

ADM-Aeolus, VAMP, Task 4: Quantification of L2B HLOS wind accuracies for typical wind shear, aerosol and cloud conditions

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: 1.0
issue date: 02-Dec-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
0.4	30-Jun-2009	Adapted to ESA comments of 03-Jun-2009 plus section 4b extension
0.5	13-Jul-2009	Updated section 3.10.4 (Altitude effects) and written section 3.10.5 (Effect of rangebin size)
0.6	07-Aug-2009	Adapted to ESA comments of 14-Jul-2009, and added Mie Core inner loop results.
1.0	02-Dec-2009	Adapted to ESA comments of 8-Oct-2009.

Contents

Contents	3
List of Tables	5
List of Figures	11
1 General introduction	12
2 Documents and acronyms	13
2.1 Applicable documents	13
2.2 Reference documents	13
2.3 Literature	14
2.4 Acronyms	15
2.5 Document preparation	15
3 Task 4a: Mie Core QC	16
3.1 Introduction	16
3.2 Setup of the tool	16
3.3 Simulation of transmission and spectrum	18
3.4 Estimation of Signal-to-Noise ratio	20
3.5 Mie Core Algorithm	21
3.6 Input data	22
3.7 Error multiplier	23
3.8 Horizontal gradients and turbulence	24
3.9 Results: LITE data; signal scaling	25
3.10 Results: CALIPSO data; data selection	28
3.11 Results: CALIPSO data; Mie Core outputs	28
3.11.1 Outer loop iterations	28
3.11.2 Inner loop iterations	30
3.11.3 Residual error	32
3.11.4 FWHM	34
3.11.5 Altitude effects	37
3.11.6 Effect of rangebin size	39
3.12 Task 4a: QC conclusions and recommendations	40
4 Task 4a: Rayleigh QC	42
4.1 Determining the impact of vertical shear in horizontal winds	42
4.2 Determining the impact of horizontal variation in horizontal winds	42
5 Task 4b: LIPAS upgrade	43
5.1 Background	43
5.2 From LIPAS to LIPAS-2009	43
5.3 Atmospheric return signal	45
5.4 Mie channel signal and wind retrieval	46
5.4.1 Mie channel Doppler shift estimate	48
5.5 Rayleigh channel signal and wind retrieval	50
5.5.1 Fabry-Perot transfer function	50
5.5.2 Rayleigh channel Doppler shift estimate	52
5.5.3 Cross-talk	54
5.6 Results	56
5.6.1 LIPAS-2009 E2S comparison; case 1 – cloud-free atmosphere	56
5.6.2 LIPAS-2009 E2S comparison; case 2 – single cloud layer	58
5.6.3 LIPAS-2009 Mie/Rayleigh wind retrieval	60
5.6.4 LIPAS-2009 E2S-L1B-L2B comparison; photon count noise	64

5.7	Wind observation representativeness error	65
5.7.1	Atmosphere induced representativeness error	67
5.7.2	Model representativeness error	69
5.7.3	HLOS wind error overview	72
5.8	Summary and conclusions	72
6	Task 4b: Quantification of the L2 HLOS information content for Mie only, Rayleigh only and collocated Mie and Rayleigh sampling	74
6.1	Introduction	74
6.2	Mie and Rayleigh channel quality control	74
6.2.1	Mie channel quality control	74
6.2.2	Rayleigh channel quality control	74
6.3	HLOS wind retrieval	75
7	Task 4c: Vertical sampling scenarios information content	87
7.0.1	PBL wind information content	87
7.0.2	Lower-troposphere wind information content	89
7.0.3	Upper-troposphere lower-stratosphere wind information content	91
7.0.4	Stratosphere wind information content	93
7.0.5	Total wind information content	95
7.0.6	Wind profile information content	97
8	Task 4c: Vertical sampling scenarios information content - new scenarios	99
8.0.7	PBL wind information content	99
8.0.8	Lower-troposphere wind information content	101
8.0.9	Upper-troposphere lower-stratosphere wind information content	103
8.0.10	Stratosphere wind information content	105
8.0.11	HLOS wind quality in tropical cirrus	107
8.0.12	Total wind information content	109
8.1	Summary and conclusions	109
A	Vertical sampling scenarios	111
A.1	Vertical sampling scenarios with 500 m Rayleigh bins near the surface	111
A.2	Vertical sampling scenarios with 1000 m Rayleigh bins near the surface	114
A.3	Vertical sampling scenarios with 1000 m Rayleigh bins near the surface and maximum bin height below 31 km	116
A	HLOS wind statistics	118
B	Task 4c: Sampling scenarios	128
C	Conclusion	129

List of Tables

1	Parameter values used to simulate the Mie spectral peak, and the Fizeau transmission curves.	18
2	Mie Core algorithm settings.	22
3	Overview of effect of maximum iteration threshold setting for the Mie Core inner loop (based on the LITE scene A dataset, consisting of 5760 valids point with SNR > 10 and Validity Flag set to true).	32
4	Overview of effect of innerloop NonLinOpt thresholds (based on the LITE scene A dataset, consisting of 5760 valid points with SNR > 10 and Validity Flag set to true).	32
5	Overview of effects of applying the different thresholds	37
6	Overview of effects of applying the different thresholds on different regions of the atmosphere.	39
7	Overview of the data quality for subsequent range bins. The only thresholds applied are SNR>10 and ValidityFlag=True.	41
8	Parameter values used for the ADM performance simulation. The ADM Phase-B design values for the Rayleigh equivalent and earth background equivalent bandwidths are obtained from an e-mail correspondence with ESA (Anne Grete Straume) on Wednesday 3 June 2009.	46
9	Parameter values used to simulate the Fizeau transmission function, see also Table 1.	48
10	Parameter values used to simulate the Fabry-Perot transfer function.	50
11	Bias and variance for Mie and Rayleigh channel wind errors for a 1000 meter vertical bin including a very thin cloud layer with one-way cloud transmission τ_c . A constant wind-shear of 0.004 s^{-1} over the bin is assumed. The table values correspond to 3 different values for the one-way cloud transmission. X means no wind retrieval	69
12	ECMWF zonal wind effective model resolution as a function of pressure level (hPa) as obtained from [RD9]. Italic values denote extrapolation by matching Table 4.4 in [RD12] with respective values in [RD9], i.e. equal missing variance at 500 and 150 hPa and linear interpolation of the missing variance from 150 to 50 hPa. The effective model resolution of 200 km at 50 hPa is a guess value.	71
13	Estimates of Mie channel HLOS wind error contributors; stddev is an abbreviation of standard deviation, repr. is an abbreviation of representativeness	72
14	Estimates of Rayleigh channel HLOS wind error contributors; stddev is an abbreviation of standard deviation, repr. is an abbreviation of representativeness	72

List of Figures

1	LOS Wind shear plotted as function of the wind velocity error, for the CALIPSO data discussed in section 3.10. 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 velocity is therefore mainly caused by hight assignment errors. Note that the report text uses the phrase wind speed error, but for historical reasons in most plots the phrase wind velocity error is still used to indicate the same property.	16
2	Mie spectral shape for a wind velocity of 50 ms^{-1} (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. Also note that the ACCD columns are numbered between 1 and 16 in this report. 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 ms^{-1} . Finally, the small blue peak below the Mie spectral shape, represents the amount of light entering accd bin 3.	19
3	Mie spectral shape for a wind velocity of 50 ms^{-1} (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.	21
4	Location of the selected LITE data.	23
5	Location of the selected CALIPSO data.	23

6	Scatterplot of 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 velocity error found by the simulations (for results with SNR>10 and ValidityFlag set). The overplotted black diamonds and ranges denote the average and standard deviation of 25 deviations along the y-axis. This illustrates that there is no clear relation between the absolute wind velocity value and the data quality, even though at first sight the large amount of data between wind velocities of 0 and 20 ms ⁻¹ might suggest otherwise.	24
7	First simulation results using LITE data. The data shows nonsense retrievals for very low SNR and a clear overflow problem.	25
8	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.	26
9	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).	27
10	Simulation results using LITE data, again using a SNR threshold of 10, and only accepting data for which the Validity flag is set to true. In this case a correct signal scaling is applied and the overflow problem disappears (left panel). To illustrate the effect of Poisson noise on the simulations the same case has been run without Poisson noise (right panel).	27
11	Simulation results using LITE data, without wind variability (using zero input wind) and without Poisson noise, to illustrate the numerical inaccuracies of the simulation program.	28
12	Simulation results using CALIPSO data, all results without selection (left panel) and after applying a SNR threshold of 10 (right panel).	29
13	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).	29
14	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.	30
15	Number of outer loop iterations, as function of the wind velocity error, for the selected CALIPSO data.	30
16	Average number of inner loop iterations, as function of the number of outer loop iterations for the selected LITE data (left panel), and as a function of residual error (right panel).	31
17	Average number of inner loop iterations, as function of the number of outer loop iterations for the selected LITE data (left panel), and as a function of residual error (right panel). For this test a maximum of 200 inner loop iterations was allowed, in stead of the default value of 20.	31
18	Residual error reported by the Mie Core algorithm as function of the wind velocity error, for the selected CALIPSO data.	33
19	Residual error reported by the Mie Core algorithm as function of the wind velocity error, plotted as contour plot. All valid results (left panel) and the region below ResErr=0.01 magnified.	33
20	Same data as shown in Fig. 19, 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 velocity error std exceeds the value of 1 ms ⁻¹ (requirement for the LOS wind velocity accuracy) can be read, which is at about 0.005.	34
21	FWHM reported by the Mie Core algorithm as function of the wind velocity error, for the selected CALIPSO data.	35
22	FWHM reported by the Mie Core algorithm as function of the wind velocity 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.	35
23	Same data as shown in Fig. 22 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).	36

24	Scatterplot of FWHM against wind velocity 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).	36
25	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.	37
26	Scatterplot of SNR against FWHM, all cases with SNR _i >10 and validity flag true (left panel), and zoomed in to the region of SNR _i >200 (right panel). The overplotted lines show the mean and one std value on both sides of the mean, for a number of SNR layers. Clearly the std increases for lower SNR and also the mean FWHM value shifts to larger values for lower SNR, because all transparent cloud cases accumulate here.	38
27	LOS wind error std for different climate zones against altitude. The blue dots and curve are for all data (global). The green, red, cyan and purple dots and curves are for tropical, subtropical, midlatitude and polar data, as is explained in the text. The black line indicates the requirement, the green dotted line a step in rangebin size and the red dotted line the ground surface. The dotsizes represent the amount of valid datapoints used to reconstruct the statistics.	40
28	Distribution of the normalized atmosphere Mie peak on the Fizeau interferometer ACCD bins for a 50 ms ⁻¹ wind velocity. The ACCD bin numbers are on the horizontal axis.	47
29	Fabry-Perot channel A (direct, blue) and B (reflected, red) transmission (left panel) and Fizeau reflection, $\tau_{\text{Fiz}}(\nu)$, (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.	51
30	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.	51
31	Rayleigh receiver response (left) and receiver sensitivity (right) for $T=273$ (K) (solid) and $T=210$ (K) (dashed). Red lines correspond to LIPAS-2009, blue lines to the L2Bp receiver response tables (file AE.TEST_AUX_RBC_L2_20071003T123456_20081002T123456.0001), L2Bp version CY35R2_Jun03, June 2009).	53
32	Rayleigh LOS wind error bias (ms ⁻¹) through Mie contamination and as a function of temperature (K). The true LOS wind is 13 ms ⁻¹ . The scattering ratio is a measure of Mie contamination in the Rayleigh channel signal.	55
33	Same as Fig. 32 but now for larger scattering ratio values up to 5 (left panel) and an extreme large wind of -110 ms ⁻¹ (right panel).	56
34	Case 1. Total signal from 14 accumulated measurements on the 16 ACCD bins, after digitisation, 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. Bin 24 is below the surface and contains no signal in both the E2S and LIPAS. See the text for further details.	57
35	Case 1. Signal, after digitisation, 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.	58
36	Relative difference between the E2S and LIPAS-2009 Fizeau (left) and Fabry-Perot (right) ACCD signals. The large difference in Fizeau bin 22 and Fabry-Perot bin 23 is explained in the text.	59
37	E2S (red) and LIPAS-2009 (black) backscatter (left) and attenuated backscatter (right).	59
38	Relative difference between E2S and LIPAS-2009 for attenuated backscatter (left) and one-way molecular transmission (right).	59
39	Case 2. Signal on the 16 ACCD bins of the Fizeau interferometer, after digitisation, 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 7-9 clearly show an increased signal from cloud particles.	60

40	Case 2. Signal on the 16 ACCD bins of the Fabry-Perot interferometer , after digitisation, for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. Bins 8-10 clearly show an increased signal due to cross-talk from cloud particles. See the text for further details. . . .	61
41	Same as Fig. 36 but now for case 2.	61
42	Atmospheric pressure (left) and temperature (right)	62
43	Atmospheric wind (left) and molecular backscatter (right). There is no aerosol backscatter. . .	62
44	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.	63
45	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) . . .	63
46	Rounding effect from floating point to integer numbers of electrons. Left, the number of electrons on the ACCD with (red) and without (black) the rounding operator. Middle, the difference of the curves: black minus red. Right, corresponding Rayleigh wind bias (blue diamonds).	64
47	Atmospheric temperature (left) and aerosol (right). The left curve/middle/right shows aerosol/molecular/cloud backscatter.	64
48	Mie (green/red) and Rayleigh (blue) HLOS wind retrieval statistics from 100000 simulations. The left/middle (zoomed in) panel show the bias (including the rounding effect explained in Fig. 46), the right panel the standard deviation. The green symbols are obtained from the simplified Mie retrieval algorithm described in MMS, the red symbols are obtained with the Mie Core algorithm.	65
49	L1Bp and L2Bp HLOS wind error statistics obtained from 94 BRCs (observation profiles) along one CALIPSO orbit; error bias (left) and error standard deviation (right). Blue and red symbols correspond to L1Bp and L2Bp winds respectively with stars corresponding to Mie winds and triangles to Rayleigh winds. The E2S was used to simulate the noisy Mie and Rayleigh channel signals.	65
50	LIPAS HLOS wind error statistics obtained from about 400 BRCs (observation profiles), see Fig. 51 along one CALIPSO orbit; error bias (left) and error standard deviation (right). Red stars correspond to Mie winds and the blue diamonds, magenta squares and green triangles to Rayleigh winds obtained with various processing options that are further discussed in section 6.3. LIPAS was used to simulate the noisy Mie and Rayleigh channel signals.	66
51	Number of observations used for the statistics in Fig. 50. Symbols are explained in Fig. 50.	66
52	Rayleigh wind error bias (ms^{-1}) (left) and variance (m^2s^{-2}) (right) as a function of one-way cloud transmission and cloud layer thickness (m). The measurement bin is 1000 m, the wind-shear is constant over the bin with a value of 0.004 s^{-1}	70
53	Half (nighttime) orbit CALIPSO level-1B attenuated backscatter at 532 nm measured on 1/1/2007 around 00:30 UTC. Top: Raw data at instrument resolution, see Table 4.2.3-1 of [RD7]. Bottom: processed data to 3300 m horizontal and 125 m vertical resolution. The spurious artefacts near 8 km are due to a change of the on-board averaging from 30 to 60 m.	76
54	Atmosphere scattering ratio along the nighttime part of one CALIPSO orbit, measured on 1/1/2007 around 00:30 UTC (top) and corresponding zonal wind (bottom) as obtained from the atmospheric database winds and ADM pointing. The horizontal resolution is 3.5 km, the vertical resolution 125 m.	77
55	Same as Fig. 54 but zoomed in to the Tropics.	78
56	Same as the bottom panel in Fig. 54 but now the HLOS wind.	79
57	Retrieved Mie (top) and Rayleigh (bottom) winds for the scene displayed in Fig. 54 and vertical sampling scenario WVM1. Rayleigh wind retrieval option 1 was used, see text for further details. The LIPAS output quality indicator flags are used to select the good quality winds for plotting. Dark blue spots denote missing winds based on the quality indicator flags. Note the different height interval along the y-axis.	80
58	Same as the bottom panel of Fig. 57 but for option 2 (top) and option 3 (bottom) for the Rayleigh wind retrieval, see text for further details.	81

59	HLOS wind error bias for Mie (red stars) and Rayleigh channel (left) for vertical sampling scenario WVM1. Three processing options for Rayleigh wind retrieval are displayed: option 1, simply accumulating all measurements (blue diamonds), option 2, application of classification (green triangles) and option 3, application of cross-talk correction (magenta squares). Right, the number of data that passed quality control and used in the statistics.	82
60	Wind and temperature statistics for scenes with tropical cirrus on 1/1/2007. Left, zonal (black) and meridional (blue) wind distribution in the tropics. solid and dashed lines denote the mean and mean +/- one standard deviation respectively. Zonal wind (middle) and temperature (right) statistics at the location of tropical cirrus. Red lines correspond to the mean value. The tropics is bounded here at 30 degrees latitude.	82
61	Scattering ratio statistics inside tropical cirrus clouds. The right plot is zoomed in to the lowest scattering ratio values. Red lines correspond to the mean value.	83
62	Same as Fig. 61 but now for the cumulative probability.	83
63	HLOS wind error standard deviation for Mie and Rayleigh channel for vertical sampling scenario WVM1. For an explanation of the symbols see Fig. 59.	84
64	Same as Fig. 57 but now for sampling scenario WVM2 and retrieval option 2 for the Rayleigh winds.	85
65	Same as Fig. 59 but now for sampling scenario WVM2	86
66	Same as Fig. 63 but now for sampling scenario WVM2	86
67	Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the PBL for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.	88
68	Same as Fig. 67 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).	88
69	Total number of good, moderate, bad and quality controlled HLOS wind obtained from accumulating the Mie (Fig. 67) and Rayleigh (Fig. 68) winds for processing option 1 (left), option 2 (middle) and option 3 (right).	88
70	Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the lower troposphere for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.	90
71	Same as Fig. 70 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).	90
72	Same as Fig. 69 but now for the lower troposphere.	90
73	Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the UTLS for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.	92
74	Same as Fig. 73 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).	92
75	Same as Fig. 69 but now for the UTLS.	92
76	Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the stratosphere for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.	94
77	Same as Fig. 76 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).	94
78	Same as Fig. 69 but now for the stratosphere.	94
79	Same as Fig. 79 but now for all atmospheric regions accumulated. These plots show the total data coverage and HLOS wind quality throughout the atmosphere.	95
80	Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the UTLS in the extra-tropics (left) and tropics (right).	95
81	Same as Fig. 80 but now for the stratosphere above 15 km.	96

82	DWL profile information content relative to radiosonde information content. Retrieval option 2 (classification) was used here for Rayleigh channel winds.	98
83	Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the PBL for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.	100
84	Same as Fig. 83 but now for Rayleigh channel winds retrieved with processing option 1 (left), option 2 (middle) and option 3 (right).	100
85	Total number of good, moderate, bad and quality controlled HLOS wind obtained from accumulating the Mie (Fig. 83) and Rayleigh (Fig. 84) winds for processing option 1 (left), option 2 (middle) and option 3 (right).	100
86	Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the lower troposphere for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.	102
87	Same as Fig. 86 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).	102
88	Same as Fig. 85 but now for the lower troposphere.	102
89	Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the UTLS for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.	104
90	Same as Fig. 89 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).	104
91	Same as Fig. 85 but now for the UTLS.	104
92	Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the stratosphere for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.	106
93	Same as Fig. 92 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).	106
94	Same as Fig. 85 but now for the stratosphere.	106
95	Quality and coverage of Mie HLOS winds in tropical cirrus scenes as a function of vertical sampling scenario: Mie HLOS wind error bias (left), standard deviation (middle) and the number of data (right).	107
96	Quality and coverage of Rayleigh HLOS winds (processing option 1) in tropical cirrus scenes as a function of vertical sampling scenario: Rayleigh HLOS wind error bias (left), standard deviation (middle) and the number of data (right).	108
97	Same as Fig. 96 but now for Rayleigh channel processing option 2.	108
98	Same as Fig. 96 but now for Rayleigh channel processing option 3.	108
99	Same as Fig. 85 but now for all atmospheric regions accumulated. These plots show the total data coverage and HLOS wind quality throughout the atmosphere.	109
100	DWL profile information content relative to radiosonde information content. Retrieval option 2 (classification) was used here for Rayleigh channel winds.	110
101	Vertical sampling default scenarios wvm1 (left) and wvm2 (right). The Mie channel reaches the ground for both scenarios thus providing the opportunity for Mie channel zero wind calibration.	112
102	Extra-tropical vertical sampling with (leftmost two) and without (rightmost two) zero wind calibration opportunity. Scenario names from left to right: et_zwc1,et_zwc2,et_nozwc1,et_nozwc2.	112
103	Tropical vertical sampling with (leftmost two) and without (rightmost two) zero wind calibration opportunity. Scenario names from left to right: tr_zwc1, tr_zwc2, tr_nozwc1, tr_nozwc2.	113
104	Same as Fig. 101 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names are wvm1_1km (left) and wvm2_1km (right).	114

105	Same as Fig. 102 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names from left to right: et_zwc1_1km, et_zwc2_1km, et_nozwc1_1km, et_nozwc2_1km.	114
106	Same as Fig. 103 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names from left to right: tr_zwc1_1km, tr_zwc2_1km, tr_nozwc1_1km, tr_nozwc2_1km.	115
107	Stratospheric scenario with 1000 m (right) Rayleigh bins near the surface. This scenario has no zero wind calibration opportunity. The scenario is denoted wmv_stratos.	115
108	Same as Fig. 101 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names are wvm1_1km (left) and wvm2_1km (right).	116
109	Same as Fig. 102 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names from left to right: et_zwc1_1km, et_zwc2_1km, et_nozwc1_1km, et_nozwc2_1km.	116
110	Same as Fig. 103 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names from left to right: tr_zwc1_1km, tr_zwc2_1km, tr_nozwc1_1km, tr_nozwc2_1km.	117
111	HLOS wind statistics for sampling scenario wvm1_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.	118
112	HLOS wind statistics for sampling scenario wvm2_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.	119
113	HLOS wind statistics for sampling scenario wvm.et_nozwc1_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.	120
114	HLOS wind statistics for sampling scenario wvm.et_nozwc2_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.	121
115	HLOS wind statistics for sampling scenario wvm.et_zwc1_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.	122
116	HLOS wind statistics for sampling scenario wvm.et_zwc2_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.	123
117	HLOS wind statistics for sampling scenario tr_nozwc1_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.	124
118	HLOS wind statistics for sampling scenario tr_nozwc2_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.	125
119	HLOS wind statistics for sampling scenario tr_zwc1_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.	126
120	HLOS wind statistics for sampling scenario tr_zwc2_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.	127

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 B. Finally this technical note is concluded with an overview of the findings in section C.

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 1.41 Software Release Note (SRN)”, 12-Jun-2009, AE-RN-ECMWF-L2BP-001.
- [RD2] “ADM-Aeolus L1B Master Algorithm Document (MAD)”, AE-SW-ASU-GS-023, issue 5.1, dated Jun-2008.
- [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 2.5, issued 27-Aug-2008.
- [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 7 (final revision), dated 15-Jun-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] “TN10.1, Modified ADM-Aeolus Operations Concept (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 10-Oct-2008, reference code: AE-TN-ECMWF-GS-101.
- [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.
- [RD11] ADM-Aeolus Science Report, April 2008, SP1311.
- [RD12] Lorenc, A.C., Graham, R.J., Dharsi, I., Macpherson, B., Ingleby, N.B., Lunnen, R.W., Study of preparation for the use of Doppler wind lidar information in meteorological assimilation systems, 1991, ESA study contract 9063/90/HGE-I.
- [RD13] Vamp TN5, “Impact of the Vertical Sampling Scenarios on NWP and Modeling of Stratospheric Circulation”, by Harald Schyberg, MetNo, and Heiner Körnich, MISU, reference code: AE-TN-KNMI-VAMP-005, version 19-Nov-2009.
- [RD14] Vaughan, J.M. and D.V. Willetts, “The influence of Atmospheric Turbulence, Shear, and Aerosol Stratification and Cloud Motion on Space Doppler Wind Lidar Signals”, DERA Technical note for ESA, contract no. 12006/96/NL/CN, DRA/EL/LPD/CR96/ESTEC/TN2/1.0.

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.
- [LR6] Paffrath, U., Lemmertz, C., Reitebuch, O., Witschas, B., The airborne demonstrator for the direct-detection Doppler wind lidar ALADIN on ADMAeolus: II. Simulations and radiometric performance, submitted to *JTECH*, 2009

2.4 Acronyms

ACCD	Accumulation Charge Coupled Device
AD	Applicable Document (see section 2.1)
ADM	Atmospheric Dynamics Mission
ANX	Ascending Node Crossing
ARTS	Anomaly Reporting and Tracking System
BRC	Basic Repeat Cycle (covering a 200 km orbit section)
DEM	Digital Elevation Model
DPM	Detailed Processing Model
E2S	End-to-End Simulator (as developed by MDA)
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
QC	Quality Control
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
UTLS	Upper Troposphere, Lower Stratosphere
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. All plots in sections 5 and 6 have been created using matlab.

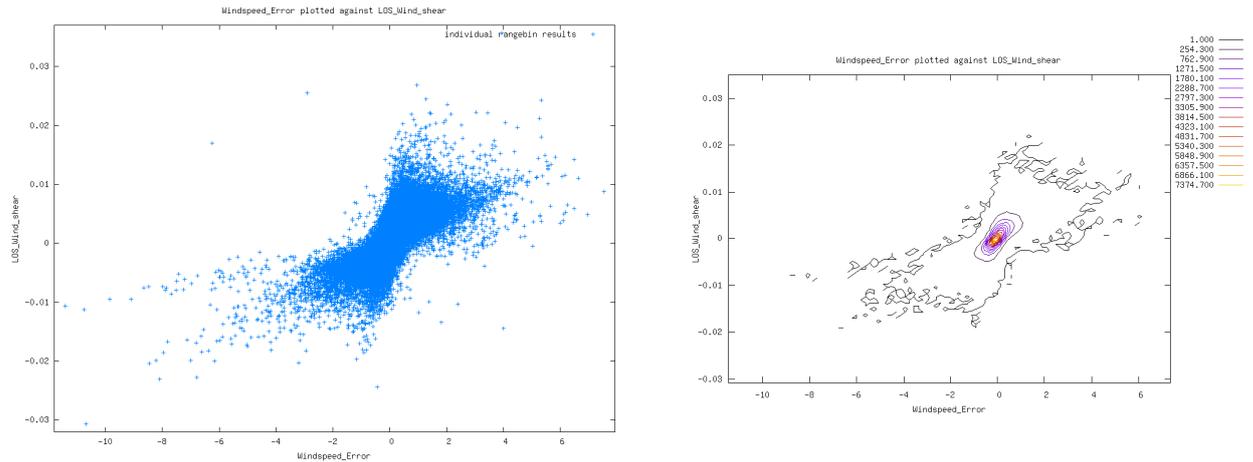


Figure 1: LOS Wind shear plotted as function of the wind velocity error, for the CALIPSO data discussed in section 3.10. Scatterplot (left panel) and contour plot (right panel). The data has been selected on $SNR > 100$, which means this result includes mainly reflections on optically thick clouds. The error in wind velocity is therefore mainly caused by height assignment errors. Note that the report text uses the phrase wind speed error, but for historical reasons in most plots the phrase wind velocity error is still used to indicate the same property.

3 Task 4a: Mie Core QC

3.1 Introduction

For this sub-task the usefulness of Mie core algorithm outputs to quantify winds errors caused by 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 1.41. 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°, using eq. 11 defined on page 45. Also an azimuth angle of 45° is assumed, this is needed since the simulation does not implement an orbit calculation. All simulations described in section 3 have been run using LOS winds as input and output. To keep the program simple, no LOS to HLOS conversion was implemented. 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 independently on the u and v component profiles to increase the wind shear
- the vertical shear in horizontal wind velocity and optical properties are calculated from the input atmosphere

¹It was delivered on 12-Jun-2009. See the eRoom folder: AEOLUS > UPLOAD.ECMWF > L2BP_Release1.41_20090612

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 a simulated 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]

- 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 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
- 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 velocity), and σ describes the width of the peak, and relates to the σ_{FWHM} as specified in eq. 2:

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

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 velocity 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 wind velocity of 50 ms^{-1} 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.

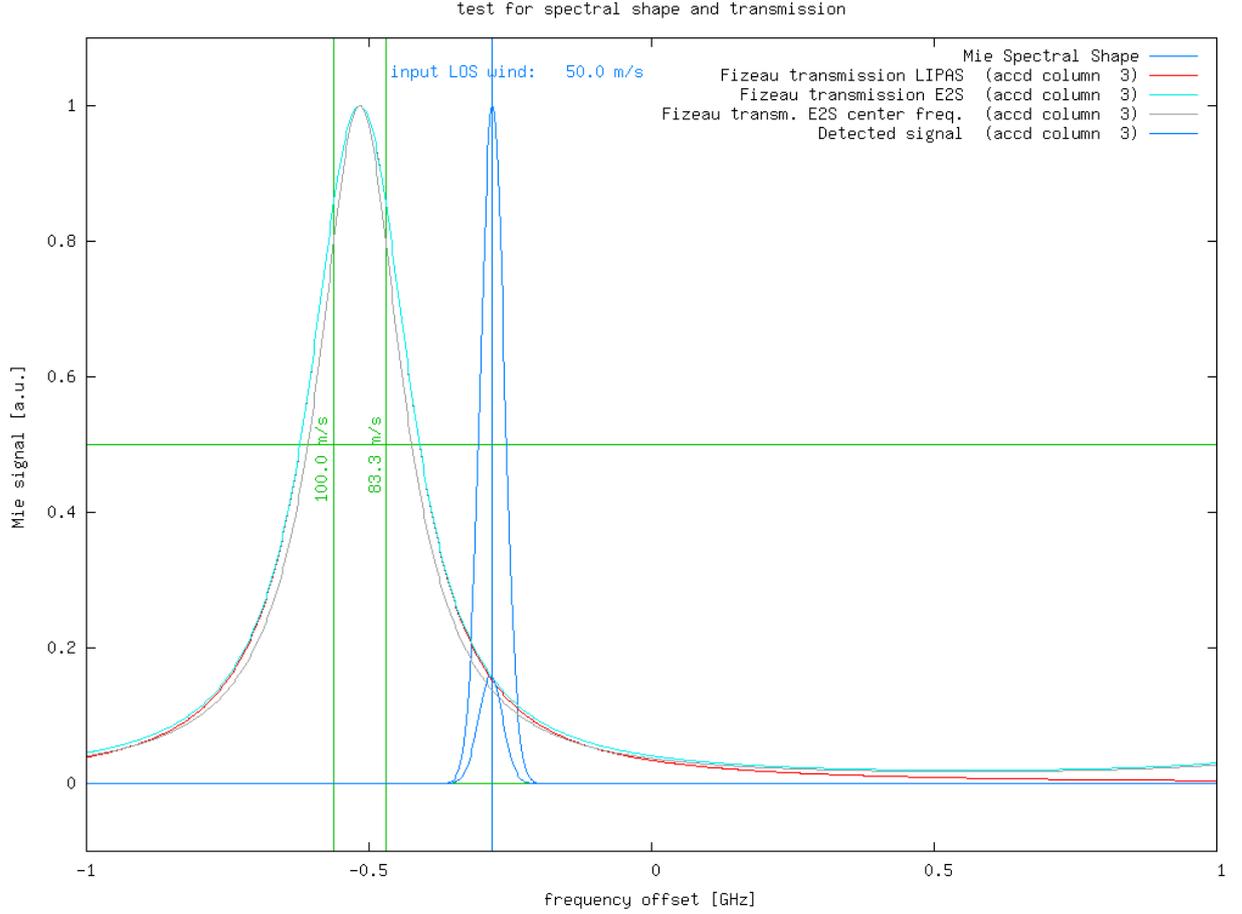


Figure 2: Mie spectral shape for a wind velocity of 50 ms^{-1} (blue), compared to the much broader Fizeau transmission curve for accd column 3 calculated in 3 different ways red: using the Lorentzian function as is done in LIPAS; cyan: using the Airy function as in done by the E2S; gray: using the Airy function but without integration over the accd bin surface. Also note that the ACCD columns are numbered between 1 and 16 in this report. 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 ms^{-1} . Finally, the small blue peak below the Mie spectral shape, represents 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 eqs. 4 to 8:

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

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

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

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

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

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 integral of each curve 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 procedure⁷. 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 Estimation of Signal-to-Noise ratio

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. Note that this is not the refined SNR which is calculated inside an updated MieCore version in newer L1bP versions, and which is not yet available in the L2bP software.

The procedure is as follows:

- sort the 16 Mie accd bin values in ascending order
- subtract the background level
- average the 4 lowest bin values to calculate s_{lowest}
- average all bin values to calculate s_{all}
- estimate SNR from $SNR = \frac{s_{all} - s_{lowest}}{\sqrt{s_{all}}}$

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.

⁷Note that due to this offset correction the actual FWHM to be used by LIPAS should be slightly smaller than the FWHM used by the E2S.

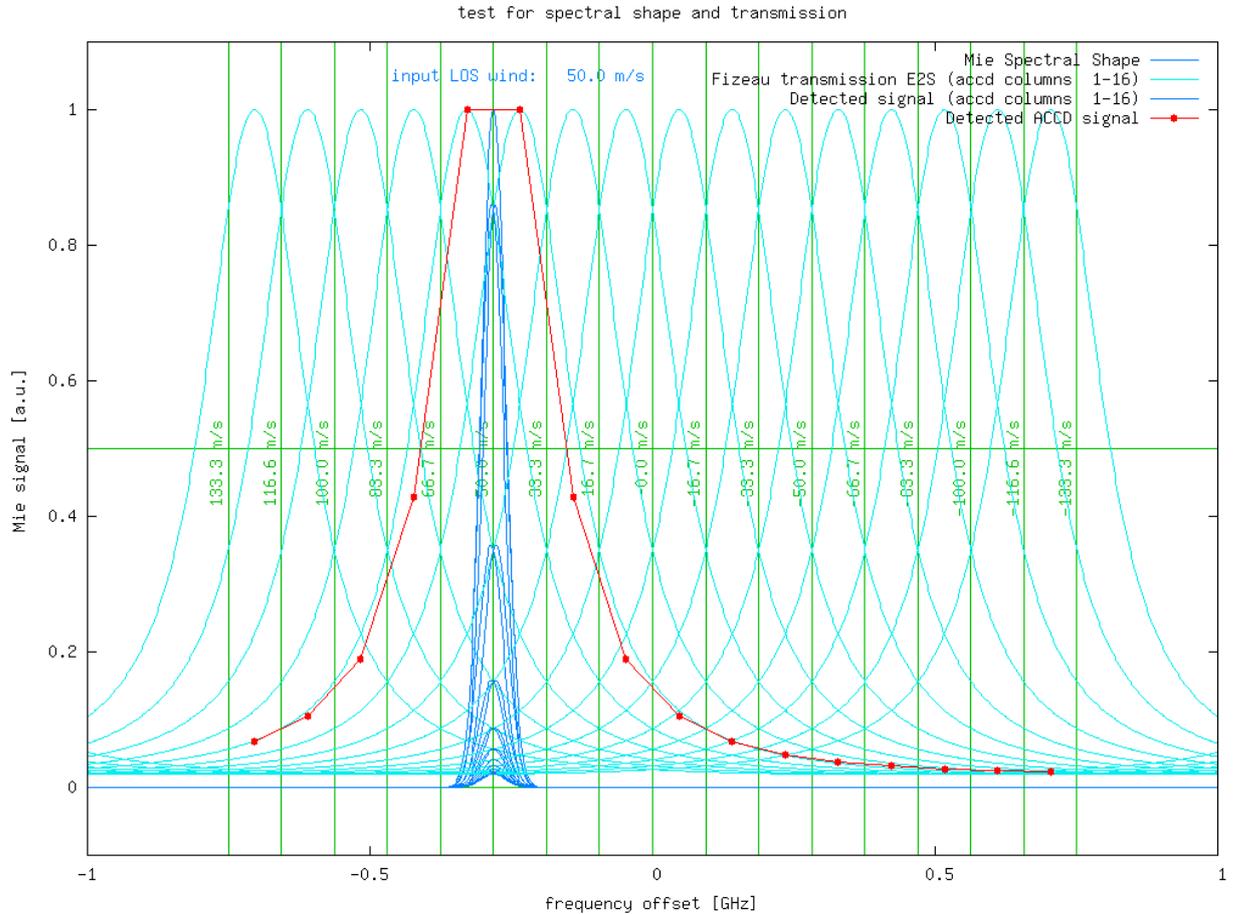


Figure 3: Mie spectral shape for a wind velocity of 50 ms^{-1} (blue), compared to the the Fizeau transmission curves for all 16 accd columns (cyan). The blue curves below the Mie peak denote the fraction of light detected by each of the accd bins. Clearly, due to the long tails of the transmission curves, all accd bins will receive some signal. The expected signal that will be measured by each accd channel is denoted by the red squares.

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

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 for this simulation

if any 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.

3.6 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, pressure, aerosol and molecular backscatter and extinction (see [RD8] and [RD14]). 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. The location of the selected LITE data is above Indonesia, see Figure 4.
- **CALIPSO**: 10 half orbits of CALIPSO data have been used to produce the statistics presented in the next sections (sections 3.10 and 3.11). This includes 70 scenarios from the atmospheric database named Calipso_Orbit[A]_Segment[B]_and_EcmwfColloc in which [A] runs from 1 to 10, 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 56.000 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. The location of the selected LITE data is above Indonesia, see Figure 5.

An example of the ECMWF input data Colocated with the CALIPSO data used is displayed in Fig. 6.

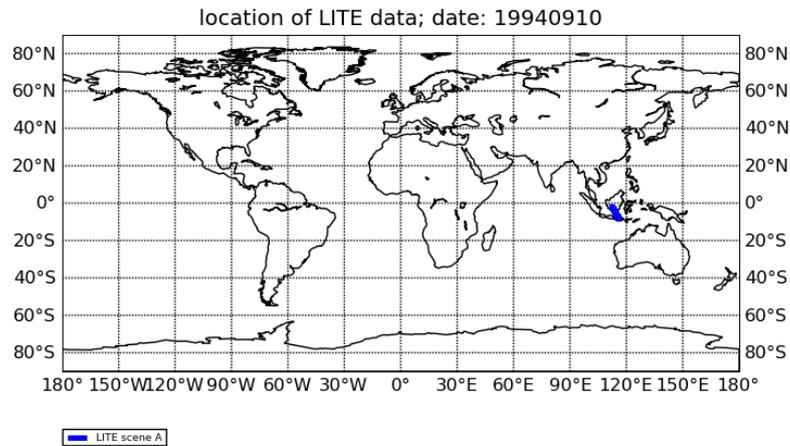


Figure 4: Location of the selected LITE data.

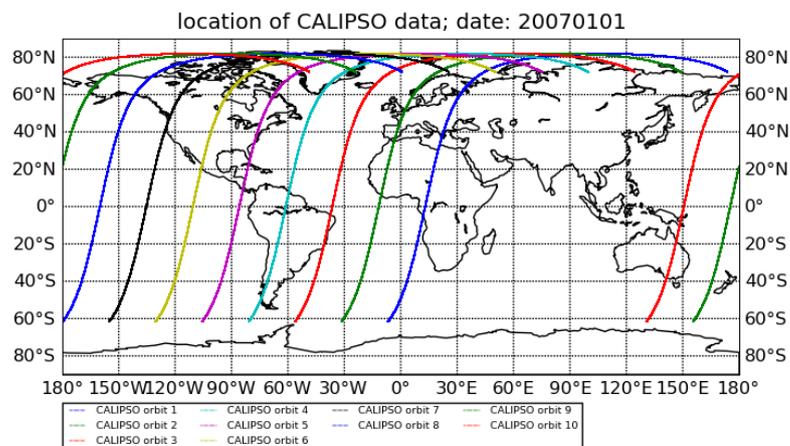


Figure 5: Location of the selected CALIPSO data.

3.7 Error multiplier

By comparing the NWP data with radio sonde measured profiles it was found, as is reported 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.43 and section 5, p.187). This multiplier adds noise to the NWP wind profiles in such a way that the statistical

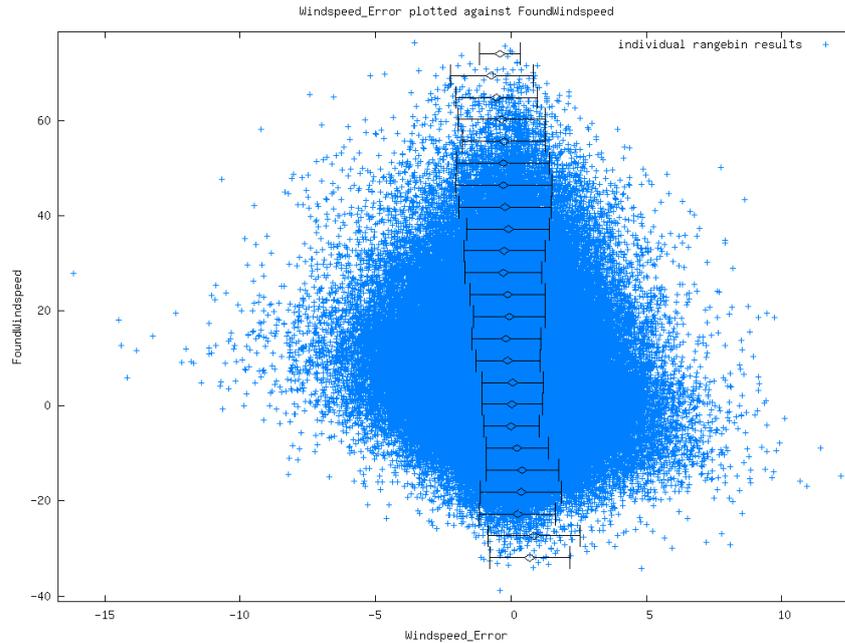


Figure 6: Scatterplot of input LOS winds derived from ECMWF u,v,w components, collocated at the locations of the CALIPSO data used for the majority of simulations, plotted against the wind velocity error found by the simulations (for results with $\text{SNR} > 10$ and ValidityFlag set). The overplotted black diamonds and ranges denote the average and standard deviation of 25 deviations along the y-axis. This illustrates that there is no clear relation between the absolute wind velocity value and the data quality, even though at first sight the large amount of data between wind velocities of 0 and 20 ms^{-1} might suggest otherwise.

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 on each wind component independently as follows:

- take vertical atmospheric slices of 1 km thickness
- 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 3 has been used)
- apply this enhanced variance to add extra noise (with a Gaussian distribution) on top of the NWP model wind profile

3.8 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 horizontal 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 ms^{-1} in standard deviation over 50 km [RD9], while the lacking wind variability in the vertical is about 2 ms^{-1} in standard deviation and much larger [RD7].

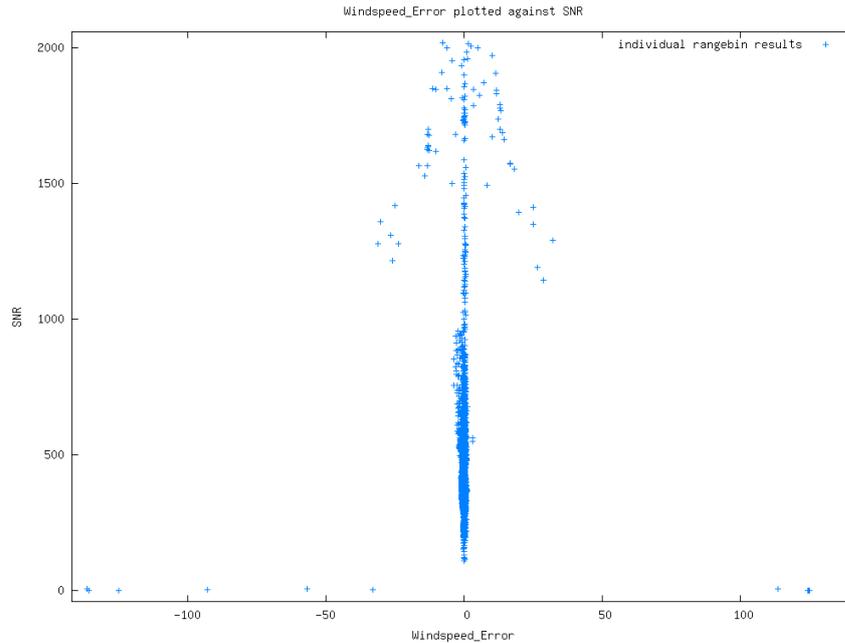


Figure 7: First simulation results using LITE data. The data shows nonsense retrievals for very low SNR and a clear overflow problem.

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.

3.9 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 velocity was determined. An interesting result was obtained when SNR was plotted against wind velocity error, see Fig. 7.

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 applied⁸ in Fig. 8 (left panel). This improves the situation, the points around SNR=0 with wind velocity errors upto 120 ms⁻¹ 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 height⁹. This causes too high photon count numbers to be stored in the 2 byte integer fields

⁸the current L1B processing uses a threshold value of SNR=10 as well.

⁹in fact measurement level signals had been scaled to match observation level E2S results, causing the peak height to be a

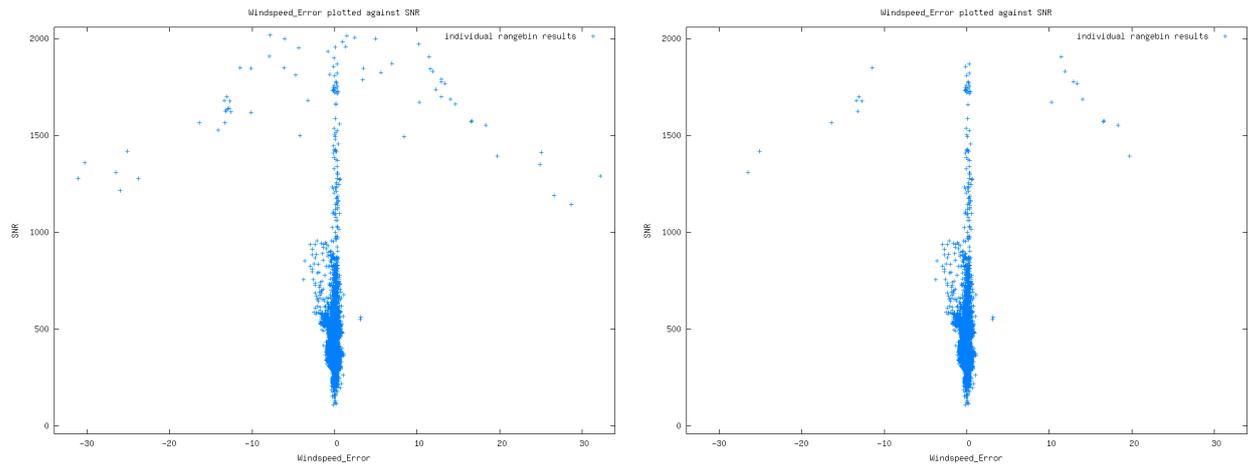


Figure 8: Simulation results using LITE data, after applying a SNR threshold of 10 (left panel), and after discarding data reported invalid by the ValidityFlag as well (right panel). The nonsense retrievals for very low SNR have disappeared now but the overflow problem remains.

available for the spectrum. As a consequence the numbers get truncated and become nonsense¹⁰

To illustrate this an example of a spectrum showing this problem for the internal reference pulses is shown in Fig. 9. 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. 8 it is clear that errors upto 15 ms^{-1} can result from this effect.

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. 8 (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. 10.

The same simulation was re-run without Poisson noise. In this case the wind velocity errors are significantly reduced, as expected to the range -0.7 to $+2 \text{ ms}^{-1}$, see Fig. 10. 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 ms^{-1} . By removing variations in wind velocity 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. 11. Typical wind velocity errors are now below 0.1 ms^{-1} . 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 ms^{-1} in this plot, especially visible for higher SNR values and accentuated by the 2 vertical lines of points is not yet explained. It is not caused by the finite sampling used on the frequency scale,

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.

¹⁰Note 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 informally to D. Huber on 08-Apr-2009, and a formal Anomaly report was filed 2-Dec-2009 by J. de Kloe as AE-IPF-81).

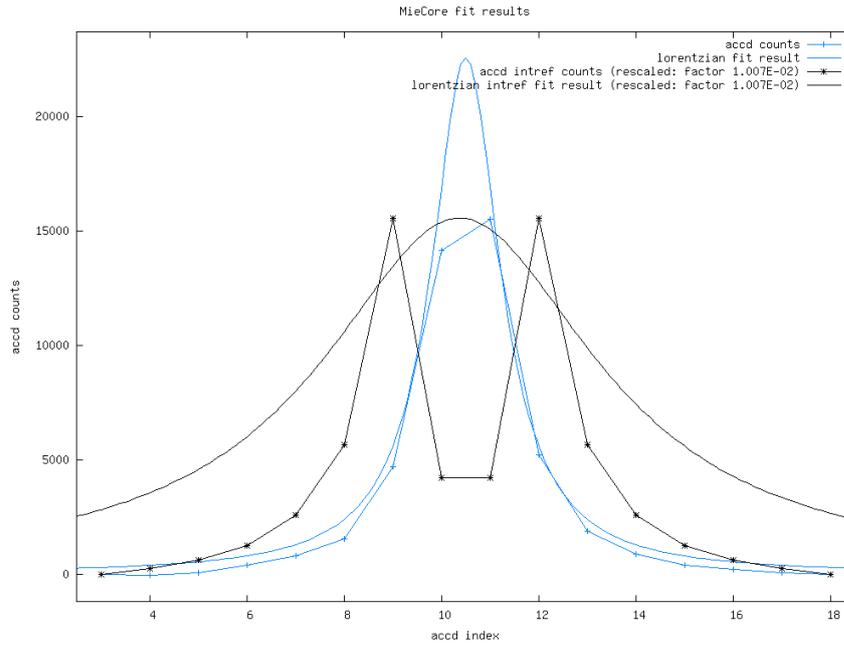


Figure 9: A case showing the effect of an integer overflow in the accd counts of the Mie channel (these are internal reference counts, and this case was generated by L1B v1.07).

