National Aeronautics and Space Administration

Algorithm Theoretical Basis Document (ATBD) for Global Precipitation Climatology Project Version 3.1 Precipitation Data

Prepared by:

George J. Huffman
NASA/GSFC
NASA/GSFC Code 612
Greenbelt, MD 20771

and Robert F. Adler, Ali Behrangi, David T. Bolvin, Eric J. Nelkin, Yang Song

Last Revised March 10, 2021

Revision History

Revision Date	*Changes*	*Author*
March 14, 2019	Original	George Huffman
July 29, 2020	Version 3.1	George Huffman
December 10, 2020	Update MCTG Reference	George Huffman
March 10, 2021	Add probability of liquid/solid formula	George Huffman

Table of Contents

Contents

Tables

Figures

1.0 Introduction

Precipitation observations are critical to many applications including drought monitoring, flash floods, crop forecasting, disease prediction, and ocean salinity studies. Surface precipitation gauges (hereafter simply "gauges") are the primary source of direct precipitation observations. Unfortunately, these gauges are point measurements and much of the globe is sparsely covered, especially in underdeveloped countries and areas of low population density. Furthermore, with the exception of a few buoy arrays, there are no precipitation gauge observations over the open ocean. Satellites seek to mitigate the limitations of rain gauge observations by estimating precipitation over land and ocean for most, or all, of the entire globe. When converted to gridded precipitation estimates, the satellite observations facilitate a multitude of studies, including those on the larger space-time scales that gauge analyses typically cannot provide. To augment the satellite-based precipitation estimates, uniformly processed gauge analyses are incorporated to improve the land-based estimates.

The Global Precipitation Climatology Project (GPCP) is a community-based activity supported by the Global Water and Energy Exchange (GEWEX) project of the World Climate Research Programme (WCRP), focused on creating a global, long-term homogeneous record of gridded precipitation estimates and ancillary information for use in climate studies and other applications. GPCP Version 3 is the successor to the highly successful GPCP V2 data set. The current GPCP V3.1 products are at the monthly resolution with the daily and 3-hourly products to be developed next.

1.1 Purpose

The purpose of this document is to describe the algorithm used to create the Global Precipitation Climatology Project (GPCP) Version 3 Monthly Satellite-Gauge (SG) analysis. It was developed primarily under a NASA Making Earth Science Data Records for Use in Research Environments (MEaSUREs) program (PI: George Huffman, NASA GSFC) and is now upgraded under a current NASA MEaSUREs project (PI: Ali Behrangi, University of Arizona). The GPCP is computed as a contribution to the World Climate Research Program (WCRP) and Global Water and Energy Exchange (GEWEX) activities, and is part of the array of data sets describing the water and energy cycles of the planet. Prior versions of the GPCP analysis have been produced by a consortium of individual scientists at various government and university institutions and most recently as part of the NOAA Climate Data Record (CDR) Program. The current GPCP Monthly SG product described here blends precipitation estimates from polar-orbit passive microwave (PMW) satellites (SSMI, SSMIS), polar orbit infrared (IR) sounders (TOVS, AIRS), and geostationary IR satellites (GOES, MeteoSat, GMS, MTSat, and Himawari); and then combines in Global Precipitation Climatology Centre (GPCC) precipitation gauge analyses, the Tropical

Combined Climatology (TCC); and the Merged CloudSat, TRMM, and GPM (MCTG) climatology products. The intent here is to provide a guide to understanding the algorithm from a scientific perspective.

1.2 Definitions

Symbols and acronyms used in the document are defined and summarized in Appendix A.

1.3 Document Maintenance

This document describes the Version 3.1 GPCP Monthly SG product.

1.4 What's New?

V3.1 is the follow-on version of the beta GPCP Version 3.0 data set. This release integrates many new enhancements and is an improvement over Version 3.0 beta, which had known limitations. The team continues to work toward improving the processing for future releases. Version 3.1 also introduces two new data fields – "gauge relative weighting" and "quality index".

2.0 Observing Systems Overview

2.1 Products Generated

This document describes the GPCP Monthly SG dataset. The primary output of this algorithm is monthly precipitation starting in January 1983 on a 0.5°, globally complete grid obtained by merging precipitation observations from satellites and gauges. The primary output product is an analysis of surface precipitation based on a systematic merger of both polar-orbit and geosynchronous satellite-based precipitation estimates with surface precipitation gauge information. The prior GPCP monthly analysis procedures are described in Huffman et al. (1997), Adler et al. (2003), Huffman et al. (2009), and Adler et al. (2016). In addition to the monthly precipitation analysis, the product provides fields of estimated random errors, precipitation phase, and several intermediate data fields.

2.2 Instrument Characteristics

The GPCP Monthly SG precipitation product is based on data from polar orbiting satellites, geostationary satellites, and gauges. The actual data used are described in section 3.3.1. The following gives information on satellite sensor characteristics that are relevant.

SSMI

The Special Sensor Microwave/Imager (SSMI) is a multi-channel PMW radiometer that flew on selected Defense Meteorological Satellite Program (DMSP) platforms beginning in mid-1987. The DMSP is placed in a sun-synchronous polar orbit with a period of about 102 minutes. The SSMI provides vertical and horizontal polarization values for 19, 22, 37, and 85.5 GHz frequencies (except only vertical at 22) with conical scanning. Pixels and scans are spaced approximately 25 km apart at the suborbital point, except the 85.5-GHz channels are collected at approximately 12.5 km spacing. The channels have resolutions that vary from 12.5x15 km for the 85.5 GHz (oval due to the slanted viewing angle) to 60x75 km for the 19 GHz.

The polar orbit provides nominal coverage over the latitudes 85°N-S, although limitations in retrieval techniques currently prevent useful precipitation estimates in cases of ice- or snow-covered land and ocean. Further details are available in Hollinger et al. (1987, 1990).

SSMIS

The Special Sensor Microwave Imager/Sounder (SSMIS) is a multi-channel PMW radiometer that has flown on selected DMSP platforms since late 2003 as a follow-on to the SSMI instrument. The DMSP is placed in a sun-synchronous polar orbit with a period of about 102 minutes. The SSMIS provides vertical and horizontal polarization values for the SSMI-like 19, 22, 37, and 91 GHz frequencies (except only vertical at 22) with conical scanning, as well as other channels with a heritage in the Special Sensor Microwave/Temperature 2 (SSMT2) sensor.

The SSMI-like frequencies use three separate feed horns: one for the 91 GHz channels, another for the 37 GHz channels and a third for the 19 and 22 GHz channels. This means that there is not a 1:1 co-location of channel values, as there is for SSMI. The SSMI-like channels have the resolutions 46.5x73.6 km (19, 22 GHz) 31.2x45.0 km (37 GHz) 13.2x15.5 km (91 GHz) with the slanted viewing angle and in-line processing determining the oval shape.

The polar orbit provides nominal coverage over the latitudes 85°N-S, although limitations in retrieval techniques prevent useful precipitation estimates in cases of ice- or snow-covered land and ocean. Further details are available in Northrup Grumman (2002) and at *https://www.wmo-sat.info/oscar/instruments/view/536.*

TOVS

The Television InfraRed Operational Satellite (TIROS) Operational Vertical Sounder (TOVS) dataset of surface and atmospheric parameters is derived from analysis of High-Resolution Infrared Sounder 2 (HIRS2) and Microwave Sounding Unit (MSU) data aboard the NOAA series of polar-orbiting operational meteorological satellites. The precipitation estimates from TOVS are derived as a secondary product utilizing various retrieved sounding parameters, including atmospheric temperature and water vapor profiles, cloud-top pressure, and radiatively effective fractional cloud cover.

For the period January 1979 - February 1999 (used January 1983 – February 1999), the TOVS estimates are based on two NOAA satellites. Beginning in March 1999, the TOVS estimates are based on a single NOAA satellite due to the failure of NOAA-11. More information can be found in Susskind et al. (1997).

AIRS

The Atmospheric Infrared Sounder (AIRS) aboard the NASA Aqua polar-orbiting satellite is the source of precipitation estimates that succeeded TOVS (in 2002, although TOVS continued to produce successively more-degraded observations into 2005). The precipitation estimates from AIRS are derived in a very similar way to those from TOVS as a secondary product utilizing various retrieved sounding parameters, including atmospheric temperature and water vapor profiles, cloud-top pressure and radiatively effective fractional cloud cover (Susskind 1997). Because the Advanced Microwave Sounding Unit (AMSU) aboard Aqua failed in late September 2016, GPCP Version 3.1 uses AIRS precipitation values for the entire record that were reprocessed and are based on infrared-only data for the sake of homogeneity.

IR data from geosynchronous satellites

The Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks – Climate Data Record (PERSIANN-CDR; Ashouri et al., 2015) is used to convert the GridSat geo-IR Tb (Knapp 2008) to precipitation rates on a 0.25°x0.25° latitude/longitude grid every 3 hours.

