GPM/DPR Level-2
Algorithm Theoretical Basis Document

Authors: Toshio Iguchi, Shinta Seto, Robert Meneghini,
Naofumi Yoshida, Jun Awaka, Minda Le, V. Chandrasekar,
Stacy Brodzik, Simone Tanelli, Kaya Kanemaru, Takeshi Masaki,
Takuji Kubota and Nobuhiro Takahashi

Revised July 2024

Revision history

Date	Note
December 2010	New
Revised October 2014	Overall update
Revised January 2015	Overall update
Revised April 2015	Overall update
Revised August 2015	Overall update
Revised February 2016	Overall update
Revised March 2016	Overall update
Revised April 2017	Overall update for V05
Revised October 2018	Overall update for V06
Revised June 2020	Overall update for V06X
Revised December 17, 2021	Overall update for V07
Revised July 2023	Overall update
Revised March 2024	Update Chapter 0 , Chapter 2 and Appendix-1 for V07C
Revised July 1, 2024	Update for GPM orbit boost

Contents

0 INTRODUCTION AND CAVEATS

This document describes the basic idea of DPR data processing. It was originally written for the algorithm used in the at-launch version (V03). The algorithm has been modified and improved since then. Although the basic idea of data processing remains the same, the actual flow of processing, in particular that in the solver module, has changed substantially. As a result, some part of description in Section 3.1 may not be relevant any more.

0.1 Minor changes in the DPR Level-2 products from Version 07B to Version 07C

The GPM Core Observatory satellite performed orbit boost maneuvers on Nov. 7 and 8, 2023 that raised its altitude from 400km to 435km. Details show at the following link.

`https://gpm.nasa.gov/missions/gpm/orbit-boost`

The orbits in which this operation was performed are shown operation status in the following link.

`https://gportal.jaxa.jp/gpr/information/product`

In Version 07C, the sensitivity, the spatial resolution and the swath width in the DPR instrument were changed. See full descriptions by the following link.

`https://www.eorc.jaxa.jp/GPM/en/boost.html`

The updates of the DPR Level 2 algorithm are shown in detail in the following link.

`https://www.eorc.jaxa.jp/GPM/doc/algorithm/GPM_Orbit-boost_DPR-note_20240304A.pdf`

0.2 Major changes in the DPR Level-2 from Version 07A to Version 07B

In Version 07B, a new NOAA/NESDIS autosnow was introduced as the input to the DPR Level-2 algorithm. A comparison of precipitation estimates between old and new versions of the autosnow shows that there were differences in specific cases, but few in number. The monthly surface precipitation rate over areas where the snow cover indexes were changed in V07B was very close to that of V07A with the difference less than 1 %.

0.3 Major changes in the DPR Level-2 from Version 06A to Version 07A

Version 07A is the first standard product to account for the Ka-band Precipitation Radar (KaPR) scan pattern change implemented on May 21, 2018. This change in scan pattern allows for a more accurate precipitation estimation method for dual frequency radar, Ku-band and Ka-band Precipitation radar (KuPR and KaPR), to be applied to the entire observation swath. On the

other hand, this led to significant changes of DPR file specifications for common file structures before and after the scan change, in addition to algorithm evolutions. See full descriptions at the following websites:

`https://arthurhou.pps.eosdis.nasa.gov/atbd.html`

`https://www.eorc.jaxa.jp/GPM/en/archives.html`

1. Format change in the product
   In the V06X (experimental product), a new format was implemented including "`FS`" which is defined as the full swath dual-frequency product with 125 m range resolution. In the V07, this FS format is applied to data taken both before and after the scan pattern change of the KaPR in May 2018 (Figure 0.3.1). Note that the DPR/FS has KuPR single frequency data in the outer swath before the scan pattern change, whereas it has DPR dual frequency data in the outer swath after the scan pattern change.

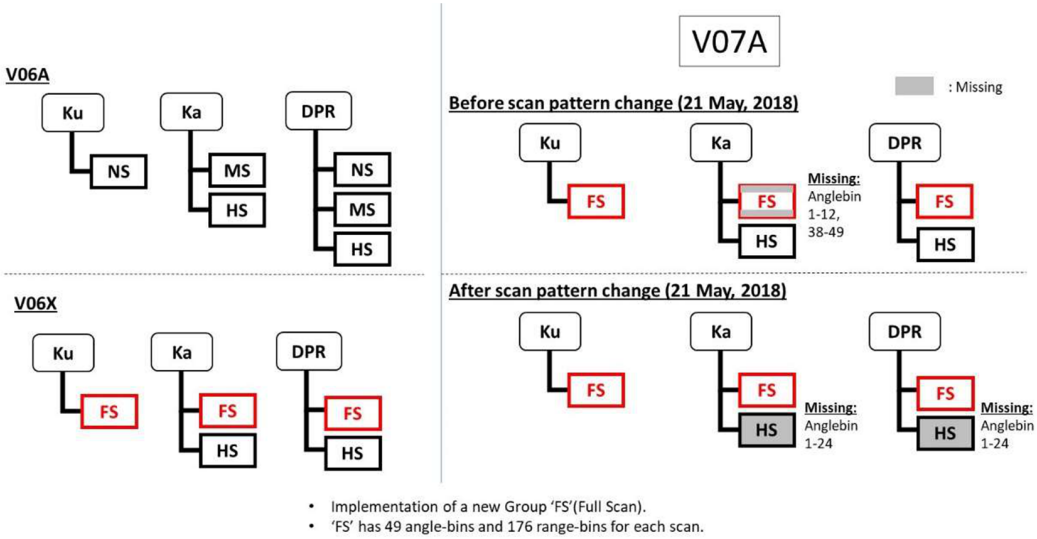


Figure 0.3.1: Changes of File structure from V06 to V07

2. In version 07, the sidelobe clutter removal routine of the single frequency (KuPR, KaPR, and PR) L2 algorithms have been improved based on the results of Kanemaru et al. (2020, 2021). In addition, a new 3-D precipitation determination method has been implemented to improve the detectability of precipitation signals. This method uses not only vertical signals, but also cross-track and along-track signals. It is expected that this method will make it easier to detect precipitation signals distributed horizontally at low altitudes that often occurs at high latitudes. For details, see Section 3.3. In addition, clutter removal in the mainlobe clutter region has been experimentally processed in version 07. Therefore, the Z_m profile in the mainlobe region included in the Level 2 product (and the received echo

signal power included in the Level 2 ancillary product) contains not only the information of the precipitation signal, but also the effect of possible clutter removal errors. Please pay attention to the validity of the data in the mainlobe region when using it.

In version 07, the measurement error of the data is taken into account in the estimation of the received echo signal power. This changes the magnitude of the signal echo power and Z_m near the noise level. See Section 3.3.

3. "`flagHail`" is newly implemented in the CSF. (Dual frequency only.)
4. The following new items are added to present information about `flagHeavyIcePrecip` in the CSF: `binHeavyIcePrecipTop`, `binHeavyIcePrecipBottom`, and `nHeavyIcePrecip`.
5. Variables of the Trigger module (TRG) are newly implemented in V07.
6. Changes in the DSD/Solver module from V06 to V07 are listed as follows.

   A) Revision of the relationship between precipitation rate R and volume-weighted mean drop size D_m (the so-called R-D_m relationship). In addition, the parameter ε in the R-D_m relationship, which is independent of range in V06, is allowed to vary with range in V07.
   B) Revision of DSD database (used for single-frequency).
   C) Soil moisture effect.

   Note that the soil moisture effect is expected to increase the precipitation amount over land as estimated by the single frequency algorithms by about 14-23%.
7. "`precipRateEsurface2`" is a surface precipitation estimate based on the a priori low-level precipitation profiles (Hirose et al., 2021, `https://doi.org/10.2151/jmsj.2021-060`). This estimate is stored as an Experimental variable.
8. Other changes in DPR L2 algorithm:

   · "`binMirrorImageL2`" is newly implemented in the Preparation module (PRE) to notify a false precipitation echo that appears at high altitude due to the mirror image.
   · Adjustment factor for KaPR's σ^0 and Z-factor is updated in the Preparation module (PRE).
   · The VER module includes new variables such as the profile of air temperature, a flag that indicates an inversion layer of air temperature, and the rain-free path-integrated attenuation estimate `piaNP`.
   · Several variable names are changed.
     - `zFactorCorrected` → `zFactorFinal`
     - `zFactorCorrectedESurface` → `zFactorFinalESurface`
     - `zFactorCorrectedNearSurface` → `zFactorFinalNearSurface`

· Bugs are fixed in several modules. Details of the description for variables are described in Appendix or in the following websites:

`https://www.eorc.jaxa.jp/GPM/en/archives.html`

`https://arthurhou.pps.eosdis.nasa.gov/GPMprelimdocs.html`

0.4 CAVEATS for DPR Level-2 V07A

0.4.1 PRE module

In the DPR/FS and KaPR/FS, the following variables have unexpected values only in angle bins 1-12 of the first scan of the granule without any quality indications by the quality flags. This issue occurred only in orbits after the scan pattern change.

- `attenuationNP`, `piaNP`, `piaNPrainFree`, `sigmaZeroNPCorrected`, and `sigmaZeroCorrected`.

0.4.2 CSF module

1. There are bugs on Heavy Ice Precipitation (HIP). The bugs appear only in the data before the antenna scan pattern change (viz., before May 21, 2018). Since the occurrence of HIP is not frequent, the effect of the bugs on the rainfall rate statistics would be small.

   (a) `flagHeavyIcePrecip` in DPR FS does not reflect the Ka-band detection of HIP and the flag value is different from what is designed. The result is based on HIP detected by Ku-band and/or DFR_m methods only. Nevertheless, a positive flag value indicates that HIP is detected.
   On the other hand, the additional HIP flags

   `binHeavyIcePrecipTop`

   `binHeavyIcePrecipBottom`

   `nHeavyIcePrecip`

   in the DPR FS structure contain correct values in the inner swath. These flags consist of two dimensional arrays in the DPR FS structure and the second array index distinguishes values obtained by the Ku, Ka, and DFR_m methods.
   In the outer swath, however, `binHeavyIcePrecipTop` and `binHeavyIcePrecipBottom` for Ka-band should contain Missing values, but the bugs make them contain values copied from Ku-band.

   As for the DPR HIP data before May 21, 2018, our suggestion is that users should examine whether `flagHeavyIcePrecip` > 0 or not. Examination of `nHeavyIcePrecip` would be a good idea.

(b) In the HS mode data, the additional HIP flags, i.e.,

`binHeavyIcePrecipTop`, `binHeavyIcePrecipBottom`, and `nHeavyIcePrecip`

do not contain expected values. The values of those for the Ka-band single frequency HS data are always zero, and those items are even non-existent in the dual frequency DPR HS data structure, the latter of which might be a toolkit bug.

2. A bug on `binBBTop`: In the V07 DPR FS processing, `binBBTop` - `binBBPeak` in the outer swath is bounded by the same parameter value as that in the inner swath. Since the apparent width of BB increases in the outer swath because of the smearing of BB peak, a different parameter value should be used. This bug will be fixed in the next V08. Incidentally, there is not such a bug in the Ku-band single frequency FS processing. The bug was introduced when transplanting the Ku-band FS processing code to the DPR FS processing code.
The effect of the bug on the rainfall rate statistics is estimated to be very small.

0.5 Major changes in the DPR Level-3 from Version 06A to Version 07A

1. Format changes in Level-3 DPR Daily/Monthly Products
For a full description of these changes see the L3 DPR V07 ATBD document. To accommodate the change in the FS/HS/MS swath structures in the L2 products, the L3 grids have been restructured as below.

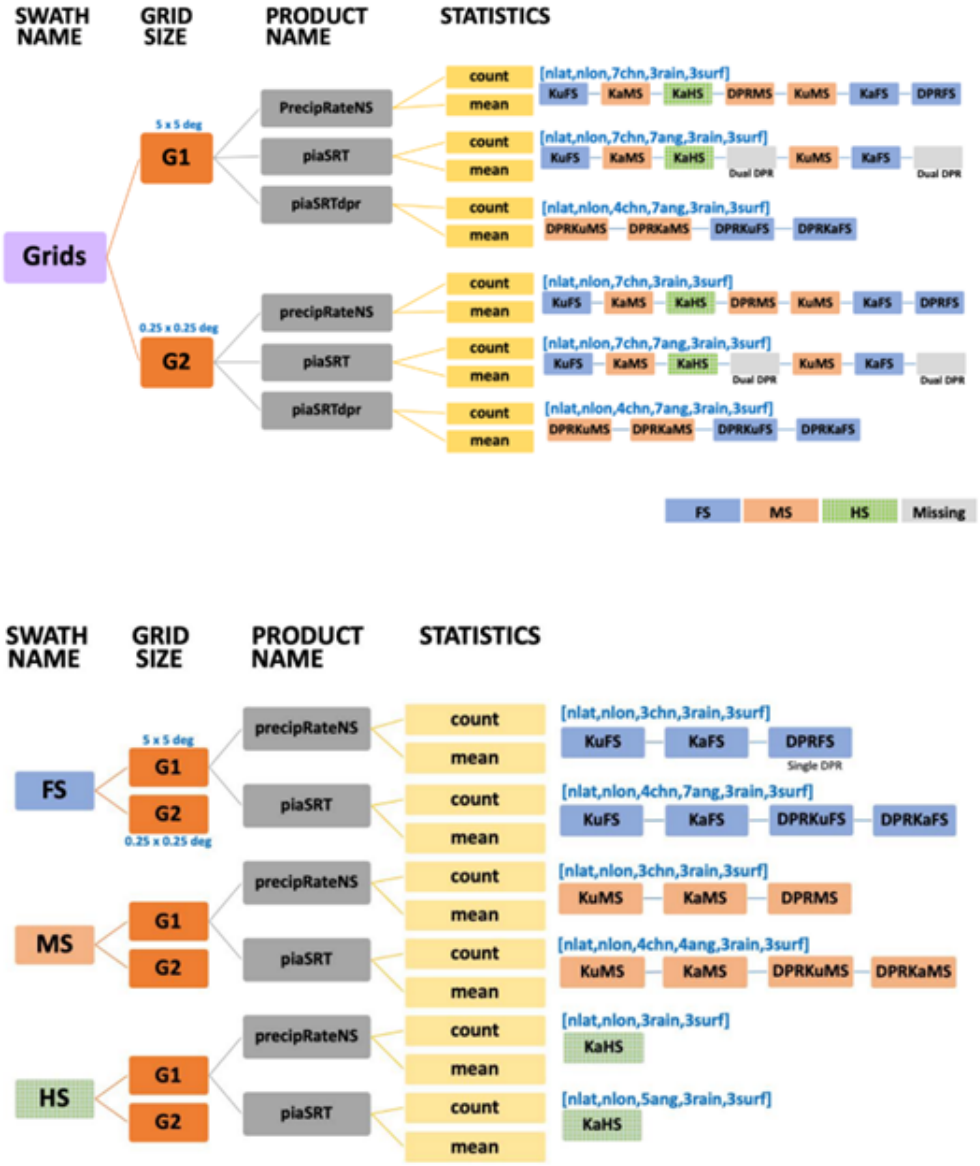


Figure 0.5.1: The organization for Level-3 in V06X (upper) and V07A (lower)

0.6 CAVEATS for DPR Level-3 V07A

The Version 7 L3 daily products 3DPR-ASC and 3DPR-DES do not accumulate FS statistics (full 49 ifov data) from 5/21/2018 to 5/28/2018 even though the Ka scanning pattern change occurred on 5/21/2018. This will be corrected in a later product version. Users should note that monthly statistics of the Ka and DPR FS/HS data do not cover the entire month of May, 2018.

0.7 CAVEATS for DPR Level-3 V06A

1. The DPR Daily Product (3DPRD) of Version 06A includes a minor bug. In the array of KuPR channel for ascending path, `stratPrecipRateNearSurfMean` and `stratPrecipRateESurfMean` stores convective precipitation (they should have stored stratiform precipitation). This issue is fixed in Version 07A.

0.8 Major changes in the DPR Level-2 from Version 06A to Version 06X

DPR L2/L3 V06X products are regarded as "experimental products", not "standard products". The V06X provides early outcomes of the full-swath coverage after the DPR scan pattern change in 21 May 2018 aiming to apply dual-frequency observation for full-swath of KuPR observation by assigning KaHS rays to outer swath. Basic concept of V06X algorithm is the same as V06A, but there are several new features of the algorithm for the V06X due to progress in algorithm development. Preliminary evaluations show that estimated precipitation in the outer swath from the dual-frequency method tends to be underestimated in V06X. Users need to be cautious about this. Mitigating this underestimation is regarded as one of the future tasks in V07. See full descriptions by the following website:

`https://www.eorc.jaxa.jp/GPM/en/archives_v6X.html`

1. Implementation of a new format. The latest TKIO supports the new format including "FS" that is full swath dual-frequency product with 125 m range resolution.
2. Algorithm updates.

0.8.1 PRE module

· A sidelobe clutter reduction method for KaHS is improved after the scan pattern change was installed in the V06X algorithm because sidelobe clutter contaminations of KaHS became more problematic in the new Ka scan pattern.

· A sidelobe clutter filter technique for the KaHS is applied also for the Ku.

· Clutter free detection in the Ku is improved in cases when the brightband (BB) is found near the surface.

0.8.2 CSF module

· The dual frequency technique in the classification module was improved to be applied in the FS.

· Some parameters in the KaHS were changed in order to accomodate with the dual frequency technique.

· New variables related to solid precipitation is implemented.

· Reclassification by the slope method is improved.

0.8.3 SRT module

· The latest temporal reference files were applied.

· SRT codes have been updated so that they now read temporal data over the full swath.

· Dual-frequency SRT and Hybrid estimates are now applicable to the full swath.

0.8.4 SLV module

· V06X uses the same dual-frequency precipitation estimation algorithm as V06A in the Solver (SLV) module.

0.9 Major changes in the DPR Level-2 from Version 05B to Version 06A

1. A new SRT code has been modified to include calculations of Hitschfeld-Bordan PIA as well as a hybrid PIA that combines HB and SRT results. As a result of applying the new PIA in SLV module, erroneously large estimates of high precipitation over ocean (near coast) are mitigated and both DPR(MS) and Ku rain estimates in V06A agree better with Ground Validation data over USA.
2. A new classification algorithm is introduced by the University of Washington (Stacy Brodzik and Robert Houze) to reclassify the stratiform rain type. The new algorithm improved an angle-bin dependence of rain classification and SLH profiles.

Minor changes in DPR L2 algorithm:

· Mitigated KaHS's sidelobe clutter and re-calculated the data base for KuPR's sidelobe clutter.

· New variables are introduced. They are `flagScanPatten` in PRE module, `PIAhb`, `PIAhybrid`, `reliabFactorHY`, `reliabFlagHY`, `stddevEff`, `stddevHY` and `zeta` in SRT module. For definitions of these variables, the reader is referred to the user manual

· Modification of surface snow index for the winter temperature inversions.

· Applied the latest SRT data base.

· Bug fix of flagEcho in dual frequency data processing.

0.10 Major changes in the DPR Level-3 products from Version 05B to Version 06A

1. New variables are introduced. They are `zFactorMeasuredNearSurface`, `DFRNearSurface`, `DFRmNearSurface`, `piaHybrid`, `piaHybridDPR`, `piaHB`, `zeta`, and `flagHeavyIcePrecip`. The definitions of these variables can be found in the user manual.

0.11 Major changes in the DPR Level-2 products from Version 05A to Version 05B

Figure 0.11.1 shows DPR's scan patterns before May 21 2018 (left) and after May 21 2018 (right). KaHS beams scan in the inner swath before May 21 2018, but now they scan in the outer swath and match with KuPR's beams. Numbers in color indicate angle bin numbers for KuPR (blue), KaMS (yellow), and KaHS (red).

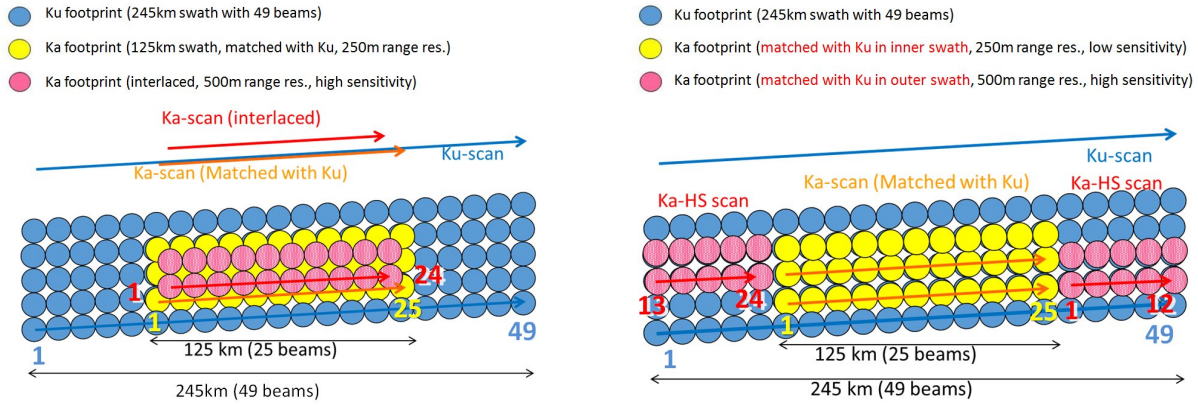


Figure 0.11.1: DPR's scan patterns before and after May 21 2018.

1. As of May 21, 2018, the scan pattern of KaHS beams was changed as shown in Figure 0.11.1. The KaHS beams scanned in the inner swath before May 21 2018, but now they scan in the outer swath and match with KuPR's beams. Please note that the range resolution of KaHS is 500m and differs from that of the KuPR or KaMS. One scan of KuPR over the full swath consists of 49 beams which are numbered from 1 to 49 in Figure 0.3.1 below. The array for one scan of KaMS data consists of 25 beams which correspond to KuPR's central beams from 13 to 37, whereas one scan of KaHS data array for 24 beams consists of two parts. The first 12 elements in the array correspond to KuPR's beams from 38 to 49, and the last 12 elements correspond to KuPR's beams from 1 to 12 of the following scan. All KaMS beams and the first 12 KaHS beams match well with KuPR's beams from 13 to 49. The magnitude of misalignment between these KuPR and KaPR beams is estimated to be less than 50 m after May 21. However, KaHS's beams from 13 to 24 are slightly shifted from the corresponding KuPR beams from 1 to 12 in the along-track direction. The magnitude of misalignment in these beams is about 400 m.

   The dual-frequency algorithm can be applied to the full swath of data after the change of scan pattern. However, because it takes time to develop a new algorithm applicable to the full swath and to adjust the necessary parameters to ensure the quality of the products, KaHS data are not processed in V05B L2 algorithm. Missing values are stored in the

output variables that use KaHS data. KaHS's received power data are available only in L1 products in V05B.

2. All beam directions of KuPR and KaPR were adjusted to match better with the nominal footprint locations. The difference between KuPR and KaPR's footprint centers is now about 30 m at nadir. It was about 300 m before May 21, 2018. This improvement of beam matching is not considered to make a big difference in the L2 output products except for very heavy rain cases because the mismatch was small (300 m) from the beginning.

0.12 Major changes in the DPR Level-2 products from Version 04A to Version 05A

All users of DPR L2 data should keep in mind the following changes in V5 products.

This document describes only the major changes in the Level-2 products. There are some changes in Level-1 that affect the Level-2 products. Please refer to "Release Notes for the DPR Level-1 products" for the details of the changes in Level-1 products.

Among several changes in Level-1 products, the important changes that affect Level-2 products substantially are the following points. The DPR's system parameters were re-examined. Based on the new calibration results, the offset parameters for the transmitting powers, receiver's gains, the beam widths, and the pulse width of both KuPR and KaPR are redefined. As a result, Z_m of KuPR has increased by about +1.3 dB, and Z_m of KaPR by about +1.2 dB. The radar surface cross section (σ^0) of KuPR has increased by about +1.2 dB and that of KaPR by about +1.1 dB, although the changes in σ^0 depend slightly on the incidence angle due to the changes in the beam widths. Because of the introduction of the adjustment factors in L2 (see below) whose magnitudes vary with time, the numbers mentioned above are not fixed numbers but change with time, especially near the beginning of the GPM mission. In fact, for example, statistics show that an average (over angle) increase in σ^0 at Ku-band is about +1.0 dB, that at KaMS is about +0.6 dB and that at KaHS is about +0.9 dB on the first 5 days in June 2014.

· The FCIF-LUT for the KuPR near the saturation was improved so that the effect of saturation near the saturation level was mitigated.

Changes in Level-2 algorithm:

· In addition to the changes in the DPR L1 calibration, adjustment factors are introduced to remove small trends in the overall system gains in KuPR and KaPR. The adjustment factors change the measured received powers only by a small fraction of dB.

· Since the FCIF-LUT for the KuPR near the saturation level was modified to mitigate the effect of saturation, the statistics of σ^0 near the saturation level in the KuPR has changed. This change affects the SRT performance as well. The side-lobe echo cancellation parameters are adjusted to cope with this change as well.

- Measured radar reflectivity factor Z_m and surface radar cross section σ^0 are calculated based on the new values of the pulse widths of both KuPR and KaPR. Accordingly, the angle bin dependence of σ^0 has changed slightly.
- A DSD database that depends on the month, region, surface type and rain type was created from the statistics of DSD parameters estimated with the dual-frequency algorithm. The R-D_m relationship in the DSD database is used as the default R-D_m relationship in the single frequency (Ku-only and Ka-only) data processing before it is modified by other constraints such as the path-integrated attenuation. The introduction of the DSD database has modified the precipitation estimates substantially when they are light (less than about 3mm/h) in many cases. Rain estimates from the Ku-only and dual-frequency algorithms now agree very well.
- New flags are introduced:
  `snowIceCover` in the preparation module,
  `flagHeavyIcePrecip` and `flagAnvil` in the classification module,
  `flagSurfaceSnowfall` and `surfaceSnowfallIndex` in the experimental module.
  For definitions of these flags, the reader is referred to the user manual.
- Winter convective storms that give large DFR_m (measured Dual-Frequency Ratio) at the storm top are flagged and the corresponding pixels are classified as convective in V5. This category only appears in the inner swath since DFR_m is available only there.

1 OBJECTIVES

The objective of the level 2 DPR algorithms is to generate from the level 1 DPR products radar-only derived meteorological quantities on an instantaneous FOV (field of view) basis. A subset of the results will be used by the level 2 combined radar-radiometer algorithm and the level 3 combined and radar-only products.

The general idea behind the algorithms is to determine general characteristics of the precipitation, correct for attenuation and estimate profiles of the precipitation water content, rainfall rate and, when dual-wavelength data are available, information on the particle size distributions in rain and snow. It is particularly important that dual-wavelength data will provide better estimates of rainfall and snowfall rates than the TRMM PR data by using the particle size information and the capability of estimating, even in convective storms, the height at which the precipitation transitions from solid to liquid.

2 BACKGROUND INFORMATION

2.1 Historical Perspective

The Dual-Frequency Precipitation Radar (DPR) on the GPM core satellite will be the second space-borne precipitation radar, following the first such radar, the Precipitation Radar (PR), launched on the TRMM satellite in November, 1997. The TRMM PR has already revolutionized the measurement of precipitation from space by providing high resolution 3-dimensional rain echoes in the tropics and subtropics. The DPR consists of Ku-band (13.6GHz) and Ka-band (35.5GHz) channels. A major source of error in the rainfall estimates from the TRMM/PR comes from the uncertainty in the conversion of radar reflectivity into rainfall rate. This uncertainty originates in the variations of the raindrop size distribution (DSD) that changes by region, season and rain type. One of the reasons for adding the Ka-band frequency channel to the DPR is to provide information on the DSD that can be obtained from non-Rayleigh scattering effects at the higher frequency.

Another reason for the new Ka-band channel is to provide more accurate estimates of the phase-transition height in precipitating systems. This information is very important not only in increasing the accuracy of rain rate estimation by the DPR itself, but in improving rain estimation by passive microwave radiometers.

The third reason for the Ka-band channel arises from the fact that the GPM core satellite will provide coverage up to about 65 degrees latitude; by increasing the sensitivity of this channel, a larger fraction of snow events will be detected.

Since the Ku-band channel of the DPR is very similar to the TRMM PR, the principal challenge in the development of the DPR level 2 algorithms is to combine the new Ka-band data with the Ku-band data to achieve the objectives mentioned above.

2.2 Instrument Characteristics

The DPR consists of the Ku-band precipitation radar and the Ka-band precipitation radar. They are abbreviated as KuPR and KaPR, respectively. These Earth-pointing KuPR and KaPR instruments will provide rain sensing over both land and ocean, both day and night. The KuPR and KaPR design specifications, with all active phased array elements functioning, are shown in Table 2.2.1. The spacecraft orbital information is shown in Table 2.2.2.

Figure 2.2.1 shows the DPR scan pattern. KuPR's scan pattern is similar to that of the TRMM PR. It has 49 footprints in a scan and the footprint size is about 5 km in diameter. The scan swath is 245 km. The KaPR also has 49 footprints, but these are divided into two types of scan. In the first type of scan (Ka_MS), the beams are matched to the central 25 beams of KuPR, providing a swath of 120 km. In the second type of scan (Ka_HS), the KaPR is operated in the

high-sensitivity mode to detect light rain and snow. In this case, its beams are interlaced within the scan pattern of the matched beams as shown in Figure 2.2.1. The KuPR and KaPR for the Ka_MS scan have the same range resolution (250 m), while the range resolution of data in Ka_HS is 500m. In both cases, radar echoes are over-sampled at twice the rate of the corresponding resolution: 125 m for the matched beams and 250 m for the Ka_HS.

Table 2.2.1: DPR Design Specification

Item	KuPR	KaPR
Swath Width	245 km	120 km
Range Resolution	250 m	250/500 m
Spatial Resolution	5.2 km (Nadir at the height of 407 km)	
Beam Width	0.71° (Center Beam)	
Transmitter	128 Solid State Amplifiers	
Peak Transmit Power[1]	1012.0 W	146.5 W
Pulse Repetition Freq.[2]	4000 to 4500 Hz	
Pulse Width	two 1.6 μs pulses	two 1.6 μs pulses (matched beams) two 3.2 μs pulses (interlaced scans)
Beam Number	49	49 (25 in matched beams and 24 in interlaced scans)
Min. measurable rain rate	0.5 mm/h	0.2 mm/h
Beam matching error	Less than 1000 m	
Observable range	19 km to Surface (to −5 km near nadir)	
Dynamic range	From −5 dB below the system noise level to +5 dB above the nominal maximum surface echo level	
Receiver power accuracy	±1 dB	
Scan Angle	±17° Cross Track	±8.5° Cross Track
Frequencies	13.597 and 13.603 GHz	35.547 and 35.553 GHz
Bandwidth	14 MHz	
Max. Mass	472 kg	336 kg
Power (max)	446 W (orbit average)	344 W (orbit average)
Science Data Rate (max)	109 kbps (The Total of KuPR and KaPR is 190 kbps)	81 kbps
Housekeeping Data Rate[3]	1 kbps (nominal)	

1. This parameter is for informational purposes in the ICD.
2. In nominal operation mode.
3. 1 kbps may increase up to 2 kbps during SCDP switch-overs.

Table 2.2.2: Spacecraft Orbital Information

Inclination	65°
Mean semi-major axis	6776.14 km
S/C Altitude Control Box	±1 km
Orbit Eccentricity	0.00010 (0-0.0005 tolerance)
Geodetic Altitude Variation Range	397 km to 419 km

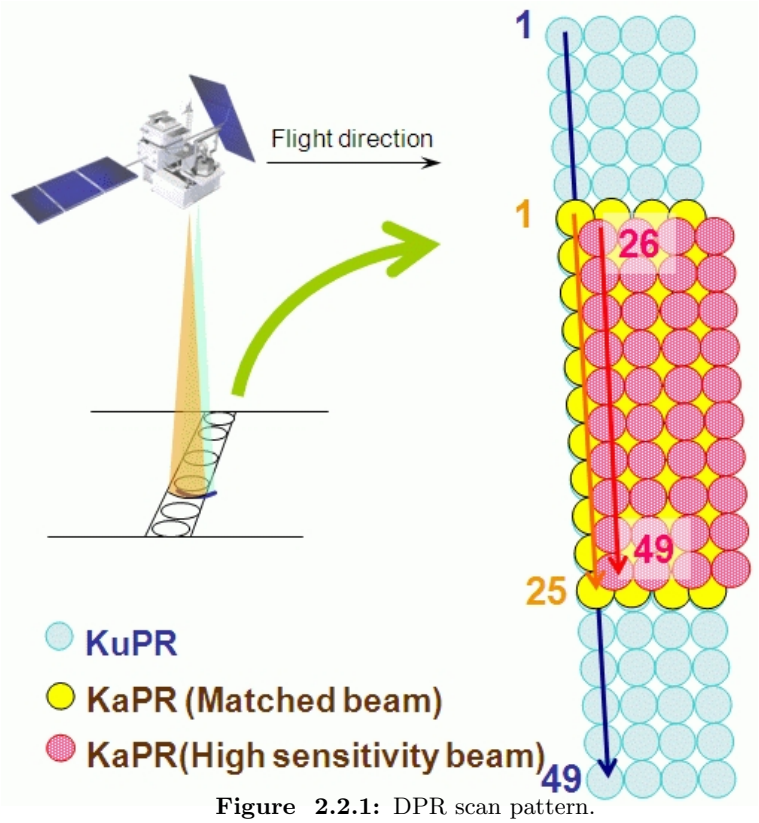


Figure 2.2.1: DPR scan pattern.

Figure 2.2.2 shows the observation range. The DPR's echo sampling is designed to cover a range that, at minimum, extends from the surface to 19 km above the sea level (or from the Ellipsoid). The pulse repetition interval is adjusted according to the satellite altitude and the angle of observation. As a result, the number of independent samples changes slightly as a function of the scan angle.

The required sensitivities of the DPR at design were 0.5 mm/h (18 dBZ) for KuPR and 0.2 mm/h (12 dBZ) for KaPR (Kojima et al.,2012, Hou et al., 2014). Masaki et al. (2020) confirmed that the sensitivities of both KuPR and KaPR were satisfied after applying the new calibration factors. In fact, using the parameters that were specified in the design requirements, the minimum detectable precipitation rate ($R_{\min}$) and radar reflectivity ($Z_{\min}$) are 0.21 mm/h and 12.17 dBZ for KuPR, 0.34 mm/h and 15.61 dBZ for Ka_MS, and 0.16 mm/h and 10.40 dBZ for Ka_HS, respectively. In the DPR design, an echo signal that is 2σ higher than the noise equivalent radar reflectivity (Z_t) is defined as precipitation in logarithmic scale, and the Z-R relationship: $Z = 200R^{1.6}$ is used to estimate the precipitation rate. Here, σ is the standard deviation of noise fluctuation.

On the other hand, in the DPR level-2 algorithm, the threshold for identification of precipitation and the assumed Z-R relationship differ from the parameters used in the DPR's design.

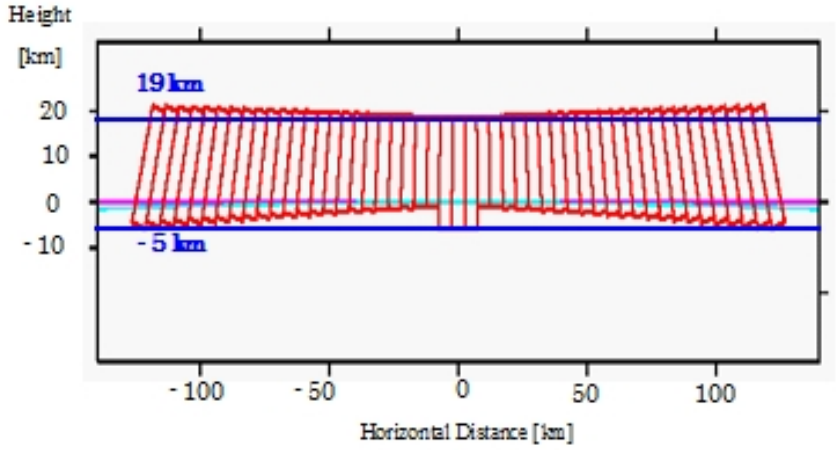


Figure 2.2.2: DPR's data sampling range

In the DPR level-2 V6 algorithm, only echo signals that exceed dBZt by 2.5σ dB are considered precipitation echoes. This more stringent criterion is used to mitigate the misidentification of noise signals as precipitation. With this threshold, KuPR's $Z_{\min}$ is 15.46 dBZ, KaMS's $Z_{\min}$ is 19.18 dBZ, and KaHS's $Z_{\min}$ is 13.71 dBZ. The actual conversion from attenuation corrected Z_e to R depends on the epsilon parameter. If the nominal Z-R relationship of $Z = 298.84R^{1.38}$ for stratiform precipitation is used to calculate the corresponding R, KuPR's $R_{\min}$ is 0.21 mm/h, KaMS's $R_{\min}$ is 0.39 mm/h, and KaHS's $R_{\min}$ is 0.16 mm/h. See section VI-D of Masaki et al. (2020) for details.

The identification of precipitation is made on a ray-by-ray basis in the DPR level-2 algorithm. Regardless of the echo intensity, as long as it exceeds the noise level, all echoes in the ray are converted to radar reflectivity. Therefore, even if the intensity of echo signal is below the threshold, it will appear in the output (e.g., "`zFactorMeasured`"). To know if the echo signal is really above the threshold, users can refer to the "`flagEcho`" in the output.

2.3 GPM Orbit Boost

A controlled re-entry is mandatory for the GPM Core satellite at the end of mission. Because solar activity in recent years turned out to be much higher than forecasted before the launch of the GPM satellite, the fuel consumption to keep the satellite at the original orbit has increased more than expected. In order to decrease the fuel consumption and to continue measurements as long as possible without substantially degrading the data quality, NASA and JAXA increased the orbit height of the satellite from nominal 407 km to 442 km in November 2023. Based on the amount of fuel remaining, the satellite is expected to be in the new orbit and collect data until around 2030.

Table 2.3.1 shows the spatial resolutions and swath width for the pre- and the post-boost

periods. The spatial resolutions at the nadir and swath width are changed from 5.04 km × 5.04 km and 255.8 km at satellite altitude of 407 km to 5.48 km × 5.48 km and 277.9 km at satellite altitude of 442 km, respectively. The spatial resolutions were analyzed at the nadir and scan edges. The minimum detectable rain rate of the DPR is changed by the orbit boost, as shown in Table 2.3.2. The detail of the boost and its influence are discussed in Kubota et al. (2024).

Table 2.3.1: Spatial resolutions and swath width for the pre-boost (altitude of 407 km) and the post-boost (altitude of 442 km) periods. Spatial resolutions were analyzed at the nadir (angle-bin No. 25) and scan edges (angle-bin No. 1, 49).

Satellite altitude	Spatial resolution		Swath width
	Nadir	Scan edge	
407 km	5.04 km × 5.04 km	5.04 km × 5.57 km	255.8 km
442 km	5.48 km × 5.48 km	5.48 km × 6.05 km	277.9 km

Table 2.3.2: Sensitivity of the DPR for the pre-boost (altitude of 407 km) and post-boost (altitude of 442 km) periods.

Satellite altitude		Minimum detectable rain rate		
		KuPR	KaPR(High sensitivity beam)	KaPR(Matched beam)
407 km	Required value	0.5 mm/h	0.2 mm/h	(blank)
	Actual value	0.30 mm/h	0.16 mm/h	0.38 mm/h
442 km (After 8 Nov. 2023)	Required value	0.5 mm/h	0.2 mm/h	(blank)
	Actual value	0.33 mm/h	0.19 mm/h	0.42 mm/h

3 ALGORITHM DESCRIPTION

3.1 Theoretical Description

3.1.1 Physical Basis of the Algorithm

The radar transmits a pulse of radio waves and receives an echo from an object. The object in the case of precipitation radar is a distribution of rain drops in the volume defined by the antenna directivity, the pulse width and the time between the transmission and the reception of the pulse. The received power P_r from rain at range r is proportional to the apparent radar reflectivity factor $Z_{m0}(r)$.

$$P_r(r) = \frac{C|K|^2}{r^2} Z_{m0}(r) \tag{3.1-1}$$

where C is the radar constant, and K is a constant defined as a function of complex refractive index m of scattering particles by the following equation.

$$K = \frac{m^2 - 1}{m^2 + 2} \tag{3.1-2}$$

When the radar electric specifications and range r are given, Z_{m0} can be calculated from P_r. Therefore, we can assume that $Z_{m0}(r)$ can be derived from a measurement of the echo power. Z_{m0} is related to the effective radar reflectivity factor Z_e by

$$Z_{m0}(r) = A(r) Z_e(r) \tag{3.1-3}$$

where A is the attenuation factor. The effective radar reflectivity factor Z_e can be expressed in terms of the backscattering cross section $\sigma_b(D)$ of a precipitation particle of diameter D and the particle size distribution $N(D)$:

$$Z_e = \frac{\lambda^4}{\pi^5 |K|^2} \int \sigma_b(D) N(D)\, dD \tag{3.1-4}$$

Here λ is the wavelength of the electromagnetic waves, and rainfall rate R can also be expressed in terms of $N(D)$,

$$R = \int V(D) v(D) N(D)\, dD \tag{3.1-5}$$

where $V(D)$ is the volume of the precipitation particle of diameter D, $v(D)$ is its falling velocity.

If $N(D)$ is characterized by a single parameter, for example D^*, then Z_e and D^* are in one-to-one correspondence. Since R and D^* is in a one-to-one correspondence as well, once Z_e is obtained, then $N(D)$ can be specified by D^* corresponding to this Z_e, and R can be calculated. In this case, therefore, the attenuation correction used to obtain Z_e from Z_{m0} is the primary problem to be solved. This is the principle of rain estimation for a single frequency radar such as the TRMM PR. In fact, however, variations of $N(D)$ in nature cannot be characterized sufficiently

by a single parameter in many cases. As a result, rain estimates from a single frequency radar often involve errors and biases.

In contrast, variations in $N(D)$ can be well represented by two parameters as far as the conversion from Z_e to R is concerned. If rain is measured by the radar at two wavelengths and if one of the wavelengths divided by 2π is smaller than or comparable to the average raindrop size, then the corresponding backscattering cross section will deviate from that of Rayleigh scattering, and Z_e at this wavelength differs from Z_e at the longer wavelength where Rayleigh scattering generally applies. This situation enables us to estimate two parameters of the model function of $N(D)$ and results in a better estimate of rainfall rate. This is the idea of rainfall estimation with a dual-wavelength radar. In other words, if $N(D)$ is characterized by two parameters N^* and D^*, Z_e at two wavelengths Z_{e1} and Z_{e2} become functions of N^* and D^*,

$$Z_{e1}(N^*, D^*) = \frac{\lambda_1^4}{\pi^5 |K|^2} \int \sigma_{b1}(D) N(D; N^*, D^*)\, dD \tag{3.1-6}$$

$$Z_{e2}(N^*, D^*) = \frac{\lambda_2^4}{\pi^5 |K|^2} \int \sigma_{b2}(D) N(D; N^*, D^*)\, dD \tag{3.1-7}$$

Once Z_{e1} and Z_{e2} are given, we can solve Eq. (3.1-6) and (3.1-7) for N^* and D^*, and R can be calculated from $N(D; N^*, D^*)$.

As in the single-wavelength case, the attenuation corrections to obtain Z_{e1} and Z_{e2} at two frequencies from measured or apparent radar reflectivity factors Z_{m1} and Z_{m2} are crucial. Since the major attenuation comes from precipitation itself, and the DSD parameters can be estimated from Z_{e1} and Z_{e2}, we look for a profile of pairs of DSD parameters (one pair per range gate) that gives the attenuation corrected profiles of Z_{e1} and Z_{e2} that are consistent with the attenuations caused by the precipitation particles whose size distributions are characterized by these DSD parameters. This part of the algorithm is the heart of the rainfall retrieval with dual-frequency radar data. Details of the algorithm are described in Seto et al. (2021).

Attenuation caused by non-precipitation particles and atmospheric gases must be compensated for beforehand. Specifically, we need to take into account the attenuation caused by cloud water, water vapor and oxygen molecules. Meteorological data and storm models are used to estimate their profiles for this purpose.

3.1.1.1 Effective radar reflectivity factor Z_e, specific attenuation k and measured radar reflectivity factor

The effective radar reflectivity factor Z_e is given by Eq. (3.1-4). Similarly, the specific attenuation k due to precipitation is given in terms of the drop size distribution $N(D)$ and the total extinction cross section $\sigma_e(D)$ by

$$k = c_k \int \sigma_e(D) N(D)\, dD \tag{3.1-8}$$

where $c_k = 0.01/\ln(10)$ if k is expressed in [dB (km^{-1})], $\sigma_e(D)$ in [mm^2], and $N(D)$ in [$\mathrm{mm}^{-1}\mathrm{m}^{-3}$]. Let r [km] denote the distance from the radar along the range. The measured radar reflectivity factor $Z_{m0}(r)$ [$\mathrm{mm}^6\mathrm{m}^{-3}$] involves the attenuation by precipitation particles $A_{\mathrm{P}}(r)$ and by non-precipitation particles $A_{\mathrm{NP}}(r)$ as is expressed in the following equation.

$$Z_{m0}(r) = Z_e(r)A_{\mathrm{NP}}(r)A_{\mathrm{P}}(r) = Z_e(r)A_{\mathrm{NP}}(r)\exp\left[-0.2\ln(10)\int_0^r k(s)\,ds\right] \tag{3.1-9}$$

where s is a dummy variable. Eq. (3.1-9) can be rewritten in decibels as shown below.

$$10\log_{10} Z_{m0}(r) = 10\log_{10} Z_e(r) + 10\log_{10} A_{\mathrm{NP}}(r) - 2\int_0^r k(s)\,ds \tag{3.1-10}$$

After attenuation correction for non-precipitation particles, $Z_{m0}(r)$ becomes $Z_m(r)$, where $Z_m(r)$ is given by.

$$Z_m(r) = Z_{m0}(r)A_{\mathrm{NP}}^{-1}(r) = Z_e(r)A_{\mathrm{P}}(r) \tag{3.1-11}$$

By substituting Eq. (3.1-11) into Eq. (3.1-9), the following equation is obtained.

$$10\log_{10} Z_m(r) = 10\log_{10} Z_e(r) - 2\int_0^r k(s)\,ds \tag{3.1-12}$$

3.1.1.2 HB method

Let us assume a power-law relation between k and Z_e as shown below.

$$k(r) = \alpha(r)Z_e^{\beta}(r) \tag{3.1-13}$$

where α and β are coefficients, α can be range dependent, but β should be constant along the range. With the help of Eq. (3.1-13), Eq. (3.1-12) can be solved for Z_e and written in the form of Eq. (3.1-14),

$$Z_e(r) = \frac{Z_m(r)}{\left[1 - 0.2\ln(10)\beta\int_0^r \alpha(s)Z_m^{\beta}(s)\,ds\right]^{1/\beta}} \tag{3.1-14}$$

Alternatively, the equation can be solved for the attenuation factor, A_{P}, giving,

$$A_{\mathrm{P}}(r) = \left[1 - 0.2\ln(10)\beta\int_0^r \alpha(s)Z_m^{\beta}(s)\,ds\right]^{1/\beta} \tag{3.1-15}$$

3.1.1.3 Path integrated attenuation (PIA)

Here, Path Integrated Attenuation (PIA) is defined as integrated attenuation caused by precipitation particles from the radar to the surface. If $r = r_s$ at the surface, PIA is given as shown below.

$$\mathrm{PIA} = -10\log_{10} A_{\mathrm{P}}(r_s) = 2\int_0^{r_s} k(s)\,ds \tag{3.1-16}$$

Eq. (3.1-12) can be rewritten as a function of the PIA.

$$10\log_{10} Z_m(r) = 10\log_{10} Z_e(r) - \text{PIA} + 2\int_r^{r_s} k(s)\,ds \tag{3.1-17}$$

By substituting Eq. (3.1-15) at $r = r_s$ into Eq. (3.1-16), the following equation is obtained.

$$\text{PIA} = -\frac{10}{\beta}\log_{10}\left[1 - 0.2\ln(10)\beta\int_0^{r_s}\alpha(s)Z_m^{\beta}(s)\,ds\right] \tag{3.1-18}$$

The PIA can be estimated by taking the difference between the surface backscattering cross sections with and without precipitation (surface reference technique).

3.1.1.4 Parameterization of drop size distribution function

From a mathematical point of view, $N(D)$ should be parameterized with at most two unknown parameters in order to characterize $N(D)$ deterministically from dual-frequency measurements. To keep the discussion fairly general, $N(D)$ is parameterized in the following way.

$$N(D) = N^* n(D; D^*), \tag{3.1-19}$$

where N^* and D^* are unknown parameters and n is a function of D. By substituting Eq. (3.1-19) into Eq. (3.1-4) and (3.1-8), Z_e and k are given as below.

$$Z_e = N^* I_b(D^*), \quad I_b(D^*) = \frac{\lambda^4}{\pi^5|K|^2}\int \sigma_b(D) n(D; D^*)\,dD \tag{3.1-20}$$

$$k = N^* I_e(D^*), \quad I_e(D^*) = \frac{0.01}{\ln(10)}\int \sigma_e(D) n(D; D^*)\,dD \tag{3.1-21}$$

which shows that Z_e and k can be decomposed into N^* and a function of D^* ($I_b(D^*)$ and $I_e(D^*)$).

3.1.1.5 Retrieval

3.1.1.5.1 Dual-frequency retrieval

If Z_m is available at two frequencies, N^* and D^* can be obtained by solving the following two equations.

$$10\log_{10} Z_{m1}(r) = 10\log_{10} N^*(r) + 10\log_{10} I_{b1}(D^*(r)) - 2\int_0^r k_1(N^*(s), D^*(s))\,ds \tag{3.1-22}$$

$$10\log_{10} Z_{m2}(r) = 10\log_{10} N^*(r) + 10\log_{10} I_{b2}(D^*(r)) - 2\int_0^r k_2(N^*(s), D^*(s))\,ds \tag{3.1-23}$$

The third term of the right hand side of Eq. (3.1-22) or (3.1-23) is equal to the 2-way attenuation from the storm top to range r, which implies that the retrieval should be performed sequentially from the top range bin to the gate of interest. This attenuation can be treated as a known quantity if the attenuation is expressed as a function of the DSD parameters derived at each

gate. Since the solution progresses from the storm top downwards, the method is called the forward retrieval method. If an independent estimate of the PIA is available, in addition to Z_m,

$$10\log_{10} Z_{m1}(r) = 10\log_{10} N^*(r) + 10\log_{10} I_{b1}(D^*(r)) - \text{PIA} + 2\int_r^{r_s} k_1(N^*(s), D^*(s))\, ds \tag{3.1-24}$$

$$10\log_{10} Z_{m2}(r) = 10\log_{10} N^*(r) + 10\log_{10} I_{b2}(D^*(r)) - \text{PIA} + 2\int_r^{r_s} k_2(N^*(s), D^*(s))\, ds \tag{3.1-25}$$

Eq. (3.1-24) and (3.1-25) instead of Eq. (3.1-22) and (3.1-23) can be used to calculate N^* and D^*. In this case, the fourth term of the right hand side of Eq. (3.1-24) or (3.1-25) is equal to the 2-way attenuation from the gate of interest (at range r) to the surface, implying that the retrieval should be done sequentially from the bottom range bin to the gate of interest. As in the forward method, if the attenuation is related to the DSD parameters at each range gate, the fourth terms can be treated as known. This method is called backward retrieval method. In reality, because of ground clutter, the fourth term cannot be calculated without some assumptions regarding the profile of the precipitation in the clutter region. Furthermore, because PIA estimates by the SRT have some error, Eq. (3.1-24) and (3.1-25) differ from Eq. (3.1-22) and (3.1-23), and the backward retrieval method yields different solutions from the forward retrieval method. However, the backward retrieval method is preferable, as it generally gives numerically stable and unique solutions, while the forward retrieval method may more often have two solutions at each range bin.

In the backward retrieval method, we define $10\log_{10} Z_{b1}$ and $10\log_{10} Z_{b2}$ by moving the third and fourth term of the right hand side to the left hand side in Eq. (3.1-24) and (3.1-25).

$$\begin{aligned} 10\log_{10} Z_{b1}(r) &\equiv 10\log_{10} Z_{m1}(r) + \text{PIA} - 2\int_r^{r_s} k_1(N^*(s), D^*(s))\, ds \\ &= 10\log_{10} N^*(r) + 10\log_{10} I_{b1}(D^*(r)) \end{aligned} \tag{3.1-26}$$

$$\begin{aligned} 10\log_{10} Z_{b2}(r) &\equiv 10\log_{10} Z_{m2}(r) + \text{PIA} - 2\int_r^{r_s} k_2(N^*(s), D^*(s))\, ds \\ &= 10\log_{10} N^*(r) + 10\log_{10} I_{b2}(D^*(r)) \end{aligned} \tag{3.1-27}$$

By taking the difference between Eq. (3.1-26) and (3.1-27), we obtain:

$$10\log_{10} Z_{b1}(r) - 10\log_{10} Z_{b2}(r) = I_{b1}(D^*(r)) - I_{b2}(D^*(r)) \tag{3.1-28}$$

As the left hand side of Eq. (3.1-28) is known, D^* can be derived. For liquid precipitation, the right-hand side of Eq. (3.1-28) has a maximum, and Eq. (3.1-28) has generally two solutions.

3.1.1.5.2 Single-frequency retrieval

If Z_m is available only at a single frequency, we need to characterize $N(D)$ with a single parameter. This means that we assume a relationship between N^* and D^*. Once such a relation is assumed, we can translate the N^*-D^* relation to the k-Z_e relation, and the attenuation correction can be carried out by Eq. (3.1-14). If a PIA estimate by SRT ($\mathrm{PIA_{SRT}}$) is substituted into PIA in Eq. (3.1-18), the equality is generally not satisfied. This inequality is caused either by an error in the SRT or by an error in the k-Z_e relation. If the SRT is correct, the k-Z_e relation can be modified to $k = \varepsilon\alpha Z_e^\beta$ where ε satisfies Eq. (3.1-29). This is called the α-adjustment method.

$$\mathrm{PIA_{SRT}} = -\frac{10}{\beta}\log_{10}\left[1 - 0.2\ln(10)\beta\varepsilon\int_0^r \alpha(s)Z_m^\beta(s)\,ds\right] \tag{3.1-29}$$

Once the ε parameter is found from Eq. (3.1-29), the Hitschfeld-Bordan solution with the modified k-Z_e relation provides the attenuation-corrected radar reflectivities. From the above attenuation correction process, k and Z_e are obtained. Then, by taking the ratio of Eq. (3.1-20) to Eq. (3.1-21), D^* can be retrieved.

$$\frac{k}{Z_e} = \frac{I_e(D^*)}{I_b(D^*)} \tag{3.1-30}$$

Generally, the right hand side of Eq. (3.1-30) is a monotonic function of D^* so that Eq. (3.1-30) has a unique solution. Since $N(D)$ is characterized by a single parameter, there is one-to-one correspondence between Z_e and D^* or k and D^*, and we can directly calculate D^* as well.

3.1.2 Overall structure of the algorithm

3.1.2.1 Algorithms

There are three kinds of Level 2 algorithms for the DPR: DPR algorithm, Ku-only (KuPR) algorithm, and Ka-only (KaPR) algorithm. The latter two are single-frequency (SF) algorithms. The DPR algorithm is a dual-frequency (DF) algorithm. The DF algorithm employs both KuPR and KaPR L2 standard products as inputs. Pixels observed by DPR can be categorized into three types: pixels in the inner swath of normal scans (observed both by KuPR and KaPR), pixels in the outer swath of normal scans (observed only by KuPR), and pixels in the interleaved scans (observed only by KaPR in the high-sensitivity mode). The KuPR algorithm is executed for pixels in both inner and outer swaths of normal scans. The KaPR algorithm is executed for pixels in the inner swath of normal scans and in the interleaved scans. The DF algorithm is executed for pixels of all the three kinds.

In the DF algorithm, pixels in the inner swath of normal scans are categorized into the dual-beam (DB) pixels. The other pixels in the DF algorithm and all the pixels in the SF algorithms are categorized into the single-beam (SB) pixels. For a SB pixel, the DF algorithm can use data in dual-frequency observations at neighboring pixels, while the SF algorithms can use data only

in the corresponding single-frequency observations. For example, by using the characteristics of the drop size distribution estimated by dual-frequency measurements at DB pixels in the inner swath, the DF algorithm can give better estimates at SB pixels in outer swath than the KuPR algorithm. Each algorithm executes data for one orbit at a time. For an orbit, only after the two SF algorithms are properly executed, the DF algorithm will be executed. The order of execution between the KuPR algorithm and the KaPR algorithm is not fixed. The algorithms are designed in such a way that the order of execution for different orbits has no restrictions within the same calendar month, but one month of data should be executed as a unit to produce proper databases.

3.1.2.2 Input/Output Files

The input and output files of the KuPR, KaPR and DF algorithms are as follows.

KuPR algorithm

(Input) KuPR L1B standard product, Weather data file, Database file

(Output) KuPR L2 standard product, KuPR L2 temporary file, KuPR L2 environment data

KaPR algorithm

(Input) KaPR L1B standard product, Weather data file, Database file

(Output) KaPR L2 standard product, KaPR L2 temporary file, KaPR L2 environment data

Dual-frequency algorithm

(Input) KuPR L2 standard product, KaPR L2 standard product, KuPR L2 temporally file, KaPR L2 temporary file, Database file, KuPR L2 environment data, KaPR L2 environment data

(Output) DPR L2 standard product, DPR L2 environment data, DPR L2 temporary file

The temporary files include some results of the SF algorithms, which are not written in the standard products. The temporary products can be treated as research products, once the DPR L2 standard product and other outputs are properly produced by the DF algorithm.

The environmental data include the profiles of atmospheric parameters assumed in the L2 algorithm. Because of its large volume, the ancillary environmental data are output into a separate file from the standard product.

The weather data file is prepared by converting the resolution and/or variables of weather analysis/reanalysis/forecast dataset in advance to L2 algorithm. Database files are explained later.

3.1.2.3 Modules

The framework of the L2 algorithm is shown in Figure 3.1.1. This framework is common to the KuPR, KaPR, and DF algorithms. The main module manages the overall procedure, and it employs seven sub-modules. The main module will open and close files, call sub-modules, and read/write all the input and output files and variables. It will also terminate the algorithm. Sub-modules can read/write files and variables as long as they are permitted. As a basic rule, they should process for all the pixels in the orbit, and the order of pixels processed is not fixed (can be determined by each module). Sub-modules cannot call other sub-modules. When a sub module is terminated, the processing is returned to the main module.

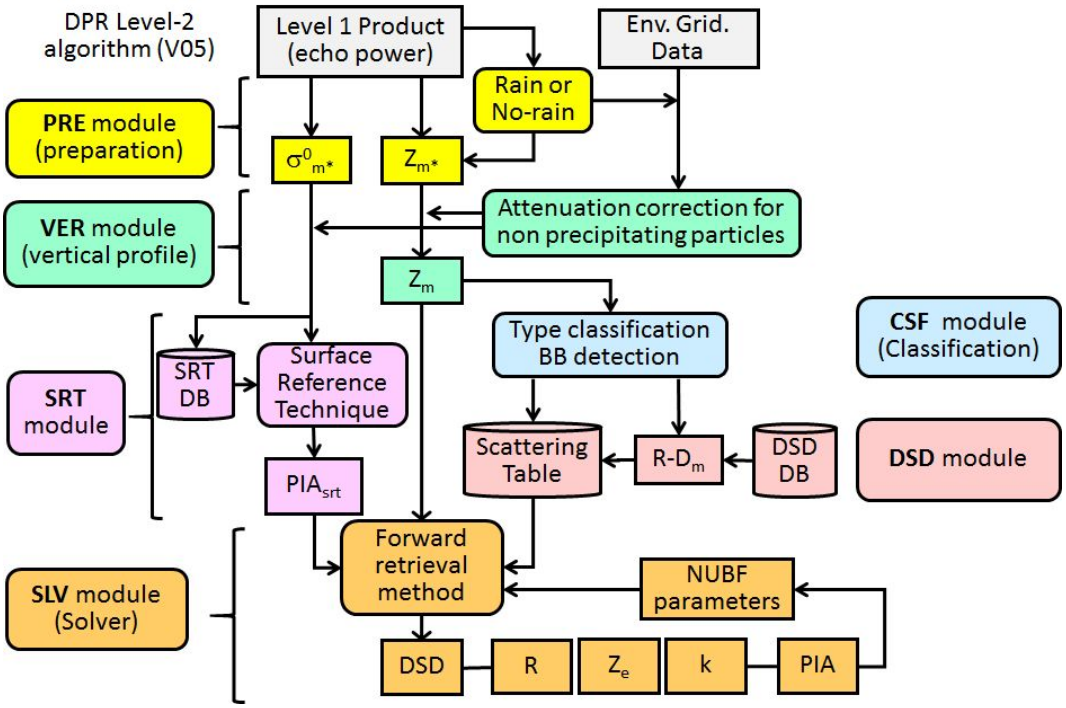


Figure 3.1.1: DPR algorithm flow

3.1.2.4 Basic roles of the modules

The L2 algorithm is to numerically solve Eq. (3.1-22) or (3.1-23) for DB pixels and Eq. (3.1-30) for SB pixels to obtain N^* and D^*. The retrieval process is carried out in the Solver module, but the preparation of equations is shared by the other six modules. The PIA in Eq. (3.1-29), (3.1-24) and (3.1-25) can be estimated in the SRT module. The DSD module is responsible for quantifying the terms in the equations such as α, β, $n(D;D^*)$, $I_b(D^*)$, and $I_e(D^*)$ based on the physical characteristics of precipitation (precipitation types and bright band information) determined by the Classification module. Z_m is converted from Z_{m0} through attenuation correction for non-

precipitation particles in the Vertical Profile module and Z_{m0} is converted from received echo power in the Preparation module.

3.1.2.5 Variables

A tentative table of variables is attached to this document. The table is used for all the three algorithms in the L2 algorithms. There are two types of variables: L1B variables and L2 variables. L1B variables are just copied from the L1B standard products. L2 variables are somehow processed in the L2 algorithms. Not all but some selected variables are written in L2 standard, temporary, or environment products. A variable with scan dependence, angle bin dependence, range bin dependence, and frequency dependence is expressed as an array in the source code. Some variables may have other dimensions than scan, angle bin, range bin, and frequency. Type (bytes) of a variable can be changed when it is written into the product to save the file size. Some variables can take different values in the SF algorithms and DF algorithm. Those variables are product dependent. Product dependence and frequency dependence are explained by examples below.

- "Location of observation" is without product dependence and without frequency dependence.
- "Measured radar reflectivity factor Z_{m0}" is without product dependence and with frequency dependence.
- "Equivalent radar reflectivity factor Z_e" is with product dependence and with frequency dependence.

In the DF algorithm, variables with product dependence should be reprocessed, but variables without product dependence can be copied from SF products. Variables without product dependence and without frequency dependence should take the same value in KuPR product and KaPR product, but there is some possibility of disagreement as the KuPR algorithm and KaPR algorithm are processed independently. In the case of disagreement, a rule (e.g., KuPR algorithm is more reliable) should be prepared to determine the value in the DF algorithm.

3.1.2.6 Databases

A database is to include useful information for the algorithm, and exists as separate files from the main body of the algorithm (source code). For example, the SRT database is used in the SRT module to increase the accuracy of the PIA estimates. The L2 algorithm can refer to and update databases. The actual updating process is done not by directly modifying the database file, but by creating an intermediate file. The intermediate files are summarized and the database file is modified off line.

3.1.2.7 Look-up tables

A look-up table has a similar role with a database, but the difference is that a look-up table will basically not be updated. Therefore, we can treat the look up tables as part of the main body of the algorithm, and do not list the look-up tables as input and output files. For example, the L2 algorithm employs a DSD look-up table and a σ^0 look-up table in the DSD module and the SRT module, respectively.

3.2 Main module

The basic procedure of the main module is described in Section 3.2.1 below. The test version before the data release in 2014 used the basic procedure. Version 3 algorithm adopts the advanced procedure Section 3.2.2.1. In near future, advanced procedures described in Section 3.2.2.2 and Section 3.2.2.3 should be tested.

3.2.1 The basic procedure

In the basic procedure, sub-modules are executed in the following order, and each module is called only once. Firstly, the Preparation module is executed. This module judges the existence of precipitation, and identifies the precipitation pixels. Secondly, the Vertical profile module is executed. The ancillary atmospheric profile data are used to correct for the attenuation caused by non-precipitation particles to obtain the surface backscattering cross sections both at precipitation and no-precipitation pixels. Thirdly, the Classification module classifies each precipitation pixel into an appropriate storm type. At no-precipitation pixels, almost no processes are taken.

Fourthly, the SRT module is executed. At no-precipitation pixels, SRT database is updated based on the measurement of surface backscattering cross sections. Note that the SRT module can be executed anytime after the Vertical profile module and before the Solver module. Finally, the Solver module is executed. At no-precipitation pixels, almost no processes are taken.

3.2.2 Advanced procedure[1]

By executing some of the modules multiple times, improvement of results is expected. Two kinds of examples are given below.

3.2.2.1 Recursive procedure

Once the basic procedure is done, the same procedure is repeated again with the parameters estimated in the first cycle. In the second cycle, for example, we can estimate the attenuation due to cloud liquid water by referring the precipitation rate estimated in the first cycle.

3.2.2.2 Parallel procedure

In the basic procedure, the sub-modules are required to determine the values of variables in charge, but sometimes they do not have enough information to estimate the values in a deterministic way. A parallel procedure allows a sub-module to give multiple estimates and to let the following sub-modules execute with multiple estimates. For example, if the precipitation type is not determined with confidence in Classification module, then the following (DSD and Solver)

[1] This subsection is obsolete.

modules are executed with multiple assumptions of precipitation types, and the main module checks which assumption gives the vertical profiles of Z_e and the corresponding precipitation rate R in accordance with the assumed precipitation type. Thus, the precipitation type can be determined afterwards.

3.2.2.3 Additional processes

Some important processes may be missing in the basic procedure, for example, non-uniform beam filling (NUBF) correction. They should be involved in the final version of the algorithm, but currently, it has not been determined which modules are in charge of additional processes; NUBF correction can be done inside the Solver module, or should be shared with the DSD module and/or the SRT module, or a new module (Texture module) should be introduced.

3.3 Preparation (PRE) module

3.3.1 Objectives and functions of the Preparation Module (PRE)

The primary purposes of the preparation module are (1) to classify the pixels into either rain or no-rain pixels, (2) to convert the received power P_r into measured reflectivity factor Z'_m without any attenuation corrections at each range bin, and (3) to calculate apparent normalized surface cross section $\sigma^{0\,\prime}_m$ without any attenuation correction at each pixel.

3.3.2 Algorithm descriptions of the PRE

In this section, the preparation module of the Ku-band level-2 algorithm is described in detail. The preparation module of the Ka-band level-2 algorithm is almost the same as that for the Ku-band. The DPR level-2 algorithm can use level-2 products and level-2 supplementary products of both Ku-only and Ka-only algorithms. As a result, they can provide necessary information to other modules of the DPR level-2 algorithm without re-calculations.

Reading of input data

The module reads data from the Ku-band level-1B product that includes not only the radar echoes but also other variables related to the measurements such as scan time (`ScanTime`), footprint position (`Latitude`, `Longitude`), local zenith angle (`scLocalZenith`), operational mode (`operationalMode`), elevation (`DEMHmean`), and data quality flag (`dataQuality`). In the case of the DPR level-2 algorithm, the module reads level-2 products and level-2 supplementary products of both Ku-band and Ka-band algorithms.

Status confirmation

The module refers to `dataQuality` of level-1B product and reflect `qualityData` of level-2 product scan by scan. If `dataQuality` is bad (not 0), the level-2 values of the corresponding scan are overwritten by missing values.

