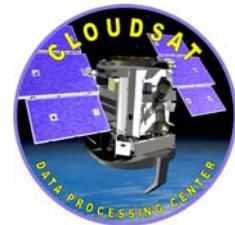# CloudSat Project

A NASA Earth System Science Pathfinder Mission

CloudSat Standard Data Products Handbook

Revised: April 25th , 2008

Cooperative Institute for Research in the Atmosphere
Colorado State University
Fort Collins, CO 80523

CloudSat Standard Data Products Handbook

1. CloudSat Mission Overview

CloudSat is a NASA Earth Sciences Systems Pathfinder (ESSP) mission.   ESSP missions are relatively low cost missions that require that the mission be built, tested and launched in a short time interval – normally within 3 years.   ESSP missions are peer-reviewed science investigations selected from proposals submitted to the NASA Office of Earth Science.   They are lead by a Principal Investigator (PI) who is often affiliated with a university. The PI for CloudSat is Dr. Graeme Stephens who is a Professor in the Colorado State University Atmospheric Science Department.

The purpose of the experimental CloudSat mission is to measure the vertical structure of clouds from space and, for the first time, will simultaneously observe cloud and precipitation. The primary CloudSat instrument is a 94-GHz, nadir-pointing, Cloud Profiling Radar (CPR).

A unique aspect of this mission is the fact that CloudSat will be flying in formation with other Earth Sciences missions.   CloudSat will be a part of a constellation of satellites that currently include NASA's EOS Aqua and Aura satellites as well as a NASA-CNES lidar satellite (CALIPSO), and a CNES satellite carrying a polarimeter (PARASOL).   A unique feature that CloudSat brings to the constellation is the ability to fly a precise orbit enabling the fields of view of the CloudSat radar to be overlapped with the lidar footprint and the other measurements of the constellation. The precision of this overlap creates a unique multi-satellite observing system for studying the atmospheric processes of the hydrological cycle. Additional information about the CloudSat mission may be found at http://cloudsat.atmos.colostate.edu

CloudSat Science and Engineering Data will be down-linked, from the CloudSat on-board data recorder, through the Air Force Satellite Communications Network (AFSCN) to the USAF RDT&E Support Center (RSC) in Albuquerque NM (fig. 1).   There the data will be decommutated, checked for transmission errors (checksum), and stored on-line as a set of binary files. These raw data will be accessed by the CloudSat Data Processing Center, via internet/ftp, for subsequent level 0 through level 2 processing.
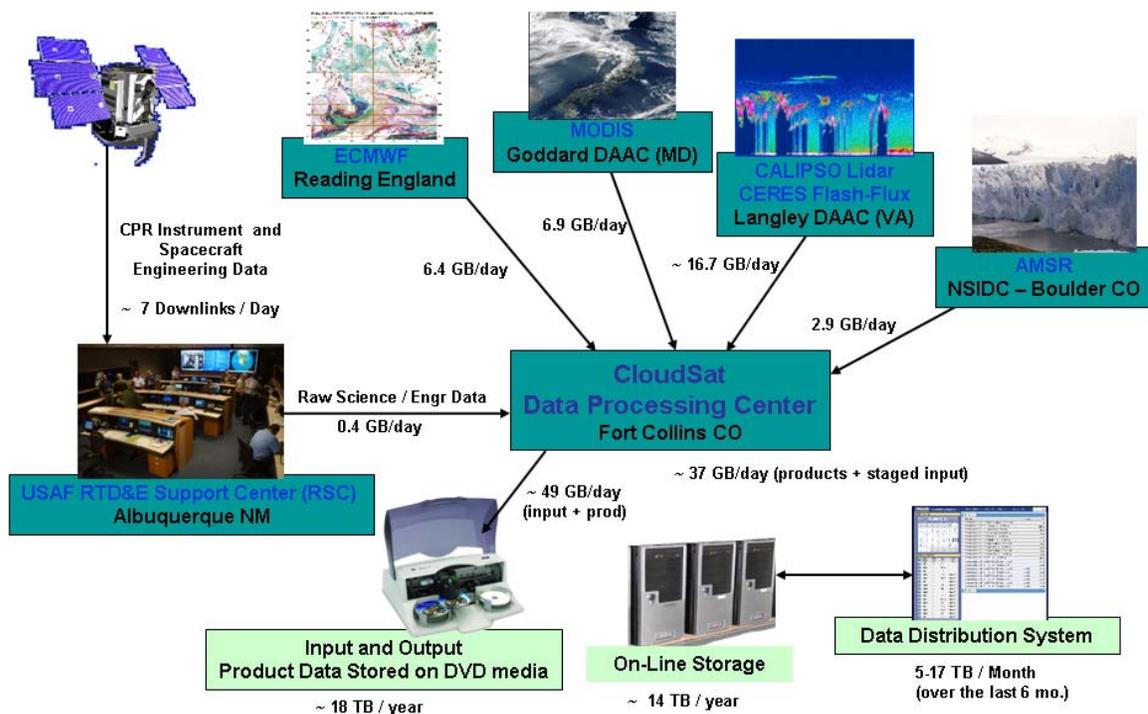

Figure 1.   CloudSat System Data Flow

All of the science data processing support for the CloudSat mission will be provided by the Cooperative Institute for Research in the Atmosphere (CIRA). All of the CloudSat standard data products will be produced at the CloudSat Data Processing Center, located in the ATS-CIRA Research Center (adjacent to CIRA and the Atmospheric Science Department).

The DPC will produce eleven CloudSat Standard Data Products, as shown in Table 1.

