CloudSat Project
A NASA Earth System Science Pathfinder Mission

Level 2C RAIN-PROFILE Product Process Description and Interface Control Document Algorithm version 0.0

Date: 20 June 2011

Address questions concerning the document to:

Matthew Lebsock
Lebsock@atmos.colostate.edu
970-491-8457

Approvals

_____Date_____
Matthew Lebsock
Data Product Algorithm Lead

_____Date_____
Tristan L'Ecuyer
Data Product Algorithm Lead

_____Date_____
Deborah Vane
Deputy Project Scientist

_____Date_____
Graeme Stephens
Project Scientist

_____Date_____
Donald Reinke
Data Processing Center Lead Engineer

Table of Contents

1. INTRODUCTION

This document provides an overview of the level 2C RAIN-PROFILE algorithm for CloudSat. The objective of the algorithm is to infer profiles of precipitation liquid and ice water content along with an associated surface rain rate from the CloudSat Profiling Radar (CPR) reflectivity profiles and a constraint on the Path Integrated Attenuation (PIA) of the radar beam. Key inputs to the algorithm flow from the 2C-PRECIP-COLUMN product, which flags profiles for precipitation, determines the freezing level, determines the precipitation type (convective/stratiform/shallow), and provides an estimate of the magnitude and uncertainty of the PIA. The algorithm further makes use of ancillary temperature and humidity estimates from the ECMWF analysis provided in the ECMWF-AUX product. The remainder of this document describes the algorithm in greater detail. Section 2 describes the physical basis upon which the algorithm is based, Section 3 lists the algorithm input, Section 4 outlines the algorithm flow, Section 5 describes the algorithm output, Section 6 shows an example retrieval, and Section 7 highlights some known caveats.

2. ALGORITHM THEORETICAL BASIS

The basis for this work has been outlined in L'Ecuyer and Stephens (2002), Mitrescu et al. (2010) and Lebsock an L'Ecuyer (2011) (Hereafter, LL2011). This section reviews the salient details outlined in those works.

2.1. RETRIEVAL FRAMEWORK

The algorithm follows an optimal estimation framework that seeks to minimize a cost function of the form,

$$\Phi = \left[\mathbf{Z} - \mathbf{Z}_{sim}\right]^{\mathbf{T}} \mathbf{S}_{\mathbf{y}} \left[\mathbf{Z} - \mathbf{Z}_{sim}\right] + \left[\mathbf{x} - \mathbf{x}_{\mathbf{a}}\right]^{\mathbf{T}} \mathbf{S}_{\mathbf{a}} \left[\mathbf{x} - \mathbf{x}_{\mathbf{a}}\right] + \frac{\left(\mathrm{PIA}_{sim} - \mathrm{PIA}\right)}{\sigma_{PIA}^2},$$

where $\mathbf{Z}$ represents a vector of radar reflectivities, $\mathbf{Z}_{sim}$ is a simulated vector of reflectivites, $\mathbf{x_a}$ represents an *a priori* estimate of the state ($\mathbf{x}$), $\mathbf{S_y}$ is the observation error covariance matrix, $\mathbf{S_a}$ is the *a priori* error covariance matrix, and σ_{PIA}^2 represents the estimated error variance in PIA. The error variances and covariances determine the relative influence of the four terms in determining the retrieved state. The cost function is minimized in a straightforward manner using Newtonian iteration until a solution is achieved that provides an optimal match to both the observations and *a priori* constraints given their relative error bounds. Because the CPR operates at the strongly attenuated frequency of 94 GHz, the attenuation constraint given by the PIA is central to the retrieval to avoid the propagation of errors discussed by Hitschfeld and Bordan (1954).

The true utility of the optimal estimation retrieval framework lies in careful error characterization, the details of which are buried in the process of defining the elements of the $\mathbf{S_y}$ and $\mathbf{S_a}$ error covariance matrices. The error variance matrices are used to (1) impose correlations between radar bins through the $\mathbf{S_a}$ terms and (2) impose a measurement error ($\mathbf{S_y}$) that increases with depth into the precipitation column to account for uncertainty in modeling the attenuation along the radar path. As shown in the example in Figure 1, both matrices account for correlated errors. A detailed description of the error characterization is provided in LL2011. Figure 1 shows one realization of these error matrices. Note that both matrices account for error correlation.