Note that prior versions of the GPCP SG Monthly product were computed using Outgoing Longwave Radiation (OLR) Precipitation Index (OPI) for 1979-1985, because the summaries of geo-IR data available at that time did not extend before 1986. However, the GridSat archive extends back to 1980. We choose to begin the GPCP V3.1 in 1983, which is the first full year of essentially complete geo-IR coverage, and thus avoid the more-approximate OPI data for the prior years. Basically, we drop four years of data to achieve a more homogeneous record.

3.0 Algorithm Description

3.1 Algorithm Overview

The algorithm to produce the monthly 0.5° GPCP product takes inputs from several different sources and merges them to create the most consistent and accurate monthly precipitation estimates. This document describes the high-level procedures, inputs, and outputs, of the GPCP V3.1 Monthly analysis. Many parts of the GPCP V3.1 techniques are built upon those used in GPCP V2, the details of which can be found in Adler et al. (2003) and the other referenced GPCP papers. The SSMI and SSMIS inputs are satellite radiances that need to be converted to precipitation estimates, while PERSIANN-CDR, TOVS, and AIRS are satellite-based precipitation estimates computed outside the merger process. The gauge-based monthly analysis from the

Global Precipitation Climatology Centre (GPCC) of the Deutscher Wetterdienst (DWD) is also an input, as well as is the Tropical Combined Climatology (TCC) and the Merged CloudSat, Tropical Rainfall Measuring Mission (TRMM), and Global Precipitation Measurement (GPM) Climatology (MCTG). All these data are merged together via a hierarchical analysis procedure meant to produce the most complete analysis with the lowest bias. The analysis procedure systematically uses relatively sparse, higher quality estimates to adjust lower quality estimates that are more frequently observed to maximize the sampling, but minimize any bias errors.

3.2 Processing Outline

Figures 1a and 1b show the steps required for GPCP processing during the SSMI/SSMIS and pre-SSMI periods, respectively. Over the period of the GPCP V3.1 (1983 to present) there have been several major upgrades in the available data that can be used as inputs. These changes in satellite availability lead to two distinct periods that have some differing inputs and necessarily have different processing requirements. Figure 1 shows the processing steps required for each epoch of GPCP processing. The first epoch stretches from 1983 until 1991 during which time the PERSIANN-CDR is the main satellite input to GPCP in the tropics and subtropics. The second epoch stretches from 1992 to the present, during which time SSMI and SSMIS data are available to routinely calibrate the PERSIANN-CDR in the tropics and subtropics.

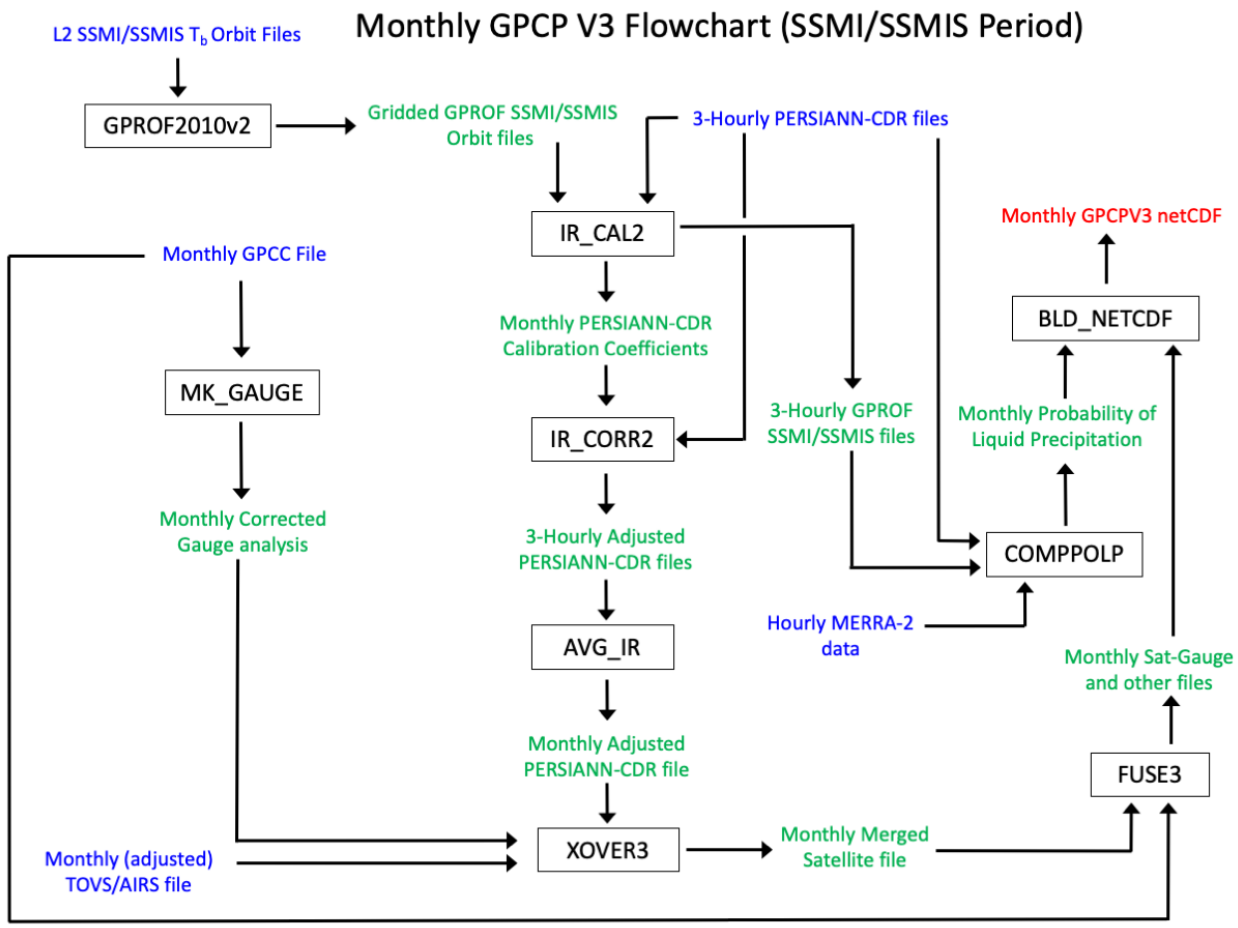

Figure 1a: Processing steps required to produce GPCP for the SSMI/SSMIS Period. Names in boxes approximate names used by code.

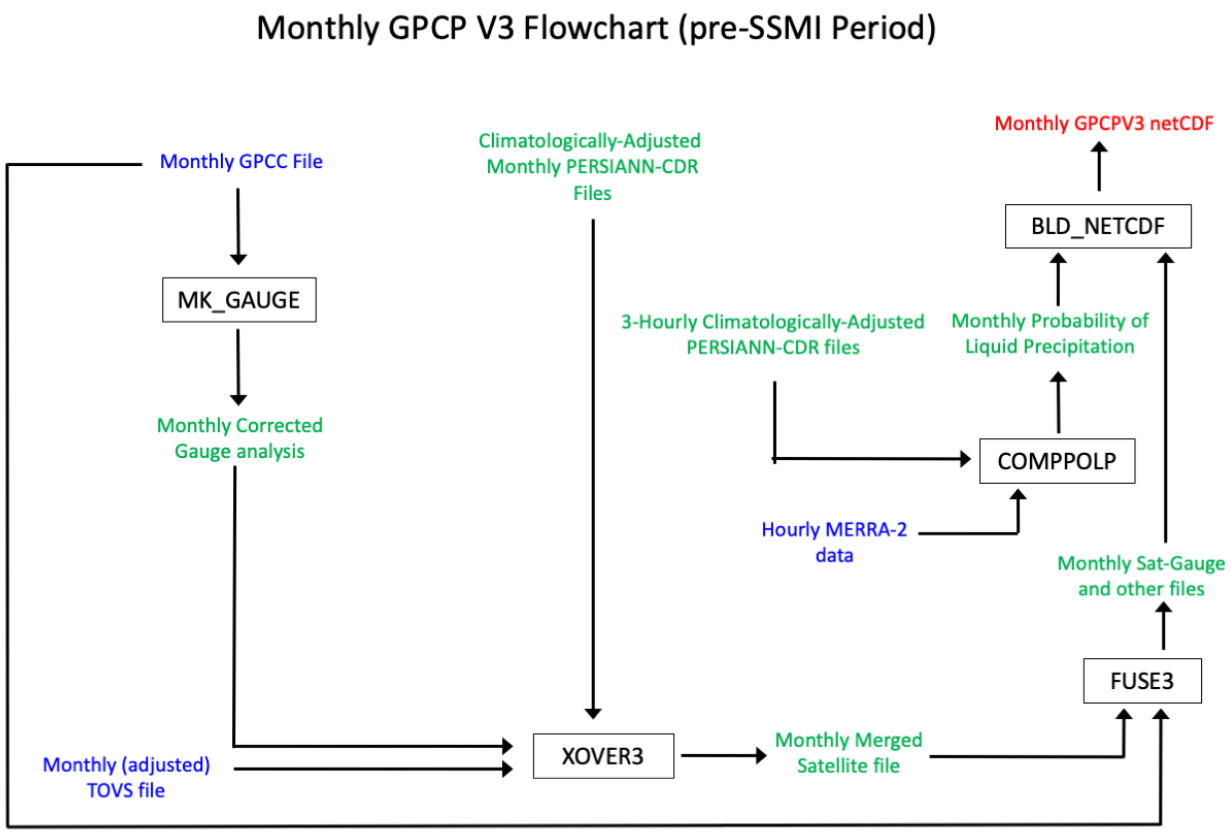

Figure 1b: Processing steps required to produce GPCP for the pre-SSMI Period. Names in boxes approximate names used by code.