Product	Description	Algorithm Developers
1B-CPR and 1B-CPR-FL	Radar Backscatter Profiles	Steve Durden (JPL)
2B-GEOPROF	Cloud Geometrical Profile	Jay Mace, Qiuqing Zhang (U of Utah)
2B-CLDCLASS	Cloud Classification	Ken Sassen (U of Alaska), Zhien Wang (U. Wyo)
2B-CWC-RO	Combined Water Content - Radar Only	Graeme Stephens, Richard Austin (CSU)
2B-TAU-OFF-N	Cloud Optical Depth - Off Nadir	Graeme Stephens, Igor Polonski (CSU)
2B-CWC-RVOD	Combined Water Content - Radar + Vis. Optical Depth	Graeme Stephens, Richard Austin (CSU)
2B-FLXHR	Fluxes and Heating Rates	Graeme Stephens, Tristan L'Ecuyer (CSU)
2B-GEOPROF-LIDAR	Cloud Geometrical Profile from CPR and CALIPSO Lidar	Jay Mace, Qiuqing Zhang (U of Utah)
2B-CLDCLASS-LIDAR	Cloud Classification from CPR and CALIPSO Lidar	Ken Sassen (U of Alaska), Zhien Wang (U. Wyo)

Table 1. CloudSat Standard Data Products

Members of the CloudSat Science Team have developed the Science algorithms and software for the CloudSat Standard Data Products. Four universities and the NASA Jet Propulsion Lab (JPL) are participants on the CloudSat algorithm development team.

During the Operational (on-orbit) Phase, the DPC will be staffed by CIRA employees, Science and Technology Corporation personnel (under a sub-contract to CIRA), and part-time CSU students. More information about the DPC can be found at http://cloudsat.cira.colostate.edu

2. CPR Instrument and Data Footprint

The Cloud Profiling Radar (CPR) is a 94-GHz nadir-looking radar which measures the power backscattered by clouds as a function of distance from the radar. The overall design of the CPR is simple, well understood, and has strong heritage from many cloud radars already in operation in ground-based and airborne applications. Most of the design parameters and subsystem configurations are nearly identical to those for the Airborne Cloud Radar, which has been flying on the NASA DC-8 aircraft since 1998.

a. CPR Instrument Sampling Characteristics

· The CloudSat radar samples at **625 kHz**. At this frequency the **burst rate = 0.16 sec / burst** (burst rate = 1/6.25). This sample interval defines a CloudSat "profile" (also referred to as a "ray")
· The CloudSat **Pulse Repetition Frequency (PRF) = 4300**. From this we can compute: (4300 pulses / sec) * (0.16 sec/burst) = **688 pulses/burst**. These pulses are averaged to produce a single reflectivity value for each profile. (Note: Approx. 21 pulses/burst of overhead are used for calibration and "dead time")
· The CloudSat antenna pattern will provide an **instantaneous footprint** of approximately **1.3 km** (at Mean Sea Level).
· The CPR instrument will be flown in a sun-synchronous orbit at an 89° inclination angle, and a nominal altitude of 705km, producing an **along-track velocity of approximately 7 km/sec**.
· Using this velocity, and the sample rate of 0.16 sec/profile, we can approximate that **a CPR profile will be generated every 1.1 km along track**.
· Each profile will have 125 vertical "bins". Each bin will be approximately 240m thick.
· As shown in Figure 1, **the footprint for a single profile will be approximately 1.3km (across-track) by 1.7 km (along-track)**

- A CloudSat data **"Granule" is defined as one orbit**. A granule starts at the first profile that falls on or past the equator on the descending node.   (note: Approximately 20 seconds, or 125 profiles, will be appended to the beginning and end of each granule).
- The **time stamp that is assigned to the granule will be the time that corresponds to the center profile.** This time is halfway between the start of the first instantaneous sample in the profile and the start of the first instantaneous sample of the next profile. (Note: the overlap was added to facilitate the generation of several CloudSat standard data products.)
- The **CloudSat data products are time stamped with the time at the middle of the 0.16 second-interval profile**.

b. **Data Footprint and Granule Size**

Granules, Profiles and Bins : CPR footprint & granule size

1 GRANULE = 1 orbit of data (~ 40,786 km / ~ 37,088 profiles)

"Data Window"

30-km

Top of Data Window

- Granule begins on descending node (night side)
- "Data Window" is 30-km high by 37,088 profiles wide

Each "Profile" has 125 vertical "BINS" (~30 km)

