ATLAS

ATLID Algorithms and Level 2 System Aspects

Algorithm Theoretical Basis Document (ATBD) for ACM-Classification

Code	:	EC-TN-KNMI-ATL-024
Issue	:	2.1
Date	:	25/04/2011
Authors	:	G-J van Zadelhoff J. Delanoe D.P. Donovan S. Berthier

This page intentionally left blank

Issue	Revision	Date	Reason for Change	Author
1.0	Baseline	2010-06-23	Baseline Issue	GJvZ, DD, SB [KNMI]
2.0	0	2011-04-08	Pre IR changes	GJvZ
2.1	1	2011-04-25	Final meeting changes	GJvZ

Table of Contents

1. Purpose and Scope

This document describes the algorithm for retrieving the synergetic lidar/radar/msi classification of the EarthCARE signals. The ATBD is set up according to the EarthCARE ATBD format. This algorithm will ingest Lidar, Radar and MSI data (L1d) as well as their respective target classifications (L2a) data in order to create a combined target mask and classification field. This product will be compatible both in terms of conventions and file format with the L2a Radar-only and L2a-Lidar-only products.

2. Applicable and Reference Documents

2.1. Applicable documents

Reference	*Code*	*Title*	*Issue*	*Date*
[MRD]	EC-RS-ESA-SY-012	EarthCARE Mission Requirements Document	5	Nov 2 2006

2.2. Reference & Related documents

Reference	*Code*	*Title*	*Issue*	*Date*
CASPER-PARD	CASPER-DMS-PARD-001	CASPER Products and Algorithms Requirement Document (PARD)	2.0	30-10-08
CASPER-FINAL	CASPER-DMS-FR-01	Cloud and Aerosol, Synergetic Products from, EarthCARE Retrievals, (CASPER), Final Report.	1.1	30-01-09
A-TC	EC-TN-KNMI-ATL-022	ATLID L2a target classification		25-04-11
A-EBD	EC-TN-KNMI-ATL-021	ATLID L2a Extinction, Backscatter & Depolarization		25-04-11
M-CM	IRMA_D5_ATBD_M_CM	MSI Cloud Mask	5.0	03-11
AM-ACD	IRMA_D5_ATBD_AM_ACD	ATLID-MSI Aerosol Column Descriptor	3.0	03-11
ATL-PARD	EC-TN-KNMI-ATL-005	ATLAS Products and Algorithms Requirements Document (PARD)	1.1	10-03-10
EarthCARE	EC-ICD-ESA-SYS-0314	EarthCARE product Table	1.3	15-06-10

2.3. References

Keyword	Reference
Delanoe 2010	Delanoe J, Hogan RJ. 2010. Combined CloudSat-CALIPSO-MODIS retrievals of the properties of ice clouds. J. Geophys. Res. 115: D00H29, qj.00010.1029/2009JD012346.
DARDAR	http://www.icare.univ-lille1.fr/projects/dardar/
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
Luke et al. 2007	Luke, Edward P., Pavlos Kollias, Karen L. Johnson, Eugene E. Clothiaux, 2008: A technique for the automatic detection of insect clutter in cloud radar returns. *J. Atmos. Oceanic Technol.*, **25**, 1498–1513. doi: 10.1175/2007JTECHA953.1
Mittermaier and Illingworth 2003	Mittermaier, M. P. and Illingworth, A. J. (2003), Comparison of model-derived and radar-observed freezing-level heights: Implications for vertical reflectivity profile-correction schemes. Quarterly Journal of the Royal Meteorological Society, 129: 83–95. doi: 10.1256/qj.02.19

3. Scientific Background of the algorithm

3.1. Algorithm history

The algorithm is based on the CloudNET classification (Hogan and O'Conner 2004) and DARDAR-mask algorithms (Delanoe 2010 & http://www.icare.univ-lille1.fr/projects/dardar/).

3.2. Algorithm introduction

This product facilitates the application of other synergetic algorithms. Principally, it identifies the nature of the targets in each pixel and highlights those that are bad or ambiguous, thereby informing subsequent algorithms where they should perform a retrieval (e.g. an ice cloud algorithm would only be applied to pixels containing ice cloud) and in some cases the confidence that they should assign to the observations at each pixel. In addition to facilitating application of other algorithms, it can also be used to derive cloud fraction and cloud overlap on arbitrary model-type grids

3.3. Physical/mathematical Background

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 defined as a decision-tree, as was recommended by the CASPER project [CASPER-FINAL] 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).

3.4. Summary of changes to the ATLAS algorithm

This version is the first of this algorithm

4. Justification for the selection of the algorithm

There are several options for storing classification, and our considerations in deciding on the system presented above have been:

1. Easy to interpret by the user
2. Flexible, allowing combinations of targets to be reported
3. Concise, so not using an excessive amount of storage space

The CloudNET "categorization" format, in which the 8 bits of a single byte were taken to indicate the presence or absence of up to 8 different types (e.g. rain, ice etc), was both sufficiently flexible and concise, yet turned out not to be easy to interpret by many users. Conversely, a simple list of possible types is usually not flexible, since usually not every possible combination of types is present in the list. Since only a single byte is required for any piece of classification information, we have chosen to present the same information in two ways, so that the user can choose which to use. Thus we have *Classifications by type*, for example the "liquid classification" variable which describes what we can say about liquid cloud in a particular pixel. In the Cloudnet system, each of these types would be represented by just a single bit, and therefore it was not possible to store, for example, when it was unknown if a certain type was present. We also have the *Summary classification* variable, which provides the main types in a form that can be easily plotted (without too many possible colours).