Calculation of range distance (`rangeDist`)

The range distance (`rangeDist`) is defined by the distance from the satellite to each range bin along the radar beam. Each `rangeDist` is calculated from the range distance from the satellite to (the center of) the highest range bin (`startBinRange`), the number of range bins with normal sampling (`echoLowResBinNumber`), the number of range bins with over-sampling (`echoHighResBinNumber`) and the range bin size (`rangeBinSize`). Specifically, the range distance (`rangeDist`) for normal sampling ranges are calculated by using the following equation.

$$\texttt{rangeDist} = R_0 + (n-1) \times \Delta R, \quad (n = 1, 2, 3, \cdots, N)$$

The `rangeDist` for over sampling ranges are

$$\texttt{rangeDist} = R_0 + \{N + (m-1)\} \times \Delta R, \qquad (m = 1, 2, 3, \cdots, M)$$

where

$R_0 =$ `startBinRange` in L1B product
$N =$ `echoLowResBinNumber` in L1B product
$M =$ `echoHighResBinNumber` in L1B product
$\delta R =$ `rangeBinSize` in L1B product

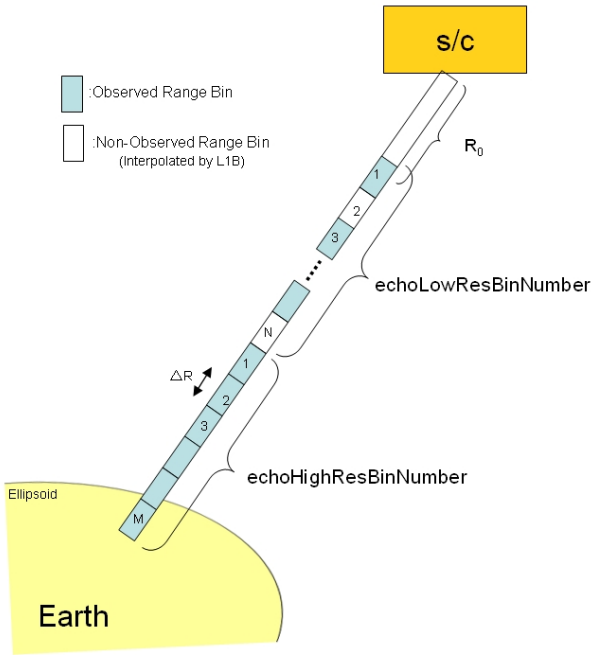


Figure 3.3.1: Definition of `rangeDist`

Calculation of height (`Height`)

The height is defined by the vertical distance from the footprint of radar beam on the ellipsoid to the range bin in question. In order to calculate `Height`, we define `ellipsoidBinOffset` as follows.

$$\begin{aligned}\texttt{ellipsoidBinOffset} =& \texttt{scRangeEllipsoid} \\ & - \{R_0 + (\texttt{binEllipsoid} - 1) \times \texttt{rangeBinSize}\}\end{aligned}$$

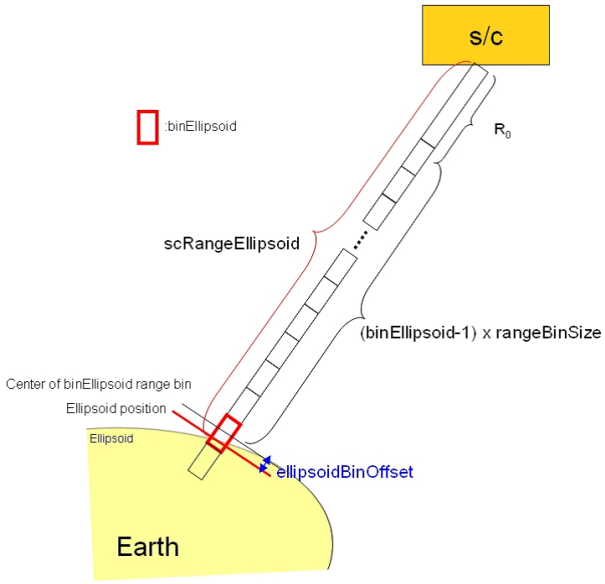


Figure 3.3.2: Definition of `ellipsoidBinOffset`

Then, `Height` of `binRangeNo` (as shown "p" in Figure 3.3.3) is calculated by the following equation.

$$\begin{aligned}\texttt{Height[binRangeNo]} = &\{(\texttt{binEllipsoid\_2A} - \texttt{binRangeNo}) \times \texttt{rangeBinSize} \\ &+ \texttt{ellipsoidBinOffset}\} \times \cos(\texttt{localZenithAngle}),\end{aligned}$$

where `binEllipsoid_2A` is the ellipsoid range bin number in level 2. `binEllipsoid_2A` is 176 for Ku and KaMS, and 88 for KaHS. `rangeBinSize` is 125 m for Ku and KaMS, and 250 m for KaHS.

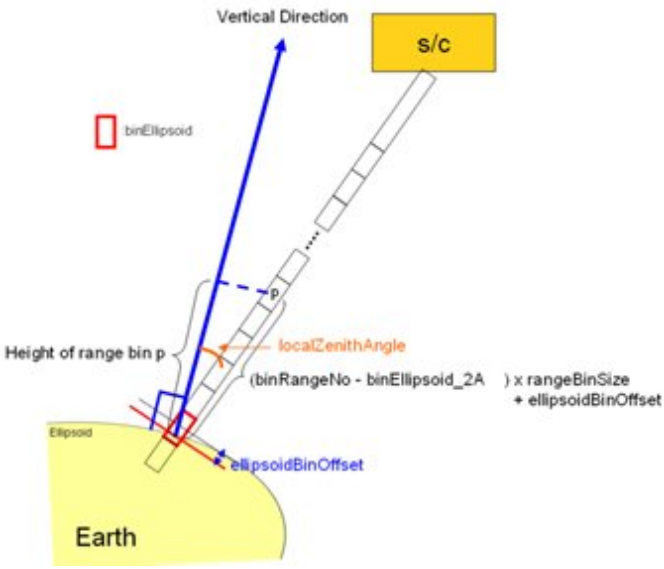


Figure 3.3.3: Definition of `Height`

Extraction of level-1B range data

`echoPower` and `echoFlag` of level-1B product are given at each range bin. The array size of data with range-bin dependence in Ku level-1B is 260. The array sizes of data with range-bin dependence in Ka_MS and Ka_HS level-1B are 260 and 130, respectively. On the other hand, the array size of data with range-bin dependence in Ku level-2 is 176, whereas the array sizes for range bins of Ka_MS and Ka_HS level-2 are 176 and 88, respectively. The preparation module renumbers the range bins relative to the ellipsoid position in level-2 so that the range bin number at the ellipsoid becomes 176 in the Ku and Ka_MS level-2 products, and 88 in the Ka_HS products. Thus, at the nadir angle bin, data from 0 km to +22 km altitude relative to the ellipsoid are extracted.

In the version 06X or later versions, 'FS' structure was introduced in all level-2 products. The FS structure in the new version of Ku level-2 product coincides NS structure in current Ku level-2 product, whereas the FS structure in the new version of Ka level-2 product (Ka_FS) stores both current Ka_MS and Ka_HS level-2 products. Note that the new version of Ka level-2 has HS structure, however, its contents are set to missing after the scan pattern change (See Appendix).

Figure 3.3.4 shows horizontal structure of the Ka_FS in the new version of Ka level-2 product, Ka_MS and Ka_HS in the current Ka level-2 products. As for the inner swath of Ka_FS, angle bin number 1 to 25 of Ka_MS at scan number n are stored at angle bin number 13 to 37 of Ka_FS at scan number n. In the outer swath of Ka_FS, angle bin number 1 to 12 of Ka_HS at scan number n are stored at angle bin number 38 to 49 of Ka_FS at scan number n, and angle bin number 13 to 24 in Ka_HS at scan number n are stored at angle bin number 1 to 12 of Ka_FS at

scan number $n+1$. In the first scan of Ka_FS, missing values are stored in angle bin number 1 to 12 because there are no corresponding angle bins of Ka_HS. Angle bin number 1 to 12 in the last scan of Ka_HS are discarded.

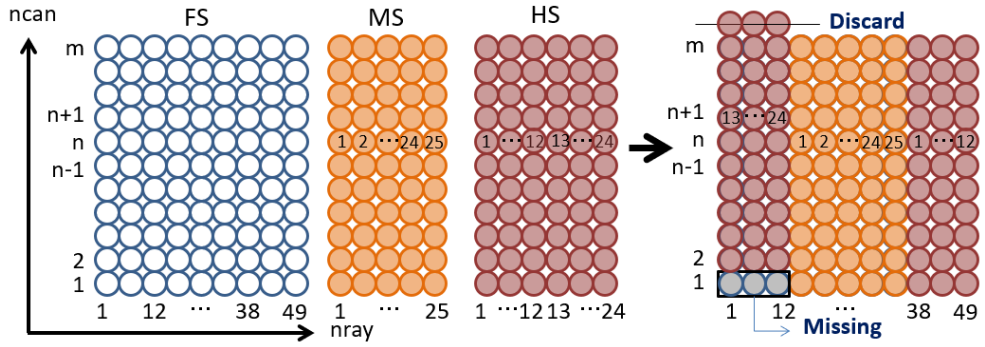
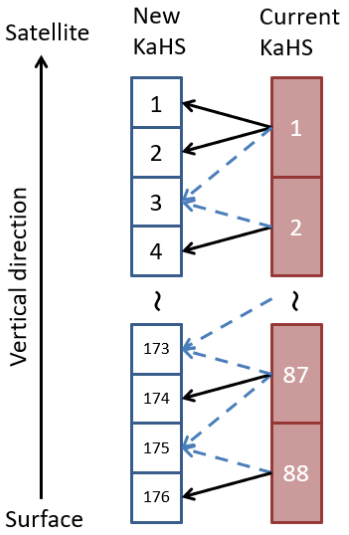


Figure 3.3.4: How to merge for horizontal structure of the Ka_MS and Ka_HS.

The vertical structure in the outer swath of the Ka_FS level-2 product differs from the current Ka_HS level-2 product. Figure 3.3.5 shows the schematic image of the difference. In the current version of Ka_HS, the array size of `echoPower` is 88. The new preparation module interpolates this 88 range bin data to make them 176 range bin data so that its vertical resolution coincides with the Ku_FS's vertical resolution for the dual-frequency algorithm.



Original
Average

New KaHS Rangebin number	Current KaHS Rangebin number
1	1
2	1
3	mean(1+2)
4	2
5	mean(2+3)
~	~
173	mean(86+87)
174	87
175	mean(87+88)
176	88

Figure 3.3.5: Interpolation for vertical structure of the Ka_HS.

Matching echo profiles between KuPR and KaPR

As described in the previous subsection, `binEllipsoid` in the level-2 products is defined with respect to the bin that includes the ellipoid surface. A small mismatching of KuPR and KaPR's beam directions may occasionaly identify different range bins in level 1 data as the ellipsoid bin in KuPR and KaPR products, because `binEllipsoid` in level-2 KuPR and KaPR products is determined individually. This inconsistent assignment of `binEllipsoid` between KuPR and KaPR causes mismatching of range data in the dual-frequency algorithm. Figure 3.3.6 depicts an example of such a mismatched case. Figure 3.3.6 (a) shows the alignment of level 1 data. The ellipsoid appears at different bins at the center of the beams of KuPR and KaPR. This is because the range bins have a certain width (i.e., range bin size) and the actual echo peaks can be stored in different range bins for KuPR and KaPR if the two beams are not matched. If range data are realigned accoding to the `binEllipsoid` individually defined by the KuPR and KaPR algorithms, the misalignment shown in Figure 3.3.6 (b) will result. To avoid this discrepancy, the preparation module in the dual frequency algorithm shifts KaPR's range bin array. As a result, the actual echo profiles of KuPR and KaPR match as shown in Figure 3.3.6 (c). This range bin adjustment is applied to variables with range bin arrays in the dual-frequency product.

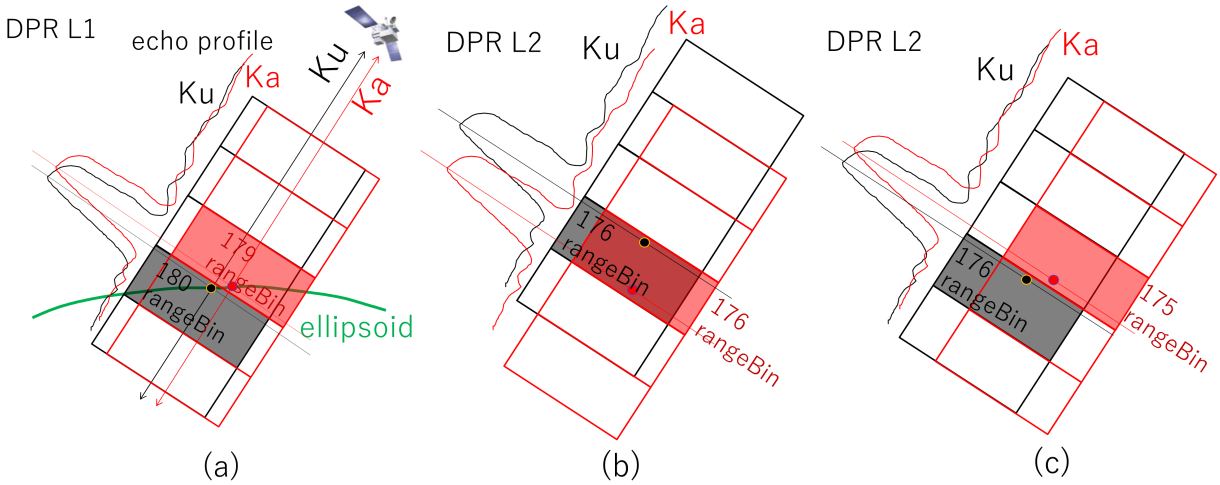


Figure 3.3.6: A concept for matching of `binEllipsoid` between KuPR and KaPR

Mirror image flag (`binMirrorImageL2`)

The mirror-image-flag (`binMirrorImage`) of level-1B product gives the lowest range bin number above which unusual false precipitation echoes may appear near the top of the vertical profile. These false echoes are caused by mirror image echoes of precipitation at very high altitude that leak into the subsequent sampling window. The preparation module calculates this range bin number based on the `binEllipsoid_2A` and the PRF (or sampling window interval) for each

footprint.

Land surface type (landSurfaceType)

The land surface type of the level 2 product is based on the land-and-ocean flag (landOceanFlag) of the level-1B product which gives land, ocean, inland water, and coast classification for each footprint. The preparation module adds surface type information for land and coast area, using footprint position (Latitude and Longitude) and an external data base. The surface type information is stored in landSurfaceType.

The landOceanFlag of level-1B product is based on the SWBD, which is SRTM-3 water body mask data. Within the SWBD range (latitude 60 S -60 N), DPR land-ocean flags are generated based on SWBDs, and missing areas are supplemented by other data bases.

On the other hand, those outside the SWBD range (south of latitude 60 S and north of 60 N) are generated in the same format using TRMM's land and sea flag data (DID) and GSHSS (A Global Self-consistent, Hierarchical, High-resolution Shoreline Database).

The DID was used in the TRMM algorithms.
https://www.eorc.jaxa.jp/TRMM/documents/PR_algorithm_product_information/pr_manual/PR_Instruction_Manual_V7_L1.pdf#page=52

Snow- or Ice-covered area (snowIceCover)

In version 5, a new flag called snowIceCover is added to indicate the areas that are covered with surface snow or sea ice. This flag is used in the SRT module when it needs to separate different surface types and also when it collects surface scattering data to create the σ^0 database. The flag is based on NOAA/NESDIS autosnow.

Surface detection (binRealSurface)

The preparation module determines the surface range bin using the received power (echoPower), and DEM related data, such as binDEM, scRangeDEM, DEMHmean, binDEMHtop, and binDEMHbottom of level-1B product. The estimated surface position is stored in binRealSurface.

Estimation of Clutter-free bottom (binClutterFreeBottom)

The range bin number of the clutter-free bottom is estimated using echoPower profiles. binRealSurface may be used as a reference of the search window. The clutter-free bottom range bin number is stored in binClutterFreeBottom.

In V5, binClutterFreeBottom is automatically raised up by one range bin (by adding a one range bin margin to binClutterFreeBottom). This automatic one range bin raise-up sometimes causes the contamination of binClutterFreeBottom with the side lobe clutter. This problem

is solved in V6 by further raising up the `binClutterFreeBottom` (by more than one range bin until the contamination disappears) when the contamination with the side lobe clutter occurs. The V5 code also has a very small bug on Ka-HS `binClutterFreeBottom`. This bug is fixed in V6.

In V07, an adjustment of the `binClutterFreeBottom` is introduced with the clutter pattern described latter. For KuPR, since the strong sidelobe clutter in angle bins 18-20 and 30-32 (1 origin) is separated only slightly from the mainlobe edge, the clutter pattern near the mainlobe edge is complicated, Some sidelobe clutter is erroneously identified as precipitation echo and `binClutterFreeBottom` is wrongly selected. The resultant overestimation of the precipitation rate in these angle bins appears as spikes in the statistics of the incident angle dependence of precipitation rate estimates. Here, the lower limits of the `binClutterFreeBottom` in angle bins 18-20 and 30-32 are given for the KuPR data. For KaPR, `mainlobeEdge`, that is used for the first guess of `binClutterFreeBottom`, is adjusted because the signal to noise ratio of the KaPR's surface clutter is lower than that of KuPR's. After the scan pattern change of the KaPR, the `mainlobeEdge` of angle bins 1-12 and 38-49 is lowered by 1 or 2 range bins.

Reduction of sidelobe clutter contamination

A routine to reduce the sidelobe clutter was installed in the KuPR/KaPR L2 algorithms originally developed by Kubota et al. (2016, 2018). In the version 07, the improved clutter removal routine was installed in the single frequency (KuPR, KaPR, and PR) L2 algorithms based on results of Kanemaru et al. (2020). The filter technique for reducing the sidelobe clutter, which was first implemented in V06X, was also improved with development of the precipitation judgement. The improved method of the clutter removal is described in the paper by Kanemaru et al. (2021) so that the current document briefly describes the changes in the clutter removal routine from version 06 to version 07. Table 3.3.1 indicates the major changes in the clutter removal routine between versions 06 and 07. The clutter signal power P_c at the sampled range r_n is obtained as an area integration of surface scattering as follows (Kozu, 1995);

$$P_c(r_n) = \frac{P_t L_p \lambda^2}{(4\pi)^3} \int_S \frac{\sigma^0 G^2(\theta,\phi)|u[2(r_n - r)/c]|^2}{r^4}\, ds$$

where P_t, L_p, λ, σ^0, G^2, θ, ϕ, $|u|^2$, c, r, S are the transmitted power, the peak loss factor, wavelength, the normalized radar cross section of the Earth surface, the two-way antenna gain, polar angle, and azimuth angle, the normalized pulse waveform, the light of speed, the range, and the illuminated surface by the pulse, respectively. Note that θ, ϕ and r are functions of the position s on the surface. The clutter power in version 06 is estimated by using the following regression formula:

$$P_{c,\mathrm{est}}(\mathrm{rb}_n)\ [\mathrm{mW}] = A_1(\mathrm{rb}_n) \times \sigma^0_m(\mathrm{ab}_1)\ [\mathrm{m}^2/\mathrm{m}^2] + A_2(\mathrm{rb}_n) \times \sigma^0_m(\mathrm{ab}_2)\ [\mathrm{m}^2/\mathrm{m}^2] + B(\mathrm{rb}_n)$$

where A_1, A_2, B are regression coefficients. Measured surface cross sections σ^0_m at two angles are used. The coefficients were generated as functions of the range-bin (rb), angle-bin (ab), satellite height, and surface type. Since the clutter databases in the version 06 were generated as a function of range-bin number corresponding to the range resolution of 125 m (250 m for KaPR's HS mode), the change in geometrical relationship between the surface and radar within the range-bin resolution was not resolved so that the clutter removal was sometimes insufficient. Moreover, the clutter databases in version 06 were generated after the adjustment of phase shifter that changed the antenna pattern slightly. The routine assumed that the clutter pattern in the same configuration was invariant. However, accumulated data showed that the clutter pattern changed with time. It was necessary to improve the clutter removal algorithm for the long-term processing during the GPM mission.

To improve clutter removal and to deal with the long-term reprocessing of the DPR data, a new improved method for reducing the clutter was developed in version 07. In version 07, the clutter signal power is estimated as follows:

$$P_{c,\mathrm{est}}(\theta_c)\ [\mathrm{dB}m] = P_{s,\mathrm{est}}\ [\mathrm{dB}m] + G_c(\theta_c)\ [\mathrm{dB}]$$

where θ_c, $P_{s,\mathrm{est}}$, G_c are the angle in the cross-track direction, the estimated surface peak power and the clutter pattern reconstructed by DPR data, respectively. The clutter patterns were generated with the resolution of the sub-range-bin (0.01 range-bin) and the cross-track angle (0.01 degrees) by using satellite height (`scAlt`) and digital elevation model data (`DEMHmean`). The cross-track angle is estimated as the angle that intersects the Earth surface with the law of cosines. The surface peak power is required, but the DPR scans with the interval of 0.71 degrees in the cross-track direction. Here, the surface peak power is estimated with interpolation with the discrete angle data of the DPR. Moreover, to deal with the temporal change in the clutter pattern, the clutter patterns were generated monthly from August 2014 to December 2020 and switched monthly in the reprocessing. In the standard processing in version 07, the clutter pattern in December 2020 is used for the clutter removal. We will plan to update the latest clutter database if necessary.

In version 07, the clutter removal is processed for KuPR, KaPR for matched scan mode and high-sensitivity mode. The clutter removal is conducted not only for the sidelobe signals but also for the mainlobe signals experimentally. Note that the clutter-free-bottom is estimated with the original level-1 power data so that the change in the clutter removal does not affect the estimates of the clutter-free-bottom.

For TRMM PR, the clutter estimates are conducted, but the clutter removal is NOT conducted. The clutter estimation in version 07 requires accurate geometrical information of radar beams with respect to the surface. The attitude error of the TRMM satellite is larger than that of the GPM core satellite so that the clutter removal for the PR is not conducted. Note that the

estimated clutter power is used for the PR's precipitation judgement described latter.

Table 3.3.1: A difference in the clutter removal routine

	Version 06 or older	Version 07 or later
Estimates in P_c	Statistical relation of σ^0-P_c	Geometrical calculation given by P_s and G_c
Resolution of the databases	Range-bin (125 m or 250 m)	Cross-track angle and sub-range-bin (0.01 degrees and 0.01 range-bin)
Temporal change in the databases	Only radar's configuration	Not only radar's configuration but also monthly
References	Kubota et al. (2016, 2018)	Kanemaru et al. (2020, 2021)

Calculation of Signal power (`echoSignalPower`)

Signal power is calculated at all range bins in L1B product. If the corresponding radar data are missing, they are flagged by `dataQuality`. In version 06 and older versions, the signal power P_s is given by subtracting the noise power P_n from the observed (or clutter-removed) received echo power P_r, that is:

$$\texttt{echoSignalPower}\ [\mathrm{dBm}] = 10\log_{10}(P_s\ [\mathrm{mW}]) = 10\log_{10}(P_r\ [\mathrm{mW}] - P_n\ [\mathrm{mW}]).$$

`echoPower` and `noisePower` are stored in dBm unit in L1B product.

$$\texttt{echoPower}\ [\mathrm{dBm}] = 10\log_{10}(P_r\ [\mathrm{mW}]) \text{ and } \texttt{noisePower}\ [\mathrm{dBm}] = 10\log_{10}(P_n\ [\mathrm{mW}]).$$

If P_s becomes negative, missing value is stored in `echoSignalPower`. In version 07, the expected value of $10\log_{10} P_s$ or $\langle 10\log_{10} P_s\rangle$ is introduced and contained in the product. If the measurement error does not exist, the equation above is exactly correct. However, the echoes from the natural target fluctuate so that averaging many radar echoes in the same configuration is required to reduce the fading noise. For the DPR, the received echo power is averaged after the logarithmic receiver so that the measurement error of the power is theoretically given by $5.57/\sqrt{N}$ where N is the independent sampling number. For the DPR, the typical number of pulses at each angle bin is about 100 so that the measurement error of the P_r or σ_r is $5.57/\sqrt{100} = 0.557$ [dB] corresponding to $\sim$14% in linear space. Since `echoPower` and `noisePower` fluctuate independently, the proper magnitude of P_s is not obtained from the first equation. We derive the expected value of P_s by taking the fading of P_r and P_n into account (Iguchi, 2017). The expected value of $10\log_{10} P_s$ is given as a conditional average when positive P_s is logarithmically averaged as follows:

$$\langle 10\log_{10} P_s\rangle = \frac{\int_x^\infty \int_{-\infty}^\infty 10\log_{10}[10^{y/10} - 10^{x/10}]p(x)p(y)\,dxdy}{\int_x^\infty \int_{-\infty}^\infty p(x)p(y)\,dxdy}$$

where $p(x)$ and $p(y)$ are the probability density functions for x = `noisePower` and y = `echoPower`, respectively, and given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma_n^2}} \exp\left[-\frac{1}{2}\left(\frac{x - 10\log_{10} P_n}{\sigma_n}\right)^2\right]$$

$$p(y) = \frac{1}{\sqrt{2\pi\sigma_r^2}} \exp\left[-\frac{1}{2}\left(\frac{y - 10\log_{10} P_n}{\sigma_r}\right)^2\right]$$

Here, σ_n is the standard deviation of `noisePower`. The standard deviations of `echoPower` and `noisePower` are obtained from the effective independent sampling numbers, N_r and N_n, as shown in Table 3.3.2.

Table 3.3.2: N_r and N_n for calculating the expected value of P_s

	KuPR / KaPR MS	KaPR HS	PR
N_r	100	100	64
N_n	800	400	256

Rain/NoRain classification (`flagEcho`, `flagPrecip`, `binStormTop`, `heightStormTop`)

Up to version 06, Rain/NoRain classification is carried out in two steps. In the first step, it is done for all range bins above the `binClutterFreeBottom` except missing data which are labeled by `dataQuality` in L1B product. If `echoSignalPower` exceeds a certain threshold, it means that rain is detected for the range bin. The results of Rain/NoRain classification for range bins are stored in `flagEcho`. In the second step, using the surrounding `flagEcho` results, the Rain/NoRain classification for angle bin is carried out for all angle bins except missing data. The results of Rain/NoRain classification for each angle bin are stored in `flagPrecip`. The modules in the downstream in the flow chart may use the `flagPrecip` to determine the target pixels for processing. The function also detects a highest rain position in each angle bin and provides `binStormTop` and `heightStormTop`.

In version 07, the new precipitation judgement is developed to improve detectability of precipitation signals. Figure 3.3.7 indicates the concept of the precipitation judgement. In version 06 and older, the precipitation judgement is conducted by searching signals above the threshod in more than 4 succesive range bins in the vertical direction (1-D judgement) so that thin precipitation layer cannot be detected even if the precipitation echoes are higher than radar's sensitivity. In version 07, we count up the successive signals higher than the certain level not only in the vertical direction but also in the cross-track and along-track directions (3-D judgement). The 3-D judgement is expected to improve the detection of thin precipitation layers distributed horizontally such as high-latitude precipitation.

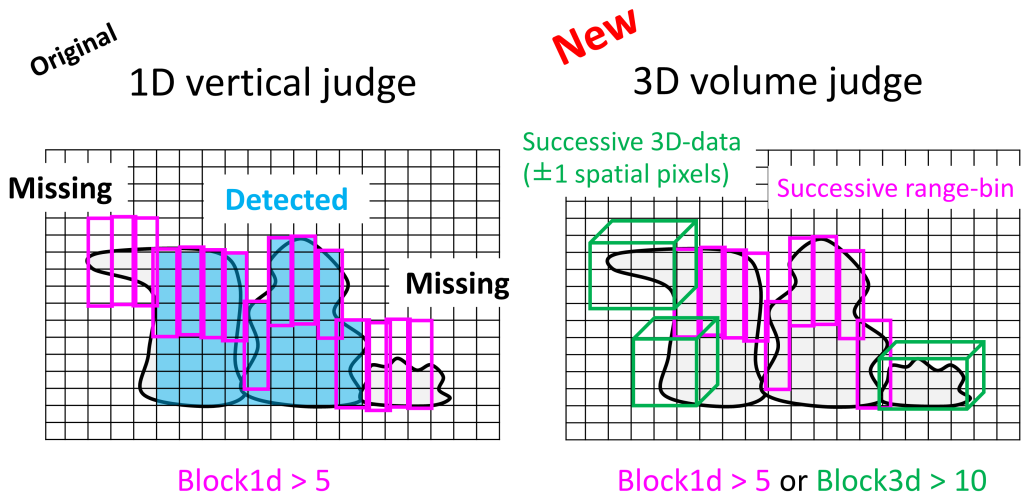


Figure 3.3.7: Precipitation judgement

The new precipitation judgement is processed according to the flowchart shown in Figure 3.3.8 and separated into some sub-process: Judge-1D, Filter-A, Judge-3D, and Filter-B.

- Judge-1D (search successive signals in the range-bin direction as block)

First, the precipitation judgement from the top of the valid range-bin to the bin of the clutter-free-bottom is processed to extract the successive signal of the `echoPower` data in the range-bin direction

$$P_r\ [\mathrm{dB}m] - P_{c+n}\ [\mathrm{dB}m] > M_{\mathrm{threshold}} \times 5.57\sqrt{\frac{1}{N_r} + \frac{1}{N_n}}\ [\mathrm{dB}]$$

where P_{c+n} is the sum of the clutter and noise powers, that is estimated as $P_{c+n}\ [\mathrm{dB}m] = 10\log_{10}(P_{c,\mathrm{est}}\ [\mathrm{mW}] + P_n\ [\mathrm{mW}])$, and $M_{\mathrm{threshold}}$ is a threshold for the precipitation judgement. In version 07, $M_{\mathrm{threshold}} = 2.33$ (equivalent to the 99% one sided confidence interval of the normal distribution). Then, the number of successive range-bin is extracted as a block in the range-bin direction (block1D).

- Filter-A (filter doubtful signal simply)

After 1D-judgement, a simple filtering is conducted to remove doubtful signals such as fading noise and clutter signals. For example, the bottom of the position of block1D higher than 15 km above the ellipsoid is judged as sidelobe signals.

- Judge-3D (count up signals in cross-track/along-track directions)

The number of signals higher than the threshold is counted up as block3D by searching surrounding ±1 cross-track/along-track/range-bin area of block1D. Thus, searched volume of the 3-D judgement is $3 \times 3 \times (\text{block1D} + 2)$.

- Filter-B (filter possible clutter signals)

Filtering is conducted with the shape of block3D. When block1D is much larger than a threshold, the filtering is skipped. By searching the surrounding area, the threshold is set to be 10. If sidelobe signals are found in the surrounding area, the threshold is 15. Note that the thresholds for KaPR'S HS mode is assumed to be 6 or 9 when the sidelobe signals are found or not found. Here, $\text{block2D}_{\text{XZ}}$ (± 1 along-track/range-bin area) and $\text{block2D}_{\text{YZ}}$ (± 1 cross-track/range-bin area), where X is the along-track direction and Y is the cross-track direction and Z is the range-bin direction, is also counted up. If $\text{block2D}_{\text{XZ}}$ or $\text{block2D}_{\text{XZ}}$ is equal to block1D (thus, the shape of block3D is long and narrow in the horizontal direction), its block is judged as residual clutter signals.

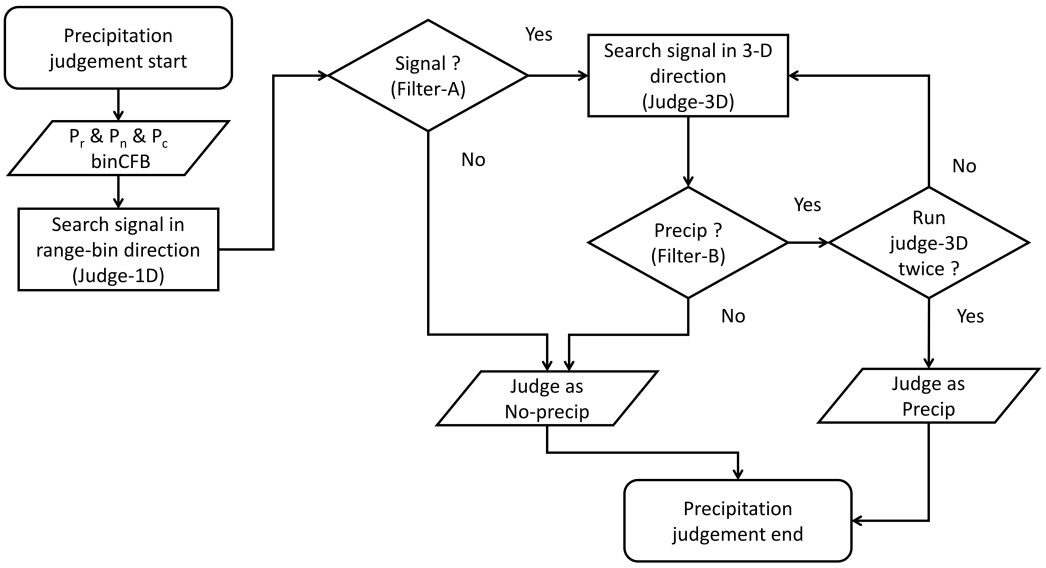


Figure 3.3.8: Flowchart of precipitation judgement

Since the 3-D judgement depends on the spatial pattern of signals, the 3-D judgement is conducted again after the processing of Filter-B. Finally, the judged signals are flagged as precipitation's bit of `flagEcho`. The top of the precipitation signal is set as `binStormTop` and its height `heightStormTop`, and `flagPrecip` is set to be 1 or 2. `flagPrecip` = 1 is set when block1D is larger than the threshold (thus, almost same as the conventional method). `flagPrecip` = 2 is set when block3D is larger than the threshold although block1D is smaller than its threshold,

which is the new definition of `flagPrecip` with the 3-D judgement. Table 3.3.3 shows a rule of `flagPrecip` for single frequency algorithms. For dual frequency algorithm, $\texttt{flagPrecip}_{\texttt{DPR}}$ is given as follows:

$$\texttt{flagPrecip}_{\texttt{DPR}} = 10 \times \texttt{flagPrecip}_{\texttt{Ku}} + \texttt{flagPrecip}_{\texttt{Ka}}$$

Table 3.3.3: A rule of `flagPrecip`. Note that the numbers in parentheses indicate the thresholds for KaPR's HS mode.

	block1D $\geq$ 5 (3)	block1D $<$ 5 (3)
block3D $\geq$ 10 (5)	flagPrecip = 1	flagPrecip = 2
block3D $<$ 10 (5)	flagPrecip = 1	No-precip

Calculation of Z'_m factor (`zFactorMeasured`)

Z'_m is defined as the measured reflectivity factor without any attenuation corrections. Z'_m is calculated using a radar equation at all range bins above the `binClutterFreeBottom` except missing data which are labeled by `dataQuality` in L1B product. The result of Z'_m is stored in `zFactorMeasured`. The radar equation is as follows.

$$\begin{aligned} P_r(r) &= C \cdot P \cdot L_r \cdot E \\ &= [P_t \cdot G_{\mathrm{at}} \cdot G_{\mathrm{ar}}] \left[\left(\frac{\lambda}{4\pi r} \right)^4 \frac{4\pi}{\lambda^2} \right] L_r \left[\frac{\pi^5}{\lambda^4} |K|^2 Z_e 10^{-18} c\tau \frac{\pi r^2 \theta_{0a} \theta_{0c}}{2^4 \ln(2)} \right] \\ &= \frac{\pi^3 c}{2^{10} 10^{18} \ln(2)} \frac{P_t G_{\mathrm{at}} G_{\mathrm{ar}} \theta_{0a} \theta_{0c} \tau |K|^2}{\lambda^2 r^2} L_r Z_e(r) \end{aligned}$$

In the preparation module, Z'_m $(= L_r Z_e)$ is calculated by using the following equation.

$$P_r(r) = \frac{\pi^3 c}{2^{10} 10^{18} \ln(2)} \frac{P_t G_{\mathrm{at}} G_{\mathrm{ar}} \theta_{0a} \theta_{0c} \tau |K|^2}{\lambda^2 r^2} Z'_m(r)$$

$$Z'_m(r) = \frac{2^{10} 10^{18} \ln(2)}{\pi^3 c} \frac{\lambda^2 r^2 P_r(r)}{G_{\mathrm{at}} G_{\mathrm{ar}} \theta_{0a} \theta_{0c} \tau |K|^2 P_t}$$

where

r	range distance (`rangeDist`)
L_r	loss factor
λ	wave length (`eqvWavelength` of L1B)
$\lvert K \rvert^2$	dielectric factor (0.9255 for KuPR and 0.8989 for KaPR)
G_t	transmitter antenna gain (`txAntGain` of L1B)
G_r	receiver antenna gain (`rxAntGain` of L1B)
θ_c	Cross-track beam width (`crossTrackBeamWidth` of L1B)
θ_a	Along-track beam width (`alongTrackBeamWidth` of L1B)
c	speed of light
τ	transmitter pulse width (`transPulseWidth` of L1B)
P_t	transmitted power (`radarTransPower` of L1B)
P_r	received signal power (`echoSignalPower`)

It should be noted that $|K|$ is a constant in the preparation module. Any adjustment of $|K|$ by temperature is controlled in the solver module if necessary.

In TRMM PR, over-sampled received powers of surface and rain echoes (e.g. 125m range resolution) are stored in the `osSurf` and `osRain`, and received powers of the normal samples (e.g. 250m range resolution) are stored in the `normalSample` (TRMM Precipitation Radar Team, 2011).

The Z'_m of TRMM PR can be calculated from P_r which is combined by `osSurf`, `osRain` and `normalSample`. To adjust the range resolution of the `normalSample` to the DPR range resolution (e.g. 125m range resolution), an interpolation is conducted in linear scale for each ray independently.

Calculation of $\sigma^{0\,\prime}_m$ (`sigmaZeroMeasured`)

$\sigma^{0\,\prime}_m$ is defined as the measured normalized surface cross section without any attenuation corrections. Calculation of $\sigma^{0\,\prime}_m$ is done for all footprints except missing data which are labeled by `dataQuality` in L1B product. The value of $\sigma^{0\,\prime}_m$ is stored in `sigmaZeroMeasured`.

$$\sigma^{0\,\prime}_m(\theta_z) = P_{rs}(r_0)\frac{512\pi^2\ln(2)}{P_t\lambda^2 G_t G_r}\frac{\mathrm{loss}}{1}\frac{\cos(\theta_z)r_0^2}{\theta_a\theta_{bp}}$$

where

$$\theta_{bp}^{-1} = \sqrt{\theta_c^{-2} + \theta_p^{-2}}, \quad \theta_p^{-1} = \frac{2}{c\tau}r_0\tan(\theta_z)$$

and

P_{rs}	received signal power (`echoSignalPower`)
r_0	range distance from the satellite to the geographic surface (ref. `binRealSurface`)
θ_z	zenith angle (`localZenithAngle`)
loss	band path filter loss

3.3.3 Interfaces to other algorithms

As to the Ku-band and Ka-band level-2 algorithms, input data for this module is Ku-band and Ka-band level-1B product, respectively. For the DPR level-2 algorithm, input data for this module are level-2 products and level-2 supplementary products of both Ku-band and Ka-band. The outputs will be used by the Solver module and other modules in the DPR algorithm.

3.3.4 Output Variables

`adjustFactor`: Adjustment factor (dB) for `zFactorMeasured` (dBZ'_m) and `sigmaZeroMeasured` (dB$\sigma^{0\,\prime}_m$). dBZ'_m = dBZ_m − `adjustFactor`, dB$\sigma^{0\,\prime}_m$ = dBσ^0_m − `adjustFactor`

`binClutterFreeBottom`: Range bin number for clutter free bottom.

`binRealSurface`: Range bin number for real surface.

`binStormTop`: Range bin number for the storm top.

`elevation`: Elevation of the measurement point.

`ellipsoidBinOffset`: Distance between the ellipsoid and a center range bin of `binEllipsoid`.

`flagPrecip`: Flag of Rain/NoRain classification.

`flagSigmaZeroSaturation`: Flag of saturation of `sigmaZeroMeasured`.

`heightStormTop`: Height of storm top.

`landSurfaceType`: Land surface type.

`localZenithAngle`: Local zenith angle of each ray.

`sigmaZeroMeasured`: Surface backscattering cross section without attenuation correction.

`snRatioAtRealSurface`: Signal/Noise ratio at real surface range bin.

`snRatioAtRealSurface`: $10 \times \log_{10}$(echoPower[mW]/noisePower[mW])

`snowIceCover`: Snow and ice cover information. It refers to the multisensor snow/ice cover maps provided by NOAA.

`zFactorMeasured`: Vertical profile of reflectivity factor without attenuation correction.

Details of the description for variables are described in Appendix or in the following websites:

`https://www.eorc.jaxa.jp/GPM/en/archives.html`

`https://arthurhou.pps.eosdis.nasa.gov/GPMprelimdocs.html`

3.4 Vertical Profile (VER) module

3.4.1 Objectives and functions of the Vertical Profile Module (VER)

The primary purposes of the VER are to read ancillary environmental data, to provide vertical profiles of the environmental parameters, to compute the path-integrated attenuation (PIA) due to non-precipitation (NP) substances, and to give radar reflectivity factors corrected for attenuation by the non-precipitation substances. The VER provides environmental information such as pressure, temperature, water vapor, and cloud liquid water at each range bin. The VER calculates the attenuation due to water vapor, molecular oxygen, and cloud liquid water.

For the single frequency algorithms (Ku, Ka, and Ka_HS), the VER is executed using the information of the pixel lat/lon and the range bin width. For the DPR, results of the Ku, Ka, Ka_HS are introduced.

3.4.2 Algorithm descriptions of the VER

Basic functions of the VER are reviewed in Kubota et al. (2020). Here, they are described in terms of "Utilization of ancillary environmental data", "Calculation of attenuation by water vapor", "Calculation of attenuation by molecular oxygen", and "Calculation of attenuation by cloud liquid water (CLW)".

Utilization of ancillary environmental data

The VER inputs ancillary environmental data, objective analysis data by Japan Meteorological Agency (JMA) named as JMA Global Analysis (GANAL) and JMA's forecast (FCST) data. In TRMM/PR processing, the VER module uses the JMA's reanalysis (JRA55) (Kobayashi et al. 2015; Harada et al. 2016), instead of the GANAL and the FCST. By reading the ancillary environmental product, the VER provides pressure, temperature, water vapor, and cloud liquid water for each range bin. In addition, the VER computes the height of 0 degree centigrade and flags the range bin corresponding to that height.

The horizontal resolutions of the JMA data are 0.5 degree latitude/longitude, respectively. The pressure levels such as 1000, 925, 850, 700, 600, 500, 400, 300, 250, 200, 150, 100 hPa, and so on are converted to the height levels before the input of the algorithm. The temporal resolution is 6 hourly as 00, 06, 12, and 18Z. The VER inputs two 6-hourly files and computes the values of parameters at the time of measurement using linear temporal interpolation. We also use a linear horizontal and vertical interpolation to calculate the parameters at the location of each range bin from the lat/lon's of the footprint and the satellite and the range bin height. Figure 3.4.1 shows how this interpolation is carried out by finding the antenna beam direction from the lat/lon information and the local zenith angle of the spacecraft and the "height" provided by the

preparation module (see Section 3.3). The VER provides pressure, temperature, water vapor, and cloud liquid water at each range bin.

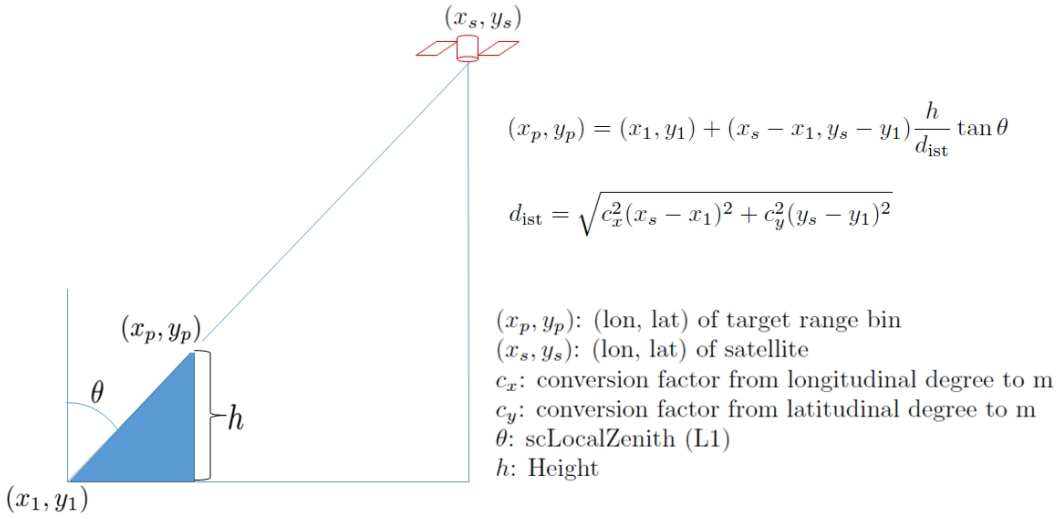


Figure 3.4.1: Calculation of the coordinates of each range bin in the VER.

Calculation of attenuation by water vapor

At a frequency less than 100 GHz, the attenuation coefficient due to water vapor, $\kappa_{\text{H2O}}(f)$ (dB/km) is expressed as follows (Waters 1976, Ulaby et al. 1981, Meneghini and Kozu 1990),

$$\kappa_{\text{H2O}}(f) = 2f^2\rho_\nu \left(\frac{300}{T}\right)^{3/2} \gamma_l \left[\left(\frac{300}{T}\right) e^{-644/T} \frac{1}{(494.4 - f^2)^2 + 4f^2\gamma_l^2} + 1.2 \times 10^{-6}\right]$$

where

f: frequency (GHz)
ρ_ν: water vapor content (g/m3)
γ_l: parameter of line width (GHz)
T: temperature (K)

The line width parameter γ_l is given as

$$\gamma_l = 2.85 \left(\frac{P}{1013}\right) \left(\frac{300}{T}\right)^{0.626} \left(1 + 0.018\frac{\rho_\nu T}{P}\right)$$

where P: pressure (hPa).

In precipitation pixels, the water vapor content provided by the GANAL may be problematic due to its relatively low spatial resolution (0.5x0.5 deg. lat/lon). Therefore, a minimum threshold is used in the VER. The relative humidity is set to 90% when the GANAL gives a value less than 90% in precipitation pixels.

Calculation of attenuation by molecular oxygen

For attenuation by molecular oxygen, the following expression is valid for frequencies less than 45 GHz (Rosenkranz 1975, Ulaby et al. 1981, Meneghini and Kozu 1990).

$$\kappa_{\mathrm{O2}}(f) = 1.1 \times 10^{-2} f^2 \left(\frac{P}{1013}\right)\left(\frac{300}{T}\right)^2 \gamma \left[\frac{1}{(f-f_0)^2+\gamma^2} + \frac{1}{f^2+\gamma^2}\right]$$

where

f: frequency (GHz)
f_0: 60 GHz
γ: parameter of line width (GHz)
T: temperature (K)

Line width parameter γ is given by

$$\gamma = \gamma_0 \left(\frac{P}{1013}\right)\left(\frac{300}{T}\right)^{0.85}$$

where

$$\gamma_0 = \begin{cases} 0.59, & \text{for } 333 \leq P(\mathrm{hPa}); \\ 0.59[1 + 3.1 \times 10^{-3}(333 - P)], & \text{for } 25 \leq P < 333(\mathrm{hPa}); \\ 1.18, & \text{for } P < 25(\mathrm{hPa}). \end{cases}$$

Calculation of attenuation by cloud liquid water (CLW)

For attenuation by cloud liquid water κ_{CLW} over non-precipitation pixels, cloud water contents in the ancillary environmental data product are utilized. In the cloud scheme of the JMA Climate Data Assimilation System (JCDAS) product (Onogi et al. 2007), the effective radius (r_e) of cloud liquid droplets is fixed at 15 microns. In this calculation, we assume cloud particle distribution $n_c(D)$ as mono-disperse, that is,

$$n_c(D) = N_c \delta(D - 2r_e)$$

where

D: Diameter
r_e: 15 microns
N_c: Number concentration of cloud liquid particles
$\delta()$: Dirac delta function

For cloud liquid water content Q (kg/m3), N_c is expressed as follows,

$$N_c = \frac{3Q}{4\pi r_e^3 \rho_w}$$

κ_{CLW} is computed due to the Rayleigh scattering using $n_c(D)$. The mono-disperse assumption implies the homogeneity of the cloud distribution, although clouds generally distribute with high inhomogeneity within the grid-size of the reanalysis such as 1.25-degree latitude/longitude. Therefore, the previous formula will underestimate κ_{CLW} over precipitation pixels. In the 2A25 algorithm for the TRMM/PR, κ_{CLW} over precipitation pixels is estimated based on the result of a numerical simulation of storms with a cloud-system resolving model (CRM) that was used to create the database for the TRMM/TMI 2A12 algorithm (Iguchi et al. 2009). In their method, the attenuation due to CLW is estimated as a function of surface precipitation rate, separately for convective columns and stratiform columns. The VER adopts a method which expands Iguchi et al. (2009) and κ_{CLW} is estimated from a CLW database as a function of surface precipitation rate, precipitation type (convective or stratiform), temperature, latitude, and land surface type (Kubota et al. 2012, 2020). The CLW database was generated by 3.5 km-mesh global simulations performed using a Nonhydrostatic ICosahedral Atmospheric Model (NICAM). The NICAM is a global cloud-system resolving model (GCRM), and explicitly calculates moist convection using a cloud microphysical scheme. The NICAM data were provided by the University of Tokyo and Japan Agency for Marine-Earth Science and Technology (JAMSTEC).

3.4.3 Input Variables

Input: from MOSS or PPS Ancillary environmental data, Two 6-hourly files

Surface: Surface pressure, Mean sea level pressure

Pressure levels: Geopotential height, temperature, water vapor, and cloud liquid water

Input: from Preparation Module

`Geolocation`
`scanTime`
`Elevation`
`landSurfaceType`
`localZenithAngle`
`flagPrecip`
`binRealSurface`
`binStormTop`
`heightStormTop`
`binClutterFreeBottom`
`sigmaZeroMeasured`
`zFactorMeasured`
`rangeBottom`
`rangeTop`
`Height`

3.4.4 Definitions of Output Variables

`airTemperature[176][49]`: Temperature

`heightZeroDeg[49]`: Height of the level of 0 degree centigrade.

`binZeroDeg[49]`: Range bin number corresponding to the level of 0 degree centigrade.

`binZeroDegSecondary[49]`: When the inversion layers are detected, the "`binZeroDegSecondary`" is used to output the `binZeroDeg` related to the inversion layer. A peak level with the ground surface of T < 0°C. (" `binZeroDeg` " is 177).

`flagInversion[49]`: Flag for an inversion layer of air temperature related to the 0 degrees C level.
-1 : the surface ground below 0°C
0 : The VER code detects a level of 0°C without an inversion layer.
>=1 : The VER code detects a level of 0°C with an inversion layer.

`attenuationNP[176][49][4]`:

$$\mathrm{Atten}_{\mathrm{NP}}(r) = \kappa_{\mathrm{WV}}(r) + \kappa_{\mathrm{O2}}(r) + \kappa_{\mathrm{CLW}}(r)$$

`attenuationNP[nbin][nray][nNP]`: includes 4 types of the nNP defined as follows.
nNP=1: Sum of Attenuations by water vapor, molecular oxygen, and cloud liquid water
nNP=2: Attenuation by water vapor (dB/km)
nNP=3: Attenuation by molecular oxygen (dB/km)
nNP=4: Attenuation by cloud liquid water (dB/km)

`piaNP[49][4]`: PIA by the non-precipitation substances.

$$\mathrm{PIA}_{\mathrm{NP}}(r) = 2\int_0^r \mathrm{Atten}_{\mathrm{NP}}(s)\,ds = 2\int_0^r \kappa_{\mathrm{WV}}(s) + \kappa_{\mathrm{O2}}(s) + \kappa_{\mathrm{CLW}}(s)\,ds$$

`piaNP[nray][nNP]`: includes 4 types of the nNP defined as follows.
nNP=1: Sum of PIA by water vapor, molecular oxygen, and cloud liquid water
nNP=2: Attenuation by water vapor (dB/km)
nNP=3: Attenuation by molecular oxygen (dB/km)
nNP=4: Attenuation by cloud liquid water (dB/km)

`piaNPrainFree[49][4]`: "Rain-free" path-integrated attenuation due to non-precipitation substances (piaNP). This is calculated assuming a simple atmosphere with less than 50% relative humidity and 0 cloud liquid water.

`piaNPrainFree[nray][nNP]`: includes 4 types of the nNP defined as follows.
nNP=1: Sum of PIA by water vapor, molecular oxygen, and cloud liquid water
nNP=2: Attenuation by water vapor (dB/km)
nNP=3: Attenuation by molecular oxygen (dB/km)
nNP=4: Attenuation by cloud liquid water (dB/km)

`zFactoNPCorrected[176][49]`: Radar reflectivity corrected for attenuation by the non-precipitation substances.

$$Z_m(r) = Z'_m(r)/A_{\mathrm{NP}}(r)$$

$$A_{\mathrm{NP}}(r) = \exp(-q\mathrm{PIA}_{\mathrm{NP}}(r)) \quad \text{where} \quad q = 0.1\ln(10)$$

`sigmaZeroNPCorrected[49]`: σ^0 corrected for attenuation by the non-precipitation substances.

$$\sigma^0 = \sigma^{0\prime}/A_{\mathrm{NP}}(r)$$

$$A_{\mathrm{NP}}(r_{\mathrm{sfc}}) = \exp(-q\mathrm{PIA}_{\mathrm{NP}}(r_{\mathrm{sfc}})) \quad \text{where} \quad q = 0.1\ln(10)$$

Details of the description for variables are described in Appendix or in the following websites:

`https://www.eorc.jaxa.jp/GPM/en/archives.html`

`https://arthurhou.pps.eosdis.nasa.gov/GPMprelimdocs.html`

3.4.5 Intermediate Files

Before the input of the algorithm, the pressure levels of the ancillary environmental data are converted to the height levels. In addition, the attenuations and PIA estimates by water vapor, molecular oxygen, and cloud liquid water are computed in advance.

3.4.6 Description of the Processing Procedure

Before the input of the algorithm, the pressure levels of the ancillary environmental data are converted to the height levels. In addition, the attenuations and PIA estimates by water vapor, molecular oxygen, and cloud liquid water are computed in advance.

For Ku, Ka, Ka_HS, the VER is executed using the information of the pixel lat/lon and the range bin width. For the DPR, calculated results of the Ku, Ka, Ka_HS are used.

3.4.7 Interfaces to other algorithms:

Input data for this algorithm is from the Preparation Module and Ancillary environmental data; the outputs are used by the Classification module, DSD module, Solver Module and others.

3.5 Classification (CSF) module

3.5.1 Introduction

In reliable rain rate retrieval using space-borne radars, information about the drop size distribution (DSD) is necessary. DSD varies depending on the rain type. Therefore, rain type classification plays an important role in the GPM DPR algorithm. There exist two distinctive rain types, one is stratiform and the other is convective. Stratiform rain is characterized by its weakness in intensity, wide extension in area, and in many cases accompanying a bright band (BB) in radar echo. Because of the last characteristic, detection of a BB can be used for determining stratiform rain.

The rain type classification is made on pixel basis. Hence, the rain type is the same along a radar beam. Then there arises the following ambiguous situation. Suppose that, when the radar reflectivity is examined along a given radar beam, the precipitation echo happens to exist only at altitude higher than 0°C height. What is the rain type for this case? Is it stratiform or convective? This is a difficult question to answer. To handle such an ambiguous case, the third category of "other" type is introduced in a similar manner to the case of the TRMM PR algorithm 2A23. The third category "other" means that there exists only cloud or possibly noise when the radar echo is examined along the radar beam.

A surprise in the early TRMM PR observation was a ubiquitous shallow isolated rain system, which may be warm rain but it still remains a puzzle what it actually is. TRMM PR observation shows that shallow isolated rain is weak. This observation does not agree with the concept of warm rain which is thought to be very strong. There also exists shallow non-isolated rain, whose statistical properties seem to be very similar to those of shallow isolated rain. Shallow rains (both shallow isolated and shallow non-isolated rain) are marked by a shallow rain flag, `flagShallowRain`, which is independent of the rain type flag, `typePrecip`.

Main features of each CSF module, from V05 to V07, are as follows.

In the V05 classification (CSF) module, new items (`flagHeavyIcePrecip`, `flagAnvil`, `surfaceSnowfallIndex`, and `flagSurfaceSnowfall`) were added to the output fields and a decision on winter convective type was also made. Section 3.5.5 describes these features.

In the V06 CSF module, reclassification of precipitation type from convective to stratiform is made by using a method developed by the UW (University of Washington) (Houze and Brodzik, 2017; private communication), which examines the slope of Z profile above the 0°C isotherm (see Section 3.5.4.1.4(e) for details).

In the V06X CSF module, the full scan (FS) Ka band single frequency data processing and the FS dual frequency data processing were made for the first time.

In the V07 CSF module, a new `flagHail` is added to the STANDARD group of the dual

frequency output. In the V07 CSF module, the winter time temperature inversion problem is solved by consulting a new flag `flagInversion` from the VER module.

Except for those and bug fixes, the V07 CSF algorithms are very close to those of V06X in the parts where the processing is made on the data taken after the Ka-band antenna scan pattern change made on May 21, 2018. Though the CSF algorithms themselves are very similar between V06X and V07, a large increase of rain type counts are observed in V07 because of the introduction of Dr. Kanemaru's three dimensional (3D) precipitation detection processing in the V07 PRE module.

3.5.2 Objectives

The CSF module detects a bright band (BB) and classifies rain into three major categories, which are stratiform, convective, and other.

3.5.3 Algorithm Principles

In single frequency CSF modules, i.e., in Ku-only and Ka-only modules, rain type classification is made by a V-method (Vertical profiling method) and by an H-method (Horizontal pattern method) (Awaka et al., 1998, 2009). The rain types by the V-method and H-method are unified, and the Classification (CSF) module outputs the unified rain type, which consists of three major categories; stratiform, convective, and other.

In the dual frequency algorithm of GPM DPR, however, rain type classification is made by a new method called measured dual frequency ratio ($\mathrm{DFR_m}$) method (Le and Chandrasekar, 2013a, 2013b) and by using the single frequency result from the Ku-only module. The $\mathrm{DFR_m}$ method is a kind of vertical profiling method, and classifies rain type into stratiform, convective and transition. Though transition is a new type, it is not a major category. The rain types by the $\mathrm{DFR_m}$ method and those by the Ku-only module are unified (Awaka et al. 2016, 2021). Three major categories, after the unification of rain types of $\mathrm{DFR_m}$ method and Ku-only results, are again stratiform, convective and other.

3.5.3.1 Single frequency: V-method

In the single frequency V-method, detection of BB is made first. Detection of a BB is made by examining the vertical profile of the radar reflectivity factor (Z) to see if the vertical profile of Z satisfies certain conditions which are typical to the profile of Z when a BB exists. When a BB is detected, the rain type is stratiform if the reflectivity factor in the rain region does not exceed a special convective threshold (46 dBZ). When a BB is not detected, and the reflectivity factor exceeds a conventional convective threshold (40 dBZ), the rain type is convective. (For

details of convective thresholds, see Section 3.5.4.1.4(a).) When rain type is neither stratiform nor convective, the rain type is "other" in the V-method.

3.5.3.2 Single frequency: H-method

In the H-method, a horizontal pattern of representative radar reflectivity factors is examined. Here the representative reflectivity factor means the maximum value of reflectivity factor in the rain region along the considering radar beam. Rain type is classified using a modified University-of-Washington convective/stratiform separation method (Steiner et al. 1995, Yuter and Houze 1997). In the H-method, rain type is again classified into three categories: stratiform, convective and other. In the H-method, detection of convective rain is made first. If the rain type is not convective, the rain type is stratiform unless the reflectivity factor is very small, being almost identical to noise. If rain type is neither convective nor stratiform, the rain type is other.

3.5.3.3 Dual frequency: $\mathrm{DFR_m}$ method

In the dual frequency module, in place of the V-method, the $\mathrm{DFR_m}$ method is used for rain type classification and for BB detection in the inner swath. The $\mathrm{DFR_m}$ method uses the difference between the measured Ku-band Z in dB and the measured Ka-band Z in dB. To make the BB detection reliable, BB detected by the $\mathrm{DFR_m}$ method and that by the Ku-only algorithm are unified in the dual frequency CSF module. Rain types by the $\mathrm{DFR_m}$ method are stratiform, convective and transition.

Before May 21, 2018, the Ka-band radar operated in the HS mode in the inner swath with the interlaced antenna scan pattern. For the data before May 21, 2018, the $\mathrm{DFR_m}$ method is also applied to the inner swath HS data by interpolation, which is made vertically and horizontally. In the vertical interpolation, the Ka-band data with a 125 m interval is interpolated from the 250 m interval HS data. In the horizontal interpolation, four Ku-band pixels adjacent to the HS pixel are used for the interpolation of the corresponding Ku-band data at the HS pixel.

3.5.4 Actual Algorithm (Data Processing)

3.5.4.1 Single frequency processing

In the Ku-only and Ka-only algorithms, BB detection is based on a search for a typical BB peak by examining the range profile of radar reflectivity along a given antenna beam. This method is effective for the Ku-band data which show clear BB peaks. In the case of Ka-band, however, since a clear BB peak is not expected, detection of BB is not effective.

3.5.4.1.1 Selection of pixels to be processed:

- Skip the process for the pixel if no precipitation echo exists in Ku-band and Ka-band data. This judgement can be made by examining a flag from the PRE module, `flagPrecip`.

- Skip the process for the pixel if the quality flag of Ku-band and/or that of Ka-band shows that the pixel is a bad pixel or that the data is missing. This judgement can be made by examining a flag from the PRE module, `qualityData`. (It is assumed here that the information about missing data is also available from the `qualityData` flag.)

3.5.4.1.2 Determination of range where precipitation echoes exist:

- Determination of the echo top and the echo bottom, where the echo top is given by `rangeStormTop` from the PRE module and the echo bottom is given by `binClutterFreeBottom` from the PRE module.

3.5.4.1.3 Detection of bright band (BB) by the single frequency method

- Set the BB search window using `rangeZeroDeg` from the VER module. The BB search window ranges from `rangeZeroDeg` -4 to `rangeZeroDeg` $+8$ with a 250 m interval, where the center of the BB window is `rangeZeroDeg` $+2$ which is about 0.5 km below the 0 degree height. This window range is reasonable from TRMM PR experience and many other radar observations (unfortunately, Ka-band radar data may not exhibit a clear BB peak, but the BB search window can be the same for the Ku-band and Ka-band). Experience shows that BB peak appears about 0.5 km below the 0 degree height.

Detection of BB is made using the NP-attenuation corrected radar reflectivity factors, `zFactorNPCorrected`, from the VER module, where NP-attenuation means attenuation due to non-precipitation substances.

Detection of BB can be made by a vertical method, which examines the profile of radar reflectivity factor. If the profile of the radar reflectivity satisfies certain conditions which characterize BB, it is determined that there exists BB. The above characteristic conditions are different between Ku-band and Ka-band. In the Ku-band, a sharp BB peak should be observed in the profile of radar reflectivity. In the Ka-band, however, a clear BB peak may not be observed, but there must be a detectable characteristic change in the slope in the radar reflectivity profile near the BB.

When BB is detected, the upper boundary of BB (i.e., `rangeBBTop`) and the lower boundary of BB (i.e., `rangeBBBottom`) are determined. The lower boundary of BB is detected first. The lower boundary of BB is defined as the point where there is the largest change in the slope of Z in the region just below the BB peak. This definition is very close to that by Fabry and Zawadzki (1995). The upper boundary of BB is determined by finding the following two points A and B:

Point A: where there is the largest change in the slope of Z in the upper region of BB peak.

Point B: where Z becomes smaller than Z at the lower boundary of BB for the first time when Z is examined upward in the upper part of BB starting from the BB peak.

When points A and B are the same, the upper boundary of BB is defined as the point A (which is the same as the point B in this case). When points A and B are different, the upper boundary of BB is defined as either A or B that is closer to the BB peak. The definition of the upper boundary is the one which is somewhere in between the definition by Fabry and Zawadzki (1995) and that by Klaassen (1998).

The width of BB (`widthBB`) is computed by the following empirical formula:

$$\texttt{widthBB} = [(\texttt{rangeBBBottom} - \texttt{rangeBBTop}) \times 125 - L\sin(\theta_z)]\cos(\theta_z) \qquad [\text{m}]$$

where θ_z is the local zenith angle, and L is given below:

$$L = L_0 F / \cos^2(\theta_z)$$

Here L_0 (=5000 [m] tentatively) is the footprint size of antenna beam, and F is an empirical factor ($F = 0.5$, tentatively). When the above equation gives $\texttt{widthBB} < 250\cos(\theta_z)$ [m], the width is set to $\texttt{widthBB} = 250\cos(\theta_z)$ [m]. When BB is detected, the following quantities are computed or given and written to the output of the single frequency CSF module.

- `flagBB`: This flag indicates that BB is detected. (`flagBB` > 0 when BB is detected.)
- `rangeBBPeak`: Range bin number of BB peak. In the case of Ku-band, it is simple and straightforward. In the case of Ka-band, BB peak may not be clear, and the peak position may be displaced from that of Ku-band. Nevertheless, a simple peak search method is used also at Ka-band.
- `rangeBBTop`: Range bin number of BB top.
- `rangeBBbottom`: Range bin number of BB bottom.
- `heightBB`: Height [m] of BB which corresponds to `rangeBBPeak`.
- `widthBB`: Width of BB [m].
- `qualityBB`: Quality flag of BB detection. Details are still TBD. The flag `qualityBB` is meaningful when BB is detected. In V06X, `qualityBB`=1 always when BB is detected (which means that the quality of all the detected BB is good).

3.5.4.1.4 Rain type classification by the single frequency method

In each single frequency CSF module, that is, in Ku-only or Ka-only CSF module, rain type classification is made by two methods: one is a V-method and the other is an H-method. V- and H-methods classify rain into three categories; stratiform, convective and other. The rain types by these two methods are unified, and each single frequency CSF module outputs the unified rain type, which consists of three major categories; stratiform, convective and other.

Each single frequency DPR L2 algorithm adopts a loop structure in the flow of data as shown in Figure 3.5.1 The data go through the loop only twice. In the 1-st loop, the CSF module uses the radar reflectivity factor, `zFactorNPCorrected`, processed by the VER module. In the 2-nd loop, the CSF module uses the updated `zFactorNPCorrected` for the detection of BB in the V-method and for the H-method processing and also uses the attenuation corrected radar reflectivity factor, `zFactorCorrected` (`zFactorFinal` in V07), processed by the SLV module for the determination of convective type in the V-method.

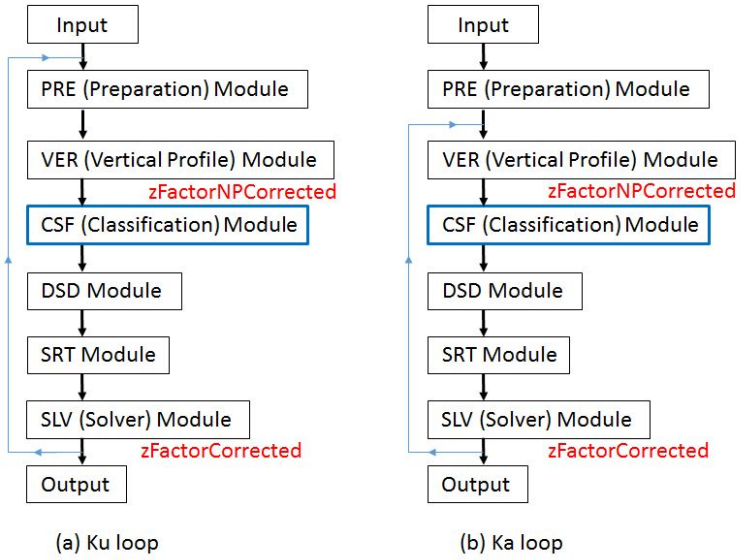


Figure 3.5.1: Loop structures of L2 data flow (single frequency algorithms only)

(a) **V-method**

In the V-method, stratiform rain is detected first.

- When BB is detected, the rain type is basically stratiform. There is an exception, however, that when the following two conditions are satisfied, it is convective even though BB is detected:

  (1) $Z_{\text{max\_below}} >$ `conv_th_BB` (=46 dBZ for both Ku- and Ka-band),

  (2) $Z_{\text{max\_below}} > Z_{\text{BBPeak}}$, where $Z_{\text{max\_below}}$ is the maximum value of `zFactorCorrected` between the height `H_BBBottom` − (about) 0.375 km and the clutter free bottom, and Z_{BBPeak} is the maximum value of `zFactorCorrected` in the BB. `H_BBBottom` is the height of the BB bottom.

- When BB is not detected, the rain type is convective if the following condition is satisfied. Otherwise, the rain type is other.

  (1) The radar reflectivity factor, `zFactorCorrected`, in the valid range between

`rangeStormTop` and `binClutterFreeBottom` exceeds a threshold, conv_th, which is 40 dBZ for both Ku-band and Ka-band.

