

# FINAL REPORT

## ATLAS

### *ATLID Algorithms and Level 2 System Aspects*

ESTEC: Contract No. 22638/09/NL/CT

Dave Donovan  
Gerd-Jan van Zadelhoff  
Sebastien Berthier



Koninklijk Nederlands  
Meteorologisch Instituut  
Ministerie van Infrastructuur en Milieu

**Code** : EC-FR-KNMI-ATL-027  
**Issue** : 1.1  
**Date** : 26/06/2011  
**Authors** : D.P. Donovan  
G-J van Zadelhoff  
S. Berthier

---

## Table of Contents

1.	Purpose and Scope .....	9
2.	Applicable and Reference Documents .....	9
2.1.	Applicable documents .....	9
2.2.	Reference & Related documents .....	9
2.3.	Scientific References .....	10
3.	Introduction .....	12
3.1.	EarthCARE .....	12
3.2.	ATLID instrument and measurement principle .....	13
3.3.	The ATLAS project .....	17
4.	ATLAS Algorithms .....	21
4.1.	ATLID- Feature Mask (A-FM): .....	21
4.1.1.	Overview .....	21
4.1.2.	Algorithm Flowchart and description .....	23
4.1.3.	Example .....	25
4.1.4.	Algorithm Status .....	28
4.2.	ATLID-only Aerosol Extinction, Backscatter and Type (A-AER) .....	29
4.2.1.	Overview .....	29
4.2.2.	Algorithm Flowchart and description .....	31
4.2.3.	Example .....	34
4.2.4.	Algorithm Status .....	35
4.3.	ATLID Target Classification (A-TC) .....	36
4.3.1.	Overview .....	36
4.3.2.	L2a Lidar classification .....	36
4.3.2.1.	Algorithm Flowchart and description .....	38
4.3.2.2.	Example .....	40
4.3.2.3.	Algorithm Status .....	41
4.3.3.	Aerosol typing .....	42
4.3.3.1.	Algorithm Flowchart and description .....	44
4.3.3.2.	Example .....	45
4.3.3.3.	Algorithm Status .....	46
4.4.	ATLID-Extinction, Backscatter and Depolarization (A-EBD) .....	47
4.4.1.	Overview .....	47
4.4.2.	Algorithm Flowchart and description .....	49
4.4.3.	Example .....	52
4.4.4.	Algorithm Status .....	54
4.5.	ATLID- Ice Cloud Properties (A-ICE): .....	55
4.5.1.	Overview .....	55
4.5.2.	Algorithm Flowchart and description .....	56
4.5.3.	Example .....	58
4.5.4.	Algorithm Status .....	59
4.6.	ATLID-CPR-MSI Synergetic Target Classification (ACM-TC) .....	60
4.6.1.	Overview .....	60
4.6.2.	Algorithm Flowchart and description .....	61
4.6.3.	Example .....	64
4.6.4.	Algorithm Status .....	65

5.	L2 System Support (WP4).....	66
5.1.	ECSIM code maintenance.....	66
5.2.	Performance evaluation comparing ATLID-UV, ATLID-Green and CALIPSO optical depth retrievals. ....	67
5.3.	On the Use of Linear vs Circular Polarization.....	70
5.4.	Investigations into the laser-spot image.....	73
5.5.	Evaluation of Hogan’s Multiple-scattering Model. ....	76

---

## List of Tables

Table 1: Summary of the main classification rules used to combine the L2a ATLID and CPR classification input masks (Columns #2 and #3) in to the combined classification (column #1). The final column shows the additional rules needed to combine the information, where  $T_{wb}$  reflects the wet-bulb temperature,  $T$  the temperature,  $C_{Z_{max}}$  the maximum reflectivity in the warm column,  $z_{TP}$  the tropopause height and  $C_Z$  the local radar reflectivity.....63

## List of Figures

Figure 1: Schematic depiction of the HSR lidar technique.....	16
Figure 2: Schematic relationship of the ATLID-algorithm chain as defined within ATLAS. Shown are the ATLAS algorithms (yellow boxes), the relevant MSI algorithms (blue boxes) and CPR algorithm (green box). .....	19
Figure 3: Schematic relationship of the current algorithm chain as is available at the end of the ATLAS project as described in the relevant ATBDs .....	20
Figure 4: Atlid Featuremask flow diagram. The red annotations along the arrows describe the relevant data flow, the blue text the Sections in the feature mask ATBD [EC-ATBD-FM] in which more information can be found. ....	24
Figure 5: Full orbit of CALIOP night time data (CAL_LID_L1_ValStage-V3-01.2010-04-18T20-52-36ZN), the top Figure represents the raw 1064nm data, the center panel the Featuremask, and the bottom panel the VFM mask .....	26
Figure 6: Zoom of an aerosol rich region from the orbit presented in Figure 5. ....	27
Figure 7: Zoom of a region from the orbit presented in Figure 5 where both ice clouds and liquid layers are present. There is an additional aerosol region in the bottom right. ....	27
Figure 8: ATLID large-scale aerosol properties retrieval algorithm flow diagram.....	31
Figure 9: Top “true” extinction from the ESCIMO scene derived from the Falcon dataset. Bottom: Extinction as derived from the EarthCARE signal. Here a horizontal window of 20 km and a vertical window of 900 meters was used. Note that structure at finer scales is visible due to the sliding window (in both the horizontal and vertical dimensions) nature of the algorithm. ....	34
Figure 10 : Corresponding to horizontal averages of the case shown in Figure 9 this plot shows (1) The Extinction as retrieved by the DLR (Red thick continuous line), (2) the Extinction from the Falcon dataset (black thick continuous line), (3) the EarthCARE extinction (Red thick dotted line), and (3) the True extinction (black thin line). Data under the blue line should not be taken in account due to ground return. ....	35
Figure 11: Sketch of the L2a Lidar classification main steps. ....	37
Figure 12: Aerosol cloud discrimination and water cloud/ice cloud discrimination procedure. More detail associated with each step can be found in [A-TC-ATBD].....	38
Figure 13: Left: Histogram built using CALIPSO observations taken from Hu et al. 2009. The upper left region corresponds to ice clouds. Middle: Points: Results of Lidar Monte-Carlo calculations for various water clouds applied to the ATLID configuration. Solid-Line Fit to earlier MC results for CALIPSO [Hu. 2007]. Right: Overlap of the other two panels. ....	39
Figure 14: Top Left: Extinction field. Top Right Extinction-to-Backscatter field, Middle-Left: Mie channel signals. Middle-Right: Cross-polar channel signals. Bottom-Left: Feature mask output. Bottom-Right: L2a Target Classification. Here Red is Aerosol, Cyan is clear sky, Orange is ice cloud and Blue is unknown (Lidar signals are attenuated too much). ....	40
Figure 15: Lidar characteristic properties for different aerosol types from DLR field measurements (ICAROHS ATBD1) preliminary results. ....	42
Figure 16: Aerosol Typing flow diagram .....	44
Figure 17: Top panel: Observed backscatter profile from the DLR Falcon HSRL. Bottom panel: retrieved aerosol type mask. Note the different horizontal scales. ....	45

Figure 18: Schematic depiction of the horizontal averaging strategy employed within this algorithm. ....	48
Figure 19: High-level structure of the main algorithm. Here the Yellow trapezoids represent input or output data sets. ....	49
Figure 20: High level simplified schematic of the core approach of the algorithm Note that within this diagram multiple-scattering is not explicitly treated. The Equation numbers refer to those used in [A-EBD-ATBD]. ....	51
Figure 21: Top-Left: 3D view of the “Fractal Cirrus” ECSIM scene. Top-Right: “True” Extinction along the lidar track. Middle-Left: Mie channel signal Middle-Right: Extinction retrieval produced by the first-guess state-vector Bottom-Left: Retrieved best estimate and Bottom-Right: Estimated error. ....	52
Figure 22: Schematic of the A-ICE algorithm. ....	56
Figure 23: 2D-Histogram of correlation between IWC as given by the Heymsfield parametrization, against the product as given inside the DARDAR product : From left to right (1) Lidar + Radar pixels, (2) Lidar only pixels, (3) Radar only pixels. ....	58
Figure 24: Flow diagram of the ACM-TC algorithm. Yellow boxes indicate input and output files, the oval reflects the configuration parameter file and the grey boxes the steps within the algorithm. The number next to the grey boxes refer to the section numbers in the ACM-TC-ATBD in which the topic is explained. ....	61
Figure 25: ACM-Target Classification example using CloudSAT and Calipso data measured on 2010-04-19. The top figure shows the result of the ACM-TC, the second is the DARDAR mask, third the CloudSAT radar data (the read line indicates the $T_{wb}=0^{\circ}\text{C}$ isotherm) and the bottom figure the 532nm CALIOP data. This scene shows a large number of different regions (light blue: ice clouds, green: rain, orange: liquid layer, red: super-cooled liquid, green within the blue: super-cooled liquid and ice and darker green: aerosols). Note that the two classification show very similar results ....	64
Figure 26: Example CASE 1 results for the UV HSRL concept. The green lines show CALIPSO lidar simulation results. The Red line in the second panel from left show the Rayleigh return in the absence of cloud. The Blue lines in the first two panels denote the cross-talk corrected signals and the black-lines show the signals before cross-talk correction. ....	68
Figure 27: Example CASE 1 results for the Green HSRL concept. The green lines show CALIPSO lidar simulation results. The Red line in the second panel from left show the Rayleigh return in the absence of cloud. The Blue lines in the first two panels denote the cross-talk corrected signals and the black-lines show the signals before cross-talk correction. ....	68
Figure 28: Statistical Error in cirrus OT retrievals for the UV HSRL concept. Here $iwin1$ denotes the vertical averaging interval above the layer while $iwin2$ denotes the below layer interval. ....	69
Figure 29: Left panel except for the Green HSRL concept. ....	69
Figure 30: As above except for Daylight conditions. ....	69
Figure 31: As above except for daylight conditions. ....	69
Figure 32: Linear (Grey) and circular (Black) depolarization ratios for a water cloud 2 km from the lidar. The dotted lines show theoretical results presented in Fig 19 of Roy and Roy (2008), while the solid lines show ECSIM Monte-Carlo results for a roughly comparable situation. ....	71
Figure 33: Mie and Rayleigh returns for an ice cloud. Black, circular depolarization. Grey linear depolarization. ....	72

Figure 34: Ice cloud depolarization ratios. Black, circular depolarization. Grey linear depolarization.....	72
Figure 35: Mie and Rayleigh returns for an water cloud. Black, circular depolarization. Grey linear depolarization.....	72
Figure 36: Water cloud depolarization ratios. Black, circular depolarization. Grey linear depolarization.....	72
Figure 37: (Top)_ laser spot image as a function of angle for different homogeneous water cloud optical depths (from 0.05 to 10.0 left to right). Here linear polarization was transmitted and the cross-polar intensity measured. An ocean surface was specified. (Bottom) Average relative image intensity as a function of angle. The dashed line shows the angular envelope of the transmitted (Gaussian) laser beam. The Max value labels in the lower panels show the value of the peak in 1/sr. ....	74
Figure 38: As Figure 35 except that the co-polar signal is measured (linear co-polar). .....	74
Figure 39: As previous except that circular polarization has been used (circular co-polar). ....	75
Figure 40: As Figure 35 except that a snow surface has been used.....	75
Figure 41: (Left) Mie co-polar and (Right) Rayleigh channel co-polar returns for a water cloud of OT 1.0 and an effective radius of 5 microns. Black solid: ECSIM results. Dashed Black: Single scattering results. Solid Grey: Hogan's model results for the true value of Ra. Dashed Grey from left to right: Hogan's model results for Ra=2.5, 5.0, 10, 20 and 25 microns respectively .....	77
Figure 42: As Figure 41 except $R_{\text{eff}}=19.8$ .....	77
Figure 43: As Figure 41 except OT=5.0 .....	77
Figure 44: As Figure 42 except OT=5.0 .....	77
Figure 45: As Figure 43 but for OT=1.0 cirrus clouds. ....	78
Figure 46: As Left figure but for larger particle sizes. ....	78
Figure 47: As Figure 45 but for OT=2.5 cirrus clouds. ....	78
Figure 48: As Left figure but for larger particle sizes. ....	78

---

## 1. Purpose and Scope

This document summarises the work carried out under the ATLAS project. The main objective of ATLAS was to develop suite of (mainly) lidar based algorithms for cloud and aerosol retrievals specifically for the EarthCARE lidar.

The structure of this document is as follows. The following Introduction section supplies a brief background to the EarthCARE mission and the ATmospheric LIDar (ATLID) instrument. The subsequent sections present a summary of the Algorithm developed during the course of the project. An additional section describes work carried out during the project that is not connected with any one specific algorithm but instead falls into the category of “L2 System Support”.

## 2. Applicable and Reference Documents

### 2.1. *Applicable documents*

<i>Reference</i>	<i>Code</i>	<i>Title</i>	<i>Issue</i>	<i>Date</i>
[MRD]	EC-RS-ESA-SY-012	EarthCARE System Requirements Document	5	Nov 2 2006

### 2.2. *Reference & Related documents*

<i>Reference</i>	<i>Code</i>	<i>Title</i>	<i>Issue</i>	<i>Date</i>
EarthCARE	EC-ICD-ESA-SYS-0314	EarthCARE product Table	1.3	15/06/2010
ATLAS-PARD	EC-TN-KNMI-ATL-005	ATLAS Products and Algorithms Requirements Document (PARD)	1.1	10/03/2010
A-AER-ATBD	EC-TN-KNMI-ATL-ATBD-A-AER-019	ATLID L2a Aerosol Extinction, Backscatter and Depolarization	2.2	27/06/2011
A-FM-ATBD	EC-TN-KNMI-ATL-ATBD-A-FM-010	ATLID L2a Feature Mask ATBD	2.2	26/05/2011
A-TC-ATBD	EC-TN-KNMI-ATL-ATBD-A-TC-022	ATLID L2a Target Classification ATBD	2.2	27/06/2011
A-EBD-ATBD	EC-TN-KNMI-ATL-ATBD-A-	ATLID L2a High Resolution Extinction, Backscatter and Depolarization	1.3	09/06/2011

<i>Reference</i>	<i>Code</i>	<i>Title</i>	<i>Issue</i>	<i>Date</i>
	EBD-021			
A-ICE-ATBD	EC-TN-KNMI-ATL-ATBD-A-ICE-023	ATLID L2a empirical IWC ATBD	1.3	27/06/2011
A-ACM-ATBD	EC-TN-KNMI-ATL-ATBD-ACM-TC-024	L2b ATLID+MSI+CPR Classification ATBD	2.1	26/05/2011
ECSIM	ECSIM-KNMI TEC-MAD01-R	ECSIM Models and Algorithm Document	1.2.4	05/11/2010
CASPER-FR	CASPER-DMS-FR-01	Cloud and Aerosol, Synergetic Products from, EarthCARE Retrievals, (CASPER), Final Report	1.1	30/01/2009

### 2.3. *Scientific References*

<b>Keyword</b>	<b>Reference</b>
Ansmann et al. 2011	Saharan Mineral Dust Experiments SAMUM-1 and SAMUM-2: What have we learned? Tellus B., submitted
Avery 2010	Avery, M., D. Winker, M. Vaughan, S. Young, R. Kuehn, Y. Hu, J. Tackett, B. Getzewich, Z. Liu, A. Omar, K. Powell, C. Trepte, and K.-P. Lee, (2010), "A first look at CALIOP/CALIPSO cloud ice water content", submitted to the 25th International Laser and Radar Conference.
Delanoë and Hogan., 2010	Delanoë, J., and R. J. Hogan (2010), Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds, <i>J. Geophys. Res.</i> , 115, D00H29, doi:10.1029/2009JD012346.
Ferrare 2007	Ferrare, R. et al., "Airborne high spectral resolution lidar aerosol measurements and comparisons with transport models", <i>Eos Trans. AGU</i> , 88(52), Fall Meet. Suppl., Abstract A14D-06, (2007).
Foot 1988	Foot, J. S. (1988), Some observations of the optical properties of clouds. Part II: Cirrus, <i>Q. J. R. Meteorol. Soc.</i> , 114, 145 – 164 .
Heintzenberg 2009	The SAMUM-1 experiment over Southern Morocco: Overview and introduction. <i>Tellus</i> 61B, 2-11, DOI: 10.1111/j.1600-0889.2008.00403.x

Keyword	Reference
Heymsfield et al., 2005	Heymsfield, A. J., D. Winker, and G.-J. van Zadelhoff (2005), Extinction-ice water content-effective radius algorithms for CALIPSO, <i>Geophys. Res. Lett.</i> , 32, L10807, doi:10.1029/2005GL022742.
Hogan and O'Conner 2004	Facilitating cloud radar and lidar algorithms: the Cloudnet Instrument Synergy/Target Categorization product Hogan, R. J., and E. J. O'Connor, CloudNET project document, 2004
Hogan 2006	Fast, approximate calculation of multiply scattered lidar returns, <i>Applied Optics</i> , 45, 5984-5992 (2006).
Hogan and Kew 2005	Hogan, R. J., and S. F. Kew: A 3D stochastic cloud model for investigating the radiative properties of inhomogeneous cirrus clouds. <i>Q. J. R. Meteorol. Soc.</i> , 131, 2585-2608 (2005)
Hu. 2007	Hu, Y., 2007: Depolarization ratio–effective lidar ratio relation: Theoretical basis for space lidar cloud phase discrimination. <i>Geophys. Res. Lett.</i> , 34, L11812, doi:10.1029/2007GL029584.
Hu. et al. 2009	Hu Y., D. Winker, M. Vaughan, B. Lin, A. Omar, C. Trepte, D. Flittner, P. Yang, S.L. Nasiri, B. Baum, W. Sun, Z. Liu, Z. Wang, S. Young, K. Stamnes, J. Huang, R. Kuehn and R. Holz, "CALIPSO/CALIOP Cloud Phase Discrimination Algorithm", <i>JTECH</i> , DOI: 10.1175/2009JTECHA1280.1
IPCC 2007	IPCC (2007). "IPCC Fourth Assessment Report: Climate Change 2007 (AR4)". Cambridge, United Kingdom and New York, NY, USA.: Cambridge University Press. Retrieved 2011-06-14
Kittaka, C et al. 2008	Kittaka, C.; Winker, D.; Omar, A.; Liu, Z.; Vaughan, M.; Trepte, C., Global Aerosol Distributions Derived From the CALIPSO Observations, American Geophysical Union, Fall Meeting 2008, abstract #A41A-0078
Klett 1985	Klett, J.D. Lidar inversion with variable backscatter/extinction ratios <i>Applied Optics</i> , Vol. 24, Issue 11, pp. 1638-1643 (1985) doi:10.1364/AO.24.001638
Muller, et al, 2011	ICAROHS final report; Recommendations for future single and multi-wavelength HSRL instruments
Omar et al. 2009	Omar, Ali H., and Coauthors, 2009: The calipso automated aerosol classification and lidar ratio selection algorithm. <i>J. Atmos. Oceanic Technol.</i> , 26, 1994–2014. doi: 10.1175/2009JTECHA1231.1
Platt 1973	Platt, C. M. R., 1973: Lidar and radiometric observations of cirrus clouds, <i>J. Atmos. Sci.</i> , 30, 1191–1204.
Rodgers 2000	Rodgers, C.D., "Inverse methods for Atmospheric Sounding; Theory and Practice", Series on Atmospheric, Oceanic and Planetary Physics-Vol. 2, World Scientific, ISBN 981-02-

Keyword	Reference
	2740-X, 2000.
Russ 2007	Russ, J.C., 2007, The Image Processing Handbook, 5 <sup>th</sup> Ed., Published by CRC Press, ISBN 0849372542, 9780849372544, ch. 4, correcting imaging defects. 214-224.
Shiplely 1983	Shiplely, S.T., D.H. Tracy, E.W. Eloranta, J.T. Trauger, J.T. Sroga, F.L. Roesler and J.A. Weinman, "A High Spectral Resolution Lidar to Measure Optical Scattering Properties of Atmospheric Aerosols, Part I: Instrumentation and Theory", Applied Optics, 23, 3716-3724, 1983
Skilling and Gull 1985	Skilling J.S., and S.F. Gull, 1985, Algorithms and Applications; Eds. C. Ray Smith and W.T. Grandy, Jr. Maximum-Entropy and Bayesian Methods in Inverse Problems, 83-132, 1985 by D. Reidel Publishing Company .
Vaughan 2009	Vaughan, Mark A., and Coauthors, 2009: Fully Automated Detection of Cloud and Aerosol Layers in the CALIPSO Lidar Measurements. J. Atmos. Oceanic Technol., 26, 2034-2050.
Weinzierl and van Zadelhoff 2011	Weinzierl, B. and van Zadelhoff, G-J, Technical Note 2: ICAROHS project, 2011

### 3. Introduction

#### 3.1. *EarthCARE*

The interactions between clouds, aerosols and solar and terrestrial radiation play key roles in the Earth's Climate. As stated by the most recent assessment report of the Intergovernmental Panel on Climate Change [IPCC 2007], clouds and aerosols are two of the largest causes of uncertainty in climate prediction: aerosols because of their uncertain direct and indirect radiative forcing of climate, and clouds because the way they will respond in response to a warmer world is one of the most poorly understood climate feedbacks.

The Earth Clouds, Aerosol and Radiation Explorer (EarthCARE) is a combined ESA/JAXA mission to be flown in 2015. EarthCARE will study the spatial (3D) distribution of clouds and aerosols and their impact on the Earth's radiative balance. To do this, the EarthCARE platform will carry a combination of active and passive sensors. The instruments that will be carried on-board EarthCARE are:

- a 94 GHz, Doppler Cloud Radar supplied by Japan (CPR)
- an advanced 353 nm High-Spectral resolution Lidar (ATLID)
- a multispectral cloud/aerosol imager measuring narrow-band TOA radiances (MSI)
- a 3-view Broad-Band Long- and Short-Wave Radiometer for TOA radiance (BBR)

An important goal of the EarthCARE mission is to use the retrieved cloud and aerosol properties along the vertical atmospheric slice defined by the active sensors in conjunction with the multi-spectral imager (MSI) to construct  $(10 \text{ km})^2$  3D retrieved atmospheric scenes at a resolution of around 500 m in both the vertical and horizontal. The retrieved properties of the  $(10 \text{ km})^2$  3D atmospheric boxes should be such that the top of atmosphere (TOA) radiative fluxes for each  $(10 \text{ km})^2$  broad-band radiometer (BBR) cell surrounding the cross-section can be estimated to within  $10 \text{ W m}^2$ . This will enable EarthCARE to deliver data that has been self-checked for consistency in a radiative sense.

In order to achieve this goal, a system of individual instrument geophysical algorithms (L2a) and (L2b) synergetic (e.g. multi-instrument algorithms) must be developed. It should also be recognized that even the L2a (single-instrument) algorithms must fit into the overall (fundamentally synergetic) overall retrieval process. One consequence of this is that the even the L2a algorithms must report their outputs in a manner that is compatible with the idea of a “Joint Standard Grid” or LID synergetic grid. This standard grid is a construction which greatly facilitates the synergetic use of the EarthCARE data.

ATLAS has focused on aspects mainly concerning the lidar instrument. However, the structure of the ATLAS processing chain and the algorithms themselves reflect the synergetic nature of the overall mission. An overview of the ATLAS processing chain and the component algorithms will be given later in Section. However, first some background concerning the ATLID instrument will be presented.

