AMSR2 PRECIPITATION

Algorithm Theoretical Basis Document

GPROF2017 Version 1 and Version 2

November 17th, 2021

TABLE OF CONTENTS

5.0 KNOWN LIMITATIONS

6.0 REFERENCES

GLOSSARY OF ACRONYMS

A

Advanced Microwave Scanning Radiometer for the Earth observing system (AMSR-E)
Advanced Microwave Scanning Radiometer 2 (AMSR2)
Advanced Microwave Sounding Unit (AMSU)
ATBD (Algorithm Theoretical Basis Document)

C

Colorado State University (CSU)

D

Dual Frequency Radar (DFR)

E

European Centre for Medium-Range Weather Forecasts (ECMWF)

G

GANAL (JMA Global ANALysis)
GEOS (Goddard Earth Observing System)
GPM (Global Precipitation Measurement)
GPM Microwave Imager (GMI)
GPM Profiling Algorithm (GPROF)
Global Data Assimilation System (GDAS)
Ground Validation (GV)

L

Land Surface Model (LSM)

N

National Centers for Environmental Prediction (NCEP)
Numerical weather prediction (NWP)

P

Precipitation Processing System (PPS)
Passive microwave retrieval (PWR)
Precipitation radar (PR)

T

Brightness temperature (T_b)
Tropical Rainfall Measuring Mission (TRMM)
Tropical Rainfall Measuring Mission - Microwave Imager (TMI)

1.0 INTRODUCTION

AMSR-E and AMSR2 are part of the GPM Constellation. The precipitation algorithm is the same as has been developed for GPM, using the same sensor calibration and precipitation algorithm. Only the version numbers which are specific to processing Centers, output fields and file formats are different in that the AMSR products defined their own specific outputs and formats and update cadence. The AMSR algorithm corresponds to GPROF 2017, which refers to the fundamental algorithm before it is adapted to specific sensors and processing Centers.

1.1 PURPOSE

This ATBD describes the parametric precipitation algorithm used by AMSR. The output parameters of the algorithm are enumerated in Table 1. Precipitation is based upon the concept that the GPM core satellite, with its Dual Frequency Radar (DPR) and GPM Microwave Imager (GMI), are be used to build a consistent *a-priori* database of cloud and precipitation profiles to help constrain possible solutions for any passive microwave radiometer. The AMSR algorithm searches this a-priori database for profiles that are consistent with its own observations.

Table 1. Key output parameters from the Level 2 Rainfall Product.

Pixel Information		
Parameter	**Units**	**Comments**
Latitude, longitude, Time	Deg.	Pixel earth coordinate position, and Pixel time
Surface Type	None	land surface emissivity class/ocean/coast/sea ice
Pixel Status	None	Identifies pixels eliminated by QC procedures
Quality Flag	None	Pixels w/o good T_b matches in database
2 meter Surface Temperature Total Column Water Vapor Surface skin temperature	^oK mm ^oK	Pass-through variables from GEOS5 Model
Surface Precipitation	mm/hr	Total Precipitation
Frozen Precipitation Convective Precipitation	mm/hr mm/hr	Frozen Precipitation – graupel and snow Convective Precipitation
Precipitation Diagnostics	None	Precip Retrieval diagnostics and uncertainties
Cloud, Rain, Ice Water Paths	Kg/m^2	Integrated water paths

The GPROF Quality Flag variable for GPROF 2017 are as follows:

Value 0: pixel is “good” and has the highest confidence of the best retrieval.

Value 1: “use with caution” . Pixels can be set to value 1 for the following reasons:

1) Sunglint is present, RFI, geolocate, warm load or other L1C ‘positive value” quality warning flags
2) All sea-ice covered surfaces
3) All snow covered surfaces
4) Sensor channels are missing, but not critical ones.

Value 2: "use pixel with extreme care over snow covered surface" This is a special value for snow covered surfaces only. The pixel is set to 2 if the probability of precipitation is of poor quality or indeterminate. Use these pixels for climatological averaging of precipitation, but not for individual storm scale daily cases.

Value 3: "Use with extreme caution". Pixels are set to value 3 if they have channels missing critical to the retrieval, but the choice has been made to continue the retrieval for these pixels.

2.0 INSTRUMENT CHARACTERISTICS

The AMSR-E instrument is a twelve channel, six frequency total power passive microwave radiometer system. It measures brightness temperatures at 6.925, 10.65, 18.7, 23.8, 36.5, and 89.0 GHz. AMSR2 is very similar instrument, with only one additional channel: 7.3 GHz used for mitigating Radio Frequency interference at 6.9265 GHz. Vertically and horizontally polarized measurements are taken at all channels.

The AMSR-E instrument, modified from the design used for the ADEOS-II AMSR, consists of a 1.6 meter diameter offset parabolic reflector (2.0 meters for AMSR2) fed by an array of six feedhorns. The reflector and feedhorn arrays are mounted on a drum, which contains the radiometers, digital data subsystem, mechanical scanning subsystem, and power subsystem. The reflector/feed/drum assembly is rotated about the axis of the drum by a coaxially mounted bearing and power transfer assembly. All data, commands, timing and telemetry signals, and power pass through the assembly on slip ring connectors to the rotating assembly.

A cold load reflector and a warm load are mounted on the transfer assembly shaft and do not rotate with the drum assembly. They are positioned off axis such that they pass between the feedhorn array and the parabolic reflector, occulting it once each scan. The cold load reflector reflects cold sky radiation into the feedhorn array thus serving, along with the warm load, as calibration references for the AMSR-E/AMSR2. Calibration of the radiometers is essential for collection of useful data. Corrections for spillover and other antenna pattern effects are incorporated in the data processing algorithms.