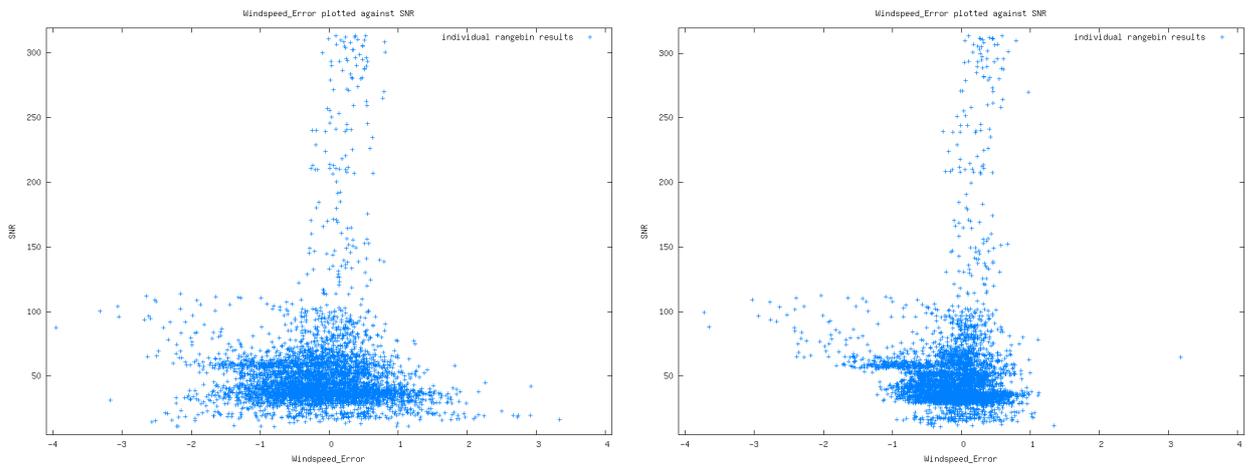


Figure 10: Simulation results using LITE data, again using a SNR threshold of 10, and only accepting data for which the Validity flag is set to true. In this case a correct signal scaling is applied and the overflow problem disappears (left panel). To illustrate the effect of Poisson noise on the simulations the same case has been run without Poisson noise (right panel).

because they remain at the same position when the frequency sampling is doubled. Clearly these remaining numerical and discretisation effects are insignificant compared to the errors that occur when Poisson noise is added, and compared to the variability that occurs when a varying wind profile (including the error multiplier) is used.

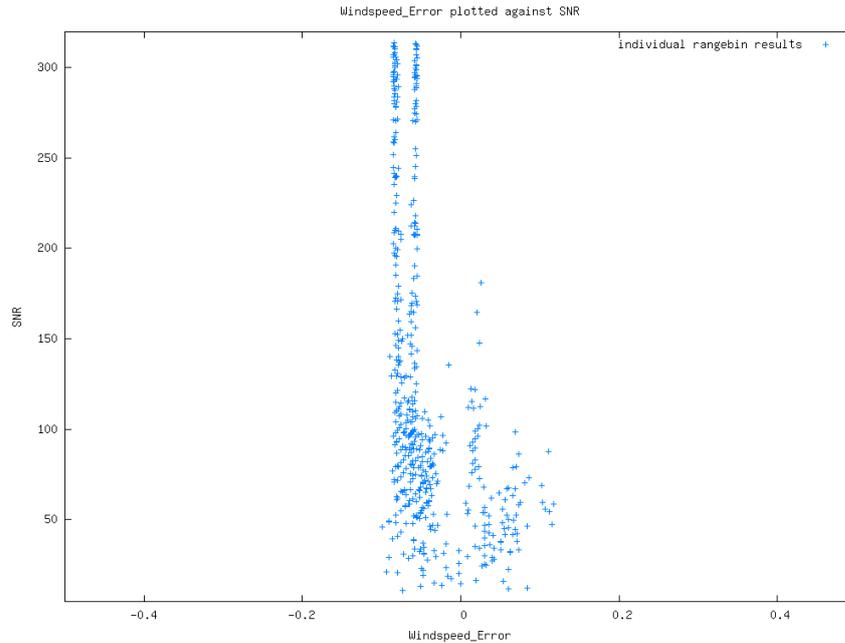


Figure 11: Simulation results using LITE data, without wind variability (using zero input wind) and without Poisson noise, to illustrate the numerical inaccuracies of the simulation program.

3.10 Results: CALIPSO data; data selection

A next series of tests that have been run using CALIPSO data (for details see section 3.6). 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 1.046.000 simulated range bins. Note that about 300.000 range bins had to be discarded already at this stage due to lack of signal. The same data selection as presented in section 3.9 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. 12. This plot displays the SNR as function of the wind velocity error found by the simulation (left panel: all data; right panel: points with $SNR < 10$ 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. 13. The actual number of points discarded by the SNR threshold of 10 is about 292.000 (28 %) and leaves about 753.000 points. The number of additional points discarded by the ValidityFlag is about 19.000 (1.8 %) and leaves about 734.000 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. 13 has been displayed as contour plot in Fig. 14. Note that the majority of points seem located near the origin of the plot, so a magnified view zooming in on the region $10 < SNR < 100$ is given in the right panel of Fig. 14. Actually about 80% of all data is located in the magnified region.

3.11 Results: CALIPSO data; Mie Core outputs

The selected CALIPSO data as described in section 3.10 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.

3.11.1 Outer loop iterations

Figure 15 shows the number of outer loop iterations used by the MieCore fitting as a function of wind velocity error. The most prominent feature in this plot is the discontinuity in the data, i.e. iterations

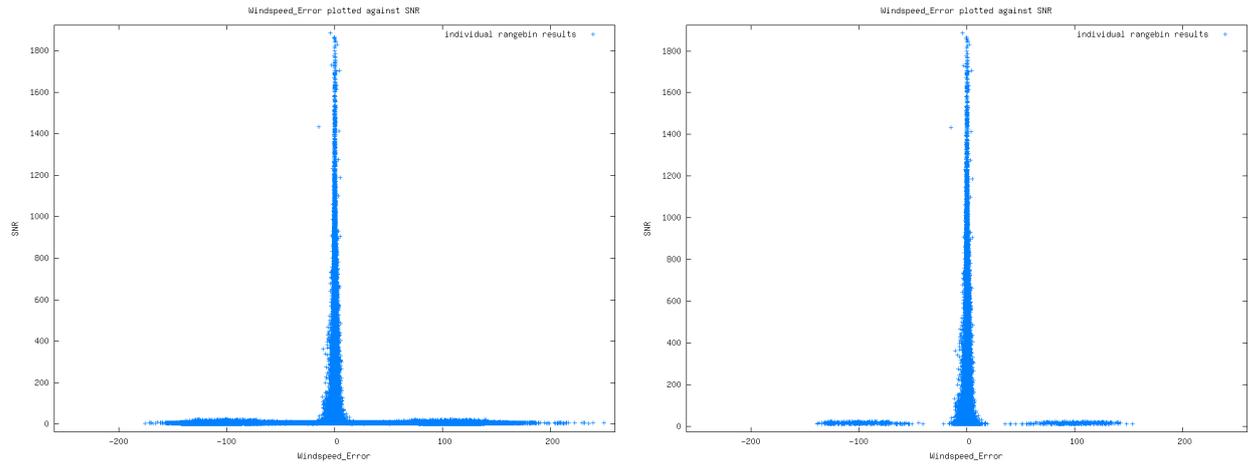


Figure 12: Simulation results using CALIPSO data, all results without selection (left panel) and after applying a SNR threshold of 10 (right panel).

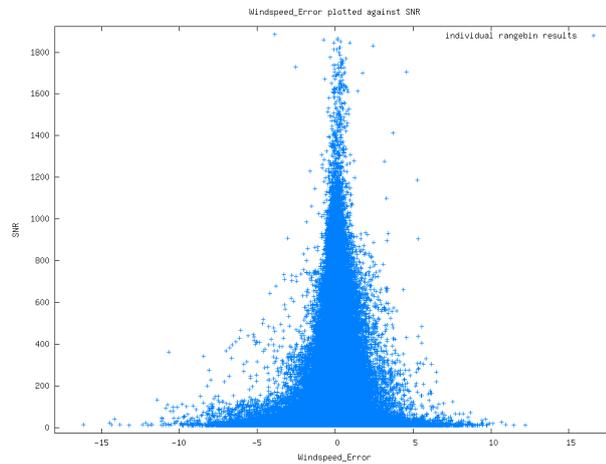


Figure 13: Simulation results using CALIPSO data, after applying a SNR threshold of 10 and after discarding data reported invalid by the ValidityFlag. as well (right panel).

between 11 and 30 seem never to occur. When looking at the actual number of points in both regions, there are about 182.000 points with 10 or less iterations (25%) and about 552.000 points with 30 iterations¹¹ (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 two¹².

Note that the current MieCore algorithm implementation does not allow easy access to the number of inner loop iterations. If time permits this will be investigated at a later stage.

¹¹Note that the threshold on the iteration count is set to 30, but the current L2B Mie Core algorithm performs 31 iterations. This has been identified as a bug in the L2BP, and has been reported using the ARTS system. It should have no effect whatsoever on the results presented in this report because convergence of the fit takes places in the first 10 iterations, or not at all.

¹²Assuming 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.

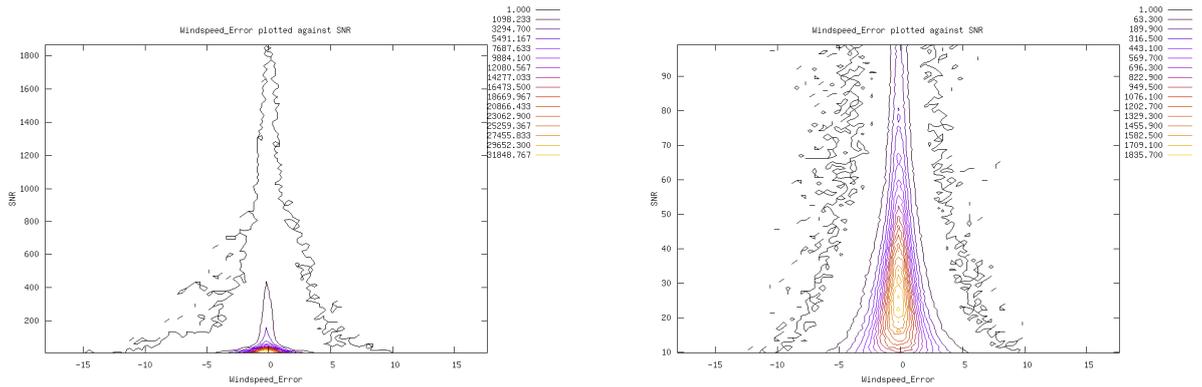


Figure 14: All valid data after discarding points with $SNR < 10$ and invalid ValidityFlag (left panel). The right panel shows a magnification of the region $10 < SNR < 100$.

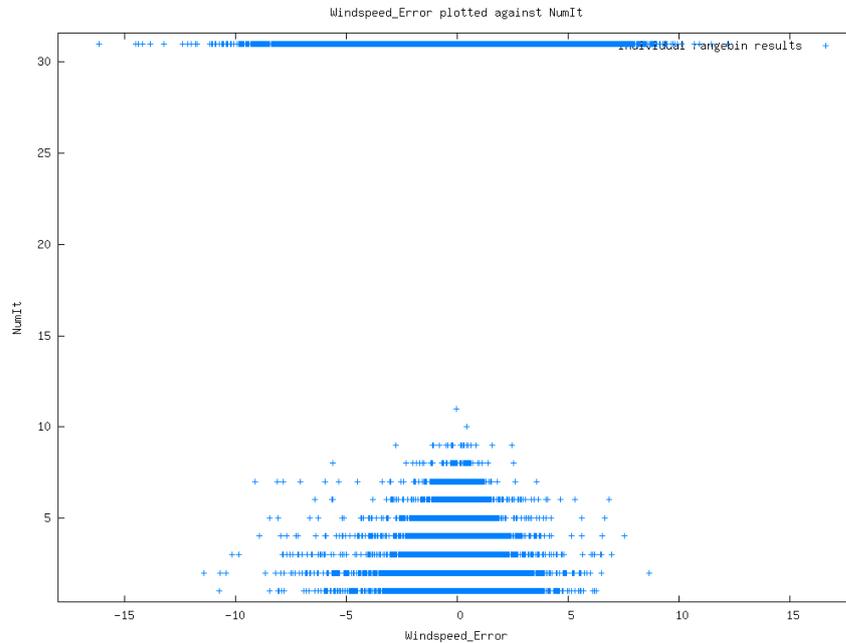


Figure 15: Number of outer loop iterations, as function of the wind velocity error, for the selected CALIPSO data.

3.11.2 Inner loop iterations

To obtain the properties of the inner loop of the Mie Core Algorithm, the Mie Core Response Fortran90 module had to be modified. This has been done only after generating all other Calipso results. Since regenerating them takes a long time, the results in this section are based on the small LITE scenario A dataset.

Figure 16 (left panel) shows the average number of inner loop iterations against the number of outer loop iterations performed (again using the selection $SNR > 10$ and validityFlag=true, which yielded 5719 valid data points). It is again very clear that the data points are mainly found in the top of the diagram. About 87.5% of all data points (4995 out of 5719 valid points) have more than 19.5 inner loop iterations

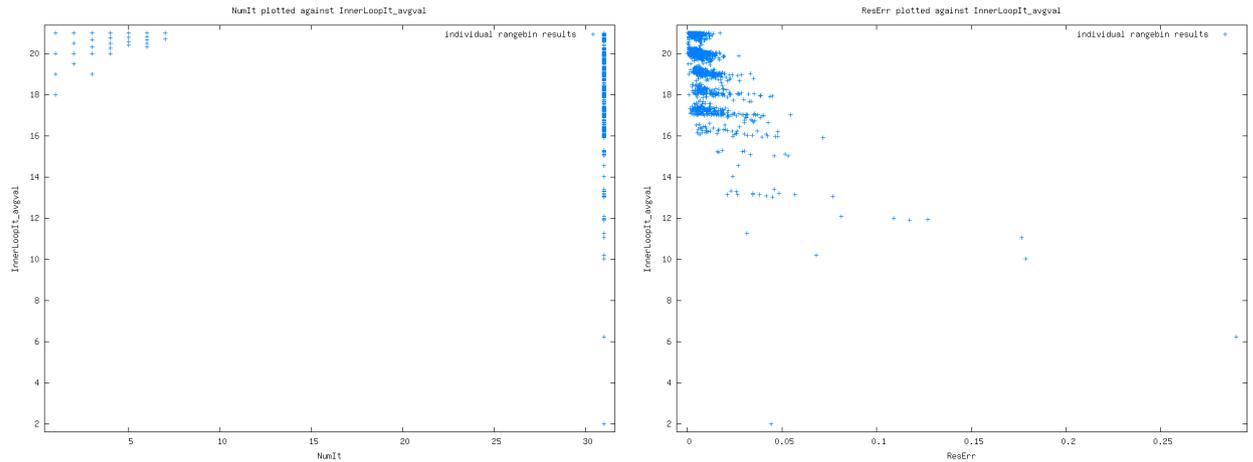


Figure 16: Average number of inner loop iterations, as function of the number of outer loop iterations for the selected LITE data (left panel), and as a function of residual error (right panel).

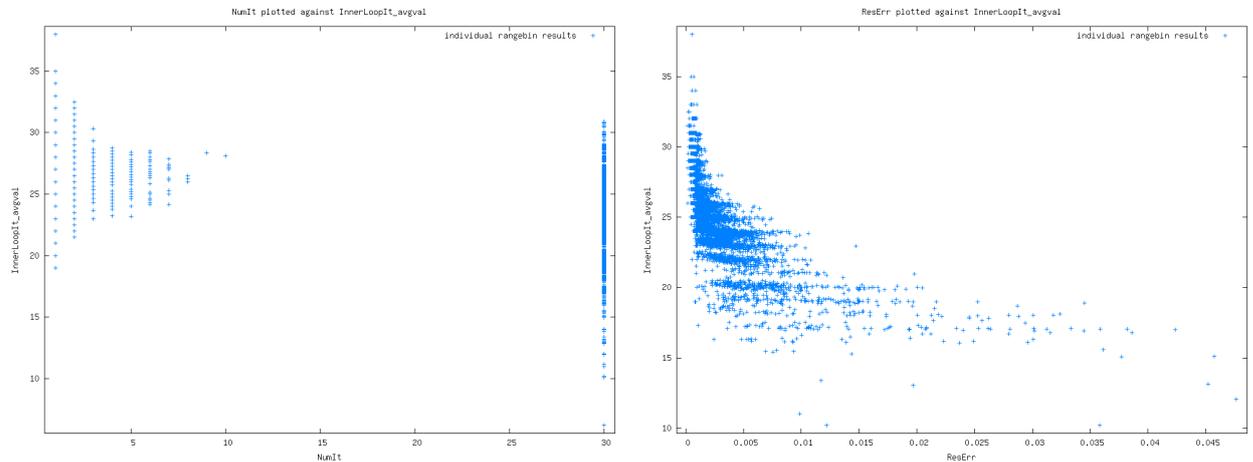


Figure 17: Average number of inner loop iterations, as function of the number of outer loop iterations for the selected LITE data (left panel), and as a function of residual error (right panel). For this test a maximum of 200 inner loop iterations was allowed, in stead of the default value of 20.

on average. For the points with less than 10 outer loop iterations (which are supposed to be the very good fits) even more than 99% of all points (807 out of 811 points) have more than 19.5 inner loop iterations on average, and only break the inner loop when the maximum number of inner loop iterations of 20 is reached.

This indicates that the convergence criterium for the inner loop is almost never reached, especially for the very good fits. This is also illustrated by plotting the average number of inner loop iterations against the residual error after the end of the fitting procedure. Clearly the data points with lower residual error show more inner loop iterations, indicating that the convergence criterium for this loop may not be optimal.

In order to find the point where the threshold on the inner loop stops the iterations, in stead of the maximum allowed iteration count, the maximum number of inner loop iterations allowed was set to 200 (so much higher than the default value of 20). This resulted in the plots show in figure 17. From this it becomes clear that 200 iterations are never needed in the inner loop. After typically 25 to 30 iterations the minimum is found and the loop is terminated. More than 35 iterations never occur in this example.

To look at the effect on the precision of the Mie Core results, the experiment was repeated with 5

Table 3: Overview of effect of maximum iteration threshold setting for the Mie Core inner loop (based on the LITE scene A dataset, consisting of 5760 valids point with SNR > 10 and Validity Flag set to true).

MaxNumIt	nr of datapoints left	Windspeed Error	nr of datapoints left for ResErr < 0.005	Windspeed Error for ResErr < 0.005
10	5720	0.7169 ms ⁻¹	4519	0.6728 ms ⁻¹
15	5720	0.6396 ms ⁻¹	4818	0.5889 ms ⁻¹
20	5719	0.6274 ms ⁻¹	4854	0.5764 ms ⁻¹
25	5719	0.6267 ms ⁻¹	4857	0.5761 ms ⁻¹
200	5719	0.6268 ms ⁻¹	4857	0.5763 ms ⁻¹

Table 4: Overview of effect of innerloop NonLinOpt thresholds (based on the LITE scene A dataset, consisting of 5760 valid points with SNR > 10 and Validity Flag set to true).

NonLinOptThresh	nr of datapoints left	WindspError	nr of datapoints left for ResErr < 0.005	Windspeed Error for ResErr < 0.005
0.01	5719	0.6268 ms ⁻¹	4857	0.5763 ms ⁻¹
0.001	5716	0.6213 ms ⁻¹	4876	0.5740 ms ⁻¹
0.0001	5713	0.6206 ms ⁻¹	4878	0.5739 ms ⁻¹

different values for the maximum number of inner loop iterations (MaxItNonLinOpt). This is tabulated in table 3. From this table it is clear that less then 20 iterations clearly results in larger std of windspeed error. However, allowing more than 20 iterations seems not to add anything significant to the retrieval quality.

Next the NonLinOptThresh threshold value has been varied. This threshold currently is set to be 0.01 by default. The result for allowing 200 inner loop iterations, and various values for NonLinOptThresh is shown in table 4. From this table a small quality improvement can be seen when the NonLinOptThresh is refined, but the effect is tiny, even when zooming in on the “very good” fits with ResErr < 0.005. The consequence of increasing this threshold is that the number of iterations needed to reach the stop threshold increases slightly. For NonLinOptThresh=0.001 the number of inner loop iterations typically is between 25 and 40, with a few exceptional points up to 44 iterations. For NonLinOptThresh=0.0001 the number of inner loop iterations typically is between 35 and 45, with a few exceptional points up to 51 iterations. Again, these high iteration counts only occur for the very good fit values at very low residual error.

Looking at these results, the settings for the inner loop of the Mie Core algorithm seem very well tuned, and changing them will only result in a minimal quality change at the cost of a significant increasing the runtime.

3.11.3 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 3.5). 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. 18. At first sight the distribution seems symmetric and centered around the origin.

When zooming in there is a clear concentration of points around 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 ResErr>0.1 is only 2%, but the number of points with ResErr>0.01 is already 35%, and the number of points with 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.

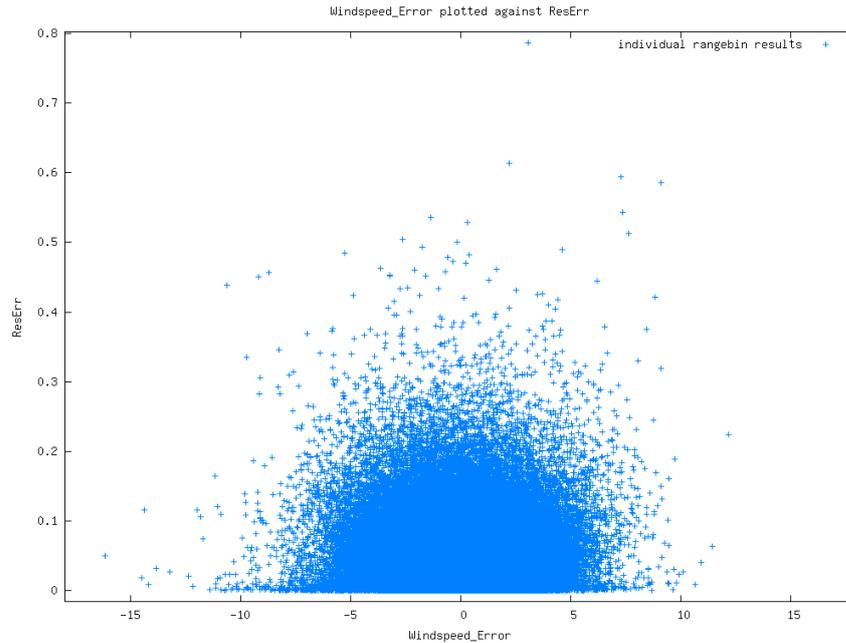


Figure 18: Residual error reported by the Mie Core algorithm as function of the wind velocity error, for the selected CALIPSO data.

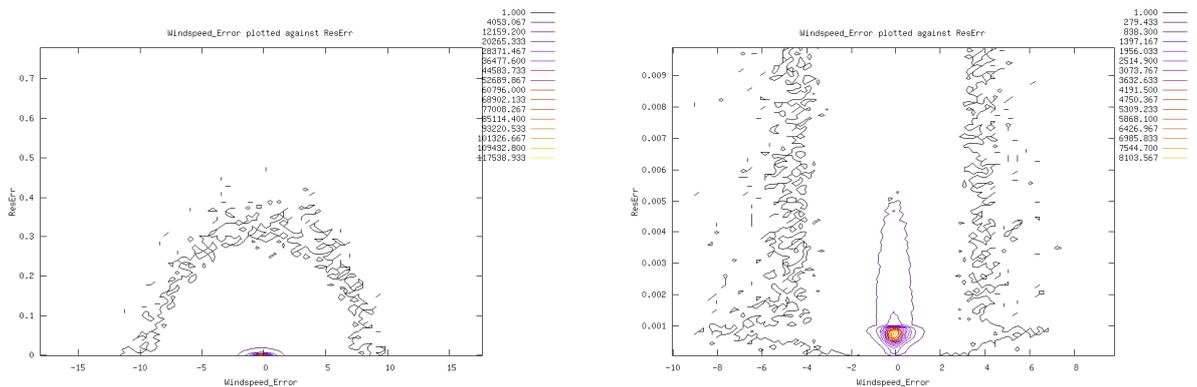


Figure 19: Residual error reported by the Mie Core algorithm as function of the wind velocity error, plotted as contour plot. All valid results (left panel) and the region below ResErr=0.01 magnified.

Since from the contours in Fig. 19 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. 20. In this scatterplot the mean and 1 standard deviation (STD) on both sides of the mean have been over plotted. The resulting wind velocity error STD values vary between 0.91 and 4.71 ms^{-1} for the plot showing all data (left panel), and between 0.54 and 1.21 ms^{-1} 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 velocity error STD exceeds the value of 1 ms^{-1} can be read, which is at about 0.005¹³.

¹³Note that the value of 1 ms^{-1} is taken just as a simple example value here. The requirement for the HLOS wind velocity accuracy is 1 m/s in the PBL, which translates to LOS wind velocity accuracy of about 0.6 m/s.

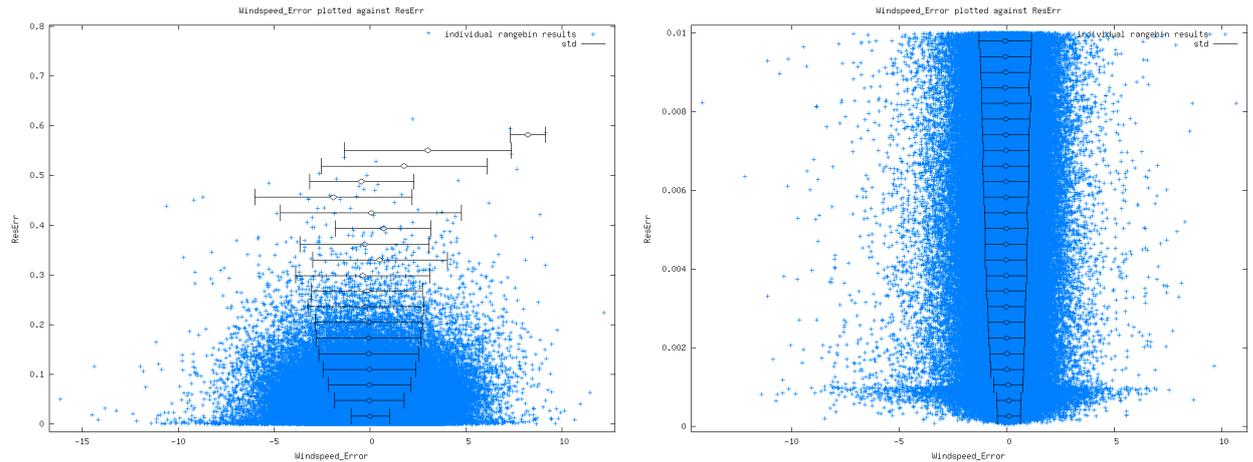


Figure 20: Same data as shown in Fig. 19, again as scatterplot, but now with over plotted lines to show the mean and one std value on both sides of the mean, for a number of Residual Error layers (full range is split in 25 sub-ranges for this calculation). From the right panel the position at which the wind velocity error std exceeds the value of 1 ms^{-1} (requirement for the LOS wind velocity accuracy) can be read, which is at about 0.005.

The overall STD of the wind velocity error distribution decreases from 1.2 to 0.75 ms^{-1} by applying this selection. This is at the cost of throwing away 365.400 points (so 50% of all data flagged valid by the SNR test and the Validity flag).

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.

3.11.4 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 velocities 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 3.8). A simple scatterplot of all selected data is shown in Fig. 21 and does not reveal much detail. It clearly has been cut-off at $\text{FWHM}=1.5$ and $\text{FWHM}=2.5$ due to the selection applied by the validity flag.

When plotted as contourplots a bit more detail becomes visible, see Fig. 22. Clearly the majority of all points are concentrated around $\text{FWHM}=2.05$.

When only points are selected close to $\text{FWHM}=2.05$, for example in the region as plotting in Fig. 22 (right panel) between 1.95 and 2.16, a slight beneficial effect can be seen on the wind velocity error distribution¹⁴. The wind velocity error std of this distribution decreases from 1.19 to 1.06 ms^{-1} . This is at the cost of throwing away 196.000 data points (so 27% of all data flagged valid by the SNR test and the Validity flag).

Since from the contours in Fig. 22 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. 23. In this scatterplot the mean and 1 standard deviation (STD) on both sides of the mean have been overplotted. The resulting wind velocity error STD values vary between 0.89 and 2.16 ms^{-1} when using all valid data (left panel) and between 0.79 and 1.38 ms^{-1} when looking at the region $1.95 < \text{FWHM} < 2.16$. From this a threshold can be defined using the 1 ms^{-1} values, as was done for the residual error. This results in the range 2.017 to 2.059.

¹⁴Note 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. 22 (right panel).

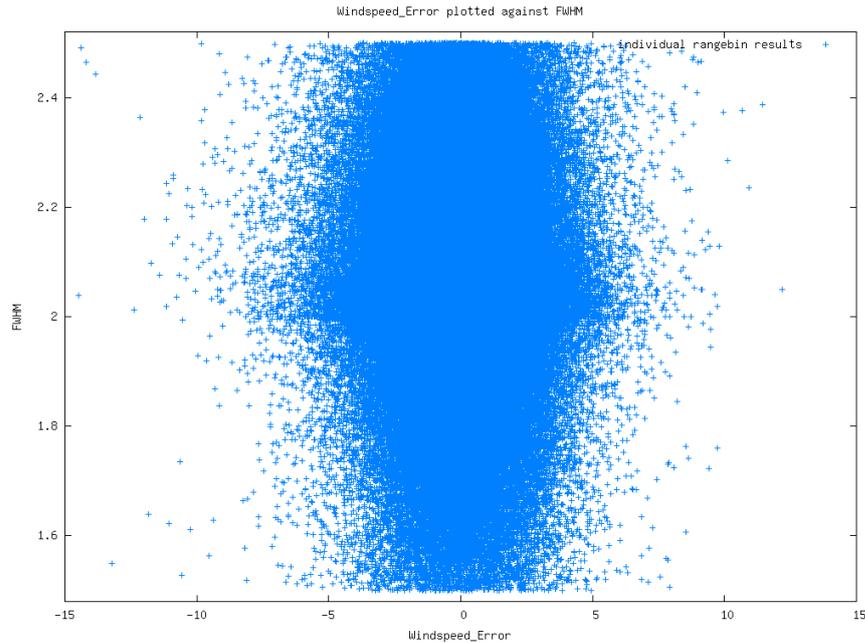


Figure 21: FWHM reported by the Mie Core algorithm as function of the wind velocity error, for the selected CALIPSO data.

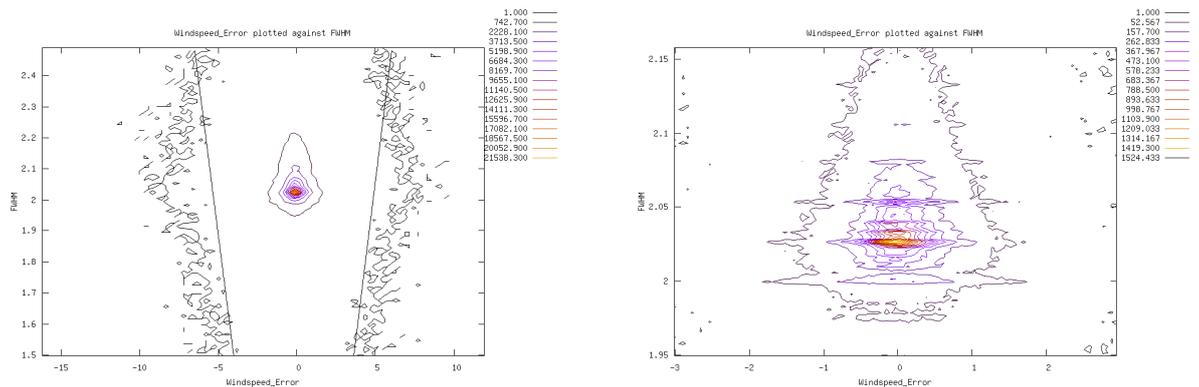


Figure 22: FWHM reported by the Mie Core algorithm as function of the wind velocity error, plotted as contourplot. All valid results (left panel), and a magnified region around FWHM=2 (right panel). The straight lines in the left panel are plotting artefacts and should be ignored. Also the striping feature in the right panel is an artefact caused by the binning of the data and probably not relevant.

When these numbers are recalculated on data with a residual error below 0.005 (so applying the residual error threshold as proposed in section 3.11.3) the result improves clearly, see Fig. 24. In this case the STD ranges from 0.62 to 1.82 and by taking a FWHM range of 1.74 upto 2.18, it seems possible to select data that has a STD in wind velocity error of 1 ms^{-1} or smaller.

By applying the different thresholds the wind velocity error can clearly be reduced. An overview of the results obtained is given in table 5. 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

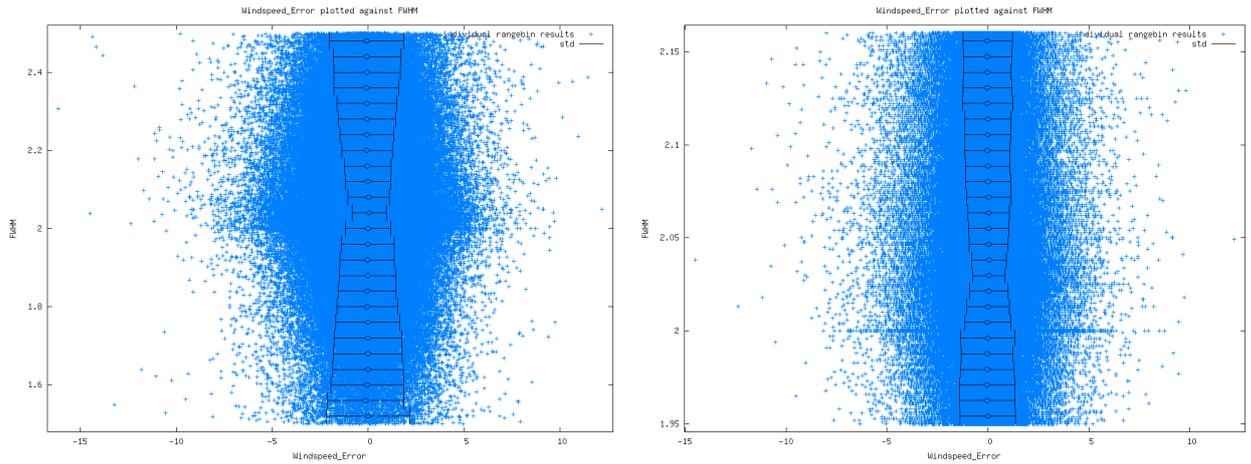


Figure 23: Same data as shown in Fig. 22 again as scatterplot, but now with overplotted lines to show the mean and one std value on both sides of the mean, for a number of FWHM layers (full range is split in 25 sub-ranges for this calculation).

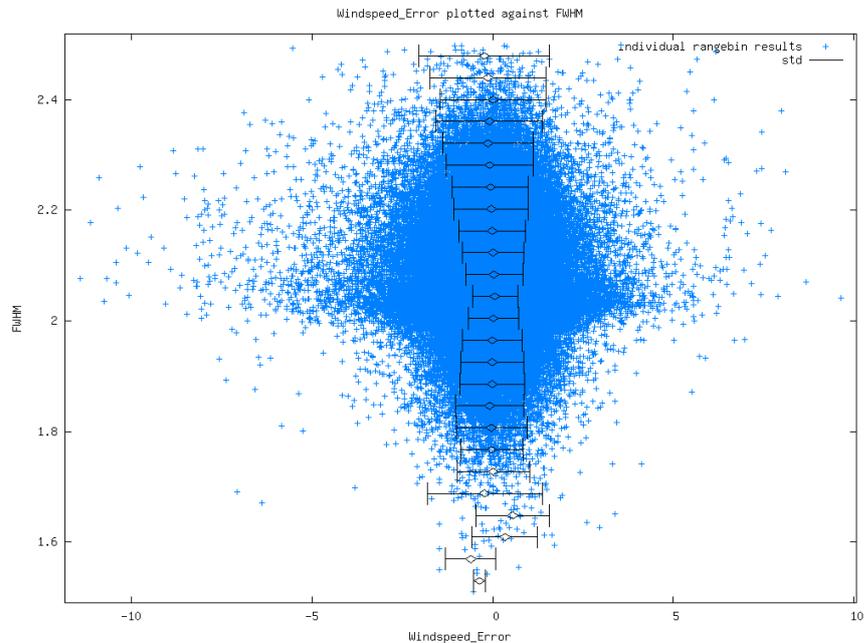


Figure 24: Scatterplot of FWHM against wind velocity error for cases with ResErr < 0.005. The overplotted lines show the mean and one std value on both sides of the mean, for a number of FWHM layers (full range is split in 25 sub-ranges for this calculation).

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 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 wind velocities 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

Table 5: Overview of effects of applying the different thresholds

Applied thresholds	nr of datapoints left	percentage	std in wind velocity error
SNR>10, ValFlag set	734091	100%	1.193 ms ⁻¹
SNR>10, ValFlag set, ResErr<0.005	368684	50%	0.745 ms ⁻¹
SNR>10, ValFlag set, 1.80<FWHM<2.35	692761	94%	1.149 ms ⁻¹
SNR>10, ValFlag set, 2.017<FWHM<2.059	241616	33%	0.897 ms ⁻¹
SNR>10, ValFlag set, 1.80<FWHM<2.35, ResErr<0.005	367143	50%	0.742 ms ⁻¹
SNR>10, ValFlag set, 2.017<FWHM<2.059, ResErr<0.005	179368	24%	0.615 ms ⁻¹

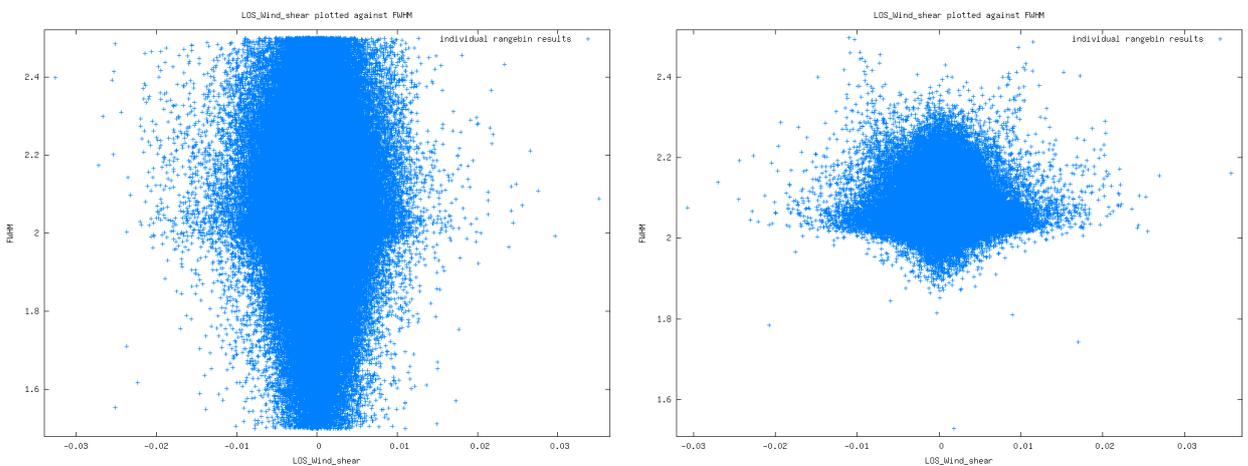


Figure 25: FWHM reported by the Mie Core algorithm as function of the wind shear in the input wind. The left plot selects data with SNR between 10 and 50, the right plot selects data with SNR above 50.

as simulation input in Fig. 25. 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).

Another way of visualising the same effect, is by plotting SNR versus FWHM, see Fig. 26. This plot clearly shows the increase of std in FWHM for lower SNR, but it also clearly shows a shift in the average FWHM value as a function of SNR. This effect can be explained by the fact that broader peaks (larger FWHM) are caused by windshear combined with partially transparent atmospheres. This automatically is correlated with a certain amount of extinction so will yield lower signals and lower SNR values. Note that it has been a mistake in previous versions of this report to assume that peak-broadening would be most easily visible for very large SNR values. This is not true, since these cases mainly are caused by reflection on cloud-tops, and the only shear related effect on these cases may be a height assignment error, not a peak broadening.

3.11.5 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.

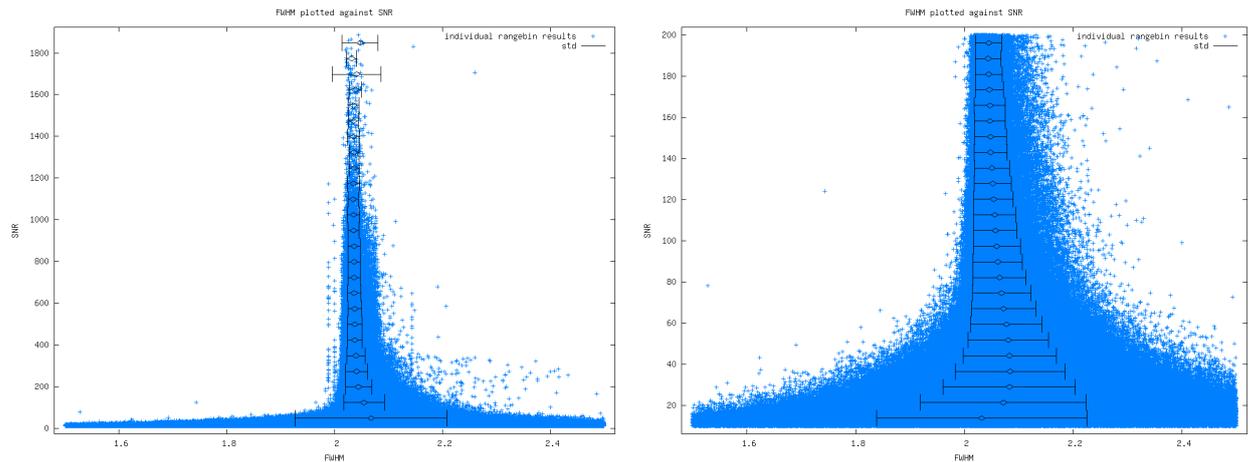


Figure 26: Scatterplot of SNR against FWHM, all cases with $SNR \geq 10$ and validity flag true (left panel), and zoomed in to the region of $SNR \geq 200$ (right panel). The overplotted lines show the mean and one std value on both sides of the mean, for a number of SNR layers. Clearly the std increases for lower SNR and also the mean FWHM value shifts to larger values for lower SNR, because all transparent cloud cases accumulate here.

- altitudes below 2.2 km (the region in which 500 m Mie rangebins are used) The required HLOS wind accuracy in this region is 1 ms^{-1} , which translates to a required LOS wind accuracy of 0.6 ms^{-1} .
- altitudes between 2.2 and 16.2 km (the region in which 1 km Mie rangebins are used) The required HLOS wind accuracy in this region is 2 ms^{-1} , which translates to a required LOS wind accuracy of 1.2 ms^{-1} .
- altitudes above 16.2 km (the region in which 2 km Mie rangebins are used) The required HLOS wind accuracy in this region is 3 ms^{-1} , which translates to a required LOS wind accuracy of 1.8 ms^{-1} .

The results of this exercise are given in table 6

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. The presented results indicate that the LOS wind requirement is met for altitudes above 2 km by applying a SNR threshold of 10, combined with a selection on the Mie Core ValidityFlag. For the PBL below 2 km the std of the LOS wind exceeds the requirement of 0.6 ms^{-1} slightly. Applying an additional threshold on either FWHM or Residual Error is sufficient to meet the requirement.

In addition to the results per rangebin, the data has also been divided into several climate zones. The climate regions are defined as follows: between plus and minus 20 latitude (tropics), between plus and minus 20 and 40 N/S (subtropics) between plus and minus 40 70 N/S (mid-latitudes) and greater/smaller than plus/minus 70 (Polar) (matching what was defined in section 4.2.1.7 of [RD7], p.37). The results are presented in Figure 27, together with the global result presented in the previous section. In this plot the different climate zones are represented by the different colors. The amount of data used to calculate the statistics for each rangebin is indicated by the size of the symbol.

From these results it is clear that there is some difference in quality between the different climate zones at some altitudes. The polar and mid-latitude regions seem to have a better quality between 2 and 7 km, and also between 16 and 20 km. For the other altitudes the quality is comparable. The comparison to the requirement leads to the same conclusion for all climate zones for the PBL. They all need additional filtering to meet the 0.6 ms^{-1} requirement. The region around 15-16 km is near the requirement for most cases. Only the sub-tropical case clearly exceeds the requirement here, and would surely need additional filtering, but the other cases are very close, so taking another statistical sample may change the situation in such a way that these also need filtering.

Table 6: Overview of effects of applying the different thresholds on different regions of the atmosphere.

Results for all altitudes combined			
Applied thresholds	nr of datapoints left	percentage	std in wind velocity error
SNR>10, ValFlag set	734091	100%	1.193 ms ⁻¹
Results for Altitude below 2.2 km			
Applied thresholds	nr of datapoints left	percentage	std in wind velocity error
SNR>10, ValFlag set	156296	21.3%	0.790 ms ⁻¹
SNR>10, ValFlag set, ResErr<0.005	116759	15.9%	0.454 ms ⁻¹
SNR>10, ValFlag set, 2.017<FWHM<2.059	77108	10.5%	0.526 ms ⁻¹
SNR>10, ValFlag set, 2.017<FWHM<2.059, ResErr<0.005	70383	9.59%	0.396 ms ⁻¹
Results for altitude between 2.2 and 16.2 km			
Applied thresholds	nr of datapoints left	percentage	std in wind velocity error
SNR>10, ValFlag set	436224	59.4%	1.147 ms ⁻¹
SNR>10, ValFlag set, ResErr<0.005	206897	28.2%	0.694 ms ⁻¹
SNR>10, ValFlag set, 2.017<FWHM<2.059	136649	18.6%	0.901 ms ⁻¹
SNR>10, ValFlag set, 2.017<FWHM<2.059, ResErr<0.005	97495	12.3%	0.631 ms ⁻¹
Results for altitude above 16.2 km			
Applied thresholds	nr of datapoints left	percentage	std in wind velocity error
SNR>10, ValFlag set	141571	19.3%	1.583 ms ⁻¹
SNR>10, ValFlag set, ResErr<0.005	45028	6.13%	1.280 ms ⁻¹
SNR>10, ValFlag set, 2.017<FWHM<2.059	27859	3.80%	1.444 ms ⁻¹
SNR>10, ValFlag set, 2.017<FWHM<2.059, ResErr<0.005	11490	1.57%	1.160 ms ⁻¹

3.11.6 Effect of rangebin size

Since the current dataset has been produced using the WVM1 range bin definition, there are currently no altitudes for which the effect of changing range bin size can be studied. However, it is possible to display the quality of the individual range bins and tabulate them against altitude. This has been done in Table 7. From this table it is clear that the std of the data of most rangebins fullfills the mission requirements without further selection (see [RD11], table 3.3 on page 42). As was mentioned above in section 3.11.5, some additional selection will be needed.

The range bins below 2 km are all above the required std of 0.6 ms⁻¹ in LOS wind velocity error. Additional selection will be needed to meet the requirement, and due to the varying quality the threshold values may be chosen to vary with altitude. To make this possible it would probably be best to make these Mie Core parameters (especially FWHM and Residual Error) available to the user in the L2B product file, for each retrieved wind result.

Especially range bin 24 has relatively large errors. Note however that this one is below the surface. Since surface reflections are not included in the current simulation, the data is reported anyway, based on extrapolated atmospheric data. The poor quality probably is due to the significant aerosol extinction in the layer just above the surface (since aerosol levels are usually highest in the atmospheric bin just above the surface, which is then extrapolated to allow calculations below the surface), and this also is reflected in the small amount of points from this layer that pass the SNR and ValidityFlag test.

