

Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation

George J. Huffman (1), David T. Bolvin (1,2), Robert Joyce (1,2), Owen A. Kelley (1,3),
Eric J. Nelkin (1,2), Jackson Tan (1,4), Daniel C. Watters (5), B. Jason West (1,6)

(1) Earth Sciences Division, NASA Goddard Space Flight Center

(2) Science Systems and Applications, Inc.

(3) George Mason University

(4) University of Maryland Baltimore County

(5) NASA Marshall Space Flight Center

(6) KBR, Inc.

13 July 2023

“Recent” News

See Section 11, “News Archive” for previous items.

11 July 2023 Starting in February 2023, the NRT MetOp 1C data stream has occasionally provided partial orbit granules that duplicate previous granules, but contain faulty data. PPS has now instituted QC procedures that can successfully screen out 12 of 16 known cases. Further work is necessary to capture the rest of the cases without mistakenly removing good granules.

2 July 2023 Early and Late switched to using the new 2 km Autosnow, replacing the hemispheric version.

14 June 2023 NOAA21 data entered Early and Late Runs as of 10:00 UTC 14 June 2023.

18 April 2023 A 2-reaction wheel safhold patch was installed on the GPM-CO. This required taking the satellite out of science operations during the period 13:13-21:16 UTC.

6 January 2023 GOES-18 began serving as GOES-West in CPC's Merged-IR products starting from 4 January 2023 at 18:00 UTC. Also, Himawari-8 was replaced by Himawari-9 starting from 12 December 2022 at 19:00 UTC.

Contents

“Recent” News

Contents

Keywords

README Summary

1. Data Set Names and General Content
2. Related Projects, Data Networks, and Data Sets
3. Data Archive Information
4. Data Set Information
5. Sensors

6. Definitions and Defining Algorithms
7. Error Detection and Correction
8. Missing Value Estimation and Codes
9. Quality and Confidence Estimates
10. Documentation
11. News Archive

Keywords (searchable as *keyword*, except as noted)

2BCMB
3B42
3B42RT
3IMERGHH
3IMERGHH data fields
3IMERGM
3IMERGM data fields
accuracy
acronyms and jargon
advice for new users
AMSR2
AMSR-E
AMSU-B
archive and distribution sites
ATMS
Autosnow
calibrated precipitation field
calibration
calibration outside the CORRA area of coverage
changes from V06 to V07
completely missing fields
Constellation sensor error detection/ correction (see “GMI, TMI, and constellation sensor error detection/ correction”)
controlling factors on dataset performance
CORRA-G (see “GPM Combined Radar-Radiometer Algorithm”)
CORRA-T (see “TRMM-era Combined Radar-Radiometer Algorithm”)
data access policy
data file layout
data providers
data set archive
data set creators
data set file names
data set inventory
data set name
data set provenance
data user registration

diurnal cycle
documentation custodian
documentation revision history
DOI
DPR
DPR and PR error detection/correction
Early Run (see Section 4)
estimate missing values
ERA5 (see “numerical analysis products”)
file date/time
final post-processing
Final Run (see Section 4)
Frequently Asked Questions (FAQ)
GEOS FP (see “numerical analysis products”)
Giovanni
GMI
GMI, TMI, and constellation sensor error detection/correction
GMI, TMI, and constellation sensor Level 2 precipitation datasets
GPM
GPCP-calibrated CORRA
GPM Combined Instrument (see “GPM Combined Radar-Radiometer Algorithm”)
GPM Combined Radar-Radiometer Algorithm
GPM constellation
GPM Core Observatory
GPM data access pages
GPROF
gray-shaded text
grid
GV
High Quality precipitation (see “merged PMW precipitation”)
HQ (see “merged PMW precipitation”)
*MPW*precipitation (see “merged PMW precipitation”)
IMERG
IMERG in the TRMM era
IMERG processing scheme
IMERG-TMPA design comparison
interannual differences driven by data set calibrators (see “controlling factors on dataset performance”)
intercomparison results
IR
IR data correction
IR sensor error detection/correction
*IR*influence(see “Kalman filter influence for IR field”)
*IR*precipitation (see “IR precipitation field”)
IR precipitation field
JMA Forecast (see “numerical analysis products”)

Kalman filter influence for IR field
Kalman Smoother framework
known data set access issues
known errors and anomalies
Late Run (see Section 4)
Latency
Merged 4-Km IR Tb data set
merged PMW precipitation
MERRA-2 (see “numerical analysis products”)
MHS
missing data (see “known errors and anomalies”)
multiple Runs, sources, and formats
MWI
navigation error for GPM and TRMM
numerical analysis products
obtaining data
period of record
PDIR
PMM
PMW satellite overpass times
PPS
PR
PR error detection/correction (see “DPR and PR error detection/correction”)
precipitation gauge adjustment
precipitation gauge analysis
precipitation phase
precipitation post-processing step (“precipitation gauge adjustment”)
precipitation variable
precipitationCal (see “calibrated precipitation field”)
precipitationUncal (see “uncalibrated precipitation field”)
probabilityLiquidPrecipitation (see “probability of liquid phase field”)
probability of liquid phase field
processing during the TRMM/GPM overlap period
processing strategy for each version
production and updates
quality index
randomError (see “random error field”)
random error field
rain gauge analysis (see "precipitation gauge analysis")
read a file of data
References
Release Notes
SAPHIR
sensors contributing to IMERG
SG combination
SHARPEN

similar data sets
smooth-fill
spatial coverage
spatial resolution
SSMI
SSMIS
standard missing value
STORM
temporal resolution
THOR
time zone
TMI
TMI error detection/correction (see “GMI, TMI, and constellation sensor error
detection/correction”)
TMPA (see “3B42”)
TMPA-RT (see “3B42RT”)
transition from TMPA to IMERG
TRMM
TRMM data access pages
TRMM end of mission issues
TRMM-era Combined Radar-Radiometer Algorithm
uncalibrated precipitation field
units of the IMERG estimates
Web Resources
X-Cal Working Group
yellow-shaded text

README Summary

The Integrated Multi-satellite Retrievals for GPM (**IMERG**) is the unified U.S. algorithm that provides the multi-satellite precipitation product for the U.S. GPM team. The precipitation estimates from the various precipitation-relevant satellite passive microwave (PMW) sensors comprising the GPM constellation (which includes those during the Tropical Rainfall Measuring Mission [TRMM] era) are computed using the 2021 version of the Goddard Profiling Algorithm (GPROF2021), then gridded, intercalibrated to the GPM Combined Radar Radiometer Analysis product (with Global Precipitation Climatology Project [GPCP] climatological calibration), and combined into half-hourly $0.1^\circ \times 0.1^\circ$ fields. These are provided to the Climate Prediction Center (CPC) Morphing-Kalman Filter (CMORPH-KF) quasi-Lagrangian time interpolation procedure and the re-calibration procedure applied to the Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks (PERSIANN) Dynamic Infrared–Rain Rate (PDIR) infrared (IR) precipitation retrievals. In parallel, the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA2) and Goddard Earth Observing System Forward Processing (GEOS FP) precipitation (PRECTOT), total precipitable liquid water (TQL), and total precipitable water vapor (TQV) fields are used in the CMORPH-KF quasi-Lagrangian time interpolation procedure. The PDIR estimates are computed (supported by an asynchronous re-

calibration cycle) and sent to the CMORPH-KF quasi-Lagrangian time interpolation procedure. The Kalman Filter (KF) quasi-Lagrangian time interpolation (supported by an asynchronous KF weights updating cycle) uses the PMW and IR estimates to create half-hourly estimates. The IMERG system is run twice in near-real time

- “Early” multi-satellite product ~4 hr after observation time and
 - “Late” multi-satellite product ~14 hr after observation time,
- and once after the monthly gauge analysis is received
- “Final” satellite-gauge product ~3.5 months after the observation month.

The baseline is for the post-real-time Final Run half-hour estimates to be calibrated so that they sum to the monthly satellite-gauge Final Run combination. For all Runs, the output contains multiple fields that provide information on the input data, selected intermediate fields, and estimation quality.

The current standard reference is Huffman et al. (2020), augmented by the IMERG ATBD (Huffman et al. 2023), posted at

<https://gpm.nasa.gov/resources/documents/imerg-v07-atbd.pdf>.

Notes:

- This document only discusses the “native” half-hourly and monthly (for Final Run) IMERG datasets. Documentation for “value-added” datasets created by PPS, GES DISC, and other archive sites is created and held by those sites.
- **gray-shaded text** denotes information that pertains to time periods and instruments not currently used in IMERG. This information is not germane to the current IMERG data sets, but is included in anticipation of subsequent upgrades. Any other shading, specifically **yellow-shaded text**, denotes information that is not yet finalized.

1. Data Set Names and General Content

The formal **data set name** is the Integrated Multi-satellite Retrievals for GPM (IMERG). Note that there are other products in the general GPM data inventory, so it is important to be specific about the product being used, including the Run (e.g., Early, Late, or Final) and the version number (e.g., V07).

.....

The **IMERG processing scheme** is as follows:

The IMERG system is run twice in near-real time

- “Early” multi-satellite product ~4 hr after observation time and
 - “Late” multi-satellite product ~14 hr after observation time
- and once after the monthly gauge analysis is received
- “Final” satellite-gauge product ~3.5 months after the observation month.

All three Runs create half-hourly $0.1^\circ \times 0.1^\circ$ products (3IMERGHH), while the Final post-real-time Run additionally provides a monthly $0.1^\circ \times 0.1^\circ$ satellite-gauge combination data set (3IMERGM). In all cases the output contains multiple fields that provide information on the input data, selected intermediate fields, and estimation quality. See “multiple Runs, sources, and formats” for more information.

In earlier versions, GPM data sets were computed for the GPM era, starting in March 2014. [However, the actual start date was the middle of the month, so we suggested ignoring data before April.] Now, the GPM data sets use TRMM as a calibrator for the period from the start of TRMM observations to several months after the launch of the GPM Core Observatory to allow GPM processing spin up and a graceful transition from the TRMM era to the GPM era (starting with June 2014). The GPM home page is located at <https://gpm.nasa.gov/>.

The current standard reference is Huffman et al. (2020), augmented by the IMERG ATBD, posted at

<https://gpm.nasa.gov/resources/documents/imerg-v07-atbd.pdf>.

.....

2. Related Projects, Data Networks, and Data Sets

The **data set creators** are G.J. Huffman, D.T. Bolvin, E.J. Nelkin, R. Joyce, and J. Tan working in the Earth Sciences Division-Atmospheres, NASA Goddard Space Flight Center, Greenbelt, Maryland; E.F. Stocker, working in the Precipitation Processing System (PPS), NASA Goddard Space Flight Center, Code 619, Greenbelt, Maryland; and the other members of the GPM Multi-Satellite Team: Dan Braithwaite Kuolin Hsu, and Phu Nguyen (Univ. of Calif. Irvine); Christopher Kidd (Univ. of Maryland College Park; NASA/GSFC 612); and Pingping Xie (NOAA/NWS/CPC).

.....

The work is being carried out as part of the Global Precipitation Measurement (**GPM**) mission, an international project by NASA and JAXA designed to provide improved estimates of precipitation over most of the globe, following the highly successful Tropical Rainfall Measuring Mission (see “TRMM”). GPM has four key components: the GPM Core Observatory, the GPM Constellation, the Precipitation Processing System (PPS), and GPM Ground Validation (GV) team. See each topic for more details. In addition, the GPM Applications and Outreach team works extensively with current and prospective IMERG users.

.....

The **GPM Core Observatory** was launched on 27 February 2014 (UTC) into an (83-day) precessing orbit with a 65° inclination, a period of about 95 min., and an altitude of 407 km. This orbit allows GPM to build up a complete view of the climatological diurnal cycle, as well as providing calibration for other precipitation-relevant sensors in Sun-synchronous orbits in the GPM constellation. The calibrations are based on data from the GMI, a passive microwave sensor

provided by NASA, and the GPM DPR, a Ku- and Ka-band electronically scanning radar provided by JAXA and are applied in the IMERG processing.

.....

The **GPM constellation** is a virtual constellation of satellites carrying precipitation-relevant sensors. The satellites are “of opportunity” from various U.S. and international agencies, with no coordination of function or orbit other than providing data to GPM as quickly as possible. The GPM constellation has been designated as an official CEOS Virtual Constellation. The constellation members are listed in the “sensors contributing to GPM” section.

.....

The Precipitation Processing System (**PPS**) is responsible for working with the GPM Science Team algorithm developers to create an end-to-end processing system that ingests raw satellite data and ancillary data, then computes and distributes the precipitation and related products.

.....

The GPM Ground Validation **GV** team is responsible for assembling the necessary data and computing validation and error statistical products. In some cases this includes organizing or participating in field experiments to collect high-quality data that are not otherwise available for particular climatic zones.

.....

The Tropical Rainfall Measuring Mission (**TRMM**) was an international project of NASA and JAXA designed to provide improved estimates of precipitation in the Tropics, where the bulk of the Earth's rainfall occurs. The TRMM satellite began recording data in December 1997 in a (46-day) precessing orbit at a 35° inclination with a period of about 91.5 min. This orbit allowed TRMM to build up a complete view of the climatological diurnal cycle, as well as providing calibration for other precipitation-relevant sensors in Sun-synchronous orbits. TRMM exhausted its station-keeping fuel in July 2014 and began to descend (due to atmospheric drag), with passivation (shutdown) in April 2015 and reentry in June 2015. The TRMM home page is located at <https://nasa.gov/missions/trmm>.

.....

Precipitation Measurement Missions (**PMM**) is the umbrella organization within NASA to pursue satellite projects that advance precipitation science. To date, the two such projects are TRMM and GPM. The PMM home page is located at <http://gpm.nasa.gov/>.

.....

IMERG draws on data from numerous **data providers**:

1. NASA/GSFC Level 1 GMI Tb's;
2. NASA/GSFC Level 2 GPM Combined Instrument (DPR-GMI) precipitation;
3. NASA/GSFC Level 1 TMI Tb's;
4. NASA/GSFC Level 2 TRMM Combined Instrument (PR-TMI) precipitation;
5. NSIDC Level 1 AMSR-E Tb's;
6. NSIDC Level 1 AMSR2 Tb's;

7. NOAA/NCDC CLASS Level 1 SSMI Tb's;
8. NOAA/NCDC CLASS Level 1 SSMIS Tb's;
9. NOAA/NCDC CLASS Level 1 (NOAA-series) AMSU-B Tb's;
10. NOAA/NCDC CLASS Level 1 (NOAA- and MetOp-series) MHS Tb's;
11. NOAA/NCDC CLASS Level 1 (NOAA- and MetOp-series) ATMS Tb's;
12. CNES ICARE (alternatively ISRO MOSDAC) Level 1 SAPHIR Tb's;
13. NESDIS Level 1 ATMS Tb's;
14. NOAA/NWS/CPC Level 3 Merged 4-Km Geostationary Satellite IR Tb Data;
15. GPCC Level 3 Full and Monitoring Precipitation Gauge Analyses;
16. NASA/GMAO MERRA-2 Reanalysis product and GEOS FP Forecast product;
17. EUMETSAT ERA5;
18. JMA forecast products; and
19. NOAA Autosnow.

See “sensors contributing to GPM” for more details.

Some of these data sets extend beyond the TRMM and GPM periods in their original archival locations.

.....

There are numerous **similar data sets**, although no other quite matches all the attributes of being routinely produced for multiple latencies, publicly available, fine-scale in space and time, quasi-global, available from June 2000 onwards, and subsequently from January 1998 when IR data issues are resolved, intercalibrated, and formed by combining multiple data sources including precipitation gauges. The IPWG data set tables at <http://www.isac.cnr.it/~ipwg/data/datasets.html> provide a good listing of other precipitation data sets. The closest include the set of estimates based on:

1. Huffman et al. (2007a, 2007b, 2010): the legacy Version 7 TMPA and TMPA-RT estimates;
2. Sorooshian et al. (2000): the PERSIANN neural network calibrates IR with microwave;
3. Joyce et al. (2011): the CMORPH morphing scheme time-interpolates microwave patterns with IR-based motion vectors, and incorporates IR-based precipitation estimates;
4. Kubota et al. (2007): the GSMaP system applies a quasi-Lagrangian time-interpolation scheme similar to CMORPH; and
5. Beck et al. (2017): MSWEP combines multiple precipitation estimations to create a “best” global estimate.

Several SSMI/SSMIS-based data sets are available as gridded single-sensor data sets, including those computed with the GPROF algorithms (based on Kummerow et al. 1996); and the NOAA Microwave Integrated Retrieval System (MiRS) algorithm (Boukabara et al. 2011) among others. Other daily, single-sensor data sets are available for open-water regions based on SSMI/SSMIS data (RSS, Wentz and Spencer 1998; HOAPS, Andersson et al. 2010), MSU data (Spencer 1993), AMSR-E, and AMSU-B/MHS data. Several daily single-sensor or combination data sets are available at the regional scale, but are not really "similar." At the monthly scale over open-water regions, SSMI/SSMIS data are used in the Chang/Chiu/Wilheit emission algorithm (Wilheit et al. 1991, Chiu and Chokngamwong 2010).

.....
The **transition from TMPA to IMERG** occurred as

1. initial production of IMERG for the GPM era,
2. extension of IMERG to the TRMM era at the V06 GPM reprocessing in Summer 2019, and
3. continued production of the Version 7 TRMM 3B42/43 and 3B42RT until IMERG was considered stable, at the end of December 2019.

An extended discussion of this topic, with current details is contained in the document *The Transition in Multi-Satellite Products from TRMM to GPM (TMPA to IMERG)*, which is available at https://gpm.nasa.gov/sites/default/files/document_files/TMPA-to-IMERG_transition.pdf. Fortunately, the scenarios under which a less graceful transition might have become necessary did not come to pass.

.....

The Version 7 TRMM product **3B42** was computed with the TMPA in post-real time, and constituted the research-grade archive of TMPA estimates. Note that the version numbering for the TMPA-RT and official TRMM products are not necessarily related, although both ended as number 7. The post-real time computation allowed several improvements in 3B42 compared to 3B42RT:

1. Data were processed starting with the first full month of TRMM data, which began 1 January 1998.
2. The IR calibration period was the calendar month in which the observation time fell, rather than a trailing 30-day accumulation.
3. The TRMM Combined Instrument product (2B31) was used as the calibrating standard month-to-month, which gave better estimates than the climatological calibration used in the TMPA-RT.
4. For each grid box, the individual 3B42 3-hourly precipitation values were scaled to sum to a combination of monthly 3B42 and gauge analysis, which was TRMM product 3B43.

The first set of reprocessed data for Version 7 was posted in May 2012, while the second was posted in December 2012.

With the end of routine TRMM PR precipitation estimates on 7 October 2014 (see “TRMM end of mission issues”), the last month that fully conforms to the processing described above is September 2014. Thereafter a climatological calibration was developed that minimizes the data discontinuity (Huffman and Bolvin 2015). Both the production and RT TMPA systems were superseded by IMERG, although they continued to be computed until the long-record IMERG was considered operational, ending with December 2019. See “transition from TMPA to IMERG” for more details.

.....

The Version 7 Real-Time TRMM product **3B42RT** was computed with the TMPA in near-real time, and constituted the most timely source of TMPA estimates. The near-real time computation required several simplifications in 3B42RT compared to 3B42:

1. The IR calibration period was a trailing approximately 30-day accumulation, rather than the calendar month in which the observation time falls.
2. Static monthly climatological calibrations were used to calibrate GPROF-TMI to TRMM Combined Instrument product (2B31) as 2B31 was not available in real time, and this climatological calibration continued to be used after TRMM PR data ceased.
3. In near-real time it was not possible to include precipitation gauge data. Rather, fixed monthly climatological corrections based on 3B43 and gauge were applied to the Version 7 real time 3B42 that varied by month and location.

The first set of retrospectively processed data for Version 7 3B42RT was released in June 2012 using the Version 7 3B42 for calibration, while the second set was posted in December 2012. This version includes the following:

1. SSMIS data were introduced based on interim calibration developed in conjunction with D. Vila (ESSIC).
2. The RT system was retrospectively processed back to 1 March 2000 using the full satellite data sets available in the Version 7 production system. The main difference from true RT processing is that the production input data records are somewhat more complete than those available in real time. The start date is driven by the start date of the then-available CPC Merged 4 Km IR data record. The TMPA-RT continued to be computed until IMERG was considered operational (eventually providing a record through the end of 2019). See “transition from TMPA to IMERG” for more details.

.....

3. Data Archive Information

Production and updates for IMERG are a joint activity of the precipitation research group in NASA Goddard Space Flight Center in the Earth Sciences Division-Atmospheres and PPS.

The latency of the various IMERG Runs is governed by the latency of the individual input products. See “multiple Runs, sources, and formats” and “latency” for more details.

Updates will be released to (1) extend the data record, (2) take advantage of improved combination techniques, or (3) correct errors. Updates resulting from the last two cases will be given new version numbers.

NOTE: The changes described in this section are typical of the changes that are required to keep IMERG abreast of current requirements and science. Users are strongly encouraged to check back routinely for additional upgrades and to refer other users to this site rather than redistributing data that are potentially out of date.

In some cases, such as the failure of AMSR-E, the end of a data record is clear. In other cases, such as the gradual failure of the NOAA-16 AMSU-B during 2010, the point at which to end use of the data is a matter of judgment. In the latter case we chose 30 April 2010 despite continued operation into early 2011 based on apparent issues caused by these data in the TRMM 3B42RT.

In V07 we provide the first fully global IMERG products. For a discussion of possible future enhancements, refer to the sections “Pre-Planned Product Improvements” and “Options for Processing” in the IMERG ATBD.

.....

Several **TRMM end of mission issues** impacted the IMERG system:

1. On 7 October 2014, routine production ended for the TRMM PR precipitation estimates due to the ongoing descent and ultimate decommissioning of the TRMM satellite (PMM, 2014). [PR data were briefly available from 12 February to 1 April 2015 as TRMM descended past its original altitude of 350 km.] Estimates from the TMI continued to be produced until it was turned off on 8 April 2015 as part of the TRMM decommissioning.
2. TMI continued to be included as one of the input data sets for IMERG until it was ended on 8 April 2015 as part of the TRMM decommissioning activities.

.....

The **archive and distribution sites** for the official release of IMERG (and all other GPM products) are listed on the GPM data access pages

<https://gpm.nasa.gov/data/directory>

which point to archives hosted at the Goddard Earth Sciences Data and Information Services Center (GES DISC) <https://disc.gsfc.nasa.gov/datasets?keywords=gpm%20imerg%2007> and the Precipitation Processing System (PPS) https://arthurhou.pps.eosdis.nasa.gov/Documents/Master_List_of_PPS_Data_Products.html, as well as several interactive sites. Please visit <https://gpm.nasa.gov/data/news> for the latest updates. Note that users are required to complete a simple, free, and automatic on-line registration in order to access the data; see “data user registration” for more details. This is required by data security considerations.

The responsible archive organizations are:

Precipitation Processing System (PPS)

NASA Goddard Space Flight Center

Code 619

Greenbelt, MD 20771 USA

Email: helpdesk@pps-mail.nascom.nasa.gov

Web site: <https://pps.gsfc.nasa.gov>

Goddard Earth Sciences Data and Information Services Center (GES DISC)

NASA Goddard Space Flight Center

Code 619

Greenbelt, MD 20771 USA

Phone: +1-301-614-5224

Fax: +1-301-614-5268

Email: gsfc-help-disc@lists.nasa.gov

Web site: <https://disc.gsfc.nasa.gov>

The Level 1C and 2 input satellite data sets are available in the PPS archive. The other input data, as well as the original (i.e., non-NASA input data) are available in independent archive and distribution sites, and contact information may be obtained through PPS. A comprehensive listing of PPS and GES DISC holdings, including value-added products, is provided on the GPM data access pages

<https://gpm.nasa.gov/data/directory>.

.....

The native **data set archive** consists of Version 5 Hierarchical Data Format (HDF5) files. Each half-hour dataset (3IMERGHH) or monthly dataset (3IMERGM) is contained in a separate file with standard self-documenting HDF5 metadata. The data are distributed via the Internet. Each 3IMERGHH file is approximately 10 MB (with internal compression), while the 3IMERGM files are each about 30 MB (with internal compression).

The full collection of IMERG Final, Late, and Early Run original files is provided and archived by PPS at <https://storm.pps.eosdis.nasa.gov/storm/> and by GES DISC at the DOI (see “DOI” for the correct values). Secondary archives of these data files exist outside of PPS and GES DISC, but users are responsible for gaining a clear understanding of the provenance of those data to assure that they are working with current, clean data.

In addition, PPS and GES DISC provide a variety of value-added products, which are listed on the GPM data access pages

<https://gpm.nasa.gov/data/directory>.

Finally, quick-look imagery can be created using the STORM and THOR applications, interactive Web-based analysis for selected IMERG precipitation estimates and related fields is provided through Giovanni, and a selection of event-based and routine visualizations is listed at <https://gpm.nasa.gov/resources>. Note particularly the “latest week of IMERG”, which is provided in a variety of formats at <https://svs.gsfc.nasa.gov/4285>. See the respective topics for details.

Note that users are required to complete a simple, free, and automatic on-line registration in order to access the data; see “data user registration” for more details. This is required by data security considerations.

.....

The short names and **DOI** list for the original IMERG HDF5 files are

| | | |
|--------------------|----------------|-------------------------------|
| Final half-hourly: | 3IMERGHH | 10.5067/GPM/IMERG/3B-HH/07 |
| Final monthly: | 3IMERGM | 10.5067/GPM/IMERG/3B-MONTH/07 |
| Early half-hourly: | 3IMERGHH_EARLY | 10.5067/GPM/IMERG/3B-HH-E/07 |
| Late half-hourly: | 3IMERGHH_LATE | 10.5067/GPM/IMERG/3B-HH-L/07 |

.....

Known data set access issues include:

1. Besides the primary data sites (see “data set archive”), several “mirror” and value-added archive sites outside of PPS and GES DISC provide the IMERG data sets and/or value-added

products in their holdings. Users availing themselves of these sites should work with the personnel in charge of those sites to iron out access problems. Also, users are responsible for gaining a clear understanding of the provenance of those data to assure that they are working with current, clean data.

2. FTP access is replaced with https pages, including <https://jsimpsonhttps.pps.eosdis.nasa.gov/> and <https://arthurhouhttps.pps.eosdis.nasa.gov/> for near- and post-real-time data at PPS.
3. The IMERG data sets are in IEEE “little-endian” floating-point format. Some Unix-based computers run in “big-endian”, meaning the data might need to be “byte-swapped” to be useful, although HDF libraries should take care of this detail.
4. Note that users are required to complete a simple, free, and automatic on-line registration in order to access the data; see “data user registration” for more details. This is required by data security considerations.
5. The dimensions of the (native) HDF5 data fields changed from (lon, lat) in V05 to (time, lon, lat) in V06 and V07, with the corresponding lengths of (1, 3600, 1800). This is done to conform to GES DISC archiving conventions. In most cases this should be transparent to application programs, but some users have reported needing to re-work code that makes individual calls to the HDF API.

.....
Data user registration is required at PPS and GES DISC to satisfy modern data security considerations. This registration is simple, free, automatic, and on-line in both cases. For PPS and GES DISC, respectively, go to

<https://registration.pps.eosdis.nasa.gov/registration/>
<https://disc.gsfc.nasa.gov/registration/registration-for-data-access>

.....
The **data set inventory** may be obtained by accessing the IMERG product listings on the GPM data access pages:

Data Access Homepage: <https://gpm.nasa.gov/data/directory>
GPM Data Directory: <https://gpm.nasa.gov/data/directory>
IMERG Early Run Homepage: <https://gpm.nasa.gov/data/directory> , Level 3 tab, Early Run
IMERG Late Run Homepage: <https://gpm.nasa.gov/data/directory> , Level 3 tab, Late Run
IMERG Final Run Homepage: <https://gpm.nasa.gov/data/directory> , Level 3 tab, Final Run

.....
Users interested in **obtaining data** can follow the links on the GPM data access pages

Data Access Homepage: <https://gpm.nasa.gov/data/directory>
GPM Data Directory: <https://gpm.nasa.gov/data/directory>
IMERG Early Run Homepage: <https://gpm.nasa.gov/data/directory> , Level 3 tab, Early Run
IMERG Late Run Homepage: <https://gpm.nasa.gov/data/directory> , Level 3 tab, Late Run
IMERG Final Run Homepage: <https://gpm.nasa.gov/data/directory> , Level 3 tab, Final Run

to determine the correct Run/format/service for their particular needs and then navigate to the correct data site. Note that users are required to complete a simple, free, and automatic on-line

registration for both PPS and GES DISC in order to access the data; see “data user registration” for more details. This is required by data security considerations.

It is possible to subset some of the data sets by parameter and/or space, as well as acquiring only the data files that correspond to the user’s time period of interest. See the documentation for the individual dataset formats/services to determine what parameter/space subsetting is supported.

As well, web-based interactive access to the IMERG and related data is provided by Giovanni; see that topic for details.

.....

The **data access policy** is "freely available" with three common-sense caveats:

1. It is an emerging best practice that the data set source should be referenced when the data are used. Unless a journal/publication has a policy on format, formal reference of the form

Huffman, G.J., E.F. Stocker, D.T. Bolvin, R. Joyce, E.J. Nelkin, Jackson Tan, 2023, last updated 2023: *<dataset identifier> Data Sets*. NASA/GSFC, Greenbelt, MD, USA, *<archive site – PPS or GES DISC>*, accessed *<enter user data access date>*, [doi:*<doi>* or at *<data landing page URL>*].

is suggested following the AMS policy statement at

<https://www.ametsoc.org/index.cfm/ams/publications/ethical-guidelines-and-ams-policies/data-policy-and-guidelines/>.