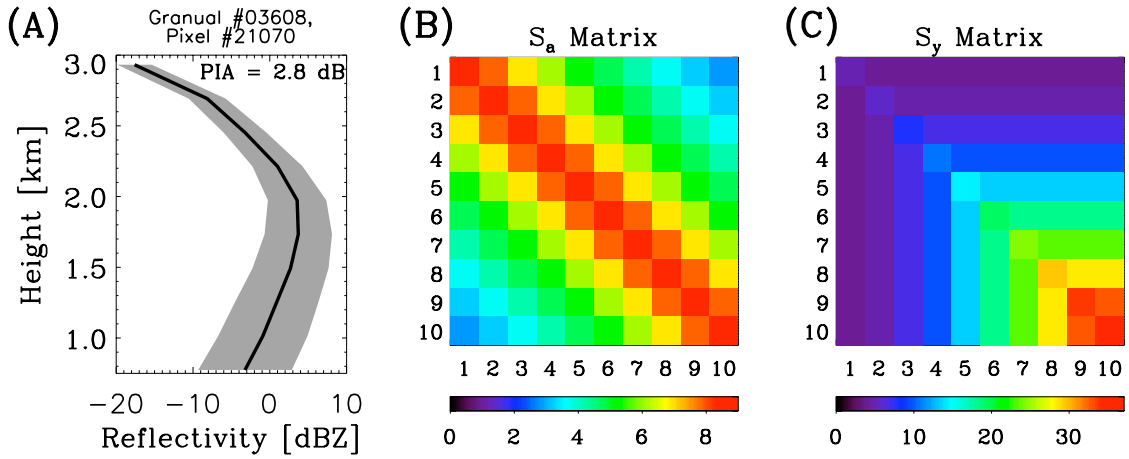

Figure 1: (A) An example reflectivity profile with gray shading indicating the estimated observational uncertainty. Also shown are the associated (B) $\mathbf{S_a}$ and (C) $\mathbf{S_y}$ error covariance matrices.

2.2. *Algorithm Implementation*

2.2.1. *Radar Model*

Simulated radar reflectivites (Z_{sim}) and Path Integrated Attenuation (PIA_{sim}) are calculated using the general expressions,

$$\begin{aligned} Z_{sim} &= Z_{ss} + \Gamma_{ms} - \Gamma_{att} \\ \mathrm{PIA}_{sim} &= \mathrm{PIA}_{ss} - \Gamma_{ms} \end{aligned} \qquad (1)$$

where Γ_{ms} represents a multiple scattering correction and Γ_{att} represents an attenuation correction both of which are defined to be greater than 0. Multiple scattering is modeled using the fast Time-Dependent Two-Stream (TDTS) method of *Hogan and Battaglia* (2008). The TDTS model has been shown to compare favorably with benchmark Monte Carlo simulations while being significantly more computationally efficient. The TDTS model is used to correct both the reflectivities and the observed PIA for multiple scattering effects. Multiple scattering always increases the apparent reflectivity therefore

these corrections always reduce the single scattering reflectivity values while increasing the PIA estimate.

2.2.2. Physical Models

The problem of estimating the rain rate from the observations is incompletely defined as posed and requires a number of simplifying assumptions. These assumptions take the form of simple physical models that are imposed upon the problem to make the necessary radiative calculations possible. These models include: (1) A model to distribute cloud water in the vertical and determine the cloud DSD; (2) A model of evaporation of rain below cloud base; (3) A model of the precipitation DSD. Descriptions of these models are described in this sub-section and (4) a description of the thermodynamic phase of the hydrometeors in each radar bin.

Cloud water must be modeled not because of its influence on the reflectivities themselves but rather due to its influence on the PIA. As a result, the location of the cloud water within the vertical profile is of second order importance. It is assumed within the algorithm that the cloud water content is uniform with height below the freezing level and the cloud water path is given by a parameterization based on LL2011.