3.3 Algorithm Input

3.3.1 Primary Sensor Data

Only the precipitation estimates generated from the SSMI and SSMIS are produced as part of the processing code, and the algorithm to do this is described in section 3.4.1.

Although SSMI data began in July 1987, the first sensor aboard the Defense Meteorological Satellite Program (DMSP) F08 platform had significant problems that prevent the available versions of the Goddard Profiling (GPROF) retrieval algorithm from providing useful data for much of its record. Thus, for the sake of homogeneity we choose to begin the SSMI/SSMIS era with 1992 the first full year of data from the SSMI aboard the DMSP F11, and run to the present, providing a nearly homogeneous set of input data and analysis techniques. The precipitation estimates provided by the SSMI/SSMIS are used to adjust the PERSIANN-CDR estimates month-by-month over both ocean and land. The TOVS/AIRS estimates provide the information at higher latitudes with a smooth transition from the microwave-driven lower latitude estimates to the gauge information-driven at higher latitudes. For the early period, before 1992, the PERSIANN-CDR estimates use a monthly climatological adjustment to the

microwave estimates. Thus, the GPCP Monthly period is extended back to 1983 using a cross-calibration overlap.

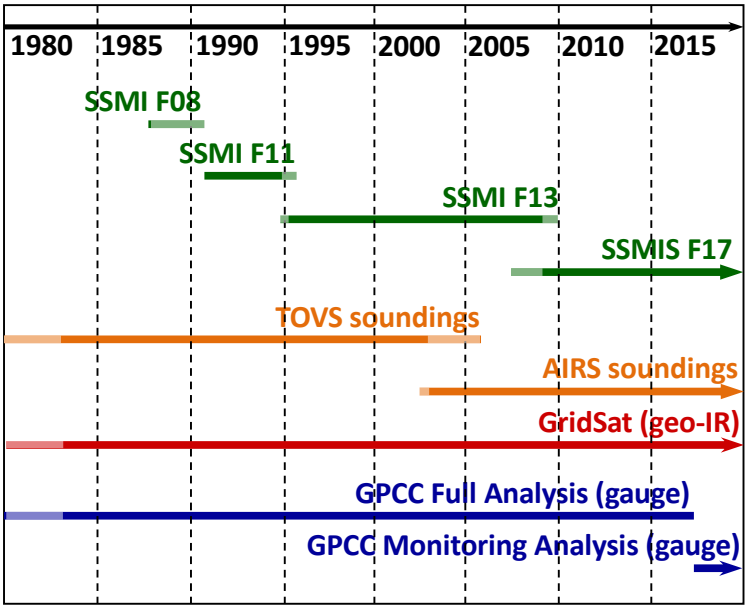

Figure 2: Inputs to GPCP by time. Lighter line segments show data periods not used in V3.1 processing for the various inputs.

Table 1: Inputs to GPCP. The current satellite sensors are highlighted in blue.

Name	Data type	GPCP time period	Data source	Notes
GPCC V2018 Full Data Reanalysis	Precip	Jan 1983 to Dec 2016	DWD (GPCC)	Climate quality 1° gridded gauge data reprocessed regularly
GPCC V6 Monitoring Product	Precip	Jan 2017 to present	DWD (GPCC)	Interim version of Full analysis intended to continue record
RSS SSMI L2 Tb	Tb	Jan 1992 to Dec 2008	RSS	F11: Jan 1992 to May 1995; F13: May 1995 to Dec 2008
RSS SSMIS L2 Tb	Tb	Jan 2009 to present	RSS	Replaces SSMI Tb; data only from F17
TOVS	Precip	Jan 1983 to Aug 2002	GSFC/SRT	
AIRS V6	Precip	Sep 2002 to present	GSFC DISC	Replaces TOVS
PERSIANN-CDR IR 3-hourly files	Precip	Jan 1983 to present	U.C.-Irvine	
TCC	Ratios	Jan 1983 to present	U. Md.	Adjusts the tropics to the TCC
MCTG	Ratios	Jan 1983 to present	U. Az.	Adjusts the mid- and high latitudes to the MCTG

In addition to the two epochs, there are several other times in which the exact inputs changed (or the version number changed). In order to be exact for the purposes of this section, these different datasets are discussed as separate datasets although they may have a single source data. An example of such a situation is the TOVS data.

RSS SSMI Tb

The Special Sensor Microwave/Imager (SSMI) was a seven channel, conically scanning microwave radiometer that flew aboard the DMSP polar orbiting satellites. The SSMI possessed four different frequencies (19, 22, 37 and 85 GHz), three of which had vertical and horizontally polarized channels (there was no 22H channel). GPCP makes use of SSMI data from the F11 and F13 satellites. SSMI Tb are obtained from Remote Sensing Systems (RSS; Santa Rosa, CA), and GPCP uses the consistently processed V7 RSS SSMI Tb. These data have been extensively quality controlled by RSS and have been intercalibrated using the methods described by Wentz (1997).

RSS SSMIS Tb

The Special Sensor Microwave Imager/Sounder (SSMIS) is a 24 channel, conically scanning microwave radiometer that replaced the SSMI and flies aboard DMSP polar orbiting satellites. Only the 7 SSMI-like channels are used for GPCP, although the 85GHz channel on SSMI was replaced with a 91GHz channel. At present, only the SSMIS data from F17 are used in GPCP. SSMIS data are also obtained from RSS and the Tb are produced with the RSS V7 calibration.

TOVS precipitation

The TOVS instrument flew aboard the NOAA series of polar-orbiting platforms. Susskind and Pfaendtner (1989) and Susskind et al. (1997) described the process for estimating precipitation from TOVS. The TOVS precipitation estimates infer precipitation from deep, extensive clouds. The technique begins with a first guess driven by a simple global numerical analysis, then uses a climatological multiple regression relationship between collocated First Global Atmospheric Research Program (GARP) Global Experiment (FGGE) precipitation gauge measurements and several TOVS-based parameters that relate to cloud volume: cloud-top pressure, fractional cloud cover, and relative humidity profile. This relationship is allowed to vary seasonally and latitudinally. Furthermore, separate relationships are developed for ocean and land.

The TOVS data are used for the early SSMI period 1983-August 2002 and are provided as monthly 1° gridded estimates for January 1983 – September 1996, then as daily 1° gridded estimates for October 1996 – August 2002. The data covering the span January 1983 - February 1999 are based on information from two satellites. For the period March 1999 - August 2002, the TOVS estimates are based on information from one satellite due to changes in satellite data format. TOVS data were obtained directly from the NASA Goddard Sensor Research Team (SRT), led by Joel Susskind.

AIRS precipitation

The AIRS instrument is flying aboard the Earth Observing System Aqua polar-orbiting satellite and has been used to succeed TOVS data. The same algorithm applied to TOVS to produce precipitation is also applied to AIRS (Susskind et al. 2003), except the first guess is a spatially and seasonally varying climatology. The AIRS data are also provided in daily 1° gridded format. AIRS data are available from September 2002 and used for the period from September 2002 - present. [The TOVS data continued to May 2005, but are considered questionable after early 2003.]

To adjust the TOVS estimates to AIRS to maintain a homogeneous record, the TOVS data were first "zoomed" from 1° resolution to 0.5° via grid box replication. A histogram-matching approach was then used to calibrate the TOVS estimates to AIRS. This was done in two parallel steps, one for the portion of the TOVS record from January 1983 through November 1999, and one for the portion from December 1999 through August 2002. The December 1999 boundary was chosen because the original, unadjusted TOVS frequency of precipitation over 70°-90°N ocean was found to be much lower from that month forward than the earlier part of the record. Note that the pre-December 1999 TOVS period comprises the entire initial monthly TOVS and 2-satellite daily TOVS records, and the first nine months of 1-satellite TOVS daily data. [It appears that this recently-discovered shift is the result of changes to how overlapping

orbit segments were combined into daily data in the original TOVS processing, but it is not currently possible to verify this hypothesis.]

