# CloudSat-GPM Coincidence Dataset

## Version 2A
## (includes CloudSat and GPM precipitation products)

**Prepared by:**

**Joe Turk**
**Jet Propulsion Laboratory, California Institute of Technology**

**13 March 2017**

## 1. **Introduction**.

Owing to the Global Precipitation Measurement (GPM) core satellite's unique asynchronous orbit, its orbital ground tracks intersect the orbital tracks of many other sun-synchronous satellites. Of particular interest are the intersections (coincidences) between the GPM core satellite and the 94-GHz (W-band) CloudSat profiling radar (CPR), within small enough time differences, such that the combination of the resulting "pseudo three-frequency" radar profiles (W-band from CPR, and Ku/Ka-band from GPM), and the 13-channel (10-183 GHz) GMI radiometer are useful for many scientific purposes. Examples include algorithm evaluation and identification of deficiencies, snow and light rain sensitivity studies, cloud process studies, radiative transfer simulations, and studying land surface effects on the radar, radiometer, or combined-sensor precipitation retrieval algorithms.

## 2. **Data Processing**.

The processing flow is driven by a pre-generated list of the ±15-minute time coincidences between the GPM and CloudSat ground tracks, obtained from orbital propagation using two-line elements (TLE), which are updated every day (sufficient accuracy for this particular use). The current dataset spans the period between 8 March 2014 (the first date that DPR data are available) and 1 January 2017 (CloudSat data partially available, Table 1). A total of 8367 orbit coincidences occurred within this time span. Owing to the fact that CloudSat only operates on the ascending (daytime) orbit, this cuts that number by approximately one-half. Accounting for various missing data periods from either satellite, etc., a total of 3982 coincidence datasets are available in this Version 2A collection.

| Dataset Name | Satellite | Description | Date Available |
|---|---|---|---|
| **2A.GPM.DPR** | GPM | DPR Ku and Ka-band radar reflectivity profile and radar-based precipitation retrievals, Version 04A | 2014/03/08-current |
| **1B.GPM.GMI** | GPM | GMI Level 1B brightness temperatures, Version 04A | 2014/03/05-current |
| **2B.GPM.DPRGMI** | GPM | DPR+GMI combined precipitation profiling algorithm, Version 04A | 2014/03/05-current |
| **2A.GPM.GMI.GPROF** | GPM | GPROF GMI-based precipitation profiling algorithm, Version 04A | 2014/03/05-2016/12/03 |
| **2B-GEOPROF** | CloudSat | CloudSat vertical reflectivity profile | 2006/06/15-2016/06/06*. Daylight only ops since Jan 2012. |
| **ECMWF-AUX** | CloudSat | ECMWF forecast analysis interpolated to each vertical CloudSat bin | 2006/06/15-2016/06/06*. |
| **2C-SNOW-PROFILE** | CloudSat | CloudSat snowfall profile | 2014/03/10-2016/05/21* |
| **2C-RAIN-PROFILE** | CloudSat | CloudSat precipitation profile | 2014/03/10-2016/05/21* |
| **2C-PRECIP-COLUMN** | CloudSat | CloudSat column-averaged precipitation rate | 2014/03/10-2016/05/21* |
| **1C.NOAA19.MHS.XCAL** | NOAA-19 | MHS Level 1C brightness temperatures, Version 04A | 2014/03/10-current** |

**Table 1.** GPM and CloudSat datasets used for creating the coincidence dataset. *The various CloudSat datasets are updated intermittently, and indicate the most current data available at the time of this update. The blue shaded rows indicate the additional CloudSat and GPM-based precipitation variables that were added in Version 2A. **MHS data not available for all coincidences, see Section 5d.

First, the CloudSat 2B-GEOPROF and ECMWF-AUX files are located that straddle the coincidence time. Next, the corresponding GPM 1B.GPM.GMI and 2A.GPM.DPR files are located. Since the CPR has finer horizontal resolution and along-track sampling (~1-km) than the GPM radars (~4-km), the profile matching procedure is intentionally a very simple method, based on a nearest-neighbor approach to avoid any spatial or temporal resampling artifacts. Future versions may consider a resampling of CloudSat to match the DPR spatial locations and vertical bin placement, but this feature is not implemented in this dataset. A number of associated solid and liquid precipitation products from Table 1 (shaded rows) are extracted for each CloudSat position (and written into a separate HDF5 group, see end of this document).

Since the GPM launch, the orbit of NOAA-19 (with the Microwave Humidity Sounder, MHS) and CloudSat are in a similar inclination (slightly different elevation) and local afternoon crossing time, so every few days there are both spatial and ±15 minute temporal alignments between CloudSat and NOAA-19. For this update cycle, 890 of these triple coincidences (MHS-CloudSat-GPM) were processed. The associated MHS 5-channel TB (matched to CloudSat locations) included in these datasets are discussed in Section 7.

*a) Along-track segments.*

The horizontal matching procedure is depicted in Figure 1. For each CPR profile (latitude, longitude) location, the (along-track, across-track) coordinates of each GPM dataset (DPR and GMI) are located, that are closest to this CPR location. For the vertical bin matching, the nearest DPR bin is located and matched to each CloudSat vertical bin height (i.e., nearest neighbor). For the DPR off-nadir positions, the DPR slant path is not considered, except to adjust the DPR elevation by the *cos*(zenith) term.

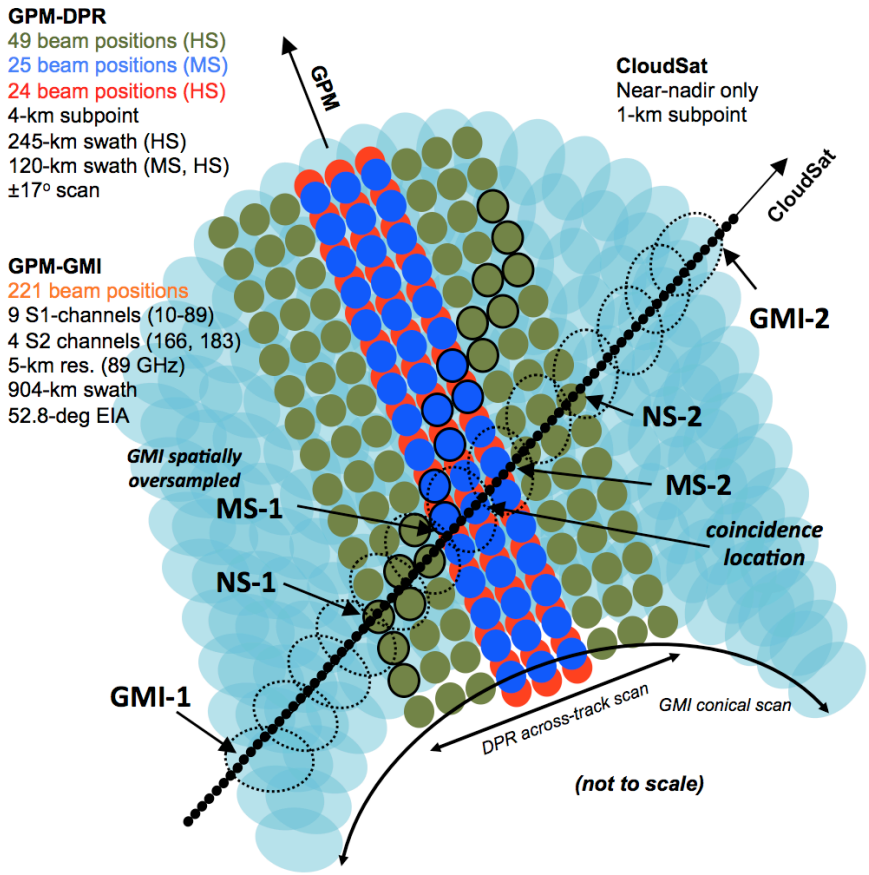


**Figure 1**. Illustration of a CloudSat-GPM coincidence segment. The location of each 1-km CPR beam position is denoted by the small black dots. The location of each DPR Ku-band swath (normal scan, NS) pixel is shown in green, the Ka-band (matched swath, MS) is in blue, and the Ka-band high-sensitivity (HS) scan is in red. The lower-resolution conically scanning GMI footprints spatially overlap across the swath, depending upon channel (light aqua). GMI-1, GMI-2 identify the GMI scan numbers that cover the coincidence segment, similarly NS-1, NS-2 and MS-1, MS-2 for the DPR NS and MS, respectively.

The 2A.GPM.DPR dataset carries three radar profiles, the Ku-band normal scan (NS), the Ka-band matched scan (MS), and the Ka-band high-sensitivity scan (HS). For each 240-m CPR vertical bin, the DPR radar profiles are searched from bottom-to-top until the first bin that lies above each CPR bin is identified, and this is reported as the associated DPR matched bin (Figure 2). This procedure is repeated for each of the NS, MS and HS profiles. Since the NS and MS profiles are reported at 125-m bin spacing, with this approach only every other NS and MS bin is used (HS is reported at a 250-m bin spacing, so all bins are used). This simple nearest bin in latitude, longitude and elevation does not modify or average any values in the native dataset. However, full-swath segments are also provided for users who desire different bin-matching methods.

