

Recommendation JPL Document No. D-xxxx

CloudSat Project

A NASA Earth System Science Pathfinder Mission

Level 2 Radar-Lidar GEOPROF Product ***VERSION 1.0***

Process Description and **Interface Control Document**

Date: 4 June 2007



Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California

Approvals

_____ Date _____
Gerald Mace
Data Product Algorithm Lead

_____ Date _____
Deborah Vane
Deputy Project Scientist

_____ Date _____
Graeme Stephens
Project Scientist

_____ Date _____
Donald Reinke
Data Processing Center Lead Engineer

_____ Date _____
TBD
“Receiver(s) of this data up the Level 2 Chain”

Questions concerning the document and proposed changes shall be addressed to

Gerald Mace
mace@met.utah.edu
(801) 585-9489

Contents

Contents	3
1. Introduction.....	4
2. Algorithm Theoretical Basis.....	5
3. Algorithm Inputs.....	11
3.1. CloudSat.....	11
3.2. Calipso Vertical Feature Mask.....	11
4. Algorithm Summary	12
4.1. Volume Fraction	12
4.2. Layer Base and Top	12
5. Data Product Output Format.....	13
6. Operator Instructions	16
7. References.....	19
8. Acronym List	20

1. Introduction

The Cloudsat cloud profiling radar (CPR) and the Calipso Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP; hereafter referred as the Lidar) are slated to fly in close coordination with one another when on orbit within the Aqua MODIS swath. The obvious synergy of this combined observational capability is significant. With the ability of the CPR to probe optically thick large-particle layers and the ability of the Lidar to sense optically thin layers and tenuous cloud tops, the two instruments have the potential of providing as complete a picture of the occurrence of cloud and aerosol as has been compiled to date. However, not only do the two instruments sense the atmosphere in different ways, the macroscopic characteristics of the observations such as vertical resolution, spatial resolution, and spatial frequency are quite different from one another. The pointing certainty of the two instruments also adds complexity to combining the two data streams. Our goal is to optimally merge these two data streams in order to produce the most accurate quantitative description of the location of hydrometeor layers in the atmosphere that is possible. Beyond this primary goal, we will also attempt to extract from the combined data streams the degree to which the volumes illuminated by the cloudsat CPR are fully or partially filled by hydrometeors.

This document describes the algorithm that will be implemented operationally to combine the CPR and Lidar data cloudmasks to meet the goals described above. The goals of the algorithm are to

- Estimate the degree to which each radar resolution volume is cloud filled.
- Identify the heights of the hydrometeor layer base and top for up to five layers in each vertical CPR profile.

2. Algorithm Theoretical Basis

2.1 Methodology

The CPR radar and Calipso lidar will provide complimentary information regarding the occurrence of hydrometeor layers in the vertical column. The radar will 1) penetrate optically thick layers that will attenuate the lidar signal; 2) observe layers of cloud-free precipitation that may not be observed by the lidar. The lidar will sense 1) tenuous hydrometeor layers that are below the detection threshold of the radar; 2) the tops of optically thin ice cloud layers that the radar will not observe; 3) layers at higher spatial (both vertically and horizontally) resolution than the radar (Table 1).