Each vertical bin is 240 m thick

Surface

1.1 km along-track

1.7 km

1.1 km

Profile Timestamp Location

1.3 km

START OF PROFILE

END OF PROFILE

SATELLITE DIRECTION OF MOVEMENT

~98.9 minutes per orbit ~14.56 orbits/day
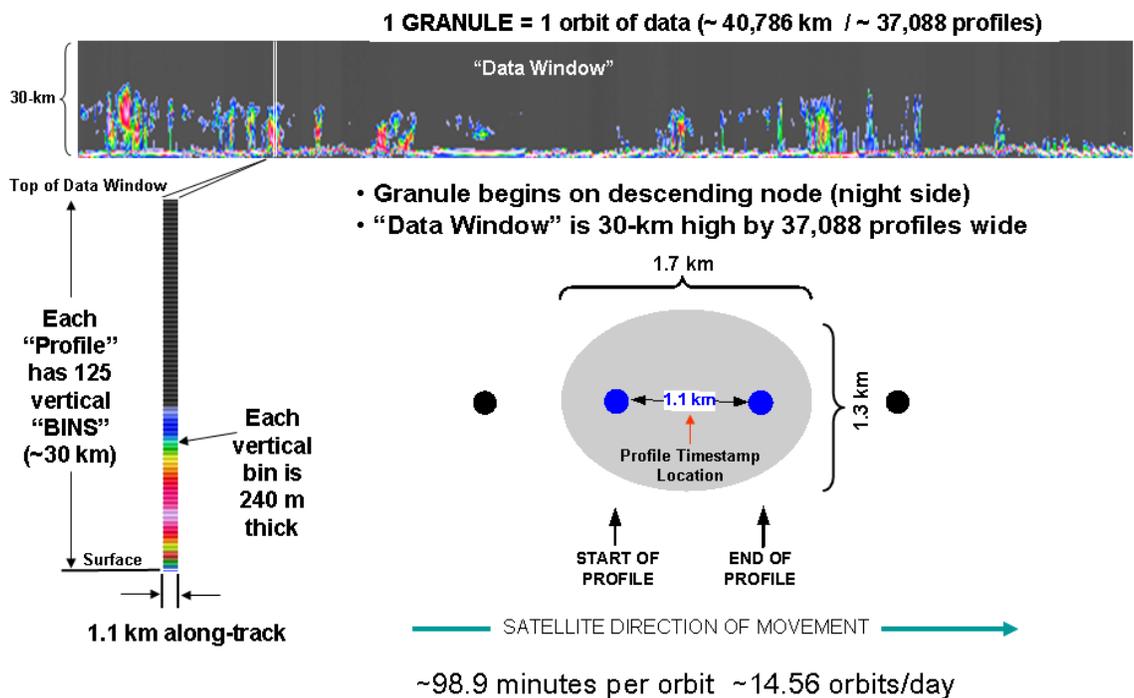
Figure 2. CloudSat Data Footprint and Granule Size

The CloudSat data footprint is approximately 1.7km along-track by 1.3km across-track.   A granule is one orbit of data beginning at the first profile on or after the equator on the descending node.   There are **125 vertical bins**, each one approximately **240m thick**.   There are approximately **36,383 profiles per granule**.

3. CloudSat Data Products Overview

NOTE: Current Data Format Specifications can be viewed via the CloudSat Data Processing Center website: http://www.cloudsat.cira.colostate.edu/dataSpecs.php . This is the "product documentation" link and is the most current information about data products and product ICD's.

CloudSat data products are generated for each CloudSat profile location. There are two types of CloudSat data products, Standard Data Products, and Auxiliary Data Products (table 1).

Product	Description	Algorithm Developers
1A-AUX	Level 1 Auxiliary Data (Engr Data, Geoloc., Time Tag)	Dale Reinke (CIRA)
ECMWF-AUX	ECMWF Auxiliary Data (State Variables)	Phil Partain (STC-Metsat / CIRA)
MODIS-AUX	MODIS Auxiliary Data (data from 22 MODIS channels)	Phil Partain (STC-Metsat / CIRA)
LIDAR-AUX	CALIPSO Lidar Auxiliary Data (Lidar Profiles)	Phil Partain (STC-Metsat / CIRA)

Table 1. Auxiliary Data Products

Auxiliary data products are produced from Ancillary data that are mapped in the horizontal to the center of the CloudSat profile footprint and, where appropriate, in the vertical to the center of each CloudSat profile bin.

Detailed specifications for each of these products can be found in their respective Interface Control Documents via the CloudSat DPC website link: http://cloudsat.cira.colostate.edu/dataICDlist.php . Users can also find input and output field specifications by going to the Data Products section of the website where data formats and sizing information can be found for each standard and auxiliary data products.

a. Hierarchical Data Processing Overview

All of the CloudSat Level 1 and Level 2 Standard Data Products (SDPs) are produced in a hierarchical scheme. Output parameters from each product, starting with the 2B-GEOPROF, are required input for successive products in the processing stream.

In addition, CloudSat will use auxiliary data, when available, for the generation of level 2 SDP's. These data include ECMWF state variables and MODIS Cloud Mask and Radiance data.

The descriptions of CloudSat Auxiliary and Standard Data Products will following in the order in which they are required for processing.

b. 1A-AUX and 1A-AUX-FL (Level 1 Auxiliary Data) (view 1A-AUX ICD)
(note: if linked ICD page does not appear refresh your browser)

CloudSat Science and Engineering Data will be down-linked, in separate data streams, from the USAF RDT&E Support Center (RSC) in Albuquerque NM. As they arrive, raw CloudSat science and auxiliary data will be pre-processed to form a time-ordered data file (with duplicate data packets removed and missing data blank-filled/flagged). These data are then stored on the local DPC mass storage system by orbit (granules).

When full granules of both CPR Level 0 data and engineering data are available, an intermediate product called the 1A-AUX-FL (First Look) will be created. The 1A-AUX-FL file will contain engineering data, time, geolocation, and elevation for each CPR level 0A data record.

Once per day the DPC will receive a Definitive Ephemeris from the RSC, which contains the time, spacecraft location (geodetic lat/lon, elevation), orbit #, and velocity/orientation vectors. This product is used to create the 1A-AUX product, which then replaces the 1A-AUX-FL.

The distinction between the 1A-AUX-FL and the 1A-AUX file is that the First Look geolocation is generated from the GPS data that are taken from the engineering data, while the geolocation for the 1A-AUX product is taken from the Definitive Ephemeris which is available at approximately 16:00 UTC (valid for 00:00 – 23:59 UTC the previous day).

