

ATLAS

ATLID Algorithms and Level 2 System Aspects

Algorithm Theoretical Basis Document (ATBD) for A-Classification (A-TC ATBD)

Code : EC-TN-KNMI-ATL-022
Issue : 2.2
Date : 24/06/2011
Authors : G-J van Zadelhoff
D.P. Donovan
S. Berthier

This page intentionally left blank

Table of Contents

1. Purpose and Scope	8
2. Applicable and Reference Documents	9
2.1. Applicable documents	9
2.2. Reference & Related documents	9
2.3. Scientific References	9
3. Scientific Background of the algorithms	11
3.1. Algorithm history	11
3.1.1. L2a Lidar classification	11
3.1.2. Aerosol Typing	11
3.2. Algorithm introduction	11
3.2.1. L2a Lidar classification	11
3.2.2. Aerosol Typing	11
3.3. Physical/Mathematical Background	12
3.3.1. L2a Lidar Classification	12
3.3.1.1. Detection of significant height boundaries	12
3.3.1.2. Cloud-Aerosol identification	13
3.3.1.3. Water-cloud/Ice-cloud separation	13
3.3.2. Aerosol Typing	17
3.3.2.1. Use of additional information	19
4. Justification for the selection of the algorithm	22
4.1. L2a Lidar Classification	22
4.2. Aerosol Typing	22
5. Mathematical algorithm Description	23
5.1. Input parameters	23
5.2. Configuration parameters	25
5.3. Output parameters	26
5.4. Algorithm flow charts	29
5.4.1. L2a Lidar Classification	29
5.4.2. Aerosol Typing	31
5.5. Algorithm Definition	32
5.5.1. L2a lidar classification	32
5.5.1.1. Detection of significant layer boundaries	32
5.5.1.2. Aerosol cloud discrimination and water cloud/ice cloud discrimination procedure	34
5.5.2. Aerosol Typing	36
5.5.2.1. Input of the data and configuration parameters	36
5.5.2.2. Extracting slices from available aerosol maps	37
5.5.2.3. Stratosphere - troposphere separation	37
5.5.2.4. Defining the Aerosol probabilities	37
5.5.2.5. Calculating the type probabilities	39
5.5.2.6. Use of additional information	40
5.5.2.7. Assignment of aerosol type	40
6. Algorithm performance, sensitivity studies, limitations	43
6.1.1. L2a lidar classification	43
6.1.2. L2a lidar aerosol typing	47

6.1.2.1. Saharan desert scene (SAMUM1)	47
6.1.2.2. Aged Saharan dust with Biomass burning (SAMUM2)	48
6.1.2.3. Anthropogenic pollution (EUCAARI).....	48
7. Validation status.....	49
7.1.1. L2a lidar classification.....	49
7.1.2. Aerosol Typing	51
Annex A: Technical implementation	51
External models	51

List of Tables

Table 1: Input data parameters L2a Lidar Classification.....	23
Table 2: Input data parameters Aerosol Typing from the <i>A-EBD and A-AER procedures</i>	24
Table 3: Aerosol probability map information	25
Table 4: Configuration parameters L2a Lidar Classification	25
Table 5: Configuration parameters Aerosol Typing	26
Table 6: Output parameters from L2a Lidar Classification procedure	26
Table 7: Operational Output parameters Aerosol Typing.....	28
Table 8: Aerosol types and the parameters defining the probability distributions for each of these aerosol types. The numbers given in this version of the ATBD are not based on observations but are for based on estimates of the values. Future validation and calibration efforts combining all available data will be needed to define realistic parameters to describe the different probability distributions. The stratospheric parameters are not set at this moment.....	38
Table 9: Rules used for defining the local aerosol type which have a high enough probability. This typing will result in a maximum combination of three aerosol types for a single observation. Whenever the S, δ combination lies outside any of the 3σ regime the type will be set to unknown (-1)	41
Table 10: Possible aerosol type assignments for the probability distributions defined in Table 8, with 5% error estimates in the retrieved lidar ratio and depolarization values as presented in Figure 16. The colors match those in the figure.	42

List of Figures

Figure 1: Schematic relationship of the algorithm described in this ATBD (red-box) with respect to other lidar-only (L2a) algorithms as well as relevant MSI and CPR synergetic (L2b) algorithms.	8
Figure 2: Sketch of the L2a Lidar classification main steps.	12
Figure 3: Sample Circular depolarization ratios (Black-line) and Linear Depolarization ratios (Grey-lines) produced by Monte-Carlo calculations for an (Top) Ice cloud and (Bottom) water cloud. Note that EarthCARE will make only linear depolarization measurements.	15
Figure 4: Water cloud integrated depolarization ratio vs integrated attenuated backscatter (Left) and (Right) vs integrated backscatter. Ice clouds would be typically limited to the upper left regions of these plots. The solid line is a fit to MC results calculated at the CALIPSO wavelength of 532 nm found earlier by Hu et al. 2007...	16
Figure 5: Histogram of depolarization-vs-attenuated backscatter values for January 2009 derived from CALIOP observations along with limits used to separate Ice, Water and Horizontally Oriented crystals (Figure taken from Hu et al. 2009). The solid Black line is the same line as shown in Figure 4.	16
Figure 6: Lidar characteristic properties for different aerosol types from DLR field measurements (ICAROHS ATBD1) preliminary results.	17
Figure 7: Type depended Aerosol probabilities for important aerosol quantities measured by HSRL in a large number of campaigns with the NASA B200 King aircraft (Ferrare et al. 2007). In case of the ATLID instrument only the top two panels can be retrieved.	18
Figure 8: Average seasonal optical depth up to 12 km based on global CALIPSO measurements (averaged in 5x5 degrees boxes) for marine aerosols (panel A) and dust (panel C) with the colour-scale presented on the right. The corresponding mean extinction at each height is shown in panels B and D respectively. The colors in the latter two profiles indicate the number of cells included in the calculation of the mean and std. deviation for all cells with a retrieved extinction.	19
Figure 9 Flexpart trajectory model results for the Eyjafjallajökull volcanic eruption in 2010 (N. Kristiansen). Forecast models results like these could be used for enabling the detection of the volcanic ash aerosol type within the model.	20
Figure 10: The ESA global detection of hot spots by ERS-2's Along Track Scanning Radiometer (ATSR-2) and Envisat's Advanced Along Track Scanning Radiometer (AATSR) from July 1996 to August 2010.	21
Figure 11: Significant layer detection procedure flow diagram	29
Figure 12: Aerosol cloud discrimination and water cloud/ice cloud discrimination procedure.	30
Figure 13: Aerosol Typing flow diagram	31
Figure 14: Aerosol probabilities for the 5 tropospheric aerosol types defined by the parameters in Table 7. The greyscales show the distributions from 1e-11 up to 1 in the centre of the distributions.	39
Figure 15: Calculating the clean continental type probability for an observed point with a lidar ratio of 30 and depolarization of 0.2. The calculation is performed at all 49 points and weighted according to the distance to the observed value [Equation (4.2)].	40

Figure 16: Normalized probability fractions for the 5 aerosol types. for three slices of depolarization. The observations are assumed to have a negligible error estimate in these cases. The line colors define the following aerosol types: Black: Marine Aerosols, Green: Pure continental, Blue: Pure smoke, light green: Dust, orange: Volcanic ash.....41

Figure 17: Detectable aerosol types in observation space assuming a fixed error estimates of 5% in lidar ratio and depolarization. The black ovals represent the 1σ contour lines for each of the individual probabilities. The different colors depict the relevant regions, within a 3σ minimum probability, e.g. the yellow region represents the 3σ dust region. All possible type combinations (incl. colours) are presented in Table 10.42

Figure 18: Top-Left: ECSIM extinction field corresponding to an idealized super cooled water layer embedded in an ice cloud. Top-Right: Corresponding backscatter signals. Bottom panels: Mie co-polar and total cross polar signals as calculated by ECSIM.44

Figure 19: Bottom Panels: Sample Backscatter Ratios and Depolarization ratios derived from a horizontal average of the data shown in Figure 17 from 0 to 1 km. The Red-lines show the retrieved optimal layering. The top panel shows the decrease in the summed goodness-of-fit parameter with the allowed number of layers.45

Figure 20: 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).....46

Figure 21: Top panel: Observed backscatter profile from the DLR Falcon HSRL. Bottom panel: retrieved aerosol type mask,.....47

Figure 22: Top panel: Observed backscatter profile from the DLR Falcon HSRL. Bottom panel: retrieved aerosol type mask. Note the different horizontal scales..48

Figure 23: Top panel: Observed backscatter profile from the DLR Falcon HSRL on May 14th 2008 during the EUCAARI campaign. Bottom panel: retrieved aerosol type mask.49

Figure 24: Left: Histogram built using CALIPSO observations taken from Hu et al. 2009. Middle: Same as Left panel of Figure 3. Right: Overlap of the other two panels.50

1. Purpose and Scope

This document describes two modules relevant to the classification of the scene. The ATBD is set up according to the EarthCARE ATBD format and the two modules will be described subsequently in the different relevant sections. 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. In practise, this means that in case of , e.g. the description of the input parameters, first a Table is presented with inputs for the main classification and secondly the Table for the aerosol typing. Both modules can be called from any of the other L2a ATLID algorithms when required, for example, see [A-AER-ATBD]. The context of these algorithms with respect to other EarthCARE algorithms is shown in Figure 1.

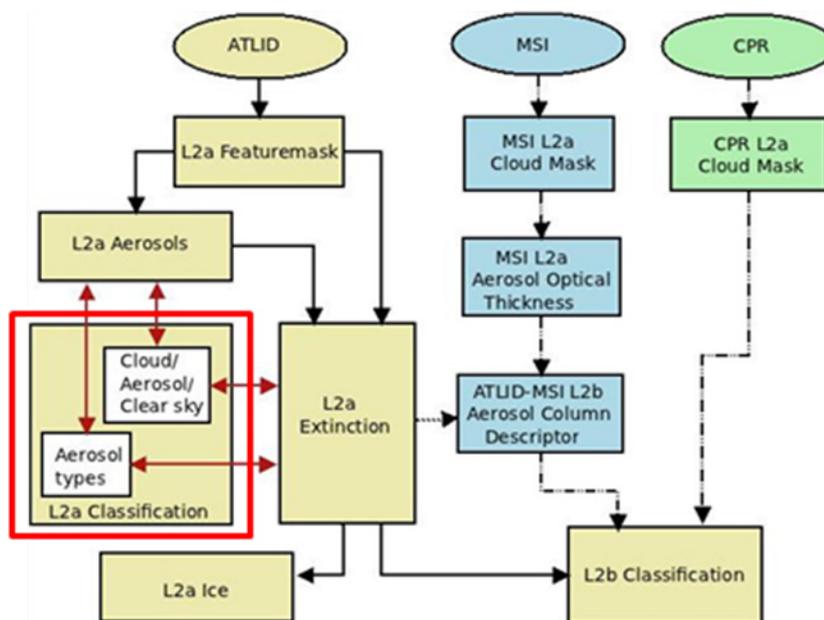


Figure 1: Schematic relationship of the algorithm described in this ATBD (red-box) with respect to other lidar-only (L2a) algorithms as well as relevant MSI and CPR synergetic (L2b) algorithms.

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 Mission 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	2.1	26/05/2011
A-FM-ATBD	EC-TN-KNMI-ATL-ATBD-A-FM-010	ATLID L2a Feature Mask ATBD	2.2	26/05/2011
A-EBD-ATBD	EC-TN-KNMI-ATL-ATBD-A-EBD-021	ATLID L2a High Resolution Extinction, Backscatter and Depolarization	1.2	26/05/2011
EC-FN-ATLAS	EC-FR-KNMI-ATL-027	ATLAS Final Report	1.0	27/05/2011

2.3. Scientific References

Keyword	Reference
Ansmann et al., 2011	Saharan Mineral Dust Experiments SAMUM-1 and SAMUM-2: What have we learned? Tellus B., submitted
Ferrare 2007	Ferrare, R. et al., "Airborne high spectral resolution lidar aerosol measurements and comparisons with transport models", Eos Trans. AGU, 88(52), Fall Meet. Suppl., Abstract A14D-06, (2007).

Keyword	Reference
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
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
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
Muller, D. Ansmann, A. and Petzoldm A. 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
Weinzierl and van Zadelhoff, 2011	Weinzierl, B. and van Zadelhoff, G-J, Technical Note 2: ICAROHS project, 2011

3. Scientific Background of the algorithms

3.1. *Algorithm history*

3.1.1. *L2a Lidar classification*

The lidar only water/ice/aerosol procedure discussed in this ATBD is new, so there is no formal previous version. However, the ideas behind the procedure are generally draw on aspects of the earlier developed CALIPSO classification procedure.

3.1.2. *Aerosol Typing*

As with the L2a Lidar target classification procedure the Aerosol typing procedure discussed in this ATBD is new, so there is no formal previous version. However, the ideas behind the procedure draw on aspects of the earlier CALIPSO aerosol typing procedure [Omar et al. 2009].

3.2. *Algorithm introduction*

3.2.1. *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.

3.2.2. *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.

