reported invalid by the ValidityFlag as well (right panel). The nonsense retrievals for very low SNR have disappeared now but the overflow problem remains.	19
7	A case showing the effect of an integer overflow in the accd counts of the Mie channel (these are internal reference counts, and this case was generated by LIB v1.07).	19
8	Simulation results using LITE data. In this case a correct signal scaling is applied and the overflow problem disappears (left panel). Note that the MieCore algorithm is applied at measurement level in this case. To illustrate the effect of Poisson noise on the simulations the same case has been run without Poisson noise (right panel).	20
9	Simulation results using LITE data, without wind variability and without Poisson noise, to illustrate the numerical inaccuracies of the simulation program.	20
10	Simulation results using CALIPSO data, all results without selection (left panel) and after applying a SNR threshold of 10 (right panel).	21
11	Simulation results using CALIPSO data, after applying a SNR threshold of 10 and after discarding data reported invalid by the ValidityFlag. as well (right panel).	22
12	All valid data after discarding points with $SNR < 10$ and invalid ValidityFlag (left panel). The right panel shows a magnification of the region $10 < SNR < 100$.	22
13	Number of outer loop iterations, as function of the wind speed error, for the selected CALIPSO data.	23
14	Residual error reported by the Mie Core algorithm as function of the wind speed error, for the selected CALIPSO data.	24

15	Residual error reported by the Mie Core algorithm as function of the wind speed error, plotted as contour plot. All valid results (left panel) and the region below ResErr=0.01 magnified.	24
16	Same data as shown in Fig. 15, again as scatterplot, but now with over plotted lines to show the mean and one std value on both sides of the mean, for a number of Residual Error layers (full range is split in 25 sub-ranges for this calculation). From the right panel the position at which the wind speed error std exceeds the value of 1 m/s (requirement for the LOS wind speed accuracy) can be read, which is at about 0.005.	25
17	FWHM reported by the Mie Core algorithm as function of the windspeed error, for the selected CALIPSO data.	25
18	FWHM reported by the Mie Core algorithm as function of the windspeed error, plotted as contourplot. All valid results (left panel), and a magnified region around FWHM=2 (right panel). The straight lines in the left panel are plotting artefacts and should be ignored. Also the striping feature in the right panel is an artefact caused by the binning of the data and probably not relevant.	26
19	Same data as shown in Fig. 18 again as scatterplot, but now with overplotted lines to show the mean and one std value on both sides of the mean, for a number of FWHM layers (full range is split in 25 sub-ranges for this calculation).	26
20	Scatterplot of FWHM against windspeed error for cases with ResErr<0.005. The overplotted lines show the mean and one std value on both sides of the mean, for a number of FWHM layers (full range is split in 25 sub-ranges for this calculation).	27
21	FWHM reported by the Mie Core algorithm as function of the wind shear in the input wind. The left plot selects data with SNR between 10 and 50, the right plot selects data with SNR above 50.	28
22	Distribution of atmosphere Mie peak on Fizeau interferometer ACCD bins for a 50 ms ⁻¹ wind velocity.	34
23	Fabry-Perot channel A (direct, blue) and B (reflected, red) transmission (left panel) and Fizeau reflection, $\tau_{Fiz}(\nu)$, (right panel). The blue solid and red dotted curves show the channel A, $\tau_{FP}^A(\nu)$, and B, $\tau_{FP}^B(\nu)$, transmission respectively as a function of wavelength offset. The blue dotted line is the reflected signal from the direct channel, $\tau_{FP}^{A-refl}(\nu)$. The solid red line is the product of the channel A reflected transmission and the channel B transmission.	38
24	Fabry-Perot Relay to focus the channel A, $\tau_{FP}^{A-relay}$ and B, $\tau_{FP}^{B-relay}$, signals on the ACCD bins (left panel). The right panel shows the relative signal detected from the A (direct) and B (reflected) channel.	38
25	Receiver response function for $T = 273$ (K) as a function of wind velocity in ms ⁻¹	40
26	Case 1. Signal on the 16 ACCD bins of the Fizeau interferometer for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. The blue curve in the lower right panel is the E2S earth radiance background signal. See the text for further details.	41
27	Case 1. Signal on the 16 ACCD bins of the Fabry-Perot interferometer for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. See the text for further details.	42
28	Case 2. Signal on the 16 ACCD bins of the Fizeau interferometer for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. The blue curve in the lower right panel is the E2S earth radiance background signal. See the text for further details. Bins 8-10 clearly show an increased signal due to cross-talk from cloud particles.	43
29	Case 2. Signal on the 16 ACCD bins of the Fabry-Perot interferometer for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. See the text for further details.	44
30	Atmospheric pressure (left) and temperature (right)	44
31	Atmospheric wind (left) and molecular backscatter (right). There is no aerosol backscatter.	45
32	Single case HLOS wind retrieval from Rayleigh (blue diamond) and Mie (red star) channel. The left panel shows the result for a noise-free instrument, the right panel does include instrument noise.	45
33	HLOS wind retrieval statistics for Rayleigh (blue diamond) and Mie (red star) channel from 100000 simulations. The left panel shows the bias, the right panel the standard deviation. The standard deviation for Mie winds above the surface is close to 85 ms ⁻¹ (not shown)	45
34	Atmospheric temperature (left) and aerosol (right). The left curve/middle/right shows aerosol/molecular/cloud backscatter.	46

35 Mie (red) and Rayleigh (blue) HLOS wind retrieval statistics from 100000 simulations. The left/middle (zoomed in) panel show the bias, the right panel the standard deviation. 46

1 General introduction

This document describes the work done and the results obtained for Work Package (WP) 4 of the Vertical Aeolus Measurement Positioning (VAMP) project. The objective of this WP is to investigate the quality of the Level-2B (L2B) Horizontal projected Line-of-Sight (HLOS) wind results, in case of separate processing of the Mie and Rayleigh channels, and in case they can be combined. Especially the benefit of Mie oversampling (i.e. having multiple Mie range bins covering a single Rayleigh range bin) will be investigated in conditions with large gradients in for example wind shear (i.e. vertical variation of the horizontal wind components) and optical properties.

This WP and thus this document has been divided in three parts. First, the utilization of Mie Core and Rayleigh processing outputs as quality indicators is described in sections 3 and 4. Then in section 5 the LIPAS tool is described in some detail, including the modifications added to it to make it suitable for use in the current project. In section 6 the L2B HLOS information content for the separate and combined Mie and Rayleigh channels is quantified. A summary of the benefits and drawbacks of the proposed sampling scenarios is discussed in section 7. Finally this technical note is concluded with an overview of the findings in section 8.

NOTE: on 17th of April a bug was discovered in the simulation code, causing almost all input LOS winds to positive. See Fig. 4 in section 3.5 for a plot illustrating this. This affects the statistical results and causes small systematic errors (biases) in the wind speed result in the order of 0.1 to 0.2 m/s. Because re-running all simulations will take about 5 days, we decided to deliver this draft document anyway, to give you the opportunity to review it and plan to present corrected results at the next progress meeting.

2 Documents and acronyms

2.1 Applicable documents

- [AD1] “ADM-Aeolus: Optimization of the Spatial and Temporal Sampling”, by: A. G. Straume-Lindner, issue 1.0, dated 20-Apr-2007, reference code: AE-SW-ESA-GS-027.

2.2 Reference documents

- [RD1] “Aeolus Level 2B processor Issue 1.40 Software Release Note (SRN)”, 28-Feb-2009, AE-RN-ECMWF-L2BP-001.
- [RD2] “ADM-Aeolus L1B Master Algorithm Document (MAD)”, AE-SW-ASU-GS-023, issue 4, dated 26-Jun-2006.
- [RD3] “Aeolus Level 1b Processor and End-to-End Simulator, End-to-End Simulator Detailed Processing Model (DPM)”, by: P. Saeedi, ADM-MA-52-1801-E2S-DPM, version 2.4, issued 12 april 2007.
- [RD4] “ALADIN - Incoherent DWL Performance model description”, TN written by F. Fabre of Matra Marconi Space, reference code: ADM/MMS/TN007.98, issue 01.00, dated 26-Nov-1998.
- [RD5] “Additional Computational Steps ADM-Aeolus L1B”, by Dorit Huber and Ines Leike, issue 1.4, dated 10-Mar-2006, reference code: AE.TN.DLR.ACS-L1B.100306.
- [RD6] “ADM-Aeolus VAMP, Vertical Aeolus Measurement Positioning, Technical note TN1”, by: J. de Kloe, G.J. Marseille, and A. Stoffelen, issue 1.0, dated 28-Mar-2007, reference code: AE-TN-KNMI-VAMP-001.
- [RD7] “Assessment of Optical and Dynamical Atmospheric Heterogeneity” (Technical note TN2), by: A. Stoffelen, H. Körnich, G.J. Marseille, K. Houchi, J. de Kloe, issue 0.5, dated 25-Mar-2009, reference code: AE-TN-KNMI-VAMP-002.
- [RD8] “The ADM Atmospheric database, Technical note TN2.4 on WP1400 (L1B study task 2 output)” by: J. de Kloe, issue 1.5, dated 25-Jun-2008, reference code: AE-TN-KNMI-L1B-001.
- [RD9] “TN5.1, Comparison of Aeolus burst and continuous mode concepts”, by: A. Stoffelen, G.J. Marseille, J. de Kloe, A. Dabas, D. Huber, O. Reitebuch, P. Flamant, D. Tan, dated 07-Oct-2008, reference code: AE-TN-KNMI-GS-051.
- [RD10] Merci executive summary, Measurement Error and Correlation impact on the Atmospheric Dynamics Mission, by: Ad Stoffelen, Pierre Flamant, Måns Håkansson, Erland Källén, Gert-Jan Marseille, Jean Pailleux, Harald Schyberg, Michael Vaughan.

2.3 Literature

- [LR1] Tan, D., Andersson, E., “Simulation of the yield and accuracy of wind profile measurements from the Atmospheric Dynamics Mission (ADM-Aeolus)”, *Q. J. R. Meteorol. Soc.* (2005), 131, pp. 1737-1757.
- [LR2] G.J. Marseille and A. Stoffelen, “Simulation of Wind Profiles from a Space-borne Doppler Wind Lidar”, *Q. J. R. Meteorol. Soc.* (2003), 129, pp. 3079-3098.
- [LR3] Stoffelen, A., Marseille, G.J., Bouttier, F., Vasiljevic, D., de Haan, S. and Cardinali, C., “ADM-Aeolus Doppler wind lidar Observing System Simulation Experiment”, *Q. J. R. Meteorol. Soc.* (2006), 132, pp 1927-1947.
- [LR4] Tan, D.G.H., Andersson, E., Fisher, M. and Isaksen, L. “Observing system impact assessment using a data assimilation ensemble technique: Application to the ADM-Aeolus wind profiling mission”, *Q. J. R. Meteorol. Soc.* (2006), 133, pp 381-390.

- [LR5] Marseille, G.J., Stoffelen, A., “Barkmeijer J., Impact Assessment of Prospective Space-Borne Doppler Wind Lidar Observation Scenarios”, *Tellus A*, 60 (2), 2008., pp. 234–248, doi: 10.1111/j.1600-0870.2007.00289.x.

2.4 Acronyms

ACCD	Accumulation Charge Coupled Device
AD	Applicable Document (see section 2.1)
ADM	Atmospheric Dynamics Mission
ANX	Ascending Node Crossing
BRC	Basic Repeat Cycle (covering a 200 km orbit section)
DEM	Digital Elevation Model
DPM	Detailed Processing Model
E2S	End-to-End Simulator
ESA	European Space Agency
FP	Fabry-Perot (spectrometer)
FSR	Free Spectral Range
Fz	Fizeau (spectrometer)
HLOS	Horizontal projection of the Line-Of-Sight (component of the wind)
IRC	Instrument Reponse Calibration
ISR	Instrument Spectral Registration
KNMI	Koninklijk Nederlands Meteorologisch Instituut (Royal Dutch Meteorological Institute)
L1Bp/L2Ap/L2Bp	Level 1B/2A/2B processor
LIPAS	Lidar Performance Analysis Simulator
LOS	Line-Of-Sight
LR	Literature Reference (see section 2.3)
LTAN	Local Time Ascending Node
MMS	Matra Marconi Systems
MRC	Mie Response Calibration
NWP	Numerical Weather Prediction
OSSE	Observation System Simulation Experiment
SOSE	Sensitivity Observing System Experiment
PBL	Planetary Boundary Layer
RD	Reference Document (see section 2.2)
RMA	Reference Model Atmosphere
RRC	Rayleigh Response Calibration
SNR	Signal to Noise Ratio
SRN	Software Release Note
STD	Standard Deviation
USR	Useful Spectral Range
VAMP	Vertical Aeolus Measurement Positioning
WGS84	World Geodetic System 1984 Earth Gravity Model 96 (EGM96)
WVM	Wind Vector Mode or Wind Velocity Measurement
ZWC	Zero-Wind-Calibration

2.5 Document preparation

This document was written using the \LaTeX typesetting system. All plots in section 3 have been created using gnuplot and the KNMI gnuplot wrapper module for Fortran90.

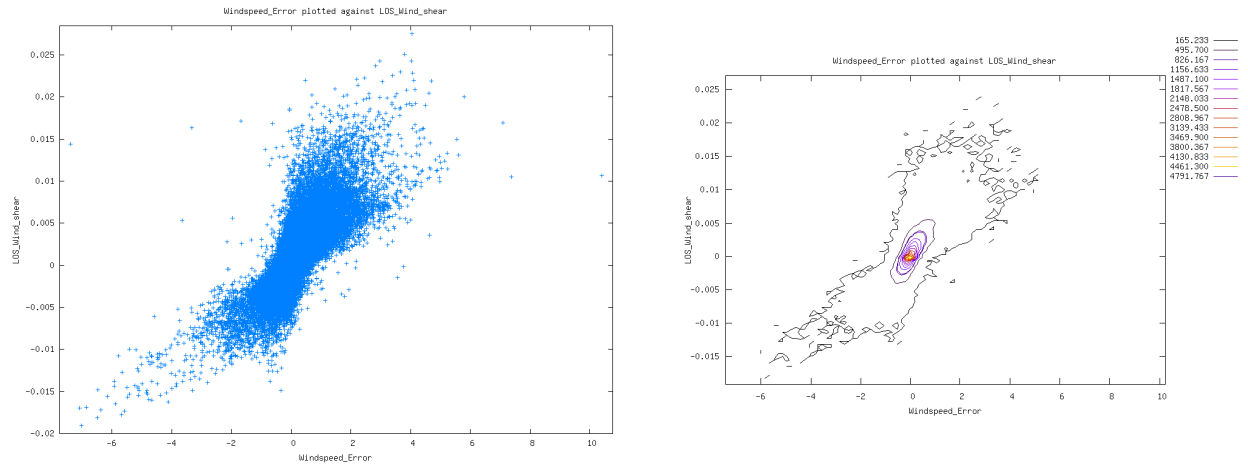


Figure 1: LOS Wind shear plotted as function of the wind speed error, for the CALIPSO data discussed in section 3.9. Scatterplot (left panel) and contour plot (right panel). The data has been selected on SNR > 100, which means this result includes mainly reflections on optically thick clouds. The error in wind speed is therefore mainly caused by height assignment errors.

3 Task 4a: Mie Core QC

3.1 Introduction

For this sub-task the usefulness of Mie core algorithm outputs to quantify wind shear, turbulence and horizontal gradients is studied.

Strong wind shear will lead to a broadened Mie return, caused by the fact that the return signal from a bin is composed of returns with different Doppler shifts. The spectral broadening may be amplified in case of strong variability in optical properties (cloud tops, aerosol layers), and will especially be present in case of transparent clouds or aerosols combined with vertical wind shear. Therefore the first quantity to be looked at is Mie line-width (FWHM), but also other fit results like Residual Error, number of outer loop iterations and number of inner loop iterations, and the Validity Flag reported by the algorithm will be looked at.

Another effect may be reflection on the top of a dense cloud, which will give an error in the altitude assignment of the wind, i.e. a wind in the top or bottom of the range bin (which may differ significantly due to wind shear) might be assigned to the center of the range bin. An example of this effect is given in Fig. 1. This will not lead to broadening of the Mie spectral signal, and thus cannot be detected by QC on the FWHM. An option to detect this is to look at the corresponding Rayleigh signal, which should be very low if the cloud top is near the top of the range bin, and higher if the cloud top is near the bottom of the range bin.

During this exercise the differences with the existing LIPAS simulation tool have been studied, in order to investigate whether it is suitable for further simulation work in the VAMP study, and what modifications it needs to make it compatible with the way the E2S performs the simulations. This is especially the case in the simulation of the transmission and Mie spectrum as discussed in subsection 3.3.

3.2 Setup of the tool

The original idea was to re-use the existing software for this study. This would mean running the matlab E2S simulator, which is very expensive in terms of run-time, then run the L1BP and the L2BP, and analyse the results. Apart from the inconvenient design of especially the E2S software (which has been circumvented partly by embedding it in a python script), the fact that it is implemented using an interpreted script language makes it inefficient in runtime, and this has proven to be a major bottleneck, which makes it impossible to generate sufficient statistics in a reasonable amount of time. Therefore a (somewhat simplified) dedicated Fortran90 simulation tool was built to perform this study. Note that surface reflections have not

been included in the current tool. This new tool runs about a factor 25 faster than the E2S matlab code, which makes it possible to run 5 half orbits worth of CALIPSO data in just 5.5 days, in stead of 140 days.

For the design of this new tool the actual algorithms and implementation used for the E2S have been taken as starting point. No additional work has been done in verifying the correctness of the E2S implementation or of the choice of functions and algorithms used by the E2S.

From this the Fizeau spectrometer transmissions have been implemented using Airy functions, and the laser and reflected Mie signal spectral lines have been modeled as Gaussian functions. The choice of Fortran90 for this new tool made it possible to re-use certain pieces of code already implemented in the L2BP software. This especially is the case for the Mie Core algorithm, which we test in this section. Other larger pieces of code that have been reused are the interpolation module, and the reader for the atmospheric database files.

The actual version of the tool used for the simulations described below has been delivered to ESA¹ as part of the L2BP software package version 1.40. It can be found in subdirectory MieCoreProcessing and is automatically compiled when the main make or build command is executed as is detailed in [RD1].

The simulation tool is called MieCoreScan and is started by calling it on the command line, followed by some command line options. Relevant for this study is the option “-do_atm_db_scan” followed by two filenames to be used as input. First the name of the atmospheric database file must be given, then the range bin definition file must be given. In addition the option “-NoPoissonNoise” can be used to switch off the simulation of Poisson noise, and the (debug) option “-AccdPrintSignal” can be used (followed by a profile and bin number) to print the actual accd counts as simulated for a given data point.

The atmospheric database files used by the tool are compatible with the database version 1.6 as delivered to ESA². The range bin definition data files are included in the L2BP software delivery in directory input_data/rangebin_defs.

The output of the simulation tool is a text file listing a number of relevant atmospheric inputs, and the Mie Core outputs for each simulated rangebin. To analyse this text file (or a number of these files glued together) a simple plot program named PlotMieCoreAtmDBScanResult has been added. The plot program reads this ASCII file, and through command line options it can be told which parameter to plot against which other parameter. Also selections on one or more variables can be applied before the data is plotted (i.e. SNR>10 etc.).