NOTE: When BB is not detected, but if the Ku- or Ka-band storm top is very high (> 15 km, for example), then the rain type may be convective but not other. (Not implemented: a future issue.)

(b) **H-method**

In the H-method, horizontal texture of the maximum value of `zFactorNPCorrected` in the rain region along each radar beam, $Z_{\max}$, is examined. A modified University of Washington convective/stratiform separation method is applied to the horizontal extent of $Z_{\max}$. When one of the following conditions is satisfied, the considering pixel is a convective center:

(1) $Z_{\max}$ in the pixel exceeds a convective threshold (40 dBZ for Ku and Ka), or

(2) $Z_{\max}$ in the pixel stands out against those in the surrounding area.

Rain type of the convective center is convective, and the rain type of pixels adjacent to the convective center is convective. If the rain type is not convective, and if $Z_{\max}$ is not small enough to be considered as noise, the rain type is stratiform. If $Z_{\max}$ is very small, being almost identical to noise, the rain type is other.

(c) **Shallow rain and small cell-size rain**

Detection of shallow rain is also made independently of the above mentioned methods of rain type classification. When the following condition is satisfied, it is judged as shallow rain, which will be marked by an internal flag:

$$\texttt{heightStormTop} < \texttt{heightZeroDeg} - \text{margin}$$

where margin is currently 1000 m.

Shallow rain is separated out into shallow isolated and shallow non-isolated by examining the horizontal extent of shallow rain. In the rain type unification, both shallow isolated and shallow non-isolated are classified as convective. (It should be noted that in the TRMM PR rain type classification algorithm 2A23 V7, all the shallow isolated is convective, but shallow non-isolated can be either stratiform or convective.)

Detection of rain having small cell size is also made independently. The rain having small cell size is classified as convective in the unification of rain type.

(d) **Unification of rain type by the single frequency method**

Rain types by the V-method and by the H-method are unified, and the CSF module outputs

the unified rain type, which consists of three major categories: stratiform, convective and other. The unified rain type is written to the flag `typePrecip`.

Unification of single frequency rain types is made in the following way.

(1) When a BB is detected by the V-method, rain type is basically stratiform (but is convective if Z close to the surface is exceptionally strong). When a BB is detected, the unified main type follows the V-method decision (and the H-method decision does not affect the unified main type).

(2) When a BB is not detected, and Z examined by the V-method is strong enough to be convective, the unified main type is convective (and the H-method decision does not affect the unified main type).

(3) If not stratiform nor convective, rain type is other by the V-method.

(4) The other type by the V-method is further classified into one of 3 main unified types by using the H-method decision.

The above (1) - (4) form the foundation of the unification of single frequency rain types.

		H-method		
		Stra.	Conv.	Other
V-method	Stra. +BB	Stra. +BB	Stra. +BB	Stra. +BB
	Conv.+BB Conv.	Conv.+BB Conv.	Conv.+BB Conv.	Conv.+BB Conv
	Other	Stra.	Conv.	Other

Figure 3.5.2: Single frequency rain type unification (basic part).
When the type is stratiform or convective by the V-method, this result is respected by the H-method. On the other hand, the type "other" by the V-method is further classified into stratiform, convective, or other by the the H-method.

The following considerations are also added to the above mentioned basic part.

(5) In the above unification, there are exceptions as follows:

(i) All the shallow rain is convective in the unified main type except for other.

(ii) All the small cell-size rain is convective in the unified main type except for other. (In the Ku-band processing, some small cell-size rain is reclassified as stratiform by the slope method but its possibility would be very small.)

(iii) When HIP is detected, the unified main type is convective unless a BB is detected. (If a BB is detected and the V-method type is stratiform, the unified main type is stratiform.)

(6) For Ku-band data, some (unified) convective type is reclassified to (unified) stratiform type by the slope method. Details of the slope method are given in the next subsection.

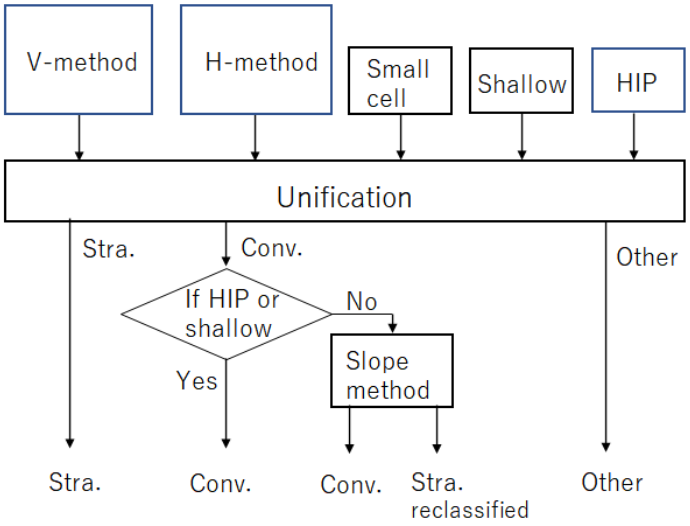


Figure 3.5.3: Detailed flow of single frequency rain type unification.
The slope method is included in the Ku-band processing only. (When the unification result shown below the very wide narrow box is convective, there is not any further processing in the case of Ka-band.)

Figure 3.5.3 shows the whole structure of the Ku-band single frequency rain type unification. The structure of the Ka-band single frequency rain type classification is similar to that of Ku-band, but the Ka-band structure does not have the slope method part.

The unified rain type is written to the flag `typePrecip`. There are many rain type subcategories depending on the combination of V- and H-method results. The unified rain type is expressed by eight digit numbers as follows:

`typePrecip` =	1xxxxxxx	for stratiform rain,
	2xxxxxxx	for convective rain,
	3xxxxxxx	for other rain,

where xxxxxxx varies depending on subcategories. Details are given in Section 3.5.9. In a computer program, the three main categories of unified rain type can be obtained easily from `typePrecip`. Let `Type` be the integer variable for the main category, for example, then the main category can be obtained by one of the following codes:

`Type` = `typePrecip` / 10000000 (in Fortran),
`Type` = `typePrecip` / 10000000; (in C)

(e) **Reclassification of convective type to stratiform type by the "slope method"**

In 2017, after examining hundreds of cross-sections of V05 Ku-band data, R. Houze and S. Brodzik of the University of Washington (UW) recognized that many convective pixels in V05 Ku-band output were classified incorrectly. They developed a method for solving the problem by examining the slope of the Ku-band Z profile above the 0 degrees C isotherm and reclassifying the type from convective to stratiform when certain conditions were satisfied. Since the method developed by the UW examines the slope of Z, we refer to it as the "slope method" (Houze and Brodzik, 2017). The V06 CSF module adopts the "slope method" for reclassifying some convective type to stratiform type at the end of the Ku-only rain type unification. The V06X and V07 CSF modules also adopt the "slope method" in the Ku-only processing. (Since the reclassified results are handed down to the DPR CSF processing, the CSF DPR module does not carry out a duplicate (i.e., redundant) reclassification using the "slope method.")

The reclassification of convective type to stratiform type by the "slope method" is carried out by using the Z factor corrected for attenuation by the nonprecipitation (NP) substances, `zFactoNPCorrected`, and the procedure is as follows:

(1) For each beam, If the unified Ku-only rain type is convective and the shallow_rain_type flag is not set, find the maximum value of Ku-band `zFactorNPCorrected` (`max_refl`) in the BB search window (from `rangeZeroDeg` $-$ 8 to `rangeZeroDeg` $+$ 16 with a 125 m range interval bounded by the highest BB height being 6.5km). At this stage, the CSF module uses the BB search window in order to keep the consistency with the detection of BB. However, further restriction is imposed later (see (3)).

(2) Starting at the `max_refl` point, move up the ray bin by bin to the last point which satisfies this condition:

$|$(`max_refl` $-$ `zFactorNPCorrected`)$|$ <= `REFL_MAX_OFFSET` (currently defined as 2 dBZ)

This is called the reference_point.

(3) If the height of the reference_point is lower than `MAX_BB_HT` (currently defined as 5km), continue. Otherwise, no reclassification will occur for this beam.

(4) Find the comparison point along the NP corrected Z profile which is 1.5 km (`DELTA_HT_FOR_SLOPE`) above the reference_point. (If `zFactorNPCorrected` at 1.5 km above the reference_point is noise, step down the ray bin by bin until a valid value of `zFactorNPCorrected` is found. That point will become the comparison point.)

(5) Calculate the slope between the reference_point and the comparison_point (defined as |`delta_zFactorNPCorrected`|/|`delta height`|.

(6) If the slope is greater than or equal to SLOPE_THRESHOLD (currently defined as 7.5 dBZ/km), then reclassify the type as stratiform.

NOTE: All parameter values for REFL_MAX_OFFSET, MAX_BB_HT, DELTA_HT_FOR_SLOPE, SLOPE_THRESHOLD were determined empirically. While they have worked well in testing, they may need to be adjusted in the future.

(f) **Problem of V06 rain type reclassification by the slope method – Solved in V06X**
Immediately after the release of V06 data, it was found that the near surface rainfall rate of the reclassified stratiform type can become very large in rare cases. For stratiform rain, a very large rainfall rate is unacceptable. This problem has occurred mainly because the V06 "slope method" examines the Z profile aloft, but not Z close to the surface.

The V06X CSF module has solved this problem by introducing a kind of filter at the end of the "slope method" so that if the near surface rainfall rate of the reclassified stratiform type exceeds a threshold, TH_slope, then the reclassified stratiform type is put back to the convective type. In V06X (and in V07), the following threshold value is used: TH_slope=20 mm/h (Awaka and Brodzik, 2019).

Though the threshold is set to be 20 mm/h, the near surface rainfall rate of the reclassified stratiform rain is found to exceed 20 mm/h in some rare cases. This occurs because the Level 2 (L2) single frequency Ku-band algorithm has a loop structure as shown in Figure 3.5.1.

Figure 3.5.4 shows the single frequency Ku-band loop in a concise form (on the left), which is the same as that shown in Figure 3.5.1, and in a detailed form (on the right). Note that the threshold is set for the 1-st loop SLV output, but the near surface rainfall rate available to the users belongs to the 2-nd loop SLV output. This is the reason why the near surface rainfall rate of the reclassified stratiform rain can exceed 20 mm/h in some rare cases; the CSF module can handle the 1-st loop SLV output but not the 2-nd loop SLV output.

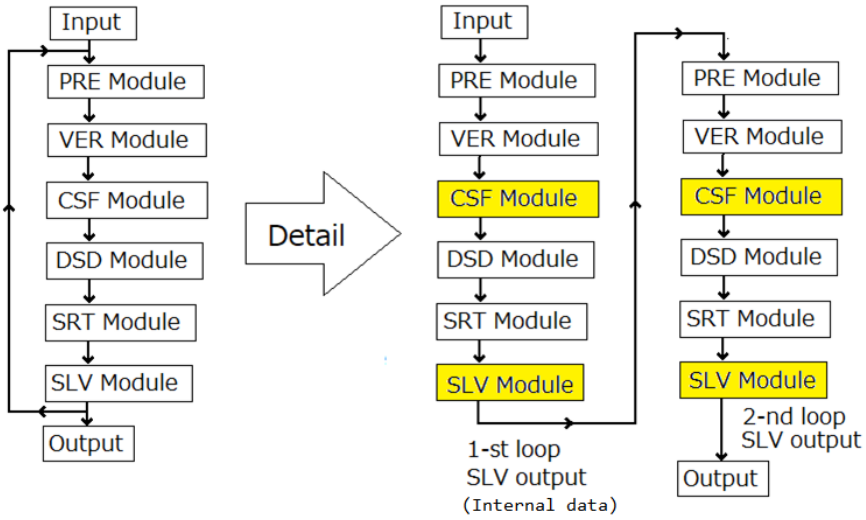


Figure 3.5.4: Loop structure detail of the single frequency Ku-band L2 algorithm.
The CSF module can use the 1-st loop output of the SLV module, but can not use the 2-nd loop output of the SLV module. The near surface rainfall rate which is available to the users is produced by the 2-nd loop SLV module. This is the reason why, in some rare cases, the near surface rainfall rate of the reclassified stratiform rain, which is obtained by the 2-nd loop SLV module, can exceed the threshold set for the 1-st loop SLV output by the CSF module.

3.5.4.2 Dual frequency processing

In the dual frequency processing, a newly developed DFR_m method is used for the detection of melting layer (ML) and rain type classification. Details of DFR_m method is given in Section 3.5.4.2.3. The concept of ML has a meaning wider than that of BB. In other words, BB which appears in the stratiform rain is a subset of ML. Since BB is a subset of ML, the DFR_m method can be used for the detection of BB. In the dual frequency processing, the DFR_m result and the Ku-only result are merged.

3.5.4.2.1 Detection of bright band (BB) by the dual frequency method

In the dual frequency method, BB detected by the Ku-only method and ML by the DFR_m method are compared. Firstly, ML detected by the DFR_m method is regarded as BB. Together with BB detected by the Ku-only method, the median of the height of BB is computed in one scan of data. When the considering height deviates from the median to a large extent, it is judged that the considering BB is not a true BB and the BB decision is rejected.

If BB is detected by both the Ku-only and the DFR_m methods at a given angle bin, it is judged that the height of BB is the height which is closer to the median of BB height.

When BB is thus detected, properties of BB such as height, width, etc., for matched beam data are then determined by using the profile of Z at Ku-band only, because Z at Ku-band shows a clear BB peak but in most cases Z at Ka-band doesn't.

Since the concept of ML is wider than that of BB, and the upper and lower bounds of ML are different from those of BB in general, the following two items are added to the output of dual frequency CSF module:

- `binDFRmMLTop`: Range bin number of ML top.

- `binDFRmMLBottom`: Range bin number of ML bottom.

3.5.4.2.2 Rain type classification by the dual frequency method

As is shown in the next Section 3.5.4.2.3, the DFR_m method classifies rain into the following three types: stratiform, convective, and transition, the last of which is a new rain type. The rain type by the DFR_m method is merged with the single frequency Ku-band rain type, and the dual frequency CSF module outputs unified rain type which consists of the following three major types: stratiform, convective and other.

Figure 3.5.5 shows the diagram for unifying rain types by the dual frequency CSF module. Rain types by the DFRm method and those of the Ku-band single frequency are unified. After the unification shown in the figure, there follows a further modification to the type by consulting the detection of HIP and the detection of winter convection as explained later.

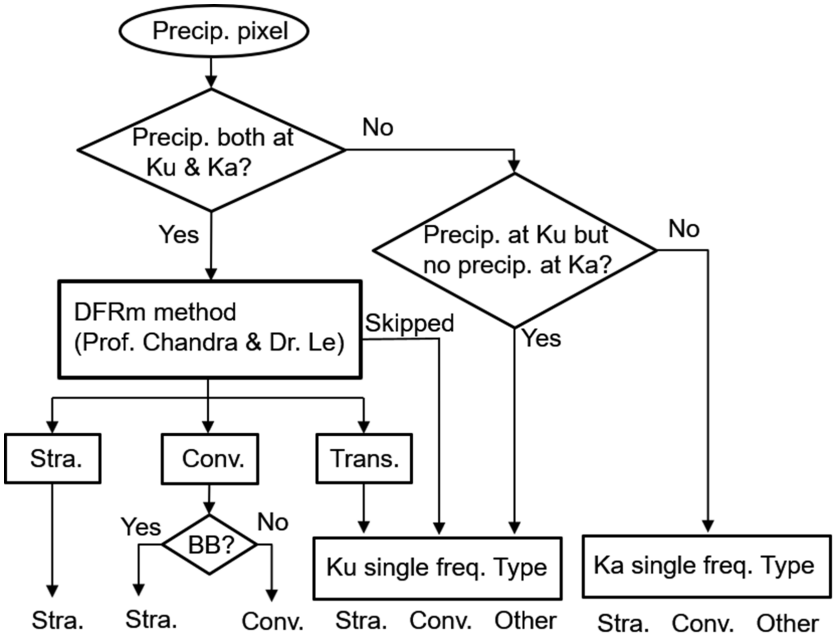


Figure 3.5.5: Unification of rain types by the dual frequency CSF module.

Some features of the unification of dual frequency rain types are as follows:

(1) When the DFR_m rain type is convective or stratiform, the DFR_m decision is respected and the DFR_m rain type is the unified rain type if a BB is not detected. If a BB is detected,

however, the unified rain type is basically stratiform.

(2) When the DFR_m rain type is transition, the single frequency Ku-band rain type is the unified rain type.

(3) When the DFR_m processing is skipped, the single frequency Ku-band rain type is the unified rain type.

(4) Though not shown in the figure, if HIP is detected or if winter convection is detected, some stratiform type is changed to convective type as explained later.

The unified rain type is expressed by eight digit numbers as follows:

`typePrecip` =	1xxxxxxx	for stratiform rain,
	2xxxxxxx	for convective rain,
	3xxxxxxx	for other rain,

where xxxxxxx varies depending on sub-categories. (See Section 3.5.9 for details.) In a computer program, the three main categories of unified rain type can be obtained easily from `typePrecip`. Let Rtype be the integer variable for the main category, for example, then the main category can be obtained by one of the following codes:

Rtype = `typePrecip` / 10000000 (in Fortran),
Rtype = `typePrecip` / 10000000; (in C).

The rain type by the DFR_m method is expressed in the 2nd digit of typePrecip. Hence the rain type by the DFR_m method, if we denote it by DFRmRtype for example, can be obtained as follows:

DFRmRtype = MOD(`typePrecip`, 10000000) / 1000000 (in Fortran),
DFRmRtype = (`typePrecip` % 10000000) /1000000; (in C).

When rain exists, DFRmRtype takes one of the following values:

1: stratiform,
2: convective,
4: transition,
5: when stratiform is changed to convective by the extended DFRm method explained in Section 3.5.5.2,
8: DFR_m method skipped at Part B in Figure 3.5.8 (shown later in Section 3.5.4.3).
9: DFR_m method skipped at Part A in Figure 3.5.8 (shown later in Section 3.5.4.3).

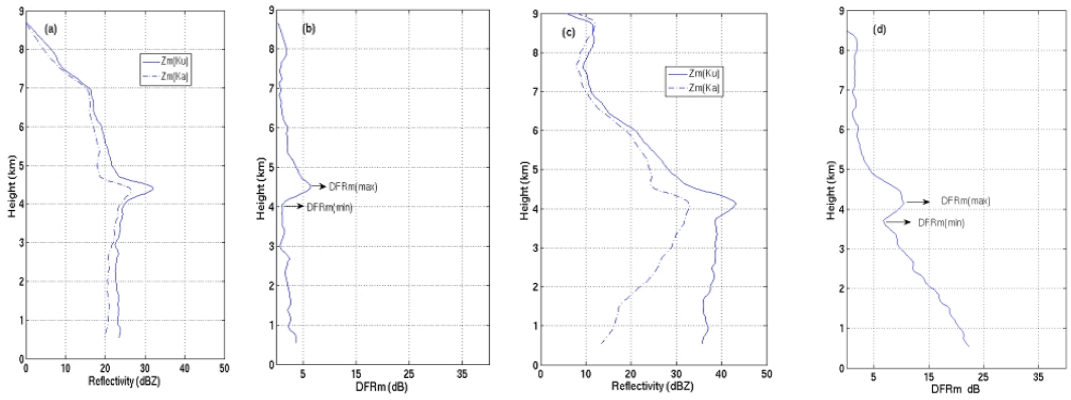
3.5.4.2.3 Detailed description of DFR_m method

Dual-frequency precipitation radar DPR offers Ku and Ka band dual-frequency observations along the vertical profile which allow us additional information to investigate the micro-physical

properties using the difference between two frequency observations or so-called measured dual-frequency ratio ($\mathrm{DFR_m}$) defined as

$$\mathrm{DFR_m} = 10\log_{10}(Z_m(K_u)) - 10\log_{10}(Z_m(K_a)) \tag{3.5-1}$$

where Z_m is measured reflectivity. $\mathrm{DFR_m}$ profile holds rich information to assist in precipitation type classification and melting layer detection. There are two main functions of $\mathrm{DFR_m}$ method in classification module: (1) precipitation type classification, and (2) melting layer detection. Correspondingly, there are two models developed for dual-frequency profile classification method. Figure 3.5.6 shows typical vertical profiles of reflectivity and $\mathrm{DFR_m}$ for stratiform and convective rain. These profiles are observations from airborne precipitation radar.



(a) (b) (c) (d)

Figure 3.5.6: Typical vertical profiles for stratiform ((a), (b)) and convective ((c), (d)) rain. (a) and (c): Measured reflectivity at Ku and Ka band; (b) and (d): $\mathrm{DFR_m}$. $\mathrm{DFR_m}$(max) and $\mathrm{DFR_m}$(min) marked on (b) and (d) are local maximum and minimum values, respectively.

3.5.4.3 Precipitation type classification model

Precipitation type classification model classifies stratiform, transition, and convective rain types. The main parameter used in the model is $\mathrm{DFR_m}$ and its vertical variation. In order to quantify the features of $\mathrm{DFR_m}$, a set of $\mathrm{DFR_m}$ indices are defined. Let $V1$ be

$$V1 = \frac{\mathtt{DFR}_{\mathrm{ml}}(\max) - \mathtt{DFR}_{\mathrm{ml}}(\min)}{\mathtt{DFR}_{\mathrm{ml}}(\max) + \mathtt{DFR}_{\mathrm{ml}}(\min)} \tag{3.5-2}$$

$\mathrm{DFR_m}$(max) and $\mathrm{DFR_m}$(min) are shown in Figure 3.5.6. $\mathtt{DFR}_{\mathrm{ml}}$ used in Eq. (3.5-2) means $\mathrm{DFR_m}$ in linear scale. Let V2 be the absolute value of the mean slope for $\mathrm{DFR_m}$ below the local minimum point

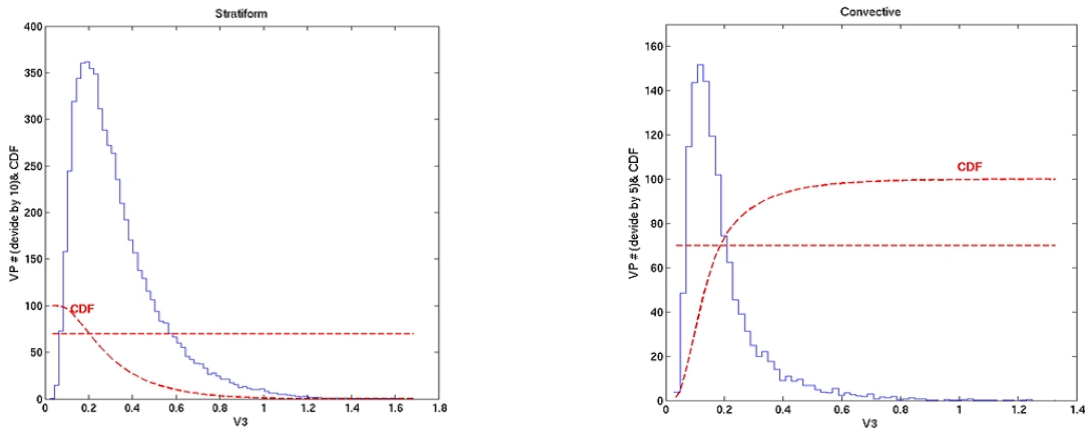
$$V2 = \mathrm{abs}(\mathrm{mean}(\mathtt{DFR}_{\mathrm{m}}\ \mathrm{slope})) \tag{3.5-3}$$

Both $V1$ and $V2$ are normalized values and not dependent on the height or depth of the melting layer. $V1$ values are normally larger for stratiform rain than for convective rain and $V2$ values

are larger for convective than for stratiform rain. To further enlarge the difference between stratiform and convective rain types, a third index $V3$ is defined as

$$V3 = \frac{V1}{V2} \tag{3.5-4}$$

The index $V3$ is an effective parameter and provides a separable threshold for performing precipitation type classifications. Extensive statistic studies are performed on index $V3$ using both airborne radar data and GPM real data. Cumulative density function (CDF) of index $V3$ is calculated for stratiform and convective database separated using Ku only classification algorithm. Separable thresholds of $C1$ and $C2$ can be found on index $V3$ for stratiform and convective rain types with 70% of CDF. In other words, for stratiform rain: $V3 > C2$; convective rain: $V3 < C1$; transition type: $C1 \leq V3 \leq C2$. $C1$ is smaller than $C2$. "Transition" type is neither a stratiform, nor a convective rain type, but a type transitioning from stratiform to convective rain. Figure 3.5.7 shows histogram of index $V3$ and its cumulative density function calculated using total of 73 storms data with 121859 vertical profiles. Table 3.5.1 shows the thresholds of $C1$ and $C2$ used in the current version. In the future, further adjustment of $C1$ and $C2$ is needed. Figure 3.5.8 illustrates the block diagram for precipitation type classification model that used in the current version.



(a) (b)

Figure 3.5.7: Histogram of DFR_m index $V3$ with CDF line.
(a) stratiform rain and (b) convective rain calculated using total of 121859 vertical profiles from GPM real data.

Table 3.5.1: $C1$ and $C2$ thresholds used in precipitation type classification model

70% CDF	121859 vertical profiles of GPM data
$C1$	0.18
$C2$	0.20

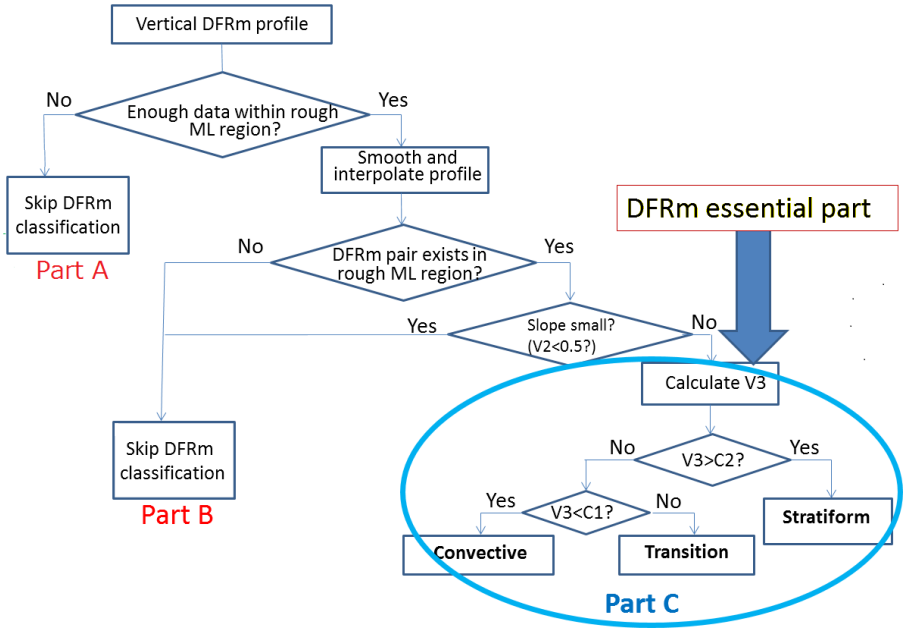


Figure 3.5.8: Block diagram of precipitation type classification model.

3.5.4.4 Melting layer detection model

Melting layer detection model detects melting layer top and bottom height for each vertical profile. The main parameter used in the model is $\mathrm{DFR_m}$ profile and its vertical variation. When $\mathrm{DFR_m}$ pair as shown in Figure 3.5.6 is detected, the melting layer top is defined as the height at which the slope of the $\mathrm{DFR_m}$ profile hits a peak value. The melting layer bottom is defined as the height at which the $\mathrm{DFR_m}$ profile has a local minimum value. The dashed lines in Figure 3.5.9 show an illustration of melting layer top and bottom detected.

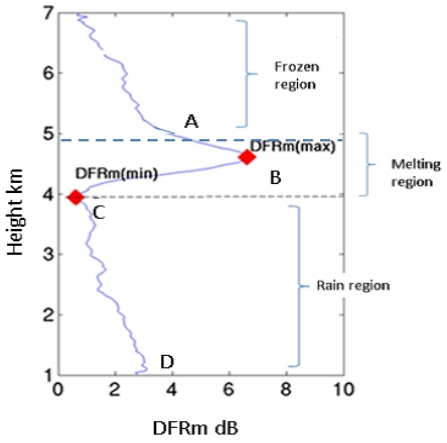
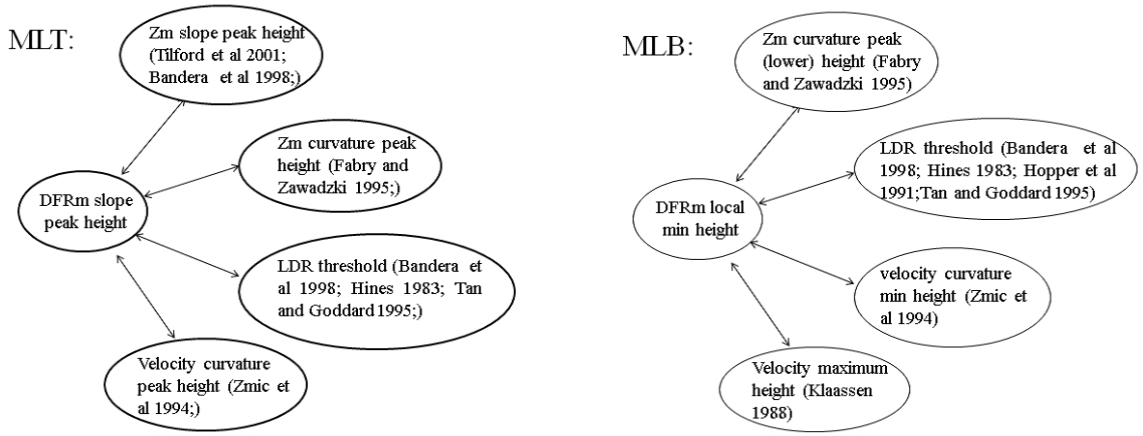


Figure 3.5.9: Schematic plot of DFR_m profile with key points A, B, C, and D. Point A: slope of DFR_m has peak value. Point B: local maximum of DFR_m. Point C: local minimum of DFR_m. Point D: DFR_mvalue near surface.

The criteria described above have been compared with other existing criteria in the literature using different radar parameters. Tilford et al. (2001) used the gradient of reflectivity (Z_m) to detect the bright band top and bottom for stratiform rain type. The linear depolarization ratio (LDR) has been pointed out by many researchers as an important signature in melting phase detection, with certain thresholds determined for different hydrometeor particles (Smyth et al., 1998; Bandera et al., 1998; Tan and Goddard, 1995; Hines, 1983). Typical vertical profiles of reflectivity as well as the corresponding velocity for stratiform and convective type were extensively studied by Fabry and Zawadzki (1994). Baldini and Gorgucci (2006) mentioned that the rapid change of the hydrometeor fall velocity is an implication of the melting layer. The curvature of velocity was used by Zrnic et al. (1994) in characterizing the melting boundaries. Klaassen (1988) found that the melting bottom can be detected by maximum of velocity. Figure 3.5.10 shows a schematic plot of the comparisons between melting layer detection criteria in DFR_m method and other existing criteria. The comparison results using airborne radar data are listed in Table 3.5.2. From the table, estimations from the DFR_m method match best with velocity based criteria with normalized bias of 1.3% and 2.2% for melting layer top and bottom respectively. The DFR_m method also compares well with the LDR criteria, with a -28 dB threshold, the bias between these two criteria is around -2.8%. Details can be found in Le and Chandrasekar (2012). Figure 3.5.11 shows a profile comparison between DFR_m method and LDR as well as velocity criteria using APR-2 observations. Figure 3.5.12 is the block diagram of melting layer detection model used in the current version of DFR_m classification method.



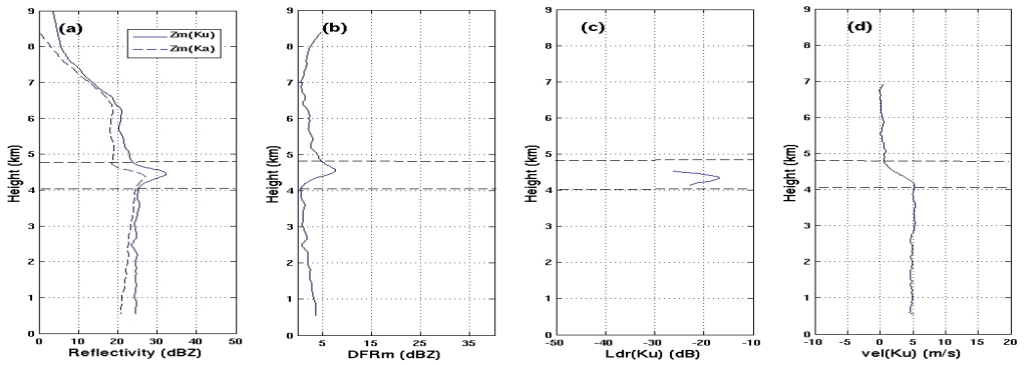
(a) (b)

Figure 3.5.10: Schematic plots of some current criteria for melting layer boundaries detection and their possible relations with melting layer detection model used in the $\mathrm{DFR_m}$ method. (a) melting layer top criteria; (b) melting layer bottom criteria.

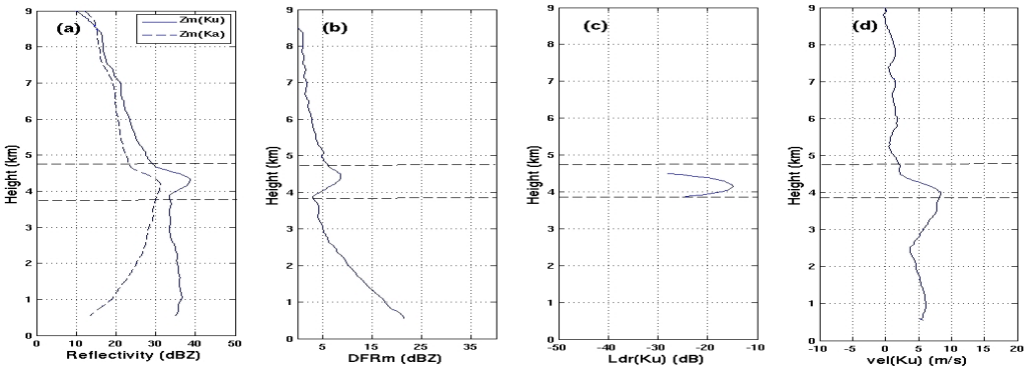
Table 3.5.2: Comparisons of melting layer boundaries between different criteria for NAMMA, GRIP and Wakasa Bay data.

	Criteria	$\mathrm{DFR_m}$ slope peak (NAMMA)		$\mathrm{DFR_m}$ slope peak (GRIP)		$\mathrm{DFR_m}$ slope peak (Wakasa Bay)	
		NB	NSE	NB	NSE	NB	NSE
Melting layer top comparison	Z_m slope peak	−2.6%	3.6%	−2.5%	3.6%	−4.9%	6.6%
	Z_m curvature peak	1.6%	3.3%	1.5%	3.0%	2.8%	5.2%
	LDR	−2.8%	4.5%	−3.3%	4.2%	−6.0%	7.2%
	Velocity curvature peak	−1.3%	3.6%	−1.4%	3.7%	−1.9%	5.6%

	Criteria	$\mathrm{DFR_m}$ slope peak (NAMMA)		$\mathrm{DFR_m}$ slope peak (GRIP)		$\mathrm{DFR_m}$ slope peak (Wakasa Bay)	
		NB	NSE	NB	NSE	NB	NSE
Melting layer bottom comparison	Z_m curvature peak	4.3%	5.5%	3.7%	5.0%	4.3%	6.9%
	LDR	4.5%	5.9%	4.0%	5.4%	5.4%	11.2%
	Velocity curvature min	2.2%	4.9%	1.7%	4.4%	−0.08%	7.0%
	Velocity max	1.6%	5.9%	1.9%	4.3%	−2.6%	13.9%



(a)



(b)

Figure 3.5.11: Sample profile from airborne radar data. Top row (stratiform rain): (a) Z_m(Ku) and Z_m(Ka); (b) $\mathrm{DFR_m}$; (c) LDR; (d) Velocity;Dashed lines are melting layer top and bottom decided by the $\mathrm{DFR_m}$ method. Bottom row (convective rain): (a) to (d) are the same as in top row.

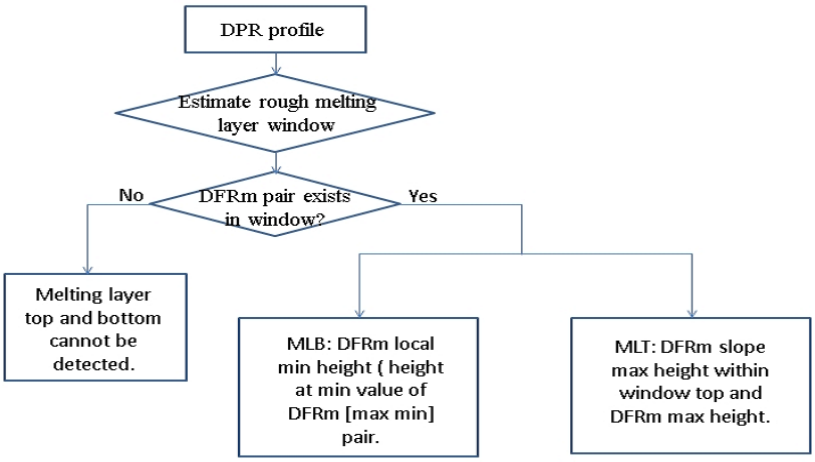


Figure 3.5.12: Block diagram of melting layer detection model for $\mathrm{DFR_m}$ method.

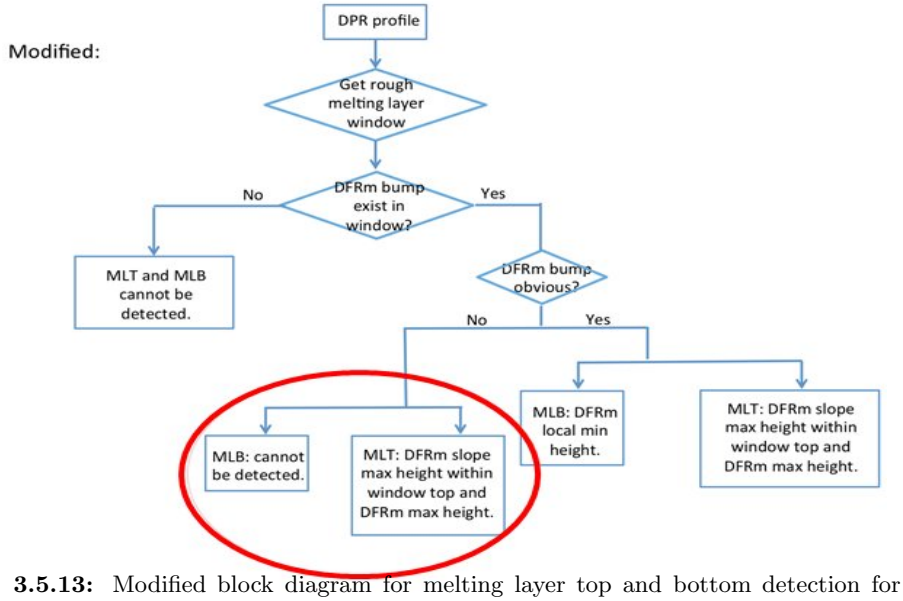


Figure 3.5.13: Modified block diagram for melting layer top and bottom detection for dual-frequency classification algorithm.

3.5.4.4.1 Modification on melting layer detection of dual-frequency classification model

One function of the dual-frequency classification module is to detect melting layer on a profile basis. Currently, the detection of melting layer top and bottom is done simultaneously. If either the top or bottom is not detected, both melting layer top and bottom are not available. In order to increase the detectability, we modify the current version of algorithm to separate the melting layer top and bottom detection to make it independent. This is reasonable since melting layer top is relatively easier to be detected than melting layer bottom.

In the modification, we have threshold1 and threshold2. When DFRm_diff > threshold2, melting layer top and melting layer bottom are both detected. When DFRm_diff > threshold1 and < threshold2, then only melting layer top is detected, not melting layer bottom. When DFRm_diff < threshold1 or not exist, neither melting layer top nor bottom is detected. The threshold1 and threshold2 are 0.5 dB and 2 dB in the modified version. "FlagML" information is temporarily stored at "`qualityTypePrecip`" for the purpose of testing the quality of melting layer detection. Independent Flag information (for both MS and HS) should be added to the output of future toolkit. Figure 3.5.13 illustrates the modified flow chart of the melting layer detection algorithm in the dual-frequency classification module.

Figure 3.5.14 shows a section (Scan range: 3032–3070) of the GPM DPR orbit #24981, where precipitation was captured over Buffalo, New York on 22 July 2018. Figure 3.5.14 (a) is Zmku at 2km, while 2 (b) is Zmka at 2 km.

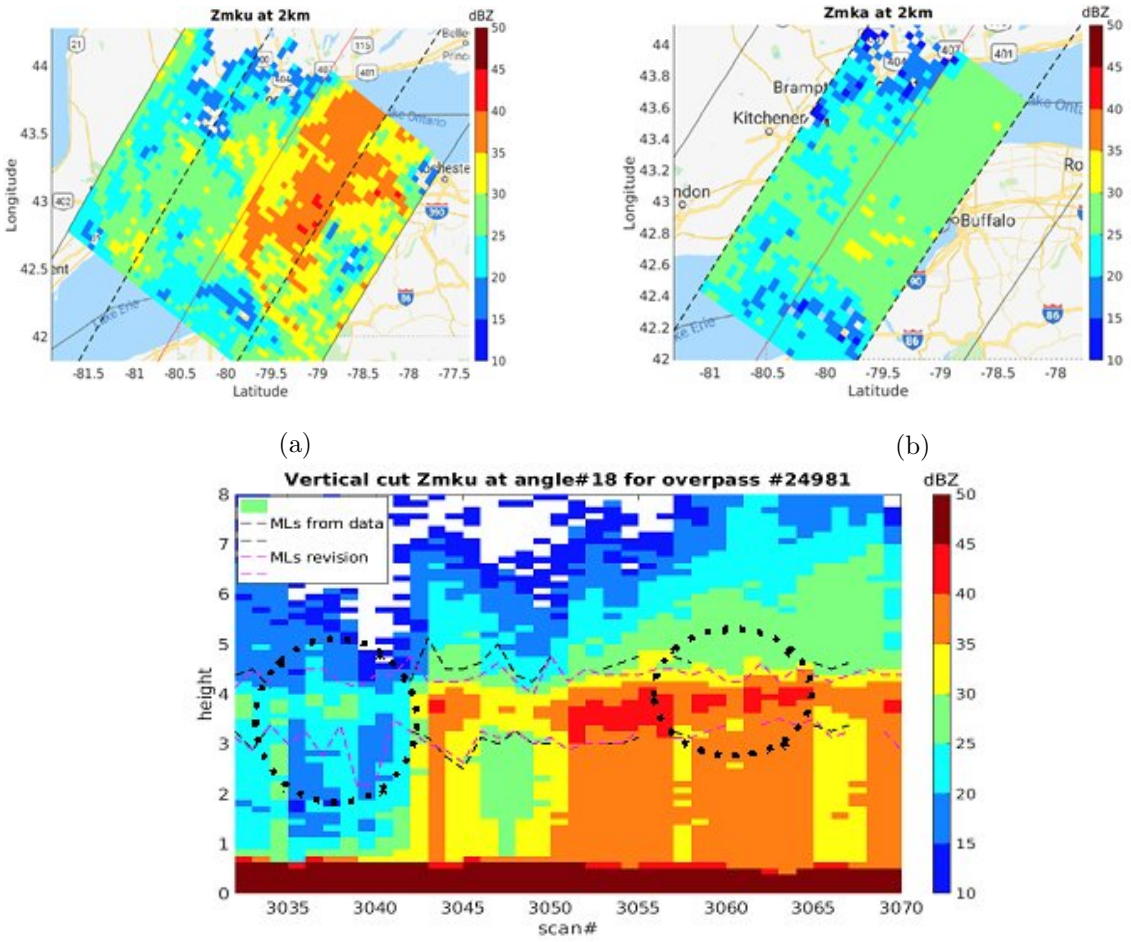


Figure 3.5.14: GPM DPR orbit #24981 with scan range of (3032–3070). (a) Zmku at 2km. (b) Zmka at 2 km. (3) Vertical cut at angle bin #18 of Zmku.

Sample vertical cut at angle bin #18 is shown in Figure 3.5.14 (c). Black dashed lines are melting layer top and bottom before modification. Pink dashed lines are melting layer top and bottom after modification. For both top and bottom, more detections are made after modification of the code, some of which are circled by the dotted lines. Overall, the percentage of melting layer detection in the scan range increases from 64.2% to 90.3% for top and from 64.2% to 86.5% for bottom.

Figure 3.5.15 illustrates the count of rain types and melting layer top detection for one day of GPM orbits on 01 August 2018. Column of "V06" indicates the results before modification and column of "new" is for after modification. Changes of stratiform/convective counts occur in the inner swath only. Those small differences between V06 and updated codes occur due to the change of threshold1 (from 2.0 to 1.5). However, melting layer top detection counts increase

largely for both MS and HS condition. For MS, during one-day orbits, the count increases from 24,262 to 35,918, with increase percentage of 48%. For the HS condition, the increase percentage is around 63%. This is a big improvement on the melting layer detection part of the dual-frequency classification module. More extensive analysis will be performed in the near future.

	V6	new
stra (MS)	132,549	132,468
conv (MS)	33,688	33,771
MLtop (MS)	24,262	35,918
stra (HS)	120,313	120,408
conv (HS)	17,880	18,148
MLtop (HS)	14,020	22,938

Figure 3.5.15: Counts of stratiform, convective rain types as well as melting layer top detection for MS and HS on one-day GPM DPR orbits of 01 August 2018.

3.5.5 Other features in the CSF module

Most of the features described in this section arise from the advantage of dual frequency data processing.

3.5.5.1 flagHeavyIcePrecip

This flag indicates that Heavy Ice Precipitation (HIP) is detected in the upper region of storm where the temperature is lower than −10 degrees in Centigrade. Here, HIP is defined as precipitation consisting of ice particles which produces a large measured Z factor (Z_m) and/or produces a large $\mathrm{DFR_m}$. Detection of HIP, proposed by Dr. Iguchi, is made by examining Ku Z_m and Ka Z_m in the single frequency modules and by additionally examining $\mathrm{DFR_m}$ in the dual frequency module (Iguchi et al. 2018). The following values of `flagHeavyIcePrecip` are devised to distinguish the decisions by which HIP is detected (see Figure 3.5.16). Default value: `FlagHeavyIcePrecip` = 0 (HIP not detected, including the case of missing).

Ku-band decision (Ku FS in V07 & V06X, Ku NS in V06):

If KuPR's $Z_m > 45$ dBZ, then `flagHeavyIcePrecip` = 12 = `0x0C`

Else if KuPR's $Z_m > 40$ dBZ, then `flagHeavyIcePrecip` = 8 = `0x08`

Else if KuPR's $Z_m > 35$ dBZ, then `flagHeavyIcePrecip` = 4 = `0x04`

Ka-band decision (Ka FS in V07 & V06X, Ka MS in V06):

If KaPR's $Z_m > 40$ dBZ, then `flagHeavyIcePrecip` = 3 = `0x03`

Else if KaPR's $Z_m > 35$ dBZ, then `flagHeavyIcePrecip` = 2 = `0x02`

Else if KaPR's $Z_m > 30$ dBZ, then `flagHeavyIcePrecip` = 1 = `0x01`

DPR decision:

The sum of (A) the above Ku-band decision value, (B) the above Ka-band decision value, and (C) the following DFR_m decision value:

KuPR's $Z_m > 27$ dBZ and $DFR_m > 7$ dB, then `flagHeavyIcePrecip` = 16 = `0x10`

Therefore, (A)+(B)+(C) is assigned to the DPR `flagHeavyIcePrecip` (V07 & V06X full swath, V06 inner swath).

Note: In V06, the flag value for the DPR outer swath is the same as that for the Ku NS outer swath.

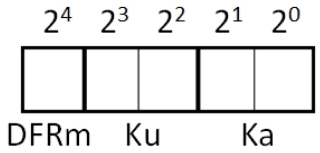


Figure 3.5.16: Bit assignment of `flagHeavyIcePrecip`

In V07 and V06X, the following items give some information about ranges where the HIP is detected:

(1) `binHeavyIcePrecipTop`, which shows the range bin corresponding to the top height where HIP is detected,

(2) `binHeavyIcePrecipBottom`, which shows the range bin corresponding to the bottom height where HIP is detected,

(3) `nHeavyIcePrecip`, which shows the number of range bins where HIP is detected.

In the C language, those new items are defined as follows:
In the Ku- and Ka- band single frequency standard output:

```
short binHeavyIcePrecipTop[49];
short binHeavyIcePrecipBottom[49];
unsigned char nHeavyIcePrecip[49];
```

In the DPR dual frequency standard output:

```
short binHeavyIcePrecipTop[49][3];
short binHeavyIcePrecipBottom[49][3];
unsigned char nHeavyIcePrecip[49][3];
```

with the second index showing the distinction among single frequency Ku, single frequency Ka, and dual frequency DPR quantities. For example, in the case of `binHeavyIcePrecipTop`

`binHeavyIcePrecipTop[i][0]` shows the Ku decision value for the i-th angle bin,

`binHeavyIcePrecipTop[i][1]` shows the Ka decision value for the i-th angle bin,

`binHeavyIcePrecipTop[i][2]` shows the DPR decision value for the i-th angle bin,

where the C index i runs from 0 to 48.

3.5.5.2 Detection of winter convection

Detection of winter convection is made by examining $\mathrm{DFR_m}$ at the storm top and its vicinity. The basic idea is that a large $\mathrm{DFR_m}$ near the storm top in winter may indicate the convective nature of precipitation. Since $\mathrm{DFR_m}$ is examined, the method in this section can be called an 'extended $\mathrm{DFR_m}$ method'. It is found that the above `flagHeavyIcePrecip` plays an important role in the detection of convective precipitation in winter. When the following conditions are satisfied, it is judged that the winter convection is detected by the extended $\mathrm{DFR_m}$ method (See DFRmRtype in Section 3.5.4.2.2).

(1) The estimated 0°C height is lower than or equal to 1 km. (Winter condition.)

(2) `flagHeavyIcePrecip` > 0 is found near the considering pixel. Here, the term "near" means the region within -3 to $+3$ scans measured from the considering pixel's scan number (see Figure 3.5.17).

(3) Examine the $\mathrm{DFR_m}$ of 5 consecutive range bins starting from storm top towards the downward direction. At three or more range bins, the following condition is satisfied:

$$\mathrm{DFR_m}(\text{at } x) \; > \frac{2.5 - 0.8}{5.0} x + 0.8$$

where x is the relative range bin number counted downward from the storm top.

(4) If Z_m at any of the examining range bins exceeds 35 dBZ, the process of detecting winter convection is skipped because such a large Z_m near the storm top may mean the existence of BB.

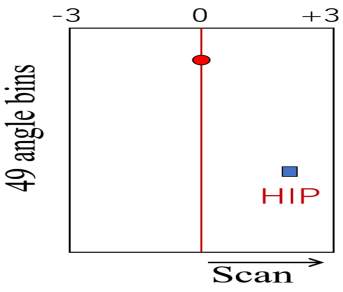


Figure 3.5.17: Nearness of the considering pixel (marked by a red circle) to the pixel of `flagHeavyIcePrecip` > 0 (marked by a blue square).

3.5.5.3 `flagAnvil`

Since the Ku-band PR turns out to have a higher capability in precipitation detection, the Ku-band PR data can be used for the detection of "Anvil". For the definition of "Anvil", we need to introduce cloud top and cloud bottom, where cloud top is identical with storm top and cloud bottom means the range bin below which no-rain bins continue to some extent. (When precipitation exists, the cloud top always exists but the cloud bottom does not always exist.) As shown in Figure 3.5.18, when the cloud bottom is 500 meters higher than the estimated 0°C height, it is defined that the "Anvil" exists. If there is no precipitation echo below the "Anvil", it is judged that the "Anvil" type 1 is detected. If precipitation echoes exist below the "Anvil", but if the no-rain interval between the cloud bottom and the top of precipitation echoes below the "Anvil" exceeds 1 km, it is judged that the "Anvil" type 2 is detected (see Figure 3.5.18(b)). The V05 CSF module detects type 1 "Anvil" and type 2 "Anvil" only. The possible values of `flagAnvil` are as follows:

`flagAnvil` =1: type 1 'Anvil' is detected (there is no rain echo below the 'Anvil'),
=2: type 2 'Anvil' is detected (there exist rain echoes below the 'Anvil'),
=0: 'Anvil' is not detected (including the case of data missing).

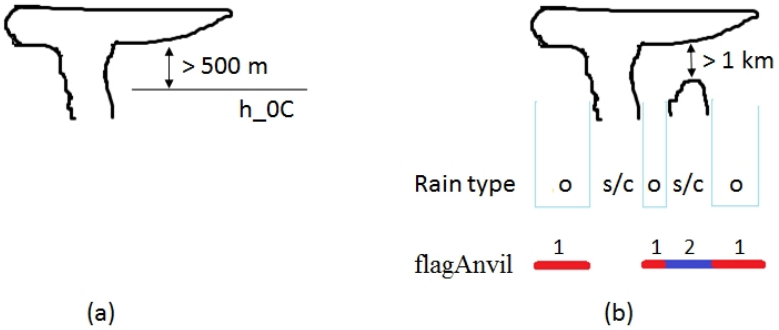


Figure 3.5.18: Definition of 'Anvil'

3.5.5.4 Surface snowfall

The V05 CSF module introduces a process which determines whether the precipitation at the surface is snow or rain. The process has been developed by Dr. Chandra's group. The process outputs the following two variables. (Though the decision is made in the CSF module, the two variables are available in the Experimental output fields of dual frequency MS.)

`surfaceSnowfallIndex`: If this index exceeds a predetermined threshold, it is judged that the surface precipitation is snow.

`flagSurfaceSnowfall`: This flag takes the following two values.

$$\texttt{flagSurfaceSnowfall} = \begin{cases} 1: & \text{surface snowfall is possible,} \\ 0: & \text{surface snowfall not detected.} \end{cases}$$

3.5.6 Snow and rain feature on measured dual-frequency ratio ($\mathrm{DFR_m}$) profile

Figure 3.5.19 (a) shows the first snow observation caught by GPM DPR with overpass #000272. On that overpass, there are snow, stratiform and convective rain precipitation within 160 scans (around 800 km) range. In the figure, A, B and C indicate the locations of snow, stratiform and convective rain. In order to study vertical profile feature for these different precipitation types, we study averaged reflectivity as well as measured dual-frequency ratio profile for snow, stratiform and convective rain. The vertical profile is calculated from linear averaging of reflectivity at corresponding heights. If data in a pixel is missing or below noise level, it is not considered in the average. Figure 3.5.19 (b) shows the averaged reflectivity profile for snow at Ku-, Ka-band and a measured dual-frequency ratio. As expected, reflectivity at Ku- band is, most of the time, below 25 dBZ. However, the difference between Ku- and Ka- band (indicated by $\mathrm{DFR_m}$) is several dBs, even when reflectivity at Ku- band remains a relatively small value. $\mathrm{DFR_m}$ values increase obviously when it comes toward the surface mainly due to the aggregation of the snow particles. Figure 3.5.19 (c) shows an averaged vertical profile for stratiform rain. The bright band is obvious from Ku- band reflectivity. Values of $\mathrm{DFR_m}$ below the melting region (or the bright band) are very small indicating that attenuation at Ka- band is very small when reflectivity at Ku- band is less than around 25 dBZ in the rain region. $\mathrm{DFR_m}$ values are quite constant with respect to height. In contrast, $\mathrm{DFR_m}$ values above melting layer is several dBs and $\mathrm{DFR_m}$ slope is noticeable, which is similar to the snow profile in Figure 3.5.19 (b). For convective rain, from Figure 3.5.19 (d), the maximum of reflectivity at Ku- band is equal to or larger than 35 dBZ, while $\mathrm{DFR_m}$ values are considerable in the rain region. Large $\mathrm{DFR_m}$ values in rain are contributed from both attenuation difference and the non-Rayleigh scattering effect, while the former factor plays a more important role. $\mathrm{DFR_m}$ slope for convective rain is large mainly due to large attenuation different change when it comes toward the surface.

One of the features for a convection storm is that it is normally formed at a higher altitude in the atmosphere (except for some warm rain or orographic rain) compared to stratiform and snow precipitation.

Based on analysis above, three ingredients of $\mathrm{DFR_m}$, maximum value of reflectivity at Ku-band, as well as storm top height become important, and they are used in the development of the algorithm to identify surface snowfall. To avoid calibration issues, we use the slope of $\mathrm{DFR_m}$ instead of $\mathrm{DFR_m}$ itself.

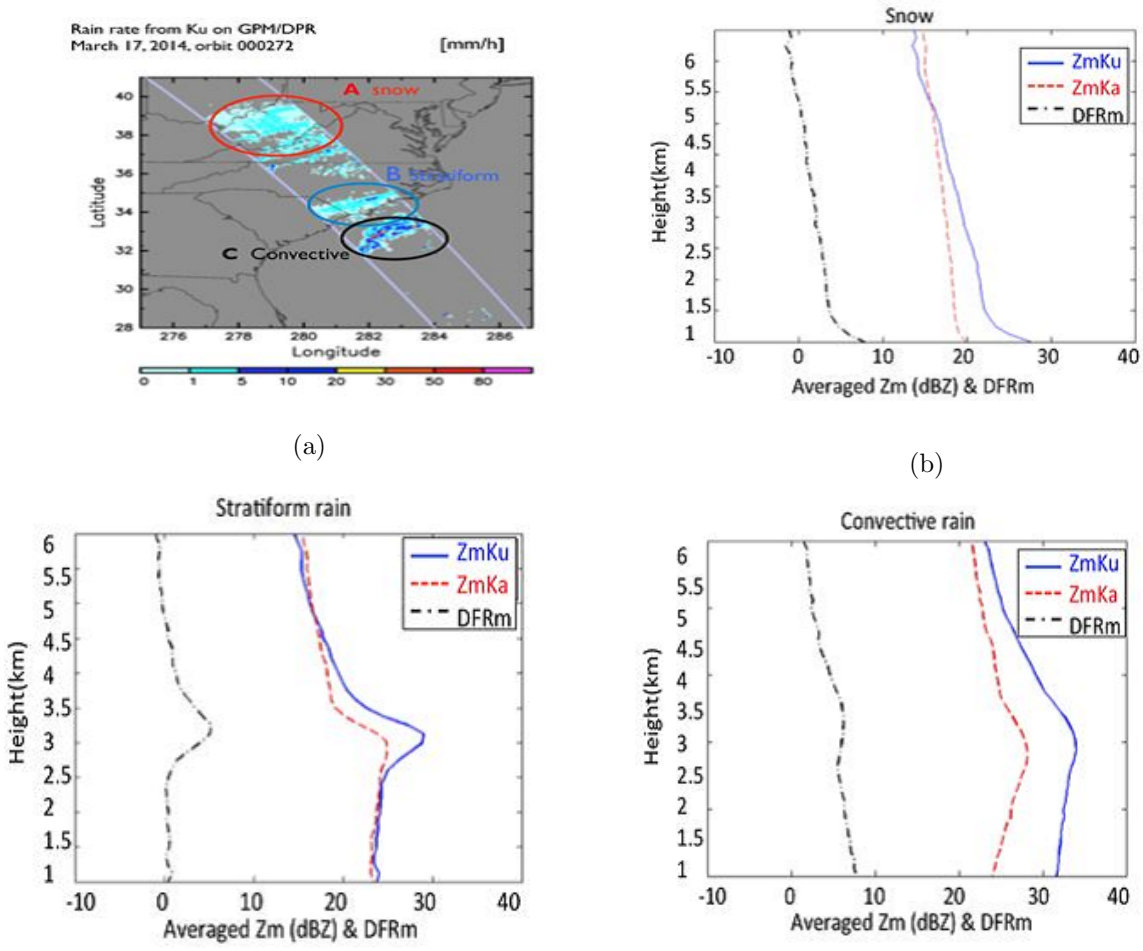


Figure 3.5.19: GPM DPR overpass of rainfall rate on March 17, 2014 (#000272). Circled A, B and C represents snow, stratiform rain, and convective rain. (b) Averaged reflectivity profiles as well as dual-frequency ratio profile for snow. (c) Same as (b) for stratiform rain. (d) Same as (b) for convective rain.

3.5.7 Algorithm description

Snow index (SI) is carefully chosen from the features discussed in Section 3.5.6. The definition of SI is

$$\mathrm{SI} = \frac{\mathrm{mean}(|\mathrm{DFRm}_{\mathrm{slope}}|)}{\mathrm{Zmku}_{\mathrm{max}} \times \texttt{Storm\_top\_height}} \tag{3.5-5}$$

$\mathrm{DFR}_{\mathrm{m}}$ slope (in dB/km) is used instead of $\mathrm{DFR}_{\mathrm{m}}$ value due to its immunity to calibration change. $\mathrm{Zmku}_{\mathrm{max}}$ (in dBZ) represents maximum of reflectivity at Ku-band along the profile. `Storm_top_height` represents altitude of storm top in km. In general, absolute value of $\mathrm{DFR}_{\mathrm{m}}$ slope in the numerator of Eq. (3.5-5) is larger for snow than for stratiform rain. $\mathrm{Zmku}_{\mathrm{max}}$ value in the denominator is larger for convective rain than for snow, while storm top height

in the denominator is lower for snow and stratiform rain than for convective rain. Therefore, SI is expected to be a larger value for snow profile than for rain. In the algorithm, we use the normalized value of Zmku_{max} and `Storm_top_height`. A statistics study of snow index is performed on large scale of GPM DPR vertical profiles. A total of 353166 rain profiles and 4935 snow profiles are used in this study. Rain profiles are chosen from tropical regions during the summer season and snow profiles are chosen from northern Europe and the northern US in the winter season. Information of 0 degree isotherm is used in profile selection. Figure 3.5.20 shows the histogram of snow index (SI) defined in Eq. (3.5-5) for snow and rain profiles respectively. In general, as expected, snow index is larger for snow profiles. The cumulative density function (CDF) is calculated for rain, and illustrated as a blue dashed curve in Figure 3.5.20, while red dashed curve represents $1 - \text{CDF}$ for snow profiles. At around 97% of CDF (or 1-CDF), SI index can separate snow and rain profiles. In other word, 97% of snow profiles have SI $>$ 17, while 97% of rain profiles have SI $<=$ 17. The statistics study considers calibration changes that will be applied to version 5 of the GPM DPR level 2 algorithm.

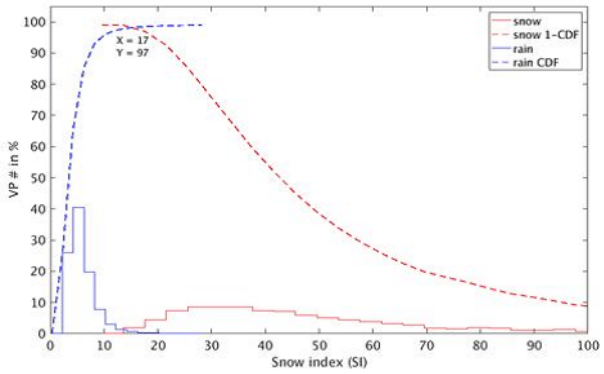


Figure 3.5.20: Large scale study of snow index using GPM DPR profiles. Histogram of snow index for rain (blue) and snow (red). Blue dashed curve is the cumulative density function (CDF) for rain. Red dashed curve is $1 - \text{CDF}$ for snow.

Figure 3.5.21 shows the flow chart of the surface snowfall identification algorithm that uses snow index. Besides snow index, other auxiliary information such as 0 degree isotherm and clutter free height are used in the algorithm. Although 0 degree isotherm (or surface temperature) plays an important role in snowfall detection, its accuracy is a challenge for space-borne weather radars such as GPM. Therefore, in this algorithm, 0 degree isotherm is not the dominant parameter but rather a constraint parameter. Snow index is calculated for profiles with 0 degree isotherm heights not higher than 10 range bins ($\approx$ 1.25 km) above clutter free height. This serves more like a sanity check to filter out nonsense profiles.

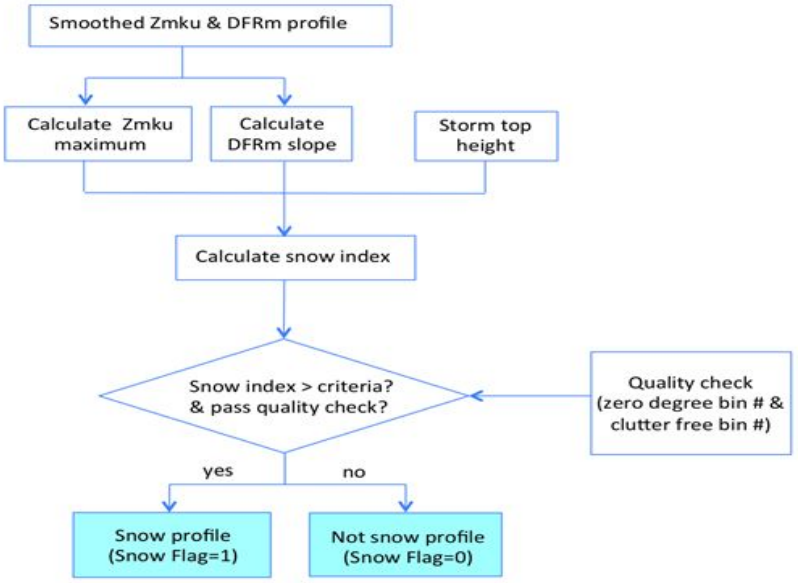


Figure 3.5.21: Flowchart to perform surface snowfall identification in profile classification module of GPM DPR level 2 algorithm.

3.5.8 Enhancement of GPM DPR level 2 dual-frequency classification module

In the current DPR level 2 classification module, the vertical profile is classified as either stratiform, convective or other type following the TRMM legacy. However, GPM orbits cover both cold and tropical regions. It is reasonable to add a snow flag to vertical profiles that indicates whether there is snow on the surface or not. The surface snowfall identification algorithm takes advantage of dual-frequency measurements from GPM DPR, thus it provides a new approach to detect snow other than the method using surface temperature or surface echo. The algorithm outputs `flagSurfaceSnowfall` and `surfaceSnowfallIndex` to each vertical profile in the DPR inner swath. This snow flag provides useful information for microphysics retrieval for other modules such as the solver module. Therefore, this algorithm enhances the current dual-frequency classification module in GPM DPR level 2 algorithm and has been implemented as an experiment version.

3.5.9 Eight digit rain type number

As mentioned in Section 3.5.4.1.4(d) and 3.5.4.2.2, the unified rain types for Ku-only, Ka-only, and DPR are expressed by eight digit numbers. Figure 3.5.22 illustrates the eight digit rain type number, which has a common structure for Ku-only, Ka-only, and DPR.

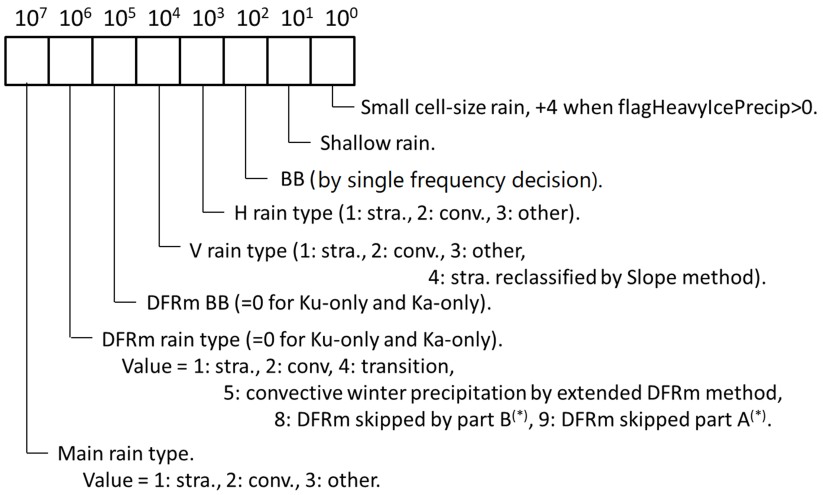


Figure 3.5.22: A summary of eight digit rain type number.
(*): For part A and part B, see Figure 3.5.8

3.5.10 New items related to the CSF module

In V06X and V07, the following two CSF related items are found in the EXPERIMENTAL group of the dual frequency output.

1. `flagGraupelHail`
2. `MixedPhaseTop`

In V07, the following new item is added to the STANDARD group of the dual frequency output.

3. `flagHail`

Descriptions of these are given in the following subsections.