3.3. Physical/Mathematical Background

3.3.1. L2a Lidar Classification

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 2. 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])

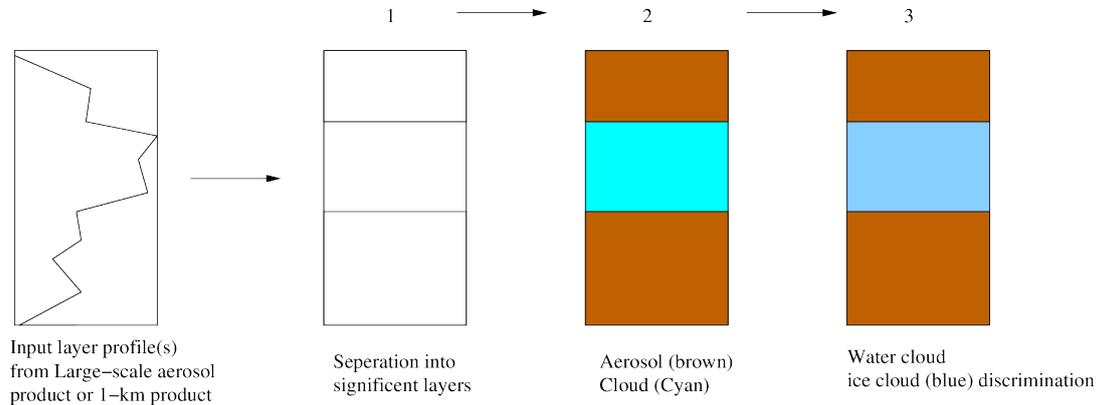


Figure 2: Sketch of the L2a Lidar classification main steps.

3.3.1.1. Detection of significant height boundaries.

The approach used here is somewhat “brute-force”. For a given trial number of layers the goodness-of-fit parameter for all possible layering structures is evaluated. The goodness of fit parameter (or reduced chi-squared parameter) here is defined (for 2 layers) as

$$\chi_{r,x}^2 = \frac{1}{nz - 1 - nl} \left[\sum_{i=1}^{nl} \sum_{j=i+1}^{nl} \left(\left(\frac{x_i - \bar{x}_1}{\sigma_{x_i}} \right)^2 + \left(\frac{x_j - \bar{x}_2}{\sigma_{x_j}} \right)^2 \right) \right] \quad (3.1)$$

where nz is the number of altitude bins being considered, nl is the number of layers being considered (2 in this example) and x_i can be, for example, the depolarization ratio δ or the Backscatter ratio R and the barred quantities represent error weighted averages. i.e.

$$\bar{x} = \frac{\sum x_i}{\sum \sigma_{x_i}^{-1}} \quad (3.2)$$

where the summations are carried out over the appropriate layer indices. After all possible goodness-of-fit parameters are determined for a fixed number of layers the best value is chosen. This is done for different number of allowed sub-layers up to a set maximum. Finally, the optimal number of layers is selected by using a criteria that involves trading-off the improvement in goodness-of-fit parameters with increasing number of layers.

3.3.1.2. Cloud-Aerosol identification

The physical information available using the lidar measurements is strictly not enough to unambiguously separate clouds from aerosol. Thus, there is no alternative but to appeal to the fact that clouds tend to be optically thicker than aerosol and to augment the decision process by using auxiliary information such as expected boundary-layer height from, for example, ECMWF forecast or analysis fields. Accordingly, the procedure described here uses a number of thresholds applied primarily to the measured backscatter.

3.3.1.3. Water-cloud/Ice-cloud separation

Once a layer has been classified as being a “cloudy” layer the depolarization measurement capability of EarthCARE provides a means to separate ice from water. Typical depolarisations associated with ice cloud particles are in the range of 0.4-0.6. In the case of spheres (water cloud droplets), the backscatter lidar signal will contain no cross-polar component *if multiple-scattering* (MS) is not occurring. Thus, if MS were not an issue the separation would be quite straight-forward.

In fact multiple-scatter from water clouds is significant and will give rise to a sometimes substantial cross-polar signal. This is demonstrated by Monte-Carlo calculations made using the ECSIM lidar forward model shown in Figure 3. Here it can be seen that linear depolarization ratios reached in water clouds can closely approach those typical of ice clouds. Thus, in this work we use as a starting point the approach taken by the CALIOP team, that is, to consider the relationship between the integrated backscatter and depolarization ratio.

The expected relationship between layer attenuated backscatter and depolarization ratio for ATLID is shown in the Left panel of Figure 4. The results come from ECSIM simulations applied to a range of water cloud cases for the indicated effective radii, extinction coefficients and cloud physical thickness ranging from 0.1 to 2 km. It can be seen that the results closely follow the fit found by Hu et al. 2007 using an independent MC code applied to the case of the CALIOP 532 nm channel. It can also

be seen that the fit of Hu et al. is inaccurate for very low optical depths.

The good match between the MC calculations made for ATLID conditions and the earlier calculations made by Hu et. al. for CALIPSO indicate that the multiple-scattering induced depolarization-vs-integrated backscatter relationship for CALIPSO also is valid for ATLID. Thus the thresholds used by CALIPSO (see Figure 5) can also be used for ATLID.

There is one issue with using the integrated attenuated backscatter that should be considered. For a classification system to be most practical it should be applicable to readily measurement quantities. Thus, in the case of CALIPSO one is limited to using the attenuated backscatter which can only be accurately determined in the case of single-layer systems (otherwise the attenuation by higher layers affects the determination of the integrated attenuated backscatter for the lower layers). However in the of ATLID the HSRL capability facilitates the direct determination of the integrated backscatter (as opposed to the integrated attenuated backscatter). The expected relationship between multiple-scattering induced depolarization and backscatter is shown in the Right panel of Figure 4. Compared to the left panel the x-axis is now stretched (so much so that some points associated with depolarization ratios above 0.2 are now off-scale). This stretching further increases the separation between water and ice regions. Thus the use of HSRL backscatter measurements may lead to even more applicable and robust water-ice discrimination compared to the use of CALIOP data.

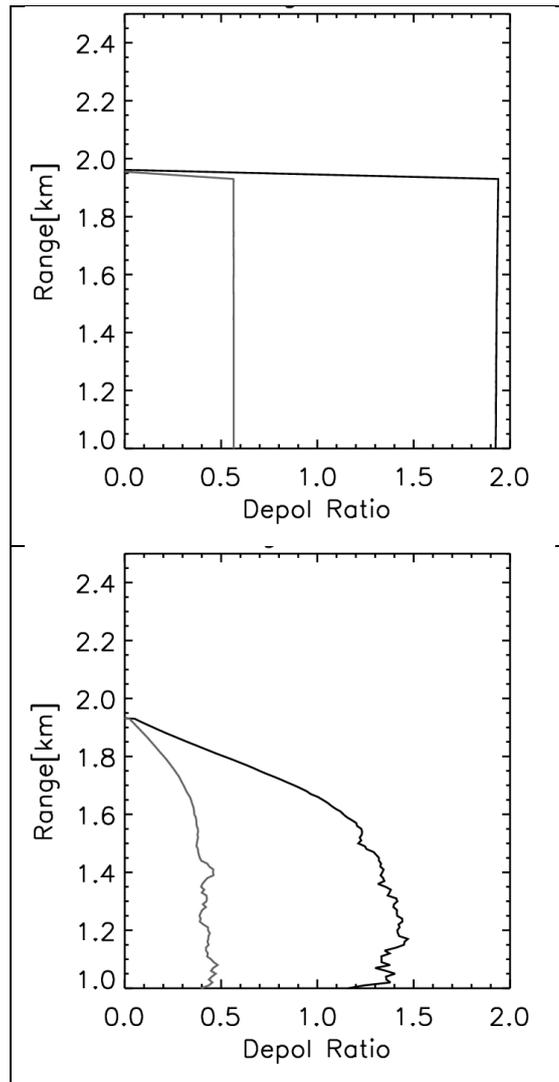


Figure 3: Sample Circular depolarization ratios (Black-line) and Linear Depolarization ratios (Grey-lines) produced by Monte-Carlo calculations for an (Top) Ice cloud and (Bottom) water cloud. Note that EarthCARE will make only linear depolarization measurements.

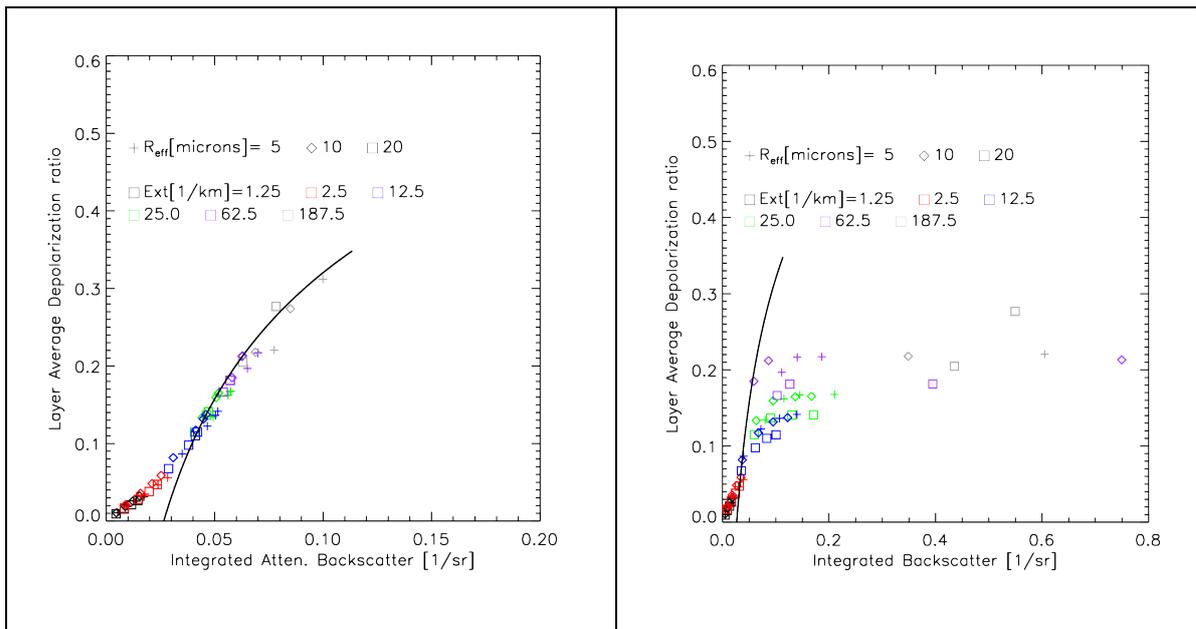


Figure 4: Water cloud integrated depolarization ratio vs integrated attenuated backscatter (Left) and (Right) vs integrated backscatter. Ice clouds would be typically limited to the upper left regions of these plots. The solid line is a fit to MC results calculated at the CALIPSO wavelength of 532 nm found earlier by Hu et al. 2007.

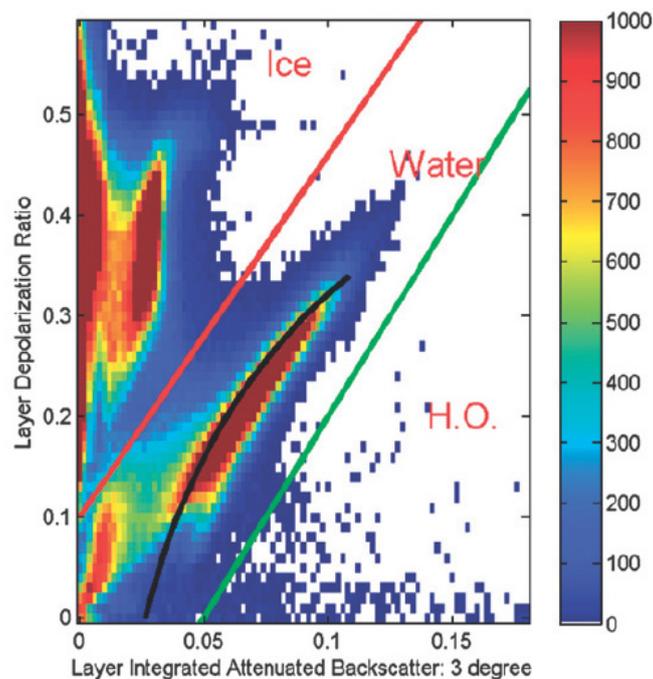


Figure 5: Histogram of depolarization-vs-attenuated backscatter values for January 2009 derived from CALIOP observations along with limits used to separate Ice, Water and Horizontally Oriented crystals (Figure taken from Hu et al. 2009). The solid Black line is the same line as shown in Figure 4.

3.3.2. Aerosol Typing

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 (Ferarre 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 6. 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.

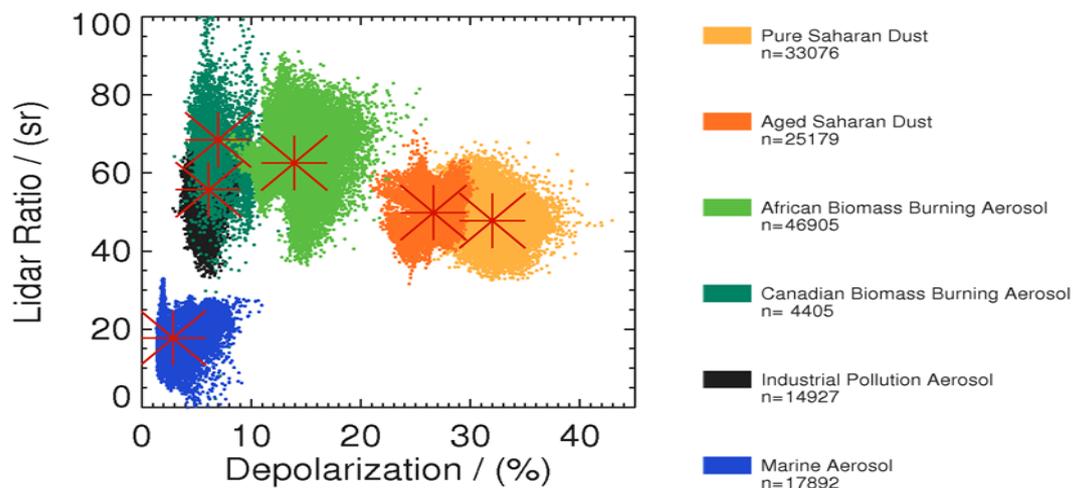


Figure 6: Lidar characteristic properties for different aerosol types from DLR field measurements (ICAROHS ATBD1) preliminary results.

