TRMM Multi-satellite Precipitation Analysis (TMPA) Monthly Precipitation

1. Intent of This Document and POC

1a) This document is intended for users who wish to compare satellite-derived precipitation estimates with climate model output in the context of the CMIP5/IPCC historical experiments. Users are not expected to be experts in satellite-derived Earth system observational data. This document summarizes essential information needed for comparing this dataset to climate model output. References are provided at the end of this document to additional information.

This NASA dataset is provided as part of an experimental activity to increase the usability of NASA satellite observational data for the modeling and model analysis communities. This particular archive of data is not a standard NASA satellite instrument product, but does represent an effort on behalf of data experts to repackage a standard product that is appropriate for routine model evaluation. The data may have been reprocessed, reformatted, or created solely for comparisons with climate model output. Community feedback to improve and validate the dataset for modeling usage is appreciated. Email comments to HQ-CLIMATE-OBS@mail.nasa.gov .

Dataset File Name (as it appears on the ESG):

obs4MIPs.NASA-GSFC.TRMM.atmos.mo

1b) Technical point of contact for this dataset:

George J. Huffman, george.j.huffman@nasa.gov

2. Data Field Description

CF variable name, units:	pr (precipitation_flux), units of kg / m^2 / s
CF variable name, units:	prSterr (precipitation_flux_standard_error), units of kg / m^2 / s
Spatial resolution:	0.25°x0.25° latitude/longitude
Temporal resolution and extent:	3-hourly snapshot (pr only) and monthly average (pr and prSterr), January 1998 – June 2011
Coverage:	latitudes 50°N – 50°S

3. Data Origin

The TMPA algorithm is used to compute the “TRMM and Other Sources” 3-hourly and monthly products, which have the identifiers “3B42” and “3B43” within the TRMM project. Within the ESG these datasets are posted with file names obs4MIPs.NASA-GSFC.TRMM.atmos.3hr and obs4MIPs.NASA-GSFC.TRMM. atmos.mo (this data set). This algorithm is designed to provide spatially-complete, consistently-calibrated 3-hourly fields of precipitation estimates for the latitude band 50°N-50°S. Data are drawn from four sources, namely TRMM precipitation radar (PR) data, passive microwave (PMW) radiances at multiple frequencies and polarizations (observed from a mixed constellation of operational and research low-earth-orbit [LEO] satellites), thermal infrared brightness temperatures (IR Tb; observed by the international constellation of geosynchronous-Earth-orbit [GEO] satellites), and surface precipitation gauge measurements. The monthly estimates contained in this data set are an optimal combination of the monthly satellite precipitation estimates and the precipitation gauge analysis (see below).

Each of the PMW data streams is processed into precipitation estimates using sensor-specific algorithms. As well, the TRMM Microwave Imager (TMI) is combined with the PR data to produce TRMM Combined Algorithm (TCI) precipitation estimates, which are then used to calibrate all the PMW estimates. The calibration is accomplished in two steps; first, climatological histogram matching is applied to each sensor type to make its precipitation record more consistent with the TMI's, with regional and seasonal dependence that varies from sensor to sensor. Then a TCI-TMI histogram matching that varies by 1°x1° gridbox and month is applied to the individual TMI-adjusted PMW precipitation estimates. The TCI is only available in the latitude band 37°N-S, but the various PW precipitation estimates are valid at higher latitudes. As a first approximation, the calibration at the limits of TCI coverage are simply used at the higher latitudes.

All of these TCI-calibrated PMW precipitation estimates are grouped into 3-hourly maps, each covering ±90 minutes from the nominal synoptic observation times (00, 03, …, 21 UTC). Where overlaps of satellite swaths occur, the TCI has the highest priority for providing the grid value, conical-scan imagers are next, and cross-track sounders have the lowest priority. For each calendar month histogram-matched calibrations of coincident IR Tb's to these merged calibrated PMW precipitation fields are computed and used to generate IR precipitation estimates for each 3-hour period. The complete 3-hourly multi-satellite precipitation estimate is composed of the calibrated PMW estimates, where available, and the IR estimates otherwise.