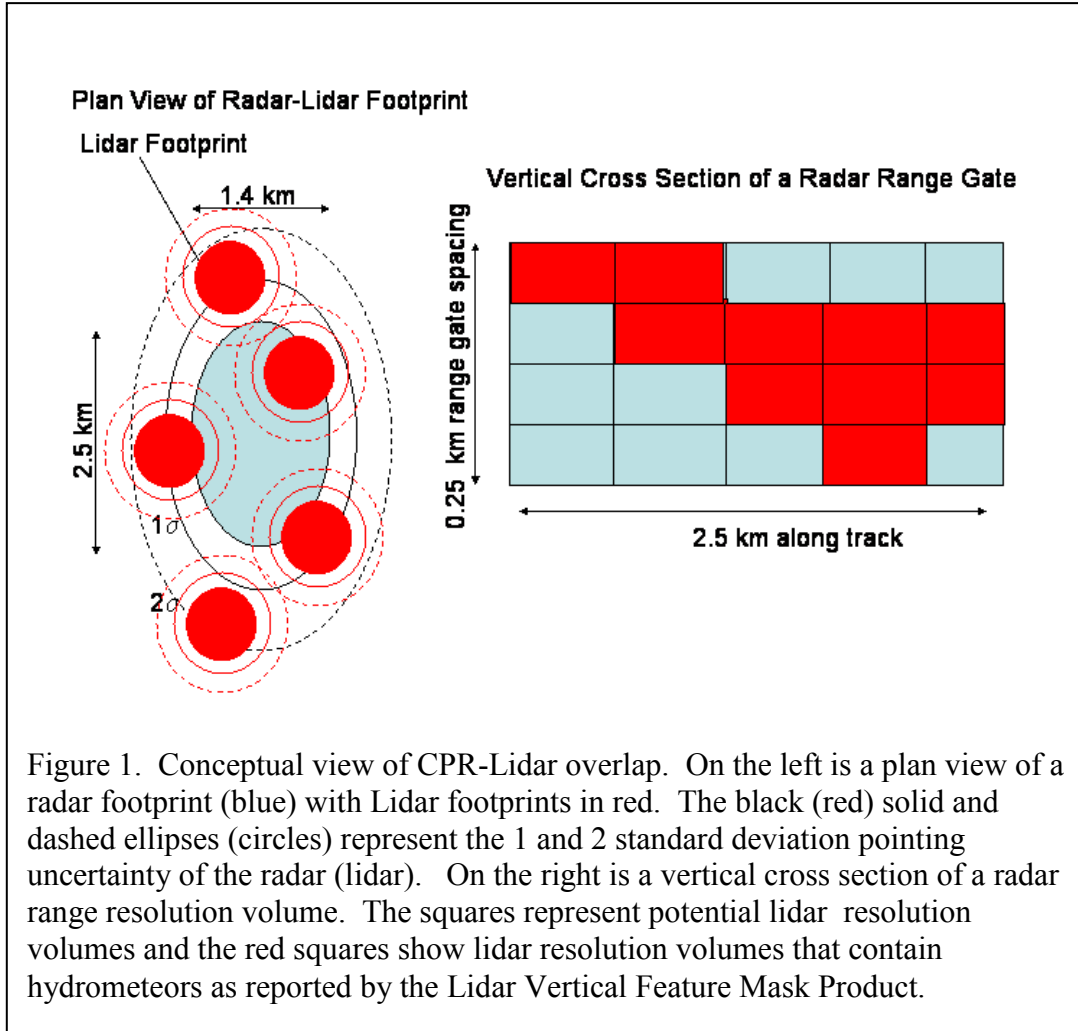
	Cross Track	Along Track	Vertical
CPR	1.4 km	2.5 km	0.25 km
Lidar	0.3 km	1 km- < 8.2 km 0.3 km > 8.2 km	0.03 km < 8.2 km 0.075 km > 8.2 km

The GEOPROF-LIDAR algorithm is designed to extract maximum information from the combined radar and lidar sensors. Our goal is to exploit the synergy between these instruments and produce a product that provides a best description of the occurrence of hydrometeor layers in the vertical column, as well as estimates the fraction of the radar range resolution volume that is cloud filled.

Since the two instruments have different spatial domain such as vertical resolution, spatial resolution and spatial frequency, the spatial domain of the output products in this algorithm is defined in terms of the spatial grid of the CPR.

Figure 1a shows a plan view of a CPR footprint in blue with some number of coincident lidar footprints in red. The solid and dashed circles surrounding the lidar footprints and the solid and dashed ellipses surrounding the radar footprint represent the 1 and 2 standard deviation pointing uncertainty in the CPR and Lidar. Figure 1b shows a vertical slice through a radar range bin. The conceptual lidar observations are shown within this range bin. Hydrometeor observations as reported by the Calipso Level 2 Vertical Feature Mask are shown in red. This rendition of the data corresponds most closely to the situation that will be encountered below 8.2 km. Above 8.2 km, the resolution of the lidar will be 1km along track and 333 meters across track and 75 meters in the vertical.