The AMSR-E rotates continuously about an axis parallel to the local spacecraft vertical at 40 rpm. At an altitude of 705 km (700km for AMSR2), it measures the upwelling scene brightness temperatures over an azimuthal range of +/- 70 degrees about the sub-satellite track, resulting in a swath width of 1500 km (1450km for AMSR2).

During a period of 1.5 seconds the spacecraft sub-satellite point travels 10 km. Even though the instantaneous field-of-view for each channel is different, active scene measurements are recorded at equal intervals of 10 km along the scan. To produce complete images at 89 GHz where the IFOV (5x5km) is smaller than the separation between scan lines, both AMSRs uses an offset 89 GHz channel (89B) that scans between the all-channel scan lines with an offset of 5 km. The half cone angle at which the reflector is fixed is 46.6 degrees which results in an Earth incidence angle of 53.8 degrees (AMSR2: 47.5 and 55.0 degrees). Table 2 lists the pertinent performance characteristics.

Table 2. AMSR-E and AMSR2 PERFORMANCE CHARACTERISTICS

AMSR-E							
Center Frequency (GHz)	6.925		10.65	18.7	23.8	36.5	89.0
Bandwidth (MHz)	350		100	200	400	1000	3000
Sensitivity (K)	0.34		0.7	0.7	0.6	0.7	1.2
IFOV(km x km)	75 x 43		51 x 29	27 x 16	32 x 18	14 x 8	6 x 4
Sampling rate (km x km)	10 x 9		10 x 9	10 x 9	10 x 9	10 x 10	5 x 5
Integration Time (msec)	2.5		2.5	2.5	2.5	2.5	1.2
Main Beam Efficiency (%)	95.3		95.0	96.3	96.4	95.3	96.0
Beamwidth (degrees)	2.2		1.5	0.8	0.92	0.42	0.19
AMSR2							
Center Frequency (GHz)	6.925	7.3	10.65	18.7	23.8	36.5	89.0
Bandwidth (MHz)	350	350	100	200	400	1000	3000
Sensitivity (K)	0.34	.43	0.7	0.7	0.6	0.7	1.4
IFOV(km x km)	62 x 35	62 x 35	42 x 24	22 x 14	26 x 15	12 x 7	5 x 3
Sampling rate (km x km)	10 x 10	10 x 10	10 x 10	10 x 10	10 x 10	10 x 10	5 x 5
Integration Time (msec)	2.6	2.6	2.6	2.6	2.6	2.6	1.3
Main Beam Efficiency (%)	>90.0	>90.0	>90.0	>90.0	>90.0	>90.0	>90.0
Beamwidth (degrees)	1.8	1.8	1.2	0.65	0.75	0.35	0.15

It should be noted that while AMSR-E and AMSR2 are very similar instruments, and use the same retrieval algorithm, the brightness temperatures provided in the Level 1B product used as input to the retrieval algorithm are not fully consistent. To avoid introducing Inconsistencies in the geophysical parameters, Tb for both AMSR-E and AMSR2 are shifted to be consistent with the GPM GMI radiometer as defined by the GPM Intercalibration Working Group (Berg et al., 2016).

3.0 ALGORITHM DESCRIPTION

The GPROF 2017 algorithm is based upon a Bayesian approach in which the GPM core satellite is used to create an *a-priori* database of observed cloud and precipitation profiles. Once a database of profiles and associated brightness temperatures is established, the retrieval employs a straightforward Bayesian inversion methodology. In this approach, the probability of a particular profile **R**, given $\mathbf{T_b}$ can be written as:

$$\Pr(\mathbf{R} \mid \mathbf{T_b}) = \Pr(\mathbf{R}) \times \Pr(\mathbf{T_b} \mid \mathbf{R}) \,, \tag{1}$$

where Pr(**R**) is the probability that a certain profile **R** will be observed and $\Pr(\mathbf{T_b} \mid \mathbf{R})$ is the probability of observing the brightness temperature vector, $\mathbf{T_b}$, given a particular rain profile **R**. The first term on the right hand side of Eqn. (1) is derived from the *a-priori* database of rain profiles established by the radar/radiometer observing systems discussed in section 3.1. The second term on the right hand side of Eqn. (1), is obtained from radiative transfer computations through the cloud model profiles. The formal solution to the above problem is presented in detail in Kummerow *et al*., (1996). In summary, the retrieval procedure can be said to compose a new hydrometeor profile by taking the weighted sum of structures in the cloud structure database that are radiometrically consistent with the observations. The weighting of each model profile in the compositing procedure is an exponential factor containing the mean square difference of the sensor observed brightness temperatures and a corresponding set of brightness temperatures obtained from radiative transfer calculations through the cloudy atmosphere represented by the model profile. In the Bayesian formulation, the retrieval solution is given by:

$$\hat{E}(R) = \sum_j R_j \frac{\exp\left\{-0.5\left(Tb_o - Tb_s\left(R_j\right)\right)^T \left(O+S\right)^{-1}\left(Tb_o - Tb_{s(R_j)}\right)\right\}}{\hat{A}} \tag{2}$$

Here, R_j is once again the vector of model profile values from the *a-priori* database model, Tb_o is the set of observed brightness temperatures, $Tb_s(x_j)$ is the corresponding set of brightness temperatures computed from the model profile R_j. The variables O and S are the observational and model error covariance matrices, respectively, and $\hat{A}$ is a normalization factor. The profile retrieval method is an integral version of the well-known minimum variance solution for obtaining an optimal estimate of geophysical parameters from available information (Lorenc, 1986, for a general discussion).