To determine what combinations are possible, it is convenient to consider that since there are essentially only two active measurements being used (radar reflectivity and lidar backscatter), the number of possible combinations must also be limited. In fact, the lidar predominantly provides information on the presence of mutually exclusive types such as aerosol and liquid cloud, while the radar predominantly provides information on the presence of mutually exclusive types such as ice and rain.

5. Mathematical algorithm Description

5.1. *Input parameters*

The L2b target classification get its input from 6 collocated files, the lidar and radar classification files, the merged L1d signals, the lidar extinction & backscatter and results from two MSI algorithms, which can be combined to create the L2b classification mask (see Figure 1 for an overview of the currently envisioned EarthCARE data stream).

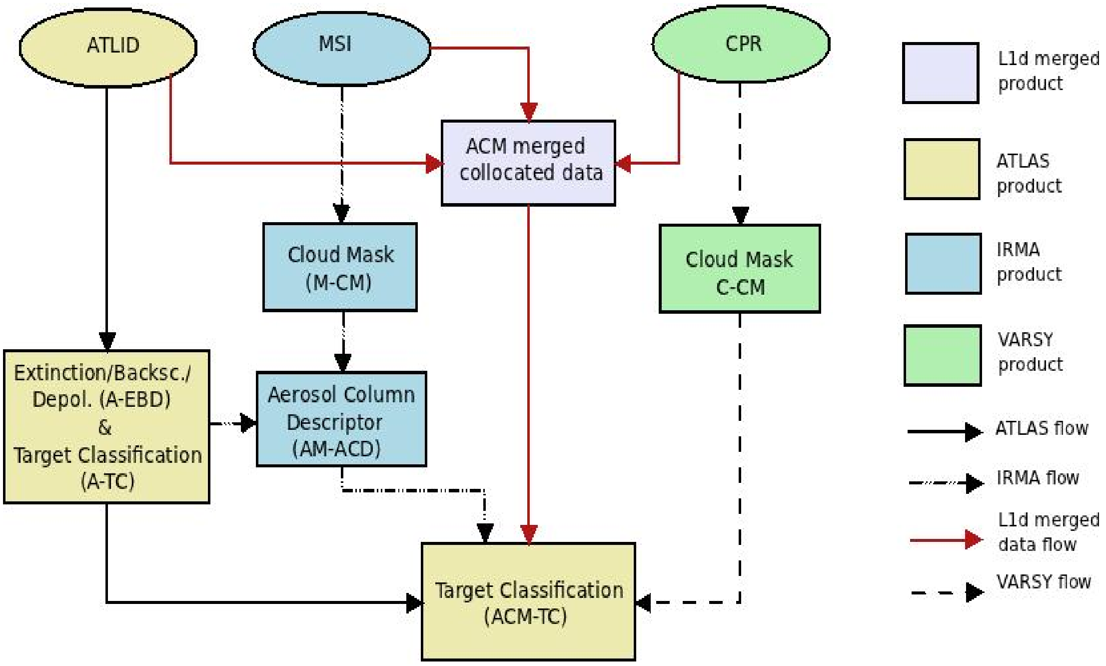

Figure 1: Overview of the data streams entering the L2b ACM-target classification. The different colours refer to the data needed from different projects.

Given in Table 1 are all the relevant input arrays needed by this algorithm. All these parameters are expected to be mapped on the collocated grid.

Table 1: Input parameters from the L1d, A-TC, M-CM, AM-ACD & C-CM output files. Some parameters come directly from the ECMWF model output, but have been previously used in another algorithm. In those cases the combination of the algorithm and ECMWF are used as source.

Variable	Symbol	Description	Unit	Dim	Source	Type
Time	t_{utc}	UTC time	S	Time	A-TC	Real*8
Latitude	Lat	Latitude of co-located CPR+ATLID footprints at the ground	degree	time	A-TC	real
Longitude	Lon	Longitude of co-located CPR+ATLID footprints at the ground	degree	time	A-TC	real
Height	z	Height of each radar/lidar gate above mean sea level	M	Time	A-TC	Real
Surface_Altitude	z_s	Height of surface above mean sea level	M	time	A-TC/ ECMWF	real
Tropopause_Altitude	Z_{TP}	Height of the tropopause above mean sea level	M	time	A-TC/ ECMWF	real
Day_night	DN	Indicating the night and day status	-	Time	A-TC	Integer
Temperature	T	Temperature	K	Time, height	ECMWF (L1d)	Real
Pressure	P	Pressure	Pa	Time, Height	ECMWF (L1d)	Real
Relative_Humidity	RH	Relative Humidity	None	Time, Height	ECMWF (L1d)	Real
Wet_bulb_Temp	T_{wb}	Wet Bulb Temperature	K	Time, Height	ECMWF (L1d)	Real

radar_detection_status	C_{det}	0 = radar not working 1 = ground detected 2 = totally extinguished 3 = clear 4 = target detected 5 = dominated by multiple-scattering 6 = don't know	none	time, height	C-CM	byte
lidar_detection_status	A_{det}	0 = lidar not working 1 = ground detected 2 = totally extinguished 3 = clear 4 = target detected 5 = molecular 6 = don't know	None	time, height	A-TC	byte
msi_detection_status	M_{det}	0 = msi not working 1 = ground detected 2 = target detected 6 = don't know	None	time	M-CM	byte
aerosol_classification	A_{aer}	Aerosol classification by type(incl auxiliary data)	None	time, height	A-TC	Int
aerosol_classification_direct	A_{aerd}	Aerosol classification by type (lidar only data)	None	time, height	A-TC	Int
Ice_classification	A_{ice}	Ice class. by sub-type	None	time, height	A-TC	Int
liquid_classification	A_{liq}	Liquid class. by sub- type	None	time, height	A-TC	Int