Note that the AMS policy states that this dataset reference should be provided in addition to references to the relevant papers on constructing the data set. This approach allows readers to find both the technical literature and the data archive.

2. New users are urged to obtain their own current, clean copy from an official archive, rather than taking a version from a third party that might be damaged or out of date.
3. Errors and difficulties in the original datasets should be reported to the dataset creators or archive sites (depending on the nature of the issue). Access and format issues at third-party sites should be directed to the respective contacts.

Note that users are required to complete a simple, free, and automatic on-line registration for both PPS and GES DISC in order to access the data; see “data user registration” for more details. This is required by data security considerations.

.....

The Tool for High-resolution Observation Review (**THOR**) is created and supported by PPS. THOR is a point-and-click program written in IDL that runs on Linux, Mac OS X, and Windows. This viewer enables displaying TRMM and GPM observations on a map of the Earth at the full instrument resolution. For use, it must be installed on each user’s computer; instructions for downloading and installing the current version of THOR are located at

<https://pps.gsfc.nasa.gov/thorrelease.html>.

As well, THOR is accessible within the PPS STORM (without downloading the application) to quickly visualize data fields.

.....
The Science Team Online Research Module (**STORM**) is a publicly available Web-based data access interface for PPS. The goals of the 3D Visualization Pages are:

1. Provide a way to visualize near-real-time data as soon as they become available on the web server. This is achieved through both the GPM Near Real Time Viewer and the STORM Event Viewer (EV)/EV Mini pages.
2. Allow users of the STORM data order interface to view products in three dimensions prior to order. For this, STORM Virtual Globe (VG) was created to complement the static images and THOR Online tools.

The home page is <https://storm.pps.eosdis.nasa.gov/storm/>,
and the user guide is posted at <https://storm.pps.eosdis.nasa.gov/storm/STORMUserGuide.pdf>.

.....

Giovanni, formerly the Geospatial Interactive Online Visualization ANd aNalysis Infrastructure, is created and supported by GES DISC. It provides a web-based resource for accessing many Earth science data sets, including IMERG Runs. It performs a variety of basic subsetting, time- and space-averaging, and output of results in plots, time series, animations, or ASCII text. The current version of Giovanni is located at

<https://giovanni.gsfc.nasa.gov/giovanni/>.

Note that an Earthdata login is required for full functionality, and registration is free and automatic.

.....

4. Data Set Information

Advice for new users: The variety of products that provide IMERG data can be daunting to a new user. These different products serve different needs, as outlined in “multiple Runs, sources, and formats”, which each user must map into their particular needs, level of computing skills, and computing tools. As well, first-time users are required to complete a simple, free, and automatic on-line registration both for PPS and for GES DISC in order to access the data; see “data user registration” for more details. This is required by data security considerations. Once a user has acquired an example dataset, it is strongly suggested that they use Giovanni (which topic see) to verify that they are correctly generating results from the data set format that they’ve chosen to use. The various IMERG dataset formats have been developed to work nicely with a range of off-the-shelf application software packages. If this is not successful, the user should contact the archive site for more information. Questions about the scientific basis for IMERG that are not answered in this technical document, the Frequently Asked Questions (FAQ), or the ATBD should be referred to the algorithm developers (see “documentation custodian”).

.....

IMERG output is available in **multiple Runs, sources, and formats**, which can lead to confusion among users. The IMERG system is run twice in near-real time

- “Early” multi-satellite product ~4 hr after observation time and
 - “Late” multi-satellite product ~14 hr after observation time,
- and once after the monthly gauge analysis is received
- “Final” satellite-gauge product ~3.5 months after the observation month.

All three Runs create half-hourly $0.1^\circ \times 0.1^\circ$ products (3IMERGHH), while the post-real-time Final Run additionally provides a monthly $0.1^\circ \times 0.1^\circ$ satellite-gauge combination data set (3IMERGM). In all cases the output contains multiple fields that provide information on the input data, selected intermediate fields, and estimation quality.

These multiple Runs are computed because of the competing demands for timeliness and completeness of data. The Early Run provides relatively quick results for flood analysis and other short-fuse applications by only employing forward morphing. The Late Run employs both forward and backward morphing with later data, and is appropriate for daily and longer applications, such as crop forecasting. The Final Run introduces monthly precipitation gauge analyses, providing more accurate results in regions with gauge information. The Final Run is considered the research-grade product.

Once computed, the IMERG data sets are hosted on both PPS and GES DISC, with each site providing value-added products based on the basic IMERG data sets, such as daily averages, GeoTIFF-format files, and WMS servers. A complete list of datasets provided by PPS and GES DISC is given on the GPM data access pages (<https://gpm.nasa.gov/data/directory>). Note that users are required to complete a simple, free, and automatic on-line registration for both PPS and GES DISC in order to access the data; see “data user registration” for more details. This is required by data security considerations.

Note that the dimensions of the (native) HDF5 data fields changed from (lon, lat) in V05 to (time, lon, lat) for V06 and V07, with the corresponding lengths of (1, 3600, 1800). This is done to conform to GES DISC archiving conventions. In most cases this should be transparent to application programs, but some users have reported needing to re-work code that makes individual calls to the HDF API.

.....

The IMERG V07 **Release Notes** provide a high-level summary of issues that users should consider in using V07 IMERG. It is located at .

.....

There are a number of **changes from V06 to V07** that are collected here for easy access:

- The dimensions of the (native) HDF5 data fields changed in V06 (and continued in V07) from (lon, lat) to (time, lon, lat), with the corresponding lengths of (1, 3600, 1800). This is done to conform to GES DISC archiving conventions. In most cases this should be transparent to application programs, but some users have reported needing to re-work code that makes individual calls to the HDF API.
- Up through V05, the “displacement vectors” used in the quasi-Lagrangian time interpolation scheme were computed from the IR data. In V06 and V07, these vectors are computed from Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2)

and Goddard Earth Observing System model (GEOS) Forward Processing (FP) data, as ingested by PPS. For details about the expanded set of inputs in V07, see “numerical analysis products” and .

- Automated IR quality control was implemented. The geostationary IR Tb field sporadically exhibits distinct artifacts (stripes and arcs) in the Early and Late Runs and required manual intervention in the Final Run. A quality control mechanism was implemented in V07 that automatically scans for anomalous features and attempts to mask out these artifacts. See item 36 in “known errors and anomalies”.
- The averaging in the Kalman filter distorts the distribution of precipitation rates; in response, SHARPEN (which topic see) was developed and implemented in V07 to mitigate this issue (Tan et al. 2021).
- For the first time, IMERG includes PMW estimates over frozen surfaces. Prior to V07, IMERG dropped PMW retrievals for grid boxes with estimated frozen surface types. See “surface coverage” for details, including the lower confidence in those regions.
- Motivated by the potential contribution of propagated precipitation (You et al. 2021), instantaneous PMW estimates are merged with propagated precipitation estimates—except the combination of instantaneous imagers and propagated sounders is not permitted over ocean—in the Kalman filter, leading to reduced inter-sensor differences and slightly improved performance in standard metrics. Furthermore, SHARPEN is now applied when the merger is performed. These changes increase the half-hourly Quality Index values on average and give a smoother visual appearance to animations.
- The Legates-Wilmott scheme for undercatch in the GPCC gauges is too large at high latitudes, so a climatological Fuchs adjustment was implemented in V07 over Eurasia north of 45°N, while Legates-Wilmott continues to be used for all other land areas. See “precipitation gauge adjustment”.
- To minimize misinterpretation of the variable names, reflect changes in the algorithm, and emphasize the most commonly accessed fields, several variables were renamed and/or moved to a group called “Intermediate” within the half-hourly HDF5 file. The V07 name/location for the variable relevant to most users is:
 - o *Grid/precipitationCal* → *Grid/precipitation*Changes in other variables are:
 - o *Grid/HQprecipitation* → *Grid/Intermediate/MWprecipitation*
 - o *Grid/HQprecipSource* → *Grid/Intermediate/MWprecipSource*
 - o *Grid/HQobservationTime* → *Grid/Intermediate/MWobservationTime*
 - o *Grid/IRprecipitation* → *Grid/Intermediate/IRprecipitation*
 - o *Grid/IRkalmanFilterWeight* → *Grid/Intermediate/IRinfluence*
 - o *Grid/precipitationUncal* → *Grid/Intermediate/precipitationUncal*See Tables 1 and 2 for definitions of the variables. The variable names for the monthly file are not changed in V07.
- In V07, GPROF 2021 is used to compute precipitation estimates for all microwave sensors as input.
- In V07, CORRA V07 is used to compute the Combined estimates, with CORRA-T and CORRA-G denoting computations in the TRMM and GPM eras, respectively.
- V07 adapted PDIR–NOW as the IR precipitation algorithm, involving dynamically shifting relationships between the IR brightness temperatures and the precipitation rate (Nguyen et al. 2020), that replaced PERSIANN-CCS. See “PDIR”.

- The V07 estimates of precipitation phase (which topic see) are uniformly calculated using thermodynamic variables from ERA5 for the Final Run, with a correction to the calculation of the wet-bulb temperature. The lookup tables used to estimate the precipitation phase also were recalculated using gridded variables from ERA5. These improvements result in lower probability of liquid phase estimates in V07 compared to V06, on average, especially near the 0°C wet-bulb temperature.
- In V07, GPROF-TMI 2021 estimates are computed for the GPM era and incorporated in the V07 IMERG datasets and incorporated as another PMW sensor, while GPROF-GMI 2021 estimates are computed for the TRMM era and incorporated as another PMW sensor.
- We continue to provide estimates from all constellation members in the microwave-only precipitation field (*MWprecipitation*) and the complete precipitation fields (*precipitation*, *precipitationUncal*) over the fully global domain (90°N-S). Note: microwave estimates are applied over frozen surfaces in the V07 *precipitation* and *precipitationUncal* data fields for the first time. This is considered somewhat uncertain, as reflected in the low half-hourly Quality Index values attributed to these areas.
- SAPHIR estimates are not used in V07 to give additional time to develop estimates that are spatially consistent with other sensors' output.
- The GPROF 2021 estimates for AMSU, ATMS, and MHS do not provide estimates for the 5, 8, and 5 footprints (respectively) at each swath edge due to algorithm issues revealed in early testing.
- At the start of the GPM era (March 2014), V07 continues to use TRMM-based calibrations through May 2014, unlike V05, in which GPM-based calibrations started immediately. This is done to allow the GPM-based calibrations to spin up before they are applied, since calibrations based on fully populated rotating match-up arrays are available from TRMM. During the overlap period for TMI and GMI, the one that is not providing calibration in a given half-hour is treated as just another (high-quality) conically scanning sensor.
- We continue the practice of thresholding input precipitation rates to adjust fractional coverage; GPROF 2021 estimates are thresholded at 0.03 mm/h in V07.
- V07 precipitation rates are rounded to the nearest 0.01 mm / h for half-hourly datasets and 0.001 mm / h for the monthly datasets. Motivated primarily by the improved dataset compression—about 50% reduction in file size—this rounding is not considered to have a practical effect as these thresholds are well below the expected uncertainties for the precipitation rates.
- The maximum permitted precipitation rate is 200 mm/h in V07, raised from 50 in V05 and an accidental mixture of 50 (for IR) and 120 (for PMW) in V06.
- A long-standing bug was fixed in the gridding code for both CORRA and GPROF that caused a spatial offset of one grid box to the east was fixed. This is a critical improvement over all previous versions of IMERG. Details are contained in Appendix 1 of the Release Notes (which topic see for the URL).
- Extensive work was done to improve the PMW intercalibration to correct for biases. Evaluations of IMERG V06 identified a persistent overestimation, which was traced to three factors that were separately addressed:
 1. The mean GMI/TMI precipitation rate is not constant along the scan, so the GMI/TMI-constellation calibration was modified to use the full-swath GMI/TMI matched to the narrower swath of CORRA.

2. The GPROF-GMI, GPROF-TMI, and CORRA footprint sizes differ so CORRA is averaged on $0.2^{\circ} \times 0.2^{\circ}$ and $0.3^{\circ} \times 0.3^{\circ}$ gridbox templates to approximately match GPROF-GMI and GPROF-TMI footprint sizes, respectively.
3. The GPCP adjustment to CORRA in V06 was not capturing the longitudinal variability over land, so the GPCP adjustment is not applied over land in V07.

The net effect of these changes was to neutralize the biases that had been driven by the calibration process; however, regime-dependent biases still remain in the GPROF retrievals.

- Procedural issues in the reprocessing prevented the intended climatological adjustment of the Early and Late Runs to the Final in V06, but it is done in V07. This should ensure more consistency in the bias characteristics among the Early, Late, and Final Runs. See “precipitation post-processing step”. We continue the practice started with V04 of calibrating 2BCMB to GPCP over ocean (at middle and high latitudes). V07 upgrades to GPCP V3.2 from V2.3.
- V06 and V07 estimate the $t=0$ (i.e., actual microwave overpass data, contained in the *HQprecipitation/MWprecipitation* field) correlations in the half-hourly Quality Index; in V05 they were identically one. See the appendix of the V07 ATBD (<https://gpm.nasa.gov/resources/documents/imerg-v07-atbd>).
- “Rippling” in the animation of IMERG data was reduced somewhat from V05 to V06, and markedly in V07, although it continues to be visible. This indicates that the different sensors continue to have different depictions of the same precipitation features due to each sensor type’s unique combination of resolution and channel selection.

.....

What’s up with all the different **calibration** approaches? It is critical to keep various data sources as comparable to each other as possible. As a result, there are intercomparisons and consequent adjustments at Level 1 (the brightness temperatures), Level 2 (the individual precipitation products), and at Level 3 (the gridded precipitation products and combinations). Pretty much all of these interventions go by the name “calibration”, so it is important for the data user to keep in mind which kind of product is being “calibrated”, and what the source and target of the calibration are at any given time. One key concept is that the calibration action is considered a success if the adjusted target dataset replicates the statistics/behavior of the source dataset to the extent that the calibration procedure can cause such changes. Other aspects of the target dataset are likely left (largely) unchanged. For example, our scheme for calibrating the half-hourly satellite time series in the Final by the GPCC monthly gauge will affect the bias of the calibrated half-hourly, but leave the occurrence of precipitation untouched.

.....

The identifier **3IMERGHH** denotes the half-hourly output from any Run of IMERG.

.....

The identifier **3IMERGM** denotes the monthly output, only computed by the Final Run of IMERG. Note that other value-added products might include monthly accumulations, which are not the same.

.....

The **file date/time** is the UTC year, month, day, and then starting and ending hour, minute, second for both the 3IMERGHH and 3IMERGM data sets. The single date provided is the first day of the month in the case of 3IMERGM; basically everything after year and month in its name is filled with nominal values. The date, start time, and end time are provided in both the metadata and file name. Within the name, start and end times are denoted by “S” and “E”. All dates and times are UTC.

.....

The template for **data set file names** for the original IMERG HDF5 files is given in https://gpm.nasa.gov/sites/default/files/document_files/FileNamingConventionForPrecipitationProductsForGPMMissionV1.4.pdf

Examples are:

3IMERGHH from the Early Run starting at 00UTC on 1 July 2014:

3B-HHR-E.MS.MRG.3IMERG.20140701-S000000-E002959.0000.V07.RT-H5

3IMERGM (always the Final Run) for July 2014:

3B-MO.MS.MRG.3IMERG.20140701-S000000-E235959.07.V07.HDF5

The prefixes for the different Runs are:

- 3B-HHR-E – half-hourly, Early Run
- 3B-HHR-L – half-hourly, Late Run
- 3B-HHR – half-hourly, Final Run
- 3B-MO – monthly, Final Run

The standard approach to version numbering in GPM is to give all "down-stream" products the same version as the inputs. Given that IMERG has numerous inputs, we choose to use the calibrator (files with names that include 2B.GPM.DPRGMI) to set the version for IMERG products. Thus, whenever 2BCMB has a version update, either major or minor (i.e., number or letter), then IMERG product identifiers will be updated as well. It also means that, if IMERG products are updated with a minor change (letter), then should 2BCMB be updated afterwards, the IMERG version will be updated once again and have a greater letter than 2BCMB. For example:

| | |
|-------------------|------------|
| 2BCMB V06D | IMERG V06D |
| IMERG updates ... | |
| 2BCMB V06D | IMERG V06E |
| 2BCMB updates ... | |
| 2BCMB V06E | IMERG V06F |

.....

The **temporal resolution** of the products is:

- 3IMERGHH: half hour (with “-E”, “-L”, and “” for the Early, Late, and Final Runs)
- 3IMERGM: month

The half-hour period for the 3IMERGHH is driven by the basic observational interval for the geo-IR data. Note that both the microwave and IR input data are snapshots. In those small regions in which two (or more) overlapping microwave scenes occur in a grid box, the conical-scan imager (CSI) closest to the mid-point of the half hour is taken, or lacking any CSIs, the cross-track-scan sounder (CTSS) closest to the mid-point. [When both a CSI and a CTSS are available, the CSI is chosen even if it is further from the mid-point than the CTSS. Recent work shows that this choice is reasonable over ocean, but less so over land.] This restriction to one snapshot makes the statistics of the data sets as comparable as possible, since the collection of two-snapshot averages of precipitation rates has a rather different PDF than single-shot averages. These snapshots might be thought of as an average rate, valid at the nominal observation time, since Villarini and Krajewski (2007) showed that the (snapshot-based) TRMM 3B42 is best correlated with radar data averaged over 60-90 minutes, not always centered on the nominal overpass time. More recently, O et al. (2017) found that IMERG-F half-hourly estimates correlated best with gauge accumulations of about 25 min that were centered 20 to 40 min. after the start of the IMERG data window. Part of the differences might be the 3-hour time window for 3B42, which is longer than IMERG’s half hour, but in either case, it is clearly not “instantaneous”.

The monthly period for 3IMERGM is driven by the typical calendar monthly period of precipitation gauge analyses, although it is also a typical period requested by many users. The precipitation value is an average rate over the calendar month.

Because the data are provided at nominal UTC half-hour intervals, each 3IMERGHH data set represents a nominal 30-minute span starting on the hour or half-hour. Thus, the first image of the day includes data for 00:00:00–00:29:59 UTC. The metadata and dataset name contain the date and start and end times.

.....

The **period of record** for TRMM-based IMERG is June 2000 <January 1998> through May 2014. Thereafter IMERG has GPM-based calibration. There is a delay (latency) of 4 hours, 14 hours, and about 3.5 months after the end of the month for the Early, Late, and Final Runs.

.....

The **grid** on which each field of values is presented is a 0.1°×0.1° lat./lon. (Cylindrical Equal Distance) global array of points. It is size 1800×3600, with X (latitude) incrementing most rapidly South to North from the southern edge, and then Y (longitude) incrementing West to East from the Dateline, as detailed in the metadata. Tenth-degree latitude and longitude values are at grid edges:

- First point center (89.95°S,179.95°W)
- Second point center (89.85°S,179.95°W)
- Last point center (89.95°N,179.95°E)

The reference datum is WGS84.

For the first time, in V07 IMERG all data fields are globally complete (to the extent that the input fields provide such coverage; IR only covers 60°N-S).

.....

The **spatial resolution** of IMERG is 0.1°×0.1° lat/lon.

.....

The **spatial coverage** of Version 07 IMERG precipitation estimates is the latitude band 90°N-S. In V07, an investigation into the performance of PMW retrievals revealed that, while having diminished performance, they still carry useful information and generally outperform IR retrievals (personal communication with P. Kirstetter, 2023). Therefore, the longstanding policy of screening PMW retrievals over frozen surfaces is lifted in V07. The result is that, except for a handful of grid boxes (mostly at the poles), IMERG has complete global coverage, with instantaneous and propagated PMW estimates covering over the polar regions. Nevertheless, users should be aware of the diminished performance of these estimates and exercise appropriate caution when interpreting the results. In particular, applications that use and evaluate these estimates over frozen surfaces, especially at high latitudes, should indicate this caveat of reduced skill in the interpretation of their results. Accordingly, the half-hourly Quality Index (QI), which is based on the correlation to the reference PMW estimate (GMI or TMI), is reduced by a factor depending on whether it is an imager or sounder and whether the grid box is over sea ice or frozen land (as indicated by the daily NOAA Autosnow product). Specifically, these factors are 0.112 for imagers and 0.275 for sounders over frozen land, and 0.495 for imagers and 0.282 for sounders over sea ice. These four factors are derived by comparing a sample of the Kalman correlations of instantaneous PMW estimates with those obtained from evaluation against KuPR by You et al. (2023).

.....

The **data file layout** for the original IMERG HDF5 files can be accessed at <ftp://gpmweb2.pps.eosdis.nasa.gov/pub/GPMfilespec/filespec.GPM.pdf>, which includes a significant quantity of metadata intended to make the files recognizable by many standard off-the-shelf applications (see <ftp://gpmweb2.pps.eosdis.nasa.gov/pub/GPMfilespec/filespecMeta.GPM.pdf>).

.....

It is possible to **read a file of data** with many standard off-the-shelf applications; any tool that reads the standard HDF5 file can be used to process IMERG files. As well, PPS provides a toolkit with C and FORTRAN APIs that allow users to write custom programs. See <ftp://gpmweb2.pps.eosdis.nasa.gov/pub/PPStoolkit/GPM> for more details. Documentation for the value-added file formats listed in the GPM data access pages at <https://gpm.nasa.gov/data/directory> should be consulted to determine the best way to read these files. Some of these formats are intended for specific functions, such as the GIS-oriented GeoTIFF files. For users of the R programming language, a package to read HDF5, called ‘hdf5r’, is located at <https://cran.r-project.org/web/packages/hdf5r/index.html>.

.....

The **3IMERGHH data fields** provide a variety of data fields for users and data developers.

Table 1. List of data fields, their variable names (in the data structure), and the data units for 3IMERGHH data files. Bold denotes user-oriented fields.

| Data field | Variable name | Units |
|---|---------------------------------------|---------------------|
| Multi-satellite precipitation estimate with gauge calibration (Final) Multi-satellite precipitation estimate (Early and Late) (recommended for general use) | <i>precipitation</i> * | mm/hr |
| Multi-satellite precipitation estimate | <i>precipitationUncal</i> | mm/hr |
| Random error for <i>precipitation</i> | <i>randomError</i> | mm/hr |
| Merged microwave-only precipitation estimate | <i>MWprecipitation</i> * | mm/hr |
| Microwave satellite source identifier | <i>MWprecipSource</i> | index values |
| Microwave satellite observation time | <i>MWobservationTime</i> | min. into half hour |
| IR-only precipitation estimate | <i>IRprecipitation</i> | mm/hr |
| Aggregate influence of the IR precipitation to the multi-satellite precipitation | <i>IRinfluence</i> | percent |
| Probability of liquid phase | <i>probabilityLiquidPrecipitation</i> | percent |
| Quality Index for <i>precipitation</i> field | <i>precipitationQualityIndex</i> | index |

* Note well that *MWprecipitation* only includes microwave data (hence “MW”), meaning it has significant gaps. *precipitation* is the complete estimate that most users will want to access.

The index values for *MWprecipSource* are:

- | | | |
|--------------------|-------------------|-------------------------------|
| 0 = no observation | 1 = TMI | 2 = (unused) |
| 3 = AMSR-2 | 4 = SSMI | 5 = SSMIS |
| 6 = AMSU | 7 = MHS | 8 = (unused) |
| 9 = GMI | 10 = (unused) | 11 = ATMS |
| 12 = AIRS | 13 = TOVS | 14 = CrIS |
| 15 = AMSR-E | 16 = SSMI F11 | 17 = Spare CSI 3 |
| 18 = Spare CSI 4 | 19 = Spare CSI 5 | 20 = SAPHIR (not used in V07) |
| 21 = NOAA-21 ATMS | 22 = Spare CTSS 3 | 23 = Spare CTSS 4 |
| 24 = Spare CTSS 5 | | |

where CSI = conical-scan imager and CTSS = cross-track-scan sounder. The time span for each 3IMERGHH field is the half-hour (in UTC) stated in the file name and metadata, covering the first or second half of a UTC hour.

The *3IMERGM data fields* provide a variety of data fields for users and data developers.

Table 2. List of data fields, their variable names (in the data structure), and the data units for 3IMERGM data files. Bold denotes user-oriented fields.

| Data field | Variable name | Units |
|--|----------------------|-------|
| Merged satellite-gauge precipitation estimate (recommended for general use) | <i>precipitation</i> | mm/hr |

| | | |
|--|--|--------------------------------|
| Random error for merged satellite-gauge precipitation | <i>randomError</i> | mm/hr |
| Weighting of gauge precipitation relative to the multi-satellite precipitation | <i>gaugeRelativeWeight</i> | percent |
| Accumulation-weighted probability of liquid phase | <i>probabilityLiquid Precipitation</i> | percent |
| Quality Index for precipitation field | <i>precipitationQuality Index</i> | Equivalent gauges per 2.5° box |

The time span for each 3IMERGM field is the calendar month stated in the file name and metadata.
.....

The **latency** of the various Runs is determined as follows:

1. The IMERG Early Run is the most constrained in terms of latency. The low-orbit microwave satellite data are downlinked once or twice an orbit (which is about 90 minutes long) to the relevant operating agency. They compute the Level 1B files, package it, and ship it around; the NOAA and EUMETSAT data are accessed through NOAA, DMSP from NRL Monterey, and the Japanese AMSR2 through NASA, and GPM DPR and GMI within GPM. The global merged geo-IR data are assembled by NOAA/NWS/CPC, and the Precipitation Processing System (PPS) accesses JMA Forecast from JAXA, GEOS FP forecast fields through the Goddard NCCS, and Autosnow from NOAA. By about 3 hours after observation time PPS typically has the geo-IR and about 85% of the microwave (depending on how computing systems feel that day) received at PPS. PPS converts the microwave data to precipitation estimates and computes the IMERG Early Run, which only uses forward propagation morphing (plus IR in the Kalman filter) by about 4 hours after observation time. [One subtle point is that the IR half-hour data fields come in pairs, so we process both halves of the hour together, meaning the second half hour almost always has shorter latency than the first. Another is that IMERG uses forecasts to get near-real-time analyses of several atmospheric fields for use in V07 to estimate system motion for the morphing, so hitches in delivery of the latest GEOS FP is not a big deal – PPS just uses the most recent forecast sequence, whatever that is. Tests indicate that even “stale” forecasts (a few days old) provide reasonable analyses for estimating the motion vectors.]
2. To compute the IMERG Late Run, which entails both forward and backward propagation morphing (plus IR), PPS has to wait long enough that the following microwave overpass has a chance to occur and then following the delivery chain described above. So, PPS waits about 11 hours to capture a reasonably complete set of “next overpass” data, and then start processing. The nominal latency is 14 hours, but in recent data it’s more like 12.
3. The IMERG Final Run uses MERRA-2 for the atmospheric fields used to estimate displacement vectors, GPCP monthly Monitoring Analysis for gauge, and “climatologically tuned” GPROF precipitation retrievals that depend on ERA5 (originally ERA-I). The pacing items are the GPCP and ERA5, usually giving a latency of about 3.5 months.

.....

5. Sensors

Some *PMW satellite overpass times** experienced significant drifting during the satellite's period of service (see <https://gpm.nasa.gov/resources/images/gpm-constellation-overpass-times> for a current time-series plot). There is no direct effect on the accuracy of the precipitation estimates, but it is possible that the systematic change in sampling time could introduce subtle shifts in the resulting collection of precipitation estimates. Furthermore, the climatological calibrations of the constellation satellites against GMI/TMI are based on a snapshot of fixed overpass times, so if the constellation overpass times drift the calibrations will "drift" as well. The plot shows that the 00/12 UTC slot entirely lacks Sun-synchronous coverage for three or more hours over the entire record. Meanwhile, the uncoordinated satellite operations end up giving a slowly varying pattern of near-duplicate data in some time slots, and totally absent coverage in others. Finally, it is routinely the case that the precessing satellites (TRMM, GPM) create a complicated pattern of overlap and fill-in for the observations and gaps in the Sun-synchronous satellites (including the 00/12 UTC void), although this diagram does not explicitly show this.

.....

The GPM Dual-frequency Precipitation Radar (*DPR**) is a pair of flat-panel phased-array weather radars in the Ku-band (13.6 GHz) and the Ka-band (35.5 GHz), based on the very successful TRMM PR in the Ku-band (13.8 GHz). The horizontal and vertical resolutions for the Ku-band unit are 5 km and 250 m, respectively, over a 245 km swath to a height above sea level of 19 km. The minimum detectable rainrate according to specification at launch is 0.5 mm/hr. The Ka-band unit had two modes up to 21 May 2018. Initially, in the standard mode, the horizontal and vertical resolutions were 5 km and 250 m, respectively, over a 125 km swath to a height above sea level of 19 km, with a minimum detectable rainrate according to specification at launch of 0.2 mm/hr. In high-resolution mode, the horizontal and vertical resolutions were 2.5 km and 250 m, respectively, over a 110 km swath to a height above sea level of 20 km. From 21 May 2018, the Ka swath was changed to match the Ku swath width of 245 km and resolution of 5 km, and dropping the high resolution mode (https://pps.gsfc.nasa.gov/Documents/Caveats_DPRL2_productV05B_fin3.pdf).