The combination of the two types most likely resulted in a relatively high depolarization in those regions dominated by biomass burning aerosols. The results from the NASA B200 King aircraft (Figure 7) show very similar results compared to those found within Europa and Africa.

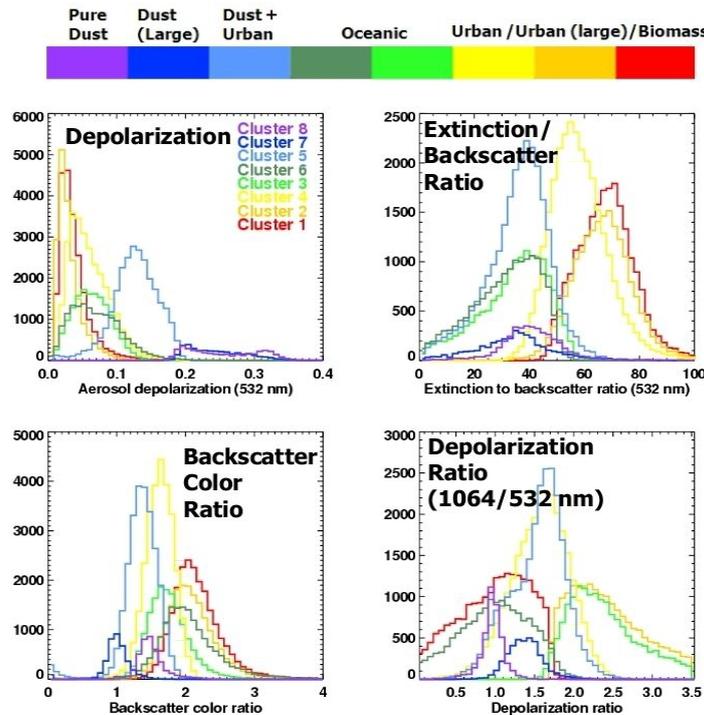


Figure 7: Type depended Aerosol probabilities for important aerosol quantities measured by HSRL in a large number of campaigns with the NASA B200 King aircraft (Ferrare et al. 2007). In case of the ATLID instrument only the top two panels can be retrieved.

In all flights particle mixtures were measured. Future validation efforts should combine the available datasets and aim to derive the pure aerosol type distributions in order to help fill the lidar ratio- depolarization parameter space.

What can be concluded from the measured probabilities are:

- 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.

3.3.2.1. Use of additional information

As the ATLID instrument has no multi wavelength capabilities not all aerosol types can be distinguished from each other. This will lead to similar probabilities for different aerosol types making it in many cases impossible to separate these. One way of including additional information is by taking into account spatial information based on models or observations. In Figure 8 seasonally averaged results from the CALIOP lidar are shown (Kittaka et al. 2008, private communications). The two types shown here are the average optical depth of marine aerosols and dust, together with their mean profile information.

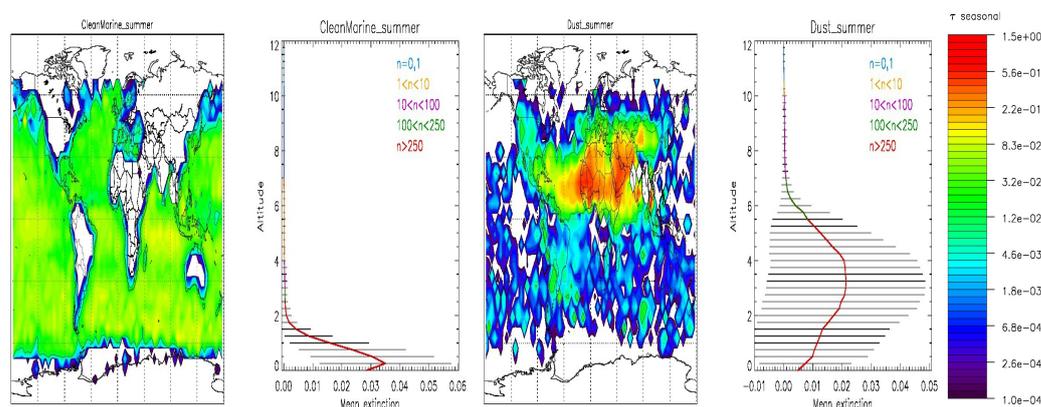


Figure 8: Average seasonal optical depth up to 12 km based on global CALIOP measurements (averaged in 5x5 degrees boxes) for marine aerosols (panel A) and dust (panel C) with the colour-scale presented on the right. The corresponding mean extinction at each height is shown in panels B and D respectively. The colors in the latter two profiles indicate the number of cells included in the calculation of the mean and std. deviation for all cells with a retrieved extinction.

There are a number of conclusions which can be drawn from these results.

- Marine aerosols follow roughly the land-ocean separation but not entirely. A map like shown here should be used to help define regions for which no marine aerosols are expected to be found
- In general marine aerosols are only available in the lowest few kilometers of the atmosphere.
- Dust has a far more patchy distribution and there are no regions which can be simple discarded as dust containing. Only night-time data is taken into account giving rise to the large region in the north without dust in the summer. Dust reaches up to 8 km in this season, so dust can also not be separated from other aerosol types based on altitude considerations.

Based on the year 2007 there are only small differences in the seasonal distribution for marine aerosols and these maps including the height distribution from CALIOP can be used as additional information for separating marine from other aerosol types. The separation of other aerosol types is more difficult since, e.g. biomass burning and

industrial pollution can reside at the same place and height and the same holds for volcanic ash and dust (see Figure 6 and Figure 7).

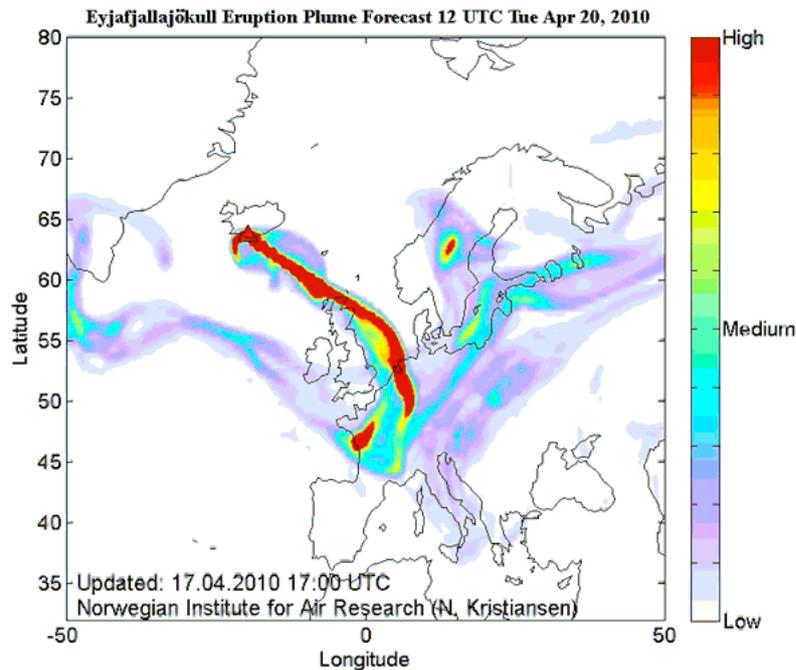


Figure 9 Flexpart trajectory model results for the Eyjafjallajökull volcanic eruption in 2010 (N. Kristiansen). Forecast models results like these could be used for enabling the detection of the volcanic ash aerosol type within the model.

To separate these remaining types one can not use simple use a climatology. Dust and volcanic ash distribution overlap each other in the lidar ratio-depolarization parameter space and as was discussed above dust particles occur globally. Most importantly volcanic eruptions can not be predicted using climatology and can only be dealt with using short-term trajectory model forecasts instead or near real time global observations. In Figure 9 an example of the Flexpart trajectory model (<http://transport.nilu.no/flexpart>) is shown for the Eyjafjallajökull volcanic eruption in 2010. The plume prediction including an additional ~5 degrees extended bounding box to account for model errors or changing wind direction can be used to enable the detection of volcanic ash in near real time. A second possibility is using the volcanic ash information retrieved from imagers (SEVIRI, GOESS, MODIS type) to generate volcanic ash probability maps. The near real time ash detection capabilities will provide valuable information on the height of the ash layer to both aviation experts (e.g. Volcanic Ash Advisory Centers (VAAC) & air traffic control (ATC) centers) and aid the groups working on the evaluation and assimilation of aerosol plume trajectories.

In the case of the distinction between biomass burning and polluted industrial one can consider using climatology in those areas where there are regular biomass burning events or in case of industrial pollution parts of the ocean regions can be excluded. However basing the typing on too strict rules could lower the correct assignment of aerosol types, e.g. it would have resulted in a low chance of finding biomass burning with the 2010 Russian wildfires compared to climatology. If there is a special need to separate these two aerosol types it is more advisable to use maps including results

from (Multi-spectral) fire detection from geostationary platforms (e.g. SEVIRI, GOESS, ESA's ATSR World Fire Atlas) or polar orbiting instruments (MODIS). These sources can be used in combination with (rough) trajectory models to limit the possible area's for which biomass burning can be detected.

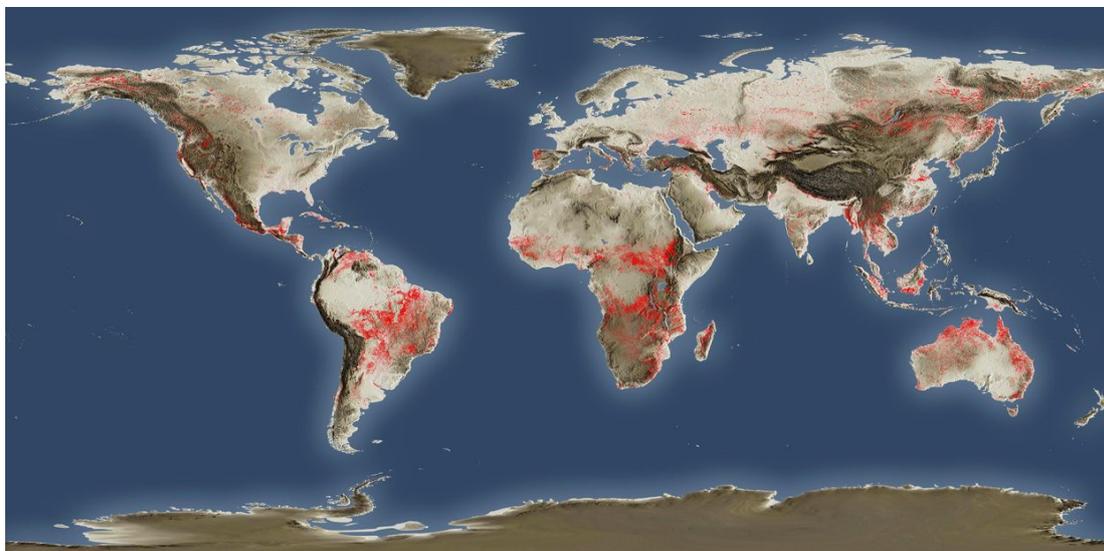


Figure 10: The ESA global detection of hot spots by ERS-2's Along Track Scanning Radiometer (ATSR-2) and Envisat's Advanced Along Track Scanning Radiometer (AATSR) from July 1996 to August 2010.

An example of biomass burning sources is shown Figure 10, where results from the ESA's ATSR World Fire Atlas is shown from July 1996 to August 2010. As the near real time measured data is available within six hours, findings from the previous day could be used to run simple trajectory models. These results can then be transformed into maps. A second option would be to create global maps of biomass burning based on UV-VIS observations (OMI/Trop-OMI, MODIS type), where results from the previous day could be combined in a rough global map to distinguish regions in which there is a chance of biomass burning detection.

For now there will be no attempt to separate biomass burning and continental pollution as this would potentially rely on external information only. If there are particular requests or future needs for these products, potential global UV-VIS retrievals available around the launch of EarthCARE would have to be tested after launch to check for any skill in separating these particular two types.

The main aerosol types which will be distinguished within the standard ATLID-Aerosol typing are:

- Marine
- Biomass burning and/or Continental Pollution
- Clean Continental
- Dust
- Volcanic ash

This typing list is the same for the ATLID only algorithm (A-TC) and ATLID-MSI algorithm (AM-ACD).

4. Justification for the selection of the algorithm

4.1. *L2a Lidar Classification*

The procedures for sub-layer determination and water cloud, ice cloud and aerosol target separation fulfil the appropriate requirements and are, particularly in the case of water-vs-ice discrimination directly physically based. Without moving towards much more sophisticated (and arguably still exotic) classification procedures such as neural net based approaches it is difficult to envision any very different approach other than the one described here.

4.2. *Aerosol Typing*

The procedures for the aerosol typing algorithm presented here fulfils the appropriate requirements. Moreover, the method is suited to the incorporation of auxiliary information into the classification process thus improving the precision of the aerosol typing. The method is also providing probabilities of occurrence to the end user which will be much more useful for i.e. model evaluation studies, than just bare type assignments.

5. Mathematical algorithm Description

5.1. Input parameters

The L1a classification procedure requires output from the Lidar Featuremask Algorithm (A-FM) the High-Resolution Extinction, Backscatter and Depolarization Ratio Algorithm (A-EBD) and the Large-scale aerosol Extinction, Backscatter and Depolarization Ratio Algorithm (A-AER)

The aerosol typing procedure receives all its information from the A-AER and A-EBD algorithms directly and when available external aerosol probability maps.