### **3.2. *ATLID instrument and measurement principle.***

ATLID is an example of a so-called High Spectral Resolution Lidar. ATLID emits a linearly polarized spectrally narrow beam (FWHM for ATLID is around 30 MHz) and there the return is detected by 3 different receiver channels. First the return is separated using a polarized beam splitter into components whose linear polarization states are either the same (co-polar) or perpendicular (cross-polar) to the transmitted beams plane of polarization. Then the co-polar channel is spectrally (imperfectly) separated into 2 channels. The so-called Mie (or more properly the elastic backscatter) channel is dominated by light scattered by aerosol or cloud particles while the Rayleigh (or more properly the inelastic return) channel is dominated by the thermally broadened return scattered from the molecules that make up the air itself. At typical atmospheric pressures and temperatures the Rayleigh return will have a FWHM on the order of 4000 MHz, which is substantially greater than the ATLID laser line width.

The process of separating the Mie and Rayleigh returns is depicted schematically in Figure 1. For ATLID the spectral separation is accomplished using an Fabry-Pérot (FP) etalon leading to an imperfect separation of the Rayleigh and Mie signals where about 16% of the Mie signal goes into the Rayleigh channel and about 30% of the

---

Rayleigh return goes into the Mie channel. Thus, a so-called cross-talk correction procedure is implemented as part of the L1 processing so that the L2 algorithms may be supplied with the “true” Rayleigh and Mie signal profiles.

The reason for separating the Rayleigh and Mie returns (the essence of the HSRL technique) may be appreciated by first considering the case of a single-wavelength elastic-backscatter lidar. Neglecting multiple-scattering processes the power received by the lidar can be written as:

$$P_{ss}(z) = \frac{C_{lid}}{r^2(z)} [\beta_M(z) + \beta_R(z)] \exp \left[ -2 \int_{z_{lid}}^z (\alpha_M(z') + \alpha_R(z')) dr' \right] \quad (1)$$

where  $C_{lid}$  is the lidar constant,  $r$  is the range from the lidar,  $z$  is the altitude,  $\beta$  is the backscatter coefficient  $\alpha$  is the extinction coefficient and the  $M$  and  $R$  subscripts are used to distinguish between Mie (cloud/aerosol) and Rayleigh scattering. Assuming that the lidar constant is known and noting that  $\alpha_R$  and  $\beta_R$  are known functions of the (presumably) known atmospheric density profile then one is still left where there are two unknowns ( $\alpha_M$  and  $\beta_M$ ) but only one measured quantity ( $P_{ss}$ ). Thus, an a priori relationship between  $\alpha_M$  and  $\beta_M$  must be imposed in order to invert Eqn. (1) so as to recover the extinction and backscatter profiles.

If, on the other hand, separate Mie and Rayleigh signals are available then, in principle, the extinction and backscatter profiles can be independently retrieved in a direct manner. The Mie signal can be written as

$$P_M(z) = \frac{C_M}{r^2(z)} \beta_M(z) \exp \left[ -2 \int_{z_{lid}}^z (\alpha_M(z') + \alpha_R(z')) dr' \right] \quad (2)$$

and the Rayleigh signal can be written as

$$P_R(z) = \frac{C_R}{r^2(z)} \beta_R(z) \exp \left[ -2 \int_{z_{lid}}^z (\alpha_M(z') + \alpha_R(z')) dr' \right]. \quad (3)$$

Since  $\beta_R$  is directly proportional to the (known) atmospheric density ( $\rho$ ) the aerosol/cloud extinction follows directly from (3), i.e.

$$\alpha_M(z) = -0.5 \frac{d}{dr} \log \left( \frac{z^2 P_R(z)}{\rho(z)} \right) - \alpha_R(z) \quad (4)$$

and the corresponding backscatter can be found by taking the ratio of Eqns. (2) and (3), i.e.

$$\beta_M(z) = \left( \frac{C_R}{C_M} \right) \left( \frac{P_M(z)}{P_R(z)} \right) \beta_R(z) \quad (5)$$

Thus, in contrast to the case of the elastic backscatter lidar, with a HSRL system the aerosol/cloud extinction and backscatter may be independently estimated. As will be explained in more detail later, this simple picture is complicated by real world signal-to-noise and resolution limitations. However, the HSRL technique does represent a notable advance over what can be achieved by elastic backscatter lidars.

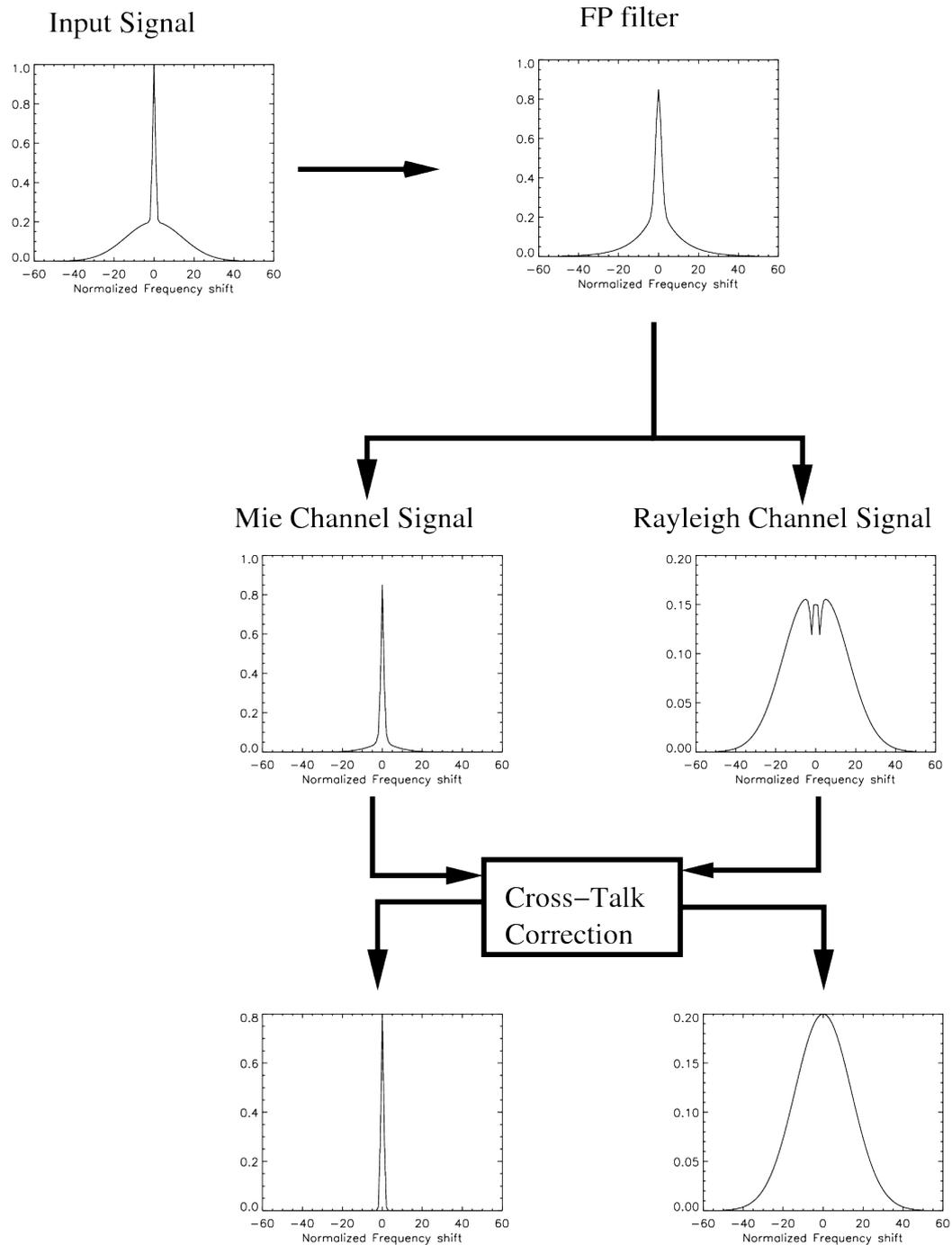


Figure 1: Schematic depiction of the HSR lidar technique.

### **3.3.           *The ATLAS project***

The ATLID Algorithms and Level 2 System Aspects (ATLAS) project started in November 2009 and is part of a number of ESA funded projects regarding the development, implementation, verification and validation of EarthCARE derived Geophysical data products (level 2 products). The current ATLAS project is part of the phase 1 research and development effort mainly focused on the ATLID only algorithms. This first phase was dedicated to the definition of the L2 data products and the identification and theoretical development of L2 algorithms. Main outputs of this phase are the L2 product and algorithm requirement specifications, documentation of the theoretical scientific basis of the selected L2 algorithms and algorithm validation.

The work within the ATLAS project was concerned with two main topics:

- 1) The definition of products and the development of prototype algorithms related to the ATLID instrument and synergistic target classification based on the ATLID, CPR and MSI instruments.
- 2) Assistance in L1 and L2 system support, dealing with general EarthCARE issues put forward by ESA during the project.

In particular, building upon the conclusions from the earlier CASPER final report [CASPER-FR], this study has been concerned with the following products/algorithms:

#### **ATLID- Feature Mask (A-FM):**

This algorithm ingests Lidar L1b signals and derives high-resolution probability estimates of the detection of “significant returns” or “features”. Classification of these features into different types of targets occurs at a different stage (Target Classification). The A-FM mask is used to define binning strategies in other relevant algorithms (A-AER & A-EBD).

#### **ATLID-only Aerosol Extinction, Backscatter and Type (A-AER):**

This algorithm provides the estimates of the aerosol-only Extinction, Backscatter and Depolarization using directly the HSRL Rayleigh method for the derivation of the extinction. Using the backscatter, extinction and depolarization data the aerosol type is determined.

---

### **ATLID-Target Classification (A-TC):**

This algorithm consists of two modules, one concerned with the classification of aerosol, water cloud, ice cloud, super-cooled water cloud etc..) and the second one with the aerosol type assignment. The modules are called by both the A-AER and A-EBD algorithms.

### **ATLID-Extinction, Backscatter and Depolarization (A-EBD):**

This algorithm is primarily concerned with the derivation of cloud/aerosol extinction and backscatter using the Rayleigh and Mie channel signals at high horizontal and vertical resolution (1km by 100m). It uses an optimal estimation approach with the A-AER retrievals supplying part of the a-priori input information. Additionally, the estimation of cloud/aerosol depolarization shall also be produced.

### **ATLID- Ice Cloud Properties (A-ICE):**

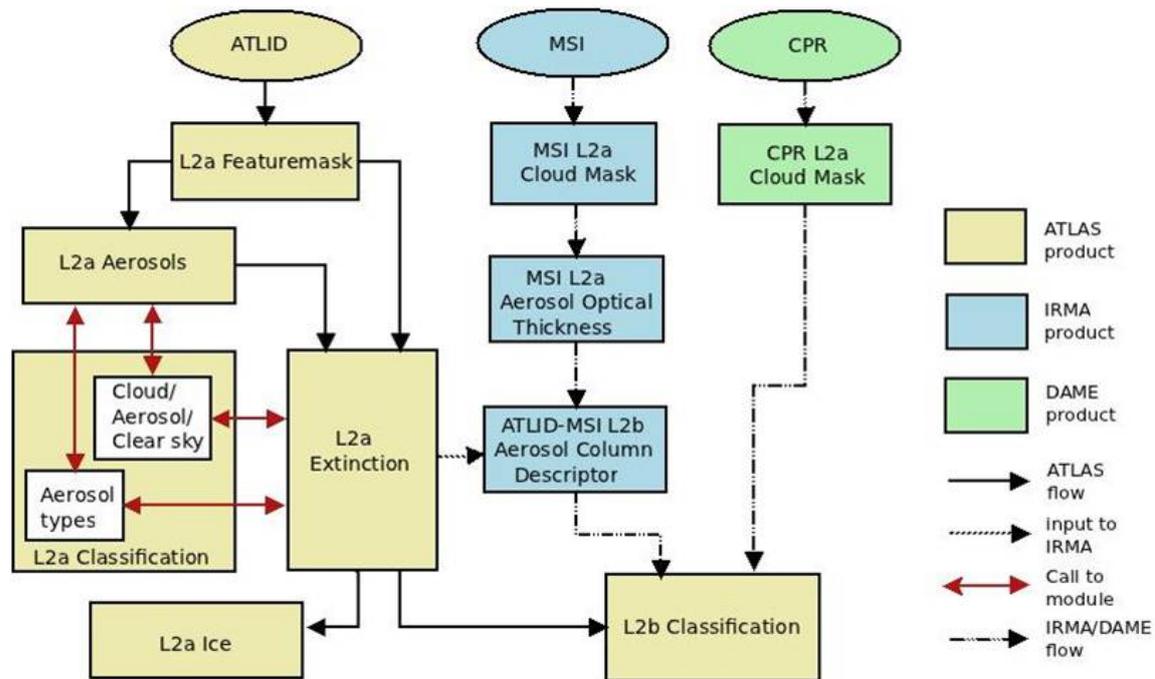
This algorithm produces the estimates of ice water content [IWC] and effective radius [ $R_{\text{eff}}$ ] via an empirical relationship using the estimated ice cloud extinction along with Temperature.

### **ATLID-CPR-MSI Synergetic Target Classification (ACM-TC):**

This algorithm ingests Lidar, Radar and MSI data (L1d) together with target classifications (L2a) data in order to create a combined target mask and classification field. This product is compatible both in terms of conventions and file format with the L2a Radar-only and L2a-Lidar-only products and is a starting point for all synergistic retrieval algorithms.

The algorithms and the logic behind the currently envisioned data flow is presented in the Algorithm Flow Diagram (Figure 2). The Algorithm Flow Diagram includes all the relevant algorithms in the ATLID, MSI and CPR algorithm chains needed for the synergistic target classification algorithm.

---



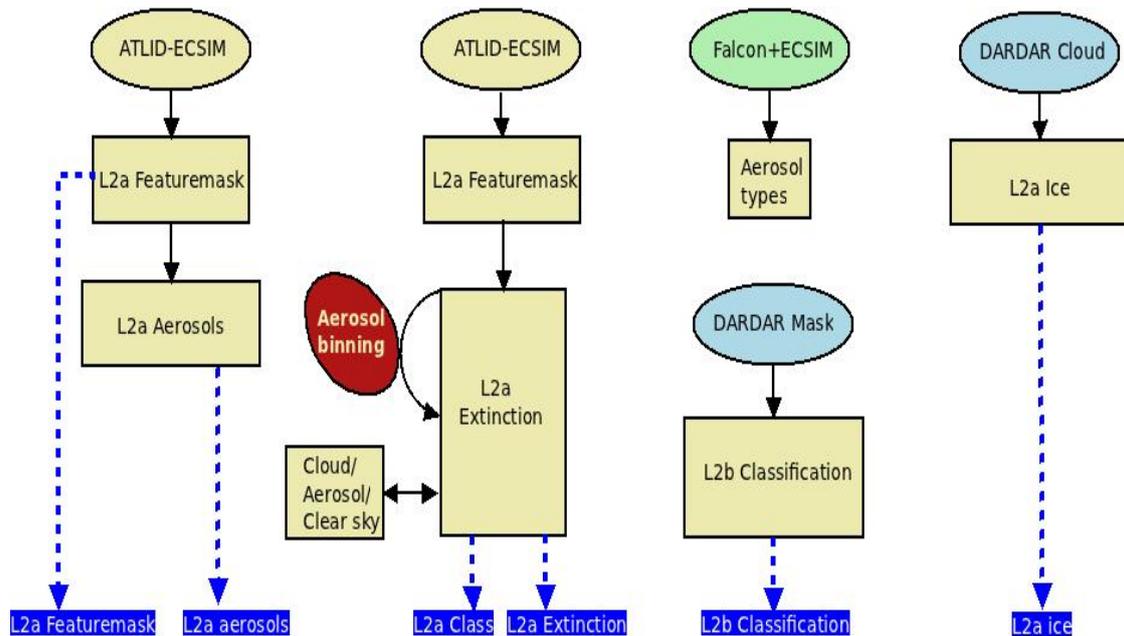
**Figure 2: Schematic relationship of the ATLID-algorithm chain as defined within ATLAS. Shown are the ATLAS algorithms (yellow boxes), the relevant MSI algorithms (blue boxes) and CPR algorithm (green box).**

The logic is based in part on the expected signal-to-noise ratio [SNR] of the ATLID instrument and the EarthCARE requirements. The current instrument design will not have sufficient SNR within a single shot to enable an extinction retrieval using the Rayleigh signals directly, instead it is expected that at least 1 km binning is required for the thicker ice clouds and 10-150 km horizontal binning for regions consisting of aerosols only. This need of multi-shot binning and low SNR requires the use of a strategy involving a masking procedure. A feature mask will enable the development of binning strategies to increase SNR and minimizing the number of shots to bin. Therefore the first step in the procedure is the retrieval of this mask. The retrieval of the extinction at a resolution of 1 km requires a reasonable a-priori input value of the local extinction to backscatter (lidar or S) ratio, this is in general, no problem for the ice clouds but the values can be very different for different types of aerosols. In order to deal with this issue two separate routines will be used derive the lidar ratio for aerosols. One will use the direct retrieval using the Rayleigh signals, which requires a horizontal binning in between 10 and 150 km and the second uses these retrieved lidar ratios as an a-priori input to the extinction retrieval at 1 km horizontal binning. Both these algorithms require a target classification to enable the correct retrieval and binning strategies. Finally, based on the high resolution extinction, backscatter, depolarization and target classification the higher order retrieval algorithm like the synergistic classification and ice microphysical retrieval can be performed.

The full integration of all the algorithms was beyond the direct scope of the current phase of ATLAS activity and the entire chain is not connected (in terms of prototype algorithm implementation) at this point in time. Within this development phase the different prototype algorithms were implemented in single product chains so as to

enable testing and development. In

Figure 3, the separate chains are shown, including the data used to validate the algorithm. In the case of ATLID-ECSIM the data is produced by the EarthCARE Simulator using the ATLID configuration settings. Falcon+ECSIM indicates the use of the DLR-Falcon HSRL data and EarthCARE simulator data using the falcon configuration settings, finally the DARDAR references refer to the CloudSAT and CALIPSO retrieved mask and ice cloud microphysics by J. Delanoe and R. Hogan [Delanoe & Hogan 2010]; <http://www.icare.univ-lille1.fr/projects/dardar/>).



**Figure 3: Schematic relationship of the current algorithm chain as is available at the end of the ATLAS project as described in the relevant ATBDs .**

All the relevant connections are available through place holders (e.g. the red box indicated by aerosol binning is a ``quick and dirty fix`` of the L2a aerosols routine) and can be subsequently connected when both algorithms are fully implemented within the EarthCARE Simulator environment [ECSIM].

## 4. ATLAS Algorithms

### 4.1. ATLID- Feature Mask (A-FM):

#### 4.1.1. Overview

The algorithm finds the feature mask based on the correlation of the data without focussing on a number of hard coded or input dependent thresholds. The main reason for this is the relatively large number of noise counts present in the ATLID signals. As the signal strength of aerosol or very optically thin ice clouds on the single shot grid can be comparable to the noise levels it was chosen to rely on image reconstruction techniques and not on signal to noise ratios and thresholds. The main reason why an image reconstruction technique can be so effective for the ATLID data is that in principle the Mie signals contain only particle backscatter and noise due to the Mie-Rayleigh cross-talk. No molecular backscatter, e.g. no variable background signal, should be present in these channels.

The present algorithm is based on earlier development carried out within the CASPER project and is based upon techniques drawn from the field of image processing. Two methods were employed in the previous CASPER algorithm to retrieve the feature mask: the median-hybrid method [Russ 2007] and the maximum entropy [Skilling and Gull 1985] method, both using the detection probability. Based on these two methods, coherent structures can be defined. The maximum entropy method however does not always converge and focuses on the stronger features in the noise and can miss some of the more tenuous widespread aerosol layers. To make sure that the algorithm is robust enough for the usage of space-based data, the algorithm has been simplified and now uses 3 or 4 convolved images instead of attempting to retrieve the rigorous maximum entropy defined image.

The features with a low signal to noise ratio within the images can be distinguished from the noise only signals by performing a Gaussian fit to the signal probability distribution. Both the high and low SNR routines implicitly use horizontal and vertical information from neighbouring pixels to define structures. A short summary of the algorithm follows while details can be found in [A-FM-ATBD] and [A-FM-PDD].

#### **Major Inputs:**

L1 Mie backscatter and associated errors

#### **Major Outputs:**

Target mask consisting of an integer scale related to the probability of a feature being detected.

#### **Output Resolution and Sampling:**

The output product is reported at the native lidar resolution, however, the actual

---

resolution of the data product will depend on the SNR of the input.

---

#### **4.1.2.        *Algorithm Flowchart and description***

A flowchart depicting the main algorithm steps is shown in Figure 4. The feature mask algorithm can be defined in a six steps procedure where each of the steps can be followed in the flow diagram. The basic idea behind the algorithm is to first extract the high signal to noise features. This is followed by a method to check for low signal to noise features by smoothing the image to an appropriate degree. The six steps can be summarized as follows:

1. Calculating the signal probabilities.
2. Applying the hybrid median edge preserving technique to retrieve the coherent high signal to noise regions.
3. Iteratively convolve the remaining low SNR signals with a 2D Gaussian smoothing kernel.
4. Calculate the probability distribution histogram for the appropriate convolved images.
5. Separate the noise from the signals by fitting a Gaussian noise peak within the histograms and applying a threshold
6. Apply the hybrid median technique to combine the results from the high SNR and low SNR results.



### **4.1.3. Example**

The algorithm has been applied to various ECSIM scenes including simple 'block' cloud scenes as well as more realistic scenes derived from actual measurements made by an airborne HSR lidar system (see [Weinzierl and van Zadelhoff 2011]). As well, a version of the algorithm has been applied to CALIOP 1064 nm data. The Rayleigh backscatter signal in the 1064nm channel is relatively weak so that the signal roughly mimics the characteristics of the ATLID Mie channel in this respect.

The CALIPSO team has a similar data product, the VFM mask [Vaughan 2009]. This mask is based on the 532nm channel and is therefore more sensitive to the very small aerosol particles compared to results based on the 1064nm channel. There is a large conceptual difference between the ATLID-Featuremask and the VFM mask in the sense that the ATLID-Featuremask is used for processing the lidar signals before these are used and the VFM mask is a higher order algorithm. The VFM mask requires a high enough signal to noise ratio data to be able to assign cloud and aerosol types and requires a larger binning of the data. This requirement of high SNR will be reflected by the larger and blockier structures found in the VFM mask compared to the ATLID-Featuremask. As this is the only available lidar dataset from space which can be directly downloaded and is kept up to date it is the ideal dataset for evaluating the algorithm presented here.

The CALIOP data changes in vertical resolution at 8.2 km from 30 to 60 m. The change in resolution can potentially change the noise structure in the image. To exclude any influence of the change in resolution the featuremask is calculated in two separate runs. One focussing on the field below 8.2 km and one dealing with the data from 8.2 km up to 19 km. This is reflected by the horizontal line in the featuremask. It is also visible in the retrieved value of the features just above and below 8.2 km.