For all three variables, only the "uncorrected" (zFactorMeasured) DPR reflectivities are carried in the coincidence dataset, to be self-consistent with the nature of the CPR reflectivity profile, and to not carry any assumptions on any particular attenuation-correction technique. However, as seen in the dataset description for this coincidence product, the coincidence files carry the

(scan, ray) values needed to return to these same CloudSat and GPM data files, or other derived-products from this same GPM orbit revolution, and extract other variables at the same DPR and GMI sensor coordinates. With the ~4-km horizontal spacing between each GPM radar location, using the nearest-neighbor approach the same DPR radar profile repeats itself 4 or 5 times, as the CPR moves across the GPM radar across-track swath. The length and time duration of the coincidence segment between CPR and GPM is longer nearest the equator, where GPM is flying "more parallel" to the CloudSat orbital direction than it does at the higher latitudes.

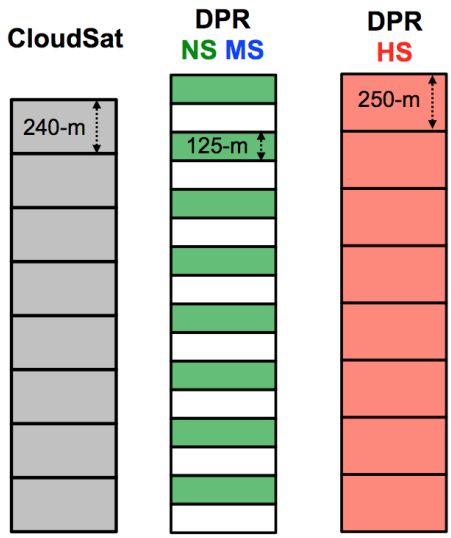


**Figure 2**. Depiction of nearest-neighbor bin matching. CloudSat is reported at 240-m vertical bin spacing, whereas at nadir DPR is reported at 125-m for the normal scan and matched scan, and 250-m for the high sensitivity scan. For each of NS, MS and HS, the "bin matched" DPR bin is reported as the bin whose top lies just above the CloudSat bin top. This is approximately every other NS and MS bin. Note that the coincidence point can be from any DPR zenith angle, but no attempt is made to account for the DPR slant path, other than to adjust the DPR elevation by the cos(zenith) term. Although more sophisticated along path averaging could have been done, this method intentionally does not alter the intrinsic values carried in each of the CloudSat and GPM native datasets. Full-swath segments are included for users who desire different matching schemes.

For the GMI radiometer, the first nine channels (10-89 GHz) are denoted as the S1 channels and are reported at 5-km spacing across the conical scan, to maintain contiguous across-track coverage at the highest resolution 89 GHz channels, with increasing spatial oversampling at the lower frequencies. The four high frequency channels (166V, 166H, 183±3, 183±8 GHz) are denoted as the S2 channels, and are also reported with this same 5-km spacing, but are sampled at slightly different locations. Similar to the DPR, the matching procedure first locates the (line, sample) coordinates of the S1 channels that are closest to each CPR beam location. Then a nearby search of S2 coordinates is done to find the S2 location that is closest to this S1 location, within no more than 5-km. Since the S2 channels are reported across a slightly narrower swath width than the S1 channels, this leaves about 15 missing S2 channels on both edges of the swath (marked as bad values in the netCDF files). Therefore, with this nearest neighbor approach, the same GMI location is reported about 4-5 times, as the CPR moves across the DPR across-track swath. No averaging or resampling is done, nor any attempt to resolve the lowest-resolution 10-37 GHz channels to a common resolution.

From the associated ECMWF-AUX file, the vertical profile of temperature, pressure and specific humidity is added, as well as the surface skin temperature and 2-m air temperature. These environmental variables useful for forward simulations of radar reflectivity profiles and passive microwave brightness temperatures.

*b) Full-swath segments.*

As mentioned above, to provide the user more flexibility in remapping or matching these two datasets, "full swath" GMI and DPR segments are added to each dataset, which cover an additional 50 scans on either side of the coincidence segment. For example, 1B.GPM.GMI files typically contain about 2930 scan lines (one orbit) and 221 pixel locations. If the CPR coincidence locations were located between GMI scan lines 1000-1050 (51 scans, denoted by "GMI-1" and "GMI-2" in Figure 1), then the "swath" segment carries the full swath of GMI TB data between scan lines 950 and 1100 (e.g., an array sized 151 x 221 x 13, along with the associated latitude and longitude arrays). A similar procedure is carried out for each of the DPR NS, MS, and HS radars. With these full resolution data segments, users are able to carry out more sophisticated pixel and bin matching procedures, or display TB imagery.

## 3. File Format and Naming Details.

The output files are written in netCDF4 format, although the files are compatible with HDF5 standards (e.g., the h5py module in python works OK) and are fairly small, averaging about 10 MB size, depending upon latitude (at the highest GPM latitudes, the GPM radar scans more in a N-S direction, so the coincidence distance segment is shorter than it is near the equator). The file names are chosen to be consistent with existing PPS standards. An example name is:

**2B.CSATGPM.COIN.25S_174W_13575_000_295_403.20140309-S013654-E013756.000146.V01A.NC**

The fifth field carries the start and end times of the coincidence segment (**20140309-S013654-E013756**), the sixth field is the GPM orbit revolution number (**000146**), and the fourth field provides specific quality information on the coincidence (**25S_174W_13575_000_295_403**). The interpretation of the six values within this field is as follows:

| Field | Value from the example | Description |
|---|---|---|
| 1 | 25S | Latitude of the coincidence, to the nearest degree |
| 2 | 174W | Longitude of the coincidence location, to the nearest degree |
| 3 | 13575 | Number of CloudSat bins where the cloud mask >= 40 |
| 4 | 000 | Percent of CloudSat profiles that are over land |
| 5 | 295 | Minimum 2-m air temperature from all CloudSat profiles in the dataset |
| 6 | 403 | Time offset (absolute value) between CloudSat and GPM |

Therefore, by simply parsing the filename, coincidences meeting specific requirements can easily be located without having to open and read the file. For example, to isolate over-land snowfall cases, a user would look for files with the 2-m air temperature near or below 273K, 100 percent over land, and with a sufficiently large number of CloudSat bins that identified cloud.

## 4. Associated Quick-Look Imagery.

For each output datafile, three associated images are created. The first is for the DPR-NS segment where GMI and all four radar profiles (CPR and the DPR NS, MS and HS) are always available. The second is for the full GMI swath segment where GMI and CPR are always available. The third for the TB from six of the GMI channels, mapped onto a polarstereo grid centered at the coincidence location. These have the same name as the datafiles, but with a "**profile.png**", "**profile_GMI.png**" and "**tb.png**" filename extension added, respectively. For the filename given as an example above, the multi-radar radar profile imagery is shown in Figure

3 from 9 March 2014, extending between 01:36:54 to 01:37:56 UTC (62 seconds). In this case, the coincidence is from a CloudSat ascending overpass in the South Pacific (25S, 174W), and with a fairly intense line of convection is easily noted in these four reflectivity profiles. Significant path-integrated attenuation and multiple scattering is evident in the CloudSat profile, which extends above 15-km elevation, and the Ka-band signal is also severely attenuated (surface reflection is lost at several locations between CloudSat beams 130 and 200). The digital elevation map (DEM) used by the CloudSat and GPM data are shown as a heavy black line near the surface (not so obvious here, but is evident for the over-land cases that are over steep terrain). The zero-degree isotherm line is shown as a blue dashed line (labeled "T=273K"), and the bottom panel shows the line-trace of the 13 GMI channels (increasing in frequency from top to bottom), in an image format. The color scale on the right indicates TB Kelvin units (blue=cold, red=warm).

Figure 4 illustrates the CloudSat profile that now covers the extent of the GMI swath (the DPR profiles cover a more limited segment of this near the middle and are not shown), from 01:35:30 to 01:39:16 UTC (226 seconds). The associated line-trace of the 13 GMI channels (note the more limited swath width of the S2 channels) is shown below.

The associated GMI TB imagery is shown Figure 5, for the 10H, 18H, 37H, 89H, 166H and 183±8 channels (Kelvin). The image is centered using the "center_lat" and "center_lon" global attribute values (see format description at the end of this document) . The heavier black line segment in each panel shows the CloudSat track, and the solid and dashed black lines depict the swath extent of the DPR NS and MS region. The beginning and end times of the DPR-NS intercept are labeled. The color scale on the bottom indicates TB Kelvin units (blue=cold, red=warm).