These 1A-AUX and 1A-AUX-FL files are subsequently combined with the 0A-CPR data to build the 1B-CPR (or 1B-CPR-First-Look) product. NOTE: The GPS data, though expected to be quite accurate, are not corrected for errors or dropouts, thus the geolocation of individual profiles in the 1B-CPR-FL product will not be as accurate as the geo-location of the 1B-CPR which is produced with the Definitive Ephemeris.   It should be noted, however the CPR data itself will be identical to the subsequent 1B-CPR product.

These intermediate products will not be made available to the general science community, as all of the information will be contained in the 1B-CPR product. They will, however, be archived for use in data reprocessing.

c.   **ECMWF-AUX (ECMWF fields mapped to CloudSat Profiles)** (view ECMWF-AUX ICD)
     (note: if linked ICD page does not appear refresh your browser)

Once the 1A-AUX product is available (with each CloudSat profile time-tagged and geolocated), several "Auxiliary" data products are generated.   The first of these is the ECMWF-AUX.   This product is built from ECMWF global data that are provided to the DPC via ftp transfer, four times per day.

The ECMWF-AUX data set is an intermediate product that contains ECMWF state variable data interpolated to each CloudSat vetical bin. These data are required for input to the 2B-GEOPROF, 2B-CLDCLASS, 2B-TAU, and 2B-FLXHR algorithms. The ECMWF-AUX product is created by the "Generic-AUX Interpolate-to-Reference" algorithm. The input data are obtained from a global dataset provided by the European Center for Medium-Range Weather Forecasting (ECMWF).
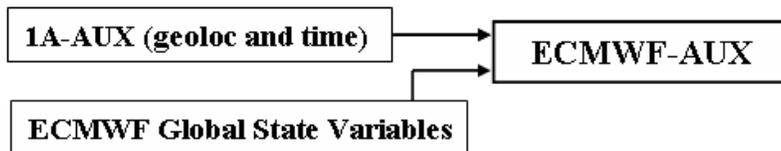
The Generic-AUX interpolate-to-reference algorithm uses two datasets: an independent dataset and a reference dataset. To produce the ECMWF-AUX product, AN-ECMWF is the independent dataset containing the atmospheric state variable data that will be interpolated to each CloudSat radar bin. The geolocation data from the 1B-CPR product comprise the reference dataset. Operating one CloudSat ray at a time, the interpolate-to-reference algorithm uses the reference and independent geolocation data to find the four bounding ECMWF grid points around the CloudSat ray. Then, for each radar bin in the ray, the height of the bin from the 1B-CPR data is used to find the bounding AN-ECMWF vertical levels at each of the grid points. A linear interpolation is performed to calculate the value of each AN-ECMWF data field of interest at the radar bin height for the four grid points. Then a bilinear interpolation is performed on the resulting four values to find the value of each data field at the location of the CPR ray. Because the AN-ECMWF data consists of grids that cover the earth for multiple forecast times, the procedure described above is performed for the two forecast times that bound the time associated with the ray in question. A linear temporal interpolation is then performed on the values obtained for each forecast time, and the final result is the value of each AN-ECMWF data field of interest interpolated to the location and time of the radar bins.

d.   MODIS-AUX (MODIS products mapped to CloudSat Profiles) (view MODIS-AUX ICD)

The MODIS-AUX data set is an intermediate product that contains a subset of ancillary MODIS radiance and cloud mask data that overlaps and surrounds each CloudSat cloud profiling radar (CPR) footprint.   The MODIS-AUX footprint will be a 3-km (along-track) by 5-km (across-track) box of MODIS data centered on each CloudSat profile location.
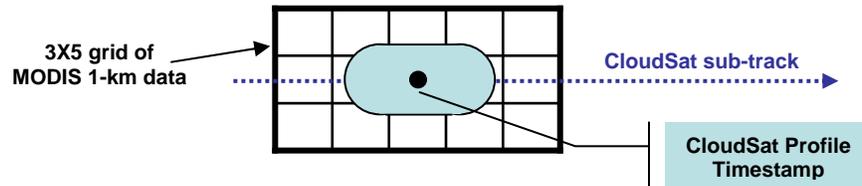

Figure 3.   MODIS-AUX footprint compared with a CloudSat profile

Input data are obtained from the 1B-CPR and AN-MODIS products, and the MODIS-AUX data are used as input to the 2B-GEOPROF, 2B-CLDCLASS, 2B-TAU, and 2B-FLXHR algorithms in the CloudSat data processing system. The MODIS-AUX product is created by the Generic-AUX Subset-to-Reference algorithm. This document describes the interfaces between the input data sets and the Generic-AUX algorithm, the format of the MODIS-AUX product, and quality assessment instructions for the Data Processing Center (DPC) operator.

The Generic-AUX subset-to-reference algorithm uses two datasets: an independent dataset and a reference dataset. To produce the MODIS-AUX product, AN-MODIS is the independent dataset containing the radiances, cloud mask, geolocation, etc. data that will be sub-set around each CloudSat ray. The geolocation data from the 1B-CPR product comprise the reference dataset. Operating one CloudSat ray at a time, the subset-to-reference algorithm uses the reference and independent geolocation data to find the closest AN-MODIS pixel to the CloudSat ray, then stores a 21-pixel across-track by 5-pixel along-track grid of each AN-MODIS parameter of interest in a 105-element vector associated with that ray. If the CloudSat geolocation for a particular ray is missing or the closest valid AN-MODIS pixel is more than 0.71 km from the CloudSat ray, the resulting MODIS geolocation data and the associated data vectors are filled with a missing value flag.