In the following figures the 1064nm raw data is shown in the top figure, the ATLID Featuremask in the center figure and the corresponding VFM mask in the bottom figure. The color scale of the Featuremask represents the chance of a pixel being molecular (0) or containing cloud or aerosol particles (10), while the VFM mask represents the classification of the pixel (1: clear sky, 2: clouds, 3:aerosols, 4:stratospheric features, 5:surface, 6: sub-surface and 7 fully attenuated). As the noise in the 1064nm channel hampers the detection of the attenuated regions the attenuated area's retrieved from the VFM mask have been added to the ATLID-Featuremask.

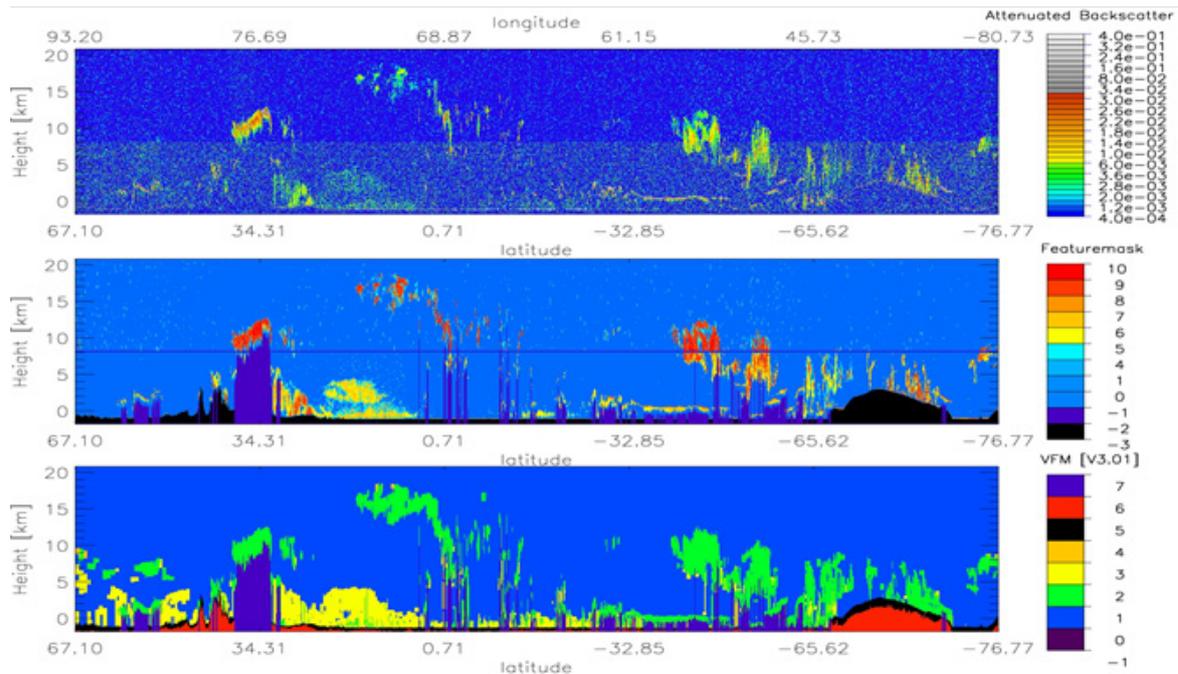


Figure 5: Full orbit of CALIOP night time data (CAL\_LID\_L1\_ValStage-V3-01.2010-04-18T20-52-36ZN), the top Figure represents the raw 1064nm data, the center panel the Featuremask, and the bottom panel the VFM mask

In Figure 5 to Figure 7 the 1064nm Calipso data from 18 April 2010 is presented (CAL\_LID\_L1\_ValStage-V3-01.2010-04-18T20-52-36ZN). In the first of these figures the entire orbit is presented, followed by two zooms into regions where a combination of aerosol, ice clouds and liquid clouds is presented.

In general all the visible features in the raw data (top figure) are present in both masks. The VFM mask fills in a lot of gaps and has in general smoother and larger features compared to the ATLID-Featuremask. The Featuremask follows the raw data structures more closely however. In the left part of the image (between 67.10 and 34.31 longitude) the VFM mask finds many features which are not available in the Featuremask. Based on the data both algorithms have issues in this region and this should be looked at in the future.

In Figure 6 and Figure 7 two zooms are presented of this orbit to visualize both the agreement of the two algorithms to detect features in the data and show the differences between the two algorithms, which are due to their specific assumptions and needs. Figure 6 shows an aerosol rich region with a thin ice cloud and some scattered liquid cloud layers. The ATLID Featuremask follows the raw data very well and finds elevated aerosol layers at -21 degrees latitude. The VFM mask retrieves a continuous aerosol layer throughout this region. The difference is most likely caused by the VFM need of a high SNR and therefore a large horizontal binning. In Figure 6 complex ice cloud structure is situated above a large number of liquid cloud layers. There is a potential aerosol layer on the right side of the figure. The median hybrid edge conserving method retrieves as much as possible the complex structure within the ice cloud. The retrieved liquid layers are thinner compared to the VFM results,

also visible by the small separation seen often between the liquid layers and the fully *attenuated regions taken from the VFM mask*.

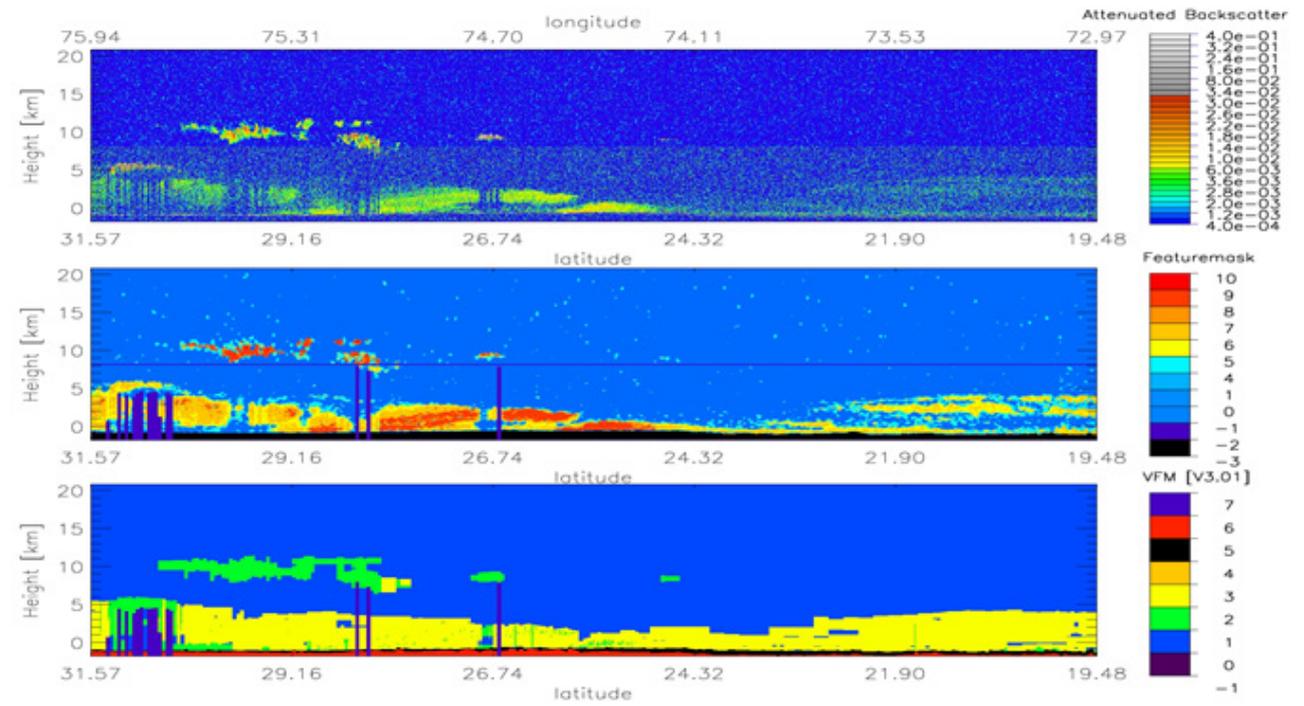


Figure 6: Zoom of an aerosol rich region from the orbit presented in Figure 5.

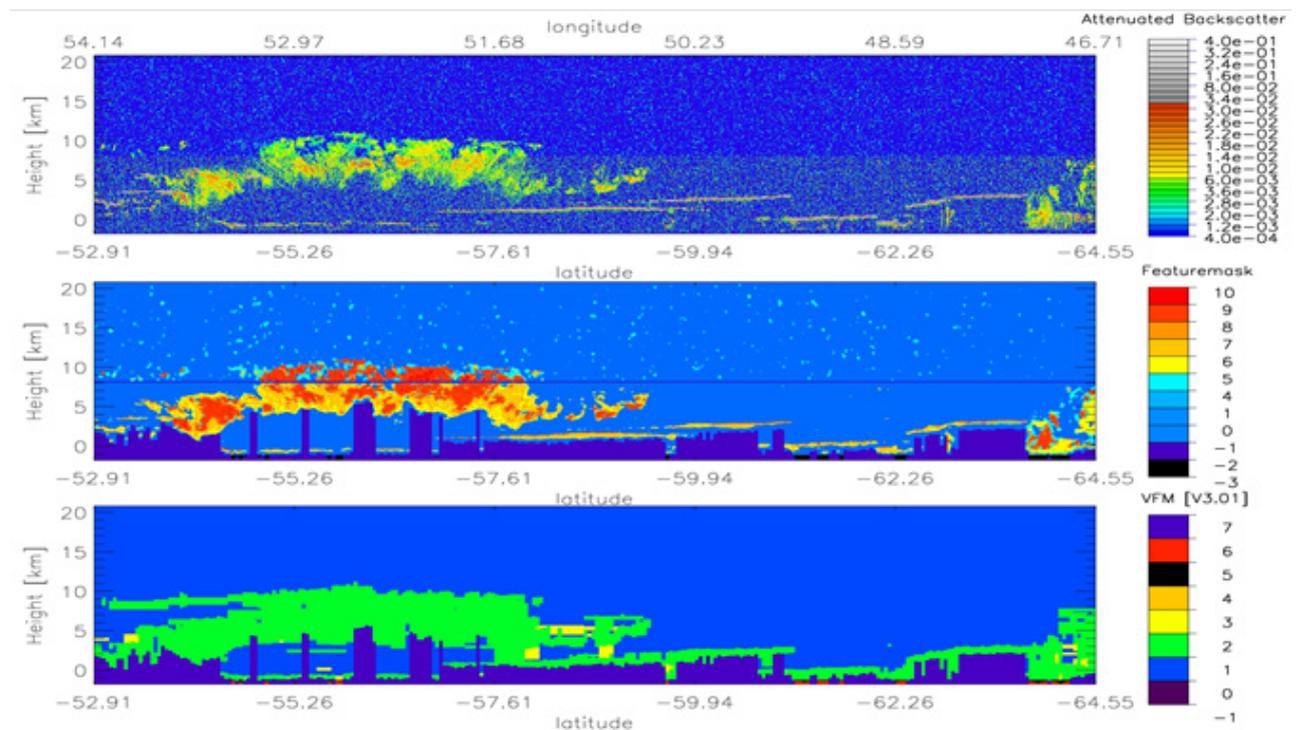


Figure 7: Zoom of a region from the orbit presented in Figure 5 where both ice clouds and liquid layers are present. There is an additional aerosol region in the bottom right.

#### **4.1.4.        *Algorithm Status***

The first version of the ATLID-Featuremask which was developed in the CASPER project solely focussed on results from the ECSIM lidar modules. At that time the background noise and cross-talk coefficients were underestimated making relatively easy to retrieve a good Featuremask.

In the presented updated version of the Featuremask the ECSIM lidar module show more realistic signals and (background) noise behaviour. The algorithm has been extensively checked using 4 night-time and 3 day-time CALIPSO orbits. The CALIOP 1064nm data is currently the only available data-set from space for this type of validation. Visually, the night time orbits seem to represent the data very well. In case of the day time orbits the algorithm settings may have to be better optimised.

#### **Future Validation and Development needs**

The most important validation still lacking for the current use of the algorithm is the determination of those regions which are fully attenuated. The current Rayleigh channel settings are based on an older version of the ECSIM lidar module and these are no longer valid for the current noise levels within the forward modelled data.

The algorithm has to be tested against more realistic ECSIM scenes such as those created in the ICAROHS project [Weinzierl and van Zadelhoff 2011]. Secondly the optimisation for CALIOP data has to be performed after which a statistical comparison can be done using the raw data, the VFM mask and the Featuremask. Thirdly; campaign data from a, as of yet unspecified, future airborne UV HSRL instrument, should be used for validation.

Finally all the settings of the algorithm have to be recalibrated in the commissioning phase of the EarthCARE satellite after launch. This will be a vital and delicate procedure for retrieving the best possible results from the ATLID instrument.

---

## **4.2. ATLID-only Aerosol Extinction, Backscatter and Type (A-AER)**

### **4.2.1. Overview**

The objective of this algorithm is to retrieve the aerosols optical properties (the extinction coefficient  $\alpha$ , the backscatter coefficient  $\beta$  and depolarization ratio  $\delta$ ) from the ATLID instrument on board the EarthCARE satellite. Unlike the 1-km scale Cloud and Aerosol extinction, backscatter and depolarization ratio described in Section 4.4, the algorithm uses the lidar signals directly and does not rely on a priori assumptions on the backscatter-to-extinction ratio. The output of this algorithm will serve as an input to the 1-km scale algorithm.

Retrieving the value of the extinction coefficient is a relatively simple procedure for signals with a sufficient SNR (Signal-to-Noise ratio). This procedure involves the determination of the vertical differentiation of the logarithm of the range corrected Rayleigh signal. However, this direct procedure cannot be applied directly in the case of the EarthCARE signals. Due to the relatively low SNR associated with the EarthCARE Rayleigh channel, the direct differentiation will lead to an amplification of the extinction variance and to an unreliable retrieval of the extinction. It is therefore important to take into account the low expected SNR of the EarthCARE signals. It should be noted that, similar arguments apply to the derivation of the backscatter, however, the extinction retrieval is more sensitive to the SNR of the input lidar signals. Different methods will be used to deal with the noisy nature of the signals.

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

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

- In the horizontal dimension, the signal is averaged using a sliding window. The width of this sliding window is variable and is adjusted according to the actual signal to noise ratio of the signals and a configurable threshold SNR of the lidar signals needed to obtain the required accuracy of the extinction and backscatter products. The lower limit of the required SNR has been determined in a sensitivity study.
- In the vertical dimension, the signal is smoothed using a linear fitting procedure, with a fixed window width, before the signal derivative is calculated.

This algorithm also calls the L1a classification and Aerosol typing routines in order to assign an aerosol type. These procedures are described within Section 4.3.

---

### **Major Inputs:**

- 1) L1b Mie and Rayleigh attenuated backscatter and associated errors.
- 2) L2a Featuremask (From A-FM)
- 3) L1D grid definition.

### **Major Outputs:**

- 1) Aerosol Extinction and associated covariance matrix.
- 2) Aerosol Backscatter and associated errors.
- 3) Aerosol Depolarization ratio and associated errors.
- 4) Aerosol type

### **Output Resolution and Sampling:**

The output product is reported on a grid corresponding to the L1D standard grid. .

The vertical resolution of the backscatter and depolarization ratio is the same as the native lidar resolution. The horizontal resolution of the extinction, backscatter and depolarization ratio products are variable (10-150 km). The vertical resolution of the extinction product is expected to be on the order of 500 m..

#### 4.2.2. Algorithm Flowchart and description

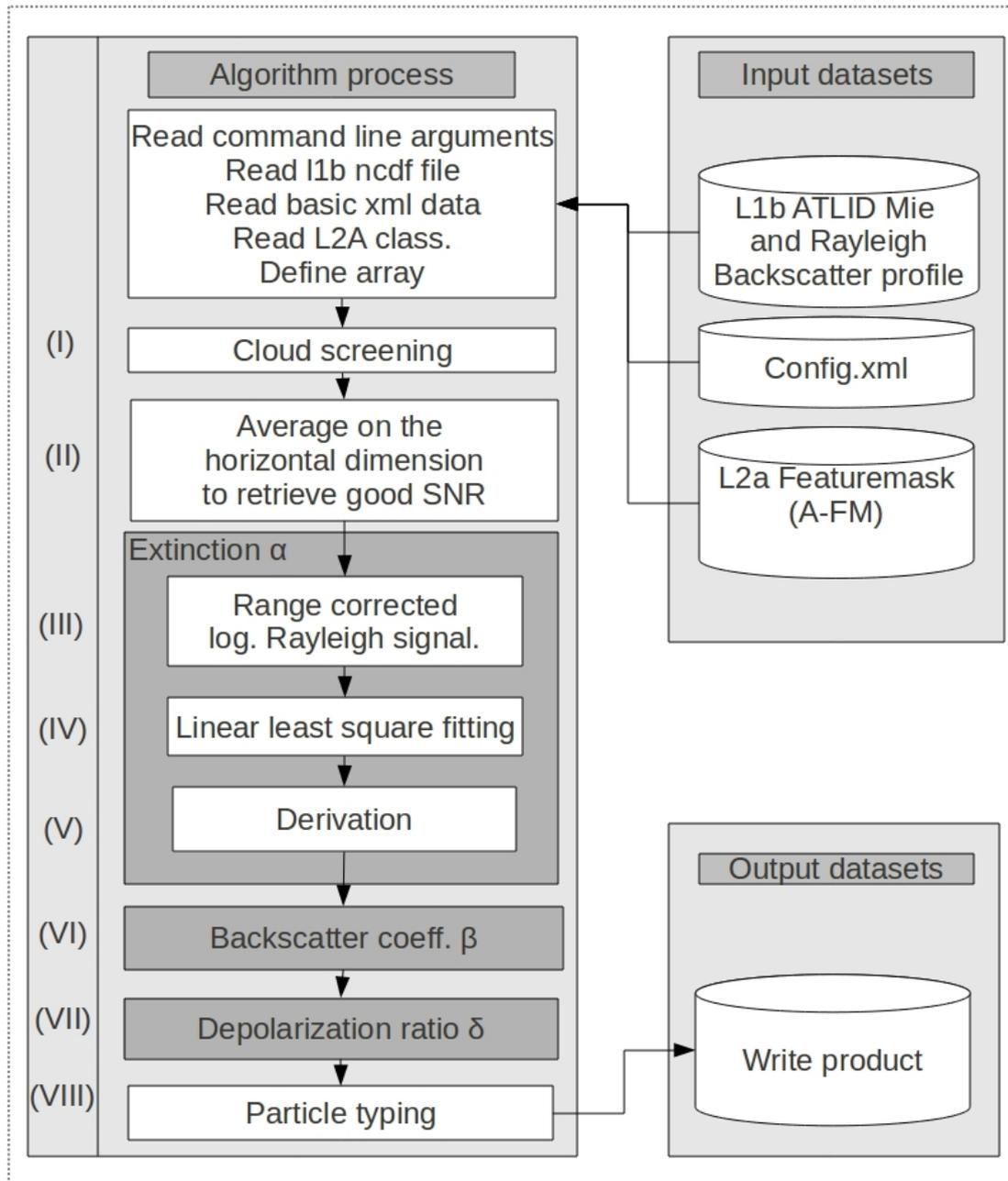


Figure 8: ATLAS large-scale aerosol properties retrieval algorithm flow diagram

The algorithm steps are schematically depicted in Figure 8.

### I) Cloud Screening

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

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

- $AM(t,z) = 0$  where  $z < Z_{cld}(t)$
- $AM(t,z) = 1$  otherwise

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

### II) Horizontal averaging in order to achieve required SNR

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

Here the following 3 steps are carried out :

1. The window width is set to its default value.
2. The average cloud scattering quantities are calculated.
3. If  $Min[SNR_{P_M}(z)]$  or  $Min[SNR_{P_R}(z)] > SNR_{min}$ , the width of the window is increased and control is passed back to step (2) until the SNR threshold is reached or the maximum allowed window width is reached.

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

### III-V) Extinction Retrieval

Here the extinction is found using the horizontally averaged cloud screened signals. The method follows a version of Eqn. (4) with the slope of the signal determined by a sliding-window linear least squares fit.

Since extinction points within the vertical fitting window width will be not be independent it is necessary to take this information into account when reporting the errors associated with the extinction. Accordingly, the covariance matrix of the extinction is calculated.

---

## **VI) Backscatter coefficient**

Here the backscatter coefficient is found using the ratio of the Mie and Rayleigh signals (i.e. Eqn. (5)). The associated errors are also calculated.

## **VII) Depolarization Ratio**

Here the depolarization ratio and its associated error estimates are found using the ratio of the cross-polar and Mie channels.

## **VIII) Particle Typing**

Using the depolarization ratio, the magnitude of the backscatter, the backscatter-to-extinction ratio and other auxiliary information such as the layer temperature and height a particle type is assigned. This is accomplished by a call to the A-TC routine.