Classification_prob	A_{prob}	Aerosol type prob. (incl auxiliary data)	None	Ntypes, time, height	A-TC	Int
Classification_prob _direct	A_{probd}	Aerosol type prob. (lidar only data)	None	Ntypes, time, height	A-TC	Int
Lid_simplified_ classification	A_{class}	Lidar classification of different particle types	None	time, height	A-TC	Int
Cloud_flag	CM_F	Cloud flag	None		M-CM	Int
Cloud_mask_type	CM_{typ}	Cloud mask and type	None		M-CM	Int
Cloud_phase_1	CP_1	Cloud thermodynamic phase	None		M-CM	Int
	F_{lm}	Land/water Flag	None		M-CM	Int
Cloud_flag_quality	F_{CM_F}	Quality flag for the cloud flag	None		M-CM	Int
Cloud_phase_quality	F_{CP_1}	Quality flag for the cloud thermodynamic phase 1	None		M-CM	Int
Cloud_mask_quality	F_{CM_typ}	Quality flag for the cloud mask and type	None		M-CM	Int
AOT_o (incl error)	AM_δ_0	Spectral aerosol optical thickness over ocean	None		AM-ACD	Real*3
AOT_l (incl error)	AM_δ_L	Spectral aerosol optical thickness over land	None		AM-ACD	Real*2
AE_o (incl error)	AM_a_o	Ångström exponent over Ocean	None		AM-ACD	Real*2
AE_l (incl error)	AM_a_L	Ångström exponent over Land	None		AM-ACD	Real
A_type	AM_t	Aerosol type	None		AM-ACD	Int

Q_A_type_o	AM_F_o	Quality flag for aerosol type over ocean	None		AM-ACD	Int*3
Q_A_type_l	AM_F_l	Quality flag for aerosol type over land	None		AM-ACD	Int*3
Rad_simplified_ classification	C_{class}	Radar cloud & precip mask	None	Time, height	C-CM	Int
Beta_ray (incl error)	L1_β_r	Measured Rayleigh backscatter	1/m/sr	Time, Height	L1d	Real
Beta_mie (incl error)	L1_β_m	Measured Mie backscatter	1/m/sr	Time, Height	L1d	Real
Beta (incl error)	A_β	backscatter	1/m	Time, Height	A-EBD	Real
Ext (incl error)	A_α	Extinction	1/m/sr	Time, height	A-EBD	Real
Depol (incl error)	A_ρ	Depolarization Ratio	None	Time, Height	A-EBD	Real
Z_e	C_Z	Radar reflectivity	dBZ	Time, Height	L1d	Real
V_dop	C_V_D	Doppler velocity measured by radar	m/s	Time, height	L1d	Real

5.2. Configuration parameters

Table 2: Configuration parameters ACM-Classification

Variable	Symbol	Description	Unit	Dim	Type
Rain_thresh	R_{th}	The rain threshold indicating the minimum required radar signal to detect rain when $T_{wb}<0^oC$	dBZ	1	real
Rain_evap	R_{ev}	Threshold to indicate if the rain will reach the surface. For $Z< R_{ev}$ the rain is evaporating and will not reach the surface	dBZ	1	real
Dz_melt	D_{melt}	Temperature regime in which the local maximum of the radar is searched for around the $T_{wb}=0^oC$ isotherm to find the melting layer	K	1	real
H_melt	H_{melt}	Horizontal width to search for retrieved melting layers in case of non-determination within the column	m	1	real

5.3. Output parameters

Table 3: Operational output parameters for the ACM-classification

Variable	Description	Units	Dim	Type	Destination
1D Coordinate variables					
Time	UTC time	S	time	double	ACM-CAP
Latitude	Latitude of co-located CPR+ATLID footprints at the ground	degree	time	real	ACM-CAP
Longitude	Longitude of co-located CPR+ATLID footprints at the ground	degree	time	real	ACM-CAP
Height	Height above mean sea level	M	height	real	ACM-CAP
Geographic information					
surface_altitude	Height of surface above mean sea level	M	time	real	ACM-CAP

Variable	Description	Units	Dim	Type	Destination
Tropopause altitude	Height of the tropopause above mean sea level	M	Time	Real	ACM-CAP
Instrument detection status (8-bit unsigned integers)					
radar_detection_status	0 = radar not working 1 = ground detected 2 = totally extinguished 3 = clear 4 = target detected 5 = dominated by multiple-scattering 6 = don't know	none	time, height	byte	ACM-CAP
lidar_detection_status	0 = lidar not working 1 = ground detected 2 = totally extinguished 3 = clear 4 = target detected 5 = molecular 6 = don't know	None	time, height	byte	ACM-CAP
msi_detection_status	0 = msi not working 1 = ground detected 2 = target detected 6 = don't know	None	time	byte	ACM-CAP
Meteorological information (*Table 4*)					

Table 4: Meteorological classification, assignment of particle types and classes

Classifications by type (8-bit unsigned integers)	
aerosol_classification	0 = ground 1 = none 2 to 8 9 = don't know
Ice_classification	0 = ground 1 = none 2 = ice or snow 3 = melting ice 9 = don't know
rain_classification	0 = ground 1 = none 2 = warm rain, originating from collision and coalescence in liquid-only clouds 3 = rain originating from melting ice 9 = don't know