A model of evaporation of rain water from cloud base to the surface is taken from Comstock et al. (2004). The evaporation fraction is based on the distance below cloud base and the mean radius of the drop size distribution.

Three precipitation DSDs are used. The default distribution is that of Marshall and Palmer (1948). As described in LL2011, for shallow clouds with tops below the freezing level two additional DSDs based on in-situ observations from Comstock et al (2004) and Snodgrass (2009).

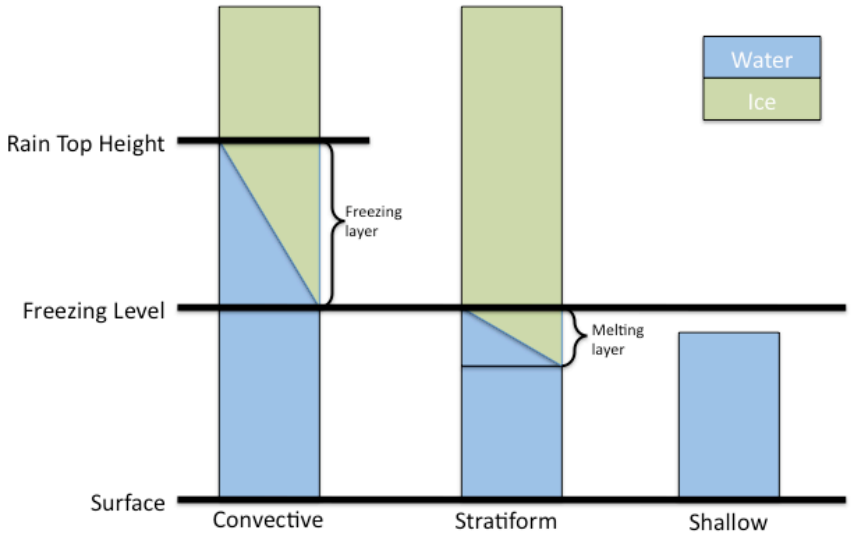

Figure 2: Description of the vertical assumption of thermodynamic phase for the various precipitation types.

The vertical structure of thermodynamic phase is based on the stratiform/convective flag input from 2C-PRECIP-COLUMN. Figure 2 illustrates the assumed vertical partitioning of thermodyanamic phase for each convective/stratiform classification. Cloud liquid water is distributed uniformly throughout the liquid portion of the profile. The optical properties of the mixed phase regions are modeled as a mixture of liquid and ice using a Maxwell-Garnett mixing formulation (Menenghini and Liao, 1996).

3. ALGORITHM INPUTS

3.1. CLOUDSAT PRODUCT INPUTS

The algorithm ingests inputs from the 2B-GEOPROF, ECMWF-AUX, and 2C-PRECIP-COLUMN products. A table of these inputs is provided in Table 1. The dimensions of the data fields are described by the variables 'nray' which is the total number of radar profiles and 'nbin', which is the number of radar range bins per profile.

Table 1: Input variables to the 2C-RAIN-PROFILE algorithm.