---

### 4.2.3. Example

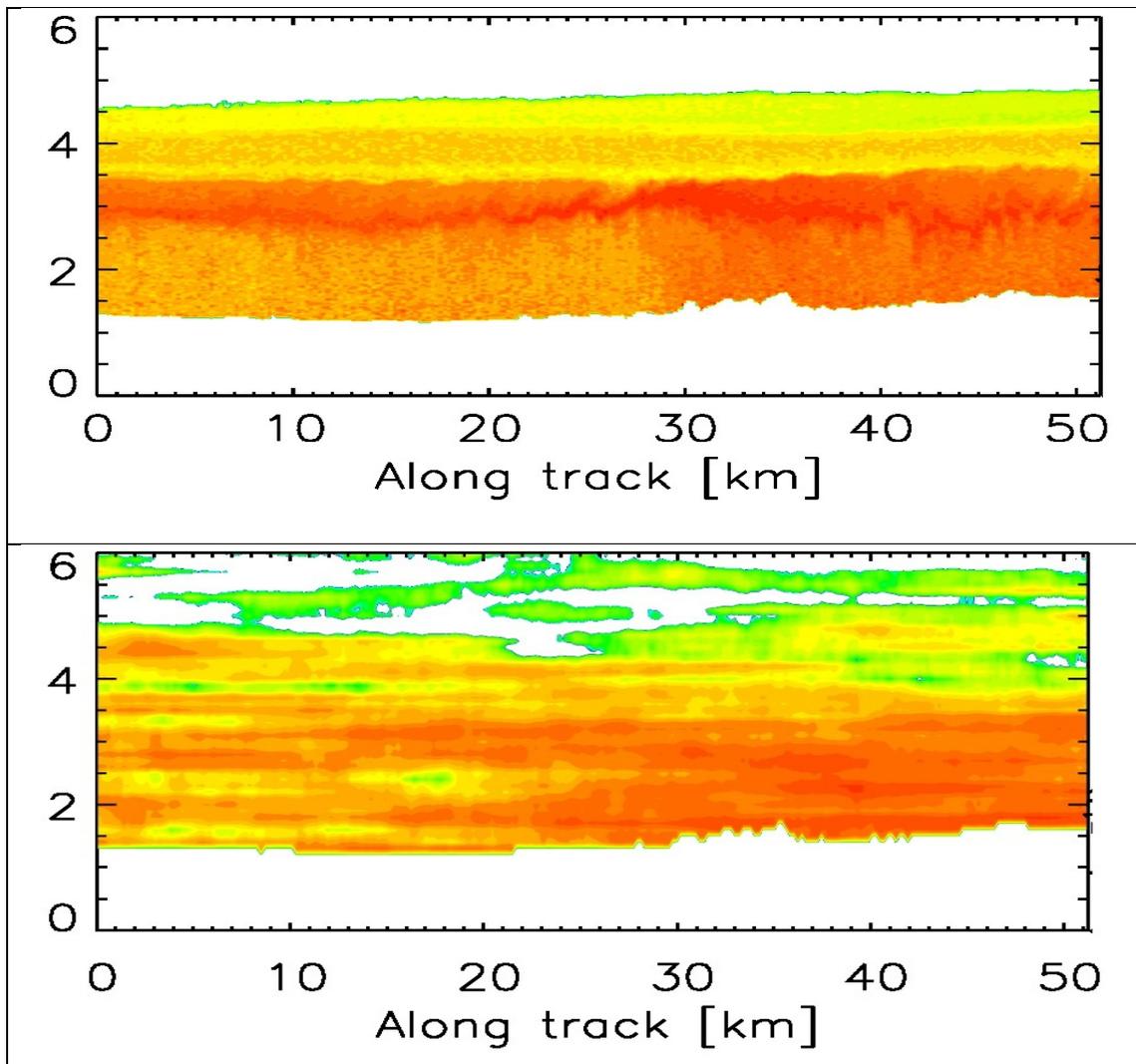
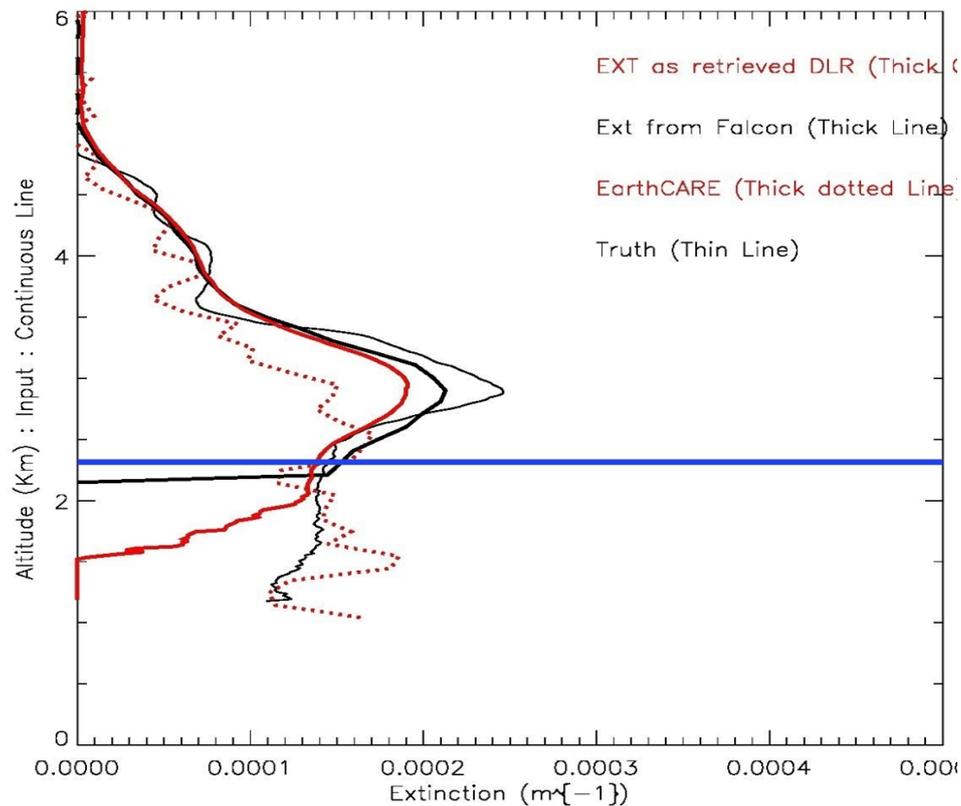


Figure 9: Top “true” extinction from the ESCIM scene derived from the Falcon dataset. Bottom: Extinction as derived from the EarthCARE signal. Here a horizontal window of 20 km and a vertical window of 900 meters was used. Note that structure at finer scales is visible due to the sliding window (in both the horizontal and vertical dimensions) nature of the algorithm.

The algorithm has been tested using several ECSIM scenes derived from actual HSRL measurements see [Weinzierl and van Zadelhoff 2011]]. An example is



**Figure 10 :** Corresponding to horizontal averages of the case shown in Figure 9 this plot shows (1) The Extinction as retrieved by the DLR (Red thick continuous line), (2) the Extinction from the Falcon dataset (black thick continuous line), (3) the EarthCARE extinction (Red thick dotted line), and (3) the True extinction (black thin line). Data under the blue line should not be taken in account due to ground return.

shown in Figure 9 and Figure 10. Here the scene corresponds conditions encountered during the SAMUM-1 campaign by the DLR Falcon on the 4<sup>th</sup> of April 2006 (Morocco during a Saharan dust event).

#### 4.2.4. Algorithm Status

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

#### Future Validation and Development needs

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

### **4.3.            *ATLID Target Classification (A-TC)***

#### **4.3.1.            *Overview***

This procedure assigns a target classification on the basis of retrieved aerosol/cloud properties (backscatter, depolarization etc.) combined with ancillary information (layer temperature, layer height). The implementation is divided into two distinct modules. The first module is the main classification module identifying water cloud, ice clouds and aerosol regions. The second module deals with the aerosol typing only.

#### **4.3.2.            *L2a Lidar classification***

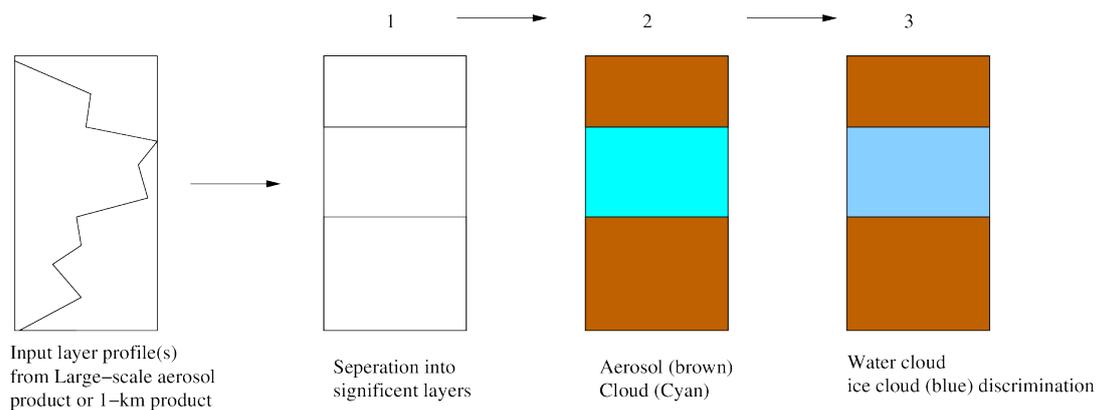
The task of the L2a Lidar classification is to process regions previously identified as containing “Targets” (see [A-FM-ATBD] and [A-EBD-ATBD]). This procedure first sub-divides an input (vertical) regions based on the associated backscatter ratio and depolarization ratio into sub regions. Then based, on a priori backscatter and depolarization thresholds as well as the observed relationship between integrated backscatter and depolarization within each sub region the sub-regions are labelled as ice, water or aerosol. The procedure also makes use of auxiliary data such as wet-bulb temperature derived from i.e. ECMWF analysis fields.

The L2a lidar classification can be decomposed into three main areas.

1.        Detection of statistically significant height boundaries.
2.        Cloud-Aerosol identification
3.        Water-cloud/Ice-cloud separation.

A high-level sketch of the process is shown in Figure 11. The procedure is intended to be applied to output from the Large-scale Aerosol Extinction Backscatter and Depolarization Algorithm [A-AER-ATBD] and also to be used as a component of the high resolution lidar Extinction, Backscatter and Depolarization product processing procedure (see Sections 3.3.6 and 5.3.2 of [A-EBD-ATBD])

---



**Figure 11: Sketch of the L2a Lidar classification main steps.**

### Major Inputs:

- 1) L2a extinction and error estimates (from A-AER or A-EBD)
- 2) L2a backscatter and error estimates (from A-AER or A-EBD)
- 3) L2a Depolarization ratio and error estimates (from A-AER or A-EBD)
- 4) Geolocation and auxiliary Met. Info.

### Major Outputs:

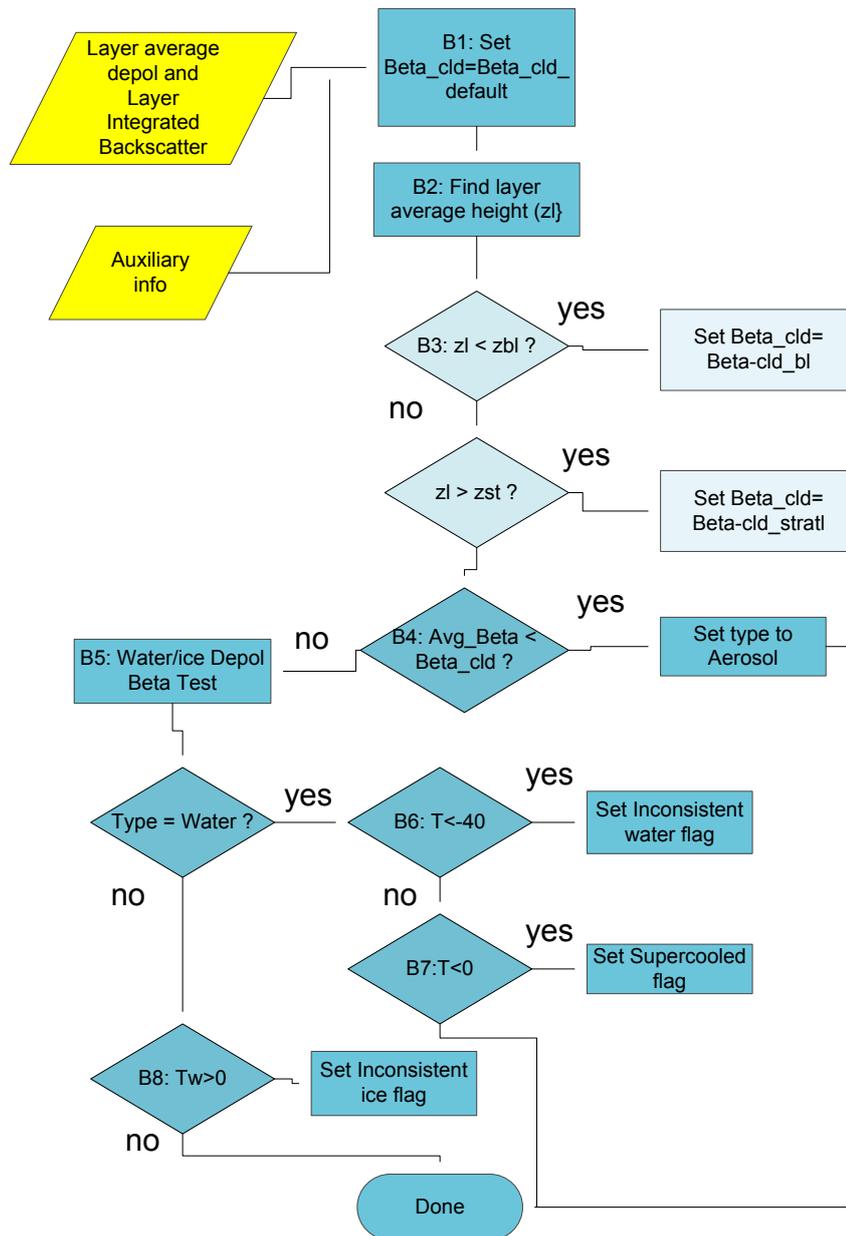
- 1) Layer averages of the input fields and associated error estimates.
- 2) Target type mask. An integer mask corresponding to different types of targets being present (Clear sky, water cloud, ice cloud, aerosol etc..). Probability/confidence indicators are also reported.

### Output Resolution and Sampling

The output grid will always correspond to the input grid.

The horizontal resolution of the output will match that of the input. However, the vertical resolution of the output will depend on the layering structure of the input data.

### 4.3.2.1. Algorithm Flowchart and description

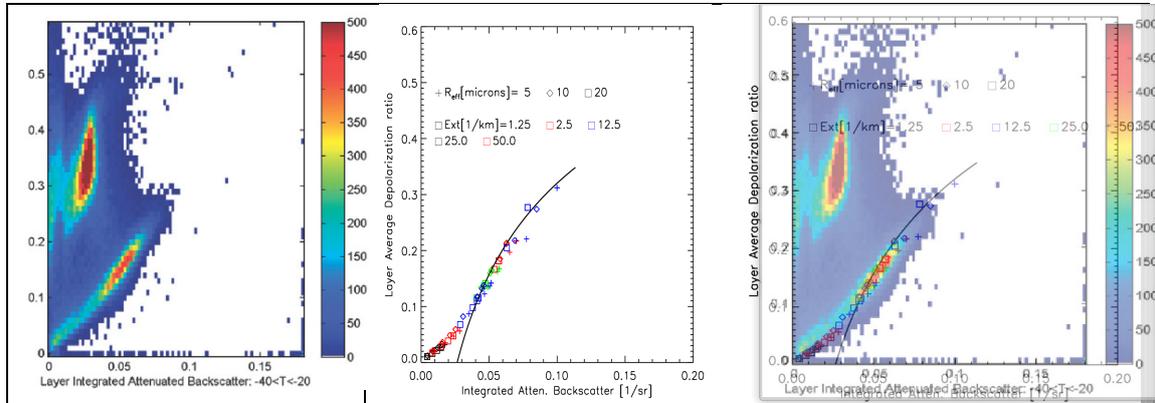


**Figure 12: Aerosol cloud discrimination and water cloud/ice cloud discrimination procedure. More detail associated with each step can be found in [A-TC-ATBD].**

A schematic depiction of the main steps involved in the ice/water/aerosol discrimination part of the classification algorithm is shown in Figure 12 (the flowchart corresponding to the layer identification part of the algorithm is not shown but may be found in [A-TC-ATBD]).

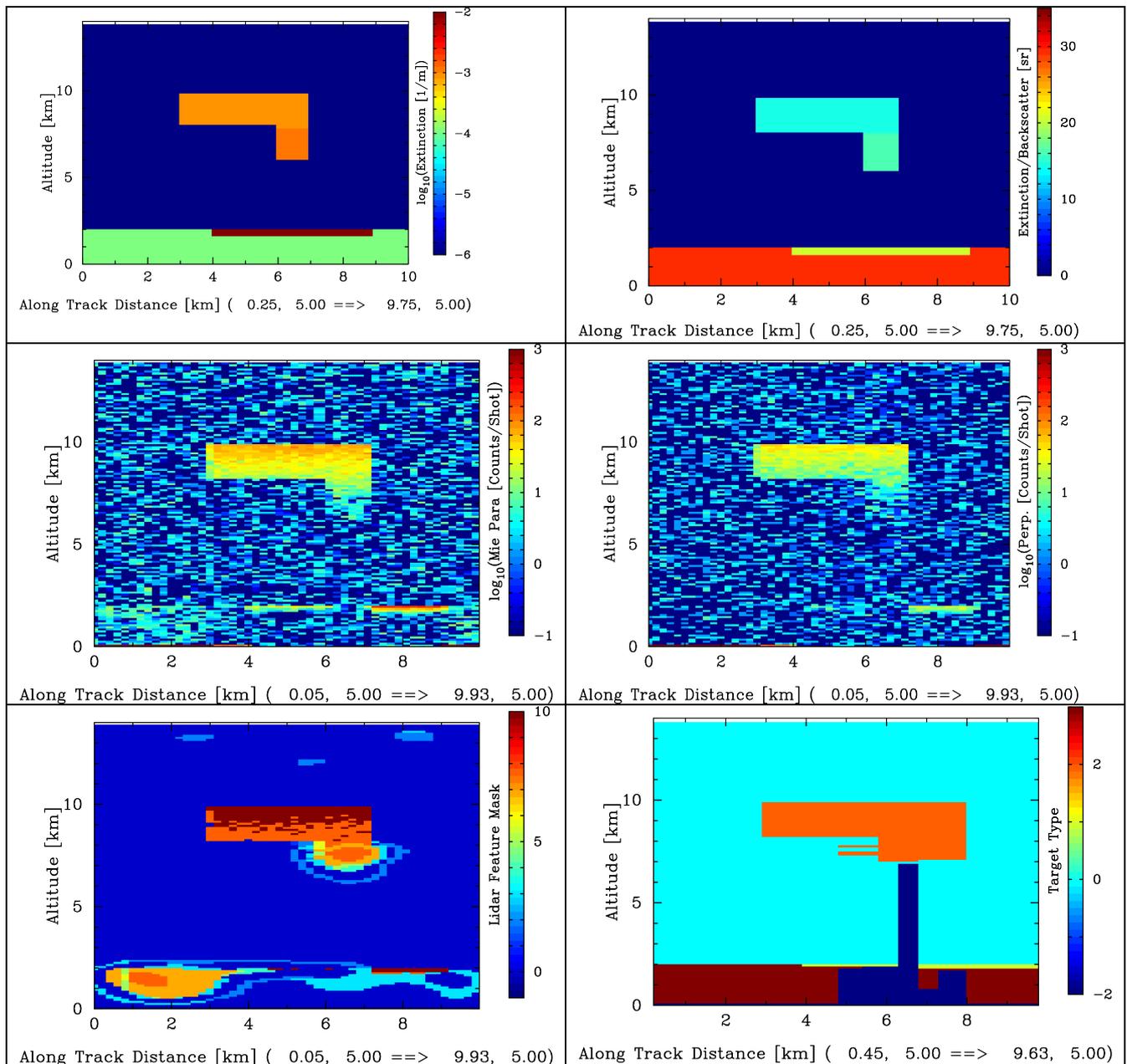
The cloud/aerosol discrimination procedure relies on a number of threshold based tests. The thresholds involving discrimination clouds from general aerosols are the most subjective and will require further refinement including likely adjustments to be

made post-launch. However, the tests used to separate ice clouds from water clouds using a combination of the backscatter and depolarization ratio are more certain. In particular, a method similar to that employed by CALIOP [Hu. et al. 2009] can be applied in the case of ATLID (See Figure 13).



**Figure 13: Left: Histogram built using CALIPSO observations taken from Hu et al. 2009. The upper left region corresponds to ice clouds. Middle: Points: Results of Lidar Monte-Carlo calculations for various water clouds applied to the ATLID configuration. Solid-Line Fit to earlier MC results for CALIPSO [Hu. 2007]. Right: Overlap of the other two panels.**

### 4.3.2.2. Example



**Figure 14: Top Left: Extinction field. Top Right Extinction-to-Backscatter field, Middle-Left: Mie channel signals. Middle-Right: Cross-polar channel signals. Bottom-Left: Feature mask output. Bottom-Right: L2a Target Classification. Here Red is Aerosol, Cyan is clear sky, Orange is ice cloud and Blue is unknown (Lidar signals are attenuated too much).**

The classification routine has been applied to several ECSIM scenes. An example is shown in Figure 14. Here the ECSIM standard scene was used. In fact this figure corresponds to the example case presented in Figures 19-21 in the High-Resolution Extinction algorithm ATBD. There are three main target areas present above 5 km a cirrus cloud is present. At 2 km there is a stratus water cloud present and below 2 km a boundary-layer aerosol field is present. Here it can be seen that the cross-polar

signals associated with the water cloud are significant and could indeed complicate the phase identification if only depolarization ratio were used. However, by using a combination of backscatter and depolarization as described earlier the water layer is correctly identified as such.

#### **4.3.2.3.     *Algorithm Status***

The layer finding and classification routines presently exist in prototype form as a combination of a stand-alone IDL routine and FORTRAN code. They have not yet been fully integrated into the ECSIM environment. The prototype procedures have been shown to perform as intended for simple cases. Verification against more realistic cases will require full ECSIM integration.

#### **Future Validation and Development needs**

The method used here in order to separate statistically significant layers is rather general in nature and has been evaluated/validated using ECSIM generated data. In the future more complex scenes based on actual observations will be used.

The method used here to distinguish between clouds and aerosols is crude and further work involving the analysis of Raman lidar and HSRL datasets will be required to appropriately set the threshold values and indeed to determine of a simple three threshold approach as currently envisioned is sufficient.

The depolarization-vs-backscatter and depolarization-vs-integrated-backscatter approaches to distinguishing between water and ice clouds, however is on a quite solid physical foundation. ECSIM lidar forward calculations (which themselves have been validated against observations see [EC-FT-ATLAS]) show that a simple and robust procedure for separating water and ice clouds using ATLID (unattenuated) backscatter and depolarization measurements can be constructed. Further, ECSIM calculations closely resemble the results of independent theoretical calculations and indeed actual CALIPSO observations of the relationship between layer integrated attenuated backscatter and depolarization (see Figure 13). This gives us a high degree of confidence in the ECSIM calculation with respect to the relationship between layer depolarization ratio and the layer (unattenuated) backscatter.

---

### 4.3.3. Aerosol typing

The aerosol typing procedure expands upon the general aerosol assignment provided by the L2a Lidar classification procedure by assigning more detailed aerosol sub-types. This procedure uses aerosol depolarization, extinction and backscatter together with auxiliary (i.e. relative humidity from ECMWF analysis) and a priori information (such as likely hood of aerosol type occurrence as a function of location and season) in order to assign a probability of occurrence for a number of suitable aerosol types.

Aerosol shapes and sizes differ from place to place and time to time, statistically however general particle properties can be defined using the measured depolarization and retrieved lidar ratio. This has been extensively shown by combining results from different measurement campaigns [Ferre et al 2007, Muller et al 2011]. An example of how the aerosol types in the troposphere depend on the aerosol backscatter-to-extinction ratio ( $S$ ) and linear depolarization are shown in Figure 15. The data used to produce this plot is based upon observations made during the LACE-98, SAMMUM 1&2 and EUCAARI campaigns. From this figure it can be seen that the most robust separation is associated with the distinction between absorbing and non-absorbing aerosols. Also shown is that African biomass burning aerosols have a higher depolarization compared to the Canadian type in spite of what is expected. The main difference is that the African measured values all come from the SAMUM2 campaign in which both biomass burning and African dust was present in the same area.