3.5.10.1 flagGraupelHail

3.5.10.1.1 Graupel and Hail identification for DPR

Similar to the product of "`flagSurfaceSnowfall`" (Le and Chandrasekar, 2019), the precipitation type index (PTI) with different threshold has been proven to be able to effectively identify graupel and hail profiles (Le and Chandrasekar, 2018). The product of "`flagGraupelHail`" (or "GH flag") has been implemented in the Experimental structure. Figure 3.5.23 illustrates a cartoon plot for the geometry of GPM product "`flagGraupelHail`". The output is a binary value for each DPR matched footprint. 1 represents "graupel and hail exists" and 0 means "graupel and hail not exists". This is for along the vertical profile, not for surface.

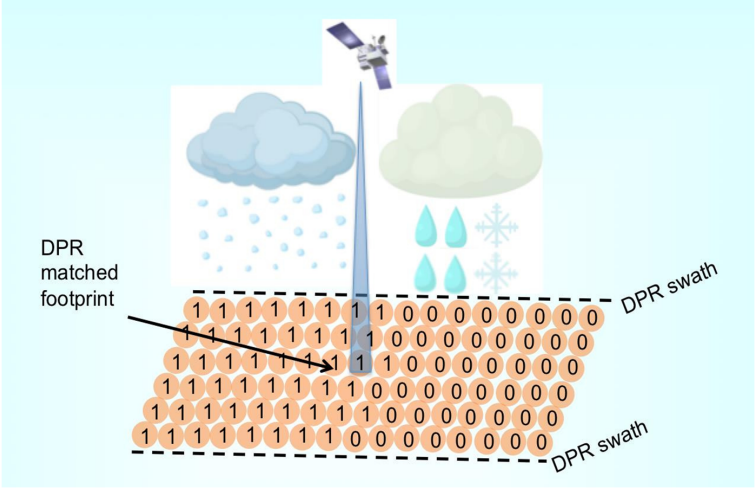


Figure 3.5.23: Cartoon illustration of GPM product "`flagGraupelHail`".

Three ingredients in calculating PTI are $\mathrm{DFR_m}$ slope, maximum Z_m value and storm top height. We collect more than 1000 DPR profiles with graupel and hail identified using WSR-88D radar network. Features of these vertical profiles are studied. Good separation can be observed which forms the fundamental of the algorithm. Figure 3.5.24 is the flowchart of the graupel and hail identification algorithm.

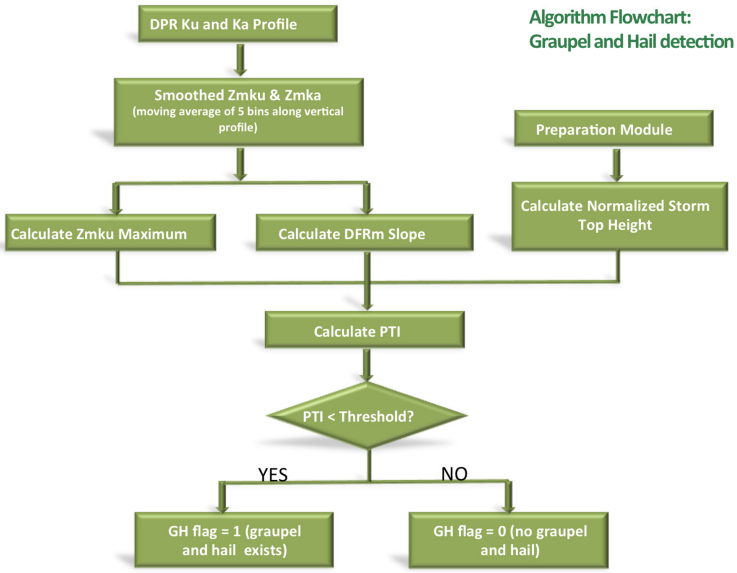


Figure 3.5.24: The flowchart of graupel and hail identification algorithm for GPM DPR.

3.5.10.1.2 Ground Validation of the Graupel and Hail Identification Algorithm

Between 2018 May to September, intense weather events are chosen from the United States for ground validation. These precipitation events are simultaneously captured by GPM-DPR and WSR-88D radars. Hydrometeor identification algorithm is performed on WSR-88D radars first. Chosen cases are all detected with Graupel or Hail. Graupel and Hail identification algorithm for GPM is applied and the product of "`flagGraupelHail`" is compared with ground radar results. More than 10 validation cases were chosen and validated showing excellent comparisons. The performance for a sample case is shown in Figure 3.5.25. In Figure 3.5.25, a strong precipitation event on Aug 24th, 2018 is simultaneously captured by GPM DPR orbit #25489, UTC 02:42:57 together with the WSR-88D radar KABR located at Aberdeen, SD. Figure 3.5.25 (a) shows the map of the overpass and the location of KABR radar. In Figure 3.5.25 (b) and (c), two PPI scans illustrate S band reflectivity of KABR at elevation of 2.42 and 6.41 degree respectively. High reflectivity value above 45 dBZ can be observed at east side as well as southwest side to the radar. "GR" and "HA" represents graupel and hail in the algorithm. Not surprisingly, we find graupel at east and southwest to the KABR radar. The orange ovals indicate the locations. Figure 3.5.25 (f) is the plot of DPR Ku band reflectivity at 2km. Dashed and solid black lines are DPR inner and outer swath. The circle is 100 km range of KABR. The result of "`flagGraupelHail`" is shown in (g), with dark green color representing Graupel and Hail exists. Comparing (d), (e) and (g), we find good agreements between two algorithms. It needs to be noticed that, in the validation cases, only inner swath data is taken for right now. The graupel and hail identification algorithm is applied to GPM DPR matched footprint for whole year 2018. Figure 3.5.26 illustrates the global distribution of "`flagGraupelHail`" count mapping to the $2^\circ \times 2^\circ$ Latitude/Longitude box for year 2018. From the plot, we spot peaks of graupel and hail occurrence at e.g., equatorial Africa, the equatorial and subtropical Americas, the Pampas of Argentina, the Himalayan Forelands, Indonesia etc. It shows great association with the World Lightening Map generated by NASA's Lightning Imaging Sensor on the Tropical Rainfall Measuring Mission satellite between 1995 and 2002, which demonstrates that "`flagGraupelHail`" is producing very reasonable results on a global scale. More validation details can be found in Le and Chandrasekar (2020).

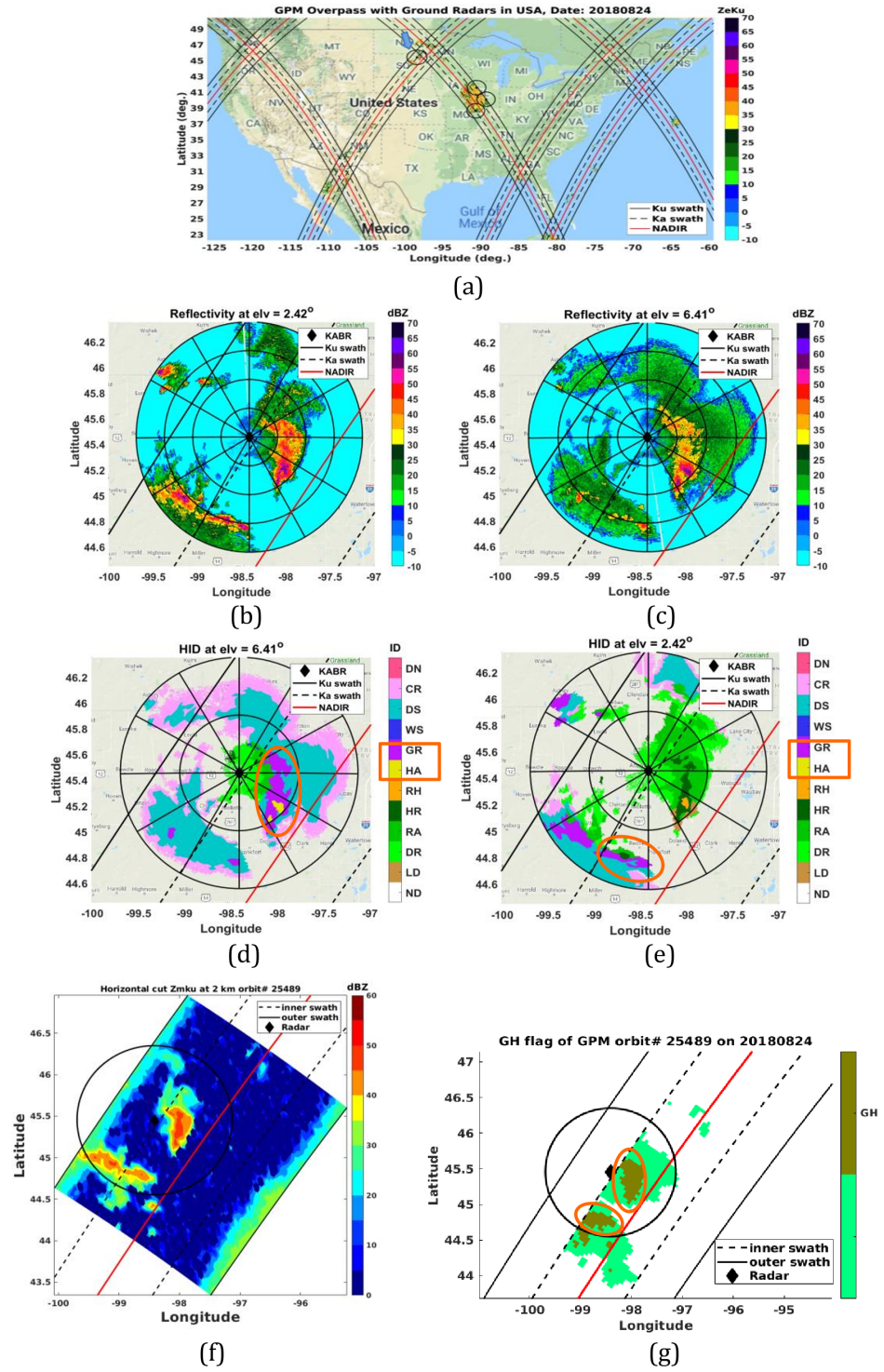


Figure 3.5.25: Validation case on 2018-08-24.
Precipitation event is observed by GPM DPR orbit #25489, UTC 02:42:57 together with WSR-88D radar KABR located at Aberdeen, SD with UTC 02:42:02. (a) Map of overpass and location of KABR radar. (b) Reflectivity of KABR at elevation of 2.42 degree. (c) Same as (b) for elevation of 6.41 degree. (d) HID of KABR at elevation of 2.42 degree. (e) Same as (d) for elevation of 6.41 degree. (f) DPR reflectivity Zmku at 2km. (g) "`flagGraupelHail`" for precipitation in (f). Dark green means Graupel and Hail exists.

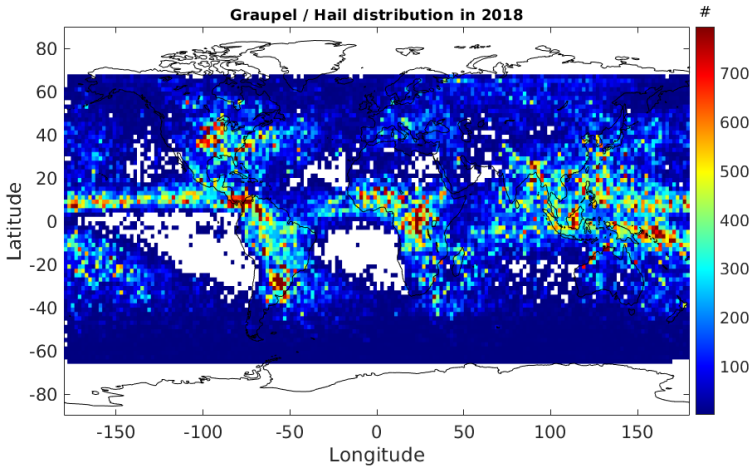


Figure 3.5.26: Global distribution of "`flagGraupelHail`" count mapping to the 2° x 2° Lat/Lon box for year 2018.

3.5.10.1.3 Hail identification algorithm for GPM DPR

With similar concept described in the graupel and hail identification section ("`flagGraupelHail`", Le and Chandrasekar, 2022a), a hail only identification algorithm has been developed for version 07 of DPR level-2 algorithm. Precipitation type index (PTI) is used to separate surface snowfall, GH and hail profiles with various thresholds. Figure 3.5.27 (a) shows the histogram of PTI for snow and GH (either graupel or hail exists) profiles. Snow profiles fall in larger value range of PTI than GH profiles. However, PTI of hail only profiles are restricted in an even smaller range as shown in Figure 3.5.27 (b). Around 85% hail PTI values are smaller than 3.3. This is mainly due to the multiple scattering features associated with hail profiles that decreases the slope of DFR_m. Meanwhile, storm top height and maximum reflectivity for hail profile are in larger value range, thus further decreasing the PTI. Due to sensitivity difference for Ka band profile for inner and outer swath, threshold of PTI in hail identification is sightly adjusted (Le and Chandrasekar, 2022b). Output of the algorithm is a Boolean product of "`flagHail`" with "1" representing hail exists along the vertical profile and "0" for not existent.

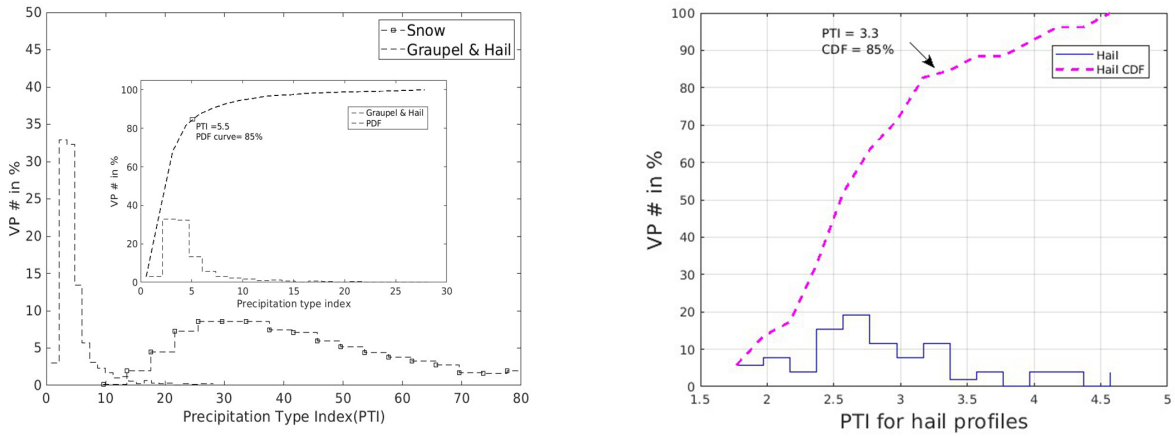


Figure 3.5.27: (a) Histogram of PTI for snow and GH profiles. (b) Histogram of PTI for hail only profiles.

Sample validation cases are illustrated here. A hail case is studied in Laviola et al. (2020) near Veneto region in Italy on 2018-07-05. The hail detection is made using a MWCC-H model (Microwave Cloud Classification method for Hail Detection). This probability-based model originally designed for AMSU-B/MHS-based (AMSU-B, Advanced Microwave Sounding Unit-B) radiometers has been fitted to the observations of all microwave radiometers onboard the satellites of the Global Precipitation Measurement (GPM) constellation. Figure 3.5.28 (a) illustrates the hail probability for this storm. There are three regions showing hail in red circles. GPM DPR (orbit #24721) captured the same storm on 2018-07-05 around UTC 17:13:35. In Figure 3.5.28 (b) and (c) depicts DPR reflectivity at Ku, and Ka band at 4 km height. Red and black lines in (b) and (c) are DPR nadir and boundaries of the swath (version 7 data with full swath). The blue lines are the location of scan # 4813 and #4801, where vertical cuts of reflectivity are illustrated in (e) and (f). Figure 3.5.28 (d) is the hail flag for the scan in (b). As expected, hail is detected with red color. The locations of hail in (d) align well with the locations at three red circles. Strong reflectivity is observed above 50 dBZ in vertical cuts of (e) and (f).

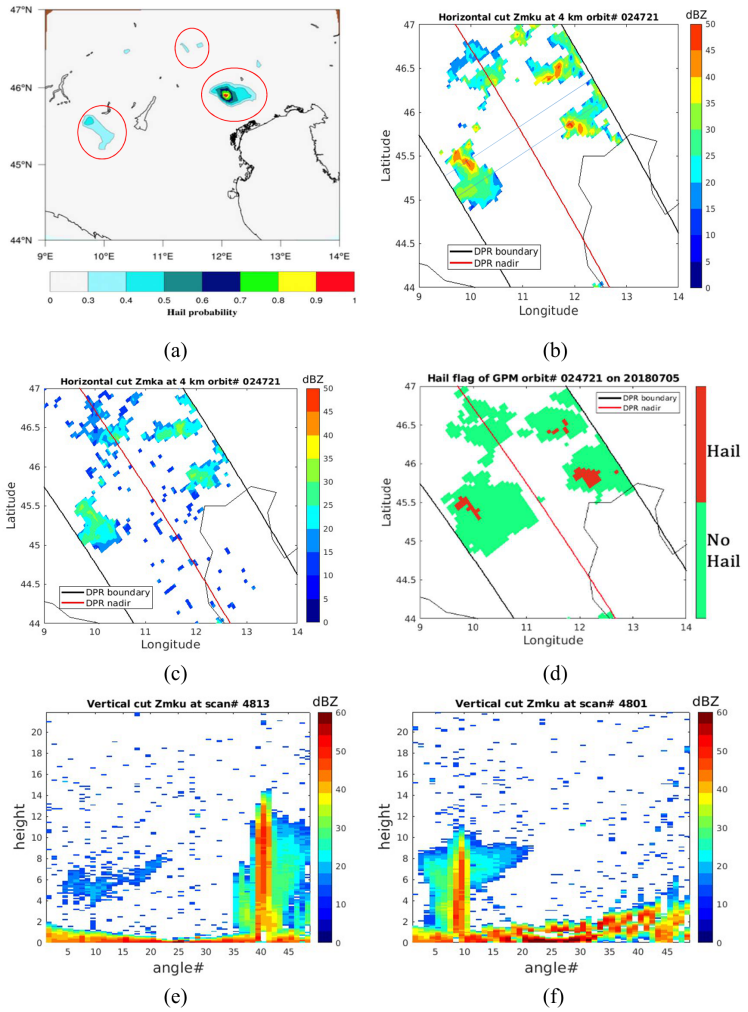


Figure 3.5.28: Hailstorm detected by DPR (orbit #24721) on 2018-07-05 near Veneto region in Italy. (a) Hail detected with MWCC-H model. (b) Reflectivity at Ku band for DPR at 4 km. (c) Reflectivity at Ka band. (d) Hail flag for scan in (b). Red indicates hail. (e) Vertical cut of DPR reflectivity Ku band at scan 4813. (f) Hydrometeor identification of scan shown in (d).

Another validation example illustrated here is with NEXRAD radar KLNX located in North Platte, NE. A strong storm was simultaneously observed by DPR (orbit #25028) and KLNX radar on 07-25-2018 around UTC 11:25:49. Figure 3.5.29 (a) illustrates the PPI reflectivity at 2.42-degree elevation angle for KLNX. The black circles are 33,66,100 km range of KLNX radar. Red line is the DPR nadir. Black dashed lines are for DPR inner swath and black solid line is the edge of DPR full swath. Figure 3.5.29 (b) is the hydrometeor type for scan in (a) using algorithm described in (Bechini and Chandrasekar, 2015). To the west of radar at about 100 km range, hail is detected, illustrated with yellow color in (b). DPR Ku reflectivity at 4 km is shown in Figure 3.5.29 (c). Black circle corresponds to 100 km range of KLNX. In (d), the hail flag is illustrated with red stands for hail locations. Comparing (b) and (d), at DPR and KLNX

overlapped region, we detect the hail at the west to the KLNX radar, exactly the same spot in (b).

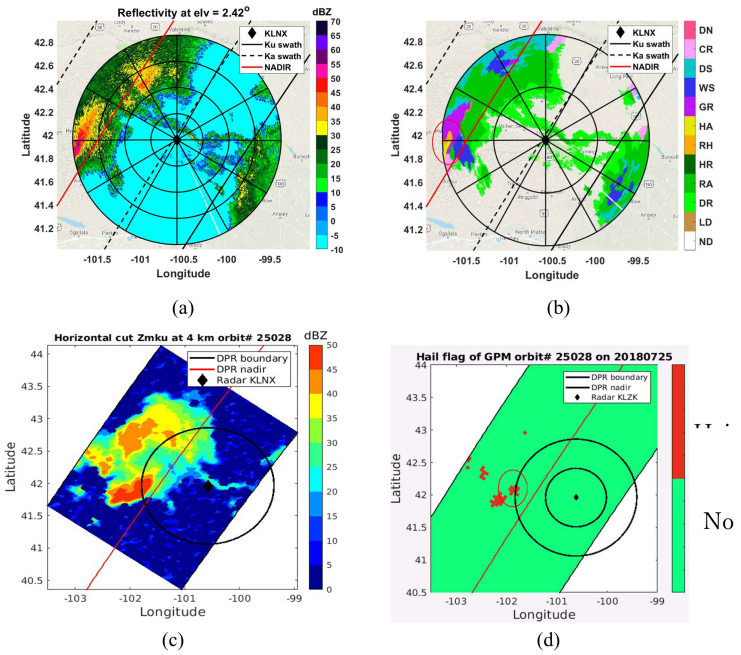


Figure 3.5.29: (a) KLNX PPI scan at elevation of 2.42 degree on July 25th, 2018 at UTC 11:25:49. (b) Hydrometeor type for scan in (a). (c) Ku band reflectivity at 4 km from DPR. (d) Hail flag for (c).

3.5.10.2 MixedPhaseTop

3.5.10.2.1 Introduction

The experimental flag `MixedPhaseTop` is included in V06X to identify the location where the phase of precipitation particles turns from solid (ice) to liquid (water). The flag is set at nearly all pixels in which precipitation is detected even when no bright band is detected.

When a bright band is detected, the range at which this `MixedPhaseTop` flag is set is generally very close to the bright-band top or the melting-layer top in the classification module.

The flag is set based on an empirical fact that the dual-frequency ratio (DFR) changes substantially when particles change their phase state. This method works well in stratiform rain in which ice particles melt into water droplets as they descend. In particular, when a bright band is formed, the average particle size increases as snow flakes form aggregates in it and the DFR increases rapidly before it decreases to a low value that corresponds to small melted water droplets.

In convective rain, however, the drop size does not change much and the method does not work well in many cases. In a convective system, a coalescence process is supposed to be happening

as liquid particles are lifted into a below-freezing temperature region by updraft. In this process, the diameters of particles, which are a decisive factor in determining DFR, do not change much. As a result, the change in DFR is mainly caused by the change in the dielectric factor of particles from water to ice when they are carried to a higher elevation. This change should be considered to be detectable. In reality, however, the change of phase state of particles does not seem to occur at a single well defined height, but it depends on the size of particles. Small particles may remain unfrozen to a relatively high altitude because of their high surface tension that requires an ambient air temperature well below 0 degree C to freeze. They can be frozen by riming to other solid particle by collisions. As a result, the phase transition height may not be well defined in a convective storm. Since liquid particles contribute to the attenuation of radar echoes that also increases the apparent DFR, detection of the phase transition by measured DFR in convective precipitation is very complicated and difficult.

The current algorithm that is implemented in V06X still assigns a flag in convective storms using the same criteria even though the flag position may not be trustworthy, simply because this is an experimental flag.

A simplified flow of the algorithm is as follows (not exactly the same as the actual flow).

1. DFR is estimated from $\mathrm{DFR_m}$ (measured apparent DFR) with small attenuation corrections.

2. Calculate the second derivative of DFR with respect to range r.

3. Find the maximum of its absolute value within the pre-defined interval around 0°C, and identify that height as the mixed-phase top height.

The advantage of this algorithm is its simplicity. It does not depend on the calibration of radar. It has higher detectability than the bright band detection algorithm implemented in the standard product. Disadvantages are that it is susceptible to noise so that we need to smooth the data before processing and that the solution depends on the estimated 0°C height in some cases.

3.5.10.2.2 Idea behind the algorithm

Both effective radar reflectivity factor Z_e and specific attenuation k are functions of N_w, D_m and f_m, where N_w is a parameter proportional to the number density of the particle size distribution, D_m is the volume-weighted mean diameter (or any mean diameter), and f_m is the melting fraction of particles. (Note that D_m here is the actual diameter without melting if the particles are not totally liquid. It is not an equivalent melted mean diameter.) They are all functions of range r.

$$Z_e(r) = Z_e(N_w(r), D_m(r), f_m(r)),$$

$$k(r) = k(N_w(r), D_m(r), f_m(r)).$$

Measured radar reflectivity factor Z_m is related to Z_e and k through

$$Z_m(r) = Z_e(r) \exp(-2q \int_0^r k(s)\, ds), \tag{3.5-6}$$

where $q = \log_e(10)/10$.

In this section, we use notations dBZ_e and dBZ_m defined by the following equations.

$$\mathrm{dBZ}_e \stackrel{\mathrm{def}}{=} 10 \log_{10}(Z_e), \qquad \mathrm{dBZ}_m \stackrel{\mathrm{def}}{=} 10 \log_{10}(Z_m).$$

With these symbols, Eq. (3.5-6) becomes

$$\mathrm{dBZ}_m(r) = \mathrm{dBZ}_e(r) - 2 \int_0^r k(s)\, ds.$$

We are concerned with radar echoes at two different frequencies. Radar variables at these two frequencies are distinguished by subscript 1 and 2. Subscript 1 is used for the lower frequency (Ku band) and 2 for the higher frequency (Ka band). No subscript is attached if the statement is applicable to both frequencies.

The dual-frequency ratio (DFR) is defined by

$$\mathrm{DFR} \stackrel{\mathrm{def}}{=} \mathrm{dBZ}_{e1} - \mathrm{dBZ}_{e2}.$$

Similarly, we define an apparent measured dual-frequency ratio ($\mathrm{DFR_m}$) by

$$\mathrm{DFR_m} \stackrel{\mathrm{def}}{=} \mathrm{dBZ}_{m1} - \mathrm{dBZ}_{m2} = \mathrm{DFR} - 2 \int_0^r (k_1(s) - k_2(s))\, ds.$$

Note that

$$\frac{\partial \mathrm{DFR}}{\partial N_w} = 0,$$

because both Z_{e1} and Z_{e2} are linearly proportional to N_w, and that

$$\frac{\partial \mathrm{DFR}}{\partial D_m} > 0 \quad \text{(for ice)} \quad \text{and} \quad \frac{\partial \mathrm{DFR}}{\partial D_m} \begin{cases} > 0 & \text{for } D_m > D_p, \\ = 0 & \text{for } D_m = D_p, \\ < 0 & \text{for } D_m < D_p. \end{cases} \quad \text{(for water)}$$

where D_p is about 1 mm if the two frequency bands are Ku and Ka bands. If particles change their phase state from solid to liquid without changing their dimensions, DFR decreases, i.e.,

$$\frac{\partial \mathrm{DFR}}{\partial f_m} < 0.$$

(Note that this inequality does not necessarily hold if the freezing ratio depends on the particle size.)

Therefore, for $D_m > D_p$, DFR changes with N_w, D_m and f_m as follows.

$$\frac{\partial \mathrm{DFR}}{\partial N_w} = 0, \quad \frac{\partial \mathrm{DFR}}{\partial D_m} > 0, \quad \text{and} \quad \frac{\partial \mathrm{DFR}}{\partial f_m} < 0.$$

On the other hand, Z_e's dependence on N_w, D_m and f_m is different.

$$\frac{\partial \mathrm{dBZ}_e}{\partial N_w} > 0, \quad \frac{\partial \mathrm{dBZ}_e}{\partial D_m} > 0, \quad \text{and} \quad \frac{\partial \mathrm{dBZ}_e}{\partial f_m} > 0.$$

In this algorithm, we look at the changes in DFR (or $\mathrm{DFR_m}$) and in dBZ_e. Their derivatives with respect to r are

$$\frac{d\mathrm{DFR}}{dr} = \frac{\partial \mathrm{DFR}}{\partial N_w}\frac{dN_w}{dr} + \frac{\partial \mathrm{DFR}}{\partial D_m}\frac{dD_m}{dr} + \frac{\partial \mathrm{DFR}}{\partial f_m}\frac{df_m}{dr} = \frac{\partial \mathrm{DFR}}{\partial D_m}\frac{dD_m}{dr} + \frac{\partial \mathrm{DFR}}{\partial f_m}\frac{df_m}{dr},$$

$$\frac{d\mathrm{dBZ}_e}{dr} = \frac{\partial \mathrm{dBZ}_e}{\partial N_w}\frac{dN_w}{dr} + \frac{\partial \mathrm{dBZ}_e}{\partial D_m}\frac{dD_m}{dr} + \frac{\partial \mathrm{dBZ}_e}{\partial f_m}\frac{df_m}{dr}.$$

In the case of space-borne radar, r increases downward. In most cases, therefore, we can safely assume that

$$\frac{df_m}{dr} \geq 0,$$

i.e., the phase of particles changes from ice to water downward. (There is a possibility that wet hail stones exist above dry hail stones in different locations within a footprint of radar. We exclude such possibilities in this algorithm.)

The signs of $\frac{d\mathrm{DFR}}{dr}$ and $\frac{d\mathrm{dBZ}_e}{dr}$ indicate which among N_w, D_m and f_m causes the major changes in DFR and dBZ_e. There are four possible cases.

1. $\frac{d\mathrm{dBZ}_e}{dr} \geq 0$ and $\frac{d\mathrm{DFR}}{dr} < 0$,
2. $\frac{d\mathrm{dBZ}_e}{dr} < 0$ and $\frac{d\mathrm{DFR}}{dr} \geq 0$,
3. $\frac{d\mathrm{dBZ}_e}{dr} < 0$ and $\frac{d\mathrm{DFR}}{dr} < 0$,
4. $\frac{d\mathrm{dBZ}_e}{dr} \geq 0$ and $\frac{d\mathrm{DFR}}{dr} \geq 0$.

In case 1, it is most likely that $\frac{df_m}{dr} \geq 0$ in that range. In other words, the condition that $\frac{d\mathrm{dBZ}_e}{dr} \geq 0$ and $\frac{d\mathrm{DFR}}{dr} < 0$ can be used to detect the interval of transition from solid to liquid phase in the vertical profile of radar echo.

If this condition is satisfied without any phase transition (i.e., $\frac{df_m}{dr} = 0$), the only other possibility is that $\frac{dD_m}{dr} < 0$ and $\frac{dN_w}{dr} > 0$ in such a way that the positive effect from $\frac{\partial \mathrm{dBZ}_e}{\partial N_w}\frac{dN_w}{dr}$ dominates the negative effect from $\frac{\partial \mathrm{dBZ}_e}{\partial D_m}\frac{dD_m}{dr}$. Such a situation may happen in a developing stage of a convective storm, but it is rather unlikely in a horizontally uniform precipitation system.

However, if there are two convective towers in the footprint and their heights are different, there is a possibility that N_w suddenly increases downward at the top of the shorter convection and that the effective D_m decreases there. (In a melting layer of stratiform rain, D_m decreases downward as particles melt, but the number density that is proportional to N_w does not increase as much to compensate for the effect of decrease in D_m.)

Note that if we assume a Gamma distribution of DSD and that the total precipitation water amount is constant (N_w changes with D_m in this case), then $\frac{\partial \mathrm{dBZ}_e}{\partial N_w}\frac{dN_w}{dr} + \frac{\partial \mathrm{dBZ}_e}{\partial D_m}\frac{dD_m}{dr}$ has the same sign as $\frac{\partial \mathrm{dBZ}_e}{\partial D_m}\frac{dD_m}{dr}$. In other words, the effect of D_m is larger than that of N_w.

In case 2, the increasing trend of DFR is caused by the increase of D_m with r (under the assumption that $\frac{df_m}{dr} \geq 0$). Therefore, if $\frac{d\mathrm{DFR}}{dr} \geq 0$ and $\frac{d\mathrm{dBZ}_e}{dr} < 0$, then $\frac{\partial \mathrm{dBZ}_e}{\partial N_w}\frac{dN_w}{dr} < 0$ and N_w decreases substantially with range. In case 3, it is likely that $\frac{dD_m}{dr} < 0$. In case 4, it is likely that $\frac{dD_m}{dr} > 0$.

In the latter three cases (cases 2, 3, and 4), it is impossible to identify the effect of $\frac{df_m}{dr}$ only from the signs of $\frac{d\mathrm{dBZ}_e}{dr}$ and $\frac{d\mathrm{DFR}}{dr}$.

3.5.10.2.3 Algorithm

This algorithm is based on the idea described above. It tries to find the range at which $d\mathrm{DFR}/dr$ changes most rapidly in case 1 above. To find its minimum, the second derivative of DFR i.e., $d^2\mathrm{DFR}/dr^2$ is sought and its range is identified as the mixed phase top.

Since neither DFR nor dBZ_e is actually a measurable variable, we need to estimate them from measurable quantities, i.e., $\mathrm{DFR_m}$ and dBZ_m. Derivatives of $\mathrm{DFR_m}$ and dBZ_e are related to those of DFR and dBZ_e through

$$\frac{d\mathrm{DFR_m}}{dr} = \frac{d\mathrm{DFR}}{dr} - 2(k_1 - k_2), \quad \text{and}$$
$$\frac{d\mathrm{dBZ}_m}{dr} = \frac{d\mathrm{dBZ}_e}{dr} - 2k$$

.

Note that $k_2 \geq k_1 \geq 0$ so that

$$\frac{d\mathrm{DFR_m}}{dr} \geq \frac{d\mathrm{DFR}}{dr} \quad \text{and} \quad \frac{d\mathrm{dBZ}_m}{dr} \leq \frac{d\mathrm{dBZ}_e}{dr}.$$

These inequalities imply that if

$$\frac{d\mathrm{DFR_m}}{dr} < 0 \quad \text{and} \quad \frac{d\mathrm{dBZ}_m}{dr} > 0,$$

then we can conclude that

$$\frac{d\mathrm{DFR}}{dr} < 0 \quad \text{and} \quad \frac{d\mathrm{dBZ}_e}{dr} > 0,$$

and that

$$\frac{df_m}{dr} > 0.$$

In other words, we can use $\mathrm{DFR_m}$ and dBZ_m instead of DFR and dBZ_e for the detection of the region where the condition $\frac{df_m}{dr} > 0$ is most likely satisfied. In fact, if we use radar reflectivity factor at Ku band, Z_{e1}, for Z_e, the attenuation can be generally neglected because it is very small for ice particles. Since we are dealing with only the radar echoes down to the mixed phase region, errors associated with attenuation correction in rain region is not critical.

Nevertheless, in order to minimize the error associated with attenuation, the following formulas are used to calculate the first derivatives of DFR and dBZ_e from those of $\mathrm{DFR_m}$ and dBZ_m.

$$\frac{d\mathrm{DFR}}{dr} = \frac{d\mathrm{DFR_m}}{dr} + 2(k_1 - k_2) \quad \text{and} \quad \frac{d\mathrm{dBZ}_{e1}}{dr} = \frac{d\mathrm{dBZ}_{m1}}{dr} + 2k_1.$$

Both k_1 and k_2 are estimated approximately from Z_{m1} (Ku band Z_m) by using a power-law k-Z_{e1} relationships with Z_{e1} replaced by Z_{m1}.

$$k_1 = \alpha_1 Z_{m1}^{\beta_1} = \alpha_1 \exp(q\beta_1 \mathrm{dBZ}_{m1}), \quad \text{and} \quad k_2 = \alpha_2 Z_{m1}^{\beta_2} = \alpha_2 \exp(q\beta_2 \mathrm{dBZ}_{m1}).$$

We assume that Z_{m1} is nearly equal to Z_{e1} because Z_{e1} suffers from attenuation only slightly down to the range where melting starts. In the case of stratiform precipitation systems, both k_1 and k_2 are small and the corrections themselves are negligibly small.

Since DFR statistically changes with Z_{e1} (and Z_{e2}) to some extent, an additional term that is proportional to $d\mathrm{dBZ}_e/dr$ is subtracted from $d\mathrm{DFR}/dr$ to minimize the effect of variation in D_m to $d\mathrm{DFR}/dr$. In other words, we try to find the range at which the following quantity changes its trend the most.

$$\frac{d\mathrm{DFR}_c}{dr} = \frac{d\mathrm{DFR_m}}{dr} + 2(k_1 - k_2) - c\frac{d\mathrm{dBZ}_{m1}}{dr},$$

where c is a proportional coefficient of $d\mathrm{DFR}/d\mathrm{dBZ}_{e1}$ at a typical value of dBZ_{e1}. (The actual value of c used in the algorithm is $c = 1/7$.) To find such a point, the range at which the second derivative of DFR_c becomes minimum is sought. The search window for the solution is between 0°C±1 km.

3.5.10.2.4 Caveats

This algorithm tries to find the minimum point of $\frac{d^2\mathrm{DFR}_c}{dr^2}$. Since derivatives are susceptible to noise, it is necessary to smooth the observed profiles of dBZ_{m1} and dBZ_{m2} substantially. A Savitzky-Golay filter (Savitzky and Golay, 1964; Press et al. 1992) is adopted to achieve this purpose. This filter has a nice property that it can preserve peaks of signal without much reduction up to a certain wavenumber (defined by the order of the filter) and that derivatives of the smoothed profiles can be obtained fast and easily. To suppress noise, 9 points (current point

+ 5 points on either side) around the point in question are used for smoothing. The number of points usable for smoothing reduces near the lower end of the valid echo profile because we cannot use data below the `binClutterFreeBottom`. As a result, the performance of detection decreases when the mixed phase top is close to the `binClutterFreeBottom`.

In an updraft in a convective storm, freezing of particles may not occur at the same height for different sizes of particles. As a result, some assumptions such as $\frac{\partial \mathrm{DFR}}{\partial f_m} < 0$ are considered to be not valid in some cases. In such cases, this algorithm may fail to identify the melting phase top height.

3.5.11 Input and Output Variables

Input Variables

From Preparation (PRE) module:

`lat`
`lon`
ScanTime information (`Year`, `Month`, `DayOfMonth`, `Hour`, `Minute`, `Second`, `Millisecond`, `DayOfYear`, `SecondOfDay`)
`elevation`
`landSurfaceType`
`localZenithAngle`
`flagPrecip`
`binRealSurface`
`binStormTop`
`heightStormTop`
`binClutterFreeBottom`
`sigmaZeroMeasured`
`rangeBottom`
`rangeTop`
`Height`
`echoSignalPower`
`zFactorMeasured`

From Vertical (VER) module:

`airTemperature`
`flagInversion`
`binZeroDeg`
`zFactorNPCorrected`
`heightZeroDeg`
`flagEcho` (R/W) (This flag marks possible sidelobe clutter positions.)
`qualityData` (R/W)

From Solver (SLV) module:

`zFactorNPCorrected` (For convective decision by the V-method in the single frequency Ku- or Ka-loop)

Output Variables

`flagBB` [4 byte integer]
`binBBPeak` [2 byte integer]
`binBBTop` [2 byte integer]
`binBBBottom` [2 byte integer]

heightBB [4 byte real]
widthBB [4 byte real]
qualityBB [4 byte integer]
typePrecip [4 byte integer]
qualityTypePrecip [4 byte integer]
flagShallowRain [4 byte integer]
binDFRmMLTop [4 byte integer] (dual frequency output only)
binDFRmMLBottom [4 byte integer] (dual frequency output only)
binHeavyIcePrecipTop [2 byte integer] – Introduced in V06X
binHeavyIcePrecipBottom [2 byte integer] – Introduced in V06X
nHeavyIcePrecip [1 byte integer (unsigned)] – Introduced in V06X
flagMLquality [1 byte integer (unsigned)] – Introduced in V06X
flagHail [1 byte integer (unsigned)] – Introduced in V07
flagHeavyIcePrecip [1 byte integer (unsigned)]
flagAnvil [1 byte integer (unsigned)] (single frequency Ku FS and dual frequency FS only)
flagSurfaceSnowfall [1 byte integer (unsigned)] (Experimental output of dual frequency FS)
surfaceSnowfallIndex [4 byte real] (Experimental output of dual frequency FS)
flagGraupelHail [1 byte integer (unsigned)] – Introduced in V06X
binMixedPhaseTop [2 byte integer (unsigned)] – Introduced in V06X

Details of the description for variables are described in Appendix or in the following websites:

https://www.eorc.jaxa.jp/GPM/en/archives.html

https://arthurhou.pps.eosdis.nasa.gov/GPMprelimdocs.html

3.5.12 Relation to other modules

The Classification (CSF) module uses data from the Preparation (PRE) module and the vertical (VER) module. Each single frequency CSF module uses data from the corresponding Solver (SLV) module in the 2-nd part of (a kind of iteration) loop for convective decision by the V-method. The output data from the CSF module is used by the Solver (SLV) module.

3.5.13 Bug fix

This section describes bug fixes made in the V07, V06X, V06, and V05 CSF modules.

(A) Bug(s) fixed in the V07 CSF module

In the V06X dual frequency CSF module, if there is rain at Ka but no rain at Ku, the unified type is other because of a bug. The bug is fixed in V07 in such a way that if there is rain at Ka but no rain at Ku, then the single frequency Ka type is used as the dual frequency type.

(B) Bugs fixed in the V06X CSF module

1) The V06 CSF module reclassifies stratiform rain by the UW "slope method." The "slope method" is applied to the Ku-band processing only. But `NearSurfacePrecipitationRate` (R_{NS}) of the reclassified stratiform rain sometimes becomes very large though the occurrence of such is rare. (R_{NS} of stratiform rain must not be very large.) The V06X CSF module fixes this problem by introducing a kind of output filter.

2) In rare cases, some BB peaks are misjudged as surface peaks in V05 by a bug in the PRE module. The effect of bug becomes appreciable in the following two cases:
   a) High BB case: When BB peaks are misjudged as the surface echo over high mountain areas, the clutter free bottom becomes dangerously too high.
   b) Low BB case: In winter, a large number of low height BB are missed.

The V06X algorithms fix this problem to some extent. Further improvement will be made in V07. (Though the bug fix is made in the PRE module, the bug affects the detection of BB in the CSF module; therefore it is mentioned here.)

(C) Bugs fixed in the V06 CSF module

The following bugs in V05 are fixed in V06.

1) Some side-lobe peaks being misjudged as HIP in V05 are fixed in V06. (The fix is applied to the Ku outer swath only.)

2) When `flagHeavyIcePrecip` > 0, rain type is expected to be convective, but the DPR rain type for this case is sometimes stratiform in V05 because of a bug. This bug is fixed in V06.

(D) Bugs fixed in the V05 CSF module

The V04 CSF codes for the detection of BB contain a large number of bugs, which are fixed in V05.

Because of the bugs, almost all V04 `widthBB`s were wrong, some V04 `heightBB`s are incorrect, and in some rare cases `binBBPeak`, `binBBTop`, and `binBBBottom` are not set properly in V04.

3.5.14 Issues

Besides the parameter tunings, the following improvements are needed in the next update.

- Improvement of BB detection, in particular the detection of low BB in winter and also the detection of BB in the outer swath.

- Extensive use of `flagEcho` in the FLG group is needed for the side-lobe clutter rejection and for the decision of other type in the CSF module.

- Improvement of CSF related EXPERIMENTAL codes (including validation of their results.)

3.6 Drop Size Distribution (DSD) module

3.6.1 Objective

The objective of DSD module is to set the physical variables of precipitation particles (especially, density, dielectric constants, and falling velocity), to parameterize $N(D)$ (in other words, to set $n(D; D^*)$), and to set the R-D_m relationship.

3.6.2 Processes

The following processes are common to single- and dual-frequency algorithms.

3.6.2.1 Target pixels and range bins

Pixels with precipitation are processed. Range bins from the storm top range bin to the land surface range bin (including those without precipitation and those with ground clutter) are processed. In dual-frequency algorithm, range bins from the higher storm top range bin of the two frequencies to the lower land surface range bin of the two frequencies are processed.

3.6.2.2 Nodes and the physical temperature and phase of the particles

Five range bins are selected as nodal bins and designated as nodes A through E. Simultaneously, the physical temperature of particles (particle temperature) and the phase of particles are set. Particle temperature is related to the dielectric constants and it is not necessarily the same with air temperature. The dielectric constant is dependent on the precipitation types and the detection of bright band.

3.6.2.2.1 Stratiform precipitation with bright band

The range bin that corresponds to the upper edge of bright band is node B, the range bin at the peak of bright band is node C, and the range bin at the lower edge of bright band is node D. Between node B and node D, particle temperature is set to 0°C. Above node B, the particle temperature is basically the same with ambient air temperature, but is 0°C if the air temperature is higher than 0°C. Below node D, the particle temperature is basically set to the same with the air temperature, but 0°C if the air temperature is lower than 0°C. The range bin with particle temperature closest to −20°C corresponds to node A, and that closest to 20°C is node E. At and above node A, the phase of precipitation is *Solid*. At and below node D, the phase of precipitation is *Liquid*. At other range bins, below node A and above node D, the phase of precipitation is *Mixed* (Solid and Liquid).

3.6.2.2.2 Stratiform precipitation without bright band

The process is the same as Section 3.6.2.2.1, but the nodes B, C, and D are at the range bin corresponding to 0°C.

3.6.2.2.3 Convective precipitation or other-type precipitation

The process is the same as Section 3.6.2.2.2.

3.6.2.3 Assumption of precipitation particles

A precipitation particle is modeled as a spherical particle composed of liquid water, solid water, and air. Its diameter (or drop size) is designated as D_b [mm], and its density as ρ_b [g/cm^3]. The volume ratios of liquid water, solid water, and air to the particle are P_w, P_i, and P_a, respectively. The following equation always holds.

$$P_w + P_i + P_a = 1 \tag{3.6-1}$$

The densities of liquid water, solid water, and air are ρ_w, ρ_i, ρ_a, respectively. Then, the density of the particle can be given as below.

$$\rho_b = P_w\rho_w + P_i\rho_i + P_a\rho_a \tag{3.6-2}$$

Generally, we can assume $\rho_w = 1.0$ (g/cm^3). ρ_i is also constant (for example, $\rho_i = 0.92$). Moreover, ρ_a can be regarded to be 0 (g/cm^3) for simplicity. With these constant values, Eq. (3.6-2) can be simplified to Eq. (3.6-3).

$$\rho_b = P_w + 0.92P_i \tag{3.6-3}$$

3.6.2.4 Setting of volume ratio and density

3.6.2.4.1 Liquid phase particle

Obviously, $P_w = 1$, $P_i = P_a = 0$, $\rho_b = 1.0$ (g/cm^3).

3.6.2.4.2 Mixed phase particle

First, P_w is set based on Awaka's model or TRMM/PR's experience. Then, ρ_b can be calculated by the following empirical equation.

$$\rho_b = \sqrt{P_w} \tag{3.6-4}$$

Then, P_i can be calculated by Eq. (3.6-2) with $\rho_a = 0$.

$$P_i = \frac{\rho_b - P_w\rho_w}{\rho_i} \tag{3.6-5}$$

3.6.2.4.3 Solid phase particle

First, ρ_b is set, for example, $\rho_b = 0.10$ (g/cm^3). As $P_w = 0$, P_i can be calculated by Eq. (3.6-5).

3.6.2.5 Dielectric constants (relative permittivity)

ε_b is calculated by mixing rule, for example, given in Eq. (3.6-6).

$$\frac{\varepsilon_b - 1}{\varepsilon_b + U} = P_w \frac{\varepsilon_w - 1}{\varepsilon_w + U} + P_i \frac{\varepsilon_i - 1}{\varepsilon_i + U} + P_a \frac{\varepsilon_a - 1}{\varepsilon_a + U} \tag{3.6-6}$$

where U is given as below,

$$U = \begin{cases} 12.0 & \text{if } \rho_b \leq 0.09 \text{ g/cm}^3, \\ 2.0e^{13.0(\rho_b - 0.09)} & \text{if } \rho_b > 0.09 \text{ g/cm}^3. \end{cases} \tag{3.6-7}$$

ε_w and ε_i are dielectric constants of water and ice and are functions of particle temperature, and ε_a is dielectric constant of air. As ε_a is approximated to 1, Eq. (3.6-6) can be simplified to Eq. (3.6-8).

$$\frac{\varepsilon_b - 1}{\varepsilon_b + U} = P_w \frac{\varepsilon_w - 1}{\varepsilon_w + U} + P_i \frac{\varepsilon_i - 1}{\varepsilon_i + U} \tag{3.6-8}$$

3.6.2.6 Falling velocity

Falling velocity V_b of the particle can be given as below.

$$V_b = 3.3 \times [\rho_b - \rho_a]^{1/2} \qquad \text{if } \rho_b \leq 0.05 \text{ g/cm}^3; \tag{3.6-9a}$$

$$V_b = 8.8 \times [0.1 D_b (\rho_b - \rho_a)]^{1/2} \qquad \text{if } 0.05 \text{ g/cm}^3 < \rho_b \leq 0.3 \text{ g/cm}^3; \tag{3.6-9b}$$

$$V_b = \frac{\rho_b^{1/3} - 0.3^{1/3}}{1.0 - 0.3^{1/3}} (V_R - V_{b,0.3}) + V_{b,0.3} \qquad \text{if } 0.3 \text{ g/cm}^3 < \rho_b \leq 1.0 \text{ g/cm}^3. \tag{3.6-9c}$$

In Eq. (3.6-9c), V_R is the fall velocity of rain drops, and $V_{b,0.3}$ is the falling velocity calculated in Eq. (3.6-9b) with $\rho_b = 0.3$ g/cm^3. The relationship between D_b and the diameter after melting D_{melt} is as follows.

$$D_b = D_{\text{melt}} \, \rho_b^{-1/3} \tag{3.6-10}$$

3.6.2.7 Drop size distribution function

3.6.2.7.1 Liquid phase

Gamma distribution function may be most commonly used for the drop size distribution function in liquid phase. The function is given as below.

$$N(D) = N_m D^{\mu} \exp\left[-\frac{(4 + \mu) D}{D_m} \right], \tag{3.6-11}$$

where D_m is the volume weighted diameter (or often called the mass weighted diameter) defined as

$$D_m = \frac{\int D^4 N(D)\, dD}{\int D^3 N(D)\, dD}.$$

N_m is the corresponding scale factor. The shape factor μ is assumed to be known and constant. N^* in Eq. (3.1-19) corresponds to N_m, D^* to D_m, and n is given as Eq. (3.6-12).

$$n(D; D_m) = D^\mu \exp\left[-\frac{(4+\mu)D}{D_m}\right], \tag{3.6-12}$$

3.6.2.7.2 Mixed phase and solid phase

It is assumed that when all the particles are melted to liquid particles, the drop size distribution obeys Eq. (3.6-11). With Eq. (3.6-10), Eq. (3.6-11) can be expressed by D_b and ρ_b.

$$N(D_b)\, dD_b = N_m\, \rho_b^{(\mu+1)/3} D_b^\mu \exp\left[-\frac{(4+\mu) D_b\, \rho_b^{1/3}}{D_m}\right] dD_b \tag{3.6-13}$$

Eq. (3.6-13) can be modified to Eq. (3.6-14), if we consider the difference of number density caused by the difference of falling velocity with a Non-break-up and Non-coalescence model (N/N model).

$$N(D_b)\, dD_b = N_m\, \rho_b^{(\mu+1)/3} D_b^\mu \frac{V_R(D_{\mathrm{melt}})}{V_b(D_b)} \exp\left[-\frac{(4+\mu) D_b\, \rho_b^{1/3}}{D_m}\right] dD_b \tag{3.6-14}$$

3.6.2.8 R-D_m relation

R-D_m relations are assumed for stratiform and convective rain separately to calculate k and Z_e for a given R at both Ka and Ku bands. In version 3 algorithm, default k-Z_e relations were given as the basic relations to constrain the DSD parameters. However, since these relations were derived from the R-Λ relation (Kozu et al. 2009) by assuming a power law between k and Z_e, the resultant k-Z_e relations did not reproduce the original R-Λ relations exactly. The discrepancy was apparent especially in heavy rainfall cases for KaPR and solid particles, and caused a serious error in the algorithm. Since the R-D_m relation is directly derived from the R-Λ relation without approximation, we have consistent DSD parameters to relate k, Z_e at both frequencies and R.

The actual R-D_m relations adopted are $R = 0.401\, \varepsilon^{4.649} D_m^{6.131}$ for stratiform precipitation and $R = 1.370\, \varepsilon^{4.258} D_m^{5.420}$ for convective precipitation, where ε is an adjustment factor and is the same with ε used in version 03 and in the TRMM/PR algorithm. The same equations are used for all range bins irrespective of its phase, while D_m is defined for particle size when melted.

3.6.3 Phase

An index called "`phase`" is introduced. "`phase`" is an integer-type variable. If "`phase`" is equal to or greater than 200, particles are in liquid `phase`, and the difference between the value of

"phase" and 200, i.e., (phase−200), indicates the particle temperature in unit of degrees C. For example, if "phase" is 210, the particle temperature is 10 degrees C. If "phase" is between 100 and 200, particles are in a bright band. 100 is for the top of bright band, 150 is for the peak of bright band, and 200 is for the bottom of bright band, where all particles are melted in liquid phase. 125 and 175 are used for intermediate positions, but other numbers between 100 and 200 are not used. If "phase" is smaller than 100, particle is partly or fully frozen, and the difference between "phase" and 100 means particle temperature in unit of degrees C. For example, if "phase" is 80, particle temperature is −20 degrees C.

3.6.4 DSD database in V05

ε used in the R-D_m relations was assumed to follow the log-normal distribution. In the version 04, the average (μ) and the standard deviation (σ) of $\log_{10}\varepsilon$ is set at 0 and 0.1, respectively. In the version 05, the statistics of ε is calculated based on the outputs of ε at dual-frequency pixels by the dual-frequency algorithm. The average and the standard deviation of $\log_{10}\varepsilon$ are calculated for each precipitation type (stratiform, convective), for each land surface type (land, ocean, all), each 5 by 5 degree grid box, and each month. Not only grid values but global values are prepared. The purpose of the DSD database is to make estimates by KuPR algorithm be closer to those by the dual-frequency algorithm. DSD database is applied only to single-frequency pixels (all pixels of the single-frequency algorithms and pixels in outer swath or in interleaved scan of the dual-frequency algorithm). In the dual-frequency pixels of the dual-frequency algorithm, μ and σ are set to 0 and 0.1, respectively, as in the version 04.

3.7 Surface Reference Technique (SRT) module

3.7.1 Objective of the algorithm

The primary purpose of the surface reference technique (SRT) is to compute the path-integrated attenuation (PIA), or simply path attenuation, using the radar return from the surface. The SRT rests on the assumption that the difference between the measurements of the normalized radar surface cross section, σ^0 or NRCS, in dB, within and outside the rain provides an estimate of the PIA.

For the dual-frequency Precipitation Radar (DPR), the method is to be applied to the Ku-band as well as the Ka-band and Ka-band high sensitivity (KaHS) data. The basic set of output products for each of the channels (Ku, Ka and KaHS) consist of the path attenuation estimate (when rain is present) and an associated reliability. As in the TRMM algorithm 2a21, version 07, the primary output for each channel will be an effective or 'best' PIA estimate. In addition, however, path attenuation estimates corresponding to specific types of surface reference data are also generated. These are described below. The standard PIA estimates will be produced by processing the data from the 3 channels (Ku, Ka, Ka_HS) independently. However, to take advantage of the correlation in the NRCS at the two frequencies, dual-frequency-derived path attenuations at Ku- and Ka-band will also be generated at fields of view where both Ku- and Ka-band returns are made. These are referred to as the DPR-Ku and DPR-Ka estimates.

For each 'channel' (Ku, Ka, KaHS, DPR-Ku, DPR-Ka), a spatially or temporally-averaged estimate of the rain-free normalized radar surface cross section (σ^0 or NRCS) is used as a reference value for computing the PIA. For the DPR channel, the rain-free reference data is obtained from a difference between the Ka- and Ku-band σ^0 data. The algorithm computes up to six estimates of path attenuation, corresponding to five different σ^0 reference estimates for each of the five channels. An effective PIA is obtained by weighting the individual estimates by a factor that is inversely proportional to the variance of the estimate. The reference estimates are described below:

Along-track Spatial Average: The along-track spatial average is obtained from an average of the N_s most recent rain-free σ^0 measurements at same angle bin and with the same surface type (nominally, $N_s = 8$) as the rain pixel of interest.

Cross-Track Spatial Average: The cross-track reference data set is generated from a quadratic fit of the along-track spatial average data over the angle bins within the cross-track swath. This is explained below in more detail. For the Ku-band, separate fits are done for the inner ($\theta \leq \theta_0$) and outer swath ($\theta > \theta_0$) where ($\theta \leq \theta_0$). This separate fitting is also done for the Ka-band data for data taken after May 2018; for earlier times, only a single fit is done for the Ka-band data in the inner swath.

Temporal Average: The temporally-averaged data, which are given in the form of look-up

tables, presently consists of statistics (sample mean, sample mean square and number of data points) of rain-free σ^0 data at $0.5^\circ \times 0.5^\circ$ latitude-longitude cells, separated into incidence angle. The look-up tables have been derived from the Ku, Ka, KaHS and differential (σ^0(Ku)- σ^0(Ka)) data categorized into 25, 13, 12, and 13 angle bins, respectively, for months up to May 2018. For months after May 2018, where KaMS and KaHS have been combined into a single full-swath set data set, KaFS, and Ku, now labeled KuFS, the number of angle bins for KuFS, KaFS and dual-frequency FS are 25 for all. Each of the files described above are further subdivided into 5 surface-dependent look-up tables that correspond to ocean, land, coast, snow-over-land, and sea-ice surface types for a total of 20 look-up tables for each 3-month period (DJF, MAM, JJA, SON). Product version 3 and earlier do not use the temporal reference data; version 4 and later versions use the standard temporal reference file but not the 'light-rain' temporal reference file. For version 5 and above, the original 3 surface types (ocean, land, coast) were expanded into 5 surface types, with the additional categories of snow-over-land and sea-ice. As noted above, because of the change in the Ka-band swath, the make-up of the tables changes from pre-scan change to post scan change. Further details can be found in Section 3.7.3.

For versions 06X and higher, the number of angle bins for Ku, Ka and differential (σ^0(Ku)-σ^0(Ka)) databases is 25, corresponding to incidence angles from 0 to 18° with angle bin interval of 0.75°. For months after May 2018, the KaHS reference data are set to missing.

Forward-Backward Processing

Different estimates for along-track and cross-track methods are obtained from forward and backward processing. This provides up to five estimates: *Forward* and *Backward Along-track*, *Forward* and *Backward Cross-track*, and *the Temporal reference* for each rain observation. Note that the cross-track estimates are only available over ocean since the fitting procedure does not work well over land. (Cross-track fitting over land might be effective for the difference field, $\sigma^0_{\mathrm{NR}}(\mathrm{Ku}) - \sigma^0_{\mathrm{NR}}(\mathrm{Ka})$. This, however, has not been implemented.) Note also that only the first of the temporal files have been implemented. The 'light-rain' temporal reference file is not defined up to and including version 07 of the code. The different reference estimates are filtered according to various criteria. For example, an along-track estimate isn't used if the location of the reference data is too far from the observed rain point. A temporal estimate is considered invalid if the number of rain-free observations at a particular angle for the reference cell is too few. The surviving estimates are weighted by the inverse of their associated variance. The weighted estimates are combined to yield an effective PIA for each of the channels, (`pathAtten_Ku`, `pathAtten_Ka`, `pathAtten_KaHS`). The individual estimates, (`PIAalt_Ku`, `PIAalt_Ka`, `PIAalt_KaHS`), and their weights, (`PIAweight_Ku`, `PIAweight_Ka`, `PIAweight_KaHS`) are included in the output product. A discussion of the weights and the effective PIA, variance, and reliability factor is given in

Section 3.7.5.

In the spatial (along-track) surface reference data set, the mean and standard deviation of the NRCS are calculated over a running window of N_s fields of view before rain is encountered (currently, $N_s = 8$). These operations are performed separately for each of the 49 incidence angles for the Ku-band data, corresponding to the cross-track scan from $-18°$ to $+18°$ with respect to nadir. The Ka and DPR and KaHS channels are treated in the same way. Prior to May 2018, the KaMS, DPR and KaHS consisted of 25, 25 and 24 angle bins, respectively. After May 2018, the KaFS and DPRFS consist of 49 angle bins, matched to those of the KuFS. Note that for the DPR or differential channel, the same procedure is used except that the reference data are formed from the $\sigma^0_{\rm NR}(\text{Ku}) - \sigma^0_{\rm NR}(\text{Ka})$ data. Prior to May 2018, the data consisted of the 25 angle bins of the inner swath; after May 2018, Ka-band data, and therefore the differential data, became available over the full swath of 49 angle bins.

Basic Processing

When rain is encountered, at incidence angle θ_i, the means and standard deviations of the reference σ^0 values are retrieved from the along-track spatial (forward or backward), the cross-track (forward or backward) and the temporal surface reference data sets. If a valid surface reference data set exists for one of the above estimates, then, denoting this by the jth estimate, the 2-way path attenuation (PIA) is computed from the equation:

$$\text{PIA}_j(\theta_i) = \langle \sigma^0_{\rm NR}(\theta_i) \rangle_j - \sigma^0(\theta_i),$$

where $\langle \sigma^0_{\rm NR}(\theta_i) \rangle_j$ is the mean of the jth rain-free reference estimate and $\sigma^0(\theta_i)$ is the value of the apparent normalized radar surface cross section at the raining field of view of interest.

To obtain information as to the reliability of the jth PIA estimate we consider the ratio of the PIA, as derived in the above equation, to the sample standard deviation associated with the jth rain-free reference data set. Labeling this standard deviation by σ_j, the reliability factor of the jth PIA estimate is defined as:

$$\texttt{reliabFactor}_j = \frac{\text{PIA}_j}{\sigma_j},$$

which is equal to the inverse of the coefficient of variation of the estimate.

The effective PIA, $\text{PIA}_{\rm eff}$, and the corresponding reliability factor, $\text{Rel}_{\rm eff}$, can be expressed in similar ways. From Section 3.7.5, we have:

$$\text{PIA}_{\rm eff} = \frac{\sum u_j \text{PIA}_j}{\sum u_j},$$

$$\text{Ref}_{\rm eff} = \frac{\sum u_j \text{PIA}_j}{\sqrt{\sum u_j}},$$

where u_j is the inverse of the variance, σ_j^2, associated with the jth reference data set:

$$u_j = \frac{1}{\sigma_j^2}.$$

Note that the summations are assumed to range over all valid reference data sets, even if the PIA's are negative. Over land, there can be a maximum of 3 valid reference data sets (forward and backward along-track and the temporal files) while over ocean there can be as many as five since the forward/backward cross-track reference data sets are defined over ocean but not land.

A generalized form of the above equations can be defined, where the various quantities are replaced by the difference between the Ka and Ku-band σ^0 values. Letting $\delta X = X(\mathrm{Ka}) - X(\mathrm{Ku})$, for any variable X, then for the differential PIA, the above equation becomes

$$\delta\mathrm{PIA}_j(\theta_i) = \langle \delta\sigma^0_{\mathrm{NR}}(\theta_i) \rangle_j - \delta\sigma^0(\theta_i).$$

The individual and effective dual-frequency estimates at Ku- and Ka-band are then derived from $\delta\mathrm{PIA}_j$ from

$$\mathrm{PIA}_j(\mathrm{Ku}, \theta_i) = \frac{\delta\mathrm{PIA}_j(\theta_i)}{p-1},$$
$$\mathrm{PIA}_j(\mathrm{Ka}, \theta_i) = \frac{p\ \delta\mathrm{PIA}_j(\theta_i)}{p-1},$$

where p is the ratio of the path attenuation at Ka-band to that at Ku-band and is nominally set to 6.

Updates for Versions 06 and later

A number of new variables have been introduced in version 06. The primary set of variables is associated with a hybrid estimate of the path attenuation that consists of a weighted sum of the SRT, Hitschfeld-Bordan (HB), and the standard dual-wavelength method (DW). For the single-frequency Ku-band and Ka-band estimates, only the SRT and HB results are used since the DW method is inapplicable.

The basic idea follows closely that used for the SRT itself: if n independent estimates of a quantity A are available and if the estimate is written in the form

$$\hat{A} = \sum_{j=1}^{n} w_j A_j,$$

then the minimum variance estimate of A is obtained by choosing the weighting factors, w_j, to be inversely proportional to the variance of A_j:

$$w_j = \frac{\tau_j}{\sum_{j=1}^{n} \tau_j}; \qquad \tau_j = \mathrm{var}(A_j).$$

It can also be shown that the variance of $\hat{A}$ is given by

$$\mathrm{var}(\hat{A}) = \frac{1}{\sum_{j=1}^{n} \tau_j}.$$

We can write explicitly the expression for the differential hybrid path attenuation as the weighted sum of the differential estimates from the SRT, HB and DW:

$$\delta A_{\mathrm{HY}} = w_{\mathrm{SRT}} \delta A_{\mathrm{SRT}} + w_{\mathrm{HB}} \delta A_{\mathrm{HB}} + w_{\mathrm{DW}} \delta A_{\mathrm{DW}}.$$

As in the SRT case, the differential PIA is converted into Ku- and Ka-band results by assuming the ratio of Ka/Ku path attenuation is 6 so that, for the DPR results, $A_{\mathrm{HY}}(\mathrm{Ku}) = 0.2 \times \delta A_{\mathrm{HY}}$ and $A_{\mathrm{HY}}(\mathrm{Ka}) = 1.2 \times \delta A_{\mathrm{HY}}$.

For the single-frequency case, the hybrid path attenuation is given by

$$A_{\mathrm{HY}} = w_{\mathrm{SRT}} A_{\mathrm{SRT}} + w_{\mathrm{HB}} A_{\mathrm{HB}}.$$

Note that the SRT estimates in the above equations should be interpreted as the effective PIA estimate described earlier, which itself is a weighted sum of the SRT estimates using different rain-free reference NRCS estimates. Since the procedures used to derive the SRT and hybrid estimates are the same, the formula for the reliability factor, the ratio of the mean of A_{HY} to the standard deviation, can be written in the same form:

$$\mathrm{Rel}_{\mathrm{eff}} = \frac{\sum \tau_j A_j}{\sqrt{\sum \tau_j}},$$

where

$$A_1 = A_{\mathrm{SRT}}; \qquad \tau_1 = \frac{1}{\mathrm{var}(A_{\mathrm{SRT}})}, \quad A_2 = A_{\mathrm{HB}}, \qquad \tau_2 = \frac{1}{\mathrm{var}(A_{\mathrm{HB}})}.$$

The A_{HB} is given by the following formula:

$$A_{\mathrm{HB}}(r) = -\left(\frac{10}{\beta}\right) \log_{10}[1 - \zeta(r)]^{1/\beta},$$

$$\zeta(r) = 0.2\beta \ln 10 \int_0^r \alpha(s) Z_m^{\beta}(s)\, ds; \quad k = \alpha Z^{\beta},$$

where the range, r, is taken to be the last clutter-free range bin above the surface and Z_m is the measured radar reflectivity factor. The expression for δA_{HB} is obtained simply by taking the difference between A_{HB} at Ka-band and at Ku-band. The only difference between these is in the values that are taken for α and β as shown in the Table 3.7.1.

An expression for the standard dual-wavelength PIA estimate, DW, can be obtained by noting that the differential path attenuation is equal to the difference between the measured and actual dual-frequency ratio, DFR

$$\delta A_{\mathrm{DW}}(r) = \mathrm{DFR}_{\mathrm{m}}(r) - \mathrm{DFR}(r).$$

Table 3.7.1: Coefficients in the k-Z relationship

	Snow	Mixed Phase	Rain
α(Ku)	5.97×10^{-5}	1.39×10^{-3}	7.60×10^{-4}
α(Ka)	1.2×10^{-3}	9.51×10^{-3}	5.47×10^{-3}
β(Ku)	0.661	0.661	0.661
β(Ka)	0.691	0.691	0.691

The first term is the difference of the measured reflectivity factors at Ku- and Ka-band and is a measured quantity. The second term in the above equation, the true DFR, is not known but we approximate it by using the measured $\mathrm{dB}Z_m(\mathrm{Ku})$ so that

$$\delta A_{\mathrm{DW}}(r) = [\mathrm{dB}Z_m(\mathrm{Ku}, r) - \mathrm{dB}Z_m(\mathrm{Ka}, r)] - \mathrm{DFR}(\mathrm{dB}Z_m(\mathrm{Ku}, r)).$$

As in the HB method, the range is evaluated at the clutter-free gate nearest the surface.