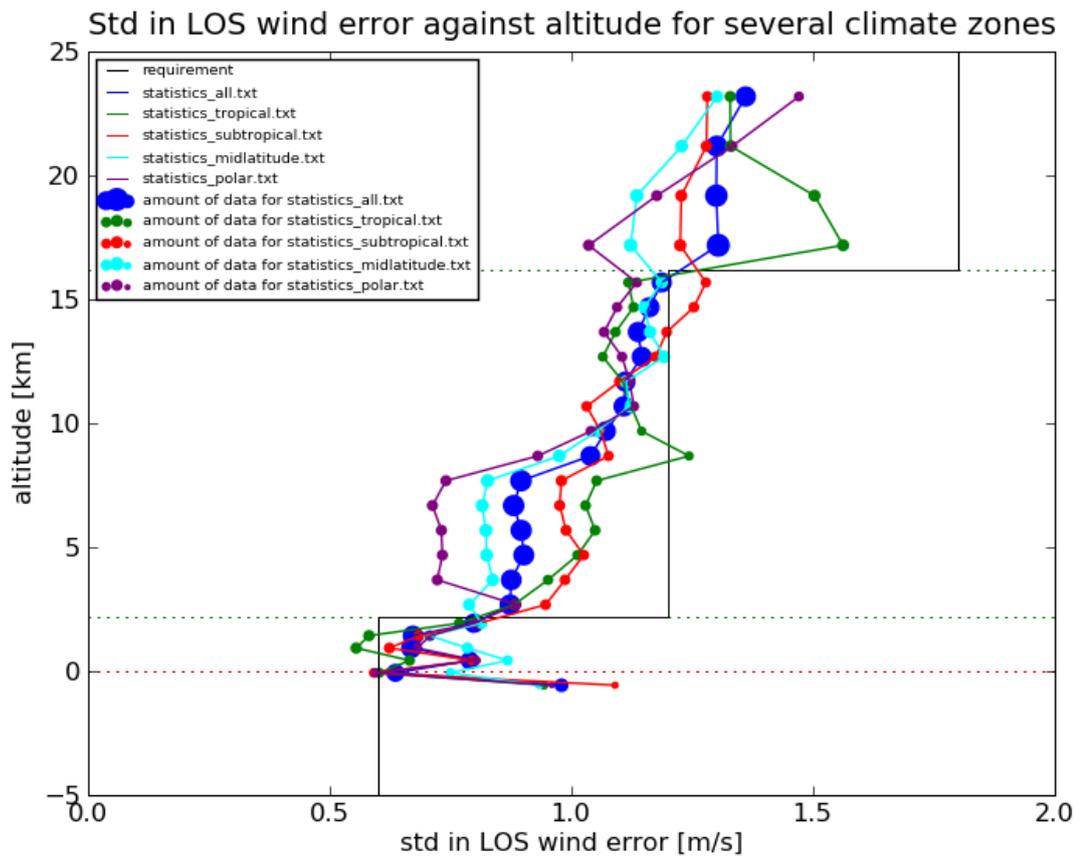


Figure 27: LOS wind error std for different climate zones against altitude. The blue dots and curve are for all data (global). The green, red, cyan and purple dots and curves are for tropical, subtropical, midlatitude and polar data, as is explained in the text. The black line indicates the requirement, the green dotted line a step in rangebin size and the red dotted line the ground surface. The dotsizes represent the amount of valid datapoints used to reconstruct the statistics.

The range bins between 2 and 16 km are mostly below the required 1.2 ms^{-1} in std of LOS wind velocity error. Only bin numbers 5 and 6 slightly exceed this requirement, so additional filtering is probably needed for these 2 range bins as well.

The bins above 16 km all perform very well, and are well below the required 1.8 ms^{-1} in std of LOS wind velocity error. For these rangebins no additional filtering is needed.

Another clear result from the numbers in this table is the fact that std of wind velocity error increases with range bin size. This is to be expected, because both height assignment problems and peak broadening due to transparent clouds show larger effects for larger range bins.

3.12 Task 4a: QC conclusions and recommendations

In section 3.11 the quality of the wind as retrieved by the Mie Core algorithm has been investigated as function of several of the Mie Core algorithm outputs. From this it can be concluded that:

- although (currently) not strictly a Mie Core output, screening on SNR is the easiest way to improve the average data quality. For all altitudes $\text{SNR} > 10$ remains a requirement to obtain good fit results.
- only data with a ValidityFlag set to true should be used

Table 7: Overview of the data quality for subsequent range bins. The only thresholds applied are $\text{SNR} > 10$ and $\text{ValidityFlag} = \text{True}$.

bin number	bin bottom	bin top	bin size	nr of datapoints	percentage	std in wind velocity error
All	-	-	-	734091	100%	1.193 ms^{-1}
1	22.18 km	24.18 km	2 km	30387	4.14%	1.551 ms^{-1}
2	20.18 km	22.18 km	2 km	33449	4.56%	1.499 ms^{-1}
3	18.18 km	20.18 km	2 km	38008	5.18%	1.604 ms^{-1}
4	16.18 km	18.18 km	2 km	39727	5.41%	1.630 ms^{-1}
5	15.18 km	16.18 km	1 km	29652	4.04%	1.313 ms^{-1}
6	14.18 km	15.18 km	1 km	30355	4.14%	1.288 ms^{-1}
7	13.18 km	14.18 km	1 km	31021	4.23%	1.245 ms^{-1}
8	12.18 km	13.18 km	1 km	29763	4.05%	1.246 ms^{-1}
9	11.18 km	12.18 km	1 km	30322	4.13%	1.252 ms^{-1}
10	10.18 km	11.18 km	1 km	30373	4.14%	1.267 ms^{-1}
11	9.18 km	10.18 km	1 km	28524	3.89%	1.221 ms^{-1}
12	8.18 km	9.18 km	1 km	28646	3.90%	1.174 ms^{-1}
13	7.18 km	8.18 km	1 km	34267	4.67%	1.022 ms^{-1}
14	6.18 km	7.18 km	1 km	33758	4.60%	0.994 ms^{-1}
15	5.18 km	6.18 km	1 km	33291	4.53%	0.999 ms^{-1}
16	4.18 km	5.18 km	1 km	32383	4.41%	1.004 ms^{-1}
17	3.18 km	4.18 km	1 km	31857	4.34%	0.971 ms^{-1}
18	2.18 km	3.18 km	1 km	32012	4.36%	0.950 ms^{-1}
19	1.68 km	2.18 km	0.5 km	28524	3.89%	0.849 ms^{-1}
20	1.18 km	1.68 km	0.5 km	31818	4.33%	0.707 ms^{-1}
21	0.68 km	1.18 km	0.5 km	31988	4.36%	0.706 ms^{-1}
22	0.18 km	0.68 km	0.5 km	25148	3.43%	0.811 ms^{-1}
23	-0.32 km	0.18 km	0.5 km	25907	3.53%	0.679 ms^{-1}
24	-0.82 km	-0.32 km	0.5 km	12911	1.76%	1.130 ms^{-1}

- a threshold on the residual error of 0.005 will select data in such a way that the overall STD of the wind velocity error will be below 0.75 ms^{-1} on average (but such a strict selection is not needed on the overall dataset). This flags about 50% of all data invalid. Note that this average value is much better than the LOS wind accuracy requirement for most of the atmosphere (1.2 ms^{-1} above 2 km and 1.8 ms^{-1} above 16 km), so it is not needed to apply this selection to meet these requirements, except for the PBL and maybe some of the highest 1 km range bins around 16 km.
- also a threshold region on the FWHM of 2.017 upto 2.059, combined with the threshold on residual error, will select data in such a way that the overall STD of the wind velocity error will be below the required 0.62 ms^{-1} on average. Again, this is not needed to meet the mission requirements on most altitudes, and may only be useful in the PBL and maybe some of the highest 1 km range bins around 16 km.

The simulations clearly show different behaviour for different altitude regions. This is largely caused by the varying rangebin definition. The variations in data quality for different altitudes but with the same range bin size is smaller and more gradual, while the variations between neighbouring rangebins with different bin size is very clear.

Looking at these results it is recommended to add the SNR, Residual Error, FWHM en Number of Iterations results of the Mie Core algorithm to the L2B product file, to allow the user to do his own data selection in case he needs a higher data quality on some or all altitude levels.

4 Task 4a: Rayleigh QC

4.1 Determining the impact of vertical shear in horizontal winds

To derive QC information related to vertical shear in horizontal winds within a range bin, the Rayleigh spectral shape must be determined with sufficient precision.

Some effort has been invested in understanding the current implementation of the Rayleigh channel simulation in both the E2S and LIPAS. From this it became clear that no wavelength dependant signal is to be expected on the different ACCD bins of the Fabri-Perot spectrometer. Therefore it is not possible to estimate the line width from these signals, and no possibilities were found to do quality control on the Rayleigh channel ACCD bin signals within an individual rangebin at measurement level.

In case when good quality absolute calibration for both the Mie and Rayleigh channel were available, the Rayleigh background from the Mie channel could in principle be used as 3rd spectral point, next to the channel A and B useful signals, and this would enable to estimate the Rayleigh line width, which then in turn might be used for QC. This would need a very precise cross-talk correction on the Rayleigh useful signals. Also, since the broadening of Rayleigh signal spectrum is a function of both temperature and wind shear, a relatively accurate knowledge of the temperature would be needed to derive the effect of wind shear on the signal intensities. This has not yet been simulated. [has this been tried with LIPAS?]

4.2 Determining the impact of horizontal variation in horizontal winds

Alternatively, the oversampling used by the instrument on observation level does provide information of the variability within the 50 km observation length for a given range bin. For cases without Mie signal this variability is an indication of varying Rayleigh peak width, which in turn may be caused by horizontal changes in wind conditions. It is assumed in this case that any temperature variation within the scene can be either neglected, or is known precisely in advance from NWP. Note however that, as soon as any aerosol or transparent cloud layer enters the scene, it will be extremely difficult to discriminate between the effect caused by this layer and the effect caused by the wind variability.

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. These algorithms are based on the the phase-A instrument hardware design and include the simulation of the emitted laser beam, atmospheric propagation, receiver blocking filters and detection of the atmosphere backscattered photons by the Fizeau (Mie) and Fabry-Perot (Rayleigh) interferometers on the ACCD, see also [LR6]. 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, i.e. the Fizeau and Fabry-Perot ACCD signals simulated by LIPAS should be compatible with the E2S and the HLOS wind quality simulated by LIPAS should be compatible with the L2B-processor output. 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 wind information content for Mie only, Rayleigh only and collocated Mie and Rayleigh sampling.

5.2 From LIPAS to LIPAS-2009

This section provides an overview of the main issues in the context of the LIPAS upgrade. A more detailed discussion of these issues is found in subsequent sections.

- *LIPAS atmosphere input.* For the description of the atmospheric scene as observed by ADM LIPAS input was limited to NWP model field parameters. NWP model fields do not include explicit atmosphere optical properties such as cloud, aerosol and molecular backscatter and extinction. Aerosol backscatter and extinction were simulated from climatology datasets and a random process to simulate backscatter variability in the vertical, e.g. [LR2]. NWP model fields also do not include cloud optical properties explicitly but implicitly through model cloud 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 in addition atmospheric profiles from the Level-2 atmospheric database can be used as input to define the atmospheric scene as observed by ADM. This atmospheric database includes explicit aerosol and cloud backscatter and extinction. Molecular backscatter and extinction are calculated from atmospheric pressure and temperature, see [LR2].
- *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 ACCD peak location that is related to the particles velocity and a Monte Carlo simulation was performed to generate a

lookup table that related Mie channel SNR to LOS wind accuracy. As part of the upgrade, the Fizeau interferometer parameters have been made compatible to the E2S and the more advanced so-called Mie Core algorithm for peak-fitting, that is also part of the L1B and L2B processors, has been implemented in LIPAS-2009 and tested against the previously implemented MMS 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, similar as in the L1B and L2B processors.
- *Cross-talk.* 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 explicitly simulates 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.
- *Atmosphere heterogeneity.* Initially LIPAS assumed homogeneous atmospheres and no wind biases in retrieved Mie and Rayleigh winds. Biases may be introduced by the atmosphere e.g. through cross-talk and atmosphere heterogeneity as well as by the hardware. During the MERCI [RD10] project, 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 calculated 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. It is assumed that biases induced by hardware are stable and can be corrected in the processing.
- *Wind error estimates.* Besides a wind observation, LIPAS provides an estimate of the wind error bias and standard deviation. These are needed to give the observation a proper weight in the atmospheric analysis during data assimilation. LIPAS already included instrument wind error estimates based on detected SNR of detected Fizeau and Fabry-Perot signals and the computation of a representativeness error that quantifies the mismatch of atmospheric scales observed by the instrument but that are not well resolved by NWP models. The representativeness error has been updated based on more recent ECMWF model spectra evaluations, see [RD9]. An additional error related to the unknown location of an aerosol or cloud layer within the vertical range bin has been introduced.
- *Quality control (QC).* Following the conclusions in section 3.12 output parameters of the Mie core algorithm are used to decide on accepting or rejecting Mie channel winds. Mie winds are accepted for i) SNR values larger than 10, ii) residual errors smaller than 0.005 and iii) a FWHM in the range 2.017-2.059. For Rayleigh channel winds quality control flags are set based on measured SNR and estimated cross-talk.
- *Hardware design parameters.* During the evolution from (pre-)phase A to phase B/C, some of the ALADIN hardware components have evolved. The corresponding parameters are updated and summarized in various tables that are discussed in the corresponding subsequent sections.

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 followed by wind retrieval from the measured data. The items described above are discussed in more detail in the corresponding sections.

5.3 Atmospheric return signal

This section describes the simulation of the atmospheric return signal that is detected on the instrument receiver. LIPAS extracts the atmospheric scene observed by ADM from the atmospheric database. For VAMP we use the same database as used to test the ADM L1B and L2B processors. The database contains profiles of pressure, temperature, zonal, meridional and vertical wind components, particle and molecular backscatter and extinction, all as a function of height from mean sea level to 40 km altitude at 125 m vertical resolution. The horizontal resolution is 3.5 km. The database profiles are interpolated to the range gates of the vertical sampling scenario. 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 and their spectral distribution is 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 [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 and 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 for an off-nadir pointing angle of 35 degrees.

The Doppler shifted wavelength, λ' , is obtained from the Doppler shifted frequency, 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 8. 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:

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

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

Table 8: Parameter values used for the ADM performance simulation. The ADM Phase-B design values for the Rayleigh equivalent and earth background equivalent bandwidths are obtained from an e-mail correspondence with ESA (Anne Grete Straume) on Wednesday 3 June 2009.

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 Mie receiver	T_p^{Fiz}	0.6	
peak transmission of the complete Rayleigh receiver	T_p^A, T_p^B	see Table 10	
equivalent spectral sampling of the detector on the Mie channel	$\Delta\lambda$	0.0394	[pm]
Rayleigh equivalent bandwidth on Fizeau ¹	$\Delta\lambda_{eq_ray}$	0.15	[pm]
earth background equivalent bandwidth on Fizeau ¹	$\lambda_{eq_bkg}^{Fiz}$	83.75	[pm]
earth background equivalent bandwidth on Fabry-Perot ¹	$\lambda_{eq_bkg}^{FP}$	297.5	[pm]
Fizeau ADC dynamic range	n_{ADC}^{Fiz}	74000	[el]
Fabry-Perot ADC dynamic range	n_{ADC}^{FP}	120000	[el]
Number of ADC bits	n_{Bits}	16	

- $R(z)$ 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.
- $T_{atm}(z)$ 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 .

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

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 8.¹ 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.

5.4 Mie channel signal and wind retrieval

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

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

¹This are phase B design values, based on Fizeau FSR = 2190 MHz and FWHM = 137 MHz. These values are no longer up to date in phase C/D.

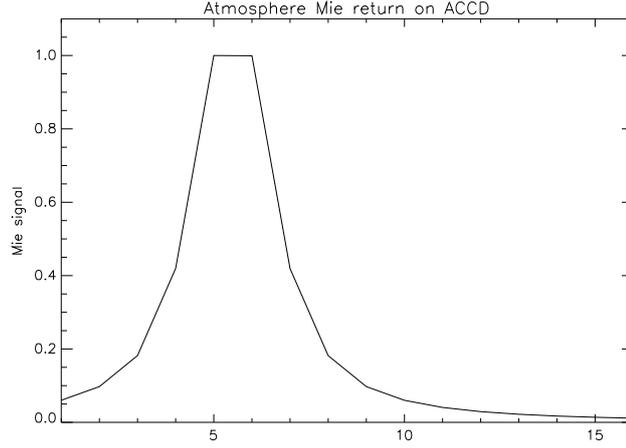


Figure 28: Distribution of the normalized atmosphere Mie peak on the Fizeau interferometer ACCD bins for a 50 ms^{-1} wind velocity. The ACCD bin numbers are on the horizontal axis.

with RS the unitless spectral response

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

Here, T is the Fizeau spectral transmission that is modelled as a Lorentzian, see Eq. (9), and with parameters defined in Table 9. $I_{\lambda_{\text{laser}}}$ is the normalized spectral distribution of the backscattered Mie signal, as discussed in the previous section. λ_i is the starting wavelength [pm] of channel i and $\Delta\lambda$ is the equivalent spectral sampling of the detector [pm], index i ranges 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, the same value as 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 28 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 bins indeed corresponds to a 50 ms^{-1} wind velocity i.e. wind velocity =

Table 9: 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	

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 MMS peak-fit method that is discussed later.

5.4.1 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,

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

with

- C_1^{e2s} a constant parameter used to make the signal intensity compatible with the E2S. This is discussed later in section 5.6
- η the detector quantum efficiency, see table 8
- 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^{Fiz} 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 8. 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

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

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

with

- C_2^{e2s} is a constant parameter used to make the signal intensity compatible with the E2S, as discussed later in section 5.6
- $S_{\text{tot.ray}}$ 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.
- $\lambda_{\text{eq.bkg}}^{\text{Fiz}}$ 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 8. No correction parameter is introduced for the earth radiance background term because its contribution to the background noise is generally negligible as compared to the Rayleigh contribution for the dawn-dusk ADM orbit. The total signal per ACCD bin i for the Mie receiver is obtained from

$$N_{\text{tot}}(i) = N_{\text{Fiz}}(i) + N_{\text{B-Ray}}(i) + N_{\text{B-BKG}}(i) \quad (24)$$

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 on-board accumulation.

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 measurement that is denoted with a prime: $N'_{\text{tot}}(i)$. As explained in section 7.9.10 of [RD3] the charge stored on each column of the ACCD is read out by analogue electronics, amplified and then digitised. The digitisation step includes the analog-to-digital conversion (ADC) as follows:

$$N_{\text{ADC}}(i) = \text{round}[N'_{\text{tot}}(i) * (2^{n_{\text{Bits}}} - 1) / n_{\text{ADC}}^{\text{Fiz}}] \quad (25)$$

The integer value $N_{\text{ADC}}(i)$ is the output of the FP detection system and is the result of the simulation of the ALADIN Fizeau spectrometer system. Typical values for the Fizeau spectrometer are found in Table 8.

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*. 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. 28. 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 Mie retrievals by this simple algorithm are still not free of biases as will be showed later in the results in section 5.6.

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

$$\hat{v}_{\text{LOS}} = -(L_c - 8.5) \times \text{USR}/16 \quad (26)$$

with USR the useful spectral range in ms^{-1} , e.g. 266.605 ms^{-1} for the value of 1502 GHz in table 9. The subsequent estimate for HLOS wind, \hat{v}_{HLOS} , is obtained from

$$\hat{v}_{\text{HLOS}} = \frac{\hat{v}_{\text{LOS}}}{\sin(\phi)} \quad (27)$$

Table 10: 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	

As an alternative to the retrieval algorithm described above, the potentially better and more advanced Mie Core has been implemented in LIPAS. The Mie Core algorithm detects the location of the Mie peak from the noisy Fizeau ACCD signal using a downhill simplex method as discussed in section 3.5. The Mie Core algorithm is also part of the ADM L1B and L2B processors. The MMS and Mie Core wind retrieval algorithms are tested in section 5.6 for a number of cases.

5.5 Rayleigh channel signal and wind retrieval

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 HLOS wind velocity accuracy (standard deviation) was implemented, see also Eq. (38). 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].

5.5.1 Fabry-Perot transfer function

Figure 29 shows the transmission curves of the direct (A), $\tau_{\text{FP}}^A(\nu)$, and reflected (B), $\tau_{\text{FP}}^B(\nu)$ 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, $\tau_{\text{Fiz}}(\nu)$.

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 30. The Fabry-Perot transfer function is then written as

$$\tau_{\text{FP}}(i, \nu) = T_p^A \tau_{\text{Fiz}}(\nu) \tau_{\text{FP}}^A(\nu) \tau_{\text{FP}}^{\text{A-relay}}(i) + T_p^B \tau_{\text{Fiz}}(\nu) \tau_{\text{FP}}^{\text{A-refl}}(\nu) \tau_{\text{FP}}^B(\nu) \tau_{\text{FP}}^{\text{B-relay}}(i) \quad (28)$$

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 10. Integrating $\tau_{\text{FP}}(i, \nu)$ for ν yields the plot in the right panel of Figure 30 and shows that the Fabry-Perot direct channel A gathers more signal than the reflected channel B.

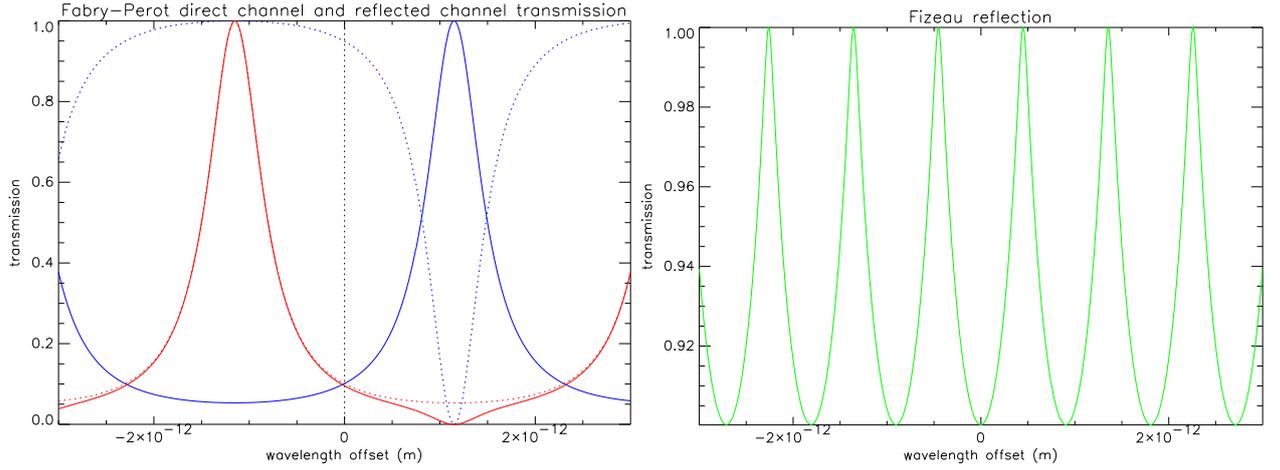


Figure 29: Fabry-Perot channel A (direct, blue) and B (reflected, red) transmission (left panel) and Fizeau reflection, $\tau_{Fiz}(\nu)$, (right panel). The blue solid and red dotted curves show the channel A, $\tau_{FP}^A(\nu)$, and B, $\tau_{FP}^B(\nu)$, transmission respectively as a function of 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.

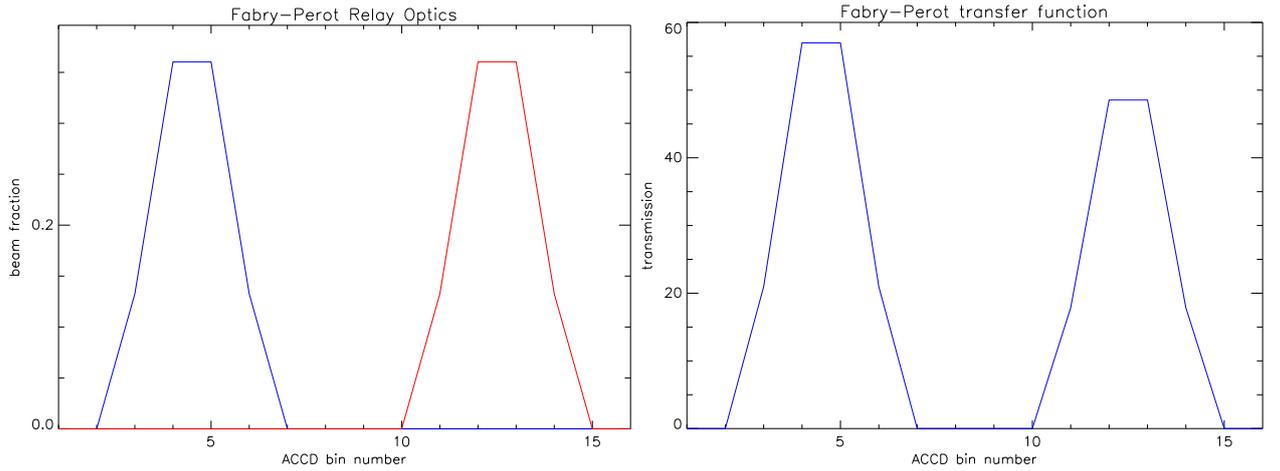


Figure 30: Fabry-Perot Relay to focus the channel A, $\tau_{FP}^{A-relay}$ and B, $\tau_{FP}^{B-relay}$, signals on the ACCD bins (left panel). The right panel shows the relative signal detected from the A (direct) and B (reflected) channel.

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

$$e_{ACCD}^{FP}(i) = \int_{\nu} \tau_{FP}(i, \nu) I_{\lambda_{laser}}(\nu) d\nu \quad (29)$$

with $\tau_{FP}(i, \nu)$ obtained from Eq. (28). Note that here $I_{\lambda_{laser}}(\nu)$ is the broad Rayleigh spectrum as defined in Eq. (10). Rather than computing the integral for each subbin and similar as for the Fizeau interferometer a lookup table for the Fabry-Perot transfer function is computed once in LIPAS-2009 by discretizing the frequency (or wavelength) range in 1201 steps, see Table 10. 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 5.5.3, 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. Following Eqs. (17,18) from MMS there are 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:

$$N_{\text{FP}}(i) = C_3^{e2s} \eta N_s S_{\text{tot.ray}} T_p^{\text{FP}} \frac{1}{2} e_{\text{ACCD}}^{\text{FP}}(i) \quad (30)$$

$$N_{\text{B,BKG}}(i) = \eta N_s S_{\text{BKG}} T_p^{\text{FP}} \Delta \lambda_{\text{eq_bkg}}^{\text{FP}} \frac{1}{2} \quad (31)$$

As for the Fizeau interferometer, no correction parameter is introduced for the earth radiance background term because its contribution to the total signal is generally negligible as compared to the Rayleigh contribution for the ADM dawn-dusk orbit.

5.5.2 Rayleigh channel Doppler shift estimate

The total signal on each of the 16 Fabry-Perot ACCD bins is obtained from

$$N_{\text{tot}}(i) = N_{\text{FP}}(i) + N_{\text{B,BKG}}(i) + N_{\text{xtalk}}(i) \quad (32)$$

The first two contributors are obtained from Eqs. (30,31), the third one is the contribution through cross-talk that is dicussed in the next section and for now assumed zero. 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 measurement. Next, the ACCD signal is digitised with an analog-to-digital converter (ADC) similar as explained in section 5.4.1 to yield $N_{\text{ADC}}(i)$ using a value for the ADC dynamic range as found in Table 8. Ignoring scene classification issues for the moment the measurements after digitisation are accumulated to obtain the signal for one observation. Here we note that the analog-to-digital conversion has no impact on Fabry-Perot receiver response, as discussed below, and the accuracy of retrieved winds. The latter is fully determined by the ACCD signal and additional Poisson noise. For this reason we continue the discussion using the ACCD signal $N'_{\text{tot}}(i)$ rather than $N_{\text{ADC}}(i)$ unless otherwise stated in the text.

The total Rayleigh signal on the Fabry-Perot direct channel, denoted A , and the total signal on the reflected channel, denoted B , are obtained from

$$A = \sum_{i=1}^8 N'_{\text{tot}}(i), \quad B = \sum_{i=9}^{16} N'_{\text{tot}}(i) \quad (33)$$

The receiver response, R , is defined as

$$R = \frac{A - B}{A + B} \quad (34)$$

The receiver response is a function of temperature, T in [K], that determines the width of the molecular spectrum, and the Doppler shifted frequency or wavelength ν in [m], that is related to mean LOS 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 Rayleigh 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 \times 3001. Figure 31 shows examples of the Rayleigh receiver response function and the receiver sensitivity as a function of temperature. The left panel shows steeper slopes for low temperatures resulting in larger wind errors than for high temperatures. This is also clear when considering the receiver sensitivity that is defined as $\partial \text{LOS} / \partial R$ and is an important measure that relates inaccuracies in the measured response (instrument noise, cross-talk, ...) to inaccuracies in the retrieved LOS wind velocity through (first order estimate) $d\text{LOS} \approx [\partial \text{LOS} / \partial R] dR$. From Eq. (34) we have for a deviation dR and its variance

$$\begin{aligned} dR &= \frac{\partial R}{\partial A} dA + \frac{\partial R}{\partial B} dB \\ \text{Var}[dR] &= \frac{4AB}{(A+B)^3} \stackrel{A \approx B}{\approx} \frac{1}{2A} \end{aligned} \quad (35)$$

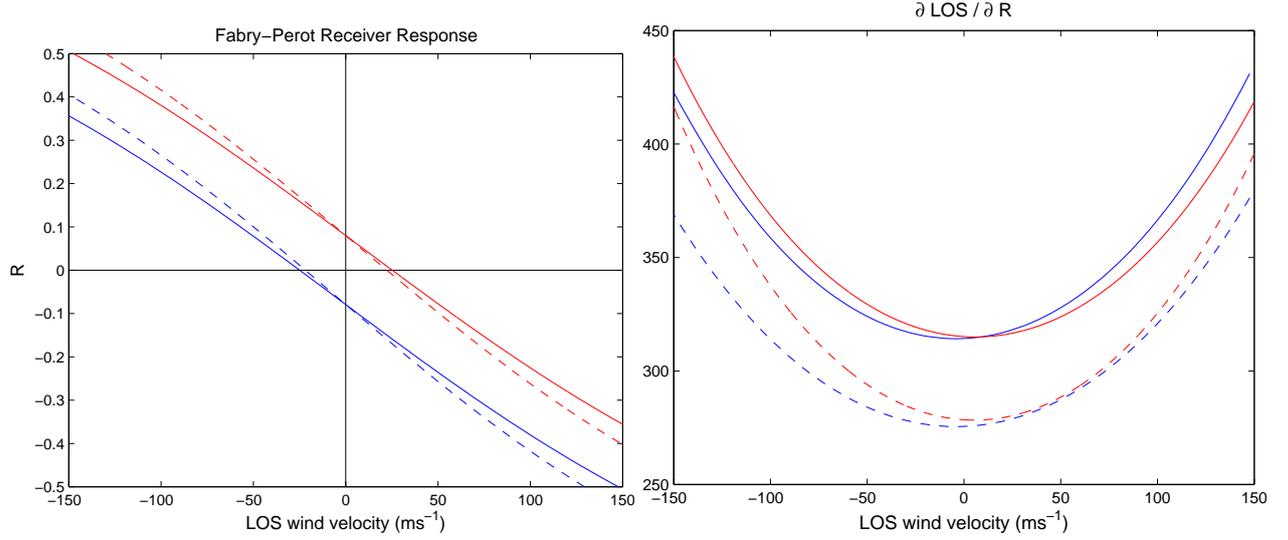


Figure 31: Rayleigh receiver response (left) and receiver sensitivity (right) for $T=273$ (K) (solid) and $T=210$ (K) (dashed). Red lines correspond to LIPAS-2009, blue lines to the L2Bp receiver response tables (file AE_TEST_AUX_RBC_L2_20071003T123456_20081002T123456_0001), L2Bp version CY35R2_Jun03, June 2009).

and thus for the (H)LOS wind error standard deviation

$$\frac{\partial \text{LOS}}{\partial R} = \frac{\partial \nu}{\partial R} \frac{c}{2\lambda} \quad (36)$$

$$\sigma_{\text{LOS}} = \frac{\partial \text{LOS}}{\partial R} \sqrt{\text{Var}[\text{d}R]} \quad (37)$$

$$\sigma_{\text{HLOS}} = \frac{\sigma_{\text{LOS}}}{\sin(\theta)} \quad (38)$$

The first equation above converts receiver response wavelength sensitivity to LOS wind sensitivity. For a given temperature (estimate), $\partial \nu / \partial R$ is obtained from the $R(T, \nu)$ lookup table. Thus the Rayleigh (H)LOS wind error standard deviation is a linear function of the receiver sensitivity and inversely proportional to the square root of the number of detected photons. From the right panel in Fig. 31 it follows that the receiver sensitivity decreases with decreasing temperature and thus increasing altitude (in general). On the other hand the number of molecules also decreases with altitude yielding a fairly constant (H)LOS wind error standard deviation throughout a particle-free atmosphere.

Figure 31 also shows that the slopes of the Rayleigh receiver response (left panel) is similar for LIPAS-2009 and the latest version of the L2Bp (June 2009). The mirroring in the origin is probably caused by a different definition in the LOS wind retrieval in the L2Bp and LIPAS-2009 (a minus when converting from frequency to wind used in the L2Bp and not in LIPAS-2009) and/or a different definition of the A (direct) and B (reflected) Rayleigh channels and their peak transmission. For a consistent use of defined parameters in the different processors and simulation tools this will not introduce biases in the retrieved winds. The receiver sensitivity in the right panel of Fig. 31 shows good similarity for high temperatures. LIPAS-2009 tends to overestimate the receiver sensitivity for low temperatures in particular for large negative winds.

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

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

From this equation we can write for ν :

$$\nu = \nu_0 + \frac{R(T, \nu) - R(T_0, \nu_0) - \frac{\partial R}{\partial T}(T_0)(T - T_0)}{\frac{\partial R}{\partial \nu}(\nu_0)} \quad (40)$$

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

$$\frac{\partial R}{\partial T}(T_0) \approx \frac{R(i+1) - R(i)}{T(i+1) - T(i)} \quad (41)$$

$$\frac{\partial R}{\partial \nu}(\nu_0) \approx \frac{R(j+1) - R(j)}{\nu(j+1) - \nu(j)} \quad (42)$$

The parameters on the right hand side of the equations above are obtained from the lookup table. Substituting the result in Eq. (40) yields the Doppler shift ν in [m]. The LOS wind component velocity is subsequently obtained from

$$\hat{v}_{\text{LOS}} = \frac{-\nu c}{2\lambda} \quad (43)$$

The HLOS wind component velocity then follows from Eq. (27).

Here we note that LIPAS does not take into account Brillouin scattering, in agreement with the E2S. Also the presented L2B results ignore the presence of Brillouin scattering.

5.5.3 Cross-talk

Particles in the measurement bin give a signal on the Fabry-Perot ACCD channels, also denoted cross-talk, yielding a bias in the Rayleigh wind because the receiver response inversion assumes a particle-free atmosphere. Cross-talk is simulated in LIPAS-2009. 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 spectrum 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. (29):

$$e_{\text{ACCD}}^{\text{FP-Xtalk}}(i) = \int_{\nu} \tau_{\text{FP}}(i, \nu) I_{\lambda_{\text{laser}}}(\nu) d\nu \quad (44)$$

$$= \int_{\nu} \tau_{\text{FP}}(i, \nu) \delta(\hat{\nu}) d\nu \quad (45)$$

$$= \tau_{\text{FP}}(i, \hat{\nu}) \quad (46)$$

Note that $I_{\lambda_{\text{laser}}}(\nu)$ now relates to the Mie spectrum. LIPAS-2009 applies a linear interpolation to the lookup table to determine $\tau_{\text{FP}}(i, \hat{\nu})$. Following Eq. (30) the total signal from cross-talk is obtained from

$$N_{\text{Xtalk}}(i) = C_4^{e2s} \eta N_s S_{\text{tot.mie}} T_p^{\text{FP}} \frac{1}{2} \tau_{\text{FP}}(i, \hat{\nu}) \quad \text{with } i=1, \dots, 16 \quad (47)$$

Effectively, cross-talk reduces the receiver response, Eq. (34), because the signal in channels A and B are increased with the same value leaving the numerator unchanged but increasing the denominator. In the extreme situation of infinite cross-talk, the receiver response equals zero yielding a bias equal to the difference between the true HLOS wind and the wind corresponding to a zero receiver response. The SNR on the other hand also goes to infinite yielding a zero wind error standard deviation. We further elaborate on this to understand the wind error statistics that is produced in section 6. Assume that the true global atmospheric wind, u_t , at a certain height bin is Gaussian distributed with mean \bar{u}_t and standard deviation σ_t . The measured wind may be expressed through $u_m(i) = [1 - f(\rho)] u_t(i) + \epsilon(i)$ with f a value between 0 and 1 as a function of the scattering ratio ρ and ϵ the wind instrument error. The latter is assumed Gaussian distributed with zero mean and standard deviation $\sigma_m(\rho)$ i.e. a function of the scattering ratio as discussed above, i.e. an increasing scattering ratio yields a larger cross-talk contribution to the Rayleigh channel signal thus increasing the SNR and reducing the wind error standard deviation. Hence $\sigma_m(\infty) = 0$.

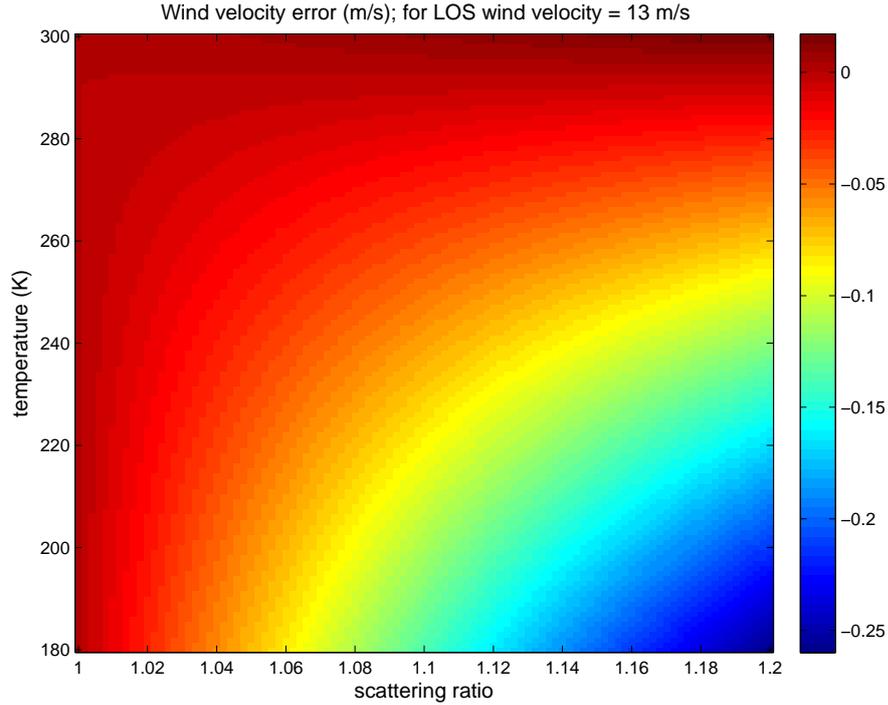


Figure 32: Rayleigh LOS wind error bias (ms^{-1}) through Mie contamination and as a function of temperature (K). The true LOS wind is 13 ms^{-1} . The scattering ratio is a measure of Mie contamination in the Rayleigh channel signal.

The scalar f depends on the instrument design and fulfills $f(1) = 0$ and $f(\infty) = 1$, i.e. for infinite scattering ratio and thus infinite cross-talk contribution, the receiver response equals zero. To simplify the discussion we assumed here that a zero receiver response corresponds to a zero wind which is in general not true, see Fig. 31. Using the formulas above we can write for the Rayleigh wind error

$$\begin{aligned} du &= u_m(i) - u_t(i) \\ &= -f(\rho)u_t(i) + \epsilon(i) \end{aligned} \quad (48)$$

The expectation value (bias) and variance of the Rayleigh wind error du equal $E[du] = -f(\rho)\bar{u}_t$ and $\text{Var}[du] = f(\rho)^2\sigma_t^2 + \sigma_m(\rho)^2$ respectively when assuming uncorrelated wind and instrument error statistics. For a scattering ratio with value 1 (molecules only, no particles) the bias equals zero and the wind error variance equals the instrument error, both independent of the wind statistics. For infinite scattering ratio the bias and variance equal \bar{u}_t and σ_t^2 respectively, i.e. fully determined by the wind statistics and not by instrument errors. Hence, the Rayleigh wind bias can still be very small, depending on the wind statistics. Also the Rayleigh wind error standard deviation can be small for small wind variability. In the extreme case of a globally non-changing wind in a certain height bin, $\sigma_t^2 = 0$ and the Rayleigh wind error standard deviation equals zero for extreme cross-talk contribution! If in addition the non-changing wind value equals zero, i.e. $\bar{u}_t = 0$, then the Rayleigh wind error statistics have zero bias and standard deviation, i.e. cross-talk contributes to the Rayleigh wind quality. This is however an unrealistic theoretical situation. In general the Rayleigh wind error statistics is a mixture of instrument error and observed wind statistics. The scattering ratio and subsequent cross-talk in the Rayleigh channel signal determines the relative contribution of both contributors.

Figure 32 shows the sensitivity of the Rayleigh wind error to the scattering ratio, that is a measure of Mie contamination. For a scattering ratio of 1 (no particles) the wind error equals zero. For increased scattering ratio the wind error bias increases, in particular for low temperatures. The Rayleigh wind bias increases substantially in case of both low temperatures and large scattering ratio values. For instance Polar Stratospheric Clouds (PSC) are formed when the temperature drops below -80 degrees Celsius. The

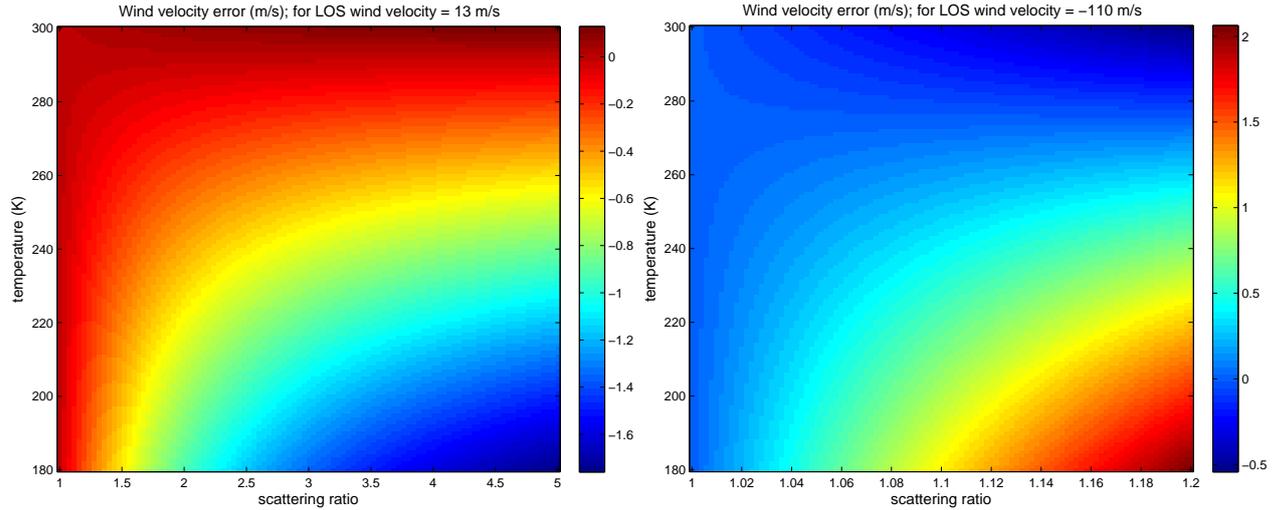


Figure 33: Same as Fig. 32 but now for larger scattering ratio values up to 5 (left panel) and an extreme large wind of -110 ms^{-1} (right panel).

typical scattering ratio for PSC's is 1.1 but values above 1.2 are often found, see [RD7]. Yet, wind biases are still below the mission bias requirement of 0.4 ms^{-1} . However, larger Rayleigh wind biases exceeding the mission requirement are found for large aerosol and cloud events in combination with low temperatures, e.g. tropospheric cirrus in the Tropics with scattering ratio values that can easily exceed 5, see the left panel in Figure 33. Also for larger wind velocities the absolute bias increases, see the right panel of Fig. 33, but the relative error stays at the same order of magnitude.

5.6 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. Here we note that LIPAS uses a different model, following MMS, for the simulation of the background terms than the E2S, i.e. the Rayleigh and earth radiance background, Eqs. (22,23) respectively for the Mie channel and the earth radiance background, Eq. (31), for the Rayleigh channel. As a result the background levels as simulated by LIPAS and the E2S may differ. The earth radiance background was switched off in the E2S and LIPAS-2009 for a useful comparison. Also, the E2S simulates a tripod obscuration, that is not available in LIPAS. The tripod obscuration was therefore switched off in the E2S simulation. For each case one observation (BRC) was simulated that is composed of 14 measurements. Each measurement is the result of 50 (on-board) accumulated shots. The constant parameters C_i^{e2s} were introduced in the equations in the previous sections to correct for small inconsistencies between the E2S and LIPAS. The values of these parameters are determined by comparing the Fizeau and Fabry-Perot ACCD signals for the cases in the following sections.

5.6.1 LIPAS-2009 E2S comparison; case 1 – cloud-free atmosphere

This case is based on the L2Bp test scenario 006, i.e. a sub-arctic winter atmosphere with the RMA aerosol backscatter profile. The HLOS wind is set to a constant value of 0 ms^{-1} throughout the profile. No instrument noise was applied and no solar background was assumed. Figure 34 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 different shapes of the E2S signals in the first couple of bins above 18 km is probably caused by rounding photon counts to integer values. LIPAS does not include rounding to integer values. To get similar amplitudes for the E2S and LIPAS-2009 signal the following correction parameter were applied: $C_1^{e2s} = 0.9$, $C_2^{e2s} = 0.9$. These values were obtained by trial and error to get a good match between the

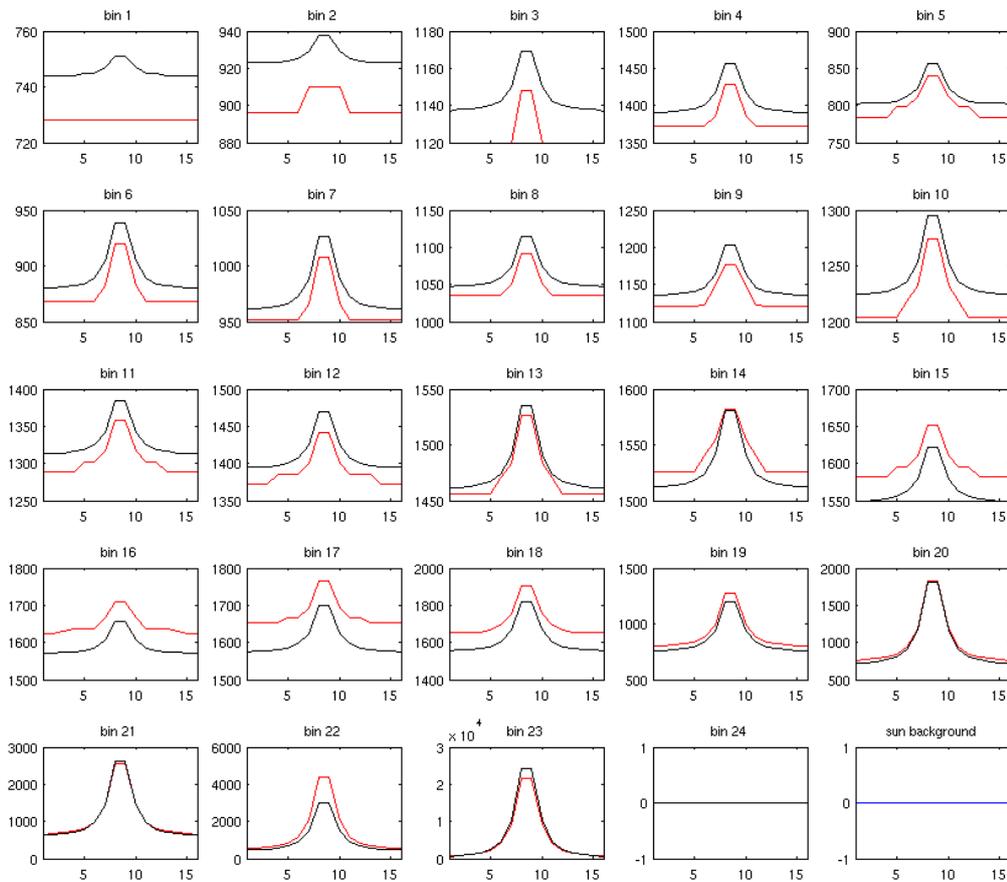


Figure 34: Case 1. Total signal from 14 accumulated measurements on the 16 ACCD bins , after digitisation, of the Fizeau interferometer for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. The blue curve in the lower right panel is the E2S earth radiance background signal. Bin 24 is below the surface and contains no signal in both the E2S and LIPAS. See the text for further details.

E2S and LIPAS signals. Figure 35 shows the signal on the 16 ACCD bins of the Fabry-Perot interferometer for 24 range gates. Range gate 24 contains the ground return. The following correction parameter were applied: $C_3^{e2s} = 1.0$, $C_4^{e2s} = 1.0$. Overall, LIPAS-2009 compares well with the E2S within 5% for the Fizeau signal and within 2% for the Rayleigh signal, see Fig. 36. A large difference is found in bin 22 for the Fizeau and bin 23 for the Fabry-Perot, i.e. corresponding to the 500 m range gate positioned between 155 and 655 m above mean sea level in sampling scenario WVM1. This difference is explained by the use of a digital elevation model (DEM) as part of the orbit simulation in the E2S. At the start of the E2S simulation the initial orbit location is determined based on the E2S input start time parameter, in this case (64N, 261E) i.e. over the North American continent. The surface in this area is not flat and the DEM computes a mean surface height for each of the 14 measurements that varies between 37 m and 195 m over the 50 km sampling length. The surface height of one of the 14 measurements exceeds the range gate bottom value of 155 m (namely 195 m). This bin thus contains the earth surface for this particular measurement which explains the larger signal in the E2S than in LIPAS for this range gate, because LIPAS does not include a DEM and obtains the orography from the input atmospheric database.