The actual steps taken during the simulation are as follows:

- the Fizeau transmission curves are pre-calculated³ and stored in a look-up table (one for each ACCD bin).
- the rangebin and atmospheric database files are read
- if needed the input los wind is calculated from the u,v,w wind components using a fixed incidence angle of 37.5°. Note that some atmospheric database files directly specify the HLOS value at the location of the u component. This is signaled to the software by setting the v component to the missing indicator.
- the error multiplier method is applied to increase the wind shear
- the vertical shear in horizontal wind speed and optical properties are calculated from the input atmosphere
- the laser light is propagated through the atmosphere using the atmosphere defined by the atmospheric database input file
- for each measurement the actual spectrum is calculated with sub-rangebin precision (using vertical layers of 50 m). The peak height is scaled to match with the numbers produced by the E2S software

¹It was delivered on 28-Feb-2009. See the eRoom folder: AEOLUS > LIBRARY_of_PROGRAMS-CFIsTOOLS > ECMWF-L2B > L2BP_Release1.40_20090228

²The database is packed into a single tar file named atm_database.1.6.tar.gz (size: 585 MB) It was delivered on 21-Aug-2008, and is available from the L1B ftp server: ftp.aeolus.esa.int path /home/aeolus-l1bp/to_estec.

³Note that the current E2S v2.08 does not store these results, but in stead it re-calculates the transmission curves again for each measurement. This is a significant waste of computing power, since these curves remain constant during an experiment.

Table 1: Parameter values used to simulate the Mie spectral peak, and the Fizeau transmission curves.

Parameter	value	unit
Number of frequency steps	1201	
sub-accd bin steps	10	
simulated frequency range	-1 to +1	[GHz]
Mie peak FWHM	50	[MHz]
Fizeau FSR	2.15	[GHz]
Fizeau USR	1.502	[GHz]
Fizeau FWHM	184	[MHz]
Incidence Angle	37.5	[deg]
Background level	5	[counts per measurement]
Peak Height multiplication factor	3555	[times]
accd channel width = USR/16	93.875	[MHz]

for a given scenario⁴

- Poisson noise is added
- the spectrum is discretised to a 2 byte integer number (as is done in the actual processing software)
- the measurements are accumulated into an observation (currently a fixed number of 20 measurements per observation is taken).
- the Signal-to-Noise-Ratio (SNR) is calculated from the spectrum using the algorithm implemented in the current L1BP, see [RD2], section 14.4.4 (page 55). eq. 79.5.
- the Mie Core algorithm as defined in the L2BP is called on the simulated spectrum
- finally, the results are written to the output file.

3.3 Simulation of transmission and spectrum

The spectrum of the reflected Mie signal is simulated as a simple Gaussian function $g(f)$, with a prescribed FWHM, as defined in eq. 1:

$$g(f) = \frac{1}{\sigma * \sqrt{2\pi}} \exp\left(\frac{-(f - f_M)^2}{2\sigma^2}\right) \quad (1)$$

in which f is the observed frequency, f_M is the center frequency of the Mie reflected signal (which depends on the local LOS wind speed), and σ describes the width of the peak, and relates to the FWHM as specified in eq. 2:

$$\sigma = \frac{FWHM}{2\sqrt{2\log(2)}} \quad (2)$$

This is as described in [RD3] (appendix B). The actual value used for FWHM is given in table 1. The frequency shift df caused by the Doppler effect for a wind speed s is given by eq. 3:

$$df = f_0 - f_M = \frac{-2sf}{c + 2s} \quad (3)$$

In which f_0 is the central frequency of the laser light, and c is the speed of light. An example of this spectral curve for a Windspeed of 50 m/s is given in figure 2.

⁴The scenario used for this calibration was single_RMA_midlat_winter_and_cirrus, and the reflection on top of the cirrus cloud has been taken as reference. In this scenario a peak height of 4039 counts (including a detection chain offset of 1444 counts), and a background level of 5 counts occurs, and the current MieCoreScan simulation tool has been tuned to produce exactly these numbers when this same scenario is taken as input.

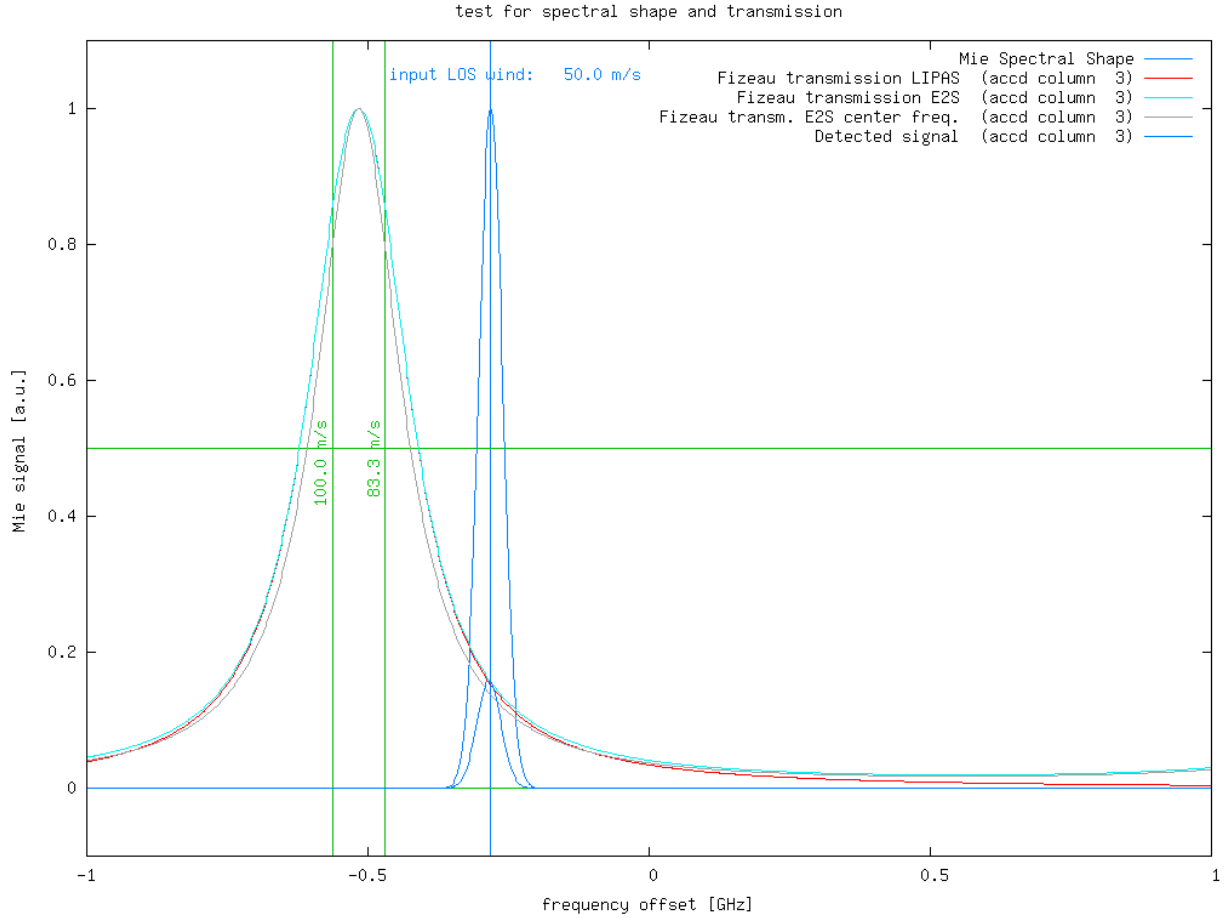


Figure 2: Mie spectral shape for a wind speed of 50 m/s (blue), compared to the much broader Fizeau transmission curve for accd column 3 calculated in 3 different ways (red: using the Lorentzian function as is done in LIPAS; cyan: using the Airy function as in done by the E2S; gray: using the Airy function but without integration over the accd bin surface). The 2 vertical green lines denote the edges of the accd bin used for this example, and the vertical blue line denotes the frequency that corresponds to the Doppler shift of a LOS wind of 50 m/s. Finally, the small blue peak below the Mie spectral shape, is the amount of light entering accd bin 3.

For the Fizeau transmission the E2S uses an Airy function $Y(\lambda)$ for each wavelength as defined in eq. 4:

$$Y(\lambda) = \frac{1}{1 + B * \sin^2\left(\frac{A}{\lambda}\right)} \quad (4)$$

$$A' = \frac{\pi}{\frac{1}{\lambda_M} - \frac{1}{\lambda_M + R_{FSR}}} \quad (5)$$

$$n = \text{floor}\left(\frac{A'}{\pi\lambda}\right) \quad (6)$$

$$A = nA'\lambda_M \quad (7)$$

$$B = \frac{1}{\sin^2(\lambda_M + R_{FWHM})} \quad (8)$$

as described in [RD3] (appendix A). Here R_{FSR} is the FSR converted to a wavelength difference, R_{FWHM} is the FWHM converted to a wavelength difference, λ is the actual observed wavelength, and λ_M is the center wavelength of the simulated accd bin.

Since each accd bin does not detect a single wavelength, but spans a range of wavelengths, this airy

function is calculated 10 times for each accd bin (identical to what is implemented in the E2S) for center wavelengths corresponding to a location that varies over the accd surface. Then the actual transmission function is constructed by averaging these 10 airy curves. The effect of this is demonstrated in figure 2, a single airy curve is plotted in grey, and the average of 10 gradually displaced airy curves is plotted in cyan. A clear broadening of the transmission curve can be seen. These curves are calculated using a frequency range of -1 upto +1 GHz (see table 1), in 1201 steps⁵.

The LIPAS software uses a simplified transmission curve by implementing it as a Lorentzian function (see [RD4], section 5.1, pages 22 and 23). This function is defined by eq. 9:

$$L(\lambda) = \frac{1}{1 + \left(\frac{\lambda - \lambda_M}{\frac{1}{2}R_{FWHM}}\right)^2} \quad (9)$$

LIPAS then applies an adaptive numerical integration over the accd width, contrary to the 10 steps used by the E2S. The shape of this function is also plotted in figure 2 (red curve). For the largest part of the peak it is very close to the Airy function (maximum difference is only 0.52% for the region with transmission above 10%). In the tail the functions clearly start to deviate because the Airy function is periodic and will have another peak at a distance of R_{FSR} from the displayed peak, while the Lorentzian approaches zero at larger distances of the peak location.

Figure 3 shows the transmission curves of all 16 accd bins, and compares them with the Mie peak. The smaller peaks plotted below the Mie peak represent the photons that are expected to be seen by each accd bin. The sum of these curves will correspond to the signal to be expected on each accd bin, and this is indicated by the red squares.

Looking at these results, there are small differences in the results from the E2S and LIPAS implementations, but these seem very small, much smaller than the typical noise expected on each bin. Typically a noise level of \sqrt{n} is to be expected for a signal level n , so for the signal on top of the cirrus cloud used for calibrating the system of 2600 counts⁶, this would lead to a noise level of 51 counts, so almost 2%, so clearly far above the difference between the Airy and Lorentzian functions in the relevant wavelength range of only 0.52%. In addition the difference between the 2 curves will mainly manifest itself as an additional offset, which is automatically corrected for by the Mie Core fitting procedure. Therefore it seems very well possible to use the LIPAS tool for further simulations, provided that the peak widths used as input are identical to the ones used by the E2S.

3.4 Mie Core Algorithm

The detected accd signals will be fitted to a Lorentzian function using the Mie Core Algorithm. Currently this routine implements a downhill-simplex method (see [RD5], section 3.2.9) to retrieve the following 4 parameters: peak position, peak height, FWHM, and offset of the Mie signal. In addition the following outputs are available: data validity flag, residual error, and number of iterations (of the outer loop).

Actually this routine implements a double loop. The peak height and offset are treated as linear parameters in the outer loop. The peak position and FWHM are the two non-linear parameters that are optimized using the Downhill Simplex method. Note that the stop condition inside the inner loop uses a threshold that needs FWHM and peak height determined in the outer loop. There is no possibility at the moment to fix one of these 4 parameters during the fitting process.

The Mie Core algorithm has an internal consistency check on the validity of the found fit result. The code checks:

- whether the peak height is within a lower and an upper bound
- whether FWHM is within a lower and an upper bound
- whether the found peak location found is within a given distance to the location of the accd bin with the highest count

⁵This is equivalent to the range -0.42 upto 0.42 pm in steps of 0.00084 pm. In comparison the E2S also uses 1201 steps (see [RD3], section 6.4, p.6-10), but calculates a much wider range of -3. upto 3 pm in steps of 0.005 pm, because the same arrays are used for simulation of the Rayleigh spectrum and transmission curves as well.

⁶this equals the total photon count minus the detection chain offset.

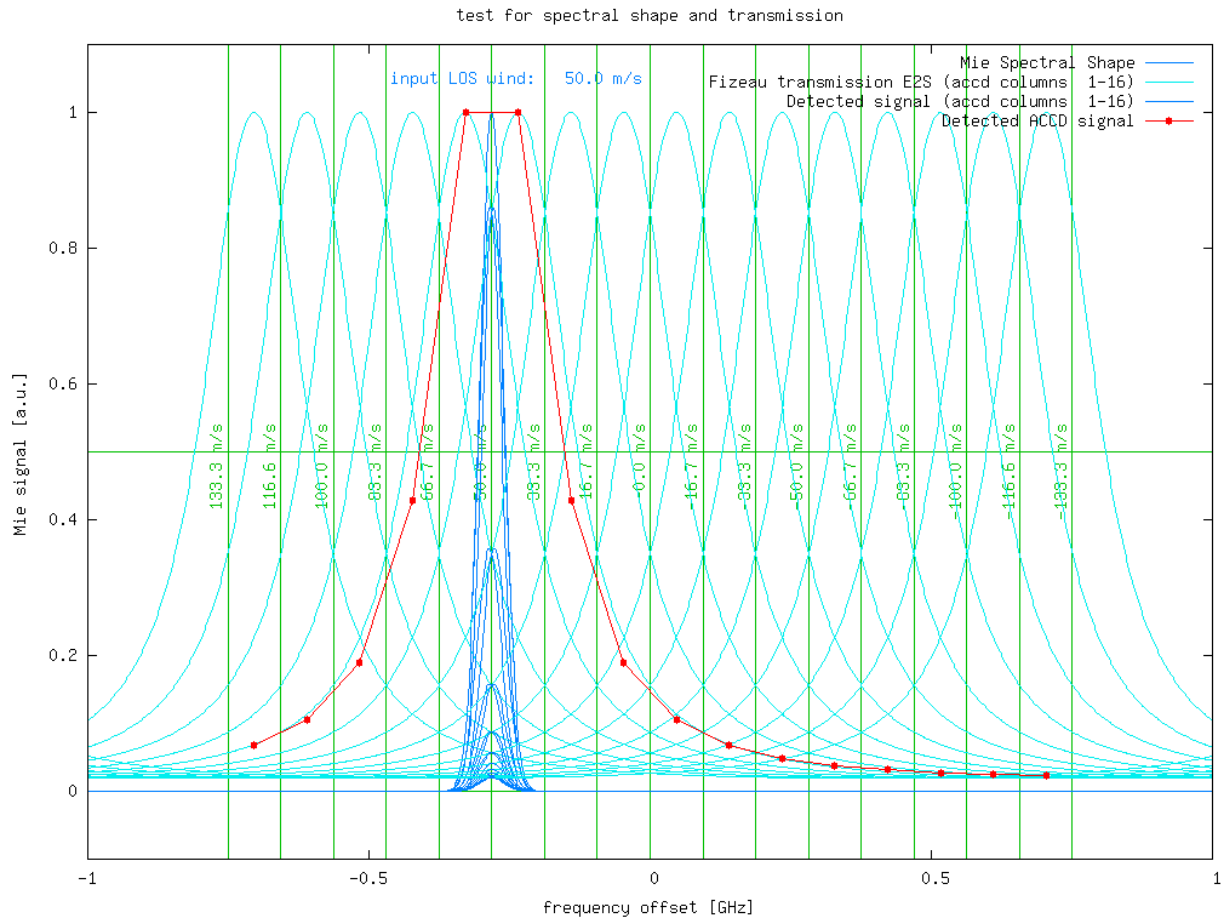


Figure 3: Mie spectral shape for a wind speed of 50 m/s (blue), compared to the the Fizeau transmission curves for all 16 accd columns (cyan). The blue curves below the Mie peak denote the fraction of light detected by each of the accd bins. Clearly, due to the long tails of the transmission curves, all accd bins will receive some signal. The expected signal that will be measured by each accd channel is denoted by the red squares.

Table 2: Mie Core algorithm settings.

Parameter	value	remark
Start FWHM	2.0	
Res.Err. Threshold	0.001	outer loop threshold
Max.Iterations Lorentz Fit	30	outer loop threshold
FWHM Upper Threshold	2.5	used for val flag
FWHM Lower Threshold	1.5	used for val flag
Peak height Upper Threshold	1.6	used for val flag
Peak height Lower Threshold	0.9	used for val flag
Peak Location Threshold	1.0	used for val flag
Non Lin. Opt. threshold	0.01	inner loop threshold
Max. It. Non Lin. Threshold	20	inner loop threshold
NumSpecSubSamp	5	
tripod obscuration correction	1.0	not used

if either one of these conditions is not fulfilled the validity flag is set to False. The thresholds used for these checks are given in table 2. In addition this table gives a start value for the FWHM, thresholds for the allowed number of iterations in the inner and outer loop, for the Residual Error Threshold (used to terminate iterating the outer loop), and for the non-linear optimization threshold (used to terminate iterating the inner loop). Finally NumSpecSubSamp determines how much points of the Lorentzian function are evaluated within a single accd bin. The tripod obscuration correction is ignored in the current study and always set to 1.

Todo: get info on inner loop out of the MieCore algo ...

3.5 Input data

The specific Mie Core tests described in this section do not depend on the actual rangebin definition. Therefore all tests have been performed with the WVM1 rangebin definition as defined in [RD6], section 6.1.1.

The following atmospheric data (as described in the atmospheric database [RD8]) has been used for testing:

- RMA: the single_RMA_midlat_winter_and_cirrus scenario consists of a single profile. It uses the climatological RMA midlatitude winter profiles for temperature and pressure and aerosol and molecular backscatter and extinction. On top it adds a cirrus layer between 12 and 14 km. Input wind is assumed to be zero. This scene has been used to calibrate the MieCoreScan simulation program with the E2S results as detailed above.
- LITE: the LiteScene_A_and_EcmwfColloc_3.6km scenario consists of 240 profiles separated by 3.5 km. It uses LITE data for aerosol and cloud backscatter and extinction. Colocated ECMWF model data is used for pressure, temperature, (and derived from this molecular backscatter and extinction), and the wind components u,v,w.
- CALIPSO: 5 half orbits of CALIPSO data have been used to produce the statistics presented in the next sections. This includes 35 scenarios from the atmospheric database named Calipso_Orbit[A]_Segment[B]_and_EcmwfColloc in which [A] runs from 1 to 5, and [B] runs from 1 to 7 for each value of [A]. Each half orbit consists of about 5600 profiles separated by 3.3 km, so in total about 28.000 profiles have been used. These scenes use CALIPSO data to define the particle backscatter and extinction (clouds and aerosols are not separated in this dataset). Colocated ECMWF model data is used for pressure, temperature, (and derived from this molecular backscatter and extinction), and the wind components u,v,w.

An example of the ECMWF input data Colocated with the CALIPSO data used is displayed in Fig. 4. This plot clearly shows the effect of the current bug in the system, which causes most LOS winds to be positive, as mentioned in the general introduction, section 1.