To implement the hybrid method, it is necessary to approximate the variances for the SRT, HB and DW. Expressions for the variance of the SRT are given above and in Section 3.7.5; for the HB and DW, measured drop size distributions are used to estimate the errors.

Updates for V06X and later

On May 21, 2018, the 24 beams of the KaHS interleaved swath were redirected to the outer swath of the KaMS swath so that, for data products since then (V06X and later), the 49 beams of the Ku and Ka-band are approximately aligned in space and time. The angle bin dimension of all Ka-band products was changed from 25 to 49 and the interleaved KaHS data are set to missing. For the temporal look-up tables, the number of angle bins for the Ka-band was increased from 13 to 25. The KaHS database exists only for months before May 2018.

Updates for V07

Updates from V06X to V07 are minor. The most significant change is that the codes were modified so that orbits both before and after the scan change in May 2018 can be processed with the same codes. Changes were also made to the expressions for the variances of the Hitschfeld-Bordan (HB), standard dual-wavelength (DW) and SRT methods by adding sampling errors. These sampling errors arise from the finite number of samples at each range gate (˜100). The effects of sampling errors on the DW and SRT are stronger than on HB so the weights are shifted slightly away from DW and SRT estimates and toward HB estimate. Despite this, the changes in weights are generally small as the sampling error is usually smaller than the other sources of error. Details can be found in Section 3.7.5.

3.7.2 Input Variables

`sigmaZeroMeasured` [real*4]: Normalized backscattering radar cross section of the surface (dB) (NRCS) at Ku-band, Ka-band or KaHS for the angle bins in the radar scan (unitless), from the Preparation Module.

`flagPrecip` [integer*2]: rain/no-rain flag from Preparation Module.

`localZenithAngle` [real*4]: Incidence angle from Preparation Module.

`landSurfaceType` [integer*2]: surface type from Preparation Module. These are modified so that only 5 surface classes are defined: ocean, land, coast, snow-over-land and sea-ice.

`lat, lon` [real*4]: latitude, longitude of the center of the FOV from Preparation Module.

`zFactorMeasured`(176/88) [real*4]: range profile of the measured radar reflectivity factor in dB from the Preparation Module. For the KuNS and KaMS the range resolution is 0.125 km with 176 gates while for the KaHS the range sampling is 0.25 km with 88 gates.

`binStormTop` [integer*2]: range bin number corresponding to the first gate at which precipitation is detected, from the Preparation Module.

`snowIceCover` [integer*2]: an indicator flag for the presence of snow-over-land or sea-ice, from the Preparation Module.

`snRatioAtRealSurface` [real*4]: signal-to-noise ratio at the surface range gate, from the Preparation Module.

`binNode`(5) [integer*2] provides the range bins corresponding to storm height, top, center and bottom of melting layer and lowest bin of the precipitation, from the DSD Module.

`binBBTop, binBBBottom` [integer*2]: bin numbers corresponding to the top and bottom of the bright-band, from the Classification Module.

3.7.3 Output Variables

Five sets of path-integrated attenuation and associated quantities are produced. These consist of: single- and dual-frequency derived Ku-band attenuation, single- and dual-frequency derived Ka-band attenuation and a single-frequency derived KaHS-band attenuation. The single-frequency Ku-band attenuation and associated products are stored in the 49 angle bins (full swath) of the Ku-NS structure.

Prior to version 06X, the single-frequency KaMS/KaHS products were stored in the KaMS/KaHS structures and were available for the 25 angle bins of the inner swath/ 24 angle bins of the interleaved swath. The dual-frequency PIA products for Ku- and Ka-band are stored in the DPR-NS and DPR-MS structures, respectively, and are produced for the 25 angle bins comprising the inner swath. (For the DPR-NS structure, the dual-frequency derived

Ku-band path attenuations are contained in angle bins 13 through 37.) Note that for versions prior to V06X, the single-frequency Ku-band results are used in the outer swath of the DPR-NS structure, covering the angle bins from 1 to 12 and from 38 to 49.

For versions 06X and later, the products are stored in the KuFS, KaFS, DPRFS (FS=full scan) structures consisting of 49 angle bins. For products generated after May 2018, all angle bins are filled. For V07 and for times before May 2018 only the central angle bins (13-37) will contain data and the remainder will be filled with missing values. The Ka-band interleaved data will be retained in the KaHS structure for data taken before May 2018. After May 2018, the variables will be filled with missing values.

The output variables for each of these channels are defined below.

`pathAtten` [real*4]: The estimated effective 2-way path-attenuation in dB where

$$\texttt{pathAtten} = 2\int_0^r k(s)\,ds,$$

where $k(s)$ is the attenuation coefficient in dB/km where the integral is taken from the storm top to the surface. The path attenuation is often designated as the PIA, the path-integrated attenuation. In the notation used above and in Section 3.7.5:

$$\texttt{pathAtten} = \mathrm{PIA}_{\mathrm{eff}} = \frac{\sum u_j \mathrm{PIA}_j}{\sum u_j},$$

where u_j is equal to the inverse of the variance associated with the jth reference data point:

$$u_j = 1/\sigma_j^2.$$

`PIAalt(6)` [real*4]: The path-integrated attenuation (dB) from the jth estimate, (PIA_j in the notation above), where

`PIAalt`(1) = PIA derived from the forward along-track spatial reference data

`PIAalt`(2) = PIA derived from the backward along-track spatial reference data

`PIAalt`(3) = PIA derived from the forward hybrid/cross-track reference data

`PIAalt`(4) = PIA derived from the backward hybrid/cross-track reference data

`PIAalt`(5) = PIA derived from standard temporal reference data

`PIAalt`(6) = PIA derived from the light-rain temporal reference data

Note that for product versions 1 through 3, the standard temporal path-attenuation estimate, `PIAalt`(5), is set to missing but is defined for versions 4 and higher. For product versions 1 through 7, the light-rain temporal estimate, `PIAalt`(6), is set to missing. Note also that the forward/backward hybrid/cross-track path attenuations are defined only over ocean and are set to missing over land.

`PIAweight(6)`[real*4]: The weights, w, of the individual PIA estimates used in deriving the effective PIA. The weight for a particular PIA estimate is proportional to the inverse

of the error variance associated with the method. The sum of the weights should equal one. As with `PIAalt`(6), `PIAweight`(6) is set to missing.

$$w_j = \frac{1}{\sigma_j^2}\frac{1}{\sum \frac{1}{\sigma_j^2}} \equiv \frac{u_j}{\sum u_j},$$

where

$$u_j = 1/\sigma_j^2,$$

$$\sum w_j = 1.$$

`reliabFlag`[integer*2]: Reliability Flag for the PIA_{eff} estimate,
= 1 if $\text{Rel}_{\text{eff}} > 3$; PIA_{eff} estimate is considered reliable
= 2 if $3 \geq \text{Rel}_{\text{eff}} > 1$; PIA_{eff} estimate is considered marginally reliable
= 3 if $\text{Rel}_{\text{eff}} \leq 1$; PIA_{eff} is unreliable
= 4 if $\text{SNR}_{\text{at surface}} < 2$ dB; provides a lower bound to the path-attenuation
= 9 (no-rain case)

`reliabFactor`[real*4]: Reliability Factor for the effective PIA estimate, `pathAtten`. This is defined as:

$$\texttt{reliabFactor} = \text{Rel}_{\text{eff}} = \left(\sum u_j\right)^{-1/2} \sum u_j \text{PIA}_j.$$

`RFactorAlt(6)`[real*4]: The reliability factors associated with the individual PIA estimates in `PIAalt`. As with `PIAalt`(6), `RFactorAlt`(6) is set to missing.

$$\texttt{RFactorAlt}(j) = \text{Rel}_j = \text{PIA}_j/\sigma_j; \qquad j = 1, \ldots, 6.$$

`refScanID(2,2)`[integer*2]: `refScanID` gives the number of scan lines between the current scan and the beginning (or end) of the along-track reference data at each angle bin. The values are computed by the equation: Current Scan Number − Reference Scan Number. The values are positive for the Forward estimates and negative for the Backward estimates. The Fortran indices are:

1,1 - Forward - Near reference
2,1 - Forward - Far reference
1,2 - Backward - Near reference
2,2 - Backward - Far reference

To illustrate, consider the following example. At a certain incidence angle assume that rain is present at scan numbers from 100 to 105 and from 110 to 120. At scan number 112,

refScanID(1,1)=3, refScanID(2,1)=16; i.e., the eight rain-free NRCS data points, used to estimate the mean and standard deviation of the rain-free NRCS, begin at scan $112 - 16 = 96$ and end at scan $112 - 3 = 109$. These numbers provide information on the distance (in terms of the numbers of scans where 1 scan $\sim$ 5 km) of the rain-free reference data from the rain pixel of interest. See Section 3.7.6 for further details.

`PIAhybrid`[real*4] (V06): For the dual-frequency output, PIAhybrid (dB) is a weighted sum of the path attenuations from the SRT, HB, and DW methods. For the single-frequency outputs, PIAhybrid is a weighted sum of the SRT and HB methods.

`reliabFlagHY`[integer*2] (V06): `reliabFlagHY` is the reliability flag for the PIAhybrid and is defined in the same way as reliabFlag where PIAeff is replaced by PIAhybrid.

`reliabFactorHY`[real*4] (V06): `reliabFactorHY` is the reliability factor associated with the PIAhybrid estimate and is defined as the mean over the standard deviation of the estimate.

`stddevHY`[real*4] (V06): `stddevHY` is the standard deviation (dB) of the hybrid estimate of path attenuation.

`PIAhb`[real*4] (V06): `PIAhb` (dB) is the path attenuation estimate derived from the Hitschfeld-Bordan equation.

`PIAdw`[real*4] (V06): `PIAdw` (dB) is the path attenuation estimate derived from the standard dual-wavelength method.

`zeta`[real*4] (V06): `zeta` (unitless) is a parameter in the Hitschfeld-Bordan equation.

`stddevEff`(3)[real*4] (V06):

`stddevEff`(1) contains the standard deviation of the `PIAeff` (i.e., the composite or effective SRT or hybrid path attenuation estimate). It is given by $\sigma_1 = \left(\sum 1/\sigma_e^2, j\right)^{-1/2} = \left(\sum u_j\right)^{-1/2}$. It is important to note that in the definition of the reliability factor, it is this standard deviation that is used. In other words, with the notation in Section 3.7.5, we have $\sigma_1 = \sigma_{\text{eff}}$.

`stddevEff`(2) is a weighted root mean square error and provides a measured of the error of the individual PIA estimates from the effective PIA estimate. It is given by $\sigma_2 = [\sum w_j (A_{\text{eff}} - A_j)^2]^{1/2}$.

`stddevEff`(3) is given by $\sigma_3 = [\sigma_1^2 + \sigma_2^2]^{1/2}$.

`PIAdw`[real*4] (V06X): `PIAdw`(dB) is the path-integrated attenuation as estimated from the standard dual-wavelength technique.

`PIAHYweight`(3)[real*4] (V06X): `PIAHYweight` contains the weights of the SRT, HB, and DF methods in the calculation of the hybrid PIA. For single-frequency results, since the DF method is inapplicable and `PIAHYweight`(3)=0.

Details of the description for variables are described in Appendix or in the following websites:

`https://www.eorc.jaxa.jp/GPM/en/archives.html`

`https://arthurhou.pps.eosdis.nasa.gov/GPMprelimdocs.html`

3.7.4 Temporal Reference Files

Temporal reference files or look-up tables (LUT) are needed to estimate the path attenuation using temporal reference data. The temporal reference data are simply the mean and standard deviations of previously measured rain-free σ^0 data that have been categorized with respect to incidence angle and location. In particular, the LUTs contain the rain-free sample mean and standard deviation of the σ^0 data at each (lat, lon, incidence angle) bin, where the incidence angles, with respect to the radar, consist of nadir, ± 0.750, ± 1.50, $\cdots$ up to the maximum angle. Also stored is the number of σ^0 data points acquired at the particular bin in question and used in computing the statistics.

For version 4 of the algorithm, a $0.5^\circ \times 0.5^\circ$ grid is used which covers the GPM-DPR latitude range of 67°S–67°N so that 268 latitude bins and 720 longitude bins are needed. Four LUTs are derived from DJF (December-January-February), MAM, JJA, SON input data. For each of these, six LUTs are prepared corresponding to ocean, land, coast, snow-over-land and sea-ice and all background types. Note that separate tables are produced for σ^0(Ku), σ^0(Ka), σ^0(KaHS), and the difference field, $[\sigma^0(\text{Ka}) - \sigma^0(\text{Ku})]$ so that a total of 20 LUT's are generated for each three-month season.

As noted earlier, two types of temporal reference files have been defined: the standard temporal reference files as described above and a light-rain reference file which, as of version 07 of the code, is undefined. The results derived from the standard temporal reference file can be found in the fifth entry in the variables `PIAalt`, `PIAweight`, `RFactorAlt`. The sixth entry for these variables, corresponding to the 'light-rain' temporal reference file, is set to missing.

For versions 06X and later, the number of angle bins needed to store the Ka-band and the differential (Ka-Ku) reference data has increased from 13 to 25 so that the number of angle bins for the Ku-, Ka- and the difference fields are the same.

3.7.5 Definitions of the Effective PIA, Variance and Reliability Factor

As noted above, multiple estimates of PIA can be generated for each channel corresponding to different surface reference estimates. Specifically, we have the following situation

$$\text{PIA}_j = \langle \sigma^0_{\text{NR}} \rangle_j - \sigma^0_R, \tag{3.7-1}$$

where the first term on the right-hand side is the jth surface reference value and the second term is the apparent NRCS in rain. Associated with the jth reference data set is a variance, σ_j^2:

$$\mathrm{var}(\mathrm{PIA}_j) = \mathrm{var}[\langle\sigma_j^0\rangle] + \mathrm{var}(\sigma_R^0) \equiv \sigma_j^2 + \sigma_s^2. \tag{3.7-2}$$

Note that the second term in Eq. (3.7-2) corresponds to the variance due to sampling error in the measurement of σ^0 in rain. With a logarithmic receiver, it is equal to $5.57^2/N$ where N is the number of independent samples used to estimate this quantity. For the DPR, N depends on the incidence angle and varies between 100 and 109. It should be noted that the sampling error was neglected versions up to V06X but is included in V07. From these PIA estimates we want to obtain an effective or 'best' PIA. We assume it can be written in the form:

$$\mathrm{PIA}_{\mathrm{eff}} = \sum w_j \mathrm{PIA}_j, \tag{3.7-3}$$

where the weights, w_j, are such that

$$\sum w_j = 1. \tag{3.7-4}$$

We assume that the individual PIA estimates are statistically independent so that the variance of $\mathrm{PIA}_{\mathrm{eff}}$ is:

$$\mathrm{var}(\mathrm{PIA}_{\mathrm{eff}}) = \sum w_j[\sigma_j^2 + \sigma_s^2]. \tag{3.7-5}$$

To simplify the equations below, we define an effective variance, $\sigma_{e,j}^2$

$$\sigma_{e,j}^2 = \sigma_j^2 + \sigma_s^2. \tag{3.7-6}$$

To minimize Eq. (3.7-5), subject to the side condition given by Eq. (3.7-4), we use the method of Lagrange multipliers where the expression

$$\sum w_j^2 \sigma_{e,j}^2 + \lambda(\sum w_j - 1) \tag{3.7-7}$$

is minimized with respect to the weights, w_j. Taking the partial derivatives of Eq. (3.7-7) with respect to w_i, then

$$2w_j\sigma_{e,j}^2 + \lambda = 0 \quad \Rightarrow \quad w_j = -\lambda/2\sigma_{e,j}^2. \tag{3.7-8}$$

Also, using Eq. (3.7-4) gives

$$\sum w_j = -(\lambda/2)\sum(1/\sigma_{e,j}^2) \quad \Rightarrow \lambda = -2/\sum(1/\sigma_{e,j}^2). \tag{3.7-9}$$

Substituting Eq. (3.7-9) into Eq. (3.7-8) gives an expression for the weights:

$$w_j = u_j/\sum u_j, \tag{3.7-10}$$

where

$$u_j = 1/\sigma_{e,j}^2. \tag{3.7-11}$$

The effective PIA is then

$$\mathrm{PIA_{eff}} = \left(\sum u_j\right)^{-1} \sum u_j \mathrm{PIA}_j. \tag{3.7-12}$$

If we define the reliability factor, Rel, as the ratio of the PIA to the standard deviation of the reference estimate, then for the jth reference estimate, we can write:

$$\mathrm{Rel}_j = \mathrm{PIA}_j/\sigma_{e,j}. \tag{3.7-13}$$

To apply this definition to the present situation, we define $\mathrm{Rel_{eff}}$ by the equation:

$$\mathrm{PIA_{eff}} = \left(\sum u_j\right)^{-1} \sum u_j \mathrm{PIA}_j \equiv \sigma_{\mathrm{eff}} \mathrm{Rel_{eff}}. \tag{3.7-14}$$

Computing $\mathrm{Rel_{eff}}$ requires a value for the standard deviation of the effective PIA. This can be found by substituting Eq. (3.7-9) into Eq. (3.7-5) and by noting that $\sigma_{\mathrm{eff}}^2 = \mathrm{var}(\mathrm{PIA_{eff}})$. This gives

$$1/\sigma_{\mathrm{eff}}^2 = \sum (1/\sigma_{e,j}^2) \quad \Rightarrow \quad \sigma_{\mathrm{eff}}^2 = \left[\sum (1/\sigma_{e,j}^2)\right]^{-1} = \left(\sum u_j\right)^{-1}. \tag{3.7-15}$$

Using Eq. (3.7-15) in Eq. (3.7-14) gives

$$\mathrm{Rel_{eff}} = \left(\sum u_j\right)^{-1/2} \sum u_j \mathrm{PIA}_j. \tag{3.7-16}$$

Eq. (3.7-10), (3.7-12) and (3.7-16) define, respectively, the weights, effective PIA, and effective reliability factor that are computed in the algorithm. In addition to the σ_{eff} error defined above in Eq. (3.7-15), a kind of RMS error can be defined by the following equation:

$$\mathrm{RMS}_{\mathrm{eff}}^2 = \sum_j w_j [\mathrm{PIA_{eff}} - \mathrm{PIA}_j]^2, \tag{3.7-17}$$

where the summation runs over the j estimates of PIA. Note that when the weighting factors are all the same then this reduces to the usual definition of RMS error.

There are several issues related to these equations. For example, what should be done if none of the reference data sets exist? This situation can occur for measurements over small islands or small bodies of water or at coastal fields of view. For example, over a small island, there may be an insufficient number of non-raining fields of view adjacent to the rain area to form a valid spatial reference. In most cases, the temporal reference data set would be used and the other reference estimates would be discarded. However, in some cases, there may be an insufficient number of data points in the temporal file to provide a valid estimate. In this case, a flag is

set indicating that no valid reference data are available and all the output variables are set to -9999.9

A somewhat different situation occurs if some of the reference data sets exist but all yield a negative PIA. For these cases, the individual variances will exist so that Rel_{eff}, PIA_{eff} and the effective variance should all exist. Note that for these cases Rel_{eff}, PIA_{eff} will be negative but the effective variance will be positive, as it should be.

A third type of situation occurs if one or more of the PIA estimates are positive and one or more of the PIA estimates are negative. In this case, the negative PIAs will be included in the definition of PIA_{eff}. In general, as long as the reference data is considered to be valid, the PIA will be used even if the value is negative.

According to these scenarios, there will be only one type of raining situation where the output variables will need to be set to some default value and this occurs when none of the reference data sets exist or are valid. This is expected to be a small fraction relative to the total number of rain cases.

3.7.6 Excluding Spatial Reference Data based on the `refScanID` Variable

This section relates to determining the circumstances under which we assume a spatial reference estimate to be valid. For the forward-going spatial reference, reference data will almost always exist. An exception is if rain is encountered at the beginning of the orbit before $N_s(= 8)$ rain-free fields of view have been measured at a particular incidence angle. A similar exception occurs for the backward spatial methods: this occurs, however, at the end of the orbit rather than the beginning. In all other cases, forward and backward spatial reference data should exist. The question is how should we exclude spatial reference estimates if the data are taken at locations far from the raining area. To make this definite we implement the following rules in the algorithm.

See definition of `refScanID`(2,2) in Section 3.7.3.

The forward along-track spatial reference at angle bin j will be assumed to be invalid if:

$$|\texttt{refScanID}(2, 1, j)| > 50.$$

Similarly, the backward along-track spatial reference at angle bin j will be assumed to be invalid if:

$$|\texttt{refScanID}(2, 2, j)| > 50.$$

The above conditions are equivalent to stating that, for a particular incidence angle, all the spatial reference data must be taken within 50 scans of the scan at which rain is encountered. The criteria for the hybrid cross-track are more complicated because two quadratic fits are used for the inner and outer portion of the swath. Nominally, we will assume that if there are 15 or

more angle bins in the inner portion of the swath for which:

$$|\texttt{refScanID}(2,1,j)| \leq 50,$$

then the forward hybrid cross-track method will be applied. Similarly in the outer portion of the swath, if there are 15 or more angle bins in this portion of the swath for which:

$$|\texttt{refScanID}(2,1,j)| \leq 50,$$

then the forward hybrid cross-track method will be applied. Application of the backward hybrid cross-track will follow the same rule, based on $\texttt{refScanID}(2,2,j)$.

3.7.7 Additional Comments and Version-dependent Variables

The presence of sea-ice is a source of error in the method since its scattering properties differ significantly from those of ocean. (Most occurrences of sea-ice are presently categorized as ocean surface type.) If sea-ice information becomes available, under both rain and rain-free conditions, an additional surface category will be defined so that the reference data are matched to the surface type encountered during rain.

Because the Ku-band has greater sensitivity than the matched-swath Ka-band, there are a number of cases for which rain is detected at Ku-band but not at Ka-band. As dual-frequency derived estimates at Ku- and Ka-band are produced when either or both channels detect rain, a number of dual-frequency derived Ka-band estimates of path attenuation will be produced without corresponding single-frequency estimates.

Versions 1 through 3 of the SRT products do not contain path attenuations based on the temporal reference data. Versions 4 and higher will include such results and which can be found in $\texttt{PIAalt}(5)$. The light-rain/wet soil temporal reference data set has yet to be derived so this product will be unavailable up to and including version 4. As a consequence $\texttt{PIAalt}(6)$ will be set to 'missing' (-9999.9). As already noted, the hybrid/cross-track reference data are derived only over ocean so that over land, $\texttt{PIAalt}(3)$ and $\texttt{PIAalt}(4)$ are set to 'missing'.

3.8 Solver (SLV) module

3.8.1 Objective

The primary objective of the Solver module is to retrieve the drop size distribution and calculate some physical variables. Details of the algorithm are described in Seto et al. (2021).

3.8.2 Algorithm Overview

The solver module employs an R-D_m plane, where the vertical axis is the precipitation rate (R) and the horizontal axis is the mass weighted mean diameter (D_m). Once an R-D_m relation (as $R = \varepsilon^r p D_m^q$) is assumed, the retrieval process goes from the top to the bottom range bin. At a "precipitating" range bin, a combination of R and D_m which satisfies both the R-D_m relation and the given Z_f is usually selected. Z_f is the radar reflectivity factor corrected for attenuation caused at the higher range bins but not for attenuation caused at the current range bin. For drawing the contour of Z_f on the R-D_m plane, scattering tables are used. After R and D_m are calculated at all "precipitating" range bins, the path integrated attenuation (PIA) is calculated. When the surface reference technique (SRT) gives a reliable PIA estimate, the calculated PIA and SRT's PIA estimate can be compared.

The R-D_m relation is linked to ε, which is originally an adjustment factor for the k-Z_e relation. The retrieval is done multiple times for different ε or different R-D_m relations corresponding ε ranging from 0.2 to 5.0. The best ε or the best R-D_m relation is selected by checking the likelihood of ε and the difference between the calculated PIA and SRT's PIA.

3.8.3 Retrieval processes

The major input data to the Solver module are radar reflectivity factor corrected for attenuation by non-precipitating substances (here it is denoted by Z_m), mean (μ) and standard deviation σ of $\log_{10}\varepsilon$, SRT's PIA estimates `PIA_SRT`, its reliability, precipitation type and the information of "phase". The retrieval algorithm first defines range bins to be processed: It mainly processes the data between the storm top and the actual surface.

From the storm top to the clutter free bottom, Z_m is usually available for the retrieval. By correcting the attenuation caused by higher range bins, Z_m is converted to Z_f. Then, the contour of the given Z_f is drawn on the R-D_m plane by reference to the scattering tables. If there is only one crossing point between the contour of Z_f and the curve for the assumed R-D_m relation, R and D_m can be determined easily. If there are multiple crossing points, a set of R and D_m which has smaller R and D_m, is preferred. Unfortunately if there are no crossing points, the point that is on the curve of R-D_m relation and gives the closest Z_f to the given value is selected. In this case, given Z_f (Z_{f1}) and the finally determined Z_f (Z_{f2}) are different.

In the main-lobe clutter region (below the clutter free bottom), Z_m is not useful for the retrieval. Here, the vertical profile of Z_e is assumed to be constant, or the same value of Z_e at the clutter free bottom is used in the clutter region. The same method is also applied for range bins in which Z_m's are smaller than the noise level if such small Z_m's are caused by large attenuation due to heavy rain above them. Specifically, if Z_m's above the point in question are reliable at 8 (4 for HS) or more liquid-phase range bins in the same pixel, Z_e is extrapolated.

In the dual-frequency algorithm, the following preference is set.

(1) Z_f of KuPR

(2) Z_f of KaPR

(3) extrapolated Z_e of KuPR

(4) extrapolated Z_e of KaPR

3.8.3.1 Determination of ε and use of SRT

Different R-D_m relations are tested by changing ε from 0.2 to 5.0. Using the following conditions, the best R-D_m relation is selected. Generally, the three conditions a) to c) are checked for the single-frequency algorithms and the four conditions a) to d) are checked for the dual-frequency algorithm.

a) A priori probability of ε; $(\log(\varepsilon) - \mu)^2$ is minimized, where μ is given by the DSD database or $\mu = 0$ if DSD database is not applied.

b) The difference of PIA calculated by the retrieval (`PIA_SLV`) and PIA estimated by SRT (`PIA_SRT`); $(\texttt{PIA\_SLV} - \texttt{PIA\_SRT})^2$ is minimized. When `PIA_SRT` is saturated, $(\min(\texttt{PIA\_SLV} - \texttt{PIA\_SRT}, 0))^2$ is minimized.

c) The difference of Z_{f1} and Z_{f2}; The sum of $(Z_{f1} - Z_{f2})^2$ is minimized.

d) If Z_m of KuPR is used for the retrieval and Z_m of KaPR is normal, estimated DSD should be also agreed with KaPR; $(\max(Z_{f2}$ of KaPR $- Z_{f1}$ of KaPR, $0))^2 + (\min(Z_{f2}$ of KaPR $- Z_m$ of KaPR, $0))^2$ should be minimized. If Z_{f2} of KaPR is between Z_m of KaPR and Z_{f1} of KaPR, the value is zero.

The following preference is set to select the use of PIA.

(1) δ`PIA_SRT`=KaPR's `PIA_SRT`$-$KuPR's `PIA_SRT` (if both are reliable and not saturated)

(2) KaPR's `PIA_SRT` (if it is reliable and not saturated)

(3) KuPR's PIA_SRT (if it is reliable and not saturated)

(4) KaPR's PIA_SRT (if it is reliable and saturated)

(5) KuPR's PIA_SRT (if it is reliable and saturated)

If none of PIA_SRT is available, or if selected PIA_SRT is saturated, abnormally high precipitation rates may result. To avoid them, condition e) is used.

e) R estimated from Z_m (not extrapolated Z_e) at liquid-phase range bins should be vertically constant; the variance of R is minimized.

PIA_SRT is set to be not "reliable" if the standard deviation of sampled surface backscattering cross section is larger than 10 dB. PIA_SRT is set to be "saturated" if the signal-to-noise ratio of measured surface backscattering cross section is smaller than 2.

From Version 06, the following condition is added. PIA_SRT is set to be not "reliable" if PIA_SRT is more than 10 times as large as PIA_HB, where PIA_HB is PIA calculated with $\varepsilon = 1$.

3.8.3.2 NUBF correction

Two NUBF parameters (p_area and rainvar) are set in the algorithm, but p_area (the ratio of rain part in an FOV) is fixed to 1.0 for the version 04 and version 05. rainvar (the square of coefficient of variation of R) is estimated by a fractal approach. In the single-frequency algorithms, NUBF correction is not applied in the first main loop. In the second (final) main loop, the NUBF parameter is estimated from piaFinal of the first main loop. In the dual-frequency algorithm, the NUBF parameter is estimated from piaFinal of the single-frequency algorithms. KuPR's piaFinal is preferred to KaPR's piaFinal. For the details, please refer to Seto et al. (2015).

3.8.4 Updates in Version 07

3.8.4.1 Estimation of the vertical profile of ε (dual-frequency algorithm)

In the dual-frequency algorithm, ε is given as a product of ε_S and ε_D. ε_S is vertically constant and ranges from 0.4 to 2.5. ε_D is dependent on the range bin and ranges from 0.5 to 2.0. At each range bin, if both the Z_m of KuPR and the Z_m of KaPR are available for the retrieval, ε_D is determined so that the value of the following equation becomes minimum. Otherwise, ε_D is set equal to that at upper range bin. If no upper range bins have the value of ε_D, ε_D is set to be 1.

$$(10 \log_{10} \varepsilon_D)^2 + \{0.5\,[Z_{f2}(\mathrm{KaPR}) - Z_{f1}(\mathrm{KaPR})]\}^2$$

When the method is applied, ε_S is used instead of ε for the condition of a) in Section 3.8.3.1 and the condition of c) in Section 3.8.3.1 is removed.

3.8.4.2 Correction of PIA_SRT by considering the soil moisture effect (single-frequency algorithms, over land)

The backscattering cross section increases as the surface soil moisture increases during the precipitation. It is called the soil moisture effect. As the SRT does not fully consider the soil moisture effect, PIA_SRT is underestimated over land. In the Solver module, PIA_SRT is corrected before it is used for the retrieval. The offset of PIA_SRT is dependent on the surface precipitation rate estimate, angle bin, location (for a 5 by 5 degree grid), and the frequency (KuPR or KaPR). In the first loop of the main module, no correction is given to PIA_SRT. In the second loop of the main module, an offset is determined by referencing the surface precipitation rate estimates in the first loop. By the correction, the surface precipitation rate estimates increase by around 10 to 20 % over land.

3.8.4.3 Parameters in R-D_m relation

The parameters in R-D_m relation are $p = 0.392$, $q = 6.131$, and $r = 4.815$. They are the same and the definition of ε is the same for all precipitation types. In the single-frequency algorithm or if the DSD database is not applied, μ is 0 for stratiform precipitation but μ is $\log_{10}(1.25)$ ≈ 0.0969 for convective precipitation. As the definition of ε for convective was changed from Version 06 to Version 07, and R-D_m relation with $\varepsilon = 1.0$ in Version 06 is similar to R-D_m relation with $\varepsilon = 1.25$ in Version 07.

Details of the description for variables are described in Appendix or in the following websites:

https://www.eorc.jaxa.jp/GPM/en/archives.html

https://arthurhou.pps.eosdis.nasa.gov/GPMprelimdocs.html

3.9 Trigger (TRG) module

3.9.1 Objectives

3.9.1.1 Purpose

This section provides the theoretical basis and computational approaches for a module integrated within the L2 DPR Algorithm: the module is called the "Trigger module".

3.9.1.2 Scope and Objectives

GPM Level 2 DPR data products are described in this ATBD. This section provides a description for one specific module integrated within the main DPR algorithm: the "Trigger" module, designed to detect the occurrence of Multiple Scattering (MS) or Non-Uniform Beam Filling (NUBF) effects in one particular radar profile, hence providing the option to trigger custom Solver modules as alternative to the standard Solver (SLV) module contained in the L2 DPR algorithm. Throughout this section, the same symbols and acronyms defined in this ATBD are adopted where applicable.

3.9.2 Background Information

The Trigger module is implemented as Experimental algorithm in DPR L2 Algorithm V07. No equivalent of the Trigger module was present in prior versions of the DPR processing, and no equivalent of the Trigger module is implemented in the Single Frequency (SF) KuPR and KaPR algorithms. The Trigger module is executed as a new module of the DPR Algorithm, immediately before the call to the Solver (SLV) module, and after the execution of the Preparation (PRE), Vertical Profile (VER), Surface Reference Technique (SRT), Classification (CSF) and Drop Size Distribution (DSD) modules whose outputs it receives as input. See Figure 3.1.1 in Section 3.1.2.3.

The Trigger module aims at detecting the occurrence of two distinct phenomena that invalidate or weaken the assumptions intrinsic to the DPR Solver Module.

These two phenomena are described at high level in this Section. Interested readers are referred to the referenced publications for more detail: in particular Battaglia et al. (2010) provides a comprehensive review on MS, while Kozu and Iguchi (1999) provides an overview of NUBF impacts.

3.9.2.1 Single Scattering (SS)

The DPR Solver module and most traditional retrieval algorithms for weather radars assume that the electromagnetic wave propagating through a medium of particles of interest (i.e., hydrometeors) is well described by the Single Scattering assumption. Under the Single Scattering

(SS) assumption, the total field scattered back to the radar by the medium of interest is simply the sum of the fields backscattered by the individual particles, and other intervening particles impact the observation only by causing attenuation along the path between the observer (radar) and the particle (hydrometeor), so that the measured profile of radar reflectivity Z_m can in principle be corrected for attenuation by estimating the single scattering properties of the hydrometeors starting from assumptions that tie the effective unattenuated reflectivity Z_e to the specific attenuation k given an assumed uniform distribution of hydrometeors inside each radar volume.

Figure 3.9.1a provides a simplified visual representation of the conceptual SS model. Each volume of resolution of the radar (purple dash boxes) contains a spatially uniform distribution of particles that interact with the incoming energy at Ku and Ka band. As result, a very small portion of the incident energy is backscattered directly at the radar (blue upward arrows, resulting in the corresponding measure of Z_m), while a some energy is absorbed by either the particles or the molecules in gas phase (and therefore contributing to the overall "attenuation" of the energy that keeps propagating downwards, green downward arrows). A third component of the propagating energy is scattered in directions that are neither exactly back nor exactly forward (yellow arrows sideways). This third component, typically modeled by the "scattering cross-section" of the particles, is assumed to be lost completely in the SS model and it therefore adds directly to the "absorption" to calculate the overall "extinction" of the signal as it propagates through the atmosphere and all the way to the surface. In retrievals based on Single Scattering models, the estimation of the two-way attenuation of measured reflectivity is therefore performed simply by adding absorption and scattering. Furthermore, under this conceptual model, time along a profile is directly associated to distance (i.e., an echo received at time T from a given reference is assumed to be generated by a target located at a slant range distance of $r = 2T/\mathrm{c}_0$ from that reference, where c_0 is the speed of light in air).

The SS assumption is almost always verified for weather radars operating in the centimetric portion of the spectrum (say, between 3 and 10 GHz). At these frequencies, most hydrometeors scatter in the Rayleigh regime, and most of the path extinction is tied to absorption (significant when liquid or mixed phase hydrometeors are present). Therefore, very little of the energy scattered by one hydrometeor in directions other than the one connecting to the observer is likely to impact a second scatterer and be finally collected by the observer.

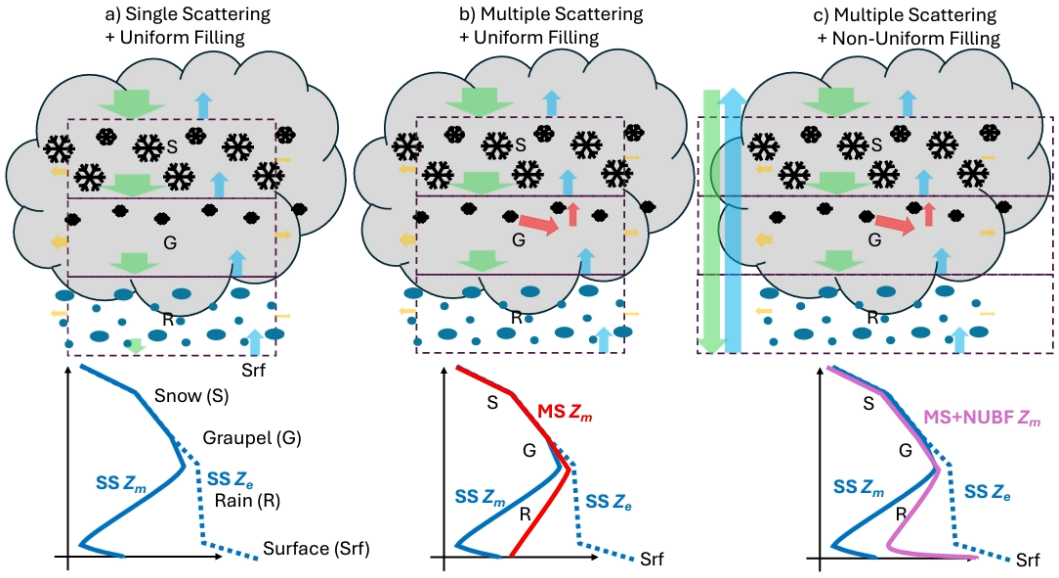


Figure 3.9.1: Simplified diagram illustrating the main impacts of Single-Scattering vs Multiple-Scattering, and Uniform vs Non-Uniform beam filling. The assumed scenario for Ka-band propagation includes (from top to bottom): a "Snow" layer, where a uniform distribution of low density ice particles results generally in low reflectivity values (i.e., low backscattering), low scattering in general, and low absorption; a "Graupel" layer, where a uniform distribution of high density ice particles results generally in higher reflectivity values, significant scattering in general, and low absorption; a "Rain" layer, where a uniform distribution of liquid drops results generally in higher reflectivity values, moderate scattering in general, and high absorption; and finally the very reflective Surface. The downward arrows represent the energy flux coming from the radar transmitter, and the blue upward arrows represent the energy backscattered towards the radar. The yellow arrows represent energy that was scattered "away" from each layer, and the red arrows represent energy that was scattered "around" within each layer.

3.9.2.2 Multiple Scattering (MS)

Contrary to the example of wavelengths of 3 cm and longer, in the optical regime (e.g., lidar) the hydrometeor scattering cross-section dominates over the absorption cross-section, and in general most of the energy scattered in all directions interacts further with other particles before eventually being received back at the observer. Radars operating at frequencies higher than 10 GHz experience a behavior that gradually transitions from SS to MS dominated, depending on the type and characteristics of the hydrometeors.

Figure 3.9.1b provides a simplified depiction of MS: a portion of energy that is scattered in directions that are neither directly back nor directly forward is not lost (represented by the red arrow). This energy interacts again with other particles (in all directions), and at each interaction a portion is directed towards the observer, while the rest is directed in other directions. Depending on the nature of the medium, this process can repeat one or many times (statistically) before it results in negligible contributions to the overall received signal. A medium with high scattering efficiency, very low absorption, and a fairly isotropic phase function is bound to produce

the most secondary scattering events for a longer time (one example of these conditions may be that of fairly large and dense dry ice particles in the upper troposphere with low amounts of water vapor and no liquid particles). A corollary of this effect is that the association between time of propagation T and slant range distance r from the observer is not valid any longer, because for some fraction of its propagation the energy was not moving directly away or towards the radar. For example in Figure 3.9.1 the measured reflectivity near the surface (bottom of the red curve) is the result of scattering events occurring above the zero isotherm in the ice layer: the measured Z_m at that location has absolutely nothing to do with the near-surface rain at that location (in this example such contribution to Z_m, in blue, is assumed to be many dB below the MS signature). This type of feature is referred to as 'ghost' or 'tail'.

It is therefore necessary to account for the fact that more than one scattering event contributes to the observation, by adopting MS radiative transfer models.

Millimeter wave (mm-wave) radars fall in an intermediate regime between lidar and centimetric radars: the contributions from MS to the observed backscatter can range from negligible to overwhelming, relative to those from SS, depending on the size, shape, phase and number concentration of the hydrometeors. Furthermore, contrary to lidar applications where the small-angle scattering dominates, wide-angle scattering dominates in mm-wave radar applications: the scattered energy is scattered more isotropically around the original angle of incidence. For this reason, the geometry of observation plays a key role: since MS contributions at mm-wave follow a complex path that extends in directions different from the direct line observer-particle, and the probability of scattering events beyond the first is governed by the transport mean free path of the medium (see eq. (2) in Battaglia et al. 2010), the higher order scattering events occur in a shaped volume around each point of first scattering. Therefore, all other conditions being equal (say, same wavelength and same medium), the dimensions of the radar volume of resolution modulate how much of the multiply-scattered energy is actually collected back by the radar and therefore contributes to the measurements. This is the reason why spaceborne cloud and precipitation radars are generally more affected by MS than their equivalent counterparts on the ground or airborne: despite the fact that they typically have smaller beamwidths, spaceborne radar geometry is dominated by the extremely long range of the observations (generally in excess of 400 km).

Why does MS matter? In a nutshell, because the classic weather radar equation, and all the standard retrieval algorithms that hinge upon it assume SS, and because MS contributions can alter the observed profiles of reflectivity macroscopically (to the extreme of providing a significant echo at a certain apparent range, whereas the SS scattering at that very same range was well below the detection threshold).

Perhaps most important for GPM, given the dependency on frequency of MS occurrence, and

given the identical footprint size of DPR's Ku and Ka band, the Ka-band observations are far more likely to be altered by it than the Ku-band, hence weakening the applicability of dual frequency approaches.

The most common situation when 'severe' MS affects DPR measurements is over deep convection, with relatively large and dense frozen hydrometeors (e.g., graupel) aloft and a mixed-phase or liquid hydrometeors below: the former are characterized by high albedo (i.e., large scattering cross-section and small absorption cross section) and the latter by high extinction (i.e., large absorption cross-section and scattering cross-section). Energy scattered in the high albedo layer at wide angles creates in essence a reverberation, while energy that penetrates the high extinction layer is absorbed. In this scenario SS predicts a sharp drop of the measured reflectivity as one moves from the higher layer to the lower layer, whereas the MS reverberation provides a much slower reduction in the profile of reflectivity with increasing range. It is important to recall that when MS occurs the classical radar association between time-of-return and range-to-target is not valid any longer: the reverberated backscatter follows a path that is not along the direction of propagation of the transmitted wave. As shown in Battaglia et al. (2015) when severe MS occurs standard solvers cannot produce reliable retrievals because the SS assumption at their foundation is not valid.

3.9.2.3 Non-Uniform Beam Filling (NUBF)

While MS has been subject of attention by the weather radar community only in the past two decades, NUBF has been understood and studied for far longer: no matter the nature of the weather radar, practical limitations often result in volumes of resolution that are fairly small in range (say 50 to 500 m) but larger in their cross-beam direction (say 500 to 5000 m). Given that clouds and precipitation are inhomogeneous over a wide range of spatial scales, quantifying the magnitude and impact of such inhomogeneities has been subject of research for several decades.

Panel c of Figure 3.9.1 provides a visual reference for an extreme situation where both MS and NUBF impact the measurement: in this case the precipitating cloud is simply assumed to be smaller than the footprint. On one hand the portion of the volume that is devoid of scatterers affects the measure of the unattenuated effective reflectivity factor, simply as $Z_e = \alpha Z_1$(where α is the fraction of the footprint uniformly filled with hydrometeors) on the other hand the SS attenuation profile is actually associated to Z_1 and not Z_e for all volumes in the precipitation column, while the surface echo is attenuated by an amount that is far less than the last near-surface rain volume (because the $1 - \alpha$ portion of the footprint doesn't contribute to the attenuation). In general, these factors alter the associations between SS properties of hydrometeors and introduce a confounding factor in the interpretation of Z_m.

Profiles of precipitation can be affected significantly only by NUBF, or only by MS, or both,

depending on the wavelength, size of the volume of resolution, and type of hydrometeors. Theoretical modeling and experimental results indicate that DPR is very rarely significantly impacted by MS at Ku-band (presence of hail and large graupel are a necessary and not sufficient condition), while it is more frequently impacted at Ka-band (especially when moderate or heavy rain is below a fairly thick layer of graupel or rimed aggregates, but also in other conditions). NUBF is somewhat ubiquitous, but its effects are typically less pronounced and various approaches to compensate for it have been applied in the past.

3.9.3 Algorithm Description

3.9.3.1 Theoretical Description

The Trigger DPR Module provides an algorithm that computes the degree to which a DPR profile is affected by Multiple Scattering (MS) and Non Uniform Beam Filling (NUBF). The two functions, MS and NUBF, are implemented as separate submodules and provide separate outputs; however their outputs should be interpreted jointly.

3.9.3.1.1 Detection of Multiple Scattering (MS)

Multiple Scattering is diagnosed based on detection of the presence of two features in the measured reflectivity fields: the "Multiple Scattering tail" and the "DFR knee".

The feature described as "**Multiple-Scattering tail**"(note the term "Ghost Echo" has also been used in the literature) where the observed backscatter power decays from a maximum Z_m value reached at an altitude typically above, or around the zero isotherm to a minimum Z_m value at an apparent altitude below the actual Earth surface following a smooth curve where the first derivative of dZ_m/dr(where r is range from radar) remains negative and the second derivative is positive (i.e., a concave shape), with no visible peak generated by the surface echo despite the minimum Z_m value being above the detection threshold of the DPR.

This behavior was first observed on the spaceborne Cloud Profiling Radar on board CloudSat (Battaglia et al. 2010), and was validated through airborne observations. It was later also observed on TRMM PR data. In both cases it was assessed to be an extremely rare occurrence for these two radars.

This feature is more frequently observed on the DPR Ka band channel as expected because of its comparatively large ratio between footprint size and mean free scattering path (Battaglia et al. 2015, 2016). This feature has been observed also in the Ku-band data, albeit extremely rarely [Tanelli STM].

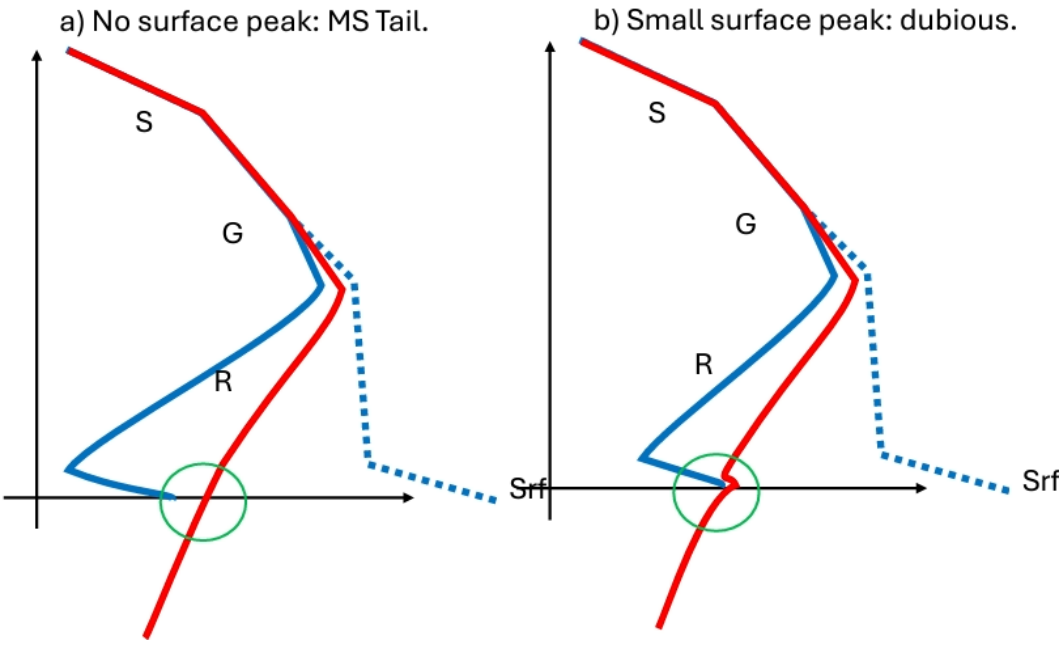


Figure 3.9.2: Simplified diagram illustrating the Multiple-Scattering "Tail" or "Ghost Echo". Red is a notional profile of Z_m, blue curves are the corresponding SS profile of Z_e (dash) and Z_m (solid).

The key revealing feature is the absence of a surface peak at the expected range intercept while the values of Z_m in the corresponding range bins are well above the noise floor both above and below the surface intercept, and still decreasing in range, as shown in the example in Figure 3.9.2a. Typically the surface peak for DPR results in very large equivalent Z_e : for example even assuming a relatively low s0 (surface normalized radar cross section) of 0 dB, the resulting equivalent Z_e is about 65 dBZ at Ku band (at its 250 m range resolution) and between 45 and 50 dBZ at Ka band (depending on the range resolution). Therefore, one would need to assume that the near surface rainfall Z_e is even larger than these levels to explain the absence of an observed peak. However, assuming a 50 dBZ (Ka-band) rain/hail echo would result in extremely high specific attenuations that would not explain why the overall Z_m is still above the noise floor above and below the range intercept of the surface. This feature is instead well modeled by MS forward models : the SS component of the signal can indeed be attenuated by several 10s' of dB at Ka-band, enough to relegate all SS echoes (from rain, hail or surface) well below the detection threshold; at the same time the MS signature, mostly generated by ice particles aloft, continues to provide echoes that appear to decrease in range only because of the recursive amounts of energy scattered away from the ice layer where it is originated. These MS echoes are never propagating through the heavily attenuating liquid layer below, and the range intercept of Earth's surface has no meaning for this type of propagation. Figure 3.9.1b, shows as similar

case where instead a small peak is visible. It is still possible that the observed profile is an MS tail, with less SS attenuation allowing for the surface peak to be visible over it. These cases can be identified as dubious on this criterion alone.

It is important to note that this criterion is particularly effective only at near-nadir angles, where the surface peak is simultaneously sharper and larger in magnitude, and more range bins are available for the sub-surface layer to detect that portion of the tail. For DPR, this criterion results in large fractions of Missed Detections for profiles at incidence angles larger than about 5°.

Two additional sources of Missed Detections for this criterion are lack of sampled range bins below the surface and still above the detection threshold, and simultaneous occurrence of significant NUBF (see example in Figure 3.9.1c, where despite possible MS occurrence in the main column of precipitation, a fraction of the field of view is almost devoid of heavy precipitation and therefore results in a very visible surface peak).

False Alarms for this criterion are mostly tied to situations when the Z_m profile has fallen at or near the detection threshold and a noise-floor contaminated profile is mistakenly interpreted as an MS Tail.

The second feature detected by this module is described as the "**DFR Knee**" (DFR = Dual Frequency Ratio) (Battaglia et al. 2014).

In a plot of the $DFR_m = Z_{m,\mathrm{Ku}} - Z_{m,\mathrm{Ka}}$ expressed in dB, the most commonly observed behavior is that it is generally non-decreasing for increasing range (because Ka-band absorption and scattering are generally larger than at Ku band). On the other hand, DPR profiles of $DFR_m = Z_{m,\mathrm{Ku}} - Z_{m,\mathrm{Ka}}$ exhibit on occasion such large reversal, where the DFR_m increases in range down to an altitude generally above the zero isotherm, and then it fairly sharply starts to decrease in range at a rate not dissimilar in modulus from the ascending branch (one example of such occurrence is provided in Figure 3.9.3). Once again, this behavior is extremely hard (when not impossible) to explain under the SS model, and it is instead naturally deriving from the MS model.

- SS-model: in fairly extreme Particle Size Distributions (PSD), where for example a small number of very large hydrometeors (e.g. large hail) are the dominant scatterer, the differential behavior can become almost null, and even reverse polarity to a small extent; however, a large reversal of polarity could only be explained by very precise and almost mono-modal PSD (such that a null in the Mie pattern of scattering at Ka-band becomes the dominant factor).

- MS-model: MS conditions are more likely to occur at Ka-band vs Ku-band (given that both frequencies have similar footprints for DPR). When MS sets in, the $Z_{m,\mathrm{Ka}}$ slope $dZ_{m,\mathrm{Ka}}/dr$ is therefore less negative than the corresponding SS profile of $Z_{m,\mathrm{Ka}}$, while the $dZ_{m,\mathrm{Ku}}/dr$

slope is still following the SS-model. The resulting slope $dDFR_m/dr$ can therefore assume a negative value when the Ka-band channel is affected by significant MS.

Since this is a feature of the difference between the two frequencies, detection of this feature relies on the availability of both Ku and Ka band at a given profile. In general, it is indication of significant Multiple Scattering originating/being generated in the upper troposphere by large frozen hydrometeors (note: forward modeling of scattering and propagation shows that graupel of only a few mm in diameter is sufficient to engender this feature given the DPR geometry and frequencies).

3.9.3.1.2 Diagnosis of Non-Uniform Beam Filling (NUBF)

NUBF at Ka-band is diagnosed for each MS vertical profile, based on two measures of horizontal variability in the measured reflectivity fields:

- the variability of the 4 partially overlapped HS profiles (limited to the period prior to May 21^{st}, 2018) and that of the 4 adjacent MS profiles;

- where Path Integrated Attenuation (PIA) estimates from Surface Reference are available at Ku and Ka band, the ratio between the two is expected to be within a range consistent with single-scattering properties of assumed Drop Size Distributions (DSD) and Particle Size Distributions (PSD), generally indicated by the $N(D)$ distribution. Typically, the range between Ka-band PIA and Ku-band PIA is expected to be between 4 and 8. Lower values of this ratio can be explained by assuming extremely large raindrops: while possible, in-situ observations of rain don't support this assumption except for extremely rare occurrences. A second explanation for such departure is offered by a simple model of non-uniformly filled columns (Tanelli et al. 2012): a simple model of NUBF has been developed to demonstrate how in the limit the ratio of the two PIA can approach 1:1 regardless of the $N(D)$. Such a model is adopted here as indicator of NUBF along the column.

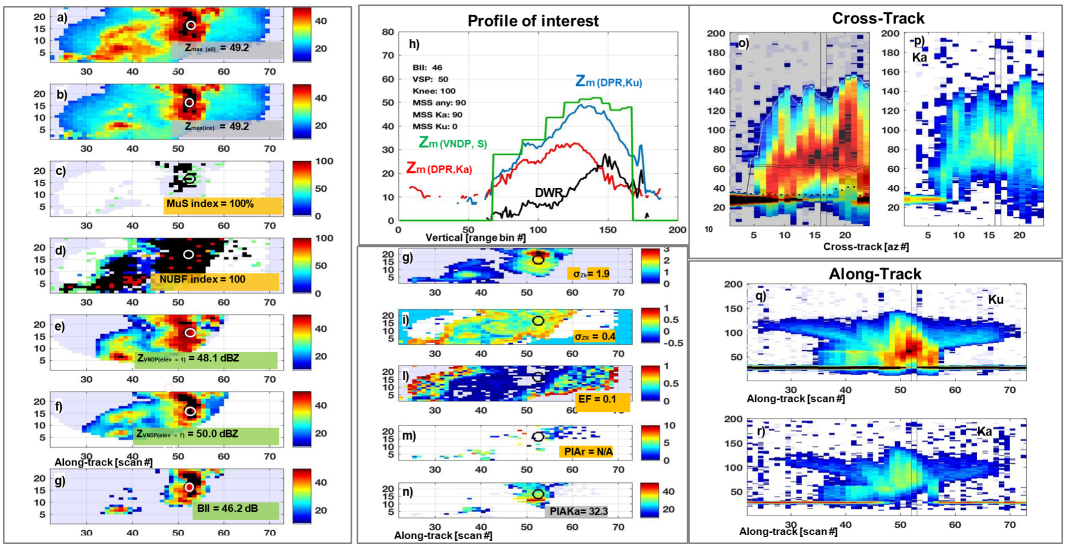


Figure 3.9.3: Example of quicklook of the Trigger module output using actual DPR data.
a) Max of $Z_{m,\mathrm{Ku}}$ along each profile of the "matched" swath of DPR;
b) Max of $Z_{m,\mathrm{Ku}}$ above the zero isotherm, along each profile of the "matched" swath of DPR;
c-d) overall Multiple Scattering and NUBF indexes (0-100) derived from the logic in Section3.9.3.2.2;
e-g and g-l) validation metrics provided by the VNDP processing of S-band ground based weather radar data;
h) profile highlighted by a small circle in all the "map" panels, the S-band profile is from VNDP, in this case the MS-Tail criterion provided a score of 90% at Ka-band and the DFR knee criterion provided a score of 100%);
m-n) PIA-related metrics mainly used to diagnose NUBF;
o-r) vertical cross sections of DPR Z_m centered on the profile of interest (cross-track and along track).

3.9.3.2 Processing Description

The code is written in C but in a MATLAB-like style where most of the common MATLAB vectorized operations have been implemented in C.

Three main data structures are used in order to conveniently store and pass data among the different methods:

1. TRG_ptInputs: Pointer to a C structure defined in TGRReader.h that contains the inputs to the Trigger code. Each quantity is prefixed by *ms* (matched scan), *ns* (normal scan) or *hs* (high sensitivity scan) depending on the acquisition mode. For instance, the $flagEcho$ from the *ns* mode is called $ns\_flagEcho$.

2. TRG_params: Pointer to a C structure defined in TGRReader.h that contains inputs parameters that allow a change in the runtime behavior of the algorithm.

3. TRGContext: Pointer to a C structure defined in TRGContext.h that contains temporary or output variables that are generated by the algorithm and that have global scope (not limited to the module in which they are created).

The use of this structure greatly simplifies the API and makes it more general and extensible since new parameters can be added to the structure without changing the signatures of the *methods* that utilize them. Note: in object-oriented programming, a *method* is a programmed procedure that is defined as part of a class and is available to any object instantiated from that class.

The Trigger module is comprised of three main submodules further described in the next subsection: an interface layer from the Trigger C code to the main DPR code (FORTRAN) ($trg\_dpr.c$), the actual Trigger algorithm ($trg\_dpr\_sub\_mod.c$), and a library ($trg\_dpr\_sub\_math.c$) of a set of function calls that mimic the MATLAB math vectorized operations such as addition, multiplication etc. and a set of data structures to represent multidimensional arrays.

3.9.3.2.1 Interface Layer between Trigger Module and Main DPR code ($trg\_dpr.c$)

This is an interface layer from the Trigger C code to the main DPR code (FORTRAN). This module provides an access point from the main DPR code to the Trigger module (C). It conforms the FORTRAN data structures to the C data structures for both input and output.

It is further divided in 5 submodules described below:

a) $initTrg$

This module allocates (if necessary) and initializes the three main structures used in the Trigger module to hold parameters (TRG_params), inputs ($TRG\_ptInputs$) and outputs ($TRGContext$). This method calls $dpr2trg$ to conform the FORTRAN data structures to the C structures.

b) $dpr2trg$

The main DPR code uses structure arrays in the form

$$struct\_name[scan].variable\_name[angle, range],$$

while the Trigger code uses a custom structure representing a multidimensional array in the form

$$variable\_name[range, angle, scans].$$

This function provides the interface between the DPR standard structures and the TRG internal structures, and fills the $TRG\_ptInputs$ structure with values from the homologous structure from the DPR code.

c) $trg2dpr$

This function is the counterpart of $dpr2trg$: it formats the outputs generated by the Trigger module in the data structure suitable for the DPR main code.

d) finalizeTrg

This module deallocates the three Trigger structures *TRG_ptInputs*, *TRG_params* and *TRGContext*.

*e) trg_dpr_*

This module orchestrates the aforementioned functions in *trg_dpr.c* and the actual algorithm functions defined in *trg_dpr_sub_mod.c* calling the necessary initializations, the Trigger processing and the finalization. [not sure I understand this.]

3.9.3.2.2 Trigger Algorithm (*trg_dpr_sub_mod.c*)

a) prepare

As the name suggests, this module prepares the inputs for the processing. In particular, the variables for the different swath modes (NS and MS) need to be collocated. The Trigger Algorithm only operates where Ku- and Ka-band data are available. Figure 0.11.1 illustrates the DPR scan pattern at the beginning of the mission, and after the scan pattern change in May 2018. It is reproduced here as Figure 3.9.4 for convenience:

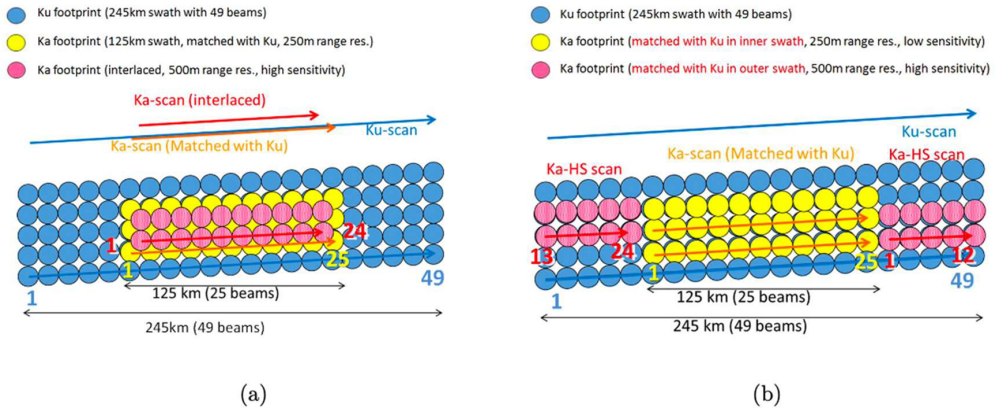


Figure 3.9.4: The DPR scan pattern: a) pre May 21, 2018; b) after May 21 2018.

To accommodate the pattern change on May 21 2018, the algorithm was modified simply by exploiting only the overlapping Ku and Ka footprints in the original inner swath, and by skipping (alt. 'bypassing') all calculations previously conducted on the HS profiles (this reduction affects only the NUBF_3D_check submodule described in Section 3.9.3.2.2).

The variables relative to the NS scan pattern have 49 beams while the MS only 25 so the beams used are 13 to 36 for the NS variables and 1 to 24 for the MS variables.

The most important variables that are collocated and stored in the *TRGContext* structure

at this stage are listed below. In parenthesis we report the original name in the L1 or L2 DPR products (full description with size and type can be found in the Appendix A of JAXA 2021 and JAXA 2022.)

- ZKU: NS measured reflectivity factor ($zFactorMeasured$ from L2 DPR).
- ZKA: MS measured reflectivity factor ($zFactorMeasured$ from L2 DPR).
- $ZKaHS$: HS measured reflectivity factor ($zFactorMeasured$ from L2 DPR).
- $Nbin$: Number of range bins.
- Naz: Number of beams (or angles in the DPR nomenclature).
- $Nscan$: Number of scans.
- $rain$: MS precipitation flag ($flagPrecip$ from DPR).
- $selRain$: Indices of $rain$ where its values are "true".
- lat: MS Latitude ($Latitude$ in DPR).
- lon: MS Longitude.
- ang: NS local zenith angle.
- isc: NS real surface bin ($binRealSurface$ from L2 DPR).
- $bcfb$: NS bin of the free clutter bottom ($binClutterFreeBottom$ from L2 DPR).
- $bcfbKA$: MS bin of the free clutter bottom ($binClutterFreeBottom$ from L2 DPR).
- $flagEchoKU$: NS flag for presence of echo ($flagEcho$ from L2 DPR).
- $flagEchoKA$: MS flag for presence of echo ($flagEcho$ from L2 DPR).
- $binnodeKU$: NS range bins corresponding to storm height, top, center and bottom of melting layer and lowest bin of the precipitation ($binNode$ from L2 DPR).
- $binnodeKA$: MS range bins corresponding to storm height, top, center and bottom of melting layer and lowest bin of the precipitation ($binNode$ from L2 DPR).
- bin_zeroC: NS bin of zero thermal ($binZeroDeg$ from L2 DPR).
- $bin\_zeroCKA$: MS bin of zero thermal ($binZeroDeg$ from L2 DPR).
- $PIAKu$: NS Path-Integrated Attenuation, i.e. PIA ($pathAtten$ from L2 DPR)
- $PIAKa$: MS Path-Integrated Attenuation ($pathAtten$ from L2 DPR).
- $PIAKuf$: NS PIA reliability factor ($reliabFactor$ from L2 DPR).
- $PIAKaf$: MS PIA reliability factor ($reliabFactor$ from L2 DPR).
- PKU: NS echo power in dB ($echoPower$ from L2 DPR).
- PKA: MS echo power in dB ($echoPower$ from L2 DPR).
- $i1ellip$: NS range bin number of the Ellipsoid ($binEllipsoid$ in L1 DPR).
- $i1surf$: NS range bin of the surface ($binEchoPeak$ in L1 DPR).
- $i1surfd$: NS range bin of the average DEM surface elevation ($binDEM$ in L1 DPR).
- $i1ellipKa$: MS range bin number of the Ellipsoid ($binEllipsoid$ in L1 DPR).

b) createMasks

Given the large size of the 3D data necessary for processing, this module uses information such as the presence of echo and the location of the zero-isotherm level to reduce the overall data only to the portion of the image where there is a potential presence of MS or NUBF.

The variables generated at this stage are

- *bin_top*: location of the top of the storm in Ku band.
- *bin_topKA*: location of the top of the storm in Ka band.
- *zmax*: ice maximum reflectivity in range in Ku band.
- *zmaxKA*: ice maximum reflectivity in range in Ka band.
- *Zmax_ice_selRain*: value of *zmax* for the bins where rain is present.
- *Zmax_ice_selRainKA*: value of *zmaxKA* for the bins where rain is present.
- *zmaxall*: maximum reflectivity in range due to all components in Ku band.
- *zmaxallKA*: maximum reflectivity in range due to all components in Ka band.
- *Zmax_all_selRain*: value of *zmaxall* for the bins where rain is present.
- *Zmax_all_selRainKA*: value of *zmaxallKA* for the bins where rain is present.

c) MS_check

The *MS_check* module applies a set of heuristics to determine the likelihood of MS presence in a profile. In particular, the algorithm tries to detect a "knee" feature in the Dual Frequency Ratio (DFR) and a "tail" feature across, or masking, the surface peak that are both indication of the presence of MS as discussed above.

Figure 3.9.5 shows an example of DFR knee detection: a piecewise linear interpolation of the DFR is performed on the "upper" and "lower" part of the maximum of the DFR. The maximum of the DFR along with the lengths of the upper and lower portion of the leg are used to establish the presence of MS.

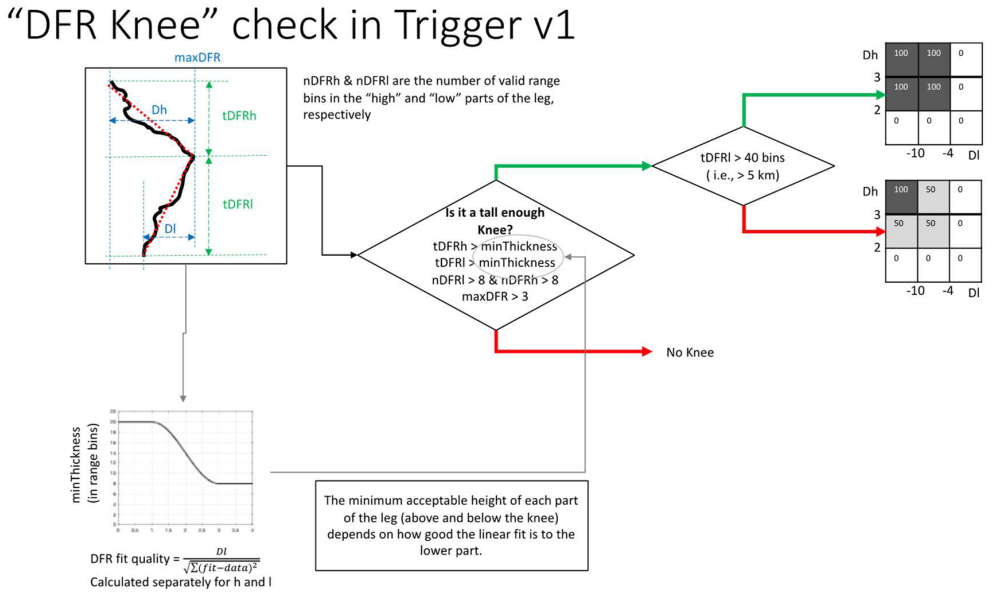


Figure 3.9.5: An example of a profile of $DFR_m(Z_{m,\mathrm{Ku}} - Z_{m,\mathrm{Ka}})$ is shown in black. This module performs an automated piecewise linear analysis of the profile and applies the logic shown in the flow diagram to detect the presence or absence of a feature consistent with a DFR Knee. The branches shown in red are negative outcomes of the parent check, the ones shown in green are the positive outcomes. Symbols used in this diagram: Dh (Dl) = Depth in dB of the high (low) portion of the Knee, maxDFR = maximum of DFR_m along the entire profile; tDFRh (tDFRl) = thickness in number of range bins of the high (low) portion of the Knee.