Other differences between the E2S and LIPAS include a varying local incidence angle with altitude in the

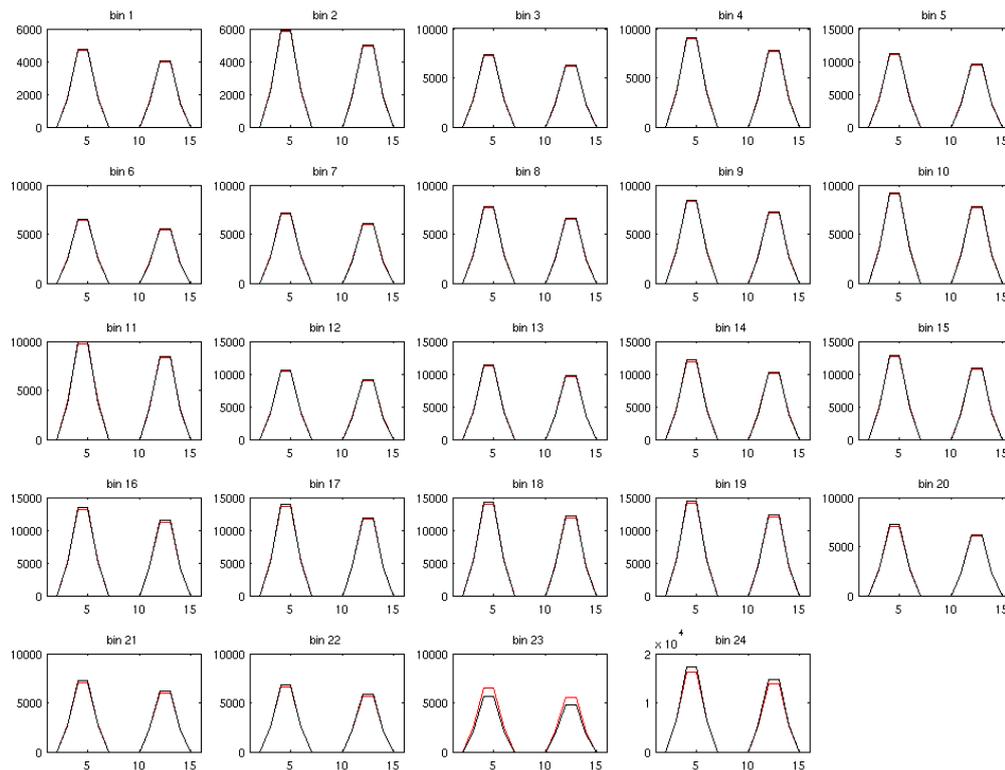


Figure 35: Case 1. Signal , after digitisation, on the 16 ACCD bins of the Fabry-Perot interferometer for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. See the text for further details.

E2S and different interpolation strategies in the E2S and LIPAS. Their implications are further discussed for case 2.

5.6.2 LIPAS-2009 E2S comparison; case 2 – single cloud layer

This case is similar to the previous one but now based on L2Bp test scenario 002, i.e. a mid-latitude winter atmosphere with the RMA aerosol backscatter profile. A constant HLOS wind of -50 ms^{-1} was imposed throughout the profile. In addition an optically very thin cloud layer between 12 and 14 km was included. No instrument noise was applied. Figure 37 shows a comparison between E2S and LIPAS used aerosol and molecular backscatter in the left panel and computed total attenuated backscatter in the right panel. The backscatter profiles almost completely overlap. The E2S extrapolates backscatter below the mean sea level, while LIPAS stops at mean sea level. The attenuated backscatter profiles in the right panel also seem to overlap, but the left panel in Fig. 38 reveals an almost 2% underestimate in the E2S (w.r.t. LIPAS) in the free-troposphere that further decreases to about 7% in the PBL. Considering a perfect backscatter match between the E2S and LIPAS, this deviation can only be explained by differences in the transmission computation. This is confirmed by the right panel in Fig. 38 that shows an underestimate of the transmission in the E2S (w.r.t. LIPAS) of between 0.5 and 1.5%. The large spikes near 14 and 12 km in the left panel correspond to the cloud top and bottom that not fully correspond in the E2S and LIPAS probably because of different atmospheric interpolation strategies.

Figure 39 is similar to Figure 34 but the cloud clearly shows up in bins 7-9. Also the peak center is shifted due to the non-zero wind velocity. Interestingly, the shape of the E2S signal in bin 22 differs substantially from LIPAS. Also, the signal amplitude is substantially higher. This deviation between the E2S and LIPAS

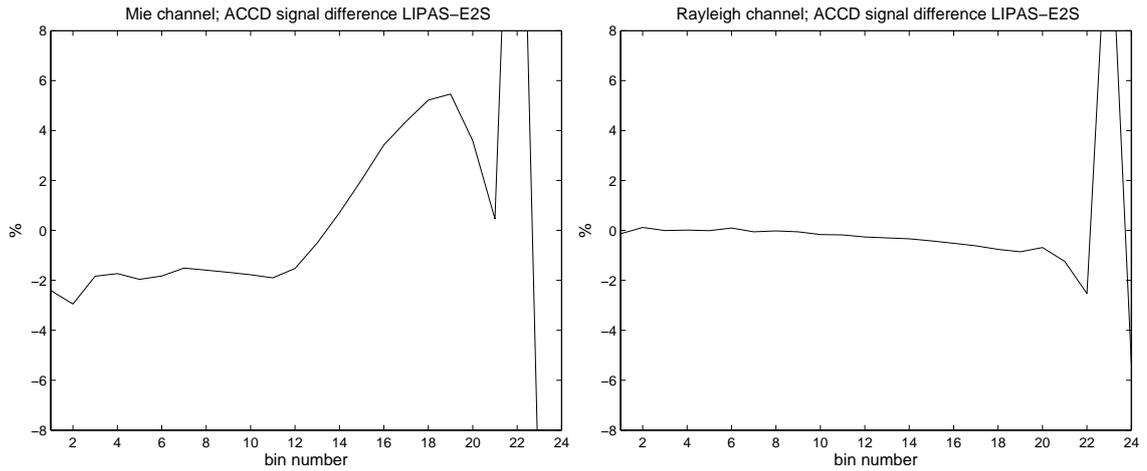


Figure 36: Relative difference between the E2S and LIPAS-2009 Fizeau (left) and Fabry-Perot (right) ACCD signals. The large difference in Fizeau bin 22 and Fabry-Perot bin 23 is explained in the text.

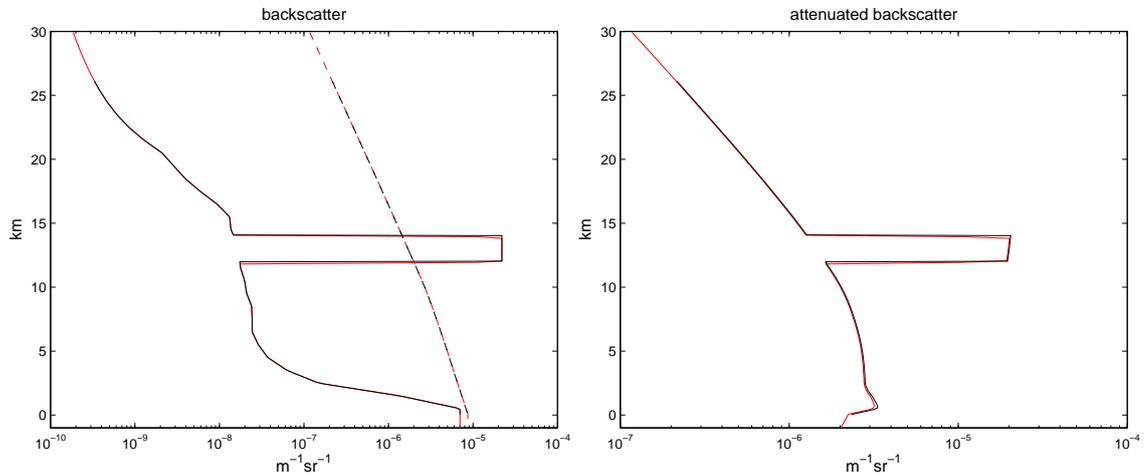


Figure 37: E2S (red) and LIPAS-2009 (black) backscatter (left) and attenuated backscatter (right).

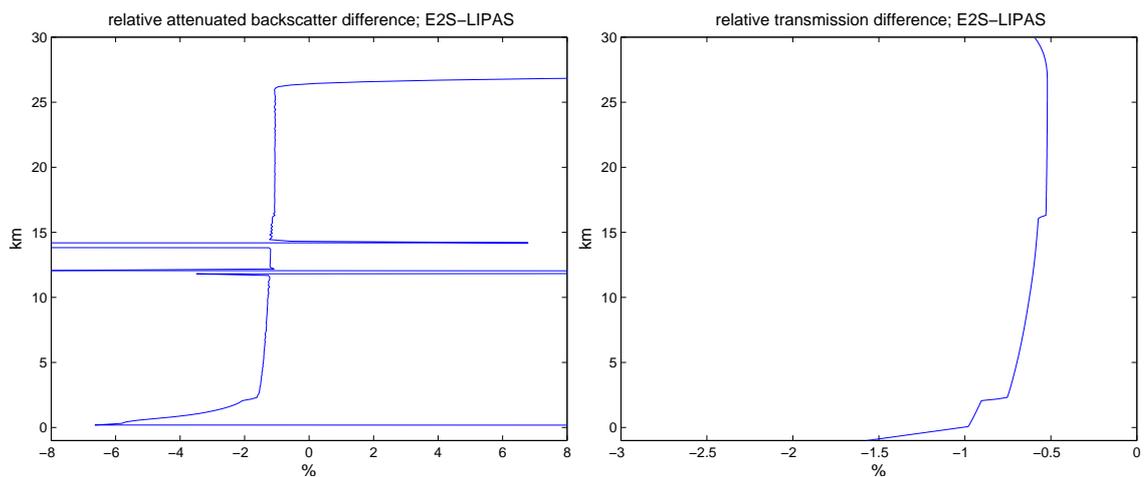


Figure 38: Relative difference between E2S and LIPAS-2009 for attenuated backscatter (left) and one-way molecular transmission (right).

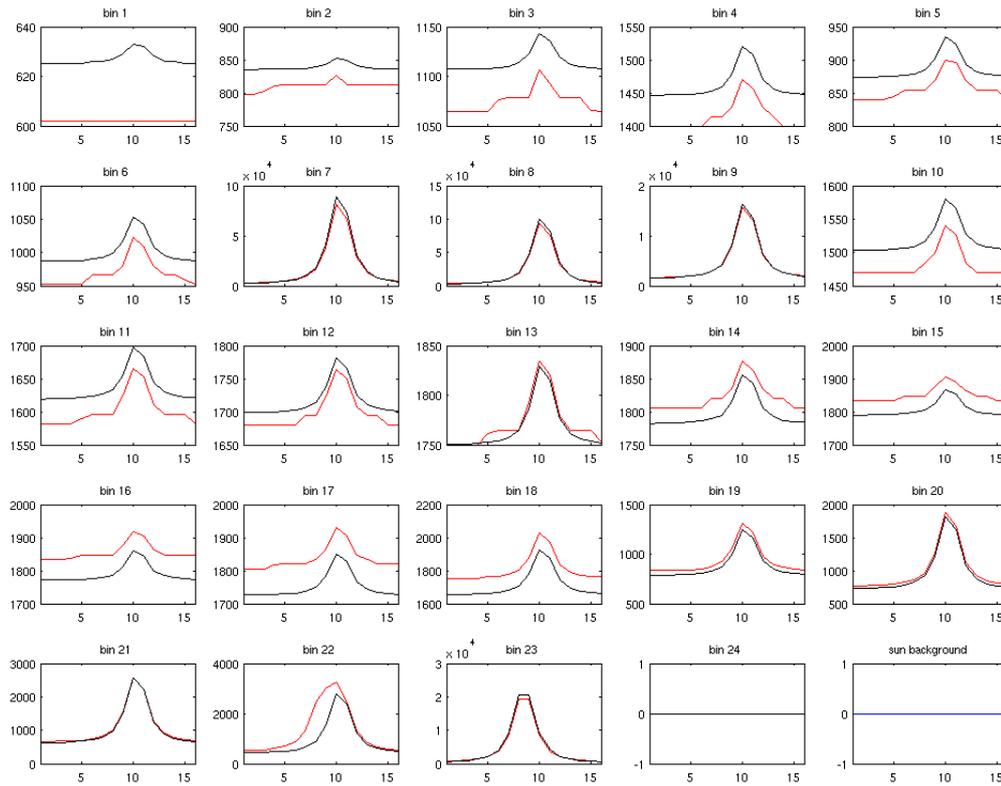


Figure 39: Case 2. Signal on the 16 ACCD bins of the Fizeau interferometer , after digitisation, for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. The blue curve in the lower right panel is the E2S earth radiance background signal. See the text for further details. Bins 7-9 clearly show an increased signal from cloud particles.

is explained by the ground return of one of the measurements in the E2S simulation as discussed in the previous section. The accumulated E2S signal is a mixture of -50 ms^{-1} winds in the PBL and a 0 ms^{-1} ground return which explains the different shape and signal strength. Figure 40 shows the signal on the 16 ACCD bins of the Fabry-Perot interferometer for 24 range gates. Range gate 24 contains the ground return.

No further effort has been spent to explain the attenuated backscatter differences between the E2S and LIPAS because this would require a complete evaluation of the E2S transmission computation implementation. The resulting differences between measured Fizeau and Fabry-Perot in Fig. 41 are within 6% for the Fizeau and 1% for the Fabry-Perot which is sufficient for VAMP purposes. The local negative 'bump' in bins 7-9 in the left panel and positive 'bump' in bins 8-10 in the right panel is explained by the small mis-match of the cloud location in the E2S and LIPAS.

5.6.3 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.

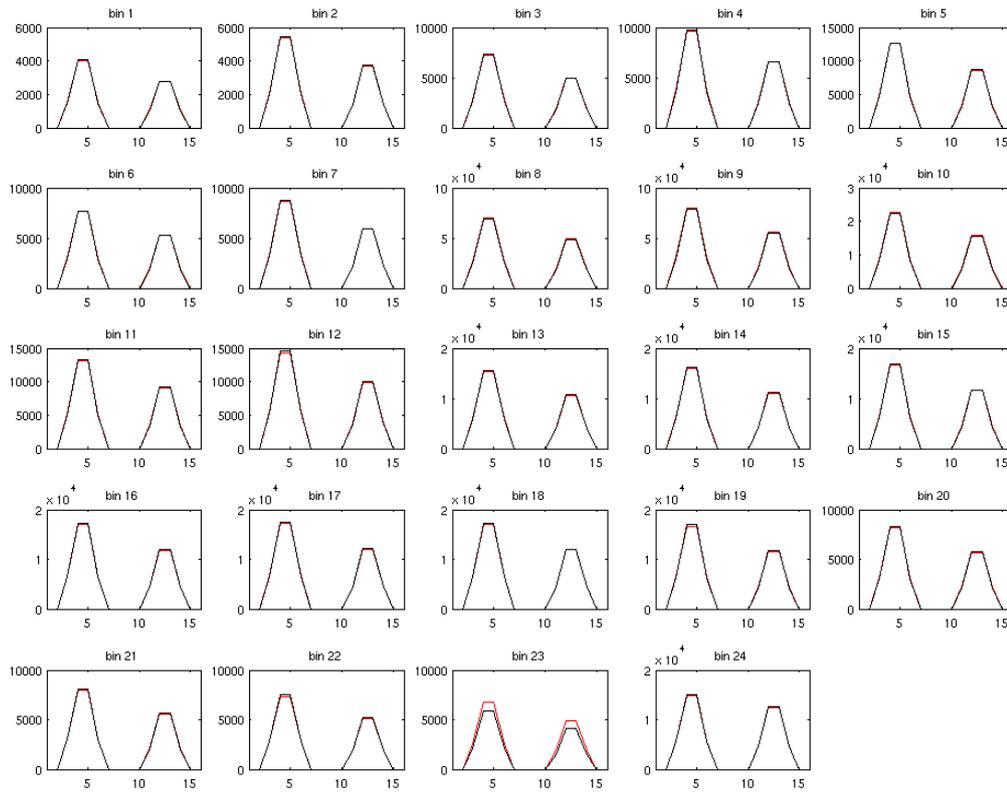


Figure 40: Case 2. Signal on the 16 ACCD bins of the Fabry-Perot interferometer, after digitisation, for 24 range gates. Red/black lines denote the E2S/LIPAS-2009 result. Bins 8-10 clearly show an increased signal due to cross-talk from cloud particles. See the text for further details.

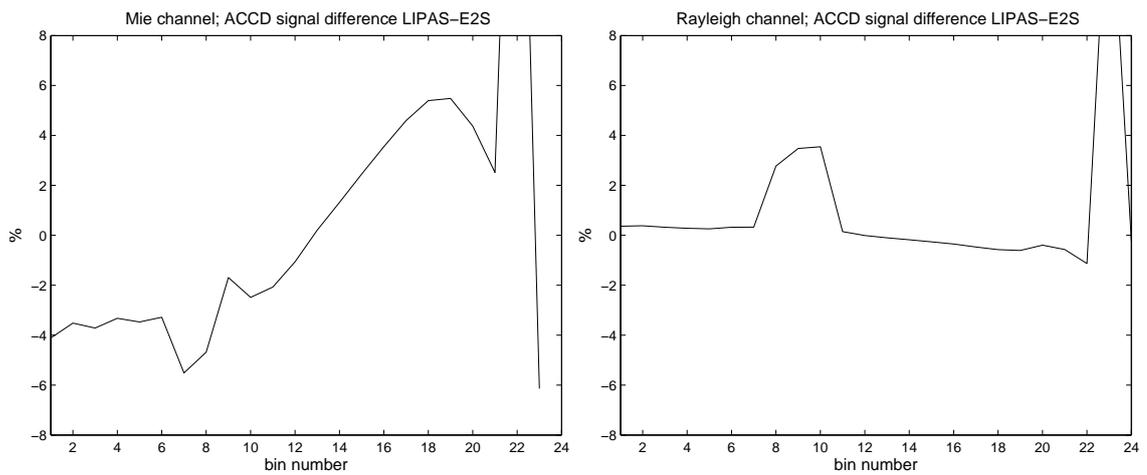


Figure 41: Same as Fig. 36 but now for case 2.

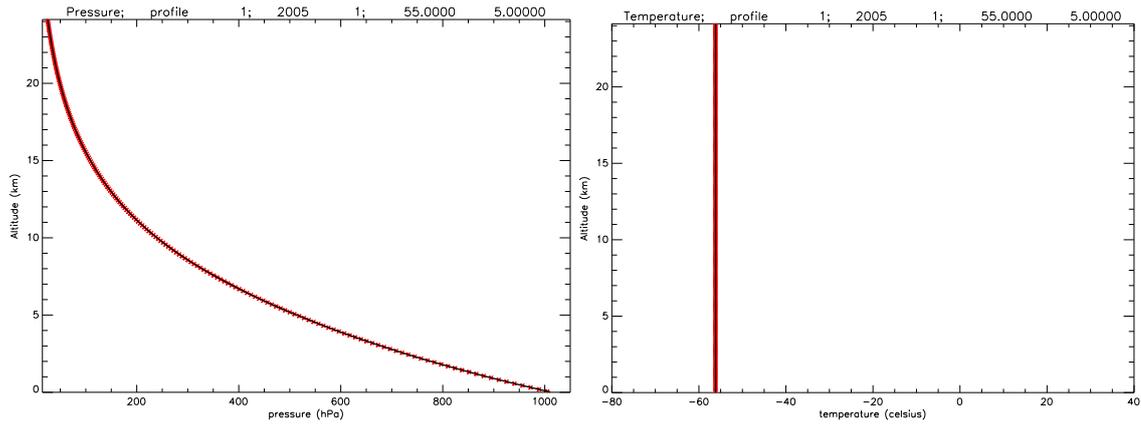


Figure 42: Atmospheric pressure (left) and temperature (right)

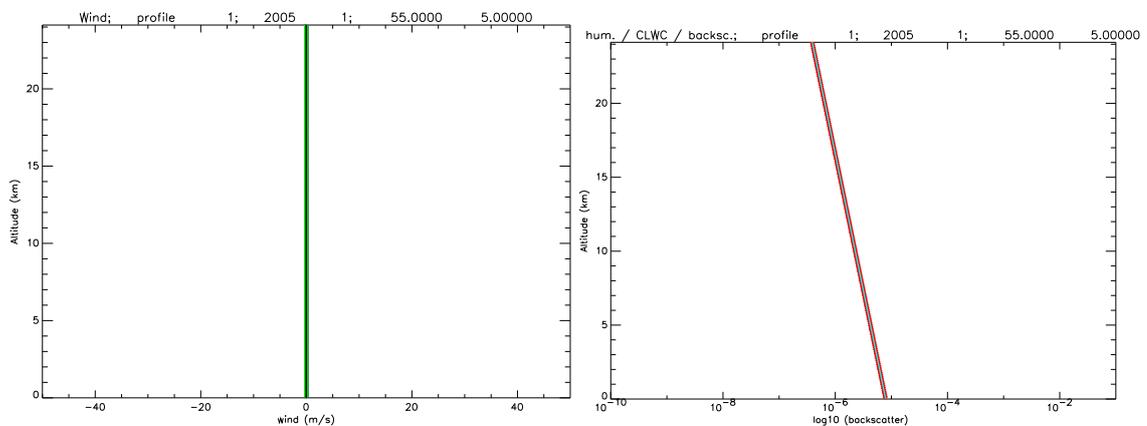


Figure 43: Atmospheric wind (left) and molecular backscatter (right). There is no aerosol backscatter.

case 1

This case is based on a synthetic atmosphere with a constant temperature, a constant HLOS wind of 0 ms^{-1} and no aerosols, see Figures 42, 43. 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 44 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 45 shows a negligible bias for Rayleigh winds. The bias in Mie winds is substantially larger because of missing aerosols. The Mie ground return standard deviation is close to zero. The standard deviation of Rayleigh winds is between 1.1 and 2 ms^{-1} .

An important note here is that the zero wind bias for Rayleigh channel winds can be achieved only if LIPAS performs the calculations with floating point numbers. Implementation of the rounding operator in Eq. (25), to yield integer numbers for the number of detected electrons on the ACCD, results in a small deviation of the actually acquired A and B signals and thus in a slightly different value for the receiver response. The left panel of Fig. 46 shows the total of number of accumulated electrons in the topmost bin

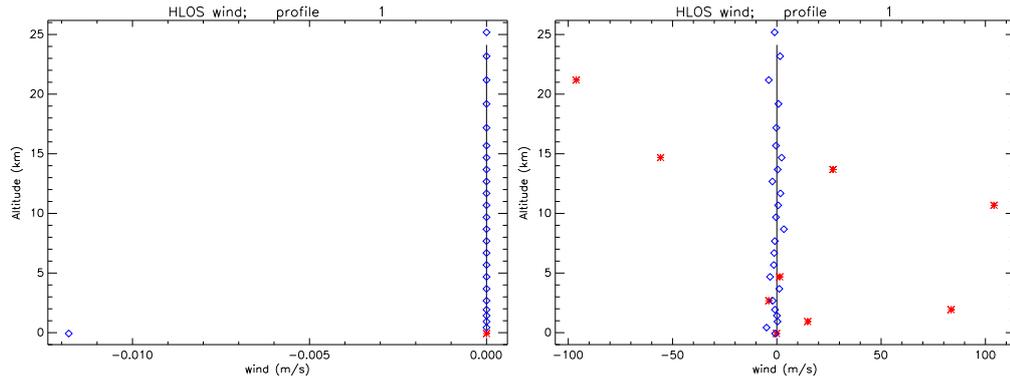


Figure 44: Single case HLOS wind retrieval from Rayleigh (blue diamond) and Mie (red star) channel. The left panel shows the result for a noise-free instrument, the right panel does include instrument noise.

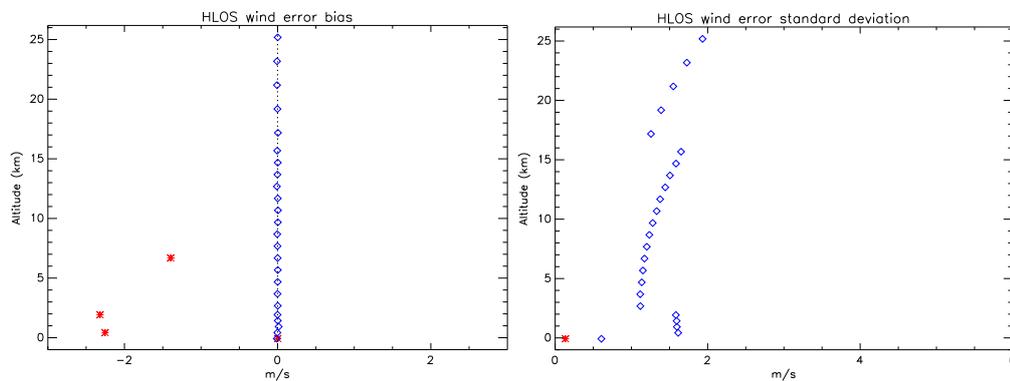


Figure 45: HLOS wind retrieval statistics for Rayleigh (blue diamond) and Mie (red star) channel from 100000 simulations. The left panel shows the bias, the right panel the standard deviation. The standard deviation for Mie winds above the surface is close to 85 ms^{-1} (not shown)

at 25 km altitude for one observation with (red) and without (black) using the rounding operator. Both curves almost overlap but the difference curve in the middle panel shows a larger deviation in channel A (leftmost 8 bins) than in channel B (rightmost 8 bins). As a result the receiver response value changes from 0.0800233 (no rounding) to 0.0792540 (rounding), i.e. a 0.96% change, corresponding to a 0.36 ms^{-1} HLOS wind bias as seen from the topmost bin in the right panel of Fig. 46. The rounding effect induces a bias in the Rayleigh winds of a couple of tenths ms^{-1} .

case 2

This case is similar to the previous case but now with a non-constant temperature, a constant HLOS wind of -50 ms^{-1} , aerosols and an optically thin cloud layer between 12 and 14 km, see Figure 48. The purpose of this test was to test the impact of cross-talk on the retrieved Rayleigh winds and the quality of Mie 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 48 shows good quality (standard deviation) Mie winds only in the first kilometer above the surface and inside the cloud layer. The relatively simple retrieval algorithm described in section 5.4.1 shows much larger biases than the Mie core algorithm both throughout the atmosphere including the particle intensive cloud layer and PBL. Also, the ground return bias, caused by aerosols with large wind velocities near the surface, is smaller for the Mie Core algorithm. 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, but the Mie wind yields a good alternative here.

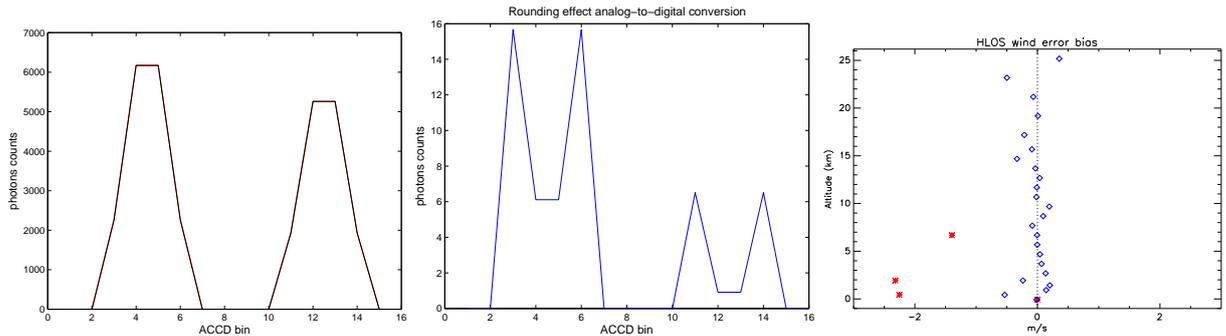


Figure 46: Rounding effect from floating point to integer numbers of electrons. Left, the number of electrons on the ACCD with (red) and without (black) the rounding operator. Middle, the difference of the curves: black minus red. Right, corresponding Rayleigh wind bias (blue diamonds).

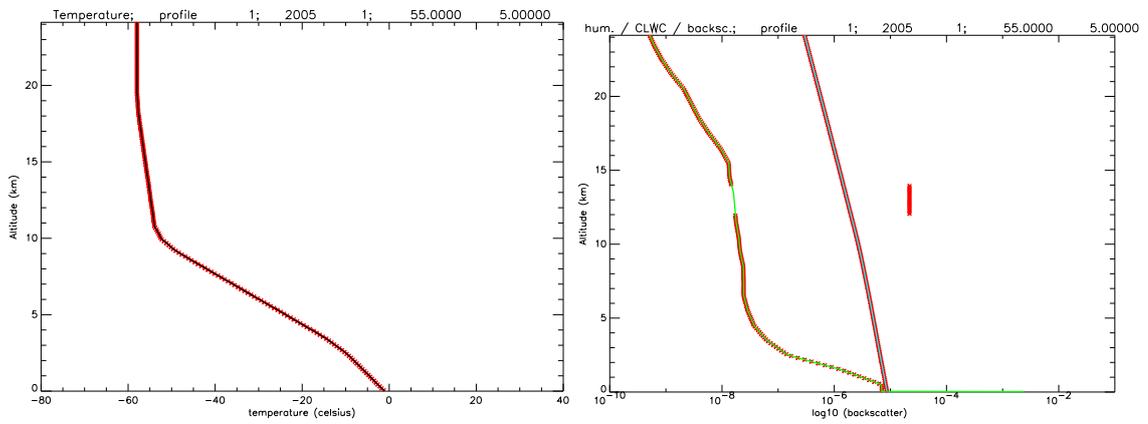


Figure 47: Atmospheric temperature (left) and aerosol (right). The left curve/middle/right shows aerosol/molecular/cloud backscatter.

5.6.4 LIPAS-2009 E2S-L1B-L2B comparison; photon count noise

In this section the quality of Mie and Rayleigh winds as obtained from the E2S-L1B-L2B chain and LIPAS are compared. Hereto HLOS wind profiles along one (nighttime, i.e. half) CALIPSO orbit were determined. The atmospheric input for both the E2S and LIPAS is identical and obtained from the atmospheric database that includes realistic atmosphere optical properties from CALIPSO and dynamics from ECMWF as described in [RD7]. Atmospheric data from 1 January 2007 along the first CALIPSO orbit was used. Vertical sampling scenario WVM1 as defined in [RD6] is used. This results in 94 wind profiles (BRCs) for ADM-Aeolus operated in burst mode. Figure 49 shows the wind statistics for Fizeau and Fabry-Perot signals simulated with the E2S and processed by the L1B and L2B processors. The bias and standard deviation of L1B Rayleigh winds is substantially larger than for L2B winds as expected due to temperature correction in the latter. Mie wind biases are similar for both processors but the standard deviation is generally larger for the L2Bp. This is unexpected because both processors use the same Mie Core algorithm for peak fitting. At the moment of simulation there were unexplained difficulties with calibration files that could explain the discrepancy for the Mie winds.

In a similar experiment LIPAS was used to simulate the Fizeau and Fabry-Perot signals and proces the data to obtain wind profiles. The burst mode was turned of in this experiment resulting in about 4 times more wind profiles along the CALIPSO than in the E2S-L1B-L2B experiment. The statistics are displayed in Fig. 50 and shows slightly smaller biases than the E2S-L1B-L2B chain that might be explained by the smaller amount of data in the latter. The Mie wind standard deviation is around 1 ms⁻¹ for LIPAS and slightly better for L1B Mie winds in the free troposphere. L2B Mie wind error standard deviations are

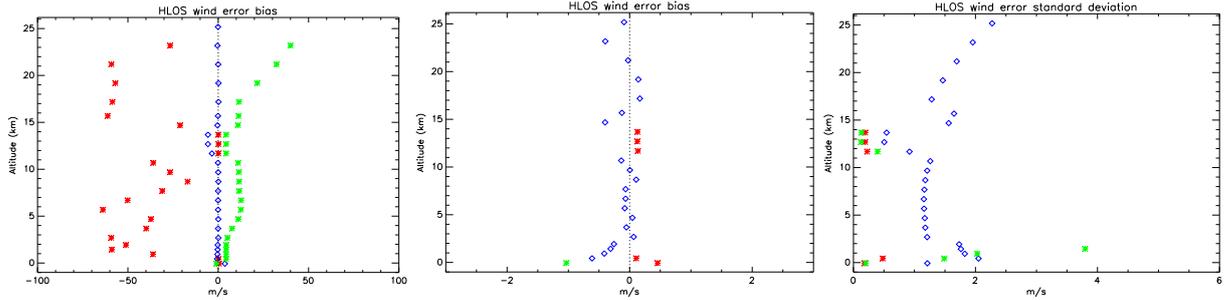


Figure 48: Mie (green/red) and Rayleigh (blue) HLOS wind retrieval statistics from 100000 simulations. The left/middle (zoomed in) panel show the bias (including the rounding effect explained in Fig. 46), the right panel the standard deviation. The green symbols are obtained from the simplified Mie retrieval algorithm described in MMS, the red symbols are obtained with the Mie Core algorithm.

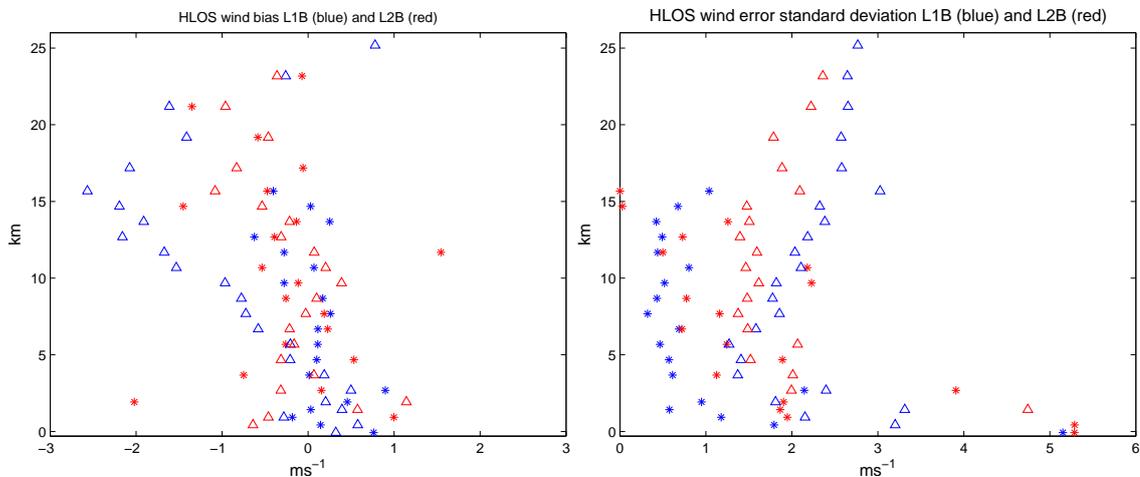


Figure 49: L1Bp and L2Bp HLOS wind error statistics obtained from 94 BRCs (observation profiles) along one CALIPSO orbit; error bias (left) and error standard deviation (right). Blue and red symbols correspond to L1Bp and L2Bp winds respectively with stars corresponding to Mie winds and triangles to Rayleigh winds. The E2S was used to simulate the noisy Mie and Rayleigh channel signals.

substantially larger as explained above. Rayleigh wind error standard deviations are close to 1.5 ms^{-1} for both L2Bp and LIPAS in the free-troposphere and increase to about 2.5 ms^{-1} in the stratosphere at 25 km. In both experiments the discontinuity at 16 km is clearly visible that marks the change of 1 km bin sizes (below 16 km) to 2 km bin sizes (above 16 km). The quality of Rayleigh winds in the PBL, below 2 km, is generally poor because of small 500 m bin sizes.

The number of used data in the LIPAS statistics is displayed in Fig. 51. The number of Mie winds decreases with altitude, because of decreasing aerosol density and cloud coverage with altitude. The number of Rayleigh winds on the other hand increases with altitude because of clouds obscuring (partly) the lower part of the atmosphere and as such reducing the number of good-quality Rayleigh winds in the lower part of the atmosphere.

5.7 Wind observation representativeness error

NWP models expect the wind observation to be representative for the mean wind inside the observation volume. For heterogeneous atmospheres with non-uniformly distributed optical properties inside the observation volume, the measured wind may substantially deviate from the mean wind in case of substantial horizontal and/or vertical wind-shear. The weight of such observations in the analysis should therefore be

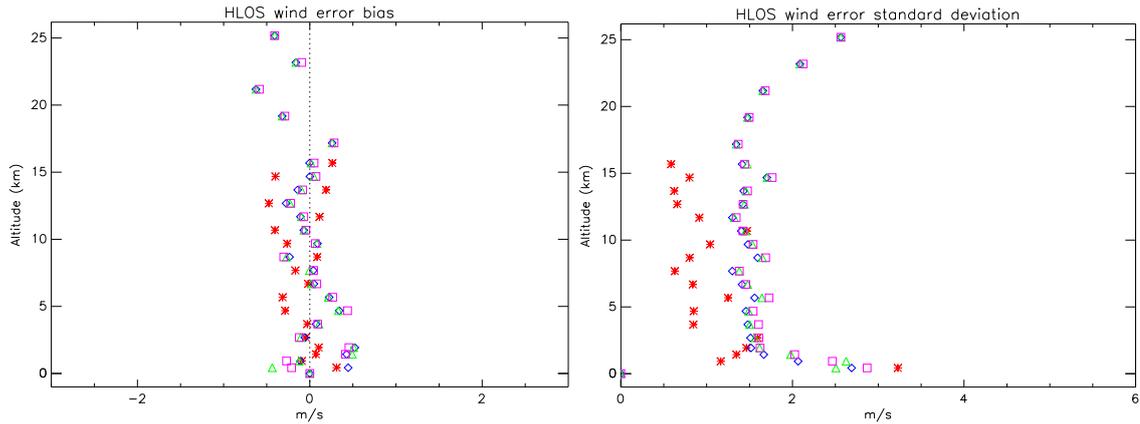


Figure 50: LIPAS HLOS wind error statistics obtained from about 400 BRCs (observation profiles), see Fig. 51 along one CALIPSO orbit; error bias (left) and error standard deviation (right). Red stars correspond to Mie winds and the blue diamonds, magenta squares and green triangles to Rayleigh winds obtained with various processing options that are further discussed in section 6.3. LIPAS was used to simulate the noisy Mie and Rayleigh channel signals.

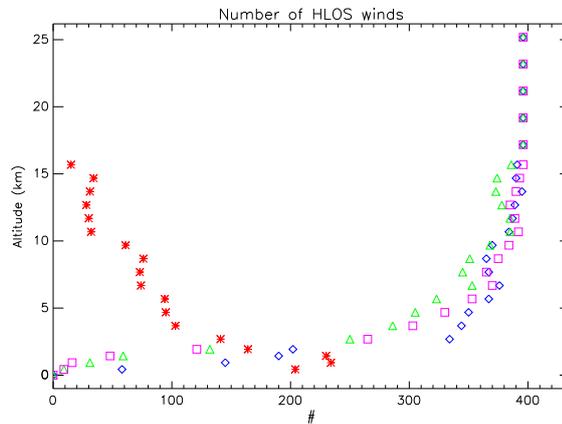


Figure 51: Number of observations used for the statistics in Fig. 50. Symbols are explained in Fig. 50.

reduced. This is done by adding a random representativeness error, further denoted the atmosphere induced representativeness error, to the instrument error. The magnitude of this error depends on the atmosphere heterogeneity.

On the other hand, there is a discrepancy between the spatial scales resolved by the observation and those resolved by the model in which the observation is ingested. If the model cannot represent (resolve) the scales measured by the observation then the impact of the observation in the assimilation is reduced by decreasing its weight in the analysis through adding a representativeness error to the instrument error. This error is in the remainder denoted the model representativeness error. The magnitude of this error is model dependent.

It is assumed that all three error sources are independent. The total wind observation error variance is then obtained from: observation error variance = instrument error variance + atmosphere induced representativeness error variance + model representativeness error variance. The instrument error has been discussed in sections 5.4.1 and 5.5.2. Both representativeness errors are discussed and quantified in the following sections.

5.7.1 Atmosphere induced representativeness error

A single ADM observation corresponds to an atmospheric slice with dimensions 50-km along track (horizontal) and l km in the vertical. Typical values for the vertical bin size l range from 0.25-2 km. The horizontal is oversampled by typically 14-50 measurements, depending on the on-board accumulation length. The L2Bp classification module provides representative weights to the measurements, based on the measured horizontal optical heterogeneity, before integration to an observation. The atmosphere induced horizontal representativeness error will therefore be negligible as compared to the other error sources.

The vertical bins are generally not oversampled. Without additional information, e.g. a priori, from other instruments or models, we lack knowledge on the distribution of particles inside the measurement bin. The measured wind inside the measurement bin, denoted u_M , can in general be written as a weighted average, denoted by w , of the wind u inside the bin of length l as follows

$$u_M = \frac{\int_{z_0}^{z_l} w(z)u(z)dz}{\int_{z_0}^{z_l} w(z)dz} \quad (49)$$

with z_0 and z_l denoting the bottom and top altitude of the vertical bin respectively. The weight function w for the Mie channel is related directly to the attenuated particle backscatter, β'_p , inside the measurement bin and for the Rayleigh channel to the attenuated molecular backscatter, β'_m . These are defined as follows

$$\beta'_p(z) = \beta_p(z)\tau(z)^2 \quad (50)$$

$$\beta'_m(z) = [\beta_m(z) + X\beta_p(z)]\tau(z)^2 \quad (51)$$

with $\tau(z)$ the total one-way atmospheric transmission between the instrument and altitude z and X denoting the contribution of the particle backscatter measured by the Fabry-Perot interferometer through cross-talk.

It is clear that the weight function very much depends on the local atmosphere optical conditions and no generic expression can be derived for the wind measurement and wind error. In the following subsections some special atmospheric situations are considered.

Very thin cloud layer We consider a constant wind-shear with value α (s^{-1}) over the measurement bin. The wind, u , inside the bin of length l can thus be modelled through $u(z) = u_0 + \alpha z$, with u_0 the wind at the bottom of the bin. The true mean wind inside the bin thus equals $\bar{u} = u_0 + \alpha l/2$. The bin includes a very thin cloud layer whose location is not known a priori. For a very thin cloud layer the backscatter can be modelled as zero throughout the measurement bin and a spike with value β_c at location z_c . The transmission of the cloud is denoted by τ_c .

For the Mie channel $w(z) = \beta_c\delta(z_c)$ with δ the delta-dirac function. Substituting in Eq. (49) yields $u_M = u_0 + \alpha z_c$. The location of the cloud inside the bin is not known a priori and we assume that it can not be determined from the measured ADM signal. The cloud location has equal probability inside the measurement bin and can thus be modelled through a uniform distribution with amplitude l^{-1} . The expectation value, μ , and variance, σ^2 , of z_c then yield

$$\mu_{z_c} = E[z_c] = l/2 \quad (52)$$

$$\sigma_{z_c}^2 = \text{Var}[z_c] = l^2/12 \quad (53)$$

The wind error, $\epsilon(z_c)$, of the measured Mie channel wind equals

$$\epsilon(z_c) = u_M - \bar{u} \quad (54)$$

that equals $\alpha(z_c - l/2)$ for the Mie channel wind and a thin cloud layer in the measurement bin. The expectation value and variance of the Mie channel wind error then equal 0 and $\alpha^2 l^2/12$ respectively.

For the Rayleigh channel the weight function is determined by the transmission inside the measurement bin, when ignoring cross-talk. The signal strength at the top of the measurement bin equals the attenuated backscatter denoted w_0 . If we also assume a constant molecular backscatter and transmission inside the

measurement bin then the weight function, $w(z)$, equals w_0 for $z_0 \leq z < z_c$ and $\tau_c^2 w_0$ for $z_c \leq z < z_l$. The Rayleigh channel wind estimate then equals

$$u_M(z) = u_0 + \frac{\int_{z_0}^{z_c} w_0 \alpha z dz + \int_{z_c}^{z_l} \tau_c^2 w_0 \alpha z dz}{\int_{z_0}^{z_c} w_0 dz + \int_{z_c}^{z_l} \tau_c^2 w_0 dz} \quad (55)$$

$$= \frac{\alpha (1 - \tau_c^2) z_c^2 + \tau_c^2 l^2}{2 (1 - \tau_c^2) z_c + \tau_c^2 l} \quad (56)$$

Substituting in Eq. (54) yields for the Rayleigh channel wind error

$$\epsilon(z_c) = \frac{\alpha}{2} \left[\frac{(1 - \tau_c^2) z_c^2 + \tau_c^2 l^2}{(1 - \tau_c^2) z_c + \tau_c^2 l} - l \right] \quad (57)$$

A first order Taylor expansion for $\epsilon(z_c)$ yields

$$\epsilon(z_c) \approx \epsilon(\mu_{z_c}) + \frac{\partial \epsilon}{\partial z}(\mu_{z_c}) [z_c - \mu_{z_c}] \quad (58)$$

The expectation value of ϵ then equals

$$\mu_{\epsilon_{z_c}} = \mathbb{E}[\epsilon(z_c)] = \epsilon(\mu_{z_c}) \quad (59)$$

$$= \frac{\alpha}{2} \left[\frac{(1 - \tau_c^2)(l/2)^2 + \tau_c^2 l^2}{(1 - \tau_c^2)(l/2) + \tau_c^2 l} - l \right] \quad (60)$$

$$= \frac{\alpha l}{2} \left[\frac{1 + 3\tau_c^2}{2(1 + \tau_c^2)} - 1 \right] \quad (61)$$

From Eq. (61) it follows that the Rayleigh channel wind has a bias with a magnitude that is related to the wind-shear over the measurement bin and the transmission of the cloud layer. For cloud transmission 1, the bias equals zero. For an opaque cloud with transmission 0 the bias is maximum with a value of $-al/4$, i.e. -1 ms^{-1} for a typical wind-shear of 0.004 s^{-1} and a binsize of 1 km. For a cirrus cloud in the tropics with typical transmission of 0.9 and a bin size of 1 km the bias equals -0.1 ms^{-1} , i.e. within the ADM-Aeolus bias requirement of 0.4 ms^{-1} for zero HLOS wind and a 0.7% slope error.

The Rayleigh wind error variance follows from Eq. 58

$$\sigma_{\epsilon_{z_c}}^2 = \text{Var}[\epsilon(z_c)] = \left[\frac{\partial \epsilon}{\partial z}(\mu_{z_c}) \right]^2 \text{Var}[z_c] \quad (62)$$

$$= \left[\frac{\alpha}{2} (1 - \tau_c^2) \frac{(1 - \tau_c^2) \mu_{z_c}^2 + l \tau_c^2 (2\mu_{z_c} - 1)}{[(1 - \tau_c^2) \mu_{z_c} + \tau_c^2 l]^2} \right]^2 \frac{l^2}{12} \quad (63)$$

using Eqs. (53,57). For $\tau_c = 1$ the Rayleigh wind error equals zero, thus also its variance. For an opaque cloud with $\tau_c = 0$ the variance of the Rayleigh wind error equals $\alpha^2 l^2 / 48$. Some numerical results are presented in Table 11. Note that $\tau_c = 1$ is equivalent to zero cloud extinction and thus zero cloud backscatter assuming a cloud lidar ratio value larger than 0. Then the Mie channel weight function equals 0. In other words, Eq. (49) has no solution which is consistent for a particle-free measurement bin.

Thick cloud layer A cloud layer of thickness δz is assumed. The cloud layer center is located at z_c . The Mie channel weight function then equals 0 for $z_0 \leq z < z_c - \delta z/2$, $\beta'_c(z)$ for $z_c - \delta z/2 \leq z \leq z_c + \delta z/2$ and 0 for $z > z_c + \delta z/2$. We assume $\beta'_c(z)$ constant for simplicity. Substituting in Eq. (49) yields $u_M = u_0 + \alpha z_c$, similar as for the very thin cloud layer in the previous section. Again, the location of the cloud layer can be modelled as a uniform distribution but is limited to the interval $[z_0 + \delta z/2, l - \delta z/2]$. The amplitude of the uniform distribution is thus $(l - \delta z)^{-1}$. The expectation value, μ and variance, σ^2 of z_c then yield

$$\mu_{z_c} = \mathbb{E}[z_c] = l/2 \quad (64)$$

$$\sigma_{z_c}^2 = \text{Var}[z_c] = (l - \delta z)^2 / 12 \quad (65)$$

Table 11: Bias and variance for Mie and Rayleigh channel wind errors for a 1000 meter vertical bin including a very thin cloud layer with one-way cloud transmission τ_c . A constant wind-shear of 0.004 s^{-1} over the bin is assumed. The table values correspond to 3 different values for the one-way cloud transmission. X means no wind retrieval

	$\tau_c=1/0.9/0$	Mie	Rayleigh
bias	(ms^{-1})	X/0/0	0/-0.1/-1
variance	(m^2s^{-2})	X/1.33/1.33	0/0.013/0.33

From Eq. (54) the wind error of the measured Mie channel wind again equals $\epsilon(z_c) = \alpha(z_c - l/2)$ with expectation value 0 and variance $\alpha^2(l - \delta z)^2/12$, i.e. smaller than for a very thin cloud. For the extreme values of δz equal to zero yields the same result as in the previous section for the very thin cloud layer, for δz equal to l , the Mie wind error variance equals zero which is consistent with a uniform weight function (attenuated scattering) over the complete measurement bin.