Table 1: Input data parameters L2a Lidar Classification

Variable	Description	Unit	Dim	Source	Type
Common inputs					
Time	UTC time	s	time	A-EBD/ A-AER	real*8
Height	Height of each radar/lidar gate above mean sea level for the altitude range being processed.	m	height. Note: the input height array here will usually be a sub-section of the full height array	A-EBD/ A-AER	real
Significant height boundary Inputs					
Ray_beta	Rayleigh Backscatter profile	1/m/sr	time,height	A-EBD/ A-AER	real
Depol	Depolarization Ratio	-	time,height	A-EBD/ A-AER	real
D_Depol	Associated Error	-	time,height	A-EBD/ A-AER	real
Beta	Backscatter	1/m/sr	time,height	A-EBD/ A-AER	real
D_Beta	Associated Error	1/m/sr	time,height	A-EBD/ A-AER	real
Ext (optional)	Aerosol/Cloud Extinction	1/m	time,height	A-EBD/ A-AER	real
C_ext (optional)	Associated error covariance matrix	1/m ²	time,height,height	A-EBD/ A-AER	real
Aerosol, ice cloud, water cloud discrimination					

Tropopause Height	Height of the tropopause	m	time	A-EBD/ A-AER	real
Boundary Layer Height	Height of the atmospheric boundary layer (from ECMWF forecast or analysis)	m	time	A-EBD/ A-AER	real
T	Temperature	K	time,height	A-EBD/ A-AER	real
Tw	Wet Bulb Temperature	K	time,height	A-EBD/ A-AER	real

Table 2: Input data parameters Aerosol Typing from the *A-EBD* and *A-AER* procedures

Variable	Symbol	Description	Unit	Dim	Source	Type
Time	t_{utc}	UTC time	S	time	A-EBD/ A-AER	real*8
Height	Z	Height of each radar/lidar gate above mean sea level	m	n_layers	A-EBD/ A-AER	real
Tropopause Height	Z_{tp}	Height of the tropopause	m	time	A-EBD/ A-AER	real
Boundary Layer Height	Z_{bl}	Height of the Atmospheric Boundary Layer	m	time	A-EBD/ A-AER	real
Longitude	Lon	Longitude	degrees	time, n_layers	A-EBD/ A-AER	real
Latitude	Lat	Latitude	degrees	time, n_layers	A-EBD/ A-AER	real
Depol_1	δ	Layer Depolarization	-	time, n_layers	A-EBD/ A-AER	real
D_Depol_1	dδ	Standard deviation of the depolarization	-	time, n_layers	A-EBD/ A-AER	real
Lidar_ratio_1	S	Layer Extinction-to-Backscatter ratio	-	time, n_layers	A-EBD/ A-AER	real
D_Lidar_ratio_1	dS	Standard deviation of the lidar ratio	-	time, n_layers	A-EBD/ A-AER	real

Table 3: Aerosol probability map information

Variable	Symbol	Description	Dim
Aerosol_type	Type	Aerosol type name	1
Date	Date	Dates/Period for which the map is valid (Julian Date)	2
Longitude	Lon	Longitude	N_lon
Latitude	Lat	Latitude	N_lat
Min_height	Z_{min}	Minimum height at which the aerosol type is expected. If set to -1 there are no aerosol expected within this column	N_lon, N_lat
Max_height	Z_{max}	Maximum height at which the aerosol type is expected. If set to -1 there are no aerosols expected within this column	N_lon, N_lat

5.2. Configuration parameters

Table 4: Configuration parameters L2a Lidar Classification

Variable	Symbol	Description	Unit	Dim	Type
Max_N_layers	$N_{l,max}$	Maximum number of layers an input layer can be subdivided into	-	-	integer
Beta_Cld_default	β_{thres}	Default backscatter threshold for cloud aerosol separation	1/m/sr	-	real
Beta_Cld_strat	$\beta_{thres, strat}$	Default backscatter threshold for cloud aerosol separation for stratospheric layers	1/m/sr	-	real
Beta_Cld_bl	$\beta_{thres, bl}$	Default backscatter threshold for cloud aerosol separation for stratospheric layers	1/m/sr	-	real
A_Depol_beta	$A_{\delta, \beta}$	Slope parameter of depol-backscatter relationship used to determine ice/water threshold	m	-	real

A_Depol_beta	$B_{\delta,\beta}$	Intercept parameter of depol-backscatter relationship used to determine ice/water threshold	-	-	real
---------------------	--------------------	---	---	---	------

Table 5: Configuration parameters Aerosol Typing

Variable	Symbol	Description	Unit	Dim	Type
N_aerosols	N_aer	Total number of aerosols for which a typing probabilities will be calculated	-	1	Integer
The following parameters will be provided for each of the aerosol types defined by N_aer					
Depol_center	δ_0	Depol center of the Gaussian distribution	-	1	Real
Depol_var	σ_δ	Gaussian width in the depol direction	-	1	Real
Lid_rat_center	S_0	Lidar Ratio center of the Gaussian distribution	-	1	Real
Lid_rat_var	σ_S	Gaussian width in the lidar ratio direction	-	1	Real
Gauss_theta	θ	Angle of the gaussian orientation. The angle is defined as the right-handed rotation assuming the depolarization at the x-axis	Deg	1	Real
Map_name	Map	Name incl. directory of the external aerosol map	-	1	Char
Strat_trop	Strat_trop	Is this an aerosol species in the troposphere (0) or stratosphere(1)	-	1	Integer

5.3. Output parameters

Table 6: Output parameters from L2a Lidar Classification procedure

Variable	Description	Unit	Dim	Type
Statistically Significant height boundary Outputs				
n_layers	Optimal number of layers	-	time	integer
Sub_Lay_Bot	Top of layer being processed	m	time,n_layers	real

Sub_Lay_Top	Bottom of layer being processed	m	time,n_layers	real
Depol_1	Layer depolarization ratio	-	time,n_layer	real
D_Depol_1	Associated standard deviation	-	time,n_layers	real
Beta_1	Layer layer average backscatter	1/m/sr	time,n_layer	real
D_Beta_1	Associated error	1/m ² /sr ²	time,n_layers	real
Ext_1 (optional)	Layer layer average extinction	1/m	time,n_layer	real
D_Ext_1 (optional)	Associated error	1/m ²	time,n_layers	Real
S_1 (optional)	Layer layer average extinction-to-backscatter ratio	sr	time,n_layer	real
D_S_1 (optional)	Associated error	Sr	time,n_layers	real
Aerosol, ice cloud, water cloud discrimination				
Target_Type	Most probable target type 1→Surface 2→Clear sky 3→Water Cloud 4→Supercooled water 5→Ice cloud 6→Aerosol	-	time,n_layers	integer
Incon_ice	Flag indicating presence of ice above melting temperature	-	time,n_layers	Logical
Incon_water	Flag indicating presence of water below homogeneous freezing temperature	-	time,n_layers	Logical
W_I_A_prob	Probabilities associated with class 3,5 and 6 (Water, ice or aerosol). For an identified target layer the sum of the probabilities should be 1.0	-	3,Time, n_layers	Real

Table 7: Operational Output parameters Aerosol Typing

Variable	Symbol	Description	Destination	Dim
Time	t_{utc}	UTC time	A-EBD/ A-AER	time
Height	z	Height of each radar/lidar gate above mean sea level	A-EBD/ A-AER	N_layers
Aerosol_type_names	names	Aerosol Type Names	A-EBD/ A-AER	Ntypes
Aerosol_type_direct	T_d	Aerosol Type	A-EBD/ A-AER	Time, n_layers
Aerosol_type	T	Aerosol Type	A-EBD/ A-AER	Time, n_layers
Aerosol_prob_direct	$P_{\text{eff,d}}$	Aerosol-probabilities using signals only	A-EBD/ A-AER	Ntypes,Time, n_layers
Aerosol_prob	P_{eff}	Aerosol-probabilities using signals and additional data	A-EBD/ A-AER	Ntypes,Time, n_layers

5.4. Algorithm flow charts

5.4.1. L2a Lidar Classification

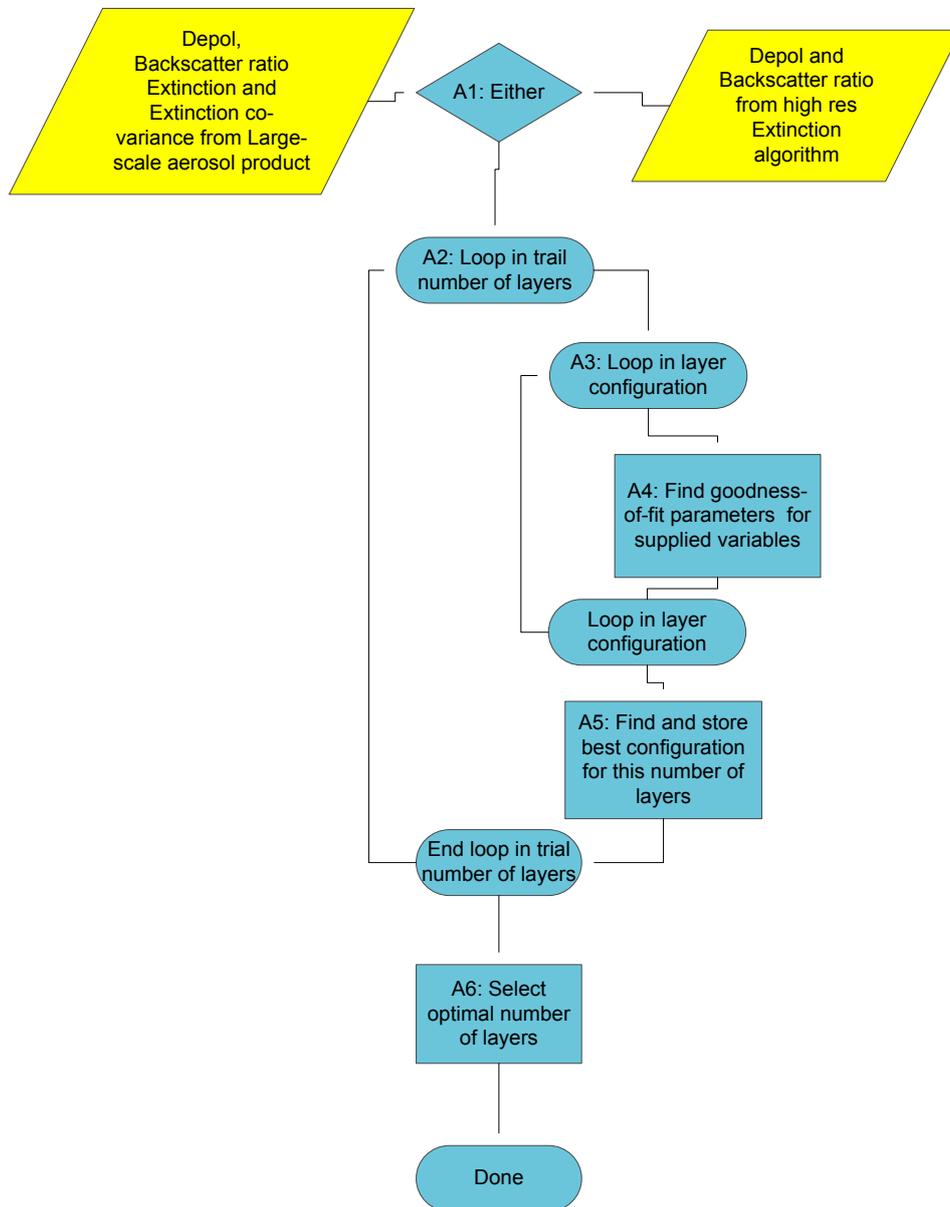


Figure 11: Significant layer detection procedure flow diagram

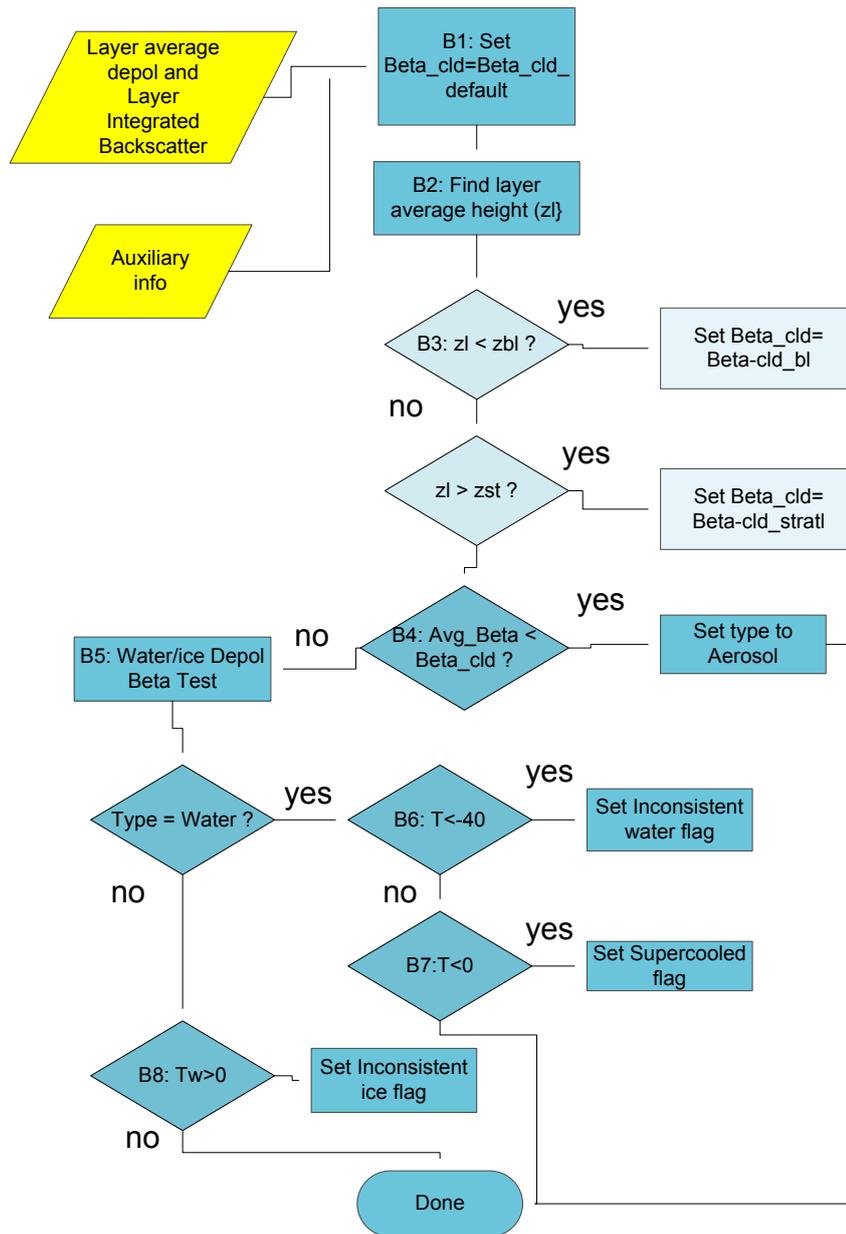


Figure 12: Aerosol cloud discrimination and water cloud/ice cloud discrimination procedure.