The plan view in figure 1 illustrates the approximate pointing uncertainty that must be considered when combining the two data streams. In combining the data, we envision the



region encompassing a radar footprint as representing some spatial field of probability that a given element of area contributes to the reflectivity profile reported by the CPR. This probability field will form an ellipse that has the highest values in the region immediately surrounding the reported geographical location of the profile. This probability will decrease outward from the point in some to-be-determined functional form. Similarly, the lidar observations that could potentially overlap the spatial region enclosing the radar observational domain out to the CPR 2-sigma boundary will be considered as potentially contributing to the spatial description of the overlap region. Below 8.2 km, as many as 9-10 separate lidar profiles will be included while above 8.2 km, 3-4 profiles will potentially contribute to the hydrometeor description. Like the CPR, we will consider the likelihood that a given element of area in the region

surrounding a reported lidar geographical point as forming some spatial probability distribution. The degree of contribution of a lidar observation to a given radar resolution volume will, therefore, be calculated in terms of the degree to which that particular observation potentially contributed to the spatial overlap in the radar observational domain. This will be calculated using a weighting scheme based on the spatial probability of overlap: $w_i = \int \int_{x,y} P_r P_l^i dy dx$ where i counts the lidar profile in a particular

radar observational domain, x and y represent spatial dimensions to form an area enclosing the radar domain, subscripts r and l represent the radar and lidar, respectively and P is the spatial probability that a particular element of area defined by x and y contributes to the observation (see figure 1). At a given level then, a reasonable lidar cloud fraction (C_l) within a radar footprint could be expressed as a weighted combination of the lidar observations within the radar probability field:

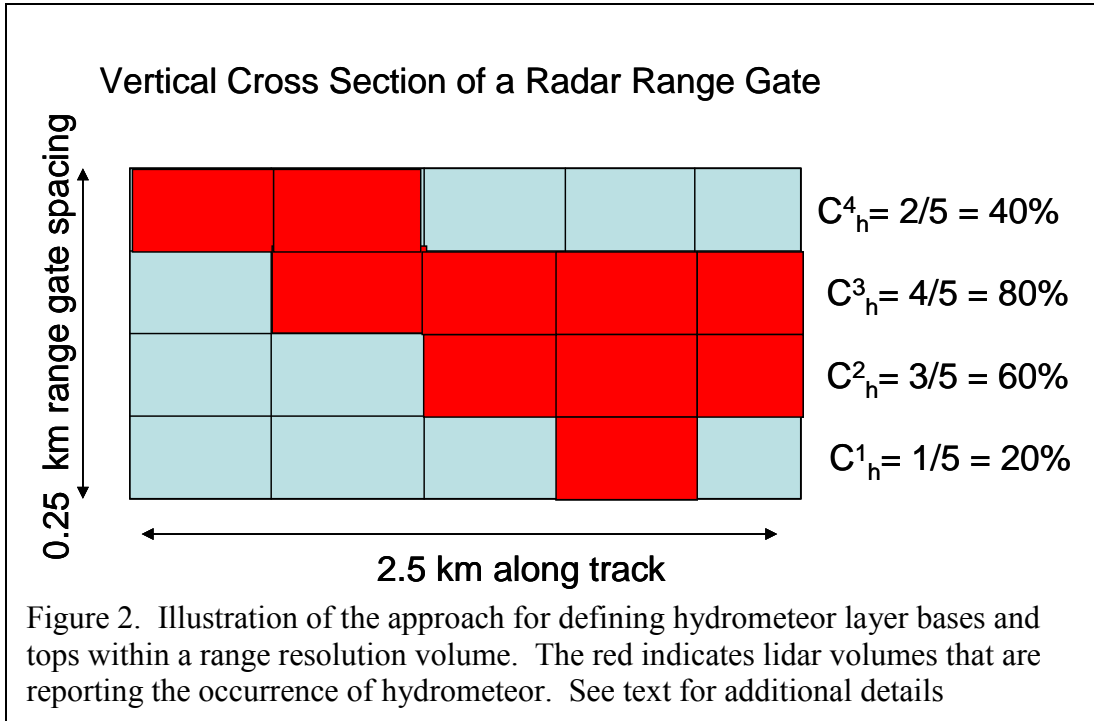
$$C_l = \frac{\sum_{i=1}^{\# \text{ of lidar obs}} w_i \delta_i}{\sum_{i=1}^{\# \text{ of lidar obs}} w_i} \quad (\text{equation 1})$$

where δ_i is the lidar hydrometeor occurrence where a

value of 1 indicates that hydrometeor exists in a given profile at a given height while a value of 0 is the non occurrence of hydrometeor. In calculating C_l the number, i , of lidar observations within a particular radar resolution volume include the total number that could potentially contribute to the overlap from all profiles and all lidar resolution volumes from those profiles. This quantity, C_l , effectively quantifies the partial filling of the radar volume by hydrometeors and will be one of the output quantities of the radar-lidar combined product. Because we can write to a reasonable approximation,

$$Z_e^{cor} = \frac{V}{V_{cor}} Z_e,$$

the quantity C_l can be used to directly correct the reported radar reflectivity for partial beam filling. The uncertainty in C_l will be determined by perturbing the observations about their reported locations.