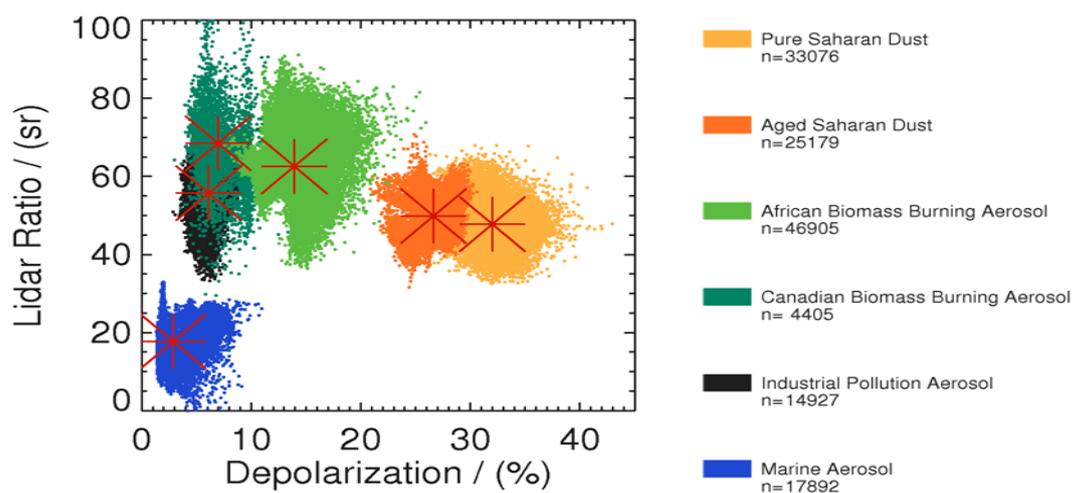


Figure 15: Lidar characteristic properties for different aerosol types from DLR field measurements (ICAROHS ATBD1) preliminary results.

What can be concluded from the measured probabilities are that:

- for each of the main aerosol types a distribution can be defined based on the observations of regions consisting of a single aerosol type (defined by the in-situ observations and/or multi wavelength HSRL/Raman lidar measurements)
- different aerosol types reside in distinct regions within the depolarization-lidar ratio parameter space

- aerosol distributions overlap and therefore the assignment of a single type is not always possible
- the observed aerosols can be a mixture themselves, showing a larger than realistic spread in parameter space
- for ATLID there will be no color ratio or depolarization ratio available making it more difficult to separate types
- additional information is needed to assist in separating different types.

For each type a Gaussian distribution will have to be defined. The two dimensional distribution needs to be able to take into account any angle dependence (correlations) in order to mimic the measured distributions best. When the main types are defined by their specific distributions the probabilities can be calculated for each individual observation.

#### **Major Inputs:**

- 1) **L2a Target Classification output (as described in Section 4.3.2).**
- 2) L2a extinction and error estimates (from A-AER or A-EBD)
- 3) L2a backscatter and error estimates (from A-AER or A-EBD)
- 4) L2a Depolarization ratio and error estimates (from A-AER or A-EBD)
- 5) Geolocation and auxiliary Met. Info.
- 6) a priori aerosol occurrence maps

#### **Major Outputs:**

- 1) Most likely dominant Aerosol type and associated probabilities

#### **Output Resolution and Sampling.**

For this procedure, the output resolution and sampling grid matches that of the input data.

---

### 4.3.3.1. Algorithm Flowchart and description

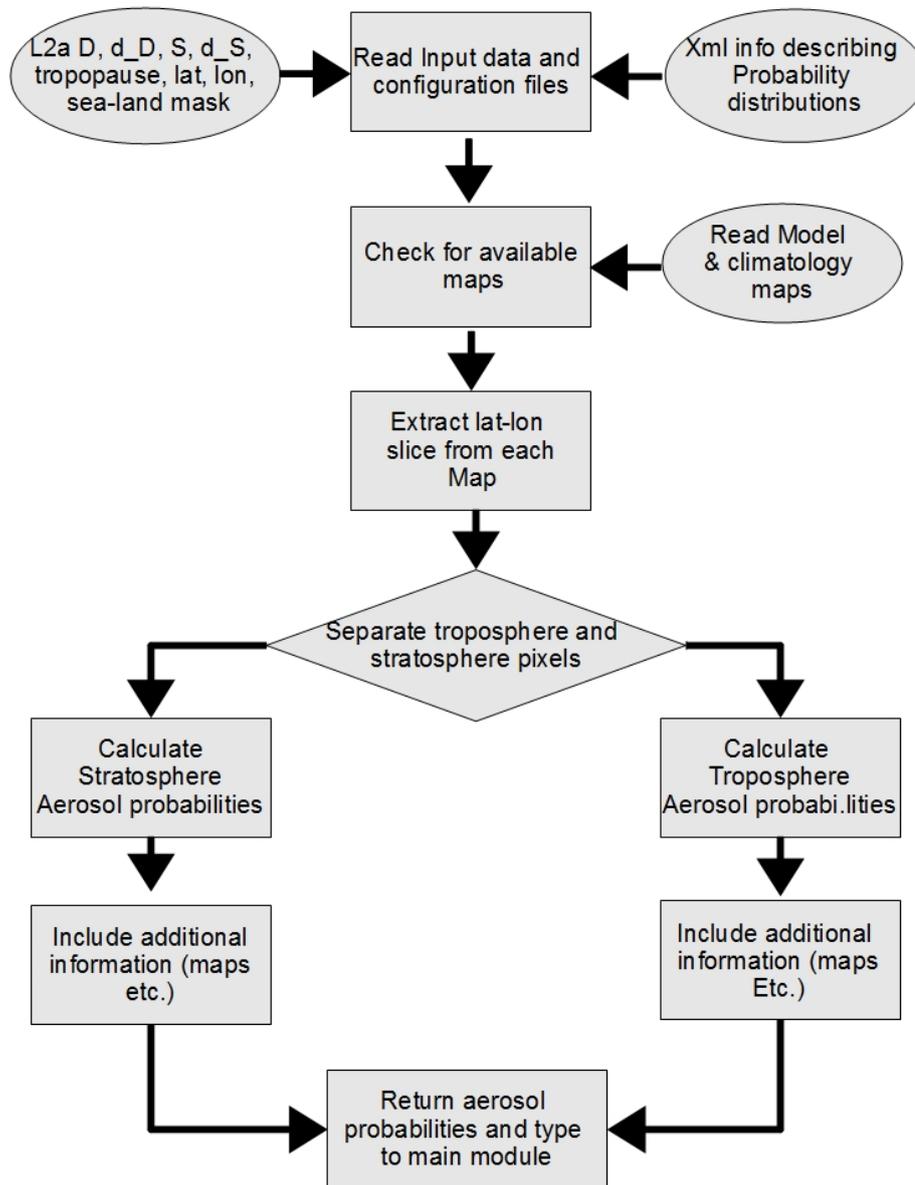
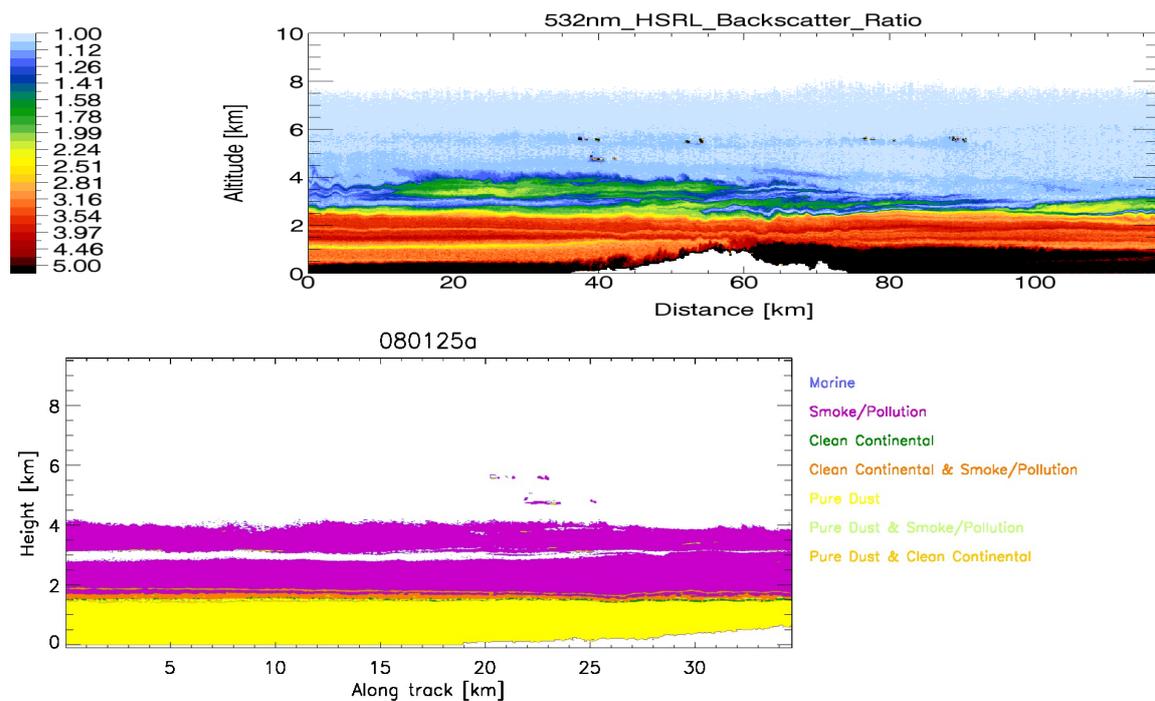


Figure 16: Aerosol Typing flow diagram

A schematic of the main steps involved in the aerosol typing procedure are shown in Figure 16. The process involves combining the information from the L2a lidar observations with a priori ‘‘aerosol map’’ information. More details can be found in [A-TC-ATBD].

### 4.3.3.2. Example

A prototype form of the aerosol typing procedure has been applied to several instances of actual HSR lidar observations. An example is shown in Figure 17 corresponding to the flight on 25 January 2008 during the SAMUM2 campaign [Heintzenberg 2009]. The SAMUM2 campaign was dedicated to a closure study in the Cape Verde area. On this day, a mineral dust layer extended from the ground up to an altitude of about 1.5 km. The dust layer was topped by a biomass burning layer which covered the altitude between 1.5 and 4 km altitude.



**Figure 17: Top panel: Observed backscatter profile from the DLR Falcon HSRL. Bottom panel: retrieved aerosol type mask. Note the different horizontal scales**

The retrieve aerosol classification indicates that the scene is dominated by dust in the lowest layer and Smoke/Pollution in the top layer. The transition layer indicates that the probabilities of both clean continental & smoke are similar suggesting that the lidar ratio is already dropping in this regime but that there are no big dust particles present to increase the depolarization. The remaining types all reside in the masked edges and surface.

### **4.3.3.3.     *Algorithm Status***

The L2a lidar aerosol typing scheme exists in prototype form and is not integrated into ECSIM at this moment. Since the algorithm is written to only calculate those pixels which are assigned as aerosols therefore it is expected that the algorithm will be fast enough for operational use. The use of additional data based on either observations or model calculations will require a pre-processing step.

#### **Future Validation and Development needs**

The most important task, related to this algorithm, will be to determine the a-priori lidar ratio-depolarization distribution for all aerosol types based on HSRL and Raman data. Also, the validation and organization of the map creation will be an important task in the future. Both of these tasks will have to be dealt with in a future cal.-val. activity and will require a dedicated effort of combining different available data bases and the design of campaigns to complement the available data. Note that at this point there are **no** aircraft UV HSR lidars and therefore all assumptions here are based on the 532nm data-sets.

#### **4.4. ATLID-Extinction, Backscatter and Depolarization (A-EBD)**

##### **4.4.1. Overview**

The Rayleigh signal from a HSRL lidar can be used to estimate the extinction profile in a rather direct manner [Shiple, 1983] (see also Eqn. (4)) by estimating the derivative of the range-corrected logarithmic signal. However, the applicability of this method is limited due to a high required SNR ratio. Thus, one can assert that accurate but low-precision extinction information is, in general, provided by the Rayleigh signal. In contrast, extinction information can also be extracted from the Mie signal channel which, in general, may be viewed as less accurate (since factors such as the extinction-to-backscatter ratio must be specified in order to “invert” the signal [Klett 1985]) but more precise (since the SNR ratio of the input data does not impact the derived extinction product to the same degree of extent.).

Taking into account these two observations, it is advantageous to formulate a retrieval procedure which simultaneously uses both the Rayleigh and Mie signals in order to combine the accurate but less precise Rayleigh channel derived information with the less-accurate but more precise information derived using the Mie channel. The essence of the algorithm is to perform a Klett-like retrieval using an extinction-to-backscatter (S) profile which yields an extinction profile which, in turn, enables an optimal reconstruction of the observed Rayleigh channel signal. It is thought that an optimal-estimation based variational approach is best suited for this purpose.

In contrast to the earlier algorithm developed during the CASPER project, the approach developed here is (to a degree) tolerant of cross-talk correction errors. In particular, the procedure in essence performs its own cross-talk correction procedure and estimates of the actually cross-talk correction coefficients are generated.

##### **Major Inputs:**

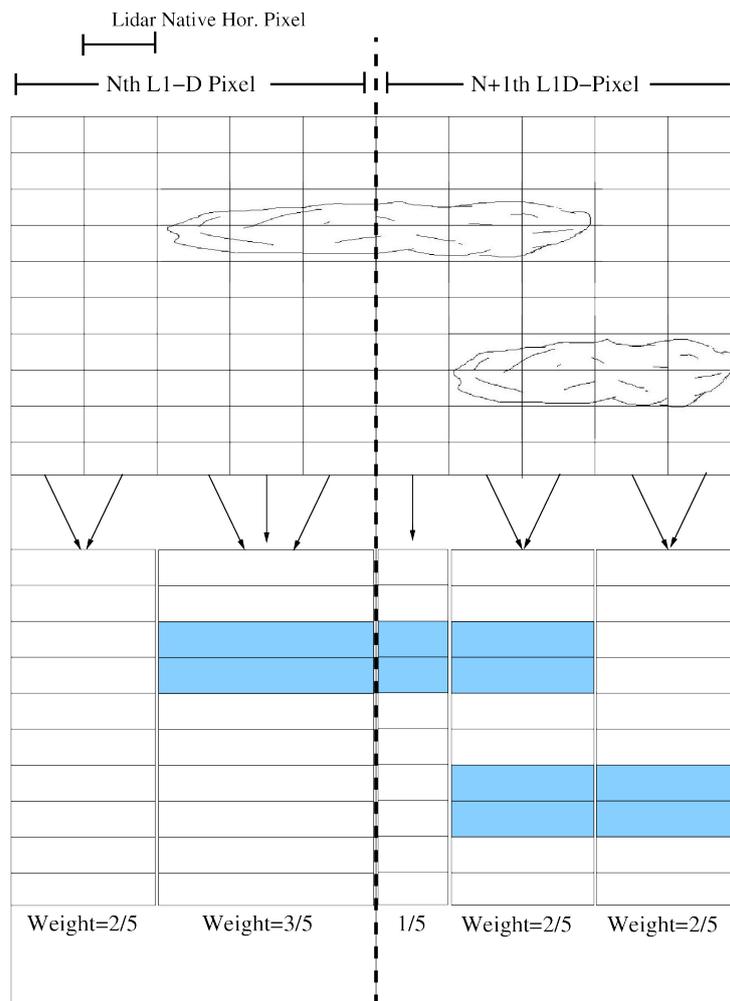
- 1) L1D Grid
- 2) L1b lidar signals and associated errors.
- 3) L1b cross-talk coefficients
- 4) Auxiliary Met data.
- 5) A priori estimates of the extinction-to-backscatter ratio per scattering type.

##### **Major Outputs:**

- 1) Extinction, Backscatter and Depolarization ratio as well as error estimates
  - 2) Target classification via calls to A-TC routines.
-

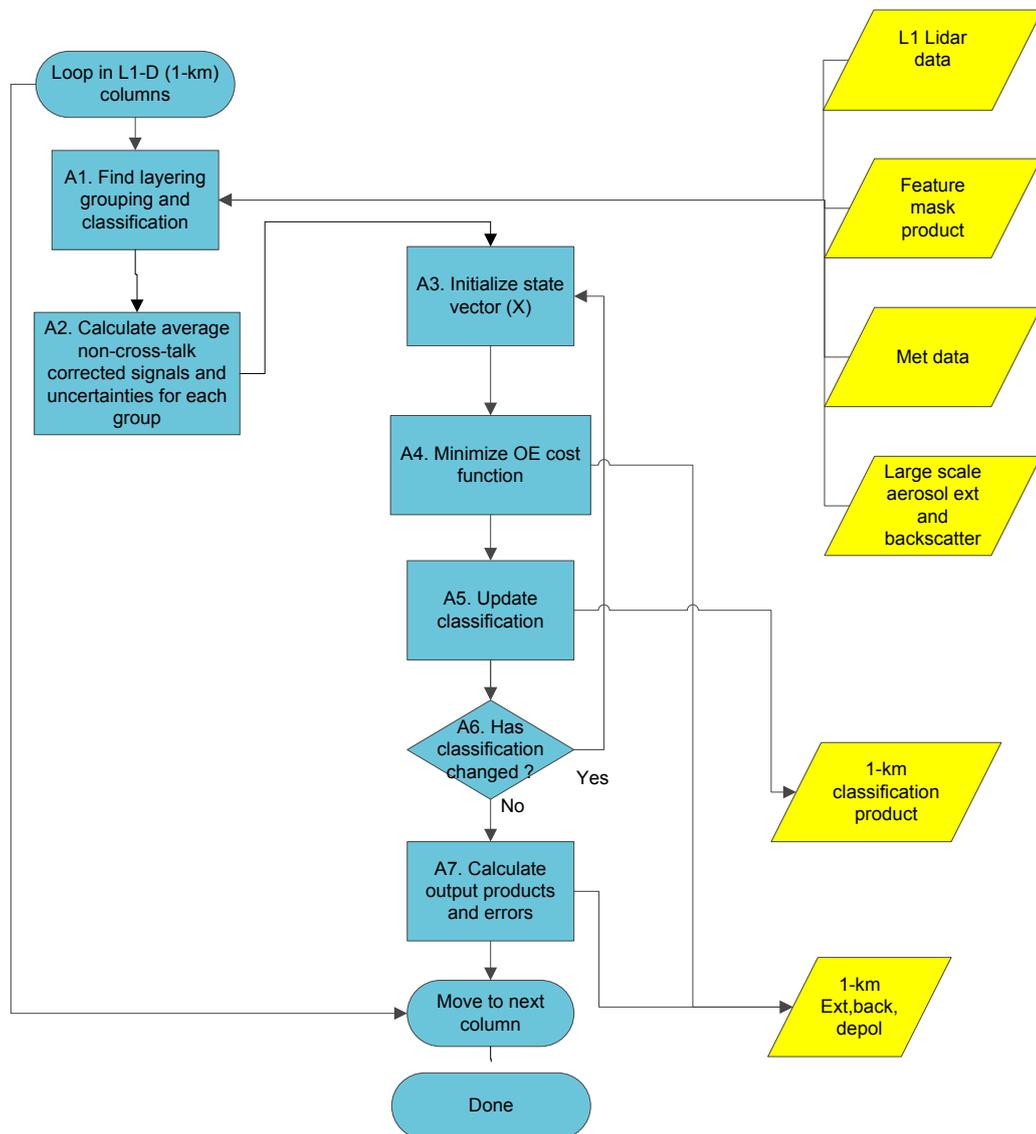
### Output Resolution and Sampling:

The output products are reported according to the L1D vertical grid. However, the horizontal sampling and resolution depends on the cloud structure (see Figure 10). This is due to the fact that it is undesirable to average cloud and non-cloud lidar profiles together. It should also be noted that extinction-to-backscatter and effective particle size variables are per-layer values.



**Figure 18: Schematic depiction of the horizontal averaging strategy employed within this algorithm.**

#### 4.4.2. Algorithm Flowchart and description



**Figure 19: High-level structure of the main algorithm. Here the Yellow trapezoids represent input or output data sets.**

A high level flowchart of the algorithm is presented in Figure 19. Each of the steps is described in detail in [A-ATLID-EBD].

The core of the algorithm involves the inversion of the lidar equation using the estimated total signal (which is estimated by appropriately summing the Mie+Rayleigh+Cross Polar) channel signals.

The extinction profile may be estimated from the total channel signal using a so-

called Klett-like inversion. In particular, it can be shown that (see [A-ATLID-EBD] Section 3.3.

$$\alpha_{Mie}(z) = -\frac{1}{2} \left[ \frac{P'_t(z)M_t(z)S(z)r^2(z)}{\frac{P'_t(z_m)M_t(z_m)S(z_m)r(z_m)^2}{\alpha_{Mie}(z_m) + S(z_m)\beta_{Ray}(z_m)} - 2 \int_{z_m}^z P'_t(z')M_t(z')S(z')r(z')^2 dr} - S(z)\beta_{Ray}(z)} \right] \quad (6)$$

where  $z$  is the altitude,  $r(z)$  is the range from the lidar,  $P'_t$  is the Total signal corrected for Rayleigh attenuation,  $z_m$  is a prescribed boundary range and  $M_t$  is the multiple-scattering factor. If  $z_m$  is chosen to be an altitude high enough so that no non-negligible amounts of aerosol/cloud exist between  $z_m$  and  $z_{lid}$  we can write

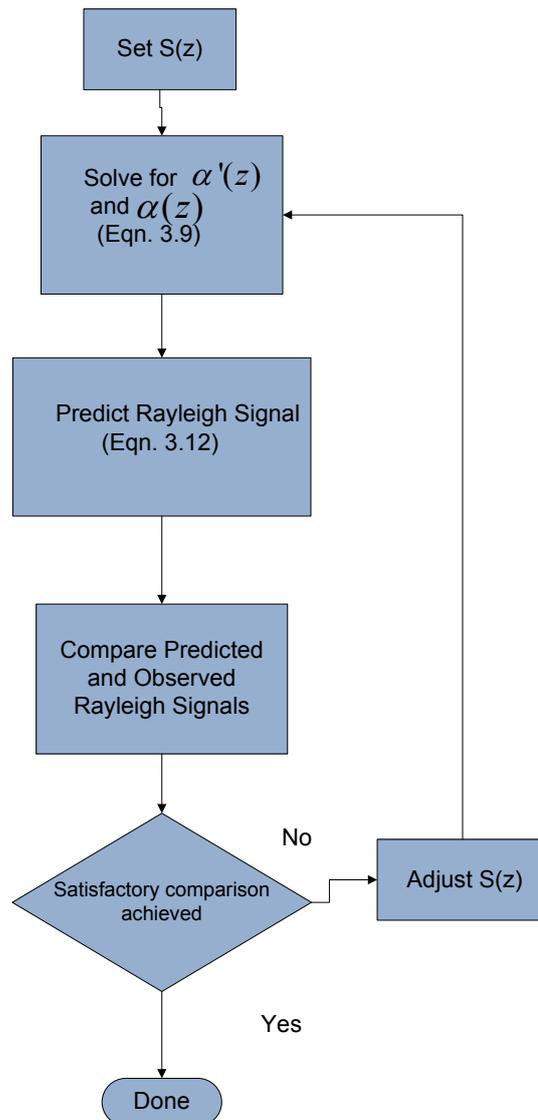
$$\alpha_{Mie}(z) = -\frac{1}{2} \left[ \frac{P'_t(z)M_t(z)S(z)r^2(z)}{\exp \left[ -2 \int_{z_m}^{z_{lid}} S(z')\beta_{Ray}(z')dr' \right] - 2 \int_{z_m}^z P'_t(z')M_t(z')S(z')r^2(z)dr'^2} - S(z)\beta_{Ray}(z)} \right] \quad (7)$$

Thus, if  $S(z)$  is specified and if  $M_t(z)$  can be accounted for then the extinction profile can be estimated. Once the extinction profile has been calculated then the Rayleigh signal can be calculated i.e.