A primary factor limiting the effectiveness of the *DFR Knee* approach is the detectability of $Z_{m,\mathrm{Ka}}$ below the zero isotherm: when $Z_{m,\mathrm{Ka}}$ drops gradually to the detection threshold above the altitude of the surface, the lower half of the "leg" often becomes too short, or the quality of the linear fit becomes so degraded that no clear "knee" can be detected in the data. In general, this approach is found to be most effective at detecting the occurrence of the most severe occurrences of MS, generated by very vertically developed storms with significant scattering and extinction in the ice phase. Profiles that exhibit this feature are the most likely to result in unstable retrievals if a standard Solver module (based on Single Scattering assumptions) is applied.

Figure 3.9.6 shows the implementation of the detection of the "MS tail" through the surface feature as described in Section 3.9.3.1.1. Four portions of the profile near the surface are identified relative to the surface intercept range (that is, the range at which the surface is expected to be encountered), two clutter free portions (denoted by 'cft' along the radar range in figure 3.9.6) above (segment a) and below (segment b) the surface intercept and two cluttered portions (denoted by 'ct') nearer the surface intercept. The slopes of the

profiles for all four segments and the extrapolation of the pre-surface-clutter free segment are computed and used to determine the presence of MS, they are indicated by S_a, S_{s1}, S_{s2} and S_b in order from segment a to segment b.

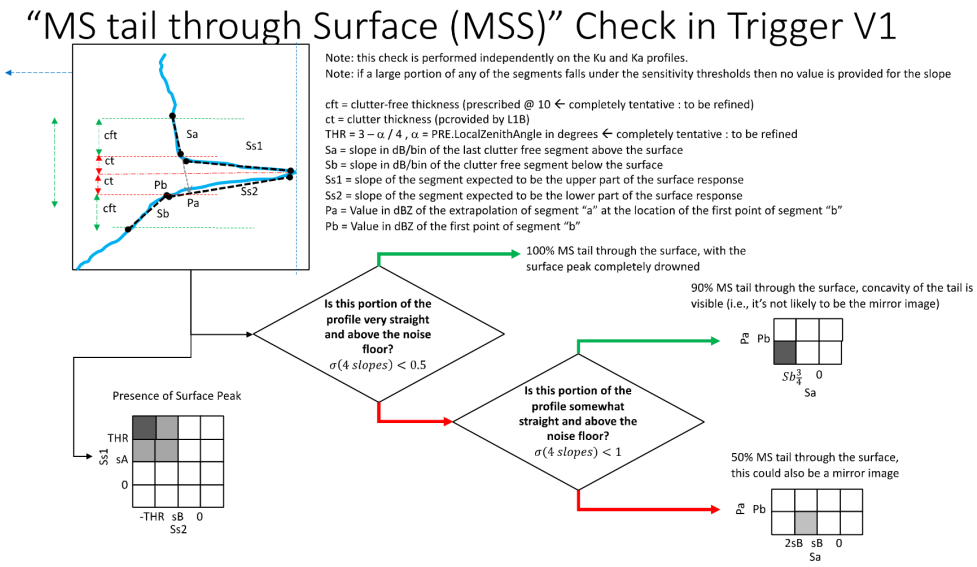


Figure 3.9.6: Example of profile of Z_m around the surface intercept is shown in blue. This module performs an automated piecewise linear analysis of the profile and applies the logic shown in the flow diagram to detect the presence or absence of a feature consistent with a surface peak. The branches shown in red are negative outcomes of the parent check, the ones shown in green are the positive outcomes.

This criterion is applied independently at Ku and Ka band, and is therefore the only one with a possibility to detect extremely rare occurrences of MS at Ku band. While on occasion it is able to detect possible occurrence of MS contamination in profiles not generated by the strongest storms (as is the case for the DFR knee) it is also more likely to generate false alarms because the detected feature can be explained also by NUBF or can be misdetected because of weak signals around the surface. Nevertheless, profiles that are flagged by this criterion are likely to be more complex than the standard uniform-filling, single-scattering assumptions adopted by the standard Solver and should be considered as candidates for more sophisticated retrieval schemes.

The main output of the MS_check module is a set of index variables that assume one of three values (0,50,100) that indicates coarsely the likelihood (in%) or the degree of the presence of a given feature. Specifically:

- $flag\_SP$: Indicates the presence of a Surface Peak for a given profile in the Ku band.

- $flag\_Spa$: Indicates the presence of a Surface Peak for a given profile in the Ka band.
- $flag\_MSS$: Indicates an initial estimate of the presence of MS tail through the surface.
- $flag\_DFR\_knee$: Indicates the presence of a Knee for a given profile.

d) NUBF_3D_check

This module provides the first element to diagnose the amount of NUBF in a given profile. The module is applied only to the Ka band channel.

Figure 3.9.7 shows the local neighborhood of an MS profile of interest (indicated in green) that plays a role in the calculations: the four partly overlapped HS profiles are marked in red, and the four adjacent MS profiles are marked in black. Use of the 4 diagonal MS profiles marked in light blue was considered in the initial stages of development of the Trigger module, but were found to provide no benefit, and therefore are not used.

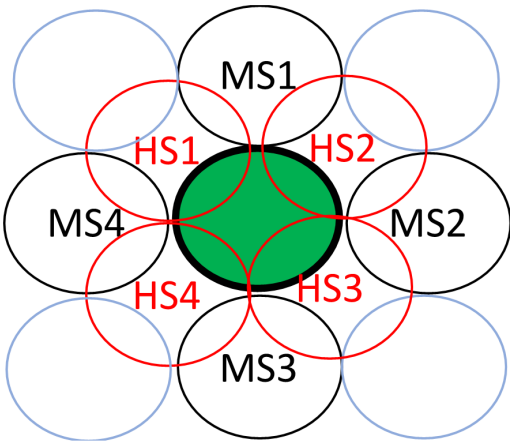


Figure 3.9.7: Neighborhood of profiles used to diagnose NUBF in the 3D check.

IMPORTANT: this module was specifically devised and tested accounting for the presence of the HS swath. In attributing relative weighting factors, a simple double-Gaussian approximation for the antenna pattern was used and the overlapping nominal energy of each neighbor with respect to the target profile was calculated: such overlap is approximately 0.18 for the 4 HS profiles and approximately 0.04 for the 4 MS profiles. This results in a relative weight of approximately 5 in favor of the HS profiles. Therefore, the reliability of this particular module should be considered generally

as **downgraded for datasets acquired after the 2018 scan pattern change**. The decrease in accuracy in the method for data taken after May 2018 has not been quantified.

Three quantities are calculated: the variability of measured reflectivity in the rain layer (σ_{ZR}), the variability of estimated attenuation (σ_{Zk}) and the empty fraction (EF).

$$\sigma_{ZR} = \frac{\log_{10} \int_0^{H_{\text{rain}}} [\sigma(Z_{\text{MS},4}) + \sigma(Z_{\text{HS},4}) * \ w_{\text{HS}}] \, dh}{(1 + w_{\text{HS}}) * H_{\text{rain}}} \tag{3.9-1}$$

where H_{rain} is the maximum height of the liquid precipitation, $\sigma(Z_{\#\#,4})$ are the standard deviations of $Z_{m,\text{Ka}}$ calculated over the target profile and the 4 neighbors from the MS and HS swaths, and w_{HS} is the weight given to the HS variability relative to the MS variability. The parameter w_{HS} was set initially to 5 to approximate the stronger overlap of the HS profiles with respect to the target profile, but was changed to 0 after the May 2018 scan pattern change to reflect that it is no longer available. This metric is best suited to reveal variability in the reflectivity field in rain, for profiles that are not heavily attenuated by convective columns extending well above the zero isotherm.

$$\sigma_{Zk} = \frac{\log_{10} \int_0^{H_{\text{top}}} [\sigma(k_{\text{MS},4}) + \sigma(k_{\text{MS},4}) * \frac{\sigma(Z_{\text{HS},4})}{\sigma(Z_{\text{MS},4})} * w_{\text{HS}}] \, dh}{(1 + w_{\text{HS}}) * H_{\text{top}}} \tag{3.9-2}$$

Where H_{top} is the echo top, k is a first-order approximation of the Ku-band specific attenuation, and $\sigma(k_{\text{MS},4})$ is the standard deviation of k calculated over the target profile and the 4 neighbors from the MS. This heuristic parameter aims at capturing the horizontal variability of the specific attenuation in situations where a vertically developed convective tower engenders most of the nonuniformity of measured reflectivity in the liquid layer via its nonuniform attenuation.

$$EF = \frac{\log_{10} \int_0^{H_{\text{top}}} \{\sum_1^4 [1 - \delta(Z_{\text{MS},i})] * w_{\text{MSEF}} + \sum_1^4 [1 - \delta(Z_{\text{HS},i})] * w_{\text{HSEF}}\} \, dh}{H_{\text{top}}} \tag{3.9-3}$$

Where $\delta(Z_{\#\#,i})$ is 1 if an echo is present at neighbor i of that set, and 0 if it is not; w_{MSEF} = 0.04 and w_{HSEF} = 0.18. Similarly to the two prior quantities, most of the value of this quantity was provided by the HS swath on account of its most significant overlap with the target profile, therefore its use for datasets acquired after May 2018 is not validated and expected to be affected by significant inaccuracies.

e) combine_MS_NUBF

This module receives as input the products of the *MS_check* and *NUBF_3D_check* modules, together with the products received directly from the DPR main module.

When the SRT estimate of path attenuation is considered reliable, this module introduces a second criterion to diagnose NUBF. That is, where the PIA reliability factor (reliabFactor is equal to the ratio of the estimated PIA to the standard deviation of the estimate) are >3 for Ku and Ka, and the PIA_{Ku} >1 dB, the PIA ratio $PIA_{\mathrm{r}} = PIA_{\mathrm{Ka}}/PIA_{\mathrm{Ku}}$ is calculated and the logic shown in Figure 3.9.8 is applied.

[We define the reliability flag, reliabFlag in the following way (reliabFactor is the ratio of the estimated PIA to the standard deviation of the estimate: PIA/std_dev(PIA)):
reliabFlag = 1 when reliabFactor >3 (reliable)
reliabFlag = 2 when 1 <reliabFactor <3 (marginally reliable)
reliabFlag = 3 when reliabFactor <1 (unreliable)
reliabFlag = 4 when surface return <2 dB above noise (in this case, the PIA estimate is a lower bound)]

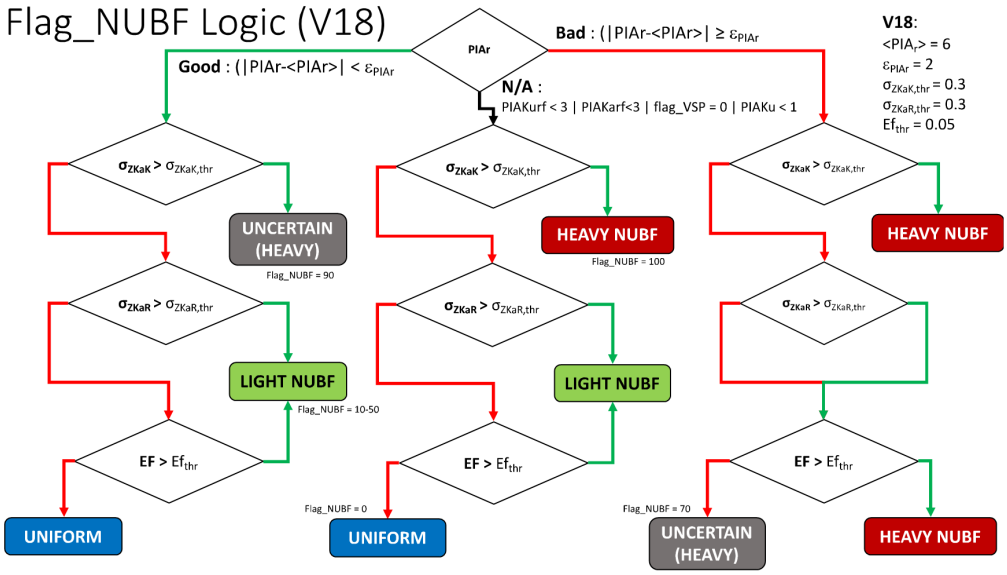


Figure 3.9.8: NUBF Reconciliation logic applied to the products of the MS_check, NUBF_3D_check and PRE, SRT and VER modules. For each decision point the green branch illustrates the positive outcome and the red branch illustrates the negative outcome.

Once NUBF is thus categorized, the interpretation of the MS flags is folded in a preliminary attribution as shown in Figure 3.9.9.

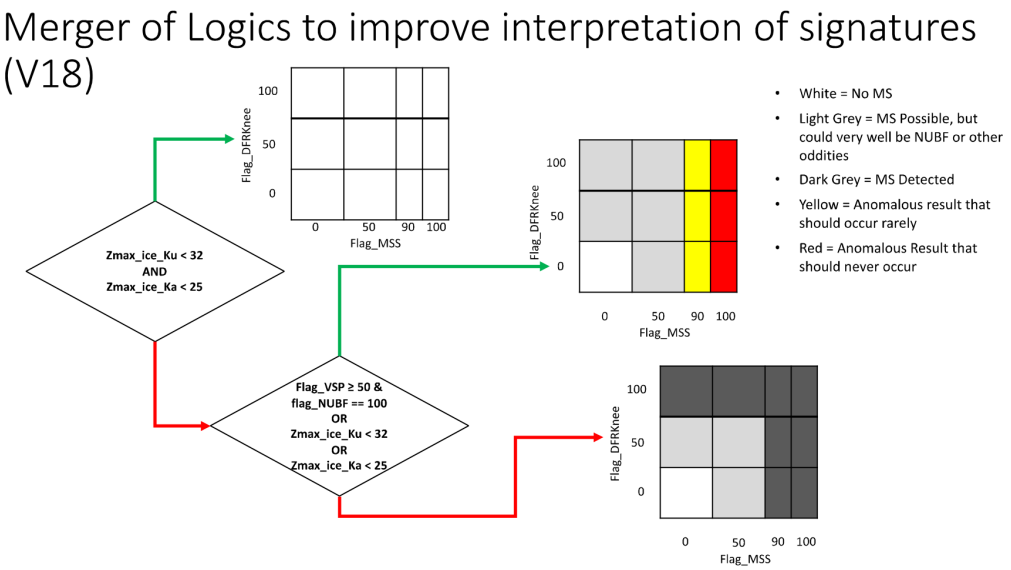


Figure 3.9.9: MS Reconciliation logic applied to the products of the MS_check, NUBF_3D_check and PRE, SRT and VER modules. For each decision point the green branch illustrates the positive outcome and the red branch illustrates the negative outcome.

3.9.3.2.3 Output variables

`MSindex`: this is the primary output for MS detection. It can assume a value of 0, 50 or 100 according to the logic shown in Figure 3.9.9.

`NUBFindex`: this is the primary output for NUBF detection. It can assume any value between 0 and 100 according to the logic shown in Figure 3.9.8.

`MSindexKu and MSindexKa`: provide additional information in regards to MS being detected only in Ka or also in Ku.

`MSsurfPeakIndexKu, MSsurfPeakIndexKa`: provide a diagnostic in regards to the presence of a visible peak in the echo attributed to the surface return according to the logic shown in Figure 3.9.6.

`MSthroughSurfIndexKu, MSthroughSurfIndexKa`: provide a diagnostic in regards to the presence of a visible MS tail through the location of the surface range according to the logic shown in Figure 3.9.6.

`MSkneeDFRindex`: provides a diagnostic in regards to the presence of a visible MS DFR knee according to the logic shown in Figure 3.9.5.

Details of the description for variables are described in Appendix or in the following websites:

`https://www.eorc.jaxa.jp/GPM/en/archives.html`

`https://arthurhou.pps.eosdis.nasa.gov/GPMprelimdocs.html`

3.9.4 Acronyms and abbreviations used in section 3.9

ADT:	Algorithm Development Team
ATBD:	Algorithm Theoretical Basis Document
API:	Application Programming Interface
Cal/Val:	Calibration and Validation
DAAC:	Distributed Active Archive Center
DFR:	Dual-Frequency Ratio
DPR:	Dual-Frequency Precipitation Radar
GSFC:	Goddard Space Flight Center
GPM:	Global Precipitation Measurement
HS:	High-sensitivity (interleaved) swath
IFOV:	Instantaneous Field Of View
JPL:	Jet Propulsion Laboratory
MS:	Multiple Scattering
NEDT:	Noise Equivalent Delta-T
NN:	Nearest Neighbor
NS:	Normal Swath
NUBF:	Non-Uniform Beam Filling
SDS:	Science Data System
TBD:	To Be Determined
TRG:	Trigger Module

4 VALIDATION (TEST AND VERIFICATION)

4.1 Pre-launch of the GPM core satellite

Before the launch of the GPM core satellite, we tested the algorithm test and physically validated it as much as possible with collaboration of the GPM GV team. In the algorithm test, algorithm mechanics and robustness were checked. More than one type of test data was needed. We generated synthetic Ka-band data in the central swath from actual TRMM PR data (Ku-band) by assuming a relationship among the DSD parameters and the attenuation by clouds and water vapor (Kubota et al. 2014). Airborne data (PR-2, APR-2) provided realistic dual-frequency radar data although with many unknown parameters (e.g., clouds). Ground-based dual-frequency radars including NASA's dual-frequency and dual-polarized Doppler radar (D3R) and JAXA's dual Ka-band radar system also provided realistic dual-frequency radar data. Synthetic data created by using a numerical model had an advantage that all parameters were known so that we could test the performance of the algorithm by checking whether the algorithm could retrieve the parameters correctly. Note that many parameters needed to reproduce radar echoes accurately were not handled well in most of the current numerical models. They included melting or partially frozen particles. Nevertheless, synthetic data created with simple assumptions were very helpful for sanity check of the algorithm. In the physical validation, various parameters of the DPR algorithm were tested with collaboration of the GPM GV team. The physical validation validated the parameters in the physical model of the precipitation system assumed in the algorithm. They included the DSD parameters, density and shape of snow and melting particles, width and structure of the melting layer, supercooled droplets, attenuation by the water vapor and cloud, inhomogeneity of rain distribution and so on.

4.2 Post-launch of the GPM core satellite

After the launch of the GPM core satellite, product validation is performed, in addition to the physical validation. Data taken nearly simultaneously by the GPM/DPR and the TRMM/PR enable us to make a direct comparison. Statistics such as averages and histograms of radar echoes taken by the DPR can be compared with the corresponding statistics of the TRMM/PR data. Ground instruments such as a dense rain gauge network and ground-radar including NASA's D3R and JAXA's dual Ka-band radar system can be used to validate the DPR products. Validations by using airborne data are desirable.

5 INTERFACE TO OTHER ALGORITHMS

The combined DPR-GMI algorithm requires outputs from the preparation module (PRE), the vertical profile module (VER), the surface reference technique module (SRT), and the classification module (CSF), in addition to the DPR radar reflectivity profiles. Precipitation detection from the PRE, pressure, temperature and humidity profiles from the VER, the PIA from the SRT, and bright band detection, altitude of the bright band (if it exists), and classification as convective or stratiform precipitation from the CSF will be used in the DPR-GMI combined algorithm.

The DPR products will be utilized also in the GPM Passive Microwave-Radar Enhanced Algorithm of an Optimal Estimation approach that uses the DPR/GMI data as the a-priori constraint.

6 REFERENCES

Awaka, J., T. Iguchi, H. Kumagai, and K. Okamoto, 1997: Rain type classification algorithm for TRMM precipitation radar, *Proc. IEEE IGARSS*, 317–319.

Awaka, J., T. Iguchi, and K. Okamoto, 1998: Early results on rain type classification by the Tropical Rainfall Measuring Mission (TRMM) precipitation radar, *Proc. 8th URSI Commission F Open Symp.*, Aveiro, Portugal, 143–146.

Awaka, J., T. Iguchi, and K. Okamoto, 2009: TRMM PR standard algorithm 2A23 and its performance on bright band detection, *J. Meteor. Soc. Japan*, **87A**, 31–52.

Awaka, J., M. Le, V. Chandrasekar, N. Yoshida, T. Higashiuwatoko, T. Kubota, and T. Iguchi, 2016: Rain Type Classification Algorithm Module for GPM Dual-Frequency Precipitation Radar, *J. Atmos. Oceanic Technol.*, **33**, 1887–1898.

Awaka, J. and S. Brodzik, 2019: Improvements of GPM DPR rain type classification algorithm, *IGARSS 2019*, 4470–4472, doi:10.1109/IGARSS.2019.8898365

Awaka, J., M. Le, S. Brodzik, T. Kubota, T. Masaki, V. Chandrasekar, and T. Iguchi, 2021: Development of precipitation type classification algorithms for a full scan mode of GPM Dual-Frequency Precipitation Radar, *J. Meteor. Soc. Japan*, **99**, 1253-1270.

Baldini, L. and E. Gorgucci, 2006: Identification of the melting layer through dual-polarization radar measurements at vertical incidence, *J. Atmos. Ocean. Technol.*, **23**, 829–839.

Bandera, J., A. D. Papatsoris, P. A. Watson, and J. W. Goddard, 1998: Method for detecting the extent of the melting layer, *Electron. Lett.*, **34(22)**, 2104–2105.

Battaglia, A., S. Tanelli, S. Kobayashi, D. Zrnic, R.J. Hogan, C. Simmer, Multiple- scattering in radar systems: a review, J. Quant. Spectrosc. Radiat. Transfer, 2010, 111 (6), 917-947.

Battaglia, A., S. Tanelli, K. Mroz, and F. Tridon (2015), Multiple scattering in observations of the GPM dual-frequency precipitation radar: Evidence and impact on retrievals, J. Geophys. Res. Atmos., 120, 4090–4101, doi:10.1002/2014JD022866.

Battaglia, A., K. Mroz, S. Tanelli, F. Tridon, and P. Kirstetter, 2016: Multiple-Scattering-Induced “ Ghost Echoes ” in GPM DPR Observations of a Tornadic Supercell. J. Appl. Meteor. Climatol., 55, 1653–1666, 10.1175/JAMC-D-15-0136.1

Battaglia, A., S. Tanelli, G. M. Heymsfield, and L. Tian, 2014: The dual wavelength ratio knee: a signature of multiple scattering in airborne ku–ka observations. J. Appl. Meteor. Climatol., 53, 1790–1808. doi:10.1175/JAMC-D-13-0341.1

Bechini, R., and V. Chandrasekar, 2015: A Semisupervised Robust Hydrometeor Classification Method for Dual-Polarization Radar Applications, *J. Atmos. Oceanic Technol.*, 32, 22–47.

Caylor, I.J., G.M. Heymsfield, R. Meneghini, and L.S. Miller, 1997: Correction of sampling errors in ocean surface cross-sectional estimates from nadir-looking weather radar. *J. Atmos. Oceanic Technol.*, **14**, 203-210.

Durden, S.L., L. Li, E. Im and S.H. Yueh, 2003: A surface reference technique for airborne Doppler radar measurements in hurricanes, *J. Atmos. Oceanic Technol.*, **20**, 269–275, 2003.

Durden, S.L., S. Tanelli, and G. Dobrowalski, 2010: CloudSat W-band radar measurements of surface backscatter, *IEEE Geosci. Remote Sens. Lett.*, vol. **7**, 401–405, doi:10.1109/LGRS.2010.2079314.

Durden S.L., S. Tanelli, and R. Meneghini, 2012: Using surface classification to improve surface reference technique over land, *Indian J. Radio & Space Physics*, **41**, 403–410.

Durden, S.L., 2018: Estimating attenuation from GPM radar reflectivity profiles without using surface reference, *IEEE Geosci. Remote Sens.*

Fabry, F. and I. Zawadzki, 1995, Long-term radar observations of the melting layer of precipitation and their interpretation, *J. Atmos. Sci.*, **52**, 838–851.

GPM/DPR Level-2 Algorithm Theoretical Basis Document, Iguchi et al. Rev Dec 17^{th}, 2021.

Harada, Y., H. Kamahori, C. Kobayashi, H. Endo, S. Kobayashi, Y. Ota, H. Onoda, K. Onogi, K. Miyaoka, and K. Takahashi, 2016: The JRA-55 Reanalysis: Representation of at-mospheric circulation and climate variability, *J. Meteor. Soc. Japan*, 94, 269-302, doi:10.2151/jmsj.2016-015.

Hines, E. L., 1983: Image processing techniques for the detection of the radar bright band, Ph.D. dissertation, Univ. Bradford, Bradford, U.K.

Hitschfeld, W. and J. Bordan, 1954: Errors inherent in the radar measurement of rainfall at attenuating wavelengths. *J. Meteor.*, **11**, 58-67.

Hou, Arthur Y., et al. "The global precipitation measurement mission." Bulletin of the American Meteorological Society 95.5 (2014): 701-722.

Iguchi, T. and R. Meneghini, 1994: Intercomparisons of single-frequency methods for retrieving a vertical rain profile from airborne or spaceborne radar data. *J. Atmos. Oceanic Technol.*, **11**, 1507–1516.

Iguchi, T., T. Kozu, J. Kwiatkowski, R. Meneghini, J. Awaka, and K. Okamoto, 2009: Uncertainties in the Rain Profiling Algorithm for the TRMM Precipitation Radar. *J. Meteor. Soc. Japan, Vol.*, **87A**, 1–30.

Iguchi, T., S. Seto, et al., 2012: An overview of the precipitation retrieval algorithm for the Dual-frequency Precipitation Radar (DPR) on the Global Precipitation Measurement (GPM) mission's core satellite. *SPIE*, Kyoto, Japan.

Iguchi, T.: Improvement of the minimum detectable precipitation echoes with the TRMM precipitation radar and the GPM dual-frequency precipitation radar, *2017 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2017, pp. 2728–2730, doi:10.1109/IGARSS.2017.8127560.

Iguchi, T., N. Kawamoto and R. Oki, 2018: Detection of intense ice precipitation with GPM/DPR. *J. Atmos. Oceanic Technol.*, **35**, 491–502, doi:10.1175/JTECH-D-17-0120.1.

JAXA 2021: GPM/DPR TRMM/PR L1 Product Format Documentation Version 4.0, https://www.eorc.jaxa.jp/GPM/doc/product/format/en/02.GPM_DPR_L1_Product_Format_Documentation_V7_E.pdf

JAXA 2022: GPM/DPR TRMM/PR L2/3 Product Format Documentation Version 5.1 https://www.eorc.jaxa.jp/GPM/doc/product/format/en/03.GPM_DPR_L2_L3_Product_Format_Documentation_V7_E.pdf

Kanemaru, K., T. Iguchi, T. Masaki, T. Kubota, 2020: Estimates of Spaceborne Precipitation Radar Pulsewidth and Beamwidth Using Sea Surface Echo Data, *IEEE Trans. Geosci. Remote Sens.*, **58(8)**, 5291–5303, doi:10.1109/TGRS.2019.2963090

Kanemaru, K., H. Hanado, K. Nakagawa, 2021: Improvement of the clutter removal method for the spaceborne precipitation radars, the IEEE Geoscience and Remote Sensing Society, *the International Geoscience and Remote Sensing Symposium 2021*, Virtual Symposium

Klaassen, W., 1988: Radar observations and simulation of the melting layer of precipitation, *J. Atmospheric Sciences*, **45**, 3741–3753.

Kobayashi, S., Y. Ota, Y. Harada, A. Ebita, M. Moriya, H. Onoda, K. Onogi, H. Kamahori, C. Kobayashi, H. Endo, K. Miyaoka, and K. Takahashi, 2015: The JRA-55 Reanalysis: General specifications and basic characteristics. *J. Meteor. Soc. Japan*, 93, 5-48, doi:10.2151/jmsj.2015-001.

Kojima, M., et al. Dual-frequency precipitation radar (DPR) development on the global "precipitation measurement (GPM) core observatory." Earth observing missions and sensors: Development, implementation, and characterization II. Vol. 8528. International Society for Optics and Photonics, 2012.

Kozu, T., 1995: A generalized surface echo radar equation for down-looking pencil beam radar. *IEICE Trans. Commun.*, **E78-B**, 1245–1248.

Kozu, T., and T. Iguchi, 1999: Nonuniform Beamfilling Correction for Spaceborne Radar Rainfall Measurement: Implications from TOGA COARE Radar Data Analysis. J. Atmos. Oceanic Technol., 16, 1722-1735, https://doi.org/10.1175/1520-0426(1999)016<1722:NBCFSR>2.0.CO;2

Kubota T., K. Kanemaru, N. Takahashi, K. Yamamoto, K. Furukawa, and T. Nio, 2024: Evaluation of effects on Dual-frequency Precipitation Radar observations due to the orbit boost of the GPM Core Observatory. *2024 IEEE International Geoscience and Remote Sensing Symposium.*

Kubota, T., T.Iguchi, M. Kojima, L. Liao, T Masaki, H. Hanado, R. Meneghini, R. Oki, 2016: A statistical method for reducing sidelobe clutter for the Ku-band precipitation radar onboard the GPM Core Observatory, *J. Atmos. Oceanic Technol.*, **33 (7)**, 1413-1428.

Kubota, T., T. Iguchi, T. Masaki, N. Yoshida, and R. Oki, 2018: Development of a statistical method for reducing sidelobe clutter in high sensitivity mode of GPM/KaPR, *Proc. IGARSS2018*, 8347-8348.

Kubota, T., M. Satoh, T. Nasuno, S. Seto, T. Iguchi, R. Oki , 2012: Development of cloud liquid water database using global cloud-system resolving model for GPM/DPR algorithm. *Proc. IGARSS 2012*, 350–353.

Kubota, T., N. Yoshida, S. Urita, T. Iguchi, S. Seto, R. Meneghini, J. Awaka, Member, H. Hanado, S. Kida, and R. Oki, 2014: Evaluation of precipitation estimates by at-launch codes of GPM/DPR algorithms using synthetic data from TRMM/PR observations. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, **7**, 3931–3944. (doi:10.1109/JSTARS.2014.2320960.)

Kubota, T., S. Seto, M. Satoh, T. Nasuno, T. Iguchi, T. Masaki, J. M. Kwiatkowski, and R. Oki, 2020: Cloud assumption of Precipitation Retrieval Algorithms for the Dual-frequency Precipitation Radar, *J. Atmos. Oceanic Technol.*, 37, 2015-2031, doi:10.1175/JTECH-D-20-0041.1.

Laviola, Sante, Giulio Monte, Vincenzo Levizzani, Ralph R. Ferraro, and James Beauchamp. 2020. "A New Method for Hail Detection from the GPM Constellation: A Prospect for a Global Hailstorm Climatology" Remote Sensing 12, no. 21: 3553.

Le, M. and V. Chandrasekar, 2013a: Precipitation Type Classification Method for Dual-Frequency Precipitation Radar (DPR) Onboard the GPM, *IEEE Trans. Geosci. Remote Sens.*, **51(3)**, 1784–1790.

Le, M. and V. Chandrasekar, 2013b: Hydrometeor Profile Characterization Method for Dual-Frequency Precipitation Radar Onboard the GPM, *IEEE Trans. Geosci. Remote Sens.*, **51(6-2)**, 3648–3658.

Le, M. and V. Chandrasekar, Graupel and Hail Identification Algorithm for the Dual-frequency Precipitation Radar (DPR) on the GPM Core Satellite. Journal of the Meteorological Society of Japan. 2021. Volume 99 Issue 1, pp.49-65.

Le, M. and V. Chandrasekar, A New Hail Product for GPM DPR Algorithm. *IGARSS', 2021*, Jul 12th 16th, Brussels.

Li, L., E. Im, L.N. Connor, and P.S. Chang, 2004: Retrieving ocean surface wind speed from the TRMM precipitation radar measurements, *IEEE Trans. Geosci. Remote Sens.*, **42**, 1271–1282.

Marzoug, M. and P. Amayenc, 1994: A class of single- and dual-frequency algorithms for rain rate profiling from a spaceborne radar. Part I: Principle and tests from numerical simulations. *J. Atmos. Oceanic Technol.*, **11**, 1480–1506.

Masaki, T., T. Iguchi, K. Kanemaru, K. Furukawa, N. Yoshida, T. Kubota, R. Oki, 2020: Calibration of the Dual-Frequency Precipitation Radar Onboard the Global Precipitation Measurement Core Observatory, IEEE Trans. Geosci. Remote Sens., 10.1109/TGRS.2020.3039978

Meneghini, R., and T. Kozu, 1990: *Spaceborne weather radar*. Artech House (Boston/London), Norwood, MA, 197pp.

Meneghini, R. and K. Nakamura, 1990: Range profiling of the rain rate by an airborne weather radar. *Remote Sens. Environ.*, **31**, 193–209.

Meneghini, R., T. Iguchi, T. Kozu, L. Liao, K. Okamoto, J.A. Jones, and J. Kwiatkowski, 2000: Use of the surface reference technique for path attenuation estimates from the TRMM Radar. *J. Appl. Meteor.*, **39**, 2053–2070.

Meneghini, R., J.A. Jones, T. Iguchi, K. Okamoto, and J. Kwiatkowski, 2004: A hybrid surface reference technique and its application to the TRMM Precipitation Radar, *J. Atmos. Oceanic Technol.*, **21**, 1645–1658.

Meneghini, R. and J.A. Jones, 2011: Standard deviation of spatially-averaged surface cross section data from the TRMM Precipitation Radar, *IEEE Geosc. Remote Sens. Letters*, **8**, 293–297.

Meneghini, R., L. Liao, S. Tanelli, and S. L. Durden, 2012: Assessment of the Performance of a Dual-Frequency Surface Reference Technique Over Ocean. *IEEE Trans. Geosci. Remote Sens.*, **50**, 2968–2977.

Meneghini, R., and L. Liao, 2013: Modified Hitschfeld-Bordan equations for attenuation-corrected rain reflectivity: Application to nonuniform beamfilling at off-nadir incidence. *J. Atmos. Oceanic Technol.*, **30**, 1149–1160.

Meneghini, R., and H. Kim, 2017: Minimizing the Standard Deviation of Spatially Averaged Surface Cross-Sectional Data From the dual-frequency precipitation radar. *IEEE Geosci. Remote Sens.*, **55**, 1709-1716.

Meneghini, R., L. Liao, T. Iguchi, and H. Kim, 2018: Hybrid estimates of path attenuation for the DPR. IGARSS 2018 Proceedings, Valencia, Spain, 23-27 July.

Meneghini, R., L. Liao, and G. M. Heymsfield, 2019: Attenuation Correction over Ocean for the HIWRAP Dual-Frequency Airborne Scatterometer. J. Atmos. Oceanic Technol., 36, 2015-2029.

Meneghini, R., H. Kim, L. Liao, J. Kwiatkowski, and T. Iguchi, 2021: Path attenuation estimates for the GPM Dual-frequency Precipitation Radar (DPR). J. Meteor. Soc. Japan, 99, 181–200, doi:10.2151/jmsj.2021-010.

Meneghini, R., L. Liao, and G.M Heymsfield, 2021: Relationship between Horizontal Wind Velocity and Normalized Surface Cross Section Using Data from the HIWRAP Dual-Frequency Airborne Radar. J. Atmos. Oceanic Technol., 38, 423-439.

Onogi, K., J. Tsutsui, H. Koide, M. Sakamoto, S. Kobayashi, H. Hatsushika, T. Matsumoto, N. Yamazaki, H. Kamahori, K. Takahashi, S. Kadokura, K. Wada, K. Kato, R. Oyama, T. Ose, N. Mannoji and R. Taira , 2007 : The JRA-25 Reanalysis. *J. Meteor. Soc. Japan*, **85**, 369–432.

Press W.H. et al., 1992: Numerical Recipes in FORTRAN, 2nd Edition. Cambridge University Press. ISBN 0 521 43064 X.

Press, W. H., Teukolsky, Flannery, S. B. P., and Vetterling W.T., 1992, "14.8 Savitzky-Golay Smoothing Filters" *Numerical Recipes: The Art of Scientific Computing*, Second Edition, 994 pp. Cambridge University Press.

Rosenkranz, P. W., 1975: Shape of the 5mm oxygen band in the atmosphere. *IEEE Trans. Ant. and Propag.*, AP-23, 498–506.

Savitzky, A., and Golay, M.J.E., 1964: Smoothing and Differentiation of Data by Simplified Least Squares Procedures. *Analytical Chemistry*. **36 (8)**, 1627–1639. doi:10.1021/ac60214a047.

Seto, S. and T. Iguchi, 2007: Rainfall-induced changes in actual surface backscattering cross sections and effects on rain-rate estimates by spaceborne precipitation radar. *J. Atmos. Oceanic Technol.*, **24**, 1693–1709.

Seto, S., and T. Iguchi, 2011: Applicability of the iterative backward retrieval method for the GPM Dual-frequency Precipitation Radar. *IEEE Transactions on Geoscience and Remote Sensing*, **49(6)**, pp.1827–1838.

Seto, S., T. Iguchi, and T. Oki, 2013: The basic performance of a precipitation retrieval algorithm for the Global Precipitation Measurement mission's single/dual-frequency radar measurements. *IEEE Transactions on Geoscience and Remote Sensing*, **51(12)**, 5239–5251.

Seto, S., T. Iguchi, T. Shimozuma, and S. Hayashi, 2015: NUBF correction method for the GPM/DPR level-2 algorithms. *Proc. IGARSS 2015*, 2612–2614, doi:10.1109/IGARSS.2015.7326347.

Seto, S., T. Iguchi, R. Meneghini, J. Awaka, T. Kubota, T. Masaki, and N. Takahashi, 2021: The Precipitation Rate Retrieval Algorithms for the GPM Dual-frequency Precipitation Radar. Journal of the Meteorological Society of Japan, 99(2), 205–237, doi:10.2151/jmsj.2021-011.

Smyth, T. J., A. J. Illingworth, and A. B. Smith, 1998: Radar estimates of rainfall rates at the ground in bright band and non-bright band events, *Q. J. R. Meteorol. Soc.*, **124(551)**, 2417–2434.

Steiner, M., R.A. Houze, Jr., and S.E. Yuter, 1995: Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data, *J. Appl. Meteor.*, **34**, 1978–2007.

Tagawa, T., 2009: Derivation of sub-footprint scale σ^0 observed by TRMM Precipitation Radar. *IEEE. Trans. Geosci. Remote Sens.*,*IGARSS 2008*, IV, 137–140.

Takahashi, N., and T. Iguchi, 2004: Estimation and correction of beam mismatch of the Precipitation Radar after an orbit boost of the Tropical Rainfall Measuring Mission Satellite, *IEEE. Trans. Geosci. Remote Sens.*, **42**, 2362–2369.

Takahashi, N., H. Hanado, and T. Iguchi, 2005: Estimation of path-integrated attenuation and its nonuniformity from TRMM/PR range profile data. *IEEE. Trans. Geosci. Remote Sens.*, **44**, 3276–3283.

Tan, J. and J.W.F. Goddard, 1995: The use of dual-polarisation techniques for bright-band detection with PPI-based radars, *Eng.-Colloq. Radar Meteorol.*, 11/1–11/6.

Tanelli, S., S.L. Durden, and E. Im, 2006: Simultaneous measurements of Ku- and Ka-band sea surface cross sections by an airborne radar. *IEEE. Trans. Geosci. Remote Sens. Letters*, **3**, 359–363.

Tanelli, S., G.F. Sacco, S.L. Durden, and Z.S. Haddad, 2012: Impact of non-uniform beam filling on spaceborne cloud and precipitation radar retrieval algorithms. *SPIE*, Kyoto, Japan.

Tanelli, S, G.F. Sacco, S. L. Durden and Z.S. Haddad, 2012: Impact of non-uniform beam filling on spaceborne cloud and precipitation radar retrieval algorithms. Proc. SPIE 8523, Remote Sensing of the Atmosphere, Clouds, and Precipitation IV, 852308; doi:10.1117/12.977375; doi:10.1117/12.977375.

Tilford, K.A., I. D. Cluckie, R. J. Griffith, and A. Lane, 2001: Vertical reflectivity characteristics and bright band correction, Proc. Radar Hydrol. Real Time Flood Forecast., *Proc. Adv. Course*, 47–65, European Communities.

TRMM Precipitation Radar Team, (2011): Tropical Rainfall Measuring Mission (TRMM) Precipitation Radar Algorithm Instruction Manual For Version 7, https://www.eorc.jaxa.jp/TRMM/documents/PR_algorithm_product_information/pr_manual/PR_Instruction_Manual_V7_L1.pdf

Ulaby, F.T., R. K. Moore, and A. K. Fung, 1981: *Microwave Remote Sensing: Active and Passive.* Vol. I. Artech House, Norwood, MA, 456pp.

Waters, J.W., 1976: Absorption and emission of microwave radiation by atmospheric gases, in Methods of Experimental Physics, M. L. Meeks, ed. 12, Part B, *Radio Astronomy*, Academic Press, Section 2.3.

Yuter, S. E., and R. A. Houze, Jr., 1997: Measurements of Raindrop Size Distributions over the Pacific Warm Pool and Implications for Z-R Relations, *J. Appl. Meteorol.*, **36(7)**, 847–867.

Zrnic, D.S., R. Raghavan, and V. Chandrasekar, 1994: Observations of copolar correlation coefficient through a bright band at vertical incidence, *J. Appl. Meteorol.*, **33(1)**, 45–52.

7 ACRONYMS

BB:	Bright band
CSF module:	Classification module
DB:	dual-beam
DF:	dual-frequency
DFR_m:	Measured Dual Frequency Ratio
DPR:	Dual-frequency Precipitation Radar
DSD:	drop size distribution
HS:	high-sensitivity
KuPR:	Ku-band precipitation radar
KaPR:	Ka-band precipitation radar
ML:	Melting layer
MS:	matched scan
NP-attenuation:	Attenuation due to non-precipitation particles
NS:	normal scan
NUBF:	non-uniform beam filling
PIA:	path-integrated attenuation
PRE module:	Preparation module
PSD:	particle size distribution
SB:	single-beam
SF:	single-frequency
SLV module:	Solver module
SRT:	surface reference technique
VER module:	Vertical module

Appendix-1 Product format

GPM/DPR
TRMM/PR
L2/3 Product Format Documentation

Version 5.3

March 2024

Japan Aerospace Exploration Agency

Revision history

revision	date	section	content, reason
Version 1.0	Sept. 2nd 2014	ALL	New
Version 2.0	Mar. 28th 2016	p.ii~vi	Correction of talble of contents
		p.56~	Addition Chapter 3 "Level 2 (2HSLH) Data Format Structure"
		p.60~	Addition Chapter 4 "Level 2(2HSLH) Contents of Objects in each Group"
		p.152~	Addition Chapter 8 " Level 3 (3GSLH) Data Format Structure "
		P.156~	Addition Chapter 9 " Level 3(3GSLH) Contents of Objects in each Group "
		p.164~	Addition Chapter 10 " Level 3 (3HSLH) Data Format Structure "
		p.168~	Addition Chapter 11 "Level 3(3HSLH) Contents of Objects in each Group "
		p.180~	A list of elements of each chapter is gathered in Chapter 12
		All page	Change of the chapter constitution
Version 3.0	May 9th 2017	P.7, P.15, P.25, P.56~P.61, P.226~227	Addition of TRG Group to MS swath of 2ADPR.
		P.11, P.41~42, P.190, P.198, P.206, P.214, P.222, P.231	Addtion of adjustFactor and snowIceCover to PRE Group.
		P.12, P.47, P.191, P.199, P.207, P.215, P.223, P.232	Addtion of flagHeavyIcePrecip and flagAnvil to CSF Group.
		P.13, P.51, P.224	Addition of flagSurfaceSnowfall and surfaceSnowfallIndex to Experimental Group.
		P.17	Addition of DOIauthority and DOIshortName to FileHeader meta data.
		P.78, P.84~85, P.141~P.150, P.248~249	Addition of the new channels to each data elements of 3DPRD.

revision	date	section	content, reason
Version 3.1	July 11th 2017	P.63, P.257~258	Change of “number of layers” for 2HSLH.
		P.64,67	In figure 3.2-1 and figure 4.1-1, addition of “AlgorithmRuntimeInfo”.
		P.72~73	In 4.2.7, addition of the value detail for rainTypeSLH
		P.160	Chang of “number of latitude”, “number of longitude” and “number of layers” for 3GSLH.
		P.161, P.165~168, P.170, P.259-260	Addition of the following variables to Grid Group for 3GSLH: shallowLHMean otherLHMean shallowQ1RMean otherQ1RMean shallowQ2Mean otherQ2Mean shallowPix otherPix
		P.175	Change of “number of latitude”, “number of longitude” and “number of layers” for 3HSLH.
		P.176,177 P.182~183, P.186,189, P.191, P.261-263	Addition of the following variables to Grid Group for 3HSLH: otherLHMean otherLHDev otherQ1RMean otherQ1RDev otherQ2Mean otherQ2RDev otherPix
Version 3.2	July 19th 2017	P.181~189, P.261~262	Change of a range of values that the variable excluding “the number of pixel” to Grid Group for 3HSLH can take.
Version 4.0	Sept. 20th 2018	Chapter, 1,3,4,8~12	Addition of TRMM/PR products. Addition of explanation of PRE Group.
		Section 1.5, 2.2	Addition of variable of CSF, SRT, FLG Group. Addition of explanation of CSF, PRE, SLV Group.
		Section 3.2, 4.2	Change of variable name of SLH.
		Section 5.2, 5.3, 5.4, 6.2	Addition of variable of PR and DPR level 3 products.

revision	date	section	content, reason
		Section 8.2, 8.3, 9.2,10.2, 11.2 Chapter 12	Addition of variable of SLH level 3 products. Addition and change of variables.
Version 4.1	Nov. 28th 2018	Section 2.2	Revise explanation of DPR Level2 products.
Version 5.0	Dec. 6th 2021	All chapter	Revise for the DPR, PR, and SLH of Level2/3 products.
Version 5.1	June 2nd 2022	Section 9.2, 11.2	Revise explanation of SLH and SLHT Level3(Orbital) and Level3(Monthly) products.
Version 5.2	July 1st 2023	All chapter Section 1.1 Section 2.2	Update “TBD” to the latest information Addition of some dimension “nfreq” and “nfreqHI”. Revise explanation of variable of navigation, PRE, CSF, SRT, DSD, Experimental, SLV and FLG.
Version 5.3	March 5th 2024	Section 2.2	Revises due to change in satellite altitude (GPM Orbit Boost) Correction of errors

Reference

(1) PRECIPITATION PROCESSING SYSTEM GLOBAL PRECIPITATION MEASUREMENT “File Specification for GPM Products”,

(2) PRECIPITATION PROCESSING SYSTEM GLOBAL PRECIPITATION MEASUREMENT “Metadata for GPM Products”,

(3) PRECIPITATION PROCESSING SYSTEM GLOBAL PRECIPITATION MEASUREMENT “File Specification for GPM Products”,

(4) NOAA NESDIS CENTER FOR SATELLITE APPLICATIONS AND RESEARCH GLOBAL 4KM MULTISENSOR AUTOMATED SNOW/ICE MAP (GMASI) ALGORITHM THEORETICAL BASIS DOCUMENT

Table of Contents

1. Level 2 Data Format Structure

In the V06X (experimental product), a new format was implemented including “FS” which is defined as the full swath dual-frequency product with 125 m range resolution. In the V07A and later version, this FS format is applied to data taken both before and after the scan pattern change of the KaPR in May 2018 (Figure 1).

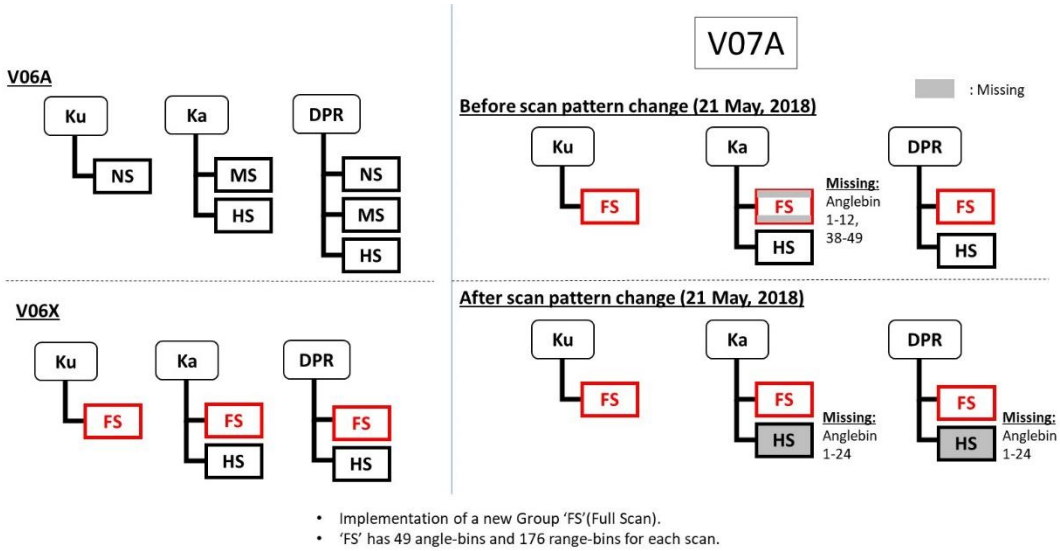


Figure 1 Changes of File structure from V06 to V07.

1.1. Dimension definition

Dimension definitions. All bin numbers are specified starting with 1.

- nscan
  - Number of scans in the granule.
- nray
  - Number of angle bins in each scan.
  - 49 angle bins in each scan in the FS swath (1–49).
  - 24 angle bins in each scan in the HS swath (1–24).
- nbin
  - Number of range bins in each ray.
  - 176 range bins in each ray in the FS swath (1–176).
  - 88 range bins in each ray in the HS swath (1–88).
- nbin SZP
  - 7 Number of range bins for sigmaZeroProfile. (FS)
  - 5 Number of range bins for sigmaZeroProfile. (HS)
- nfreq
  - 2 Number of frequency dependency (2ADPR FS).
- nfreqHI
  - 3 Number of frequency dependency (2ADPR FS).
- nNP
  - 4 Number of NP kinds.
- nearFar
  - 2 Near reference, Far reference.
- foreBack
  - 2 Forward, Backword.
- method
  - 6 Number of SRT methods.
- nNode
  - 5 Number of binNode.
- nDSD
  - 2 Number of DSD parameters. Parameters are N0 and D0.
- LS
  - 2 Liquid, Solid.
- nNUBF
  - 3 Number of NUBF parameters.
- nsdew

- 3 Number of SRT parameters.

"FS" is called as Full scan Swath in 2AKu, 2AKa, 2ADPR and 2APR.

"HS" is called as High sensitivity beam scan Swath in 2AKa and 2ADPR respectively.

1.2. Data Format Structure for 2AKu and 2APR

The Ku Level-2A product, 2AKu, “Ku precipitation”, is defined as a swath structure, which is called “FS”. The PR Level-2A product, 2APR, “Ku precipitation”, is the same with 2AKu and there are no differences between 2AKu and 2APR.

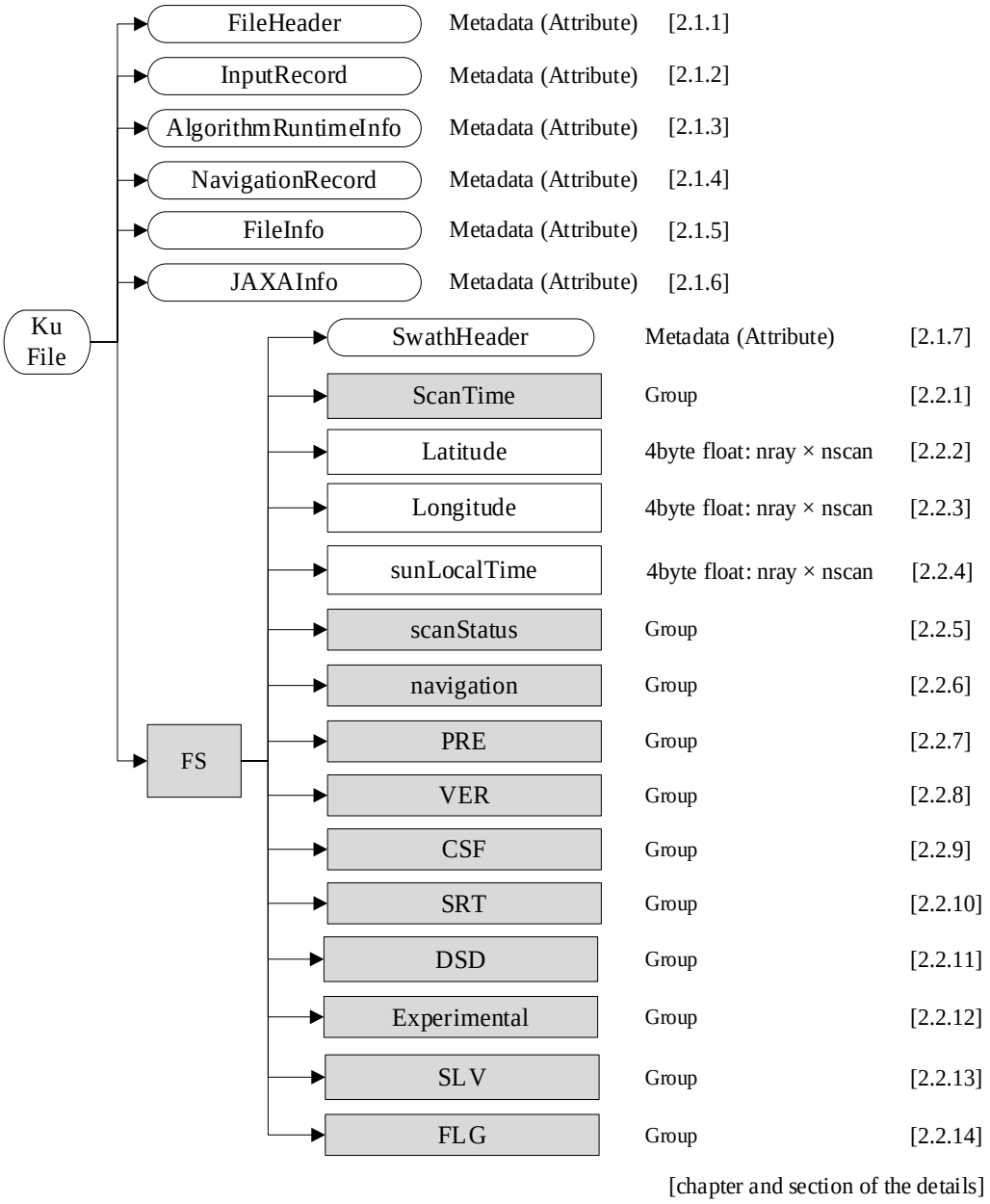


Figure 1.2-1 Data Format Structure for 2AKu and 2APR

1.3. Data Format Structure for 2AKa

The Ka Level-2A product, 2AKa, “Ka precipitation”, is defined as two-swath structures, which are called “FS” and “HS”.

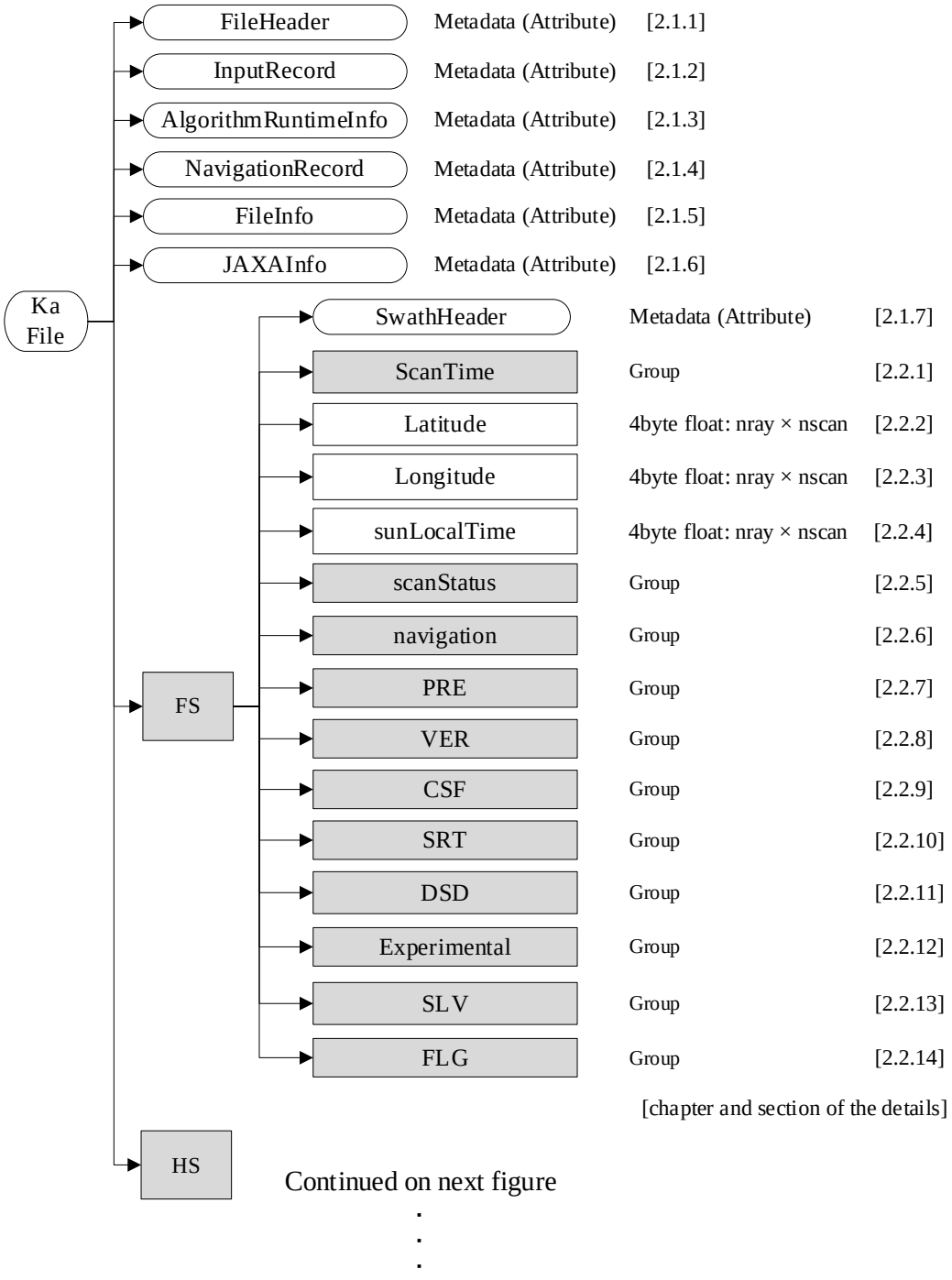


Figure 1.3-1 Data Format Structure for 2AKa

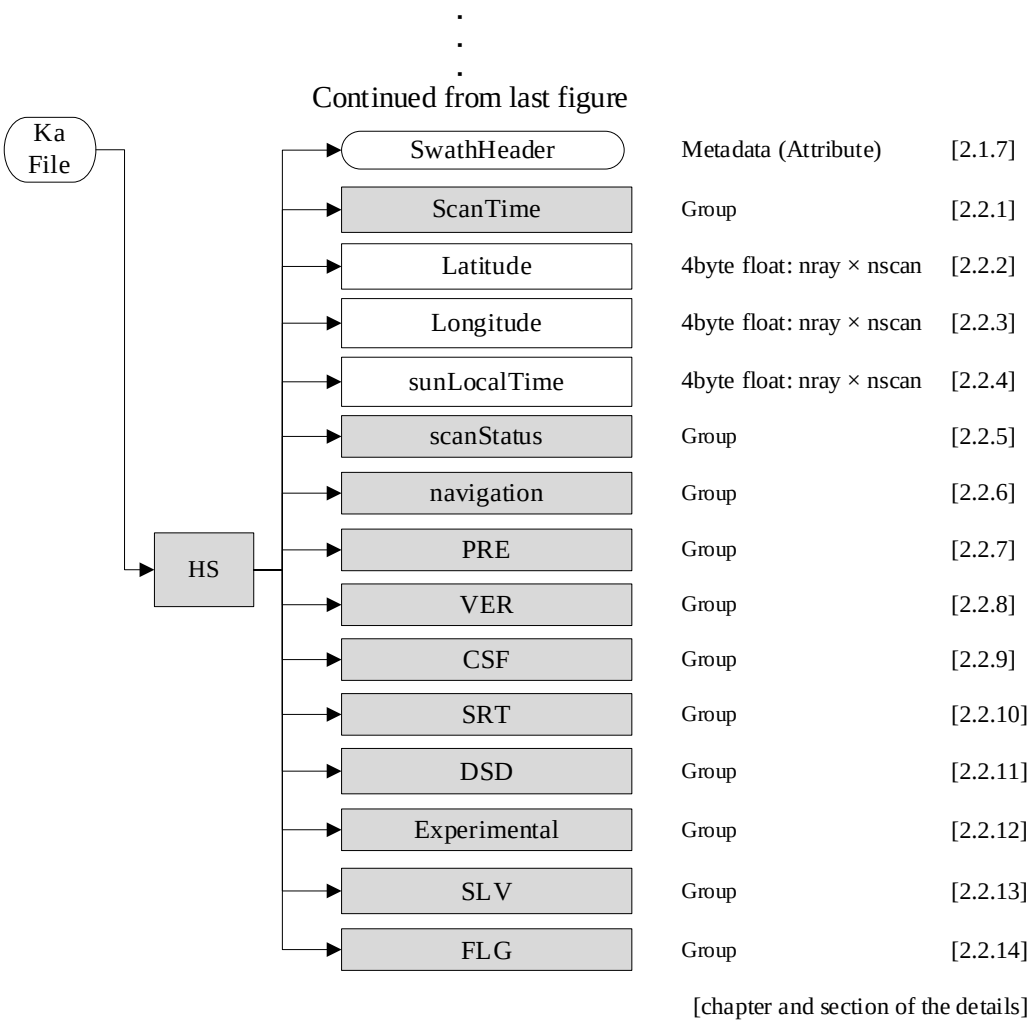


Figure 1.3-2 Data Format Structure for 2AKa

1.4. Data Format Structure for 2ADPR

The DPR Level-2A product, 2ADPR, “DPR precipitation”, is defined as two swath structures, which are called “FS” and “HS”. Some variables which have a frequency dependency have an array of “nfreq” and “nfreqHI”. See the description of each variable for details.

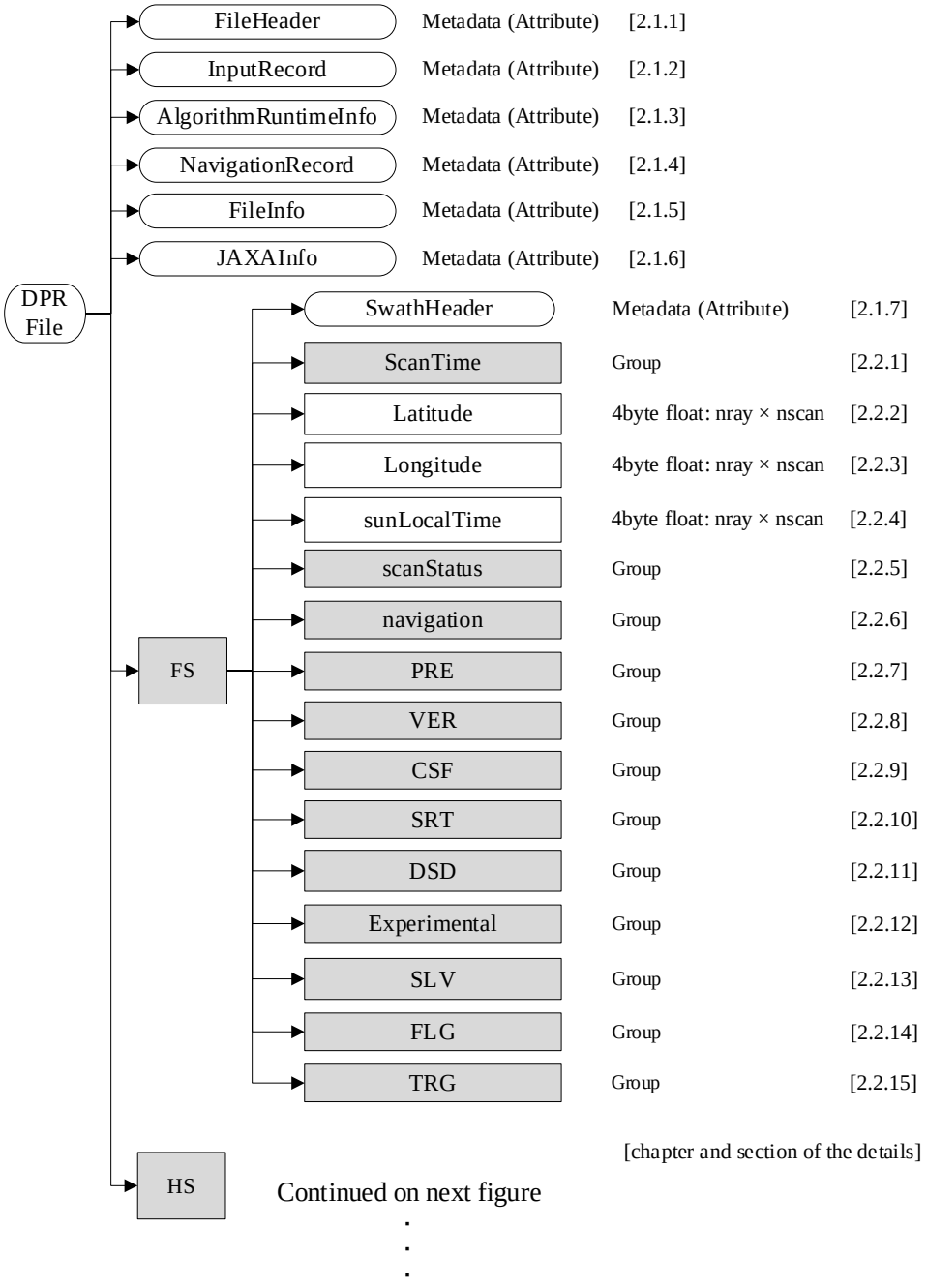


Figure 1.4-1 Data Format Structure for 2ADPR

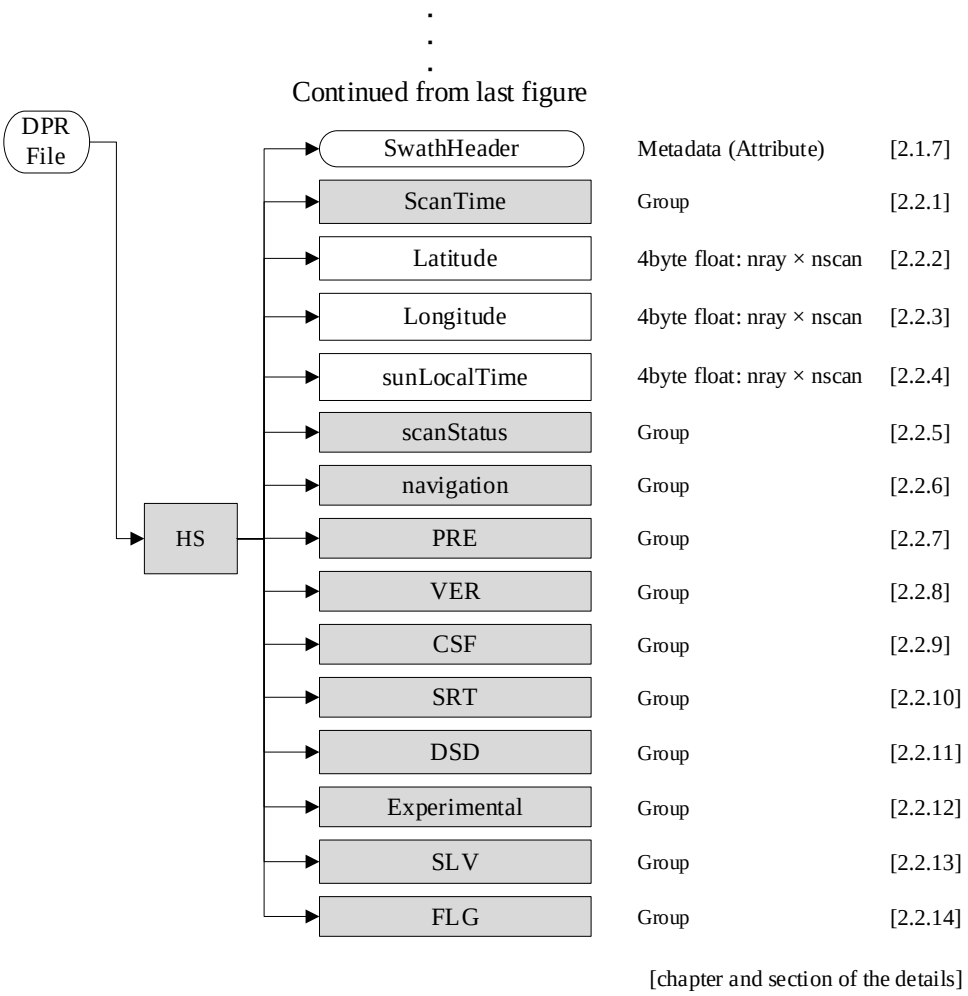


Figure 1.4-2 Data Format Structure for 2ADPR

1.5. Data Format Structure for each Group

Each group's structure is shown in this section. Structures in each grid are common. However, the number of rays and range bins are different as shown in section 1.1.