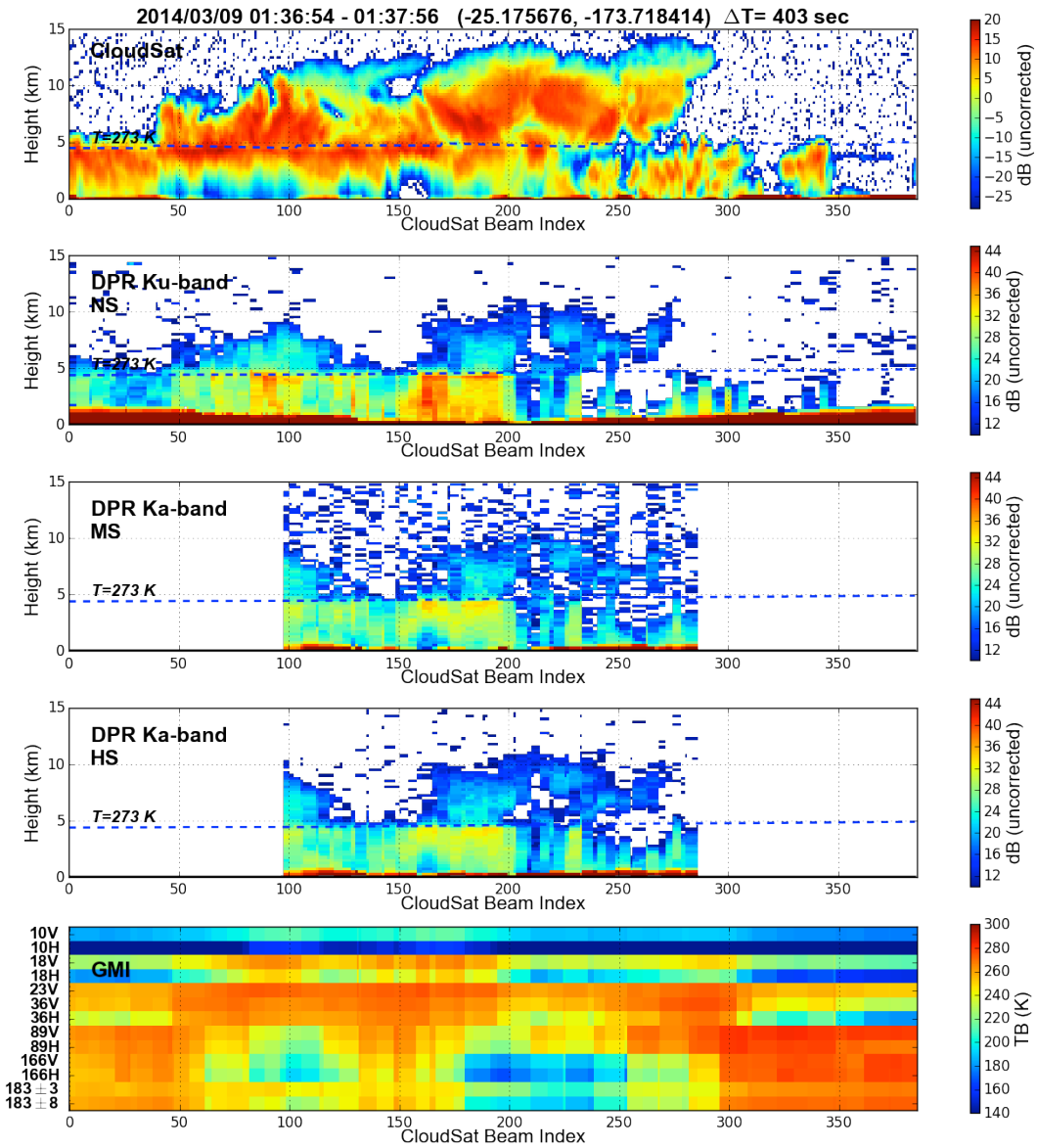


**Figure 3**. Example quick-look radar time-height profile imagery, for the CloudSat and GPM NS, MS and HS scans (from top to bottom). Each CloudSat beam is about 1-km apart. The digital elevation map (DEM) used by the CloudSat and GPM data are shown as a heavy black line near the surface (not so obvious here, but is evident for the over-land cases that are over steep terrain). The zero-degree isotherm line is shown as a blue dashed line (labeled "T=273K"), and the bottom panel shows the line-trace of the 13 GMI channels (increasing in frequency from top to bottom), in an image format. The color scale on the right indicates TB Kelvin units (blue=cold, red=warm), and equivalent radar reflectivity factor (dB) for the four radar profiles.

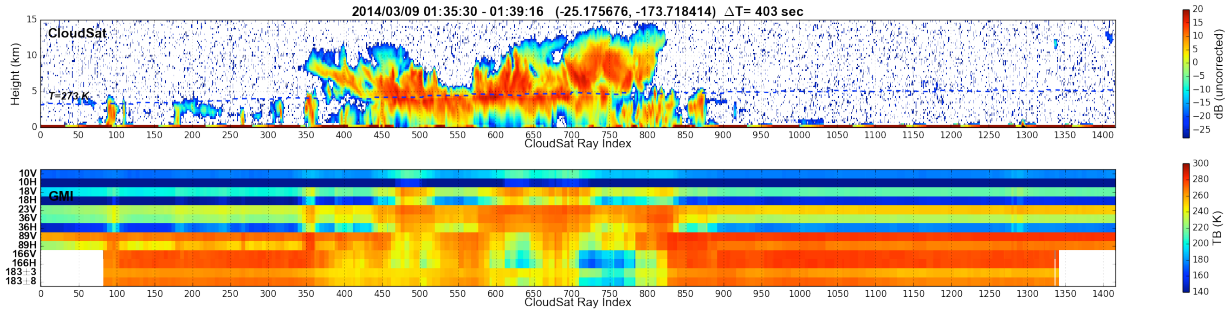


**Figure 4**. Same as Figure 3, except the horizontal axis extends across the full GMI swath, and the three DPR radar beams are not shown. The slightly limited swath extent of the four GMI S2 channel is evident. The actual image size is much larger, shown here resized to fit to page width.

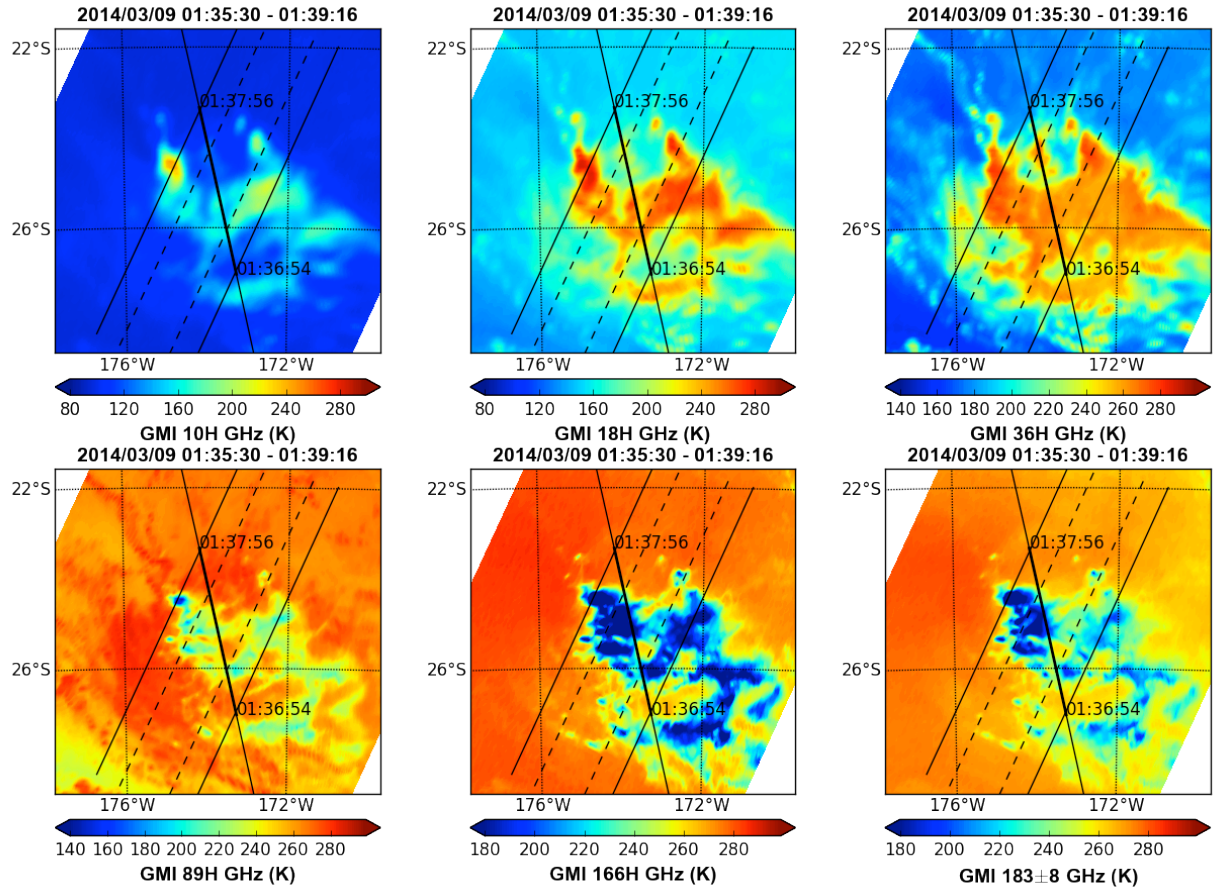


**Figure 5**. GMI TB imagery corresponding to the radar profiles shown Figure 3. From upper left, 10H, 18H, 37H, 89H, 166H and 183±8 channels (Kelvin). The image is centered using the "center_lat" and "center_lon" global attribute values (see example at the end of this document). The heavier black line segment in each panel shows the CloudSat track, and the solid and dashed black lines depict the swath extent of the DPR NS and MS region. The beginning and end times of the DPR-NS intercept are labeled. The color scale on the bottom indicates TB Kelvin units (blue=cold, red=warm).