We will also use the combined cloud masks to derive a best estimate description of the hydrometeor layers in the vertical column along the spatial dimension defined by the CPR. The output of this layer product will be the base and top heights of up to five distinct hydrometeor layers and some indication as to whether those layers were observed by the radar, the lidar or both the radar and lidar. A layer boundary is defined as the first encounter of a cloudy range level (either radar or lidar) following the occurrence of a cloud-free range level (either radar or lidar). Figure 2 illustrates the general approach. Within a radar range resolution, we will use the technique described earlier for equation 1 except that calculation will be performed for a given lidar range for the profiles overlapping a footprint. These hydrometeor fractions are denoted in Figure 2 by C_h^n where n denotes a lidar range number. Within the horizontal domain of a radar footprint, we will define a lidar range as cloudy if the value $C_h^n \geq 0.5$. In figure 2, the first cloudy layer that would be encountered moving from bottom to top would be the second layer. This would be then be reported as the layer base and the third layer would be reported as the layer top. The following conventions are defined:

- Due to its finer vertical resolution, the lidar will always be deferred to in reporting a layer boundary.
- If a layer top were to be identified by the lidar and not the radar and the lidar were to attenuate before a distinct layer base were identified, the layer base would then be defined by the radar observations. In this case, the layer top indicator would show that only the lidar observed the boundary while the layer base would show that only the radar observed the base.
- If a layer boundary were to be observed by the lidar and the radar range gate containing the lidar-defined boundary indicated the presence of significant echo within the range bin, then the boundary height would be defined by the lidar

observation and the contribution flag would show that both instruments observed the boundary.

2.2 Known Data Quality Issues

The Cloudsat and Calipso data sets are novel in many respects. Therefore since we are only approximately 1 year into the missions as of this writing, the user of the data must keep in mind the experimental aspect of the data sets and the products derived from them. Both the Cloudsat and Calipso teams continue to evaluate the data products from which the Radar-Lidar Geometrical Profile Product is derived. Full reprocessing of the data sets is planned by both teams. We summarize here several known issues that users of the Radar-Lidar Geometrical Profile Product should be aware of.

The following is quoted directly from the Calipso Data Quality web site at, http://eosweb.larc.nasa.gov/PRODOCS/calipso/Quality_Summaries/#seeLidarLevel2CloudProfile regarding the Vertical Feature Mask that we use as input to the Radar-Lidar Geometrical Profile Product:

“Overall, the algorithm performance is fairly good at labeling cloud as cloud and somewhat less successful in labeling aerosol as aerosol. Several types of misclassifications are fairly common and should be watched for. The most common misclassification is portions of dense aerosol layers being labeled as cloud. The algorithm operates on individual profiles, so small regions within an aerosol layer are sometimes labeled as cloud. These misclassifications are often apparent from study of Level 1 browse images. Actual clouds occurring within aerosol layers appear to be correctly classified as cloud most of the time. Additionally, portions of the bases of some cirrus clouds are mislabeled as aerosol, and some tropospheric polar clouds are erroneously labeled as aerosol. Improvements to the cloud/aerosol discrimination algorithm are underway and misclassifications should be greatly reduced in future data releases. “

Inspection of the Cloud-Aerosol mask has shown that layers are correctly identified as cloud or aerosol about 90% of the time. Several types of mis-classification are fairly common and should be watched for. The most common mis-classification is when portions of dense aerosol layers are labeled as cloud. The algorithm can identify clouds embedded within aerosol layers but because the algorithm treats each profile independently, small regions within an aerosol layer can be mis-identified as cloud. Clouds located at the top of aerosol layers may cause the aerosol under the cloud to also be classified as cloud. Actual clouds occurring within aerosol layers appear to be correctly classified as cloud most of the time. Additionally, portions of the bases of some cirrus clouds are mislabeled as aerosol. In the polar regions, there appear to be some systematic mis-classifications. We have noticed that in Antarctic winter clouds near the surface are classified as aerosol. In the Arctic, boundary layer aerosols are consistently classified as cloud in certain conditions. Many of these mis-classifications are apparent

from study of Level 1 browse images. Improvements in the cloud-aerosol discrimination algorithm are underway and mis-classification should be greatly reduced in future data releases.