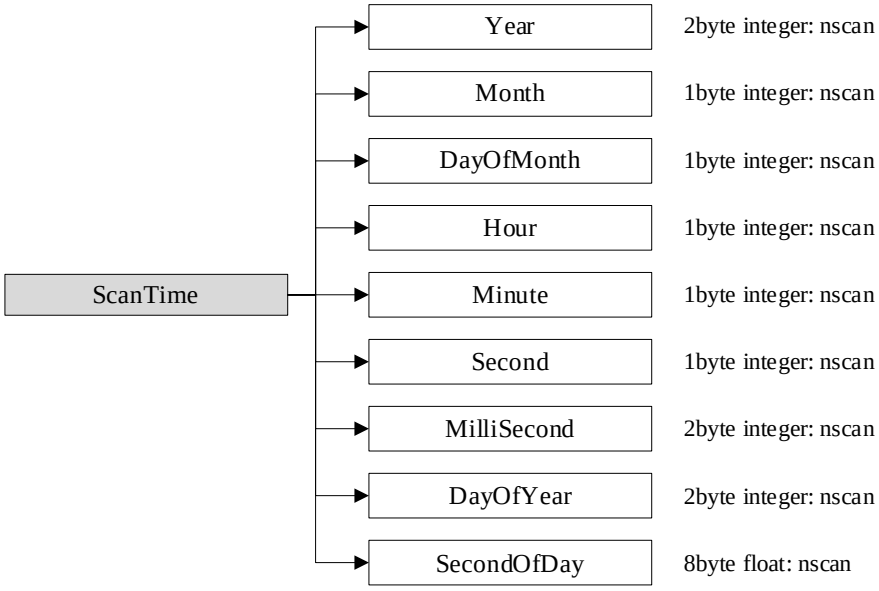


Figure 1.5-1 Data Format Structure for ScanTime Group

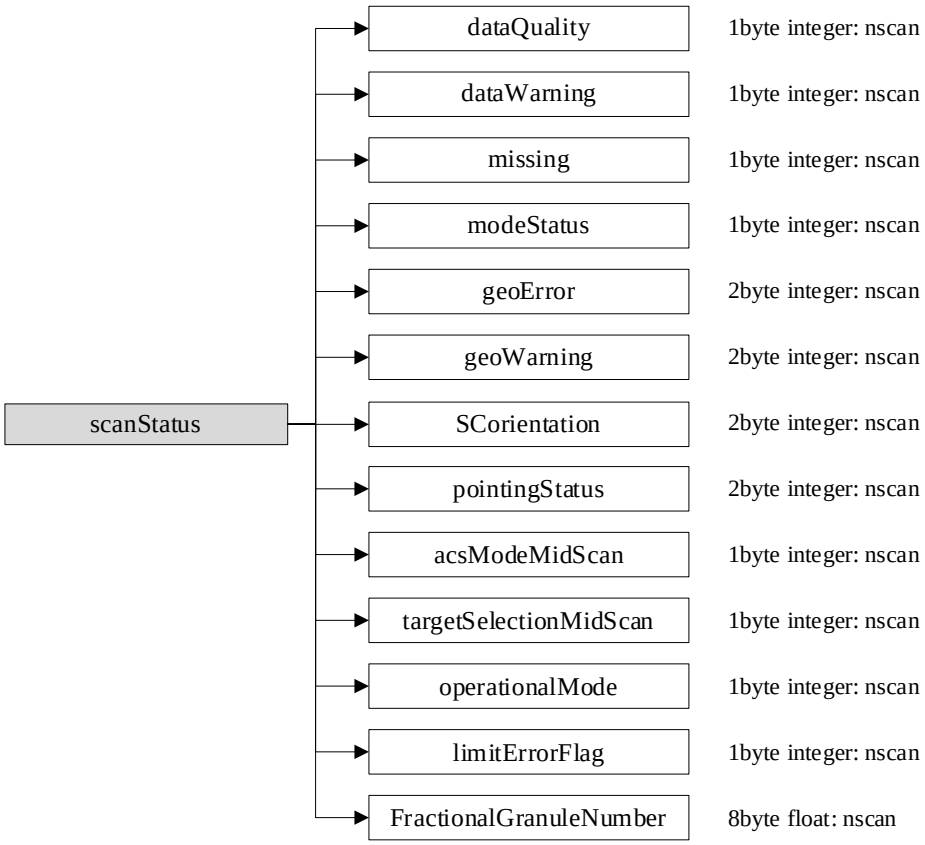


Figure 1.5-2 Data Format Structure for scanStatus Group

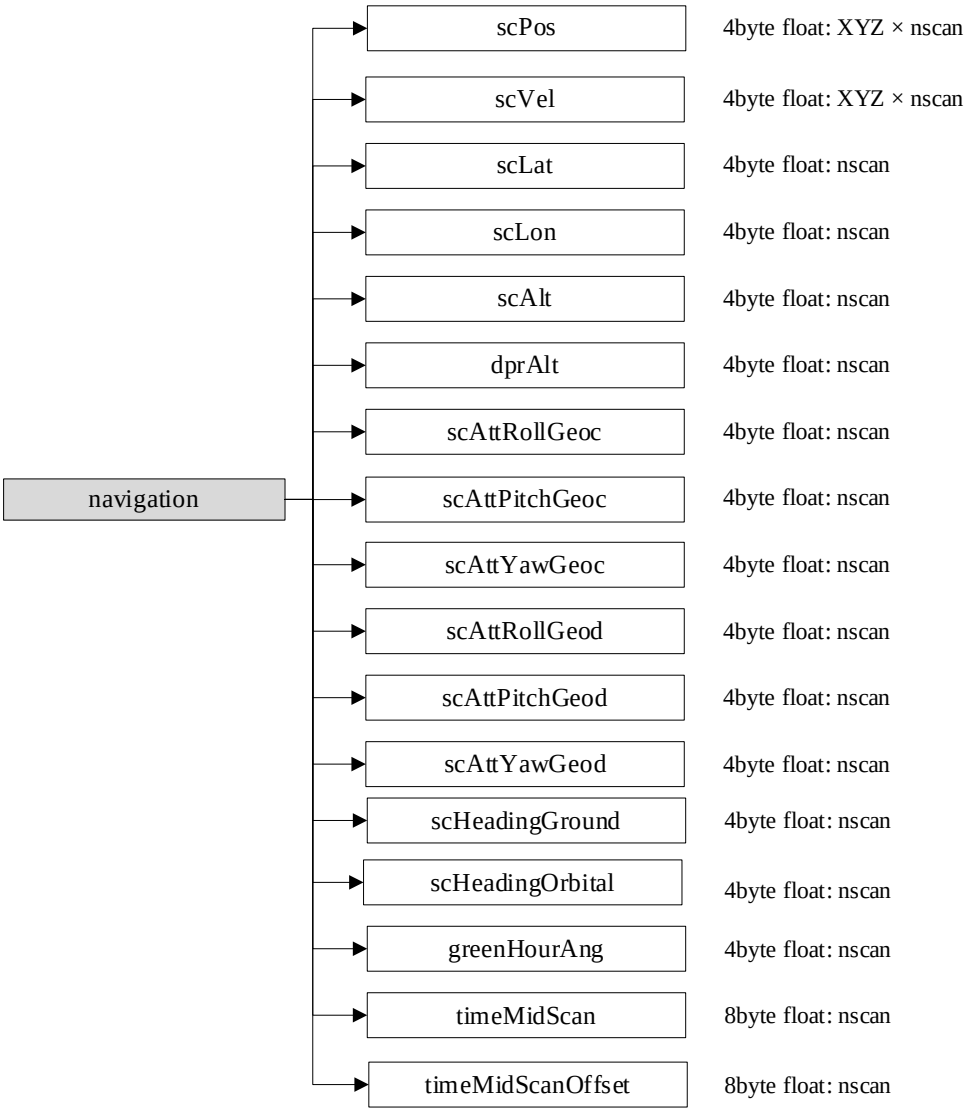


Figure 1.5-3 Data Format Structure for navigation Group

Figure 1.5-4 Data Format Structure for PRE Group

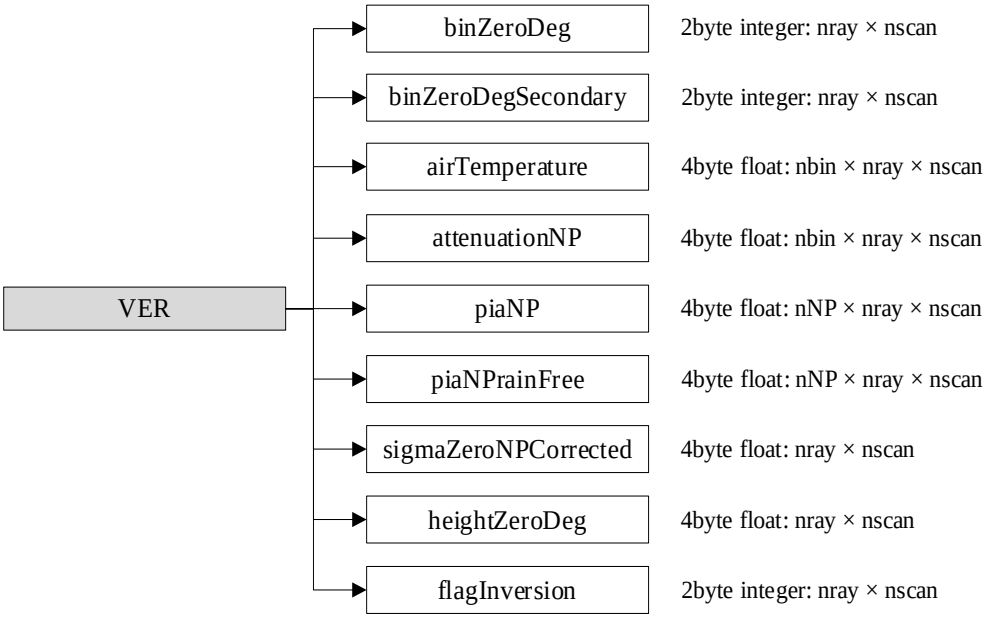


Figure 1.5-5 Data Format Structure for VER Group

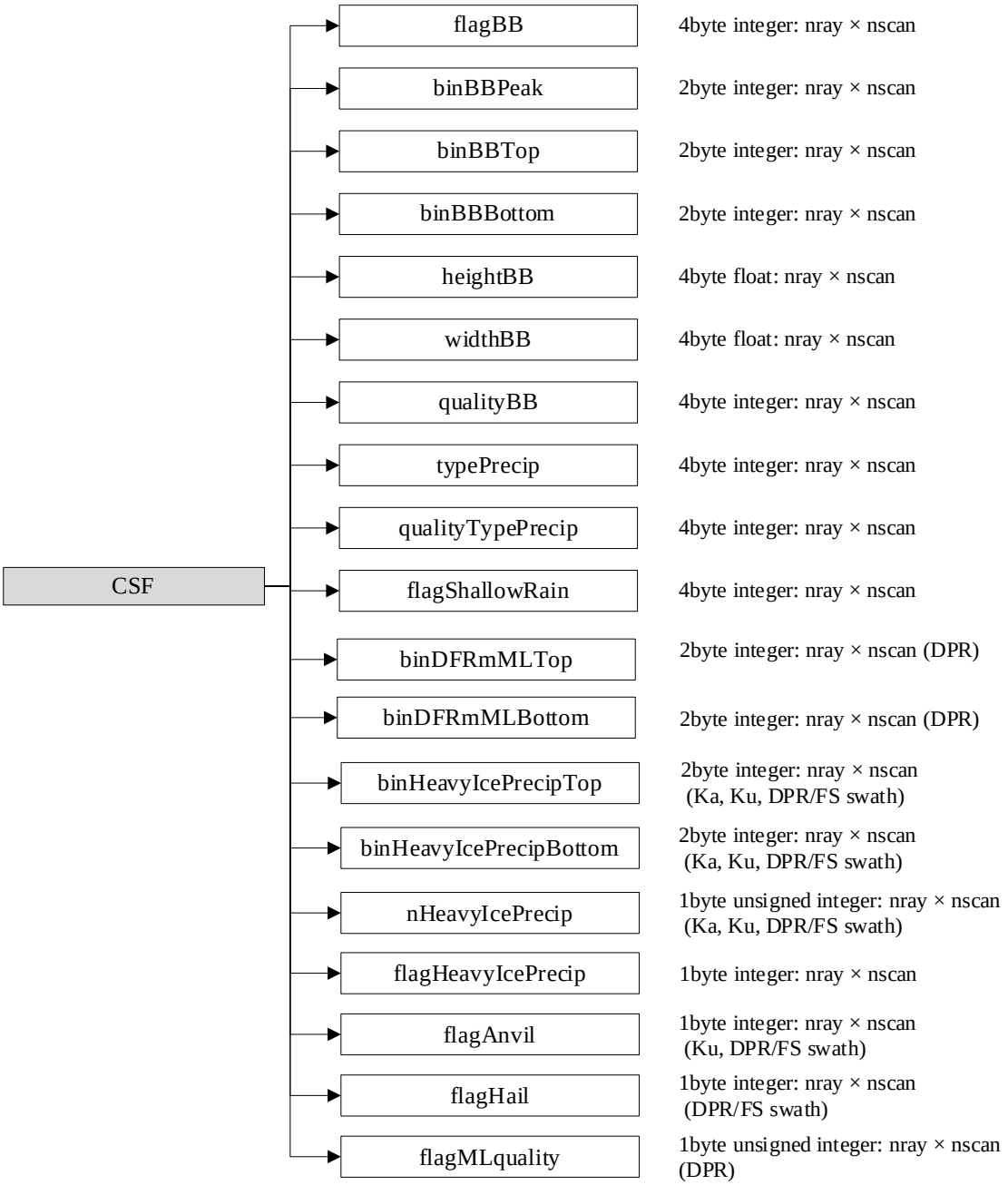


Figure 1.5-6 Data Format Structure for CSF Group

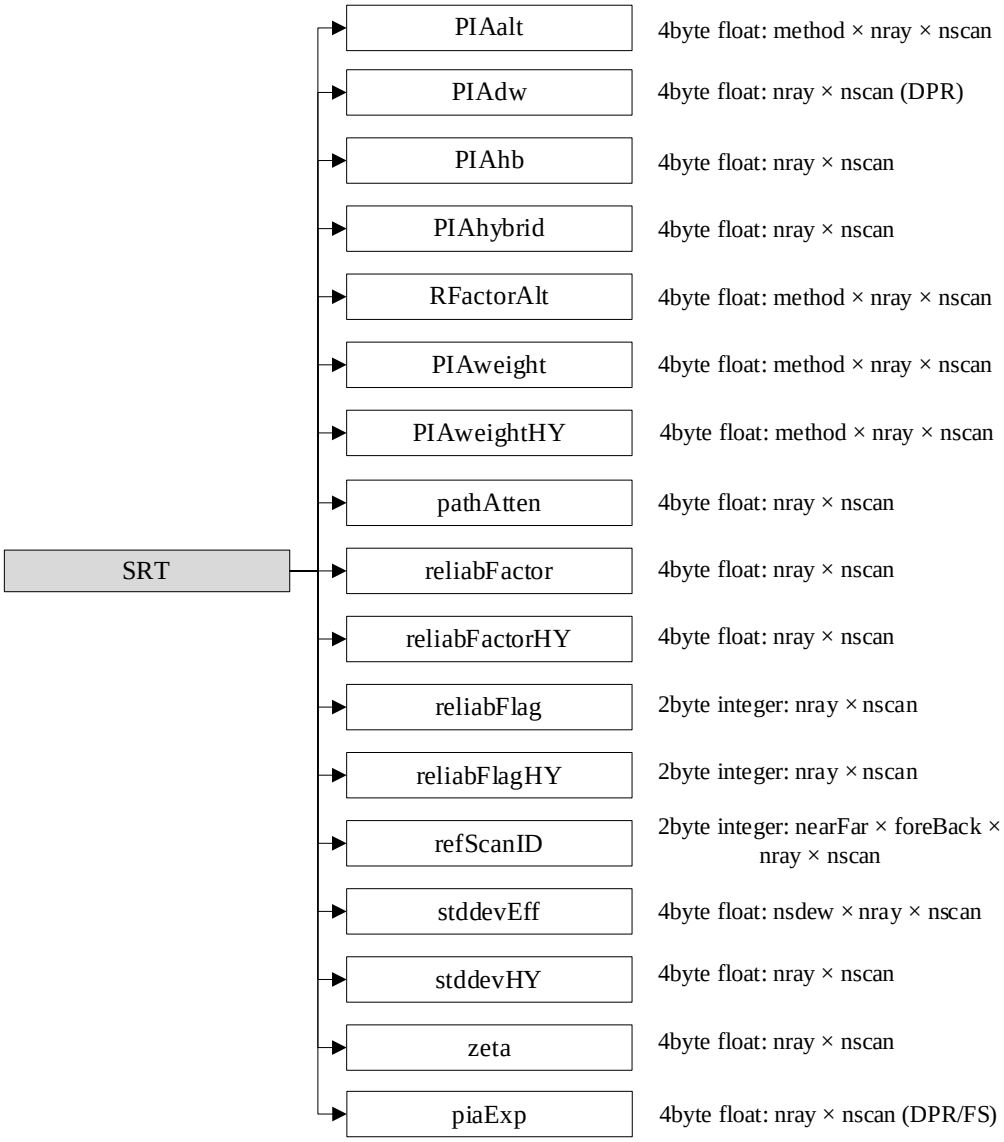


Figure 1.5-7 Data Format Structure for SRT Group

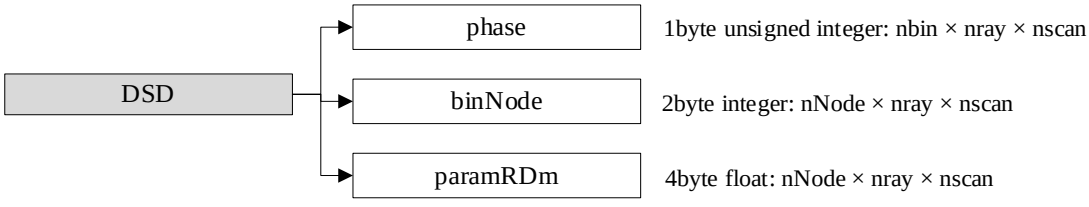


Figure 1.5-8 Data Format Structure for DSD Group

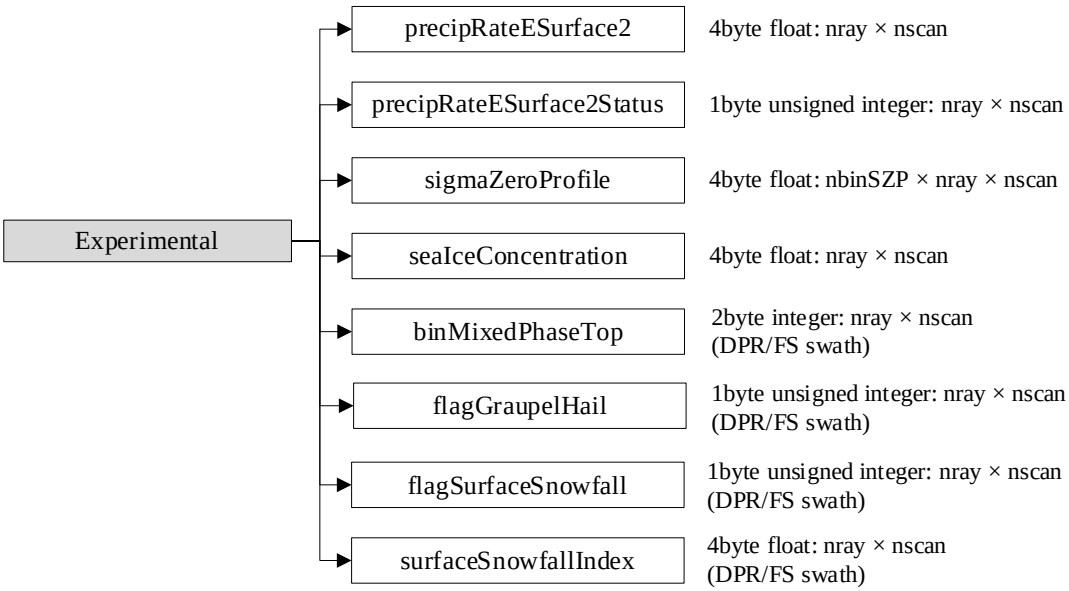


Figure 1.5-9 Data Format Structure for Experimental Group

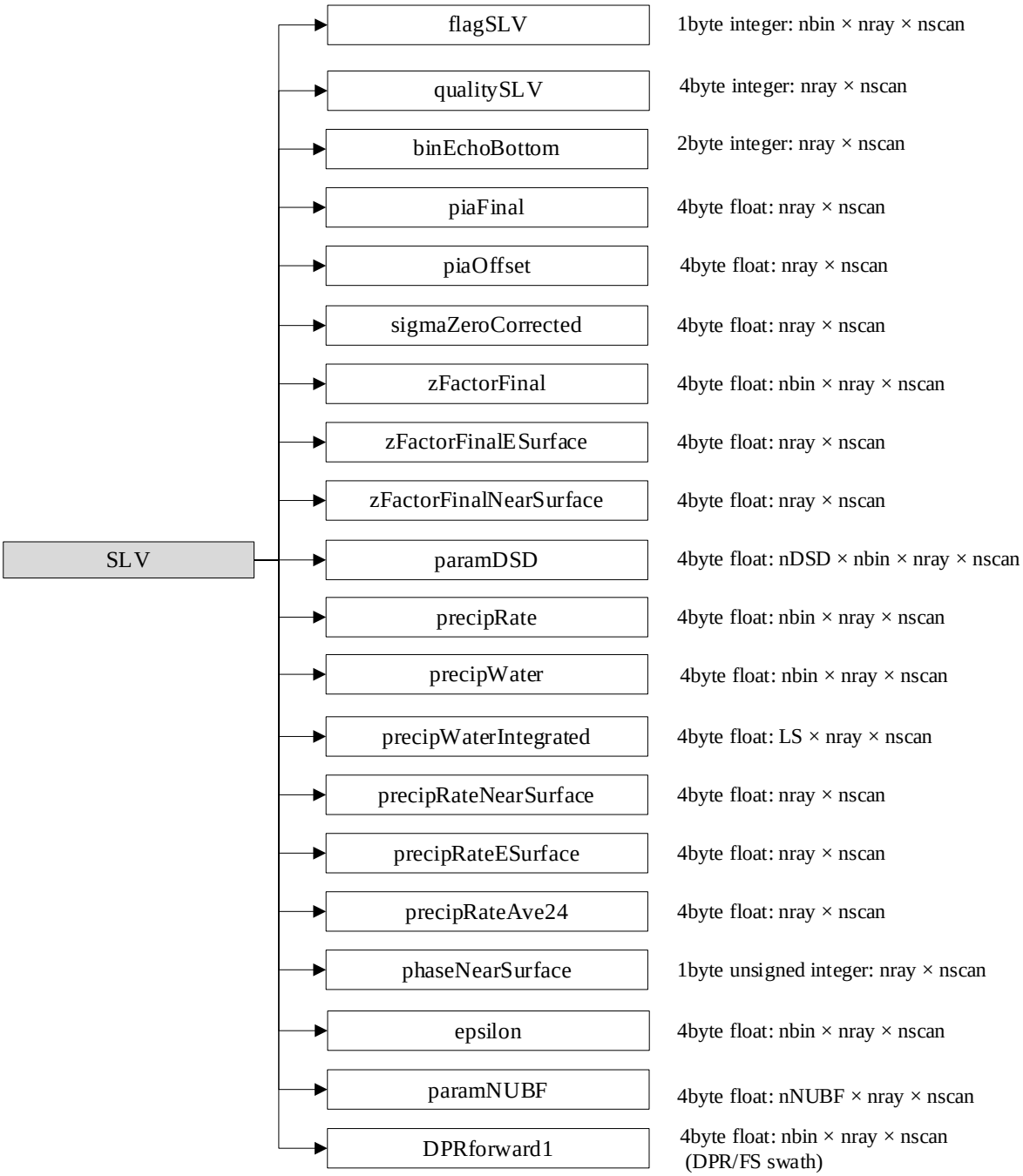


Figure 1.5-10 Data Format Structure for SLV Group

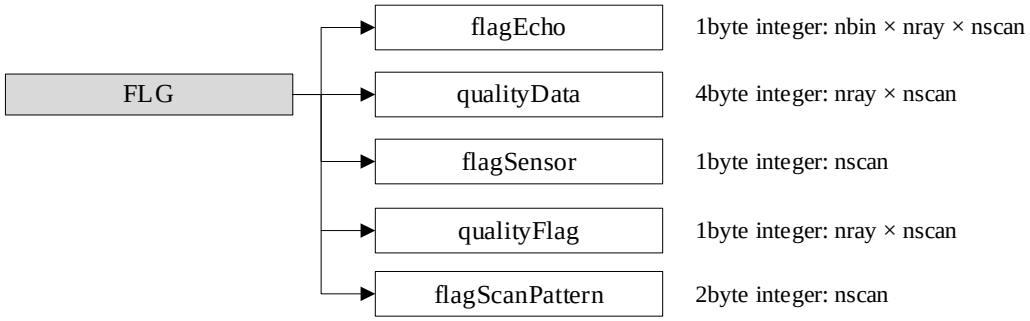


Figure 1.5-11 Data Format Structure for FLG Group

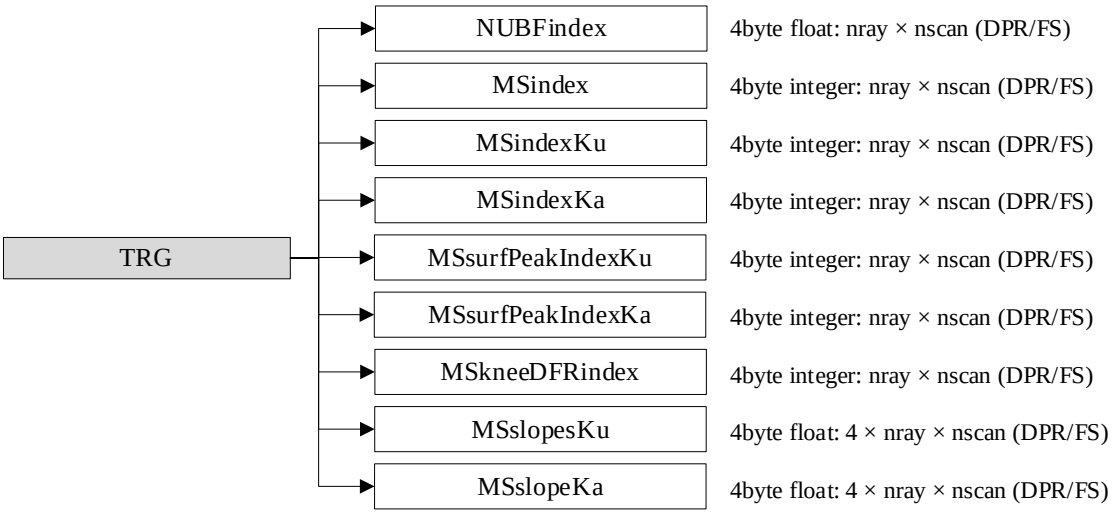


Figure 1.5-12 Data Format Structure for TRG Group

2. Level 2 Contents of Objects in each Group

2.1. Metadata

Metadata has seven elements. Figure 2.1-1 shows metadata structure.

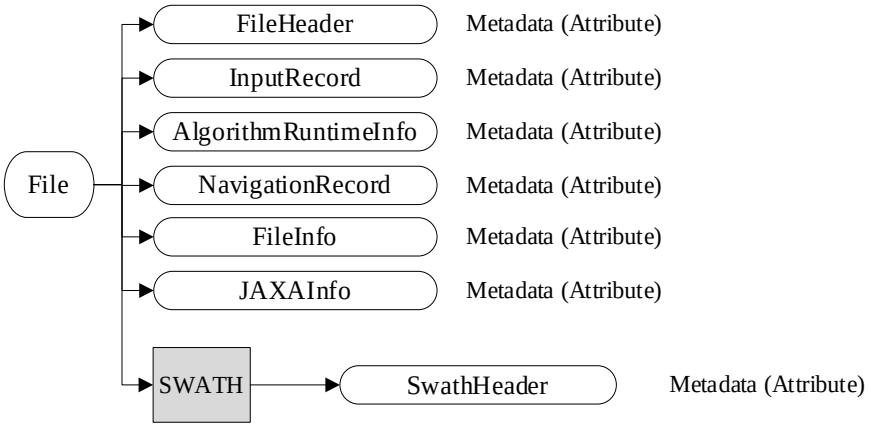


Figure 2.1-1 L2 Metadata

2.1.1. FileHeader

FileHeader contains metadata of general interest. This group appears in all data products. Table 2.1-1 shows each metadata elements in FileHeader.

Table 2.1-1 FileHeader Elements

No	Element	Description	Data size (bytes)
1	DOI	Digital Object Identifier. *Value is blank currently.	256
2	DOIauthority	Digital Object Identifier Authority.	256
3	DOIshortName	Digital Object Identifier Short Name. *Value is blank currently.	256
4	AlgorithmID	The algorithm that generated this product, e.g., 2A12.	50
5	AlgorithmVersion	The version of the algorithm that generated this product.	50
6	FileName	The file name of this granule.	50

No	Element	Description	Data size (bytes)
7	SatelliteName	Values are: TRMM GPM MULTI F10 ... F18 AQUA GCOMW1 CORIOLIS MT1 NOAA15 ... NOAA19 METOPA NPP. More values will be added as they are known.	10
8	InstrumentName	Values are: PR TMI VIRS PRTMI KU KA DPR GMI DPRGMI MERGED SSMI SSMIS AMSRE AMSR2 WINDSAT MADRAS AMSUA AMSUB SAPHIR MHS ATMS. More values will be added as they are known.	10
9	GenerationDateTime	The date and time this granule were generated. The format is YYYY-MM-DDTHH:MM:SS.sssZ, where YYYY is 4-digit year, MM is month number, DD is day of month, T is "T", HH is hour, MM is minute, SS is second, sss is millisecond, and Z is "Z". All fields are zero-filled. The missing value is constructed by replacing all digits with 9, i.e., 9999-99-99T99:99:99.999Z.	50
10	StartGranuleDateTime	The start time defining this granule. The format is the same as GenerationDateTime. DETAILS: An orbital granule starts when the satellite is at the position defined by GranuleStart. Thus, the start time is not the first scan time. Some algorithms have overlap scans in the file before the start time as defined in SwathHeader. A monthly granule starts on the first ms of the month, for example March 1998 would be 1998-03-01T00:00:00.000Z.	50
11	StopGranuleDateTime	The stop time defining this granule. The format is the same as GenerationDateTime. DETAILS: An orbital granule stops when the satellite is at the position defined by GranuleStart. Thus, the stop time is not the last scan time. Some algorithms have overlap scans in the file after the stop time as defined in SwathHeader. A monthly granule stops on the last ms of the month, for example March 1998 would be 1998-03-31T23:59:59.999Z.	50
12	GranuleNumber	The number of this granule, which starts as in GranuleStart. If the GranuleStart is identical to the orbit start, then the GranuleNumber will be the same as the orbit number. The GranuleNumber will have 6 digits, including leading zeroes, for example 001234.	50
13	NumberOfSwaths	The number of swaths in this granule.	50
14	NumberOfGrids	The number of grid structures in this granule.	50
15	GranuleStart	The starting place in the orbit of this granule. Currently defined values are “SOUTHERNMOST LATITUDE” and “NORTHBOUND EQUATOR CROSSING”.	50

No	Element	Description	Data size (bytes)
16	TimeInterval	The time interval covered by this granule. Values are “ORBIT”, “HALF ORBIT”, “HALF HOUR”, “HOUR”, ”3 HOUR”, “DAY”, ”DAY ASC”, DAY DES”, ”MONTH”, ”CONTACT”.	50
17	ProcessingSystem	The name of the processing system, e.g., ”PPS”, “JAXA”.	50
18	ProductVersion	The data version assigned by the processing system.	50
19	EmptyGranule	Whether a granule is empty. Values are "EMPTY" or "NOT EMPTY".	50
20	MissingData	The number of missing scans.	50

2.1.2. InputRecord

InputRecord contains a record of input files for this granule. This group appears in Level1, Level 2, and Level 3 orbital data products. Level 3 times averaged products have the same information separated into 3 groups since they have many inputs. Table 2.1-2 shows each metadata elements in InputRecord.

Table 2.1-2 InputRecord Elements

No	Element	Description	Data size (bytes)
1	InputFileNames	A list of input file names for this granule.	1000
2	InputAlgorithmVersions	A list of algorithm versions of the input files for this granule.	1000
3	InputGenerationDateTimes	A list of generation date times of the input files for this granule. The format is the same as GenerationDateTime.	1000

2.1.3. AlgorithmRuntimeInfo

AlgorithmRuntimeInfo contains text runtime information written by the algorithm. This group is a “Long Metadata Group”, which has no elements. This group appears in products if the algorithm developer asks for it.

2.1.4. NavigationRecord

NavigationRecord contains navigation metadata for this granule. This group appears in Level 1, Level 2, and Level 3 orbital data products. Table 2.1-3 shows each metadata elements in NavigationRecord.

Table 2.1-3 NavigationRecord Elements

No	Element	Description	Data size (bytes)
1	LongitudeOnEquator	The longitude where the satellite crosses the equator going from south to north.	50
2	UTCDateTimeOnEquator	The UTC time when the satellite crosses the equator going from south to north. The format is the same as GenerationDate Time.	50
3	MeanSolarBetaAngle	The average solar beta angle in this granule.	50
4	EphemerisFileName	Name of the ephemeris file input for processing.	50
5	AttitudeFileName	Name of the attitude file input for processing.	50
6	GeoControlFileName	Name of the GeoTK Control Parameters File input for processing.	50
7	EphemerisSource	Values are ”0 CONSTANT INPUT TEST VALUE”, ”1 GROUND ESTIMATED STATE (GES)”, ”2 GPS FILTERED SOLUTION (GEONS)”, ”3 GPS POINT SOLUTION (PVT)”, ”4 ON BOARD PROPAGATED (OBP)”, ”5 OEM GROUND EPHEMERIS FILE”, ”6 GEONS WITH FALLBACK AS FLAGGED”, ”7 PVT WITH FALLBACK AS FLAGGED”, ”8 OBP WITH FALLBACK AS FLAGGED”, ”9 GES WITH FALLBACK AS FLAGGED”.	50
8	AttitudeSource	values are ”0 CONSTANT INPUTS FOR TESTING”, ”1 ON BOARD CALCULATED PITCH ROLL YAW”	50
9	GeoToolkitVersion	Version of the GeoToolkit.	50

No	Element	Description	Data size (bytes)
10	SensorAlignmentFirstRotationAngle	Alignment angle, first rotation, in degrees. Rotation adjustment from sensor coordinates to the Attitude Control System Flight Coordinates.	50
11	SensorAlignmentSecondRotationAngle	Alignment angle, second rotation, in degrees.	50
12	SensorAlignmentThirdRotationAngle	Alignment angle, third rotation, in degrees.	50
13	SensorAlignmentFirstRotationAxis	Euler rotation sequence, first rotation axis. Values are "1","2", "3" (representing X, Y, Z).	50
14	SensorAlignmentSecondRotationAxis	Euler rotation sequence, second rotation axis. Values are "1","2", "3" (representing X, Y, Z).	50
15	SensorAlignmentThirdRotationAxis	Euler rotation sequence, third rotation axis. Values are "1","2", "3" (representing X, Y, Z).	50

2.1.5. FileInfo

FileInfo contains metadata used by the PPS I/O Toolkit. This group appears in all data products. Table 2.1-4 shows each metadata elements in FileInfo.

Table 2.1-4 FileInfo Elements

No	Element	Description	Data size (bytes)
1	DataFormatVersion	The version of the data format used to write this file. This version is separate for each AlgorithmID. The order is: "a" "b" ... "z" "aa" "ab" ... "az" "ba" "bb".	50
2	TKCodeBuildVersion	Usually TK CodeBuildVersion is "1". If the I/O routines built by TKIO change even though the DataFormatVersion is unchanged, then TK CodeBuildVersion increments to "2", "3", ...If subsequently DataFormatVersion changes, TKCodeBuildVersion becomes "1" again.	50
3	MetadataVersion	The version of metadata used to write this file. This version is separate for each AlgorithmID. The order is: "a" "b" ... "z" "aa" "ab" ... "az" "ba" "bb" ...	50
4	FormatPackage	The underlying format of this granule. Values are “HDF4”, “HDF5”, “NETCDF”, “TKBINARY”.	50
5	BlueprintFilename	The filename of the primary blueprint file that defined the format used to write this file.	50
6	BlueprintVersion	The BlueprintVersion of the format definition.	50
7	TKIOVersion	The version of TKIO used to create I/O routines to write this file. TKIOVersion does not define the format used to write this file.	50
8	MetadataStyle	The style in which the metadata was written, e.g., "PVL"."PVL" means < parameter >=< value >.	50
9	EndianType	The endian type of the system that wrote this file. Values are "BIG ENDIAN" and "LITTLE ENDIAN".	50

2.1.6. JAXAInfo

JAXAInfo contains metadata requested by JAXA. Used by DPR algorithms and GSMaP. Table 2.1-5 shows each metadata elements in JAXAInfo.

Table 2.1-5 JAXAInfo Elements

No	Element	Description	Data size (bytes)
1	GranuleFirstScanUTCDateTime	The date and time of first scan (incl. missing scan). The format is YYYY-MM-DDTHH:MM:SS.sssZ, where YYYY is 4-digit year, MM is month number, DD is day of month, T is "T", HH is hour, MM is minute, SS is second, sss is millisecond, and Z is "Z". All fields are zero-filled. The missing value is constructed by replacing all digits with 9, i.e., 9999-99-99T99:99:99.999Z.	50
2	GranuleLastScanUTCDateTime	Granule Last Scan UTC Date. Date is a 24 character string. The format is YYYY-MM-DDTHH:MM:SS.sssZ, where YYYY is 4-digit year, MM is month number, DD is day of month, T is ”T”, HH is hour, MM is minute, SS is second, sss is millisecond, and Z is ”Z”. All fields are zero-filled.	50
3	TotalQualityCode	The total quality of product is defined based on the quality of input data. Quality meaning are (a) GPM KuPR/KaPR TRMM PR L2 product Good: The total quality of input data (Ku/Ka/PR L1B) is Good. Fair: The GPM KuPR/KaPR L2 is not JMA's global weather forecast (FCST) or JMA's Global ANALsis model data (GANAL) but weather DB file. EG (Empty Granule): The total quality of input data (Ku/Ka/PR L1B) is EG (b) GPM DPR L2 product Good: The total quality of both Ku L2 and Ka L2 is Good. Fair: (i)The total quality of either Ku L2 or Ka L2 is EG (ii)The input data used in GPM DPR L2 is not JMA's global weather forecast (FCST) or JMA's Global ANALysis model data (GANAL) but weather DB file. EG (Empty Granule): The total quality of both Ku L2 and Ka L2 is EG. (c) GPM DPR SLH L2 product Good: The total quality of input data (DPR L2) is Good Fair: The total quality of input data is Fair. EG (Empty Granule): The total quality of input data (DPR L2) is EG.	50

No	Element	Description	Data size (bytes)
4	FirstScanLat	Latitude of orbit first scan.	50
5	FirstScanLon	Longitude of orbit first scan.	50
6	LastScanLat	Latitude of orbit last scan.	50
7	LastScanLon	Longitude of orbit last scan.	50
8	NumberOfRainPixelsFS	Number of rain pixels in the FS swath, judged at DPR L2 algorithm. At DPR L1, value is "-9999".	50
10	NumberOfRainPixelsHS	Number of rain pixels in the HS swath, judged at DPR L2algorithm. At DPR L1, value is "-9999".	50
11	ProcessingSubSystem	The name of the processing sub-system, e.g., "ALGORITHM","PCS".	50
12	ProcessingMode	The name of the processing mode, e.g., "STD","NRT".	50
13	Lightspeed	Constant value of light speed.	50
14	DielectricFactorKu	The dielectric factor $\|K^2\|$ at Ku.	50
15	DielectricFactorKa	The dielectric factor $\|K^2\|$ at Ka.	50

2.1.7. SwathHeader

SwathHeader contains metadata for swaths. This group appears in Level 1 and Level 2 data products. Table 2.1-6 shows each metadata elements in SwathHeader.

Table 2.1-6 SwathHeader Elements

No	Element	Description	Data size (bytes)
1	NumberScansInSet	The scans read by TKreadScan are a ”set”. For single swath data, one scan is read so NumberScansInSet=1. For multiple swath data, one TKreadScan may read more than one scan. For example, for SSM/I data one TKreadScan reads one low frequency scan and two high frequency scans. Therefore NumberScansInSet=1 for the low frequency swath and Number-ScansInSet=2 for the high frequency swath.	50
2	MaximumNumberScansTotal	The maximum allowed number of total scans in this swath. Total scans = overlap scans before granule + scans in granule + overlap scans after granule.	50
3	NumberScansBeforeGranule	The number of overlap scans before the first scan of the granule in this swath.	50
4	NumberScansGranule	The number of scans in the granule in this swath.	50
5	NumberScansAfterGranule	The number of overlap scans after the last scan of the granule in this swath.	50
6	NumberPixels	The number of IFOV in each scan in this swath.	50
7	ScanType	The type of scan in this swath. Values are: "CROSSTRACK" and "CONICAL".	50

2.2. Data Group

Elements of data group are explained in detail in this section. Each swath has 11 data group (12 data group for FS swath of 2ADPR) and 3 data (Latitude, Longitude and sunLocalTime) commonly. Figure 2.2-1 shows data group structure.

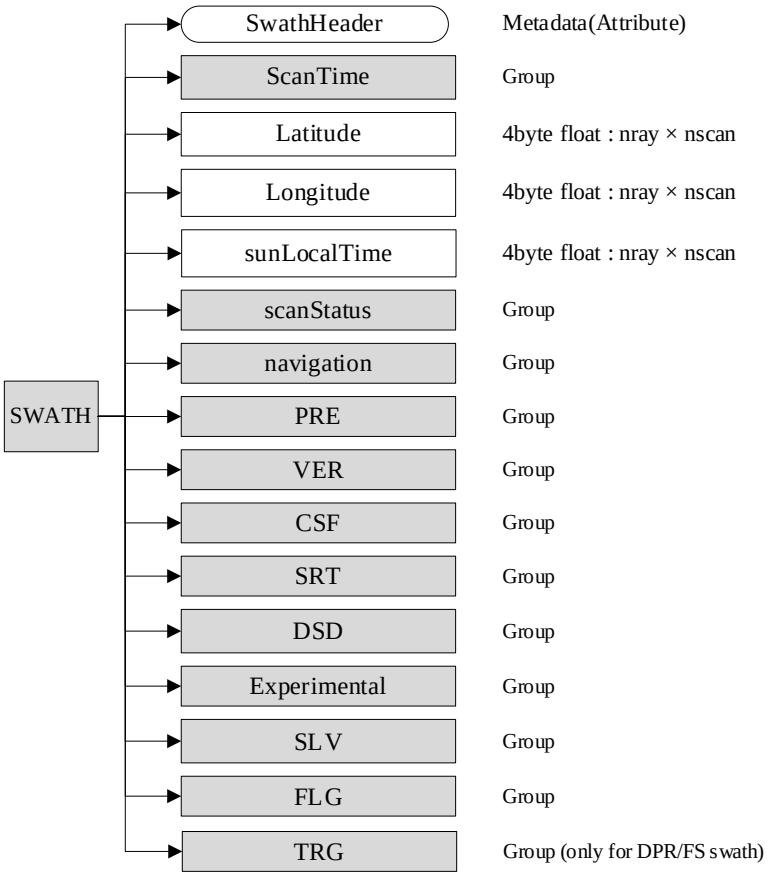


Figure 2.2-1 Data Format Structure for Data Group

2.2.1. ScanTime (Group)

(1) Year

Type	Array	Unit
2-byte integer	nscan	year

4-digit year, e.g., 1998. Values range from 1950 to 2100 years:

Missing Value: -9999

(2) Month

Type	Array	Unit
1-byte integer	nscan	month

Month of the year. Values range from 1 to 12 months.

Missing Value: -99

(3) DayOfMonth

Type	Array	Unit
1-byte integer	nscan	day

Day of the month. Values range from 1 to 31 days.

Missing Value: -99

(4) Hour

Type	Array	Unit
1-byte integer	nscan	hour

UTC hour of the day. Values range from 0 to 23 hours.

Missing Value: -99

(5) Minute

Type	Array	Unit
1-byte integer	nscan	minute

Minute of the hour. Values range from 0 to 59 minutes.

Missing Value: -99

(6) Second

Type	Array	Unit
1-byte integer	nscan	s

Second of the minute. Values range from 0 to 60 s.

Missing Value: -99

(7) MilliSecond

Type	Array	Unit
2-byte integer	nscan	ms

Thousandths of the second. Values range from 0 to 999 ms.

Missing Value: -9999

(8) DayOfYear

Type	Array	Unit
2-byte integer	nscan	day

Day of the year. Values range from 1 to 366 days.

Missing Value: -9999

(9) SecondOfDay

Type	Array	Unit
8-byte float	nscan	s

A time associated with the scan. It is expressed as the UTC seconds of the day.

Values range from 0 to 86400 s.

Missing Value: -9999.9

2.2.2. Latitude

(1) Latitude

Type	Array	Unit
4-byte float	nray * nscan	degree

The earth latitude of the center of the IFOV at the altitude of the earth ellipsoid. Latitude is positive north, negative south.

Values range from -90 to 90 degrees.

Missing Value: -9999.9

2.2.3. Longitude

(1) Longitude

Type	Array	Unit
4-byte float	nray * nscan	degree

The earth longitude of the center of the IFOV at the altitude of the earth ellipsoid. Longitude is positive east, negative west. A point on the 180th meridian has the value -180 degrees.

Values range from -180 to 180 degrees.

Missing Value: -9999.9

2.2.4. sunLocalTime

(1) sunLocalTime

Type	Array	Unit
4-byte float	nray * nscan	hour

The local hour angle of the Sun at the pixel location, where 0 is midnight and 12 is local noon when the Sun crosses the local meridian. Also known as apparent solar time at any location. In V07 TMI and GMI products will have values but partner products will be filled with the missing value.

Values range from 0 to 24 hours.

Missing Value: -9999.9

2.2.5. scanStatus (Group)

(1) dataQuality

Type	Array	Array (2ADPR FS)	Unit
1-byte integer	nscan	nfreq * nscan	N/A

A summary of data quality in the scan. Unless this is 0 (normal), the scan data is meaningless to higher precipitation processing. Bit 0 is the least significant bit (i.e., if bit i = 1 and other bits = 0, the unsigned integer value is 2**i).

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Bit Meaning

0: missing

5: geoError is not zero

6: modeStatus is not zero

(2) dataWarning

Type	Array	Array (2ADPR FS)	Unit
1-byte integer	nscan	nfreq * nscan	N/A

Flag of data warning for each scan. Bit Meaning is below.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Bit Meaning

0: beam Matching is abnormal

1: VPRF table is abnormal

2: surface Table is abnormal

3: geoWarning is not Zero

4: operational mode is not observation mode.

5: GPS status is abnormal

(3) missing

Type	Array	Array (2ADPR FS)	Unit
1-byte integer	nscan	nfreq * nscan	N/A

Indicates whether information is contained in the scan data. The values are as follows.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Bit Meaning

0: Scan is missing

1: Science telemetry packet missing

2: Science telemetry segment within packet missing

3: Science telemetry other missing

4: Housekeeping (HK) telemetry packet missing

5: Spare (always 0)

6: Spare (always 0)

7: Spare (always 0)

(4) modeStatus

Type	Array	Array (2ADPR FS)	Unit
1-byte integer	nscan	nfreq * nscan	N/A

A summary of status modes. If all status modes are routine, all bits in modeStatus = 0. Routine means that scan data has been measured in the normal operational situation as far as the status modes are concerned. modeStatus does not assess geolocation quality. modeStatus is broken into 8 bit flags. Each bit = 0 if the status is routine but the bit = 1 if the status is not routine. Bit 0 is the least significant bit (i.e., if bit i = 1 and other bits = 0, the unsigned integer value is 2**i). The non-routine situations follow:

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Bit Meaning
0: Spare (always 0)
1: SCorientation not 0 or 180
2: pointingStatus not 0
3: Non-routine limitErrorFlag
4: Non-routine operationalMode (not 1 or 11)
5: Spare (always 0)
6: Spare (always 0)
7: Spare (always 0)

(5) geoError

Type	Array	Array (2ADPR FS)	Unit
2-byte integer	nscan	nfreq * nscan	N/A

A summary of geolocation errors in the scan. geoError is used to set a bit in dataQuality. A zero integer value of geoError indicates 'good' geolocation. A non-zero value broken down into the bit flags below indicates the specified reason, where bit 0 is the least significant bit (i.e., if bit i = 1 and other bits = 0 the unsigned integer value is 2**i). Bits 0, 4, 5, 8 and 9 are per pixel error flags. If the number of bad pixels (for any of the reasons specified by these flags) is greater than the threshold then bit 7 = 1 and each of these flags is set to 1 if any pixel is bad for that reason. At launch this threshold is zero, so data is flagged if any pixel is bad. If the number of bad pixels is less than or equal to the threshold then bit 7 = 0 and all of these flags are also 0.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Bit Meaning

0: Latitude limit exceeded for viewed pixel locations

1: Negative scan time, invalid input

2: Error getting spacecraft attitude at scan mid-time

3: Error getting spacecraft ephemeris at scan mid-time

4: Invalid input non-unit ray vector for any pixel

5: Ray misses Earth for any pixel with normal pointing

6: Nadir calculation error for subsatellite position

7: Pixel count with geolocation error over threshold

8: Error in getting spacecraft attitude for any pixel

9: Error in getting spacecraft ephemeris for any pixel

10: Spare (always 0)

11: Spare (always 0)

12: Spare (always 0)

13: Spare (always 0)

14: Spare (always 0)

15: Spare (always 0)

(6) geoWarning

Type	Array	Array (2ADPR FS)	Unit
2-byte integer	nscan	nfreq * nscan	N/A

A summary of geolocation warnings in the scan. geoWarning does not set a bit in dataQuality. Warnings indicate unusual conditions. These conditions do not indicate bad geolocation but are flagged as a warning that further review of the data may be useful. A zero integer value indicates usual geolocation. A non-zero value broken down into the following bit flags indicates the following, where bit 0 is the least significant bit (i.e., if bit i = 1 and other bits = 0 the unsigned integer value is 2**i):

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Bit Meaning

0: Ephemeris Gap Interpolated

1: Attitude Gap Interpolated

2: Attitude jump/discontinuity

3: Attitude out of range

4: Anomalous Time Step

5: GHA not calculated due to error

6: SunData (Group) not calculated due to error

7: Failure to calculate Sun in inertial coordinates

8: Fallback to GES ephemeris

9: Fallback to GEONS ephemeris

10: Fallback to PVT ephemeris

11: Fallback to OBP ephemeris

12: Spare (always 0)

13: Spare (always 0)

14: Spare (always 0)

15: Spare (always 0)

(7) SCorientation

Type	Array	Unit
2-byte integer	nscan	degree

The positive angle of the spacecraft vector (v) from the satellite forward direction of motion, measured clockwise facing down. We define v in the same direction as the spacecraft axis +X, which is also the center of the GMI scan. If SCorientation is not 0 or 180, a bit is set to 1 in modeStatus.

Value Meaning
0: +X forward (yaw 0)
180: -X forward (yaw 180)
-8000: Non-nominal pointing
-9999: Missing

(8) pointingStatus

Type	Array	Array (2ADPR FS)	Unit
2-byte integer	nscan	nfreq * nscan	N/A

It is provided by the GeoTK. A value of zero means the pointing is good. Non-zero values indicate non-nominal pointing. If pointingStatus is non-zero, a bit in modeStatus is set to 1.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Value Meaning
0: Nominal pointing in Mission Science Mode
1: GPS point solution stale and PVT ephemeris used
2: GEONS solution stale and GEONS ephemeris used
-8000: Non-nominal mission science orientation
-9999: Missing

(9) acsModeMidScan

Type	Array	Unit
1-byte integer	nscan	N/A

It is provided by the GeoTK as taken from Attitude Control System telemetry and is provided in this format for information only.

Value Meaning
0: LAUNCH
1: RATENULL
2: SUNPOINT
3: GSPM (Gyro-less Sun Point)
4: MSM (Mission Science Mode)
5: SLEW
6: DELTAH
7: DELTAV

(10) targetSelectionMidScan

Type	Array	Unit
1-byte integer	nscan	N/A

It is provided by the GeoTK as taken from Attitude Control System telemetry and is provided in this format for information only.

Value Meaning
0: S/C Z axis nadir, +X in flight direction
1: Flight Z axis nadir, +X in flight direction
2: S/C Z axis nadir, -X in flight direction
3: Flight Z axis nadir, -X in flight direction
4: +90 yaw for DPR antenna pattern calibration
5: -90 yaw for DPR antenna pattern calibration
-99: Missing

(11) operationalMode

Type	Array	Array (2ADPR FS)	Unit
1-byte integer	nscan	nfreq * nscan	N/A

The operational mode of KuPR/KaPR/PR stored in science telemetry basically. However, if science telemetry is not made like as stand-by mode, KuPR/KaPR L1B algorithm decides it using HK telemetry.PR L1B algorithm stored missing value. The values range is 1 to 20 for KuPR/KaPR. The values range is 1 to 3, 5, 6, 10, -99. The value meaning is shown below.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Value Meaning
1: Ku/Ka/PR Observation
2: Ku/Ka/PR External Calibration
3: Ku/Ka/PR Internal Calibration
4: Ku/Ka SSPA Analysis
5: Ku/Ka/PR LNA Analysis
6: Ku/Ka/PR Health-Check
7: Ku/Ka Standby VPRF Table OUT
8: Ku/Ka Standby Phase Out
9: Ku/Ka Standby Dump Out
10: Ku/Ka Standby (No Science Data)
11: Ku/Ka/PR Independent Observation
12: Ku/Ka Independent External Calibration
13: Ku/Ka Independent Internal Calibration
14: Ku/Ka Independent SSPA Analysis
15: Ku/Ka Independent LNA Analysis
16: Ku/Ka Independent Health-Check
17: Ku/Ka Independent Standby VPRF Table OUT
18: Ku/Ka Independent Standby Phase Out
19: Ku/Ka Independent Standby Dump Out
20: Ku/Ka Independent Standby (No Science Data)
-99: PR missing value (No Science Data)

(12) limitErrorFlag

Type	Array	Array (2ADPR FS)	Unit
1-byte integer	nscan	nfreq * nscan	N/A

It has 2 error information. One is as for noise power limit, another one is as for binEllipsoid limit. The former is defined that if there are more than 2 overlimited rays in a swath, limitErrorFlag (at 0bit) is adapted. On the other hand, the latter is defined that if there is even an overlimited ray, limitErrorFlag (at 1bit) is adapted. Then, limitErrorFlag is used in modeStatus, dataQuality in scanStatus Group picks it up consequently.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

The values are

Bit Meaning
0: noise power limit error
1: binEllipsoid is missing
2: Spare (always 0)
3: Spare (always 0)
4: Spare (always 0)
5: Spare (always 0)
6: Spare (always 0)
7: Spare (always 0)

(13) FractionalGranuleNumber

Type	Array	Unit
8-byte float	nscan	N/A

The floating-point granule numbers. The granule begins at the Southern-most point of the spacecraft's trajectory. For example, FractionalGranuleNumber = 10.5 means the spacecraft is halfway through granule 10 and starting the descending half of the granule. Values range from 0 to 100000. In Near Real Time (NRT) process, granule number is stored only '0', so Fractional Granule Number less than 1.0.

Missing value: -9999.9

2.2.6. navigation (Group)

(1) scPos

Type	Array	Unit
4-byte float	XYZ * nscan	m

For 2AKu, 2AKa and 2ADPR, the position vector (m) of the spacecraft in Earth-Centered Earth Fixed (ECEF) Coordinates at the Scan mid-Time (i.e., time at the middle pixel/IFOV of the active scan period). For 2APR, the position vector (m) of the spacecraft in True of Date (TOD) Earth-Centered Inertial (ECI) Coordinates at the Scan mid-Time. Values range from -10000000 to 10000000 m.

Missing value: -9999.9

(2) scVel

Type	Array	Unit
4-byte float	XYZ * nscan	m/s

For 2AKu, 2AKa and 2ADPR, the velocity vector (m/s) of the spacecraft in ECEF Coordinates at the Scan mid-Time. For 2APR, the velocity vector (m/s) of the spacecraft in TOD ECI Coordinates at the Scan mid-Time. Values range from -10000000 to 10000000 m/s.

Missing value: -9999.9

(3) scLat

Type	Array	Unit
4-byte float	nscan	degrees

The geodesic latitude (decimal degrees) of the spacecraft at the Scan mid-Time. Values range from -70 to 70 degrees.

Missing value: -9999.9

(4) scLon

Type	Array	Unit
4-byte float	nscan	degrees

The geodesic longitude (decimal degrees) of the spacecraft at the Scan mid-Time. Values range from -180 to 180 degrees.

Missing value: -9999.9

(5) scAlt

Type	Array	Unit
4-byte float	nscan	m

The altitude (m) of the real spacecraft above the Earth Ellipsoid at the Scan mid-Time. It is computed by GeoTK. Values range from 350000 to 500000 m. Note that originally scAlt has been almost the same value as dprAlt, but after the GPM altitude change operated in November 2023, only scAlt represents the altitude of the satellite.

Missing value: -9999.9

(6) dprAlt

Type	Array	Unit
4-byte float	nscan	m

The altitude (m) of the spacecraft above the Earth Ellipsoid at the Scan mid-Time from DPR science telemetry for DPR operation. This is empty in non-DPR products. It is stored 'GPS Altitude Data' with LSB equal to 10m in DPR science telemetry. Values range from 350000 to 500000 m.

After the change of satellite altitude operated in November 2023 (GPM Orbit Boost), dprAlt is stored as scAlt minus 35 km. The following equation shows the relation between dprAlt and scAlt.

$$dprAlt = scAlt - 35000$$

Missing value:

-9999.9: at missing scan and internal calibration mode.

(7) scAttRollGeoc

Type	Array	Unit
4-byte float	nscan	degrees

The geocentric satellite attitude Euler roll angle (degrees) at the Scan mid-Time. The order of the components in the file is roll, pitch, and yaw. However, the angles are computed using a 3-2-1 Euler rotation sequence representing the rotation order yaw, pitch, and roll for the rotation from Orbital Coordinates to the spacecraft body coordinates. Orbital Coordinates represent an orthogonal triad in Geocentric Inertial Coordinates where the Z-axis is toward the geocentric nadir, the Y-axis is perpendicular to the spacecraft velocity opposite the orbit normal direction, and the X-axis is approximately in the velocity direction for a near circular orbit. Note this is geocentric, not geodetic, referenced, so that pitch and roll will have twice orbital frequency components due to the onboard control system following the oblate geodetic Earth horizon. Note also that the yaw value will show an orbital frequency component relative to the Earth fixed ground track due to the Earth rotation relative to inertial coordinates. Values range from -180 to 180 degrees.

Missing value: -9999.9

(8) scAttPitchGeoc

Type	Array	Unit
4-byte float	nscan	degrees

The geocentric satellite attitude Euler pitch angle (degrees) at the Scan mid-Time. Values range from -180 to 180 degrees.

Missing value: -9999.9

(9) scAttYawGeoc

Type	Array	Unit
4-byte float	nscan	degrees

The geocentric satellite attitude Euler yaw angle (degrees) at the Scan mid-Time. Values range from -135 to 225 degrees.

Missing value: -9999.9

(10) scAttRollGeod

Type	Array	Unit
4-byte float	nscan	degrees

The geodetic satellite attitude Euler roll angle (degrees) at the Scan mid-Time. The order of the components in the file is roll, pitch, and yaw. However, the angles are computed using a 3-2-1 Euler rotation sequence representing the rotation order yaw, pitch, and roll for the rotation from Geodetic Coordinates to the spacecraft body coordinates. Geodetic Coordinates represent an orthogonal triad in Geocentric Inertial Coordinates where the Z-axis is toward the geodetic nadir, the Y-axis is perpendicular to the spacecraft velocity opposite the orbit normal direction, and the X-axis is approximately in the velocity direction for a near circular orbit. Values range from -180 to 180 degrees.

Missing value: -9999.9

(11) scAttPitchGeod

Type	Array	Unit
4-byte float	nscan	degrees

The geodetic satellite attitude Euler pitch angle (degrees) at the Scan mid-Time. Values range from -180 to 180 degrees.

Missing value: -9999.9

(12) scAttYawGeod

Type	Array	Unit
4-byte float	nscan	degrees

The geodesic satellite attitude Euler yaw angle (degrees) at the Scan mid-Time. Values range from -135 to 225 degrees.

Missing value: -9999.9

(13) scHeadingGround

Type	Array	Unit
4-byte float	nscan	degrees

The spacecraft ground track heading measured about the geodetic nadir with respect to North at the scan mid-Time. This is the apparent direction of spacecraft motion over the Earth's surface, accounting for Earth rotation effects.

Values range from -180 to 180 degrees.

Missing value: -9999.9

(14) scHeadingOrbital

Type	Array	Unit
4-byte float	nscan	degrees

The spacecraft orbital reference heading measured about the geodetic nadir with respect to North at the subsatellite point at the scan mid-Time. This is the apparent direction of the inertial velocity and the zero-yaw angle reference direction for spacecraft control.

Values range from -180 to 180 degrees.

Missing value: -9999.9

(15) greenHourAng

Type	Array	Unit
4-byte float	nscan	degrees

The rotation angle (degrees) from Geocentric Inertial Coordinates to Earth Fixed Coordinates. Values range from 0 to 390 degrees.

Missing value: -9999.9

(16) timeMidScan

Type	Array	Unit
8-byte float	nscan	s

The Scan mid-Time in GPS Atomic time, namely the seconds since 0000 UTC,6 Jan 1980.timeMidScan is used as the reference time for the scPos and scVel values. Values range from 0 to 10000000000 s.

Missing value: -9999.9

(17) timeMidScanOffset

Type	Array	Unit
8-byte float	nscan	s

Offset from the secondary header packet time to the timeMidScan. Values range from 0 to 100 s.

Missing value: -9999.9

2.2.7. PRE (Group)

(1) elevation

Type	Array	Unit
4-byte float	nray * nscan	m

Elevation of the measurement point. It is a copy of DEMHmean of level 1B product.

In the 2ADPR, it is estimated by dual-frequency algorithm.

Missing Value: -9999.9

(2) landSurfaceType

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Land surface type.

In the 2ADPR, it is estimated by dual-frequency algorithm.

The values are

Value Meaning
0-99: Ocean
100 - 199: Land
200 - 299: Coast
300 - 399: Inland water
-9999: Missing

(3) localZenithAngle

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	degree

Local zenith angle of each ray. It is a copy of scLocalZenith of level 1B product.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(4) flagPrecip

Type	Array	Unit
4-byte integer	nray * nscan	N/A

The values estimated by single frequency algorithm are

Value Meaning

0: No precipitation (including the case of missing)

1: Precipitation (1-D judgement)

2: Precipitation (3-D judgement)

To know the detailed data quality of the pixel, it is helpful to read /FLG/qualityData.

In the 2ADPR, the values are defined as:

$$\text{flagPrecip}_{\text{DPR}} = 10\text{flagPrecip}_{\text{Ku}} + \text{flagPrecip}_{\text{Ka}}$$

Values Meaning:

$\text{flagPrecip}_{\text{DPR}}$	$\text{flagPrecip}_{\text{Ku}}$	$\text{flagPrecip}_{\text{Ka}}$
0	0 (No precipitation)	0 (No precipitation)
1	0 (No precipitation)	1 (1-D judgement)
2	0 (No precipitation)	2 (3-D judgement)
10	1 (1-D judgement)	0 (No precipitation)
11	1 (1-D judgement)	1 (1-D judgement)
12	1 (1-D judgement)	2 (3-D judgement)
20	2 (3-D judgement)	0 (No precipitation)
21	2 (3-D judgement)	1 (1-D judgement)
22	2 (3-D judgement)	2 (3-D judgement)

(5) binRealSurface

Type	Array	Array (2ADPR FS)	Unit
2-byte integer	nray * nscan	nfreq * nray * nscan	

Range bin number for real surface.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by dual-frequency algorithm.

Missing Value: -9999

(6) binStormTop

Type	Array	Unit
2-byte integer	nray * nscan	

Range bin number for the storm top.

In the 2ADPR, it is estimated by dual-frequency algorithm.

Missing Value: -9999

(7) binMirrorImageL2

Type	Array	Unit
2-byte integer	nray * nscan	

The lowest range bin number where a mirror image echo may appear. For FS swaths, the bin numbers are 1-based ranging from 1 at the top of the data window with 176 at the Ellipsoid. For HS swaths, the bin number are 1-based ranging from 1 at the top of the data window with 88 at the Ellipsoid.

Missing Value: -9999

(8) height

Type	Array	Unit
4-byte float	nbin * nray * nscan	m

Height of each received echo.

In the 2ADPR, it is estimated by KuPR single-frequency algorithm.

Missing Value: -9999.9

(9) heightStormTop

Type	Array	Unit
4-byte float	nray * nscan	m

Height of storm top.

In the 2ADPR, it is estimated by dual-frequency algorithm.

Missing Value: -9999.9

(10) binClutterFreeBottom

Type	Array	Unit
2-byte integer	nray * nscan	

Range bin number for clutter free bottom.

In the 2ADPR, it is estimated by KuPR single-frequency algorithm.

Missing Value: -9999

(11) sigmaZeroMeasured

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dB

Surface backscattering cross section without attenuation correction (as measured).

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Ka-band sigmaZeroMeasured estimated by dual-frequency algorithm.

Missing Value: -9999.9

(12) zFactorMeasured

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nbin *nray * nscan	nfreq * nbin * nray * nscan	dBZ

Vertical profile of reflectivity factor (Z) without attenuation correction (as measured).

10log10(Z) where Z is in mm^6/m^3.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(13) ellipsoidBinOffset

Type	Array	Unit
4-byte float	nray * nscan	m

Distance between the ellipsoid and a center range bin of binEllipsoid defined by level 1B algorithm.

$$\text{ellipsoidBinOffset} = \text{scRangeEllipsoid} - (\text{startBinRange} + (\text{binEllipsoid} - 1) \times \text{rangeBinSize})$$

scRangeEllipsoid: Distance between a sensor and the ellipsoid [m]

startBinRange: Distance between a sensor and a center of the highest observed range bin [m]

binEllipsoid: Range bin number of the Ellipsoid (1 - 260)

rangeBinSize: Range bin size [m]

Missing Value: -9999.9

(14) echoCountRealSurface

Type	Array	Array (2ADPR FS)	Unit
1-byte unsigned integer	nray * nscan	nfreq * nray * nscan	N/A

Echo count at a surface position (binRealSurface).

In 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: 0

(15) snRatioAtRealSurface

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	N/A

Signal/Noise ratio at real surface range bin.

$$\text{snRatioAtRealSurface} = 10 \times \log_{10}(\text{echoPower [mW]}/\text{noisePower [mW]})$$

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999

(16) adjustFactor

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dB

Adjustment factor (dB) for zFactorMeasured (dBZm') and sigmaZeroMeasured (dBs0m'). dBZm' and dBs0m' are used and stored as follows:

$$\mathrm{dBZm}' = \mathrm{dBZm} - \mathrm{adjustFactor}$$

$$\mathrm{dBs0m}' = \mathrm{dBs0m} - \mathrm{adjustFactor}$$

The adjustment factor is the sum of 3 components:

base adjustment for instrument dependency,

angle-bin adjustment for angle-bin dependency, and

temporal adjustment for orbit number dependency.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value:

-9999.9

(17) snowIceCover

Type	Array	Unit
1-byte integer	nray * nscan	N/A

Snow and ice cover information. It refers to ancillary data (multisensor snow/ice cover maps) provided by NOAA.