Figures 6,7 and 8 show another example, an over-land (CloudSat ascending) case in the

Himalayas where the terrain effects clearly evident above the DEM level. The relative sensitivities of the Ku- and Ka-band radars, including the HS scan, are more apparent.

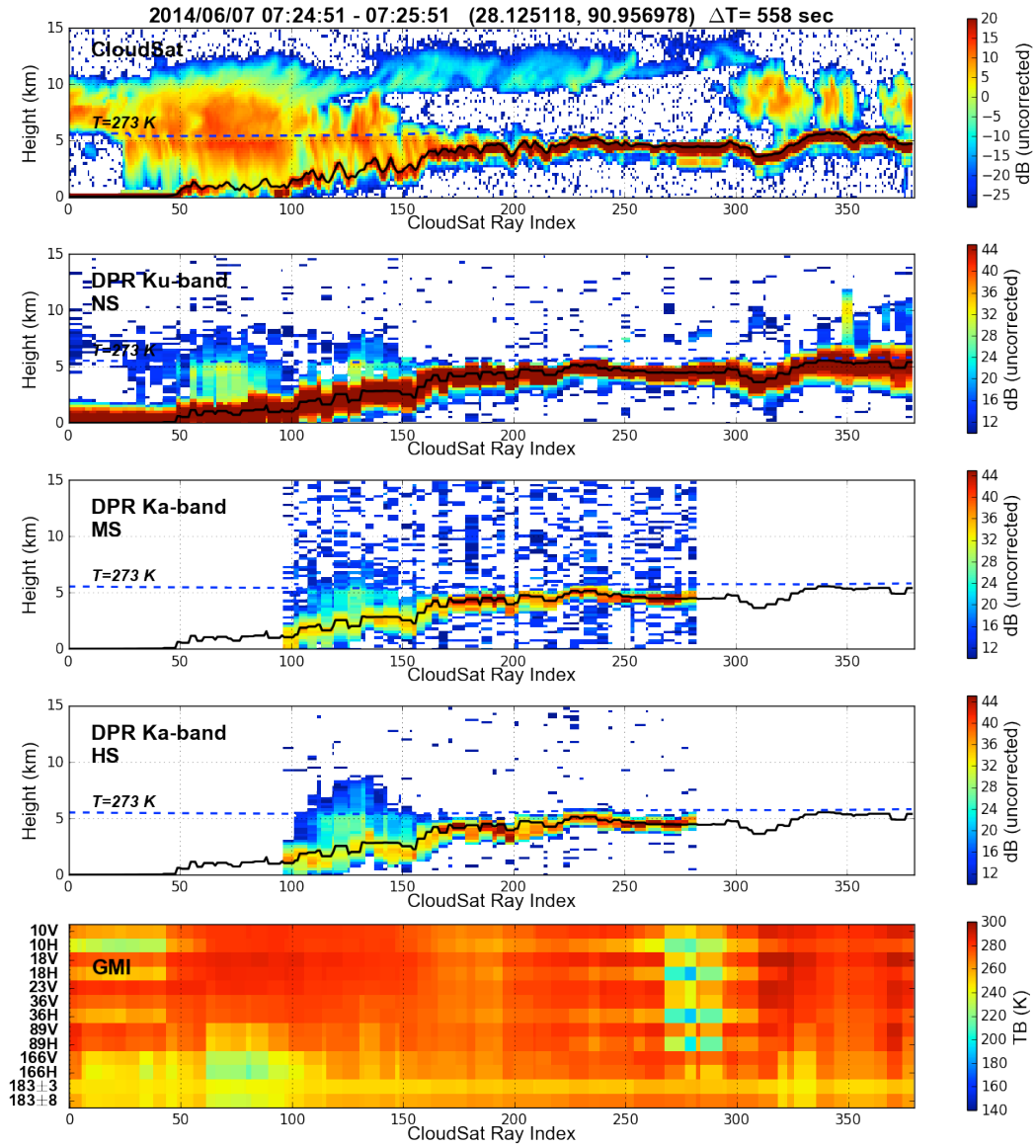


**Figure 6**. Same as Figure 3, but for the ascending CloudSat overpass on 7 June 2014, near 28N 91E (crossing the Himalayas).

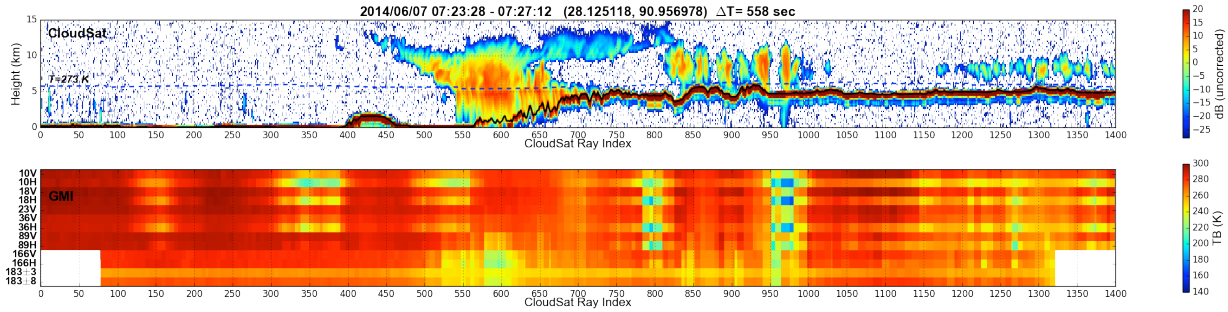


**Figure 7.** Same as Figure 4, but for the ascending CloudSat overpass on 7 June 2014, near 28N 91E (crossing the Himalayas).

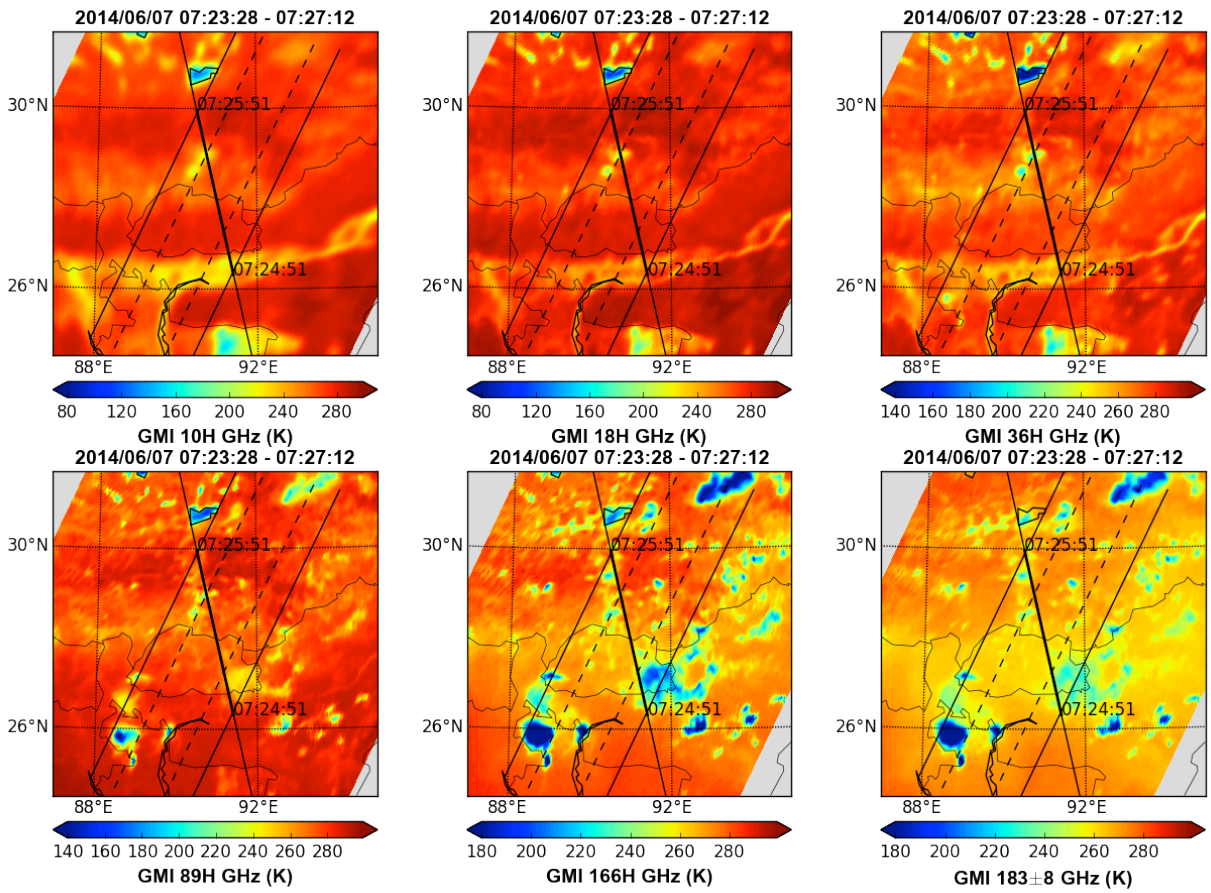


**Figure 8**. Same as Figure 5, but for the ascending CloudSat overpass on 7 June 2014, near 28N 91E (crossing the Himalayas).