Liquid_classification	0 = ground 1 = none 2 = warm liquid cloud (wet-bulb temperature > 0°C) 3 = warm but only inferred from radar (lidar extinguished) 4 = supercooled 9 = don't know (lidar attenuated and T > –40°C)
Insect_classification	0 = ground 1 = none 2 = yes 9 = don't know
convective_classification	0 = ground 1 = none 2 = warm convective core (wet-bulb temperature > 0°C) 3 = cold convective core 9 = don't know
Class Probabilities (array of integer (integers) Dimension (nx, nz, n_classes)	
Classification probabilities_direct	0-100 for each allowed lidar only class Allowed classes are: 1. Surface 2. No target present 3. Water 4. Super Cooled Water 5. Ice 6. Aerosol type 1 7. Aerosol type 2 8. . 9. Etc For any height bin, a negative entry present in each category means that no estimate is possible (i.e. lidar signal is completely attenuated)

Classification probabilities *(aerosol classes derived using a priori aerosol type map information. Other classes are preserved for consistency)*	0-100 for each allowed lidar only class Allowed classes are: 10. Surface 11. No target present 12. Water 13. Super Cooled Water 14. Ice 15. Aerosol type 1 16. Aerosol type 2 17. . 18. Etc For any height bin, a negative entry present in each category means that no estimate is possible (i.e. lidar signal is completely attenuated)
Summary classifications (8-bit unsigned integers)	
Classification	0 = ground 1 = clear sky 2 = liquid cloud 3 = ice only 4 = ice and supercooled liquid 5 = ice but can't tell if liquid too (lidar extinguished) 6 = melting ice 7 = rain 8 = rain and liquid cloud 9 = aerosol 10 = insects 11 = stratospheric 12 = convective core 13 = don't know

5.4. Algorithm flow charts

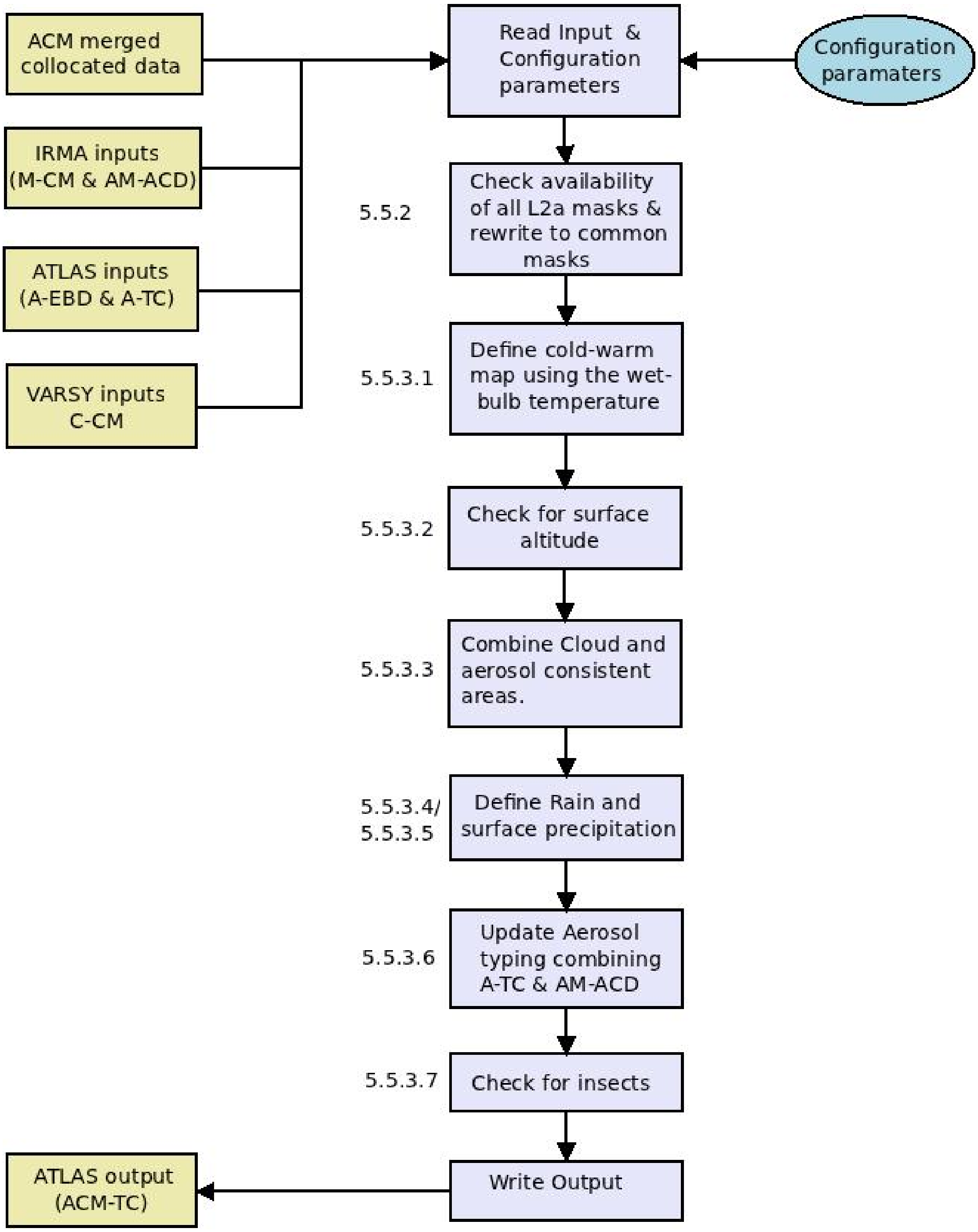

Figure 2: Flow diagram of the ACM-TC algorithm. Yellow boxes indicate input and output files, the oval reflects the configuration parameter file and the grey boxes the steps within the algorithm. The number next to the grey boxes refer to the section numbers in which the topic is explained.

5.5. *Algorithm Definition*

5.5.1. *Instrument availability*

From each of the L2a target masks there will be an instrument detection status (A_{det}, C_{det} and M_{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 can not 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.