5.4.2. Aerosol Typing

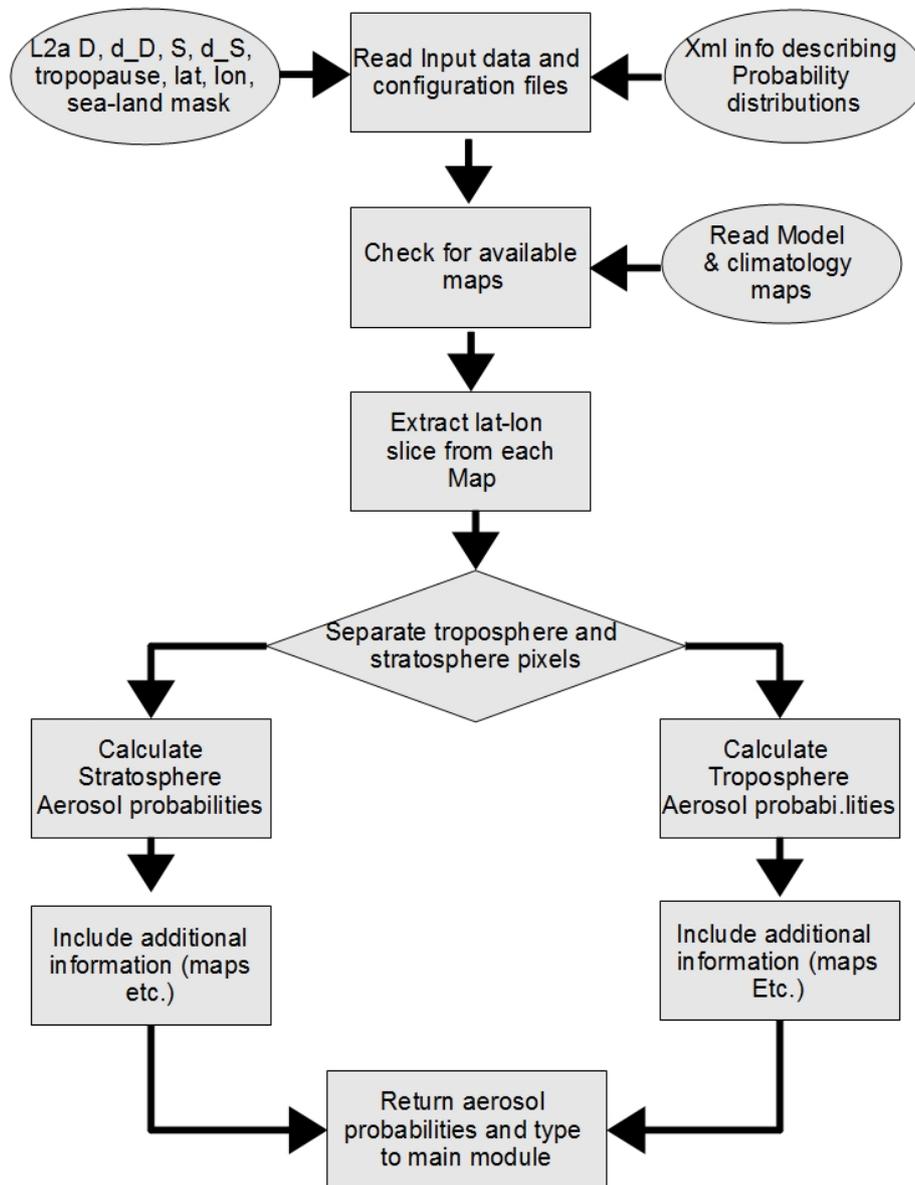


Figure 13: Aerosol Typing flow diagram

5.5. *Algorithm Definition*

5.5.1. *L2a lidar classification*

5.5.1.1. *Detection of significant layer boundaries*

The schematic flow diagram corresponding to this procedure is given in Figure 11. Each of the main steps is explained in turn

A1: The procedure is intended to be called with different inputs either output generated from the Large-scale aerosol extinction algorithm, in which case a single a profile of backscatter, depolarization and extinction together with error estimates (in the case of the extinction profile the error covariance matrix will also be imported) . When the routine is called as part of the internal processing of the High-Resolution Extinction algorithm extinction and the associated error information may or may not be passed (depending on the stage the routine described here is called in the High-Resolution algorithm).

A2: Loop in trial number of layers

The number of trial layers starts with one and goes to a set maximum or the number of height bins in the layer, whichever is smaller.

A3: Loop in layer configuration

For the specified number of layers the routine loops through all possible layering possibilities. For example, for two layers one layer would span from $iz_1 = 1, nz$ while the second layer will span from $iz_2 = iz_1, nz$. For 3 layers then we have $iz_3 = iz_2, nz$ etc.

A4: Find Goodness of fit parameters

For each layer the Goodness-of-fit parameter is calculated and the values are summed across the layers. For uncorrelated quantities, for example the depolarization ratio, for a single layer we have

$$\chi_{red,\delta}^2 = \frac{1}{nz - 1 - nl} \left[\sum_{i=iz_{bot}}^{iz_{top}} \left(\frac{\delta_i - \bar{\delta}_1}{\sigma_{\delta_i}} \right)^2 \right] \quad (3.3)$$

where

$$\bar{\delta} = \frac{\sum_{i=z_{bot}}^{i=z_{top}} \delta_i}{\sum_{i=z_{bot}}^{i=z_{top}} \sigma_{\delta_i}^{-1}} \quad (3.4)$$

For quantities, such as the layer extinction-to-backscatter ratios, the correlation in the supplied extinction profiles must be accounted for, i.e. The layer Extinction-to-Backscatter ratio is given by

$$S_l = \frac{OT_l}{I\beta_l} \quad (3.5)$$

where α is the extinction coefficient and β is the backscatter coefficient

$$OT_l = \sum_{i=z_{bot}}^{z_{top}} \alpha_i \Delta z_i \quad (3.6)$$

and

$$I\beta_l = \sum_{i=z_{bot}}^{z_{top}} \beta_i \Delta z_i \quad (3.7)$$

Since the extinction values will be correlated in altitude the error in OT_l is given by

$$\sigma_{OT_l}^2 = \sum_{i=z_{bot}}^{z_{top}} \sigma_{\alpha}^2(z) \Delta z_i + \sum_{i=z_{bot}}^{z_{top}} \sum_{j=z_{bot}}^{z_{top}} [C_{\alpha}]_{i,j} \Delta z_j \Delta z_i \Big|_{i \neq j} \quad (3.8)$$

and (assuming that the error in the integrated the backscatter is uncorrelated with the extinction error) we will have

$$\sigma_{S_l}^2 = S_l^2 \left(\left(\frac{\sigma_{OT_l}^2}{OT_l} \right) + \left(\frac{\sigma_{I\beta_l}^2}{I\beta_l} \right) \right) \quad (3.9)$$

For each supplied quantity the layer integrated quantities and their error estimates are stored for use in later steps.

A5: Find and store best configuration

For each number of trial layers the layering configuration with the best total summed goodness-of-fit is identified and stored in an array. The values and error estimates for the layer integrated backscatter, depolarization etc. corresponding to the best configuration are also stored in an array.

A6: Select optimal number of layers

The array of best goodness-of-fit values is normalized by its smallest value (which almost always will be the value associated with the highest number of levels) and the optimal number of layers is selected by choosing the smallest number of layers which is still within a factor of 25% of the minimum value.

The values and error estimates for the layer integrated backscatter, depolarization etc. corresponding to the best configuration are stored and passed as part of the routines output.

5.5.1.2. *Aerosol cloud discrimination and water cloud/ice cloud discrimination procedure.*

B1: Set Beta_cld

The value of the layer average backscatter coefficient used to distinguish clouds from aerosols is set to its default value

B2: Find layer average height

The value of the layer average height for the layer being processed is found.

B3: (which covers all of the Light-Cyan boxes)

Here the height of the layer is used to adjust the value of **Beta_cld**. In particular, alternative values of **Beta_cld** are used for stratospheric layers and layers below the expected atmospheric boundary layer height. The height of the stratosphere and the atmospheric boundary layer are supplied by auxiliary information (i.e. ECMWF forecast or analysis fields).

B4: Avg_beta < Beta_Cld ?

Layers are classified as being aerosol layers if their associated average backscatter is below the **Beta_Cld** threshold. The associated probability given the estimate error in the layer mean beta (calculated earlier as part of the layering procedure by standard quadratic summation of the individual beta errors) that the threshold has been exceeded is also calculated assuming Gaussian statistics. In particular, the cloud occurrence probability is given by

$$P_{cld}(\beta_{thres}; \beta_l, \sigma_{\beta_l}) = 1 - 0.5 \left[1.0 + \operatorname{erf} \left(\frac{\beta_{thres} - \beta_l}{\sigma_{\beta_l} \sqrt{2}} \right) \right] \quad (3.10)$$

where

$$\beta_l = I\beta(z_{top} - z_{bot})^{-1} \quad (3.11)$$

and

$$\sigma_{\beta_l} = \text{sqrt} \left(\frac{\sum_{i=z_{bot}}^{i=z_{top}} \sigma_{\beta_i}^2}{\left(\frac{(z_{top} - z_{bot})}{\Delta z} \right)} \right) \quad (3.12)$$

where Δz is the range resolution.

B5: Water/ice Depol Beta Test

A threshold between water-cloud and ice-cloud is constructed by using a linear relationship between layer average depolarization ratio and the layer averaged backscatter i.e.

$$\delta_{thres} = A_{\delta,\beta} I\beta_l(z_{top} - z_{bot})^{-1} + B_{\delta,\beta} \quad (3.13)$$

Eqn.(3.13) is evaluated for the layer integrated backscatter. If the observed value of the layer depolarization is greater than the resulting value of δ_{thres} then the layer is classified as being ice otherwise it is classified as being water.

For quality control and in order to track the confidence of the classification the quantities $f_{\Delta\delta}$ and $f_{\Delta I\beta}$ are also calculated and stored. These quantities measure the distance normalized by the error that the given layer is from the threshold line along the depolarization and average backscatter axes respectively.

$$f_{\Delta I\beta} = \left(\frac{I\beta_l(z_{top} - z_{bot})^{-1} - \left(\frac{\delta_{thres} - B_{\delta,I\beta}}{A_{\delta,I\beta}} \right)}{\sigma_{I\beta_l}} \right) \quad (3.14)$$

and

$$f_{\Delta\delta} = \left(\frac{\delta_l - \delta_{thres}}{\sigma_{\delta_l}} \right). \quad (3.15)$$

The probability that the layer is ice is then determined by the following equation

$$P_{ice}(x) = 1 - 0.5 \left[1.0 + \operatorname{erf} \left(\frac{-x}{\sqrt{2}} \right) \right] \quad (3.16)$$

where here $x = \operatorname{sqr}t(f_{\Delta\beta}^2 + f_{\Delta\delta}^2)$.

B6: T < -40C ?

Here is the layer temperature is less than the homogeneous freezing point of water but has been assigned as being a water layer in step **B5** then the *Inconsistent water flag* is set for this level. NOTE: This is only applicable for tropospheric layers. Liquid phase PSCs are known to exist at much lower temperatures. The identification of PSC types is expected to occur as a special TBD branch within the “Aerosol typing” procedure.

B7: T < 0 ?

If the water layer temperature is below the freezing temperature (but above -40 C) then the *Supercooled water flag* is set.

B8: Tw > 0

If the wet-bulb temperature of an ice layer is above 0 (the melting temperature) then the *Inconsistent Ice flag* is set for this layer.

5.5.2. Aerosol Typing

The procedures within the aerosol typing module are mathematically relatively simple. There is a straightforward route to follow as is shown in the flow diagram (Figure 13). 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.

5.5.2.1. Input of the data and configuration parameters

The aerosol typing algorithm will be structured as a module. It will therefore not be a stand-alone algorithm. The main reason for this decision is related to the need of this information in two of the envisioned algorithms, the A-EBD and A-AER algorithms, within the ATLID only algorithm chain. The module can be called following a call to the L2a Lidar classification procedure, defining the aerosol regions. The two algorithms will need to provide the retrieved lidar ratio and depolarization and their respective error estimates, the latitude, longitude and altitude information and finally the tropopause height and land-water mask. The configuration parameters describe the properties of the different aerosol probability distributions and provide the locations of additional global maps when available.

5.5.2.2. *Extracting slices from available aerosol maps*

If any aerosol a-priori probability of occurrence maps are provided, 2-D along-track x height slices are extracted for the section of the orbit being processed. For each profile the corresponding column within the map is determined. In the case of marine aerosols, this will also include a maximum height around 2km (to be determined based on CALIPSO observation derived statistics) . The slice will be set to 1 for regions where there is a chance of detecting the specific aerosol type and 0 for those regions where no aerosols of this type should be expected. One can envision an intermediate regime where values are in between 0 and 1 but the definition of any rules to do so would require a more detailed study of the possibilities and calibrations.