[This data product will be made available, however the requestor must first obtain a release from the CloudSat PI.]

e.   1B-CPR and 1B-CPR-FL (view 1B-CPR ICD view 1B-CPR-FL ICD)
(note: if linked ICD page does not appear refresh your browser)

The CloudSat Cloud Profiling Radar (CPR) measures the backscattered power as a function of distance from the radar. The backscattered power is sampled every 240 m; there are 125 range bins, for a total vertical window of 30 km. The raw, Level 0 data will be converted to calibrated Level 1B data using pre-launch and in-flight calibration measurements. The Level 1 B CPR Process Description and Interface Control Document presents the calibration theory and defines the Level 0 data contents, the Level 1B contents, and the conversion algorithm.

During normal processing, to produce the level 1B-CPR product, the Definitive Ephemeris is used after it arrives, to determine the Geolocation of each CloudSat profile.   The Definitive Ephemeris contains the latitude, longitude and altitude of the satellite at 1-second intervals.   One disadvantage of using the Definitive Ephemeris for Geolocation is that it is only available once per day at approximately 16 UTC. This file contains the orbital parameters for 0000 UTC through

2359 UTC on the previous Julian day. Thus the standard 1B-CPR data product will not be available until 16-40 hours after the data are actually measured.

The distinction between the 1B-CPR-FL and the 1B-CPR file is that the First Look geolocation is generated from the GPS data that are taken from the engineering data, while the geolocation for the 1A-AUX product is taken from the Definitive Ephemeris which is available at approximately 16:00 UTC, valid for 00:00 – 23:59 UTC the previous day.

These 1A-AUX and 1A-AUX-FL files are subsequently combined with the 0A-CPR data to build the 1B-CPR (or 1B-CPR-First-Look) product. NOTE: The GPS data, though expected to be quite accurate, are not corrected for errors or dropouts, thus the geolocation of individual profiles in the –FL product could be in error – however the CPR data itself will be identical to the subsequent 1B-CPR product.

The advantage of the 1B-CPR-FL products is that the DPC should be able to provide it within 10 minutes from the time that data are received at the DPC. The normal time lag, between the actual CPR measurement and the time the 1B-CPR-FL product is generated will, of course, will bel longer and is based on the combined lag times of the data downlink, transfer of data from the downlink site to the RSC, processing at the RSC, transfer of data to the DPC, and processing at the DPC.   The total lag time is as follows:

     a.   Minimum Data Latency = 45 minutes
     b.   Maximum Data latency = 28.5 hours
     c.   Expected Data Latency = 4.5 hours

The CloudSat Level 1A Auxiliary Data Process Description and Interface Control Document contains a description of the data processing that takes place at the DPC after the CPR Science Data, SSOH data, and Definitive Ephemeris data arrive.

f. **2B-GEOPROF (Cloud Geometrical Profile)** (view 2B-GEOPROF ICD)

(note: if linked ICD page does not appear refresh your browser)

The Geometrical Profile algorithm reads a CloudSat 1B-CPR granule and the corresponding MODIS-AUX and ECMWF-AUX granules to produce an estimate of the radar reflectivity factor for those levels, in the vertical column, which contain a significant echo (Stephens et al, 2001). This product can also be called the "Cloud Mask" as it provides the first product to provide a field that indicates the presence of a cloud inside a CPR bin.
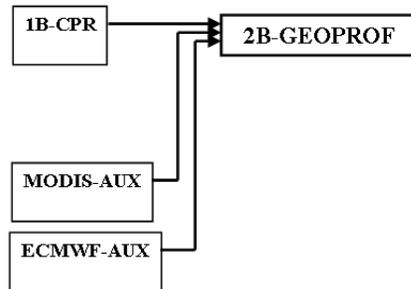
The presence or absence and vertical location of cloud layers impose powerful constraints on the radiative properties of an atmospheric column and, thereby, modulate strongly the radiative heating rate of the column. Comprehensive information on the vertical distribution of cloud layers have been largely missing from analyses of conventional passive-sensor satellite radiometers that observe only the emitted and reflected radiance from the atmosphere and surface. Conventional satellite data have only allowed us to crudely estimate the location and vertical extent of clouds.

Active remote sensors such as CloudSat, on the other hand, are uniquely adapted to observe the vertical location of hydrometeor layers.

The CloudSat orbit will follow closely the orbit of the EOS PM1 satellite (Aqua) on which a number of advanced passive remote sensors will observe the earth. This synergistic association can add significantly to our understanding of the CloudSat geometrical profile and we will use this data source to our advantage. In particular, the Moderate-Resolution Imaging Spectroradiometer (MODIS) will provide information regarding the occurrence and horizontal distribution of clouds within the CloudSat footprint. The MODIS cloud mask product (Ackerman et al., 1998) will be available for use in the operational CloudSat processing stream. The MODIS cloud mask contains not only an estimate of the likelihood of clouds in 1km MODIS pixels, but also contains the results of a number tests with MODIS-observed spectral radiances that will provide limited information on the nature of the clouds in the vertical column. We will use the results of these tests to help us evaluate the characteristics of the CPR hydrometeor returns by performing comparisons of the CPR hydrometeor mask and the MODIS spectral tests.