5.5.2. *L2a Target classifications available*

Locally the information from the different L2a target classifications will be stored presented in Table 5, Table 6 and Table 7.

Table 5: Lidar Input classification

Classification	Name	Lidar
surface/sub-surface	Surface	0
clear sky	Clear	1
liquid cloud	Liquid	2
ice only	Ice	3
Aerosol	Aerosol	9
Stratospheric	Strat	11
Don’t know/fully attenuated	unknown	13
No data	Nodata	14

Table 6: Radar input classification

Classification	Name	Radar
surface/sub-surface	Surface	0
No clouds	Clear	1
Cloud or Rain	Cloud	2
Unknown	unknown	13
No data	Nodata	14

Table 7: MSI input based on th*e* L2a-M-CM (Cloud map), note that some of the data availability depends on the day-night flag. The information within the input files will be summarized as:

Classification	Name	MSI
No clouds	Clear	1
Cloud	Cloud	2
Aerosols	aerosol	7
Unknown	unknown	13
No data	Nodata	14

5.5.3. Combining the L2a classifications

At this point it is assumed that the different L2a masks exist or, in the case of a missing instrument, all values for the respective mask are set to nodata. In the following subsections the information from the different classifications is combined into the L2b classification. 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 results in a more detailed typing during day-time only. In Table 8 the combined classification results from the different combinations from the lidar Table 5 and radar Table 6 inputs are presented. These results reflect the 'simple categorization' as defined in the PDD and Table 4. From this the complete categorization can be defined by filling in the sub-classifications (e.g. rain_classification).

Table 8: 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. The rules are explained in more details in the Sections 5.5.3.2 up to 5.5.3.7.

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 K
Liquid	Unknown	Cloud	T_{wb} > 273 K
Ice	Ice	Cloud or Clear	T_{wb} < 273 K
Ice	Unknown	Cloud	T_{wb} < 273 K
Supercooled	Liquid	Clear	T_{wb} < 273 K
Supercooled + ice	Liquid	Cloud	T_{wb} < 273 K
Aerosols	Aerosols	Clear	-
Rain	Unknown	Cloud	C_Z_{max} >-17dBZ, T_{wb} >273K
Liquid cloud +Rain	Liquid	Cloud	C_Z_{max} >-17dBZ, T_{wb} >273K

Insects	Liquid	Cloud	C_Z< -20 dBZ, T > 283K
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 distinguish the liquid cloud from ice cloud regimes. As this distinction between cold and warm areas is needed for a number of tasks it will be discussed first (Section 5.5.3.1)

Since the lidar and radar show the vertical layering the basic type characterization is defined by these two data streams. As most of the information is available at the input level, the combined typing is basically the result of a number of 'where' statements. The specific statements and reason are explained in the Sections starting with 5.5.3.2.

5.5.3.1. Cold & warm regions

At the start of the procedure the warm_cold field is created. This field indicates the regions where the wet-bulb temperature (T_{wb}) is smaller or larger than 273.15 K.
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 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.

The warm_cold field is initially defined solely by the wet-bulb temperature T_{wb}, however in order to cope with isothermal layers, it is implemented such that all pixels below the highest 0^oC isotherm in the profile are deemed to be “warm”, since melted ice precipitation is unlikely to refreeze. This field is then refined using the radar to locate the melting layer more precisely in stratiform precipitation. The radar reflectivity profile usually provides a distinct step (there is not always a “distinct step” but the Doppler velocity can be used as well to identify the melting region since the melted ice particles clearly accelerate) at 94 GHz where ice particles melt (see Mittermaier and Illingworth 2003) . The radar reflectivity profile is checked around the T_{wb}=273.15K region, if the radar profile is continuous and there is a local maximum in the radar reflectivity within *1K ($D_{melt:}$: to be defined after validation)* degree the warm_cold field is adjusted to the detected melting layer.
As a future improvement the horizontal continuity of the 0^oC isotherm has to be taken into account. For those profiles for which no melting layer are determined an interpolated value can be derived by taking into account the local radar maxima in surrounding profiles within H_{melt} km.

5.5.3.2. Surface definition

Surface information comes from three sources. The surface altitude is defined from:

- Lidar surface return
- Radar surface return
- Database surface altitude

where the order defines its importance, e.g. the lidar has the smallest vertical width information and any surface detection is therefore the most accurate, the signal however will be very often extinguished by clouds. The centre of the radar surface return is used when the lidar signal does not reach the surface and finally the input longitude, latitude surface map is used when the radar is either fully attenuated or can not be trusted due to enhanced multiple scattering within the beam.

The L2b-class field is set to "Surface" for all pixels up to the defined surface altitude

.

5.5.3.3. Check for Cloud & aerosol consistent areas and cloud phase definition.

The lidar and radar probe different parts of the particle size regimes, e.g radar can only detect larger particles (R_{eff}>~15 µm), whereas the lidar is attenuated very rapidly. The cloud regions, detected by the respective instruments, overlap at best but neither can determine all clouds solely. The combination of the two signals is needed to classify the different particle type regimes. To enable this classification a number of logical checks are implemented within the algorithm, which should be followed in the presented order as some regions will be overwritten when additional data is added.

The checks performed can be summarized using the following statements (see Table 8). The first step is to check for all the clear sky (molecular only) pixels:

***L2b class is set to Clear*: where lidar_mask=clear and radar_mask=clear or unknown**

Next, the different cloud regimes have to be checked. In general, liquid clouds will not be detected by the radar and should be considered as liquid layers with precipitation, however by setting them as liquid clouds first and later overwrite the relevant pixels as rain will results in defining both liquid layers and rain correctly. In the case of liquid layers detected by lidar the T_{wb} will determine if these are liquid clouds or supercooled layers.