In the 2ADPR, it is estimated by KuPR single-frequency algorithm.

In the 2APR, 0 is set everywhere since the current TRMM/PR algorithm does not input the ancillary data provided by the NOAA.

The values are

Value Meaning
0: Open water
1: Land, no snow
2: Snow cover on land
3: Ice on water
-99: Missing

(18) flagSigmaZeroSaturation

Type	Array	Array (2ADPR FS)	Unit
1-byte unsigned integer	nray * nscan	nfreq * nray * nscan	N/A

A flag to show whether echoPower is under a saturated level or not at a range bin with a calculation of sigmaZeroMeasured.

The values are as follows:

Value Meaning
0: Normal (under saturated level)
1: Possible saturated level at real surface
2: Saturated level at real surface
99: Missing

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

2.2.8. VER (Group)

(1) binZeroDeg

Type	Array	Unit
2-byte integer	nray * nscan	

Range bin number with 0 degrees C level.

Missing Value: -9999

(2) binZeroDegSecondary

Type	Array	Unit
2-byte integer	nray * nscan	

When the inversion layers are detected, the “binZeroDegSecondary ”is used to output the binZeroDeg related to the inversion layer. A peak level with the ground surface of T < 0 deg.C. (“binZeroDeg” is 177). A missing value when there are no inversion layers and T of the ground surface of T > 0 deg.C.

Missing Value: -9999

(3) airTemperature

Type	Array	Unit
4-byte float	nbin *nray * nscan	K

Air Temperature.

Missing Value: -9999.9

(4) attenuationNP

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nbin *nray * nscan	nfreq * nbin *nray * nscan	dB/km

Vertical profile of attenuation by non-precipitation particles (cloud liquid water, cloud ice water, water vapor, and oxygen molecules).

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(5) piaNP

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nNP * nray * nscan	nfreq * nNP * nray * nscan	dB

Path integrated attenuation caused by non-precipitation particles (cloud liquid water, cloud ice water, water vapor, and oxygen molecules).

nNP (1): Total (sum of 2, 3, and 4)

nNP (2): Water Vapor

nNP (3): Oxygen molecules

nNP (4): Cloud liquid water

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(6) piaNPrainFree

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nNP * nray * nscan	nfreq * nNP * nray * nscan	dB

“Rain-free” path-integrated attenuation due to non-precipitation (piaNP).

nNP (1): Total (sum of 2, 3, and 4)

nNP (2): Water Vapor

nNP (3): Oxygen molecules

nNP (4): Cloud liquid water

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(7) sigmaZeroNPCorrected

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dB

Surface backscattering cross section with attenuation correction only for non-precipitation particles.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(8) heightZeroDeg

Type	Array	Unit
4-byte float	nray * nscan	m

Height of freezing level (0 degrees C level) Values are in m.

Missing Value: -9999.9

(9) flagInversion

Type	Array	Unit
2-byte integer	nray * nscan	N/A

Flag of inversion layers of air temperature related to 0 degrees C level.

-1: The surface ground below 0 deg.C

0: The VER code detect a level of 0 deg.C without the inversion layers.

>=1: The VER code detect a level of 0 deg.C with the inversion layers.

Missing Value: -9999

2.2.9. CSF (Group)

In some variables in the CSF module, no-rain values are used for the missing. To know the detailed data quality of the pixel, it is helpful to read /FLG/qualityData.

(1) flagBB

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Bright band (BB) exists or not.

In case of 2AKu and 2AKa,

The values are

Value Meaning
0: BB not detected
1: BB detected
-1111: No rain value (including the case of missing)
-9999: Missing

In case of 2ADPR,

The values are

Value Meaning
0: BB not detected
≥ 1: BB detected
1: BB detected by both single-frequency (Ku-band) and dual-frequency algorithm.
2: BB detected by single-frequency (Ku-band) algorithm only.
3: BB detected by dual-frequency algorithm only.
-1111: No rain value (including the case of missing)
-9999: Missing

(2) binBBPeak

Type	Array	Unit
2-byte integer	nray * nscan	

Range bin number for the peak of bright band.

Missing Value: -9999

(3) binBBTop

Type	Array	Unit
2-byte integer	nray * nscan	

Range bin number for the top of bright band.

Missing Value: -9999

(4) binBBBottom

Type	Array	Unit
2-byte integer	nray * nscan	

Range bin number for the bottom of bright band.

The values are

Value Meaning
0: BB not detected
-1111: No rain value (including the case of missing)
-9999: Missing

(5) heightBB

Type	Array	Unit
4-byte float	nray * nscan	m

Height of bright band.

The values are

Value Meaning
0.0: BB not detected
-1111.1: No rain value (including the case of missing)
-9999.9: Missing

(6) widthBB

Type	Array	Unit
4-byte float	nray * nscan	m

The width of bright band.

The values are

Value Meaning
0.0: BB not detected
-1111.1: No rain value (including the case of missing)
-9999.9: Missing

(7) qualityBB

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Quality of the bright band.

The values are

Value Meaning
1: Good
0: BB not detected in the case of rain
-1111: No rain value (including the case of missing)
-9999: Missing

(8) typePrecip

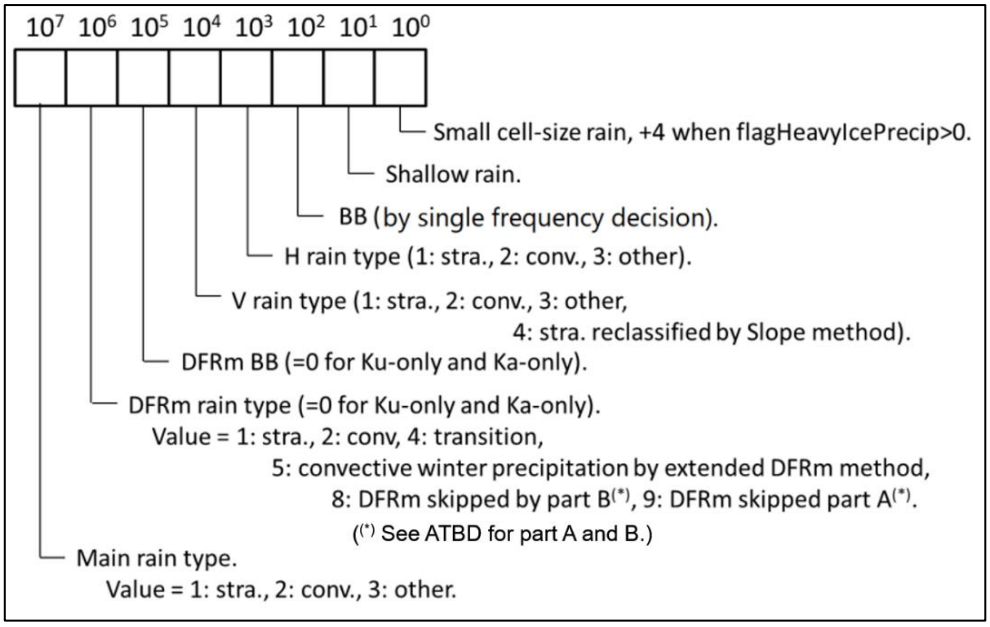
Type	Array	Unit
4-byte integer	nray * nscan	N/A

When positive, typePrecip shows precipitation type by an 8-digit number as shown in the next box. The negative typePrecip means as follows:

-1111: No rain,

-9999: Missing.

Details of 8-digit typePrecip:



The three major rain categories, stratiform, convective, and other, can be obtained from typePrecip as follows:

When typePrecip is greater than zero,

Main rain type = typePrecip/10000000

1: stratiform

2: convective

3: other

In the DPR product, rain type by the CSU's DFRm (measured dual frequency ratio) method is also included in typePrecip and can be obtained as follows:

DFRm rain type = (typePrecip%10000000)/1000000 in C

DFRm rain type = (MOD(typePrecip,10000000)/1000000 in FORTRAN

DFRm rain type

1: stratiform

2: convective

4: transition

5: Decided winter precipitation as convective by extended DFRm method.

8: DFRm method cannot be applicable at Part B (in this case the conventional method determines the major rain type)

9: DFRm method cannot be applicable at Part A (in this case the conventional method determines the major rain type)

-1111: No rain value (including the case of missing)

-9999: Missing

(9) qualityTypePrecip

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Quality of the precipitation type.

The values are

Value Meaning

1: Good

-1111: No rain value (including the case of missing)

-9999: Missing

(10) flagShallowRain

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Type of shallow rain.

The values are

Value Meaning

0: No shallow rain

10: Shallow isolated (maybe)

11: Shallow isolated (certain)

20: Shallow non-isolated (maybe)

21: Shallow non-isolated (certain)

-1111: No rain value (including the case of missing)

-9999: Missing

(11) binDFRmMLTop (2ADPR FS, 2ADPR HS)

Type	Array	Unit
2-byte integer	nray * nscan	N/A

The DFRm method detects melting layer (ML) the meaning of which is wider than that of BB. Since ML and BB are different, new output item binDFRmMLBottom and binDFRmMLTop are added to FS and HS data.

Range bin number for ML top detected by the DFRm method.

The values are

Value: Meaning
> 0: Range bin number when ML top is detected
0: ML top is not detected
-1111: Value for no rain in FS (HS) mode at Ka band (including the case of missing)
-9999: Missing

(12) binDFRmMLTBottom (2ADPR FS, 2ADPR HS)

Type	Array	Unit
2-byte integer	nray * nscan	

The DFRm method detects melting layer (ML) the meaning of which is wider than that of BB. Since ML and BB are different, new output item binDFRmMLBottom and binDFRmMLTop are added to FS and HS data.

Range bin number for ML bottom detected by the DFRm method.

The values are

Value: Meaning
> 0: Range bin number when ML bottom is detected
0: ML bottom is not detected
-1111: Value for no rain in MS (HS) mode at Ka band (including the case of missing)
-9999: Missing

(13) binHeavyIcePrecipTop (except for 2ADPR HS)

Type	Array	Array (2ADPR FS)	Unit
2-byte integer	nray * nscan	nfreqHI * nray * nscan	N/A

Range bin number for the top of heavy ice precip. For FS swaths, bin numbers are 1-based ranging from 1 at the top of the data window with 176 at the Ellipsoid. For HS swaths, bin numbers are 1-based ranging from 1 at the top of the data window with 88 at the Ellipsoid.

Value: Meaning
(A) In the single frequency 2AKu,
>0: Range bin corresponding to the Ku-band top height for HIP when it is detected
0: When HIP is not detected
(B) In the single frequency 2AKa,
>0: Range bin corresponding to the Ka-band top height for HIP when it is detected
0: When HIP is not detected
(C) In the 2ADPR,
nfreqHI (1): Same to the above (A)
nfreqHI (2): Same to the above (B)
nfreqHI (3):
>0: Range bin corresponding to the top height for HIP detected by the dual frequency method
0: When HIP is not detected by the dual frequency method

-1111: no precipitation is present.
-9999: Missing

(14) binHeavyIcePrecipBottom (except for 2ADPR HS)

Type	Array	Array (2ADPR FS)	Unit
2-byte integer	nray * nscan	nfreqHI * nray * nscan	N/A

Range bin number for the bottom of heavy ice precip. For FS swaths, bin numbers are 1-based ranging from 1 at the top of the data window with 176 at the Ellipsoid. For HS swaths, bin numbers are 1-based ranging from 1 at the top of the data window with 88 at the Ellipsoid.

Value: Meaning
(A) In the single frequency 2AKu,
>0: Range bin corresponding to the Ku-band bottom height for HIP when it is detected
0: When HIP is not detected
(B) In the single frequency 2AKa,
>0: Range bin corresponding to the Ka-band bottom height for HIP when it is detected
0: When HIP is not detected
(C) In the 2ADPR,
nfreqHI (1): Same to the above (A)
nfreqHI (2): Same to the above (B)
nfreqHI (3):
>0: Range bin corresponding to the bottom height for HIP detected by the dual frequency method
0: When HIP is not detected by the dual frequency method
-1111: no precipitation is present.
-9999: Missing

(15) nHeavyIcePrecip (except for 2ADPR HS)

Type	Array	Array (2ADPR FS)	Unit
1-byte unsigned integer	nray * nscan	nfreqHI * nray * nscan	N/A

Range bin number for the bottom of heavy ice precip.

Value: Meaning

(A) In the single frequency 2AKu,

>0: The total number of bins where Ku-band HIP is detected

0: When Ku-band HIP is not detected or No rain or Missing

(B) In the single frequency 2AKa,

>0: The total number of bins where Ka-band HIP is detected

0: When Ka-band HIP is not detected or No rain or Missing (in FS and HS before the antenna scan pattern change of the V07 data)

255: Missing (in HS of the V07 data after the antenna scan pattern change)

(C) In the 2ADPR,

nfreqHI (1): Same to the above (A)

nfreqHI (2): Same to the above (B)

nfreqHI (3):

>0: The total number of bins where HIP is detected by the dual frequency method

0: When HIP is not detected by the dual frequency method or No rain or Missing

(16) flagHeavyIcePrecip

Type	Array	Unit
1-byte signed integer	nray * nscan	N/A

This flag denotes detection of solid ice hydrometeors which cause severely strong Z factor or huge DFRm in the sky less than -10 degrees C temperature.

Value Meaning

(A) The case of Ka band FS (and HS before the antenna scan pattern change):

1 (=0x01): 35 dBZ >= Zm (Ka) > 30 dBZ

2 (=0x02): 40 dBZ >= Zm (Ka) > 35dBZ

3 (=0x03): Zm (Ka) > 40 dBZ

(B) The case of Ku band FS:

4 (=0x04): 40 dBZ >= Zm (Ku) > 35 dBZ

8 (=0x08): 45 dBZ >= Zm (Ku) > 40 dBZ

12 (=0x0c): Zm (Ku) > 45 dBZ

(C) The case of DPR FS:

If Zm (Ku) > 27 dBZ and DFRm > 7 dB, the following value is added in addition to (A) and (B).

16 (=0x10)

(However, before the antenna scan pattern change, this rule applies to inner swath only; outer swaths are same as (B).)

0: Not detected（including the case of missing）

Negative value: Missing in 2ADPR HS of V07 data after the antenna scan pattern change

(17) flagAnvil (2AKu FS, 2ADPR FS)

Type	Array	Unit
1-byte signed integer	nray * nscan	N/A

flagAnvil is positive values when anvil precipitation is detected by the Ku-band radar. 0 when anvil precipitation is not detected.

The values are

Value Meaning
(A) The case of 2AKu FS:
1: Type 1 (without rain downward)
2: Type 2 (with rain downward)
0: Not detected (which includes the case of missing)
(B) The case of 2ADPR FS:
1: Type 1 (without rain downward)
2: Type 2 (with rain downward)
0: Not detected
-99: Missing

(18) flagHail (2ADPR FS)

Type	Array	Unit
1-byte signed integer	nray * nscan	N/A

Value Meaning
0: Hail not detected (including the case of missing)
1: Hail detected

(19) flagMLquality (2ADPR FS, 2ADPR HS)

Type	Array	Unit
1-byte unsigned integer	nray * nscan	N/A

This flag indicates the quality of detected ML by the following values.

Value Meaning
0: ML is not detected (including the case of missing).
1: ML top and bottom is detected by the standard DFRm method.
2: ML top and bottom is detected by an extension, but the result is not used in the V07 rain type decision.
255: Missing in 2ADPR HS of V07 data after the antenna scan pattern change.

2.2.10. SRT (Group)

(1) PIAalt

Type	Array	Array (2ADPR FS)	Unit
4-byte float	method * nray * nscan	nfreq * method * nray * nscan	dB

The path-integrated attenuation (dB) from the jth estimate, (PIAj in the notation above), where

PIAalt (j=1) = PIA derived from the forward along-track spatial reference data

PIAalt (j=2) = PIA derived from the backward along-track spatial reference data

PIAalt (j=3) = PIA derived from the forward hybrid/cross-track reference data

PIAalt (j=4) = PIA derived from the backward hybrid/cross-track reference data

PIAalt (j=5) = PIA derived from standard temporal reference data

PIAalt (j=6) = PIA derived from the light-rain temporal reference data

Note that for product versions 1 through 3, the standard temporal path-attenuation estimate, PIAalt (5), is set to missing but is defined for versions 4 and higher. For product versions 1 through 6, the light-rain temporal estimate, PIAalt (6), is set to missing. Note also that the forward/backward hybrid/cross-track path attenuations are defined only over ocean and are set to missing over land.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(2) PIAdw(2ADPR FS, 2ADPR HS)

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dB

PIAdw (dB) is the path attenuation estimate derived from the standard dual-wavelength method.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(3) PIAhb

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dB

PIAhb (dB) is the path attenuation estimate derived from the Hitschfeld-Bordan equation.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(4) PIAhybrid

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dB

For the dual-frequency output, PIAhybrid (dB) is a weighted sum of the path attenuations from the SRT, HB, and DW methods. For the single-frequency outputs, PIAhybrid is a weighted sum of the SRT and HB methods.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(5) RFactorAlt

Type	Array	Unit
4-byte float	method * nray * nscan	N/A

The reliability factors associated with the individual PIA estimates in PIAalt. As with PIAalt (6), RFactorAlt (6) is set to missing.

$$RFactorAlt_j = \mathrm{Rel}_j = PIA_j/\sigma_j; j = 1,\ldots,6$$

method (j): The reliability factor associated with PIAalt (j): j=1…,6

Missing Value: -9999.9

(6) PIAweight

Type	Array	Unit
4-byte float	method * nray * nscan	dB

The weights, w, of the individual PIA estimates used in deriving the effective PIA. The weight for a particular PIA estimate is proportional to the inverse of the error variance associated with the method. The sum of the weights should equal one. As with PIAalt (6), PIAweight (6) is set to missing.

$$w_j = \frac{1}{\sigma_j^2}\frac{1}{\Sigma\frac{1}{\sigma_j^2}} \equiv \frac{u_j}{\Sigma u_j}$$

where

$$u_j = 1/\sigma_j^2$$

$$\sum w_j = 1$$

method (j): The weights of the PIAalt (j): j=1…, 6

Missing Value:

-9999.9

(7) PIAweightHY

Type	Array	Unit
4-byte float	method * nray * nscan	dB

The weights of the individual PIA Ku estimates used in deriving the effective path attenuation estimate, pathAtten. The sum of the weights should equal one. Where j is method and sigma j is the standard deviation of reference data for method j.

PIAweight_j = 1/sigma j^2 (1/Sum_j(1/$sigma\_j^2$)).

method (1): SRT method

method (2): HB method

method (3): DF method (2ADPR FS only)

Missing Value: -9999.9

(8) pathAtten

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dB

The estimated effective 2-way path-attenuation in (dB) where

$$\text{pathAtten} = 2\int_0^{r_s} k(s)ds$$

where k(s) is the attenuation coefficient in dB/km where the integral is taken from the storm top to the surface. The path attenuation is often designated as the PIA, the path-integrated attenuation. In the notation used above and in ATBD:

$$pathAtten = PIA_{eff} = (\Sigma u_j)^{-1}\Sigma u_j PIA_j$$

Where u_j is equal to the inverse of the variance associated with the jth reference data point:

$$u_j = 1/\sigma_j^2$$

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(9) reliabFactor

Type	Array	Unit
4-byte float	nray * nscan	N/A

Reliability Factor for the effective PIA estimate, pathAtten. This is defined as:

$$reliabFactor = \mathrm{Rel}_{eff} = (\Sigma u_j)^{-1/2} \Sigma u_j PIA_j$$

Missing Value: -9999.9

(10) reliabFactorHY

Type	Array	Unit
4-byte float	nray * nscan	N/A

reliabFactorHY is the reliability factor associated with the PIAhybrid estimate and is defined as the mean over the standard deviation of the estimate.

Missing Value: -9999.9

(11) reliabFlag

Type	Array	Unit
2-byte integer	nray * nscan	N/A

The reliability flag for the effective PIA estimate (pathAtten) based on the reliability factor (Rel_eff) in reliabFactor. Reliability Flag is:

= 1 if Rel_eff > 3 ; PIAeff estimate is considered reliable

= 2 if 3 >= Rel_eff > 1 ; PIAeff estimate is considered marginally reliable

= 3 if Rel_eff ≦1 ; PIAeff is unreliable

= 4 if SNR at surface < 2dB ; provides a lower bound to the path-attenuation

= 9 (no-rain case)

Missing Value: -9999

(12) reliabFlagHY

Type	Array	Unit
2-byte integer	nray * nscan	N/A

reliabFlagHY is the reliability flag for the PIAhybrid and is defined in the same way as reliabFlag where PIAeff is replaced by PIAhybrid.

Missing Value: -9999

(13) refScanID

Type	Array	Unit
2-byte integer	nearFar * foreBack * nray * nscan	Number

refScanID gives the number of scan lines between the current scan and the beginning (or end) of the along-track reference data at each angle bin. The values are computed by the equation: Current Scan Number - Reference Scan Number. The values are positive for the Forward estimates and negative for the Backward estimates. The Fortran indices are:

Bit Meaning
1, 1: Forward - Near reference
2, 1: Forward - Far reference
1, 2: Backward - Near reference
2, 2: Backward - Far reference
-9999: Missing

To illustrate, consider the following example. At a certain incidence angle assume that rain is present at scan numbers from 100 to 105 and from 110 to 120. At scan number 112, refScanID (1, 1) = 3, refScanID (2, 1) = 16; i.e., the eight rain-free NRCS data points, used to estimate the mean and standard deviation of the rain-free NRCS, begin at scan 112-16 = 96 and end at scan 112-3=109. These numbers provide information on the distance (in terms of the numbers of scans where 1 scan ~5 km) of the rain-free reference data from the rain pixel of interest. See section 6 for further details.

(14) stddevEff

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nsdew * nray * nscan	nfreq * nsdew * nray * nscan	N/A

stddevEff (1) contains the standard deviation of the PIAeff (i.e., the composite or effective SRT or hybrid path attenuation estimate). It is given by

$$\sigma_1 = (\sum 1/\sigma_{e,j}^2)^{-1/2} = (\sum u_j)^{-1/2}$$

It is important to note that in the definition of the reliability factor, it is this standard deviation that is used. In other words, with the notation in ATBD, we have $\sigma_1 = \sigma_{eff}$.

stddevEff (2) is a weighted root mean square error and provides a measured of the error of the individual PIA estimates from the effective PIA estimate. It is given by

$$\sigma_2 = [\sum w_j (A_{eff} - A_j)^2]^{1/2}$$

stddevEff (3) is given by

$$\sigma_3 = \left[\sigma_1^2 + \sigma_2^2\right]^{1/2}$$

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(15) stddevHY

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	N/A

stddevHY is the standard deviation (dB) of the hybrid estimate of path attenuation.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(16) zeta

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	N/A

zeta (unitless) is a parameter in the Hitschfeld-Bordan equation.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(17) piaExp (2ADPR FS)

Type	Array	Unit
4-byte float	nfreq * nray * nscan	N/A

TBD

nfreq (1): TBD

nfreq (2): TBD

Missing Value: -9999.9

2.2.11. DSD (Group)

(1) phase

Type	Array	Unit
1-byte unsigned integer	nbin * nray * nscan	N/A

Phase state of the precipitation. As an unsigned byte value this represents:

phase < 100 Temperature(C)=phase-100

phase > 200 Temperature(C)=phase-200

phase = 100 Top of the bright band

phase = 200 Bottom of the bright band

phase = 125 is used for the range bins between

the top and peak of bright band

phase = 175 is used for the range bins between

the peak and bottom of bright band

Integer values of phase/100 =

0 - solid

1 - mixed phase

2 - liquid

Missing Value: 255

(2) binNode

Type	Array	Unit
2-byte integer	nNode * nray * nscan	N/A

The bin number of the 5 nodes defined as:

1 - Bin number of storm top.

2 - Stratiform: 500m above center of bright band.

Convective: 750m above 0deg C level.

3 - Stratiform: center of bright band.

Convective: 0deg C level.

4 - Stratiform: 500m below center of bright band.

Convective: 750m below 0deg C level.

5 - Bin number of real surface equal to binRealSurface in PRE group.

For FS swaths, bin numbers are 1-based ranging from 1 at the top of the data window with 176 at the Ellipsoid. For HS swaths, bin numbers are 1-based ranging from 1 at the top of the data window with 88 at the Ellipsoid.

Missing Value: -9999

(3) paramRDm

Type	Array	Unit
4-byte float	nNode * nray * nscan	N/A

The parameters in R-Dm relation are p = 0.392, q = 6.131, and r = 4.815. They are the same and the definition of ε is the same for all precipitation types. In the single-frequency algorithm or if the DSD database is not applied, μ is 0 for stratiform precipitation but μ is $\log_{10}(1.25)$ ~ 0.0969 for convective precipitation. The details of these parameters are described in ATBD.

nNode (1): p-value in R-Dm relation.

nNode (2): q-value in R-Dm relation

nNode (3): r-value in R-Dm relation

nNode (4): μ-value for stratiform and convective precipitation

nNode (5): σ-value for stratiform and convective precipitation

Missing Value: -9999

2.2.12. Experimental (Group)

(1) precipRateESurface2

Type	Array	Unit
4-byte float	nray * nscan	mm/hr

Surface precipitation estimate based on an a priori low-level precipitation profiles, based upon Hirose et al. (2021, https://doi.org/10.2151/jmsj.2021-060).

Missing Value: -9999.9

(2) precipRateESurface2Status

Type	Array	Unit
1-byte unsigned integer	nray * nscan	N/A

Status of the estimated surface precipitation using alternate method.

Missing Value: 255

(3) sigmaZeroProfile

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nbinSZP * nray * nscan	nfreq * nbinSZP * nray * nscan	dB

The profiles of surface backscattering cross section around the binDEM (the range-bin number of averaged DEM). nbinSZP =7 for FS and nbinSZP = 5 for HS.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(4) seaIceConcentration

Type	Array	Unit
4-byte float	nray * nscan	%

Sea ice concentration (30.0 – 100.0%) estimated by Ku.

Missing Value: -9999.9

(5) binMixedPhaseTop (2ADPR FS)

Type	Array	Unit
2-byte integer	nray * nscan	N/A

Range bin number of MixedPhaseTop when it is detected. Definition of the MixedPhaseTop is described in ATBD.

Missing Value: -9999

(6) flagGraupelHail (2ADPR FS)

Type	Array	Unit
1-byte unsigned integer	nray * nscan	N/A

This flag takes the following value:

1: Flag = 1 indicates graupel or hail exists along vertical profile.

0: Flag = 0 indicates no graupel or hail exists along profile.

Missing Value: 255

(7) flagSurfaceSnowfall (2ADPR FS)

Type	Array	Unit
1-byte unsigned integer	nray * nscan	N/A

This flag takes the following value:

1: when surface snowfall exists (on surface, not aloft).

0: when surface snowfall doesn’t exist.

Missing Value: 255

(8) surfaceSnowfallIndex (2ADPR FS)

Type	Array	Unit
4-byte float	nray * nscan	N/A

flagSurfaceSnowfall is 1 when this index exceeds the defined threshold.

When no rain or skipped, the value is 0.0.

Missing Value: -9999.9

2.2.13. SLV (Group)

(1) flagSLV

Type	Array	Unit
1-byte integer	nbin * nray * nscan	N/A

A flag for each range bin data. At rain range bins, flagSLV is positive. At no-rain range bins, flagSLV is 0. If a range bin is located below ESurface, flagSLV is negative (-64). When the retrieval is abnormally terminated or data quality is bad, flagSLV is negative (-128).

The values are

flagSLV%2

0: no rain

1: rain

flagSLV%4

3: Zm is used for the retrieval

(1: extrapolated Ze is used for the retrieval)

(0: no rain)

flagSLV%16

(0-3: no rain)

4-7: only KuPR is used for the retrieval

8-11: only KaPR is used for the retrieval

12-15: Both KuPR and KaPR's Zm are used for the retrieval

flagSLV%64

0-15: Dm takes normal value (or no-rain)

16-31: Dm takes the minimum value

32-47: Dm takes the maximum value

48-63: Dm takes an abnormal value

flagSLV% 256

0-63: R takes a normal value (or no-rain)

64-127: R takes the maximum value

(128: bad data quality)

(192: below ESurface)

Missing Value: -99

(2) qualitySLV

Type	Array	Unit
4-byte integer	nray * nscan	N/A

A flag to show methods in which precipRateNearSurface is retrieved. See the Appendix of Seto et al. (2021, JMSJ)

Special values are defined as:

Negative value indicates processing error.

Missing Value: -9999

(3) binEchoBottom

Type	Array	Unit
2-byte integer	nray * nscan	N/A

The bin number of bottom of echo. For FS swaths, bin numbers are 1-based ranging from 1 at the top of the data window with 176 at the Ellipsoid. For HS swaths, bin numbers are 1-based ranging from 1 at the top of the data window with 88 at the Ellipsoid.

Missing Value: -9999

(4) piaFinal

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dB

The final estimates of path integrated attenuation caused by precipitation particles. It is calculated from the retrieved DSD profiles. It includes the attenuation only by precipitation.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(5) piaOffset

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dB

Increments of PIA estimates by SRT for considering the soil moisture effect (over land only).

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(6) sigmaZeroCorrected

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dB

Surface backscatter cross section with attenuation correction.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(7) zFactorFinal

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nbin * nray * nscan	nfreq * nbin * nray * nscan	dBZ

Vertical profile of reflectivity factor calculated from DSD estimates.

10 log10(Z) where Z is in mm^6/m^3.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(8) zFactorFinalESurface

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dBZ

Reflectivity factor calculated from DSD estimates at estimated surface.

10 log10(Z) where Z is in mm^6/m^3.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(9) zFactorFinalNearSurface

Type	Array	Array (2ADPR FS)	Unit
4-byte float	nray * nscan	nfreq * nray * nscan	dBZ

Reflectivity factor calculated from DSD estimates at near surface.

10 log10(Z) where Z is in mm^6/m^3.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

Missing Value: -9999.9

(10) paramDSD

Type	Array	Unit
4-byte float	nDSD * nbin * nray * nscan	nDSD (1): 10 $\log_{10}$(Nw) nDSD (2): mm

Parameters of DSD functions, Nw and Dm. Nw in $1/m^3$ mm

nDSD (1): $10\log_{10}$(Nw)

nDSD (2): Dm

Missing Value: -9999.9

(11) precipRate

Type	Array	Unit
4-byte float	nbin * nray * nscan	mm/hr

Precipitation rate.

Missing Value: -9999.9

(12) precipWater

Type	Array	Unit
4-byte float	nbin * nray * nscan	g/m^3

The amount of precipitable water. Note that the incorrect unit of "precipWater" is defined as "kg/m^3" in the HDF products. "g/m^3" is correct.

Missing Value: -9999.9

(13) precipWaterIntegrated

Type	Array	Unit
4-byte float	LS * nray * nscan	g/m^2

Precipitation water vertically integrated.

For LS=1, sum of liquid water (phase >= 200)

For LS=2, sum of non-liquid water (phase < 200)

Missing Value: -9999.9

(14) precipRateESurface

Type	Array	Unit
4-byte float	nray * nscan	mm/hr

Precipitation rate for the estimated surface, i.e., at binRealSurface.

Missing Value: -9999.9

(15) precipRateNearSurface

Type	Array	Unit
4-byte float	nray * nscan	mm/hr

Precipitation rate for the nearSurface bin, i.e., at binClutterFreeBottom.

Missing Value: -9999.9

(16) precipRateAve24

Type	Array	Unit
4-byte float	nray * nscan	mm/hr

Average of precipitation rate for 2 to 4km height.

Missing Value: -9999.9

(17) phaseNearSurface

Type	Array	Unit
1-byte unsigned integer	nray * nscan	N/A

Value of the Phase parameter in the DSD module at binClutterFreeBottom (nearSurface bin).

Missing Value: 255

(18) epsilon

Type	Array	Unit
4-byte float	nbin * nray * nscan	N/A

Epsilon is the indication of the adjustment away from the initial drop size distribution, epsilon = 1 is no adjustment.

Missing Value: -9999.9

(19) paramNUBF

Type	Array	Unit
4-byte float	nNUBF * nray * nscan	N/A

The parameter to adjustment of None Uniform Beam Filling (NUBF).

paramNUBF (1) is σ_T^2 where $\sigma_T = \sqrt{\frac{\sigma^2+1}{p} - 1}$.

paramNUBF (2) is σ^2 where σ is the coefficient of variation of Nw.

paramNUBF (3) is p where p is the ratio of the raining area to the total area in FOV. (Currently p is set to 1.)

Missing Value (in the pixel with zero precipitation or in the case of poor data quality): -9999.9

(20) DFRforward1 (2ADPR FS)

Type	Array	Unit
4-byte float	nbin * nray * nscan	N/A

Difference of Zf (calculated by attenuation correction) between KaPR and KuPR.

Missing Value: -9999.9

2.2.14. FLG (Group)

(1) flagEcho

Type	Array	Unit
1-byte integer	nbin * nray * nscan	N/A

The values are

Bit Meaning

0: For L2 Ku/PR: Precipitation judged by L2 Ku algorithm (copy of bit 2)

0: For L2 Ka: Precipitation judged by L2 Ka algorithm (copy of bit 3)

0: For L2 DPR: Precipitation judged by L2 DPR algorithm (copy of bit 1)

1: Precipitation judged by L2 DPR algorithm

2: Precipitation judged by L2 Ku algorithm

3: Precipitation judged by L2 Ka algorithm

4: Main lobe clutter judged by L2 Ku algorithm

5: Main lobe clutter judged by L2 Ka algorithm

6: Side lobe clutter judged by L2 Ku algorithm

7: Side lobe clutter judged by L2 Ka algorithm

(2) qualityData

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Normal data gives ”0”. Non-zero values mean the kinds of errors

The values are

The 2 bit flag for each module has values: [higher bit lower bit]

[0 0]: Good

[0 1]: Warning but usable

[1 0]: NG or error

The bits of qualityData are assigned as follows:

0 - 7: Copy of dataQuality in level 1B product

8 - 9: Flag by input module

10 - 11: Flag by preparation module

12 - 13: Flag by vertical module

14 - 15: Flag by classification module

16 - 17: Flag by SRT module

18 - 19: Flag by DSD module

20 - 21: Flag by solver module

22 - 23: Flag by output module

24 - 31: Spare

-9999: Missing

(3) flagSensor

Type	Array	Array (2ADPR FS)	Unit
1-byte integer	nscan	nfreq * nscan	N/A

Flag of input Ku/Ka data condition

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

The values are

1: Valid

-99: Invalid (judged by dataQuality)

(4) qualityFlag

Type	Array	Array (2ADPR FS)	Unit
1-byte integer	nray*nscan	nfreq * nary*nscan	N/A

qualityFlag is a sample flag generated by qualityData.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

The values are

0: High quality. No issues.

1: Low quality. (DPR modules had warnings but still made a retrieval)

2: Bad. (DPR modules had errors or dataQuality is bad and retrieval is missing)

-99: Missing value

(5) flagScanPattern

Type	Array	Array (2ADPR FS)	Unit
2-byte integer	nscan	nfreq * nscan	N/A

Flag of scan pattern information. Ku and PR are always “0”.

In the 2ADPR FS,

nfreq (1): Estimated by KuPR single-frequency algorithm.

nfreq (2): Estimated by KaPR single-frequency algorithm.

The values are

0: Original scan pattern. (from the beginning of the mission until May 21, 2018)

1: KaHS outer swath scan pattern (After May 21, 2018)

-99: Others or Missing

2.2.15. TRG (Group)

(1) NUBFindex

Type	Array	Unit
4-byte float	nray * nscan	N/A

Qualitative index for presence of NUBF. Not to be used, just for our development and debugging.

Values range from 0 to 100.

Missing value: -9999.9

(2) MSindex

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Index for presence of MS based on Ku and Ka return signal.

Values [0,50,100]

(3) MSindexKu

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Index for presence of MS based on Ku return signal only.

Values [0,50,100]

(4) MSindexKa

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Index for presence of MS based on Ka return signal only.

Values [0,50,100]

(5) MSsurfPeakIndexKu

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Index for detection of surface peak for Ku return signal.

Values [0,50,100].

(6) MSsurfPeakIndexKa

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Index for detection of surface peak for Ka return signal.

Values [0,50,100]

(7) MSkneeDFRindex

Type	Array	Unit
4-byte integer	nray * nscan	N/A

Index for detection of knee feature in DFR.

Values [0,50,100].

(8) MSslopesKu

Type	Array	Unit
4-byte float	4 * nray * nscan	N/A

Slope values in different portion of the Ku signal around the surface peak.

Values [0,50,100]

(9) MSslopesKa

Type	Array	Unit
4-byte float	4 * nray * nscan	N/A

Slope values in different portion of the Ka signal around the surface peak.

Values [0,50,100]

Appendix-2 Product format of ENV file

1. Level 2 Data Format Structure

1.1. Dimension definition

Dimension definitions:

- nscan
  - Number of scans in the granule.
- nray
  - 49 Number of angle bins in each scan. (FS)
  - 24 Number of angle bins in each scan. (HS)
- nbin
  - 176 Number of range bins in each ray. (FS)
  - 88 Number of range bins in each ray. (HS)

"FS" is called as Full scan Swath in 2AKu, 2AKa and 2ADPR.

"HS" is called as High sensitivity beam scan Swath in 2AKa and 2ADPR.

1.2. Data Format Structure for 2AKu ENV

The Ku Level-2A Environment product, 2AKu ENV, is defined as a swath structure, which is called “FS”.

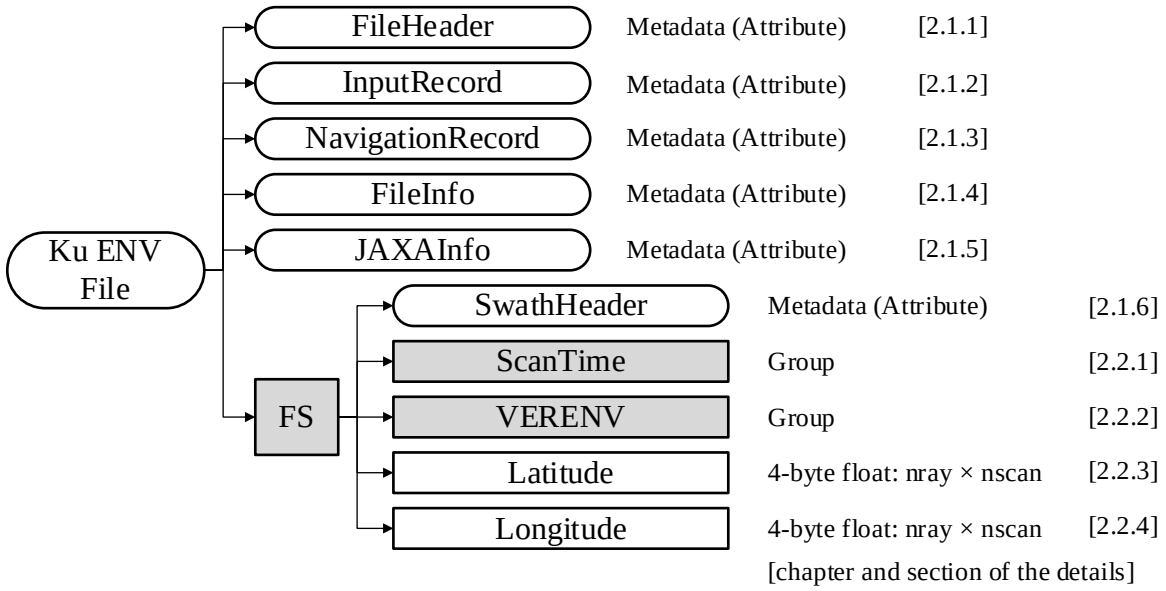


Figure 1.2-1 Data Format Structure for 2AKu ENV

1.3. Data Format Structure for 2AKa ENV

The Ka Level-2A Environment product, 2AKa ENV, is defined as two-swath structures, which are called “FS” and “HS”.

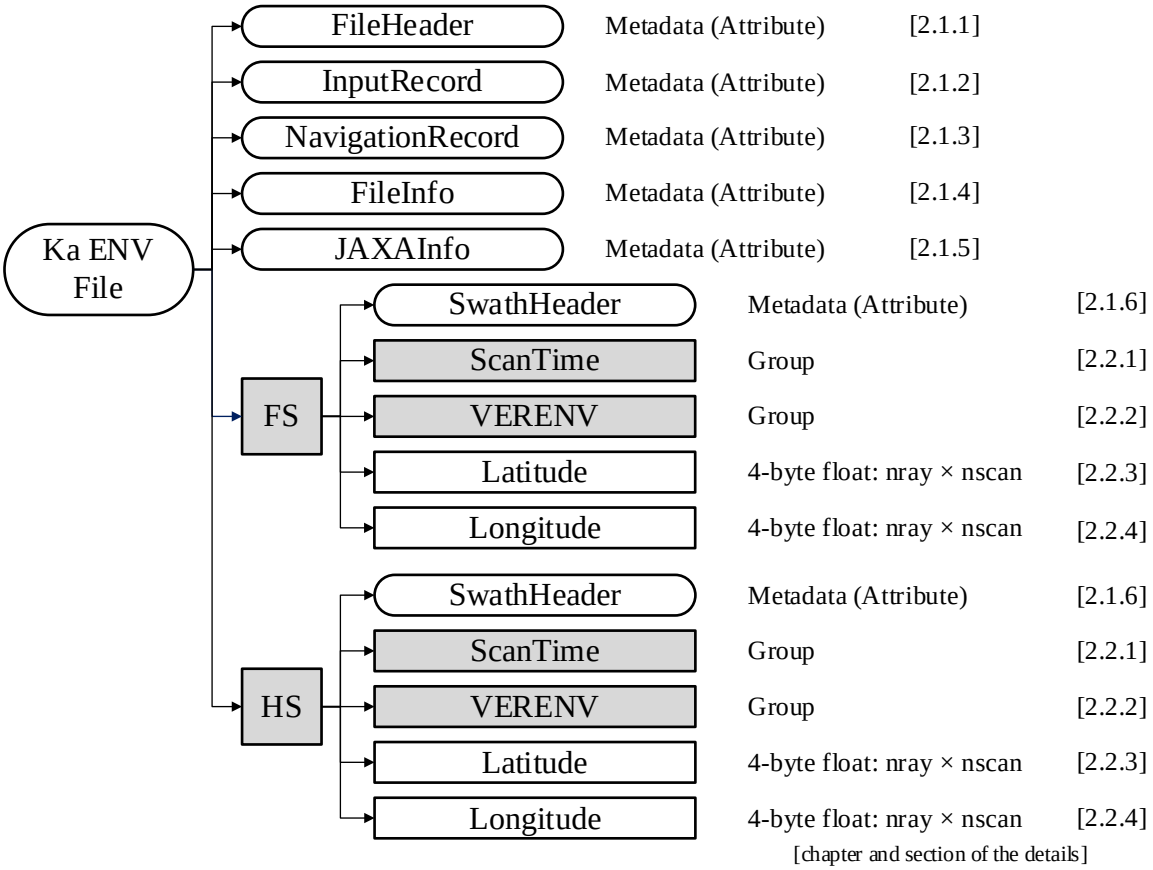


Figure 1.3-1 Data Format Structure for 2AKa ENV

1.4. Data Format Structure for 2ADPR ENV

The DPR Level-2A Environment product, 2ADPR ENV, is defined as two-swath structures, which are called “FS” and “HS”.

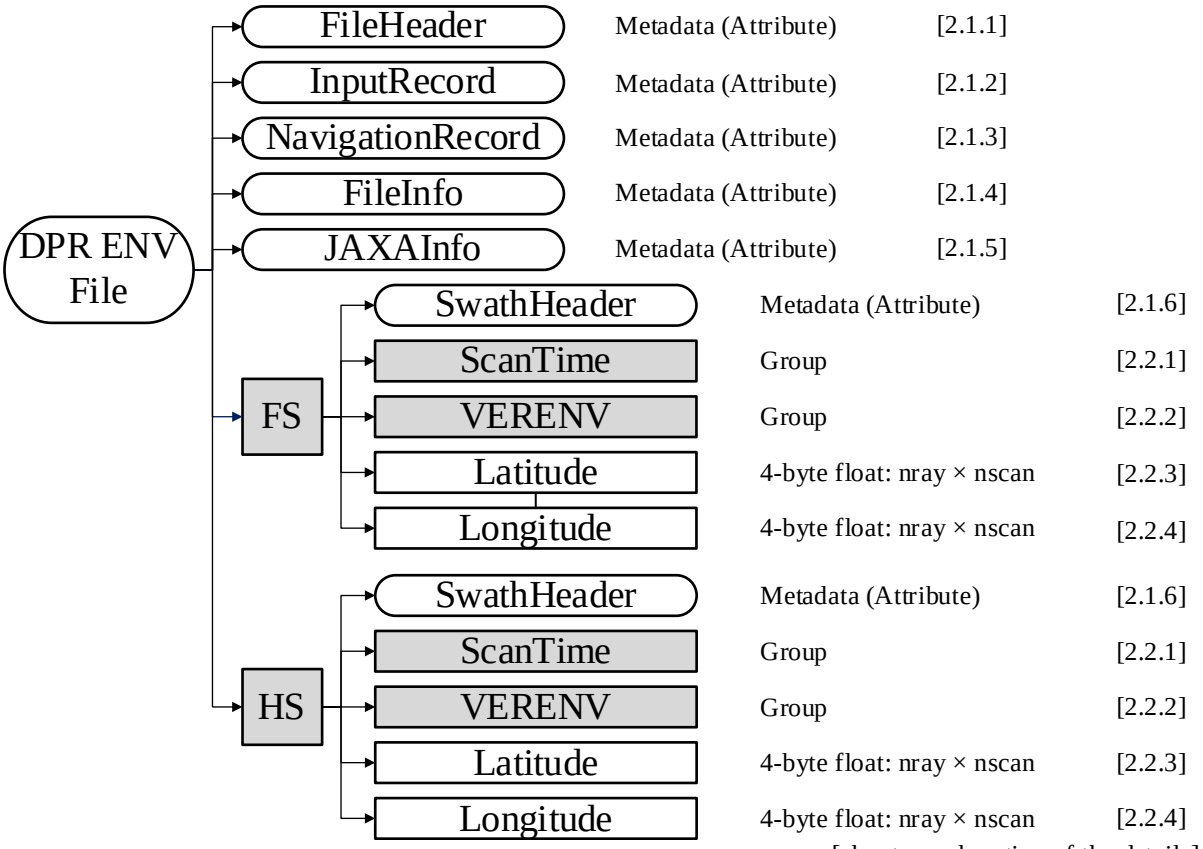


Figure 1.4-1 Data Format Structure for 2ADPR ENV

1.5. Data Format Structure for each Group

Each group's structure is shown in this section. Structures in each grid are common. However, the number of rays and range bins are different as shown in section 1.1.

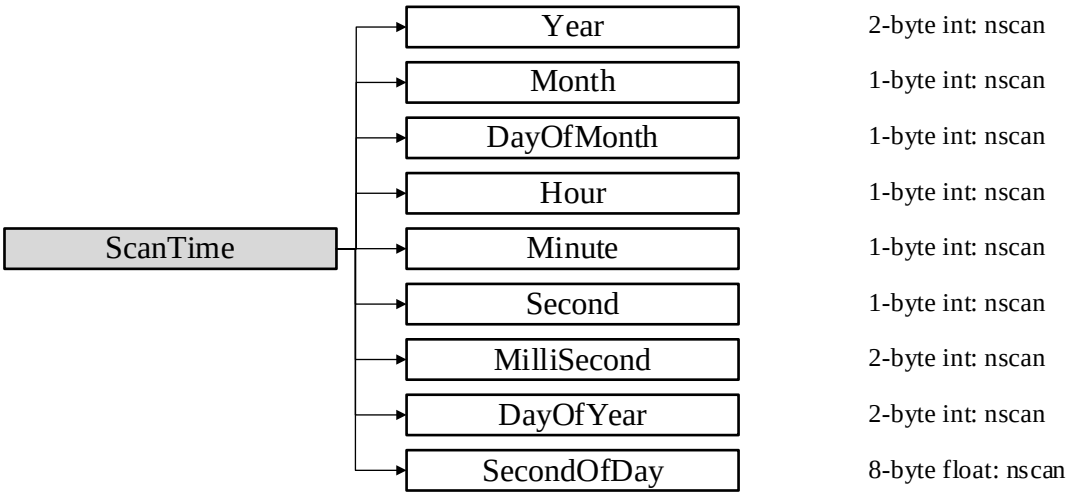


Figure 1.5-1 Data Format Structure for ScanTime Group

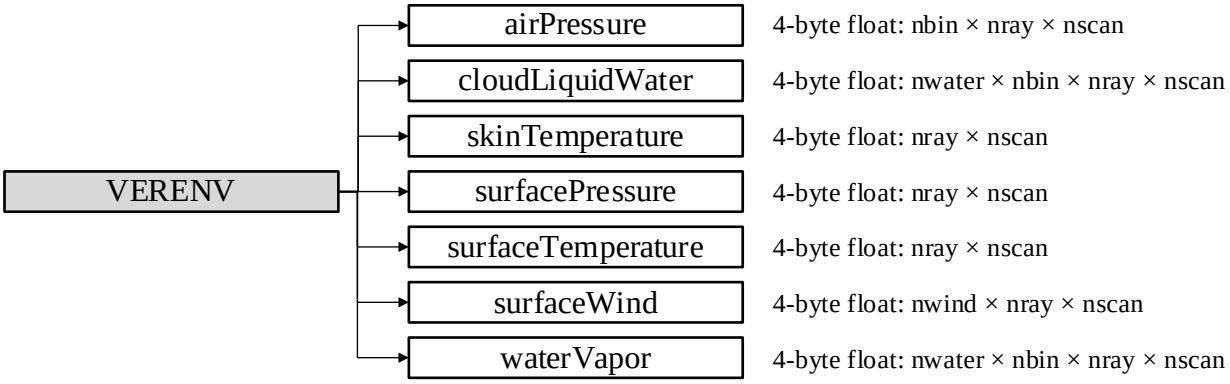


Figure 1.5-2 Data Format Structure for VERENV Group

2. Level 2 Contents of Objects in each Group

2.1. Metadata

Metadata has six elements. Figure 2.1-1 shows metadata structure.

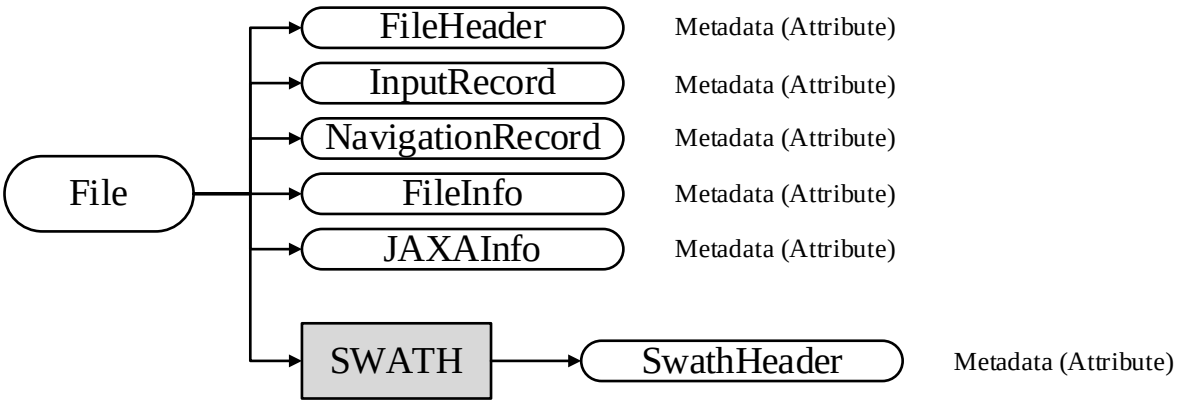


Figure 2.1-1 L2 ENV Metadata

2.1.1. FileHeader

FileHeader contains general metadata. Table 2.1-1 shows each metadata elements in FileHeader.

Table 2.1-1 FileHeader Elements

No	Element	Description	Data size (bytes)
1	DOI	Digital Object Identifier. *Value is blank currently.	256
2	DOIauthority	Digital Object Identifier Authority.	256
3	DOIshortName	Digital Object Identifier Short Name. *Value is blank currently.	256
4	AlgorithmID	The algorithm that generated this product, e.g., 2A12.	50
5	AlgorithmVersion	The version of the algorithm that generated this product.	50
6	FileName	The file name of this granule.	50
7	SatelliteName	Values are: TRMM GPM MULTI F10 ... F18 AQUA GCOMW1 CORIOLIS MT1 NOAA15 ... NOAA19 METOPA NPP. More values will be added as they are known.	10

No	Element	Description	Data size (bytes)
8	InstrumentName	Values are: PR TMI VIRS PRTMI KU KA DPR GMI DPRGMI MERGED SSMI SSMIS AMSRE AMSR2 WINDSAT MADRAS AMSUA AMSUB SAPHIR MHS ATMS. More values will be added as they are known.	10
9	GenerationDateTime	The date and time this granule was generated. The format is YYYY-MM-DDTHH:MM:SS.sssZ, where YYYY is 4-digit year, MM is month number, DD is day of month, T is "T", HH is hour, MM is minute, SS is second, sss is millisecond, and Z is "Z". All fields are zero-filled. The missing value is constructed by replacing all digits with 9, i.e., 9999-99-99T99:99:99.999Z.	50
10	StartGranuleDateTime	The start time defining this granule. The format is the same as GenerationDateTime. DETAILS: An orbital granule starts when the satellite is at the position defined by GranuleStart. Thus the start time is not the first scan time. Some algorithms have overlap scans in the file before the start time as defined in SwathHeader. A monthly granule starts on the first ms of the month, for example March 1998 would be 1998-03-01T00:00:00.000Z.	50
11	StopGranuleDateTime	The stop time defining this granule. The format is the same as GenerationDateTime. DETAILS: An orbital granule stops when the satellite is at the position defined by GranuleStart. Thus the stop time is not the last scan time. Some algorithms have overlap scans in the file after the stop time as defined in SwathHeader. A monthly granule stops on the last ms of the month, for example March 1998 would be 1998-03-31T23:59:59.999Z.	50
12	GranuleNumber	The number of this granule, which starts as in GranuleStart. If the GranuleStart is identical to the orbit start, then the GranuleNumber will be the same as the orbit number. The GranuleNumber will have 6 digits, including leading zeroes, for example 001234.	50
13	NumberOfSwaths	The number of swaths in this granule.	50
14	NumberOfGrids	The number of grid structures in this granule.	50
15	GranuleStart	The starting place in the orbit of this granule. Currently defined values are ”SOUTHERNMOST LATITUDE” and ”NORTHBOUND EQUATOR CROSSING”.	50
16	TimeInterval	The time interval covered by this granule. Values are ”ORBIT”, ”HALF ORBIT”, ”HALF HOUR”, ”HOUR”, ”3 HOUR”, ”DAY”, ”DAY ASC”, DAY DES”, ”MONTH”, ”CONTACT”.	50
17	ProcessingSystem	The name of the processing system, e.g.,”PPS”, ”JAXA”.	50
18	ProductVersion	The data version assigned by the processing system.	50
19	EmptyGranule	Whether a granule is empty. Values are "EMPTY" or "NOT EMPTY".	50
20	MissingData	The number of missing scans.	50

2.1.2. InputRecord

InputRecord contains a record of input files for this granule. Table 2.1-2 shows each metadata elements in InputRecord.

Table 2.1-2 InputRecord Elements

No	Element	Description	Data size (bytes)
1	InputFileNames	A list of input file names for this granule.	1000
2	InputAlgorithmVersions	A list of algorithm versions of the input files for this granule.	1000
3	InputGenerationDateTimes	A list of generation date times of the input files for this granule. The format is the same as GenerationDateTime.	1000

2.1.3. NavigationRecord

NavigationRecord contains navigation metadata for this granule. Table 2.1-3 shows each metadata elements in NavigationRecord.

Table 2.1-3 NavigationRecord Elements

No	Element	Description	Data size (bytes)
1	LongitudeOnEquator	The longitude where the satellite crosses the equator going from south to north.	50
2	UTCDateTimeOnEquator	The UTC time when the satellite crosses the equator going from south to north. The format is the same as GenerationDate Time.	50
3	MeanSolarBetaAngle	The average solar beta angle in this granule.	50
4	EphemerisFileName	Name of the ephemeris file input for processing.	50
5	AttitudeFileName	Name of the attitude file input for processing.	50
6	GeoControlFileName	Name of the GeoTK Control Parameters File input for processing.	50
7	EphemerisSource	Values are ”0 CONSTANT INPUT TEST VALUE”, ”1 GROUND ESTIMATED STATE (GES)”, ”2 GPS FILTERED SOLUTION (GEONS)”, ”3 GPS POINT SOLUTION (PVT)”, ”4 ON BOARD PROPAGATED (OBP)”, ”5 OEM GROUND EPHEMERIS FILE”, ”6 GEONS WITH FALLBACK AS FLAGGED”, ”7 PVT WITH FALLBACK AS FLAGGED”, ”8 OBP WITH FALLBACK AS FLAGGED”, ”9 GES WITH FALLBACK AS FLAGGED”.	50
8	AttitudeSource	values are ”0 CONSTANT INPUTS FOR TESTING”, ”1 ON BOARD CALCULATED PITCH ROLL YAW”	50
9	GeoToolkitVersion	Version of the GeoToolkit.	50
10	SensorAlignmentFirstRotationAngle	Alignment angle, first rotation, in degrees. Rotation adjustment from sensor coordinates to the Attitude Control System Flight Coordinates.	50
11	SensorAlignmentSecondRotationAngle	Alignment angle, second rotation, in degrees.	50
12	SensorAlignmentThirdRotationAngle	Alignment angle, third rotation, in degrees.	50
13	SensorAlignmentFirstRotationAxis	Euler rotation sequence, first rotation axis. Values are "1","2", "3" (representing X, Y, Z).	50
14	SensorAlignmentSecondRotationAxis	Euler rotation sequence, second rotation axis. Values are "1","2", "3" (representing X, Y, Z).	50
15	SensorAlignmentThirdRotationAxis	Euler rotation sequence, third rotation axis. Values are "1","2", "3" (representing X, Y, Z).	50

2.1.4. FileInfo

FileInfo contains metadata used by the PPS I/O Toolkit (TKIO). This group appears in all data products. Table 2.1-4 shows each metadata elements in FileInfo.

Table 2.1-4 FileInfo Elements

No	Element	Description	Data size (bytes)
1	DataFormatVersion	The version of the data format used to write this file. This version is separate for each AlgorithmID. The order is: "a" "b" ... "z" "aa" "ab" ... "az" "ba" "bb".	50
2	TKCodeBuildVersion	Usually TK CodeBuildVersion is "1". If the I/O routines built by TKIO change even though the DataFormatVersion is unchanged, then TK CodeBuildVersion increments to "2", "3", ...If subsequently DataFormatVersion changes, TKCodeBuildVersion becomes "1" again.	50
3	MetadataVersion	The version of metadata used to write this file. This version is separate for each AlgorithmID. The order is: "a" "b" ... "z" "aa" "ab" ... "az" "ba" "bb" ...	50
4	FormatPackage	The underlying format of this granule. Values are ”HDF4”,”HDF5”, ”NETCDF”, ”TKBINARY”.	50
5	BlueprintFilename	The filename of the primary blueprint file that defined the format used to write this file.	50
6	BlueprintVersion	The BlueprintVersion of the format definition.	50
7	TKIOVersion	The version of TKIO used to create I/O routines to write this file. TKIOVersion does not define the format used to write this file.	50
8	MetadataStyle	The style in which the metadata was written, e.g., "PVL"."PVL" means < parameter >=< value >.	50
9	EndianType	The endian type of the system that wrote this file. Values are "BIG ENDIAN" and "LITTLE ENDIAN".	50

2.1.5. JAXAInfo

JAXAInfo contains metadata requested by JAXA. Used by DPR algorithms and GSMaP. Table 2.1-5 shows each metadata elements in JAXAInfo.

Table 2.1-5 JAXAInfo Elements

No	Element	Description	Data size (bytes)
1	GranuleFirstScanUTCDateTime	The date and time of first scan (incl. missing scan). The format is YYYY-MM-DDTHH:MM:SS.sssZ, where YYYY is 4-digit year, MM is month number, DD is day of month, T is "T", HH is hour, MM is minute, SS is second, sss is millisecond, and Z is "Z". All fields are zero-filled. The missing value is constructed by replacing all digits with 9, i.e., 9999-99-99T99:99:99.999Z.	50
2	GranuleLastScanUTCDateTime	Granule Last Scan UTC Date. Date is a 24 character string. The format is YYYY-MM-DDTHH:MM:SS.sssZ, where YYYY is 4-digit year, MM is month number, DD is day of month, T is ”T”, HH is hour, MM is minute, SS is second, sss is millisecond, and Z is ”Z”. All fields are zero-filled.	50
3	TotalQualityCode	The total quality of product is defined based on the quality of input data. Quality meaning are (a) GPM KuPR/KaPR L2 product Good: The total quality of input data (Ku/Ka L1B) is Good. Fair: The GPM KuPR/KaPR L2 is not JMA's global weather forecast (FCST) or JMA's Global ANALsis model data (GANAL) but weather DB file. EG (Empty Granule): The total quality of input data (Ku/Ka L1B) is EG (b) GPM DPR L2 product Good: The total quality of both Ku L2 and Ka L2 is Good. Fair: (i)The total quality of either Ku L2 or Ka L2 is EG (ii)The input data used in GPM DPR L2 is not JMA's global weather forecast (FCST) or JMA's Global ANALysis model data (GANAL) but weather DB file. EG (Empty Granule): The total quality of both Ku L2 and Ka L2 is EG. (c) GPM DPR SLH L2 product Good: The total quality of input data (DPR L2) is Good Fair: The total quality of input data is Fair. EG (Empty Granule): The total quality of input data (DPR L2) is EG.	50
4	FirstScanLat	Latitude of orbit first scan.	50
5	FirstScanLon	Longitude of orbit first scan.	50
6	LastScanLat	Latitude of orbit last scan.	50
7	LastScanLon	Longitude of orbit last scan.	50
8	NumberOfRainPixelsFS	Number of rain pixels in the FS swath, judged at DPR L2 algorithm. At DPR L1, value is "-9999".	50

No	Element	Description	Data size (bytes)
9	NumberOfRainPixelsHS	Number of rain pixels in the HS swath, judged at DPR L2algorithm. At DPR L1, value is "-9999".	50
10	ProcessingSubSystem	The name of the processing sub-system, e.g., "ALGORITHM","PCS".	50
11	ProcessingMode	The name of the processing mode, e.g., "STD","NRT".	50
12	Lightspeed	Constant value of light speed.	50
13	DielectricFactorKa	The parameter of dielectric for Ka.	50
14	DielectricFactorKu	The parameter of dielectric for Ku.	50

2.1.6. SwathHeader

SwathHeader contains metadata for swaths. Table 2.1-6 shows each metadata elements in SwathHeader.

Table 2.1-6 SwathHeader Elements

No	Element	Description	Data size (bytes)
1	NumberScansInSet	The scans read by TKreadScan are a ”set”. For single swath data, one scan is read so NumberScansInSet=1. For multiple swath data, one TKreadScan may read more than one scan. For example, for SSM/I data one TKreadScan reads one low frequency scan and two high frequency scans. Therefore NumberScansInSet=1 for the low frequency swath and Number-ScansInSet=2 for the high frequency swath.	50
2	MaximumNumberScansTotal	The maximum allowed number of total scans in this swath. Total scans = overlap scans before granule + scans in granule + overlap scans after granule.	50
3	NumberScansBeforeGranule	The number of overlap scans before the first scan of the granule in this swath.	50
4	NumberScansGranule	The number of scans in the granule in this swath.	50
5	NumberScansAfterGranule	The number of overlap scans after the last scan of the granule in this swath.	50
6	NumberPixels	The number of IFOV in each scan in this swath.	50

No	Element	Description	Data size (bytes)
7	ScanType	The type of scan in this swath. Values are: "CROSSTRACK" and "CONICAL".	50

2.2. Data Group

Elements of data group are explained in detail in this section. Each swath has 2 data group and 2 data (Latitude and Longitude) commonly. Figure 2.2-1 shows data group structure.

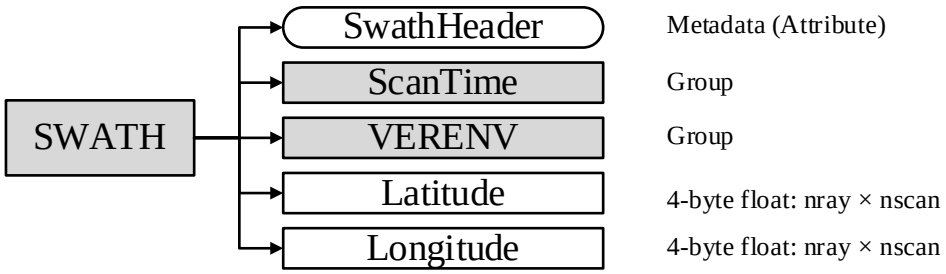


Figure 2.2-1 Data Format Structure for Data Group

2.2.1. ScanTime (Group)

(1) Year

Type	Array	Unit	Missing value
2-byte integer	nscan	year	-9999

4-digit year, e.g., 1998. Values range from 1950 to 2100 years:

(2) Month

Type	Array	Unit	Missing value
1-byte integer	nscan	month	-99

Month of the year. Values range from 1 to 12 months.

(3) DayOfMonth

Type	Array	Unit	Missing value
1-byte integer	nscan	day	-99

Day of the month. Values range from 1 to 31 days.

(4) Hour

Type	Array	Unit	Missing value
1-byte integer	nscan	hour	-99

UTC hour of the day. Values range from 0 to 23 hours.

(5) Minute

Type	Array	Unit	Missing value
1-byte integer	nscan	minute	-99

Minute of the hour. Values range from 0 to 59 minutes.

(6) Second

Type	Array	Unit	Missing value
1-byte integer	nscan	s	-99

Second of the minute. Values range from 0 to 60 s.

(7) MilliSecond

Type	Array	Unit	Missing value
2-byte integer	nscan	ms	-9999

Thousandths of the second. Values range from 0 to 999 ms.

(8) DayOfYear

Type	Array	Unit	Missing value
2-byte integer	nscan	day	-9999

Day of the year. Values range from 1 to 366 days.

Missing Value-9999

(9) SecondOfDay

Type	Array	Unit	Missing value
8-byte float	nscan	s	-9999.9

A time associated with the scan. It is expressed as the UTC seconds of the day.

Values range from 0 to 86400 s.

2.2.2. VERENV (Group)

(1) airPressure

Type	Array	Unit	Missing value
4-byte float	nbin * nray * nscan	hPa	-9999.9

Vertical profile of air pressure inserted from the ancillary data.

(2) cloudLiquidWater

Type	Array	Unit	Missing value
4-byte float	nwater * nbin *nray * nscan	kg/m^3	-9999.9

Vertical profile of cloud liquid water.

nwater = 0: a value diagnosed by the algorithm.

nwater = 1: a value inserted from the ancillary data.

(3) skinTemperature

Type	Array	Unit	Missing value
4-byte float	nray * nscan	K	-9999.9

Surface skin temperature inserted from the ancillary data.

(4) surfacePressure

Type	Array	Unit	Missing value
4-byte float	nray * nscan	hPa	-9999.9

Surface pressure inserted from the ancillary data.

(5) surfaceTemperature

Type	Array	Unit	Missing value
4-byte float	nray * nscan	K	-9999.9

Surface (2m) air temperature inserted from the ancillary data.

(6) surfaceWind

Type	Array	Unit	Missing value
4-byte float	nwind * nray * nscan	m/s	-9999.9

Surface wind.

nwind = 0: zonal direction.

nwind = 1: meridional direction.

(7) waterVapor

Type	Array	Unit	Missing value
4-byte float	nwater * nbin * nray * nscan	kg/m^3	-9999.9

Vertical profile of water vapor.

nwater = 0: a value diagnosed by the algorithm.

nwater = 1: a value inserted from the ancillary data.

2.2.3. Latitude

Type	Array	Unit	Missing value
4-byte float	nray * nscan	degrees	-9999.9

The earth latitude of the center of the IFOV at the altitude of the earth ellipsoid. Latitude is positive north, negative south. Values range from -90 to 90 degrees.

2.2.4. Longitude

Type	Array	Unit	Missing value
4-byte float	nray * nscan	degrees	-9999.9

The earth longitude of the center of the IFOV at the altitude of the earth ellipsoid. Longitude is positive east, negative west. A point on the 180th meridian has the value -180 degrees. Values range from -180 to 180 degrees.