3.6 Error multiplier

By comparing the NWP data with radio sonde measured profiles it was found in [RD7] that the variability of the profiles in the NWP model underestimates the real situation significantly. To compensate for this an error multiplier method has been proposed in [RD7] (see section 4.2.1.8, p.47 and section 5, p.190) to compensate for this. This multiplier adds noise to the NWP wind profiles in such a way that the statistical properties of the vertical shear in the modified NWP profiles approach the properties of the measured radio sonde profiles. This method has been adopted in the MieCoreScan simulation program and has been applied to the CALIPSO data. In short the error multiplier is implemented as follows:

- take slices of 1 km atmosphere
- for each slice calculate the vertical shear profile using 30 m steps
- calculate the variance of this shear within the slice
- multiply this variance with the error multiplier (currently a constant multiplier value of 2 has been used)

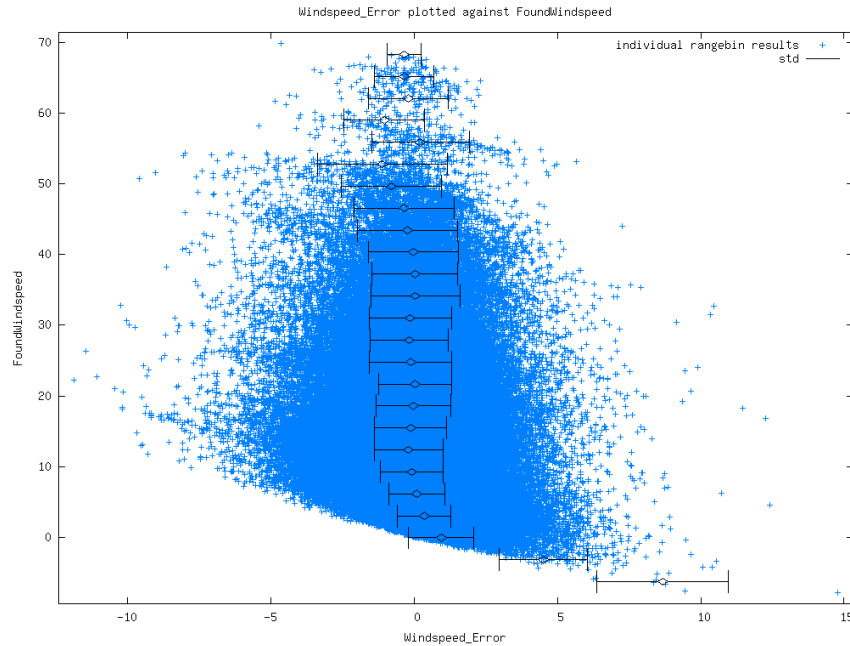


Figure 4: Input LOS winds derived from ECMWF u,v,w components, colocated at the locations of the CALIPSO data used for the majority of simulations, plotted against the wind speed error found by the simulations (for results with SNR>10 and ValidityFlag set).

- apply this enhanced variance to add extra noise (with a Gaussian distribution) on the wind profile

Finally it must be noted that a small bug was discovered in the implementation of this error multiplier method in the MieCoreScan simulation tool. After comparing the statistics of the correct version and the version implemented in this tool it was found that the amount of shear has been increased a bit to much (between 30 and 40 %). Since this means that the span of values used for these simulations at least covers the range expected in a realistic atmosphere it will most likely not affect the conclusions.

3.7 Horizontal gradients and turbulence

The ECMWF winds used in the simulations here, lack both vertical and horizontal wind variability ([RD9] and [RD7]). The current simulations use a single profile for all measurements within a BRC. Each measurement therefore sees the same atmosphere, and the same realization of the enhanced shear from applying the error multiplier method, and only the noise realizations differ from measurement to measurement. This allows for simulation of vertical shear in the HLOS wind because the resolution of the atmospheric profile (125 m) is finer than the rangebin sizes (500 m, 1 km and 2 km).

The lacking wind variability in the horizontal is about 0.5 m/s in standard deviation over 50 km [RD9], while the lacking wind variability in the vertical is about 2 m/s [To Check This] in standard deviation and much larger [RD7]. Horizontal wind variability has thus not been simulated here, since its contribution is relatively small, while this would make the simulation program much more complicated.

Moreover, no effort has been put in implementing horizontal variability because the effect will be exactly the same when a single rangebin is considered (which is what we do for this Mie Core testing). After all horizontal accumulation of measurements having variable wind into an observation, will give the exact same result as vertical accumulation of sub-rangebin layers having variable wind into one rangebin, as long as the (observation scale) rangebin itself is processed independently from its neighbours.

Turbulence represents coherent vertical and horizontal wind variability, and has not been simulated for the reasons defined above. Enhanced turbulence associated with the formation of cumulus clouds will be correlated with the optical properties of the atmosphere. However, in [RD7] it was established that such updrafts are often largely hidden below the non-transparent cloud tops.

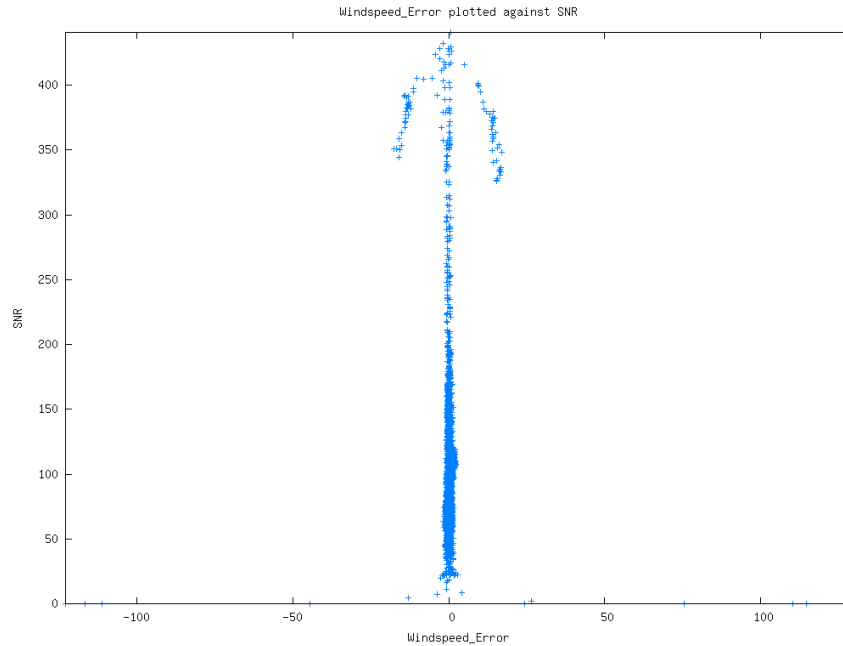


Figure 5: First simulation results using LITE data. The data shows nonsense retrievals for very low SNR and a clear overflow problem.

3.8 Results: LITE data; signal scaling

During development of the MieCoreScan simulation program a number of preliminary tests have been done using LITE data. These tests used LITE scene A consisting of 240 profiles, which resulted in simulation of 5760 range bins. For each rangebin the Mie Core algorithm was applied as detailed above, the SNR was determined and the difference between input and retrieved los wind speed was determined. An at first sight weird but interesting result was obtained when SNR was plotted against wind speed error, see Fig. 5.

The very large wind errors for SNR close to zero can easily be understood since these are cases without significant signal level, for which one cannot expect a good fit result. These can be easily discarded by setting a threshold on the SNR value. A value of SNR=10 has been applied⁷ in Fig. 6 (left panel). This improves the situation, the points around SNR=0 with wind speed errors upto 120 m/s have disappeared.

However, the umbrella like shape for large SNR values needs an explanation. At closer inspection it was found that this was caused by a too high signal level assumed for this test, caused by a wrong calibration of the peak height⁸. This causes too high photon count numbers to be stored in the 2 byte integer fields available for the spectrum. As a consequence the numbers get truncated and become nonsense⁹.

To illustrate this an example of a spectrum showing this problem for the internal reference pulses is shown in Fig. 7. The blue lines give the measurement data, which is correct and reasonably well fitted in this example. The black lines give the internal reference data, which clearly have been wrapped around. The peak location is just slightly wrong for this example, but from Fig. 6 it is clear that errors upto 15 m/s can result from this effect.

⁷the current L1B processing uses a threshold value of SNR=10 as well.

⁸in fact measurement level signals had been scaled to match observation level E2S results, causing the peak height to be a factor of 67 too high. Note that this preliminary test assumed 67 measurements per observation. The remainder of this study will use 20 measurements per observation.

⁹Note that this has been observed on the actual L1BP v1.07 output as well, but then for the internal reference accd counts, which have higher signal levels and are much more vulnerable for this kind of problem. That particular problem has been solved now, but the software may experience this phenomenon again in case signal levels are higher than anticipated (for example due to some wrong magnification setting). It must be noted that this may happen without any warnings or errors issued by the L1BP code, since this is implemented by means of a simple cast operator (i.e. $x = (\text{unsigned short}) y.$) in line 4875 of `aeolus/lobproc/LBC/src/LBC.WVMProcessor.cc`, even in the most current L1BP version (5.03 when writing this) (this issue has been reported to Dorit Huber on 08-Apr-2009).

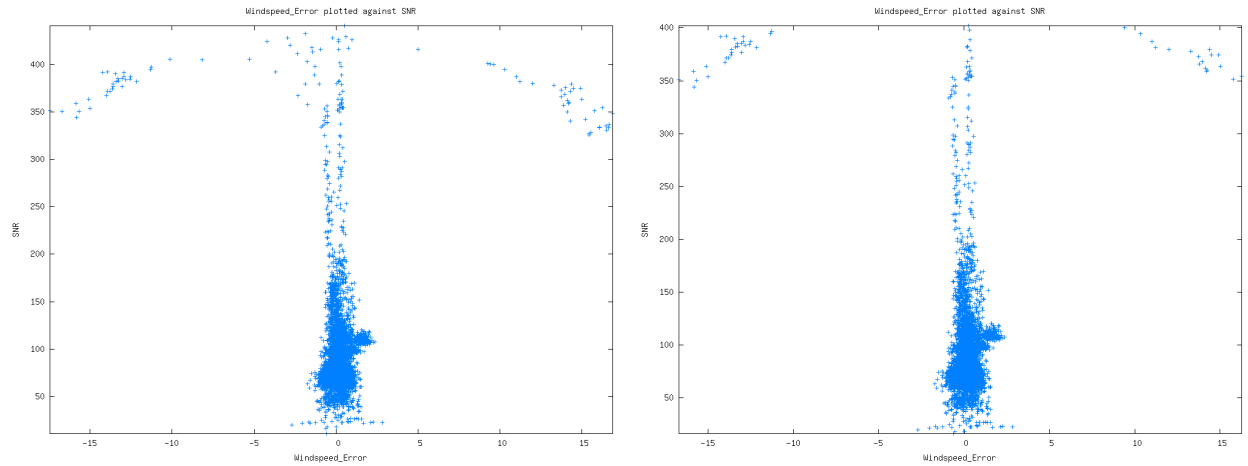


Figure 6: Simulation results using LITE data, after applying a SNR threshold of 10 (left panel), and after discarding data reported invalid by the ValidityFlag as well (right panel). The nonsense retrievals for very low SNR have disappeared now but the overflow problem remains.

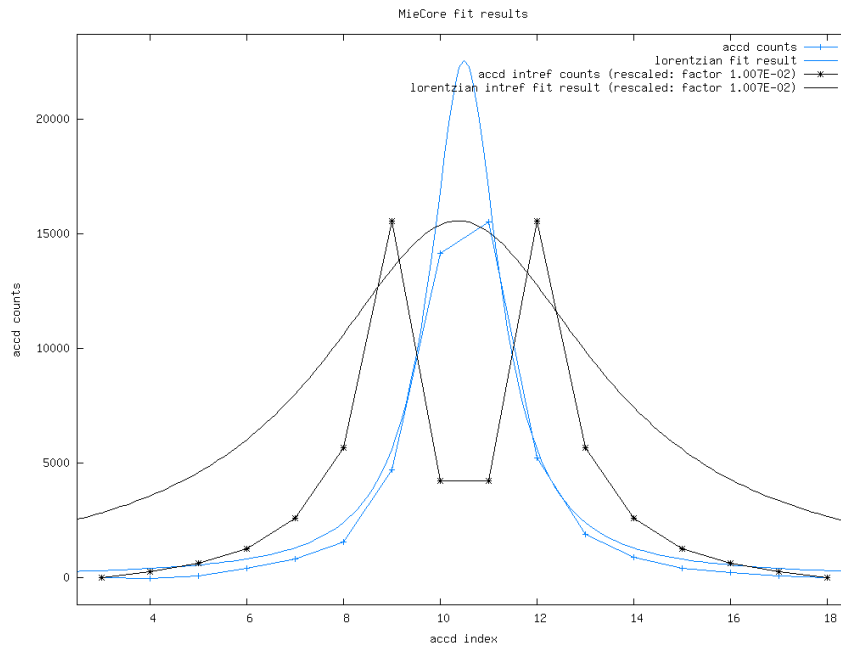


Figure 7: A case showing the effect of an integer overflow in the accd counts of the Mie channel (these are internal reference counts, and this case was generated by L1B v1.07).

A first and obvious approach to solve this processing problem was to discard all data reported invalid by the ValidityFlag which is one of the MieCore algorithm outputs. The result is displayed in Fig. 6 (right panel). In this way a few bad points can be filtered out, but many problematic points remain.

Only when the signal levels were adapted to the proper values the overflow problem disappeared, see Fig. 8. Note here that these results are simulated at measurement level, which explains the rather large errors of 5 m/s and more found for the lower SNR values. As is shown in the next section, this will improve significantly when simulations are done at observation level.

The same simulation was re-run without Poisson noise. In this case the wind speed errors are significantly

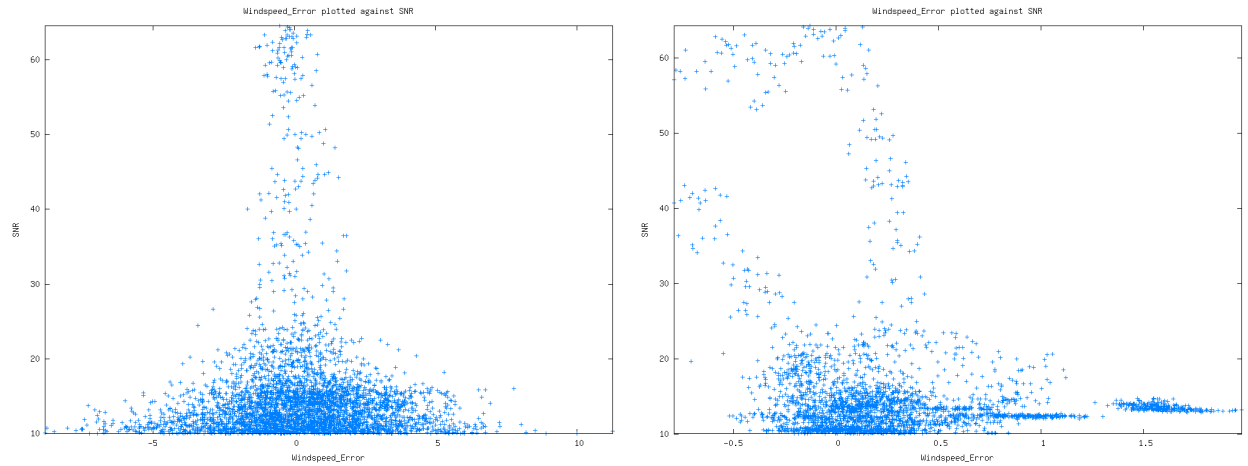


Figure 8: Simulation results using LITE data. In this case a correct signal scaling is applied and the overflow problem disappears (left panel). Note that the MieCore algorithm is applied at measurement level in this case. To illustrate the effect of Poisson noise on the simulations the same case has been run without Poisson noise (right panel).

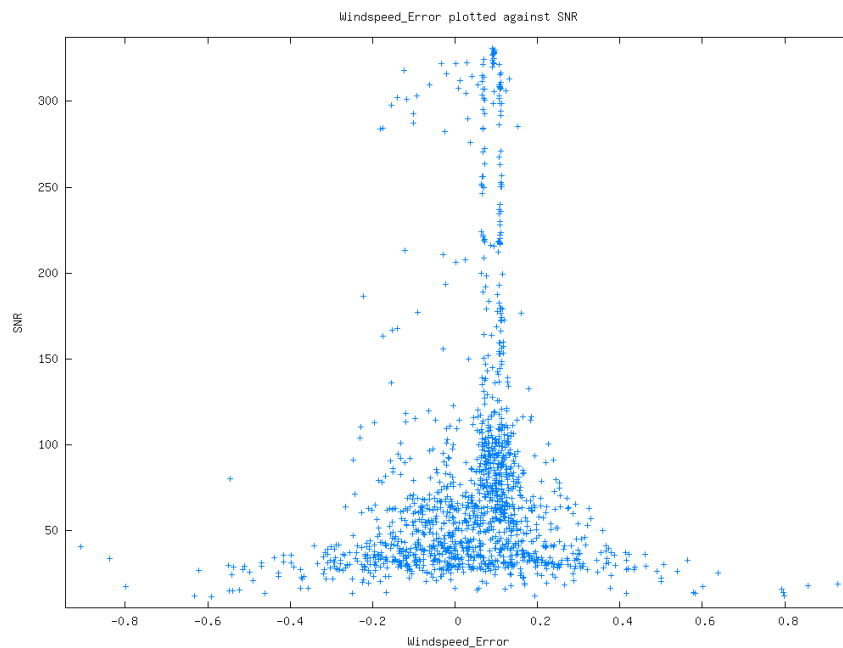


Figure 9: Simulation results using LITE data, without wind variability and without Poisson noise, to illustrate the numerical inaccuracies of the simulation program.

reduced, as expected to the range -0.7 to $+2$ m/s, see Fig. 8. The remaining errors are mainly caused by the atmospheric variability within a single rangebin, although also some numerical inaccuracy in the simulation might enter.

To highlight the influence of numerical inaccuracies on the simulations a case was simulated with the same LITE data but with constant input wind of 50 m/s. By removing variations in wind speed also all variations in optical properties should become insignificant for the shape of the spectrum (they may only influence the magnitude of the signal in this case). The results of this tests are shown in Fig. 9 Typical wind

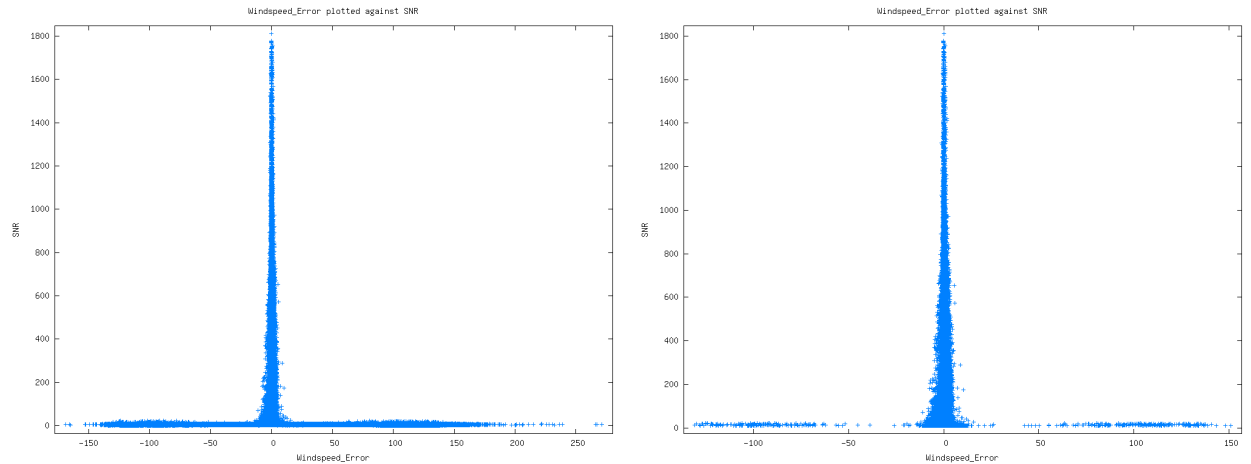


Figure 10: Simulation results using CALIPSO data, all results without selection (left panel) and after applying a SNR threshold of 10 (right panel).

speed errors are now below 0.2 m/s although for low SNR values slightly higher values do occur. However, note that a significant part of these errors are actually caused by the discretisation of the signal to a 2-byte integer number (to simulate the discrete number of photons that will be detected), and there are no further tricks possible to remove this term. This discretization also is probably the reason for the two vertical lines that appear in this plot. **The bias of around 0.1 m/s in this plot, especially visible for higher SNR values and accentuated by the 2 vertical lines of points are due to the bug mentioned in the general introduction, section 1.** Clearly these remaining numerical and discretisation effects are insignificant compared to the errors that occur when Poisson noise is added.