## 5. Description of File Contents.

The output files are written in netCDF4 format, although the files are compatible with HDF5 standards (e.g., the h5py module in python works fine) and are fairly small, averaging about 10 MB size. For CloudSat, which does not scan, the along-track dimension is referred to as a *ray*.

For DPR, the along-track and across-track dimensions are referred to as a *scan* and *ray*, respectively. For GMI, the along-track and across-track dimensions are referred to as a *scan* and *pixel*, respectively. This maintains close to the native dataset nomenclature.

Running "**`ncdump -h`**" command on a file will lists every variable contained in the file, type, units, scale/offset if used, and a brief description. Whenever possible, the same variable name as the input dataset is used. Each file has six groups: **CS, CS_PRECIP, S1, NS, MS and NS**.

### *a) Curtain Coincidence Variables (type and dimensions)*

"**CS**" refers to the CloudSat curtain group for all non-retrieved variables (as of Version 2A, all precipitation retrievals are in the **CS_PRECIP** group, explained in the next section). Variables within this group are either 1-D or 2-D (the exception is the GMI TB, written as a 3-D variable, below). For example, in the file named,

**`2B.CSATGPM.COIN.57N_142E_05696_032_279_653.20140615-S032452-E032726.001672.V01B.HDF5`**

the dimensions of the variables in the **CS** group are:

```
nray_CS = 967          /* number of CloudSat rays in the GMI coincidence segment */
nlev_CS = 125          /* number of CloudSat bins, same as 2B-GEOPROF */
ntb_GMI = 13           /* number of GMI S1+S2 channels, same as 1B.GPM.GMI */
```

Several global attributes are added, which provide the names of the input datasets, and other relevant information, including the latitude and longitude of the subpoint coincidence location (useful to center a map projection), total number of over-land and over-ocean rays, number of CloudSat bins that exceed a cloud mask value of 30 (weak cloud detection) or 40 (positive cloud detection). For the example file mentioned above, the global attributes are:

```
2B-GEOPROF = "2014166022039_43252_CS_2B-GEOPROF_GRANULE_P_R04_E06.hdf"
2A.GPM.DPR = "2A.GPM.DPR.V5-20140827.20140615-S025634-E042905.001672.V03B.HDF5"
1B.GPM.GMI = "1B.GPM.GMI.TB2014.20140615-S025634-E042905.001672.V03B.HDF5"
ECMWF-AUX = "2014166022039_43252_CS_ECMWF-AUX_GRANULE_P_R04_E06.hdf"
2C-SNOW-PROFILE = "2014166022039_43252_CS_2C-SNOW-PROFILE_GRANULE_P_R04_E06.hdf"
2C-RAIN-PROFILE = "2014166022039_43252_CS_2C-RAIN-PROFILE_GRANULE_P1rc1_R04_E06.hdf"
2C-PRECIP-COLUMN = "2014166022039_43252_CS_2C-PRECIP-COLUMN_GRANULE_P2_R04_E06.hdf"
2B.GPM.DPRGMI = "2B.GPM.DPRGMI.CORRA2015.20140615-S025634-E042905.001672.V04A.HDF5"
2A.GPM.GMI.GPROF = "2A.GPM.GMI.GPROF2014v2-0.20140615-S025634-E042905.001672.V04A.HDF5"
production_date = "2016/01/14 22:43:25"  /* date that this file was created */
start_date = "2014/06/15 03:24:52" /* Time of the first CS ray in the GMI segment */
end_date = "2014/06/15 03:27:26"   /* Time of the last CS ray in the GMI segment */
start_date_NS = "2014/06/15 03:25:47"    /* Time of the first CS ray in the DPR-NS segment */
end_date_NS = "2014/06/15 03:26:29"      /* Time of the last CS ray in the DPR-NS segment */
center_lat = "56.979408"           /* latitude of the subpoint coincidence */
center_lon = "142.399323"          /* longitude of the subpoint coincidence */
CS_minus_NS_time_diff_seconds = "-653"   /* time offset at the subpoint coincidence */
ray_index_range_NS = "347 610"     /* indices of the segment that intercept the DPR-NS swath */
ray_index_range_MS = "414 542"     /* indices of the segment that intercept the DPR-MS swath */
CS_bin_height_in_meters = "240"    /* CloudSat bin size */
NS_bin_height_in_meters = "125"    /* DPR normal scan bin size */
MS_bin_height_in_meters = "125"    /* DPR matched scan bin size */
HS_bin_height_in_meters = "250"    /* DPR high-sensitivity scan bin size */
CS_total_bins_mask_ge_30 = "5800"  /* total bins 2B-GEOPROF cloud mask >= 30 */
CS_total_bins_mask_ge_40 = "5696"  /* total bins 2B-GEOPROF cloud mask >= 40 */
CS_nray_land = "318"               /* CS rays that are over land or coast */
CS_nray_ocean = "649"              /* CS rays that are over water */
N_snow_profiles = "0"              /* Number of CS rays that detect snow */
```

For example, in this file there are 967 CloudSat rays that cover the full GMI swath. Of these, ray numbers 347 to 610 cover the DPR-NS swath (i.e., "NS-1" and "NS-2" in Figure 1), and ray numbers 414-542 cover the narrower DPR-MS swath (i.e., "MS-1" and "MS-2" in Figure 1).

The variables that are associated with the coincidence curtain shown in Figures 3 and 6 are described below.

**double time (nray_CS)**
Time in seconds since 1 Jan 1970, created from the "TAI_start" variable from 2B-GEOPROF.

**integer DEM (nray_CS)**
Digital elevation map (DEM), in meters above the reference surface, copied from 2B-GEOPROF.

**integer SurfaceHeightBin (nray_CS)**
CloudSat surface height bin, copied from 2B-GEOPROF.

**integer land_sea_flag (nray_CS)**
CloudSat land_sea_flag, copied from 2B-GEOPROF, where 1=land, 2=ocean, 3=coast.

**float Latitude (nray_CS)**
CloudSat latitude in degrees, copied from 2B-GEOPROF.

**float Longitude (nray_CS)**
CloudSat longitude in degrees, ranging from -180 to 180, copied from 2B-GEOPROF.

**float along_track_dist (nray_CS)**
Accumulated distance, in km along track, starting from first CloudSat ray. This is included for setting the x-axis scale to distance (rather than ray number) in creation of any figures.

**float localZenithAngle_NS (nray_CS)**
DPR normal scan (NS) local zenith angle, in degrees from nadir, copied from 2A.GPM.DPR, that is closest to the CloudSat ray.

**integer elevation_NS (nray_CS)**
DPR normal scan (NS) elevation, in meters above the reference surface, copied from 2A.GPM.DPR, that is closest to the CloudSat ray.

**integer heightZeroDeg_NS (nray_CS)**
Height of the 0-deg Celsius level, in meters above the reference surface, copied from 2A.GPM.DPR, that is closest to this CloudSat ray.

**integer ray_index_CS (nray_CS)**
Value of the CloudSat along track index, from 2B-GEOPROF. Typically there are 37800 CloudSat rays per orbit. For example, if the coincidence segment starts at CloudSat ray 100 and nray_CS=500, this value would range from 100 to 599. This is included in case the user wants to go back to the 2B-GEOPROF file and perform any additional processing.

**integer scan_index_NS (nray_CS)**
Value of the DPR NS along track index, from 2A.GPM.DPR, that is closest to the CloudSat ray. Typically there are approximately 7900 DPR scans per orbit. For example, if the coincidence segment starts at DPR NS ray 100 and nray_CS=500, and since each DPR ray "repeats" about four times (Figure 1), this value would range from 100 to about 224. This is included in case the user wants to go back to the 2A.GPM.DPR file and perform any additional processing.

**integer ray_index_NS (nray_CS)**
Value of the DPR NS across track index, from 2A.GPM.DPR, that is closest to the CloudSat ray. There are 49 NS ray positions per scan. This is included in case the user wants to go back to the 2A.GPM.DPR file and perform any additional processing.

**integer scan_index_S1 (nray_CS)**
Value of the GMI S1 along track scan index, from 1B.GPM.GMI, that is closest to this CloudSat

ray. Typically there are approximately 2900 GMI scans per orbit. This is included in case the user wants to go back to the 1B.GPM.GMI file and perform any additional processing.

**integer pix_index_S1 (nray_CS)**
Value of the GMI S1 across track scan index, from 1B.GPM.GMI, that is closest to this CloudSat ray. There are 221 GMI pixels per scan. This is included in case the user wants to go back to the 1B.GPM.GMI file and perform any additional processing.

**float Temperature_2m (nray_CS)**
ECMWF 2-m air temperature, in Kelvin, copied from ECMWF-AUX. This is useful for discriminating precipitation phase near the surface.

**float Skin_temperature (nray_CS)**
ECMWF surface temperature, in Kelvin, copied from ECMWF-AUX.

**float Surface_pressure (nray_CS)**
ECMWF surface pressure, in Pa, copied from ECMWF-AUX.

**integer height_273K (nray_CS)**
Height where the ECMWF temperature crosses 0-deg Celsius, determined from ECMWF-AUX. This is approximately the same value as **heightZeroDeg_NS**, but since CloudSat and GPM use different sources for their ancillary model data, the two values may be slightly different.