Input Source	*Variable Name*	*Dimensions*	*Units*
2B-GEOPROF	Latitude	nray	degrees
2B-GEOPROF	Longitude	nray	degrees
2B-GEOPROF	Height	nbin, nray	m
2B-GEOPROF	SurfaceHeightBin	nray	
2B-GEOPROF	Gaseous_Attenuation	nbin, nray	dBZe
2B-GEOPROF	Radar_Reflectivity	nbin, nray	dBZe
2B-GEOPROF	DEM_elevation	nray	m
2B-GEOPROF	Navigation_land_sea_flag	nray	
2B-GEOPROF	Sigma_zero	nray	dB*100
2B-GEOPROF	CPR_Echo_Top	nray	
2B-GEOPROF	Data_quality	nray	
ECMWF-AUX	Temperature	nbin, nray	K
ECMWF-AUX	Pressure	nbin, nray	Pa
ECMWF-AUX	Temperature_2m	nray	K
ECMWF-AUX	Specific_humidity	nbin, nray	kg/kg
2C-PRECIP-COLUMN	Precip_rate	nray	mm/hr
2C-PRECIP-COLUMN	Precip_flag	nray	
2C-PRECIP-COLUMN	Freezing_level	nray	km
2C-PRECIP-COLUMN	Status_flag	nray	
2C-PRECIP-COLUMN	Conv_strat_flag	nray	
2C-PRECIP-COLUMN	PIA_hydrometeor	nray	dB
2C-PRECIP-COLUMN	PIA_uncertainty	nray	dB
2C-PRECIP-COLUMN	Lowest_sig_layer_top	nray	km
2C-PRECIP-COLUMN	Rain_top_height	nray	km
2C-PRECIP-COLUMN	Surface_type	nray	
2C-PRECIP-COLUMN	Precip_rate_min	nray	mm/hr
2C-PRECIP-COLUMN	Precip_rate_max	nray	mm/hr
2C-PRECIP-COLUMN	Diagnostic Retrieval Info	nray	

3.2. CONTROL AND CALIBRATION

The algorithm does not require control and calibration data.

4. ALGORITHM SUMMARY

```
Read 2B-GEOPROF
Read ECMWF-AUX
Read 2C-PRECIP-COLUMN

Loop over number of profiles
        Check for valid data
        If (data = valid) then
                Check for Rainfall
                If (PIA = saturated) then
                        Pass through 2C-PRECIP-COLUMN rate
                Else
                        Assign thermodynamic phase profiles
                        Determine DSD
                        Call retrieval algorithm
                        If (retrieval = converged) Then
                                Set output
                        Else
                                Pass through 2C-PRECIP-COLUMN rate
                        End
                End
        End
End

Write 2C-RAIN-PROFILE output
```

5. OUTPUT DATA

5.1. OUTPUT DATA FORMAT

Table 2 provides a description of the structure of the 2C-RAIN-PROFILE data file description. The dimensions of the data fields are described by the variables ‘nray’ which is the total number of radar profiles and ‘nbin’, which is the number of radar range bins per profile.

Table 2: Level 2C-RAIN-PROFILE HDF-EOS Data File Structure.

Data Granual			*Variable Name*	*Dimensions*	*Units*
	Swath Data	Geolocation fields	Profile_time	nray	seconds
			UTC_start	scalar	seconds
			TAI_start	scalar	seconds
			Latitude	nray	degrees
			Longitude	nbin, nray	degrees
			Height	nray	m
			Range_to_intercept	nray	km
			DEM_elevation	nray	m
			Vertical_binsize	scalar	m
			Pitch_offset	scalar	degrees
			Roll_offset	scalar	degrees
		2B-GEOPROF pass through fields	Data_quality	nray	
			Data_status	nray	
			Data_targetID	nray	
			Navigation_land_sea_flag	nray	
		2C-RAIN-PROFILE data fields	Precip_flag	nray	
			Rain_quality_flag	nray	
			Rain_status_flag	nray	
			Rain_rate	nray	mm/hr
			Rain_rate_uncertainty	nray	%
			Modeled_PIA_hydrometeor	nray	dB
			Surface_ms_correction	nray	dB
			Precip_liquid_water	nbin, nray	g/m^3
			Precip_ice_water	nbin, nray	g/m^3
			Cloud_liquid_water	nbin, nray	g/m^3
			PWC_uncertainty	nbin, nray	%
			Modeled_reflectivity	nbin, nray	dBZ
			Attenuation_correction	nbin, nray	dB
			MS_correction	nbin, nray	dB

5.2. OUTPUT VARIABLE DESCRIPTIONS

Those variables that are generated by the 2C-RAIN-PROFILE algorithm are described below:

- **precip_flag:** Precipitation occurrence flag. This flag is determined using input from the 2C-PRECIP-COLUMN product. Only pixels that are determined to contain certain surface precipitation as reported as precipitating.