**g. 2B-CLDCLASS (Cloud Classification)** (view 2B-CLDCLASS ICD)

(note: if linked ICD page does not appear refresh your browser)

The 2B-CLDCLASS algorithm receives input from 2B-GEOPROF and the corresponding MODIS-AUX and ECMWF-AUX granules to identify eight basic cloud types (that are recognized by surface observers internationally).
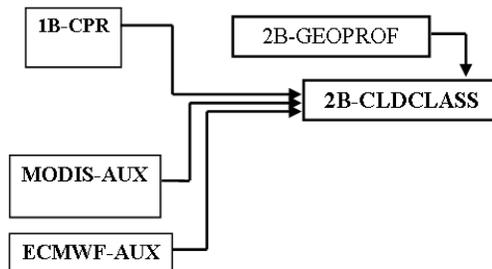
h. 2B-CWC-RO and 2B-CWC-RO (Radar-Only Combined Water Content)

NOTE: These two intermediate products are generated to provide first-guess CWC parameters for the Level 2B-TAU algorithm. The 2B-TAU output is, in turn, fed into the 2B-CWC algorithm for the generation of the CWC-RVOD Standard Data Product.
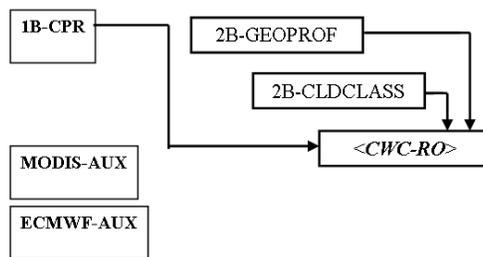
i. 2B-TAU (Cloud Optical Depth – Off Nadir) (view 2B-TAU ICD)

(note: if linked ICD page does not appear refresh your browser)

The 2B-TAU algorithm reads the to produce estimates of the optical depth and its uncertainty for each radar profile determined to contain cloud.
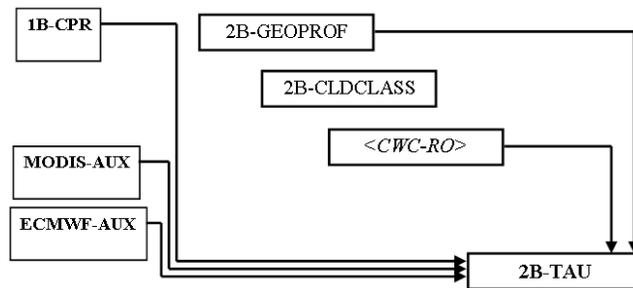
The algorithm retrieves the optical depth and effective radius using radar reflectivity measured by the CloudSat CPR, upwelling top of the atmosphere radiance measured by MODIS in several channels, and other ancillary data (eg. geolocation, time, and surface albedo). The retrieved values are spectral and represent the optical depth and effective 4adius in the short wave. Using CPR information about the distribution of cloud in the atmospheric column, estimates are then produced of the profile of cloud optical depth versus height in the column.

j.   2B-CWC (Combined Water Content) (view 2B-CWC-RVOD ICD)
(note: if linked ICD page does not appear refresh your browser)
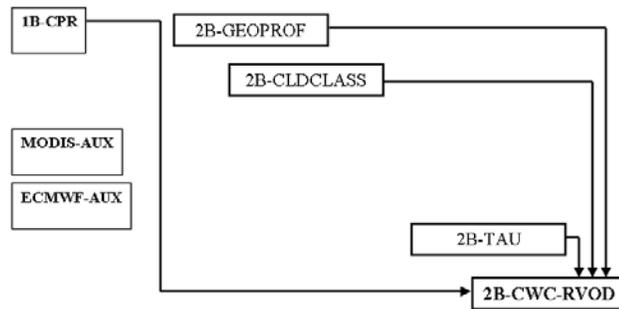
This product is calculated using data from the CloudSat Profiling Radar and other sources, including radiance measurements from the MODIS instrument on the Aqua platform, with which CloudSat will orbit in close formation. In practice, 2B-CWC simply obtains visible optical depth from the 2B-TAU product; details of the workings of 2B-TAU are not relevant to the operation of 2B-LWC as long as 2B-TAU provides an estimate of visible optical depth and its uncertainty

A radar-only version of the retrieval has been implemented for use during the nighttime half of CloudSat's orbit when visible optical depth information will not be available. For continuity, the radar-only retrieval will also be run during the daytime half of the orbit.

The algorithm will use the 2B-CLDCLASS product to handle various cloud types. 2B-CWC results will be output for any cloud not classified as ``ice-only''; this will undoubtedly include many mixed-phase clouds and scenes where there are separate ice and liquid layers. At the present time, no CloudSat standard product explicitly handles these more complex cloud scenarios.

k.   2B-FLXHR (Fluxes and Heating Rates) (view 2B-FLXHR ICD)

(note: if linked ICD page does not appear refresh your browser)

The objective of the FLXHR algorithm is to make use of liquid and ice water content estimates to produce broadband fluxes and heating rates for each radar profile. For a particular provide, upwelling and downwelling longwave and shortwave flux profiles are calculated at discrete levels of the atmosphere. Corresponding heating rates are inferred from these fluxes. In order to perform these calculations, the algorithm makes use of a combination of atmospheric stat variables obtained from ECMWF reanalysis data, profiles f cloud ice and liquid water content obtained from the CloudSat LWC and IWC products, and surface albedos obtained from seasonally-varying maps of surface reflectance properties.
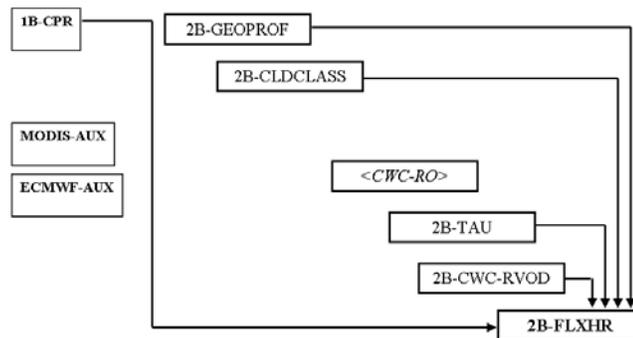
l.   2B-GEOPROF-Lidar (Cloud Geometrical Profile from combined CPR and CALIPSO Lidar) (view 2B-GEOPROF-Lidar ICD)
(note: if linked ICD page does not appear refresh your browser)