$$P_{Ray}r(z)^2 = \beta_{\pi, Ray}(z)M_{Ray}(z) \times \exp \left[ -2.0 \int_{z_{lid}}^z (\alpha_{Mie}(z') + \alpha_{Ray}(z')) dr' \right] \quad (8)$$

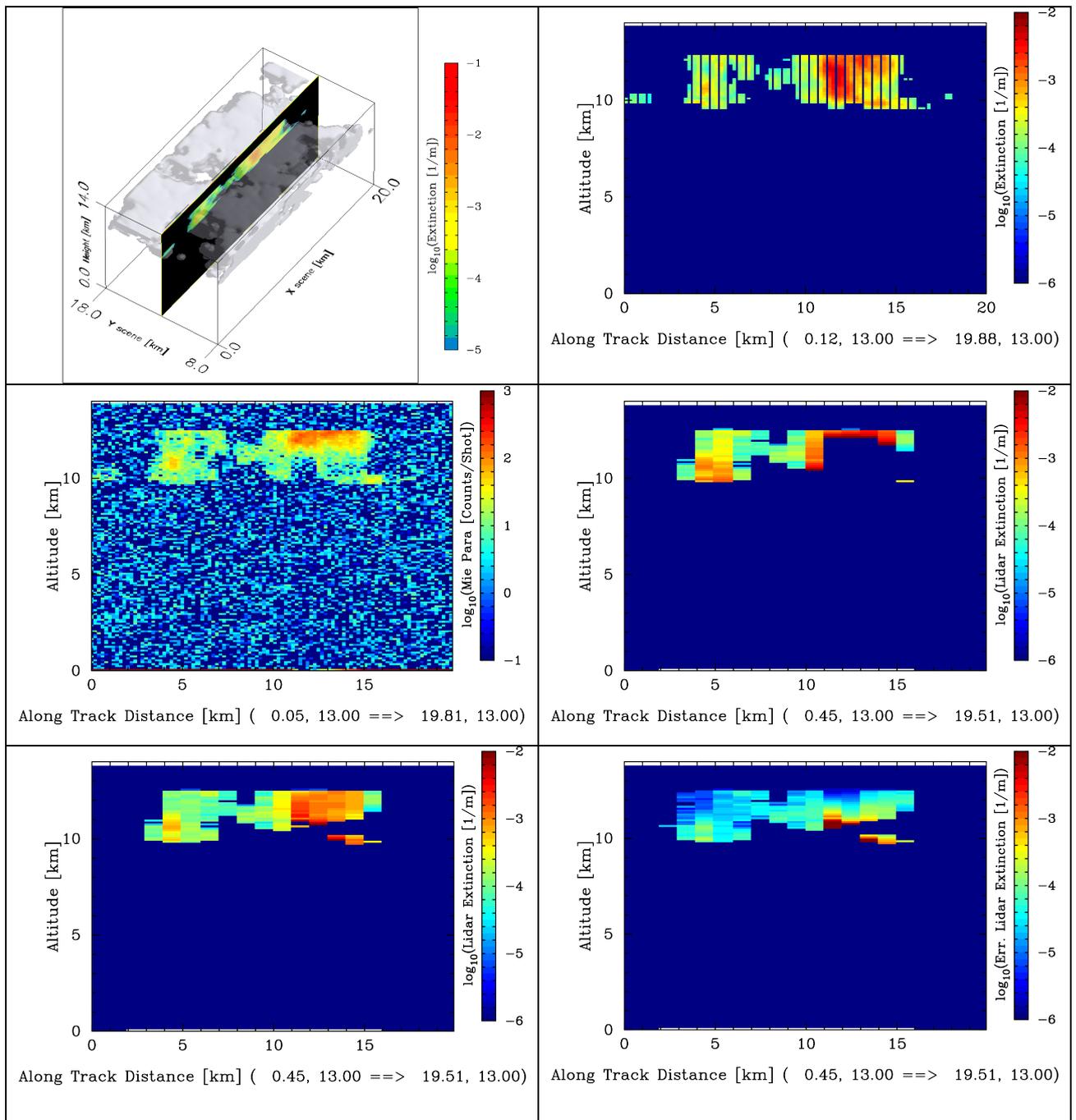
The observed and predicted Rayleigh signal may then be compared against each other and the  $S$  profile adjusted iteratively (using some sort of non-linear minimization solver) to achieve an optimal fit (see Figure 20).

In the full procedure, an optimal estimation approach is used [Rodgers 2000] and the cost function contains terms related (amongst others) to the a priori  $S$  profile. The cost function also contains a term which avoids problems associated with “forward inversion instability” associated with the solution of Eqn. (7). Uncertainty in the calibration and cross-talk parameters is also accounted for and multiple scattering effects are accounted for with the aid of the lidar multiple scattering model due to [Hogan, 2006].



**Figure 20: High level simplified schematic of the core approach of the algorithm Note that within this diagram multiple-scattering is not explicitly treated. The Equation numbers refer to those used in [A-EBD-ATBD].**

### 4.4.3. Example



**Figure 21: Top-Left: 3D view of the “Fractal Cirrus” ECSIM scene. Top-Right: “True” Extinction along the lidar track. Middle-Left: Mie channel signal Middle-Right: Extinction retrieval produced by the first-guess state-vector Bottom-Left: Retrieved best estimate and Bottom-Right: Estimated error.**

The prototype procedure has been applied to several ECSIM scenes of various complexity. For example it has been applied to cirrus cases generated by ECSIM aided by the use of the “fractal cloud generator” developed by [Hogan and Kew 2005]. A 3D view of the scene in terms of the extinction field is shown in the top-left panel of Figure 21. The scene has an inhomogeneous cloud structure and the Extinction-to-Backscatter ratio varies between about 11 and 17.

The Mie signals corresponding to a satellite overpass through the middle of the scene are shown in the Middle-Left panel Figure 21 and the corresponding extinction field is shown in the Top-Left panel retrieval results are shown by the other panels of the figure. Here an a priori value for  $S$  of 20 was used with an uncertainty of  $\pm 50\%$ . Here it can be seen that the optimal extinction retrieval is a marked improvement over the first-guess retrieval with a seemingly good correspondence between the true values and the retrievals in spite of the biased a priori  $S$  specification. Also, the extinction error estimates in this case seem appropriate and only reach large values in areas in where the cirrus optical depth is near to fully attenuating the signal. With respect to the  $S$  retrievals it can be seen that the retrievals are biased towards the a priori in the thinner parts of the cloud. However, the agreement is better in the thicker parts. Encouragingly, the error estimates associated with the thinner areas are correspondingly larger than for the thicker parts.

#### **4.4.4.        *Algorithm Status***

A prototype implementation of the algorithm has been largely integrated into the ECSIM environment.

It has been demonstrated that a real improvement in retrieved Extinction and Extinction-to-Backscatter ratios is achieved by using an optimal estimation base approach. This is compared with the case of Klett-based retrievals, and though not discussed in any depth here, much more precise than direct Rayleigh signal derivative based inversions on high resolution horizontal and vertical scales. In short, the goal of wedding the high precision (but low accuracy) of Klett-based approaches with the high accuracy (but low precision) of Rayleigh signal derivative based methods has arguably been convincingly demonstrated.

#### **Future Validation and Development needs**

It goes almost without saying that much validation and development remains before the procedure described in this work reaches maturity. Since there is no available (or foreseen, barring ATLID itself) space-based lidar signals similar enough to EarthCARE simulations, such as those generated by ECSIM will continue to play a key role in any further development work. In the future the focus will shift towards the use of larger scenes derived from observations (i.e. for example, those built as part of the ICARHOS project [ICAROHS-TN2]) and those built using cloud resolving atmospheric models.

In addition, any practical implementation this algorithm requires the specification of a number of parameters which depend on the desired output resolution and the scale of the cloudy features themselves. In particular, the maximum number of allowed layers and allowed sub-columns must be set. As these parameters will impact the algorithm performance, including data product storage size, it will be desirable to optimize the choices for these two and other parameters. This may be accomplished by suitable analysis of CALIPSO data.

---

## **4.5. ATLID- Ice Cloud Properties (A-ICE):**

### **4.5.1. Overview**

The objective of this algorithm is to retrieve, by using the lidar only datasets (backscatter and extinction as retrieved in the L2 datasets), the Ice cloud properties. Using Extinction and temperature parameters, The Ice Water Content (IWC) profiles are retrieved using an existing parametrization given in [Heymsfield et al., 2005]. Using the [Foot, 1988] parametrization and the IWC values previously processed, the effective radius ( $R_{eff}$ ) profiles can also be estimated.

#### **Major Inputs:**

- 1) L1D grid
- 2) Extinction and Classification information from L2a A-EBD
- 3) Temperature from ECMWF analysis

#### **Major Outputs:**

- 1) Empirical estimates of IWC and associated errors
- 2) Empirical estimates of ice cloud  $R_{eff}$  and associated errors

#### **Output Resolution and Sampling:**

The output resolution and sampling is the same as for the as the A-EBD product.

---

#### 4.5.2. Algorithm Flowchart and description

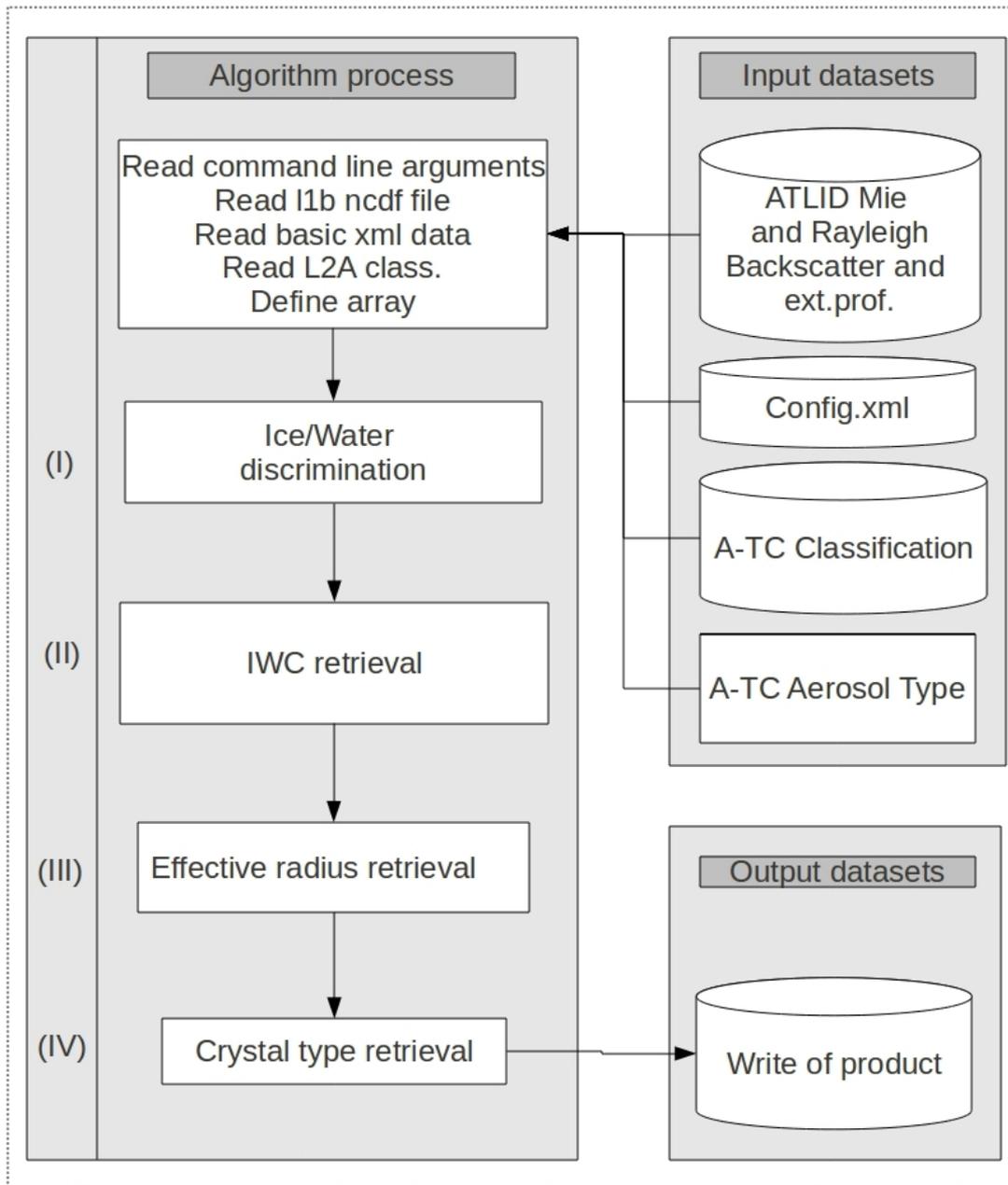


Figure 22: Schematic of the A-ICE algorithm.

The main steps involved in the algorithm are shown in .

- I) The process of ice-water discrimination is not described further in this document. This procedure is described in detail in a separate document [ATLID Target classification, A-TC ATBD]. From this product, the **Simplified classification** variable is used to localize ice layers. For the

purposes of the IWC algorithm, an Ice/Water mask (noted IWM) is constructed, and is defined by setting IWM=1 where the input target mask indicates the presence of ice clouds and 0 otherwise.

- II) The IWC product is the result of a simple parametrization :

$$IWC(k) = C_0 \frac{\alpha(k)^{C_1}}{1000} \cdot IWM(k) \quad (9)$$

Where the IWM parameter represent the Ice/Water mask, and the parameters  $C_0$  and  $C_1$  are respectively equal to  $C_0 = 89 + 0.62204 \cdot T(C)$  and  $C_1 = 1.02 - 0.00281 \cdot T(C)$ .

Additionally, given the uncertainty in the input extinction, the corresponding error in the IWC estimate is calculated.

- III) The effective radius  $R_{eff}$  product is produced by the simple parameterization :

$$R_{eff}(k) = C \cdot \left( \frac{IWC(k)}{\alpha(k)} \right) \cdot IWM(k) \quad (10)$$

where  $C=1.64$ ,  $k$  the considered altitude level, and IWM the Ice/Water mask. Also included in this step is the calculation of the uncertainty in the effective radius estimate.

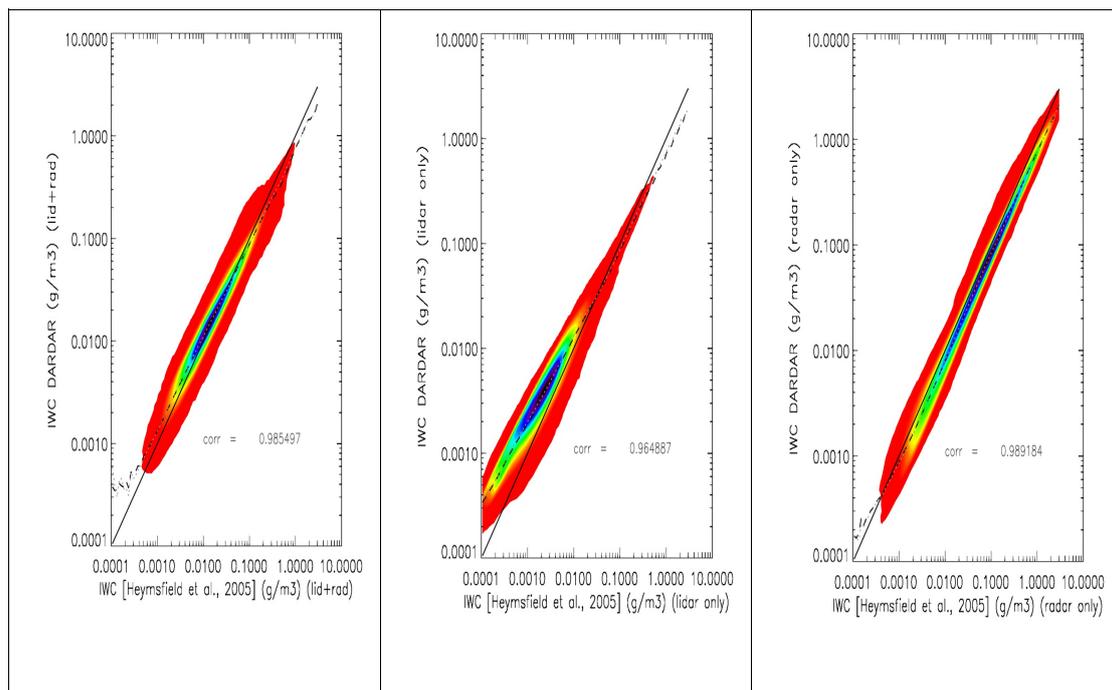
- IV) The determination of the ice crystal type has not yet been included.
-

### 4.5.3. Example

To evaluate the used parametrization the retrieved results are compared to ones that have been retrieved in previous works.

The only existing dataset that can provide this type of information based on space based data is given by the DARDAR-CLOUD product [see Delanoe and Hogan, 2010]. This algorithm uses a variational method for retrieving profiles of visible extinction coefficient, Ice water content and effective radius in ice cloud using the combination of radar reflectivity, lidar attenuated backscatter and Infrared Radiances in the water-vapour window.

One entire day has been retrieved (8<sup>th</sup> November 2009) in this study. This amount of data is sufficient to reach a good statistics on the considered parameters. Example results are shown in Figure 23. Here the DARDAR-CLOUD values of IWC are compared to the Heymsfield parameterization results using the DARAR retrieved extinction values. Here it can be seen that when radar data is available the DARAR retrieved IWC values are well correlated to the values obtained by applying the Heymesfield parameterization. However, for cloudy regions where only lidar observations are available (i.e. thin low-reflectivity cirrus) the results begin to diverge with the DARAR results becoming increasingly higher with respect to the Heymesfield values with lower IWC values. The reason for this is unclear at this point.



**Figure 23: 2D-Histogram of correlation between IWC as given by the Heymsfield parametrization, against the product as given inside the DARDAR product : From left to right (1) Lidar + Radar pixels, (2) Lidar only pixels, (3) Radar only pixels.**

#### **4.5.4.        *Algorithm Status***

The prototype algorithm has been implemented as an IDL routine and is not integrated into the ECSIM environment. The application of the IWC-Extinction parameterization to the DARDAR extinction output shows that the resulting IWC values are largely consistent with the DARDAR existing IWC product. However, , at this point unexplained, inconsistencies may be present for low reflectivity (low IWC) clouds.

#### **Future Validation and Development needs**

A full error assessment is needed in the retrieval for all the retrieved parameter. The evaluation of IWC and effective radius estimates will require comparisons with aircraft borne in-situ measurements.

The possibility of retrieving the ice crystal type must be determined also in the future. This work will likely evolve after the launch of the EarthCARE mission.

---

## **4.6. ATLID-CPR-MSI Synergetic Target Classification (ACM-TC)**

### **4.6.1. Overview**

This product facilitates the application of other synergetic algorithms by assigning a target classification to each of the pixels. A “target classification” field indicates the occurrence of the following types of targets, or combinations thereof: liquid water droplets, ice particles, rain or drizzle drops, aerosol particles detectable by the lidar, molecular scattering detectable by the lidar (or put this in “detection status”), insects detectable by the radar (if possible), radar surface echo, lidar surface echo. As the individual instruments fail to see the entire atmosphere the synergetic combination will give a superior classification compared to a single instrument classification.

The algorithm is based on the CloudNET classification [Hogan and O’Conner 2004] and the DARDAR-mask algorithms [Delanoe and Hogan 2010 & <http://www.icare.univ-lille1.fr/projects/dardar/>] and consequently the algorithm is defined as a decision-tree. Due to the lack of other type of mature algorithms. The algorithm deals mostly with the individual instrument classifications from the L2a data-streams, only in some cases the backscatter signals from radar or lidar are required (e.g. in the case of rain).

#### **Major Inputs:**

- 1) Lidar Classification (A-TC)
- 2) Radar Cloud Mask (C-CM)
- 3) L1d lidar, radar and MSI signals.
- 4) L1d grid
- 5) Lidar Extinction & Backscatter (A-EBD)
- 6) MSI Cloud Mask (M-CM)
- 7) ATLID-MSI Aerosol Column Descriptor (AM-ACD)

#### **Major Outputs:**

- 1) Synergistic lidar-radar-MSI classification (ACM-TC)
- 2) Pixel detection status

#### **Output Resolution:**

The output is reported on the combined L1d grid (1km horizontal and ~100m vertical).

---

#### 4.6.2. Algorithm Flowchart and description

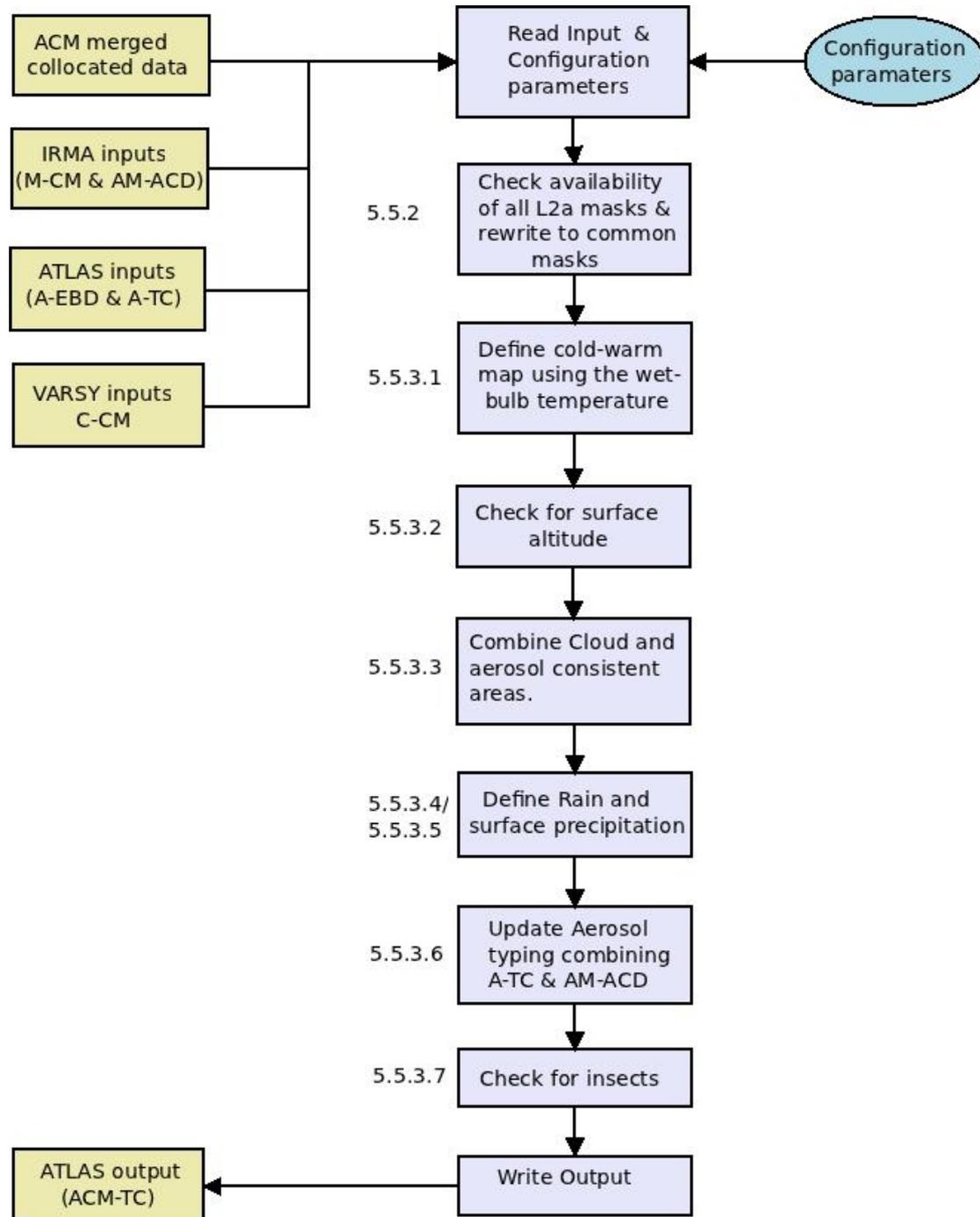


Figure 24: Flow diagram of the ACM-TC algorithm. Yellow boxes indicate input and output files, the oval reflects the configuration parameter file and the grey boxes the steps within the algorithm. The number next to the grey boxes refer to the section numbers in the ACM-TC-ATBD in which the topic is explained.