***L2b class is set to Liquid*: where (lidar_mask=liquid and radar_mask=clear, cloud or unknown) or (lidar_mask=unknown and radar_mask=cloud) and $T_{wb}>0^oC$**

***L2b class is set to Ice*: where (lidar_mask=ice and radar_mask=cloud or clear) and $T_{wb}<0^oC$**

***L2b class is set to ice)*: where (lidar_mask=unknown and radar_mask=cloud) and $T_{wb}<0^oC$**

In the latter case the liquid_classification bit is set to 'don’t know (lidar attenuated and T > –40°C)'

***L2b class is set to ice clouds+super cooled liquid*: where (lidar_mask =liquid and radar_mask=cloud) and $T_{wb}<0^oC$ and $T>-40^oC$**

There is an additional check which can be used to check for supercooled layers. Experience from the CloudSAT and Calipso missions shows that some supercooled layers can be identified as ice only if embedded in the cloud when lidar has lost most of its signal. A possible way to avoid missing this layer is to check the neighbouring profiles to check for a supercooled layers around the same altitude. This procedure will have to be investigated in a future version of the algorithm.

***L2b class is set to aerosols* : where (lidar_mask=aerosols and radar_mask=clear)**

5.5.3.4. Rain

The above set of rules mainly distinguishes the different types of clouds but so far neglects precipitation and its determination. There are two types of precipitation which are distinguished in the rain-typing. Cold-rain and warm-rain, both referring to the type of clouds in which the rain originates. In both cases the lidar will be extinguished so fast that only a liquid layer can be determined. Rain assignment will therefore depend on radar signals only.

In the warm-rain case it is relatively easy, the liquid water cloud itself has small droplets, for which the top is detected by the lidar, if the beam is not extinguished by higher cloud layers. The radar will start to detect the droplets when they grow into rain-droplets. The main issue is what the radar reflectivity threshold should be to determine what is rain or not. The current version of the DARDAR mask assumes a threshold of -17 dBZ (R_{th}), which comes from private communications between R. Hogan and J. Delanoe. The same value is used for now in the current algorithm. To determine the availability of rain first the maximum reflectivity (C_Z_{max}) within the warm region of the column is determined. If this maximum is larger than R_{th} it is assumed that the region below the $T_{wb}=0^oC$ isotherm is raining.

***L2b class is set to rain*: where (lidar_mask=unknown and radar_mask=cloud) and $C\_Z_{max}>-17$ dBZ and $T_{wb}>0^oC$**

The top of the warm rain-clouds are set to liquid and rain as long as the lidar is not fully attenuated.

L2b class is set to rain and liquid cloud: **where (lidar_mask=liquid and radar_mask=cloud) and C_Z_{max}>-17 dBZ) and $T_{wb}>0^oC$**

In the cold-rain case the rain falls from ice-clouds, where the melting layer defines the place at which the ice particles melt. As the ice cloud shows high radar reflectivity values, similar to the rain droplets one cannot simple use the threshold value directly. Instead the cold-warm area defined by the $T_{wb}=0^oC$ and the local radar maximum is needed, as presented in Section 5.9.1.1.

L2b class is set to rain**: where (lidar_mask=unknown and radar_mask=cloud) , $T_{wb}>0^oC$ and shows a continuous profile with an ice cloud layer in the , $T_{wb}<0^oC$ region.**

5.5.3.5. Surface rain and cloud boundaries

In the case where the lidar is completely extinguished and does not reach the surface there can be no detection of rain and clouds down to the surface due to the large radar pulse length. The large radar return from the surface dominates all other cloud & rain returns up to ~750 meter [z(Z_surf)] in the case of CloudSAT and approximately 500m for EarthCARE. The return from the surface can be very well described by a Gaussian (see Figure 3). By fitting the surface return, using a Gaussian fit, the radar reflectivity in the pixels above the surface return can be checked. When the three smoothed pixels, adjacent to the Gaussian fit, show an excess to the Gaussian fit > 5 (tbd) dBZ, the rain or ice cloud is extended all the way to the surface. In the case of an ice cloud it is assumed to be snow when reaching the surface.

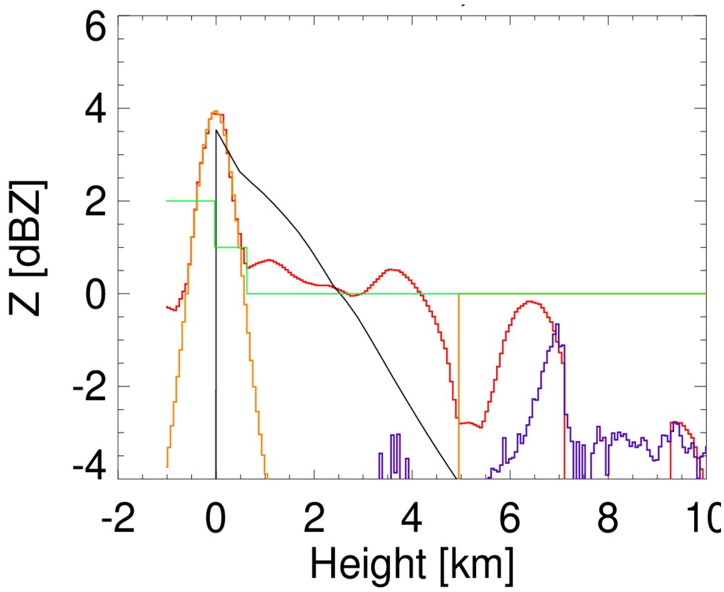

Figure 3: Single profile of CloudSAT data (red) divided by 10 and over plotted the collocated log(532nm CALIOP data) (blue). The black line indicates the T_{wb} profile, the orange line a Gaussian fit to th e surface return and green the surface mask. Shown is a profile of a supercooled layer (at ~7km) with an ice cloud layer down to ~2.5km (T_{wb}<0). From the ice cloud it rains down to the surface.