Several issues must be kept in mind regarding the Cloudsat Geometrical Profile Product and the Cloudsat data in general. These are discussed in Mace et al. (2007) and Marchand et al. (2007). Additional information can be found at <http://www.cloudsat.cira.colostate.edu/dataHome.php>

- The minimum detectable signal of the Cloud Profiling Radar is approximately -31 dBZ_e. Therefore, some fraction of high thin cirrus, and non precipitating water cloud such as altocumulus and continental stratus will be below the detection threshold of the CPR
- Due to reflection from the surface and the 1 km pulse length of the CPR, sensitivities in the lowest 1 km near the surface are reduced. The lowest approximately 500 m of each profile will contain no identifiable hydrometeor signal. Between approximately 500 m and 1 km, hydrometeors can be detected at a reduced sensitivity. Research indicates that the sensitivity between 500 m and 1 km will be largely recovered in reprocessing of the data.
- We make no effort in this product to separate cloud from precipitation.

Finally, the user should be aware that we have found substantial difference between the day and night cloud occurrence statistics. While in some circumstances, diurnal variability is known to occur such as in the subtropical boundary layer stratocumulus regimes, the differences we are finding in the upper troposphere suggest that differences in detection thresholds are resulting in more thin clouds being found at night when the lidar data are much less noisy. Day versus night comparisons will not provide physically reasonable comparisons until we use extinction coefficient as a filter. This will be added in future versions of the algorithm.

3. Algorithm Inputs

3.1. CloudSat Level 2B GEOPROF Science Data

2B-GEOPROF-LIDAR algorithm requires the following inputs from the CPR Level 2B GEOPROF product:

- Spacecraft latitude
- Spacecraft longitude
- Height of each radar bin
- Radar significant echo mask

3.2 CALIPSO Science Data

2B-GEOPROF-LIDAR algorithm requires the following inputs from CALIPSO Level 2 Vertical Feature Mask product:

- Latitude of lidar profile
- Longitude of lidar profile
- Lidar vertical feature mask

4. Algorithm Summary

The GEOPROF-LIDAR algorithm is described in Section 2.

5. Data Product Output Format

5.1. Format Overview

Using the radar and lidar cloud masks, the CPR Level 2 GEOPROF-LIDAR Product produces the cloud volume fraction, uncertainty of cloud volume fraction, number of hydrometeor layers, height of cloud layer base, height of cloud layer top, as well as the contribution flags of the base and top.

The format chosen for the GEOPROF-LIDAR data consists of metadata, which describes the data characteristics, and swath data, which includes the cloud fraction (C_1), C_1 uncertainty, number of hydrometeor layers, height of layer base, height of layer top, contribution flag for each layer base and top. The following schematic illustrates how GEOPROF-LIDAR data is formatted using HDF EOS. The variable nray is the number of radar profiles (frames, rays) in a granule.

Table 2. CPR Level 2 GEOPROF-LIDAR HDF-EOS Data Structure

Data Granule	CloudSat Metadata		TBD
	CPR Metadata		TBD
Swath Data	Latitude		nray, 4-byte float
	Longitude		nray, 4-byte float
	CloudFraction (C_1)		$125 \times$ nray, 1-byte integer
	UncertaintyCF		$125 \times$ nray, 1-byte integer
	CloudLayers		nray, 1 byte integer
	LayerBase		$5 \times$ nray, 4-byte float
	LayerTop		$5 \times$ nray, 4-byte float
	FlagBase		$5 \times$ nray, 1-byte integer
	FlagTop		$5 \times$ nray, 1-byte integer

5.2. CPR Level 2 GEOPROF-LIDAR HDF-EOS Data Contents

The contents of the metadata are still TBD and so the sizes are also TBD.

- **CloudSat Metadata** (Attribute, Size TBD)

TBD by CIRA

- **CPR Metadata** (Attribute, Size TBD)

TBD by CIRA

- **Profile Time** (Vdata data, array size: nray, record size: 10 byte):

Seconds since the start of the granule for each profile. The first profile Time is 0.