For the Rayleigh channel wind we assume a linear weight function (cloud transmission) in the cloud layer. Again, the total one-way cloud transmission is denoted τ_c . The Rayleigh channel weight function then equals w_0 for $z_0 \leq z < z_c - \delta z/2$, $az + b$ for $z_c - \delta z/2 \leq z \leq z_c + \delta z/2$ with $a = -w_0(1 - \tau_c^2)/\delta z$ and $b = w_0 - a(z_c - \delta z/2)$ and $\tau_c^2 w_0$ for $z > z_c + \delta z/2$. Substituting in Eq. (49) yields for the measured Rayleigh wind

$$u_M(z) = \frac{\alpha}{2} \frac{(1 - \tau_c^2)z_c^2 + \tau_c^2 l^2 + \frac{1}{12}(1 - \tau_c^2)\delta z^2}{(1 - \tau_c^2)z_c + \tau_c^2 l} \quad (66)$$

and for the Rayleigh wind error

$$\epsilon(z_c) = \frac{\alpha}{2} \left[\frac{(1 - \tau_c^2)z_c^2 + \tau_c^2 l^2 + \frac{1}{12}(1 - \tau_c^2)\delta z^2}{(1 - \tau_c^2)z_c + \tau_c^2 l} - l \right] \quad (67)$$

This yields for the expectation value

$$\mu_{\epsilon_{z_c}} = \frac{\alpha l}{2} \left[\frac{1 + 3\tau_c^2}{2(1 + \tau_c^2)} + \frac{(1 - \tau_c^2)\delta z^2}{6l^2(1 + \tau_c^2)} - 1 \right] \quad (68)$$

i.e. a smaller bias than for a very thin cloud. For the Rayleigh wind error variance we get

$$\sigma_{\epsilon_{z_c}}^2 = \left[\frac{\alpha}{2} (1 - \tau_c^2) \frac{(1 - \tau_c^2)\mu_{z_c}^2 + l\tau_c^2(2\mu_{z_c} - 1) - (1 - \tau_c^2)\delta z^2/12}{[(1 - \tau_c^2)\mu_{z_c} + \tau_c^2 l]^2} \right]^2 \frac{l^2}{12} \quad (69)$$

i.e. smaller than the variance for a very thin cloud. Figure 52 shows that the maximum bias values are obtained for optically thick and geophysically thin clouds. The maximum error of -1 ms^{-1} corresponds to the optically thick and geophysical very thin cloud layer discussed in the previous section. Such a cloud also has a maximum variance of $0.33 \text{ m}^2\text{s}^{-2}$ as found in the previous section. For a cloud covering half the bin and with transmission 0.5 the bias and variance are about 0.55 ms^{-1} and $0.25 \text{ m}^2\text{s}^{-2}$ respectively.

5.7.2 Model representativeness error

Nowadays NWP models generally do not well represent atmospheric structures as observed by ADM. This is true in particular for global models such as ECMWF that does not well resolve spatial structures below 250 km, see [RD9]. The small scale variability which is sampled by individual observations, but which the model is incapable of representing, is referred to as the representativeness error. For a DWL this error mainly relates to a discrepancy of the spatial scales resolved by the observation and those resolved by the model in which the observation is ingested. If the model cannot represent (resolve) the scales measured by the observation then the impact of the observation in the assimilation is reduced by decreasing its weight in the analysis through adding a representativeness error to the observation error, i.e., the observation error variance = instrument error variance + representativeness error variance.

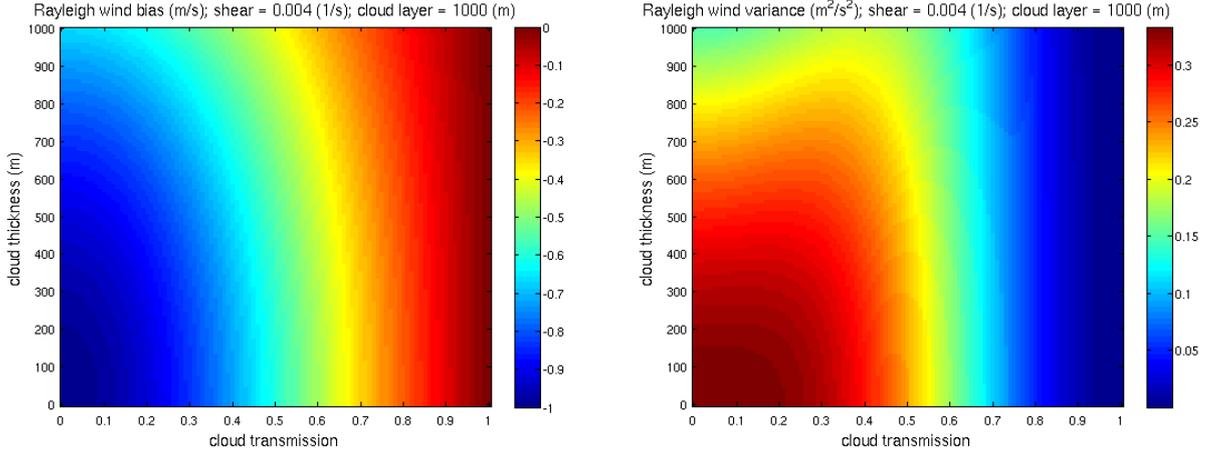


Figure 52: Rayleigh wind error bias (ms^{-1}) (left) and variance (m^2s^{-2}) (right) as a function of one-way cloud transmission and cloud layer thickness (m). The measurement bin is 1000 m, the wind-shear is constant over the bin with a value of 0.004 s^{-1} .

Following, section 6.2.1 of [RD9], we distinguish the along-track and across-track HLOS wind component variability, where across-track is taken in the direction of the HLOS. We assume that the along- and across-track variabilities are independent, but of equal size in amplitude, i.e., $0.5r^2$ with r^2 the wind variance in the model resolution cell (see [RD9]), i.e. the representativeness error variance for a point measurement such as a radiosonde. The model resolution cell size is also denoted effective model resolution. In the across-track direction we only obtain one measurement location, and as such the wind variability within a resolution cell in this direction is not resolved. The HLOS wind component representativeness error thus amounts to $r_{\text{across}}^2 = 0.5r^2$. In the along-track direction we may obtain up to N accumulations over the observation sampling length. If the N samples are uniformly distributed over the observation sampling length then the representativeness error variance over this distance is reduced by a factor N . The remaining along-track wind variance in the resolution cell that is not sampled equals $0.5r^2$ minus the sampled variance. The along-track error variance contribution then equals (see [RD9])

$$r_{\text{along}}^2 = 0.5 \left[1 - \left(\frac{L_{\text{sample}}}{L_{\text{cell}}} \right)^{2/3} \right] r^2 + 0.5 \left(\frac{L_{\text{sample}}}{L_{\text{cell}}} \right)^{2/3} \frac{r^2}{N} \quad (70)$$

with L_{sample} the sample length and L_{cell} the size of the square effective model resolution cell. Table 12 is a copy of Table 3 from [RD9] showing the effective ECMWF model resolution. For instance, in burst mode the observation sampling length, L_{sample} , is 50 km, $N=14$ for a 3.5 km on-board accumulation length. At 500 hPa the effective model resolution L_{cell} equals 263 from Table 12 with corresponding wind variance $r^2 = 3.3$. Substituting in Eq. (70) yields $r_{\text{along}}^2 = 0.33r^2 + (0.17/14)r^2$, or a reduction of the along-track representativeness error of about 30% as compared to a point measurement. Note that the number of accumulated measurements may be smaller than $N = 14$ in heterogeneous atmospheres, e.g. partly cloudy scenes such as convective cumulus clouds. In LIPAS we assume that the valid measurements are clustered. Then the sample length needs to be adjusted accordingly, i.e. in case of $M < N$ valid measurements, N is replaced by M in Eq. (70) and the sample length $L_{\text{sample}} = 50 \times (M/N)$.

Table 12 and Eq. (70) have been implemented in LIPAS-2009 to compute the HLOS wind representativeness error. Linear interpolation with altitude is applied to get the effective model resolution and atmospheric variance in intermediate points. Extrapolation to higher altitudes is done by fixing the 50 hPa values.

Table 12: ECMWF zonal wind effective model resolution as a function of pressure level (hPa) as obtained from [RD9]. Italic values denote extrapolation by matching Table 4.4 in [RD12] with respective values in [RD9], i.e. equal missing variance at 500 and 150 hPa and linear interpolation of the missing variance from 150 to 50 hPa. The effective model resolution of 200 km at 50 hPa is a guess value.

pressure (hPa)	effective model resolution (km)	unresolved wind variance (m^2s^{-2})
1000	340	3.94
925	232	3.06
850	253	3.24
700	312	3.72
500	263	3.30
400	229	3.03
250	312	3.72
<i>150</i>	<i>263</i>	<i>3.30</i>
<i>50</i>	<i>200</i>	<i>2.52</i>

5.7.3 HLOS wind error overview

Tables 13 and 14 provide an overview of the estimates of Mie and Rayleigh HLOS wind errors as discussed in the previous sections. These error estimates are needed in data assimilation to give a proper weight to the observations in NWP model analyses.

Table 13: Estimates of Mie channel HLOS wind error contributors; stddev is an abbreviation of standard deviation, repr. is an abbreviation of representativeness

error source	bias/stddev (ms^{-1})	comment
instrument error	0 / 0.5-1.5 m^{-1}	obtained from Tables 6/7 of section 3.11. only winds with $\text{SNR} \geq 10$ are used. Other parameters that impact the wind quality are the ResErr and peak FWHM
repr. error atmosphere	0 / $\frac{\alpha^2(l-\delta z)^2}{12}$	Error due to unknown location of the particle layer of thickness δz in range gate of length l . HLOS wind shear α [s^{-1}] is obtained from climatology and δz [m^{-1}] is estimated by the L2B optical properties code
repr. error NWP model	0 / Eq. (70)	For the total error $r_{\text{across}}^2 = 0.5r^2$ must be added, see section 5.7.2. The number of useful measurements is obtained from the L2B classification

Table 14: Estimates of Rayleigh channel HLOS wind error contributors; stddev is an abbreviation of standard deviation, repr. is an abbreviation of representativeness

error source	bias/stddev (ms^{-1})	comment
instrument error	0 / Eq. (38)	An estimate of channel A/B signals is obtained from the L2B auxiliary meteorological data
repr. error atmosphere	Eq. (68) / Eq. (69)	see comment in Table 13
repr. error NWP model	0 / Eq. (70)	see comment in Table 13

5.8 Summary and conclusions

In this section the upgrade of the lidar performance analysis simulator, LIPAS, has been discussed. The main objective was to make the Fizeau and Fabry-Perot digitised ACCD output signals as simulated by LIPAS compatible with the signals simulated by the End-to-End simulator E2S. These instrument output signals are input to the L1B and L2B processors (L1Bp and L2Bp) for HLOS wind retrieval. The second main objective was to demonstrate compatibility of LIPAS and L2Bp retrieved winds.

The upgraded LIPAS can process data from the atmospheric database that is input for the E2S. LIPAS includes all relevant ADM hardware aspects that are simulated in the E2S, including Fizeau and Fabry-Perot transfer functions and analog-to-digital ACCD signal conversion. It was shown that the digitised ACCD signal on the Fizeau interferometer in LIPAS is within 5% of the E2S. For the Fabry-Perot signal the difference is even smaller to within 2%. Cross-talk simulation has been implemented and impact on Rayleigh wind biases validated with earlier results (Météo France ILIAD study).

Wind retrieval algorithms have been implemented in LIPAS similar to those used in the L1B and L2B processors. This includes the Mie core peak fitting algorithm for Doppler shift retrieval from the Mie channel signal and the inverse receiver response method for Doppler shift retrieval from the Rayleigh channel signal. It was found that the rounding effect from floating point to an integer value of the number of detected electrons, as part of the analog-to-digital conversion of the ACCD signal, yields a bias in Rayleigh channel winds in the order of a couple of tenths ms^{-1} .

Both the E2S-L1B-L2B chain and LIPAS were used in an experiment to retrieve simulated winds along a half orbit and using realistic atmospheric scenes composed of CALIPSO atmospheric optical properties and ECMWF atmospheric dynamics and temperature. The experiment included photon count (Poisson) noise simulation, but solar background and tripod obscuration were ignored. Overall, the results compared very well, with slightly smaller wind error biases from LIPAS, slightly smaller Mie wind error standard deviations from the L1Bp and close correspondence of L2Bp and LIPAS Rayleigh wind error standard deviations throughout the atmosphere.

Finally, various types of representativeness errors for HLOS wind observations have been identified, discussed and quantified. These include an atmosphere induced representativeness error and a numerical weather prediction model representativeness error. The former relates to the error made due to incorrect knowledge of an aerosol or cloud layer within the measurement bin. The latter relates to the spatial scales resolved by the wind observations but that are not represented by the model into which the data are assimilated. Quantification of these representativeness errors is done by LIPAS and needed in addition to knowledge of the instrument error to give the observations a proper weight in the NWP analysis. This is further discussed in section 6.

The close correspondence of LIPAS and the E2S-L1B-L2B has been demonstrated in this section. LIPAS will be used in the next section to simulate winds in realistic atmospheric conditions for the vertical sampling scenarios defined in [RD6] and to quantify the information content of the sampling scenarios.

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

6.1 Introduction

The information content of a HLOS wind can be considered as inversely proportional to its error variance. This is in line with the use of unbiased observations in data assimilation where the weight of observations in the analysis is proportional to the inverse of the observation error covariance matrix. As discussed in the previous section, there are three contributors to the total wind error: instrument noise, atmosphere induced representativeness errors through optical heterogeneity and model horizontal and vertical representativeness errors. For Rayleigh winds also cross-talk adds to the wind error. The amplitude of these errors is related to the vertical sampling scenario. For instance, reducing the bin length reduces the atmosphere induced and vertical representativeness errors on one hand but increases the instrument error because of loss of SNR.

The upgraded LIPAS, discussed in section 5, and the atmospheric database of collocated CALIPSO optical and ECMWF meteorological parameters, that is described in [RD7], are used here to assess the information content of the vertical sampling scenarios as defined in [RD6]. The information content of a vertical sampling scenario is quantified here in two ways. First by its data coverage, i.e. counting the number of good AND bad quality winds at various altitudes and as a function of climate zone. It is important to note that a small number of bad quality winds can be detrimental for NWP analyses. The number of erroneous winds that can not be identified by quality control therefore has a large impact on the information content of the sampling scenario, even if the number of good quality winds is large. A second measure of information content makes use of the wind error variance estimate to relate DWL wind profile information to radiosonde profile information. Because quality control is an important aspect of the sampling scenario information content the next section first discusses Mie and Rayleigh channel quality control indicators.

6.2 Mie and Rayleigh channel quality control

There are a couple of ways to control the quality of observations, e.g. from instrument characteristics such as SNR or the evaluation of output parameters from data processing tools or by comparing with other independent observations or short-term NWP forecasts. This section focuses on quality indicators related to the SNR and processing output parameters.

6.2.1 Mie channel quality control

Quality control of Mie channel winds is based on output parameters of the Mie Core algorithm that has been discussed extensively in section 3. Currently LIPAS only includes the SNR indicator, i.e. only Fizeau signals with SNR values exceeding 10 are processed. In addition winds obtained in bins that contain the ground surface are flagged and may be used at a later stage for ground calibration purposes.

6.2.2 Rayleigh channel quality control

Quality control of Rayleigh channel winds is based on the measured signals in both channels denoted A and B. Erroneous Rayleigh winds may be related to i) ground returns, ii) low SNR, e.g. below clouds or iii) large cross-talk, e.g. inside cloud layers. We further elaborate on the latter two by considering the true (free of noise and cross-talk) and measured (including noise and eventually cross-talk) receiver response signals, R_t and R_o respectively and their difference ΔR :

$$\begin{aligned}
R_t &= \frac{A_m - B_m}{A_m - B_m} \\
R_o &= \frac{A_m + dA_m - (B_m + dB_m) + A_x + dA_x - (B_x + dB_x)}{A_m + dA_m + B_m + dB_m + A_x + dA_x + B_x + dB_x} \\
\Delta R &= R_o - R_t \\
&= 2 \frac{B_m (A_x + dA_x + dA_m) - A_m (B_x + dB_x + dB_m)}{(A_m + dA_m + B_m + dB_m + A_x + dA_x + B_x + dB_x) (A_m + B_m)} \tag{71}
\end{aligned}$$

with subscripts m and x denoting the molecular and cross-talk contribution respectively and terms $d\bullet$ denoting the noise contributions from the respective signals. Assuming noise with zero mean, the bias and variance of the receiver response error equal

$$\begin{aligned}
E[\Delta R] &= 2 \frac{B_m A_x - A_m B_x}{(A_m + B_m + A_x + B_x)(A_m + B_m)} & (72) \\
\text{Var}[\Delta R] &= \left(\frac{\partial \Delta R}{\partial dA_m} \right)^2 \text{Var}[dA_m] + \left(\frac{\partial \Delta R}{\partial dB_m} \right)^2 \text{Var}[dB_m] \\
&+ \left(\frac{\partial \Delta R}{\partial dA_x} \right)^2 \text{Var}[dA_x] + \left(\frac{\partial \Delta R}{\partial dB_x} \right)^2 \text{Var}[dB_x] \\
&= \left[\frac{1}{2(A_m + A_x)} \right]^2 (2A_m + 2A_x) \\
&= \frac{1}{2(A_m + A_x)} & (73)
\end{aligned}$$

where we assumed $B_m = A_m$ and $B_x = A_x$ in the variance computation for simplicity. Eq. (73) is the receiver response error variance when no cross-talk correction is applied and is similar to Eq. (35), but now including the cross-talk term. The variance decreases for increasing cross-talk but also the bias generally increases from Eq. (72). In case that cross-talk correction can and has been applied $A_x = B_x = 0$ which simplifies the equations. Note that although cross-talk can potentially be corrected for, the corresponding noise can not be corrected for, i.e. dA_x dB_x do not equal zero after cross-talk correction. The expectation value (bias) of ΔR after cross-talk correction equals zero, from Eq. (72), and for the variance we get

$$\text{Var}[\Delta R] = \frac{1}{2A_m} \left(1 + \frac{A_x}{A_m} \right) \quad (74)$$

This means that the receiver response error variance will generally be large in Rayleigh bins with strong cloud and aerosol layers, even when their cross-talk contribution can be well estimated. The Rayleigh wind quality will therefore generally be poor for these conditions.

Quality control (QC) indicators (flags) have been defined and are tagged to the retrieved winds for later processing. These QC indicators include

1. Ground return. If the height bin includes ground returns within the 50 km sampling length then the flag is set to the value 1.
2. SNR check. From Eq. (37) it follows that for winds meeting the mission requirement

$$\frac{\partial \text{LOS}}{\partial R} \sqrt{\text{Var}[dR]} \leq \sigma_{\text{LOS}}^T \quad (75)$$

with σ_{LOS}^T the wind error standard deviation threshold or mission requirement value. $\text{Var}[\Delta R]$ is obtained from Eq. 35 with A and B the total signal including cross-talk. The sensitivity term is obtained from the receiver response table, see section 5.5.2. The threshold LOS wind error standard deviation is a function of altitude according to the mission requirement. If the threshold is not met then the flag is raised and set to the value 2.

3. If cross-talk correction is applied then the increased error through cross-talk noise is calculated. The same procedure as in the previous item is followed. From Eq. (75) and Eq. (74) it then follows that $A_x \leq \left[2A_m [\sigma_{\text{LOS}}^T / (\partial \text{LOS} / \partial R)]^2 - 1 \right] A_m$ to meet the mission requirement. Note that A_x and A_m are not directly available from the measured signal, but can be estimated from the AUX_MET data (A_m) and the scattering ratio estimate (A_x). If A_x exceeds the threshold value then the flag is raised and set to a value of 3.

6.3 HLOS wind retrieval

LIPAS has been used to retrieve Mie and Rayleigh channel winds along CALIPSO nighttime orbits. The atmosphere optics and dynamics are obtained from the atmospheric database. Figure 53 shows an example

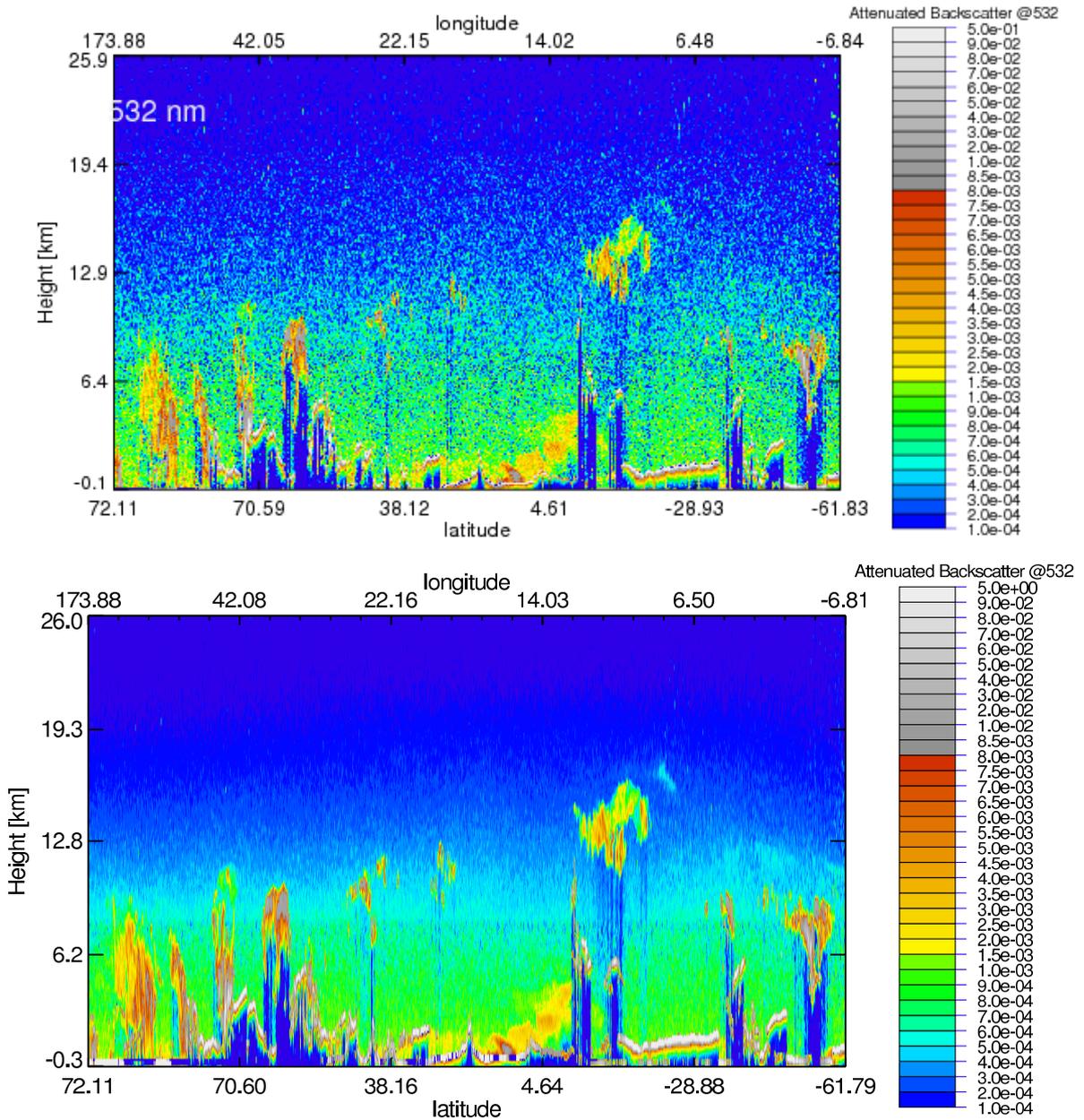


Figure 53: Half (nighttime) orbit CALIPSO level-1B attenuated backscatter at 532 nm measured on 1/1/2007 around 00:30 UTC. Top: Raw data at instrument resolution, see Table 4.2.3-1 of [RD7]. Bottom: processed data to 3300 m horizontal and 125 m vertical resolution. The spurious artefacts near 8 km are due to a change of the on-board averaging from 30 to 60 m.

of raw CALIPSO data along one half orbit. The top panel in Figure 54 shows an example of the scattering ratio along half a CALIPSO orbit as obtained from atmospheric database molecule and aerosol backscatter. The bottom panel in Figure 54 shows the collocated zonal wind along the orbit. Figure 55 zooms in on the tropical part of the orbit showing a cirrus cloud near 15 km with a scattering ratio of about 4 but with locally increased values exceeding 10. Underneath the cirrus cloud optically thick clouds show up. The left hand-side of the figure shows enhanced aerosol loading up to 4 km altitude. Figure 56 shows the HLOS wind rather than the zonal wind as in the bottom panel of Fig. 54. The HLOS wind is obtained from

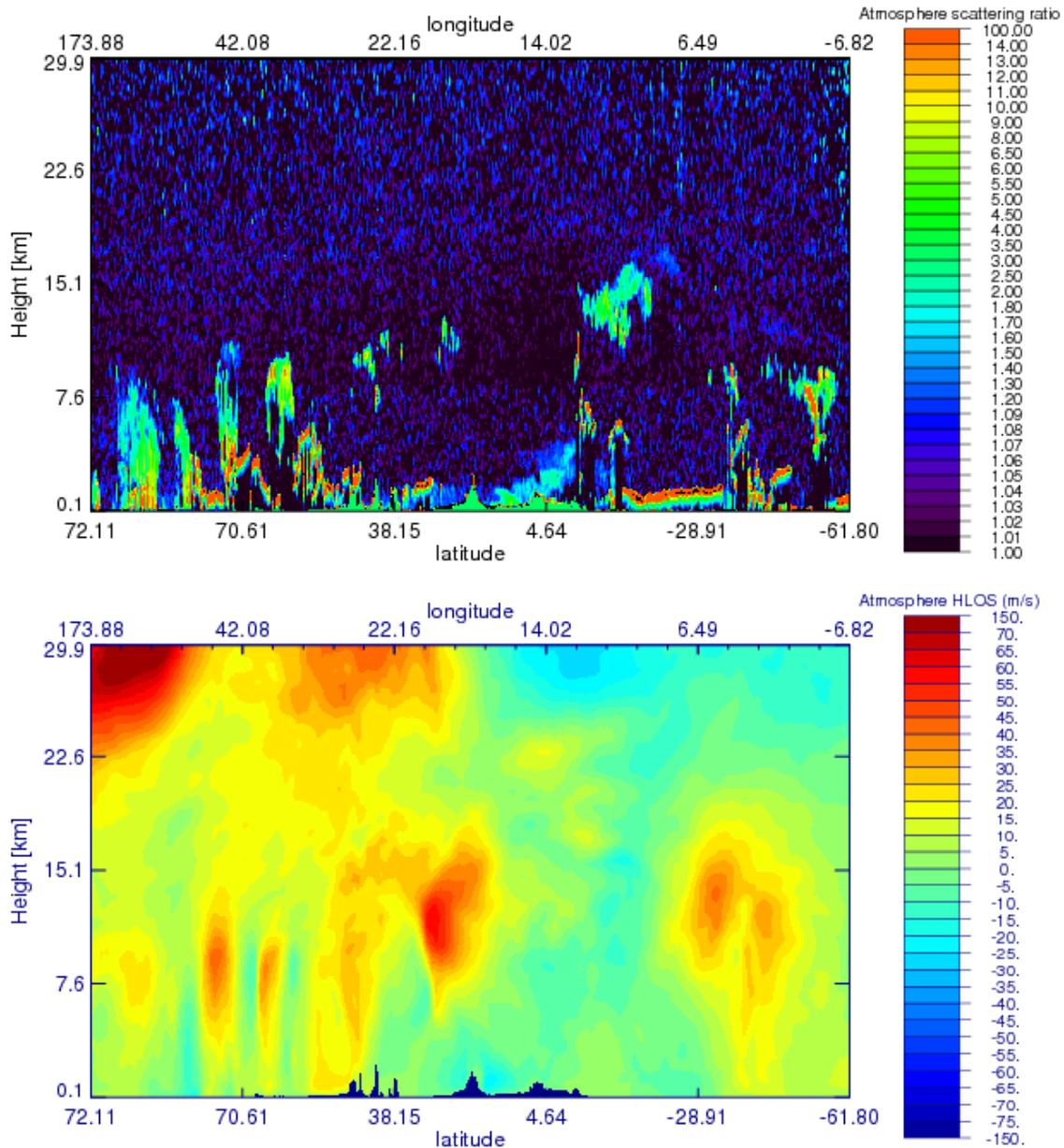


Figure 54: Atmosphere scattering ratio along the nighttime part of one CALIPSO orbit, measured on 1/1/2007 around 00:30 UTC (top) and corresponding zonal wind (bottom) as obtained from the atmospheric database winds and ADM pointing. The horizontal resolution is 3.5 km, the vertical resolution 125 m.

the zonal and meridional wind components taking into account the pointing of ADM through the azimuth angle, see Eqs. (11) and (27) or Eq. (4-7) of [RD7]. The LOS direction is oriented close to the west along most of the half orbit yielding a HLOS wind close to minus the zonal wind. The discontinuity near the Pole is explained by a change of the azimuth angle from 270 to 90 degrees in this area. Figure 57 shows the Mie and Rayleigh channel retrieved winds, that passed the quality control, for 50 km adjacent along track samples corresponding to the ADM observation accumulation length. The ADM burst mode 150 km gaps are not considered, because we focus on observation information content here and not on exact ADM sampling. Vertical sampling scenario WVM1 was used. Appendix A provides an overview of all sampling

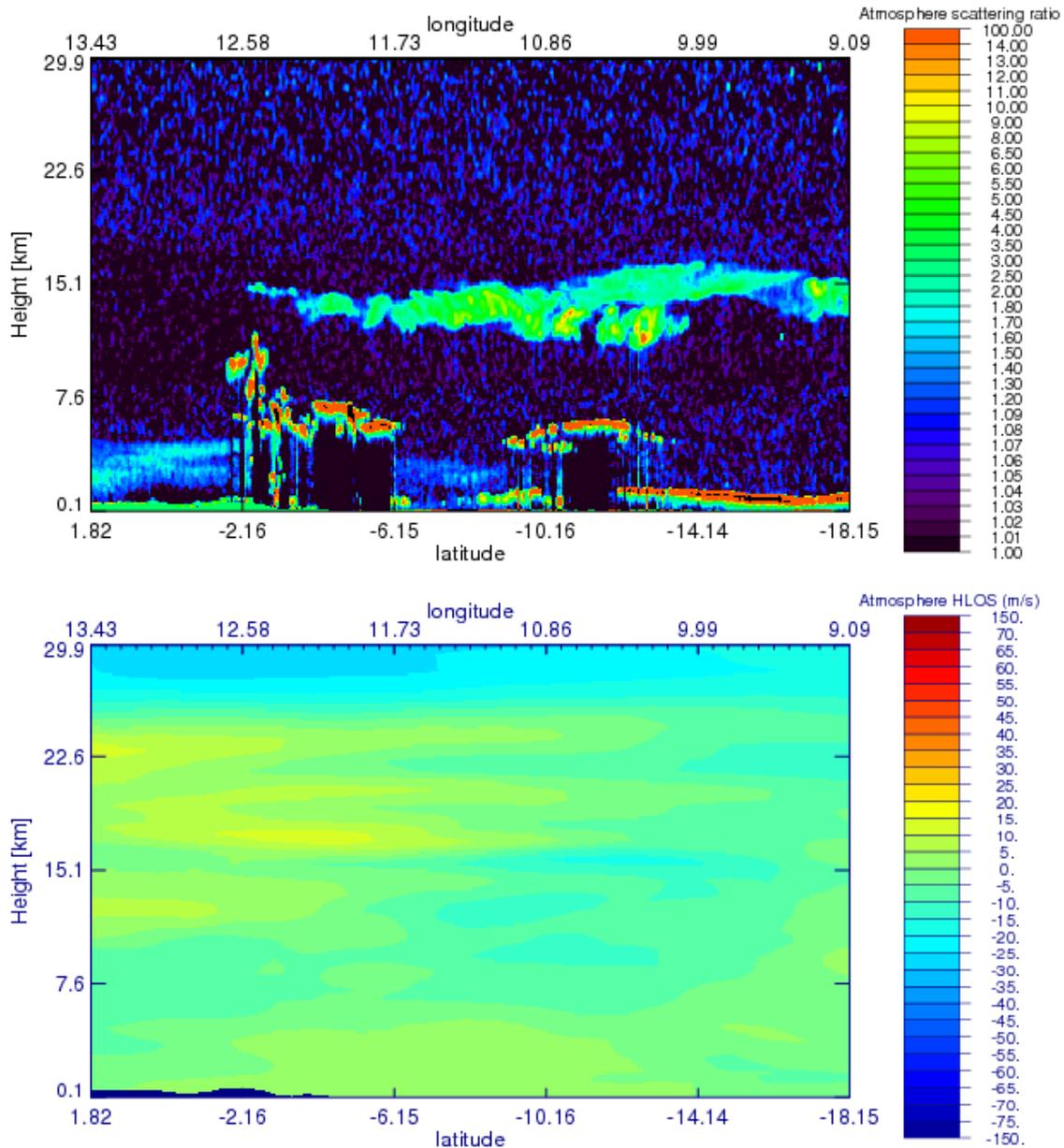


Figure 55: Same as Fig. 54 but zoomed in to the Tropics.

scenarios used in this document. Mie winds are obtained only in particle-rich areas where the SNR on the Fizeau ACCD bins exceeds the threshold value of 10. This threshold value is also used in the L1B and L2B processors. The SNR computation is part of the Mie Core algorithm. For the Rayleigh wind retrieval currently three options have been implemented in LIPAS-2009:

1. Accumulate the signals from all measurements, without signal classification and ignoring cross-talk. This option mimics *Rayleigh-only sampling* where no Mie information is available for classification or to correct for cross-talk.
2. Apply classification, i.e. accumulate only measurements with small particle presence. A scattering ratio threshold value of 1.5 is used to discriminate between particle and "particle-free" measurements. Currently the scattering ratio is obtained from the input database. At a later stage the collocated

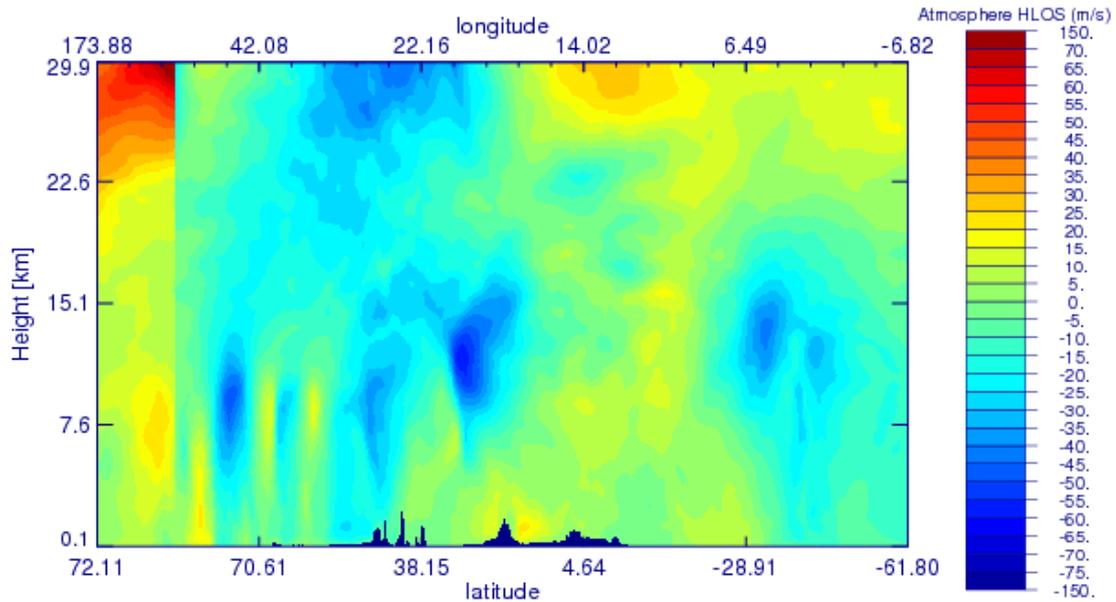


Figure 56: Same as the bottom panel in Fig. 54 but now the HLOS wind.

Mie channel signal or optical properties code should provide this information. Currently, classification can only be applied to Rayleigh bins with colocated Mie bins. This option thus mimics *collocated Rayleigh-Mie sampling*.

3. Apply cross-talk correction, i.e. accumulate the signals from all measurements after cross-talk correction. For each measurement the cross-talk contribution is estimated and subtracted from the measured signal before accumulation. Currently the exact cross-talk contribution (as obtained from the atmospheric database aerosol properties) is subtracted. At a later stage the scattering ratio estimate from the Mie channel or the optical properties code is used for cross-talk correction. Currently, cross-talk correction can only be applied to Rayleigh bins with colocated Mie bins. This option thus also mimics *collocated Rayleigh-Mie sampling*.

LIPAS has been applied to 1 day of data in the database. This covers 15 (half) orbits and yields about 6000 (adjacent) HLOS wind profiles for both Mie and Rayleigh channels. For the Rayleigh channel 3 wind solutions have been determined based on the processing options described above. The left panel in Fig. 59 shows the error bias of retrieved Mie and Rayleigh winds. Only retrieved winds that pass the quality control are used in the statistics. The bias in the Mie winds is generally below 0.3 ms^{-1} throughout the troposphere. At high altitudes the 3 Rayleigh wind processing options yield about the same bias as expected because of negligible particle presence. The bias at these altitudes is substantial with values up to 0.7 ms^{-1} . This may be explained by a larger molecular density in the lower part of the 2 km vertical bins. In combination with a generally positive wind-shear, this yields a negative bias in the retrieved winds. [In addition the rounding to integer-valued electrons as part of the ADC, see section 5.6.3, may contribute to biases, although this effect is expected to be small in the statistics when averaged over many different wind scenes.](#)

From the right panel in Fig. 59 it is seen that the number of Rayleigh winds for all three processing options is almost identical above 20 km because of negligible aerosol content. The number of Mie winds is thus also negligible. Near 15 km tropical cirrus scenes reduce the number of Rayleigh winds for options 2 and 3 because some or all of the measurements are classified as "particle" and not used in the accumulation, thus reducing the observation SNR (option 2) or cross-talk correction that reduces the SNR (option 3). The small amount of Rayleigh winds in the lowest 2 km is explained by the small 500 m Rayleigh bins. The number of Mie winds on the other hand increases substantially in the PBL due to increased aerosol loading.

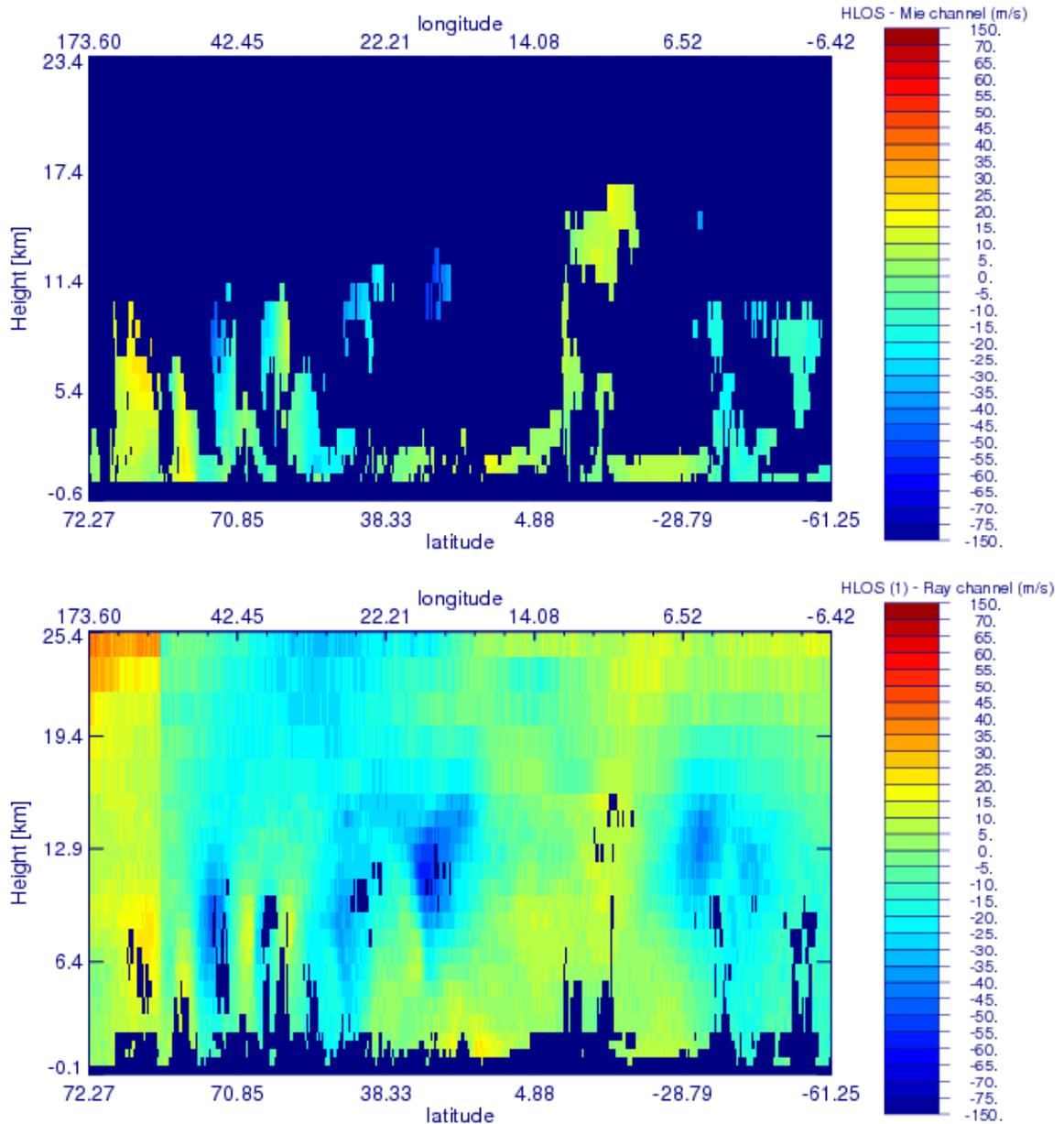


Figure 57: Retrieved Mie (top) and Rayleigh (bottom) winds for the scene displayed in Fig. 54 and vertical sampling scenario WVM1. Rayleigh wind retrieval option 1 was used, see text for further details. The LIPAS output quality indicator flags are used to select the good quality winds for plotting. Dark blue spots denote missing winds based on the quality indicator flags. Note the different height interval along the y-axis.

Interestingly, also at lower altitudes with particle presence through clouds and aerosols, the biases for all three Rayleigh options are similar. This may be surprising, in particular for option 1 because it is known from section 5.5.3 that cross-talk may induce substantial biases. We further elaborate on this.

One explanation for similar biases in option 1 as compared to the other options is that biases for individual HLOS wind estimates may be both positive and negative and thus cancel in the average statistics. This is true if there is no prevailing wind direction. However, for instance for the prevailing zonal wind in the Northern Hemisphere extra-tropics (and considering only nighttime (half) orbits) the majority of biases

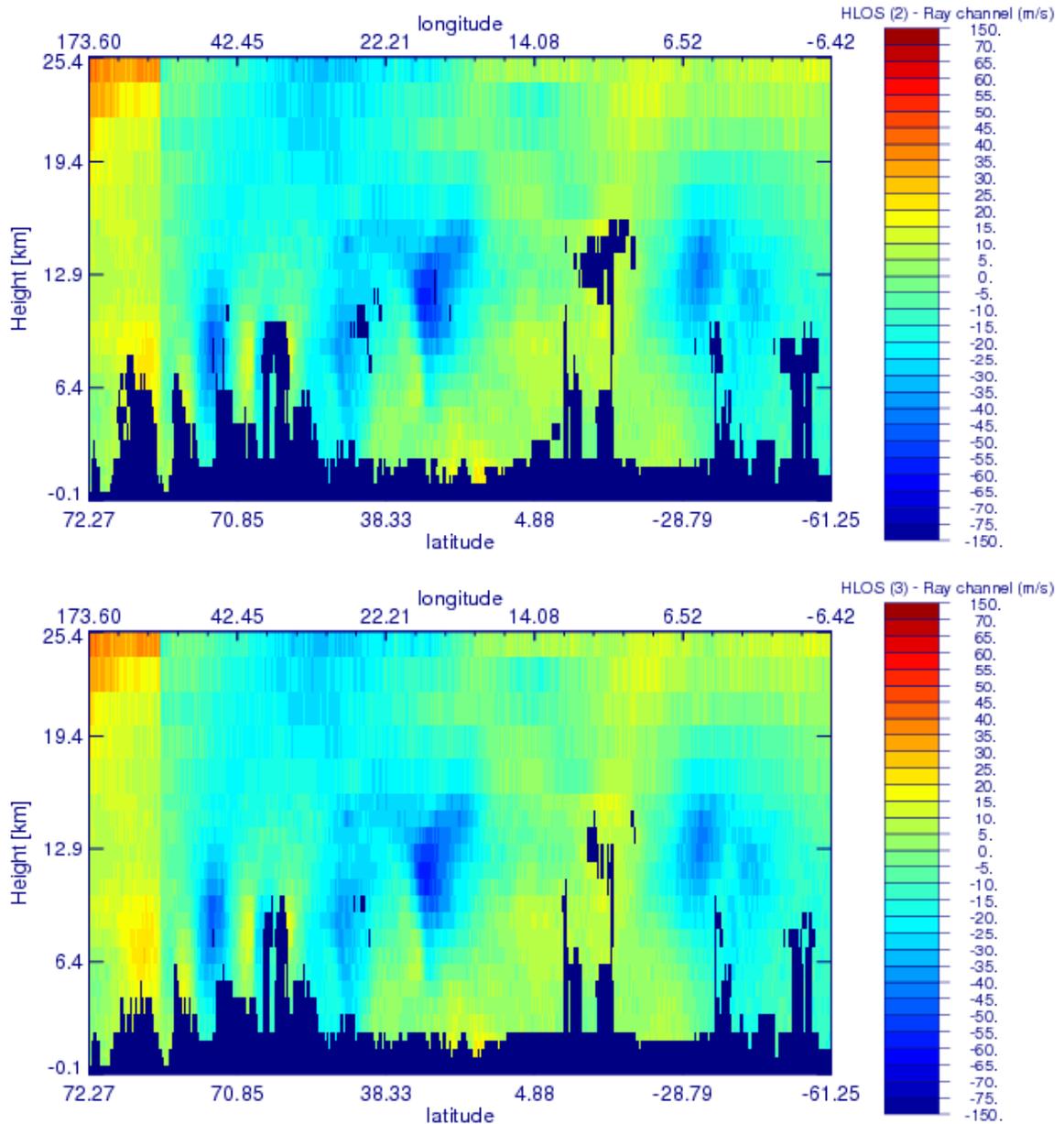


Figure 58: Same as the bottom panel of Fig. 57 but for option 2 (top) and option 3 (bottom) for the Rayleigh wind retrieval, see text for further details.

have the same sign and cancelation does not take place. Another more plausible explanation is that the bias through cross-talk is generally small. As explained in section 5.5.3 the HLOS wind bias through cross-talk is a function of temperature, scattering ratio and the actual wind. Bias through cross-talk is substantial for low temperatures, a large scattering ratio and large winds. All three conditions should apply at the same time. If one of these conditions does not apply, then the bias through cross-talk is small. This is the case, for instance, for temperatures above zero degrees celsius that applies for most parts of the globe in the lower part of the atmosphere.

A potential challenging region is the upper troposphere in the tropics with low temperatures and the presence of tropical cirrus with substantial scattering ratio values. To investigate this region scenes with

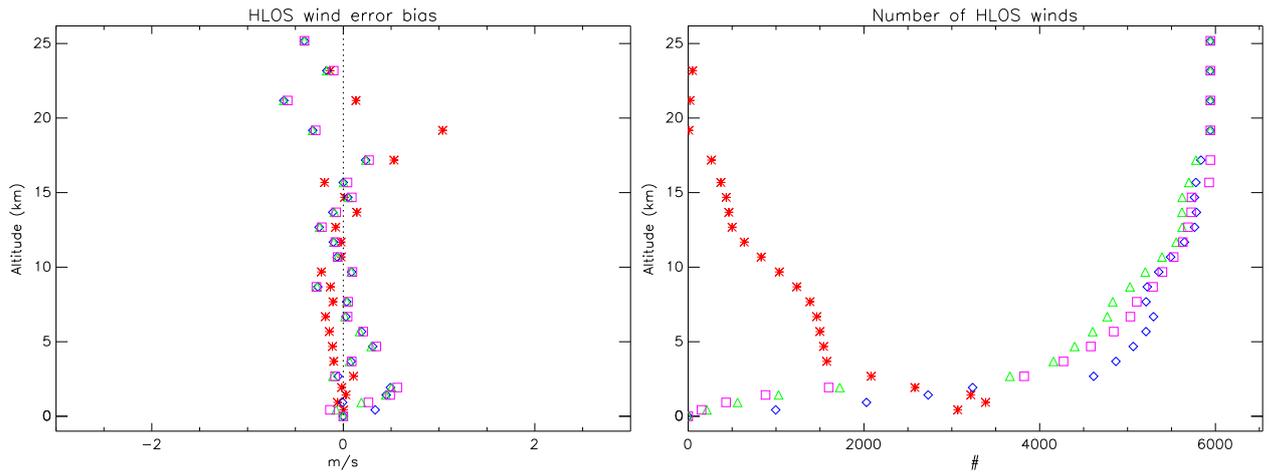


Figure 59: HLOS wind error bias for Mie (red stars) and Rayleigh channel (left) for vertical sampling scenario WVM1. Three processing options for Rayleigh wind retrieval are displayed: option 1, simply accumulating all measurements (blue diamonds), option 2, application of classification (green triangles) and option 3, application of cross-talk correction (magenta squares). Right, the number of data that passed quality control and used in the statistics.