3.9 Results: CALIPSO data; data selection

A next series of tests that have been run using CALIPSO data (for details see section 3.5). These tests have been run at observation level, i.e. each rangebin was constructed by simulating 20 measurements with differing noise realizations, and subsequent accumulation of the spectral counts. This yielded about 670.000 simulated range bins. The same data selection as presented in section 3.8 above has been applied, i.e. a threshold on SNR of 10 and a selection of results with ValidityFlag set by the MieCore algorithm. To illustrate the effect of these selections, the raw results are displayed in Fig. 10. This plot displays the SNR as function of the wind speed error found by the simulation (left panel: all data; right panel: points with $\text{SNR} < 10$ discarded).

Clearly a number of bad results still are present in this dataset. Fortunately these are all flagged by the ValidityFlag result reported by the MieCore algorithm. If only points are selected that pass this test, this results in Fig. 11. The actual number of points discarded by the SNR threshold of 10 is about 189.000 (28 %) and leaves about 481.000 points. The number of additional points discarded by the ValidityFlag is about 12.000 (1.8 %) and leaves about 469.000 valid range bins.

Since interpreting this kind of scatter plots is often easier when the data is displayed as a contour plot, the same data as in Fig. 11 has been displayed as contour plot in Fig. 12. Note that the majority of points seem located near the origin of the plot, so a magnified view zooming in on the region $10 < \text{SNR} < 100$ is given in the right panel of Fig. 12. Actually about 80% of all data is located in the magnified region.

3.10 Results: CALIPSO data; Mie Core outputs

The selected CALIPSO data as described in section 3.9 has been used to look at the available outputs of the MieCore algorithm. This includes the number of outer loop iterations, residual error and FWHM.

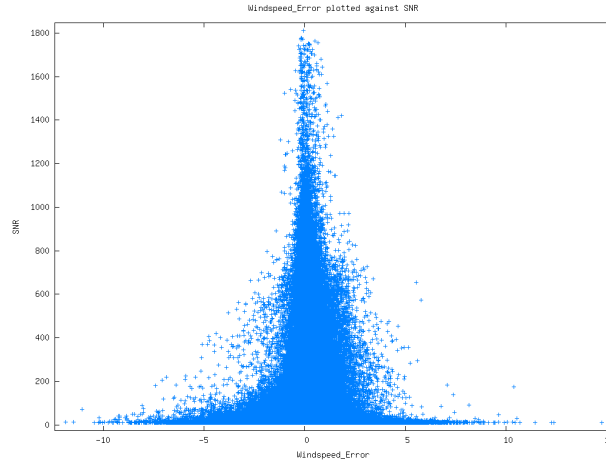


Figure 11: Simulation results using CALIPSO data, after applying a SNR threshold of 10 and after discarding data reported invalid by the ValidityFlag. as well (right panel).

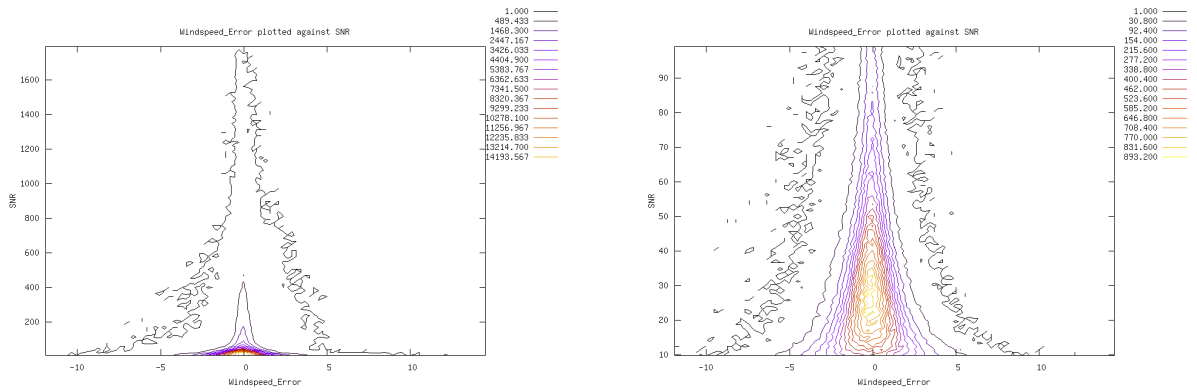


Figure 12: All valid data after discarding points with $SNR < 10$ and invalid ValidityFlag (left panel). The right panel shows a magnification of the region $10 < SNR < 100$.

3.10.1 Outer loop iterations

Figure 13 shows the number of outer loop iterations used by the MieCore fitting as a function of wind speed error. The most prominent feature in this plot is the discontinuity in the data, i.e. iterations between 11 and 30 seem never to occur. When looking at the actual number of points in both regions, there are about 116.000 points with 10 or less iterations (25%) and about 352.000 points with 30 iterations¹⁰ (75%). Clearly the Max.Iterations Lorentz Fit threshold can be set tighter, to about 12. Although it has not been investigated, a clear improvement in runtime may be expected from this change. A simple estimation results in about a factor of two¹¹.

Note that the current MieCore algorithm implementation does not allow easy access to the number of

¹⁰Note that due to the applied threshold the iteration count is set to 31, but the 31st iteration is never executed, therefore the number of iterations for all points on this line actually 30.

¹¹Assuming the inner loop needs a fixed amount of time T (which probably is too simple), and assuming 2 states to represent the 2 regions of the plot, one with 30 and one with 8 iterations, this would mean a total runtime of: $0.75 \cdot 30T + 0.25 \cdot 8T = 26.5T$. For the changed iteration threshold this would be: $0.75 \cdot 12T + 0.25 \cdot 8T = 13T$. The runtime would thus decrease by about a factor of 2.

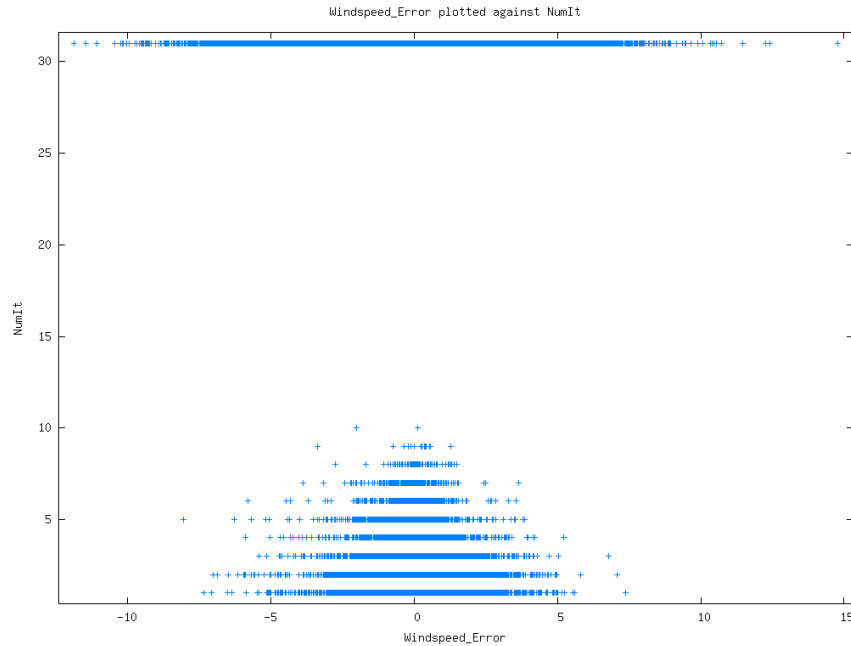


Figure 13: Number of outer loop iterations, as function of the wind speed error, for the selected CALIPSO data.

inner loop iterations. If time permits this will be investigated at a later stage.

3.10.2 Residual error

The residual error reported by the Mie Core algorithm is closely related to the number of outer loop iterations, since it is used as criterium to stop iterating in the outer loop. Currently a threshold of 0.001 is used (see section 3.4). As soon as the residual error drops below this value the algorithm stops iterating the outer loop.

The actual found values for the residual error are shown in Fig. 14. At first sight the distribution seems symmetric and centered around the origin.

When zooming in there is a clear concentration of points around $\text{ResErr}=0.001$ and below. These correspond to the range bins that did converge within 11 outer loop iterations. Clearly these converged results are much better than the remaining ones.

When looking at the actual occurrences this still is a minority of all points due to the small surface of the spot. The number of points with $\text{ResErr}>0.1$ is only 2%, but the number of points with $\text{ResErr}>0.01$ is already 33%, and the number of points with $\text{ResErr}>0.001$ is already 75% (this corresponds to the 75% of cases that did not converge within 30 iterations as presented in the previous subsection). Therefore applying a threshold on the number of outer loop iterations of 12, or on the residual error of 0.001 would significantly improve the quality of the data, but would also reject the majority of all data points.

Since from the contours in Fig. 15 it is not obvious to see the actual distribution of the data for a given residual error, due to the variation of the number of data points, a new figure was made, see Fig. 16. In this scatterplot the mean and 1 standard deviation (STD) on both sides of the mean have been over plotted. The resulting wind speed error STD values vary between 0.98 and 4.79 m/s for the plot showing all data (left panel), and between 0.57 and 1.20 m/s for the plot which selects data with a residual error of 0.01 or lower (right panel). From the right panel the position at which the wind speed error STD exceeds the value of 1 m/s (requirement for the LOS wind speed accuracy) can be read, which is at about 0.005.

The overall STD of the wind speed error distribution decreases from 1.2 to 0.75 m/s by applying this selection. This is at the cost of throwing away 227.500 points (so 48% of all data flagged valid by the SNR test and the Validity flag).

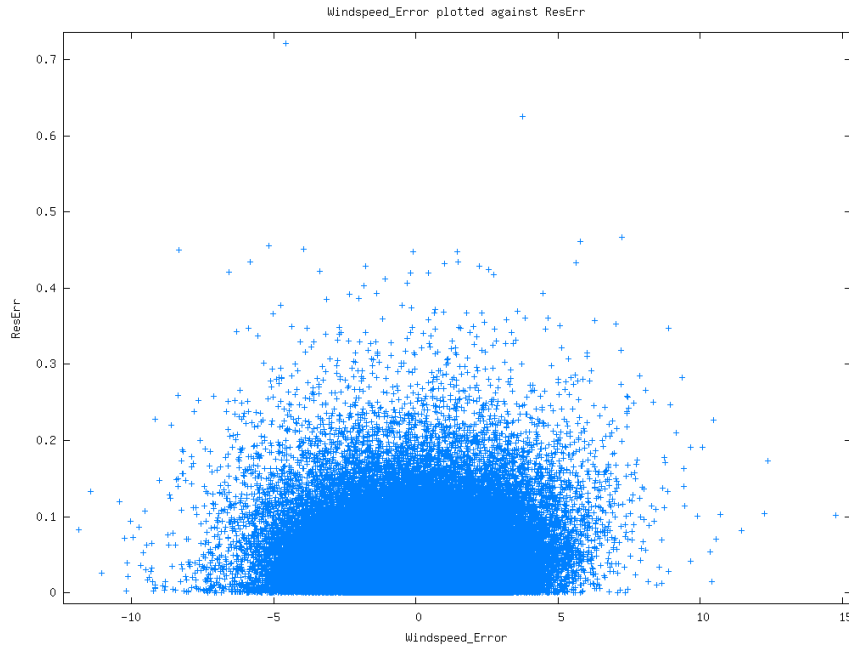


Figure 14: Residual error reported by the Mie Core algorithm as function of the wind speed error, for the selected CALIPSO data.

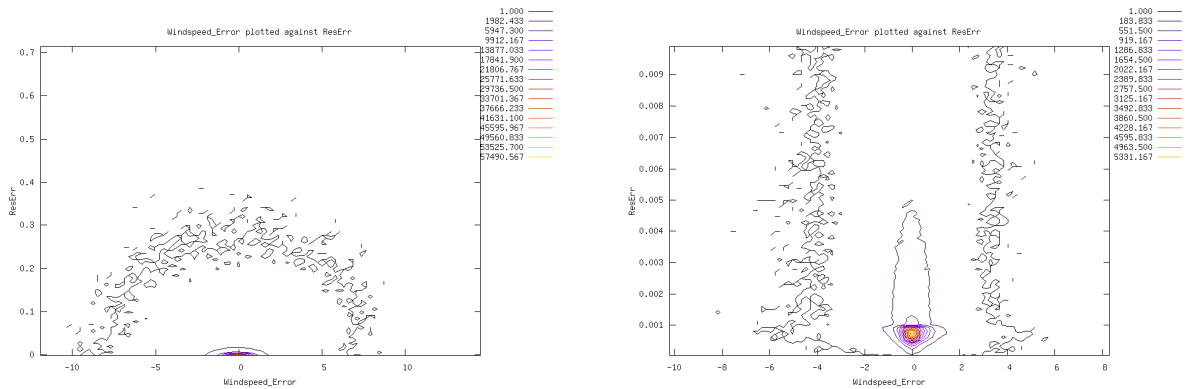


Figure 15: Residual error reported by the Mie Core algorithm as function of the wind speed error, plotted as contour plot. All valid results (left panel) and the region below ResErr=0.01 magnified.

Therefore the residual error shows a clear variation of the expected error and could be used to screen or downweight the Mie Core wind results.

3.10.3 FWHM

A last variable to look at is the FWHM of the Mie Core fit result. This variable is expected to be larger when reflected signals from different wind speeds are mixed within a single rangebin. This typically may happen when vertical wind shear occurs in combination with transparent clouds or aerosol layers (as argued in section 3.7). A simple scatterplot of all selected data is shown in Fig. 17 and does not reveal much detail. It clearly has been cut-off at FWHM=1.5 and FWHM=2.5 due to the selection applied by the validity flag.

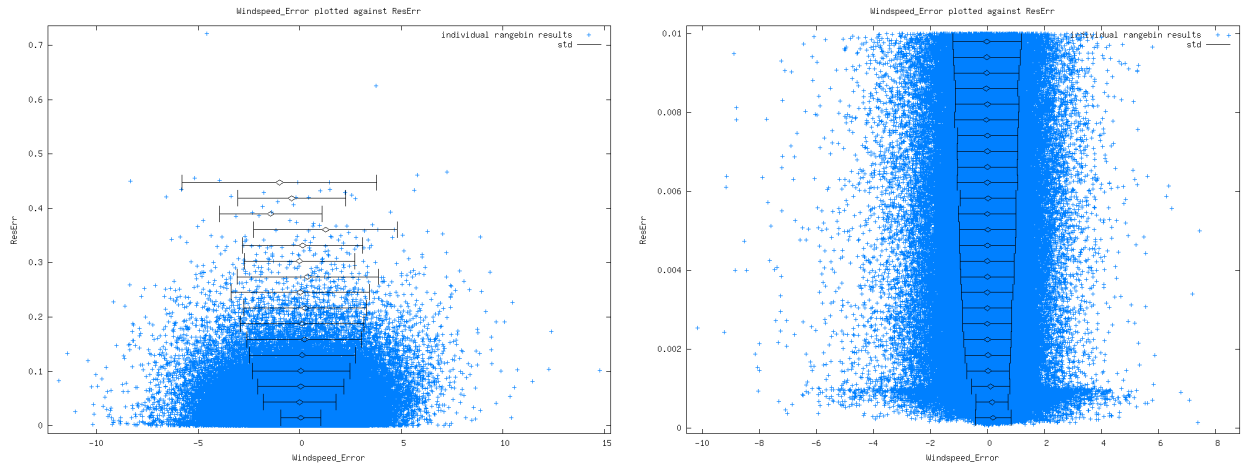


Figure 16: Same data as shown in Fig. 15, again as scatterplot, but now with over plotted lines to show the mean and one std value on both sides of the mean, for a number of Residual Error layers (full range is split in 25 sub-ranges for this calculation). From the right panel the position at which the wind speed error std exceeds the value of 1 m/s (requirement for the LOS wind speed accuracy) can be read, which is at about 0.005.

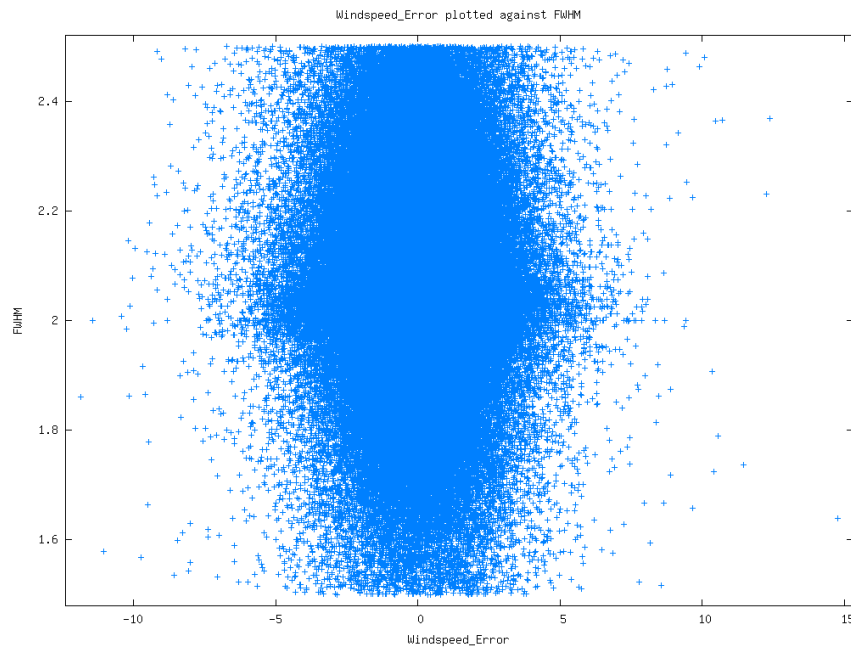


Figure 17: FWHM reported by the Mie Core algorithm as function of the windspeed error, for the selected CALIPSO data.

When plotted as contourplots a bit more detail becomes visible, see Fig. 18. Clearly the majority of all points are concentrated around FWHM=2.05.