The DPR is an operational sensor, so the data record suffers the usual gaps in the record due to processing errors, down time on receivers, etc.

The 65° inclination provides nominal coverage over the latitudes 67°N-S for the Ku-band unit, and to 66°N-S for the Ka-band unit, initially, then changing on 21 May 2018 to 67°N-S.

Further details are available in Iguchi et al. (2010).

.....

The GPM Microwave Imager (*GMI**) is a multi-channel passive microwave radiometer. The GMI provides vertical and horizontal polarization values for 10.7, 18.7, 23.8, 36.5, 89.0, 165.5, 183.3 ± 3 , and 183.3 ± 7 GHz frequencies (except only vertical at 23 and the 183's) with conical scanning, similar to the SSMIS. The channels have effective fields of view that vary from 4×7 km for the 89 GHz (oval due to the slanted viewing angle intersecting the surface at 51°) to 19×32 km for the 10 GHz. Channels above 89 GHz are resolved at the 89 GHz footprint size. For practical

reasons, as with SSMIS, two separate feed horns collect the lower-frequency channels (36.5 and below) and the remaining higher-frequency channels. As a result, each set of channels has a separate navigation, and the footprints are not collocated. At Level 1C PPS provides both the original, separately navigated channels and a special product in which the high-frequency channels are remapped to the low-frequency footprint locations. The 89 GHz and higher frequency channels are undersampled near nadir, and the lower-frequency channels are more or less oversampled (see <https://www.star.nesdis.noaa.gov/mirs/gpmgmi.php>). At the swath edge even the higher-frequency channels are oversampled.

The GMI is an operational sensor, so the data record suffers the usual gaps in the record due to processing errors, down time on receivers, etc.

The 65° inclination provides nominal GMI coverage over the latitude band 70°N-S, although limitations in retrieval techniques limit the accuracy, or even availability, of precipitation estimates in cases of frozen surface.

Further details are available at <https://gpm.nasa.gov/sites/default/files/2020-06/GPMGMILIBATBDV4.0.pdf>.

.....

The TRMM Precipitation Radar (**PR**) was a flat-panel phased-array weather radar in the Ku-band (13.8 GHz), the first flown in space. The horizontal and vertical resolutions were 4 km and 250 m, respectively, over a 220 km swath to a height above sea level of 20 km. The minimum detectable signal was 17 (18) dBZ before (after) the TRMM orbit boost in August 2001.

The PR was an operational sensor, so the data record suffered the usual gaps in the record due to processing errors, down time on receivers, etc. There were outages for an operational anomaly in May 2000, and the boost to a higher orbit during the first part of August 2001. The PR suffered an electronics failure on 29 May 2009. Data were lost until the “B-side” electronics were activated on 19 June 2009. Some residual differences are noticeable. The descent of TRMM following fuel exhaustion in July 2014 prevented useful retrievals after 7 October 2014, except for a short period in 2015 from 12 February to 1 April as TRMM descended past its original at-launch altitude of 350 km.

The 35° inclination provided nominal PR coverage over the latitudes 37°N-S.

Further details are available in Kummerow et al. (1998).

.....

The TRMM Microwave Imager (**TMI**) was a multi-channel passive microwave radiometer that flew on TRMM. The TMI provided vertical and horizontal polarization values for 10.65, 19.35, 21.3, 37.0, and 85.5 GHz frequencies (except only vertical at 21) with conical scanning, similar to the SSMI. The channels had effective fields of view that varied from 4.6×6.9 km for the 85 GHz (oval due to the slanted viewing angle intersecting the surface at 51°) to 29.1×55.2 km for the 10 GHz. Consequently, the 85 GHz was undersampled near nadir, and all other channels are more or less oversampled. At the swath edge even the 85.5 GHz was oversampled.

The TMI was an operational sensor, so the data record suffered the usual gaps in the record due to processing errors, down time on receivers, etc. There were outages for an operational anomaly in May 2000 and the boost to a higher orbit during the first part of August 2001. The TMI continued to be operated during the descent of TRMM following fuel exhaustion in July 2014, until passivation (shutdown of the satellite) on 8 April 2015. During the descent there was some effect due to a gradually changing Earth Incidence Angle (EIA) of the sensor, but the retrievals are considered useful.

The 35° inclination provided nominal TMI coverage over the latitudes 40°N-S, although limitations in retrieval techniques curtail the accuracy, or even availability, of precipitation estimates in cases of frozen surface (which is unlikely in this latitude range).

Further details are available in Kummerow et al. (1998).

.....

The Advanced Microwave Scanning Radiometer Version 2 (**AMSR2**) is a multi-channel passive microwave radiometer provided by the Japan Aerospace Exploration Agency that has flown on GCOM-W since mid-2012. GCOM-W is placed in a Sun-synchronous polar orbit with a period of about 102 min. The AMSR2 provides vertical and horizontal polarization values for 6.9, 10.65, 18.7, 23.8, 36.5, and 89.0 GHz frequencies (except only vertical at 23.8) with conical scanning, similar to the SSMI. Pixels and scans are spaced 10 km apart at the suborbital point, except the 89-GHz channels are collected at 5 km spacing. Every other high-frequency pixel was co-located with the low-frequency pixels, starting with the first pixel in the scan and the first scan in a pair of scans. The channels have resolutions that vary from 3×5 km for the 89 GHz (oval due to the slanted viewing angle) to 35×62 km for the 6 GHz.

The polar orbit provides nominal AMSR2 coverage over the latitudes 85°N-S, although limitations in retrieval techniques limit the accuracy, or even availability, of precipitation estimates in cases of frozen surface.

Further details are available at

https://suzaku.eorc.jaxa.jp/GCOM_W/w_amsr2/whats_amsr2.html.

.....

The Advanced Microwave Scanning Radiometer for the Earth Observing System (**AMSR-E**) was a multi-channel passive microwave radiometer provided by the Japan Aerospace Exploration Agency that flew on Aqua from mid-2002 until the sensor failed in late 2011. The useful data period is 19 June 2002 – 3 October 2011. Aqua was placed in a Sun-synchronous polar orbit with a period of about 102 min. The AMSR-E provided vertical and horizontal polarization values for 6, 10, 18, 23, 36, and 89 GHz frequencies (except only vertical at 23) with conical scanning, similar to the SSMI. Pixels and scans were spaced 10 km apart at the suborbital point, except the 89-GHz channels were collected at 5 km spacing. However, the B-scan sensor, which provides the 89 GHz scan between the lower-frequency scans, failed around 4 November 2004. Every other high-frequency pixel was co-located with the low-frequency pixels, starting with the first pixel in the

scan and the first scan in a pair of scans. The channels had resolutions that vary from 4×6 km for the 89 GHz (oval due to the slanted viewing angle) to 43×74 km for the 6 GHz.

The polar orbit provided nominal AMSR-E coverage over the latitudes 85°N-S, although limitations in retrieval techniques limit the accuracy, or even availability, of precipitation estimates in cases of frozen surface.

The AMSR-E was an operational sensor, so the data record suffered the usual gaps in the record due to processing errors, down time on receivers, etc. Over time the coverage improved as the operational system matured. As noted above, the B-scan sensor failed around 4 November 2004.

Further details are available at <https://global.jaxa.jp/projects/sat/aqua/topics.html>.

.....

The Special Sensor Microwave Imager/Sounder (**SSMIS**) is a multi-channel passive microwave radiometer that has flown on selected Defense Meteorological Satellite Program (DMSP) platforms since late 2003. The DMSP is placed in a Sun-synchronous polar orbit with a period of about 102 min. 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 (SSM/T2) sensor: 150, 183±1, 183±3, and 183±7 GHz. Unlike SSMI, every SSMIS scan observes at all channels: pixels and scans are respectively spaced 25 and 12.5 km apart at the suborbital point for channels below 91 GHz, but 12.5 km for both pixel and scans for 91 GHz and above. Thus, the high-frequency channels have twice as many footprints per scan as the lower-frequency channels. For practical reasons, as with GMI, multiple separate feed horns are used to collect sets of channels. As a result, each set of channels has a separate navigation, and the footprints are not collocated. The SSMI-like channels have the resolutions

42.4×70.1 km (19, 22 GHz)

27.5×44.2 km (37 GHz)

13.1×14.4 km (91 GHz)

while the “sounding” channels have the resolutions

13.1×14.4 km (150 GHz)

13.1×14.4 km (183±1, 183±3, 183±7 GHz)

with the slanted viewing angle and in-line processing determining the oval shape.

Operational and design problems early in the program raised serious obstacles to use of the data. Accordingly, the useful periods of record (below) start relatively long after launch for F16 and F17. These dates are based on the start of the first publicly available SSMIS as determined by NRL/FNMOC through the Shared Processing Program with NESDIS.

The polar orbit provides nominal SSMIS coverage over the latitudes 85°N-S, although limitations in retrieval techniques limit the accuracy, or even availability, of precipitation estimates in cases of frozen surface.

The SSMIS is an operational sensor, so the data record suffers the usual gaps in the record due to processing errors, down time on receivers, etc. In particular, for the periods 5 April – 18 May 2016 and 9 March 2017 to the present the F17 SSMIS 37V channel experienced intermittent noisy values and the values have been set to missing.

Further details are available in Northrup Grumman (2002) and at <https://en.wikipedia.org/wiki/SSMIS> . Note that the acronym was originally “SSMI/S”, but “SSMIS” has since come into common use.

Table 3. The inventory of SSMIS data used in IMERG, period of record, and sensor status.

| <i>DMSP</i> | <i>Period of Record</i> | <i>Status</i> |
|-------------|------------------------------------|---------------|
| F16 | 20 November 2005 - ongoing | active |
| F17 | 19 March 2008 - ongoing | active |
| F18 | 8 March 2010 - ongoing | active |
| F19 | 1 December 2014 - 11 February 2016 | inactive |

Note that F20 was built, but never launched.

The Special Sensor Microwave/Imager (*SSMI*) was a multi-channel passive microwave radiometer that began flying on selected Defense Meteorological Satellite Program (DMSP) platforms in mid-1987. The DMSP was placed in a Sun-synchronous polar orbit with a period of about 102 min. The SSMI provided vertical and horizontal polarization values for 19, 22, 37, and 85 GHz frequencies (except only vertical at 22) with conical scanning. Pixels and scans were spaced 25 km apart at the suborbital point, except the 85-GHz channels were collected at 12.5 km spacing. Every other high-frequency pixel was co-located with the low-frequency pixels, starting with the first pixel in the scan and the first scan in a pair of scans. The channels had resolutions that vary from 12.5×15 km for the 85 GHz (oval due to the slanted viewing angle) to 60×75 km for the 19 GHz.

The polar orbit provided nominal SSMI coverage over the latitudes 85°N-S, although limitations in retrieval techniques limit the accuracy, or even availability, of precipitation estimates in cases of frozen surface.

The SSMI was an operational sensor, so the data record suffered the usual gaps in the record due to processing errors, down time on receivers, etc. Over time the coverage improved as the operational system matured. As well, the first 85 GHz sensor to fly (on F08) degraded quickly due to inadequate solar shielding. After launch in mid-1987, the 85.5 GHz vertical- and horizontal-polarization channels became unusable in 1989 and 1990, respectively. Another issue arose on 14 August 2006: DoD activated the RADCAL beacon on the F15 DMSP, which interfered with the 22V and 85.5V channels and prevented reliable estimates using then-current GPROF code. It remains to be determined whether additional calibration work might allow F15 to be used in a future reprocessing.

Further details are available in Hollinger et al. (1987, 1990). Note that the acronym was originally “SSM/T”, but “SSMI” has since come into common use.

Table 4. The inventory of SSMI data used in IMERG, period of record, and sensor status.

| <i>DMSP</i> | <i>Period of Record</i> | <i>Status</i> |
|-------------|-----------------------------------|-------------------------------|
| F11 | 1 January 1998 - 16 May 2000 | inactive |
| F13 | 1 January 1998 - 30 November 2009 | inactive |
| F14 | 1 January 1998 - 23 August 2008 | inactive |
| F15 | 23 February 2000 - 14 August 2006 | inactive, unusable thereafter |

The Sounder for Atmospheric Profiling of Humidity in the Intertropics by Radiometry (*SAPHIR*) is a multi-channel passive microwave radiometer that flew on the Megha-Tropiques platform from 12 October 2011 to 23 March 2022. The satellite was placed in an orbit with an inclination of 20° and a period of about 102 min. SAPHIR contained 6 channels, around 183 GHz, with cross-track scanning. Pixels and scans are spaced 10 km apart at nadir, with the pixels increasing in size and changing from circular to elongated in the cross-track direction as one moves away from nadir.

The 20° inclined orbit provides excellent coverage over the deep tropics.

SAPHIR was an operational sensor, so the data record suffered the usual gaps in the record due to processing errors, down time on receivers, etc. Over time the coverage has improved as the operational system has matured. However, difficulties with accessing the data on the satellite began on 15 December 2018 and successful data transmissions were re-established on 22 March 2019, when it was discovered that overheating of the electronics was the problem. As a result, only a few orbits could be transmitted per day. Additional problems with Earth Incidence Angle computations delayed return to service until 27 June 2019.

Further details are available at https://megha-tropiques.cnes.fr/en/MEGHAT/lien2_sat.htm.

Table 5. Inventory of SAPHIR data used in IMERG, period of record, and sensor status. The PRPS estimates do not provide estimates for the five footprints at each swath edge due to lower retrieval quality.

| <i>Satellite</i> | <i>Period of Record</i> | <i>Status</i> |
|------------------|--|---------------|
| Megha-Tropiques | 13 October 2011 – 15 December 2018, 27 June 2019 – 23 March 2022 intermittently | inactive |

The Advanced Temperature and Moisture Sounder (*ATMS*) is a multi-channel passive microwave radiometer that has flown on the Suomi National Polar-orbiting Partnership platform since 28 October 2011, on NOAA-20 since 29 November 2017, and on NOAA-21 since 1 February 2023. The satellite is placed in a Sun-synchronous polar orbit with a period of about 102 min. ATMS contains 22 channels, some being similar to AMSU-B. These channels cover the

frequencies 89.0, 157.0, 183.311±1 and 3, and 190.311 GHz, with cross-track scanning. Pixels and scans are spaced 16.3 km apart at nadir, with the pixels increasing in size and changing from circular to elongated in the cross-track direction as one moves away from nadir.

The polar orbit provides nominal coverage over the entire globe, although limitations in retrieval techniques limit the accuracy, or even availability, of precipitation estimates in cases of frozen surface.

The ATMS is an operational sensor, so the data record suffers the usual gaps in the record due to processing errors, down time on receivers, etc. Over time the coverage has improved as the operational system has matured.

Further details are available at <https://www.star.nesdis.noaa.gov/mirs/snppatms.php>.

Note that the GPROF 2021 estimates for ATMS do not provide estimates for the 8 footprints at each swath edge due to algorithm issues as revealed in early testing.

Table 6. Inventory of ATMS data used in IMERG, period of record, and sensor status. The Version 07 GPROF estimates for ATMS do not provide estimates for the 8 footprints at each swath edge due to lower retrieval quality.

| <i>Satellite</i> | <i>Period of Record</i> | <i>Status</i> |
|------------------|----------------------------|---------------|
| SNPP | 10 December 2011 – current | active |
| NOAA-20 | 29 November 2017 – current | active |
| NOAA-21 | 1 February 2023 – current | active |

The Microwave Imager (*MWI*) is a multi-channel passive microwave radiometer that will fly on the EUMETSAT Polar Satellite – Second Generation platform in the mid-2020’s. The satellite is assigned to a Sun-synchronous polar orbit with a period of about 102 min. MWI contains The MWI provides vertical and horizontal polarization values for 18.7, 23.8, 31.4, 50.3, 52.8, 53.24, 53.75, 89, 118.75±3.2, 118.75±2.1, 118.75±1.4, 165.5±0.725, 183.1±7, 183.1±6.1, 183.1±4.9, 183.1±3.4, and 183.1±2 GHz frequencies (except only vertical at/above 118.75 with conical scanning. Pixels and scans are spaced xx km apart at the suborbital point, except the xx-GHz channels are collected at xx km spacing. Every other high-frequency pixel is co-located with the low-frequency pixels, starting with the first pixel in the scan and the first scan in a pair of scans. The channels have resolutions that vary from 50×XX km for the 18.7 GHz (oval due to the slanted viewing angle) to 10×XX km for all channels at/above 89 GHz.

The polar orbit provides nominal coverage over all latitudes, although limitations in retrieval techniques limit the accuracy, or even availability, of precipitation estimates in cases of frozen surface.

The MWI is an operational sensor, so the data record will suffer the usual gaps in the record due to processing errors, down time on receivers, etc.

Further details are available in Mattioli et al. (2019).

Further details are available at <web URL> .

Table 7. Inventory of MWI data used in IMERG, period of record, and sensor status.

| <i>Satellite</i> | <i>Period of Record</i> | <i>Status</i> |
|------------------|------------------------------|------------------|
| EPS-SG B1 | [January 25] – [December 32] | not yet launched |
| EPS-SG B2 | [January 32] – [December 39] | not yet launched |
| EPS-SG B3 | [January 39] – [December 46] | not yet launched |

.....

The Microwave Humidity Sounder (*MHS*) is a multi-channel passive microwave radiometer that has flown on NOAA platforms since mid-2005 as a follow-on to AMSU-B and on the EUMETSAT MetOp platforms since late 2006. The satellites are placed in Sun-synchronous polar orbits with periods of about 102 min. The MHS contains 5 channels, similar to AMSU-B. These channels cover the frequencies 89.0, 157.0, 183.311±1 and 3, and 190.311 GHz, with cross-track scanning. Pixels and scans are spaced 16.3 km apart at nadir, with the pixels increasing in size and changing from circular to elongated in the cross-track direction as one moves away from nadir.

The polar orbit provides nominal MHS coverage over the entire globe, although limitations in retrieval techniques limit the accuracy, or even availability, of precipitation estimates in cases of frozen surface.

The MHS is an operational sensor, so the data record suffers the usual gaps in the record due to processing errors, down time on receivers, etc. Over time the coverage has improved as the operational system has matured.

Further details are available in the NOAA KLM User's Guide (August 2014 update) at https://www.star.nesdis.noaa.gov/mirs/documents/0.0_NOAA_KLM_Users_Guide.pdf, specifically Section 3.9.

Note that the GPROF estimates in Version 07 for MHS do not provide estimates for the 5 footprints at each swath edge due to algorithm issues as revealed in early testing.

Table 8. Inventory of MHS data used in IMERG, period of record, and sensor status. The Version 07 GPROF estimates for MHS do not provide estimates for the 5 footprints at each swath edge due to lower retrieval quality.

| <i>Satellite</i> | <i>Period of Record</i> | <i>Status</i> |
|------------------|------------------------------------|---------------|
| NOAA-18 | 25 May 2005 – 20 October 2018 | inactive |
| NOAA-19 | 25 February 2009 – current | active |
| MetOp-A | 5 December 2006 – 30 November 2021 | inactive |
| MetOp-B | 1 April 2013 – current | active |
| MetOp-C | 10 November 2019 – current | active |

Note: METOP-A experienced a satellite anomaly 17:00 30 July 30 – 12:00 9 August 2019. However, PPS received files for a subset of this time span, 09:57 5 August – 12:00 9 August 2019. Checks done by PPS and the Merged Algorithm Team show that most of the data in this period appear to be OK, except for orbits 066384 and 066385, which were replaced with Empty granules.

Starting with NOAA-20 MHS has been superseded by ATMS on the NOAA series.

.....

The Advanced Microwave Sounding Unit B (**AMSU-B**) was a multi-channel passive microwave radiometer that flew on selected National Oceanic and Atmospheric Administration (NOAA) platforms from early 2000 to 2011. The NOAA satellites are placed in Sun-synchronous polar orbits with periods of about 102 min. The complete AMSU contained 20 channels, the first 15 referred to as AMSU-A, and the last 5 as AMSU-B. These channels (identified as 16 through 20) covered the frequencies 89.0 ± 0.9 , 150.0 ± 0.9 , and 183.31 ± 1 , 3, and 7 GHz, with cross-track scanning. Pixels and scans were spaced 16.3 km apart at nadir, with the pixels increasing in size and changing from circular to elongated in the cross-track direction as one moves away from nadir.

The polar orbit provided nominal AMSU-B coverage over the entire globe, although limitations in current retrieval techniques prevent useful precipitation estimates in cases of frozen surface.

The AMSU-B was an operational sensor, so the data record suffered the usual gaps in the record due to processing errors, down time on receivers, etc. NOAA-15, the first AMSU-B, had RFI problems early in the mission, so even though CLASS posts data starting 26 October 1998 the considered advice from NOAA is to start use with 1 January 2000. Over time the coverage improved as the operational system matured. As well, the NOAA-17 50-GHz channel failed in late October 2003, apparently due to solar flare activity. Finally, NOAA-16 gradually failed during 2010, and eventually it was determined that IMERG should stop using the data at the end of April 2010.

Further details are available in the NOAA KLM User's Guide (August 2014 revision) at https://www.star.nesdis.noaa.gov/mirs/documents/0.0_NOAA_KLM_Users_Guide.pdf, specifically in Section 3.4.

Note that the GPROF estimates in Version 07 for AMSU-B do not provide estimates for the 5 footprints at each swath edge. This is due to algorithm issues as revealed in early testing.

Table 9. The inventory of AMSU-B data used in IMERG, period of record, and sensor status. The Version 07 GPROF estimates for AMSU do not provide estimates for the 5 footprints at each swath edge due to lower retrieval quality.

| <i>Satellite</i> | <i>Period of Record</i> | <i>Status</i> |
|------------------|--|---------------|
| NOAA-15* | 1 January 2000 - 14 September 2010 | inactive |
| NOAA-16 | 4 October 2000 - 16 February 2011 (last used for 30 April 2010) | inactive |

| | | |
|---------|---------------------------------|----------|
| NOAA-17 | 28 June 2002 - 17 December 2009 | inactive |
|---------|---------------------------------|----------|

* NOAA-15 starts 26 October 1998, but RFI issues prevent use before 1 January 2000.

.....

The infrared (**IR**) data are collected from a variety of sensors flying on the international constellation of geosynchronous-orbit meteorological satellites – the Geosynchronous Operational Environmental Satellites (GOES, United States); the Geosynchronous Meteorological Satellite (GMS, Japan), subsequently Multi-functional Transport Satellite, (MTSat, Japan), followed by Himawari (Japan); and the Meteorological Satellite (Meteosat, European Community). There are usually two GOES platforms active covering the eastern and western regions of the Americas, two Meteosats covering the Europe/Africa and Indian Ocean sectors, and the Japanese series over east Asia. The geosynchronous IR data are collected by scanning (parts of) the earth's disk. By international agreement, all satellite operators collect full-disk images at the synoptic observing times (00, 03, ..., 21 UTC) at a minimum. Most of the time the operators provide substantially complete coverage for their respective regions every half hour, but the Japanese satellite tends to provide Southern Hemisphere data every hour, and GOES-W tends to cut off the higher latitudes over the South Pacific Ocean.

Subsequent processing is described in "Merged 4-Km IR Tb data set".

The various IR instruments are operational sensors, so the data record suffers the usual gaps in the record due to processing errors, down time on receivers, sensor failures, etc., as well as choices by the operators on observing strategies. Most notably, there was no geo-IR coverage in the Indian Ocean sector until 06 UTC 16 June 1998, although high-zenith-angle observations from adjacent geo-satellites are used to largely cover the gap. As well, GMS-5 was replaced by GOES-9 starting 01 UTC 22 May 2003, which introduced slightly different instrument characteristics. MTSat-1R went operational starting 19 UTC 17 November 2005, and was replaced by MTSat-2 on 1 July 2010. Himawari 8 took over from MTSat-2 effective 0200 UTC on 7 July 2015. Meteosat 8 replaced Meteosat 7 as of 1 February 2017.

Further details are available in Janowiak and Arkin (1991).

.....

The **precipitation gauge analysis** that is used in IMERG is produced by the Global Precipitation Climatology Centre (GPCC) under the direction of Andreas Becker and Udo Schneider, located in the Deutscher Wetterdienst, Offenbach a.M., Germany (Becker et al. 2013, Schneider et al. 2014, 2018). [Note: Throughout, we are clearly dealing with all forms of precipitation, but we follow the customary practice here of referring to precipitation gauges as “rain gauges”.] Rain gauge reports are archived from a time-varying collection of over 86,000 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. The lack of complete consistency in period of record for individual stations

has been shown to be less important than the gain in detail, particularly in complex terrain. Then for each month, the individual gauge reports are converted to deviations from the monthly climatology, are analyzed into gridded values using a variant of the SPHEREMAP spatial interpolation routine (Willmott et al. 1985), and then added to the gridded monthly climatology. In the latest version, grid boxes that are “too distant” from any stations are given anomaly values of zero. Finally, the month’s analysis is produced by superimposing the anomaly analysis on the month’s climatology.

The GPCC creates multiple products, and two are used in IMERG. The Full Data Reanalysis (currently Version 2022) is a retrospective analysis that covers the period 1891-2020, and it is used in IMERG for the span <1998> 2000-2020. Thereafter we use the GPCC Monitoring Product (currently Version 2022), 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 GPCC analyses are updated, we switch to the new version and continue processing forward, even though (in principle) there is a small inhomogeneity as a result. We continue our long-standing practice of correcting all gauge analysis values for climatological estimates of systematic error due to wind effects, side-wetting, evaporation, etc., following Legates (1987). The Legates-Wilmott scheme for undercatch in the GPCC gauges is too large at high latitudes, so a climatological Fuchs adjustment is implemented in V07 over Eurasia north of 45°N, while Legates-Wilmott continues to be used for all other land areas.

In use, we make several adjustments to the GPCC analyses for IMERG:

1. The Monitoring product is provided at 1° resolution, so we make a bilinear interpolation to the IMERG 0.1° grid. This does not preserve precipitation volume in individual grid boxes, but prior schemes that did resulted in unphysical blockiness at the scale of the GPCC grid. The volume is accounted for, but only by considering the surrounding GPCC grid box area.
2. The GPCC analysis is trimmed to the coastline at the IMERG grid scale. This results in a somewhat unphysical jump in values across coastlines, but seems preferable to extending the GPCC values over the water, and giving sometimes-unphysical jumps on the edges of the coarser GPCC grid. In addition, small islands are masked off because they are generally unrepresentative of the ocean that occupies the same (and nearby) grid boxes. The mask was created in a side study of the previous GPCC analyses that also ignored small islands.

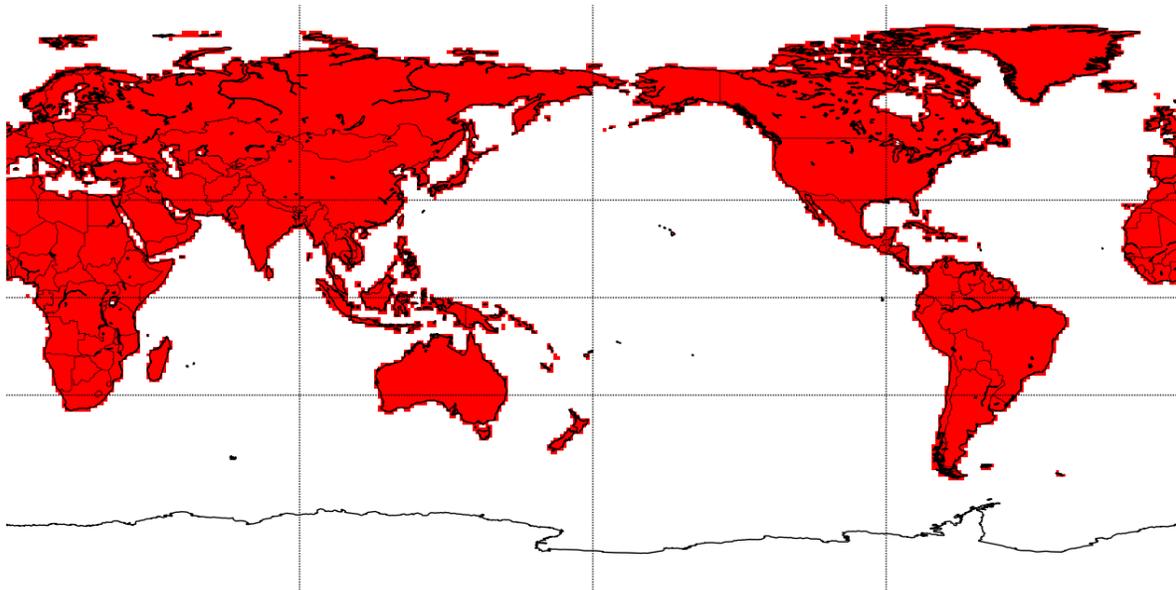


Fig. 1 GPCCC 1° island mask. Red denotes locations for which the GPCCC gauge analysis is used.

3. The GPCCC values are adjusted for precipitation gauge undercatch. See “precipitation gauge adjustment” for a summary.

.....

The inventory of **sensors contributing to IMERG** is summarized here for convenience; refer to the individual sensor descriptions for additional details.