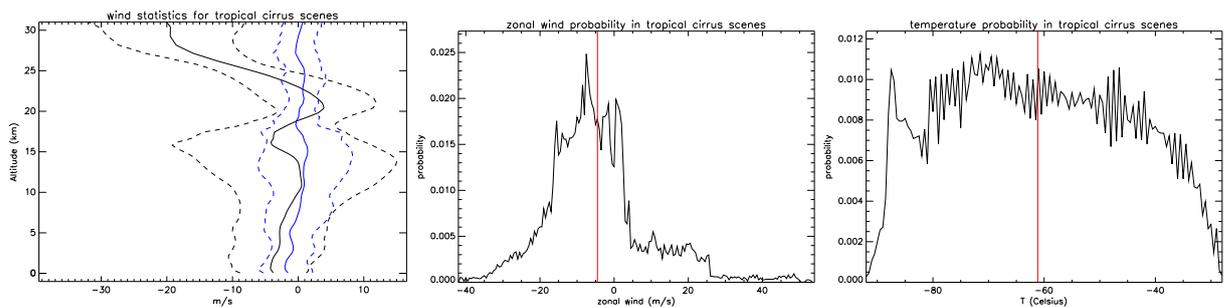


Figure 60: Wind and temperature statistics for scenes with tropical cirrus on 1/1/2007. Left, zonal (black) and meridional (blue) wind distribution in the tropics. solid and dashed lines denote the mean and mean +/- one standard deviation respectively. Zonal wind (middle) and temperature (right) statistics at the location of tropical cirrus. Red lines correspond to the mean value. The tropics is bounded here at 30 degrees latitude.

tropical cirrus were selected from the results. Figure 60 shows the wind statistics for scenes containing tropical cirrus. Note that the HLOS wind is close to minus the zonal wind in the Tropics. From the right panel, the mean zonal wind in the region of tropical cirrus presence between 10 and 20 km is close to -5 ms^{-1} with most of the winds found between -20 and $+5 \text{ ms}^{-1}$, that is between -12 and 3 ms^{-1} for the LOS wind component. The temperature inside tropical cirrus (right panel) is between -90 and -30 degrees Celsius with a mean value close to -60 degrees Celsius. Statistics for the scattering ratio inside cirrus clouds are found in Figs. 61 and 62 with a mean value of 4 but with a large spread. In section 8.0.11 it is shown that processing option 1 generally has larger biases than the other processing options in cirrus regions between 10 and 20 km.

From the statistics of the tropical cirrus parameters and from the discussion in section 5.5.3 we conclude that temperature and scattering ratio fulfil the conditions for expected substantial bias through cross-talk, but the LOS wind is generally small in this region yielding biases of less than 0.1 ms^{-1} in general.

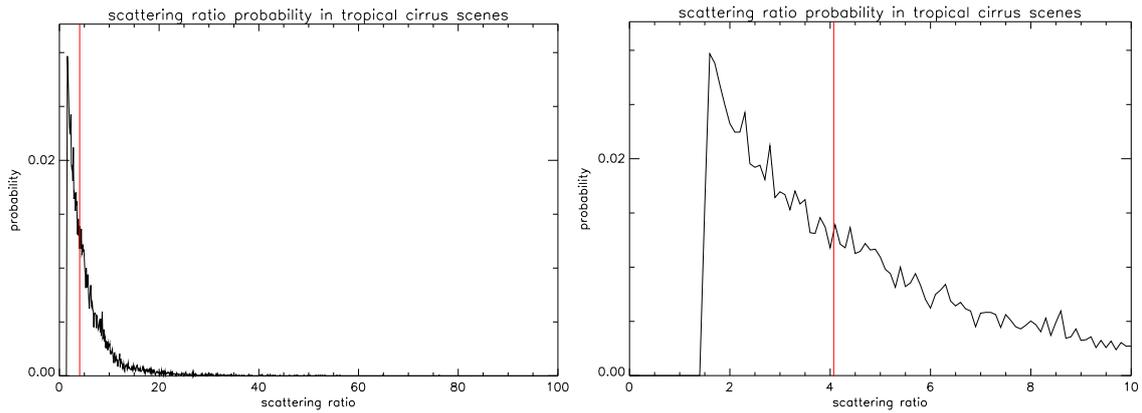


Figure 61: Scattering ratio statistics inside tropical cirrus clouds. The right plot is zoomed in to the lowest scattering ratio values. Red lines correspond to the mean value.

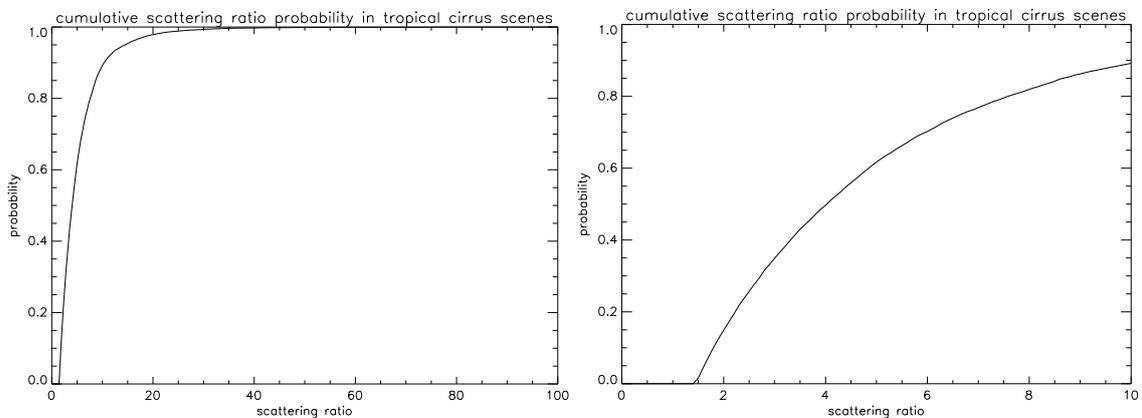


Figure 62: Same as Fig. 61 but now for the cumulative probability.

Figure 63 shows that the Mie wind error standard deviation is generally smaller than the Rayleigh wind error standard deviation with values of 1-1.5 ms^{-1} for Mie winds and 1.5-2.5 ms^{-1} for Rayleigh winds throughout most of the troposphere and lower stratosphere. The standard deviation of Mie and Rayleigh winds is within the mission requirement throughout the troposphere and lower stratosphere. The Rayleigh wind error statistics are quite similar for all three processing options with slightly smaller errors for option 1 in the PBL. From these results it appears that processing option 1 is favourable for Rayleigh wind retrieval, although caution is needed in areas where the conditions for substantial cross-talk biases apply.

Figure 64 through Fig. 66 show similar results but now for sampling scenario WVM2 with more Mie bins in the PBL and less in the UTLS.

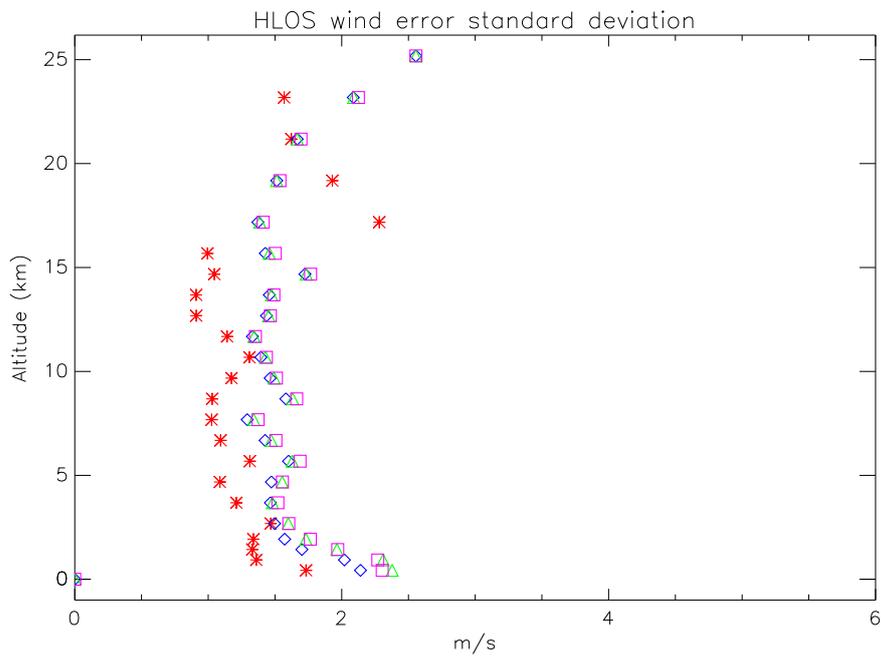


Figure 63: HLOS wind error standard deviation for Mie and Rayleigh channel for vertical sampling scenario WVM1. For an explanation of the symbols see Fig. 59.

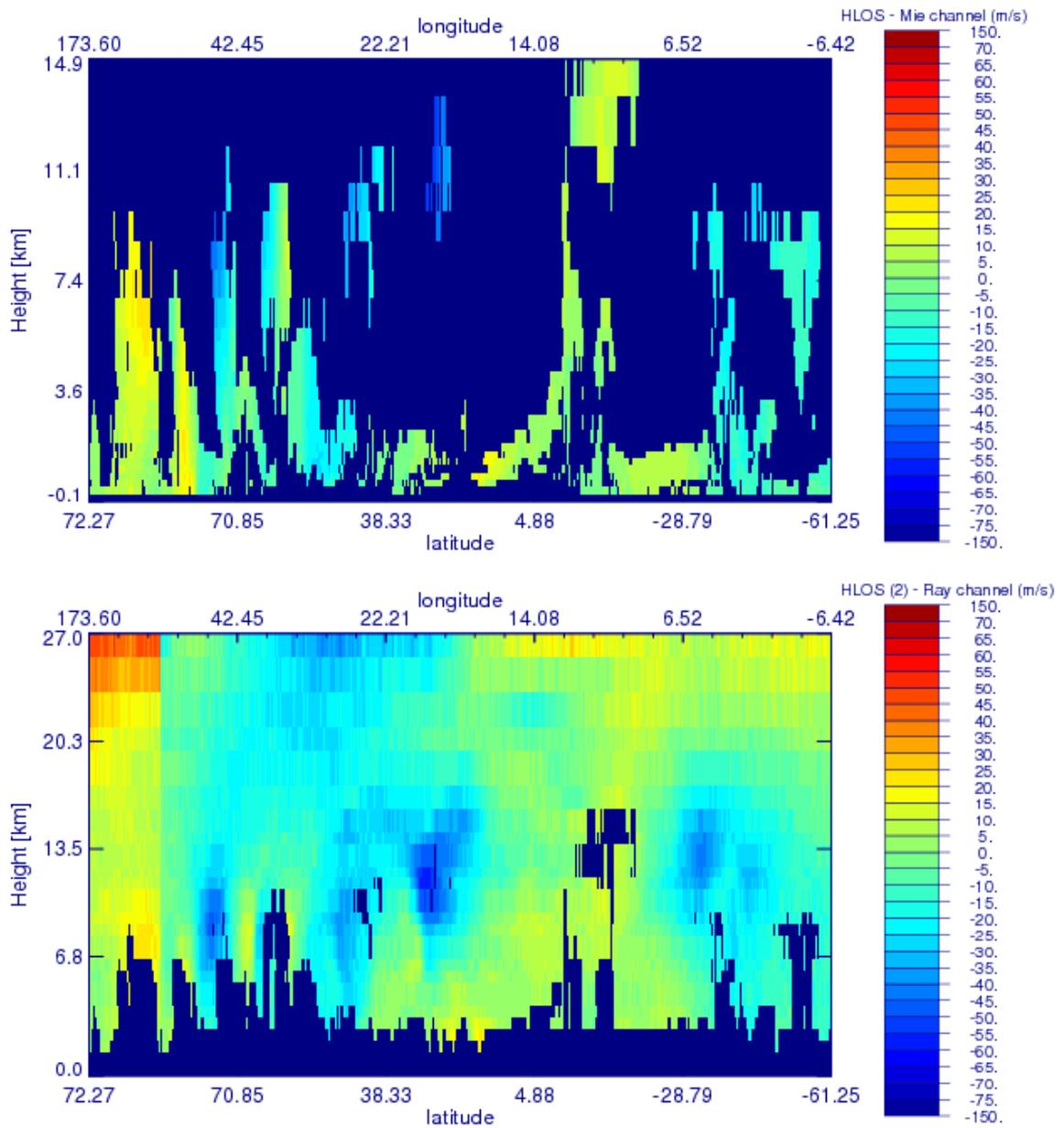


Figure 64: Same as Fig. 57 but now for sampling scenario WVM2 and retrieval option 2 for the Rayleigh winds.

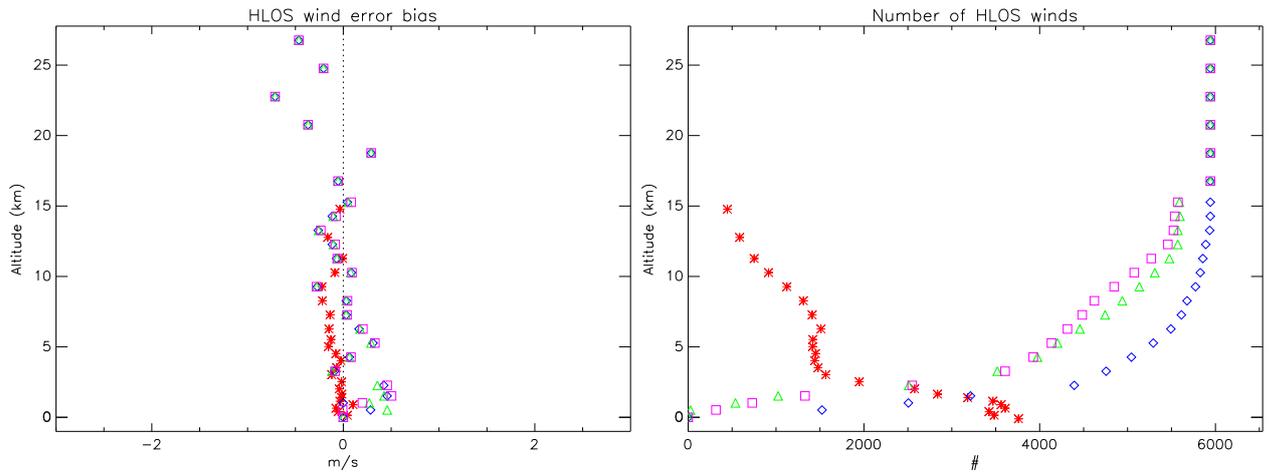


Figure 65: Same as Fig. 59 but now for sampling scenario WVM2

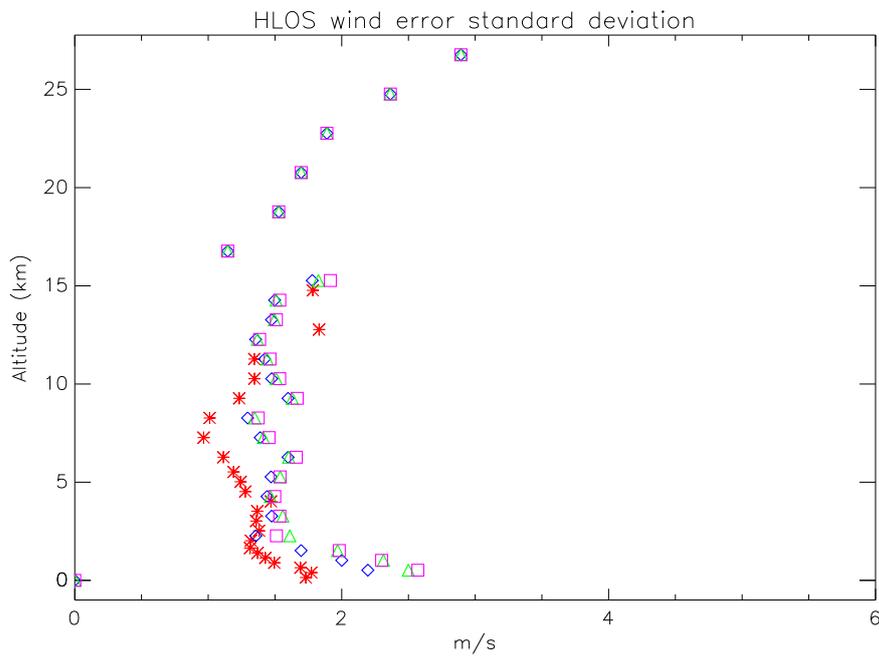


Figure 66: Same as Fig. 63 but now for sampling scenario WVM2

7 Task 4c: Vertical sampling scenarios information content

One way to quantify the information content of a vertical sampling scenario is by its data coverage, i.e. counting the number of good AND bad quality winds at various altitudes and as a function of climate zone. It is important to note that a small number of bad quality winds can be detrimental for NWP analyses. The number of erroneous winds that can not be identified by quality control therefore has a large impact on the information content of the sampling scenario, even if the number of good quality winds is large.

The atmosphere is subdivided in four regions: the PBL (0-2 km), lower-troposphere (2-7 km), upper troposphere lower stratosphere (UTLS) (7-15 km) and stratosphere (above 15 km). For each of these regions the number of good, moderate, bad and quality-controlled winds from both Mie and Rayleigh channels are determined as a function of sampling scenario. Qualification of HLOS winds is performed as follows. First, the quality control flag value, see sections 6.2.1 and 6.2.2, is considered. If the flag is set, meaning that the quality of the retrieved HLOS wind is suspicious, then the HLOS wind is quality-controlled and not further processed. Otherwise, the retrieved HLOS wind is valid and verified against the true HLOS wind. Three classes for the HLOS wind error are defined: 1) good quality if the HLOS wind error is less than σ_r , 2) moderate quality if the HLOS wind error is between σ_r and $2\sigma_r$ and 3) bad quality if the HLOS wind error is larger than $2\sigma_r$. Here σ_r is the mission requirement HLOS wind error standard deviation, not including the model representativeness error. The HLOS wind error standard deviation mission requirement for HLOS winds including the representativeness error is obtained from Table 3.3 of the ADM-Aeolus Science Report¹ and equals 2 ms^{-1} for the PBL, 3 ms^{-1} for the lower-troposphere and UTLS and 5 ms^{-1} for the stratosphere as defined above. The model representativeness error standard deviation, discussed in section 5.7.2, is a function of altitude with values obtained from the line error column of Table 3.2 in the ADM-Aeolus Science Report, i.e. $1.5/1.7/2.0/1.5 \text{ ms}^{-1}$ in the PBL/lower troposphere/UTLS/stratosphere respectively. It is noted here that these values are more optimistic, i.e. smaller, than the values found in section 5.7.2. The mission requirement excluding the model representativeness error then equals $1.32/2.47/2.24/4.77 \text{ ms}^{-1}$ respectively.

In this section the information content of the scenarios in sections A.1 and A.2 are considered. These scenarios were originally defined in [RD6]. At a later stage an additional maximum height constraint of 31 km was considered. These scenarios are found in section A.3 and are considered state of the art at the time of writing. Results from these scenarios are presented in section 8. For the sake of completeness and to illustrate their drawbacks the originally defined scenarios are discussed in this section.

7.0.1 PBL wind information content

Figure 67 shows the number of Mie winds measured in the PBL by the various vertical sampling scenarios displayed in the first two sections of Appendix A. The colours correspond to the winds quality. The relative large number of bad quality data is explained by a more relaxed quality control in the PBL than at higher altitudes. This is motivated by the fact that the representativeness error in the PBL is close to the mission requirement, meaning that almost all Mie (and Rayleigh) winds would be flagged as suspicious. For this reason HLOS winds in the PBL were flagged as suspicious when the estimated accuracy was less than 2 ms^{-1} rather than 1.32 ms^{-1} . However, for the quality assignment the more stringent value of $\sigma_r = 1.32 \text{ ms}^{-1}$ was used. From Fig. 67 it is clear that many Mie winds are flagged which is explained mainly by insufficient SNR in cloudy regions. Figure 68 shows a similar plot but now for PBL Rayleigh channel winds. The data coverage of all three retrieval options is obviously identical with more flagged winds for options 2 and 3. Relaxing the quality control in the PBL implies that Rayleigh winds may pass the quality control for 500 m bins, but yielding a relative small number of good quality winds. Option 1 provides both more good and bad quality winds and this latter property makes this option less favourable. Figure 69 shows the combined Mie and Rayleigh data coverage in the PBL obtained by adding the numbers in Figs. 67 and 68. There may be redundancy in this plot in case of both good quality Mie and Rayleigh winds in the same bin, although this can only be the case for processing option 1.

¹http://esamultimedia.esa.int/docs/SP-1311_ADM-Aeolus_FINAL_low-res.pdf

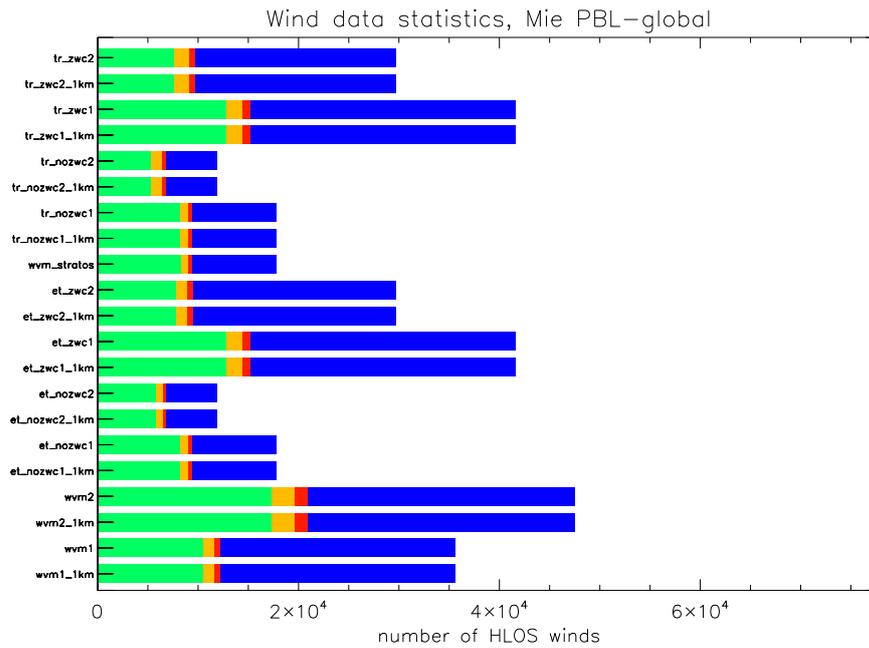


Figure 67: Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the PBL for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.

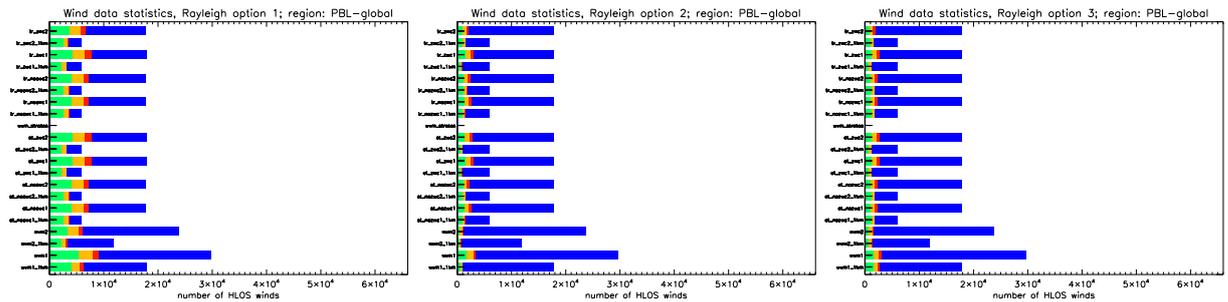


Figure 68: Same as Fig. 67 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).

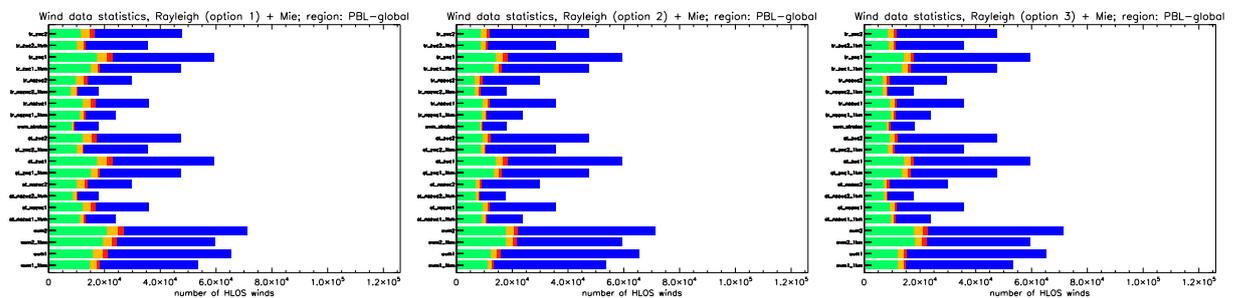


Figure 69: Total number of good, moderate, bad and quality controlled HLOS wind obtained from accumulating the Mie (Fig. 67) and Rayleigh (Fig. 68) winds for processing option 1 (left), option 2 (middle) and option 3 (right).

7.0.2 Lower-troposphere wind information content

Figure 70 shows the number of Mie winds measured in the lower troposphere by the various vertical sampling scenarios displayed in Appendix A. Note the substantially smaller number of bad quality winds as compared to the PBL, because no quality control relaxation was applied in this region. From Fig. 70 it is clear that many Mie winds are flagged which is explained by the generally small aerosol loading above the PBL. Most Mie winds are obtained in cloudy regions. Figure 71 shows a similar plot but now for lower troposphere Rayleigh channel winds. The data coverage is equal for all scenarios because the bin size has a constant value of 1000 m for all scenarios in this region. Again, more winds are flagged for options 2 and 3 than for option 1. Figure 72 shows the combined Mie and Rayleigh data coverage in the lower troposphere obtained by adding the numbers in Figs. 70 and 71.

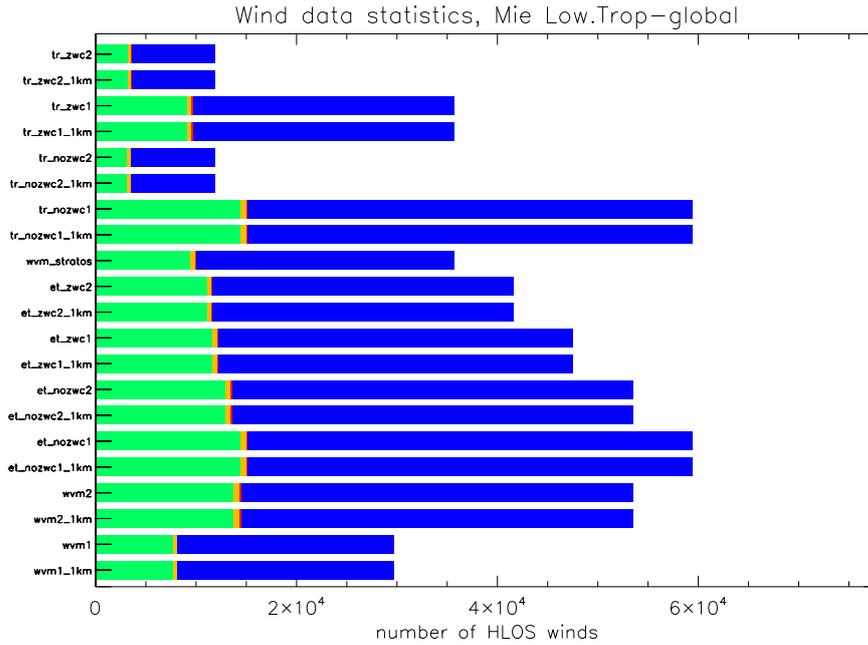


Figure 70: Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the lower troposphere for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.

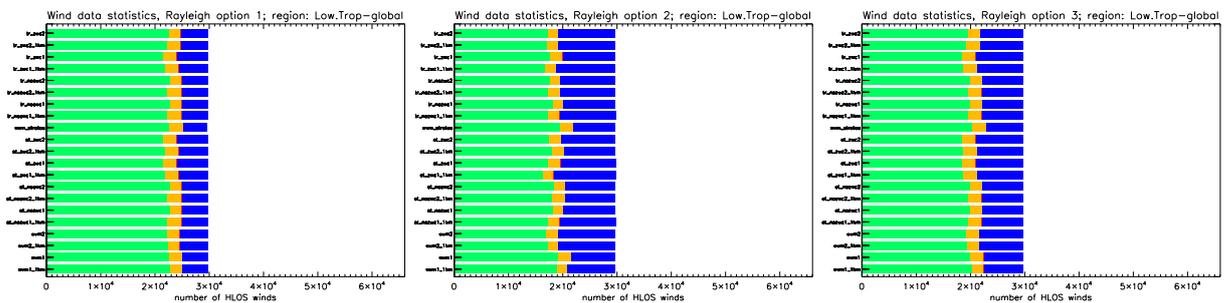


Figure 71: Same as Fig. 70 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).

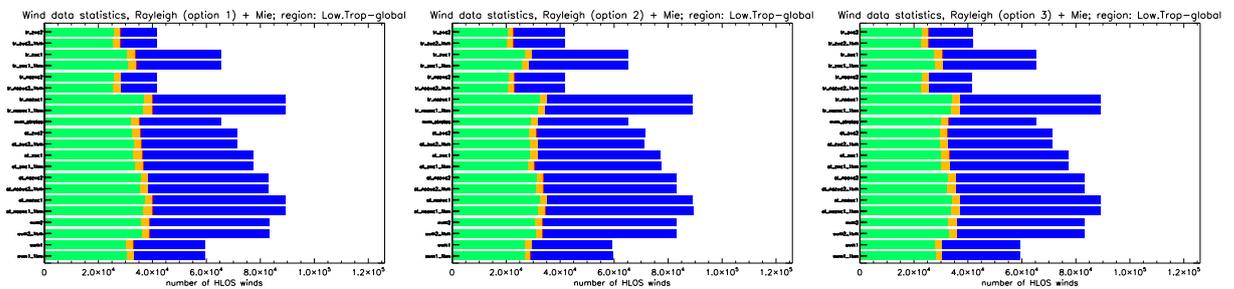


Figure 72: Same as Fig. 69 but now for the lower troposphere.

7.0.3 Upper-troposphere lower-stratosphere wind information content

Figure 73 shows the number of Mie winds measured in the UTLS by the various vertical sampling scenarios displayed in Appendix A. From Fig. 73 it is clear that many Mie winds are flagged which is explained by the generally negligible aerosol loading in this region. Virtually all Mie winds are obtained in cloudy regions. Figure 74 shows a similar plot but now for UTLS Rayleigh channel winds. The data coverage is equal for all scenarios because the bin size has a constant value of 1000 m for all scenarios in this region. Again, more winds are flagged for options 2 and 3 than for option 1. Figure 75 shows the combined Mie and Rayleigh data coverage in the UTLS obtained by adding the numbers in Figs. 73 and 74

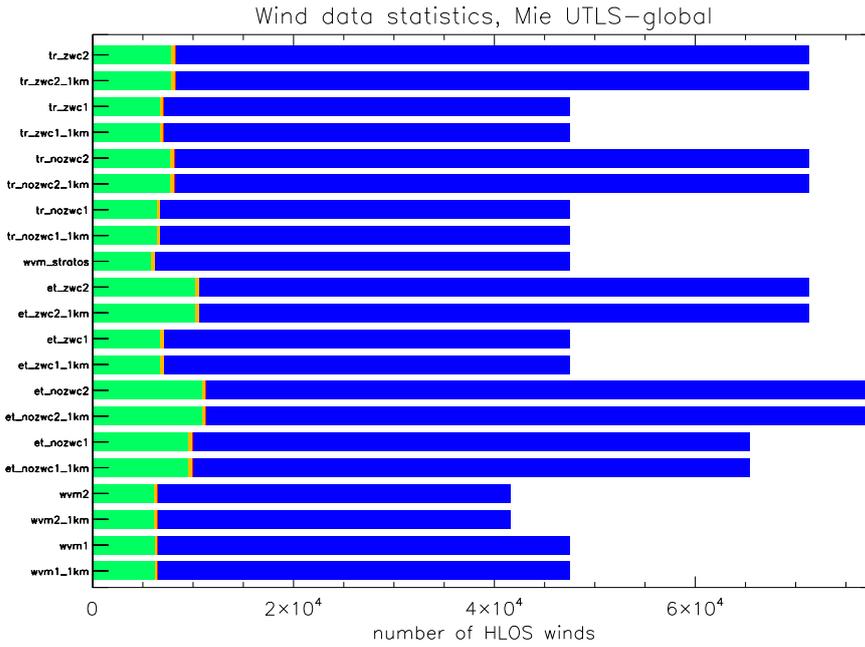


Figure 73: Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the UTLS for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.

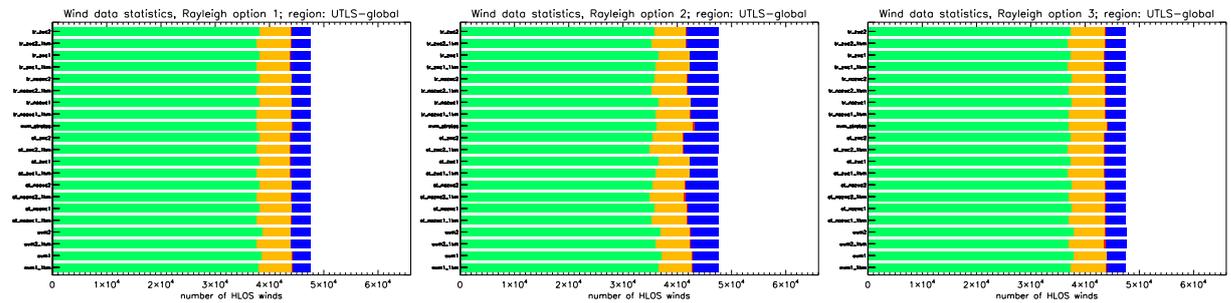


Figure 74: Same as Fig. 73 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).

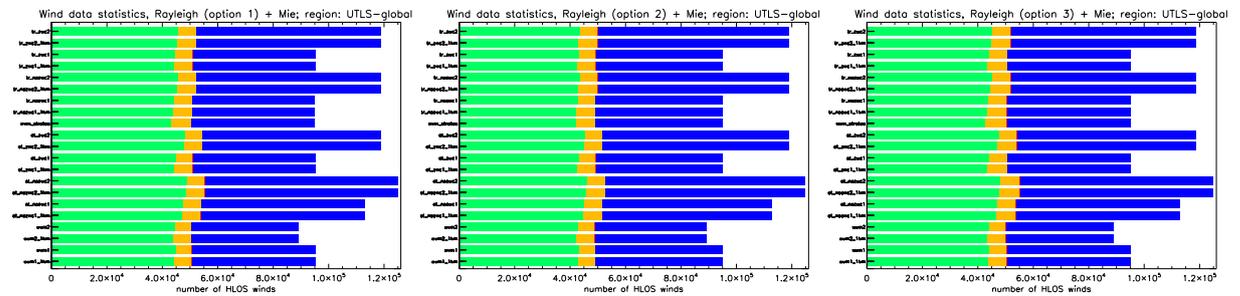


Figure 75: Same as Fig. 69 but now for the UTLS.

7.0.4 Stratosphere wind information content

Figure 76 shows the number of Mie winds measured in the stratosphere above 15 km by the various vertical sampling scenarios displayed in Appendix A. For some scenarios the Mie channel does not reach this altitude. The Mie channel is generally not very efficient in this region, except in the tropics in case of cirrus clouds. From Fig. 76 it is clear that many Mie winds are flagged which is explained by the generally negligible aerosol and cloud loading at these altitudes. Virtually all Mie winds are obtained in cloudy regions in the tropics. Figure 77 shows a similar plot but now for stratosphere Rayleigh channel winds. The statistics for all 3 options are similar and virtually all winds have good quality with only a small number of flagged winds because the atmosphere is free of particles. Figure 78 shows the combined Mie and Rayleigh data coverage in the stratosphere obtained by adding the numbers in Figs. 76 and 77.

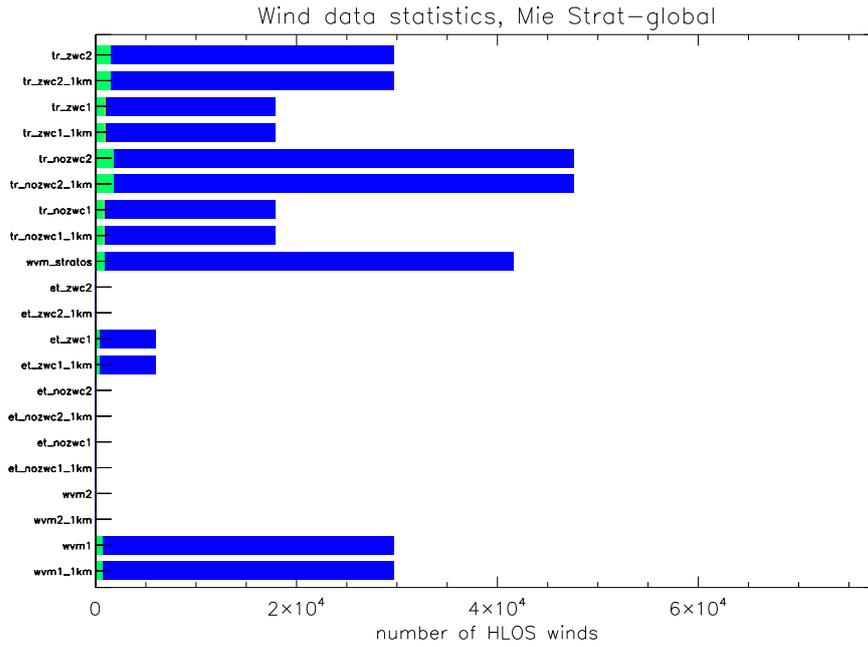


Figure 76: Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the stratosphere for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.

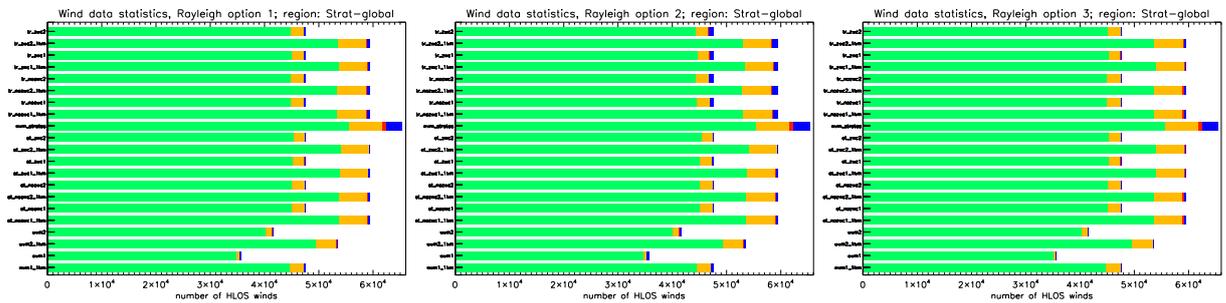


Figure 77: Same as Fig. 76 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).

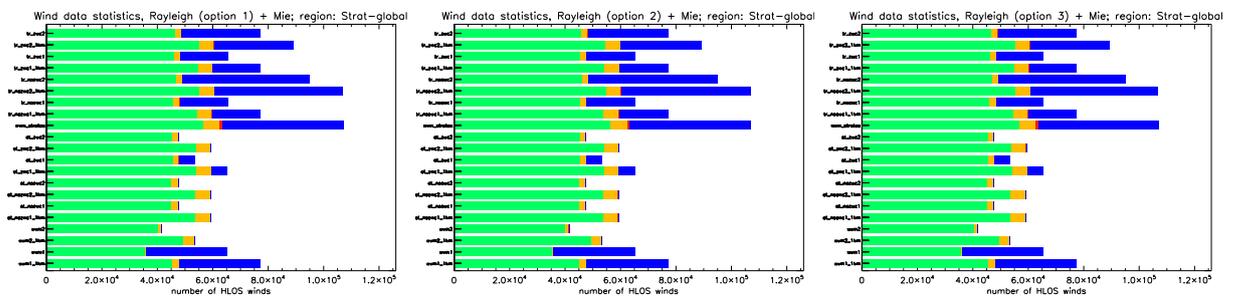


Figure 78: Same as Fig. 69 but now for the stratosphere.

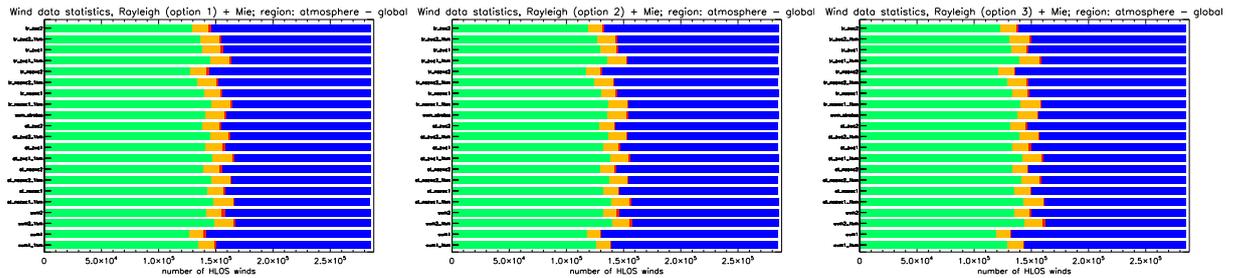


Figure 79: Same as Fig. 79 but now for all atmospheric regions accumulated. These plots show the total data coverage and HLOS wind quality throughout the atmosphere.

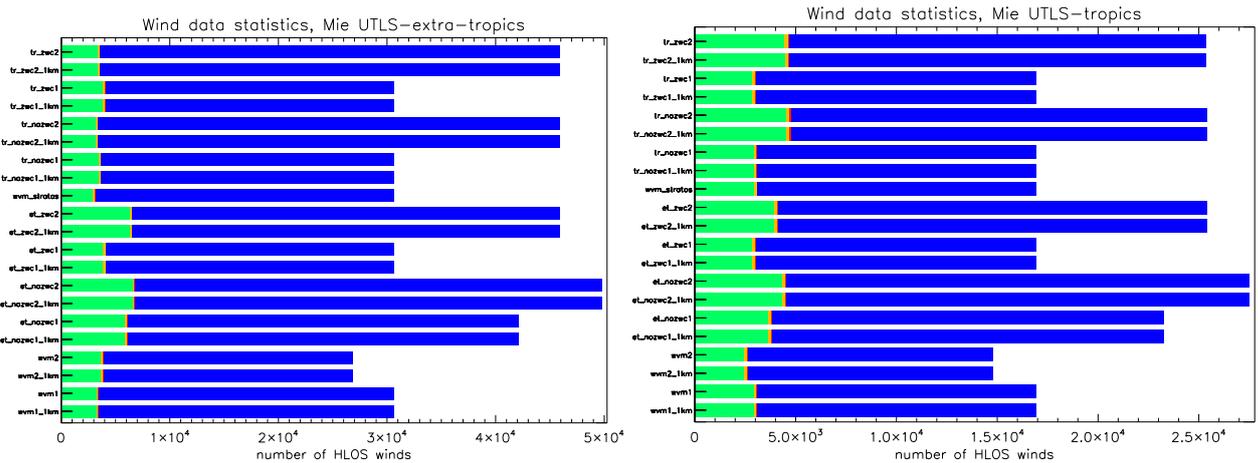


Figure 80: Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the UTLS in the extra-tropics (left) and tropics (right).

7.0.5 Total wind information content

Figure 79 shows the overall result of combined Mie and Rayleigh data coverage from all atmospheric regions obtained by adding the numbers in Figs. 69, 72, 75 and 78. The total number of wind observations equals $6000 \text{ (profiles)} \times 24 \text{ (bins)} \times 2 \text{ (channels)} \approx 2.88 \cdot 10^5$. From this figure it is concluded that about half of the retrieved HLOS winds is flagged as suspicious mainly because of lack of SNR. It is clear that processing option 1 yields most good quality winds and processing option 3 yields slightly more good quality winds than processing option 2. It is also clear that all scenarios with 500 m Rayleigh bins yield less good-quality winds than their 1000 m bins (denoted with `_1km`) counterparts. Also the increased Mie resolution in the UTLS in the tropical scenarios `tr_zwc2` and `tr_nozwc2` does not pay off in the global picture and is at the expensive of Mie winds in other height levels. However, when zooming in on the tropical region, `tr_zwc2_1km` shows more good quality winds than e.g. `wvm1_1km` (not shown). Also note that these plots do not take into account that wind information at a certain height level may be more informative for a NWP model than winds at other height levels.

The relative merit of some scenarios over others is most pronounced by the location of the Mie channel bins in the UTLS and stratosphere. In Fig. 80 we separate tropical and extra-tropical UTLS winds. It is seen that the extra-tropical scenarios (denoted with `et_`) are indeed more favourable in the extra-tropics. Similarly the tropical scenarios (denoted with `tr_`) are more favourable in the tropics. This is even more pronounced for the stratosphere in Fig. 81 where none of the scenarios has valid winds in the extra-tropics but all tropical scenarios still have some good quality winds in the tropics.

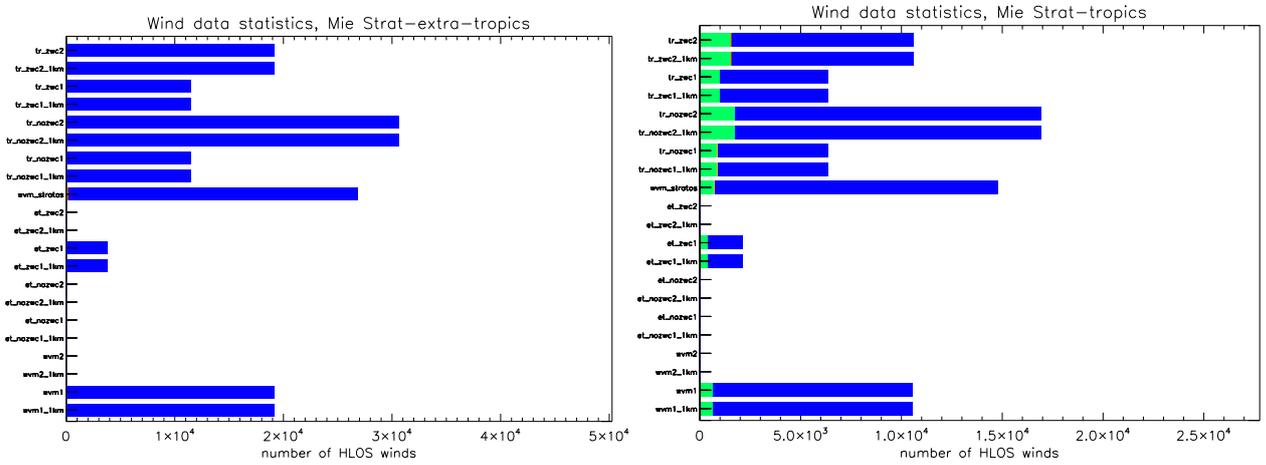


Figure 81: Same as Fig. 80 but now for the stratosphere above 15 km.

7.0.6 Wind profile information content

Counting of the number of good/bad/moderate winds is a measure of information content but has some drawbacks. For instance scenarios without zero wind calibration measure more winds. Reduced information content through missing calibration was not taken into account in the previous section thus overestimating the information content of these scenarios. Second, there may be overestimation in case of overlapping Mie and Rayleigh bins where both channels produce a good quality wind. Both are counted but only one is used in data assimilation (DA) in practice.

Another measure of information content that is more in analogy with DA is discussed here. In DA the weight (or information content) of an observation in the analysis is proportional to the inverse of its error variance estimate. Using this definition the information content for each DWL wind profile can be computed by adding the information content of all individual Mie and Rayleigh winds within the wind profile and assuming no correlation between observations for simplicity. To assign a quality measure to such a DWL profile the information content can be related to the information content of a synthetic radiosonde at the same location, with the same amount of vertical bins and measuring a single wind component. The information content of each vertical sampling scenario is then quantified relative to a corresponding synthetic radiosonde.

To account for the calibration issue the method described in [RD13] has been adopted. There, a constant bias value was added to all elements of the observation error covariance matrix, i.e. the lack of absolute calibration behaves as a correlated error component with an infinitely large correlation scale. Equivalently this also means that the information on the deepest vertical eigenfunction, reflecting the vertical average in the observation, is lost. In other words, a singular vector decomposition of the observation error covariance matrix only changes the first eigenvalue after adding the constant bias term. In terms of information content this means that the information of one bin is lost, or a loss of about 4% for a 24 bin profile. This correction factor is applied to sampling scenarios that lack zero wind calibration opportunities.