Rain does not always reach the surface. To check for evaporating rain the radar backscatter above the Gaussian fit to the surface return has to be larger than -25dBZ (R_{ev}). For $C\_Z<R_{ev}$ the classification is set to unknown.

5.5.3.6. Additional MSI information on the aerosol typing and check for missing layers.

In most of the cases the aerosol regions found by the A-EBD and A-TC algorithms will define the regions within the L2b-classification. However there are two checks which can result in an improvement of the aerosol typing by using the results from the ATLID-MSI L2b Aerosol Column Descriptor (AM-ACD) algorithm. Within this algorithm the aerosol type is defined for different layers using the aerosol optical thickness (AOT) and the Angstrom exponent (AE), which is the characteristic aerosol parameter containing information on the mean size of the particles. Small particles result in AE>1 and larger particles (dust and ash) AE<0.5. The AE can only be assigned during daytime as it uses the MSI visible channels at 670 and 865nm. Within the AM-ACD ATBD the following look-up table is used:

Table 9: Aerosol typing look-up table as defined in the AM-ACD ATBD

	AE ≤ 0.6	**0.6 < AE < 1.0**	**≥1.0**
AOT ≤ 0.15	Clean marine (t = 1)	No specific type (t = 0)	Clean continental (t = 2)
0.15 < AOT < 0.3	No specific type (t = 0)	No specific type (t = 0)	No specific type (t = 0)
AOT ≥ 0.3	Desert dust (t = 3)	No specific type (t = 0)	Biomass burning, pollution (t = 4)

This information can help separate the clean marine and clean continental for relatively large lidar ratios (~25-30) and the biomass burning and desert dust. The polluted continental and biomass burning is also in this case to difficult to distinguish from each other.

To keep the aerosol typing procedure equivalent between the L2a and L2b classification the four regimes from Table 9 should be represented as probability distributions. The distributions can be multiplied (relatively) to the L2a type probabilities and check whether this would result in a better aerosol typing.

A second and maybe more important check is to see if an aerosol type is detected when the L2a classification determined clear sky. In those cases where this happens the AM-ACD type will be set in the ACM-classification.

5.5.3.7. Detection of insects

The increased sensitivity of the EarthCARE CPR in comparison to the CLOUDSAT radar might give rise to the detection of insects. Insect radar returns at 35GHz range from -35dBZ to 0dBZ (Luke et al. 2007) at the ARM-SGP site, showing a large variability of the signal strength. At 94GHz the insect radar return was found to be almost 20dBZ lower, giving values < -20 dBZ. The big difference between the observed radar reflectivity's can be explained by the non-Rayleigh scattering of the radar signals by insects. As precipitation can be found in the same regime, insects can only be found by combining the lidar and radar results. The lidar return from insects will be, due to their size, very small and hardly any attenuation is expected in this regime. To be able to trust the non-detection one would need to see a clear surface return from the lidar.

In previous work it was found that the local temperature can be defining for the availability and height to which insects rise (e.g. Luke et al 2007). In general it was found that the insect-layer top height coincides with the 10°C isotherm, with some exceptions to the case.

***L2b class is set to insects*: where (lidar surface return is detected and radar_mask=cloud and lidar_mask=clear and T>10°C)**

In case of insects above clouds the lidar signals should detect no Mie signals and no decrease within the Rayleigh signals. Also the smearing of the radar signal due to the long pulse length has to be taken into account before assigning the insect flag to the radar return. In general more than one pixel is needed to be sure that insects are detected.

6. Algorithm performance, sensitivity studies, limitations

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 4 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.