While the mechanics of Bayesian inversions are fairly well understood, four important issues are discussed separately in the following sections. The first concerns the use of ancillary data such as Surface Temperature and Total Column Water Vapor (TCWV or TPW) to search only appropriate portions of the a-priori database. Previous studies such as Berg et al., (2006) have shown that searching only over the appropriate SST and TCWV over oceans constrains the solution in a significant and positive manner. An important step is, therefore, to select the appropriate *a-priori* profiles in the Bayesian inversion. In the current version of the algorithm, the a-priori database is subsetted by 2-meter temperature (T2M), TCWV and Land Surface Classification. For this search to work, the ancillary data must be added to both the retrieval as well as the a-priori database. It is therefore discussed first in section 3.1. This is followed by the construction of the *a-priori* database itself in section 3.2. The next section (3.3) then deals with the uncertainties that are assigned to each channel in the Bayesian retrieval framework. Section 3.4 explains the probability of precipitation while section 3.5 deals with the phase of precipitation (liquid vs solid)

3.1 ANCILLARY DATA

Ancillary data is attached to each of the observed pixels in the sensor preprocessors. These parameters are included computed to each database pixel, and observed pixel: Two Meter Temperature (T2M), Total Column Water Vapor (TCWV), Surface Type, and Wet Bulb Temperature. Other parameters available, but not used for this version of GPROF are: surface skin temperature. The source of ancillary data determines the output product type. Real-time data products require forecast model output from the GEOS model close to the time of satellite data collection, while the standard product uses the analysis as it contains more observations. As described in section 4, this is handled in the sensor pre-processor portion of the algorithm to minimize changes to the retrieval code.

3.1.1 Creating the Surface Class Specification

The GPROF algorithm sub-sets the a-priori database by three different parameters: T2m, and TCWV (from model output) and 14 different surface classifications derived from a CSU surface classification scheme which intends to separate surfaces with similar microwave emissivity. Surface type classification begins with ten land classifications, ocean (or inland large water bodies), sea-ice and the two different boundaries that are possible between the interfaces (land-ocean and ocean-sea ice).

To create the land surface classes Prigent *et al*. 1997, estimated land surface emissivities from all available SSM/I observations from 1993 to 2008, under clear sky conditions. The dataset has been extensively analyzed and evaluated, by comparisons with both related surface parameters and model outputs. It has been shown to provide robust emissivity calculations, *i.e*., radiative transfer simulations using the emissivities are closer to the satellite observations. These estimates of the emissivities for all SSM/I frequencies are available with a spatial resolution of 0.25°×0.25° at the equator (equal-area grid) at monthly averaged intervals.

The seven dimensional emissivity space of mean SSM/I emissivities have been clustered using a K-means or Kohonen method. The emissivity classes are static but are applied on a monthly basis so that a single point can change classes as a function of time. In Fig. 3, the globe is classified into 10 land classes for January (Prigent *et al*. 2008). In this example, classes 1 to 5 are for increasing vegetation cover classes, 6 to 9 are for increasing snow/ice-covered pixels, and class 10 is for inland water-covered pixels. The TELSEM tool was used to analyze the correlation structure and the covariance matrices for each class, and each pixel location.

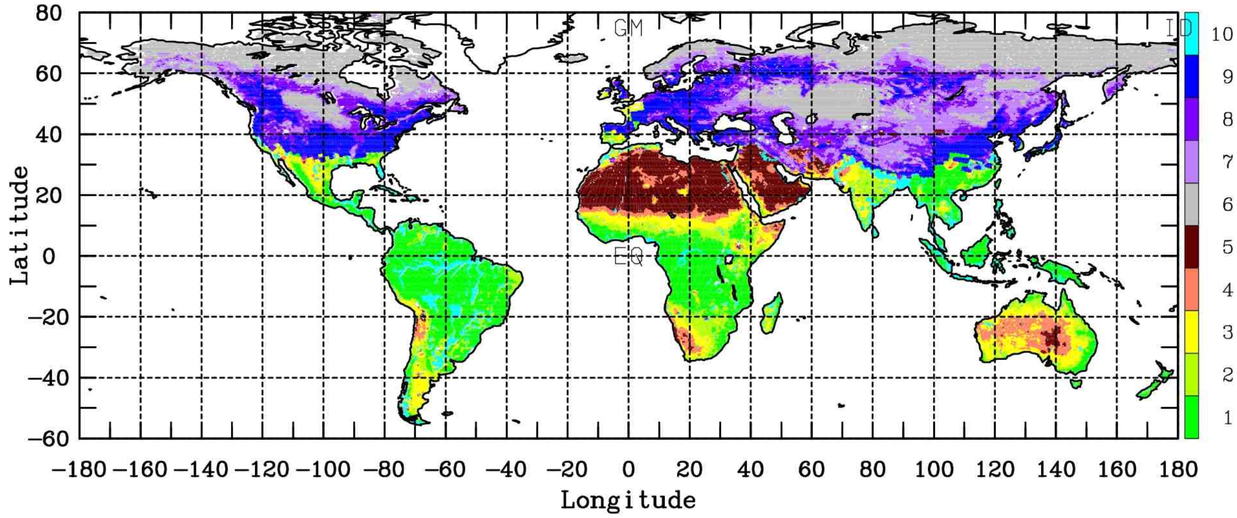

Fig. 3. Clustering of the SSM/I classes in ten self-similar emissivity classes. The ten classes have been defined corresponding to: five classes with increasing vegetation, four classes with increasing snow and ice and a class of what appears to be rivers/standing water/ estuaries.