At the monthly time scale, all available 3-hour estimates in the month are averaged in each gridbox to generate a monthly multi-satellite field. Meanwhile, monthly accumulations for the available precipitation gauge data are analyzed and gridded by other organizations – see "surface precipitation gauges" in Section 6. Each month, the gauge analysis is used to remove large-area biases in the satellite data, then combined with the (debiased) multi-satellite data using optimal weighting by the inverse (estimated) error variances to form the TMPA monthly satellite-gauge combination, which has the TRMM product number 3B43 (ESG dataset obs4MIPs.NASA-GSFC.TRMM.atmos.mo). After that, the 3-hourly fields in a month are linearly scaled gridbox-by-gridbox so that they (approximately) sum to the gridbox's monthly value. This adjusted 3-hourly field is TRMM product number 3B42 (ESG dataset obs4MIPs.NASA-GSFC.TRMM.atmos.3hr). In the CMIP5 collection the precipitation is referred to as field pr (precipitation_flux). The formal reference for 3B42/3B43 is Huffman et al. (2007), while the detailed technical documentation (Huffman and Bolvin 2011) is posted at ftp://precip.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_doc.pdf. Version 6 of the 3B42 and 3B43 products finished with June 2011, and the new Version 7 is due around the start of 2012.

For most of the period of record essentially every grid box has a value, so sampling is not typically an issue. However, the quality of the estimates varies widely due to the heterogeneous sources of data, both for individual data sets and across years as the available constellation of sensors varies. Experience shows that quality in the monthly product is relatively unaffected by the rapid changes in input data source, since the most important factor is that they are all calibrated to a single standard. The primary sampling issue is that the alternative IR data set which is necessary for months before mid-February 2000 only covers the latitude band 40°N-40°S. Thus, coverage outside that band is only provided by the PMW data. Over land, one would expect the combination with gauge analyses to largely resolve this problem at the monthly time scale, but oceanic regions should see reduced quality. As well, the Indian Ocean sector lacked GEO-IR data before July 1998. As a partial offset, we employed Geosynchronous Orbit Environmental Satellite (GOES) Precipitation Index (GPI) data computed from LEO-IR data, but even in combination with the PMW data the sampling is reduced. The sampling for May 1998 is shown in Fig. 1. Apart from the loss of one IR image over Africa, sampling is generally complete, except for the areas discussed in the text. The minor variations in sampling outside 40°N-40°S are the result of variations in microwave overpasses, while the more substantial reductions, including along the Himalaya Mountains, north of Japan, and around the Canadian Maritime Provinces are the result of drop-outs in the microwave due to snow/ice masking.

The precipitation research group in the NASA/GSFC Mesoscale Atmospheric Processes Laboratory is responsible for technical development and maintenance for the TMPA. Data set processing and reprocessing are the responsibility of the Precipitation Processing System (PPS) at NASA/GSFC, while archival activities are supplied by the Goddard Earth Sciences Data and Information Services Center. Gerald L. Potter developed the conversion routines to CMIP-standard files.