□ **Geolocation**

- **Latitude** (SDS, array size: nray, record size: 4-byte float)

Spacecraft geodetic latitude. Latitude is positive north, negative south.

- **Longitude** (SDS, array size: nray, record size: 4-byte float)

Spacecraft geodetic longitude. Longitude is positive east, negative west. A point on the 180th meridian is assigned to the western hemisphere.

- **Height** (SDS, array size: nray, record size: 2-byte integer)

Height of the radar range bins in meters above mean sea level.

- **Range to intercept** (SDS, array size: nray, record size: 4-byte float)

Range from the spacecraft to the CPR boresight intercept with the geoid.

- **DEM elevation** (SDS, array size: $125 \times \text{nray}$, record size: 2-byte integer)

Elevation is in meters above Mean Sea Level. A value of -9999 indicates ocean. A value of 9999 indicates an error in calculation of the elevation.

□ **Swath data**

- **CloudFraction** (SDS, array size: $125 \times \text{nray}$, record size: 1-byte integer)

The CloudFraction (C_1) reports the fraction of lidar volumes in a radar resolution volume that contains hydrometeors. It is recorded per ray and per bin as a 1-byte integer variable. It is a percentage from 0 to 100.

- **UncertaintyCF** (SDS, array size: $125 \times \text{nray}$, record size: 1-byte integer)

The UncertaintyCF is the uncertainty of cloud volume fraction. It is recorded per ray and per bin as 1-byte integer. Its value ranges from 0 to 100.

- **CloudLayers** (SDS, array size: nray, record size: 1-byte integer)

CloudLayers is a description of the number of the observed hydrometeor layers in the vertical column of the radar footprint. It is recorded per ray as 1-byte integer. Its value is from 0 to 5. A maximum of 5 layers are recorded.

- **LayerBase** (SDS, array size: $5 \times n_{ray}$, record size: 4-byte float)

LayerBase is a description of the height of the observed hydrometeor layer base. It is recorded per ray as 4-byte float. Its value is from 0 to 25000. The units are meters.

- **LayerTop** (SDS, array size: $5 \times n_{ray}$, record size: 4-byte float)

LayerTop is a description of the height of the observed hydrometeor layer top. It is recorded per ray as 4-byte float. Its value is from 0 to 25000. The units are meters.

- **FlagBase** (SDS, array size: $5 \times n_{ray}$, record size: 1-byte integer)

FlagBase is the contribution flag for each layer base. It tells which instrument has been used to identify the base height. It is recorded per ray as 1-byte integer. The value is from 0 to 3.

- 0 means that neither radar nor lidar finds a layer base.
- 1 indicates that only the radar has found the base.
- 2 indicates that only the lidar has found the base.
- 3 indicates that both radar and lidar have found the base.
- 9 corresponds to missing data.

- **FlagTop** (SDS, array size: $5 \times n_{ray}$, record size: 1-byte integer)

FlagTop is the contribution flag for each layer top. It tells which instrument is finding the top height. It is recorded per ray as 1-byte integer. The value is from 0 to 3.

- 0 means that neither radar nor lidar find a top.
- 1 indicates that only the radar has found the top.
- 2 indicates that only the lidar has found the top.
- 3 indicates that both radar and lidar have found the top.
- 9 corresponds to missing data.

6. Operator Instructions

The Level 2 GEOPROF-LIDAR product processing software is integrated into the CloudSat Operational and Research Environment (CORE). It is called using the standard CORE procedure for calling modules to operate on data files. The output is in the form of an HDF-EOS structure in memory, which can be saved by CORE and passed on to other Level 2 processing.

For quality assessment purposes, images of the CPR radar mask and Calipso lidar cloud mask are created (Figure 3). These two plots show the original input of the radar mask and lidar mask. We also create images of the combined radar-lidar cloud layer information and volume fraction (Figure 4). The GEOPROF-LIDAR algorithm is not working well if figure 4 shows unreasonable cloud layer or volume fraction. And by comparing figure 3 and figure 4, we are able to verify the GEOPROF-LIDAR algorithm. If figure 3 shows cloud mask at some place but figure 4 doesn't show cloud layer, the algorithm is not working. Vice versa, if figure 4 shows cloud layer at some place but the cloud masks in figure 4 don't show cloud, the algorithm is not working.