Table 10. Alphabetized lists of contributing data sets for IMERG, broken out by sensor type. Data sets with start dates of Jan 98 extend before that time, but these data are not relevant to IMERG. Square brackets ([]) indicate an estimated date. The geosynchronous IR data are processed into “even-odd” files at NESDIS. All data are computed to Level 2 (scan/pixel) precipitation from Level 1 data by PPS, except for the precipitation gauge analyses (Level 3) and IR data (Level 1).

| <i>Merged Radar – Passive Microwave Conical-Scan Imager Products</i> | |
|--|-------------------------|
| <i>Product</i> | <i>Period of Record</i> |
| GPM DPR-GMI | Apr 14 - [Dec 29] |
| TRMM PR-TMI | Jan 98 - Sep 14 |

| <i>Conical-Scan Passive Microwave Imagers and Imager/Sounders</i> | | |
|---|-------------------------|---------------|
| <i>Sensor</i> | <i>Period of Record</i> | <i>Agency</i> |
| Aqua AMSR-E | Jun 02 - Oct 11 | JAXA |
| DMSP F13 SSMI | Jan 98 - Nov 09 | U.S. DoD |
| DMSP F14 SSMI | Jan 98 - Aug 08 | U.S. DoD |
| DMSP F15 SSMI | Feb 00 - Aug 06 | U.S. DoD |
| DMSP F16 SSMIS | Nov 05 - [Dec 24] | U.S. DoD |

| | | |
|-----------------|---------------------|----------|
| DMSP F17 SSMIS | Mar 08 - [Dec 25] | U.S. DoD |
| DMSP F18 SSMIS | Mar 10 - [Mar 26] | U.S. DoD |
| DMSP F19 SSMIS | Dec 14 - Feb 16 | U.S. DoD |
| GCOMW1 AMSR2 | Jun 12 - [May 26] | JAXA |
| GOSAT-3 AMSR3 | [Feb 24] - [Jan 32] | JAXA |
| GPM GMI | Mar 14 - [Jan 29] | NASA |
| METOP-SG B1 MWI | [Jan 24] - [Dec 34] | EUMETSAT |
| METOP-SG B2 MWI | [Jan 29] - [Dec 39] | EUMETSAT |
| METOP-SG B3 MWI | [Jan 36] - [Dec 46] | EUMETSAT |
| TRMM TMI | Jan 98 - Apr 15 | NASA |
| WSF-M 1 MIS | [Jan 24 - Jan 30] | U.S. DoD |
| WSF-M 2 MIS | [Jan 28 - Jan 34] | U.S. DoD |

| <i>Cross-Track-Scan Passive Microwave Sounders</i> | | |
|--|-------------------------|---------------|
| <i>Sensor</i> | <i>Period of Record</i> | <i>Agency</i> |
| JPSS-3 ATMS ** | [Jan 26] - [Jan 32] | NOAA |
| JPSS-4 ATMS ** | [Jan 31] - [Jan 37] | NOAA |
| METOP-2/A MHS ** | Dec 06 - [Jan 20] | EUMETSAT |
| METOP-1/B MHS ** | Apr 13 - [Aug 26] | EUMETSAT |
| METOP-3/C MHS ** | Nov 19 - [Apr 27] | EUMETSAT |
| M-T SAPHIR * | Oct 11 - Mar 22 | CNES/ISRO |
| NOAA-15 AMSU **+ | Jan 00 - Sep 10 | NOAA |
| NOAA-16 AMSU ** | Oct 00 - Apr 10 | NOAA |
| NOAA-17 AMSU ** | Jun 02 - Dec 09 | NOAA |
| NOAA-18 MHS ** | May 05 - Oct 18 | NOAA |
| NOAA-19 MHS ** | Feb 09 - Sep 20 | NOAA |
| NOAA-20 ATMS** | Nov 17 - [Jan 24] | NOAA |
| NOAA-21 ATMS ** | [Feb 23] - [Jan 30] | NOAA |
| SNPP ATMS ** | Dec 11 - [Dec 24] | NOAA |

** The Version 07 GPROF estimates for AMSU, ATMS, and MHS do not provide estimates for the 5, 8, and 5 footprints (respectively) at each swath edge.

+ NOAA-15 starts with 26 October 1998, but RFI issues prevent use before 1 January 2000.

| <i>Geosynchronous Infrared Imagers</i> | | |
|--|---------------------------------------|---------------|
| <i>Satellite</i> | <i>Sub-sat. Lon.</i> | <i>Agency</i> |
| EWS-G1 (old GOES-13) | 161°E | U.S. DoD/NOAA |
| GMS, MTSat, Himawari series | 140°E | JMA |
| GOES-E series | 75°W | NESDIS |
| GOES-W series | 135°W | NESDIS |
| Meteosat prime series | 0°E | EUMETSAT |
| Meteosat repositioned series | 63°E from Jul 98, 41°E from Oct 16 | EUMETSAT |

Precipitation Gauge Analyses

| <i>Analysis</i> | <i>Period of Record</i> | <i>Institution</i> |
|-------------------------|-------------------------|--------------------|
| Full Version 2022 | Jan 98 – Dec 2020 | DWD/GPCC |
| Monitoring Version 2022 | Dec 2020 - ongoing | DWD/GPCC |

Strictly speaking, the products provided by the Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2), Goddard Earth Observing System model Version 5 (GEOS-5) Forward Processing (FP), European Centre for Medium Range Weather Forecasting (ECMWF) Reanalysis Version 5 (ERA5), Japan Meteorological Agency forecast product (JMA Forecast), and the Autosnow product are not sensors, so they are described in “numerical analysis products”.

.....

Several **numerical analysis products** contribute to IMERG. MERRA-2 is a global atmospheric reanalysis beginning in 1980 using the GEOS-5 data assimilation system and numerical model (Bosilovich et al. 2016; Gelaro et al 2017). GEOS FP is a global forecast analysis using the GEOS-5 system at every 6 hr (Lucchesi 2017). Both datasets are processed and released by the NASA Global Modeling and Assimilation Office. ERA-5 is a global atmospheric reanalysis beginning in 1950 (Hersbach et al. 2020) from the ECMWF. JMA Forecast is the then-operational numerical weather prediction output, delivered to PPS. See “numerical analysis products” for more information. Autosnow is a daily NOAA product that provides surface coverage of snow and ice from calibrated satellite data (Romanov 2016), and is delivered to PPS. See “Autosnow” for more details. In Version 07, the surface precipitation rate (PRECTOT), the total precipitable liquid water (TQL), and the total precipitable water vapor (TQV; also called total column water vapor) are used in the sequence listed to compute the displacement vectors for morphing the PMW data. This change is motivated by the fact that the TQV-based vectors in V06 are less accurate over orography. The motion vectors are computed from these three variables in a hierarchical fashion. First, motion vectors are computed from PRECTOT pixels with at least 0.1 mm / h. In regions where PRECTOT fails to produce a motion vector (usually because there is no model precipitation in the region) or if the vector is identified as “errant”, the computation falls to the next variable. Errant vectors are identified by comparing them pairwise to their four neighbors; a vector is considered “errant” if any of the pairwise differences is greater $0.6^\circ / \cos(\text{latitude})$ in the longitude direction or 0.4° in the latitude direction. After PRECTOT, the next variable is TQL with a threshold of 0.01 kg / m², with the errant vector screening applied as well. Finally, the motion vector computation falls back to TQV (no threshold), which based on our experience in V06 produces a smooth and complete vector field. PPS extracts all four data fields from daily files of hourly data at $0.5^\circ \times 0.625^\circ$ latitude/longitude for MERRA-2, and from hourly files at $0.25^\circ \times 0.3125^\circ$ latitude/longitude for GEOS FP, feeding into post-real-time and retrospective processing, and near-real-time processing, respectively.

The Probability of Phase (PLP) is estimated as a function of surface wet-bulb temperature, with separate curves for land and ocean. The surface temperature, humidity, and pressure information needed to compute the surface wet-bulb temperature are taken from the JMA NWP forecast for the Early and Late Runs, and the ECMWF Reanalysis Version 5 (ERA5) for the Final Run.

ERA5 and JMA Forecast are episodically upgraded as model development progresses. The IMERG team considers the impact of each such upgrade, but in general the computations of the displacement vectors and PLP are robust to minor changes in input data.

.....

The NOAA **Autosnow** product is used to provide global data on daily surface coverage by snow and ice. Currently, information on the snow cover is derived from the Advanced Very High Resolution Radiometer (AVHRR) onboard METOP satellites, imagers onboard Geostationary Operational Environmental Satellites (GOES) East and West, Spinning Enhanced Visible and Infrared Imager (SEVIRI) onboard Meteosat second Generation (MSG), and Special Sensor Microwave Imager/Sounder (SSMIS) onboard Defense Meteorological Satellite Program (DMSP) satellites. Ice cover is derived from the METOP AVHRR and DMSP SSMIS data until a new version (instituted in June 2023 and adopted by PPS in July 2023) stopped using SSMIS. Both snow and ice are identified in satellite images using threshold-based decision tree image classification algorithms. Information on snow and ice cover derived from observations in the visible/infrared and the microwave bands is combined to generate continuous (gap-free) maps on a daily basis. The main output product of the system is a daily global snow and ice cover map generated on a 0.04° lat/lon grid (Plate Carree), which is about 4×4 km at the Equator for the period June 2000 – June 2023. In July 2023, the Autosnow output changed to a higher resolution (0.02° lat/lon grid). See Romanov (2016) for more details.

As of 14 June 2019, it was discovered that Autosnow sometimes provides incorrect values other than “sea ice” above 89°N. This error has been corrected in V07.

.....

6. Definitions and Defining Algorithms

The **precipitation variable** is computed as described under the individual product headings. All precipitation products have been converted from their original units to mm/hr. Throughout this document, “precipitation” refers to all forms of precipitation, including rain, drizzle, snow, graupel, and hail (see “precipitation phase”).

.....

The **time zone** for these data sets is Universal Coordinated Time (UTC, also as GMT or Z).

.....

The Goddard Profiling Algorithm (**GPROF**) is based on Kummerow et al. (1996), Olson et al. (1999), and Kummerow et al. (2001), Passive Microwave Algorithm Team Facility (2021). GPROF is a multi-channel physical approach for retrieving rainfall and vertical structure information from satellite-based passive microwave observations, both conical-scan imagers and imagers/sounders (AMSR-E, AMSR2, GMI, SSMI, SSMIS, TMI) and cross-track-scan sounders (AMSU, ATMS, MHS). 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. The GPROF2021 library depends on compilations of GPM Combined Instrument data. 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 strong heterogeneity of the surface emission signal in the lower frequencies when the underlying surface is other than entirely water.

There is a known problem with low bias in GPROF at high latitudes due to the low sensitivity of the DPR to light rain and snow. Accordingly, in Version 07 both the Combined and the GPROF library are modified to give more weight to GMI in these passes. These changes did not fully correct the low bias; see “GPCP-calibrated CORRA”. V07 GPROF also features an orographic adjustment to precipitation rates based on vertical air motion forced by the orography and on horizontal moisture convergence, stratified into convective and stratiform.

Note that the GPROF estimates in Version 07 for AMSU-B, ATMS, and MHS do not provide estimates for the 5, 8, and 5 footprints (respectively) at each swath edge. This is due to algorithm issues as revealed in early testing. Taken together, these limitations somewhat reduce the amount of microwave-based data contained in Version 07 IMERG.

.....
The Version 07 PPS datasets for **GMI, TMI, and constellation sensor Level 2 precipitation datasets** contain Level 2 (scan-pixel) GPROF2021 estimates of precipitation based on the sensor named in each dataset. These are provided by PPS. Each file contains an orbit of estimates, except the real-time 2AGPROFGMI is in nominally 5-minute granules as down-linked from the Core Observatory. The data have had some quality control, and are converted from sensor units to Ta, then to Tb, then to precipitation at PPS.

These data are used as input to IMERG processing.

Note that the Version 07 GPROF estimates for AMSU, ATMS, and MHS do not provide estimates for the 5, 8, and 5 footprints (respectively) at each swath edge. This is due to algorithm issues as revealed in V06 testing. Taken together, these limitations somewhat reduce the amount of microwave-based data contained in Version 07 IMERG.

.....
The **GPM-era Combined Radar-Radiometer Algorithm** (CORRA-G; also known as the “GPM Combined Instrument”) provides, in principle, the most accurate, high resolution estimates of surface precipitation rate and vertical precipitation distributions that can be achieved with the GPM Core Observatory’s GMI and DPR instruments. CORRA is based upon a hybrid ensemble Kalman filtering / variational approach for inverting the DPR reflectivities and GMI brightness temperatures to estimates of precipitation profiles (Olson et al. 2016). This architecture is largely consistent with the successful TRMM Combined Instrument (TCI; product 2B31) algorithm design, but it has been updated and modularized to take advantage of improvements in the representation of physics, new climatological background information, and model-based analyses available during the GPM mission. CORRA estimates are provided on both the Ka swath and the Ku swath. It is important to note that IMERG uses the Ku swath product which has no influence from the Ka radar.

There is a known problem with low bias in CORRA-G at high latitudes due to the low sensitivity of the DPR to light rain and snow. Accordingly, starting in Version 05 both CORRA-G and the GPROF library were modified to give more weight to GMI in these passes.

In V07 there were significant changes in the mapping procedure of CORRA-G to better fit the 19 GHz estimation frequency FOV of the GPROF GMI precipitation retrieval to yield closer precipitation distribution frequencies between CORRA-G and GMI. In V06 the nearest CORRA-G retrieval was mapped to the 0.1° IMERG grid. For spatial consistency with typical GPROF Level 2 footprints, in V07 a spatial 8×8 average of the native-resolution CORRA-G estimates was performed before mapping to the 0.1° IMERG grid.

.....

The **TRMM-era Combined Radar-Radiometer Algorithm** (CORRA-T) provides, in principle, the most accurate, high resolution estimates of surface precipitation rate and vertical precipitation distributions that can be achieved with TRMM’s TMI and PR instruments. The CORRA-T algorithm is a modification of the CORRA-G based upon a hybrid ensemble Kalman filtering / variational approach for inverting the PR reflectivities and TMI brightness temperatures to estimates of precipitation profiles (Olson et al. 2016), and tuned to (approximately) match CORRA-G. This architecture addresses the same considerations as the successful TRMM Combined Instrument (TCI; product 2B31) algorithm design, but there are considerable differences that take advantage of improvements in modularity, the representation of physics, new climatological background information, and model-based analyses that available during the GPM mission.

The known problem with low bias in CORRA-G at high latitudes does not affect CORRA-T due to its limited latitudinal coverage (35°N-S), but overall it shows a low bias compared to CORRA-G.

In a similar manner in V07 there were significant changes in the mapping procedure of CORRA-T to better fit the 19 GHz estimation frequency FOV of the GPROF TMI precipitation retrieval to yield closer precipitation distribution frequencies between CORRA-T and TMI. In V06 the nearest CORRA retrieval was mapped to the 0.1° IMERG grid. For spatial consistency with typical GPROF Level 2 footprints, in V07 a spatial 12×12 average of the native resolution CORRA-T estimates was performed before mapping to the 0.1° IMERG grid.

.....

PPS dataset **2BCMB** contains Level 2 (scan-pixel) output from the Version 07 CORRA, computed at PPS for both the GPM and TRMM eras. Each file contains an orbit of GPM Combined Instrument rain rate and path-integrated attenuation at 5 km horizontal and 250 m vertical resolutions over a 250 km swath. More information is available at

https://gpm.nasa.gov/sites/default/files/document_files/Combined_algorithm_ATBD.V04.rev_.pdf

and Olson et al. (2016).

These Level 2 data are used as input to IMERG processing.

The **GPCP-calibrated CORRA** was introduced in Version 04 because, even though CORRA is considered to be the best estimator, in Version 04 all of the GPM individual-sensor precipitation products were biased low in the high latitude oceans compared to the GPCP Satellite-Gauge product and estimates based on CloudSat (Behrangi et al. 2014 and subsequent revisions to the approach). There is some variation between the latter two, but the GPCP can provide regionally varying seasonal climatological adjustment factors, while the CloudSat-based data currently only have an annual latitudinal profile over ocean. Thus, we made a simple ratio adjustment to bring the CORRA-calibrated constellation estimates in line with what is considered a more reasonable estimate, using ratios computed from seasonal averages for the years 2017-19. We apply these corrections outside of a somewhat subjectively chosen, seasonally varying low-latitude zone where the ratio of CORRA to GPCP is one:

- DJF 30°N-20.1°S
- MAM 25°N-30°S
- JJA 30.4°N-30.4°S
- SON 29.9°N-25.7°S

Note that in V06 the GPCP adjustment was applied to both ocean and land. In V07, the land adjustment, as currently implemented, had the potential to provide unrealistic estimates so it was removed.

.....

Calibration outside the CORRA area of coverage in the GPM era is provided by extrapolating the northernmost and southernmost GMI-CORRA correction curves at 60° N and S to the poles using a 50-pass iterative smooth-fill technique to obtain corrections in the latitude bands 60°-90° N and S. [See the description of the **smooth-fill** for details.] To add stability to the extrapolated correction curves, the zonal mean of the corrections is computed at 60° N and S and assigned to the 90° N and S zonal bands, respectively.

In the TRMM era, background monthly climatological GPM-era GMI-CORRA histograms and correction curves are used to fill in the corrections for the latitude bands 33°-90° N and S. This fill-in is performed as follows:

1. Compute the count-normalized precipitation volumes for CORRA-G and CORRA-T in the latitude bands 0°-33° N and S using the TMI-CORRA month and GMI-CORRA background histograms. To build more stability, in V07 the monthly climatological GMI-CORRA background histograms were increased to a 3-month seasonal histogram collection centered on each calendar month.
2. Compute volume adjustment factors for northern and southern hemisphere ocean and land.
3. Apply these volume adjustment factors to the GPM background correction curves. Insert the volume-adjusted GPM correction curves into the TMI-CORRA correction curves in the region 25°-90° N and S, performing a weighted average of the two corrections in the latitude bands 25°-33° N and S.
4. Unfortunately, in V06 the following month of both the GPM background histograms and GPM correction curves were accidentally used to volume adjust and insert into the TMI-CORR-T correction curves. For example, June GPM background histograms were used to volume adjust

May CORRA-T and June GPM correction curves were inserted into the May TMI-CORRA-T correction curves outside the TRMM coverage. This monthly mis-match issue has been corrected in V07.

The resulting corrections are based on TRMM for the latitude band 25°N-25°S, volume-adjusted GPM for the latitude bands 33°-90° N and S, and a blend of the two in the latitude bands 25°-33° N and S. The goal is to provide the GPM correction structure outside the TRMM coverage area, but with the volume of the TRMM estimates. This was necessary, as directly extrapolating the TRMM region corrections to the poles created artificially low corrections.

.....

The **smooth-fill** algorithm provides smoothly varying values for one or more grid boxes with “missing” values that match the non-missing boxes around the edge of the “missing” hole. At each step of the iterative process, each originally missing boxes is filled with the 3×3-gridbox average of the non-missing box values in the previous iteration. Preliminary analysis showed that the innovations in the missing grid boxes substantially settle down when the number of iterations is 6 times the number of iterations it takes to give each missing grid boxes a value. For the CED grid, all grid boxes at the pole are given the same value.

.....

The **random error field**, data field *randomError*, is computed for both the half-hourly (3IMERGHH) and monthly (3IMERGM) datasets. The units are mm/hr.

The monthly random error is computed as in Huffman (1997), with appropriate adjustments for the difference in gridbox size. The half-hour random error is similarly computed, even though the approach is not entirely consistent for such fine scales. [The Huffman (1997) approach assumes that the various input data are statistically independent.] As a partial compensation, the half-hour random errors are approximately scaled to aggregate to the monthly random error, assuming that the half-hour values are statistically independent. This is also not strictly true, but the overall result appeared somewhat useful.

Work is currently underway by several research groups to develop more-appropriate estimators for random error, and to introduce estimates of bias error. We expect that more sophisticated error fields will be incorporated as part of IMERG in the future, for example providing additional information on the error quantiles following Maggioni et al. (2014), Li et al. (2023), or the correlation parameter computed in the Kalman filter methodology. In such a case, the critical problem is to limit the number of time/space-varying parameters that consequently require the insertion of additional parameter fields in each dataset. In such a case, the critical problem is to limit the number of time/space-varying parameters that consequently require the insertion of additional parameter fields in each dataset.

.....

The **merged PMW precipitation** estimate, data field *MWprecipitation* in 3IMERGHH data files provides a global Level 3 (0.1°×0.1°-gridded) half-hourly combination of all currently available

GMI, TMI, SSMI, SSMIS, AMSR-E, AMSR2, AMSU-B, MHS, and ATMS precipitation estimates as a field in the 3IMERGHH files:

1. The radiometer estimates and 2BCMB are gridded to a $0.1^\circ \times 0.1^\circ$ grid for a half-hour period starting at full and half hours, UTC.
2. After the 2BCMB estimates are gridded (to a $0.1^\circ \times 0.1^\circ$ grid for a half-hour period starting at full and half hours, UTC) the GPCP adjustment is then applied at the grid box level.
3. The gridded 2AGPROFGMI is calibrated to the gridded 2BCMB using a matched histogram correction computed afresh for a 45-day period once a pentad. In the Early and Late Runs this period is necessarily trailing, but for the Final we take advantage of the delay in waiting for the gauge analysis and use a centered 45-day period updated once a pentad. Because the 2AGPROFGMI and 2BCMB estimates vary significantly by region and time of year, the correction is computed and applied on a $1^\circ \times 1^\circ$ grid using a $3^\circ \times 3^\circ$ template. Experience in Version 03 showed that regions with differing gradients in GMI and GCI resulted in blocky patterns when the calibrations were used on the original 1° grid. So, starting in Version 04, the calibrations have been distance-weighted interpolations of the four surrounding 1° calibration values.
4. In V07 significant changes were made in both the CORRA-T to TMI and CORRA-G to GMI calibrations. In Version 06 IMERG, 2BCMB(T) – GMI(TMI) calibrations only used 30% of the GMI/TMI retrievals that coincided with the narrower 2BCMB swath. When this calibration was applied to GMI/TMI retrievals from the full radiometer swath, the histograms of calibrated precipitation from GMI/TMI often did not match up well with the histograms of 2BCMB over both land and ocean, leading to an overestimation in IMERG V06B precipitation rates. The same issue continued to be present in early versions of Version 07. The IMERG team worked with PPS and the GPROF team to determine that the original GMI/TMI precipitation rate histograms within the radar swath exhibited small but systematic differences from the histograms computed for just the parts of the GMI/TMI swath outside the radar swath. To resolve these biases, we compute the calibrations using the full GMI/TMI swath and the narrower 2BCMB retrievals. The spatial resolution of the 2BCMB(T) – GMI(TMI) match-up collection process was increased to a $9^\circ \times 5^\circ$ template in order to resolve issues of lack of complete spatial coincidence in this approach. The (climatological) 2AGPROFGMI-calibrated radiometer estimates are calibrated to 2BCMB using the 2AGPROFGMI/2BCMB adjustment coefficients.
5. The rain rate in each grid box is the calibrated conical-scan microwave radiometer estimate (i.e., GMI, TMI, SSMI, SSMIS, AMSR-E, AMSR2) contributing during the half hour, or the cross-track-scan microwave sounder estimate (i.e., AMSU-B, MHS, ATMS) if a conical-scan imager isn't available. Most of the time, during the dataset's half-hour window there is only one overpass of whatever the "best" type of sensor is; when more than one of that type is available, the one closest to the center of the period (15 or 45 minutes into the hour) is used. This selection is made because the histogram of rain rates is sensitive to averaging one, two, or three overpasses in a half-hour period. Very occasionally, two swaths of the same type sensor will coincide and their time fields will be essentially the same. Since the times are only recorded to the nearest minute, a complicated interleaved pattern can result, where the first sensor that is processed in a grid box is taken when the times are the same (to the nearest minute), and the second is taken when it is closer to the mid-point of the time period for that grid box. The same can also happen if one sensor is ascending and one is descending, in which

case the time comparison chooses the closer of the absolute values of the difference, but keeps the first sensor processed if the absolute values of the difference are the same.

6. Additional merged PMW fields in the intermediate data file describe the PMW merger include the instrument type producing the estimate, and the time of the instrument's overpass.
7. All of the merged PMW estimates are less accurate in regions with frozen or icy surfaces.

Note well that *MWprecipitation* only includes microwave data (hence "MW"), meaning it has significant gaps. The *precipitation* data field is the complete estimate that most users will want to access.

.....

The **Merged 4-Km IR Tb data set** is produced by the Climate Prediction Center (CPC), NOAA National Centers for Environmental Prediction, Washington, DC under the direction of P. Xie. Each cooperating geostationary (geo) satellite operator (the Geosynchronous Operational Environmental Satellites [GOES], United States; the Geosynchronous Meteorological Satellite [GMS], followed by the Multi-functional Transport Satellite [MTSat] and then Himawari, Japan; and the Meteorological Satellite [Meteosat], European Community) forwards infrared (IR) imagery to CPC. Then global geo-IR are zenith-angle corrected (Joyce et al. 2001), re-navigated for parallax, and merged on a global grid. In the event of duplicate data in a grid box, the value with the smaller zenith angle is taken. The data are provided on a 4-km-equivalent latitude/longitude grid over the latitude band 60°N-S, with a total grid size of 9896×3298.

The data set was first produced in late 1999, but the current uniformly processed record is available starting 17 February 2000. CPC is working to extend the record back to January 1998.

All 5 geo-IR satellites are used, with essentially continuous coverage. GMS-5 was replaced by GOES-9 starting 01 UTC 22 May 2003, which introduced slightly different instrument characteristics. Then starting 19 UTC 17 November 2005 the new Japanese MTSat-1R took over, followed by MTSat-2 on 1 July 2010. Himawari 8 took over effective 0200 UTC on 7 July 2015. There is an extended data dropout in the presently available IR dataset for Japanese sector from late on 17 November 2005 to the middle of 22 March 2006 due to issues in NOAA coping with the format of the then-newly introduced MTSat-1. Consequently, during that period a small sector over Japan lacks all data, and the adjacent IR are entirely based on high-zenith-angle data from the METEOSAT to the west and GOES-W to the east.

Each UTC hour file contains 2 data fields. All geo-IR images with start times within 15 minutes of the UTC hour are accumulated in the "on-hour" field. Images with start times within 15 minutes of the UTC hour plus 30 minutes are accumulated in the "half-hour" field. The nominal image start times for the various satellites and their assignment to half-hour fields are shown in Table 10.

Table 11. Nominal sub-satellite longitude (in degrees longitude) and image start time (in minutes past the hour) for the various geosynchronous satellites. The start times are displayed according to their assignment to either the on-hour or half-hour fields in the CPC Merged 4-Km IR Tb data set. Full-disc views are guaranteed only at 00, 03, ..., 21 UTC. These appear in the on-hour field except MTSat appears in the previous half-hour for all hours. For images not at these times, a satellite's "image" may be assembled from

various operator-specified regional sectors. MTSat provides N. Hemisphere sectors (only) on-hour, except S. Hemisphere sectors (only) at 00, 06, 12, 18 UTC.

| Satellite | Sub-sat. Lon. | on-hour | half-hour |
|--|--|---------|-----------|
| Himawari-8 (formerly MTSAT-2, MTSat-1R, GMS) | 140°E | 00 | 30 |
| GOES-E (8, 12, 13, now 16) | 75°W | 45 | 15 |
| GOES-W (10, 11, 15, 17, now 18) | 135°W | 00 | 30 |
| Meteosat-11 (formerly 5, 7, 8, 9, 10) | 9.5°E | 00 | 30 |
| Meteosat-9 (formerly 5, 7, 8) | 45.5°E (63°E for 5, 7; 41.5°E for 8) | 00 | 30 |

Table 12. Geo-IR satellites contained in the Merged 4-km IR Tb dataset provided by NOAA/NWS/CPC.

| | GMS | GOES-WEST | GOES-EAST | METEOSAT | INDOEX |
|------|----------------------|---------------------|--------------------|-----------------|-----------------|
| 1998 | GMS-5 | GOES-10 | GOES-8 | MET-7 | MET-5 |
| 1999 | GMS-5 | GOES-10 | GOES-8 | MET-7 | MET-5 |
| 2000 | GMS-5 | GOES-10 | GOES-8 | MET-7 | MET-5 |
| 2001 | GMS-5 | GOES-10 | GOES-8 | MET-7 | MET-5 |
| 2002 | GMS-5 | GOES-10 | GOES-8 | MET-7 | MET-5 |
| 2003 | GMS-5/ GOES9 | GOES-10 | GOES-8/ GOES-12 | MET-7 | MET-5 |
| 2004 | GOES-9 | GOES-10 | GOES-12 | MET-7/ MSG-1 | MET-5 |
| 2005 | GOES-9/ MTSAT-1R* | GOES-10 | GOES-12 | MSG-1 | MET-5 |
| 2006 | MTSAT-1R* | GOES-10/ GOES-11 | GOES-12 | MSG-1 | MET-5 |
| 2007 | MTSAT-1R | GOES-11 | GOES-12 | MSG-1/ MSG-2 | MET-5/ MET-7 |
| 2008 | MTSAT-1R | GOES-11 | GOES-12 | MSG-2 | MET-7 |
| 2009 | MTSAT-1R | GOES-11 | GOES-12 | MSG-2 | MET-7 |
| 2010 | MTSAT-1R | GOES-11 | GOES-12/ GOES13 | MSG-2 | MET-7 |
| 2011 | MTSAT-1R | GOES-11/ GOES-15 | GOES-13 | MSG-2 | MET-7 |
| 2012 | MTSAT-1R | GOES-15 | GOES-13 | MSG-2/ MSG-3 | MET-7 |
| 2013 | MTSAT-1R | GOES-15 | GOES-13 | MSG-3 | MET-7 |
| 2015 | MTSAT-1R/H-8 | GOES-15 | GOES-13 | MSG-3 | MET-7 |
| 2016 | H-8 | GOES-15 | GOES-13 | MSG-3 | MET-7 |

| | | | | | |
|------|-----|---------------------|---------------------|-------|-----------------|
| 2017 | H-8 | GOES-15 | GOES-13/ GOES-16 | MSG-3 | MET-7/ MSG-1 |
| 2018 | H-8 | GOES-15 | GOES-16 | MSG-3 | MSG-1 |
| 2019 | H-8 | GOES-15/ GOES-17 | GOES-16 | MSG-3 | MSG-1 |
| 2020 | H-8 | GOES-17 | GOES-16 | MSG-3 | MSG-1 |
| 2021 | H-8 | GOES-17 | GOES-16 | MSG-3 | MSG-1 |
| 2022 | H-8 | GOES-17 | GOES-16 | MSG-3 | MSG-2 |
| 2023 | H-8 | GOES-18 | GOES-16 | MSG-3 | MSG-2 |

* The beginning of the MTSat-1 record, 19 UTC on 17 November 2005 to the middle of 22 March 2006 is not available in the current IR record due to format issues at the time the data were introduced.