The Colorado State University Land surface definition is a sensor specific classification starting with these 10 surface classes. Then we define the AMSR land/ocean boundary using a roughly one km MODIS/SeaWIFS / Ocean Color land mask. This was initially generated in 1993 and based on the World Vector Shoreline (WVS) database. In October 1997, shortly after the SeaWiFS launch, the land mask database was modified to include inland waterways, based on the World Data Bank (WDB) information. The final result is a 1/128th degree global grid that specifies either land or water. The GPROF sensor specific land mask files are on a 1/16th degree grid, derived at the nominal sensor footprints of 19GHz.

In the GPROF classification, the above land mask defines 2 additional classes to add to the 10 land surface classifications: Ocean and large inland water, and Land/Ocean Boundaries (coast). The coast class is defined as the percentage of land falling within the 1/16th degree grid box to be between 2 and 95%.

Finally, the areas of snow and sea-ice are defined from the GEOS5 model. The final CSU GPM surface classifications, and the ones output from the GPROF retrieval are numbered as:

1 : Ocean / Large inland Water
2 : Sea Ice
3 – 7 : Decreasing Vegetation Covered (3=Amazon, 7 = Sahara Desert)
8 – 11: Decreasing Snow Covered (8=Antarctica, 11 =lightly snow covered)
12 = : Inland Water / Rivers/ Estuaries
13 = : Coastlines (land/ocean boundary)
14 : Ocean / Sea-ice Boundary

3.1.2 Global Model Parameters

The GEOS5 FP-IT forecast model (GEOS5 Fcst) is used for the GPROF near real-time (NRT) processing. Model data is assimilated every 6 hours for profile and surface parameters. Forecast global fields are available and used at the closest observation time, until the subsequent analysis is completed.

For the GPROF Standard product (usually available within 12 hours of real-time) GEOS5 FP (GEOS5) is used.

Retrieved surface parameters for GEOS5 are : surface pressure, SL pressure, U and V component winds at 10 meters, total column water vapor, 2 meter temperature. 2 meter specific humidity, and Skin Temperature. The vertical profiles on constant pressure surfaces are: Temperature and Specific Humidity. The spatial resolution of the GEOS5 global grids is 0.5 X 0.625 degrees and of the GEOS5 Fcst it is 0.25 X 0.3125 degrees.

3.2 THE *A-PRIORI* DATABASES

The GPM V5 a-priori databases are constructed from one year (September 2014 – August 2015) of matched GMI/hydrometeor observations. One year is used to in order to capture the annual cycle of precipitation in the majority of meteorological conditions. For GPM V5, GMI radiances are matched with 3 hydrometeor sources- Combined MS, DPR Ku, and Multi-Radar/Multi-Sensor System (MRMS) ground-based radars.

First Source: The Combined MS product profiles of hydrometeors and surface precipitation derived from the DPR are used (parenthesis are GPROF surface types) over oceans (1), sea-ice (2), and sea-ice/ocean boundary (14).

Second Source: Version 4 of the Combined algorithm appeared to significantly overestimate precipitation over land. Therefore for V4 and also V5 the a-priori databases were constructed using DPR Ku surface precipitation for GPROF surfaces: vegetated (3-7), inland water (12), and coastlines (13). However, the Combined MS vertical hydrometeor profiles are still used for these surface types.

Third Source: Over land, the US based MRMS (ground-based radar) data was used to build a-priori databases for snow covered surfaces (8-11) for each of the constellations radiometers. Two years of MRMS data were matched up Tbs from each satellite sensor overpasses. This greatly improved the low bias that GPROF V4 had over snow covered surfaces. Because the MRMS data was only 2D and did not contain the vertical hydrometeor profiles, no vertical profile information is available from the GPROF retrieval over snow covered surfaces.

GPROF 2017 V5 uses the previous version (i.e. V4) of the Combined and DPR-Ku products for its databases. Future versions of GPROF, because of its need for existing GPM products to construct it's a-priori database, will always be one version behind the Combined algorithm. In GPROF 2017, we nonetheless improved some of the ice hydrometeor simulations in order to get better agreement between computed and simulated brightness temperatures. This leads to smaller bias adjustments in the radiometer simulations and leads to an overall better fit between the radiometer retrievals and both the Combined products as well as ground validation data.

Examination of the database simulations using the initial ice amounts and spherical ice particles from the V4 Combined DPR/GMI Algorithm product suggested that, though successful in matching observed Tbs at frequencies of 89 GHz and below, it was not possible to obtain agreement in a multi-specular way at the higher frequencies using the spherical particles. The ice was therefore adjusted as follows: ice contents from 2BCMB are used in the database simulations as a first guess. Ice particles are assumed to be non-spherical, and are represented as an ensemble of particle shapes, with single scatter parameters derived from the database produced at Florida State and described in Liu 2008. A simple exponential PSD is used to calculate the bulk optical properties. Resulting Tbs are then compared to GMI observations and adjusted iteratively to optimize agreement at 166 GHz. Residual high-biased disagreement at 183 GHz +/-3 is additionally improved by the addition of small cloud ice particles.

This adjustment process is designed to decrease simulation biases in the higher frequency channels. It is anticipated that in future versions of the algorithm, such adjustments will be unnecessary as the 2BCMB product begins to utilize the more realistic non-spherical ice particles within the DPR/GMI retrieval system starting with version 5.