  -1 = missing data input or land surface
  0 = non precipitating (corresponds to PRECIP_COLUMN flag = [0,1,2,4, or 6])

1 = certain rain (corresponds to PRECIP_COLUMN flag = 3)
2 = certain snow/mixed precipitation (No intensity estimate made) (corresponds to PRECIP_COLUMN flag = [5 or 7])

- **rain_quality_flag:** Flag indicating the quality of the rain rate estimate. Flagging is based on the modeled multiple scattering correction, estimate of the uncertainty in the Path Integrated Attenuation (PIA) and the magnitude of the estimated PIA. Increasing values of confidence are indicative of lower uncertainty in the PIA and smaller multiple scattering effects.

  -1 = missing data input or land surface
  0 = no confidence
  1 = very low confidence
  2 = low confidence
  3 = moderate confidence
  4 = high confidence

  The following pseudo code illustrates the algorithm logic that determines the value of the rain quality flag. Here (δ_{PIA}) is the uncertainty in the PIA and (Γ_{MS}) is the surface multiple scattering correction

  If (PIA = saturated signal) Then
  Rain_quality_flag = 0
  Else If (algorithm did not converge) Then
  Rain_quality_flag = 1
  Else If ((δ_{PIA}<2.5)&(Γ_{MS}<5)) Then
  Rain_quality_flag = 4
  Else If ((δ_{PIA}<2.5)&(Γ_{MS}<10)) OR ((δ_{PIA}<5)&(Γ_{MS}<5)) Then
  Rain_quality_flag = 3
  Else If ((δ_{PIA}<2.5)&(Γ_{MS}<15)) OR ((δ_{PIA}<5)&(Γ_{MS}<10)) Then
  Rain_quality_flag = 2
  Else
  Rain_quality_flag = 1
  End

- **rain_status_flag:** Status indicating the retrieval method used for the rain intensity estimate.

  -1 = missing data input/land surface.
  0 = non-raining pixel or rain rate derived from the profile algorithm.
  1 = rain rate passed through from 2C-PRECIP-COLUMN because the profile algorithm did not converge to a valid solution.
  2 = rain retrieval not possible due to extreme attenuation. Negative rain rate is passed through from the 2C-PRECIP-COLUMN product. The absolute value of this rain rate should be interpreted as the minimum possible rain rate.

- **rain_rate:** Surface rain rate.

  NOTE 1: Negative rain rates indicate a high rain rate where the radar signal has been saturated. In this situation the absolute value of the rain rate should be interpreted as the minimum possible rain rate.

  NOTE 2: Profiles determined as snow or mixed phase report rain_rate = 0.

- **rain_rate_uncertainty:** 1-sigma uncertainty estimate in the surface rain rate.

- **modeled_PIA_hydrometeor:** The PIA from hydrometeors (cloud/rain/ice) that is modeled by the algorithm. This quantity does not include a multiple scattering correction. To compare this quantity to the observed PIA one most subtract the surface_MS_correction variable.

- **surface_MS_correction:** The multiple-scattering correction at the surface that is modeled by the algorithm. This quantity should be added to the modeled hydrometeor PIA to derive the uncorrected hydrometeor PIA.

- **precip_liquid_water:** The liquid precipitation water content.

- **precip_ice_water:** The precipitation ice water content.

- **cloud_liquid_water:** The cloud liquid water.

- **PWC_uncertainty:** 1-sigma uncertainty in the precipitation (liquid + ice) water content.

- **modeled_reflectivity:** The modeled reflectivity profile Including modeling of attenuation and multiple scattering effects.

- **attenuation_correction:** total attenuation correction (gas + hydrometeor). The retrieved single scattered reflectivity can be computed as: [Zss = modeled_reflectivity + attenuation_correction - MS_correction]

- **MS_correction:** Total multiple-scattering correction. The retrieved single scattered reflectivity can be computed as: [Zss = modeled_reflectivity + attenuation_correction - MS_correction]

6. AN EXAMPLE

Figure 3 shows an example retrieval scene composed primarily of startiform precipitation. The example shows that both the observed PIA and reflectivites are matched to within there uncertainty bounds by the retrieval. An brightband is evident below the freezing level below which a strong attenuation signal is evident. The retrieval produces a reflectivity field that corrects for both attenuation and multiple scattering demonstrating the large influence of attenuation in the observed CloudSat radar reflectivities. The regions highlighted by the black oval are areas in which the PIA signal is either saturated or the algorithm does not converge. The algorithm in it's current manifestation does not produce profile information in these situations.