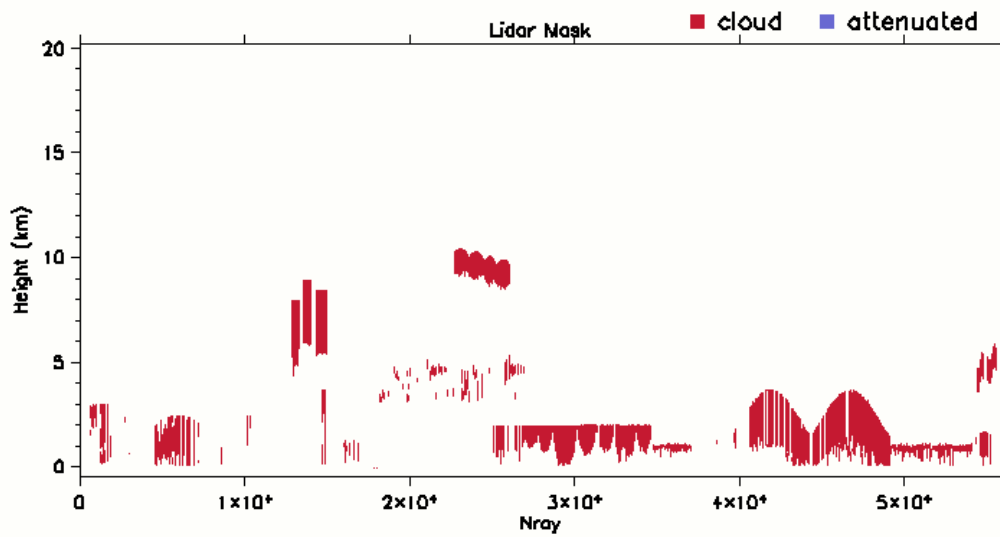
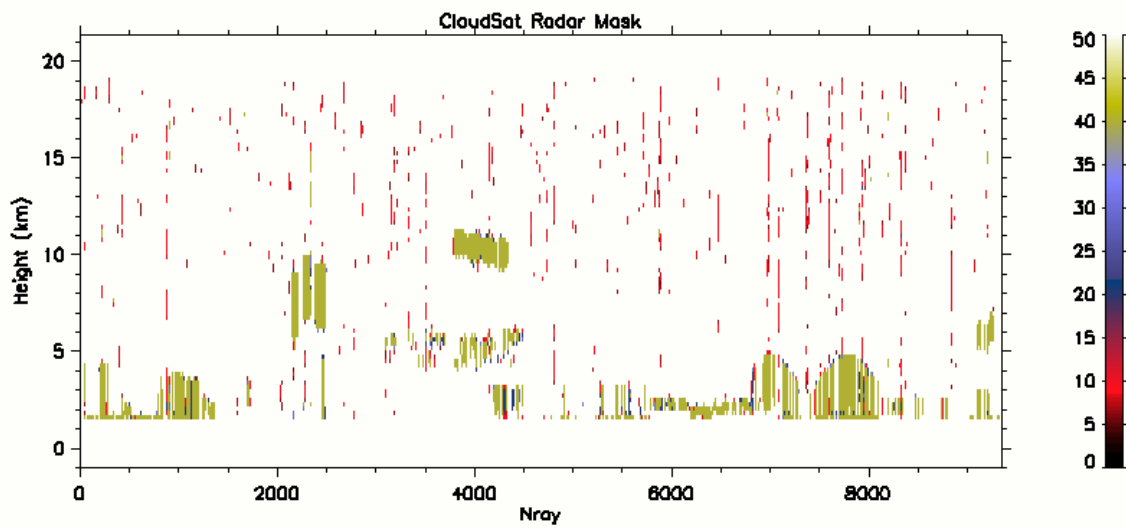


Figure 3. (a) CPR radar mask, (b) Calipso lidar cloud mask. These images are created to show the original input of the radar and lidar mask.