These data are used as input to IMERG processing.

.....

The Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN) Dynamic Infrared–Rain rate model (*PDIR*, Nguyen et al. 2020a,b) is based on the framework of the PERSIANN-Cloud Classification System (PERSIANN-CCS, Hong et al. 2004), which is in turn based on the original PERSIANN framework (Hsu et al. 1997, Sorooshian et al. 2000). PERSIANN and PERSIANN-CCS, along with PERSIANN-Climate Data Record (PERSIANN-CDR, Ashouri et al. 2015), are satellite-based data-driven precipitation estimating systems that use machine-learning techniques to establish a link between GEO cloud-top brightness temperature (Tb) and RR. PDIR advances the framework of the PERSIANN-CCS system by (i) improving the capture of warm precipitation by including a higher temperature threshold; (ii) introducing gradient filtering and applying morphological filters to improve the watershed method used for segmentation; (iii) expanding the cloud classification system to include monthly sets of cloud types, thus improving the algorithm’s ability to distinguish between different rainfall regimes; (iv) improving the skill of the Tb–RR curves with IMERG Merged PMW precipitation dataset; and (v) utilizing gridded climatology data from WorldClim version 2 (Fick & Hijmans 2017), and PERSIANN-CDR to create a dynamical Tb–RR curve model, optimized by the shuffled complex evolution optimization algorithm (SCE-UA; Duan et al. 1994). The central hypothesis behind PDIR is that characteristics related to the land surface below the clouds (topography, elevation, climate type, temperature climatology, etc.) play a role in the cloud-top temperature–rain rate (Tb–RR) relationship of clouds. In order to create a dynamic relationship between input IR-measured Tb and output RR, rainfall climatology values influenced by the unique land surface characteristics inherent to a region, are used to “slide” the Tb–RR curve along the Tb axis. This results in identical clouds with identical Tb intensities and distributions having different pixel RR values when occupying spaces with differing rainfall climatology.

.....

The **IR precipitation field**, data field *IRprecipitation* in 3IMERGHH data files, provides a global Level 3 (0.1°×0.1°-gridded) half-hourly IR estimate computed with PDIR. The units are mm/hr.

For the period 1 January 1998 - 03 UTC 16 June 1998 geo-IR data were not available in the Indian Ocean sector, but high-zenith-angle data from adjacent geo-satellites are generally sufficient for fill-in.

.....

The **precipitation phase**, namely whether it is liquid, solid, or mixed, is not currently provided as a satellite-based calculation by the standard PMW precipitation algorithms, so we must use ancillary data sets to create the estimate. [Note: the DPR does estimate phase using the radar data.] 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. [Note: a fraction between 0 and 100% gives a probability that the precipitation is liquid; it does *not* denote a mixture in which the fraction gives the precipitation that is liquid and the converse of which is therefore solid. By far, the most likely event is either “all liquid” or “all solid”, and the probability of each is the fraction and its converse, respectively.]

For the half-hourly data, we adopt the Liu scheme (personal communication, 2013; Sims and Liu 2015). The V06 form was a simple look-up table for probability of liquid phase (*probabilityLiquidPrecipitation*) as a function of wet-bulb temperature, with separate curves for land and ocean. In V07, the parameterization is tuned to ERA-5 and the ocean and land look-up tables were converted to functions for stability. 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 PLP for all grid boxes, including those with zero estimated precipitation. [This raises the possibility that the IMERG *probabilityLiquidPrecipitation* 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 JMA NWP forecast for the Initial Processing of Early and Late Runs, and the ECMWF Reanalysis Version 5 (ERA5) for the Final Run and the Retrospective Processing of Early and Late Runs. Caution should be exercised when comparing retrospectively processed and initially processed Early and Late Run data due to differences that result from the two input data sources (ERA5 vs. JMA NWP). These forecast and reanalysis fields are retrieved and reformatted by PPS.

GPROF2021 retrieves total hydrometeor mass in the atmospheric column and then implicitly correlates 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 estimates are calibrated to the passive microwave retrievals, again, without reference to precipitation phase. V07 is the first time that IMERG provides PMW data over frozen surfaces, and so directly provides calibration data for the IR. Previously, the IR calibrations were in-filled from surrounding areas in the frozen-surface areas, where PMW was previously masked out due to perceived low quality.

At the monthly scale the *probabilityLiquidPrecipitation* value 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. The monthly probability of liquid is computed as the precipitation-rate-weighted average of all half-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: $liquid\ precip = probabilityLiquidPrecipitation \times precipitation$, and $solid\ precip = (100 - probabilityLiquidPrecipitation) \times precipitation$.

Note well 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).

.....

The **probability of liquid phase field**, data field *probabilityLiquidPrecipitation*, provides a global Level 3 (0.1°×0.1°-gridded) half-hourly estimate in 3IMERGHH data files based on the implementation of the Liu wet-bulb temperature diagnostic discussed in “precipitation phase”. For the 3IMERGM data files, this field provides the precipitation-rate-weighted average of all half-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. See “precipitation phase” for more discussion of the monthly fields. In both cases, the field is globally complete as long as the input ancillary forecast/reanalysis data are complete. Regardless of phase, the precipitation units are depth of liquid, which for solid precipitation is usually referred to as Snow Water Equivalent (SWE), as discussed in “precipitation phase”. The temperature, humidity, and surface pressure fields used to compute the *probabilityLiquidPrecipitation* field are provided by the then-operational JMA NWP forecast for the Early and Late Runs. Note: the mis-match between description (“phase”) and field name (“precipitaiton”) will be rectified in V08.

.....

Under the **Kalman Smoother framework** as originally developed in CMORPH-KF and adapted here, the precipitation analysis for a grid box is defined in three steps (Joyce et al. 2011). First, the half-hourly PMW precipitation estimates closest to the target analysis time in both the forward and backward directions are propagated from their observation times to the analysis time using the storm system motion vectors computed from numerical analyses of precipitation (PRECTOT), total liquid water (TQL), and total vapor (TQV). The “prediction” of the precipitation analysis is then defined by compositing the forward- and backward-propagated PMW estimates with weights proportional to the mean correlations computed using three months of overlapping PMW and propagated estimates, updated once per month. If the gap between the analysis time and the closer (in time) PMW observation is longer than 30 min, the final "analysis" is defined by updating the “forecast” with IR-based precipitation observations with weights proportional to the observation correlation. This threshold is due both to the natural timescale of precipitation at these fine scales

and to the retrieval errors in the microwave algorithms. In practice, the actual weights used are the square of the correlations, also referred to as the coefficient of determination.

The post-real-time Final Run analysis, as well as all retrospective analyses are carried out with Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2) numerical assimilations which are available hourly on a $0.5^\circ \times 0.625^\circ$ lat/lon global grid. In near-real time MERRA-2 is not yet available, so the Early and Late Runs use TQV from Goddard Earth Observing System model Forward Processing (GEOS FP), which shares computational elements with MERRA-2 and provides TQV on an hourly $0.25^\circ \times 0.3125^\circ$ global grid. See Tan et al. (2019b) for more details.

See Tan et al. (2019b) for a discussion of morphing. To summarize, the storm system motion vectors used to propagate the PMW estimates are calculated by computing the pattern correlation between spatially lagged arrays from two consecutive images of various moisture-related fields. The spatial displacement with the highest correlation is used to define the cloud motion vector within each field. The storm system motion vectors are defined for each 2.5° lat/lon grid box using TQV data over a 5° lat/lon domain centered on the target grid box. Vectors are computed on an hourly scale. Then the first half-hour's values are the hourly vector divided by two, while the second half-hour's values are interpolated from adjacent hourly vectors. Finally, a spatial bilinear interpolation is used to shift to the 0.1° precipitation grid boxes.

Initially (in V06), we switched from GEO-IR to TQV to work with a globally complete field that was more closely related to precipitation. However, near sharp topography, such as the Andes, the TQV results are sometimes degraded. So, we choose to compute motions for PRECTOT and TQL as well, and apply them sequentially. Both PRECTOT and TQL are intermittent fields that are very closely and rather closely related to storm motion, respectively, so the hierarchy is to use PRECTOT vectors where they exist, else use TQL vectors where they exist, else use TQV vectors.

Errors for the individual satellite estimates are calculated by comparison against TMI/GMI estimates in the TMI/GMI eras, with all conical-scan imagers (CSI) lumped into one correlation table and all cross-track-scan sounders (CTSS) lumped into another. Because of the simple form used (see following), the error table for the TMI is the lumped CSI table in the TRMM era. During the GPM era, the TMI is treated as “just” another CSI and included in computing the lumped CSI error table (against GMI), and the lumped CSI error table is taken as representative of GMI. Expressed in the form of correlation, the errors for the propagated PMW estimates are defined as regionally dependent and seasonally changing values depending on sensor type (CSI, CTSS, IR) and the length of propagation time. Over land, the error coefficients are computed for each 10° latitude band using data collected over a 30° -wide latitude band centered on the target band. No zonal differences in the error are considered due to the limited sampling of the data. Over ocean, the error coefficients are defined for each $20^\circ \times 20^\circ$ lat/lon box using data over a $60^\circ \times 60^\circ$ lat/lon region centered on the target box. Over both land and ocean, the error functions are calculated for each month using data over a three-month period, trailing for Early and Late, and the current and previous two months for Final, to account for the seasonal variations. The comparisons against TMI/GMI are updated monthly.

These data are computed as part of IMERG processing.

.....

The **Kalman filter influence for IR field**, data field *IRinfluence* in 3IMERGHH data files, provides a global Level 3 (0.1°×0.1°-gridded) half-hourly listing of the relative weight (in percent) that the IR data get in the Kalman filter step described in “Kalman Smoother framework” compared to the total weighting from the three inputs, namely the forward-propagated PMW, backward-propagated PMW (not available for Early), and IR.

.....

The Scheme for Histogram Adjustment with Ranked Precipitation Estimates in the Neighborhood (**SHARPEN**) module was developed to address unphysical smoothing that Rajagopal et al. (2021) discovered in analyzing rainrate histograms of mesoscale convective systems. The data were collected during the Convective Processes EXperiment (CPEX), a NASA-sponsored field program conducted in May-June 2017 over the waters surrounding Florida. The issue is that the Kalman filter’s weighted average tends to both increase the areal coverage by light precipitation, due to averaging zero and non-zero values, and suppress high rainrates, due to averaging peak values with lower values. As described in Tan et al. (2021), SHARPEN performs a quantile-mapping-inspired computation for a local region. A PDF of Kalman filter (KF) estimates is assembled on an N×N template centered on a grid box, as well as a PDF being assembled from all of the forward, backward (if used), and IR inputs on the same template. These input values carry the KF weights with which they are tagged in the KF processing, which automatically gives more importance to the closest PMW overpass (or the IR if no PMW is close in time). In use, one finds the quantile rank of the central grid box value in the (weighted) KF estimates PDF, then one uses that quantile rank in the merged (weighted) PDF of inputs to find a revised value. When implemented in IMERG, precipitation estimates from SHARPEN exhibit a distribution that resembles that of the original instantaneous retrievals, with matching precipitating area and peak precipitation rates. Case studies demonstrate improved maps that bridge more realistically between the parent precipitation fields. Evaluation against ground observations reveals a distinct improvement in precipitation detection skill, but also a slightly reduced correlation, likely because of a sharper precipitation field. The increased computational demand of SHARPEN can be mitigated by striding over multiple grid boxes, which only marginally degrades the accuracy of the estimates.

SHARPEN is introduced in Version 07, with a template size of 25×25 (covering 2.5°×2.5° and a stride of 3).

.....

The monthly satellite-gauge, or **SG combination**, is computed as follows, following the TMPA approach for infusing monthly gauge information into the fine-scale satellite precipitation estimates (Huffman et al. 2007a, 2007b, 2010):

1. The original (i.e., before the scaling step) half-hourly KF-filtered (satellite-only) estimates are summed for the calendar month.
2. The GPCC monthly precipitation gauge analysis is undercatch-corrected (see “precipitation gauge adjustment”).

3. The monthly precipitation gauge analysis is used to create a large-scale (37×37 0.1° grid boxes) bias adjustment to these satellite-only estimates in regions where the gauge stations are available, mostly land. Note that analysis values distant from any gauges are not used.
4. The monthly gauge-adjusted satellite-only estimate is combined directly with the precipitation gauge analysis using inverse error variance weighting.

The monthly random error is also computed at this point as a separate field following Huffman (1997). See “random error” for a summary.

These data are computed in the Final Run alone.

.....

IMERG ends with a **precipitation post-processing step** that introduces gauge information into the multi-satellite-only half-hourly data (carried as field *precipitationUncal*). For the Final Run, the ratio between the monthly accumulation of half-hourly multi-satellite-only fields and the monthly satellite-gauge field is computed, then each half-hourly field of multi-satellite-only precipitation estimates in the month is multiplied by the ratio field to create the final half-hourly calibrated IMERG precipitation estimates.

The ratio between the monthly satellite-gauge and the monthly accumulation of half-hourly multi-satellite-only fields is limited to the range [0.2,3]. The cap of 3 was chosen because the value 2 (used in TRMM V6) was too restrictive. The value was moved to 3 because it did a better job of matching the two accumulations, while testing showed that 4 started introducing unrealistic shifts to the histogram of half-hourly precipitation rates for the month. Early in TRMM the lower bound was set to 0.5, but it was argued that a smaller value allows matching between the two accumulations without creating the egregious high snapshot values that result when the upper bound is expanded too far.

One important improvement compared to the TMPA is that CORRA is computed in real time for GPM. This contrasts to the situation in TRMM where the TCI was not computed in real time and we had to use a climatological monthly TCI calibration. This allows a simplification in the IMERG Early and Late Runs in that the spatially varying monthly climatological calibration to the Final’s monthly satellite-gauge analysis is not applied as in TMPA, since CORRA already has a climatological correction to GPCP. “Climatological” means all Januarys get the January adjustment, all Februarys get the February adjustment, and so on.

Note that in V06 we did not do this last calibration because we lacked the long record of Final at the time that the Early and Late started V06 Initial Processing, but it is done in V07.

The half-hourly random error is also computed at this point as a separate field following Huffman (1997). See “random error” for a summary.

.....

The **uncalibrated precipitation field**, data field *precipitationUncal*, provides a global Level 3 ($0.1^\circ \times 0.1^\circ$ -gridded) half-hourly field in 3IMERGHH data files of the morphed and KF-combined estimates computed in the Kalman filter step described in “Kalman Smoother framework”, which

is recorded before the last post-processing step. Note that in V06 we did not do this last calibration because we lacked the long record of Final at the time that the Early and Late started V06 Initial Processing (computed as new data are received), but it is done in V07. For the Final Run the ratio field for the month described above is applied as described in “precipitation post-processing step”. In 3IMERGM data files it provides the global Level 3 (0.1°×0.1°-gridded) monthly accumulation of uncalibrated 3IMERGHH data described in “precipitation post-processing step”. Differences in *precipitationUncal* values among Runs are driven by the following differences for Initial Processing:

1. The primary factor is which satellites have reported data by the time the Run is computed. The Early currently has a cut-off time of 3 hours after observation time, which means it lacks data for the backward part of the morphing and even some of the forward-propagated data. Late should have more data for forward propagation, and importantly, provides data for backward propagation by waiting almost 13 hr. Final might have more data in both the forward and backward propagation.
2. The CORRA calibration and Kalman statistics are trailing for Early and Late, but centered for Final.
3. The Early and Late use GPROF that applies JMA model output as ancillary data, while the Final applies ERA5 data.
4. The Final uses 7-hr forward propagation chunks for each half hour versus the sequential forward propagation used in the Early and Late.
5. The Early and Late use motion vectors computed from GEOS FP, while the Final uses vectors based on MERRA-2.

Retrospective Processing (RP; applied to previously received data) lacks differences 1, 3, and 5 – everything is computed as is done for the Final (except the Early continues to use forward propagation alone), leaving 2 and 4 as the drivers for the RP period.

.....

The **calibrated precipitation field**, data field *precipitation*, provides global Level 3 (0.1°×0.1°-gridded) half-hourly listing in 3IMERGHH data files of the calibrated data described in “precipitation post-processing step”. In 3IMERGM data files it provides the global Level 3 (0.1°×0.1°-gridded) monthly SG combination data described in “SG combination”.

.....

At user request, a **quality index** that provides a simple way of judging the utility of data products was introduced in Version 05 in data field *precipitationQualityIndex*. While the goal is reasonable, there is no agreement about how this quantity should be defined. After some discussion within the team, two distinctly different quality indices were chosen for the half-hourly and monthly data fields for implementation. It is a matter of investigation to determine if users find these insightful, or if different quality indices should be developed for future releases. Details are provided in the document “IMERG Quality Index”, available as an appendix in the ATBD document <https://gpm.nasa.gov/resources/documents/imerg-v07-atbd.pdf>.

Quality Index for Half-Hourly Data (QIh)

At the half-hourly scale, the best metric is some measure of the relative skill that might be expected from the fluctuating mix of different passive microwave- and infrared-based precipitation

estimates. The Kalman smoother used in IMERG (and originated in the CPC Morphing [CMORPH] algorithm, Joyce et al. 2011) routinely recomputes estimates of correlation between GMI and each of the other satellite estimates in coarse blocks across the entire domain (90°N-S) and then uses these correlation coefficients (squared) to provide weights for use in the combination of forward-propagated passive microwave, backward-propagated passive microwave, and current-time infrared precipitation estimates. In V07 the polar caps have the same treatment without IR data. However, the formalism never provides an overall correlation for the combined estimate, so one approach is provided here.

The correlation coefficients developed for the Kalman smoother are substituted for inverse error variance to compute the approximate correlation coefficient of the merged precipitation estimate. Furthermore, the correlation coefficients are transformed with the Fisher (1915) z statistic before the computation and back-transformed afterwards, which is a simple variance stabilization transformation. Formally, the Fisher transformation requires that the two variables being correlated follow a bivariate normal distribution. While this is not true for precipitation, we adopt this approach as a first approximation to computing the correlation coefficient of the combined precipitation estimate because its use as a Quality Index seems reasonable and useful. The units are non-dimensional correlation coefficients.

There is one additional issue: we lack the $t=0$ (zero half-hour) correlation of each constellation member to the GMI for computational reasons in the current implementation of IMERG and need an approximate value. Lacking strong justification for alternatives, in V05 we set the correlation to 1, but in V06, and now V07, we choose to compute a set of baseline monthly $t=0$ correlations using the data span December 2014 – November 2015 from the Level 2 (GPROF) passive microwave estimates after intercalibration. These baseline correlations are then dynamically adjusted based on nearby-in-time correlations. These are expected to be slightly higher than if they had been computed from the $t=0$ outputs of the morphing scheme due to the lack of equivalent post-processing of the Kalman correlations.

In V07, an investigation into the performance of PMW retrievals revealed that, while having diminished performance, they still carry useful information and generally outperform IR retrievals (personal communication with P. Kirstetter, 2023). Therefore, the longstanding policy of screening PMW retrievals over frozen surfaces is lifted in V07. The result is that, except for a handful of grid boxes (mostly at the poles), IMERG has complete global coverage, with instantaneous and propagated PMW estimates covering over the polar regions. Nevertheless, users should be aware of the diminished performance of these estimates and exercise appropriate caution when interpreting the results. The QI_h is based on regional correlation computed against GMI/TMI and then extrapolated to global coverage, which leads to unrealistically high values over the polar latitudes that do not reflect the actual (diminished) skill over frozen surfaces. Therefore, the QI value over frozen surfaces is reduced by a factor depending on whether it is an imager or sounder and whether the grid box is over sea ice or frozen land (as indicated by the daily NOAA Autosnow product). Specifically, these factors are 0.112 for imagers and 0.275 for sounders over frozen land, and 0.495 for imagers and 0.282 for sounders over sea ice. These four factors are derived by comparing a sample of the Kalman correlations of snapshot PMW estimates with those obtained from evaluation against KuPR by You et al. (2023)

The suggested “stoplight” ranges are:

- 0-0.4 = “red” questionable
- 0.4-0.6 = “yellow” fair quality
- 0.6-1 = “green” good

Quality Index for Monthly Data

At the monthly scale, a relatively well-founded metric exists for random error, 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°×2.5° of latitude/longitude, even though IMERG 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.

Unlike QIh, QIm was not changed in frozen-surface regions because the random error coefficients are tuned to rain, not snow. Both gauges and satellite are less certain for snowfall, so it is yet clear how the relationship might change.

The suggested “stoplight” ranges for QIm are:

- 0-2 = "red" equivalent to the gauge coverage in regions such as central Africa, where the lack of data in a gauge-only analysis a critical problem
- 2-10 = "yellow" the mid-range has enough gauge data to ensure reasonable bias adjustment, but still require interpolation to fill in gaps several grid boxes wide between stations more or less routinely
- 10+ = "green" these are developed areas with good-to-excellent gauge networks

.....

The **processing strategy for each version** is as follows: the shift to a new version begins with DPR, then the Combined, then GPROF for GMI, and subsequently GPROF for all partner radiometers. The IMERG team follows these developments and folds them into advancements in the combination process, making adjustments and calibrations as needed to give the most realistic data sets. While IMERG is still on the “old” version, but the GPROF input is on the “new” version, calibrations are instituted as necessary to approximate the “old” GPROF. In common with the other algorithms, once IMERG starts processing with the new version, “initial processing” results (input data being computed for the first time) are run in parallel with “retrospective processing” results (input data being used for dates that already have an old-version dataset).

Retrospective processing is key to creating consistent archives of data for users. This is true for users of all three Runs of IMERG, so all three will be reprocessed. By design, the Production processing system, which computes the Final Run, supports reprocessing. The RT processing system, which computes the Early and Late Runs, does not support reprocessing. This issue is addressed by installing clones of the Early and Late systems (“Early Clone” and “Late Clone”)

Correspondingly, the TRMM-era processing for IMERG V07 includes GPM as a partner satellite and cross-calibrates it with TRMM.

PPS Configuration and Processing

Several of the IMERG calibration steps use a “centered” approach. Centering uses data forward and backward relative to the observation time as input to the processing. In order to complete the TRMM-era IMERG through 31 May 2014 processing for some of the IMERG algorithm groups will extend beyond 31 May 2014 and overlap the GPM-era processing. Similarly, some GPM-era processing will commence in March 2014 and overlap the TRMM-era processing.

Table 13 details the IMERG algorithm groups and processing date ranges for each era. Different product versions are applied to the intermediate files to avoid duplication of file names for any overlapping processing dates and to ensure that TRMM and GPM calibrations are “pure”, using a specific version of the intermediate products as input. Algorithm groups processed in the DATA_PREP stream are not inter- or cross-calibrated and may be run once and used for input across the combined TRMM and GPM eras.

Table 13. Layout of practical details for IMERG algorithm groups, including file version nomenclature designed to ensure unique file names during the TRMM-GPM overlap period.

| Processing Stream | IMERG Algorithm Group (#) | Process Boundary | Start of Processing* | End of Processing | Version |
|--------------------------|--|-------------------------|-----------------------------|--------------------------|----------------|
| TRMM_FINAL | Grid <i>MWprecipitation</i> (1) | Day | 1997-12-08 | 2014-02-28 | V07A |
| TRMM_FINAL | Inter-calibrate <i>MWprecipitation</i> (2) | Pentad | 1997-12-12 | 2014-06-29 | V07T |
| TRMM_FINAL | Merge <i>MWprecipitation</i> (3) | Day | 1998-12-08 | 2014-06-14 | V07T |
| TRMM_FINAL | Propagation <i>MWprecipitation</i> (6) | Day | 1998-12-08 | 2014-05-31 | V07T |
| TRMM_FINAL | Compute <i>IRprecipitation</i> (7-1 & 7-2) | Pentad | 1998-12-12 | 2014-06-14 | V07T |
| TRMM_FINAL | Compute <i>IRprecipitation</i> (7-3) | Day | 1998-12-08 | 2014-05-31 | V07T |
| TRMM_FINAL | Compute Kalman Filter Precip (8) | Month | 1998-01-01 | 2014-05-31 | V07T |
| TRMM_FINAL | Compute Sat-Gauge Precip (9) | Month | 1998-01-01 | 2014-05-31 | V07A |
| GPM_FINAL | Grid <i>MWprecipitation</i> (1) | Day | 2014-03-01 | TBD | V07A |
| GPM_FINAL | Inter-calibrate <i>MWprecipitation</i> (2) | Pentad | 2014-03-12 | TBD | V07G |
| GPM_FINAL | Merge <i>MWprecipitation</i> (3) | Day | 2014-03-09 | TBD | V07G |
| GPM_FINAL | Propagation <i>MWprecipitation</i> (6) | Day | 2014-03-09 | TBD | V07G |
| GPM_FINAL | Compute <i>IRprecipitation</i> (7-1 & 7-2) | Pentad | 2014-03-12 | TBD | V07G |
| GPM_FINAL | Compute <i>IRprecipitation</i> (7-3) | Day | 2014-03-09 | TBD | V07G |

| Processing Stream | IMERG Algorithm Group (#) | Process Boundary | Start of Processing* | End of Processing | Version |
|-------------------|----------------------------------|------------------|----------------------|-------------------|---------|
| GPM_FINAL | Compute Kalman Filter Precip (8) | Month | 2014-06-01 | TBD | V07G |
| GPM_FINAL | Compute Sat-Gauge Precip (9) | Month | 2014-06-01 | TBD | V07G |
| DATA_PREP | Compute Propagation Vectors (4) | Day | 1997-12-01 | TBD | V07A |
| DATA_PREP | Compute Snow Grid (5) | Day | 1997-12-01 | TBD | V07A |
| DATA_PREP | Convert ECMWF Model Data (10) | Day | 1997-12-01 | TBD | V07A |

* The start and end of processing in the table is relative. The specific processing date range is set according to the availability of the constellation data compared to start/end times listed in the table.

Table 14. Estimates of file counts and sizes used in IMERG across the entire TRMM-GPM era. The letters *i*, *o*, *t*, *a*, *s* in “Module Relation” indicate input, output, transfer (between modules or within a module), accumulator, and static, respectively.

| INTERMEDIATE/STATIC | | | | |
|---------------------|---------|-----------|-----------------|-----------------|
| ID | # Files | Size (MB) | Total Size (MB) | Module Relation |
| CORRA-G accum | 1 | 1385 | 21905 | a2 |
| TMI-CORRA accum | 1 | 1385 | 11795 | a2 |
| GMI-other cal | 21 | 25 | 1685 | s2 |
| Surface type | 1 | 25 | 25 | s2 |
| TMI-other cal | 21 | 25 | 1685 | s2 |

| INPUT | | | | | |
|-------------------|------------|-------------|-------------------|-----------------|-----------------|
| ID | # Granules | Granularity | Granule Size (MB) | Total Size (MB) | Module Relation |
| AIRS | 1 | One orbit | 11 | 11 | i2 |
| AMSR x 2 | 2 | One orbit | 56 | 112 | i2 |
| AMSU x 3 | 3 | One orbit | 56 | 168 | i2 |
| ATMS x 2 | 2 | One orbit | 56 | 112 | i2 |
| CORRA-G | 1 | One orbit | 328 | 328 | i2 |
| CORRA-T | 1 | One orbit | 328 | 328 | i2 |
| CrIS | 1 | One orbit | 11 | 11 | i2,i4 |
| GEOS FP | 1 | One hour | 2.5 | 2.5 | i4 |
| GMI | 1 | One orbit | 56 | 56 | i2 |
| GPCC Full | 1 | 120 months | 178 | 178 | i11 |
| GPCC Monitoring | 1 | One month | 2 | 2 | i11 |
| Autosnow (binary) | 1 | One day | 40 | 40 | |
| Autosnow (netCDF) | 1 | One day | 5 | 5 | |

| | | | | | |
|--------------------------------|---|--------------|-----|-----|--------------|
| IR | 2 | One hour | 65 | 130 | i1,t1-4,t1-8 |
| JMA Forecast Tsfc, RHsfc, Psfc | 1 | forecast run | 117 | 117 | i11,i12 |
| ERA-5 Tsfc, RHsfc, Psfc | 1 | analysis | 226 | 226 | i11,i12 |
| MERRA-2 | 1 | One day | 60 | 60 | i4 |
| MHS x 5 | 5 | One orbit | 56 | 280 | i2 |
| SSMI x 3 | 3 | One orbit | 56 | 168 | i2 |
| SSMIS x 4 | 4 | One orbit | 56 | 224 | i2 |
| TMI | 1 | One orbit | 56 | 56 | i2 |
| TOVS | 1 | One orbit | 11 | 11 | i2 |

| <i>OUTPUT</i> | | | | | |
|--------------------------------|-------------------|--------------------|--------------------------|------------------------|------------------------|
| <i>ID</i> | <i># Granules</i> | <i>Granularity</i> | <i>Granule Size (MB)</i> | <i>Total Size (MB)</i> | <i>Module Relation</i> |
| Half-hourly IMERG | 1 | 30 minutes | 8 | 8 | o11,o12 |
| Monthly IMERG (Final Run only) | 1 | One month | 20 | 20 | o11 |

| <i>TRANSFER</i> | | | | | |
|--------------------------------------|-------------------|--------------------|--------------------------|------------------------|------------------------|
| <i>ID</i> | <i># Granules</i> | <i>Granularity</i> | <i>Granule Size (MB)</i> | <i>Total Size (MB)</i> | <i>Module Relation</i> |
| Cloud classification | 1 | 30 minutes | 6900 | 6900 | s8 |
| CORRA-G cal | 1 | One orbit | 259 | 259 | t2 |
| CORRA-T cal | 1 | One orbit | 259 | 259 | t2 |
| Gridded MW | 14* | One orbit | 103 | 1442 | t2 |
| Intermediate IR | 1 | 30 minutes | 52 | 52 | t8 |
| Intermediate MW | 1 | 30 minutes | 106 | 106 | t9 |
| Merged PMW/IR (uncal) | 1 | 30 minutes | 149 | 149 | t7-11,t7-12 |
| Motion Vectors | 2 | 30 minutes | 0.1 | 0.2 | t4,t4-6 |
| Kalman Filter weights | 1 | One month | 2600 | 2600 | t6-7 |
| PMW | 2 | 30 minutes | 106 | 212 | t2-5,t2-5,t2-9 |
| PMW forward and backward propagation | 2 | 30 minutes | 38 | 76 | t5-7 |
| PDIR precip (unadjusted) | 1 | 30 minutes | 54 | 54 | t8-10 |
| PDIR precip (adjusted) | 1 | 30 minutes | 54 | 54 | t10 |

* Although many sources of “high-quality” satellite data are listed under “INPUT”, it is assumed that no more than 14 such satellites will be available at any given time.