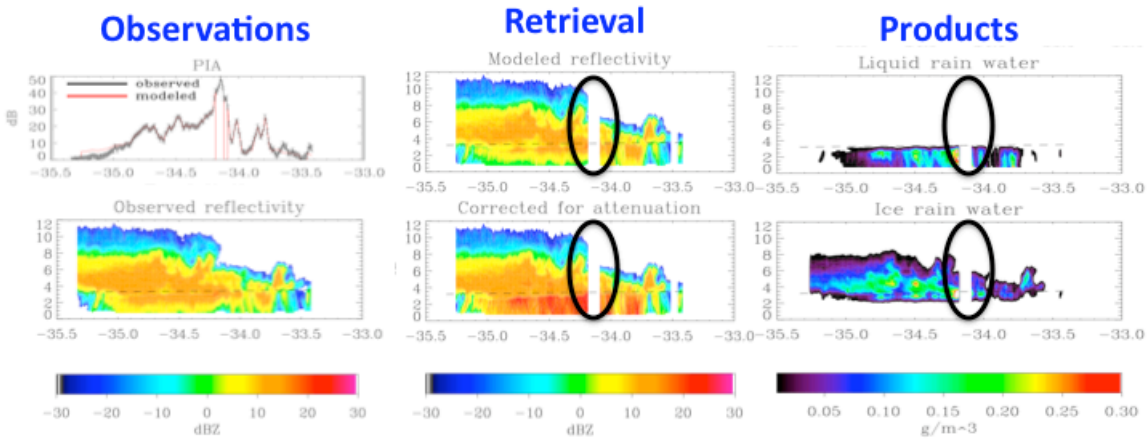

Figure 3: An example retrieval scenario.

7. CAVEATS AND KNOWN ISSUES

- *Land Surfaces:* Profiles over land surfaces are currently treated as missing pixels by the algorithm due to the inherent difficulty in estimating the PIA over land. Future versions of the algorithm may attempt to address these profiles. An estimate of precipitation incidence is provided in the 2C-PRECIP-COLUMN product.

- *Saturation of the surface return*: In the heaviest precipitation, the surface return may be completely saturated making an estimate of the surface cross section impossible. In this situation only an estimate of the lower bound on the PIA may be made. In this case the 2C-RAIN-PROFILE algorithm is not run, and a negative rain rate is passed through from the 2C-PRECIP-COLUMN product. The absolute value of this rain rate may be interpreted as a minimum possible rain rate. In this situation the rain_quality flag is set to 0 (no confidence) and the rain status flag is set to 2. To screen these cases check for rain status = 2. Care should be taken

when working with a dataset in which the fraction of pixels that meet this criteria exceeds 0.5% because this condition is set in the heaviest rainfall events that contribute the most to accumulations. Figure 4 shows a map of the frequency of occurrence of this condition.

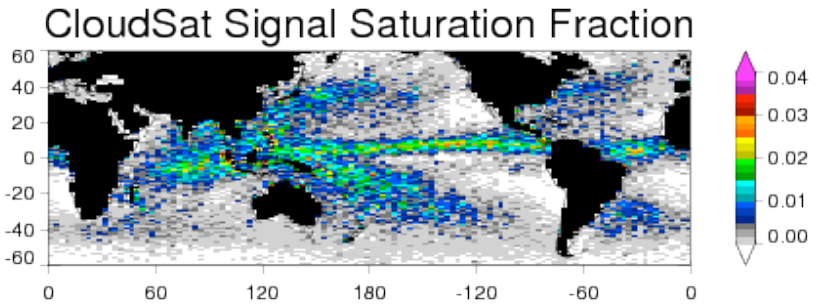

Figure 4: The fraction of pixels in which the surface signal is saturated and rain_status = 2.