When only points are selected close to FWHM=2.05, for example in the region as plotting in Fig. 18 (right panel) between 1.95 and 2.16, a slight beneficial effect can be seen on the windspeed error distribu-

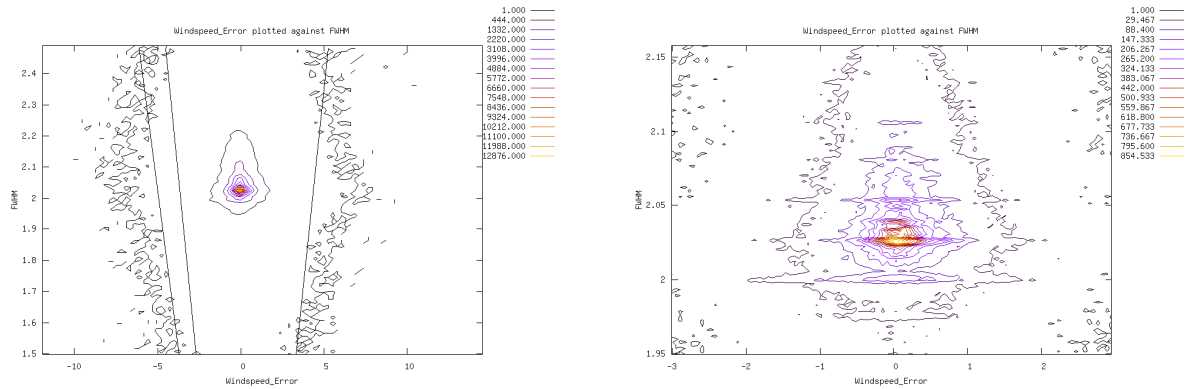


Figure 18: FWHM reported by the Mie Core algorithm as function of the windspeed error, plotted as contourplot. All valid results (left panel), and a magnified region around FWHM=2 (right panel). The straight lines in the left panel are plotting artefacts and should be ignored. Also the striping feature in the right panel is an artefact caused by the binning of the data and probably not relevant.

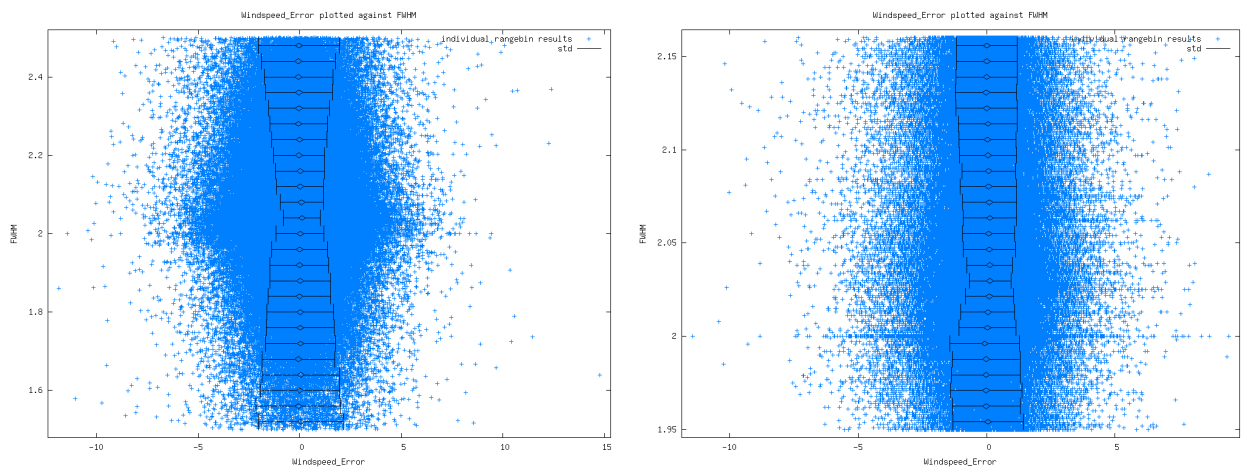


Figure 19: Same data as shown in Fig. 18 again as scatterplot, but now with overplotted lines to show the mean and one std value on both sides of the mean, for a number of FWHM layers (full range is split in 25 sub-ranges for this calculation).

tion¹². The windspeed error std of this distribution decreases from 1.19 to 1.06 m/s. This is at the cost of throwing away 126.000 data points (so 27% of all data flagged valid by the SNR test and the Validity flag).

Since from the contours in Fig. 18 it is not obvious to see the actual distribution of the data for a given FWHM, due to the variation of the number of datapoints, a new figure was made, see Fig. 19. In this scatterplot the mean and 1 standard deviation (STD) on both sides of the mean have been overplotted. The resulting windspeed error STD values vary between 0.91 and 2.09 m/s when using all valid data (left panel) and between 0.81 and 1.40 m/s when looking at the region $1.95 < \text{FWHM} < 2.16$. From this a threshold

¹²Note that taking 2.16 as boundary has been done on purpose here to prevent the gnuplot contouring routine to crash. This happens when 2.15 is taken as boundary, and is probably related to the binning artefacts that can be seen in Fig. 18 (right panel).

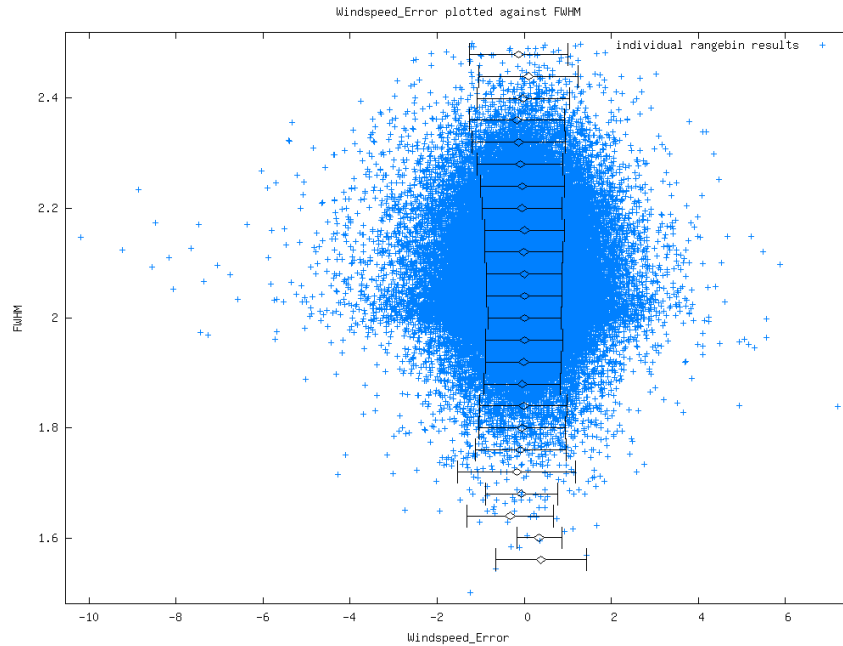


Figure 20: Scatterplot of FWHM against windspeed error for cases with ResErr<0.005. The overplotted lines show the mean and one std value on both sides of the mean, for a number of FWHM layers (full range is split in 25 sub-ranges for this calculation).

Table 3: Overview of effects of applying the different thresholds

Applied thresholds	nr of datapoints left	std in windspeed error
SNR>10, ValFlag set	469206	1.188 m/s
SNR>10, ValFlag set, ResErr<0.005	241674	0.748 m/s
SNR>10, ValFlag set, 1.80<FWHM<2.35	443868	1.145 m/s
SNR>10, ValFlag set, 2.015<FWHM<2.045	122124	0.885 m/s
SNR>10, ValFlag set, 1.80<FWHM<2.35, ResErr<0.005	240730	0.746 m/s
SNR>10, ValFlag set, 2.015<FWHM<2.045, ResErr<0.005	92118	0.622 m/s

can be defined using the 1 m/s values, as was done for the residual error. This results in the range 2.015 to 2.045.

When these numbers are recalculated on data with a residual error below 0.005 (so applying the residual error threshold as proposed in section 3.10.2) the result improves clearly, see Fig. 20. In this case the STD ranges from 0.86 to 1.13 and by taking a FWHM range of 1.80 up to 2.35, it seems possible to select data that has a STD in windspeed error of 1 m/s or smaller.

By applying the different thresholds the windspeed error can clearly be reduced. An overview of the results obtained is given in table 3. The residual error threshold clearly has the strongest effect, and throws away about 50% of all data. The wider FWHM thresholds have only minimal effect. The smaller FWHM thresholds can be used to reduce the error a bit further, but at the cost of discarding a larger fraction of about 75% of all data.

From the results presented above it can be seen that small improvements by applying a tight selection

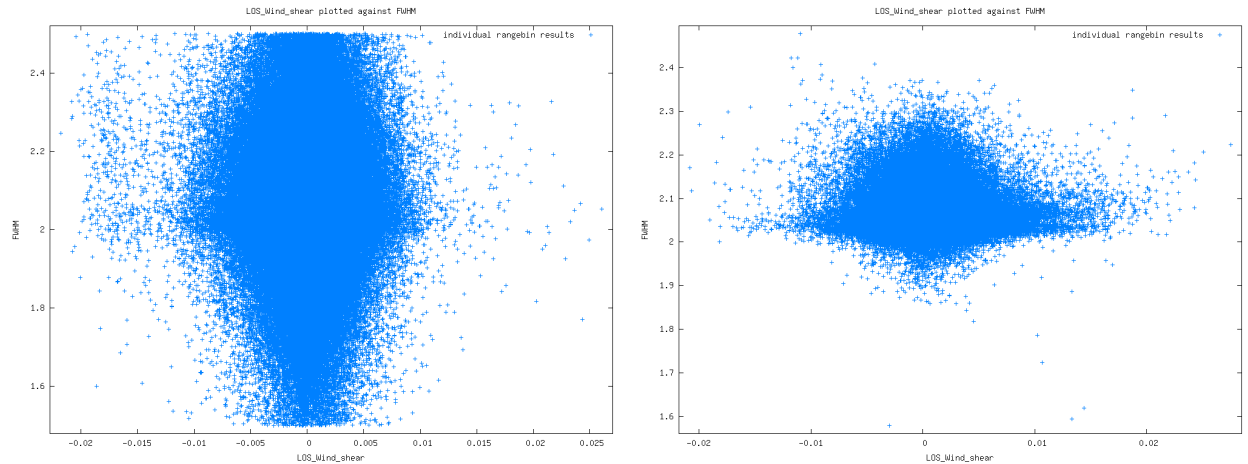


Figure 21: FWHM reported by the Mie Core algorithm as function of the wind shear in the input wind. The left plot selects data with SNR between 10 and 50, the right plot selects data with SNR above 50.

on the reported FWHM can be obtained. Especially when combined with a threshold on residual error.

A special case to be looked at is the occurrence of wind shear combined with transparent clouds or aerosol layers. Accumulation of different windspeeds may lead to significant broadening of the Mie peak in this special case. To make this visible FWHM has been plotted as a function of LOS wind shear used as simulation input in Fig. 21. This data shows a strong effect. Data with high SNR is clearly centered around FWHM=2.05, while data with lower SNR shows both much smaller FWHM values (due to noise) and larger FWHM values (due to noise and to broadening of the Mie peak).

3.10.4 Altitude effects

Most good Mie returns are expected in the PBL and on cloud tops. On the other hand problems due to windshear are also expected near the tropopause level. Therefore simulation properties will probably differ for different altitude classes. To study this the atmosphere has been divided into 3 regions, taking the WVM1 rangebin definition as basis.

- altitudes below 2.2 km (the region in which 500 m Mie rangebins are used)
- altitudes between 2.2 and 16.2 km (the region in which 1 km Mie rangebins are used)
- altitudes above 16.2 km (the region in which 2 km Mie rangebins are used)

The results of this exercise are given in table 4

From this table it is clear that the PBL gives much better results than the free troposphere or the UTLS. Since this was anticipated, and the wind requirements of the mission differ for these different regions, it might be useful to define separate thresholds for residual error and FWHM for each region.

3.11 Task 4a: QC conclusions and recommendations

In section 3.10 the quality of the wind as retrieved by the Mie Core algorithm has been investigated as function of several of the MieCore algorithm outputs. From this it can be concluded that:

- although not strictly a Mie Core output, screening on SNR>10 remains a requirement to obtain good fit results
- only data with a ValidityFlag set to true should be used
- a threshold on the residual error of 0.005 will select data in such a way that the overall STD of the windspeed error will be below the required 1 m/s. This flags about 50% of all data invalid.

Table 4: Overview of effects of applying the different thresholds on different regions of the atmosphere.

Results for Altitude below 2.2 km		
Applied thresholds	nr of datapoints left	std in windspeed error
SNR>10, ValFlag set	95766	0.799 m/s
SNR>10, ValFlag set, ResErr<0.005	73318	0.492 m/s
SNR>10, ValFlag set, 2.015<FWHM<2.045	38101	0.530 m/s
SNR>10, ValFlag set, 2.015<FWHM<2.045, ResErr<0.005	35272	0.408 m/s
Results for Altitude between 2.2 and 16.2 km		
Applied thresholds	nr of datapoints left	std in windspeed error
SNR>10, ValFlag set	282844	1.164 m/s
SNR>10, ValFlag set, ResErr<0.005	137746	0.735 m/s
SNR>10, ValFlag set, 2.015<FWHM<2.045	70009	0.901 m/s
SNR>10, ValFlag set, 2.015<FWHM<2.045, ResErr<0.005	51007	0.662 m/s
Results for Altitude above 16.2 km		
Applied thresholds	nr of datapoints left	std in windspeed error
SNR>10, ValFlag set	90596	1.540 m/s
SNR>10, ValFlag set, ResErr<0.005	30610	1.137 m/s
SNR>10, ValFlag set, 2.015<FWHM<2.045	14014	1.403 m/s
SNR>10, ValFlag set, 2.015<FWHM<2.045, ResErr<0.005	5839	1.071 m/s

- a threshold region on the FWHM of 2.015 upto 2.045, combined with the threshold on residual error, will select data in such a way that the overall STD of the windspeed error will be below the required 1 m/s.

The simulations clearly show different behaviour for different altitude regions. This because of the properties of the atmosphere, and because of the varying rangebin definition. Therefore The above mentioned thresholds may be finetuned depending on altitude.

4 Task 4a: Rayleigh QC

t.b.d. -investigate the Rayleigh line-width to be retrieved from the A and B signals. It is assumed the optical properties are well known to allow reconstruction of the peak height as well.

5 Task 4b: LIPAS upgrade

5.1 Background

LIPAS, the Lidar Performance Analysis Simulator, has been used for ADM simulation studies and impact experiments for many years. Its development started during the ADM pre-phase A period in the late 1990's. At that time, the selection for the laser wavelength was not definite and LIPAS was made flexible to cover many wavelengths ranging from the UV to the near infrared. Most preferred wavelengths included 0.355, 0.532, 1.064, 2.06 and 10.6 micron. This wide range of optional wavelengths required the simulation of both coherent and incoherent measurement techniques. Before the selection in 1999 as a Core Explorer Mission the wavelength of ADM-Aeolus was fixed at 0.355 micron and further LIPAS developments were dedicated to this wavelength. The algorithms for the simulation of the 0.355 micron combined Mie/Rayleigh system that were implemented in LIPAS were developed by Matra Marconi Systems and are found in [RD4], in the remainder denoted MMS. A description of the main characteristics of LIPAS is also found in [LR2]. LIPAS was used for the simulation of ADM and to assess its impact for NWP in many experiments including OSSE [LR3], assimilation ensemble experiments [LR1],[LR4] and SOSE [LR5], the latter mainly focusing on the impact of optional post-ADM scenarios.

Parallel to LIPAS other ADM hardware simulation tools were developed mainly the E2S [RD3] by Astrium-MDA and the DE2S by DLR. The E2S is currently set as the standard. However, as already explained in section 3.2 the E2S is programmed in Matlab and computationally not efficient. For the assessment of the vertical sampling scenarios, defined in [RD6], statistics from large observation samples need to be generated, i.e. ADM needs to be simulated for many orbits. It was concluded that the E2S is not suitable for this purpose. Alternatively it was decided to upgrade LIPAS and to make it compatible with the E2S-L1B-L2B chain. The next section discusses the LIPAS upgrade. In a subsequent section the upgraded LIPAS, in the remainder denoted LIPAS-2009, is used for the assessment of the vertical sampling scenarios, more in particular the quantification of their L2 HLOS information content for Mie only, Rayleigh only and collocated Mie and Rayleigh sampling.

5.2 From LIPAS to LIPAS-2009

Assumptions and simplifications made in LIPAS are reviewed in this section. Below is an overview of items that are reviewed in the context of the LIPAS upgrade. A more detailed description of the reviewed items is found in subsequent sections.

- LIPAS input was limited to NWP model field parameters. NWP model fields do not include aerosol. Aerosol backscatter and extinction were simulated from climatology datasets and a random process to simulate backscatter variability, e.g. [LR2]. NWP model fields also do not include cloud optical properties explicitly but implicitly through model parameters such as cloud cover and cloud liquid water and ice content. A cloud overlap model, that assumes maximum spatial correlation of layered clouds, and randomization were implemented to decide on cloud encounters on a shot-by-shot basis and parameterizations to convert cloud model parameters to cloud optical properties (backscatter and extinction), see [LR2]. These options are maintained in LIPAS-2009, but the upgrade LIPAS can read atmospheric profiles from the Level-2 atmospheric database in addition. This atmospheric database includes explicit aerosol and cloud backscatter and extinction.
- LIPAS simulation of the Fizeau interferometer, also denoted Mie channel. LIPAS does already include the simulation of the number of electrons on the ACCD bins of the Fizeau interferometer. A simple algorithm, described in chapter 7 of MMS, was implemented to retrieve the peak location that is related to the particles velocity. In the past a monte Carlo simulation was performed to generate a lookup table that related Mie channel SNR to LOS wind accuray. As part of the upgrade, the Fizeau interferometer parameters have been made compatible to the E2S and the peak-fit method is now more advanced than in MMS. The L1B and L2B processors include the so-called Mie Core algorithm to find the peak on the Fizeau bins. The Mie Core algorithm will be implemented in LIPAS-2009 and tested against the currently implemented peak-fit method.
- LIPAS simulation of the Fabry-Perot interferometer, also denoted Rayleigh channel. LIPAS did not include explicit simulation of the number of electrons detected on the Fabry-Perot ACCD bins. Instead,

Eq. (46) from MMS that relates Rayleigh receiver signal SNR to LOS wind accuracy was implemented. LIPAS-2009 explicitly simulates the numbers of electrons detected on the Fabry-Perot ACCD bins. It includes the simulation of the Fabry-Perot transfer function that maps the incoming atmospheric return signal on the ACCD bins. The receiver response is computed from the signals in the Fabry-Perot channels A and B and related to a receiver response lookup table to retrieve the Doppler shift and corresponding LOS wind velocity.

- LIPAS ignored cross-talk, i.e. contamination of the Rayleigh channel signal by particles in the measurement bin. It was assumed that this could be resolved by the L2Bp, once available. LIPAS-2009 will explicitly simulate cross-talk. This also paves the way for LIPAS-2009 to test L2Bp algorithms to correct for cross-talk. These algorithms are currently under investigation and development, for instance the L2Bp optical properties code.
- Initially LIPAS assumed homogeneous atmospheres and no wind biases in retrieved Mie and Rayleigh winds. Biases may be introduced by the atmosphere e.g. cross-talk and atmosphere heterogeneity as well as by the hardware. At a later stage, for MERCI [RD10], a post-processing was developed to introduce biases in retrieved winds related to atmospheric heterogeneity. LIPAS-2009 explicitly simulates atmosphere induced biases by subdividing the atmosphere in bins that are much smaller than the ADM vertical bins. For each of these bins the atmospheric return signal and spectrum are determined in addition to their distribution over the Fizeau and Fabry-Perot interferometer bins. Integrating the signals of these atmospheric bins to the ADM vertical bins followed by wind retrieval from the integrated signal will yield biases in case of heterogeneous atmospheres. Biases induced by instrument hardware will be ignored for the moment.