For the later TOVS record, histograms were developed using a 24-month period of daily data. We chose to create the calibration using independent data due to the very limited overlap period between TOVS and AIRS (September-December 2002). Priority for the 24 months was assigned to selecting two of each calendar month, then to the more ENSO-neutral conditions among the available choices. Since the resulting histogram relationships are applied from December 1999 forward, which falls entirely within the TOVS 1-satellite period (that began in March 1999), the 24 months of TOVS used to build the histograms were chosen from the 1-satellite period. These months are July 2000-June 2002. The 24 AIRS months of daily data consisted of September 2012-August 2014, except March 2015 replaced March 2014 due to six missing days in the latter.

One global land histogram was generated for those grid boxes containing 0-15% ocean. One global "mixed" histogram was generated for grid boxes containing 15-75% ocean. For grid boxes with 75+% ocean, a set of 34 histograms were generated, in 15° latitude bands with 10-degree overlap (i.e., 90°-75°N, 85°-70°N, …, 75°-90°S). The resulting adjustment was then applied to the period December 1999-August 2002, and these daily adjusted TOVS files were accumulated to monthly.

For the earlier TOVS record, a similar approach was used. However, here the histogram development was generated by calibrating monthly TOVS data from January 1983-September 1996 against an equal number of months of AIRS data spanning January 2003-September 2016. The resulting adjustment was then applied to the period January 1983-November 1999. Note that this covers the entire TOVS monthly record as well as a portion of the TOVS daily record, which began in September 1996.

As the final step, the entire monthly TOVS/AIRS record is scaled by the monthly climatological MCTG ratios, capped at a maximum value of 2, to obtain the best estimate of precipitation using TRMM, GPM, and CloudSat information. This is especially important at latitudes higher than 60°N-S where TOVS/AIRS is the only source of satellite precipitation estimates and the CloudSat calibration dominates. The MCTG is described later in this section.

PERSIANN-CDR precipitation

The PERSIANN-CDR is summarized in Section 3.4.1.

The PERSIANN-CDR is provided as a monthly product available from NCEI. The "background" 3-hour PERSIANN-CDR products were provided by UC-Irvine by request to enable grid box calibration by the SSMI/SSMIS estimates. Because SSMI estimates are not consistently available until January 1992, a climatological ratio adjustment based on the overlapping (January 1993 – December 2008) SSMI/SSMIS-corrected PERSIANN is applied during the period January 1983 – December 1991. Because February 1992 lacks a sufficient number of PERSIANN-CDR estimates to create a stable monthly calibration by SSMI, the above climatological calibration is used.

The SSMI/SSMIS-corrected PERSIANN is then adjusted by the monthly climatological merged TCC/MCTG ratios to obtain the best estimate of precipitation in the tropics and mid-latitudes. The merged TCC/MCTG is described later in this section.

Table 2. Geosynchronous satellites contributing to the GridSat IR data archive that PERSIANN uses. The current satellites are highlighted in blue. [Table courtesy K. Knapp, NCEI.]

Satellite	Launch Date	Primary Location*	Period of Record	Bands Available
GOES-5	1981/05/22	-75°	1983/07/01-1984/07/30	Visible & IR
GOES-6	1983/04/28	-135°	1983/09/30-1989/01/24	Visible & IR
GOES-7	1987/02/26	-100°	1987/03/25-1996/01/17	Visible & IR
GOES-8	1994/04/13	-75°	1994/11/30-2003/03/31	Visible, WV & IR
GOES-9	1995/05/23	-135°	1996/01/01-1998/07/27	Visible, WV & IR
GOES-9[1]	1995/05/23	155°	2003/04/25-2005/10/01	Visible, WV & IR
GOES-10	1997/04/25	-135°	1998/07/21-2006/06/21	Visible, WV & IR
GOES-11	2000/05/03	-135°	2006/06/21-2011/12/06	Visible, WV & IR
GOES-12	2001/07/23	-75°	2003/04/01-2010/04/14	Visible, WV & IR
GOES-13	2006/05/24	-75°	2010/04/14-2017/01/08	Visible, WV & IR
GOES-15	2010/03/04	-135°	2011/12/06-current	Visible, WV & IR
GOES-16	11/19/2016	-75.2°	2018/01/09-current	Visible, WV & IR
MET-2	1981/06/19	0°	1983/08/31-1988/08/24	Visible, WV & IR
MET-3	1988/06/15	0°	1988/08/24-1989/06/19	Visible, WV & IR
MET-3[2]	1988/06/15	-70°	1992/09/11-1995/01/29	Visible, WV & IR
MET-4	1989/03/06	0°	1989/06/19-1994/02/04	Visible, WV & IR
MET-5	1994/04/13	0°	1994/01/01-1997/02/01	Visible, WV & IR
MET-5[3]	1994/04/13	63°	1998/06/01-2007/04/16	Visible, WV & IR
MET-6	1993/11/20	0°	1997/02/13-2000/01/20	Visible, WV & IR
MET-7	1997/09/02	0°	1998/05/31-2005/01/01	Visible, WV & IR
MET-7[3]	1997/09/02	63°	2007/04/16-2017/04/04	Visible, WV & IR
MET-8	2002/07/28	0°	2005/01/01-2007/05/10	Visible, WV & IR
MET-9	2005/12/21	0°	2007/05/10-2013/01/20	Visible, WV & IR
MET-10	2012/07/05	0°	2013/01/20-2018/02/20	Visible, WV & IR
MET-11	2015/07/15	0°	2018/02/20-current	Visible, WV & IR
MET-8	1997/09/02	56°	2017/04/05-current	Visible, WV & IR
GMS-1	1977/07/14	140°	1984/01/21-1984/06/29	Visible & IR
GMS-2	1981/08/11	140°	1983/06/30-1984/09/26	Visible & IR
GMS-3	1984/08/03	140°	1984/09/27-1989/12/03	Visible & IR
GMS-4	1989/09/06	140°	1989/12/04-1995/06/12	Visible & IR
GMS-5	1995/03/18	140°	1995/06/13-2003/05/21	Visible, WV & IR
MTS-1	2005/02/26	140°	2005/10/01-2010/06/30	Visible, WV & IR
MTS-2	2006/02/18	140°	2010/06/30-2015/07/06	Visible, WV & IR
HIM-8	2014/10/07	141°	2015/07/07-present	Visible, WV & IR

GPCC gauge precipitation analysis

The gauge analysis used in GPCP is produced by the Global Precipitation Climatology Centre (GPCC) under the direction of Andreas Becker, located in the Deutscher Wetterdienst, Offenbach a.M., Germany (Schneider et al. 2017; Schneider et al. 2014; Becker et al. 2013). Precipitation gauge reports are archived from a time-varying collection of over 70,000 stations around the globe, both from Global Telecommunications System (GTS) reports and from other world-wide or national data collections. An extensive quality-control system is run, featuring an automated screening and then a manual step designed to retain legitimate extreme events that characterize precipitation. This long-term data collection and preparation activity feeds into an analysis that is done in two steps. First, a long-term climatology is assembled from all available gauge data, focusing on the period 1951-2000. For each month, the individual gauge reports are converted to deviations from climatology, and are analyzed into gridded values using a variant of the SPHEREMAP spatial interpolation routine (Willmott et al. 1985). Finally, the analysis is produced by superimposing the monthly anomaly analysis on the climatology for that month. Two GPCC products are used in GPCP; the Full Data Reanalysis (currently Version 2018) is a retrospective analysis that covers the period 1901-2016, and it is used in GPCP for the span 1983-2016. Thereafter we use the GPCC Monitoring Product (currently Version 6), which has a similar quality control and the same analysis scheme as the Full Data Reanalysis, but whose data source is limited to GTS reports. When the Full Data Reanalysis is updated to a longer record we expect to reprocess the GPCP datasets to take advantage of the improved data.

Merged TCC/MCTG Climatology

As part of Version 3, information from the Tropical Rainfall Measuring Mission (TRMM) and the Global Precipitation Measurement (GPM) mission is used to adjust the monthly GPCPV3.1 SSMI/SSMIS-adjusted PERSIANN-CDR estimates prior to merging with MCTG-adjusted AIRS and the gauge analysis. The Tropical Composite Climatology (TCC) used in this GPCP V3.1 is constructed using a mean of the estimates from the PMW, radar, and combined products from both TRMM and GPM. The TCC for the period 1998-2018 for tropical latitudes is an update of the earlier versions using just TRMM data (Adler et al. 2009; Wang et al. 2014). In this application the TCC month-of-the-year climatological values are compared with the overlapping SSMI/SSMIS-adjusted PERSIANN-CDR to compute the monthly climatological adjustment ratios. The goal is to force the GPCP V3.1 estimates to be as consistent as possible with the TCC, which is assumed to be the “best” current tropical climatology available. The TCC adjustment procedure is only applied over ocean, where gauge adjustment is not possible.