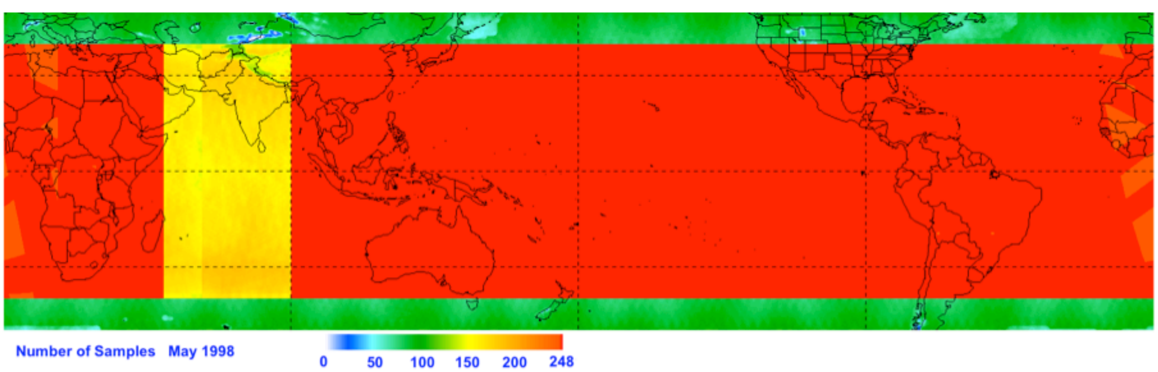

Fig. 1. Number of samples available in 3B42 Version 6 for May 1998.

4. Validation and Uncertainty Estimate

The TMPA has been subjected to a number of validation studies, primarily focusing on the fine scales. Generally, the full-resolution (3-hourly 0.25°) validation results show low skill, while averaging improves the results. It is a major result of the validations that the TMPA's use of gauge analyses is key to controlling the bias that tends to typify satellite estimates over land, and other dataset producers are starting to adopt this practice. This result extends well beyond the accidental inclusion of the same gauges in both the TMPA and the validation, but rather reflects the fact that on the monthly time scale adjacent gauges tend to be sufficiently correlated that after the first few relatively well-distributed gauges are incorporated, additional gauges will tend to simply confirm the analysis (Rudolf and Schneider 2005; Bolvin et al. 2009). Furthermore, hydrological modeling tends to show that the influence of the gauge data extends well down into the sub-monthly time scales, even while the satellite estimates completely determine the sequence of precipitation events and their relative sizes.

Over ocean there is no such gauge control, and indeed validation studies with gauge data from isolated atolls and the ATLAS II buoys tend to show a low bias in the TMPA (see, e.g., Table 1 in Huffman et al. 2007; Adler et al. 2009). Within the framework of the TMPA, we expect the monthly average to track rather closely with the final calibrator, which is the TCI over ocean and the gauge analysis over land. Looking at the global ocean average over the latitude band 30°N-30°S, we see in Fig. 2 that the 3B43 product has a negative bias in the early and mid-2000's. The GPCP monthly analysis is included as a fairly independent measure of the average. The initial sag occurs as AMSU estimates come on-line, then sags more in 2003 with an algorithm revision by the data producer, then essentially disappears in 2007 with an additional AMSU algorithm upgrade. The behavior over the central U.S. (Fig. 3, left) is typical: the TMPA follows the original GPCC Monitoring Product up to April 2005, has a negative bias with respect to the Version 2.1 GPCP monthly product, which uses the Version 4 GPCC Full Analysis over this time period. Thereafter, the TMPA follows the CAMS, which not only has a negative bias, but also shows a stronger seasonal cycle. To the extent that the newer GPCC analysis is considered superior, these differences are considered inaccurate on the whole.

The debiased month-to-month fluctuations about the true mean are termed "random error" in the TMPA and computed for each grid box for each month following Huffman (1997). This field is included in the TMPA monthly file as the field prSterr (precipitation_flux_standard_error). It is awkward to quote bulk error statistics because the random error behaves roughly like the square root of the monthly rain rate, so neither it nor the random error normalized by the mean (which approximately varies as the inverse square root of the mean) yields a nice linear statement. The sensor type also enters the picture, primarily (in the monthly) providing more accurate estimates over land than ocean. Inspecting Table 1 in Huffman et al. (2007), one seems RMS differences on the order of 32 to 75 % of the mean, depending on mean rainrate and location. It is still a matter for research to characterize the random (and bias) errors in a more complete way.