5.5.2.3. *Stratosphere - troposphere separation*

One of the required inputs is the tropopause height. An internal array (strat_trop) is created. This strat_trop consists of the following 5 integers:

- 9 surface and below surface
- 1 no lidar data
- 0 No aerosols
- 1 Troposphere
- 2 Stratosphere

This array is used when determining which particle type probabilities need to be determined for the cells. For all the cells defined by 1 or 2, the troposphere or stratosphere type probabilities are calculated respectively.

5.5.2.4. *Defining the Aerosol probabilities*

The aerosols types are assigned based on distributions based on observations from regions consisting of a single aerosol type. The aerosol types currently under consideration are provide in Table 7, including the parameters defining the 2D-Gaussian probability distribution as described in Equation (4.1). The aerosol types and center values are compliant with the values used in the IRMA AM-ACD ATBD (private communications). In the future the types and values should, when applicable, be the same for the AM-ACD and A-TC algorithms. All these parameters will be defined in the configuration file to make the validation and updating of the algorithm to new parameterizations straightforward. The 1σ distribution values for each of the particle types are estimated for now.

Table 8: Aerosol types and the parameters defining the probability distributions for each of these aerosol types. The numbers given in this version of the ATBD are **NOT based on observations but are for based on estimates of the values. Future validation and calibration efforts combining all available data will be needed to define realistic parameters to describe the different probability distributions. The stratospheric parameters are not set at this moment.**

#	Name	Angle (rad)	δ_0	σ_δ	S	σ_S
Troposphere						
1	Marine	0	0.03	0.05	20	8
2	Biomass burning/ Continental Pollution	0	0.03	0.05	60	15
3	Clean Continental	0.2	0.03	0.05	40	10
4	Dust	0	0.35	0.06	55	10
5	Volcanic	0	0.35	0.06	70	10
Stratosphere						
6	Volcanic	x	x	x	x	x
7	PSC	x	x	x	x	x
8
..						

The bivariate Gaussian probability densities for the different types are described by:

$$\begin{aligned}
 A &= \frac{\cos(\theta)^2}{2\sigma_\delta^2} + \frac{\sin(\theta)^2}{2\sigma_S^2} \\
 B &= \frac{-\sin(2\theta)}{4\sigma_\delta^2} + \frac{\sin(2\theta)}{4\sigma_S^2} \\
 C &= \frac{\sin(\theta)^2}{2\sigma_\delta^2} + \frac{\cos(\theta)^2}{2\sigma_S^2} \\
 P(\delta, S) &= P_0 \exp(-1(A(\delta - \delta_0)^2 + B(\delta - \delta_0)(S - S_0) + C(S - S_0)^2))
 \end{aligned} \tag{4.1}$$

with θ the rotation of the 2D-Gauss, assuming the depolarization on the x-axis. S is the lidar ratio, δ the depolarization σ_S and σ_δ their respective variances. Each distribution is normalized to a maximum of 1.

In Figure 14 the distributions described by the parameters in Table 7 are shown. The probabilities are shown in between the range of 1e-11 up to 1.

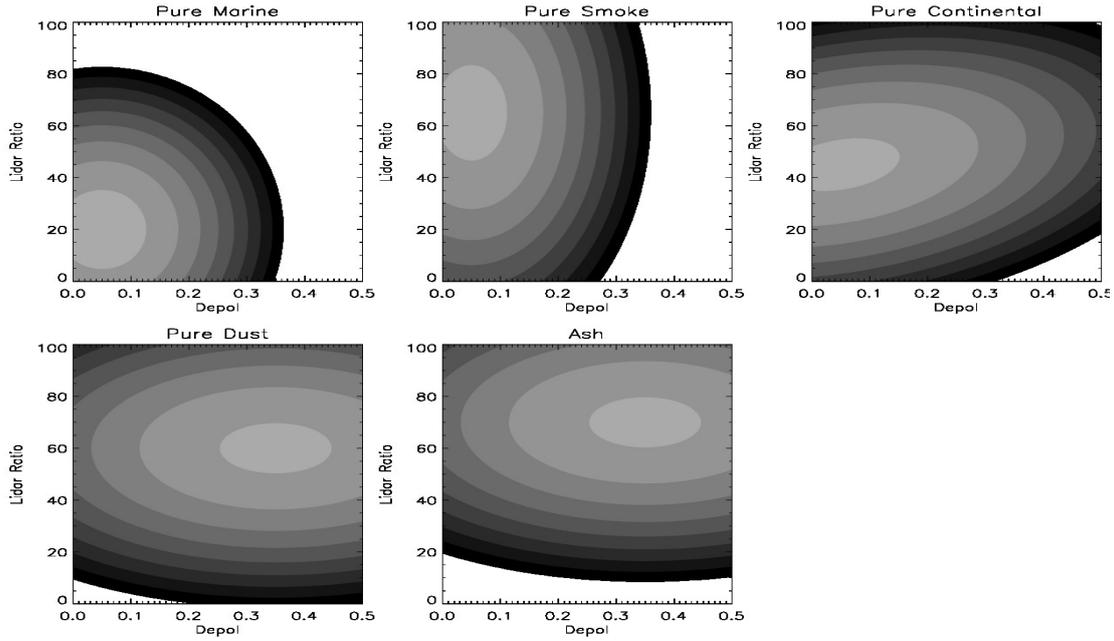


Figure 14: Aerosol probabilities for the 5 tropospheric aerosol types defined by the parameters in Table 7. The greyscales show the distributions from 1e-11 up to 1 in the centre of the distributions.

5.5.2.5. Calculating the type probabilities

For each observed combination of lidar ratio and depolarization, the probability can be calculated for all the aerosol types being considered. If this calculation were to be done rigorously, one would need to calculate the convolution of the 2D observed Gaussian with the 2D bivariate Gaussian probability density functions. This calculation becomes an enormous expression when solved analytically and therefore the calculation is carried out approximately by retrieving the probability 49 times (7x7) between plus and minus 3σ in both the probability and depolarization (see Figure 15 for an example). The effective probability ($P_{\text{eff},d}$) is calculated by adding the weighted 49 probabilities (P) according to:

$$P_{\text{eff},d}^l(\delta_k, S_k) = \sum_{i=1}^7 \sum_{j=1}^7 P(\delta_k + (i-4)\sigma_{\delta,k}, S_k + (j-4)\sigma_{S,k}) \bullet W(i, j) \quad (4.2)$$

$$W(i, j) = \frac{1}{2\pi} \exp\left(-\frac{(i-4)^2}{2} - \frac{(j-4)^2}{2}\right)$$

with S_k and δ_k the observed lidar ratio and depolarization with their respective error estimates ($\sigma_{S,k}, \sigma_{\delta,k}$), $P_{\text{eff},d}^l$ the probability of the observation k for type l . The weighting function (W) assumes an even spread between -3σ and $+3\sigma$ and P is the probability distribution as was defined by Eqn. (4.1).

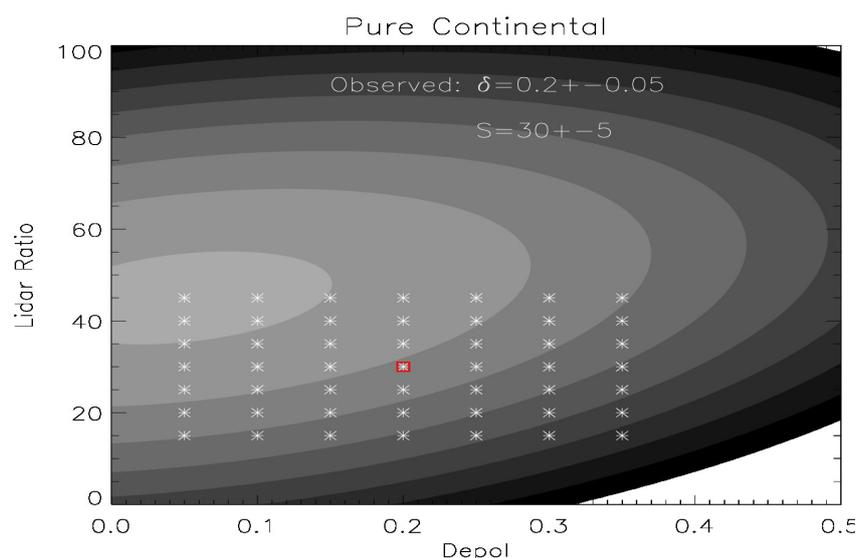


Figure 15: Calculating the clean continental type probability for an observed point with a lidar ratio of 30 and depolarization of 0.2. The calculation is performed at all 49 points and weighted according to the distance to the observed value [Equation (4.2)]

The resulting probabilities for all the relevant types is the first direct result and will be reported in the output file ($P_{\text{eff,d}}$ [**Aerosol_prob_direct**]).

5.5.2.6. *Use of additional information*

As stated above the direct probabilities will very often not result in a single or combination of 2 most likely types. Additional information can be added to lower the probabilities of certain types. The slices, containing 0 or 1 created from the (climatology) maps will be multiplied to the directly retrieved probabilities. The new resulting probabilities (P_{eff} [**Aerosol_prob**]) will be reported in the output as well.

5.5.2.7. *Assignment of aerosol type*

Even though the aerosol type probabilities are the main results from the procedure, the most probable aerosol type will be provided in the output. Defining a single aerosol type will only be possible in those cases where a single type probability is much higher than any of the other types. In Figure 16, three slices through the parameter space are plotted, each at a constant depolarization. Shown are the percentages of the aerosol probabilities to the sum of the probabilities at each point. As all the aerosol types considered here (except for marine aerosols) are centred around relatively high lidar ratios the marine aerosol type is always the preferred type for lidar ratio's below 10, even at high depolarisations. However, it should be noted that the absolute probability of marine aerosols can be very low at the high depolarisations. This indicates that a minimum probability is needed to define the type. As the algorithm

deals with Gaussian distributions the minimum is set to the 3σ level (which includes 99.73% of all data), or a probability of ($e^{-9/2} = 0.0111$).

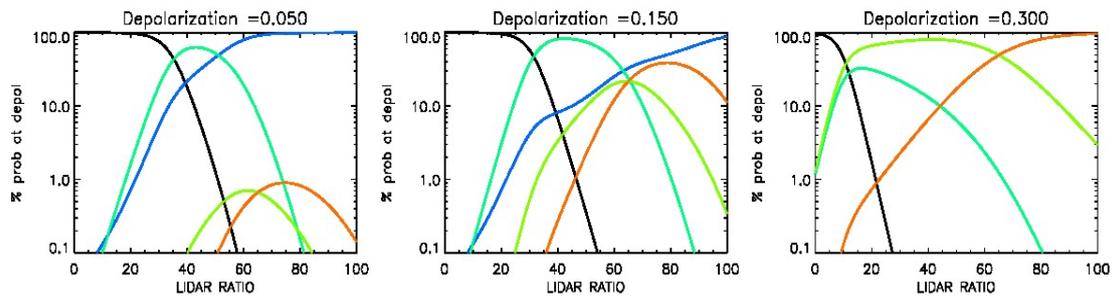


Figure 16: Normalized probability fractions for the 5 aerosol types. for three slices of depolarization. The observations are assumed to have a negligible error estimate in these cases. The line colors define the following aerosol types: Black: Marine Aerosols, Green: Pure continental, Blue: Pure smoke, light green: Dust, orange: Volcanic ash

In case of the volcanic and dust probabilities overlap in a large region and it would be beneficial to use additional tools to separate these types (Figure 16). As there are regions where more than one probability is of comparable height, multiple types will be retrieved for these regions. This does not specify that both types are present, it indicates that more than one type has a high probability in this region.

The type assignment is calculated according to the following rules (Table 9), where P_{i1} ($P_{i1} = \frac{P_{\max}}{\sum P_i}$) is the local highest normalized probable type, P_{i2} the second highest etc. The index i indicates the corresponding aerosol number as was defined in Table 7 (first column).

Table 9: Rules used for defining the local aerosol type which have a high enough probability. This typing will result in a maximum combination of three aerosol types for a single observation. Whenever the S, δ combination lies outside any of the 3σ regime the type will be set to unknown (-1)

Rules	Types included	Type value
$P_{\max} > e^{-9/2}$ & $P_{i1} > 0.55$	i1	$1+2^{(i1)}$
$P_{\max} > e^{-9/2}$ & $P_{i1} < 0.55$ & $P_{i2} > 0.3$	i1,i2	$1+2^{(i1)}+2^{(i2)}$
$P_{\max} > e^{-9/2}$ & $P_{i1} < 0.55$ & $P_{i2} < 0.3$	i1,i2,i3	$1+2^{(i1)}+2^{(i2)}+2^{(i3)}$
No aerosols	0	0
Lidar ratio > 0 & $P_{\max} < e^{-9/2}$	0	-1

These indicated rules need to be validated and refined using available HSRL and in-situ datasets. As an example, if one assumes that these rules apply, it would result in the following type assignments in observation parameter space, when assuming a constant 5% error estimate in the lidar ratio and depolarization (Figure 17). When additional information concerning the marine and continental seasonal spread and height information is used or by taking into account additional information on available volcanic dust, the assignment of a single type will be far more common.