The MCTG is a composite monthly climatology based on CloudSat, TRMM-and GPM-based precipitation estimates using a concept similar to that described in Behrangi and Song (2020) and Behrangi et al. (2012; 2014). CloudSat precipitation frequency is first adjusted for the spatial resolution of TRMM and GPM radar footprint (~5km) and is used as a constraint for combining CloudSat precipitation (rainfall plus snowfall), precipitation rates from the TCC (within 25°N-S), the TRMM 2BCMBT V06 (within 35°N-S), and the GPM 2BCMB V06 (within 35° – 65°N-S). Poleward of 65°N-S the precipitation climatology is based on CloudSat as it does not have a signal saturation issue (Behrangi et al. 2012) and GPM lacks coverage. The GPM 2BCMB

is used as it provides an overlapping period with CloudSat (2007-2010) when both day- and night-time observations are available.

The TCC provides the best monthly climatological estimates in the region 25°N-S while the MCTG provides the best monthly climatological estimates in the regions 25°-90°N,S. Separate monthly climatological TCC and MCTG ratios were computed for SSMI/SSMIS-adjusted PERSIANN, and MCTG ratios were also computed for TOVS/AIRS. To take advantage of this information, merged monthly TCC/MCTG climatological adjustment ratios were developed and applied at the grid box level to the SSMI/SSMIS-adjusted PERSIANN estimates for the entire record. Because the TCC and MCTG have radically different means in the tropics, an overall bias adjustment is necessary to seamlessly merge the TCC and MCTG ratios. The merging process starts with computing separate NH and SH precipitation-weighted TCC/precipitation-weighted MCTG zonal average ratios for the zonal bands 22°N – 23°N and 22°S – 23°S. Weighting by precipitation is necessary to prevent large ratios in low-precipitation regions from dominating the zonal mean ratios. The MCTG is then capped at 2 to avoid undue influence of large ratios in regions of light precipitation and to maintain the regionality of the MCTG adjustment. Furthermore, the capping of ratios at 2 is consistent with the TCC. At latitudes 35°N and 35°S the precipitation-weighted TCC/precipitation-weighted MCTG zonal ratios are set to one as this is the extent of the TRMM influence in MCTG. These ratios are smooth-filled over the region 23°N-35°N and 23°S-35°S and then applied directly to the capped MCTG to provide a smooth transition from the TCC to the MCTG. The final step is to perform inverse linear weighting of TCC and adjusted MCTG across 20°-23.5°N,S to mitigate discontinuities in the ratio field.

The resulting ratio field is:

- 20°N – 20°S: TCC
- 20°N – 23.5°N and 20°S – 23.5°S: Weighted TCC/adjusted MCTG average
- 23.5°N – 35°N and 23.5°S – 35°S: Adjusted capped MCTG
- 35°N – 58°N and 35°S – 58°S: Capped MCTG

Note that the TOVS/AIRS estimates are globally adjusted to the MCTG prior to merging with the adjusted PERSIANN estimates for consistency.

3.3.2 Ancillary Data

MERRA-2 hourly, instantaneous, single-level assimilation data (“inst1_2d_asm”) were obtained from the Goddard Earth Sciences Data and Information Services Center for the entire GPCP V3.1 record. The fields used are surface pressure, 2-meter temperature, and 2-meter specific humidity. These data are used to compute the probability of liquid precipitation phase (Section 3.4.3).

3.4 Theoretical Description

The bulk of the GPCP code merges the various inputs to obtain the GPCP Satellite-Gauge estimate. Most of the input data are obtained as rain rates having been processed by other

groups. The exception to this is the PMW SSMI and SSMIS estimates computed with GPCP processing using the GPROF algorithm.

3.4.1 Physical and Mathematical Description

GPROF precipitation

GPROF is based on Kummerow et al. (1996), Olson et al. (1999), and Kummerow et al. (2001). GPROF is a multi-channel physical approach for retrieving rainfall and vertical structure information from satellite-based PMW observations. The version used here, GPROF 2010v2, applies to both SSMI and SSMIS. GPROF applies a Bayesian inversion method to the observed microwave brightness temperatures using an extensive library of profiles relating hydrometeor profiles, microwave brightness temperatures, and surface precipitation rates. GPROF includes a procedure that accounts for inhomogeneities of the rainfall within the satellite field of view. Over land and coastal surface areas the conical-scan imager library largely reduces to a scattering-type procedure using only the higher-frequency channels. This loss of information arises from the physics of the emission signal in the lower frequencies when the underlying surface is other than entirely water.

PERSIANN-CDR precipitation

PERSIANN-CDR rain rate estimates provided by UC-Irvine are generated at the 0.25° resolution product that are calibrated to the monthly 2.5° GPCP V2.3 SG product and cover the region 60°N-S. GridSat-B1 IRWIN IR data are used as input to create an intermediate output product called PERSIANN-B1. Precipitation is assigned based on an off-line training set of low-Earth-orbit PMW precipitation samples that create coefficients that vary across a coarse spatial grid. The PERSIANN-B1 estimates are then thresholded to filter out noisy pixels. These noisy pixels are generally associated with pixels where the rain rate is “zero” but the Neural Network model estimates a very small nonzero value. These resulting noisy pixels can lead to a very large number of "rainy" days (rain rate > 0 mm/day). The 3-hour PERSIANN-B1 data are accumulated to the month and bias-corrected by the 2.5° GPCP V2.3 SG estimate.

For GPCP V3, these 3-hour PERSIANN estimates are averaged to the 0.5° resolution and then calibrated by gridded SSMI/SSMIS estimates for the month using regional matched histograms. The matched histograms are accumulated on a 3x3 grid for the month to enhance sampling and then smoothed using a 5x5 boxcar. A precipitation rate correction look-up table is computed and applied to the 3-hour PERSIANN estimates.

3.4.2 Data Merging Strategy

The V3.1 GPCP SG monthly is created by merging precipitation estimates in several different stages. Each of these stages leads to the creation of intermediate products, some of which are provided as outputs.

Adjusted PERSIANN merger with TOVS/AIRS

The coverage of the adjusted PERSIANN precipitation estimates is limited to the region 60°N-S. To obtain global coverage, the regions outside this latitude band are filled using the globally complete TOVS/AIRS data. The adjusted PERSIANN data are used without any additional TOVS/AIRS data for the latitude band 35° N-S. Where there are holes in PERSIANN coverage as the result of data dropouts, the TOVS/AIRS data are inserted after the following adjustments: For small gaps, the TOVS/AIRS data are adjusted to the zonally averaged mean bias of the adjusted PERSIANN and inserted. For large gaps (listed in item o. of Section 6), the map of the ratio of PERSIANN to TOVS/AIRS (both at the monthly scale) is smoothed with a 5x5 boxcar filter, all holes are smooth-filled[1] using a 9x3 (X x Y) template, and the resulting ratio is multiplied by the TOVS/AIRS value, gridbox-by-gridbox. [This design presumes that zonal structure and the ratio near the hole are the important factors. All gaps will be smooth-filled in a future revision.] In the zone from 35° to 58° in each hemisphere, the adjusted PERSIANN and TOVS/AIRS estimates are weighted such that adjusted PERSIANN is heavily weighted at 35° N/S and TOVS/AIRS is heavily weighted at 58° N/S with a smooth transition in weighting in between. In the Northern Hemisphere, this bias adjustment is anchored on the equatorward side by the zonal average of the averaged adjusted PERSIANN and TOVS/AIRS values at 58° N. The bias adjustment over land on the polar side is anchored by the zonal average of the monthly precipitation gauge data at 70°N, with a smooth linear variation in between. The gauge's zonal average includes only grid boxes for which the gauge quality index is greater than zero. From 70°N to the North Pole, TOVS/AIRS data are adjusted to the bias of the same monthly precipitation gauge value average at 70°N. The same procedure is applied in the Southern Hemisphere, except the annual climatological precipitation gauge values are zonally averaged at 70°S. The monthly values are not used in the Antarctic as the lack of sufficient land coverage there yields unstable results. Furthermore, the current GPCC analysis lacks data over Antarctica, so this climatological adjustment is from a previous GPCC Monitoring Product. All seasonal variations in this description were developed in off-line studies of typical dataset variations, with the driving criterion being choosing a transition that ensures reasonable performance.

Satellite-gauge combination

After the multi-satellite (merged adjusted PERSIANN-TOVS/AIRS) estimate is created, it is merged with the gauge analysis. Prior to combination, the GPCC gauge analysis values are adjusted to remove systematic error due to wind effects, side-wetting, evaporation, etc., following Legates (1987). The process to combine the gauge analysis with the multi-satellite is a two-step process following Huffman et al. (1995).

In the first step the multi-satellite is adjusted to have the same bias as the gauges for all pixels with at least 35% land. This large-area adjustment is calculated and applied based on a 21x21 grid box average centered on the grid box of interest. The adjusted multi-satellite is then

[1] Smooth-filling is an iterative process. On each pass, the value in every gridbox that originally was "missing" is replaced by the average of (non-missing) values on the stated template (here, 9 gridboxes in the X direction and 3 in the Y). This continues until the data field (approximately) converges.

combined with the gauges using a weighted average where the weights are the inverse (estimated) error variances.