One error in the Version 6 data set is that the random error is mis-labeled "relative error". A second issue is that the lack of IR data outside the latitude band 40°N-S for 1998 to mid-February 2000 drives larger random error in those mid-latitude regions. In a few grid boxes in winter land regions large artifacts appear due to deficiencies in the microwave retrievals, including in northern Russia. Both problems will be fixed in Version 7.

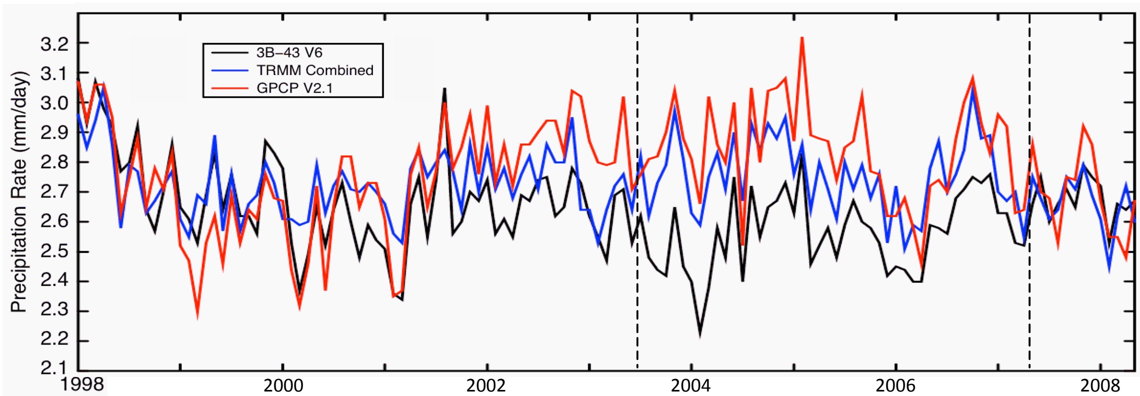

Fig. 2. Time series of precipitation averaged over ocean regions in the latitude band 30°N-30°S from 3B43 (black), TCI (blue), and GPCP monthly SG Version 2.1 (red). Dashed lines denote transitions in the AMSU input data's algorithm. Units are mm/d.