.....

The **units of the IMERG estimates** are listed in the tables in “3IMERGHH data fields” and “3IMERGM data fields” and in the file metadata. Recall that the half-hourly data are best thought

of as snapshots, valid during the stated half hour, since Villarini and Krajewski (2007) and O et al. (2017) found that 3B42 and IMERG-F (respectively) estimates correlated best with gauge accumulations over 10's of minutes. The monthly values are average rates over the month, although the probability of liquid phase is weighted by the precipitation rate accompanying each half hour's probability.

.....

Several **processing considerations** arise due to the multi-layered nature of the code system. We have not enforced consistency on the various stages in the sense that some steps might be programmed as multiple modules, while others will be computed in a single module. The data flow between modules, and between executions of the same module, is carried out using files, which typically have fixed names. Input and output datasets necessarily have names that reflect the time sequencing of the data that they contain.

Calibrations

The satellite-satellite calibrations, which include the PMW intercalibrations to a TRMM/GPM standard, IR-PMW precipitation calibration for the IR estimates, and the Kalman filter weights, are conceptually asynchronous with the actual half-hourly precipitation dataset processing. It is a matter of computational choice within PPS as to whether the calibrations are run sequentially or in parallel, but the system is designed to be very forgiving of occasional missed calibration match-ups – without significant loss of skill it can run with the then-current calibration files, as long as the dropouts do not become too severe. The heritage TMPA system computed the PMW intercalibration on a calendar month basis, while the PDIR and CMORPH-KF run the IR-PMW and KF weights, respectively, on trailing accumulations of match-ups. For IMERG we necessarily run all the Early and Late calibrations on trailing accumulations of match-ups. The post-real-time Final run has to wait for the GPCC precipitation gauge analysis and the ECMWF ancillary data, so we accumulate the match-ups with a sufficient delay after real time that the Final calibrations are approximately centered, with the exception of the Kalman statistics, which consist of the current and previous two months of data.

The only important difference between near- and post-real-time runs comes in the last calibration, which could be computed for the near-real-time as climatological adjustments to the Final product, although in V06 no calibration was applied, and for the post-real-time as calendar-month adjustments to, and combination with monthly gauge analyses.

We compute three runs of the algorithm, namely the “Early”, “Late”, and “Final” Runs at about 4 hours, 14 hours, and 3.5 months after observation time, respectively. The simplest approach is chosen in Initial Processing (IP), namely to maintain three entirely separate sets of files and to compute everything in each run. This eliminates dependencies between runs and facilitates retrospective processing. On the other hand, in Retrospective Processing (RP) we choose to compute the Final Run first, save selected intermediate files, and then employ these in computing the RP for the Early and Late Runs. One side effect of this choice is that the Early and Late RP datasets are based on a superset of what was actually available in near-real time, as well as using the CLIMO versions of precipitation estimates. A second is that each newly computed version has a gap between the newly computed Early and Late RP record and the start of their IP record, which

is only filled as subsequent months of Final data are computed as IP with the (well after real time) arrival of the ancillary and gauge data.

Parallelization

The forward V05 retrospective processing revealed that the original CMORPH-KF propagated time series was programmed to be efficient in computational resources but required serial processing. This constituted a significant impediment to timely completion for a long RP record. Thus, in V06, and now V07, the forward propagated time series of PMW was re-worked to be computed from scratch for each half hour. Taking a cue from the backward propagation, which of necessity is recomputed anew for each half hour, the HQ data in the previous 7 hours is used to develop the forward-propagated PMW field, since the KF correlations beyond 7 hours are negligible. While less efficient than the original scheme in total compute time, the shift allows parallel computation for each half hour's forward propagation and significantly decreases wall clock time.

As well, the rotating accumulation files used for calibration impose a serial requirement. It is possible, although not done in V07, to have several "chunks" of years, each starting from scratch and just filling accumulation files until they are "full", then computing results. This would allow a coarse-scale parallelization that could reduce reprocessing wall clock time by a factor of the number of chunks (plus extra overhead).

.....

An *IMERG-TMPA design comparison* begins with noting that both TMPA and IMERG use a constellation of passive microwave satellites, and within the general umbrella groups of "sounder" and "imager" the inputs are much the same, although at the end of the TRMM era the TMPA was not upgraded to include the newer satellites. The direct inputs of the TMI and GMI are swamped by the amount of data from the rest of the microwave sensors, so the absence of TMI in the last 4.5 years of TMPA was not a major problem. At the back end of the multi-satellite algorithms, both TMPA and IMERG use the same scheme for combining satellite data with the GPCC analysis, although IMERG uses the GPCC Final analysis up through 2020, which tends to be more accurate than the GPCC Monitoring analysis that the TMPA used for the last ~9 years. What's different? The algorithms for the Combined products are very different (2B31 for TMPA and CORRA for IMERG), and that is what provides calibration. The GPROF algorithm has been upgraded multiple times for use in IMERG – still Bayesian, but with the libraries of profiles sourced and organized differently. The IR scheme has shifted from VAR to PDIR – very different approaches. Compared to the simple chunking of data into 3-hour intervals in TMPA, note the massive amount of time spent in IMERG on morphing and the Kalman filter. The goal is two-fold:

- 1) Provide a finer time resolution so that system evolution is more accurately captured, compared to the 3-hour interval in TMPA. This improved evolution not only provides more-frequent data values, but it should also make the IMERG time averages (such as daily) more accurate, since precipitation changes so rapidly in space and time.
- 2) Reduce the use of IR estimates, which have lower quality, by time-interpolating the microwave estimates, which have higher quality. The hard part here is that the interpolation has to be done in a quasi-Lagrangian framework because the rain systems move. So, IMERG's morphing/Kalman framework is intended to minimize the IR contribution, even though IR is still seen as necessary in regions with long microwave gaps.

.....

7. Error Detection and Correction

Even before launch, GPM instituted the GPM Cross-Calibration Working Group (**X-Cal Working Group**), which was charged with developing a consistent “Level 1C” calibration of all constellation radiometers to the GMI. Their work identified a range of problems and developed very careful radiometric calibrations for use in GPM (and the community at large). See PPS et al. (2016) for details.

.....

Navigation error for GPM and TRMM has been carefully addressed by the X-Cal Working Group and PPS. In general, the pointing errors for both satellites are a fraction of a typical TMI/GMI or PR/DPR footprint.

.....

DPR and PR error detection/correction has several parts. The performance of the various radar components, including transmit power and Low Noise Amplifiers, are monitored. An active ground calibration target is episodically viewed, and surface Z_0 is routinely monitored. See https://pps.gsfc.nasa.gov/Documents/PR_Manual_JAXA_V6.pdf for more information about PR.

Accuracies in the radar data are within the uncertainties of the precipitation estimation techniques.

The TRMM satellite altitude boost in August 2001 introduced some changes in detectability for which the PR algorithm is supposed to approximately account. The TRMM PR electronics failure on 29 May 2009 resulted in a switch to the “B-side” electronics. Some residual differences are noticeable despite careful work to harmonize the record.

Anomalies in the DPR performance are listed in

https://gpmweb2https.pps.eosdis.nasa.gov/tsdis/AB/docs/gpm_anomalous.html

while anomalies in the PR performance are listed in

https://gpmweb2https.pps.eosdis.nasa.gov/tsdis/AB/docs/partner_anomalous.html

.....

GMI, TMI, and constellation sensor error detection/correction has several parts. The SSMI is typical: built-in hot- and cold-load calibration checks are used to convert counts to antenna temperature (T_a). An algorithm has been developed to convert T_a to brightness temperature (T_b) for the various channels (eliminating cross-channel leakage). Differences between the T_a -to- T_b conversions employed by the various data providers imply that uncertainties in the T_a -to- T_b conversion are much larger than any other known uncertainty. Consequently, GPM developed the concept of a Level 1C, which applies radiometric corrections for all constellation sensors to GMI. See “X-Cal Working Group” for more information on this intercalibration.

Accuracies in the Tb's are within the uncertainties of the precipitation estimation techniques. Trending data are tracked for individual satellites, particularly near their end of life to determine whether sensor degradation is introducing unacceptable error.

GPM, TRMM, and Megha-Tropiques fly in orbits that precess over 83-, 46-, and 51-day periods, respectively. There is no direct effect on the accuracy of their sensor data, but the continually changing diurnal sampling causes significant systematic fluctuations in the resulting GMI-, TMI-, and SAPHIR-only precipitation estimates.

Some constellation satellites experienced significant drift in their (nominally) Sun-synchronous overpass time during their period of service (see <https://gpm.nasa.gov/resources/images/gpm-constellation-overpass-times> for a current time-series plot). There is no direct effect on the accuracy of the precipitation estimates, but it is possible that the systematic change in sampling time could introduce subtle shifts in the resulting collection of precipitation estimates. Additionally, the climatological calibrations of the constellation satellites against GMI/TMI are based on a snapshot of fixed overpass times, so if the constellation overpass times drift the calibrations will “drift” as well.

Anomalies in the GMI performance are listed in

https://gpmweb2https.pps.eosdis.nasa.gov/tsdis/AB/docs/gpm_anomalous.html

while anomalies in the TMI and constellation sensor performance are listed in

https://gpmweb2https.pps.eosdis.nasa.gov/tsdis/AB/docs/partner_anomalous.html.

Starting in February 2023, the NRT MetOp 1C data stream has occasionally provided partial orbit granules that duplicate previous granules, but contain faulty data. PPS instituted QC procedures as of 11 July 2023 that can successfully screen out 12 of 16 known cases. Further work is necessary to capture the rest of the cases without mistakenly removing good granules.

.....

The dominant **IR data correction** is for slanted paths through the atmosphere. Referred to as "limb darkening correction" in polar-orbit data, or "zenith-angle correction" (Joyce et al. 2001) in geosynchronous-orbit data, this correction accounts for the fact that a slanted path through the atmosphere increases the chances that (cold) cloud sides will be viewed, rather than (warm) surface, and raises the altitude dominating the atmospheric emission signal (almost always lowering the equivalent Tb). The slant path also creates an offset to the geolocation of the IR pixel due to parallax. That is, the elevated cloud top, viewed from an angle, is located closer to the satellite than the point at which the line of sight intersects the Earth's surface. Pixels are moved according to a standard height-Tb-zenith angle profile, at the price of holes created when tall clouds are moved farther than shallow clouds behind them. In addition, the various sensors have a variety of sensitivities to the IR spectrum, usually including the 10-11 micron band. Inter-satellite calibration is computed with GOES-E as the standard. The satellite operators are responsible for detecting and eliminating navigation and telemetry errors.

.....

The **IR sensor error detection/correction** addresses two issues. The IR Tb screening routine was developed to remove artifacts in the IR field that manifest as arcs of low Tb. These arcs occur randomly in time, but their fixed locations suggest imperfections in stitching the observations from the various geosynchronous satellite observations. Since this merged IR product is not produced by the IMERG team, our approach is to develop a screening routine that removes the artifacts. This was first implemented in V06.

Since it was observed that these arcs all occur adjacent to pixels with missing value, the screening routine identifies Tb pixels below 182 K that have a missing value (designated with a value of 330 K) in any of its eight surrounding pixels. Upon detection of such a pixel, a region stretching 95 pixels in both latitude directions and 50 pixels in both longitude directions is masked. Note that each pixel in the IR Tb field is about 0.0363° . These parameters are determined by trail-and-error to best cover the arcs while minimizing the masking of actual data.

The CPC 4-km infrared data are susceptible to an issue with the cooling system aboard the GOES-17 (GOES-W) satellite, which results in episodic periods of noise in some channels. See <https://www.goes-r.gov/users/GOES-17-ABI-Performance.html> for details. While NOAA has processes in place to attempt to mitigate the issue, the IR temperature values remain nonphysical at scattered locations in the eastern Pacific and over Alaska and Canada at four times of year, generally February, April, August, and October, for periods ranging from one to ten days, concentrated in the 11:30-15:00 UTC window. These conditions cause IMERG to produce artifacts in the form of line segments of precipitation.

Beginning with data from February 15, 2021, a screening technique was implemented in the post-real-time IMERG V06, and now V07, to prevent these artifacts from affecting subsequent processing. At each pixel, the sum of the absolute value differences between the pixel and its eight surrounding neighbors is computed. If the total difference exceeds 150 K, the pixel is changed to a missing value of 330 K. This process is applied to every half-hourly IR field at every pixel. Nominally, a small number of pixels are masked at every half-hour throughout the year, as it does not appear possible to fine-tune the technique to trap only the problematic cases. However, this method has proven quite effective in eliminating the unrealistic precipitation line segments that are triggered by these episodes without materially affecting the resulting precipitation fields.

Anomalies in the GMI performance are listed in

https://gpmweb2https.pps.eosdis.nasa.gov/tsdis/AB/docs/gpm_anomalous.html

while anomalies in the TMI and constellation sensor performance are listed in

https://gpmweb2https.pps.eosdis.nasa.gov/tsdis/AB/docs/partner_anomalous.html.

.....

We apply a **precipitation gauge adjustment** to account for undercatch. That is, precipitation gauges suffer a variety of errors in collecting precipitation, including evaporation, splashing, side wetting, and wind effects, with all resulting in a low bias for most gauge configurations. The wind effects occur because the air has to flow around the opening of the gauge and hydrometeors tend to follow the air flow. This is most true for the hydrometeors that fall the most slowly, namely drizzle and snowflakes. Undercatch ranges from 5% in heavy rain situations to 100%, 200%, or

more for snow, and depends on the design of the gauge (Legates 1987; Sevruk 1989). Until recently, the state of the art was a set of monthly maps of climatological adjustment ratios computed by Legates and Willmott (1990), and these are used to adjust the gauge analyses in this work. The GPCP computes daily adjustments based on station meteorological data for the Monitoring Analysis starting in 1982 (Schneider et al. 2014) based on Fuchs et al. (2001); a climatology of these estimates proved useful for adjusting GPCP fields in northern Eurasia compared to GRACE data (Behrangi et al. 2019), so coefficients based on monthly Fuchs et al. (2001) climatology are substituted for Legates and Willmott (1990) values north of 45°N in Eurasia.

.....

A number of **known errors and anomalies** are contained in part or all of the current IMERG archive. They have been uncovered by visual inspection and other diagnostics. In some cases, corrections will be applied in the next reprocessing, while for others no satisfactory correction is possible. Other items will be included in future reprocessing cycles as possible. For ease of document maintenance, some of the following items imply the known error by stating what upgrade was applied.

General

1. In general, GPROF2021 has excessive coverage by light precipitation, even considering that the estimates are made on the scale of the ~19 GHz channel footprint, yet at the monthly scale the correct accumulation appears to be given by including this unrealistic drizzle. At present, IMERG performs thresholding at 0.03 mm/hour on individual footprints, then rescales the data to approximate the monthly accumulation. In most locations this rescale is a small change from the original values.
2. The previous GPROF2017 estimates had a variety of artifacts associated with coastal regions that are sensor- and scene-dependent; we are studying GPROF2021's behavior, but suspect it is similar. In GPROF2017: For conical-scan imagers, inland water bodies in the Southeastern U.S. (Tian and Peters-Lidard 2007), Lake Nasser in Egypt, and desert coastal regions show anomalous high precipitation, while oceanic coastal regions in a variety of rainy situations tend to be deficient in precipitation. For cross-track-scan sounders the land/water interface in coasts provokes somewhat high estimates.
3. The Version 07 GPROF estimates for AMSU, ATMS, and MHS do not provide estimates for the 5 and 8 footprints (respectively) at each swath edge. This is due to algorithm issues as revealed in V05 GPROF testing. Taken together, these limitations somewhat reduce the amount of microwave-based data contained in Version 07 IMERG.
4. The current satellite-gauge combination scheme is trimmed to the coast to prevent coastal gauge data from "bleeding" into coastal waters, up to 1° away from the coast. This is particularly noticeable where there is heavy precipitation in the gauge analysis, but modest values off-shore, but seems preferable to a blocky transition along the coast. As well, isolated island stations are masked out as likely unrepresentative of the ocean that fills most of the grid box.
5. The GPROF retrievals, and the subsequent IMERG calibrations for them, are not entirely consistent between sensors. Accordingly, users will observe "flashing" in the precipitation fields as successive overpasses in a particular location are populated by different sensors.

However, the SHARPEN and t=0 KF routines introduced in V07 largely eliminate the artifacts that were apparent in previous versions.

6. The IR estimates sometimes exhibit “flashing” because successive images in a given region are successively populated with data from different geo-IR satellites, usually with one having a near-nadir view, but then dropping out and being replaced by high-zenith-angle data from an adjacent geo-IR satellite.
7. More generally, animations of V07 show minor "ripples" in the diurnal cycle, although these are considerably reduced from V06. Explicating the source of this variation is a matter of current research, but the leading candidates are differences in the various Sun-synchronous PMW overpasses and variations between the PMW- and IR-based estimates. See “diurnal cycle” for more discussion.
8. The complement of constellation satellites has changed over time, which affects the overall performance of the algorithm in two ways. First, the relative weighting of conical-scan imagers versus cross-track-scan sounders shifts, and second, the relative proportion of IR-based estimates changes. The PMW sensor inventory is shown in “sensors contributing to IMERG”. See Behrangi et al. (2014) for more discussion of sensor performance for legacy algorithms.
9. As of 14 June 2019, it was discovered that Autosnow sometimes provides incorrect values other than “sea ice” above 89°N. These values cause IMERG to provide precipitation values in these boxes, when in fact all gridbox values above 89°N should be “missing” for the *precipitationUncal* and *Precipitation* variables. This issue is corrected in V07 IMERG and will be corrected in a future reprocessing of Autosnow.
10. UTC progresses in lockstep with International Atomic Time (TAI), except when UTC is occasionally adjusted with a leap second to account for changes in the rotation of the Earth, which is slowly decreasing. For example, a leap second was inserted between 23:59 UTC 31 December 2016 and 00:00 UTC 1 January 2017. The events that happened in that second are attributed to 23:60 UTC 31 December 2016, but TAI ascribes the events to 00:00 1 January 2017. Thus, the date/time for TAI gets further ahead of UTC with each leap second.

| clock-tick seconds | N | N+1 | N+2 | N+3 | N+4 | N+5 | |
|--------------------|------------------------|------------------------|------------------------|----------------------|----------------------|-----|----------------|
| “TAI” | 2016-12-31 23:59:58 | 2016-12-31 23:59:59 | 2017-1-1 00:00:00 | 2017-1-1 00:00:01 | 2017-1-1 00:00:02 | | IMERG metadata |
| UTC | 2016-12-31 23:59:58 | 2016-12-31 23:59:59 | 2016-12-31 23:59:60 | 2017-1-1 00:00:00 | 2017-1-1 00:00:01 | | Input data |

leap second
inserted
into UTC

Fig. 3 Example of TAI and UTC relation across a leap second insertion. For simplicity, “TAI” is set equal to UTC before the insertion for illustrative purposes.

IMERG’s input datasets have date/times in UTC. But, the half-hour chunking of the orbits doesn’t account for leap seconds, meaning it’s in TAI. As a result, the input data may be behind the IMERG half-hourly bounds by a few seconds. Internally, IMERG converts date/time to seconds-since-base-date/time to permit computations. So, the offset between UTC and TAI depends on how many leap seconds have occurred since the base date/time. It is also the case that different parts of IMERG have different base date/times, so the (very small) time offsets differ in different parts of the processing. All of these effects are so small compared to

uncertainties such as the most representative time interval for the passive microwave and IR estimates that we ignore the UTC-TAI differences.

Specifically, the metadata providing the date/time and date/time in seconds since a base date/time does not reflect leap seconds. Although this does not affect the strictly increasing number of seconds with time within IMERG, it creates an offset from the number of seconds carried in the TAI time scale since the base time. The difference grows with time and, while it is small, some applications may be sensitive to it.

See <https://www.nist.gov/pml/time-and-frequency-division/time-realization/leap-seconds> for a current list of leap seconds.

| Date | Date | Date | Date |
|------------|------------|------------|------------|
| 2016-12-31 | 1998-12-31 | 1989-12-31 | 1979-12-31 |
| 2015-06-30 | 1997-06-30 | 1987-12-31 | 1978-12-31 |
| 2012-06-30 | 1995-12-31 | 1985-06-30 | 1977-12-31 |
| 2008-12-31 | 1994-06-30 | 1983-06-30 | 1976-12-31 |
| 2005-12-31 | 1993-06-30 | 1982-06-30 | 1975-12-31 |
| | 1992-06-30 | 1981-06-30 | 1974-12-31 |
| | 1990-12-31 | | 1973-12-31 |
| | | | 1972-12-31 |
| | | | 1972-06-30 |

11. In V06 we did not calibrate the Early and Late products to the Final because we lacked the long record of Final at the time that the Early and Late started V06 Initial Processing, but it is done in V07.

Specific

12. For the period 1 January 1998 - 03 UTC 16 June 1998 there were no geo-IR data available in the Indian Ocean Sector. High-zenith-angle observations from adjacent satellites are used for fill-in as available.
13. GMS data are missing for 21 UTC 4 January 1998 – 21 UTC 8 January 1998. Since the Meteosat data over the Indian Ocean sector do not begin until mid-1998, this results in a lack of IR data over the East Asia sector.
14. PPS maintains a listing of microwave constellation anomalies at https://gpmweb2https.pps.eosdis.nasa.gov/tsdis/AB/docs/gpm_anomalous.html.
15. The TRMM orbital altitude was raised from 350 to 401.5 km in August 2001 to extend the life of the mission by reducing the amount of fuel needed to maintain the orbit. This caused small changes in footprint size and minimum detectable precipitation rates. The GPROF2021 algorithm is supposed to account for these changes, but tests show small unavoidable differences that are still being researched.
16. The TRMM PR suffered an electronics failure on 29 May 2009. Data were lost until the “B-side” electronics were activated on 19 June 2009. Small residual differences remain between A-side and B-side data.
17. At 2045 UTC on 21 March 2012 GOES-15 (WEST) suffered a “bad momentum unload” and ceased recording data. Imaging was restored at 1722 UTC on 23 March 2012. In the interim

GOES-13 (EAST) was shifted to recording full-disk images. Use of higher-zenith-angle GOES-13 and MTSat-1 data largely covers the gap caused by the GOES-15 drop-outs.

18. In previous versions, but not in GPROF2021, F17 SSMIS has anomalously high precip values for a few scans over Brazil in the 21Z 26 April 2013 3B42 HQ and multi-satellite precip fields.
19. A TRMM spacecraft anomaly resulted in the loss of most TRMM sensor data for the period 02-14 UTC on 12 November 2013, and additional issues resulted in data gaps during the period 20-23:30 UTC. This reduces the data content in IMERG somewhat, but is not a serious issue overall.
20. Snow accumulation on the receiving antenna prevented reception of MTSAT-2 data from 1832 UTC on 14 February 2014 to 1232 UTC on 15 February 2014. The data were lost.
21. Metop-A experienced an anomaly that prevented data collection for 1400 UTC 27 March 2014–0746 UTC 21 May 2014.
22. The GMI instrument on the GPM Core Observatory went into safe mode on 22 October 2014. It was returned to operations on 24 October 2014 after determining that the cause was a faulty thermistor that provides information on the GMI's environment, but is not critical to its operation. Subsequently, non-critical thermistors have been removed from the Core Observatory's spacecraft health and safety alarm conditions.
23. For the periods 5 April – 18 May 2016 and 9 March 2017 to the present the F17 SSMIS 37V channel experienced intermittent noisy values and is not used in GPROF.
24. The data flow from the Himawari-8 GEO satellite was interrupted for 57 hours, 05 UTC 7 May 2016 – 14 UTC 9 May 2016, and during that time there is a continuous zone of missing values in the *IRprecipitation* data in the center of the Himawari-8 sector (over Japan), where data from the adjoining satellites are unable to fill the gap.
25. SSMIS data stopped flowing starting with data timestamps around 00 UTC on 8 September 2016. By late on 8 September FNMOC commenced backfilling the data.
26. Anomalies in the DMSF F-16 input data began appearing on 1 December 2013, 24 April 2015), and 1 May 2015 (183 GHz, 91V GHz, and 150 GHz, respectively), but then 183 and 150 GHz became useful again after 26 August 2015.
27. A malfunction in the AMSR2 data recorder aboard the GCOM-W1 satellite resulted in the loss of data from 16:34 UTC on 27 September 2017 to 15:39 on 28 September 2017.
28. At about 4:57 UTC on 9 October 2017 the NOAA-19 MHS precipitation estimates started displaying artifacts, which was eventually traced to the instrument going into safe mode without shutting down data delivery. Since this happened over the long Columbus Day weekend, it took until 01:47 UTC on 10 October 2017 to shut down the data stream, so the Early and Late IMERG have these (very obvious) artifacts for almost 24 hours. No reprocessing is planned. The sensor resumed operations at 17:31:08 UTC on 16 October 2017, but because the basis for the safe mode was unknown, GPM chose to monitor the data stream at first; GPROF-MHS estimates from NOAA-19 were restarted for the orbit segment beginning 10:32:49 UTC on 25 October 2017.
29. NOAA18 started showing degraded channel data at 18:00 UTC 20 October 2018 and through 21 October due to irregular function of the stepper motor.
30. Megha-Tropiques data transmissions became unusable on 15 December 2018. This arose due to a problem with the electronics the SAPHIR data during this time, and were followed by an issue with computing Earth Incidence Angle that was resolved on 27 June 2019. Data during this period are not available and presumed lost.

31. Communication issues led to an outage of AMSR2 data from 14:50:55 UTC 18 March 2019 to 05:42:23 UTC 19 March 2019.
32. Note: METOP-A experienced a satellite anomaly 17:00 30 July 30 – 12:00 9 August 2019. However, PPS received files for a subset of this time span, 09:57 5 August – 12:00 9 August 2019. Checks done by PPS and the Merged Algorithm Team show that most of the data in this period appear to be OK, except for orbits 066384 and 066385, which were replaced with Empty granules.
33. Beginning with 19 August 2019 at 13:00 UTC, noise in the form of random horizontal lines over the Eastern Pacific began appearing, with most of the problems appearing each following day in the 12:00, 12:30, and 13:00 half-hours, with at least one instance of 13:30. This was traced to noise in the GOES-W Tb's (as the result of the faulty sensor cooling system), which made it past the filtering at CPC, where the Merged Global 4-km IR dataset is assembled. The V06 Early and Late datasets contained these errors, which were dealt with in the Final computation and all V07 Runs.
34. Beginning with 8 September 2019 and becoming progressively worse, to almost complete failure by 9 September, Channel 2 on NOAA-19 began failing.
35. Anomalous SAPHIR longitude data were discovered starting in orbit 9904 and ending in 10001, which covers most of the period 12 September 2013 04:00 UTC to 19 September 2013 02:29 UTC. The bad data have been set to missing, although they are included in all Runs of V06B IMERG.
36. A problem with the cooling equipment on board GOES-17 (serving as GOES-W) causes some channels to become noisy at times in February, April, August, and October due to solar illumination. This affects the Tb's provided for the GOES-W sector around 12-13 UTC as noisy nearly-horizontal streaks during the specified times. CPC has implemented a correction scheme, but the IMERG group considers it unsatisfactory and continues to develop its own QC scheme. The date/times that appear affected in IMERG were found by quality-controlling a loop of the *IRprecipitation* field in the IMERG-E files, and they include:

August 2019

| | | | | | | | | |
|----|--|-------|-------|-------|-------|-------|-------|--|
| 19 | | | | 13:00 | | | | |
| 20 | | | 12:30 | 13:00 | | | | |
| 21 | | | 12:30 | 13:00 | 13:30 | | | |
| 22 | | | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 23 | | | 12:30 | 13:00 | 13:30 | | | |
| 24 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 25 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 26 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 27 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 28 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 29 | | 12:00 | | | 13:30 | 14:00 | 14:30 | |
| 30 | | 12:00 | | | 13:30 | 14:00 | 14:30 | |
| 31 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | | |

February 2020

| | | | | | | | | |
|----|--|--|-------|-------|-------|-------|--|--|
| 15 | | | | 13:00 | 13:30 | | | |
| 16 | | | | 13:00 | 13:30 | 14:00 | | |
| 17 | | | 12:30 | 13:00 | 13:30 | 14:00 | | |

| | | | | | | | | |
|----|-------|-------|-------|--------|--------|-------|-------|-------|
| 18 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | |
| 19 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | |
| 20 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | |
| 21 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | 15:00 |
| 22 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | 15:00 |
| 23 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | 15:00 |
| 24 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | 15:00 |
| 25 | 11:30 | 12:00 | 12:30 | 13:00* | 13:30* | 14:00 | 14:30 | 15:00 |
| 26 | | | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | |

* Widespread artifacts.