Random error estimates

The random error of the combined satellite/gauge estimate is based on the technique of Huffman (1997) and is supplied as an output field. The bias error is neglected compared to the random error (both physical and algorithmic) so that the estimated error variance of an average over a finite set of observations, VAR, can be expressed.

$$\mathrm{VAR} = \frac{\mathrm{H}\left(\bar{r}+S\right)\left[24+49\sqrt{\bar{r}}\right]}{N_i}$$

where H and S are assumed constant, $\bar{r}$ is the average precipitation rate in mm day^{-1}, N_i is the number of independent samples in the set of observations, and the expression in square brackets is a parameterization of the conditional precipitation rate based on work with the Goddard Scattering Algorithm, Version 2.1 (Adler et al. 1994) and fitting to the Surface Reference Data Center analyses (Huffman 1997). The "constants" H and S are set for each of the data sets for which error estimates are required by comparison of the data set against the SRDC and GPCC analyses and tropical Pacific atoll gauge data (Morrissey et al., 1995).

For the independent data sets $\bar{r}$ is taken to be the independent estimate of precipitation itself. However, when these errors are used in the combination, theory and tests show that the result is a low bias. $\bar{r}$ needs to have the same value in all the error estimates contributing to the combination, so it is estimated as the simple average of all precipitation values contributing to the combination. Note that this scheme is only used in computing errors used in the combination.

The formalism mixes algorithm and sampling error, and should be replaced by a more complete method when additional information is available from the single-source estimates. However, when Krajewski et al. (2000) developed and applied a methodology for assessing the expected random error in a gridded precipitation field, their estimates of expected error agreed rather closely with the errors estimated for the multi-satellite and satellite-gauge combinations.

3.4.3 Precipitation Phase

GPROF2010v2 retrieves total hydrometeor mass in the atmospheric column (except the conical-scan imager PMW retrievals only consider total solid hydrometeor mass over land and coast) and then implicitly correlate it to surface precipitation in any phase. Given these facts, the “precipitation” reported in this document refers to all forms of precipitation, including rain, drizzle, snow, graupel, and hail. The IR retrievals are calibrated to the PMW retrievals, again, without reference to precipitation phase. These IR calibrations are in-filled from surrounding areas in the snowy/icy-surface areas where PMW cannot provide estimates.

Since the precipitation phase, namely whether it is liquid, solid, or mixed, is not currently

provided as a satellite-based calculation by the precipitation algorithms used in GPCP V3.1, we must use ancillary data sets to create the estimate. Formally, there should be separate estimates for each phase. However, mixed-phase cases tend to be a small fraction of all cases, and we consider the estimation schemes to be sufficiently simplistic that estimating mixed phase as a separate class seems unnecessary. Some users need information on the occurrence of the solid phase, both due to the delays it introduces in moving precipitation water mass through hydrological systems, and due to the hazardous surface conditions that snow and ice create. Accordingly, we lump together liquid and mixed as "liquid" and compute a simple probability of liquid phase.

For the swath data, we adopt the Liu scheme (personal communication, 2013; Sims and Liu 2015), which was developed for the Radiometer Team. The present (pre-publication) form is a simple look-up table for probability of liquid precipitation as a function of wet-bulb temperature, with separate curves for land and ocean. This is a current area of research, so we anticipate changes as research results are reported. Since this diagnostic is independent of the estimated precipitation, we choose to report the probability of liquid phase for all grid boxes, including those with zero estimated precipitation. [This raises the possibility that the "probability of liquid precipitation" field can be applied to any other global precipitation field for estimating phase.] The surface temperature, humidity, and pressure information needed to compute the surface wet-bulb temperature are taken from the MERRA-2 reanalysis.

At the monthly scale the probability could either be the fraction of the time that the precipitation is liquid or the fraction of the monthly accumulation that fell as liquid. The latter seems to be what most users will want, so this is the parameter computed whenever at least 97% of the 3-hourly PERSIANN samples are available for the month. [This limit excludes cases where more than a day of data is missing, either in one run or a scattered series]. The monthly probability of liquid is computed as the precipitation-rate-weighted average of all 3-hourly probabilities in the month, except for grid boxes where zero precipitation is estimated for the month, in which case it is the simple average of all available probabilities in the month. This approach assumes that the occurrence of liquid and solid over the month will approximately conform to the percentages given in the specification equation, so that the weighted probability of liquid approximates the fraction of amount of precipitation: liquid precipitation = probability * precipitation, and solid precipitation = (100 – probability) * precipitation. In the regions where PERSIANN is merged with TOVS/AIRS, the 3-hourly probabilities for PERSIANN are used. The fallback, which occurs when PERSIANN has more missings, and whenever TOVS/AIRS estimates are used, is to average all the 3-hourly probabilities in the month.

Note that the assignment of phase does not change the units of precipitation, which is the depth of liquid. In the case of solid precipitation, this is usually referred to as snow water equivalent (SWE). The depth of fallen snow that corresponds to this SWE depends on the density of the snow. Typically, it takes about 10 mm of fallen snow to yield 1 mm of SWE, but the ratio depends on location, meteorological regime, time of year, and elevation. There is an excellent discussion of how Environment Canada is addressing this in Wang et al. (2017).

3.4.4 Gauge Relative Weighting

The GPCC gauge analysis and the satellite-only estimates are merged based on inverse error variance weighting as described in Section 3.4.2. In areas of dense gauge coverage, the gauge analysis is assigned more weight due to the higher confidence, and is therefore weighted more heavily in the satellite-gauge estimate. In areas with sparse gauge coverage, the satellite-only estimate is weighted more heavily. Using these relative weights, the percent weighting of the gauge analysis can be computed. The gauge relative weighting is a direct reflection of the gauge population for a given month and can be used as guidance to assess the quality of any given grid box.

3.4.5 Quality Index

The quality index is based on Huffman's (1997) analysis of sampling error for a particular data source for a month. The general form of the relationship is simplified to a relationship that can be inverted to give the number of samples. When all the constants on the right-hand side are set for the gauge analysis, but final satellite-gauge values are used for the estimated precipitation and random error values, the number variable is defined as the equivalent number of gauges. Following Huffman (1997), the interpretation is that this is the approximate number of gauges required to produce the estimated random error, given the estimated precipitation. The units are gauges per area, and in the current implementation the area is carried as 2.5°x2.5° of latitude/longitude, even though GPCPV3 is computed on a much finer scale, in order to facilitate interpretation in large-error regions. Note that this formulation only addresses random error, not bias.

3.4.6 Algorithm Output

The GPCP Version 3.1 Satellite-Gauge monthly precipitation data set covers the period January 1983 through December 2019 (with additional months planned). The primary product in the dataset is a combined observation-only dataset, that is, a gridded analysis based on gauge measurements and satellite estimates of precipitation.

The data set archive consists of monthly netCDFs, with each file having the eight following fields:

(1) merged satellite-gauge precipitation estimate (mm/d),
(2) merged satellite-gauge precipitation random error estimate (mm/d),
(3) satellite-only precipitation estimate (mm/d),
(4) satellite source field (IR = 0, IR/TOVS/AIRS blend = 2, TOVS/AIRS = 4)
(5) undercatch-corrected gauge analysis precipitation (mm/d),
(6) probability of liquid-phase precipitation (%),
(7) gauge relative weighting, and
(8) quality index.

Each file occupies almost 7 MB. The grid on which each field of values is presented is a 0.5°x0.5° latitude--longitude (Cylindrical Equal Distance) global array of points. It is size 720x360, with X (longitude) incrementing most rapidly West to East from the International

Dateline, and then Y (latitude) incrementing North to South. Grid edges are placed on whole- and half-degree values:

First point center = (89.75°N, -179.75°W)
Second point center = (89.75°N, -179.25°W)
Last point center = (89.75°S, 179.75°E)

4.0 Test Datasets and Output

4.1 Test Input Datasets

Initial test input data sets consisted of several months of data inputs.

4.2 Test Output Analysis

4.2.1 Reproducibility

Reproducibility has been shown by comparison with results of parallel runs.

4.2.2 Precision and Accuracy

As described in Section 3.4.2, a reasonable estimate of random error is provided for each grid box based on a parameterized equation that depends on precipitation rate and sampling. At present, there is no comparable parameterization for the bias error. Section 3.4.2 uses the merged TCC/MCTG climatology to approximately adjust the GPCP V3.1 estimates to what is considered the "best" available climatology, but this remains a research topic.

4.2.3 Error Budget

It is extremely challenging to develop an error budget for satellite retrievals. The discussion in Section 4.2.2 represents the state of the art.

The Probability of Liquid Phase Precipitation field is being characterized. Early results tend to show that low(high) probabilities are low(high). That is, the statements of phase are generally too confident.

5.0 Practical Considerations

5.1 Quality Assessment and Diagnostics

Diagnostic analyses and plots of the output products are computed for each month, with comparison to climatologies and the prior GPCP V2.3 datasets. Particular attention is paid to the time series of various large-area averages for possible deviations as data boundaries are encountered. Anomalies are identified and analyzed to determine their origins.