The total error variance for a radiosonde wind component is the sum of the instrument error variance and point measurement representativeness error variance. The former is assumed constant throughout the profile with a value of $1 \text{ m}^2\text{s}^{-2}$. Figure 82 shows the mean DWL profile information content relative to that of (synthetic) radiosondes (at the same location). The information content of all scenarios is close to 0.7 meaning that the information content of one DWL profile is on average 70% of the information content of a radiosonde. The, on average, smaller information content of a DWL profile relative to a radiosonde is mainly explained by clouds obscuring the lower part of the atmosphere for DWL but not for radiosonde. A limitation of the proposed information content calculation is that DWL measurements above 20 km yield little information relative to a radiosonde because the instrument error increases rapidly for DWL at higher altitudes, but not for radiosondes, while the representativeness error of DWL winds is only slightly smaller than for radiosondes.

Overall, Fig. 82 is in close agreement with Fig. 79. Based on these global results it is hard to make a definite selection of most informative scenarios. Each of the scenarios shows merit in different global regions. An alternative measure for information content is to use the model error as a reference rather than a radiosonde. This is discussed in [RD13].

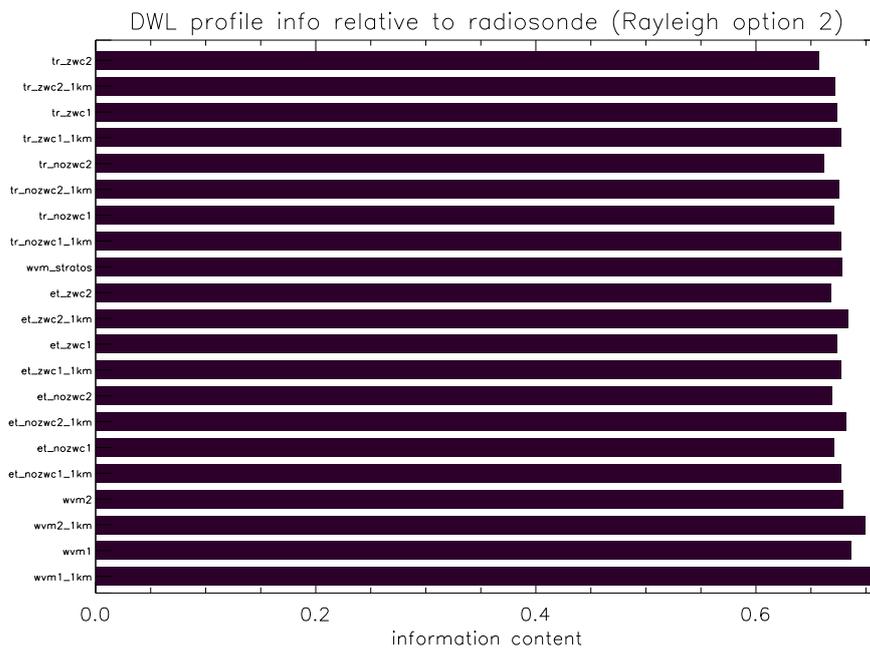


Figure 82: DWL profile information content relative to radiosonde information content. Retrieval option 2 (classification) was used here for Rayleigh channel winds.

8 Task 4c: Vertical sampling scenarios information content - new scenarios

In this section the information content of the scenarios in section A.3 are considered similar as in the previous sections. The scenarios presented in section A.3 and discussed in this section are considered to be state of the art, fulfilling the requirements defined in [RD6]. Before presenting the results the reader is referred to Appendix A for an overview of the overall statistics of Mie and Rayleigh winds for all sampling scenarios.

8.0.7 PBL wind information content

Figure 83 shows the number of Mie winds measured in the PBL by the various vertical sampling scenarios displayed in section A.3. From Fig. 83 it is clear that many Mie winds are flagged which is explained mainly by insufficient SNR in cloudy regions. Figure 84 shows a similar plot but now for PBL Rayleigh channel winds. The data coverage of all three retrieval options is obviously identical with more flagged winds for options 2 and 3. Relaxing the quality control in the PBL implies that Rayleigh winds may pass the quality control for 500 m bins, but yielding a relative small number of good quality winds. Processing option 1, and to a lesser extent processing option 3, provides both more good and bad quality winds and this latter property makes this option less favourable. Figure 85 shows the combined Mie and Rayleigh data coverage in the PBL obtained by adding the numbers in Figs. 83 and 84. There may be redundancy in this plot in case of both good quality Mie and Rayleigh winds in the same bin, although this can only be the case for processing option 1.

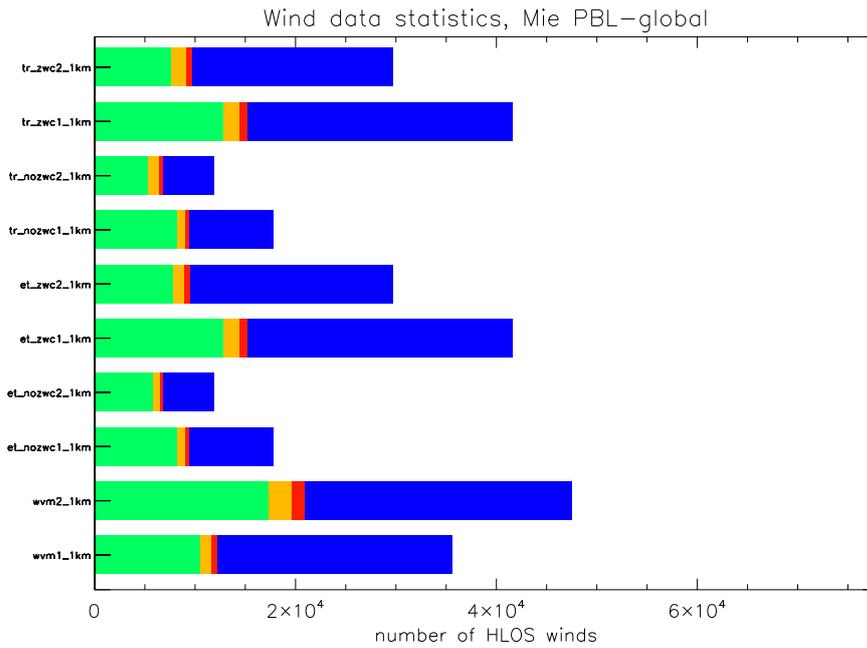


Figure 83: Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the PBL for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.

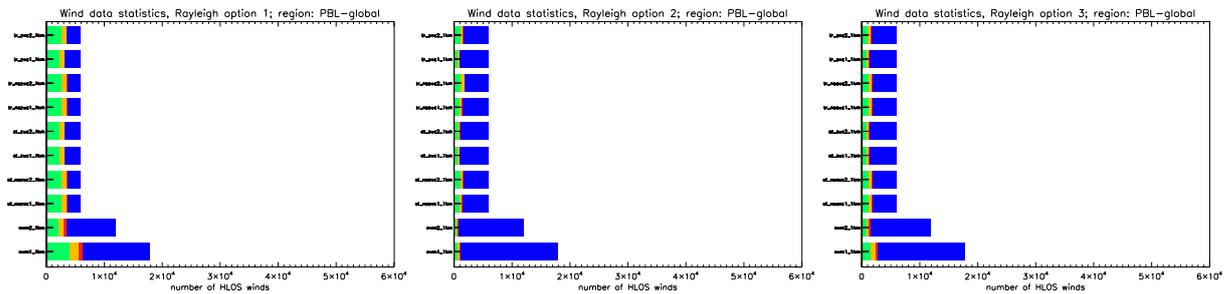


Figure 84: Same as Fig. 83 but now for Rayleigh channel winds retrieved with processing option 1 (left), option 2 (middle) and option 3 (right).

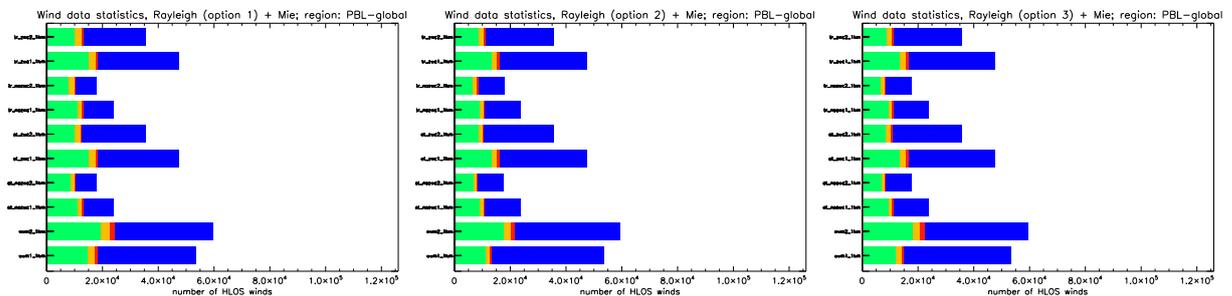


Figure 85: Total number of good, moderate, bad and quality controlled HLOS wind obtained from accumulating the Mie (Fig. 83) and Rayleigh (Fig. 84) winds for processing option 1 (left), option 2 (middle) and option 3 (right).

8.0.8 Lower-troposphere wind information content

Figure 86 shows the number of Mie winds measured in the lower troposphere by the various vertical sampling scenarios displayed in Appendix A. Note the substantially smaller number of bad quality winds as compared to the PBL, because no quality control relaxation was applied in this region. From Fig. 86 it is clear that many Mie winds are flagged which is explained by the generally small aerosol loading above the PBL. Most Mie winds are obtained in cloudy regions. Figure 87 shows a similar plot but now for lower troposphere Rayleigh channel winds. The data coverage is equal for all scenarios because the bin size has a constant value of 1000 m for all scenarios in this region. Again, more winds are flagged for options 2 and 3 than for option 1. Figure 88 shows the combined Mie and Rayleigh data coverage in the lower troposphere obtained by adding the numbers in Figs. 86 and 87.

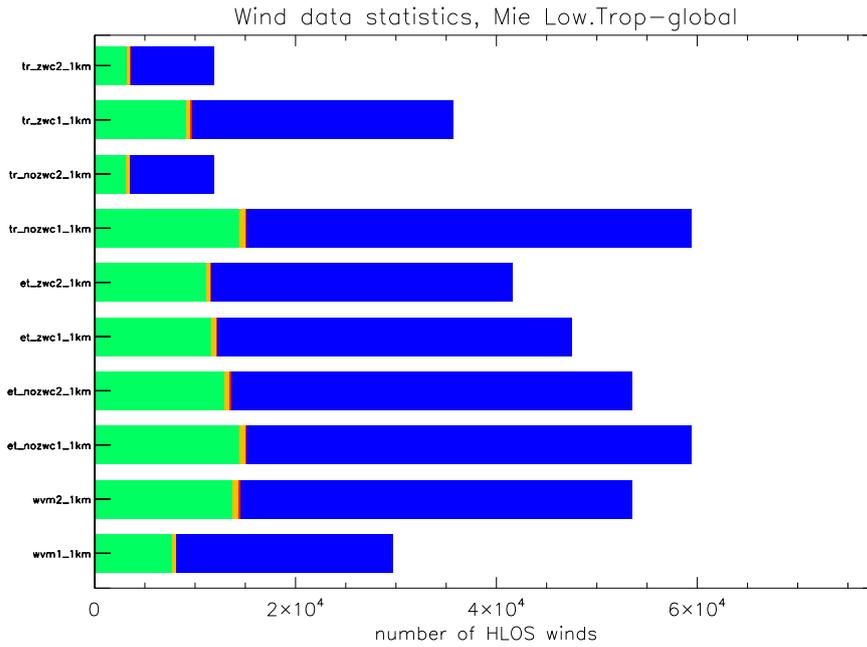


Figure 86: Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the lower troposphere for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.

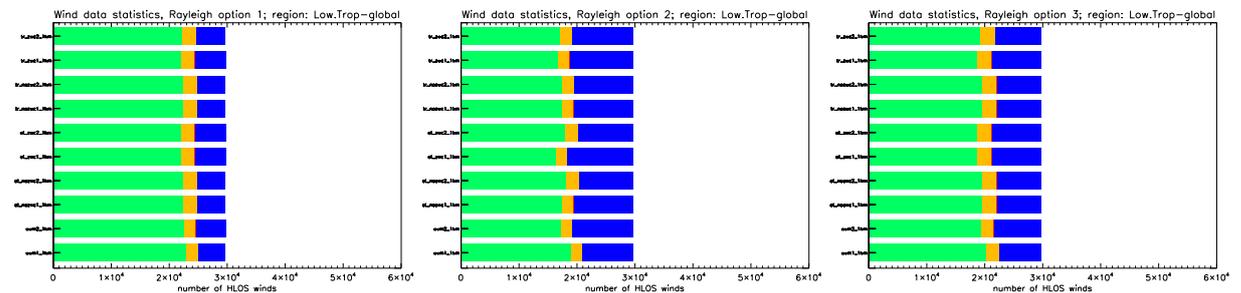


Figure 87: Same as Fig. 86 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).

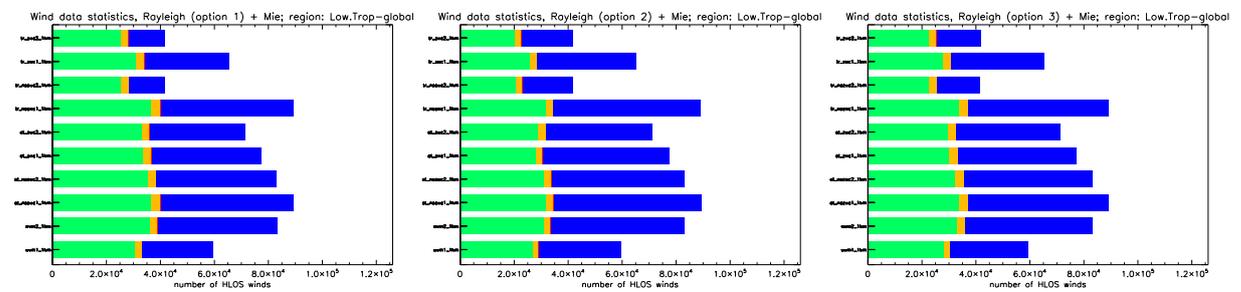


Figure 88: Same as Fig. 85 but now for the lower troposphere.

8.0.9 Upper-troposphere lower-stratosphere wind information content

Figure 89 shows the number of Mie winds measured in the UTLS by the various vertical sampling scenarios displayed in Appendix A. From Fig. 89 it is clear that many Mie winds are flagged which is explained by the generally negligible aerosol loading in this region. Virtually all Mie winds are obtained in cloudy regions. Figure 90 shows a similar plot but now for UTLS Rayleigh channel winds. The data coverage is equal for all scenarios because the bin size has a constant value of 1000 m for all scenarios in this region. Again, more winds are flagged for options 2 and 3 than for option 1. Figure 91 shows the combined Mie and Rayleigh data coverage in the UTLS obtained by adding the numbers in Figs. 89 and 90

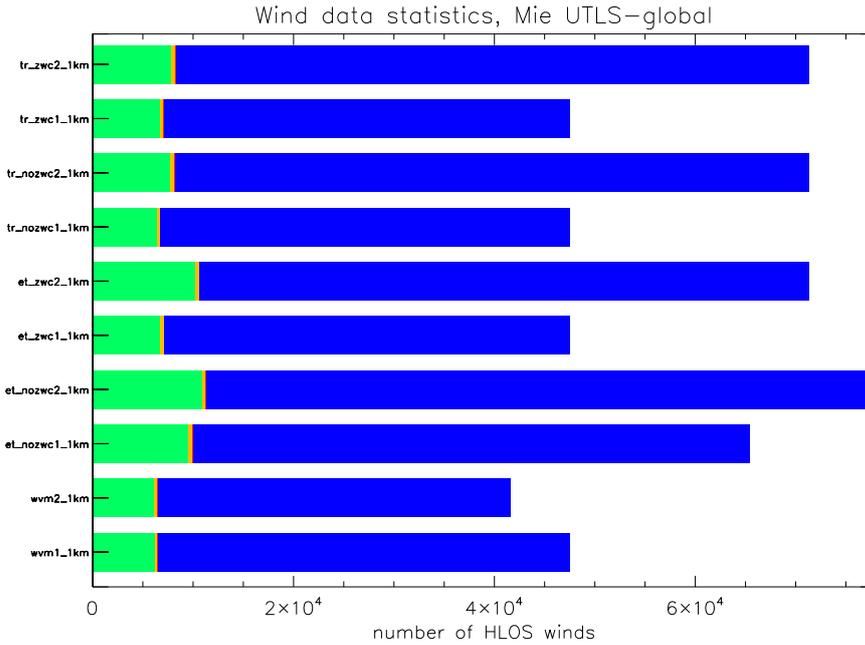


Figure 89: Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the UTLS for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.

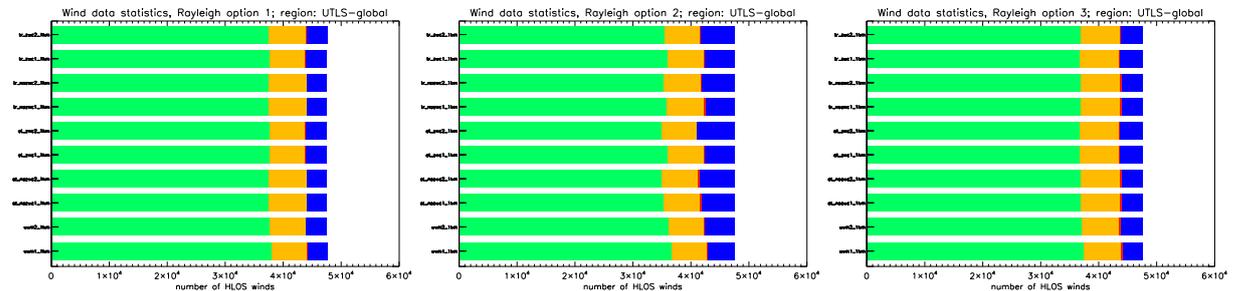


Figure 90: Same as Fig. 89 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).

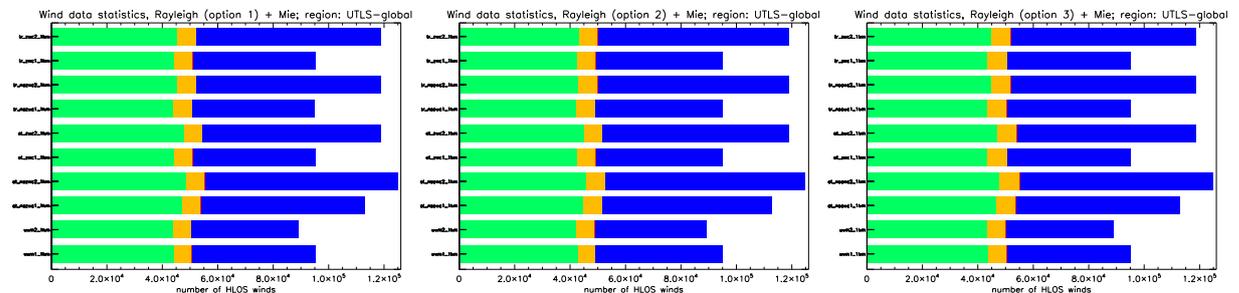


Figure 91: Same as Fig. 85 but now for the UTLS.

8.0.10 Stratosphere wind information content

Figure 92 shows the number of Mie winds measured in the stratosphere above 15 km by the various vertical sampling scenarios displayed in Appendix A. For some scenarios the Mie channel does not reach this altitude. The Mie channel is generally not very efficient in this region, except in the tropics in case of cirrus clouds. From Fig. 92 it is clear that many Mie winds are flagged which is explained by the generally negligible aerosol and cloud loading at these altitudes. Virtually all Mie winds are obtained in cloudy regions in the tropics. Figure 93 shows a similar plot but now for stratosphere Rayleigh channel winds. The statistics for all 3 options are similar and virtually all winds have good quality with only a small number of flagged winds because the atmosphere is free of particles. Figure 94 shows the combined Mie and Rayleigh data coverage in the stratosphere obtained by adding the numbers in Figs. 92 and 93.

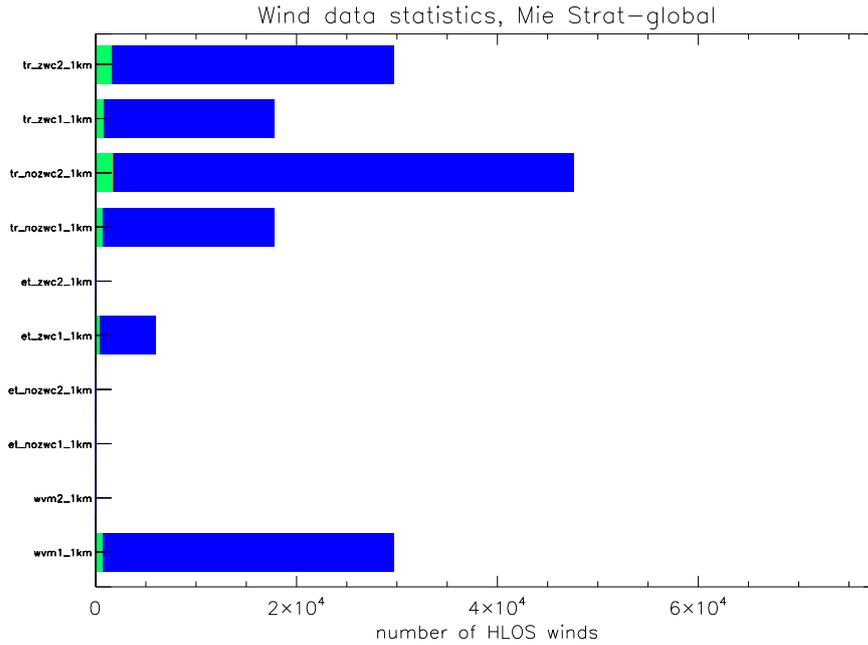


Figure 92: Number of good (green), moderate (orange), bad (red) and quality controlled (blue) Mie channel HLOS winds in the stratosphere for the sampling scenarios defined along the y-axis. Data are obtained from one day (1/1/2007) covering 15 (half CALIPSO) orbits and about 6000 HLOS wind profiles.

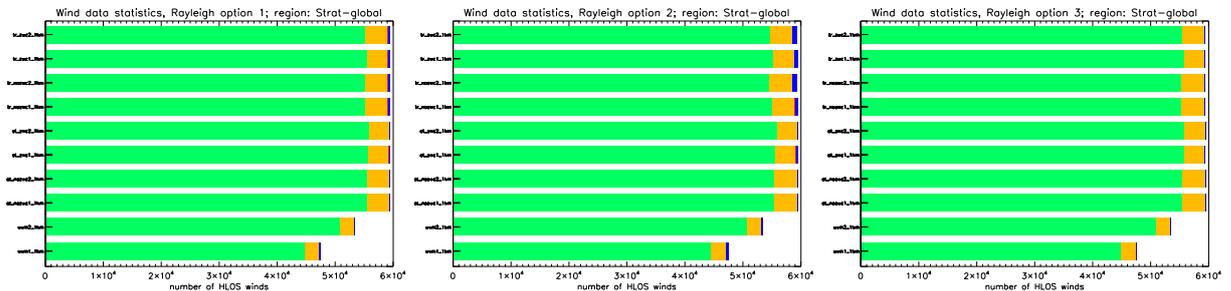


Figure 93: Same as Fig. 92 but now for Rayleigh channel winds retrieved with option 1 (left), option 2 (middle) and option 3 (right).

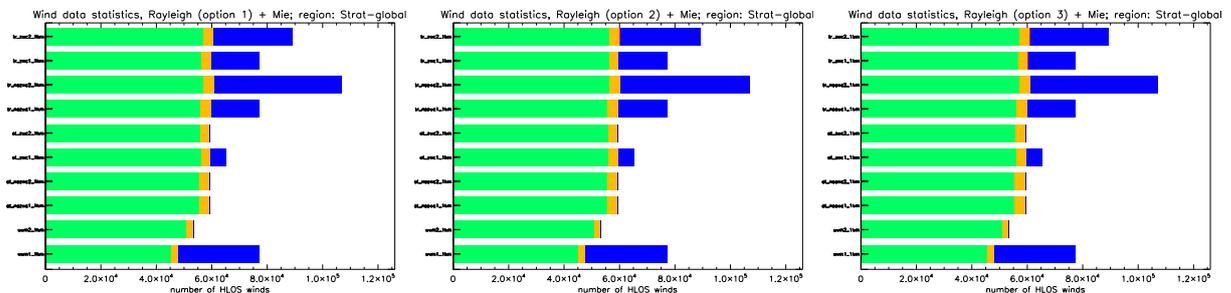


Figure 94: Same as Fig. 85 but now for the stratosphere.

8.0.11 HLOS wind quality in tropical cirrus

In this section we zoom in to scenes of tropical cirrus and determine the quality and coverage of HLOS winds for the vertical sampling scenarios displayed in Appendix A.3. Figure 95 shows the quality of Mie winds in scenes with tropical cirrus between 10 and 20 km altitude. The Mie channel resolution and bin locations is different for all scenarios in this region. The Mie wind bias is small for all scenarios with a maximum of about 0.2 ms^{-1} for scenario `tr_zwc1` that has relative low resolution in this region. Also the standard deviation is largest for scenarios with relative resolution: `tr_zwc1`, `tr_nozwc1` and `wvm2`. Scenarios `tr_zwc2` and `tr_nozwc2` with maximum resolution in this region perform best with the largest coverage (right panel) and smallest standard deviation. The data coverage of the extra-tropical scenarios is smaller than for the tropical scenarios because of lower maximum Mie channel altitude.

Figures 96, 97 and 98 show the quality and coverage of Rayleigh channel winds for processing options 1, 2 and 3 respectively. The bias is largest for processing option 1 as expected because of cross-talk. Cross-talk correction (processing option 3) reduces the HLOS wind bias substantially by about 50%, but the standard deviation also increases by about 50% from 1.5 ms^{-1} for option 1 to about 2.25 ms^{-1} for option 3, which is explained by the reduced SNR after cross-talk correction. As a result the quality of Rayleigh winds in cirrus scenes is generally worse than in cloud-free scenes. The data coverage for option 1 is larger than for option 3 as expected because more winds are flagged as suspicious due to reduced SNR for option 3. The data coverage of option 2 is substantially smaller, because no Rayleigh winds are obtained for processing option 2 if all measurements are classified as "particle". For partly cloudy scenes a wind estimate may be obtained from the "particle-free" measurements, but in many cases the SNR will be too low and the wind solution is flagged as suspicious. The bias of the remaining winds is larger than for option 3 but smaller than for option 1. The standard deviation is close to option 3 and smaller than for option 1 because only a reduced number of measurements are accumulated thus reducing the SNR.

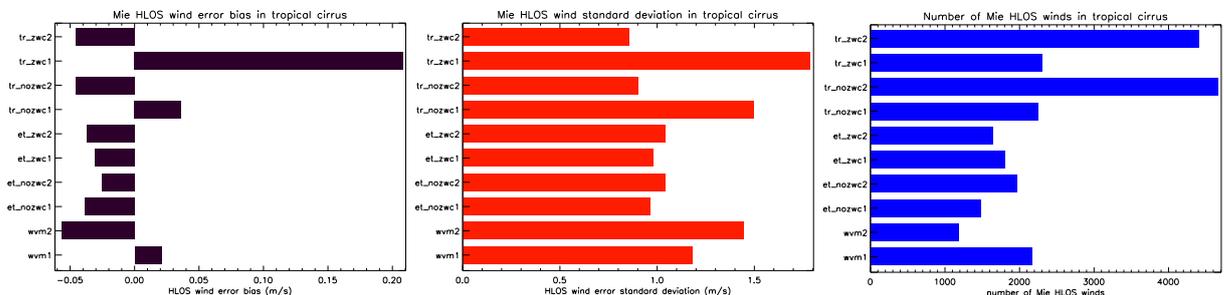


Figure 95: Quality and coverage of Mie HLOS winds in tropical cirrus scenes as a function of vertical sampling scenario: Mie HLOS wind error bias (left), standard deviation (middle) and the number of data (right).

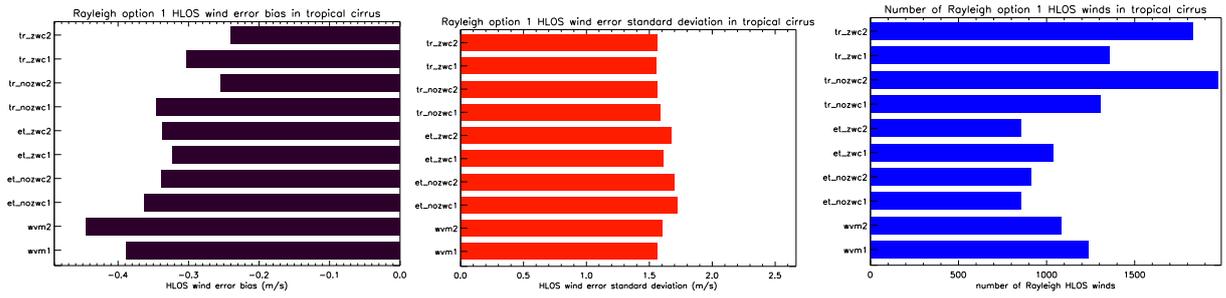


Figure 96: Quality and coverage of Rayleigh HLOS winds (processing option 1) in tropical cirrus scenes as a function of vertical sampling scenario: Rayleigh HLOS wind error bias (left), standard deviation (middle) and the number of data (right).

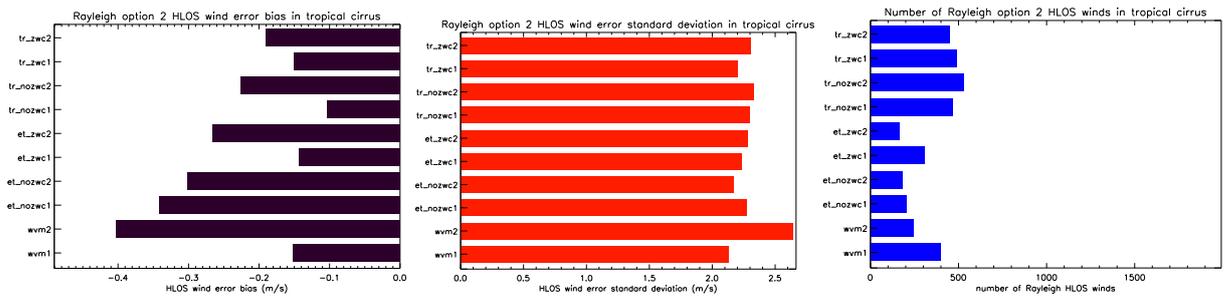


Figure 97: Same as Fig. 96 but now for Rayleigh channel processing option 2.

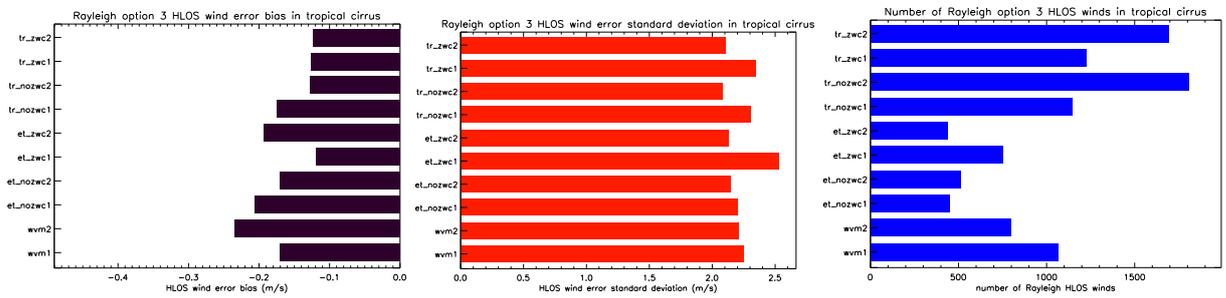


Figure 98: Same as Fig. 96 but now for Rayleigh channel processing option 3.

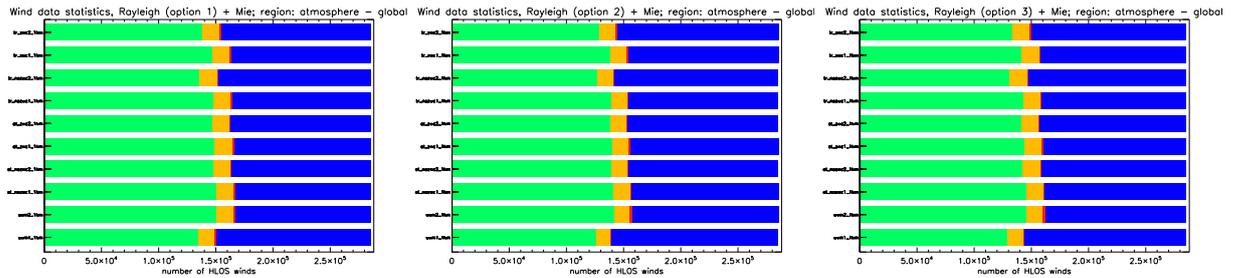


Figure 99: Same as Fig. 85 but now for all atmospheric regions accumulated. These plots show the total data coverage and HLOS wind quality throughout the atmosphere.

8.0.12 Total wind information content

Figure 99 shows the overall result of combined Mie and Rayleigh data coverage from all atmospheric regions obtained by adding the numbers in Figs. 85, 88, 91 and 94. The total number of wind observations equals 6000 (profiles) \times 24 (bins) \times 2 (channels) $\approx 2.88 \cdot 10^5$. From this figure it is concluded that about half of the retrieved HLOS winds is flagged as suspicious mainly because of lack of SNR. It is clear that processing option 1 yields most good quality winds and processing option 3 yields slightly more good quality winds than processing option 2. It is also clear that all scenarios with 500 m Rayleigh bins yield less good-quality winds than their 1000 m bins (denoted with `_1km`) counterparts. Also the increased Mie resolution in the UTLS in the tropical scenarios `tr_zwc2` and `tr_nozwc2` does not pay off in the global picture and is at the expensive of Mie winds in other height levels. However, when zooming in on the tropical region, `tr_zwc2_1km` shows more good quality winds than e.g. `wvm1_1km` (not shown). Also note that these plots do not take into account that wind information at a certain height level may be more informative for a NWP model than winds at other height levels.

Figure 100 shows the mean DWL profile information content relative to that of (synthetic) radiosondes (at the same location). The information content of all scenarios is close to 0.7 meaning that the information content of one DWL profile is on average 70% of the information content of a radiosonde. The, on average, smaller information content of a DWL profile relative to a radiosonde is mainly explained by clouds obscuring the lower part of the atmosphere for DWL but not for radiosonde. A limitation of the proposed information content calculation is that DWL measurements above 20 km yield little information relative to a radiosonde because the instrument error increases rapidly for DWL at higher altitudes, but not for radiosondes, while the representativeness error of DWL winds is only slightly smaller than for radiosondes.

Overall, Fig. 100 is in close agreement with Fig. 99. Based on these global results it is hard to make a definite selection of most informative scenarios. Each of the scenarios shows merit in different global regions. An alternative measure for information content is to use the model error as a reference rather than a radiosonde. This is discussed in [RD13].

8.1 Summary and conclusions

Quality control parameters have been defined for Mie and Rayleigh winds that are used to flag suspicious wind retrievals, e.g. in case of low SNR or large Rayleigh channel cross-talk. The Rayleigh channel bin size should be at least 1000 m for sufficient SNR.

Three processing options for Rayleigh channel wind retrieval were discussed. All three options have been implemented in LIPAS. The options differ in the strategy of accumulating measurement signals before wind retrieval. The first option ignores particle contamination and simply accumulates all measurements. In the second option classification is done before accumulating particle-free classified measurements. In the third option contaminated measurements are corrected for cross-talk before accumulation.

The bias in Rayleigh channel winds through cross-talk is generally negligible and is only substantial in case that the following three conditions apply at the same time: i) low temperatures (below -20 degrees Celsius), ii) large scattering ratio values (exceeding a value of 4) and iii) large wind speeds along the laser LOS (more than 20 ms^{-1}). As a result bias through cross-talk is generally negligible in the PBL (temperature

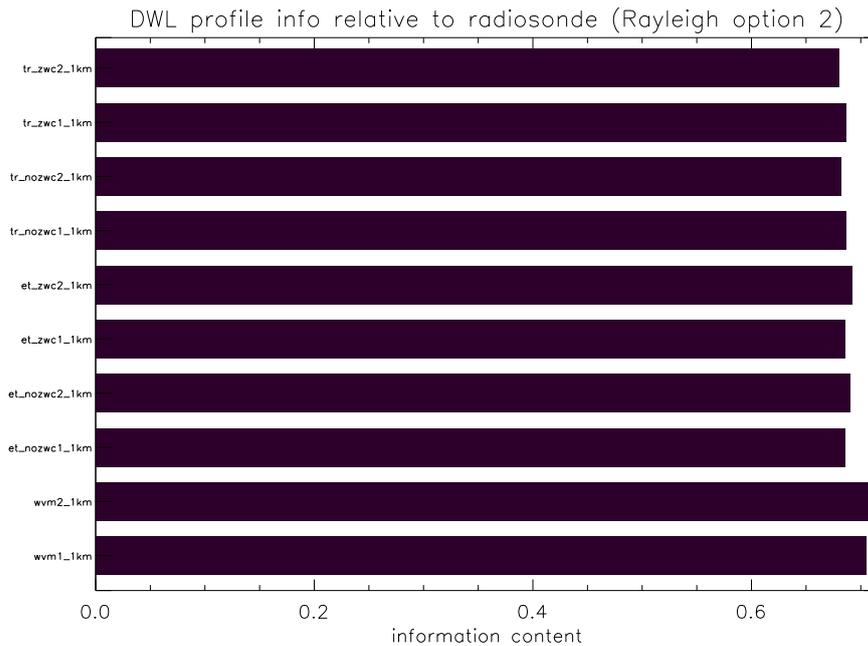


Figure 100: DWL profile information content relative to radiosonde information content. Retrieval option 2 (classification) was used here for Rayleigh channel winds.

condition), in cloud-free regions above the PBL because of generally small aerosol loading (scattering ratio condition) and in the tropical region because of generally small zonal winds (wind condition).

Rayleigh channel cross-talk correction is possible in principle if an estimate of the scattering ratio is available, e.g. from a collocated Mie bin. Cross-talk correction is at the expense of SNR that may be too low for wind retrieval in scenes with large aerosol or cloud loading. An investigation for tropical cirrus has been performed that showed that cross-talk correction reduces the biases in HLOS winds by 50%. Yet, also the HLOS wind error standard deviation increased by about 50% but is still within the mission requirement for many cases.

For tropical cirrus scenes it was shown that increasing the Mie channel resolution improves both the data coverage and quality of Mie winds in this region. Also the quality of Rayleigh channel winds after cross-talk correction was slightly improved in this region, yielding a larger data coverage.

Based on these findings we conclude that cross-talk generally does not detriment Rayleigh channel winds and causes substantial biases only in a limited number of atmospheric scenes. Cross-talk correction for these scenes is needed to reduce the wind bias and increasing the resolution of the Mie channel in these regions both improves the Mie and Rayleigh channel winds quality and coverage. The data coverage from processing option 3 (cross-talk correction) is slightly better than for option 2 (classification) in general, but substantially better for cirrus clouds. Here we note that we used prior knowledge on the cross-talk contribution that is generally not available in the real world. It remains to be investigated whether the optical properties can provide an adequate estimate of the Rayleigh channel cross-talk contribution.

It has been confirmed that scenarios defined for dedicated regions such as the tropics and extra-tropics indeed show increased performance in these regions. Scenario wvm2_1km is most informative on average when focusing on data coverage. Scenario wvm1_1km is most informative on average based on the DA information content measure, although the total number of good quality data is less than for wvm2_1km. For typical tropical conditions wvm_stratos has largest good quality data and largest DA information content although tr_nozwc2_1km has most Mie winds in the UTLS and stratosphere. For the extra-tropics there is no clear preference for either of _1km scenarios, but et_zwc2_1km and et_nozwc2_1km have most winds in the UTLS.

A Vertical sampling scenarios

This appendix provides an overview of the vertical sampling scenarios defined in [RD6] that are further studied in this document. Three sets of scenarios are defined: 1) with 500 m Rayleigh bins near the surface, 2) with 1000 m Rayleigh bins near the surface and 3) with 1000 m Rayleigh bins near the surface and a maximum bin height at or below 31 km.

A.1 Vertical sampling scenarios with 500 m Rayleigh bins near the surface

Figures 101, 102 and 103 show the scenarios with 500 m bins near the surface for the Rayleigh channel. This bin size is generally too small for the Rayleigh channel for good quality winds. For this reason a new set of scenarios was defined with at least 1000 m Rayleigh bins, see Figs. 108, 109 and 110. Although obsolete it is illustrative to consider to 500 m Rayleigh bin scenarios as well. Finally, one additional scenario with maximum stratospheric coverage is considered in Fig. 107.

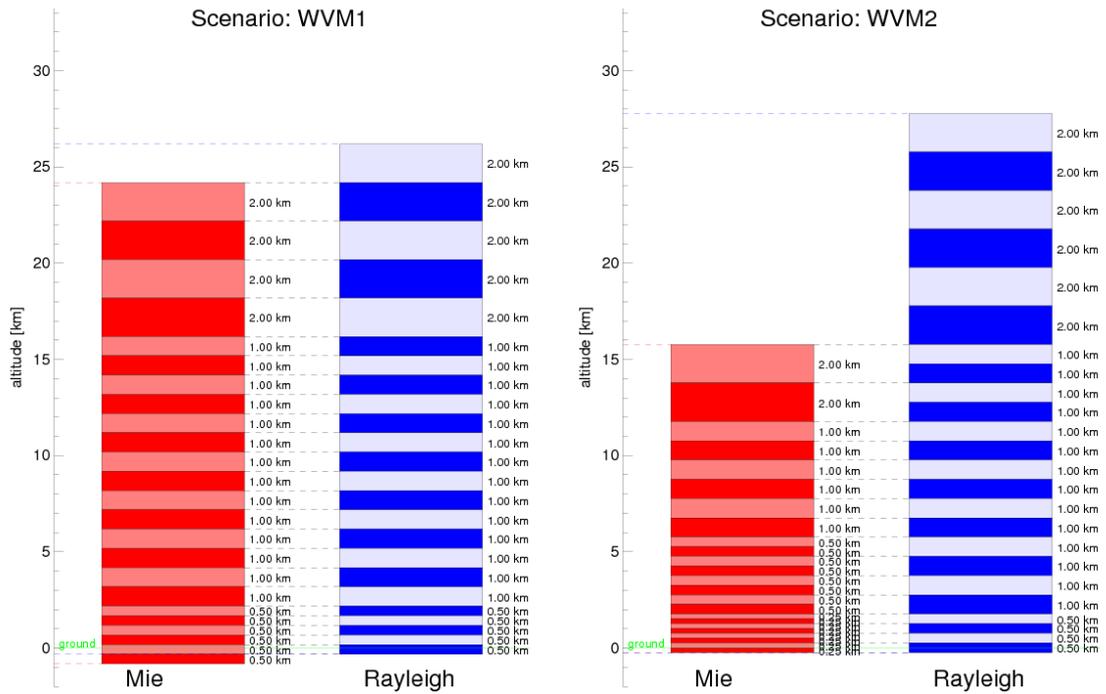


Figure 101: Vertical sampling default scenarios wvm1 (left) and wvm2 (right). The Mie channel reaches the ground for both scenarios thus providing the opportunity for Mie channel zero wind calibration.

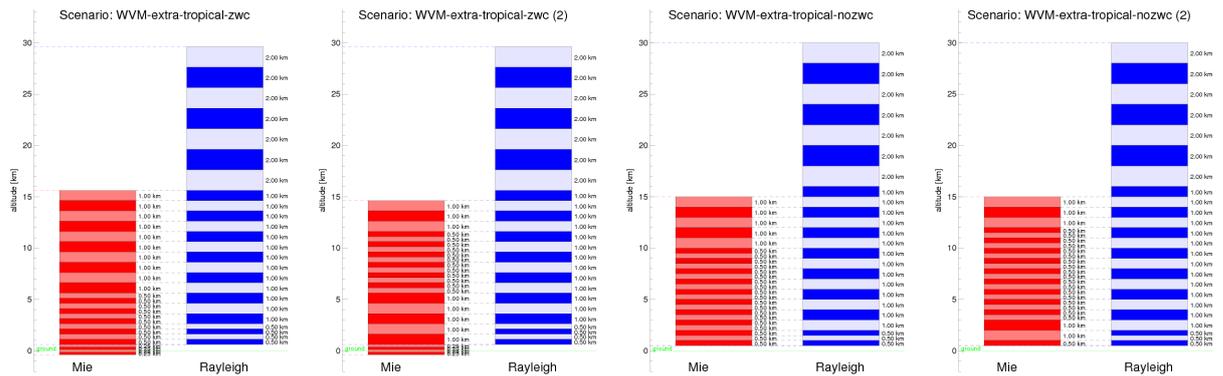


Figure 102: Extra-tropical vertical sampling with (leftmost two) and without (rightmost two) zero wind calibration opportunity. Scenario names from left to right: et_zwc1,et_zwc2,et_nozwc1,et_nozwc2.

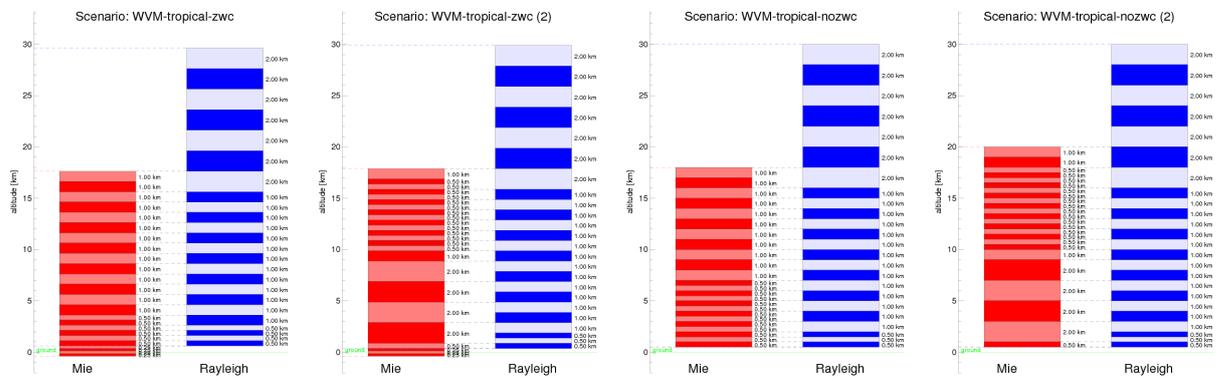


Figure 103: Tropical vertical sampling with (leftmost two) and without (rightmost two) zero wind calibration opportunity. Scenario names from left to right: tr_zwc1, tr_zwc2, tr_nozwc1, tr_nozwc2.

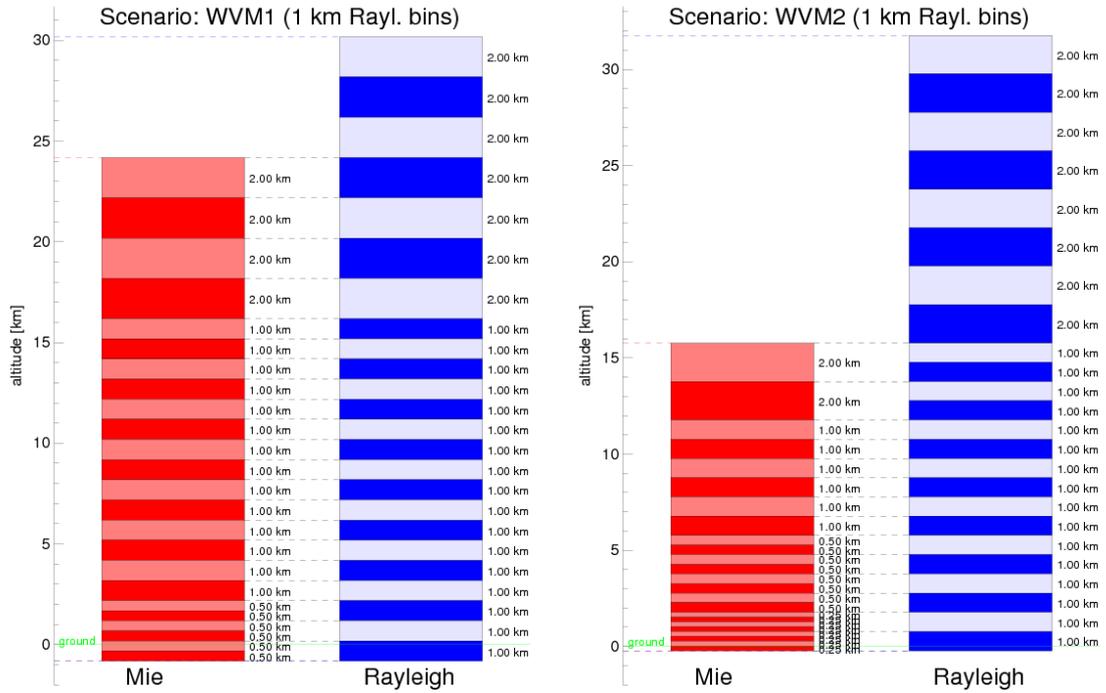


Figure 104: Same as Fig. 101 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names are wvm1_1km (left) and wvm2_1km (right).