April 2020

| | | | | | | | | |
|----|--|-------|-------|-------|-------|--------|--|--|
| 12 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 13 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 14 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 15 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00* | | |
| 16 | | 12:00 | 12:30 | 13:00 | 13:30 | | | |
| 17 | | | 12:30 | 13:00 | 13:30 | | | |
| 18 | | | 12:30 | 13:00 | | | | |
| 19 | | | 12:30 | 13:00 | | | | |

* Just an isolated bad line segment or two

August 2020

| | | | | | | | | |
|----|--|--|-------|-------|-------|--|--|--|
| 25 | | | 12:30 | 13:00 | | | | |
| 26 | | | 12:30 | 13:00 | 13:30 | | | |
| 27 | | | 12:30 | 13:00 | 13:30 | | | |
| 28 | | | 12:30 | 13:00 | 13:30 | | | |
| 29 | | | 12:30 | 13:00 | 13:30 | | | |
| 30 | | | 12:30 | 13:00 | 13:30 | | | |

October 2020

| | | | | | | | | |
|----|-------|--|--|--|--|--|--|--|
| 11 | 11:30 | | | | | | | |
|----|-------|--|--|--|--|--|--|--|

February 2021

| | | | | | | | | |
|----|--|-------|-------|-------|-------|-------|-------|-------|
| 17 | | | | 13:00 | 13:30 | 14:00 | | |
| 18 | | | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 19 | | | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 20 | | | 12:30 | 13:00 | 13:30 | 14:00 | | |
| 21 | | | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | |
| 22 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | |
| 23 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | |
| 24 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | |
| 25 | | 12:00 | 12:30 | 13:00 | 13:30 | 14:00 | 14:30 | 15:00 |
| 26 | | | | 13:00 | 13:30 | 14:00 | | |

April 2021

| | | | | | | | | |
|----|-------|-------|-------|-------|-------|--|--|--|
| 17 | 11:30 | 12:00 | 12:30 | 13:00 | 13:30 | | | |
|----|-------|-------|-------|-------|-------|--|--|--|

October 2021

| | | | | | | | | |
|----|--|--|-------|-------|-------|--|--|--|
| 16 | | | 12:30 | 13:00 | 13:30 | | | |
|----|--|--|-------|-------|-------|--|--|--|

37. More discussion of the problem is available at <https://www.goes-r.gov/users/GOES-17-ABI-Performance.html>. NOAA announced that GOES-17 ABI commenced permanent operational implementation of the Mode 3 Cooling Timeline starting 26 February 2020 through 1 March 2020 in order to mitigate the number of saturated images resulting from the loop heat pipe (LHP) temperature regulation anomaly. The timeline occurs for 6 hours, centered on spacecraft midnight from 0600 UTC to 1200 UTC each day. The same timeline occurred seasonally in operations for four periods each year. The timeline operated in Mode 3 (“Flex Mode 3” uploaded as timelineID=3) where GOES-17 ABI generated a single full disk once per 15 minutes and generated one mesoscale domain sector each minute. Alternating MESO domains were collected one time each per two-minute period. The CONUS domain was not scanned during the timeline as those periods were used for cooling. This operation continued until GOES-18 took over the GOES-W slot in January 2023. GPM Core Observatory went into safe-hold ~1100 UTC 19 August 2020 over the Southern Indian Ocean due to a single electromagnetic event that affected the spacecraft processor. Science data were lost until operations resumed on 26 August 2020, starting ~1400 UTC for GMI and ~2300 UTC for DPR. IMERG continued production throughout the outage, but without the GMI data or accumulating calibration information from CORRA-G.
38. GOES-17 (GOES-W) went into safe mode around 05:37 UTC 22 July 22, 2021. ABI-L1B data were restored by 15:06 UTC 23 July 2021. For most half-hours, this caused small triangles of missing IR data in the Gulf of Alaska and the corresponding location in the Southern Hemisphere, as well as less reliable estimates on either side of that region due to higher zenith angles for the adjacent Himawari and GOES-Central data.
39. A mis-set error counter on the Core Observatory provoked a safe hold at 13:31:29 UTC on 23 March 2022, shutting down DPR and GMI observations. DPR and GMI were restored at 10:54 and 14:54 UTC, respectively on 24 March 2022. No data were recoverable from the outage period.
40. An outage at FNMOC prevented transmission of DMSP (SSMIS) from 15:51 UTC 23 March to 00:47 UTC 24 March 2022. Gaps in the data are: F16: 14:34:07-19:29:39, F17: 15:36:42-19:46:48, F18: 14:47:47-19:47:28.
41. Files for the 16:00 and 17:00 UTC 25 October CPC 4 km Merged Global IR failed to be downloaded. They were later retrieved, but all 4 half hours were not available in the IMERG Early and Late Runs.
42. Installation of the two-reaction wheel safehold patch on 18 April 2023 prevented data collection by the GMI and DPR from 13:13 to 21:16 UTC.
43. Starting in February 2023, the NRT MetOp 1C data stream has occasionally provided partial orbit granules that duplicate previous granules, but contain faulty data. PPS instituted QC procedures as of 11 July 2023 that can successfully screen out 12 of 16 known cases. Further work is necessary to capture the rest of the cases without mistakenly removing good granules.

.....

Computing **IMERG in the TRMM Era** required significant effort to ensure that there is reasonable homogeneity in the IMERG record given the constantly evolving satellite constellation.

This includes ensuring that the Level 1 (brightness temperature) and Level 2 (precipitation) data are as consistent as possible with the GMI across the constellation of passive microwave radiometers during the GPM era, and correspondingly TMI in the TRMM era. Then IMERG enforces calibration to the perceived “best” core satellite estimate, which is the combined radar-radiometer algorithm in the TRMM era (CORRA-T), as well as selective climatological calibration to a recognized standard (the Global Precipitation Climatology monthly satellite-gauge combined product) where the combined product is known to be low in high-latitude oceans. Despite this intercalibration, users should note that variations remain across sensors and between the TRMM and GPM eras. Specifically:

- Version 06 was the first time that CORRA and IMERG were computed for both the TRMM and GPM era. Lessons learned were applied in developing the V07 code.
- CORRA-T is based on fewer channels of microwave and radar data than for the CORRA-G, so one would expect the underlying calibrations to shift across the transition boundary (May/June 2014). This affects the histogram of precipitation rates more than the mean rates.
- CORRA-T only covers the latitude band 35°N-S, while it is 65°N-S for GPM. Thus, calibrations to CORRA-T outside 35°N-S are necessarily approximate, based on the (monthly climatological) “shape” of the CORRA-G calibration field, scaled to match the precipitation volume for each month of TRMM-era calibration for northern and southern hemisphere ocean and land separately in the region 33°N-S. See "TRMM-era Combined Radar-Radiometer Algorithm" for more details.
- The early TRMM era contains estimates from SSMI and AMSU-B sensors, which are less capable than the subsequent SSMIS and MHS sensors that were phased in starting in 2005. This implies that a shift in estimation skill is likely.
- As in the GPM era, residual “rippling” is visible in the animation of IMERG data in the TRMM era. In addition, the early TRMM era lacked microwave observations for several hours around the 0/12 and 04/16 UTC orbital times. As a result, there is more use of IR-based estimates in the first few years. IR estimates are generally of lower quality and potentially introduce a systematic regional offset in the timing of precipitation systems. [IR Tb’s tend to lag precipitation occurrence, but IMERG uses the PDIR algorithm, which considers other factors, such as spatial texture in the IR Tb’s.] Furthermore, the Japanese GMS observations only provide hourly data for certain periods of the day, and in the missing half-hours the adjacent METEOSAT and GOES-W IR values are used to the extent possible, but in some cases the necessary zenith-angle parallax corrections apparently yield very different IR values, leading to dithering between two precipitation estimate scenes in successive half hours. Before 2005 there are large gaps in the available microwave data, which makes the behavior of the IR estimates more important. In addition, the importance of the propagated PMW and IR estimates at a given location can fluctuate rapidly when large gaps between PMW overpasses occur.
- There is an extended data dropout in the presently available IR dataset for Japanese sector from late on 17 November 2005 to the middle of 22 March 2006 due to issues in NOAA coping with the format of the then-newly introduced MTSat-1. Consequently, during that period a small sector over Japan lacks all IR data, and depends on (potentially) long-interval propagated PMW estimates. The IR data in adjacent regions are entirely based on high-zenith-angle data from the METEOSAT to the west and GOES-W to the east.

.....

8. Missing Value Estimation and Codes

There is generally no effort to **estimate missing values** in the single-source input data sets.

.....

All products in IMERG use the **standard missing value** "-9999.9" or "-9999" for 4-byte floats or 2-byte integers, respectively. These values are carried in the metadata.

.....

All **completely missing fields** of a product result from completely absent input data for the given time. If the input file(s) is(are) available, the product file is created, even if it lacks any valid data.

.....

9. Quality and Confidence Estimates

The **accuracy** of the precipitation products can be broken into systematic departures from the true answer (bias) and random fluctuations about the true answer (sampling), as discussed in Huffman (1997). The former are the biggest problem for climatological averages, since they will not average out. However, for short averaging periods the low number of samples and/or algorithmic inaccuracies tend to present a more serious problem for individual microwave data sets. That is, the sampling is spotty enough that the collection of values over, say, one day may not be representative of the true distribution of precipitation over the day. For the IR estimates, the sampling is good, but the algorithm likely has substantial RMS error due to the weak physical connection between IR Tb's and precipitation.

Accordingly, the "random error" is assumed to be dominant, and estimates could be computed as discussed in Huffman (1997). Random error cannot be corrected.

The "bias error" is likely smaller, or at least contained. This is less true over land, where the lower-frequency microwave channels are not useful for precipitation estimation with our current state of knowledge. The state of the art for satellites at the monthly scale is reflected in the studies by Smith et al. (2006) and Adler et al. (2012). One study of the sub-monthly TMPA bias is provided by Tian et al. (2009).

Bolvin et al. (2020) provide a statistical study of the V06B Final Run over the tropical Pacific Ocean using gauge observations from 38 low-lying atolls archived in the Pacific Rainfall Database (PACRAIN) for the period June 2000 – January 2019. Over the analysis period, IMERG estimates are higher by 7.8% with a correlation of 0.68. Seasonally, DJF and JJA offer comparable statistics with biases and correlations of 7.7% and 0.72, and 6.2% and 0.69, respectively. IMERG and the atolls show the best agreement in MAM with a bias of 2.3%, while SON exhibits the worst performance at 14.3% bias. Scatterplots of IMERG versus atolls show that IMERG is consistently low(high) for light(intense) precipitation, with best agreement at intermediate rates. The highest precipitation rates in IMERG dictate the overall positive bias. This bias varies over time but does not exhibit any discernable trend or dependence on atoll population. The PACRAIN atoll gauges

are not wind-loss corrected, so application of an appropriate adjustment would increase the amounts (order 5%), bringing them more in line with the IMERG estimates. A corresponding study is underway for V07.

.....

The *precision* of the data depends on the data field. Some data fields have always been carried as 2-byte integers (sometimes scaled) to reflect their modest accuracy and save storage. Starting in Version 07, all the fields in units of precipitation and QI are rounded to what we believe is a reasonable extent. In particular, the precipitation rounding is what we need for low precipitation amounts. We consider the next 5 or 6 floating-point digits to not be meaningful, and this approach makes for more efficient compression.

| 3IMERGHH | |
|---------------------------------------|--------------------------|
| <i>MWobservationTime</i> | (INTEGER*2; no rounding) |
| <i>MWprecipitation</i> | 0.01 mm/hour |
| <i>MWprecipSource</i> | (INTEGER*2; no rounding) |
| <i>IRinfluence</i> | (INTEGER*2; no rounding) |
| <i>IRprecipitation</i> | 0.01 mm/hour |
| <i>precipitation</i> | 0.01 mm/hour |
| <i>precipitationQualityIndex</i> | 0.001 |
| <i>precipitationUncal</i> | 0.01 mm/hour |
| <i>probabilityLiquidPrecipitation</i> | (INTEGER*2; no rounding) |
| <i>randomError</i> | 0.01 mm/hour |
| 3IMERGM | |
| <i>gaugeRelativeWeighting</i> | (INTEGER*2; no rounding) |
| <i>precipitation</i> | 0.001 mm/hour |
| <i>precipitationQualityIndex</i> | 0.01 |
| <i>probabilityLiquidPrecipitation</i> | (INTEGER*2; no rounding) |
| <i>randomError</i> | 0.001 mm/hour |

.....

The IMERG *intercomparison results* are just starting to be developed. The time series of the global images shows good continuity in time and space. Overall, the analysis approach appears to improve over the Version 06 IMERG results. Some preliminary results for Version 07 are discussed in the release notes document, available as Appendix 2 in

<https://gpm.nasa.gov/resources/documents/imerg-v07-release-notes.pdf>

and this discussion will be expanded as additional results are computed.

Some validation studies will be conducted under the auspices of the International Precipitation Working Group (IPWG) in Australia, the continental U.S., western Europe, parts of South America, and Japan (Ebert et al. 2007). Respectively, the web sites for these activities are:

<http://cawcr.gov.au/projects/SatRainVal/validation-intercomparison.html> (inactive)
http://cics.umd.edu/ipwg/us_web.html

http://meso-a.gsfc.nasa.gov/ipwg/ipwgeu_home.html (inactive)
<http://cics.umd.edu/~dvila/web/SatRainVal/dailyval.html> (inactive)
http://www-ipwg.kugi.kyoto-u.ac.jp/IPWG/sat_val_Japan.html (through 2016)

.....

The *diurnal cycle* depicted in IMERG is affected by the particular mix of satellite sensors at any given time and place, in common with all other such satellite-based precipitation estimation systems. The diurnal cycle phase produced by IR estimates, which respond to cloud tops, is known to lag the phase of surface observations in many locations. The IR lag is highly variable, but frequently reported as up to 3 hours. The passive microwave estimates over land depend on the solid hydrometeors, which typically are confined to the upper reaches of clouds. This dependence also leads to lags compared to surface observations. Tan et al. (2018) reported that V06 IMERG lags are mostly positive and generally less than 20 min; there was not a strong peak in the precise lag value between either the metrics considered or the networks (see Fig. 2 in that paper). Over ocean the passive microwave estimates are driven by the full vertical profile of precipitation for conical-scan imagers, but primarily by solid hydrometeors for cross-track-scan sounders. Thus, there is a mix of typical lags, minimal for conical-scan imagers and presumably similar to the land values for cross-track-scan sounders (since they only consider ice hydrometeors). When you consider the regional variability in the lags of the individual sensor types and the variable mix of sensors contributing to the diurnal cycle during different epochs of satellite coverage, the general statement is that lags are more likely early in the dataset, before many passive microwave satellites were available, and are more likely over land. The GPM DPR and TRMM PR, being radars, give relatively unbiased estimates of the diurnal cycle, but their sampling is so sparse that it takes several years of data to allow a reasonable estimate to appear out of the sampling noise. See Kikuchi and Wang (2008), although their study with Version 6 TMPA has larger lags due to concentrating early in the record and has fewer passive microwave satellites than Version 7 TMPA has for the bulk of their study period. As well, variation between successive sensors leads to a non-physical pulsing superimposed on the real diurnal cycle. This issue has been successively improved in V06 and particularly V07. See Tan et al. (2019a, 2021) and Rajagopal et al. (2021) for more discussion.

.....

The *controlling factors on dataset performance* critically depend on the calibration approach, since the time series of the completed data tend to follow the time series of the calibrator, at least on the large scale. All of the global precipitation data sets have some calibrating data source, which is necessary to control bias differences between contributing satellites. Otherwise, shifts in the contributing set of satellites at any given time can cause unphysical shifts in the behavior of the precipitation estimates. However, this calibration plays a large role in determining the interannual variation that the various data sets display. Experience shows that datasets/regions with passive microwave calibration (oceans for GPCP monthly SG and the previous 3B42 after August 2014, and all regions for the previous 3B42RT and the IMERG Early and Late) tend to have similar interannual fluctuations, while datasets/regions with combined passive/active microwave calibration (oceans for the previous 3B42/43 and IMERG Runs) tend to show a variation in the tropical oceans that leads the passive microwave-calibrated datasets by 3-6 months. Climatological calibrations might change the mean bias or even the seasonal cycle, but they do not

change the interannual variations or long-term trends other than amplifying or diminishing the overall values.

Analyses of monthly surface gauge data add another layer of calibration over land in some datasets. The same multi-sensor precipitation research team at Goddard has major responsibility for the GPCP monthly SG combined product, the former 3B43 monthly product, and the IMERG Final Run monthly product. In each case the multi-satellite data are averaged to the monthly scale and combined with the GPCP monthly surface precipitation gauge analysis. Furthermore, the multi-satellite data are adjusted to the large-area mean of the gauge analysis, where available (mostly over land), and then combined with the gauge analysis using a simple inverse estimated-random-error variance weighting. In both data sets the gauge analysis has an important or dominant role in determining the final combined value for grid boxes in areas with "good" gauge coverage. [See Bolvin et al. (2009) for an example with GPCP.] Regions with poor gauge coverage, such as central Africa, have a higher weight on the satellite input that has been corrected to the large-area bias of the gauges. The oceans are mostly devoid of gauges and therefore mostly lack such gauge input.

In contrast, the short-interval (as opposed to monthly, above) GPCP is the V3.2 Daily, the previous short-interval TMPA was 3B42 (3-hourly), and the short-interval IMERG is the Final Run half-hourly datasets. In each case the short-interval data are adjusted with a simple, spatially varying ratio to force the multi-satellite estimates to approximately average up to the corresponding monthly satellite-gauge product, with controls on the ratios to prevent unphysical results. Thus, monthly averages of the short-interval data should be close to the values in the monthly datasets, which the developers consider more reliable than the short-interval datasets. In fact, compared to datasets that lack the adjustment to the monthly satellite-gauge estimates, the GPCP V3.2 Daily, previous 3B42, and IMERG Final half-hourly datasets tend to score better at timescales longer than a few days. This is presumably because the random error begins to cancel out as more samples are averaged together, leaving only the bias error. Of course, the short-interval datasets and regions that lack month-to-month surface gauge data input are more clearly driven by the behavior of the satellite input data.

.....

The **data set provenance** has several layers. Most importantly, the entire record for all three Runs is computed with Version 07 IMERG. [And, all earlier versions are considered obsolete, as are the 3B42, 3B43, and 3B42RT.] All of the Combined, DPR, and Level 1C microwave data used to feed IMERG V07 are computed with Version 07 algorithms. The time history of available microwave satellites is provided in Table 10, and the particular satellite contributing to a half-hour value is given by the *MWprecipSource* field. However, a detailed listing of the satellites contributing to each morphed value is not specified. The IR data are provided by the NOAA 4-km Global Merged IR dataset as computed by the version of the system that was operational at the time of observation. It is assumed that the IR dataset is homogeneous, but it is possible to determine which geosynchronous satellite provided a particular data value by accessing the IR dataset's data source field (separately available). The GPCP data used to calibrate CORRA are drawn from Version 3.2, specifically using averages for seasons in the years 2017-2020. The GPCP gauge analysis currently in use for the GPM era is the Version 2022 Monitoring Product. The TRMM era is covered by the Full Data Reanalysis (currently Version 2022), a retrospective

analysis that covers the period 1891-2020. [See “precipitation gauge analysis”.] The temperature, humidity, and surface pressure fields used to compute the *probabilityLiquidPrecipitation* field are provided by the then-operational JMA NWP forecast for the Early and Late Runs. ERA5 is used for the Final Run. See “probability of liquid phase field” for a discussion.

.....

10. Documentation

The **documentation custodian** is:

Dr. George J. Huffman
Code 612
NASA Goddard Space Flight Center
Greenbelt, MD 20771 USA
Phone: +1-301-614-6308
Fax: +1-301-614-5492
Internet: george.j.huffman@nasa.gov
GPM Web Page: <http://gpm.nasa.gov/>

.....

**Documentation revision history*:*

| | | |
|------------------|-------------|---|
| 4 December 2014 | Version 1 | by GJH, DTB, EJM |
| 15 January 2015 | Version 1.1 | by GJH, DTB, EJM; revisions for first Final Run release |
| 20 January 2015 | Version 1.2 | by GJH; re-release due to “missing” definition |
| 9 April 2015 | Version 1.3 | by GJH; more-uniform reference to Runs; retrospective processing in 2017; end of TRMM |
| 19 June 2015 | Version 1.4 | by DTB; reading HDF5 using the R language |
| 7 July 2015 | Version 1.5 | by GJH; revise field descriptions; Himawari 8 starts |
| 1 March 2017 | Version 2 | by GJH; Version 04 |
| 19 April 2017 | Version 2.1 | by GJH; update links |
| 9 November 2017 | Version 3 | by GJH; Version 05 |
| 10 November 2017 | Version 3.1 | by GJH; no GPROF-TMI; trimmed-down MHS, ATMS swaths |
| 17 November 2017 | Version 3.2 | by GJH; V05 re-released; shift V05A to V05B |
| 7 February 2018 | Version 3.3 | by GJH; V05B retrospective processing finished; geo-IR satellites by year |
| 13 March 2019 | Version 4 | by GJH, DTB, EJM, JT; Version 06 |
| 1 May 2019 | Version 4.1 | by GJH; updates to the start of Early and Late |
| 4 June 2019 | Version 4.2 | by GJH; updates for V06B |
| 8 October 2019 | Version 4.3 | by GJH; updates for Uncal and Cal being the same for Early and Late; upgrading ERA-I to ERA5 |
| 6 October 2020 | Version 4.4 | by GJH; correct input description for retrospective Early, Late; add GOES-W IR error tables; updates to web links |
| 13 July 2023 | Version 5 | by GJH, DTB, EJM, JT, RJ; Version 07 |

.....

**References*:*

- Adler, R.F., G. Gu, G.J. Huffman, 2012: Estimating Climatological Bias Errors for the Global Precipitation Climatology Project (GPCP). *J. Appl. Meteor. and Climatol.*, **51**, 84-99. doi:10.1175/JAMC-D-11-052.1
- Andersson, A., K. Fennig, C. Klepp, S. Bakan, H. Graßl, J. Schulz, 2010: The Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite Data - HOAPS-3. *Earth Syst. Sci. Data*, **2**, 215-234. doi:10.5194/essd-2-215-2010
- Beck, H.E., A.I.J.M. van Dijk, V. Levizzani, J. Schellekens, D.G. Miralles, B. Martens, A. de Roo, 2017: MSWEP: 3-Hourly 0.25° Global Gridded Precipitation (1979–2015) by Merging Gauge, Satellite, and Reanalysis Data. *Hydrol. Earth Syst. Sci.*, **21**, 589-615. doi:10.5194/hess-21-589-2017
- 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., A. Singh, Y. Song, M. Panahi, 2019: Assessing Gauge Undercatch Correction in Arctic Basins in Light of GRACE Observations. *Geophys. Res. Lett.*, **46**, 10 pp., doi:10.1029/2019GL084221
- Behrangi, A., G. Stephens, R.F. Adler, G.J. Huffman, B. Lambrigtsen, M. Lebsock, 2014: An Update on Oceanic Precipitation Rate and Its Zonal Distribution in Light of Advanced Observations from Space. *J. Climate*, **27**, 3957-3965. doi:10.1175/JCLI-D-13-00679.1
- 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**, 1843–1857. doi:10.1175/2009JAMC2147.1
- Bolvin, D.T., G.J. Huffman, E.J. Nelkin, J. Tan, 2021: Comparison of Monthly IMERG Precipitation Estimates with PACRAIN Atoll Observations. *J. Hydrometeorol.*, **22**, 1745-1753. doi:10.1175/JHM-D-20-0202.1
- Bosilovich, M. G., R. Lucchesi, M. Suarez, 2016: MERRA-2: File Specification. GMAO Office Note No. 9 (Version 1.1), 73 pp, available from <https://gmao.gsfc.nasa.gov/pubs/docs/Bosilovich785.pdf>
- Boukabara, S.A., K. Garrett, W.C. Chen, F. Iturbide-Sanchez, C. Grassotti, C. Kongoli, R.Y. Chen, Q.H. Liu, B. Yan, F.Z. Weng, R. Ferraro, T.J. Kleespies, H. Meng, 2011: MiRS: An All-Weather 1DVAR Satellite Data Assimilation and Retrieval System. *IEEE Trans. Geosci. Rem. Sens.*, **49**, 3249-3272. doi:10.1109/tgrs.2011.2158438
- Chiu, L.S., R. Chokngamwong, 2010: Microwave Emission Brightness Temperature Histograms (METH) Rain Rates for Climate Studies: Remote Sensing Systems SSM/I Version-6 Results. *J. Appl. Met. Clim.*, **49**, 115-123. doi:10.1175/2009JAMC2204.1
- Duan, Q., S. Sorooshian, V. K. Gupta, 1994: Optimal Use of the SCE-UA Global Optimization Method for Calibrating Watershed Models. *J. Hydrol.*, **158**, 265–284, doi:10.1016/0022-1694(94)90057-4
- Ebert, E.E., J.E. Janowiak, C. Kidd, 2007: Comparison of Near-Real-Time Precipitation Estimates from Satellite Observations and Numerical Models. *Bull. Amer. Meteor. Soc.*, **88**, 47-64. doi:10.1175/BAMS-88-1-47
- Fick, S.E. R.J. Hijmans, 2017: WorldClim 2: New 1-km Spatial Resolution Climate Surfaces for Global Land Areas. *Internat. J. Climatol.*, **37**, 4302-4315, doi:10.1002/joc.5086

- Fisher, R.A., 1915: Frequency Distribution of the Values of the Correlation Coefficient in Samples of an Indefinitely Large Population. *Biometrika*, Biometrika Trust, **10**, 507–521. doi:10.2307/2331838
- Fuchs, T., J. Rapp, F. Rubel, B. Rudolf, 2001: Correction of Synoptic Precipitation Observations due to Systematic Measuring Errors with Special Regard to Precipitation Phases. *Phys. Chem. Earth (B)*, **26**, 689-693. doi:10.1016/S1464-1909(01)00070-3
- Gelaro, R., W. McCarty, M.J. Suárez, R. Todling, A. Molod, L. Takacs, C.A. Randles, A. Darmenov, M.G. Bosilovich, R. Reichle, K. Wargan, L. Coy, R. Cullather, C. Draper, S. Akella, V. Buchard, A. Conaty, A.M. da Silva, W. Gu, G.-K. Kim, R. Koster, R. Lucchesi, D. Merkova, J.E. Nielsen, G. Partyka, S. Pawson, W. Putman, M. Rienecker, S.D. Schubert, M. Sienkiewicz, B. Zhao, 2017: The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *J. Climate.*, **30**, 5419–5454. doi:10.1175/JCLI-D-16-0758.1
- Hersbach, H., B. Bell, P. Berrisford, S. Hirahara, A. Horányi, J. Muñoz-Sabater, J. Nicolas, C. Peubey, R. Radu, D. Schepers, A. Simmons, C. Soci, S. Abdalla, X. Abellan, G. Balsamo, P. Bechtold, G. Biavati, J. Bidlot, M. Bonavita, G. De Chiara, P. Dahlgren, D. Dee, M. Diamantakis, R. Dragani, J. Flemming, R. Forbes, M. Fuentes, A. Geer, L. Haimberger, S. Healy, R.J. Hogan, E. Hólm, M. Janisková, S. Keeley, P. Laloyaux, P. Lopez, C. Lupu, G. Radnoti, P. de Rosnay, I. Rozum, F. Vamborg, S. Villaume, J.-N. Thépaut, 2020: The ERA5 Global Reanalysis. *Quart. J. Roy. Meteor. Soc.*, **146** (Part A), 1999-2049. doi:10.1002/qj.3803
- Hollinger, J., R. Lo, G. Poe, J. Pierce, 1987: *Special Sensor Microwave/Imager User's Guide*, Naval Res. Lab., Washington, DC, 182 pp., ftp://rain.atmos.colostate.edu/FCDR/doc/SSMI_general/SSMI_User'sGuide_searchable.pdf
- Hollinger, J.P., J.L. Pierce, G.A. Poe, 1990: SSM/I Instrument Evaluation. *IEEE Trans. Geosci. Remote Sens.*, **28**, 781-790. doi:10.1109/36.58964
- Hong, K.L. Hsu, S. Sorooshian, X. Gao, 2004: Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System. *J. Appl. Meteor.*, **43**, 1834–1852. doi:10.1175/JAM2173.1
- Hsu, K.-L., X.Gao, S. Sorooshian, H.V. Gupta, 1997: Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks. *J. Appl. Meteor.*, **36**, 1176-1190, doi:10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2
- Huffman, G.J., 1997: Estimates of Root-Mean-Square Random Error Contained in Finite Sets of Estimated Precipitation. *J. Appl. Meteor.*, **36**, 1191-1201. doi:10.1175/1520-0450(1997)036<1191:EORMSR>2.0.CO;2
- Huffman, G.J., R.F. Adler, D.T. Bolvin, G. Gu, E.J. Nelkin, K.P. Bowman, Y. Hong, E.F. Stocker, D.B. Wolff, 2007a: The TRMM Multi-satellite Precipitation Analysis: Quasi-Global, Multi-Year, Combined-Sensor Precipitation Estimates at Fine Scales. *J. Hydrometeor.*, **8**, 38-55. doi:10.1175/JHM560.1
- Huffman, G.J., R.F. Adler, D.T. Bolvin, E.J. Nelkin, 2010: The TRMM Multi-satellite Precipitation Analysis (TMPA). Chapter 1 in *Satellite Rainfall Applications for Surface Hydrology*, F. Hossain and M. Gebremichael, Eds. Springer Verlag, ISBN: 978-90-481-2914-0, 3-22. doi:10.1007/978-90-481-2915-7_1
- Huffman, G.J., R.F. Adler, S. Curtis, D.T. Bolvin, E.J. Nelkin, 2007b: Global Rainfall Analyses at Monthly and 3-hr Time Scales. Chapter 23 of *Measuring Precipitation from Space: EURAINSAT and the Future*, V. Levizzani, P. Bauer, and F.J. Turk, Eds., Springer Verlag