GPROF 2017 also made additional changes to retrievals of high latitude oceanic drizzle and snowfall over land. Both these changes were made because the DPR sensitivity of 12 dBZ was shown to miss substantial amounts of drizzle and light to moderate snowfall events. Because the GPM radars do not have signal in these cases, they are not addressed in the newer versions of the Combined and Radar products either. Drizzle was addressed in the a-priori database by setting a threshold in the cloud liquid water retrieval from GMI (done before the DPR or Combined rainfall is inserted into the scene), to match the CloudSat based probability of rainfall. This is done for each temperature and Total Column Water Vapor (TCWV) bin used to subset the a-priori database. While this assumes that higher cloud liquid water amounts correspond to precipitation, the assumption is generally thought to be reasonable. Additional cloud water beyond the CloudSat determined threshold was partitioned between Cloud- and rain water similar to the procedure used by Hilburn and Wentz (2008). This increases rain water at high altitudes to agree better with CloudSat and ERA and MERRA re-analyses but continues to be low relative to these estimates. More work in ongoing to assess high latitude drizzle from different sources.

Some construction facts about the database:

- There are approximately 400 million matched GMI / Surface precip profiles from all sources
- Combined MS database was for September 2014 – August 2015 uses only the middle 21 pixels
- For MRMS matchups over the US, 2 winters were used, April 2014 – May 2016.
- The a-priori database is subsetted by : 0-78mm TCWV, 220 – 320 T2m, and 14 surface types
- Tbs are at their native resolution

- GEOS5 databases are used for defining pixels TCWV, and T2m, and Surface Wet Bulb Temperature
- CSU created surface maps are for 14 surface types

3.2.1 Matching Sensor Tbs to the Database Profiles

The first step in the a-priori database construction for AMSR is to use the DPR surface precipitation and layered hydrometer profiles in a forward model radiative transfer calculation to compute satellite radiances for the Tb frequency sets AMSR-E and AMSR-2. The DPR liquid and ice water content profiles are calculated for 28 vertical levels using the 2A25 Z-M relationship coefficients and the 2A25 freezing level information. The PR hydrometeors are then averaged into the 23GHz footprint using the nominal cross-track and down-track resolution for each sensor while the Tbs are kept at their native resolution. There are approximately 290 million DPR profiles used in the time period : September 2014 – August 2015.

3.2.2 Ancillary Data Added to the Profile Pixel

The next step uses the time and location of each DPR pixel to add 'ancillary' data to the pixel. These are added parameters necessary to group the pixels by common meteorological, and surface emissivity conditions. Depending on whether the database is to be used for the GPROF near-real time or Standard products, data from the GEOS5 FP-IT or GEOS5 FP model output is added to each footprint averaged DPR pixel. These parameters include two-meter surface Temperature (T2m) and Total Column Water Vapor (TCWV) and a calculated surface web bulb temperature. Interpolation of the model is then performed to an hourly resolution and smoothing to 0.25 degrees using a boxcar averaging. This greatly helps in eliminating GEOS5 model artifacts from the final precipitation products.

3.2.3 Final Clustering of Binned Profiles

For each T2m/TCWV/Surface bin a k-means clustering routine was used to decrease the number of profiles. A maximum of 800 profiles were kept along with a frequency of occurrence of the profiles within a given cluster (for each TCWV / T2m sorting bin). The clustering broadens the TCWV range up to +/- 4 mm. This has the intended effect of smoothing the profiles in the each clustered bin, and also increasing the total number of profiles in each bin. The TCWV broadening happens gradually until a maximum of 300,000 profiles is included in each of the 800 profile clusters within the bin. The final individual sensor GPROF database is made of 14 files, one for each surface type. GPROF2017 uses these 14 files to run the final the Bayesian rainfall retrieval.

3.3 CHANNELS AND CHANNEL UNCERTAINTIES

Uncertainties in physical inversions come from a combination of sensor noise and forward assumptions and errors. As described in Stephens and Kummerow (2010), rainfall retrieval errors tend to be dominated by the forward model assumptions. That is the case here as well and is particularly true when surface characteristics are not well known

In the GPROF 2017 V1 retrieval, the uncertainty is determined from the fit between the observed dataset and the CRM T_b that ultimately make the a-priori database. But there are modifications to this simplistic approach primarily due to foot print size based on the number of DPR pixels averaged into the sensors 23 GHz footprint. For AMSR which has comparable footprints to GMI, the channel error used in the retrieval is approximately the same as GMI.

3.4 PRECIPITATION PROBABILITY THRESHOLD

New in GPROF 2017 V1 (GPM V5) is the application of a probability of precipitation (POP) directly to retrieved precipitation values. The POP thresholds for rain/no-rain are calculated for each sensor, are bin (TCWV/T2m/Surface class) dependent, and are calculated as described below.

For each of the bins in the database the number raining and non-raining profiles are calculated. Precipitation above the 0.01 mm/hr threshold is counted as raining and below to non-raining. This rain fraction ratio for each bin is calculated and stored. As example is that in warm/moist bins the rain/(rain+norain) ratio is near 0.70 (70%) while in cooler dryer climates it might be 0.05 (5%). We then create a year of GPROF rain retrieval and without any probability thresholds applied. From that year of retrievals, we collate the frequency of rain, and a histogram of POP for each bin. We then calculate the needed POP threshold that matched the database rain frequency, and set a POP threshold for that bin. Finally, we compute the fraction of rain that is removed using this POP threshold. The POP threshold, and the fractional rain that is removed by the threshold is output and saved in a file that's read into the GPROF retrieval.