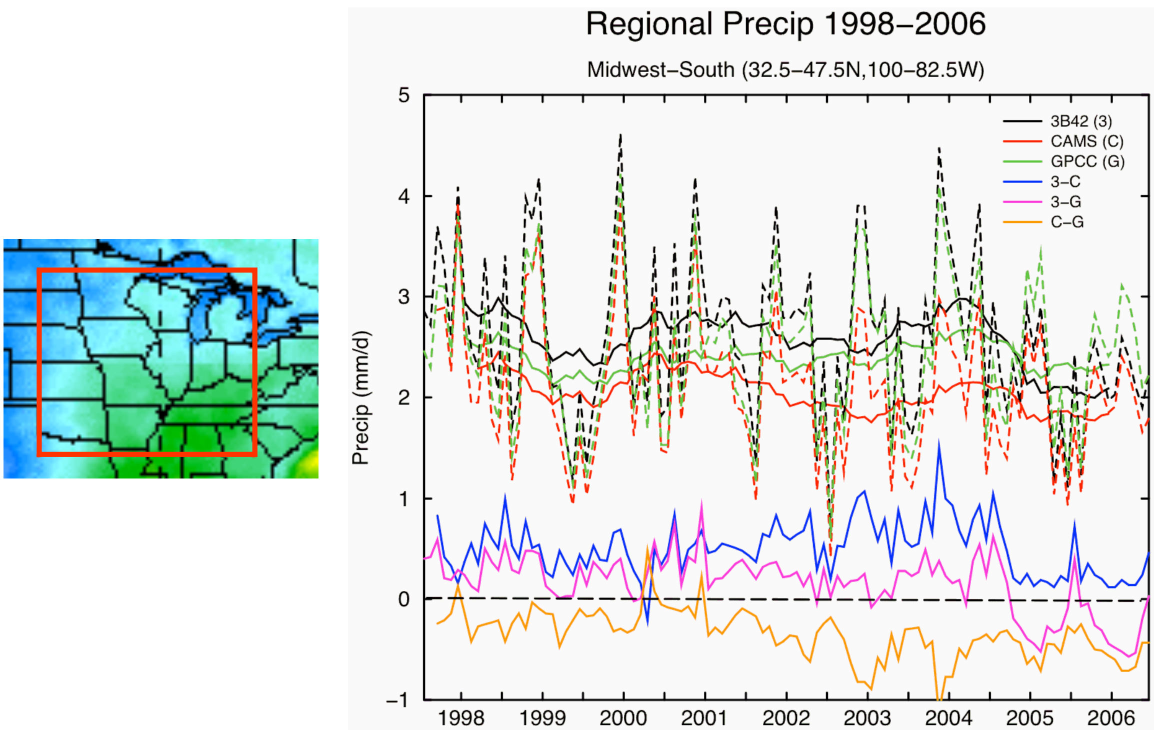

Fig. 3. (left) Map of the area used for averaging. (right) Time series of various precipitation estimates averaged over the central U.S.: 3B43 (black), CAMS (red), and GPCC Version 4 (green). The full monthly time series is dashed, while the solid lines have a running 13-month filter applied. As well, differences are show as: 3B43-CAMS (blue), 3B43-GPCC (magenta), and CAMS-GPCC (orange). All units are mm/d.

5. Considerations for Model-Observation Comparisons

Collecting the issues raised in other parts of this document:

- There is a change in bias over 4 epochs in the ocean due to AMSU algorithms, with the most negative occurring in 2003 to early 2007.
- There is a change in character for the land due to the switch from GPCC to CAMS analysis starting in May 2005. The effects vary from region to region, but tend to show a larger seasonal cycle.
- The latitude band 40°-50° in both hemispheres lacks IR data before mid-February 2000 and the Indian Ocean sector lacks GEO-IR data before July 1998, increasing the uncertainty in those regions, particularly over ocean and in winter land areas that lack gauge.
- TCI calibrations for the PMW precipitation estimates at latitudes outside the latitude band 37°N-S are approximated as the value at the north or south boundary, as appropriate.

As well, a few additional factors should be noted:

- Coastal zones present special challenges for retrievals due to the heterogeneity of the surface scene. GPROF, in particular, seems to have trouble detecting precipitation in near-coastal waters for certain weather/surface configurations, and sometimes generates artifacts in near-coastal deserts (both within about 50 km of the coastline).
- Orographic enhancement of precipitation is sometimes a challenge for the satellite schemes. The issue arises when the enhancement takes place (mostly) in the liquid phase, which current PMW algorithms cannot "see" over land.
- Current PMW schemes cannot make retrievals over snowy or frozen surfaces, which yield signals similar to frozen precipitation. This is a problem both because it denies direct use of PMW estimates in the dataset and because it denies use of the PMW estimates in the IR calibration. The TMPA contains work-arounds for such situations, but the monthly average will contain many fewer PMW estimates <u>and</u> the resulting IR estimates are of lower quality, even while they comprise the bulk of the monthly estimate. As a result, statistics over cold-season land situations should be examined for possible degradation by these snow effects.

6. Instrument Overview

The TMPA is a standard product of the TRMM project, which is a joint activity of NASA and the Japanese Aerospace Exploration Agency (JAXA). TRMM was launched in late 1997 to study tropical rainfall. The instrument complement includes three instruments focused specifically on retrieving rainfall: the Precipitation Radar (PR), the first precipitation radar in space; the TRMM Microwave Imager (TMI), a conically scanning multi-channel dual-polarization PMW radiometer; and the Visible and InfraRed Scanner (VIRS), an optical sensor providing visible and infrared imagery. Two related instruments round out TRMM's instrumentation: a Cloud and Earth Radiant Energy Sensor (CERES), an Earth radiation budget sensor that failed about six months after launch; and the Lightning Imaging Sensor (LIS), a staring imager that locates and detects lightning within individual storms.