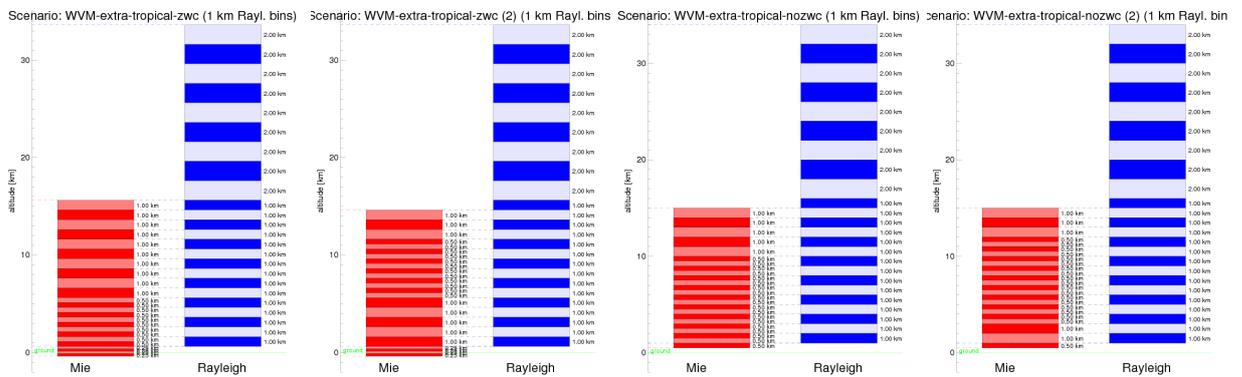


Figure 105: Same as Fig. 102 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names from left to right: et_zwc1_1km, et_zwc2_1km, et_nozwc1_1km, et_nozwc2_1km.

A.2 Vertical sampling scenarios with 1000 m Rayleigh bins near the surface

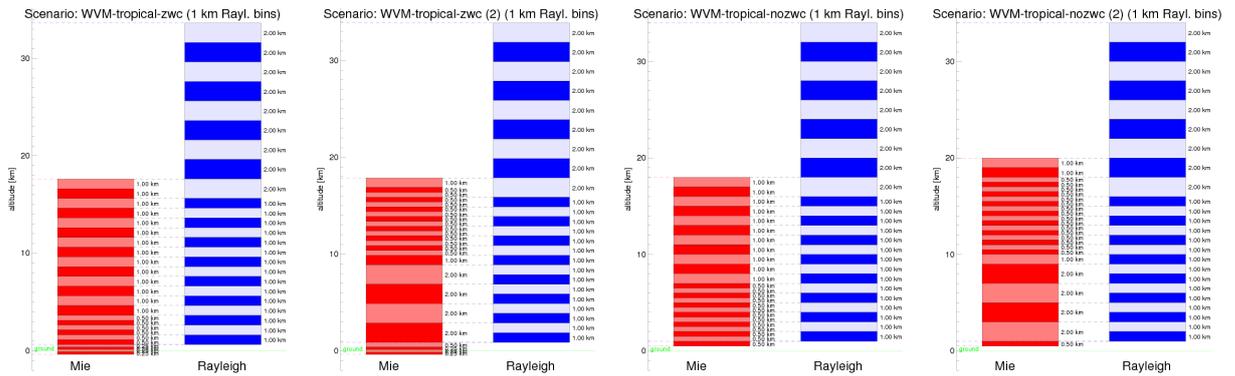


Figure 106: Same as Fig. 103 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names from left to right: tr_zwc1_1km, tr_zwc2_1km, tr_nozwc1_1km, tr_nozwc2_1km.

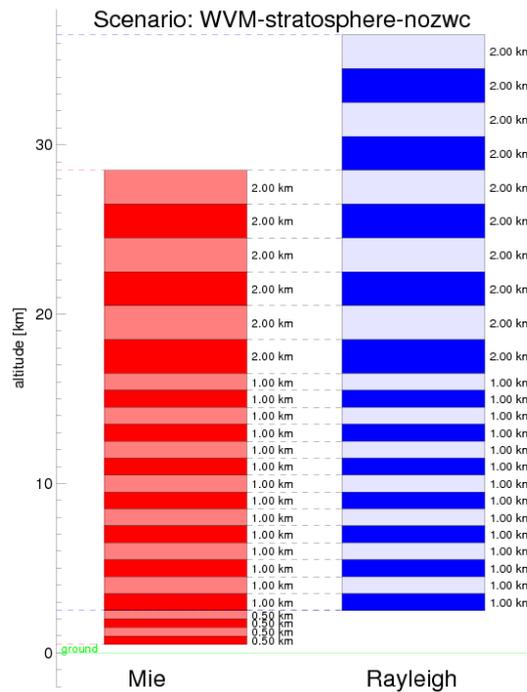


Figure 107: Stratospheric scenario with 1000 m (right) Rayleigh bins near the surface. This scenario has no zero wind calibration opportunity. The scenario is denoted wmv_stratos.

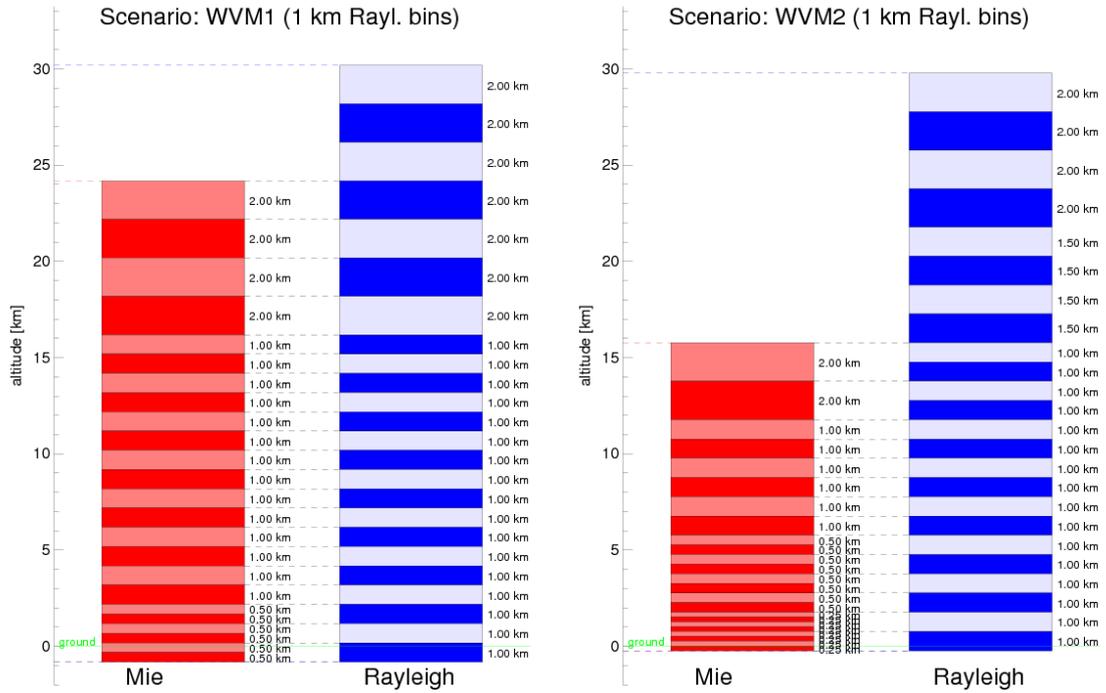


Figure 108: Same as Fig. 101 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names are wvm1_1km (left) and wvm2_1km (right).

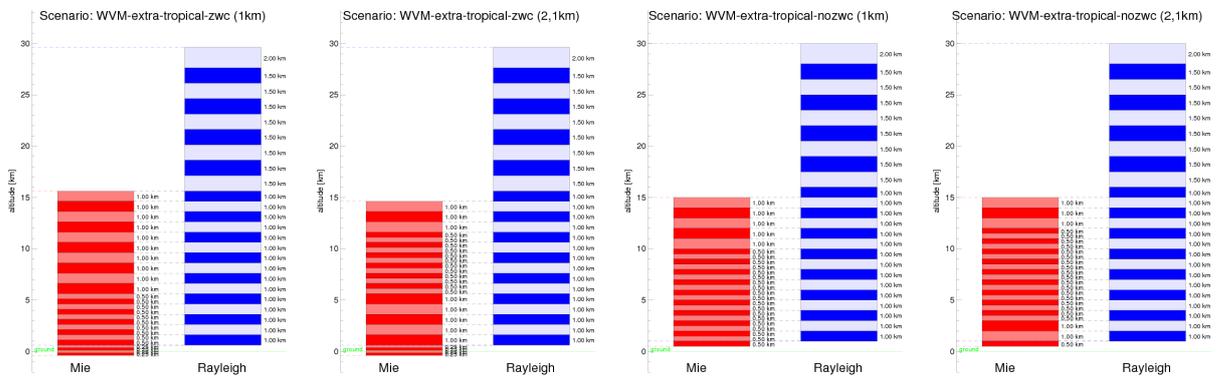


Figure 109: Same as Fig. 102 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names from left to right: et_zwc1_1km, et_zwc2_1km, et_nozwc1_1km, et_nozwc2_1km.

A.3 Vertical sampling scenarios with 1000 m Rayleigh bins near the surface and maximum bin height below 31 km

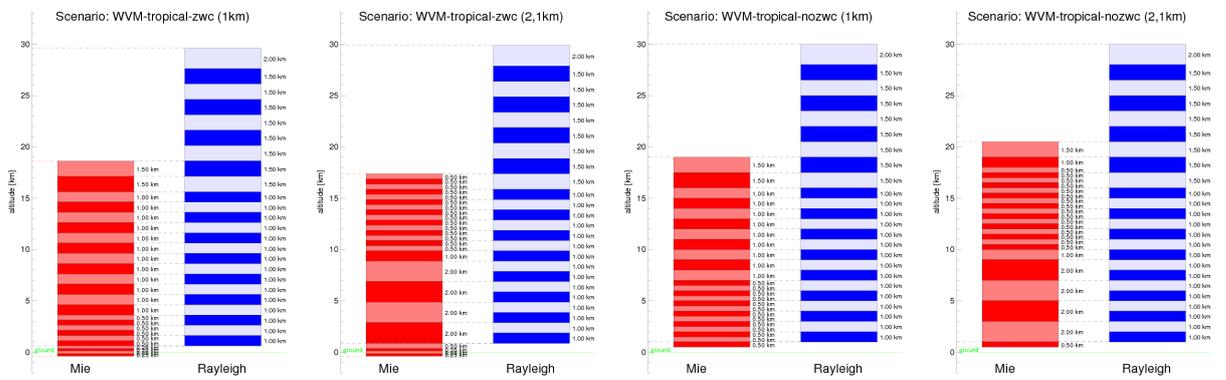


Figure 110: Same as Fig. 103 but now with 1000 m range bins for the Rayleigh channel near the surface. Scenario names from left to right: tr_zwc1_1km, tr_zwc2_1km, tr_nozwc1_1km, tr_nozwc2_1km.

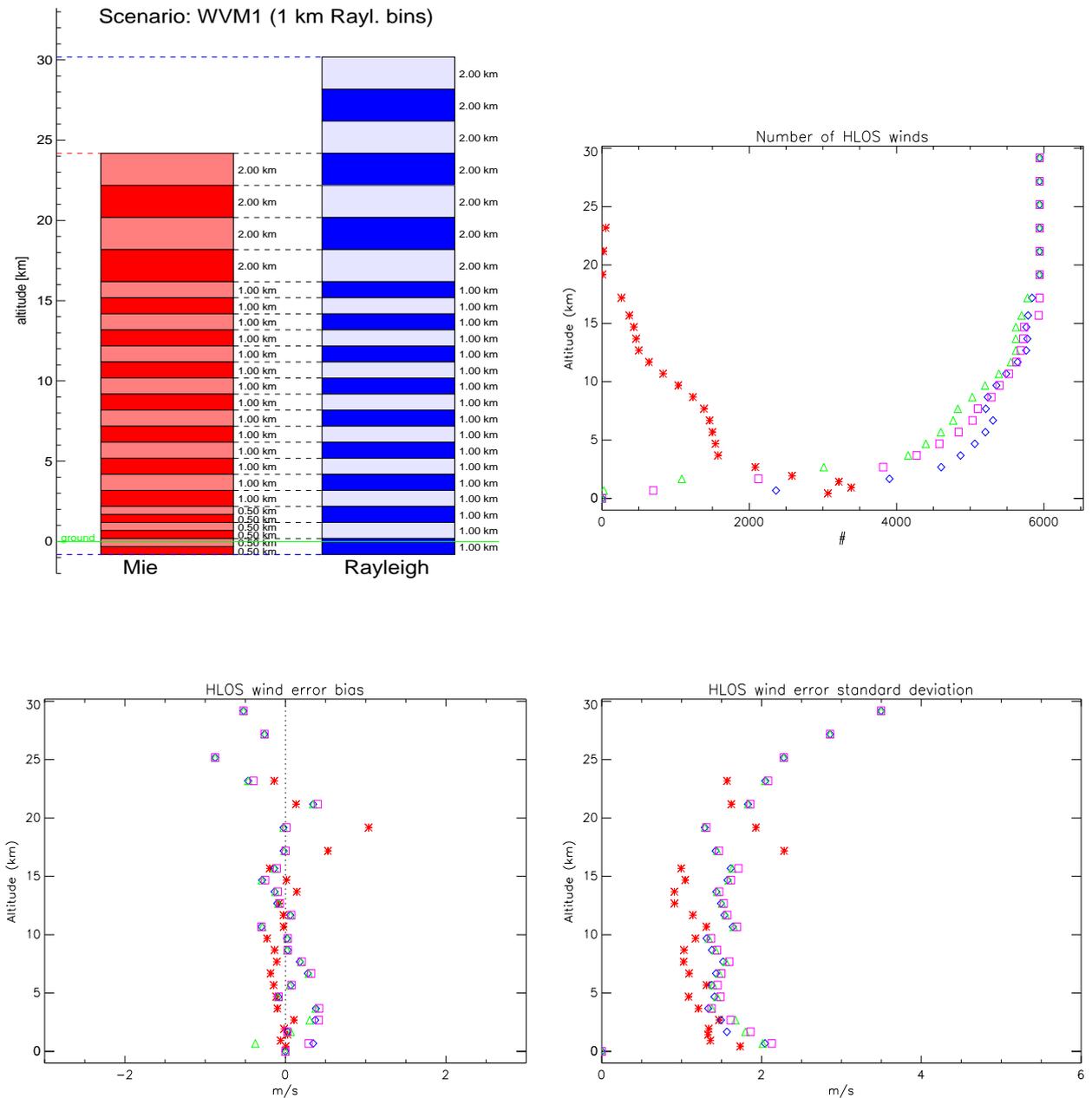


Figure 111: HLOS wind statistics for sampling scenario wvm1_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.

A HLOS wind statistics

This section provides an overview of the overall statistics of Mie and and Rayleigh winds for all sampling scenarios defined in Appendix A.3.

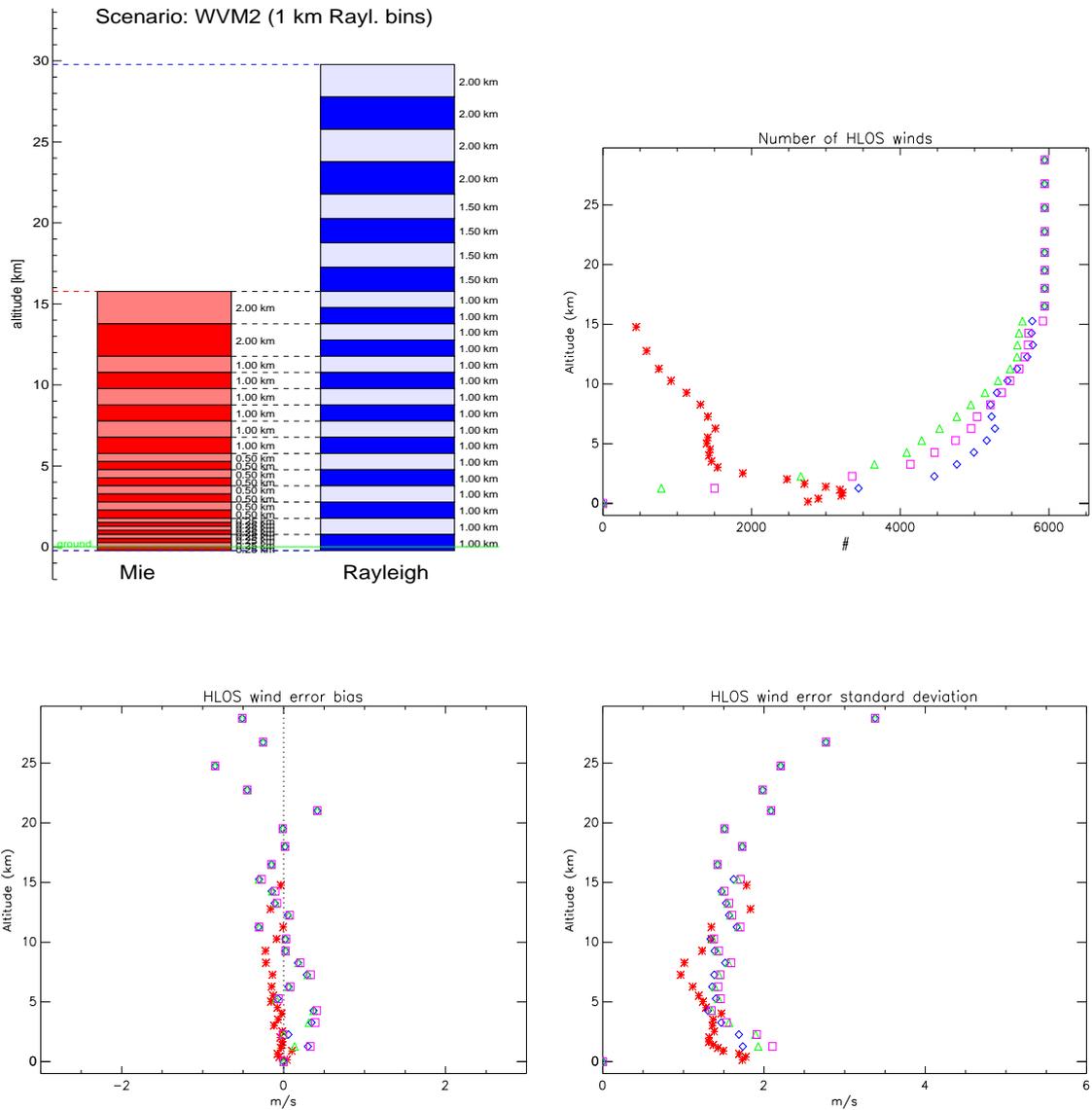


Figure 112: HLOS wind statistics for sampling scenario wvm2_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.

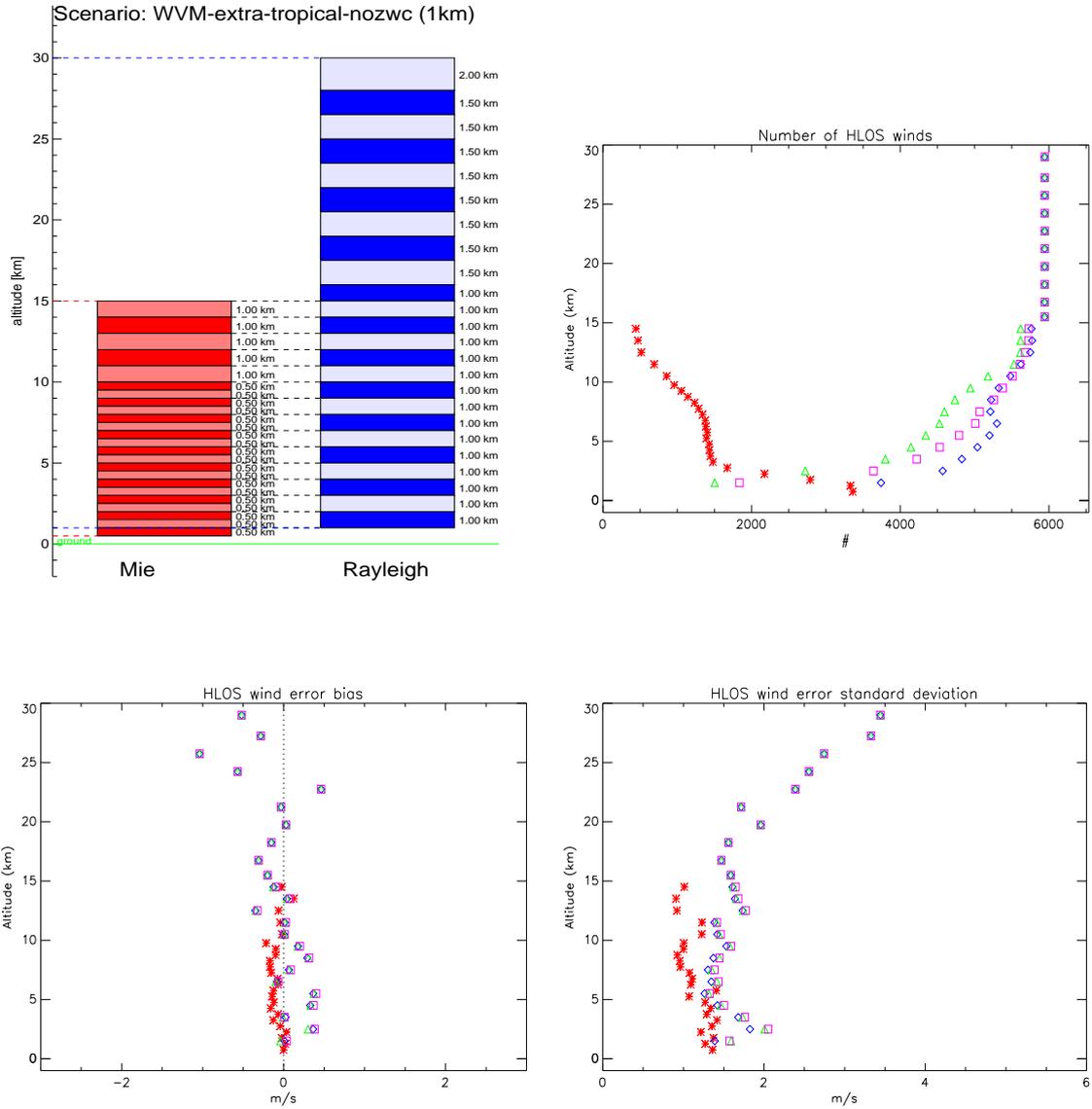


Figure 113: HLOS wind statistics for sampling scenario wvm_et_nozwc1_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.

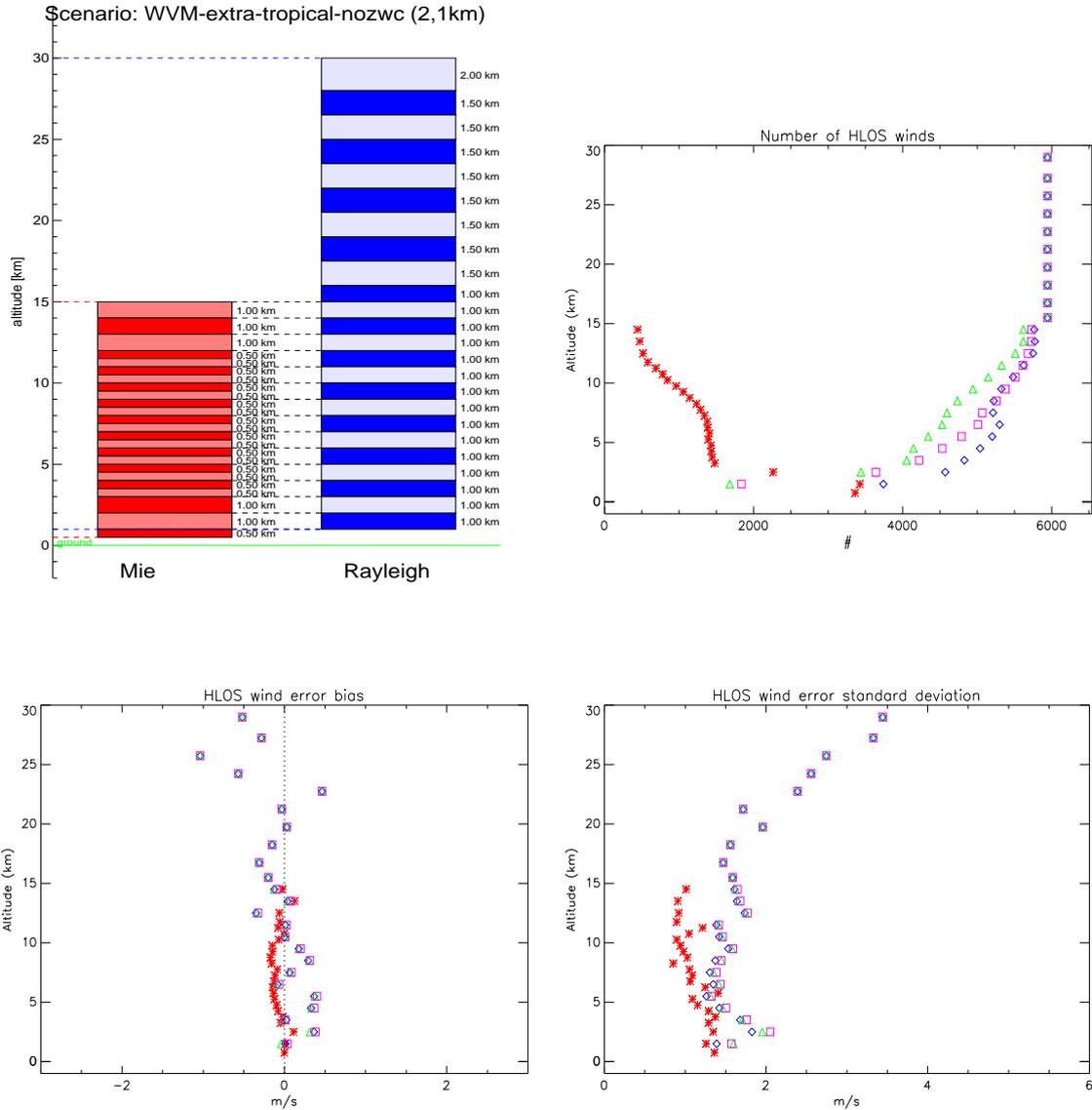


Figure 114: HLOS wind statistics for sampling scenario wvm_et_nozwc2_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.

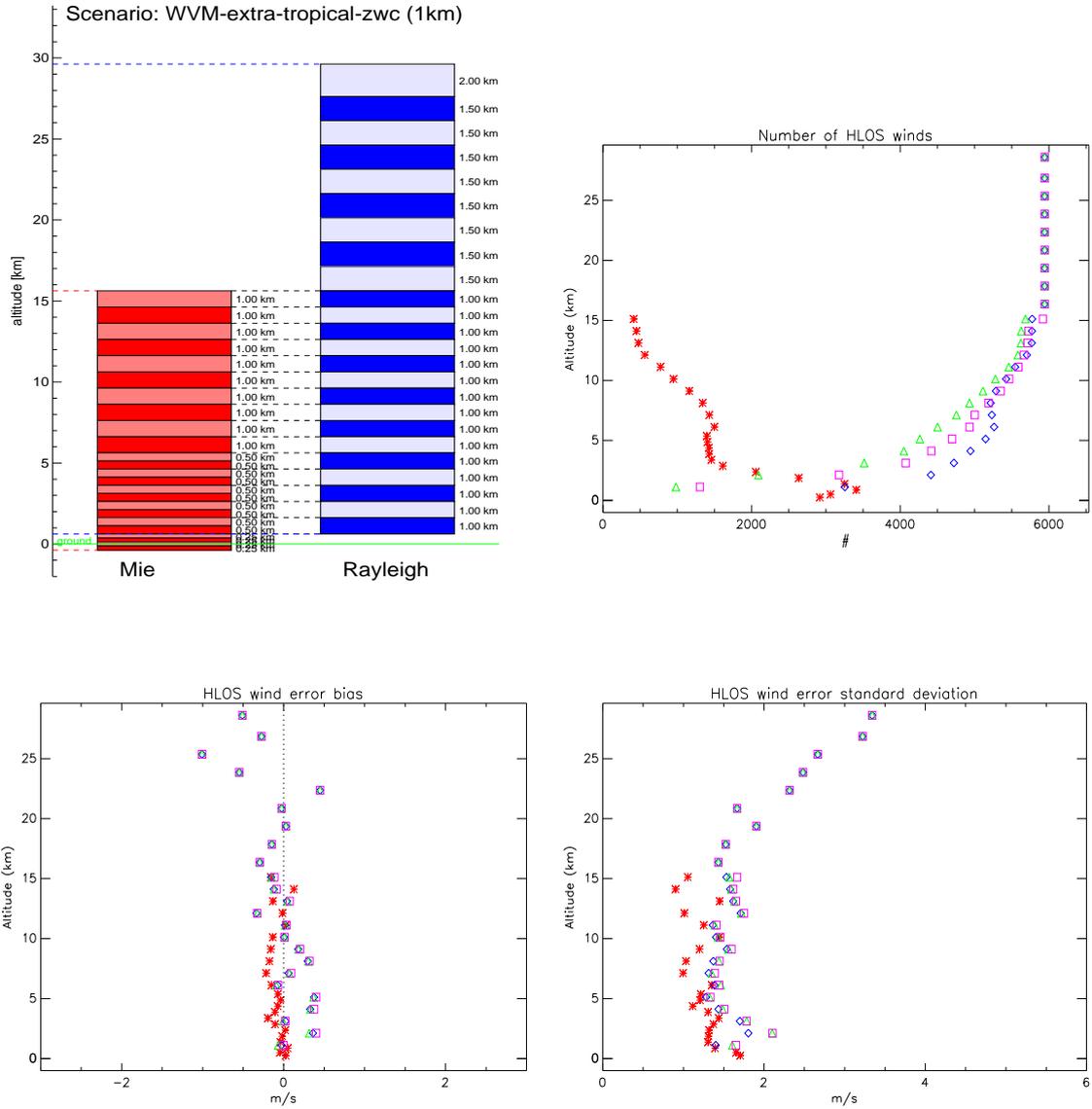


Figure 115: HLOS wind statistics for sampling scenario wvm.et.zwc1.1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.

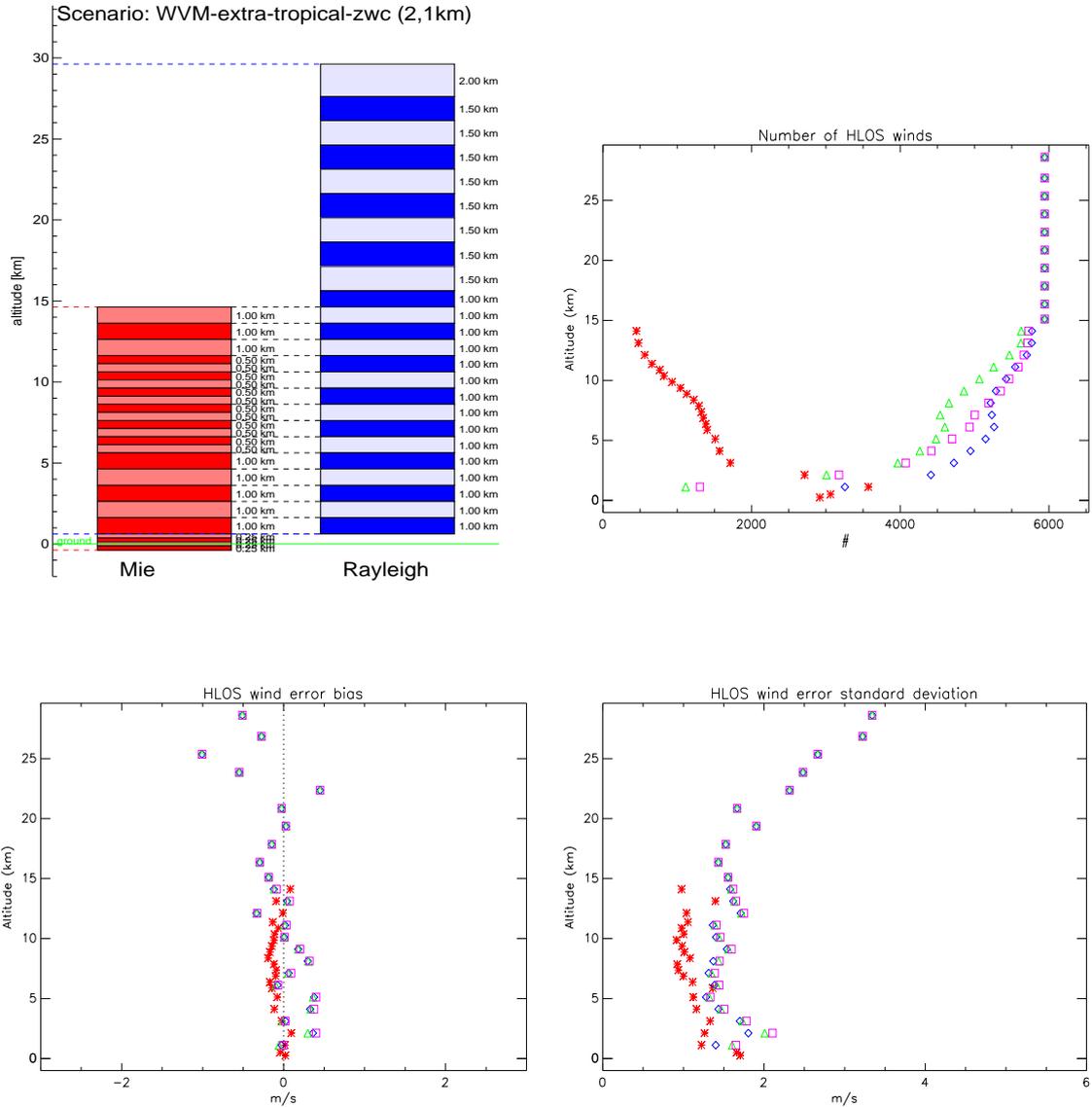


Figure 116: HLOS wind statistics for sampling scenario wvm.et.zwc2.1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.

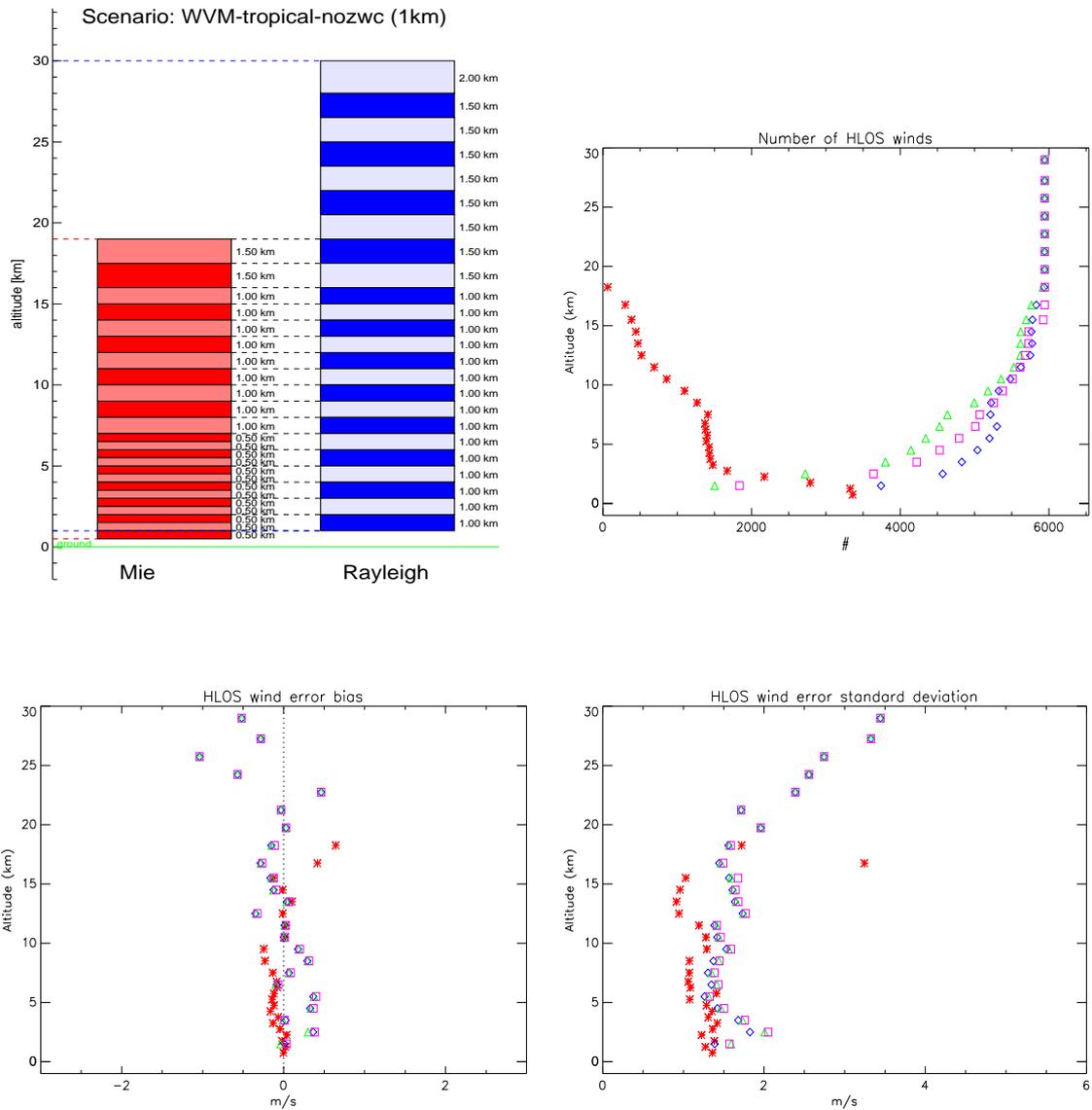


Figure 117: HLOS wind statistics for sampling scenario tr_nozwc1.1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.

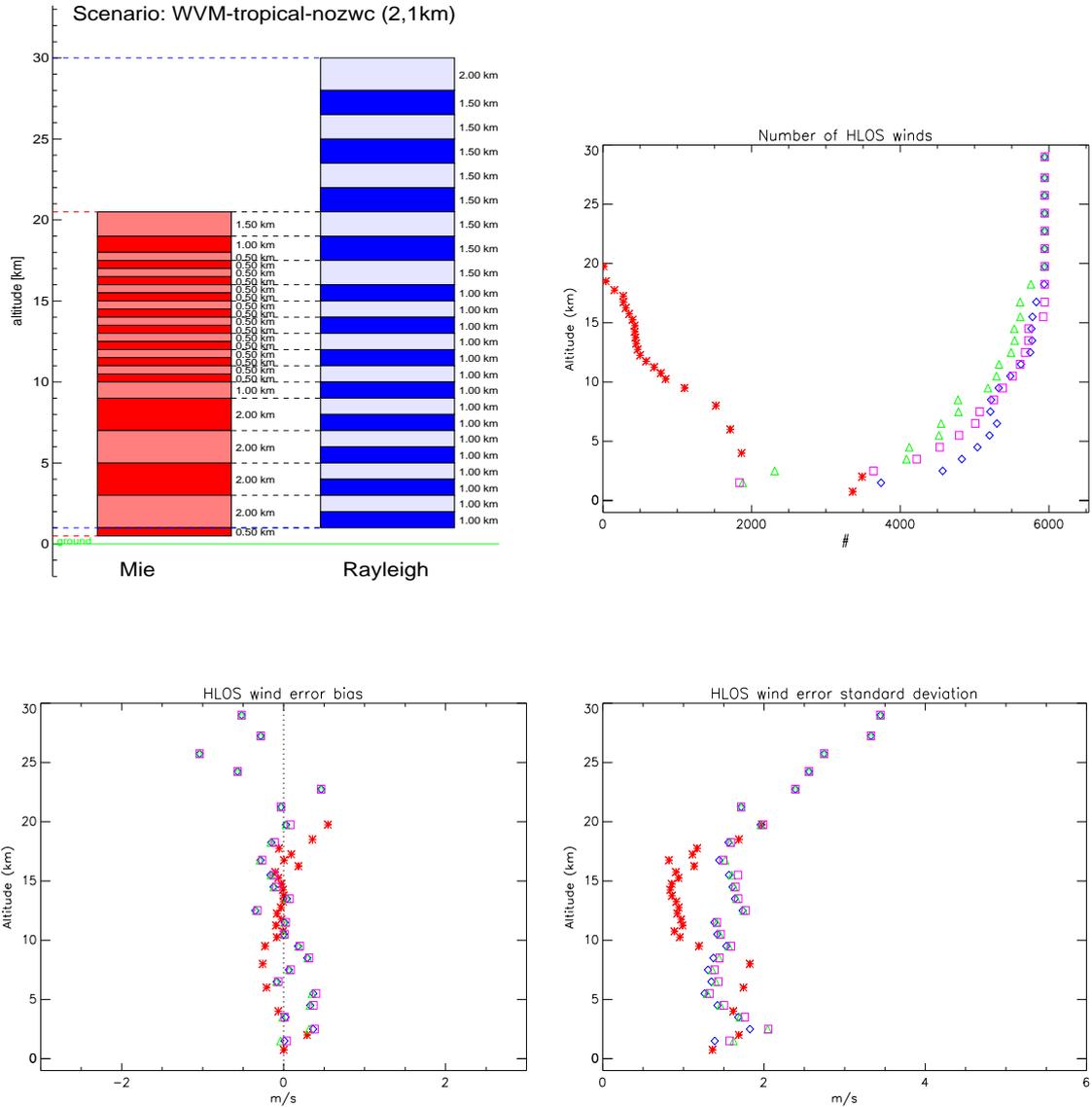


Figure 118: HLOS wind statistics for sampling scenario tr_nozwc2.1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.

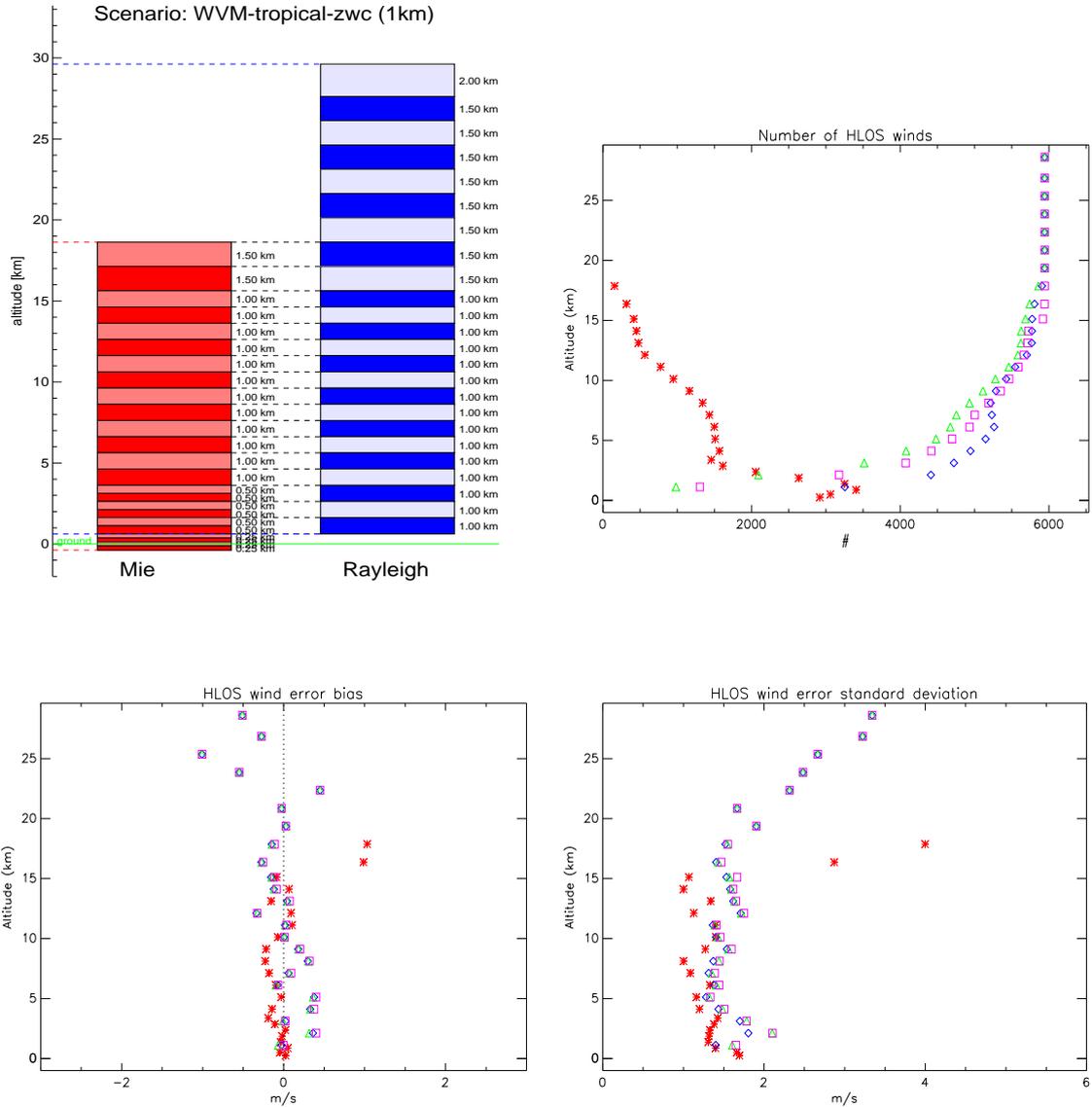


Figure 119: HLOS wind statistics for sampling scenario tr_zwc1_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.

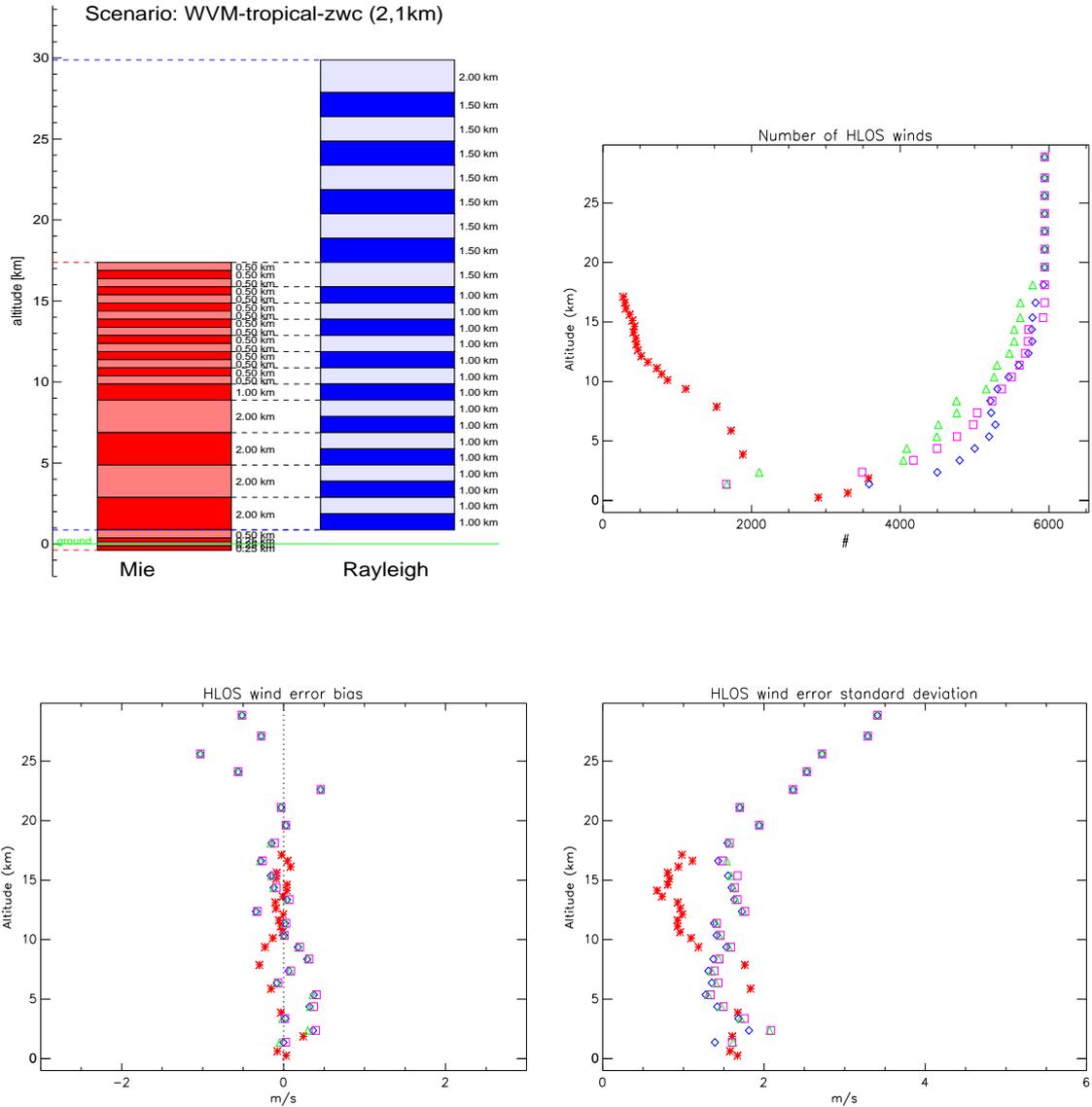


Figure 120: HLOS wind statistics for sampling scenario tr_zwc2_1km (upper left): the number of (non-flagged) winds (upper right), HLOS wind error bias (bottom left) and HLOS wind error standard deviation (bottom right). Statistics are based on 6000 wind profiles.

B Task 4c: Sampling scenarios

t.b.d.

C Conclusion

to be written.

- end of document -