During the GPROF retrieval process, precipitation with probability below the threshold is set to 0.0. Retrieved Precipitation about the threshold is increased by the fraction of removed precipitation. This method conserves the total precipitation of the retrieval, an important step in the Bayesian statistical scheme.

3.5 PRECIPITATION TYPE (Liquid vs. Frozen) DETERMINATION

The GPROF Bayesian retrieval which matches the satellite radiances with hydrometeors from the 3 input sources retrieves the total surface precipitation, not the phase of the precipitation. An additional piece of information is needed to categorize the retrieval as raining or snowing - the surface dew point temperature. From (Sims and Liu, 2015) we use a lookup table of the fractional precipitation that is liquid at certain dew point temperatures. At -6.5 ^{o}C and below, 100% of the precipitation is frozen, while above 6.5 ^{o}C the precipitation is all liquid. Between these extremes, the precipitation is mixed and each type is output in the GPROF output file. There are 2 different lookup tables, one for ocean and one for land. The look tables which use the surface wet bulb temperature and used in GPROF are graphically described in Fig. 6.

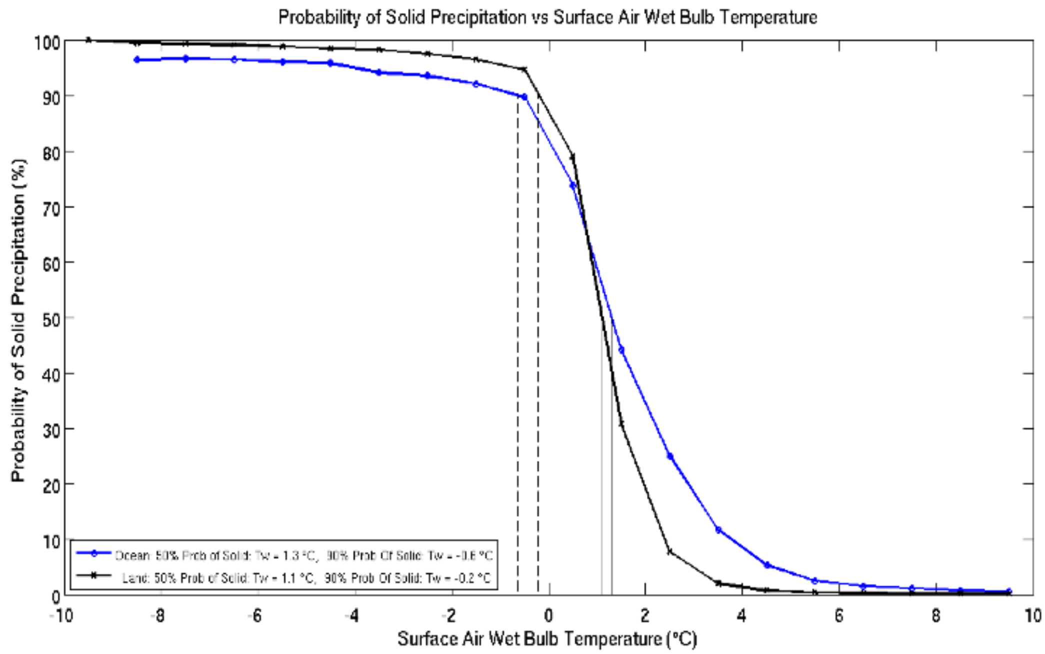

Figure 6. Solid Vs. Liquid Precipitation and Wet Bulb (Sims and Liu, 2015). Ocean is in blue, land in black.

4.0 ALGORITHM INFRASTRUCTURE

The code to ingest T_bs and ancillary files, perform quality control, assign surface types and decide on channel selection is written and maintained at Colorado State University in Fort Collins, Colorado, USA. The architecture is open to all team members as well as outside parties. We will use the input and output sections of this ATBD as a living document that is intended not only for the user of the precipitation product, but also for the algorithm developers to convey precise information about procedures, methods and formats.

The algorithm itself consists of Fortran 90 code that's self-contained in the GPROF_2017_V1 algorithm directory. All parameter fields and static databases must be accessible from this directory location as well as the dynamic ancillary data fields.

4.1 ALGORITHM INPUT

The AMSR algorithm implementation run by the AMSR Science Team described here uses L1R brightness temperatures provided by JAXA. In additional many 'ancillary' parameters are added that are required by the GPROF retrieval. These include individual sensor land masks, snow, and land surface classification, to create the surface classification. There are also the GEOS5 Forecast or GEOS5 model global grids used.

4.2 PROCESSING OUTLINE

Four processes are run to complete the GPROF Precipitation Algorithm. The following sections present a short description of each.

4.2.1 Preprocessor

The preprocessor is the interface between the orbital data files (L1R format) and the GPROF2017. The GPM sensor specific preprocessors read from the L1R HDF files and create the standard input file format. The preprocessor assigns all the ancillary data to each observed pixel along with the pixel's T_{bs}, latitude/longitudes, and sensor specifications. Also here in the preprocessor, the emissivity class, land masks, sea-ice, and model surface 2-meter temperature are used to create a surface classification for each pixel. Other parameters are output including the names and locations of the ancillary data directories, and profile databases - everything the GPROF 2017 needs to run the rainfall retrievals. A description of the output parameters of the preprocessor is given in Section 4.3.

4.2.2 AMSR Rainfall Processing Algorithm - GPROF2017

GPROF 2017 starts by reading the Standard Input file produced from the preprocessor. This includes all the ancillary data needed to match the T2m/TCWV/SurfaceType in the profile databases. This 3-dimensional matching is used to subset the entire set of database profiles for the Bayesian precipitation and profile retrieval. The width of the search in T2m/TCWV space is variable in TCWV space in the database creation depending on the number of binned profiles, in the retrieval +/- 1 K is used in the T2m bin space.