The goal of the TMPA dataset is to use "all" available quasi-global precipitation estimates from the international constellation of precipitation-relevant satellites to create a High-Resolution Precipitation Product with complete coverage over the chosen domain and period of record (50°N-50°S, 1998-present). Fig. 4 summarizes the periods of record for the various inputs:

- PR, a phased-array Ku-band (13.8 GHz, 2.2 cm) radar on the TRMM satellite, which originally flew at 350 km, but was boosted to 401.5 km in 2001 with an orbital inclination of 35°. Primarily as a result of the inclination, TRMM precesses through the diurnal cycle in about 46 days.
- PMW radiometers in two different flavors: conical-scan imagers and cross-track-scan sounders. The former include TMI, Advanced Microwave Scanning Radiometer for Earth Observations (AMSR-E; on Aqua), and Special Sensor Microwave/Imager (SSM/I; on the Defense Meteorological Satellite Program [DMSP] series); feature multiple channels and dual polarization well-suited to estimating precipitation; provide constant footprint sizes, although these sizes differ for different channels; and are processed with sensor-specific versions of the Goddard Profiling algorithm (GPROF; Kummerow et al. 1996, Olson et al. 1999). The latter includes Advanced Microwave Sounding Unit (AMSU; on the National Oceanic and Atmospheric Administration [NOAA] series) and Microwave Humidity Sounder (MHS; on the NOAA series and Operational Meteorological Satellite A [MetOp-A]); features multiple channels relevant to precipitation; provides footprints that vary from circular at nadir to highly elliptical at the limb; and are processed with the NOAA ice water path algorithm (Zhao and Weng 2002, Weng et al. 2003). All of the PMW satellites except TRMM fly in sun-synchronous orbits at about 800 km. The periods of record used in the TMPA are sometimes shorter than the full record of the sensor, reflecting the TRMM period of record, sensor degradation (F15 SSM/I, NOAA16 AMSU), and internal restrictions on changing code within a version of the TRMM datasets (MHS).
- GEO-IR imagers, whose data are ingested as analyses from two different sources. The first 25.5 months are provided by NOAA Climate Prediction Center (CPC) as 24-class histograms of Tb's accumulated on a 1°x1° latitude/longitude grid every 3 hours over the latitude band 40°N-40°S. The dataset also provides LEO-IR-based GPI on the same grid, which must be used over the Indian Ocean sector prior to July 1998, at which time GEO-IR coverage was made available there. In mid-February 2000 CPC began providing Global Merged 4-km IR datasets on a 4-km-equivalent latitude/longitude grid every half hour over the latitude band 60°N-60°S. Recalling that the minimum international agreement for data exchange is a full-disk image every three hours, the images on the major synoptic hours (00, 03, …, 21 UTC) usually have nearly complete coverage, and these are the images used in the TMPA. The remaining images not used by the TMPA have highly variable coverage.
- Surface precipitation gauges, whose measurements are ingested as analyses from two different sources. Up through April 2005 the then-current Global Precipitation Climatology Centre (GPCC) Monitoring Analysis was used. The original gauge data, mostly gleaned from the Global Telecommunications System (GTS) as received by the Deutscher Wetterdienst, are extensively quality-controlled and gridded to 1°x1° latitude/longitude with spherical-coordinate distance- and direction-weighted

interpolation. Thereafter, the NOAA/CPC Climate Analysis and Monitoring System (CAMS) analysis is used. It is based on GTS data received by NOAA, subjected to an automated quality control, and gridded to 0.5°x0.5° with the same interpolation scheme. CAMS was used for timeliness, since it is produced about two weeks after the month, versus 10 weeks for the GPCC. However, as noted above, the lower level of quality control seems to introduce additional errors.