**short height (nray_CS, nlev_CS)**
Height above the reference surface, in meters, for each CloudSat bin, copied from 2B-GEOPROF.

**short Radar_Reflectivity (nray_CS, nlev_CS)**
CloudSat radar reflectivity profile, in dB scaled by 100, copied from 2B-GEOPROF.

**short zFactorMeasured_NS (nray_CS, nlev_CS) ;**
DPR normal scan (NS) profile of uncorrected Ku-band reflectivity (zFactorMeasured), in dB scaled by 100, from 2A.GPM.DPR, that is closest to this CloudSat ray.

**short zFactorMeasured_MS (nray_CS, nlev_CS) ;**
DPR matched scan (MS) profile of uncorrected Ka-band reflectivity (zFactorMeasured), in dB scaled by 100, from 2A.GPM.DPR, that is closest to this CloudSat ray.

**short zFactorMeasured_HS (nray_CS, nlev_CS) ;**
DPR high sensitivity scan (HS) profile of uncorrected Ka-band reflectivity (HS zFactorMeasured), in dB scaled by 100, from 2A.GPM.DPR, that is closest to this CloudSat ray.

**float Tb (nray_CS, ntb_GMI)**
Values of the GMI TB, in Kelvin, from 1B.GPM.GMI, that is closest to the CloudSat ray. The closest S2 channels are mapped to each S1 position. There are nine S1 channels and four S2 channels (**ntb**=13). Channel order is: 10V 10H 18V 18H 23V 36V 36H 89V 89H 166V 166H 183+/-3 183+/-8 GHz.

**float Pressure (nray_CS, nlev_CS)**
ECMWF pressure profile, in Pa, copied from ECMWF-AUX.

**float Temperature (nray_CS, nlev_CS)**
ECMWF temperature profile, in Kelvin, copied from ECMWF-AUX.

**float Specific_humidity (nray_CS, nlev_CS) ;**
ECMWF specific humidity profile, in kg/kg, copied from ECMWF-AUX.

### *b) Curtain Precipitation Retrievals (type and dimensions)*

From Table 1, there are a number of precipitation-related retrievals written into the "**CS_PRECIP**" group, originating from six datasets, three from GPM and three from CloudSat. The names of these six datasets are given in the global attributes in the example above. As above, the same variable name as the name in the original source file was maintained for ease of

recognition. The GPM-related products use the same nearest-neighbor mapping strategy as described above, and are therefore pixel-matched (in other words, the products at the same index in the **CS** and **CS_PRECIP** groups can be directly compared). Each of these are described below. Note that all six of these datasets are not always available, especially for CloudSat. For example, for some coincidences the 2C-PRECIP-COLUMN is available, but the 2C-SNOW-PROFILE is not. In that case, the global attributes won't list an entry for the 2C-SNOW-PROFILE file, and the corresponding arrays (shown below) for 2C-SNOW-PROFILE variables will contain all missing data default values.

### Variables from 2A.GPM.DPR (*DPR-only algorithm*)

**float piaFinal_NS_ku(nray_CS)**
Ku-band Path integrated attenuation in dB from DPR NS swath, copied from piaFinal in 2A.GPM.DPR.

**float piaFinal_MS_ka(nray_CS)**
Ku-band Path integrated attenuation in dB from DPR MS swath, copied from piaFinal in 2A.GPM.DPR.

**float precipRateNearSurface_NS(nray_CS)**
Surface precipitation rate in mm/hr from DPR NS swath, copied from precipRateNearSurface in 2A.GPM.DPR.

**float precipRateNearSurface_MS(nray_CS)**
Surface precipitation rate in mm/hr from DPR MS swath, copied from precipRateNearSurface in 2B.GPM.DPR.

### Variables from 2B.GPM.DPRGMI (*DPR+GMI combined algorithm*)

**float pia_NS_ku(nray_CS)**
Ku-band Path integrated attenuation in dB from DPRGMI combined NS swath, copied from pia in 2B.GPM.DPRGMI.CORRA.

**float pia_MS_ku(nray_CS)**
Ku-band Path integrated attenuation in dB from DPRGMI combined MS swath, copied from pia in 2B.GPM.DPRGMI.CORRA.

**float pia_MS_ka(nray_CS)**
Ka-band Path integrated attenuation in dB from DPRGMI combined MS swath, copied from pia in 2B.GPM.DPRGMI.CORRA.

**float surfPrecipTotRate_NS(nray_CS)**
Surface precipitation rate in mm/hr from DPRGMI combined NS swath, copied from surfPrecipTotRate in 2B.GPM.DPRGMI.CORRA.

**float surfPrecipTotRate_MS(nray_CS)**
Surface precipitation rate in mm/hr from DPRGMI combined MS swath, copied from surfPrecipTotRate in 2B.GPM.DPRGMI.CORRA.

### Variables from 2A.GPM.GMI.GPROF *(GMI-only precipitation rate*)

**float surfacePrecipitation(nray_CS)**
Surface precipitation rate in mm/hr from GPROF GMI swath, copied from surfacePrecipitation in 2A.GPM.GMI.GPROF.

**float probabilityOfPrecip(nray_CS)**
Probability of precipitation in percent from GPROF GMI swath, copied from probabilityOfPrecipitation in 2A.GPM.GMI.GPROF.

**float liquidPrecipFraction(nray_CS)**

Fraction of surface precipitation (0-100) that is liquid from GPROF GMI swath, copied from liquidPrecipFraction in 2A.GPM.GMI.GPROF.

## Variables from CS-SNOW-PROFILE (*CloudSat snow profile*)

**float snowfall_rate_sfc(nray_CS)**
Surface snowfall rate in mm/hr of liquid water, copied from 2C-SNOW-PROFILE.

**short snow_retrieval_status(nray_CS)**
Snowfall retrieval status, copied from 2C-SNOW-PROFILE. A one-byte (8-bit) field for retrieval status flags. The following table gives the bit position and the condition indicated if that bit is set. The table starts from the least significant bit (position 0).
0: Snow detected in the profile
1: Snow detected at the surface
2: Retrieval completed but produced large chi-square value
3: The estimated snowfall rate at the base of the snow layer is substantially larger than that in the bin immediately above.
4: Required inputs from surface data are missing or in error
5: Required inputs from the profile data are missing or in error
6: Retrieval attempted and converged to invalid values
7: Retrieval attempted and failed to converge

**float snowfall_rate(nray_CS, nlev_CS) ;**
Snow rate profile in mm/hr of liquid water, copied from 2C-SNOW-PROFILE.

**float snow_water_content(nray_CS, nlev_CS)**
Snow water content profile in g/m^3, copied from 2C-SNOW-PROFILE.

## Variables from CS-RAIN-PROFILE (*CloudSat rain profile)*

**float rain_rate(nray_CS)**
Surface rain rate in mm/hr, copied from 2C-RAIN-PROFILE. Negative rain rates indicate a high rain rate where the radar signal has been saturated. In this situation, the absolute value of the rain rate should be interpreted as the minimum possible rain rate. Profiles determined as snow or mixed phase report rain_rate=0.

**float rain_rate_uncertainty(nray_CS)**
1-sigma uncertainty estimate in the surface rain rate, copied from 2C-RAIN-PROFILE.

**short rain_quality_flag(nray_CS)**
Flag indicating the quality of the rain rate estimate, copied from 2C-RAIN-PROFILE. Flagging is based on the modeled multiple scattering correction, estimate of the uncertainty in the Path Integrated Attenuation (PIA) and the magnitude of the estimated PIA. Increasing values of confidence are indicative of lower uncertainty in the PIA and smaller multiple scattering effects. -1 = missing data input or land surface, 0= no confidence, 1-4= increasing levels of confidence.

**short precip_ice_water(nray_CS, nlev_CS)**
Precipitation ice water profile in g/m^3 scaled by 1000, copied from 2C-RAIN-PROFILE.

**short precip_liquid_water(nray_CS, nlev_CS)**
Precipitation liquid water profile in g/m^3 scaled by 1000, copied from 2C-RAIN-PROFILE.

**short PWC_uncertainty(nray_CS, nlev_CS)**
1-sigma uncertainty in the precipitation (liquid+ice) water content, copied from 2C-RAIN-PROFILE.

## Variables from CS-PRECIP-COLUMN (*CloudSat precipitation column*)

**float PIA_hydrometeor(nray_CS)**

Two-way PIA hydrometeor in dB, copied from 2C-PRECIP-COLUMN.

**float PIA_uncertainty(nray_CS)**
Two-way PIA uncertainty in dB, copied from 2C-PRECIP-COLUMN.

**float Precip_rate(nray_CS)**
Precipitation rate in mm/hr, copied from 2C-PRECIP-COLUMN. If the maximum retrievable precipitation rate is encountered, Precip_rate is set to a negative number. The absolute value of this number is the minimum precipitation rate for the profile; the actual precipitation rate is probably higher (and may be much higher) but cannot be resolved. When Precip_rate is negative, Retrieval_info_flag is 50.