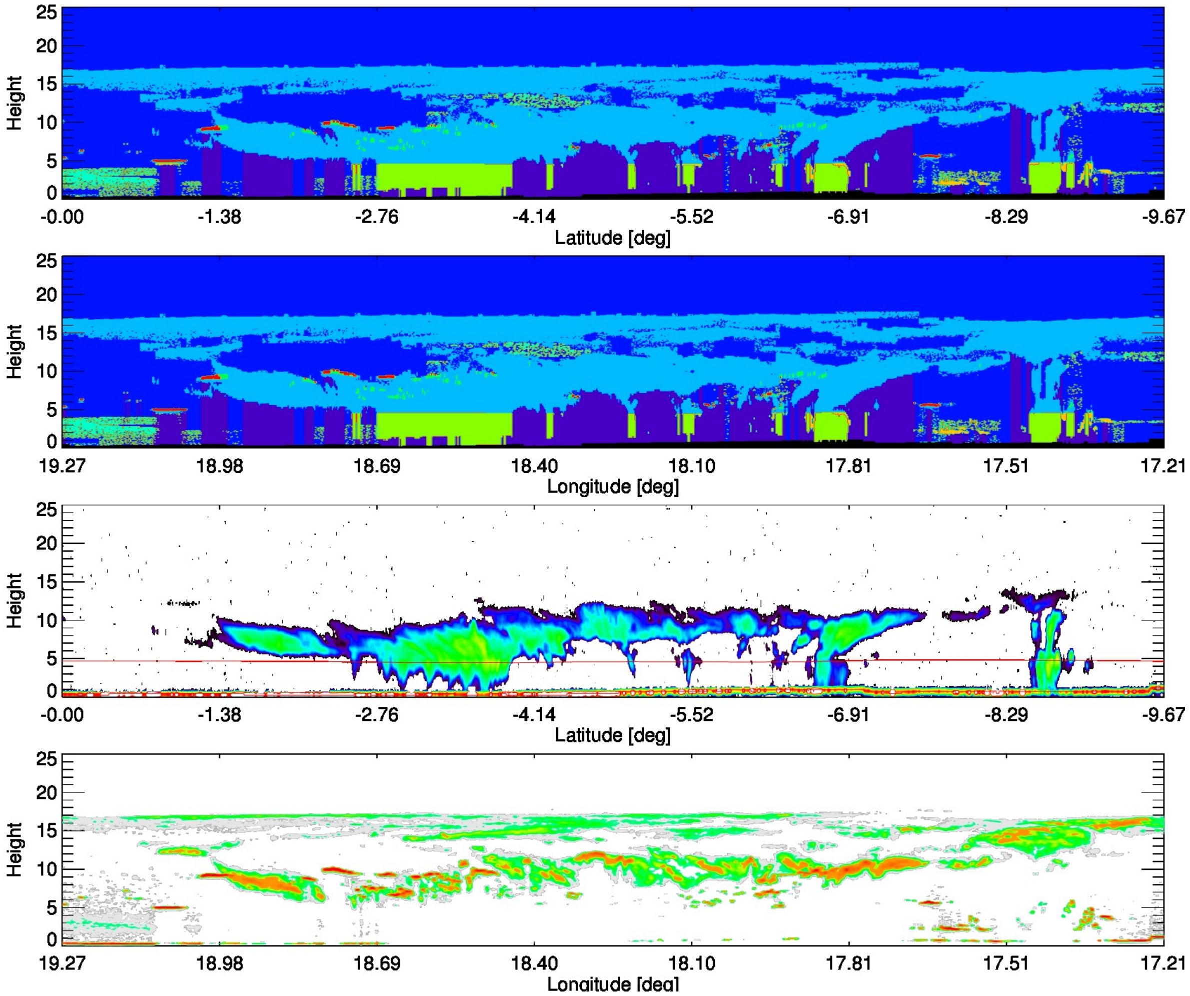

Figure 4: ACM-Target Classification example using CloudSAT and Calipso data measured on 2010-04-19. The top figure shows the result of the ACM-TC, the second is the DARDAR mask, third the CloudSAT radar data (the read line indicates the $T_{wb}=0^oC$ isotherm) and the bottom figure the 532nm CALIOP data. 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

A second example (Figure 5) focuses on a mid-latitude system (latitude between -31 to -33) off the coast of South Africa. There are no high lying cirrus layers nor large aerosol regions. The cloud system has super-cooled layers along the entire top (red when there is no radar, light green when there is additional radar data). There is a small liquid cloud layer on the left side of the main system, on the right side (lat<-32.6) the T_{wb} remains < 0 (see ref line) and is classified as ice+supercooled.

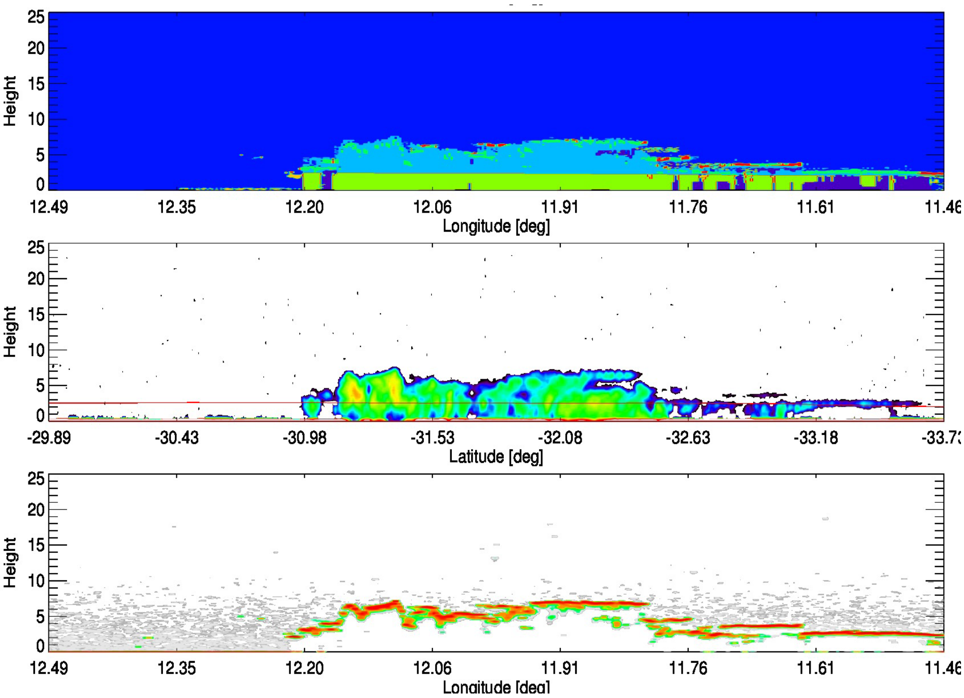

Figure 5: ACM- Target Classification example using CloudSAT and Calipso data measured on 2010-04-19. The top figure shows the result of the ACM-TC, the second is the DARDAR mask, third the CloudSAT radar data (the read line indicates the T_{wb}=0°C isotherm) and the bottom figure the 532nm CALIOP data.

6.1.1. Limitations

As is the case with all classification procedures uncertainties may be difficult to quantify rigorously (i.e. confidence flags or limits may be more qualitative in nature rather than quantitative).

7. Validation status

The L2b lidar classification scheme described in this document exists in prototype form and is not at the time of this writing (Apr 2011) integrated into ECSIM. Since the procedures them selves 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.

Future validation needs:

The current version has been based on the CloudSAT and CALIPSO experience. So far only 1 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 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.

Annex A: Technical implementation

External models