The TMPA technical document (Huffman and Bolvin 2011) provides expanded summaries for each sensor and references to relevant documentation.

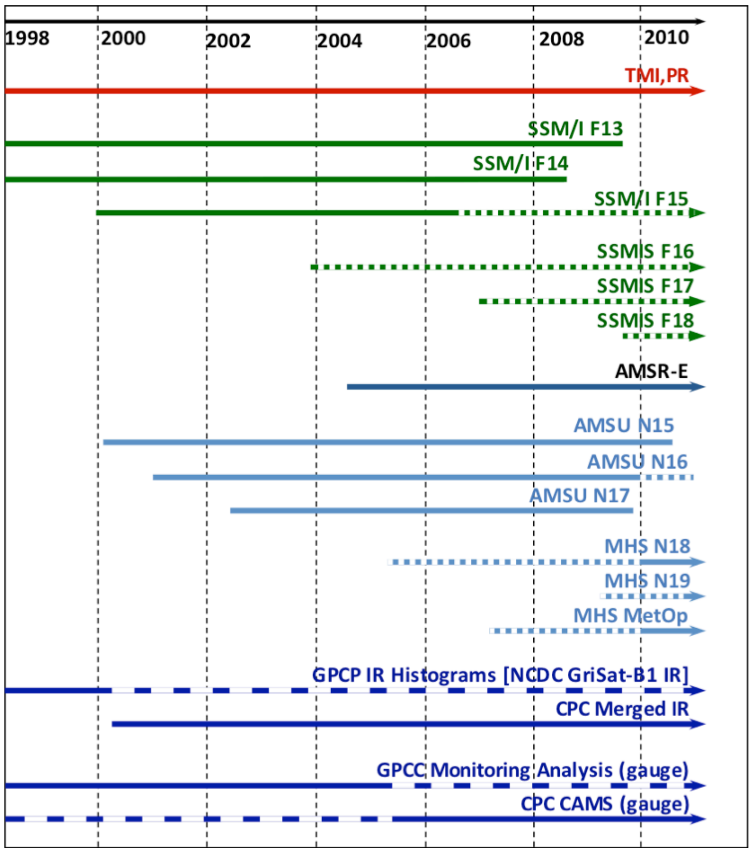

Fig. 4. Periods of record for the various data sets used in computing the TMPA (solid lines). Some of these sensors' periods of record extend beyond these periods of use, shown in dashed.

7. References

The International Polar Year (IPY) Data policy guidelines (http://.ipydis.org/data/citations.html) suggest a formal reference for data sets of the form

> Huffman, G.J., R.F. Adler, D.T. Bolvin, E.J. Nelkin, 2004, last updated 2011: *TRMM Version 6 3B42 and 3B43 Data Sets*. NASA/GSFC, Greenbelt, MD. Data set accessed <date> at http://mirador.gsfc.nasa.gov/cgi-bin/mirador/presentNavigation.pl?tree=project

&dataset=3B43:%20Monthly%200.25%20x%200.25%20degree%20merged%20TRMM%20and%20other%20sources%20estimates&project=TRMM&dataGroup=Gridded&version=006&CGISESSID=d12193016851737673840d0848b43592.

As an "Acknowledgment", one possible wording is: "The TMPA data were provided by the NASA/Goddard Space Flight Center's Laboratory for Atmospheres and PPS, which develop and compute the TMPA as a contribution to TRMM."