**float Rain_top_height(nray_CS)**
Estimated max height at which liquid precip is found in the column, in km, copied from 2C-PRECIP-COLUMN.

**short Precip_flag(nray_CS)**
Precipitation status flag, copied from 2C-PRECIP-COLUMN.
0 - no precip detected
Flags for rain:
1 - possible
2 - probable
3 - certain
Flags for snow:
4 - possible
5 - certain
Flags for mixed:
6 - possible
7 - certain
9 - uncertain, see Status_

**short Conv_strat_flag(nray_CS)**
Convective stratiform status flag, copied from 2C-PRECIP-COLUMN. Classification of rain into convective, stratiform, or shallow based on rain top height.
-2 - no determination possible due to shallow profile
-1 - no determination possible due to bad input data
0 - no certain precipitation present
1 - convective precipitation
2 - stratiform precipitation
3 - shallow precipitation

**short Status_flag(nray_CS)**
Status flag, copied from 2C-PRECIP-COLUMN.
0 - both the quantitative precip rate and occurrence retrievals were successful
1 - only the precip occurrence retrieval was successful; no precip rate was retrieved
The following values indicate an error condition occurred:
12 - reflectivity profile bad
13 - gaseous attenation missing
14 - SST missing
15 - surface wind speed missing
16 - sigma_zero could not be determined
17 - cloud base could not be determined
18 - near surface reflectivity missing
19 - freezing level could not be determined

### *c) Full Swath Variables (type and dimensions)*

In addition to the DPR and GMI variables being provided along the coincidence segment, the "full swath" data are also provided for GMI and each of the three DPR radar datasets, in case the user wants to perform any additional processing, apply a more rigorous bin matching, or look at the full native data within the vicinity of the coincidence location. These data are located

within the groups S1, NS, MS and HS, respectively.

In this same example file, these four groups have dimensions, respectively:

```
nscan_GMI = 148        /* number of GMI scans in the full-swath segment */
npix_GMI = 221         /* number of GMI pixels per scan, same as 1B.GPM.GMI */
ntb_GMI = 13           /* number of GMI S1+S2 channels, same as 1B.GPM.GMI */

nscan_DPR = 142        /* number of DPR scans in the full-swath segment */
nray_DPR_NS = 49  /* number of DPR NS rays per scan, same as 2A.GPM.DPR */
nlev_DPR_NS = 176      /* number of DPR NS bins, same as 2A.GPM.DPR */

nscan_DPR = 142        /* number of DPR scans in the full-swath segment */
nray_DPR_MS = 25  /* number of DPR MS rays per scan, same as 2A.GPM.DPR */
nlev_DPR_MS = 176      /* number of DPR MS bins, same as 2A.GPM.DPR */

nscan_DPR = 142        /* number of DPR scans in the full-swath segment */
nray_DPR_HS = 24  /* number of DPR HS rays per scan, same as 2A.GPM.DPR */
nlev_DPR_HS = 88  /* number of DPR HS bins, same as 2A.GPM.DPR */
```

In the case of the GMI, the minimum and maximum values of **scan_index_S1** were determined, and 50 scans were carried on either side. In the file example above, scan_index_S1 ranged from 1288 to 1335 (48 S1 scans total covering the coincidence segment). Therefore, **nscan_GMI**= 48+50+50= 148. The value of **pix_index_S1** provide the across-track pixel index, which ranges from 0 to 220.

Similarly for DPR-NS, the minimum and maximum values of **scan_index_NS** were determined, and 60 scans carried on either side. In the file example above, **scan_index_NS** ranged from 3456 to 3477 (22 NS scans covering the coincidence segment). Therefore, **nscan_DPR**= 22+60+60= 142. In the NS group, the variable **scan_index_NS_swath** is therefore an array of length 142, and its array values range from 3456-60 (3396) to 3477+60 (3537). The values of `nray_DPR_NS` and `nlev_DPR_NS` are the same as in the original DPR datasets, with values of 49 and 176 for NS and MS, and 24 and 88 for HS. The associated variables **binRealSurface**, **binClutterFreeBottom**, **localZenithAngle**, and **localZenithAngle** are copied unchanged from the 2A.GPM.DPR file, needed to properly place the radar profile bins above the surface.

The variables that are associated with each of the S1, NS, MS and HS groups are described below.

**integer scan_index_S1 (nscan_GMI)**
Value of the GMI S1 along track scan index, from 1B.GPM.GMI, for the full swath GMI segment. Typically there are approximately 2900 GMI scans per orbit. This is included in case the user wants to perform any additional processing or bin matching.

**float Tb (nscan_GMI, npix_GMI, ntb_GMI)**
GMI full swath TB. The closest S2 channels are mapped to each S1 position. Channel order is 10V 10H 18V 18H 23V 36V 36H 89V 89H 166V 166H 183+/-3 183+/-8.

**float Latitude (nscan_GMI, npix_GMI)**
GMI S1 full swath latitude, in degrees, copied from 1B.GPM.GMI.

**float Longitude (nscan_GMI, npix_GMI)**
GMI S1 full swath longitude, in degrees between -180 and 180, copied from 1B.GPM.GMI.

**integer scan_index_NS (nscan_DPR)**
Value of the DPR NS along track index, from 2A.GPM.DPR, for the full-swath DPR segment. Typically there are approximately 7900 DPR scans per orbit. For example, if the coincidence segment starts at DPR NS ray 100 and nray_CS=500, and since each DPR ray "repeats" about four times (Figure 1), this value would range from 100 to about 224. This is included in case the user wants to perform any additional processing or bin matching.

**short zFactorMeasured (nscan_DPR, nray_DPR_NS, nlev_DPR_NS)**
zFactorMeasured for NS Ku-band full swath, in dB scaled by 100, copied from 2A.GPM.DPR.

**float Latitude (nscan_DPR, nray_DPR_NS)**
latitude for NS Ku-band full swath, in degrees, copied from 2A.GPM.DPR.

**float Longitude (nscan_DPR, nray_DPR_NS)**
longitude for NS Ku-band full swath, in degrees between -180 and 180, copied from 2A.GPM.DPR.

**short binRealSurface (nscan_DPR, nray_DPR_NS)**
binRealSurface for NS full swath, copied from 2A.GPM.DPR.

**short binClutterFreeBottom (nscan_DPR, nray_DPR_NS)**
binClutterFreeBottom for NS full swath, copied from 2A.GPM.DPR.

**float localZenithAngle (nscan_DPR, nray_DPR_NS)**
localZenithAngle for NS full swath, in degrees, copied from 2A.GPM.DPR.

**integer elevation (nscan_DPR, nray_DPR_NS)**
elevation above the reference surface for NS full swath, in meters, copied from 2A.GPM.DPR.

**short zFactorMeasured (nscan_DPR, nray_DPR_MS, nlev_DPR_MS)**
zFactorMeasured for MS Ku-band full swath, in dB scaled by 100, copied from 2A.GPM.DPR.

**float Latitude (nscan_DPR, nray_DPR_MS)**
latitude for MS Ku-band full swath, in degrees, copied from 2A.GPM.DPR.

**float Longitude (nscan_DPR, nray_DPR_MS)**
longitude for MS Ku-band full swath, in degrees between -180 and 180, copied from 2A.GPM.DPR.

**short binRealSurface (nscan_DPR, nray_DPR_MS)**
binRealSurface for MS full swath, copied from 2A.GPM.DPR.

**short binClutterFreeBottom (nscan_DPR, nray_DPR_MS)**
binClutterFreeBottom for MS full swath, copied from 2A.GPM.DPR.

**float localZenithAngle (nscan_DPR, nray_DPR_MS)**
localZenithAngle for MS full swath, in degrees, copied from 2A.GPM.DPR.

**integer elevation (nscan_DPR, nray_DPR_MS)**
elevation above the reference surface for MS full swath, in meters, copied from 2A.GPM.DPR.

**short zFactorMeasured (nscan_DPR, nray_DPR_HS, nlev_DPR_HS)**
zFactorMeasured for HS Ku-band full swath, in dB scaled by 100, copied from 2A.GPM.DPR.

**float Latitude (nscan_DPR, nray_DPR_HS)**
latitude for HS Ku-band full swath, in degrees, copied from 2A.GPM.DPR.

**float Longitude (nscan_DPR, nray_DPR_HS)**
longitude for HS Ku-band full swath, in degrees between -180 and 180, copied from 2A.GPM.DPR.