The algorithm steps are schematically depicted in Error! Reference source not found..

At the start of the procedure the availability of the different data streams is checked. For each of the L2a target masks there will be an instrument detection status ( $A_{\text{det}}$ ,  $C_{\text{det}}$  and  $M_{\text{det}}$  for the ATLID, CPR and MSI respectively). This status will indicate if the instrument was working (status > 0). For all cases where the status of one of the instruments is 0:

- Only those decisions which are not based on this instrument are used.
- Those pixels for which decisions cannot be used will result in a classification based on the remaining instruments. I.e. in the case of missing lidar or radar data the classification will be based on the radar respectively lidar data only. The MSI may infer that there are aerosol regions or retrieve cloud phase but there is no vertical information to place the undetected layers. In the case of missing MSI data the mask would remain the same except for a few small changes in the aerosol typing.

Next the combined classification can be derived. Most of the pixels will be defined by the lidar and radar only. In the case of the aerosol typing and possibly stratospheric clouds the wavelength information (Angstrom component) will be added to the typing information. This will result in a more detailed typing during day-time only. In Table 1 the combined classification results from the different combinations from the lidar and radar inputs are presented. These results reflect the 'simple categorization' as defined in the PDD. From this the complete categorization can be defined by filling in the sub-classifications (e.g. aerosol classification and rain classification).

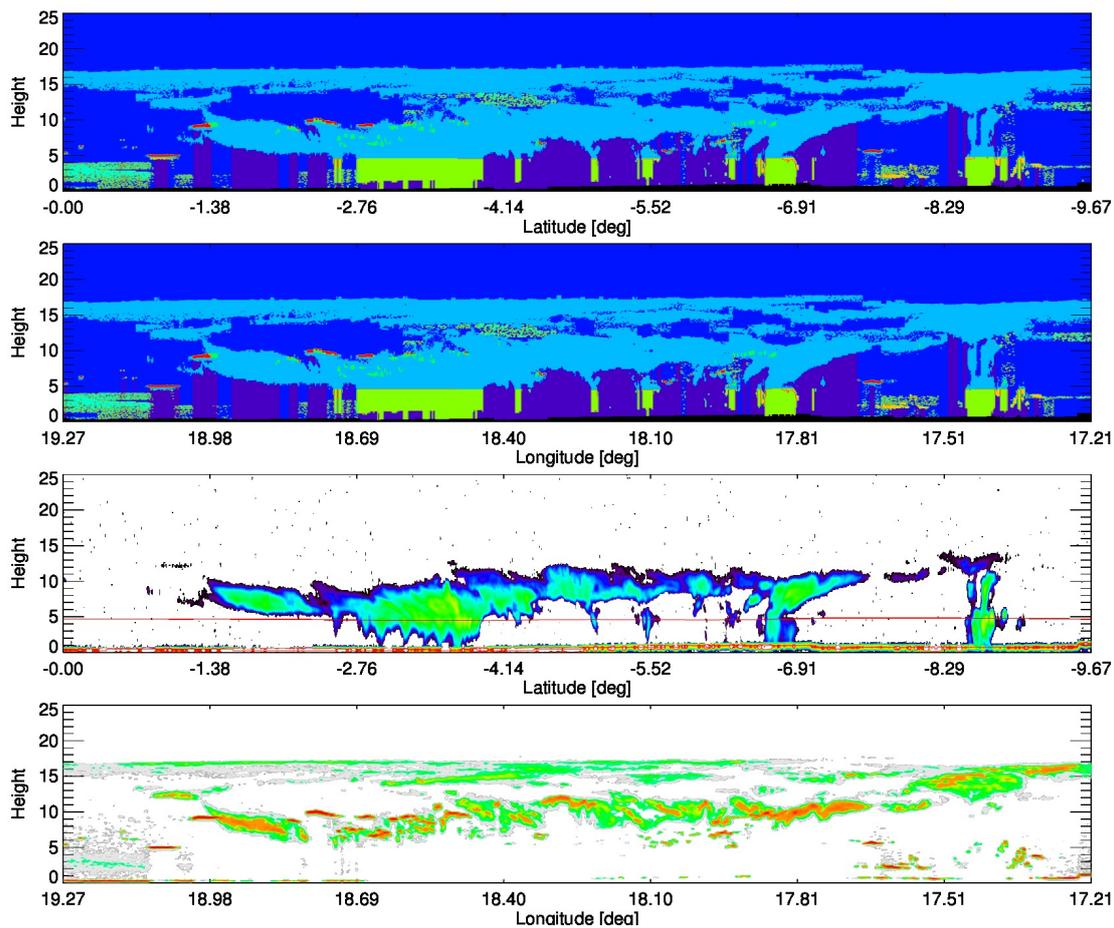
**Table 1: Summary of the main classification rules used to combine the L2a ATLID and CPR classification input masks (Columns #2 and #3) in to the combined classification (column #1). The final column shows the additional rules needed to combine the information, where  $T_{wb}$  reflects the wet-bulb temperature,  $T$  the temperature,  $C_{Z_{max}}$  the maximum reflectivity in the warm column,  $z_{TP}$  the tropopause height and  $C_Z$  the local radar reflectivity.**

L2b classification	Lidar	Radar	Additional Info
Surface	Surface or unknown	Surface or unknown	ECMWF surface
Clear	Clear	Clear/unknown	-
Liquid	Liquid	Clear/cloud/unknown	$T_{wb} > 273 \text{ K}$
Liquid	Unknown	Cloud	$T_{wb} > 273 \text{ K}$
Ice	Ice	Cloud or Clear	$T_{wb} < 273 \text{ K}$
Ice	Unknown	Cloud	$T_{wb} < 273 \text{ K}$
Supercooled	Liquid	Clear	$T_{wb} < 273 \text{ K}$
Supercooled + ice	Liquid	Cloud	$T_{wb} < 273 \text{ K}$
Aerosols	Aerosols	Clear	-
Rain	Unknown	Cloud	$C_{Z_{max}} > -17 \text{ dBZ}$ , $T_{wb} > 273 \text{ K}$
Liquid cloud +Rain	Liquid	Cloud	$C_{Z_{max}} > -17 \text{ dBZ}$ , $T_{wb} > 273 \text{ K}$
Insects	Liquid	Cloud	$C_Z < -20 \text{ dBZ}$ , $T > 283 \text{ K}$
Stratosphere	Stratospheric	Cloud or unknown	$z > z_{TP}$
Unknown	Unknown	Unknown	-

The only information needed beside the input classifications are the separation between cold and warm regions [reflected in the Table by the wet-bulb temperature ( $T_{wb}$ )]. This is needed for separating ice from rain and in order to distinguish the liquid cloud from ice cloud regimes. The wet bulb temperature is always lower than the dry bulb temperature ( $T$ ) but will be identical with 100% relative humidity (the air is saturated). The reason for adopting  $T_{wb}$  is to indicate where “falling” particles are likely to be composed of ice rather than liquid as falling ice melts when  $T_{wb}$ , rather than  $T$ , becomes positive.

It is assumed that  $T_{wb}$  will be provided through the ECMWF model output, if this is not available the temperature can be easily computed using the available temperature, pressure and humidity. This requires a number of steps in which the dew-point temperature, environmental vapour pressure,  $e(T)$  and the saturation vapour pressure  $e_s(T)$  are calculated.

### 4.6.3. Example



**Figure 25: ACM-Target Classification example using CloudSAT and Calipso data measured on 2010-04-19. The top figure shows the result of the ACM-TC, the second is the DARDAR mask, third the CloudSAT radar data (the read line indicates the  $T_{wb}=0^{\circ}\text{C}$  isotherm) and the bottom figure the 532nm CALIOP data. This scene shows a large number of different regions (light blue: ice clouds, green: rain, orange: liquid layer, red: super-cooled liquid, green within the blue: super-cooled liquid and ice and darker green: aerosols). Note that the two classification show very similar results**

In order to evaluate the current version of the algorithm CloudSAT and CALIPSO data will be used. The direct use of these L1b data streams would involve a large effort in order to collocate the two instrument signals. J. Delanoe performed this work for his DARDAR-mask algorithm and makes this available on the ICARE data base (<http://www.icare.univ-lille1.fr/projects/dardar/>). As this algorithm is based on the experience from the DARDAR mask using this data-set as an input to the ACM-TC algorithm will also directly enable to compare the results to each other.

In Figure 25: ACM-Target Classification example using CloudSAT and Calipso data measured on 2010-04-19. The top figure shows the result of the ACM-TC, the second is the DARDAR mask, third the CloudSAT radar data (the  $T_{wb}=0^{\circ}\text{C}$  isotherm) and the bottom figure the 532nm CALIOP data. This scene shows

a large number of different regions (light blue: ice clouds, green: rain, orange: liquid layer, red: super-cooled liquid, green within the blue: super-cooled liquid and ice and darker green: aerosols). Note that the two classification show very similar results part of an orbit on 19 April 2010 (DARDAR: 2010109002854\_21140) is presented. This scene shows most of the possible classification types possible. There is a very thin cirrus top layer missed completely by the radar, when the particles grow big enough there are many super cooled liquid layers (red and thin green thin layers when there is both super cooled layers and ice clouds). There are a number of rain events with a few reaching the surface and on the far left and top right two aerosol layers. The DARDAR mask and ACM-TC classification are very similar throughout the scene, there are a few differences, especially in the height of the warm-cold mask, but these are near impossible to see in the large scale as presented in the image. The similarity is not surprising since the ACM-TC algorithm is based on the DARDAR and CloudNET schemes and uses the same single instrument target classifications as an input.

#### **4.6.4.            *Algorithm Status***

The ACM-TC scheme described here exists in prototype form and is not at the time of this writing (May 2011) integrated into ECSIM. Since the procedures themselves are applied by directly comparing different input streams it is expected that the brute-force but simple algorithm will be fast enough for operational use without any special developments.

The algorithm uses the DARDAR-mask files as an input and has therefore only been tested on CloudSAT and CALIOP data and not data based on the EarthCARE instrument settings.

#### **Future validation and Developments needs:**

First of all the algorithm has to be integrated into ECSIM, while keeping the DARDAR input data possibility open. So far only 1 CloudSAT & CALIPSO orbit has been retrieved. A large number of orbits (day and night) would have to be retrieved in order to validate if the above set-rules result in consistent target classification masks for all different seasons. This evaluation should be performed statistically and by manually checking the individual profiles.

A combined detection status flag has to be created based on the already available single instrument detection status.

A future version of the algorithm will have to be evaluated with ECSIM scene signals and all input files will have to be created by realistically retrieving the different L2a algorithms input data. A true validation will have to be performed by using real EarthCARE data and, if possible, related campaign data using a HSRL UV lidar and 94GHz radar.

---

## 5. L2 System Support (WP4)

In this section, the activities mainly carried out under WP4 of the ATLAS project. The exception is the presentation of extra material related to the evaluation of the analytical lidar Multiple-Scattering model used in the A-EBD algorithm.

### 5.1. ECSIM code maintenance

As ECSIM [ECSIM] is used by more users for applications outside of the initial testing scenarios it is inevitable that new features are requested and that bugs will be found and fixed. As part of the activities within WP 4 several features have been added to ECSIM and several bug fixes have been made. Most of these issues have been confined to the lidar forward and instrument models. The new features and bug-fixes have been made available to ESA.

In Summary, the most important changes have been:

- The addition of extra background information to **lid\_filter** and **lidar**. In particular, look-up tables have been built so that the lidar background can be quickly and automatically estimated as a function of surface type, cloud optical depth and solar position.
  - The addition of extra variables to the output of **lid\_filter** (As requested by ASTRIUM).
  - The finding and fixing of various background related bugs in lidar.
  - The finding and fixing of bugs related to the calculation of multiple-scattering induced depolarization in **lid\_filter**.
  - Extending the capabilities of the **extract\_quantity** tool.
  - Improved treatment of the surface reflections in **SW\_rad**.
  - Improved treatment of the surface reflections in **lid\_filter**.
  - Assisting in the production of an updated ECSIM release (V1.5).
-

## **5.2. Performance evaluation comparing ATLID-UV, ATLID-Green and CALIPSO optical depth retrievals.**

Due to technical problems encountered during ATLID development other alternatives to the 353 nm HSRL design have been considered, namely, the possibility of implementing a 532 nm HSRL design. As part of WP 4, a performance evaluation using a number of simple scenarios was carried out comparing the old baseline 353 nm HSRL design with a preliminary 532 nm concept. The 532 nm concept was roughly comparable in terms of the product of the power, collection area and collection efficiency when compared to the UV system. For completeness, the performance of the CALIPSO lidar was also considered. Four difference scenarios were considered.

**Case 1:** A homogeneous cirrus layer between 8 and 9 km. The signals were simulated using ECSIM and the total OT estimated by comparing the Rayleigh returns above and below the cloud layer (averaged vertically over 1 km). The simulation was conducted assuming a range of optical thickness and horizontal averaging lengths and assuming night-time conditions.

**Case 2:** Same as Case 1, except that daylight (solar zenith angle of 60 Deg) and an ocean surface was considered.

**Case 3:** A homogeneous aerosol layer from the ground to a height of 2km. In the case of the HSRL designs, the total aerosol OT was estimated by comparing the return above the layer (averaged over 1-km) with the Rayleigh signal averaged over the lowest 300 meters. For the simulated CALIPSO lidar signals, the retrievals were performed using a forward inversion assuming no calibration error but assuming a 50% uncertainty in the assumed backscatter-to-extinction (S) ratio.

**Case 4:** As Case 3 except that daytime conditions were assumed.

Examples corresponding to Case 1 are shown in Figure 26 and Figure 27. Here it can be seen that the retrieval cirrus OT is much more robust for the UV HSRL case than for the Green HSRL concept. Example results for the UV and Green HSRL concepts are summarized in Figures 28 through Figure 31 for the cirrus cases. Similar results but for the retrieval of aerosol backscatter were also obtained. For the aerosol cases the GHRL and UV baseline results were much closer.

The results of this study were presented at the Dec 2009 JMAG meeting. The main conclusions that can be drawn from this exercise are may be stated as follows.

In spite of improved SNR in the Mie channel in the cases of both aerosol and cirrus clouds, the preliminary Green HSRL lidar concept is markedly inferior when it comes to the retrieval of aerosol and cirrus cloud extinction. As a result, the scientific utility of the Green HSRL concept evaluated in this work is not suitable given the scientific goals of EarthCARE. In order to be suitable for EarthCARE a green HSRL system would have to possess a power-aperture product several times greater than the concept evaluated here. Arguably, the Green HSRL concept would be inferior to the existing

---

CALIPSO lidar.

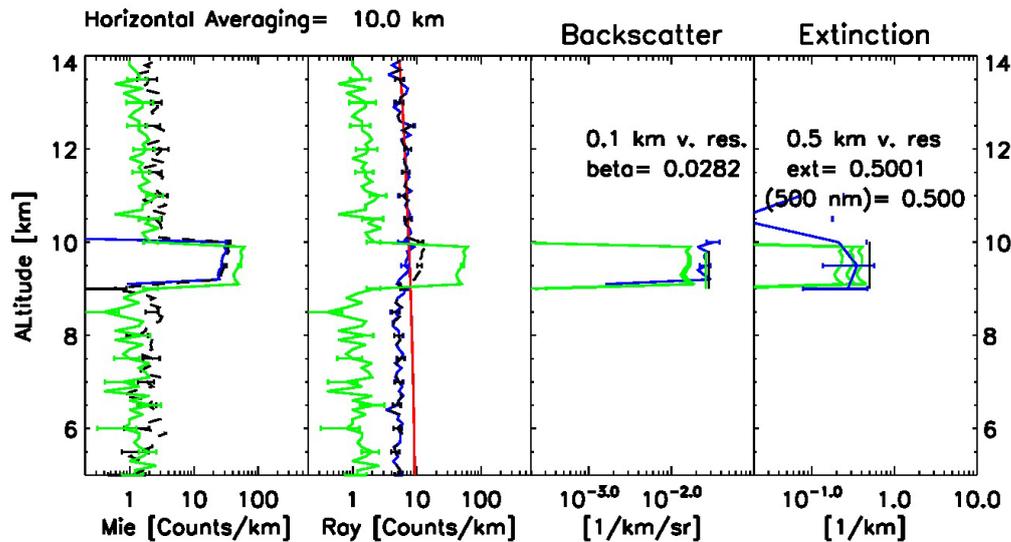


Figure 26: Example CASE 1 results for the UV HSRL concept. The green lines show CALIPSO lidar simulation results. The Red line in the second panel from left show the Rayleigh return in the absence of cloud. The Blue lines in the first two panels denote the cross-talk corrected signals and the black-lines show the signals before cross-talk correction.

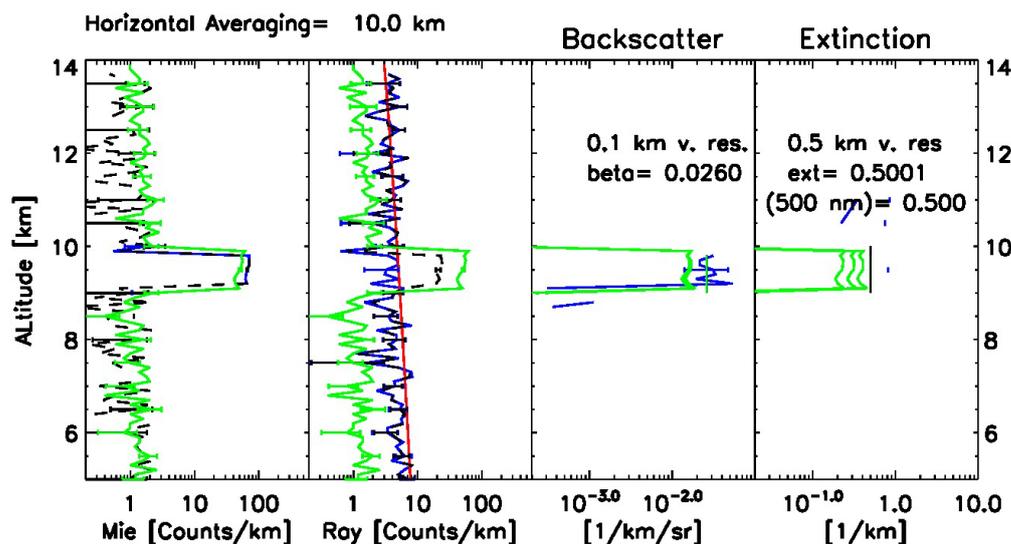


Figure 27: Example CASE 1 results for the Green HSRL concept. The green lines show CALIPSO lidar simulation results. The Red line in the second panel from left show the Rayleigh return in the absence of cloud. The Blue lines in the first two panels denote the cross-talk corrected signals and the black-lines show the signals before cross-talk correction

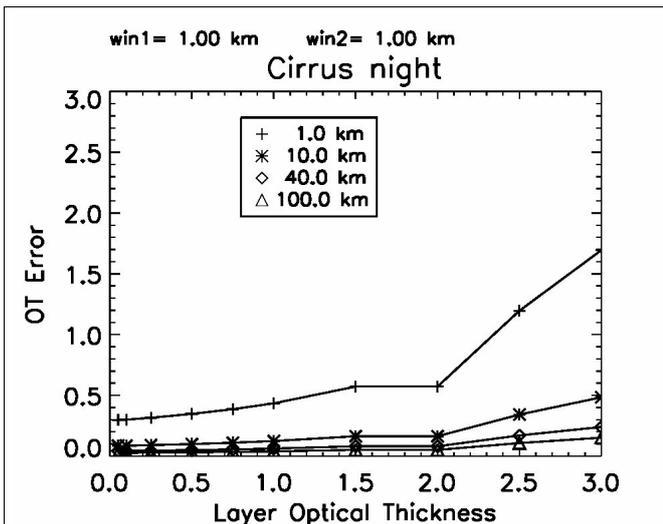


Figure 28: Statistical Error in cirrus OT retrievals for the UV HSRL concept. Here  $iwin1$  denotes the vertical averaging interval above the layer while  $iwin2$  denotes the below layer interval.

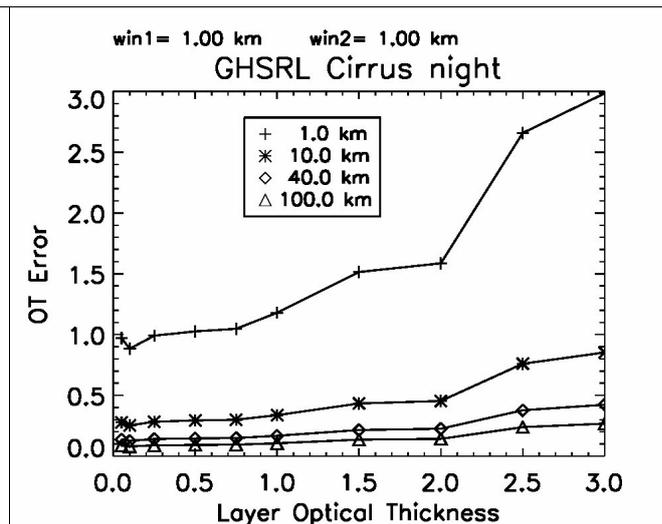


Figure 29: Left panel except for the Green HSRL concept.

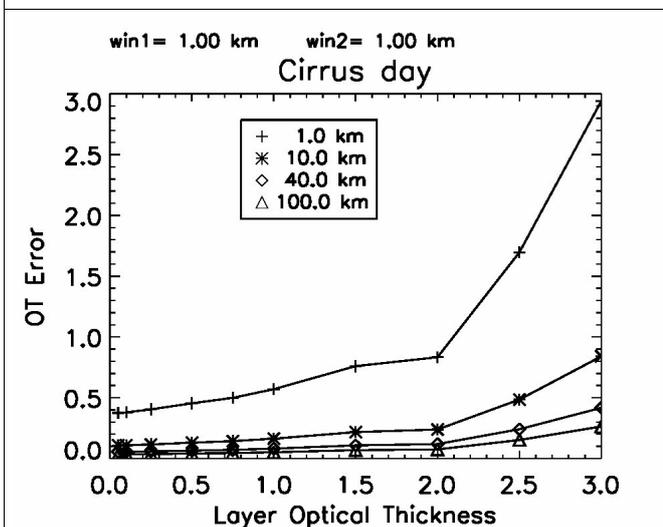


Figure 30: As above except for Daylight conditions.