Additional details: At frequencies below about 37 GHz the radiative transfer signal in PMW sensor channels is primarily a combination of emission from the surface and then from the overlying atmosphere, including cloud and precipitation liquid water. At higher frequencies the useful signal results from scattering of the upwelling radiant energy out of the line of sight. Unfortunately, the land surface is radiometrically emissive and heterogeneous, so current-generation algorithms, including GPROF, can only use the emission channels over ocean. The restriction to frozen hydrometeors alone over land is an issue because they only represent the upper reaches of clouds, while the liquid phase tells about precipitation nearer the surface. Thus, conical-scan radiometers, which span both radiometric regimes, provide better answers over ocean than land. This is also the basis for the issues with retrievals over snowy/frozen surfaces and when orographic enhancement is in the liquid phase. Cross-track scanners largely depend on the scattering channels over both land and ocean, so they are less accurate than conical-scan imagers over ocean. As well, the priority for populating the PMW data field is also affected by resolution; typically the imagers have resolutions at or below 12 km, uniformly across the entire swath, while the sounders tend to start at circular footprints of size 16 km at nadir, but then stretching to 25x50 km at the limb.

Data source:

http://mirador.gsfc.nasa.gov/cgi-bin/mirador/presentNavigation.pl?tree=project&dataset=3B43:%20Monthly%200.25%20x%200.25%20degree%20merged%20TRMM%20and%20other%20sources%20estimates&project=TRMM&dataGroup=Gridded&version=006&CGISESSID=d12193016851737673840d0848b43592

Adler, R.F., J.-J. Wang, G. Gu, G.J. Huffman, 2009: A Ten-year Tropical Rainfall Climatology Based on a Composite of TRMM Products. *J. Meteor. Soc. Japan*, **87A**, 281-293.

Bolvin, D.T., R.F. Adler, G.J. Huffman, E.J. Nelkin, J.P. Poutiainen, 2009: Comparison of GPCP Monthly and Daily Precipitation Estimates with High-Latitude Gauge Observations. *J. Appl. Meteor. Climatol.,* **48**(9), 1843–1857.

Huffman, G.J., 1997: Estimates of Root-Mean–Square Random Error for Finite Samples of Estimated Precipitation. *J. Appl. Meteor.*, **36**, 1191-1201.

Huffman, G.J., R.F. Adler, D.T. Bolvin, G. Gu, E.J. Nelkin, K.P. Bowman, E.F. Stocker, D.B. Wolff, 2007: The TRMM Multi-satellite Precipitation Analysis: Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scale. *J. Hydrometeor.*, **8**, 33-55.

Huffman, G.J., D.T. Bolvin, 2011: Version 6 TRMM and Other Data Precipitation Data Set Documentation. ftp://precip.gsfc.nasa.gov/pub/trmmdocs/3B42_3B43_V6_doc.pdf, 29 pp.

Kummerow, C., W.S. Olson, L. Giglio, 1996: A Simplified Scheme for Obtaining Precipitation and Vertical Hydrometeor Profiles from Passive Microwave Sensors. *IEEE Trans. Geosci. Remote Sens.,* **34,** 1213–1232.

Olson, W.S., C.D. Kummerow, Y. Hong, W.-K. Tao, 1999: Atmospheric Latent Heating Distributions in the Tropics Derived from satellite Passive Microwave Radiometer Measurements. *J. Appl. Meteor.,* **38,** 633–664.

Rudolf, B., and U. Schneider, 2005: Calculation of Gridded Precipitation Data for the Global Land-Surface Using In-Situ Gauge Observations. *Proc. Second Workshop of the Internat. Precip. Working Group*, Monterey, CA, EUMETSAT, 231-247.

Weng, F., L. Zhao, R. Ferraro, G. Poe, X. Li, N. Grody, 2003: Advanced Microwave Sounding Unit Cloud and Precipitation Algorithms. *Radio Sci.,* **38,** 8068–8079.

Zhao, L., and F. Weng, 2002: Retrieval of Ice Cloud Parameters Using the Advanced Microwave Sounding Unit. *J. Appl. Meteor.,* **41,** 384–395.

8. Revision History

Rev 0 – 12/16/2011 - This is a new document/dataset [G.J. Huffman]