**short binRealSurface (nscan_DPR, nray_DPR_HS)**
binRealSurface for HS full swath, copied from 2A.GPM.DPR.

```
short binClutterFreeBottom (nscan_DPR, nray_DPR_HS)
binClutterFreeBottom for HS full swath, copied from 2A.GPM.DPR.

float localZenithAngle (nscan_DPR, nray_DPR_HS)
localZenithAngle for HS full swath, in degrees, copied from 2A.GPM.DPR.

integer elevation (nscan_DPR, nray_DPR_HS)
elevation above the reference surface for HS full swath, in meters, copied from 2A.GPM.DPR.
```

### *d) NOAA-19 MHS coincidence data*

For each of the CloudSat-GPM coincidences, whenever NOAA-19 also flies within 15-minutes of the GPM-CloudSat coincidence locations, the nearest (i.e., nearest to CloudSat) Microwave Humidity Sounder (MHS) pixel is located, and the 5-channel MHS Tc values from the associated 1C.NOAA19.MHS file are added into the **CS** group. To identify these cases, the files will have two extra global attributes added:

```
1C.NOAA19.MHS = "1C.NOAA19.MHS.XCAL2015-V.20140913-S145535-E163738.028846.V04A.HDF5"
CS_minus_NS_time_diff_seconds = "-233" ;
```

The first is the names of the 1C.NOAA19.MHS XCAL file used. The user can check for the existence of the CloudSat-minus-DPR time difference attribute, and if it exists, the additional **Tc_MHS** variable will be found in the CS group, as well as the associated MHS incidence angle (**ntb_MHS**=5):

```
float Tc_MHS(nray_CS, ntb_MHS) ;
Tc_MHS:units = "Kelvin" ;
Tc_MHS:long_name = "MHS Tc that is closest to this CloudSat ray. Channel order is 91 157 183+/-1
183+/-3 183+/-8" ;

float incidenceAngle_MHS(nray_CS) ;
incidenceAngle_MHS:units = "degrees" ;
incidenceAngle_MHS:long_name = "MHS incidenceAngle, copied from 1C.NOAA19.MHS, that is closest to
this CloudSat ray" ;
```

These values were added in to assist the high-frequency sounder (MHS, ATMS, SAPHIR) algorithm in interpreting the light rain and snow cases, where the precipitation structure is weak and poorly identified in the Ku/Ka-band DPR. Two added quick-look images are created, with the extensions **tb_MHS.png** and **profile_GMI_MHS.png**, shown below. The first (Figure 9) shows three channels from GMI (166H, 183.31±3 and 183.31±8), and below these the nearest corresponding channels from MHS (157, 183.31±3 and 190.31). Even though MHS is an across-track scanning sounder with coarser resolution than GMI, the correspondence between each is obvious. The dashed line that is more or less "parallel" to CloudSat is the NOAA19 subtrack.

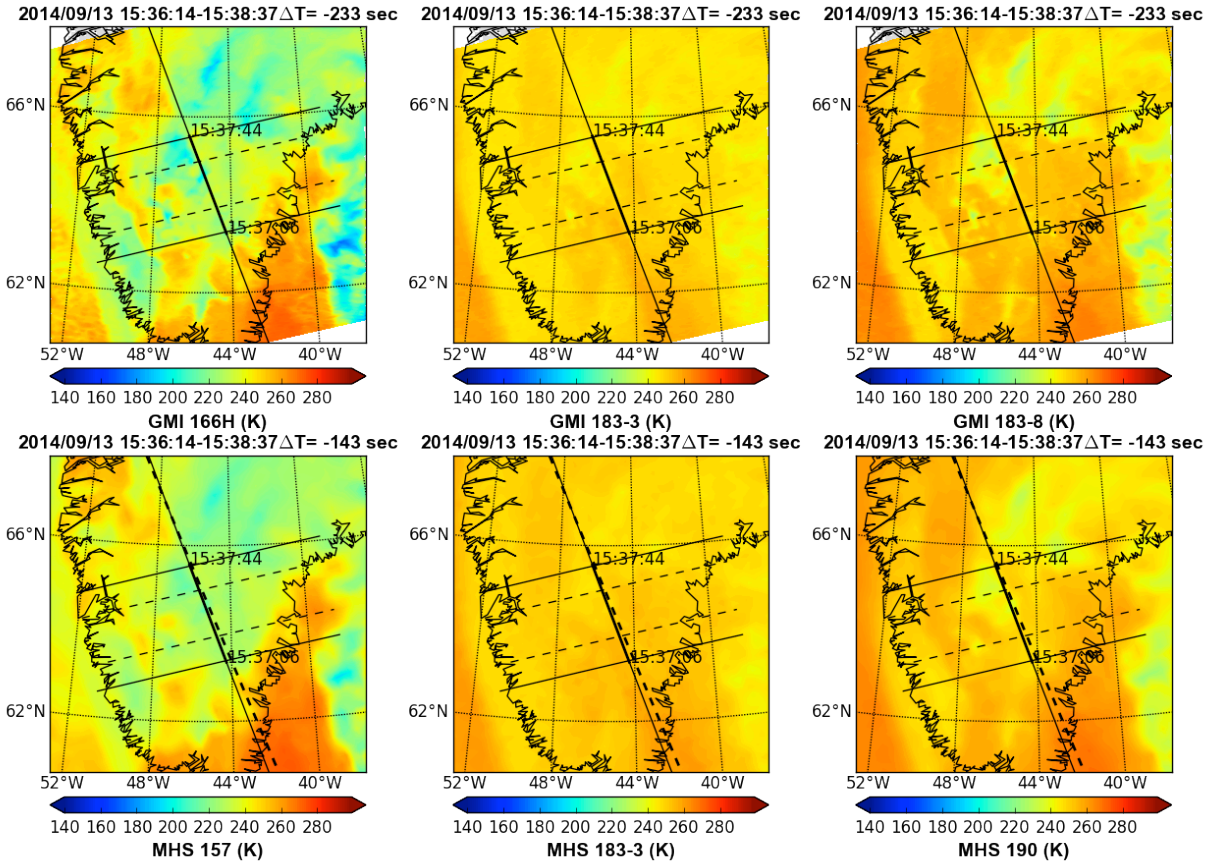


**Figure 9.** (**Top row**) Three channels from GMI (166H, 183.31±3 and 183.31±8). (**Bottom row**) The nearest pixel corresponding channels from MHS on NOAA-19 (157, 183.31±3 and 190.31).

The second quick-look image (Figure 10) is similar to Figure 7, where the trace of **Tc_MHS** (5 channels) is also shown below the GMI and the CloudSat profile. In this example shown (late summer, 13 September 2014), a snowfall event over Greenland is shown, with the high-frequency (166 GHz and above) TB signatures owing to snowfall difficult to discern from the already radiometrically cold environment. The abnormally low values of 10 GHz TB (and its polarization signature) in the high central area of Greenland could potentially be due to the presence of surface snow melting.

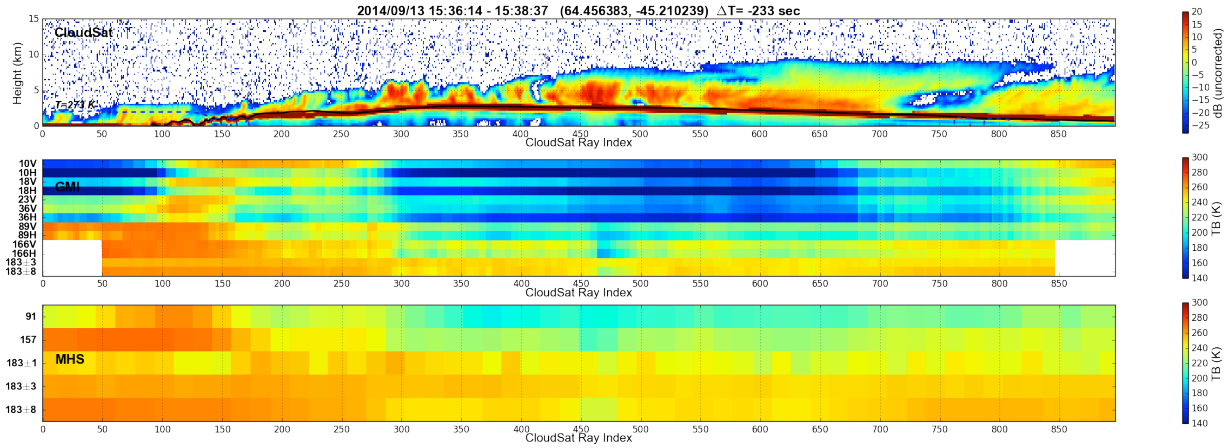


**Figure 10.** Similar to Figure 7, where the trace of **Tc_MHS** (5 channels) is also shown below the GMI and the CloudSat profile.

## 6. References

Further information on the CloudSat data products are available from the CloudSat Data Processing Center, located at the Cooperative Institute for Research in the Atmosphere (CIRA) at Colorado State University, Fort Collins, CO. The assistance by Phil Partain (CloudSat DPC) is acknowledged.

CloudSat Data Processing Center:
http://www.cloudsat.cira.colostate.edu

NASA Precipitation Processing System Science Team Online Request Module (STORM):
https://storm.pps.eosdis.nasa.gov/storm/

## 7. Change Log

**3 Feb 2017:** Version V02A (updated previous V01C)

Added a number of CloudSat and GPM related precipitation products. These are written into a new "**CS_PRECIP**" group. All other products from V01C are maintained.

**25 Jul 2016:** Version V01C (updated previous V01B)

a) Updated processing interval (GPM launch through 6 June 2016). Even though the date of the latest coincidence is 6 June 2016, there are days with missing CloudSat orbits prior to this, that will eventually get reprocessed by the CloudSat DPC and included in the next update.

b) Use of the latest Version 04A GPM datasets for 2A.GPM.DPR and 1B.GPM.GMI.

c) Addition of 5-channel 1C.NOAA19.MHS Tc (X-Cal intercalibrated TB).