The output is to a native binary formatted file, and a description of the output parameters from the GPM PA is given in Section 4.4.

4.2.3 AMSR Post-Processor

The GPM Post-processor reads the CSU native binary output and attaches additional metadata from the original orbital L1R file and writes out a final formatted HDF5 EOS formatted - GPROF2017 file.

4.3 PREPROCESSOR OUTPUT

4.3.1 Preprocessor Orbit Header

satellite	Character*12
sensor	Character*12
preprocessor version	Character*12
original radiometer file	Character*128
profile database file	Character*128
calibration file	Character*128

granule number	integer*4
number of scans	integer*4
number of pixels in scan	integer*4
number of channels with data	integer*4
channel frequencies	real(10)*4
comment	Character*40

note : channel_freq describes the exact frequencies of the channels, but they must be in the following order: 10v, 10h, 19v, 19h, 23v, 23h, 37v, 37h, 89v, 89h

4.3.2 Preprocessor Scan Header

ScanDate (6)	integer(6)*2	: year,month,day,hour,min,sec
Spacecraft latitude	real*4	
Spacecraft longitude	real*4	
Spacecraft altitude	real*4	
tai93	real*8	

4.3.3 Preprocessor Data Record

Latitude		real*4
Longitude		real*4
Brightness temperatures		real(10)*4
Earth Incident angles		real(10)*4
Wet Bulb Temperature		real*4
Total Column Water Vapor		real*4
Skin temperature		real*4
2 Meter temperature		real*4
L1R quality Flag		integer*4
Sunglint angle	I	integer*2
Surface type code		integer*1
Spare		integer*2

4.4 AMSR PRECIPITATION ALGORITHM OUTPUT

Whether in the native, HDF5 EOS formats, the output parameters will be equivalent. This following format description is for the GPM native output binary format file.

4.4.1 Orbit Header (at beginning of each file) (described in section 4.4.4)

Satellite	Character*12
Sensor	Character*12
Pre-processor Version	Character*12
Algorithm Version	Character*12

Profile Database Filename	Character*128
Original Radiometer Filename	Character*128
File Creation Date/Time(6)	integer*2
Granule Start Date/Time(6)	integer*2
Granule End Date/Time(6)	integer*2
Granule Number	integer*4
Number of Scans in Granule	integer*2
Number of Pixels/Scan	integer*2
Spares	52 bytes

4.4.2 Scan Header (at beginning of each scan, described in section 4.4.5)

Spacecraft latitude	real*4
Spacecraft longitude	real*4
Spacecraft altitude (km)	real*4
Scan Date/Time (yr,mon,day,hour,min,sec,millisec)	integer*2
tai93time	real*8

4.4.3 Pixel Data (for each pixel in scan, described in section 4.4.6)

Pixel Status	integer*1 (one byte)
Quality Flag	integer*1
L1R Quality Flag	integer*1
Surface Type Index	integer*1
Total Col Water Vapor Index	integer*1
Probability of Precip	integer*1
2-meter Temperature	integer*2
Sunglint Angle	integer*1
Spare(3)	integer*1
Latitude	real*4
Longitude	real*4
Surface Precipitation	real*4
Frozen Precipitation	real*4
Convective Precipitation	real*4
Rain Water Path	real*4
Cloud Water Path	real*4
Ice Water Path	real*4

4.4.4 Orbit Header Variable Description – total of 400 bytes

Satellite

A character string indicating EOS-Aqua or GCOM-W

Sensor

Satellite Sensor, currently:
AMSR-E or AMSR2

PreProcessor Version

Pre-Processor version number.

Algorithm Version

GPROF Processing Algorithm Version which produced the output file.

Profile Database Filename

File name of the profile database. May be expanded to include multiple databases.

Original Radiometer Filename

File Name of the original, satellite observation input data file.

File Creation Date/Time

Start date and time of file creation. Defined as the date/time structure which holds six integer*2 values - year, month, day, hour, minute, second.

Granule Start Date/Time, End Date/Time

Start and End dates and times of first and last scan in file. Defined as the date/time structure, which holds year, month, day, hour, minute, second.

Granule Number

Generally this is defined as the satellite orbit number since launch.

Number of Scans in Granule, Number of Pixels per Scan

Number of sensor scans in the file, Number of pixels per scan for this sensor

Profile Structure Flag

Flag defining whether GPM Profiling Algorithm was run with vertical profiles of the hydrometeors. No structure = 0, with vertical structure = 1.

Spares

52 spare bytes for additional parameters.

4.4.5 Scan Variable Descriptions

Spacecraft latitude, Spacecraft longitude, Spacecraft altitude (km)

Satellite sub-point earth coordinate position and altitude

Scan Date/Time (7)

Time at the beginning of the scan including milliseconds. (Year, month, day, hour, min, sec, millisecs.

tai93time

International Atomic Time 1993 i.e. time in seconds since 1 January 1993

4.4.6 Pixel Data Variable Descriptions

Pixel Status – a full list of these can only be created once the algorithm is finalized.

If there is no retrieval at a given pixel, pixelStatus explains the reason.

0 : Valid pixel
1 : Pixel out of Latitude/Longitude defined area
2 : Tbs out of range
3: Surface code / histogram mismatch
4: Missing tcwv,T2m, surface code from the preprocessor
5: no Bayesian solution possible

Quality Flag

The GPROF Quality Flag variable for GPM V5 can be 0,1,2,3

Value 0: pixel is "good" and has the highest confidence of the best retrieval.