5.2 Exception Handling

Errors in input data or processing are corrected when possible, or documented if they are not fixable.

5.3 Algorithm Validation

The diagnostics and exception handling described above form the first line of validation. Further evaluation will include validation against independent satellite-based precipitation data sets as well as in situ observations such as those from the Multi-Radar/Multi-Sensor System and Pacific atoll rain gauges.

5.4 Processing Environment and Resources

The computer used to process the GPCP monthly product is a CentOS 6 Linux Server. The programming languages and software include: C shell scripts to run the processing code, FORTRAN programs to perform the calculations, and xmgrace and Python to make the diagnostic plots and visualizations.

6.0 Assumptions and Limitations

There are a number of known issues that are relevant for a CDR-like data set. The GPCP team has worked hard to ameliorate these issues:

a. This second release, labeled V3.1, is a significant improvement over the 3.0 beta, but has known limitations. Specifically, the TOVS/AIRS record is not as homogeneous as we expect for a CDR. The team continues to work toward improving these issues in a future release.
b. The GridSat data contains residual intersatellite differences at the boundaries between the areas of coverage of the geo-IR sensors, and these boundaries are partially carried through into the PERSIANN-CDR output. Formally, the boundaries shift from one three-hour snapshot to the next, but they tend to recur in the same locations, and so are more or less visible in the monthly accumulations.
c. Unlike the previous Version 1 and 2 GPCP SG datasets, the Version 3.1 monthly accumulation of (PERSIANN-CDR) IR-based input is built from the time series of lowest-zenith-angle observations for each three-hour period. Variations in the time series (including “missing” values due to the lack of a secondary observation when the primary is missing) are driven by the availability of each geo-IR satellite.

d. Unlike the previous Version 1 and 2 GPCP SG datasets, the IR Tb data for V3.1 GPCP are provided in a consistent data format for the entire record.
e. Beginning with January 2009, SSMIS precipitation estimates replaced the SSMI estimates because the F13 SSMI failed in September 2009 and we wanted to both avoid possible degraded performance late in the SSMI record and to establish a whole-year data boundary to aid in diagnosing possible biases. The SSMIS data have been adjusted to match the large-scale bias of the SSMI to maintain homogeneity across the data boundary.
f. The TOVS precipitation estimates for the SSMI period January 1983 – February 1999 are based on two satellites. For February 1999 – August 2002 (and with a degrading record extending April 2005), the TOVS estimates are based on only one satellite. However, the recently discovered, significant shift in high-latitude TOVS estimates starting with December 1999 are a much stronger data boundary and lead to the choices made in item i.
g. TOVS data were partially denied for the period 10-18 September 2001 and cannot be recovered. As well, various operational issues caused partially or completely missing days of TOVS data.
h. Beginning with September 2002, AIRS precipitation estimates replaced the TOVS estimates at high latitudes because of TOVS product degradation later in 2003.
i. The TOVS monthly precipitation estimates up through November 1999 are calibrated to approximately match the zonal average AIRS monthly using two separate independent 13.75-year periods as the calibration periods; for December 1999-August 2002, TOVS daily are calibrated to the zonal average AIRS daily using two separate independent ENSO-neutral 24-month periods. For both of these, regional differences remain.
j. Every effort has been made to preserve the homogeneity of the Version 3.1 data record. However, the regional variances inherent in the climatologically calibrated PERSIANN-CDR data are typically smaller than those encountered in the SSMI/SSMIS data, so the statistical nature of the fields will be different for the pre-microwave and microwave eras (the latter starting January 1992).
k. The precipitation gauge data used in the Version 3.1 analysis consists of GPCC Full for the period 1983 – 2016 and GPCC Monitoring for the period 2017 – present. Although there is strong consistency in analysis scheme, quality control, and data sources between the two analyses, there can be a discernible change in statistics where the gauge population is different between the two datasets.
l. Every attempt has been made to create an observation-only based precipitation data set. However, the TOVS estimates (but not AIRS) rely on numerical model data to initialize the estimation technique. The greatest chance of model influence is at high latitudes, where the retrievals more often fail to converge, and so fall back on the first guess. As well, the precipitation phase variable is a diagnostic based strictly on the MERRA-2 global analysis. This is believed to have only modest impact from numerical effects, since temperature and humidity are typically well-constrained by observations in MERRA-2.
m. Some polar-orbiting satellites have experienced significant drifting of the equator-crossing time during their period of service. There is no direct effect on the accuracy of the retrievals, but it is possible that the systematic change in sampling time could introduce biases in the resulting precipitation estimates. It is unlikely that this issue affects the SSMI/SSMIS data used for calibration because the sequence of single satellites used have all

stayed within ±1 hour of the nominal 6 a.m. / 6 p.m. overpass time. Satellites carrying the TOVS sensors did drift, and a diurnal correction was applied to the data by the SRT. The Aqua satellite carrying the AIRS sensor has been station-keeping at 1:30 p.m. as part of the A-Train.

n. The PERSIANN-CDR for February 1992 is missing data for the span 11 18Z – 25 09Z. Because of this, there is inadequate data to provide a stable monthly calibration with the SSMI estimates, so the fallback pre-SSMI climatological calibration is used.
o. The PERSIANN-CDR for the months February, May, June, July 1984; February, May, June, September, November, December 1989; February 1990; and February 1992 are missing partial or entire sectors of precipitation. The nominal zonal average correction for the TOVS, used to fill in the sectors, proved inadequate. The backup scheme using TOVS/AIRS data is described in Section 3.4.2.

7.0 Future Enhancements

The immediate plan is to compute additional months as they occur. We will continue addressing the known anomalies in V3.1, namely inhomogeneities in the TOVS/AIRS record, and seams in the IR input. As well, we plan to shift to the smooth-fill scheme for adjusting TOVS/AIRS data for filling all gaps in the PERSIANN (Section 3.4.2). In addition, there is a plan to create a daily data set for the entire 1983-present period on the same 0.5° grid as for the monthly. The fields will be globally complete for the period when daily TOVS/AIRS data are available (October 1996 to the present), and over the latitude band 60°N-S for earlier times. As well, a 3-hourly product is anticipated, likely to be created by rescaling the Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement (GPM) mission (IMERG; Huffman et al. 2019). This short-interval product is likely to be computed at the same 0.5° grid over the global domain for June 2000 – present (and shifting to January 1998 when all necessary inputs become available). Refining the methods for uncertainty quantification is also among the future tasks. Finally, the MEaSUREs-2017 project led by Dr. A. Behrangi (University of Arizona), with which V3.1 is associated, is working toward improved high-latitude estimates. This can include implementation of the new retrievals and sensors as they become available and revisiting the gauge undercatch correction factors.

8.0 References

Adler, R.F., G.J. Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, P. Arkin, E. Nelkin 2003: The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *J. Hydrometeor.*, **4**,1147-1167.

Adler, R.F., G.J. Huffman, P.R. Keehn, 1994: Global Tropical Rain Estimates from Microwave-Adjusted Geosynchronous IR Data. *Remote Sensing Rev*., **11**, 125-152.

Adler, R., M. Sapiano, G. Huffman, D. Bolvin, J.-J. Wang, G. Gu, E. Nelkin, P. Xie , L. Chiu, R. Ferraro, U. Schneider, A. Becker, 2016: New Global Precipitation Climatology Project Monthly Analysis Product Corrects Satellite Data Drifts. *GEWEX News*, **26**, 7-9.

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

Ashouri, H., K.-L. Hsu, S. Sorooshian, D.K. Braithwaite, K.R. Knapp, L.D. Cecil, B.R. Nelson, O.P. Prat, 2015: PERSIANN-CDR: Daily Precipitation Climate Data Record from Multisatellite Observations for Hydrological and Climate Studies. *Bull. Amer. Meteor. Soc.*, **96**, 69-83. *doi:10.1175/BAMS-D-13-00068.1*

Becker, A., P. Finger, A. Meyer-Christoffer, B. Rudolf, K. Schamm, U. Schneider, M. Ziese, 2013: A Description of the Global Land-surface Precipitation Data products of the Global Precipitation Climatology Centre with Sample Applications including Centennial (Trend) Analysis from 1901-present. *Earth System Science Data. doi:10.5194/essd-5-71-2013.* http://www.earth-syst-sci-data.net/5/71/2013/essd-5-71-2013.html

Behrangi, A., Stephens, G., Adler, R.F., Huffman, G.J., Lambrigtsen, B., Lebsock, M., 2014: An Update on the Oceanic Precipitation Rate and Its Zonal Distribution in Light of Advanced Observations from Space. *Journal of Climate,* **27**, 3957-3965. *doi:10.1175/jcli-d-13-00679.1*

Behrangi, A., Lebsock, M., Wong, S., Lambrigtsen, B., 2012: On the Quantification of Oceanic Rainfall using Spaceborne Sensors. *Journal of Geophysical Research: Atmospheres*, **117**. *doi:10.1029/2012jd017979*

Behrangi, A., and Y. Song, 2020: A new estimate for oceanic precipitation amount and distribution using complementary precipitation observations from space and comparison with GPCP. *Environmental Research Letters*, **15**. *doi:10.1088/1748-9326/abc6d1*

Hollinger, J., R. Lo, G. Poe, J. Pierce, 1987: *Special Sensor Microwave/Imager User's Guide*, Naval Res. Lab., Washington, DC, 120 pp.