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.

m.   2B-GEOPROF-Lidar (Cloud Geometrical Profile from combined CPR and CALIPSO Lidar) (view 2B-CLDCLASS-Lidar ICD)

(note: if linked ICD page does not appear refresh your browser)

A great strength of microwave radar measurements of clouds and precipitation is the ability to retrieve quantitative content data from the radar reflectivity factor Z. This is made possible by devising algorithms based on empirical relationships between Z and various microphysical parameters, such as ice water content IWC or rainfall rate. However, because of the diversity of microphysical conditions found in the atmosphere, algorithms need to be applied only to those conditions for which they are considered valid. In other words, it is first necessary to identify the target and then select an appropriate algorithm. The algorithm selection process depends on such basic factors as cloud phase, and also the hydrometeor density, shape, and size distribution. For example, although cirrus, altostratus, and the upper portions of cumulonimbus clouds are all predominantly ice phase clouds, it is not possible to apply a single algorithm for retrieving IWC in these targets: cirrus generally contain only single ice crystals, altostratus likely contain low-density ice crystal aggregates at the warmer temperatures, and cumulonimbus may combine ice crystals, snowflakes, rimed particles, graupel, and even hailstones. Due to the different radiative forcings of various cloud types (Hartmann et al. 1992; Chen et al. 2000), classifying clouds into categories based on type is also an important task for cloud remote sensing and global cloud climatology studies.

As the first step in converting the vertical profiles of Z from CloudSat into meaningful microphysical data quantities, we are developing an algorithm for identifying cloud type and precipitation. However, measurements of cloud radar alone cannot provide necessary information for cloud scenario classification. The formation fly (A-train) of Aqua, CloudSat and CALIPSO provides other cloud information from lidar and passive radiometer measurements. We identify eight basic cloud types that are recognized by surface observers internationally by combining information available mainly from the CloudSat and CALIPSO satellites.

Combining lidar and radar measurements provide better cloud detection and characterization because of their unique compatible capabilities. Now combining radar and lidar measurements are widely used for cloud studies from cloud macrophysical and microphysical properties. CloudSat and CALIPSO satellites will provide us first opportunity to study cloud from space by combining lidar and radar. There are more advantages to combining lidar and radar measurements from space (than from the ground) because lidar is able to detect more cloud layers from a nadir view from space in the presence of multi-layer cloud systems. In general, cloud optical thickness decreases with altitude, thus lidar has a better chance of penetrating high and mid-level cloud than low-level clouds.

4. File Naming Convention and Configuration Control (Versioning)

   a. File Naming Convention

   YYYYDDDHHMMSS_NNNNN_CS_2B-TAU_GRANULE_S_RVV_EVV.hdf

   YYYYDDDHHMMSS = Year, Julian day, hour, minute, second of the first data contained in the file (UTC)
   NNNNN = Orbit number
   CS = Literal "CS" indicating the CloudSat mission
   2B-TAU = Product name (any product listed above)
   GRANULE = Indicates the data subset ("GRANULE" if the data are not subset)
   S = Data Set identifier ("P" for production release)
   RVV = Reprocessing version number for this product
   EVV = Epoch number for this product and reprocessing version
   .hdf = Suffix denoting HDF-EOS file

   Example:

   2007233124508_24388_CS_2B-TAU_GRANULE_P_R04_E02.hdf

   b. Configuration Control

   Data contained in a particular file produced by the DPC are influenced by a number of factors: the algorithm (code) that generated it, the structure of the product (field names, scale factors, etc.), changes that occur throughout the mission (radar calibration or other static input data), the number of times the data have been reprocessed, and any of these that affect products that are used as input to the algorithm that produced the file (an "upstream" product).

   The configuration control mechanism that has been implemented at the DPC enables us to track changes to the system and allows you to know exactly what changes have occurred that might affect the data in your file. To understand your data, you need to know the product name, data set version, and epoch number (see the file naming convention above).

   The product name is self-explanatory. The data set version keeps track of how many times data in the file have been reprocessed. The epoch number ties the two together and lets you know if any changes have occurred during the processing of the data set for that product.

   For example: you have a 2B-TAU product file for the R04 data set. From the DPC website you know that R04 is the most recent reprocessing run for the 2B-TAU product. If the epoch number is 04 you know that four changes have been made to the system since the beginning of this reprocessing run that affect the data in your file. From the DPC website or the documentation that accompanied your data order you can find out what those changes were and when they occurred in addition to the product version is for your file (e.g., 2B-TAU 004). The changes might include a product version upgrade, a calibration change, or a change that was made to an upstream product that affected the data in your file.

5. Product Data Volume

Table 2 lists the file size, per granule, for each of the CloudSat standard data products as of algorithm release 4.0.   File sizes are given in both compressed and uncompressed state.   Files transferred via ftp will be compressed.