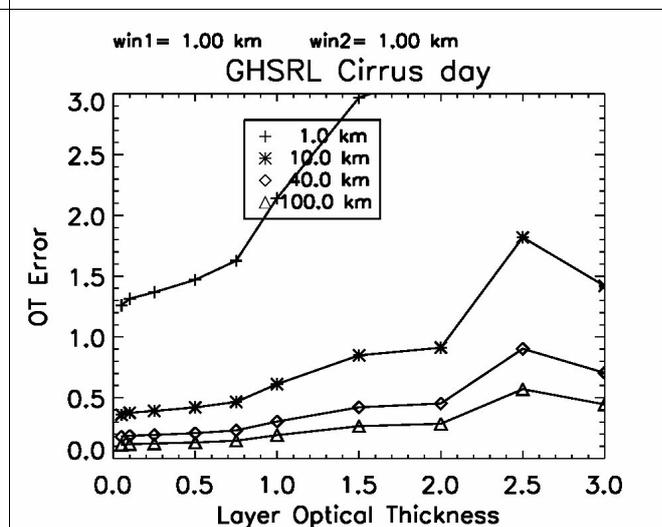


Figure 31: As above except for daylight conditions.

### 5.3. On the Use of Linear vs Circular Polarization

Changes to the proposed design of ATLID means that a bi-static configuration instead of a mono-static configuration is being considered. Thus, the technical justification for using circularly polarized light may become irrelevant and the opportunity to reconsider the relative merits of using linear as opposed to circularly polarized light has presented itself.

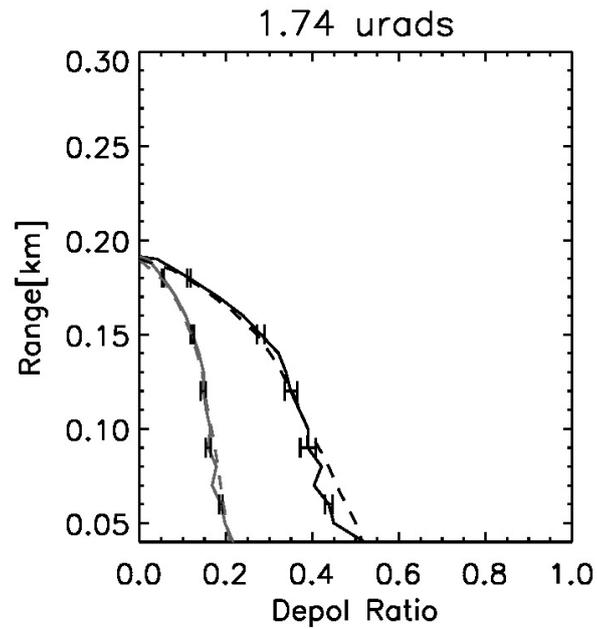
For randomly oriented cloud and aerosol particles which possess a plane of symmetry a well-defined relationship between the circular and linear depolarization ratios exist. In particular,

$$\delta_c = \frac{2\delta_l}{1-\delta_l}. \quad (11)$$

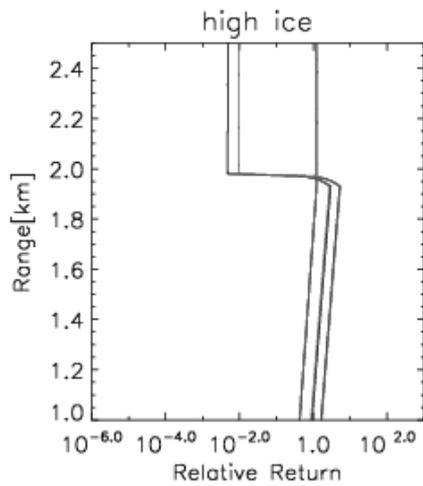
Thus, circular depolarization ratios in clouds are expected to be larger than the corresponding linear values. This means that smaller depolarization ratios associated with aerosol may be easier to measure than linear ratios. With regards to the calculation of circular and linear depolarization ratios using ECSIM additional validation work has been performed in order to verify the correctness of the calculation of the polarization state of the lidar return. In particular, comparisons with independent published theoretical and observed results (i.e. see Figure 32)

In spite of the fact that circular depolarization ratios are larger there may be advantages to using circular depolarization with respect to water and ice phase discrimination on the basis of depolarization ratio. This is illustrated in Figure 33- Figure 36 where the circular and linear depolarization ratios that may be expected from a thin ice cloud and an optically thicker water cloud are compared. For the water cloud the depolarization signal is due to multiple-scatter processes. Thus the depolarization ratio increases with increasing penetration depth into the cloud. By comparing the water cloud and ice cloud depolarization values, it can be seen that the water cloud linear depolarization ratios closely approach those expected for ice clouds. However, a larger separation in circular depolarization ratios between the ice cloud and water cloud is maintained. This indicates that the determination of cloud phase solely using depolarization ratio should be more robust in the case where circular depolarization is used. However, it should be noted that in the case of linear depolarization. Using the ratio of depolarization to backscatter should also be a robust method for ice/water discrimination.

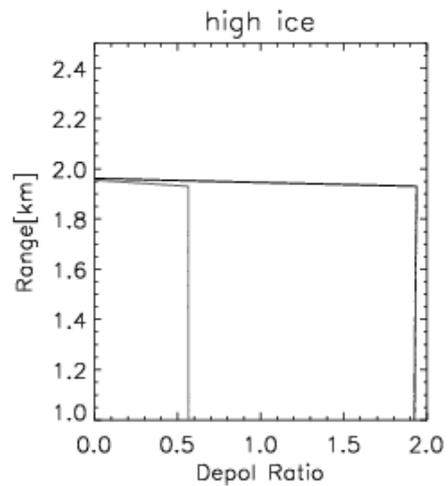
---



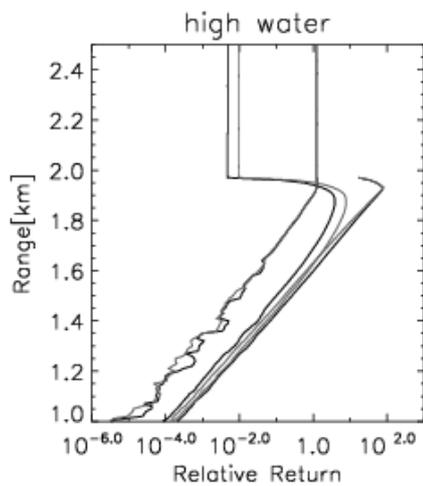
**Figure 32: Linear (Grey) and circular (Black) depolarization ratios for a water cloud 2 km from the lidar. The dotted lines show theoretical results presented in Fig 19 of Roy and Roy (2008), while the solid lines show ECSIM Monte-Carlo results for a roughly comparable situation.**



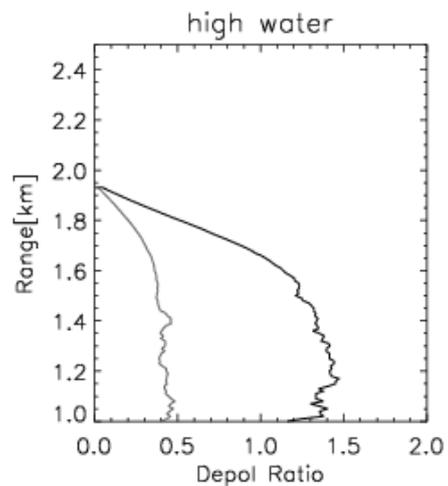
**Figure 33: Mie and Rayleigh returns for an ice cloud. Black, circular depolarization. Grey linear depolarization.**



**Figure 34: Ice cloud depolarization ratios. Black, circular depolarization. Grey linear depolarization.**



**Figure 35: Mie and Rayleigh returns for an water cloud. Black, circular depolarization. Grey linear depolarization.**



**Figure 36: Water cloud depolarization ratios. Black, circular depolarization. Grey linear depolarization.**

#### **5.4. Investigations into the laser-spot image.**

Also related to possible changes in the ATLID design, within WP4 a specific study was initiated in response to the possible shift in ATLID instrument concept from a mono-static design to a bi-static design. The impact of this shift is greatest with respect to the ATLID system transmitter-receiver alignment. With the mono-static concept, since much of the optical path is common between the transmit and receive paths proper alignment was “guaranteed”. With the bi-static system much of the commonality between the receive and transmit paths is lost and an active beam steering mechanism is deemed to be necessary. As part of this system a camera will be fitted to one of the detection channels. This camera will image the laser spot so that action can be taken to maintain the alignment between the laser and receiver fields of view.

An obvious issue connected with this issue is the form the spot image will take under different conditions (polarization state, cloud optical depth, surface type etc.). An “unofficial” modified form of the ECSIM **lid filter** module was created. This version was modified to calculate the integrated “image” that would be seen from an alignment camera. Sample results are shown in Figures 16 to 18. Standard EarthCARE conditions were assumed (i.e. orbit of 400km) using homogeneous water clouds present from 1.0 to 2.0 km with an effective particle radius of 10 microns and varying optical thicknesses. The images were obtained by time integrating the lidar signal (in terms of attenuated backscatter coefficient) from the top of the atmosphere to (including) the surface level such that the images are in units of 1/sr.

Here it can be seen that for the case of linear polarization that the image is dependent on the azimuth (clear in the cross-polar image but not as obvious in the co-polar case) while, as expected, there is no azimuthal dependence for the case where circular polarization is used. It can also be seen that the image intensity beyond 1 micron increases with optical depth up to about 10 but then decreases for higher optical thicknesses. This may be physical and due to the effect of increasing optical extinction of the clouds limiting the degree of horizontal transport, or, it may be an artifact due to the limited number of scattering orders used in the calculation (10).

An example where a bright surface was assumed is shown in Figure 19. Here a Pinty-RPV BRDF surface model was assumed with the parameter values for a “snow” surface were chosen. Here, In comparison with the dark ocean surface cases the “tails” in the case of cloud optical thicknesses less than 2.0 are more pronounced. This may be due to the fact that the beam reflectance of the snow surface is more diffuse than the ocean nadir reflectances and the tails receive a significant contribution from surface scattered photons for low cloud optical thicknesses.

With regards to the surface properties, in this study it has been assumed that the surfaces completely depolarize (known to be incorrect) the signal and the snow BRDF has not been validated against actual 353 nm lidar measurements. However, in all

---

cases examined, it can be seen that the vast bulk of the spot intensity is confined to the beam itself. This seems to be true for a range of plausible surface depolarization characteristics and surface return magnitudes. This result indicates that multiple scattering processes should not pose a significant problem when determining the beam center.

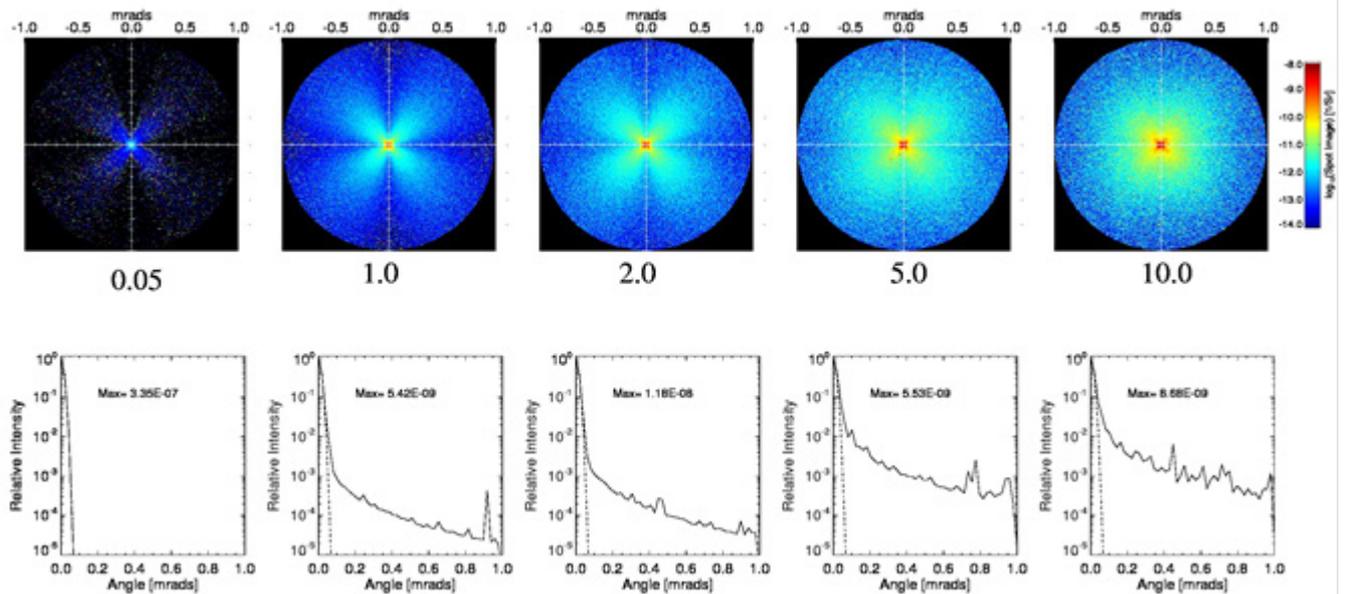


Figure 37: (Top) laser spot image as a function of angle for different homogeneous water cloud optical depths (from 0.05 to 10.0 left to right). Here linear polarization was transmitted and the cross-polar intensity measured. An ocean surface was specified. (Bottom) Average relative image intensity as a function of angle. The dashed line shows the angular envelope of the transmitted (Gaussian) laser beam. The Max value labels in the lower panels show the value of the peak in 1/sr.

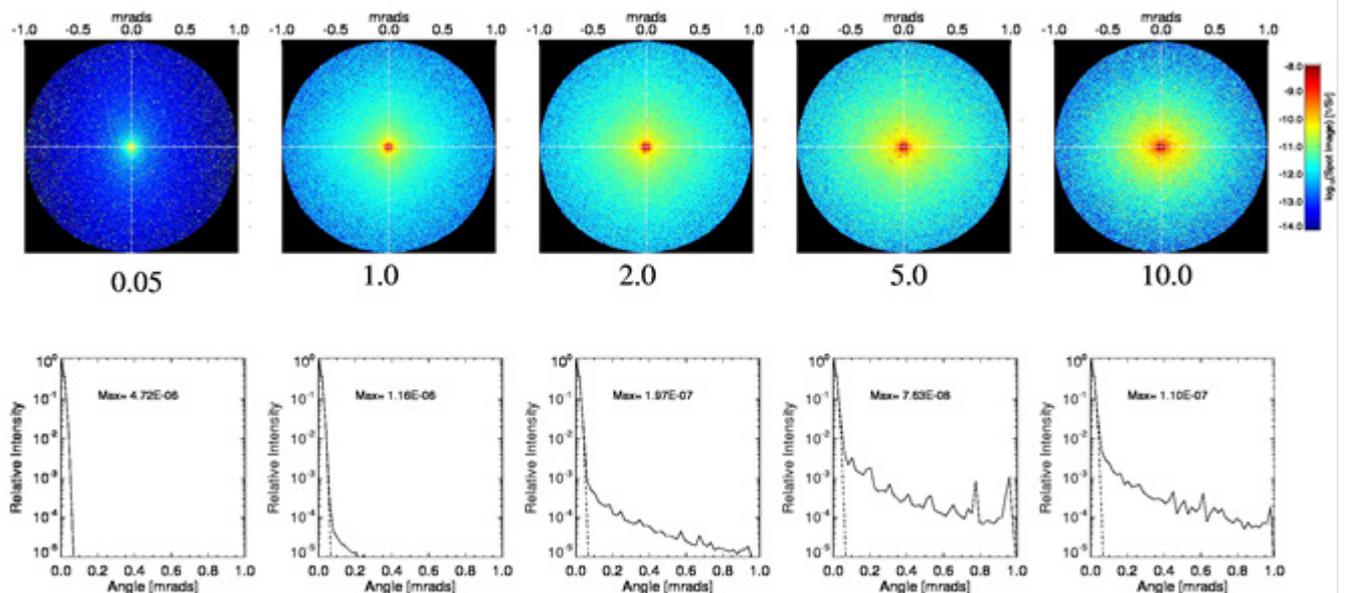


Figure 38: As Figure 35 except that the co-polar signal is measured (linear co-polar).

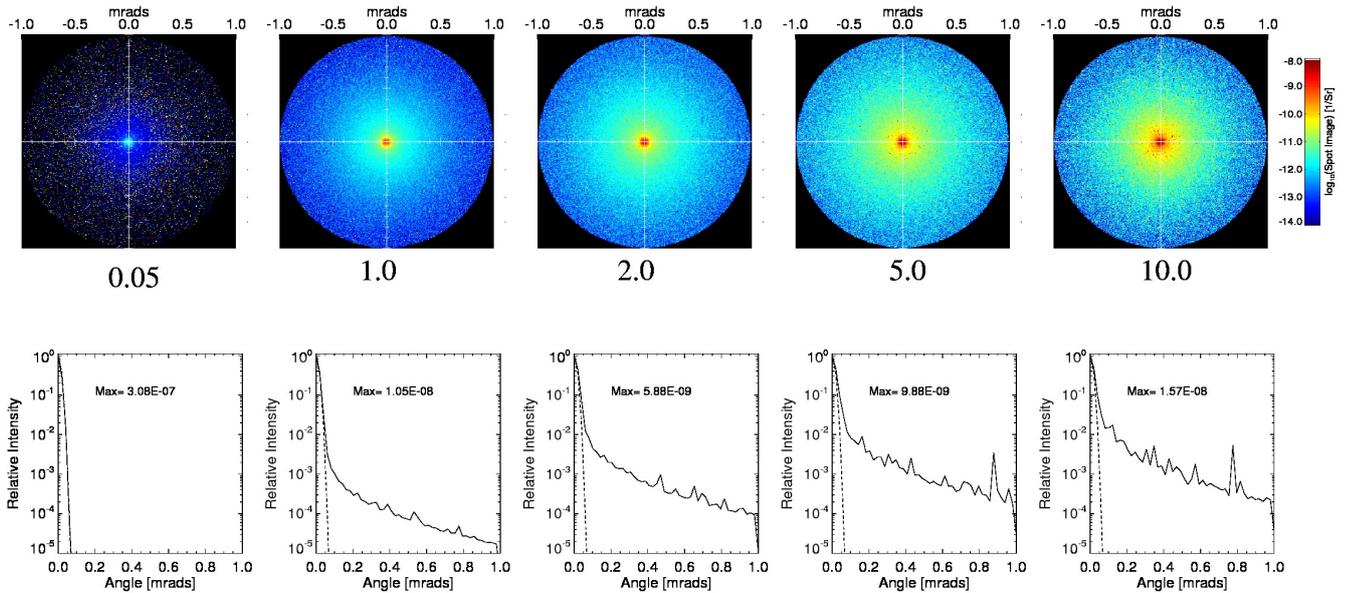


Figure 39: As previous except that circular polarization has been used (circular co-polar).

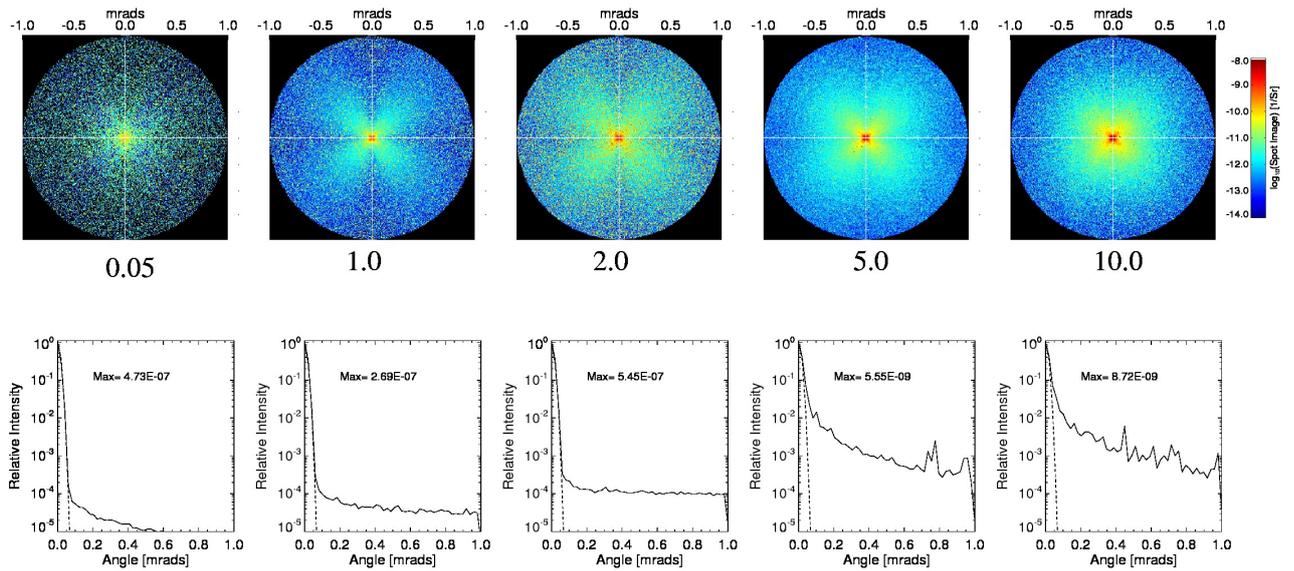


Figure 40: As Figure 35 except that a snow surface has been used.

### **5.5. Evaluation of Hogan's Multiple-scattering Model.**

In order to retrieve accurate quantitative information from ATLID the effects of Multiple-Scattering (MS) must be taken into account. Monte-Carlo methods can be very accurate but are, in general, too computationally expensive to be incorporated into routine inversion algorithms. Thus, one must appeal to less accurate but much faster analytical approaches. One such approach (the one used in the A-EBD algorithm) is the parameterized approach due to [Hogan, 2006].

As part of ATLAS a systematic evaluation of the accuracy of Hogan's approach has been conducted covering a range of idealized cloud scenes using ECSIM. Sample illustrative comparison results are shown in Figure 41 through Figure 44 for water clouds and Figure 45 through Figure 48 for ice clouds. Here  $R_a$  is the extinction-weighted equivalent area radius. It can be seen that in general the model of Hogan can, in general, well can reproduce the effect of MS on the signals. However, it can also be seen that while the choice of particle size does not greatly influence the effect of MS within the cloud (i.e. Mie channel signals) that below the cloud the choice of cloud particle size has a substantive effect. Further, it can also be seen that for large particle sizes, it appears that the MS approach of Hogan systematically underestimates the effect of MS of the signal. It appears that this effect cannot be ameliorated by varying the cloud particle size (unlike the case with the below cloud Rayleigh return). The reason for this is unclear at this point and may be a matter of resolution or other non-fundamental issue.

Based on the result of the MS analysis carried out, it appears that using a fixed particle size as is used in the CASPER algorithm is not a sufficient assumption. Thus, within the ATLAS algorithm a mean particle size per layer is added to the state vector in order to better account for MS effects. Indeed, recent work based on the CASPER algorithm has shown that the algorithm may fail due to the inability to accurately model the Rayleigh signal return below cloud base.