Hollinger, J.P., J.L. Pierce, G.A. Poe, 1990: SSMI Instrument Evaluation. *IEEE Trans. Geosci. Remote Sens*., **28**, 781-790.

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

Huffman, G.J, R. F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, U. Schneider, 1997: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset. *Bull. Amer. Meteor. Soc*., **78**, 5–20.

Huffman, G.J, R.F. Adler, D.T. Bolvin, G. Gu 2009: Improving the Global Precipitation Record: GPCP Version 2.1. *Geophys. Res. Lett*., **36**, L17808. *doi:10.1029/2009GL040000*

Huffman, G.J., R.F. Adler, B. Rudolf, U. Schneider, P.R. Keehn, 1995: Global Precipitation Estimates Based on a Technique for Combining Satellite Data, Raingauge Analysis and Model Precipitation Information. *J. Climate*, **8**, part 2, 1284-1295.

Huffman, G.J., D.T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, C. Kidd, E.J. Nelkin, S. Sorooshian, J. Tan, P. Xie, 2019: Algorithm Theoretical Basis Document (ATBD) Version 6.2 for the NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (I-MERG). GPM Project, Greenbelt, MD, 36 pp. *https://gpm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf*

Knapp, K.R., 2008: Scientific Data Stewardship of International Satellite Cloud Climatology Project B1 Global Geostationary Observations. *J. Appl. Remote Sens., B*, 023548. *doi:10.1117/1.30434 61*

Krajewski, W. F., G. J. Ciach, J. R. McCollum, C. Bacotiu, 2000: Initial Validation of the Global Precipitation Climatology Project Monthly Rainfall over the United States. *J. Appl. Meteor*., **39**, 1071–1086.

Kummerow, C., Y. Hong, W.S. Olson, S. Yang, R.F. Adler, J. McCollum, R. Ferraro, G. Petty, D-B. Shin, T.T. Wilheit, 2001: The Evolution of the Goddard Profiling Algorithm (GPROF) for Rainfall Estimation from Passive Microwave Sensors. *J. Appl. Meteor*., **40**, 1801–1820. *doi:/10.1175/1520-0450(2001)040<1801:TEOTGP>2.0.CO;2*

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.

Legates, D.R, 1987: A Climatology of Global Precipitation. *Pub. in Climatol*., **40**, U. of Delaware, 85 pp.

Morrissey, M. L., M. A. Schafer, S. E. Postawko, B. Gibson, 1995: The Pacific Rain Gage Rainfall Database. *Water Resour. Res*., **31**, 2111–2113.

Northrup Grumman, 2002: *Algorithm and Data User Manual (ADUM) for the Special Sensor Microwave Imager/Sounder (SSMIS)*, Northrup Grumman Electronic Systems, Azusa, CA, Report 12621, 69 pp. *ftp://rain.atmos.colostate.edu/pub/msapiano/SSMI_SSMIS_Archive_Docs/SSMIS_general/Algorithm_and_Data_User_Manual_For_SSMIS_Jul02.pdf*

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.

Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, B. Rudolf, 2014: GPCC's New Land Surface Precipitation Climatology based on Quality-controlled In Situ Data and its Role in

Quantifying the Global Water Cycle. *Theor. Appl. Climatol.,* **115**, 15, *doi:10.1007/s00704-013-0860-x*

Schneider, U., P. Finger, A. Meyer-Christoffer, E. Rustemeier, M. Ziese, and A. Becker, 2017: Evaluating the Hydrological Cycle over Land Using the Newly-Corrected Precipitation Climatology from the Global Precipitation Climatology Centre (GPCC). *Atmosphere*, **8**, 52, *doi:10.3390/atmos8030052*

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*

Susskind, J. C.D. Barnet, J.M. Blaisdell, 2003: Retrieval of Atmospheric and Surface Parameters from AIRS/AMSU/HSB Data in the Presence of Clouds. *IEEE Trans. Geosci. Rem. Sens*., **41**, 390-409. *doi:10.1109/TGRS.2002.808236*

Susskind, J., J. Pfaendtner, 1989: Impact of Interactive Physical Retrievals on NWP. *Report on the Joint ECMWF/EUMETSAT Workshop on the Use of Satellite Data in Operational Weather Prediction: 1989–1993*, *Vol. 1*, T. Hollingsworth, Ed., ECMWF, Shinfield Park, Reading RG2 9AV, U.K., 245-270.

Susskind, J., P. Piraino, L. Rokke, L. Iredell, A. Mehta, 1997: Characteristics of the TOVS Pathfinder Path A Dataset. *Bull. Amer. Meteor. Soc*., **78**, 1449-1472.

Wang, J.-J., R.F. Adler, G.J. Huffman, D.F. Bolvin, 2014: An Updated TRMM Composite Climatology of Tropical Rainfall and Its Validation. *J. Climate*, **27**, 273-284. *doi:10.1175/JCLI-D-13-00331.1*

Wang, X., Y. Feng, E. Mekis, 2017: Adjusted Daily Rainfall and Snowfall Data for Canada. *Atmos.-Ocean*, **55**, 155-168. *doi:10.1080/07055900.2017.1342163*

Willmott, C. J., C. M. Rowe, W. D. Philpot, 1985: Small-Scale Climate Maps: A Sensitivity Analysis of Some Common Assumptions Associated with Grid-Point Interpolation and Contouring. *Amer. Cartographer*, **12**, 5–16. *doi:10.1559/152304085783914686*

Appendix A. Acronyms and Abbreviations

AIRS	Atmospheric Infrared Sounder
AMSU	Advanced Microwave Sounding Unit
ATBD	Algorithm Theoretical Basis Document
CDR	Climate Data Record
DMSP	Defense Meteorological Satellite Program
DWD	Deutscher Wetterdienst (German Weather Service)
FGGE	First Global Atmospheric Research Program (GARP) Global Experiment
GES DISC	Goddard Earth Science Data and Information Services Center

GEWEX	Global Water and Energy Exchange project
GHz	gigahertz
GMS	Geosynchronous Meteorological Satellite
GOES	Geosynchronous Operational Environmental Satellite
GPCC	Global Precipitation Climatology Centre
GPCP	Global Precipitation Climatology Project
GPM	Global Precipitation Measurement
GPROF	Goddard Profiling retrieval algorithm
GSFC	Goddard Space Flight Center
GTS	Global Telecommunications System
HIRS2	High-Resolution Infrared Sounder 2
IR	infrared
MEaSUREs	Making Earth Science Data Records for Use in Research Environments
MERRA2	Modern-Era Retrospective analysis for Research and Applications Version 2
MeteoSat	Meteorological Satellite
MCTG	Merged CloudSat, Tropical Rainfall Measurement Mission (TRMM), and Global Precipitation Measurement (GPM) misson Climatology
MSU	Microwave Sounding Unit
MTSat	Multi-functional Transport Satellite
NASA	National Aeronautics and Space Administration
NCEI	National Centers for Environmental Information
NetCDF	Network Common Data Format
NOAA	National Oceanic and Atmospheric Administration
OPI	Outgoing Longwave Radiation (OLR) Precipitation Index
PERSIANN	Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks
PMW	passive microwave
SG	Satellite-Gauge
SRT	Sensor Research Team
SSMI	Special Sensor Microwave/Imager
SSMIS	Special Sensor Microwave Imager/Sounder
SSMT2	Special Sensor Microwave/Temperature 2
SWE	snow water equivalent
TCC	Tropical Combined Climatology
TOVS	Television InfraRed Operational Satellite (TIROS) Operational Vertical Sounder
TRMM	Tropical Rainfall Measuring Mission
WCRP	World Climate Research Programme

Appendix B. Data Set Sources

GPCC Full and Monitoring Gauge Analyses
https://opendata.dwd.de/climate_environment/GPCC/html/download_gate.html

3-Hour PERSIANN-CDR
ftp://persiann.eng.uci.edu/CHRSdata/PERSIANN-CDR/adj_3hB1

TOVS
Legacy data currently only available upon request from George Huffman (george.j.huffman@nasa.gov)

AIRS
https://disc.gsfc.nasa.gov/datasets/AIRG2SSD_IRonly_006/summary

SSMI/SSMIS GPROF2010v2
http://rain.atmos.colostate.edu/RAINMAP10v2/

MERRA-2
https://disc.gsfc.nasa.gov/datasets/M2I1NXASM_5.12.4/summary?keywords=merra-2

The authors would like to acknowledge the effort of Dan Braithwaite of UC-Irvine for making available the background 3-hour PERSIANN-CDR files during the development of the GPCPV3 monthly product.