Note: Sizes are approx. as the contents affect both hdf and compressed sizing …	SIZE PER ORBIT UNCOMPRESSED HDF Format (MB)	SIZE PER ORBIT COMPRESSED ZIP Format (MB)	TOTAL PER DAY COMPRESSED ZIP Format (GB)
1B-CPR	20.0	8.0	0.1
2B-GEOPROF	33.0	13.0	0.2
2B-CLDCLASS	19.0	7.0	0.1
2B-CWC-RO	160.0	99.2	1.4
2B-TAU	9.0	5.6	0.1
2B-CWC-RVOD	160.0	99.2	1.4
2B-FLXHR	34.0	21.1	0.3
ECMWF-AUX	90.0	53.0	0.8
MODIS-AUX	51.0	7.0	0.1
2B-GEOPROF-Lidar	33.0	13.0	0.2
2B-CLDCLASS-Lidar	19.0	7.0	0.1
	628.0	**333.1**	**4.8**

Table 2.   CloudSat Standard Data Product File Sizes

6. Data Access

During the operations phase, CloudSat data will be stored and distributed by CIRA. CIRA will provide an on-line data access system that will allow users to view browse images and order data by date/time interval and geographic location.   Ordered data will be placed on an ftp server for retrieval by the requestor.

Data will not routinely be provided to users via magnetic or optical media.   Larger data volumes (for example a user who places an order toward the end of the mission for every orbit of a group of products for the entire mission) will be provided on media at the expense of the requestor.   Note: if that same user requests that data be provided routinely during the course of the mission, those data can easily be transferred via ftp.   The intent is to encourage users to request and acquire data in smaller increments over the life of the mission.

a. Data Release and Availability

Data will be released, in three phases, following an initial "Launch and Early Orbit" checkout period. The first data release will be to the CloudSat Algorithm developers to perform their initial assessment of the data products.   Following their checkout, data products will be released to the CloudSat science team for further validation and review.   Approximately 40 days later, the first data products will be released to the general science community.   After the release of data by the PI and Science Team, "current" data products will be released, along with all previous data products that have been produced prior to the release date.

i. Launch and Early Orbit Checkout Period

The first CloudSat data dump will occur between 30 and 50 days after launch.   The algorithm checkout period will extend for approximately 4 months beyond that date.

ii. Normal Operations

After the initial checkout period, when data products are released by the Principal Investigator and CloudSat Science Team, data will be made available via the CloudSat DPC Data Ordering System, on the following schedule:

PRODUCT	ESTIMATED AVAILABILITY TO GENERAL SCIENCE COMMUNITY (average number of hours or days after data sampling time)	NOTES
1B-CPR-First-Look	8.5 hours	Geoloc. based on GPS data
1B-CPR	1 day	Geoloc. based on Def. Ephem.
2B-GEOPROF	2.5 days	Products generated after an approx. 2-day lag for MODIS data transfer to the CloudSat DPC
2B-CLDCLASS	2.5 days	
2B-TAU-OFF-N	2.5 days	
2B-LWC	2.5 days	
2B-IWC	2.5 days	
2B-FLXHR	2.5 days	
2B-GEOPROF-LIDAR	~2 days	Lag in receipt of CALIPSO data at the CloudSat DPC is not yet determined …
2B-CLDCLASS-LIDAR	~2 days	

Table 3. Data Lag Time

These data availability times include the data ingest, processing, and quality assessment review at the CloudSat Data Processing Center.

b. Data Ordering System

Several features of the data ordering system will allow users to request subsets of the larger CloudSat data product list – thus minimizing the volume of data that they will pull over to their local system. Users will be allowed to specify specific fields within a data product that they need for their research. They will also be able to create "custom products" that will combine fields from multiple Standard Data Products into a single file. In addition, users will be able to set filters on the data to select, for example, orbits that contain a higher percentage of water (vs. land) points, or granules that contain less than 30% cloud cover, etc.

All of the CloudSat data products will be available in HDF-EOS (HDF-4) format and can also be output in a user selectable binary format.

Figure 4 shows several screen shots from the main page of the Data Ordering System. The interactive system contains context help and a comprehensive help system with keyword search and indexing.

Figure 4. Screen shot of the CloudSat Data Ordering System main page.   Users have the option of selecting a range of dates, geographic location boundaries, product type, individual fields, and selecting on criteria such as percent cloud, day/night, percent "good" data, for example.   The interface works like a standard web-based shopping cart with data placed into a unique directory for ftp access by the data requestor.

The CloudSat Data Distribution System interface and help file can be accessed directly using the link http://www.cloudsat.cira.colostate.edu/data_dist/orderdata.php or from the "Data Access" tab on the DPC website at http://www.cloudsat.cira.colostate.edu/ .

7. References

Ackerman, S. A., K. I. Strabala, W. P. Menzel, R. A. Frey, C. C. Moeller, L. E. Gumley, 1998: Discriminating clear sky from clouds with MODIS. *J. Geophys. Res.,* **103,** 32141-32157.

Stephens, G. A., D. Vane, and many others, 2001: The CloudSat mission and the EOS constellation: A new dimension to space-based observations of clouds and precipitation. Submitted to *Bull. Amer. Meteoro. Soc.*

8. Interface Specifications / Data Formats

Current Data Format Specifications can be viewed via the CloudSat Data Processing Center website: http://www.cloudsat.cira.colostate.edu/dataSpecs.php .   This is the "product documentation" link and is the most current information about data products and product ICD's.