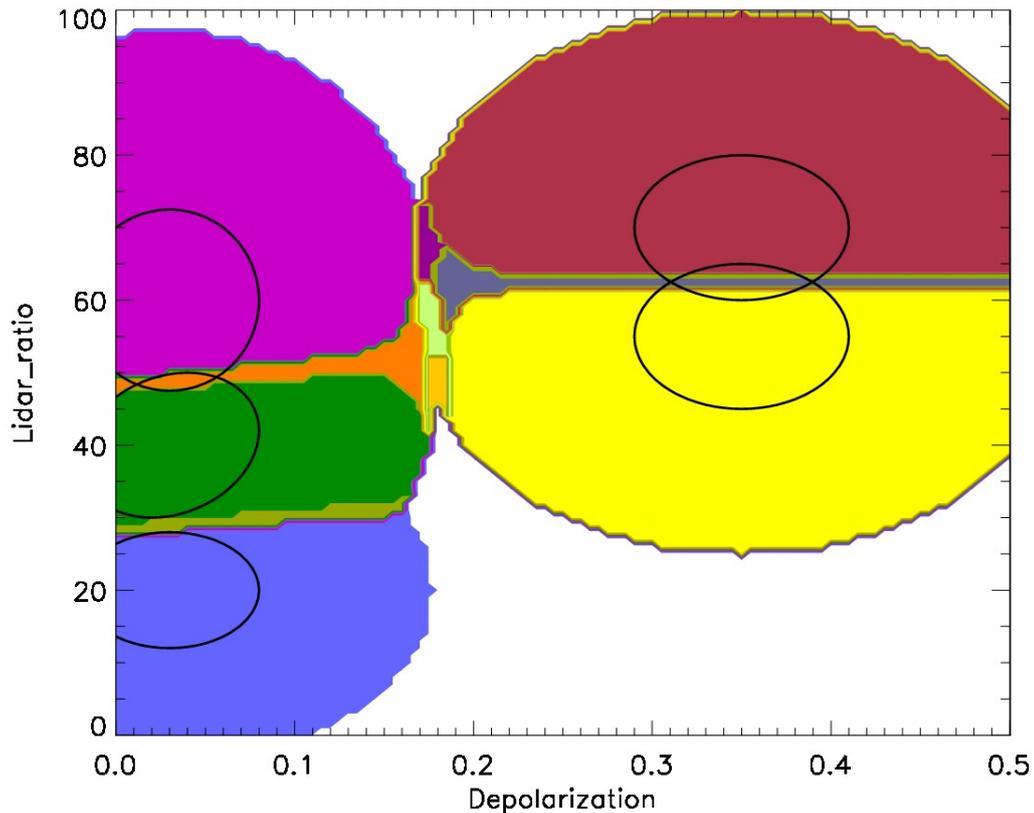


Figure 17: Detectable aerosol types in observation space assuming a fixed error estimates of 5% in lidar ratio and depolarization. The black ovals represent the 1σ contour lines for each of the individual probabilities. The different colors depict the relevant regions, within a 3σ minimum probability, e.g. the yellow region represents the 3σ dust region. All possible type combinations (incl. colours) are presented in Table 10.

The possible aerosol type combinations for which the minimum probability exceeds the 3σ levels are indicated in Table 10 (colours refer to Figure 17), with the integers indicating the byte combination.

Table 10: Possible aerosol type assignments for the probability distributions defined in Table 8, with 5% error estimates in the retrieved lidar ratio and depolarization values as presented in Figure 17. The colors match those in the figure.

1	Marine
2	Smoke/Pollution
4	Clean Continental
5	Clean Continental & Marine
6	Clean Continental & Smoke/Poll
8	Pure Dust
10	Pure Dust & Smoke/Pollution
12	Pure Dust & Clean Continental
16	Ash
18	Ash & Smoke/Pollution
24	Ash & Pure Dust

This does not mean that in case of an assignment of, e.g. 12 (2^3+2^2), there is per definition a combination of dust and smoke/pollution, it indicates that both the probabilities of dust and smoke/pollution are similarly large.

6. Algorithm performance, sensitivity studies, limitations

6.1.1. *L2a lidar classification*

The layer separation procedure has been coded into a prototype form. It is yet not fully integrated into the A-EBD algorithm. However, the prototype has been applied to a number of ECSIM test cases. A particularly relevant example is shown in Figure 18. Here an idealized thin supercooled water layer cloud embedded in a ice cloud depicted. The ice cloud has an extinction coefficient of 0.2 km^{-1} and the water cloud an extinction coefficient of $1. \text{ km}^{-1}$. Example backscatter ratio and depolarization ratio profiles derived from a one km horizontal average of the data are shown in Figure 19 along with the retrieved optimal layering structure. By looking at the top panel it can be seen that the improvement in the goodness-of-fit measure is rapid going from 1 to two layers, however, going from 2 to 3 layers yields only a relatively small improvement thus 3 layers was chosen as the optimal number of layers. It can also be seen that the optimal 3-layer structure correctly captures the separation between the ice layers and the water layer.

An example of the classification procedure output is shown in Figure 20. Here the ECSIM standard scene was used. In fact this figure corresponds to the example case presented in Figure 19-21 in the A-EBD-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.

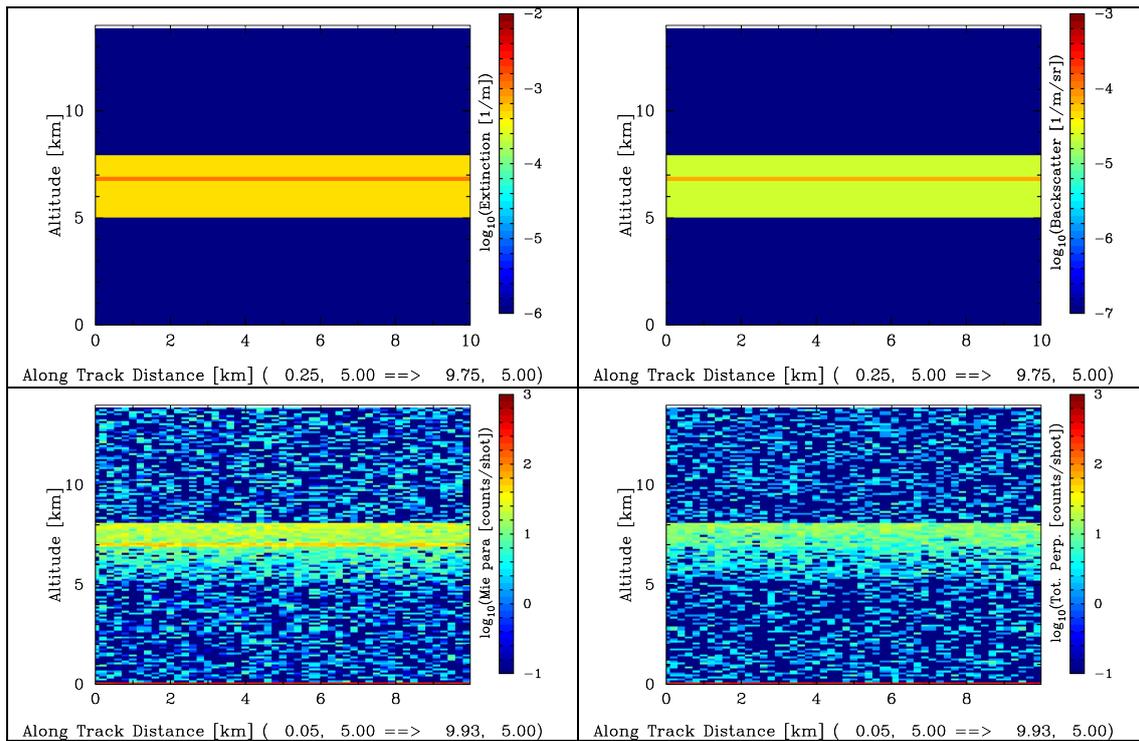


Figure 18: Top-Left: ECSIM extinction field corresponding to an idealized super cooled water layer embedded in an ice cloud. Top-Right: Corresponding backscatter signals. Bottom panels: Mie co-polar and total cross polar signals as calculated by ECSIM.

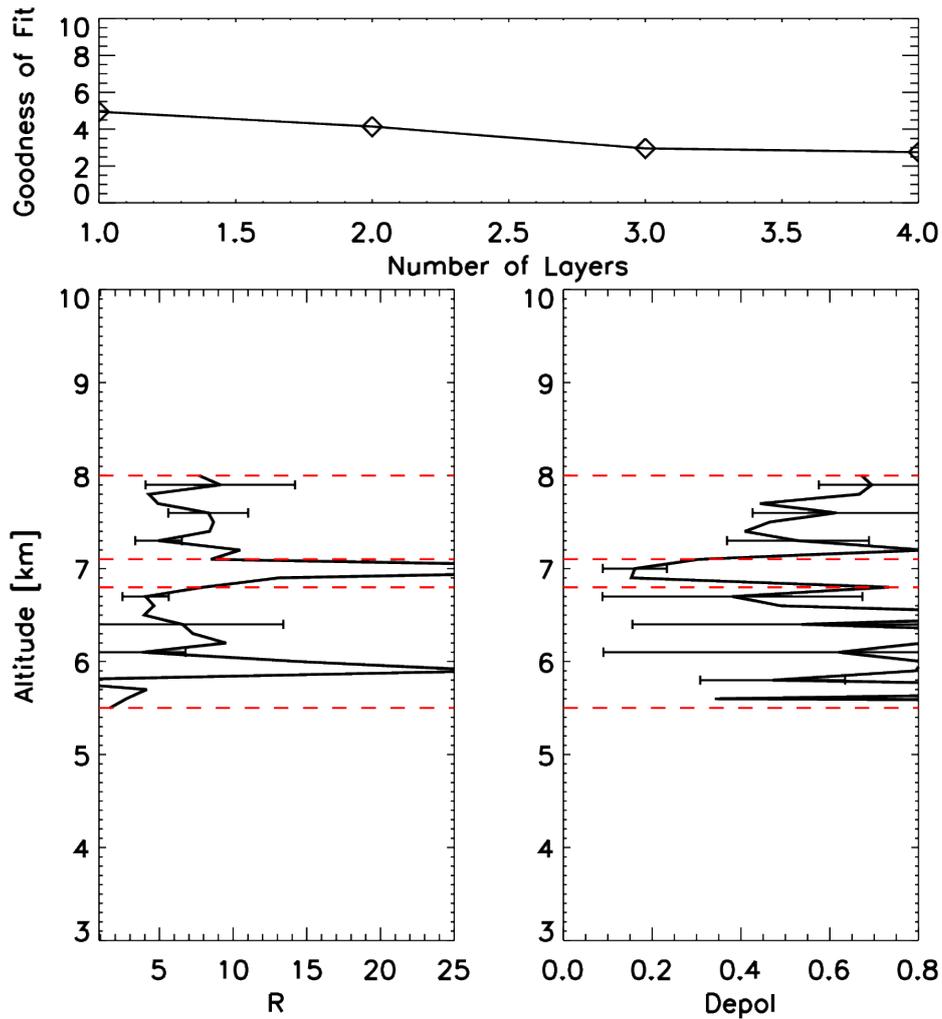


Figure 19: Bottom Panels: Sample Backscatter Ratios and Depolarization ratios derived from a horizontal average of the data shown in Figure 17 from 0 to 1 km. The Red-lines show the retrieved optimal layering. The top panel shows the decrease in the summed goodness-of-fit parameter with the allowed number of layers.

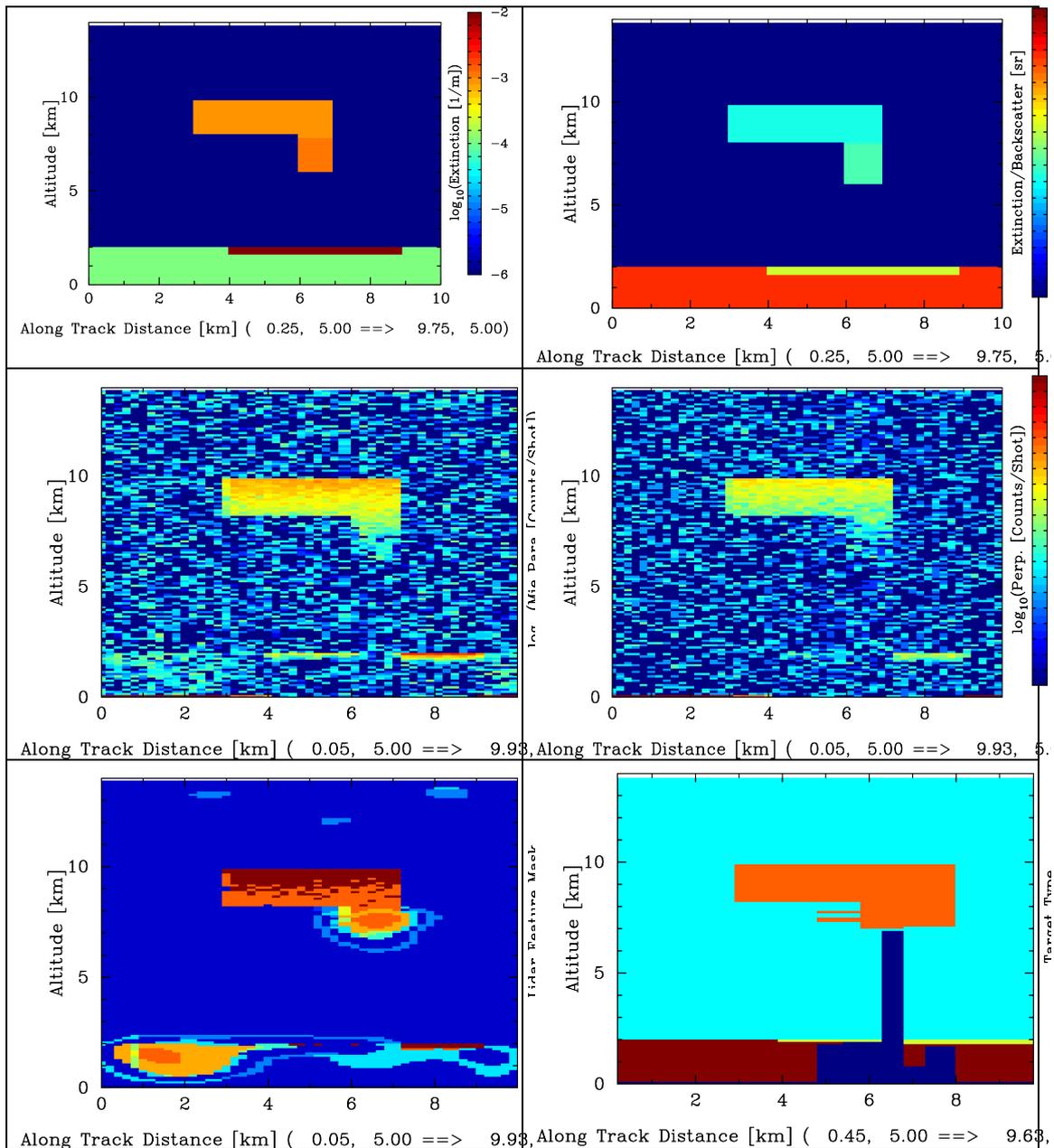


Figure 20: Top Left: Extinction field. Top Right Extinction-to-Backscatter field, Middle-Left: Mie channel signals. Middle-Right: Cross-polar channel signals. Bottom-Left: Feature mask output. Bottom-Right: L2a Target Classification. Here Red is Aerosol, Cyan is clear sky, Orange is ice cloud and Blue is unknown (Lidar signals are attenuated too much).