The following subsections describe the simulation of the atmospheric return signal in LIPAS-2009 and the propagation of the atmospheric return signal through the instrument detection chain. This includes the simulation of the number of detected electrons on the Fizeau and Fabry-Perot interferometers and wind retrieval from the measured data. The items described above are discussed in more detail in the corresponding sections.

5.2.1 Atmospheric return signal

This section describes the simulation of the atmospheric return signal that is detected on the instrument receiver. Each range gate of the vertical profile of ADM is subdivided in M equidistant subbins. LIPAS-2009 uses $M = 20$. The atmospheric return signal from particle (Mie) and molecular (Rayleigh) backscatter is simulated for each of these subbins. The spectral distribution for both Mie and Rayleigh signals is then expressed as follows, following MMS Eq. (3):

$$S(\lambda, z) = S_{\text{tot}}(\lambda, z) I_{\lambda_{\text{laser}}}(\lambda) \quad (10)$$

with $I_{\lambda_{\text{laser}}}(\lambda)$ the normalised spectral distribution of the signal in $[\text{pm}^{-1}]$ and S_{tot} the total signal in photons [ph]. The spectral distribution I is modelled as a Gaussian distribution for both Mie and Rayleigh signals, see also Eq. (3) of MMS i.e. the equivalent of Eqs. (1,2). Note that LIPAS uses units in [m] instead of [Hz]. The peak location is at the Doppler shifted wavelength $\lambda' = \lambda_{\text{laser}} + \nu$ with λ_{laser} the laser wavelength of 0.355 micron ν denotes the Doppler shift in [m] in the remainder. The latter is related directly to the LOS wind velocity within the sampled subbin. For each subbin the true atmospheric u (zonal), v (meridional) and w (vertical) wind components are obtained from the atmospheric database. Following Eq. (2) from [RD7], the true LOS wind velocity, v_{LOS} equals

$$v_{\text{LOS}} = [-u \sin(\psi) - v \cos(\psi)] \sin(\phi) - w \cos(\phi) \quad (11)$$

with ψ the azimuth angle defined as the angle (clockwise) between the laser beam direction and geographical north and ϕ the local incidence angle at the earth surface. Although the local incidence angle slightly changes with altitude along the profile LIPAS-2009 assumes it constant with a value of 37.56 degrees.

The Doppler shifted wavelength, λ' , is obtained from the frequency Doppler shifted, df (Hz), as follows

$$\lambda' = \lambda_{\text{laser}} - \frac{c}{\frac{c}{\lambda} + df} \quad (12)$$

with c the speed of light and df defined in Eq. (3) with s replaced by v_{LOS} defined in Eq. (11). Note that the second term in Eq. (12) is the Doppler shift ν in [m].

The peak location of the spectral distribution is identical for Mie and Rayleigh signals but the width of the peak at FWHM differs substantially. The Mie spectral shape is considered as a duplication of the laser pulse spectral shape with FWHM given in Eq. (1) of MMS, see also table 7. The Rayleigh spectral shape is much broader due to the Brownian motion of molecules. The FWHM of the molecular peak is a function of temperature and obtained from Eq. (2) of MMS. The Doppler shifted wavelength and the Mie/Rayleigh peak FWHM define the normalized spectral distribution of the backscattered Mie/Rayleigh signal I_{laser} .

The total signal at receiver input S_{tot} is obtained following Eq. (4,5) of MMS:

$$S_{\text{tot}}(\lambda, z) = \frac{E\lambda}{hc} \frac{\beta(\lambda, z)}{R(z)^2} \frac{\pi D^2}{4} \Delta h_{\text{LOS}} T_{\text{atm}}(z) T_{\text{RX}} T_{\text{TX}} \quad [ph/\text{range gate}] \quad (13)$$

with

- E is the laser pulse energy [J]
- λ is the laser wavelength [m]
- h is Plank constant [Js]
- c is the speed of light [ms^{-1}]
- $\beta(\lambda, z)$ is the backscatter coefficient [$\text{m}^{-1}\text{sr}^{-1}$] for both Mie and Rayleigh signals
- $R(z)$ is the range from satellite to the atmospheric layer [m]
- Δh_{LOS} is the range gate that equals $\Delta h / \cos(\phi)$ with ϕ the local incidence angle.
- $T_{\text{atm}}(z)$ is the 2-way atmospheric transmission between the satellite, at altitude h_{sat} and the observed atmospheric layer at altitude z and is determined as follows from the molecular extinction, α_m , and particle extinction, α_p .

$$T_{\text{atm}}(z) = \exp \left[\frac{-2}{\cos(\phi)} \int_z^{h_{\text{sat}}} \{\alpha_m(\lambda, z) + \alpha_p(\lambda, z)\} dz \right] \quad (14)$$

Extinction is the result of molecule and/or particle scattering and absorption.

- D is the telescope diameter [m]
- T_{TX} is the transmit optics transmission
- T_{RX} is the receive optics transmission

Values for these parameters used in the simulations are found in table 7. The following subsections describe the propagation of the atmospheric return signal through the instrument detection chain and the resulting signals on the Fizeau and Fabry-Perot interferometers.

5.2.2 Fizeau interferometer

Following Eqs. (9,11) of MMS, the distribution of the atmospheric return signal, described in the previous section, on the 16 spectral channels (also denoted ACCD bins) of the Fizeau interferometer is obtained from

$$\text{ACCD}(i) = \int_{\lambda_i}^{\lambda_i + \Delta\lambda} RS_{\lambda_{\text{laser}}}(\lambda) d\lambda \quad (15)$$

with RS the unitless spectral response

$$RS_{\lambda_{\text{laser}}}(\lambda) = \int_x T(x, \lambda) I_{\lambda_{\text{laser}}}(x) dx \quad (16)$$

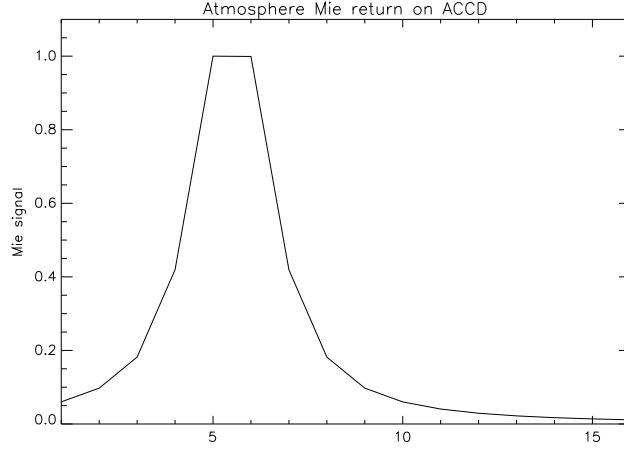


Figure 22: Distribution of atmosphere Mie peak on Fizeau interferometer ACCD bins for a 50 ms^{-1} wind velocity.

Here, T is the Fizeau spectral transmission that is modelled as a Lorentzian, see eq. (9), and with parameters defined in Table 5. $I_{\lambda_{\text{laser}}}$ is the normalized spectral distribution of the backscattered Mie signal, as discussed in the previous section. λ_i is the starting wavelength of channel i [pm] and $\Delta\lambda$ is the equivalent spectral sampling of the detector [pm], i is from 1 to 16. Instead of performing the double integration in Eqs. (15,16) for each atmospheric bin, a discretization similar as in the E2S is applied as follows

$$\text{ACCD}(i) \approx \sum_{j=1}^{N_j} RS(\lambda_j) \Delta\tilde{\lambda} \quad (17)$$

$$= \sum_{j=1}^{N_j} \sum_{k=1}^{N_k} T(x_k, \lambda_j) I(x_k) \Delta x \Delta\tilde{\lambda} \quad (18)$$

$$= \Delta\tilde{\lambda} \Delta x \sum_{k=1}^{N_k} I(x_k) \sum_{j=1}^{N_j} T(x_k, \lambda_j) \quad (19)$$

with $\lambda_j = \lambda_i + (j-1)\Delta\tilde{\lambda}$ and $\Delta\tilde{\lambda} = \Delta\lambda/N_j$. Typical values are $N_j = 10$ meaning that each ACCD bin is discretized in 10 subbins. This value is also used in the E2S. The useful spectral range (USR), denoted with x , is discretized in 1000 steps, i.e. $N_k = 1000$. Thus $x_k = -\text{USR}/2 + (k-1)\Delta x$ with $\Delta x = \text{USR}/N_k$. Eq. (19) follows from rearranging the summations of the equation above. The second summation in Eq. (19) is only a function of the Fizeau interferometer characteristics and independent from the atmosphere. For each x_k this summation is done only once in LIPAS-2009 and the result is saved in a lookup table. This procedure is performed for all 16 ACCD bins, yielding a lookup table matrix, \mathbf{F} , with dimensions $16 \times N_k$. The distribution of electrons, e , of the atmospheric return spectrum on the ACCD bins of the Fizeau interferometer is then obtained from

$$e_{\text{ACCD}}^{\text{Fiz}} = \mathbf{F} I_{\lambda_{\text{laser}}} \quad (20)$$

where $e_{\text{ACCD}}^{\text{Fiz}}$ and $I_{\lambda_{\text{laser}}}$ are vectors now of dimensions 16 and N_k respectively.

Figure 22 shows an example of the normalized distribution of a Mie peak over the 16 Fizeau bins, similar as in Figure 2 for a 50 ms^{-1} wind velocity. The useful spectral range (USR) is 1502 MHz or 266.605 ms^{-1} , meaning that the width of each ACCD bin corresponds to 16.66 ms^{-1} . The location of the center of the Mie peak on the ACCD is close to 5.5. Note that for zero wind velocity the peak is located exactly between bin 8 and 9 at 8.5. A shift of -3 indeed corresponds to a 50 ms^{-1} wind velocity from wind velocity = shift*USR/16. Also note that although the width of the Mie peak is relatively small, its signal is distributed over all bins and not limited to the bins close to the corresponding wind velocity. This is important for the peak-fit method that is discussed later.

Table 5: Parameter values used to simulate the Fizeau transmission function ,see also Table 1.

Parameter	value	unit
peak transmission	0.60	
FWHM	184	[MHz]
Useful Spectral Range (USR)	1502	[MHz]
Free Spectral Range (FSR)	2150	[MHz]
number of ACCD columns	16	

Mie channel Doppler shift estimate The previous section discussed the distribution of the atmospheric Mie return signal on the Fizeau interferometer for one atmospheric subbin. Next, LIPAS-2009 accumulates the signals from the M subbins to get the signal for one ADM range gate. The total number of detected electrons per range gate for a number of accumulated laser shot, N_s , on each of the 16 ACCD bins of the Fizeau interferometer from particles equals, following Eqs. (11) of MMS,

$$N_{\text{Fiz}}(i) = C_1^{e2s} \eta N_s S_{\text{tot.mie}} T_p \frac{1}{N} \frac{2}{\pi} \frac{1}{\Delta\lambda} e_{\text{ACCD}}^{\text{Fiz}}(i) \quad [e] \quad (21)$$

with

- C_1^{e2s} a constant parameter used to make the signal intensity compatible with the E2S. This is discussed later in section 5.2.4
- η the detector quantum efficiency, see table 5
- N_s is the number of accumulated shots
- $S_{\text{tot.mie}}$ is the total signal return from particle scattering per laser shot [ph], see Eq. (13)
- T_p is the peak transmission of the complete receiver
- N is the number of ACCD bins
- $2/\pi$ is the pupil truncation ratio
- $\Delta\lambda$ is the equivalent spectral sampling of the detector
- $e_{\text{ACCD}}^{\text{Fiz}}(i)$ is obtained from Eq. (20)

Values for these parameters used in the simulations are found in table 7. Besides the Mie return, electrons are detected from the Rayleigh return and the earth radiance background. For the former, Eq. (20) could be used, however LIPAS-2009 follows MMS by assuming a flat Rayleigh spectrum within the Fizeau USR. Similarly, the distribution of the earth radiance background on the Fizeau ACCD bins is assumed uniform. Following Eqs. (11,13,14) from MMS, the signal on each of the 16 ACCD bins of the Fizeau interferometer from molecules, $N_{\text{B.Ray}}$ and the earth radiance background, $N_{\text{B.BKG}}$ is obtained from

$$N_{\text{B.Ray}}(i) = C_2^{e2s} \eta N_s S_{\text{tot.ray}} \left[\frac{\Delta\lambda_{\text{eq.ray}} + \Delta\lambda}{\Delta\lambda_{\text{FWHM.ray}}} \right] \sqrt{\frac{\ln(2)}{\pi}} T_p \frac{2}{\pi N} \quad (22)$$

$$N_{\text{B.BKG}}(i) = \eta N_s S_{\text{BKG}} T_p \Delta\lambda_{\text{eq.bkg}} \frac{2}{\pi N} \quad (23)$$

with

- C_2^{e2s} is a constant parameter used to make the signal intensity compatible with the E2S, see section 5.2.4
- $S_{\text{tot.ray}}$ is the total signal return from molecular scattering per laser shot [ph], see Eq. (13)
- $\Delta\lambda_{\text{eq.ray}}$ is the Rayleigh equivalent bandwidth.

- $\lambda_{\text{FWHM_ray}}$ is the FWHM of the width of backscattered Rayleigh spectrum, see Eq. (2) of MMS.
- $\lambda_{\text{eq_bkg}}$ is the background equivalent bandwidth.
- S_{BKG} is the total earth radiance signal available at receiver input, see Eq. (6) in MMS.

Values for these parameters used in the simulations are found in table 7. The total signal per ACCD bin i for the Mie receiver is

$$N_{\text{tot}}(i) = N_{\text{Fiz}}(i) + N_{\text{B_Ray}}(i) + N_{\text{B_BKG}}(i) \quad (24)$$

The resulting signal N_{tot} corresponds to one range gate for a number of accumulated laser shots, called a *measurement*. The accumulation strategy is not fixed at the time of writing but typical values are $N_s = 14$ for a 1-kilometer and $N_s = 50$ for a 3.5-kilometer along track accumulation.

Assuming no advanced classification between aerosol and cloud measurements, the signals from all measurements along a 50 km track are accumulated to yield the signal for one ADM *observation*. Next, Poisson noise, dark current and the earth radiance background signal, following Eq. (26) from MMS are added to yield the noisy ACCD signal for one observation. The location of the center of the Mie peak on the ACCD corresponds to the LOS wind velocity. This procedure is done for all 24 ADM vertical range gates to yield the Mie channel LOS wind profile.

In chapter 7 of MMS a relative simple algorithm is proposed to estimate to Mie peak center location from the noisy signal on the 16 ACCD Fizeau bins. The spectrum on the Fizeau interferometer shows a peak, from the detected particles, on top of a flat spectrum, from the detected molecules. It is assumed that the Mie peak extends over 5 adjacent bins. The other bins contain noise only. A procedure is proposed to discriminate between bins containing signal plus noise and those containing noise only. From the latter an estimate of the noise level is obtained that is subtracted from the signals of all bins. Next, the bin with maximum signal is determined followed by a weighted average over 5 bins centered at the bin with maximum signal. This yields the location of the peak center. This procedure is very sensitive to the noise level estimate. In addition, the assumption that the Mie signal is found in 5 adjacent bins is too simplistic as observed from Fig. 22. These drawbacks make the method sensitive to wind biases. Improved algorithms to estimate the noise level and peak center, taking into account the (broad) distribution of the Mie signal over all 16 bins, have been implemented in LIPAS-2009 yielding better results, but still Mie retrievals by this simple algorithms are still not free of biases as will be showed later in the results.

It is expected that the Mie Core algorithm that was discussed in section 3.4 will yield better wind retrievals from the Mie channel. **This to To-be-done.**

The estimate of the LOS velocity, \hat{v}_{LOS} , within the measured ADM range gate is obtained from the estimated location of the center of the Mie peak, L_c , on the ACCD as follows

$$\hat{v}_{\text{LOS}} = -(L_c - 8.5)\text{USR}/16\text{eqios2hlos} \quad (25)$$

with USR the useful spectral range in ms^{-1} , e.g. 266.605 ms^{-1} for the value of 1502 GHz in table 1. The subsequent estimate for HLOS wind, \hat{v}_{HLOS} , is obtained from

$$\hat{v}_{\text{HLOS}} = \frac{\hat{v}_{\text{LOS}}}{\sin(\phi)} \quad (26)$$

5.2.3 Fabry-Perot transfer function

LIPAS did not include the explicit computation of the number of electrons on the 2 Fabry-Perot channels denoted A (direct) and B (reflected). Instead, the theoretical formula, Eq. (46) of MMS, that relates Fabry-Perot SNR to LOS wind velocity accuracy (standard deviation) was implemented. The LIPAS upgrade includes the explicit simulation of the signal on the A and B channels, similar as implemented in the E2S. This includes the computation of the signal on the 20 ACCD bins based on the modelling of the Fizeau reflection and transmission-reflection of the Fabry-Perot channel A and B, followed by the simulation of the focusing of both channel A signal and channel B signal on the ACCD. A more detailed explanation is found in [RD2].

Figure 23 shows the transmission curves of the direct (A), $\tau_{\text{FP}}^{\text{A}}(\nu)$, and reflected (B), $\tau_{\text{FP}}^{\text{B}}(\nu)$ channels as a function of the wavelength offset ν . Also the reflected signal from the direct channel, $\tau_{\text{FP}}^{\text{A-refl}}(\nu)$, is shown. The right panel shows the reflection transmission curve of the Fizeau interferometer, $\tau_{\text{Fiz}}(\nu)$.

Table 6: Parameter values used to simulate the Fabry-Perot transfer function.

Parameter	value	unit
simulated frequency range	-3 to +3	[GHz]
Number of frequency steps	1201	
peak transmission direct channel (T_p^A)	0.68	
peak transmission reflected channel (T_p^B)	0.61	
crossover wavelength	355	[nm]
peak separation wavelength	2.3	[pm]
FWHM direct channel	1666	[MHz]
FWHM reflected channel	1666	[MHz]
FSR direct channel	10950	[MHz]
FSR reflected channel	10950	[MHz]
beam spot centre pixel x -direction direct channel	8.5	[pxl]
beam spot centre pixel y -direction direct channel	4.5	[pxl]
beam spot centre pixel x -direction reflected channel	8.5	[pxl]
beam spot centre pixel y -direction reflected channel	12.5	[pxl]
beam spot diameter direct channel	3.3	[pxl]
beam spot diameter reflected channel	3.3	[pxl]
number of ACCD rows	16	
number of ACCD columns	16	

The Fabry-Perot transfer function, $\tau_{\text{FP}}(i, \nu)$, for ACCD bin i is the sum of the transfer functions of the direct and reflected channels. This includes the focusing of the channel A and B signals by the Relay optics, $\tau_{\text{FP}}^{\text{A-relay}}(i)$ and $\tau_{\text{FP}}^{\text{B-relay}}(i)$ respectively, that are displayed in the left panel of Figure 24. The Fabry-Perot transfer function is then written as

$$\tau_{\text{FP}}(i, \nu) = T_p^A \tau_{\text{Fiz}}(\nu) \tau_{\text{FP}}^{\text{A}}(\nu) \tau_{\text{FP}}^{\text{A-relay}}(i) + T_p^B \tau_{\text{Fiz}}(\nu) \tau_{\text{FP}}^{\text{A-refl}}(\nu) \tau_{\text{FP}}^{\text{B}}(\nu) \tau_{\text{FP}}^{\text{B-relay}}(i) \quad (27)$$

The first term on the right hand side corresponds to the transfer function of the direct channel, the second term to the transfer function of the reflected channel. The peak transmission parameters T_p^A and T_p^B are found in Table 6. Integrating $\tau_{\text{FP}}(i, \nu)$ for ν yields the plot in the right panel of Figure 24 and shows that the Fabry-Perot direct channel A gathers more signal than the reflected channel B.