- *Profiles:* Output profiles are only reported for the lowest cloud layer beginning at the first instance of that layer exceeding -15 dBZ. The product does not therefore provide a complete description of the entire profile of liquid and ice hydrometeors. Furthermore, Although an estimate of the minimum possible rain rate is provided when rain status = 2, profiles are not output when this condition is set. This results in a systematic data loss for heavy precipitation cases. Profiles are also not available when the algorithm does not converge to a valid solution (rain_status = 1)

- *Accumulations*: The global distribution of mean rain rate is provided in Figure 5 for reference. To integrate the surface rainfall estimates into accumulations:
  1. Screen pixels in which precip_flag = -1 which indicates missing data.
  2. If rain_status_flag = 2 then use the absolute value of the reported rain rate. This is a minimum possible rain rate. Care must be taken in interpreting results if a large (> 0.5% of pixels) fraction of the data meets this condition.

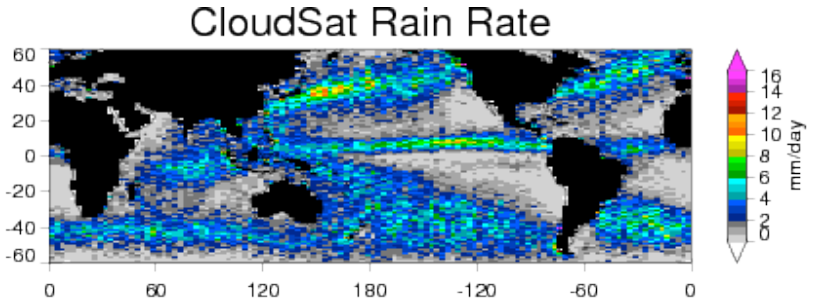

Figure 5: The global distribution of mean rain rate for the years 2007-2008 using the above screening criteria.

8. OPERATOR INSTRUCTIONS

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

9. REFERENCES

Comstock, K. K., R. Wood, S. E. Yuter, and C. Bretherton (2004), Reflectivity and rain rate in and below drizzling stratocumulus**,** *Quart. J.Roy. Meteorol. Soc., 130*, 2891-2919. DOI:10.1256/qj.03.187.

Hitschfeld, W., and J. Bordan (1954), Errors inherent in the radar measurement of rainfall at attenuating wavelengths. *J. Meteor., 11***,** 58-67.

Hogan, R. J. and A. Battaglia (2008), Fast Lidar and Radar Multiple-Scattering Models. Part II: Wide-Angle Scattering Using the Time-Dependent Two-Stream Approximation. *J. Atmos. Sci.*, *65*, 3636-3651, DOI:10.1175/2008JAS2643.1.

L'Ecuyer T.S. and G. L. Stephens (2002), An Estimation-Based Precipitation Retrieval Algorithm for Attenuating Radars, *J. Appl. Meteor., 41*, 272-285, DOI:10.1175/1520-0450.

Lebsock, M. and T.S. L'Ecuyer (submitted), The Retrieval of Warm Rain from CloudSat, J. Geophys. Res.

Marshall J. S. and W. McK. Palmer (1948), The Distribution of Raindrops with Size, *J. Atmos. Sci., 5*, 165-166, DOI: 10.1175/1520-0469.

Menenghini, R., and L. Liao (1996) Comparisons of cross sections for melting hydrometeors as derived from dielectric mixing formulas and a numerical method. *J. Appl. Meteor.*, 35, 1658–1670.

Mitrescu, C., T. L'Ecuyer, J. Haynes, S. Miller and J. Turk (2010), CloudSat Precipitation Profiling Algorithm-Model Description. *J. Appl. Meteor. Climatol.*, *49*, 991-1003, DOI:10.1175/2009JAMC2181.1.

Snodgrass, E. R., L. Di Girolamo, R. M. Rauber (2009), Precipitation Characteristics of Trade Wind Clouds during RICO Derived from Radar, Satellite, and Aircraft Measurements. *J. Appl. Meteor. Climatol.*, *48*, 464-483, DOI:10.1175/2008JAMC1946.1.