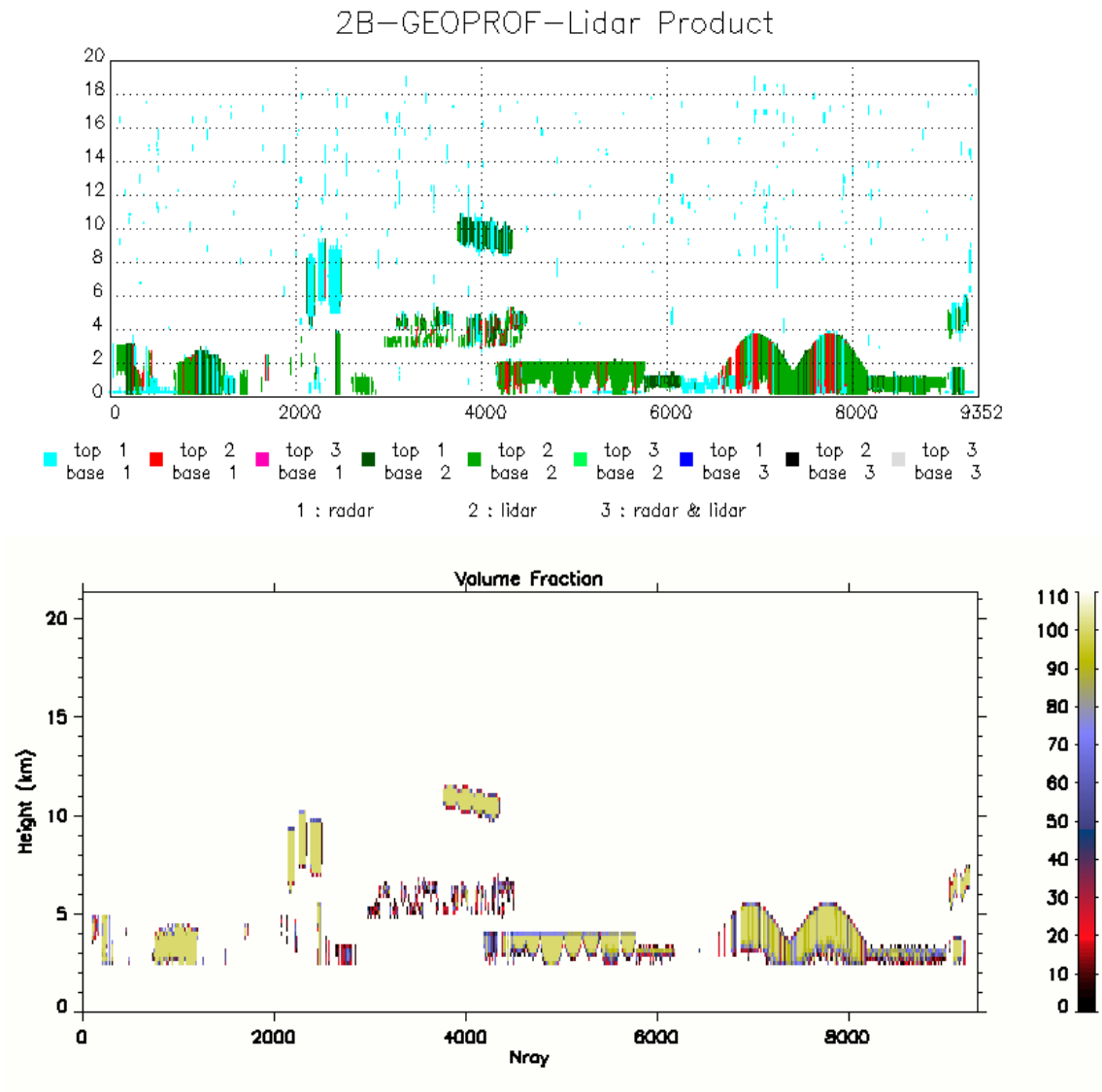


Figure 4. Upper panel shows the cloud layer information. Lower panel is the volume fraction.

- **References**

Mace, G. G., R. Marchand, Q.Zhang, and G. Stephens, 2007: Global hydrometeor occurrence as observed by Cloudsat; Initial observations from Summer 2006, *Geophys. Research Letters*, 34, doi: 10.1029/2006GL029017.

Marchand, R., T. Ackerman, G. G. Mace, and G. Stephens, 2007: Hydrometeor Detection using Cloudsat - an earth orbiting 94 GHz Cloud Radar. Submitted to *Journal of Oceanic and Atmospheric Technology*.

- ***Acronym List***