The Fabry-Perot transfer function is computed once in LIPAS-2009 by discretizing the frequency (or wavelength) range in 1201 steps, see Table 6. The resulting lookup table matrix of dimensions 16×1201 is saved. The explicit simulation of the A and B channels enables the simulation of cross-talk, as discussed in section 5.2.3, that was ignored in MMS.

LIPAS-2009 simulates the Fabry-Perot signal on the 16 ACCD bins for one ADM range gate similar as for the Fizeau interferometer by discretizing the atmosphere in M subbins, computing the contribution the atmospheric contribution on the ACCD for each subbin and then accumulating the signals from all subbins to get the total signal for one ADM range gate.

The distribution of atmosphere Rayleigh return signal on the Fabry-Perot interferometer bins is obtained from

$$e_{\text{ACCD}}^{\text{FP}}(i) = \int_{\nu} \tau_{\text{FP}}(i, \nu) I_{\lambda_{\text{laser}}}(\nu) d\nu \quad (28)$$

with $\tau_{\text{FP}}(i, \nu)$ obtained from Eq. (27). Note that here $I_{\lambda_{\text{laser}}}(\nu)$ is the broad Rayleigh spectrum as defined in Eq. (10)

Following Eqs. (17,18) from MMS we have two contributors to the signal on each of the 16 ACCD bins; N_{FP} from detected molecules and $N_{\text{B.BKG}}$ from the earth radiance background:

$$N_{\text{FP}}(i) = C_3^{e2s} \eta N_s S_{\text{tot.ray}} T_p \frac{2}{\pi} e_{\text{ACCD}}^{\text{FP}}(i) \quad (29)$$

$$N_{\text{B.BKG}}(i) = \eta N_s S_{\text{BKG}} T_p \Delta \lambda_{\text{eq.bkg}} \frac{2}{\pi N} \quad (30)$$

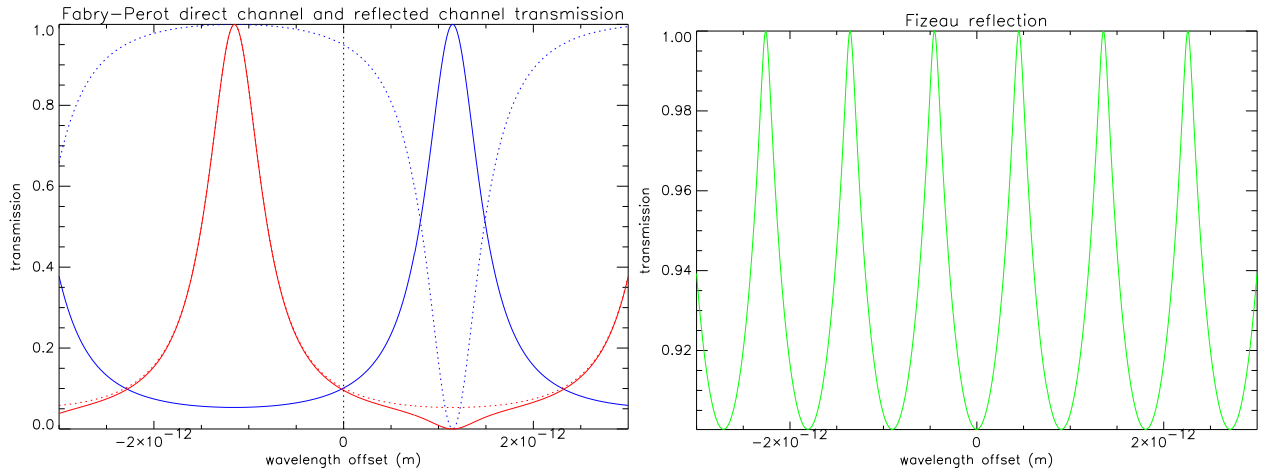


Figure 23: Fabry-Perot channel A (direct, blue) and B (reflected, red) transmission (left panel) and Fizeau reflection, $\tau_{\text{Fize}}(\nu)$, (right panel). The blue solid and red dotted curves show the channel A, $\tau_{\text{FP}}^{\text{A}}(\nu)$, and B, $\tau_{\text{FP}}^{\text{B}}(\nu)$, transmission respectively as a function of wavelength offset. The blue dotted line is the reflected signal from the direct channel, $\tau_{\text{FP}}^{\text{A-refl}}(\nu)$. The solid red line is the product of the channel A reflected transmission and the channel B transmission.

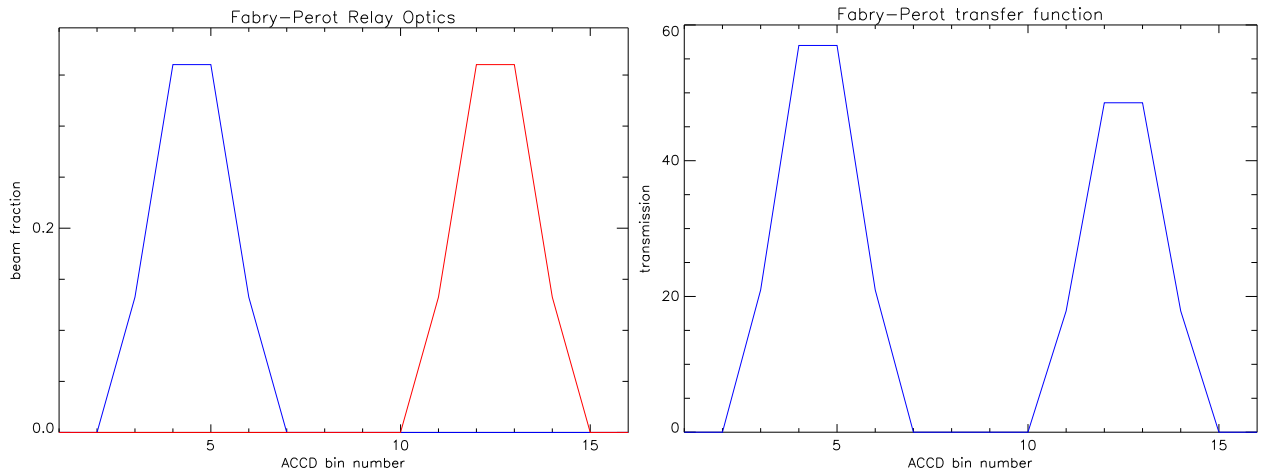


Figure 24: Fabry-Perot Relay to focus the channel A, $\tau_{\text{FP}}^{\text{A-relay}}$ and B, $\tau_{\text{FP}}^{\text{B-relay}}$, signals on the ACCD bins (left panel). The right panel shows the relative signal detected from the A (direct) and B (reflected) channel.

Note. I think I made an error in the equations above. I am not sure if the $2/\pi$ should be there. I'll check this. For the results later in the report this is no problem because C_3^{e2s} corrects for this bug.

Cross-talk Particles in the measurement bin give a signal on the Fabry-Perot ACCD channels. This contamination is called cross-talk. The peak of the Mie spectrum is much smaller than the broad tails of the Fabry-Perot transfer function. LIPAS-2009 therefore simulates the Mie peak as a delta-dirac function. Defining the Doppler shifted wavelength by $\hat{\nu}$ the distribution of the Mie peak over the Fabry-Perot ACCD bins then follows from Eq. (28):

$$e_{\text{ACCD}}^{\text{FP-Xtalk}}(i) = \int_{\nu} \tau_{\text{FP}}(i, \nu) I_{\lambda_{\text{laser}}}(\nu) d\nu \quad (31)$$

$$= \int_{\nu} \tau_{\text{FP}}(i, \nu) \delta(\hat{\nu}) d\nu \quad (32)$$

$$= \tau_{\text{FP}}(i, \hat{\nu}) \quad (33)$$

Note that $I_{\lambda_{\text{laser}}}(\nu)$ now relates to the small Mie peak. LIPAS-2009 applies a linear interpolation to the lookup table to determine $\tau_{\text{FP}}(i, \hat{\nu})$. The total signal from cross-talk is then obtained from

$$N_{\text{Xtalk}}(i) = C_4^{e2s} \eta N_s S_{\text{tot_mie}} T_p \frac{2}{\pi} e_{\text{ACCD}}^{\text{FP-Xtalk}}(i) \quad (34)$$

with $i=1, \dots, 16$. **Note.** Similar as the note above. I think I made an error in the equations above. I am not sure if the $2/\pi$ should be there. I'll check this. Again, for the results later in the report this is no problem because C_4^{e2s} corrects for this bug.

Rayleigh channel Doppler shift estimate The total signal on each of the 16 Fabry-Perot ACCD bins is obtained from

$$N_{\text{tot}}(i) = N_{\text{FP}}(i) + N_{\text{Xtalk}}(i) + N_{\text{B.BKG}}(i) \quad (35)$$

Next, Poisson noise, dark current and the earth radiance background signal are added, following Eq. (28) from MMS, to yield the noisy ACCD signal, $N'_{\text{tot}}(i)$, for one observation. The total Rayleigh signal on the Fabry-Perot direct channel is denoted A and the total signal on the reflected channel is denoted B , i.e.

$$A = \sum_{i=1}^8 N'_{\text{tot}}(i), \quad B = \sum_{i=9}^{16} N'_{\text{tot}}(i) \quad (36)$$

The receiver response, R , is defined as

$$R = \frac{A - B}{A + B} \quad (37)$$

The receiver response is a function of temperature, T , that determines the width of the molecular spectrum, and the Doppler shifted frequency ν in [m], that is related to mean wind velocity inside the sampled volume, i.e. $R = R(T, \nu)$.

LIPAS-2009 first generates a lookup table for $R(T, \nu)$. This is done only once and the table is saved. If Raleigh channel input parameters change, then the Fabry-Perot transfer function changes and a new lookup table must be generated. The temperature range is from 170-340 Kelvin with 1 degree steps. The Doppler shift range corresponds to LOS wind velocities in the range -150 to +150 ms^{-1} with 0.1 ms^{-1} steps. For values of temperature and Doppler shift frequency in the ranges defined above the signals on A and B and subsequent R are computed. The resulting receiver response lookup table has dimensions 171×3001 . Figure 25 shows an example of the receiver response function for one specific temperature value.

To retrieve the Doppler shifted frequency ν from the measured R and the lookup table the following procedure, that mimics the L2Bp, is followed. Similar as for the L2Bp a temperature estimate is extracted from the auxiliary meteorological file. The closest table value, T_0 is determined. Given T_0 the value ν_0 is determined such that $R(T_0, \nu_0)$ is closest to the measured $R(T, \nu)$. A first order estimate for $R(T, \nu)$ is obtained from

$$R(T, \nu) \approx R(T_0, \nu_0) + \frac{\partial R}{\partial T}(T_0)(T - T_0) + \frac{\partial R}{\partial \nu}(\nu_0)(\nu - \nu_0) \quad (38)$$

From this equation we can write for ν :

$$\nu = \nu_0 + \frac{R(T, \nu) - R(T_0, \nu_0) - \frac{\partial R}{\partial T}(T_0)(T - T_0)}{\frac{\partial R}{\partial \nu}(\nu_0)} \quad (39)$$

Denoting the indices of T_0 and ν_0 in the R matrix with i and j respectively, then the partial derivatives can be approximated through

$$\frac{\partial R}{\partial T}(T_0) \approx \frac{R(i+1) - R(i)}{T(i+1) - T(i)} \quad (40)$$

$$\frac{\partial R}{\partial \nu}(\nu_0) \approx \frac{R(j+1) - R(j)}{\nu(j+1) - \nu(j)} \quad (41)$$

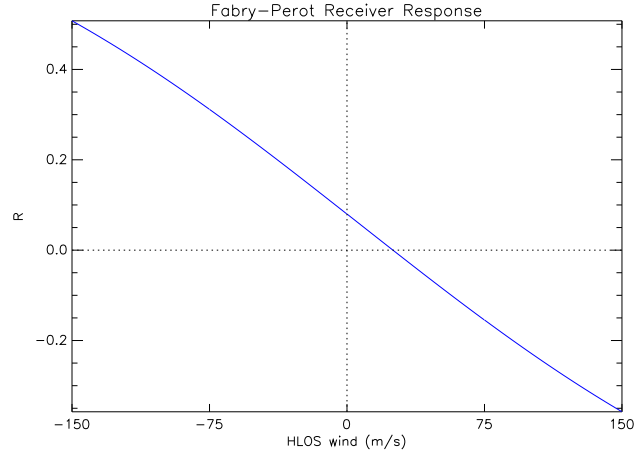


Figure 25: Receiver response function for $T = 273$ (K) as a function of wind velocity in ms^{-1} .

The parameters on the right hand side of the equations above are obtained from the lookup table. Substituting the result in Eq. (39) yields the Doppler shift ν in [m]. The LOS wind component velocity is subsequently obtained from

$$\hat{v}_{\text{LOS}} = \frac{-\nu c}{2\lambda} \quad (42)$$

The HLOS wind component velocity then follows from Eq. (26).

5.2.4 Results

The simulation of the signal on the Fizeau and Fabry-Perot interferometers in LIPAS-2009 described above are compared with the E2S for a couple of cases.

case 1 This case is based on a with a constant temperature, a constant HLOS wind of 0 ms^{-1} and the RMA aerosol backscatter profile. No instrument noise was applied. Figure 26 shows the signal on the 16 ACCD bins of the Fizeau interferometer for 24 range gates (in the figure denoted bins). The vertical sampling is the WVM1 rangebin definition as defined in [RD6] section 6.1.1. Range gate 23 contains the ground return. The E2S result looks noisy feature although the noise flag was explicitly turned off. Comparing to the lower right earth radiance background signal, it appears that all E2S bins are contaminated with this signal. To get similar amplitudes for the E2S and LIPAS-2009 signal the following correction parameter were applied: $C_1^{\text{e2s}} = 0.8$, $C_2^{\text{e2s}} = 1.9$, $C_3^{\text{e2s}} = 2.0$, $C_4^{\text{e2s}} = 2.0$. Figure 27 shows the signal on the 16 ACCD bins of the Fabry-Perot interferometer for 24 range gates. Range gate 24 contains the ground return.

case 2 This case is the same as case 1 but now the temperature is not constant but obtained from the scene denoted single_RMA_profile_midlat_winter_MF in the atmospheric database. A constant HLOS wind of -50 ms^{-1} was imposed. In addition an optically very thin cloud between 12 and 14 km was included. Figure 28 is similar as Figure 26 but the cloud clear shows up in bins 7-9. Also the peak center is shifted due to the non-zero wind velocity. Figure 29 shows the signal on the 16 ACCD bins of the Fabry-Perot interferometer for 24 range gates. Range gate 24 contains the ground return.

LIPAS-2009 Mie/Rayleigh wind retrieval The results in the previous subsections show that a correct setting of the correction parameters C_i^{e2s} make LIPAS-2009 signals on the Fizeau and Fabry-Perot interferometers very similar to the E2S. Next, the statistics of retrieved Mie and Rayleigh winds were determined for a couple of cases. The results for 2 cases are presented here.

case 1

This case is based on a synthetic atmosphere with a constant temperature, a constant LOS wind of 0 ms^{-1}

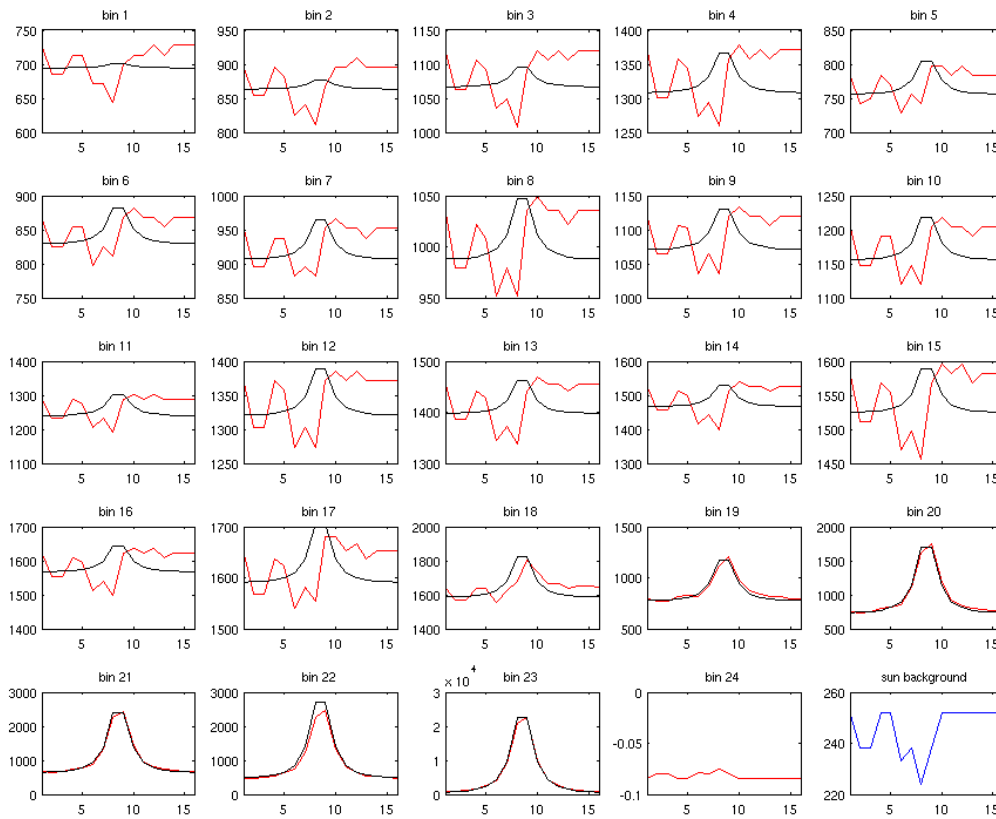


Figure 26: Case 1. Signal on the 16 ACCD bins of the Fizeau interferometer for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. The blue curve in the lower right panel is the E2S earth radiance background signal. See the text for further details.

and no aerosols, see Figures 30, 31. The purpose of this test was to check whether LIPAS-2009 could perfectly retrieve the input wind from the Rayleigh channel (no cross-talk, no interpolation of receiver response lookup table). A total of 100000 simulation were performed for this case, each with a different noise realization in the computation of the ACCD signals. Figure 32 shows the Mie and Rayleigh wind retrieval without and with instrument noise. The Rayleigh wind retrieval for a noise-free instrument in the left panel is about perfect as expected because there is no cross-talk (no aerosols) and no interpolation needed to retrieve the Doppler shift from the receiver response lookup table (constant temperature and zero wind). The ground return yields a zero Mie wind. The Rayleigh wind at the ground is not exactly zero but very close (about 0.012 ms^{-1}). Including instrument noise shows Mie winds scattered all over the place in the right panel as expected because of missing aerosols in the atmosphere. Only the ground return gives a correct Mie wind estimate. The Rayleigh winds scatter around the true wind. Figure 33 shows a small bias for Rayleigh winds. The bias in Mie winds is larger because of missing aerosols. The Mie ground return standard deviation is close to zero. The standard deviation of Rayleigh winds is between 1.75 and 3 ms^{-1} .

case 2

This case is similar to the previous case ut now with a non-constant temperature, a constant LOS wind of 50 ms^{-1} , aerosols and an optically thin cloud layer between 12 and 14 km, see Figure 35. The purpose of this test was to test the impact of cross-talk on the retrieved Rayleigh winds and the quality of Mie