6.1.2. L2a lidar aerosol typing

The aerosol typing algorithm has been coded into a prototype form. It is not fully integrated with the A-AER and A-EBD algorithms. Instead the algorithm is evaluated using ECSIM scenes which were created within the ICAROHS project. Presented below are three examples from different ICAROHS scenes which were measured during three different campaigns, SAMUM1, SAMUM2 and EUCAARI (Weinzierl and van Zadelhoff 2011). These scenes were created using HSR lidar observations and in-situ measurements from the DLR Falcon and imported into ECSIM. From these scenes the retrieved lidar ratio and measured depolarisations are fed into the Aerosol typing algorithm. For each of the scenes a short description is presented with the backscatter ratio and resulting aerosol type images.

6.1.2.1. Saharan desert scene (SAMUM1)

The Saharan Mineral Dust Experiment (SAMUM) was dedicated to the understanding of the microphysical and optical properties of desert dust, and the impact of desert dust on the global climate system [Heintzenberg, 2009; Ansmann et al., 2011]. The dust layer measured on June 4th 2006 presented a clear horizontal and vertical gradient with a 3-layer structure over Ouarzazate. The different layers over Ouarzazate coincided with different potential temperature gradients and wind direction. The aerosols were identified as fresh dust measured close to its origin. The retrieved aerosol type mask indicates that the scene is dominated by dust (with some edges and surface classified as clean continental).

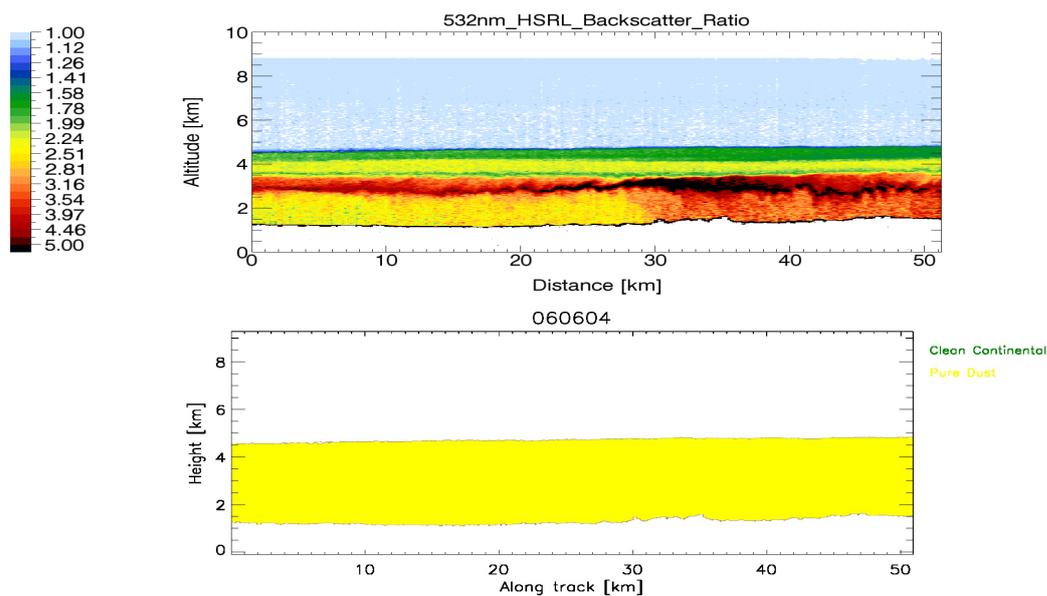


Figure 21: Top panel: Observed backscatter profile from the DLR Falcon HSRL. Bottom panel: retrieved aerosol type mask,

6.1.2.2. *Aged Saharan dust with Biomass burning (SAMUM2)*

The flight on 25 January 2008 during 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. The first 40km of this scene has been modelled within ECSIM.

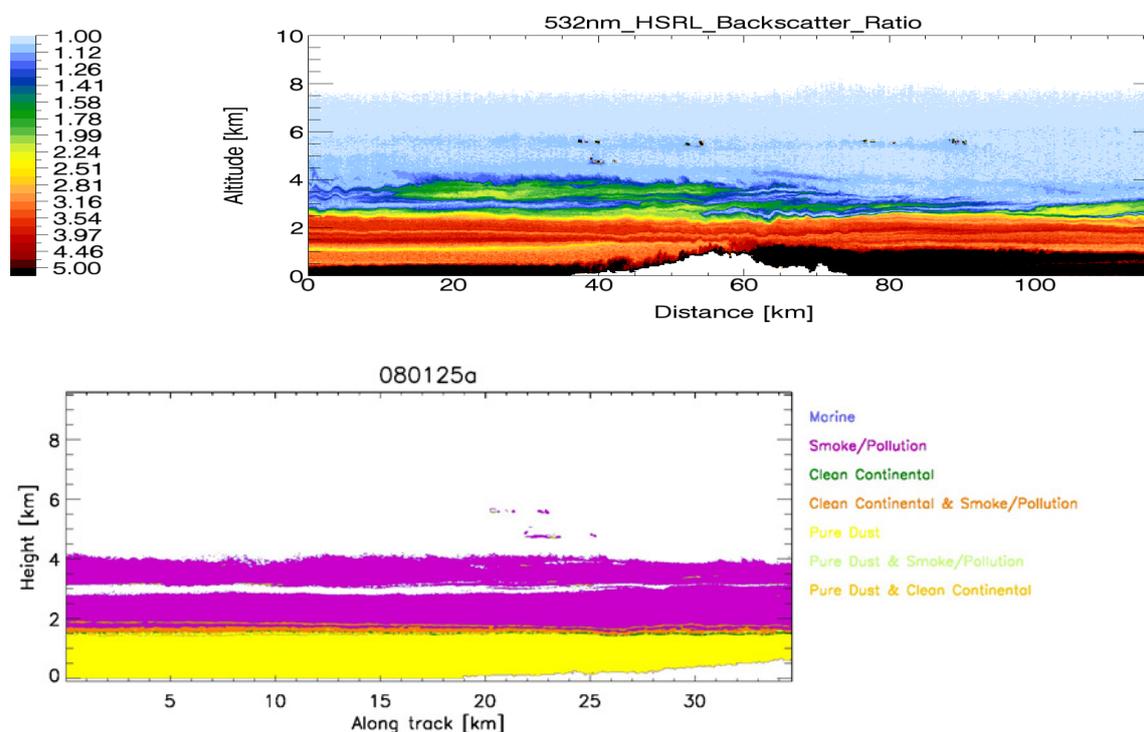


Figure 22: Top panel: Observed backscatter profile from the DLR Falcon HSRL. Bottom panel: retrieved aerosol type mask. Note the different horizontal scales..

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.

6.1.2.3. *Anthropogenic pollution (EUCAARI)*

In the morning of 14 May 2008, aged European pollution was measured by the Falcon and also by the British Aircraft BAe-146 southwest of Ireland. The DLR Falcon 20 operated in the upper troposphere to detect the pollution layers with lidar measurements. Subsequent to the measurements, the DLR Falcon measured a horizontal in-situ profile over the Atlantic southwest of Ireland flying into the pollution layers (Figure 23). The aerosol typing based on lidar ratio and depolarization detected two main regions, a marine layer in the bottom 200 m and an extended Smoke/Pollution layer on top as was indicated by the in-situ measurements.

The liquid cloud layers are either identified as marine aerosols (small droplets) or pure dust (high depolarization due to multiple scattering). These liquid layers would normally be flagged by the L2a classification as non-aerosols and would not be assigned a type. The clean continental detection is found within the noise below the detected liquid layers and should normally be flagged as not to be trusted by the L2a classification scheme.

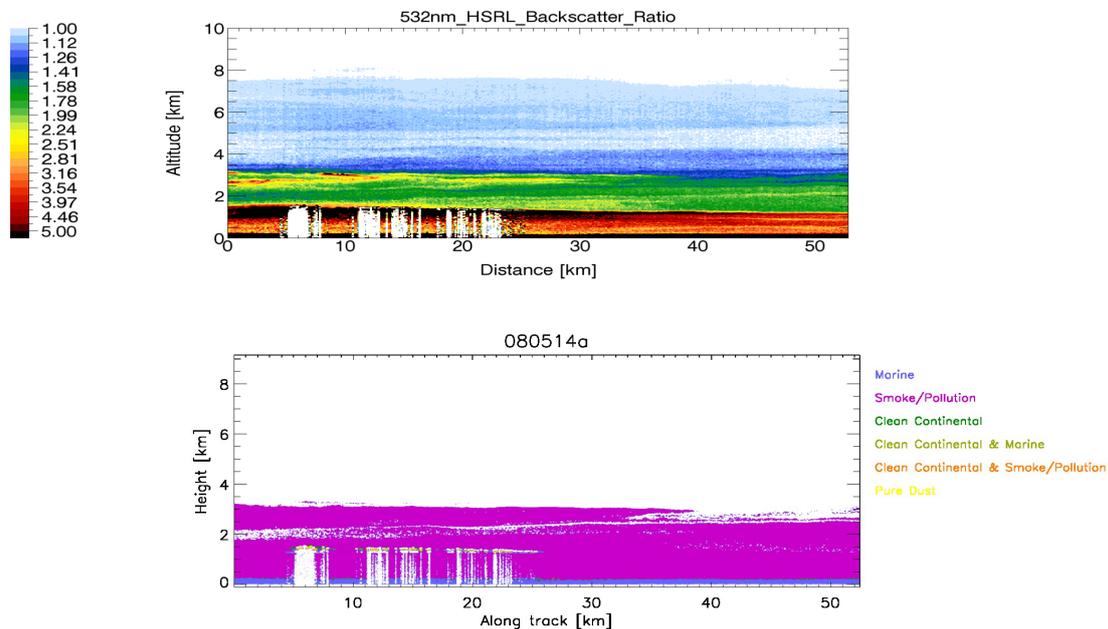


Figure 23: Top panel: Observed backscatter profile from the DLR Falcon HSRL on May 14th 2008 during the EUCAARI campaign. Bottom panel: retrieved aerosol type mask.

7. Validation status

7.1.1. *L2a lidar classification*

The method used here in order to separate statistically significant layers is rather general and likely 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 24). 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.

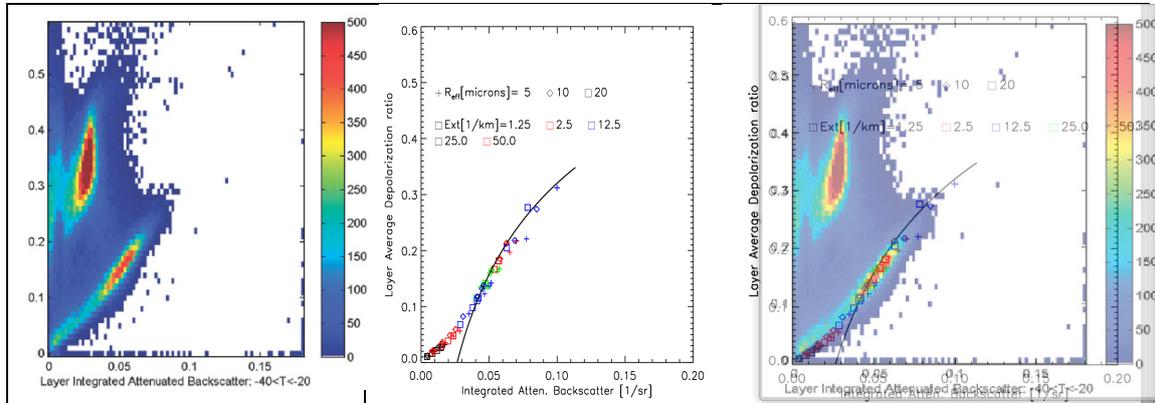


Figure 24: Left: Histogram built using CALIPSO observations taken from Hu et al. 2009. Middle: Same as Left panel of Figure 3. Right: Overlap of the other two panels.

7.1.2. *Aerosol Typing*

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.

Annex A: Technical implementation

L2a lidar Classification

The L2a lidar classification scheme described in this document exists in prototype form not quite at the time of this writing (Apr 2011) integrated into ECSIM. Since the procedure is not applied to the full vertical extent of the lidar signals but is only applied only to areas(layers) where targets have already been identified by the Target Mask algorithm or the large-scale aerosol algorithm it is expected that the brute-force but simple algorithm will be fast enough for operational use without any special developments.

Aerosol Typing

The L2a lidar aerosol typing scheme described in this document 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.

External models

TBD as development continues.