Value 1: "use with caution" . Pixels can be set to value 1 for the following reasons:

5) Sunglint is present, RFI, geolocate, warm load or other L1C 'positive value" quality warning flags
6) All sea-ice covered surfaces
7) All snow covered surfaces
8) Sensor channels are missing, but not critical ones.

Value 2: "use pixel with extreme care over snow covered surface" This is a special value for snow covered surfaces only. The pixel is set to 2 if the probability of precipitation is of poor quality or indeterminate. Use these pixels for climatological averaging of precipitation, but not for individual storm scale daily cases.

Value 3: "Use with extreme caution". Pixels are set to value 3 if they have channels missing critical to the retrieval, but the choice has been made to continue the retrieval for these pixels.

L1R Quality Flag

Within the Preprocessor, checks are made of the L1R quality flag. These quality flags have both fatal error and non-fatal warnings. These are checks and a new code 0,1,3 is passed through to GPROF. Generally:

0 = do retrieval
1 = some problems, or channels are missing
3 = retrieval has many channels missing and probably shouldn't be used.

If the L1R quality flag have a fatal error, the pixel's Tbs are set to missing.

Surface Type Index

Surface type codes are: 1 : Ocean, 2 : Sea ice, 3-7 : Decreasing vegetation, 8-11 : decreasing snow cover, 12: standing water, 13 : land/ocean or water coast, 14 : sea-ice edge

Total Column Water Vapor Index

The integer of the model total precipitable water (mm) used to select the correct database profiles. It's the nearest integer of the actual real*4 Model output value.

Probability of Precipitation

A diagnostic variable, in percent, defining the fraction of raining vs. non-raining Database profiles that make up the final solution. Values range from 0 to 100%.

2 Meter Temperature Index

The integer value of the model 2 meter temperature -used to select the correct database profiles. It's the nearest integer of the actual real*4 Model output value.

Sun Glint Angle

Conceptually, the angle between the sun and the instrument view direction as reflected off the Earth's surface. sunGlintAngle is the angular separation between the reflected satellite view vector and the sun vector. When sunGlintAngle is zero, the instrument views the center of the specular (mirror-like) sun reflection. Values range from 0 to 180 degrees. If this angle is < 10 degrees, the pixel is affected by sunglint and the pixels Quality Flag is lowered.

Latitude, Longitude

Pixel latitude and longitude.

Surface Precipitation

The instantaneous precipitation rate at the surface. Check pixelStatus for a valid retrieval. Values are in mm/hr.

Frozen Precipitation, and Convective Precipitation

The instantaneous frozen and convective precipitation at the surface. Values are in mm/hr.

Rain Water Path, Cloud Water Path, and Ice Water Path

Total integrated rain water, cloud liquid water, and ice water in the vertical atmospheric column.

5.0 KNOWN LIMITATIONS

a) There are no profiles retrieved over snow covered land surfaces due to Satellite Radiometer matching with only the MRMS (ground radar) surface precipitation. At the time no profile of hydrometeors was available.

b) Light rain and drizzle in the high latitude are still poorly observed. Though we did increase the high latitude precipitation over ocean it probably is still low.

c) Precipitation phase at high elevations is often incorrect (frozen precipitation is reported as rain). As shown in the Olympex experiment winter higher elevation precipitation had the wrong phase. The model (GEOS5) 2 meter wet bulb temperature is used as the phase discriminator, and the crude resolution (often 50km) cannot accurately capture the cold mountain temperatures.

d) Precipitation in coastal regions continue to be lower in quality. When the satellite pixel encompasses both land and water in the field-of-view (coast surface class), the microwave surface emissivity depends greatly on the relative percentages of each.

6.0 REFERENCES

Berg, W., T. L'Ecuyer, and C. Kummerow, 2006. Rainfall climate regimes: The relationship of regional TRMM rainfall biases to the environment, *J. Appl. Meteor. Climatol.*, **45**, 434–454.

Kummerow, C., W.S. Olson and L. Giglio, 1996. A simplified scheme for obtaining precipitation and vertical hydrometeor profiles from passive microwave sensors, *IEEE, Trans on Geoscience and Remote Sensing,* **34**, 1213-1232, doi: 10.1109/36.536538.

Liu, G., 2008. A Database of Microwave Single-Scattering Properties for Nonspherical Ice Particles, *Bulletin of the American Meteorological Society*, ***89***(10), 1563-1570.

Lorenc, A. C., 1986: Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **112**, 1177-1194.

Prigent, C., C., W. B. Rossow, and E. Matthews, 1997: Microwave land surface emissivities estimated from SSM/I observations. *J. Geophys. Res*., **102**, 21 867 – 21 890.

Prigent, C., E. Jaumouille, F. Chevallier, and F. Aires, 2008: A parameterization of the microwave land surface emissivity between 19 and 100 GHz, anchored to satellite-derived estimates*, IEEE Transaction on Geoscience and Remote Sensing* , **46**, 344-352, doi:10.1109/TGRS.2007.908881.

Sims, E.M. and G. Liu, 2015: A Parameterization of the Probability of Snow–Rain Transition. *J. Hydrometeor.,* **16**, 1466–1477, doi: 10.1175/JHM-D-14-0211.1.

Stephens, G. L., & C. D. Kummerow, 2007. The Remote Sensing of Clouds and Precipitation from Space: A Review, *Journal of the Atmospheric Sciences*, ***64***(11), 3742-3765.