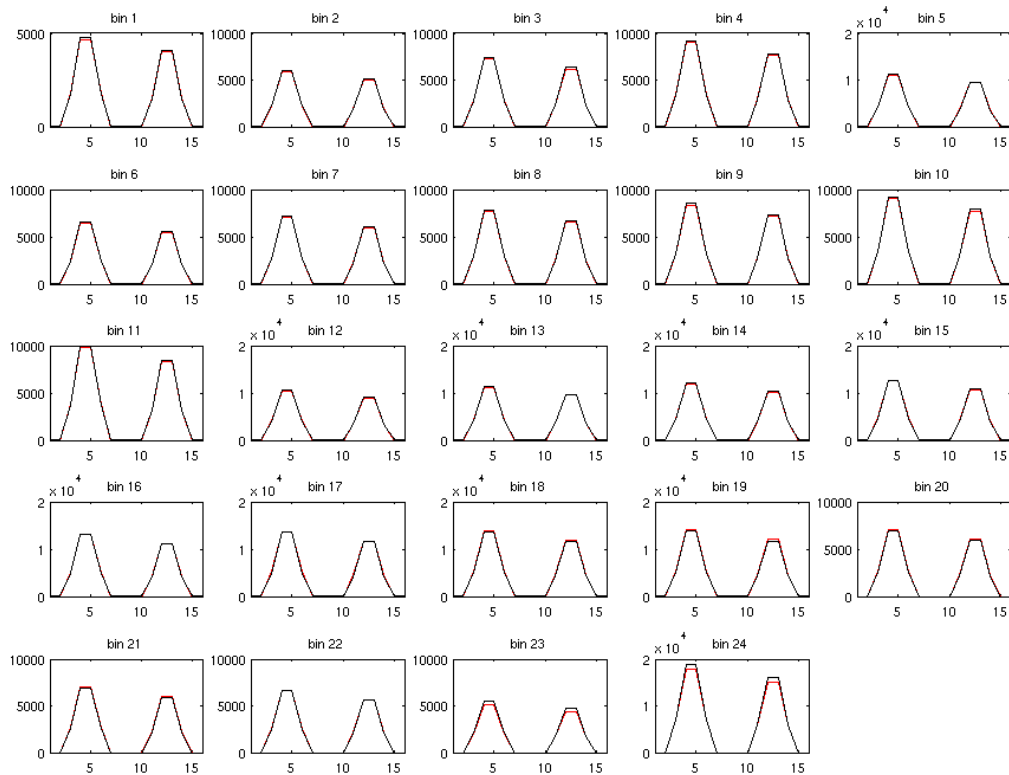


Figure 27: Case 1. Signal on the 16 ACCD bins of the Fabry-Perot interferometer for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. See the text for further details.

winds in case of aerosols and cloud. A total of 100000 simulation were performed for this case, each with a different noise realization in the computation of the ACCD signals. Figure 35 shows good quality (standard deviation) Mie winds only in the first kilometer above the surface and inside the cloud layer. However, the bias of Mie winds in the cloud layer is substantial. This is explained by the relatively simple retrieval algorithm described in section 5.2.2. It is expected that the Mie core algorithm will provide better wind retrievals. Also, the ground return show a bias of about 1 ms^{-1} most probably caused by aerosols with large wind speeds near the surface. The Rayleigh winds show generally good quality, although a bias up to 0.5 ms^{-1} is observed in the PBL due to cross-talk. Also a bias of about 5.6 ms^{-1} is observed inside the cloud layer.

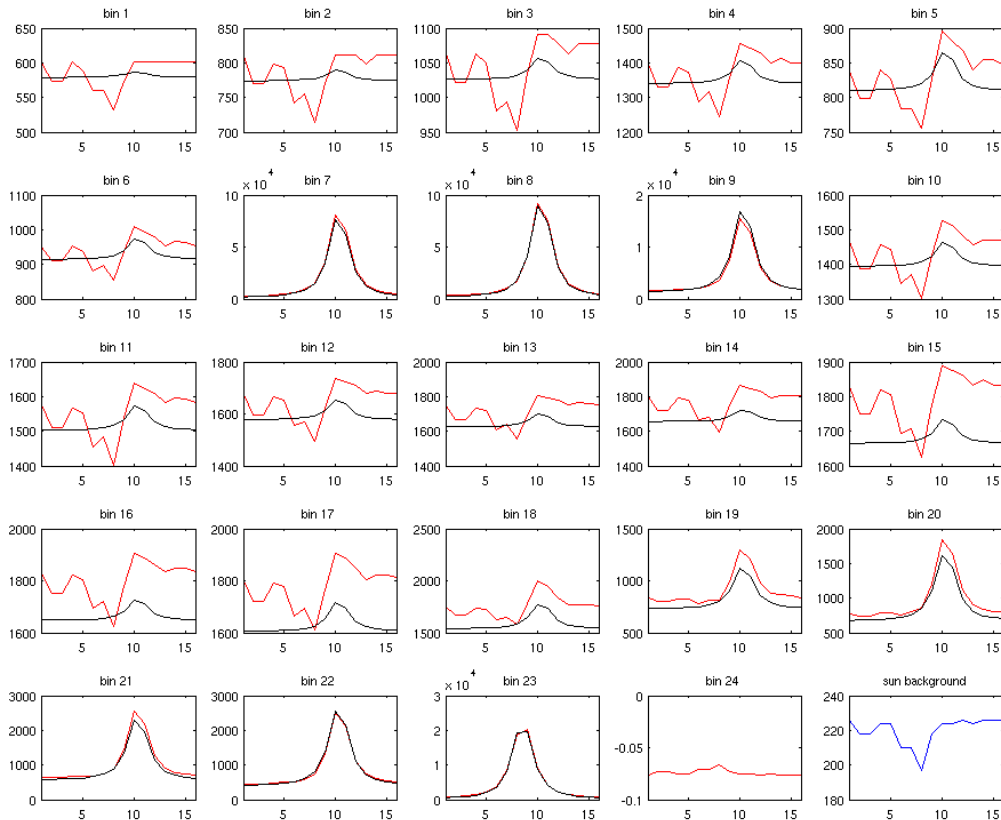


Figure 28: Case 2. Signal on the 16 ACCD bins of the Fizeau interferometer for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. The blue curve in the lower right panel is the E2S earth radiance background signal. See the text for further details. Bins 8-10 clearly show an increased signal due to cross-talk from cloud particles.

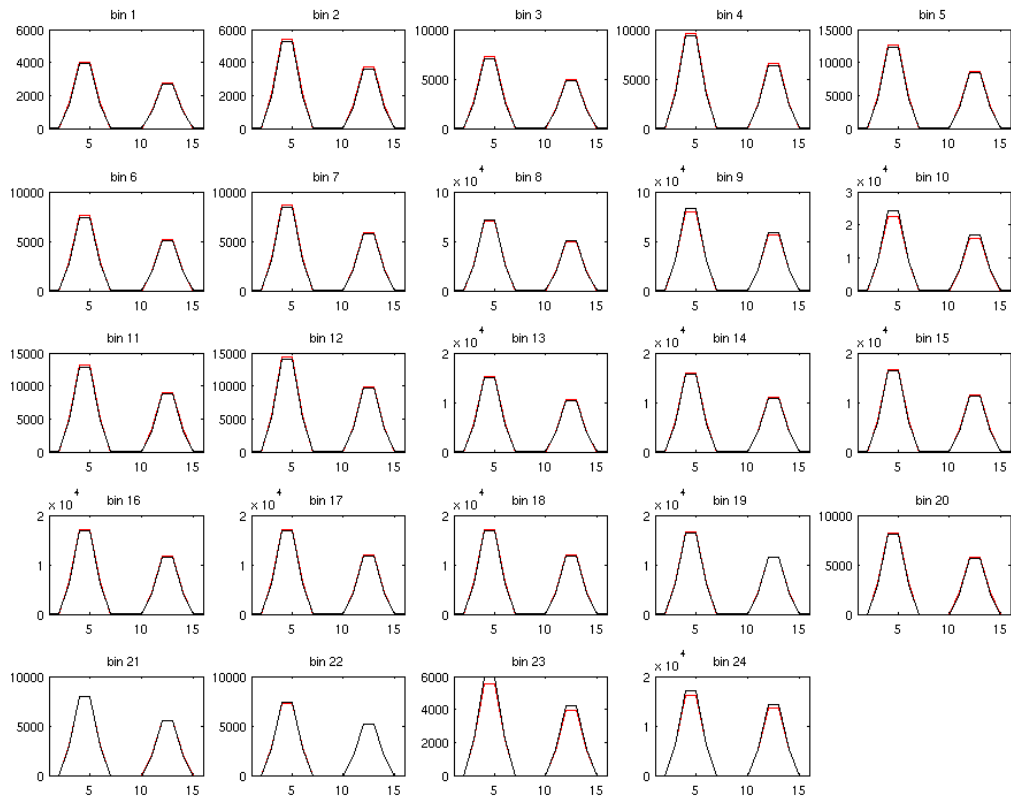


Figure 29: Case 2. Signal on the 16 ACCD bins of the Fabry-Perot interferometer for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. See the text for further details.

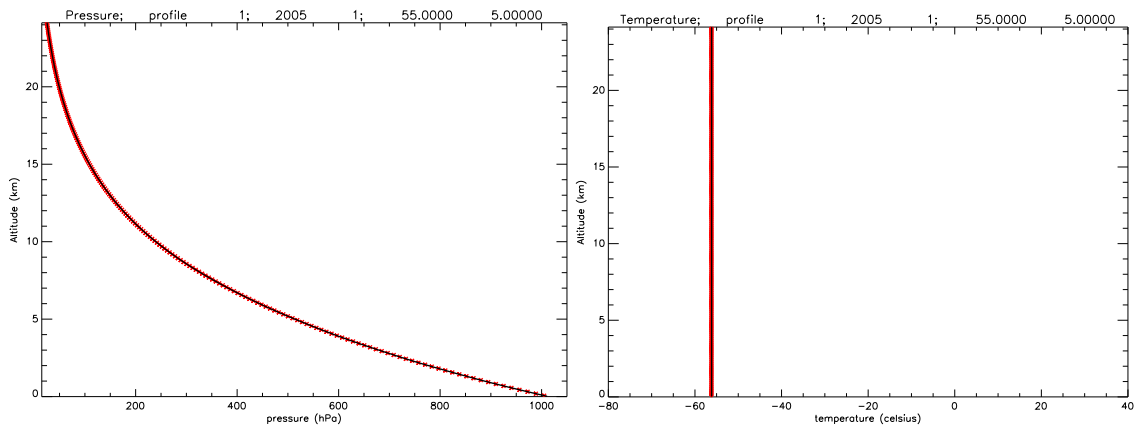


Figure 30: Atmospheric pressure (left) and temperature (right)

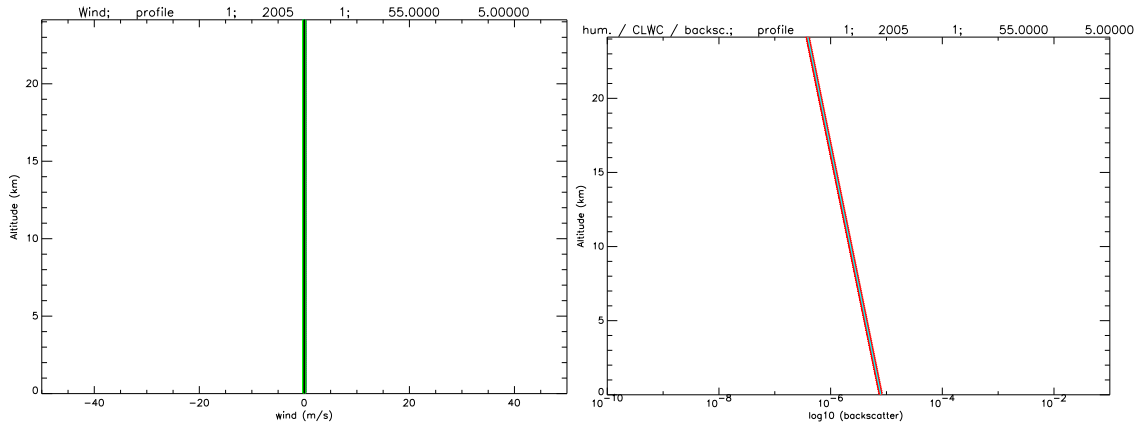


Figure 31: Atmospheric wind (left) and molecular backscatter (right). There is no aerosol backscatter.

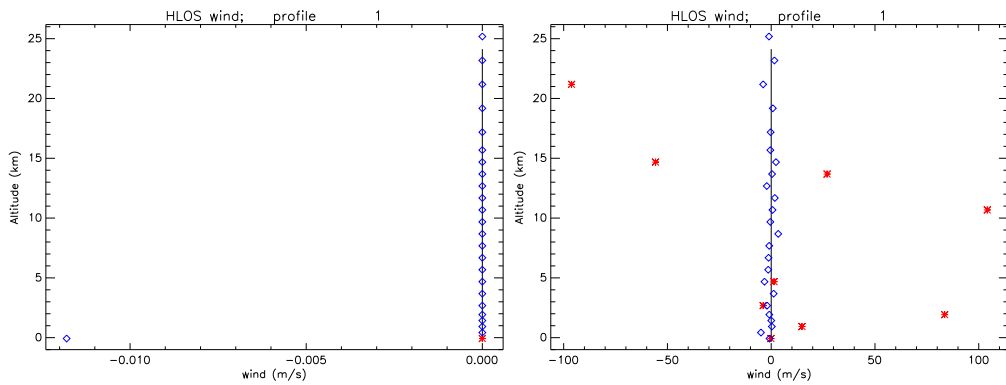


Figure 32: Single case HLOS wind retrieval from Rayleigh (blue diamond) and Mie (red star) channel. The left panel shows the result for a noise-free instrument, the right panel does include instrument noise.

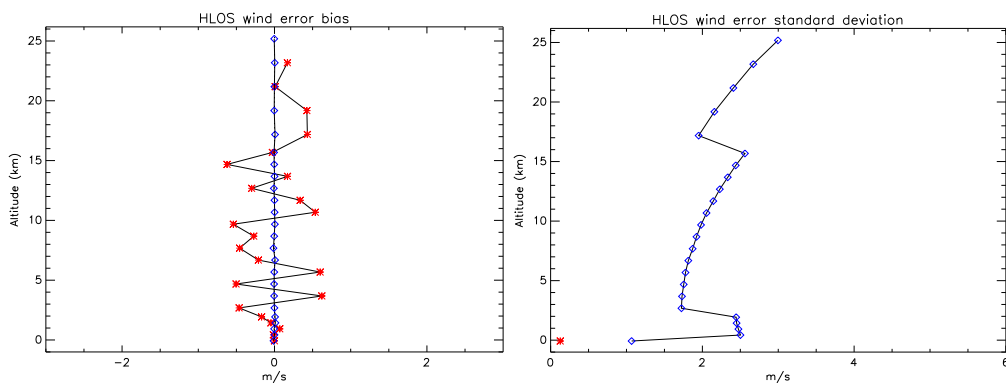


Figure 33: HLOS wind retrieval statistics for Rayleigh (blue diamond) and Mie (red star) channel from 100000 simulations. The left panel shows the bias, the right panel the standard deviation. The standard deviation for Mie winds above the surface is close to 85 ms^{-1} (not shown)

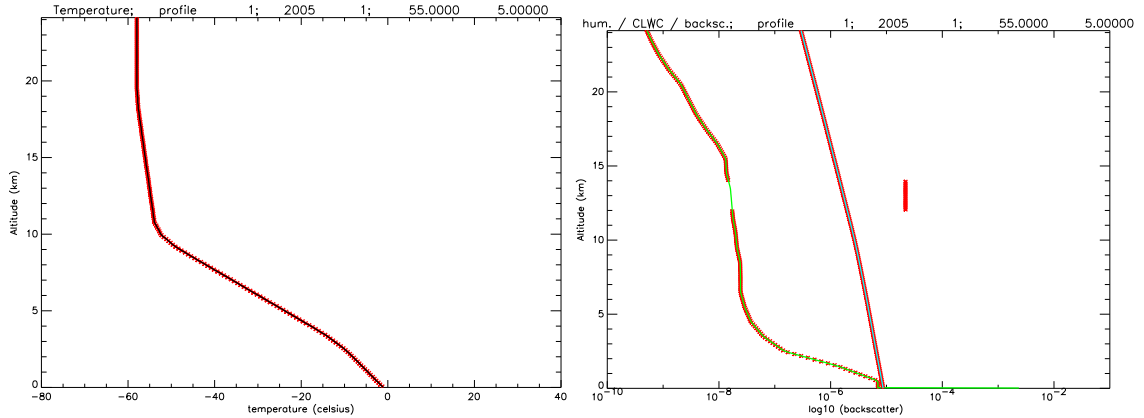


Figure 34: Atmospheric temperature (left) and aerosol (right). The left curve/middle/right shows aerosol/molecular/cloud backscatter.

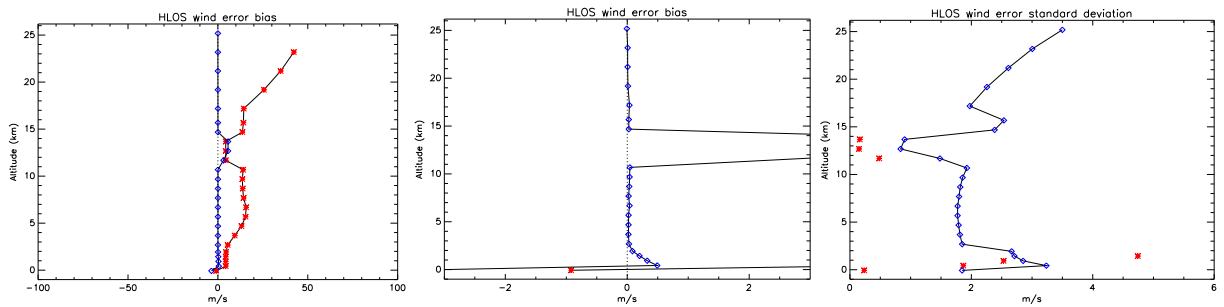


Figure 35: Mie (red) and Rayleigh (blue) HLOS wind retrieval statistics from 100000 simulations. The left/middle (zoomed in) panel show the bias, the right panel the standard deviation.

Table 7: Parameter values used for the ADM performance simulation.

Parameter description	symbol	value	unit
Emitter laser			
pulse repetition frequency		100	[Hz]
linewidth (FWHM)		50	[MHz]
laser pulse energy	E	0.12	[J]
Optics			
transmit optics transmission	T_{TX}	0.66	
receive optics transmission	T_{RX}	0.42	
telescope diameter	D	1.5	[m]
Detector/Receiver			
detector quantum efficiency	η	0.82	
peak transmission of the complete receiver	T_p	0.6	
equivalent spectral sampling of the detector	$\Delta\lambda$	0.0394	[pm]
Rayleigh equivalent bandwidth	$\Delta\lambda_{eq_ray}$	0.044	[pm]
background equivalent bandwidth	λ_{eq_bkg}	4.3	[pm]

6 Task 4b: Quantification of the L2 HLOS information content for Mie only, Rayleigh only and collocated Mie and Rayleigh sampling

6.1 Assessment of vertical sampling scenarios

7 Task 4c: Sampling scenarios

t.b.d.

8 Conclusion

to be written.

- end of document -