- (Kluwer Academic Pub. B.V.), Dordrecht, The Netherlands, 291-306.
<https://link.springer.com/content/pdf/10.1007/978-1-4020-5835-6.pdf> [Invited paper]
- Huffman, G.J., D.T. Bolvin, 2015: Transition of 3B42/3B43 Research Product from Monthly to Climatological Calibration/Adjustment, 11 pp.,
https://gpm.nasa.gov/sites/default/files/document_files/3B42_3B43_TMPA_restart.pdf
- Huffman, G.J., D.T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, C. Kidd, E.J. Nelkin, S. Sorooshian, E.F. Stocker, J. Tan, D.B. Wolff, P. Xie, 2020: Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) mission (IMERG). Chapter 19 in *Adv. Global Change Res., Vol. 67, Satellite Precipitation Measurement*, V. Levizzani, C. Kidd, D. Kirschbaum, C. Kummerow, K. Nakamura, F.J. Turk (Ed.), Springer Nature, Dordrecht, ISBN 978-3-030-24567-2 / 978-3-030-24568-9 (eBook), 343-353. doi:/10.1007/978-3-030-24568-9_19
- 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 5.2 for the NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (I-MERG). GPM Project, Greenbelt, MD, 38 pp.
https://gpm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf
- Iguchi, T., S. Seto, R. Meneghini, N. Yoshida, J. Awaka, T. Kubota, 2010: GPM/DPR Level-2 Algorithm Theoretical Basis Document. PPS, NASA/GSFC, 72 pp.
http://pps.gsfc.nasa.gov/Documents/ATBD_GPM_DPR_n3_dec15.pdf
- Janowiak, J.E., P.A. Arkin, 1991: Rainfall Variations in the Tropics during 1986-1989. *J. Geophys. Res.*, **96**, 3359-3373. doi:10.1029/90JD01856
- Joyce, R.J., J.E. Janowiak, G.J. Huffman, 2001: Latitudinal and Seasonal Dependent Zenith Angle Corrections for Geostationary Satellite IR Brightness Temperatures. *J. Appl. Meteor.*, **40**, 689-703. doi:10.1175/1520-0450(2001)040<0689:LASDZA>2.0.CO;2
- Joyce, R.J., P. Xie, J.E. Janowiak, 2011: Kalman Filter Based CMORPH. *J. Hydrometeor.*, **12**, 1547-1563. doi:10.1175/JHM-D-11-022.1
- Kikuchi, K., B. Wang, 2008: Diurnal Precipitation Regimes in the Global Tropics. *J. Climate*, **21**, 2680-2696. doi:10.1175/2007JCLI2051.1
- Kubota T., S. Shige, H. Hashizume, K. Aonashi, N. Takahashi, S. Seto, M. Hirose, Y.N. Takayabu, K. Nakagawa, K. Iwanami, T. Ushio, M. Kachi, K. Okamoto, 2007: Global Precipitation Map Using Satellite-Borne Microwave Radiometers by the GSMaP Project: Production and Validation. *IEEE Trans. Geosci. Remote Sens.*, **45**, 2259-2275. doi:10.1109/TGRS.2007.895337
- 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. doi:10.1109/36.536538
- Kummerow, C., W. Barnes, T. Kozu, J. Shiue, J. Simpson, 1998: The Tropical Rainfall Measuring Mission (TRMM) Sensor Package. *J. Atmos. Oceanic Tech.*, **15**, 809-817. doi:10.1175/1520-0426(1998)015<0809:TTRMMT>2.0.CO;2
- Legates, D.R., 1987: A Climatology of Global Precipitation. *Pub. in Climatol.*, **40**, U. of Delaware, 85 pp.

- Legates, D.R., C.J. Willmott, 1990: Mean Seasonal And Spatial Variability In Gauge-Corrected, Global Precipitation. *Internat. J. Climatol.*, **10**, 111-127. doi:10.1002/joc.3370100202
- Lucchesi, R., 2017: File Specification for GEOS-5 FP. GMAO Office Note No. 4 (Version 1.1), 61 pp. https://gmao.gsfc.nasa.gov/products/documents/GEOS_5_FP_File_Specification_ON4v1_1.pdf.
- Maggione, V., M.R.P. Sapiano, R.F. Adler, Y. Tian, G.J. Huffman, 2014: An Error Model for Uncertainty Quantification in High-Time Resolution Precipitation Products. *J. Hydrometeorol.*, **15**. doi:10.1175/JHM-D-13-0112.1, 1274-1292
- Mattioli, V., C. Accadia, J. Ackermann, S. Di Michele, H. Imke, P. Schlüssel, P. Colucci, A. Canestri, 2019: The EUMETSAT Polar System - Second Generation (EPS-SG) Passive Microwave and Sub-mm Wave Missions, 2019 Photonics & Electromagnetics Research Symposium - Spring (PIERS-Spring), Rome, Italy, 2019, pp. 3926-3933, doi:10.1109/PIERS-Spring46901.2019.9017822
- Nguyen, P., M. Ombadi, V.A. Gorooh, E.J., Shearer, M. Sadeghi, S. Sorooshian, K. Hsu, D. Bolvin, M.F. Ralph, 2020a: PERSIANN Dynamic Infrared–Rain Rate (PDIR-Now): A Near-Real-Time, Quasi-Global Satellite Precipitation Dataset. *J. Hydrometeorol.*, **21**, 2893–2906, doi:10.1175/JHM-D-20-0177.1
- Nguyen, P., E.J. Shearer, M. Ombadi, V.A. Gorooh, K. Hsu, S. Sorooshian, W.S. Logan, M. Ralph, 2020b: PERSIANN Dynamic Infrared–Rain Rate Model (PDIR) for High-Resolution, Real-Time Satellite Precipitation Estimation. *Bull. Amer. Meteor. Soc.*, **101**, E286-E302, doi:10.1175/BAMS-D-19-0118.1
- Northrup Grumman Electronic Systems, Azusa, CA, Report 12621, 69 pp. 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/FCDR/doc/SSMIS_general/Algorithm_and_Data_User_Manual_For_SSMIS_Jul02.pdf
- O, S., U. Foelsche, G. Kirchengast, J. Fuchsberger, J. Tan, W.A. Petersen, 2017: Evaluation of GPM IMERG Early, Late, and Final Rainfall Estimates Using WegenerNet Gauge Data in Southeastern Austria. *Hydrol. and Earth Sys. Sci.*, **21**, 6559-6572. doi:10.5194/hess-21-6559-2017
- 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. Meteorol.*, **38**, 633-664. doi:10.1175/1520-0450(1999)038<0633:ALHDIT>2.0.CO;2
- Olson, W.S., H. Masunaga, the GPM Combined Radar-Radiometer Algorithm Team, 2016: GPM Combined Radar-Radiometer Precipitation Algorithm Theoretical Basis Document (Version 4). PPS, NASA/GSFC, 29 pp. https://gpm.nasa.gov/sites/default/files/document_files/Combined_algorithm_ATBD.V04.rev_.pdf
- Parsons, M.A., R. Duerr, J.-B. Minster, 2010: Data Citation and Peer Review. *EOS*, **91**(34), 297-298. doi:10.1029/2010EO340001
- Passive Microwave Algorithm Team Facility, 2017: GPROF (Level 2) Algorithm Theoretical Basis Document (ATBD), Version 1.0. PPS, NASA/GSFC, 63 pp.

- https://gpm.nasa.gov/sites/default/files/document_files/ATBD_GPM_GPROF_June1_2017.pdf
- PMM, 2014: Goodbye to TRMM, First Rain Radar in Space. <https://pmm.gsfc.nasa.gov/articles/goodbye-trmm-first-rain-radar-space>
- PPS, GPM Intercalibration (X-CAL) Working Group, 2016: NASA GPM Level 1C Algorithms Algorithm Theoretical Basis Document (ATBD). PPS, NASA/GSFC, 55 pp. https://gpm.nasa.gov/sites/default/files/document_files/L1C_ATBD_v1.6_V04.pdf
- Rajagopal, M., E. Zipser, G.J. Huffman, J. Russell, and J. Tan, 2021: Comparisons of IMERG Version 06 Precipitation At and Between Passive Microwave Overpasses in the Tropics. *J. Hydrometeor.*, **22**, 2117-2130. doi:10.1175/JHM-D-20-0226.1
- Romanov, P., 2016: Global 4km Multisensor Automated Snow/Ice Map (GMASI) Algorithm Theoretical Basis Document. NOAA NESDIS Cent. for Sat. Appl. and Res., 61 pp. http://www.star.nesdis.noaa.gov/smcd/emb/snow/documents/Global_Auto_Snow-Ice_4km_ATBD.pdf
- Schneider, U., A. Becker, P. Finger, A. Meyer-Christoffer, M. Ziese, B. Rudolf, 2014: GPCP'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**, 15-40. doi:10.1007/s00704-013-0860-x . <http://link.springer.com/article/10.1007%2Fs00704-013-0860-x>
- Schneider, U., P. Finger, A. Meyer-Christoffer, M. Ziese, A. Becker, 2018: Global Precipitation Analysis Products of the GPCP. GPCP Internet Publication, DWD, 17 pp. ftp://ftp-anon.dwd.de/pub/data/gpcp/PDF/GPCP_intro_products_2008.pdf.
- Sevruk, B., 1989: Reliability of Precipitation Measurements. *Proc. WMO/IAHS/ETH Workshop on Precipitation Measurements*, St. Moritz, Switzerland, World Meteor. Org., 13–19.
- Sims, E.J., 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
- Smith, T.M., P.A. Arkin, J.J. Bates, G.J. Huffman, 2006: Estimating Bias of Satellite-Based Precipitation Estimates. *J. Hydrometeor.*, **7**, 841-856. doi:/10.1175/JHM524.1
- Sorooshian, S., K.-L. Hsu, X. Gao, H.V. Gupta, B. Imam, D. Braithwaite 2000: Evaluation of PERSIANN System Satellite-Based Estimates of Tropical Rainfall. *Bull. Amer. Meteor. Soc.*, **81**, 2035-2046, doi:10.1175/1520-0477(2000)081<2035:EOPSSE>2.3.CO;2
- Spencer, R.W., 1993: Global Oceanic Precipitation from the MSU During 1979-92 and Comparisons to Other Climatologies. *J. Climate*, **6**, 1301-1326. doi:10.1175/1520-0442(1993)006<1301:GOPFTM>2.0.CO;2
- Tan, J., G.J. Huffman, D.T. Bolvin, E.J. Nelkin, 2019a: Diurnal Cycle of IMERG V06 Precipitation. *Geophys. Res. Lett.*, **46**, 13,584-13,592. doi:10.1029/2019GL085395
- Tan, J., G.J. Huffman, D.T. Bolvin, E.J. Nelkin, 2019b: IMERG V06: Changes to the Morphing Algorithm. *J. Atmos. Ocean. Technol.*, **36**(12), 2471-2482. doi:10.1175/JTECH-D-19-0114.1
- Tan, J., G.J. Huffman, D.T. Bolvin, E.J. Nelkin, M. Rajagopal, 2021: SHARPEN: A Scheme to Restore the Distribution of Averaged Precipitation Fields. *J. Hydrometeor.*, **22**, 2105-2216. doi:10.1175/JHM-D-0225.1
- Tan, J., W.A. Petersen, G. Kirchengast, D.C. Goodrich, D.B. Wolff, 2018: Evaluation of Global Precipitation Measurement Rainfall Estimates against Three Dense Gauge Networks. *J. Hydromet.*, **19**, 517-532. doi:10.1175/JHM-D-17-0174.1

- Tian, Y., C.D. Peters-Lidard, 2007: Systematic Anomalies over Inland Water Bodies in Satellite-Based Precipitation Estimates. *Geophys. Res. Lett.*, **34**. doi:10.1029/2007GL030787
- Tian, Y., C. Peters-Lidard, J. Eylander, R. Joyce, G. Huffman, R. Adler, K.-L. Hsu, F. J. Turk, M. Garcia, J. Zeng, 2009: Component Analysis of Errors in Satellite-Based Precipitation Estimates. *J. Geophys. Res., - Atmos.*, **114**, D22104. doi:10.1029/2009JD011949
- Villarini, G., W.F. Krajewski, 2007: Evaluation of the Research-Version TMPA Three-Hourly 0.25°x0.25° Rainfall Estimates over Oklahoma. *Geophys. Res. Lett.*, **34**. doi:10.1029/2006GL029147
- 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
- Wentz, F.J., and R.W. Spencer, 1998: SSMI Rain Retrievals within a Unified All-Weather Ocean Algorithm. *J. Atmos. Sci.*, **55**, 1613-1627. doi:10.1175/1520-0469(1998)055<1613:SIRRWA>2.0.CO;2
- Wilheit, T., A. Chang, L. Chiu, 1991: Retrieval of Monthly Rainfall Indices from Microwave Radiometric Measurements Using Probability Distribution Function. *J. Atmos. Ocean. Tech.*, **8**, 118-136. doi:10.1175/1520-0426(1991)008<0118:ROMRIF>2.0.CO;2
- 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

.....

**Web resources*:*

2BCMB ATBD:

https://gpm.nasa.gov/sites/default/files/document_files/Combined_algorithm_ATBD.V04.rev_.pdf

AMS data citation policy: <https://www.ametsoc.org/index.cfm/ams/publications/ethical-guidelines-and-ams-policies/data-policy-and-guidelines/>

AMSR-E instrument: <https://global.jaxa.jp/projects/sat/aqua/topics.html>

AMSR2 instrument: https://suzaku.eorc.jaxa.jp/GCOM_W/w_amsr2/whats_amsr2.html

AMSU-B instrument: MHS is detailed in Section 3.9 and AMSU-B in Section 3.4 in the NOAA KLM User's Guide (August 2014 revision):

https://www.star.nesdis.noaa.gov/mirs/documents/0.0_NOAA_KLM_Users_Guide.pdf

ATMS: <https://www.star.nesdis.noaa.gov/mirs/snppatms.php> or

<https://www.star.nesdis.noaa.gov/jps/ATMS.php>

DPR anomalies: https://gpmweb2https.pps.eosdis.nasa.gov/tsdis/AB/docs/gpm_anomalous.html

DPR release notes:

https://pps.gsfc.nasa.gov/Documents/Caveats_DPRL2_productV05B_fin3.pdf

FAQ: <https://disc.gsfc.nasa.gov/information/faqs?keyword=precipitation&page=1>

GES DISC home: <https://disc.gsfc.nasa.gov>

Giovanni: <https://giovanni.gsfc.nasa.gov/giovanni/>

GMI channel and Level 1B processing:

https://gpm.nasa.gov/sites/default/files/document_files/GPMGMIL1BATBDV2.3.pdf

GMI anomalies: https://gpmweb2https.pps.eosdis.nasa.gov/tsdis/AB/docs/gpm_anomalous.html

GMI scanning basics: <https://www.star.nesdis.noaa.gov/mirs/gpmgmi.php>

GOES-17 anomalies: <https://www.goes-r.gov/users/GOES-17-ABI-Performance.html>

GPM / TRMM data access: <https://gpm.nasa.gov/data/directory>

IMERG Early Run Homepage: <https://gpm.nasa.gov/data/directory> , Level 3 tab, Early Run
 IMERG Late Run Homepage: <https://gpm.nasa.gov/data/directory> , Level 3 tab, Late Run
 IMERG Final Run Homepage: <https://gpm.nasa.gov/data/directory> , Level 3 tab, Final Run
 GPM data file naming:
https://gpm.nasa.gov/sites/default/files/document_files/FileNamingConventionForPrecipitationProductsForGPMMissionV1.4.pdf
 GPM home: <https://gpm.nasa.gov/>
 GPROF2017 ATBD:
https://gpm.nasa.gov/sites/default/files/document_files/ATBD_GPM_GPROF_June1_2017.pdf
 IMERG ATBD: <https://gpm.nasa.gov/resources/documents/imerg-v07-atbd.pdf> .
 IMERG HDF5 file layout:
<https://gpmweb2https.pps.eosdis.nasa.gov/pub/GPMfilespec/filespec.GPM.pdf>
 IMERG HDF5 metadata:
<https://gpmweb2https.pps.eosdis.nasa.gov/pub/GPMfilespec/filespecMeta.GPM.pdf>
 IMERG Morphing: <https://gpm.nasa.gov/sites/default/files/2020-06/MorphingInV06IMERG.pdf>
 IMERG Quality Index: appendix in <https://gpm.nasa.gov/resources/documents/imerg-v07-atbd.pdf>
 IMERG Release Notes: <https://gpm.nasa.gov/resources/documents/imerg-v07-release-notes.pdf>
 IPWG data set tables: <http://www.isac.cnr.it/~ipwg/data/datasets.html>
 IPWG Validation for Australia: <http://cawcr.gov.au/projects/SatRainVal/validation-intercomparison.html> (inactive)
 IPWG Validation for U.S.: http://cics.umd.edu/ipwg/us_web.html (inactive)
 IPWG Validation for western Europe: http://meso-a.gsfc.nasa.gov/ipwg/ipwgeu_home.html (inactive)
 IPWG Validation for South America: <http://cics.umd.edu/~dvila/web/SatRainVal/dailyval.html> (inactive)
 IPWG Validation for Japan: http://www-ipwg.kugi.kyoto-u.ac.jp/IPWG/sat_val_Japan.html
 MHS instrument: MHS is detailed in Section 3.9 and AMSU-B in Section 3.4 in the NOAA KLM User's Guide (August 2014 revision):
https://www.star.nesdis.noaa.gov/mirs/documents/0.0_NOAA_KLM_Users_Guide.pdf
 PPS near-real-time data archive: <https://jsimpsonhttps.pps.eosdis.nasa.gov/>
 PPS post-real-time data archive: <https://arthurhouhttps.pps.eosdis.nasa.gov/>
 PPS home: <https://pps.gsfc.nasa.gov>
 PPS toolkit:<https://gpmweb2https.pps.eosdis.nasa.gov/pub/PPStoolkit/GPM/>
 PR anomalies:
https://gpmweb2https.pps.eosdis.nasa.gov/tsdis/AB/docs/patrner_anomalous.html.
 PR Manual: https://pps.gsfc.nasa.gov/Documents/PR_Manual_JAXA_V6.pdf
 R language: <https://cran.r-project.org/web/packages/hdf5r/index.html>
 Registration for the GES DISC data access: <https://urs.earthdata.nasa.gov/users/new>
 Registration for PPS data access: <https://registration.pps.eosdis.nasa.gov/registration/>
 Release notes for IMERG Version 07: <https://gpm.nasa.gov/resources/documents/imerg-v07-release-notes.pdf>
 SAPHIR: https://megha-tropiques.cnes.fr/en/MEGHAT/lien2_sat.htm
 Satellite overpass times graphic: <https://gpm.nasa.gov/resources/images/gpm-constellation-overpass-times>

SSMIS description: <https://en.wikipedia.org/wiki/SSMIS>
 STORM home: <https://storm.pps.eosdis.nasa.gov/storm/>
 STORM User Guide: <https://storm.pps.eosdis.nasa.gov/storm/STORMUserGuide.pdf>
 TMI and constellation sensor anomalies:
https://gpmweb2https.pps.eosdis.nasa.gov/tsdis/AB/docs/partner_anomalous.html
 THOR download site: <https://pps.gsfc.nasa.gov/thorrelease.html>
 The Transition in Multi-Satellite Products from TRMM to GPM (TMPA to IMERG)
https://gpm.nasa.gov/sites/default/files/2020-01/TMPA-to-IMERG_transition_201002.pdf
 TRMM data access: <https://disc.gsfc.nasa.gov/datasets?keywords=trmm&page=1>
 TRMM home page: <https://gpm.nasa.gov/missions/trmm>

.....
 Acronyms and Jargon:

| | |
|------------|---|
| 2BCMB | GPM Combined Instrument data set (also called CORRA or GCI) |
| 2B31 | TRMM Combined Instrument data set (also called TCI) |
| 2AGPROFGMI | GPM GPROF2021 precipitation estimates using GMI |
| 3B42 | production 3-hourly TMPA data set |
| 3B42RT | real-time 3-hourly TMPA data set |
| 3B43 | production monthly TMPA data set |
| 3IMERGHH | half-hourly GPM IMERG data set (Final Run; the Early and Late have an appended “e” and “l”, respectively) |
| 3IMERGM | monthly GPM IMERG data set (Final Run only) |
| ASCII | American Standard Code for Information Interchange (i.e., text) |
| AIRS | Atmospheric Infrared Sounder |
| AMSR2 | Advanced Microwave Scanning Radiometer model 2 |
| AMSR-E | Advanced Microwave Scanning Radiometer for Earth Observing System |
| AMSU | Advanced Microwave Sounding Unit |
| Aqua | satellite (Latin for “water”) |
| ATBD | Algorithm Theoretical Basis Document |
| ATMS | Advanced Technology Microwave Sounder |
| C | programming language |
| CEOS | Committee on Earth Observation Satellites |
| CMORPH | CPC MORPHing algorithm |
| CMORPH-KF | Kalman Filter version of CMORPH |
| CNES | Centre National d'Etudes Spatiales |
| CONUS | Contiguous U.S. |
| CORRA | GPM Combined Radar-Radiometer Algorithm |
| CORRA-G | GPM Combined Radar-Radiometer Algorithm-GPM era |
| CORRA-T | GPM Combined Radar-Radiometer Algorithm-TRMM era |
| CPC | Climate Prediction Center |
| CrIS | Cross-track Infrared Sounder |
| CSI | Conically-Scanning Imager |
| CTSS | Cross-Track-Scanning Sounder |
| dBZ | decibels of reflectivity factor |
| DMSP | Defense Meteorological Satellite Program (U.S.) |
| DOI | Digital Object Identifier |

| | |
|----------|---|
| DWD | Deutscher Wetterdienst |
| DPR | (GPM) Dual-frequency Precipitation Radar |
| ERA5 | ECMWF Reanalysis Version 5 |
| ESSIC | (University of Maryland College Park) Earth System Science Interdisciplinary Center |
| EUMETSAT | EUropean organization for the exploitation of Meteorological Satellites |
| EV | Event Viewer (in STORM) |
| FAQ | Frequently Asked Questions |
| FORTTRAN | programming language |
| FTP | File Transfer Protocol |
| GCI | GPM Combined Instrument algorithm (also called CORRA or 2BCMB) |
| GEOS FP | Goddard Earth Observing System model Version 5 [GEOS-5] Forward Processing [FP] |
| GeoTIFF | Georeferenced Tagged Image File Format |
| GES DISC | Goddard Earth Sciences Data and Information Services Center |
| GHz | Gigahertz |
| GIS | Geographical Information System |
| GMI | GPM Microwave Imager |
| GMS | Geosynchronous Meteorological Satellite ((EUMETSAT) |
| GOES | Geosynchronous Operational Environmental Satellite (U.S.) |
| GPCC | Global Precipitation Climatology Centre |
| GPM | Global Precipitation Measurement mission (U.S.-Japan) |
| GPROF | Goddard Profiling algorithm |
| GSFC | Goddard Space Flight Center (NASA) |
| GSMaP | Global Satellite Map of Precipitation (JAXA) |
| GTS | Global Telecommunications System |
| GV | Ground Validation |
| HDF | Hierarchical Data Format |
| Himawari | Japanese series of meteorological geo-satellites |
| HOAPS | Hamburg Ocean Atmosphere Parameters and Fluxes from Satellite data set |
| HQ | High Quality (passive microwave) precipitation |
| ICARE | Interactions Clouds Aerosols Radiations Etc. (or in more exact English, Cloud-Aerosol-Water-Radiation Interactions) (Japan) |
| IDL | Interactive Data Language |
| IEEE | Institute of Electrical and Electronics Engineers |
| IMERG | Integrated Multi-satellitE Retrievals for GPM (NASA) |
| IP | Initial Processing |
| IPWG | International Precipitation Working Group |
| IR | Infrared |
| ISRO | Indian Space Research Organisation |
| JAXA | Japan Aerospace Exploration Agency |
| JMA | Japan Meteorological Agency |
| Ka | microwave band; 26.5–40 GHz |
| KF | Kalman Filter |
| Ku | microwave band; 12–18 GHz |
| lat/lon | latitude/longitude |

| | |
|--------------|---|
| leo | Low Earth orbit |
| MB | megabytes |
| MERRA-2 | Modern-Era Retrospective Analysis for Research and Applications, Version 2 |
| Meteosat | Meteorological Satellite (EUMETSAT) |
| MetOp | Operational Meteorological satellite (EUMETSAT) |
| MHS | Microwave Humidity Sounder |
| MiRS | Microwave Integrated Retrieval System |
| MOSDAC | Meteorological and Oceanographic Satellite Data Archival Centre |
| MSU | Microwave Sounding Unit |
| MTSat | Multifunctional Transport Satellite (Japan) |
| MWI | Microwave Imager (EUMETSAT) |
| NASA | National Aeronautics and Space Administration NESDIS National Environmental Satellite Data and Information Service |
| NRL/FNMOC | Naval Research Laboratory / Fleet Numerical Meteorological and Oceanographic Center |
| NOAA | National Oceanic and Atmospheric Administration; also a leo-satellite series |
| NWP | Numerical Weather Prediction |
| NWS | National Weather Service |
| PDF | Probability Density Function |
| PDIR | PERSIANN Dynamic Infrared–Rain Rate |
| PERSIANN | Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks |
| PERSIANN-CCS | PERSIANN with Cloud Classification System |
| PMM | Precipitation Measurement Missions |
| PMW | Passive Microwave |
| PPS | Precipitation Processing System |
| PR | (TRMM) Precipitation Radar |
| Q1h | precipitationQualityIndex data field for half-hourly files |
| Q1m | precipitationQualityIndex data field for monthly files |
| RMS | Root Mean Square |
| RP | Retrospective Processing |
| SAPHIR | Sounder for Atmospheric Profiling of Humidity in the Intertropics by Radiometry |
| SG | Satellite-Gauge combined data set |
| SHARPEN | Scheme for Histogram Adjustment with Ranked Precipitation Estimates in the Neighborhood |
| SNPP | Suomi National Polar Partnership satellite (U.S.) |
| SSMI | Special Sensor Microwave/Imager |
| SSMIS | Special Sensor Microwave Imager-Sounder |
| SSM/T2 | Special Sensor Microwave/Temperature 2 |
| STORM | Science Team Online Research Module |
| Ta | Antenna Temperature |
| Tb | Brightness Temperature |
| TCI | TRMM Combined Instrument algorithm (product 2B31) in previous TRMM processing |
| THOR | Tool for High-resolution Observation Review |

| | |
|---------|--|
| TKIO | Input/Output Toolkit |
| TMI | TRMM Microwave Imager |
| TMPA | TRMM Multi-satellite Precipitation Analysis |
| TMPA-RT | Real-Time TMPA |
| TRMM | Tropical Rainfall Measuring Mission (U.S.-Japan) |
| URL | Universal Resource Locator (usually the web address) |
| UTC | Universal Coordinated Time (same as GMT, Z) |
| Vxx | Version number xx; for previous TRMM products, xx was one digit, for GPM it's two, and for GPCP it's two separated by a period |
| WMS | Web Map Service |
| X-Cal | Intersatellite Calibration (working group) |
| Zo | Surface reflectivity |

.....

Frequently Asked Questions (FAQ) lists are assembled and maintained by the Goddard Earth Science Data and Information Services Center (GESDISC) and Precipitation Measurement Missions (PMM). They are respectively posted at:

<https://disc.gsfc.nasa.gov/information/faqs?keyword=precipitation&page=1>
<https://gpm.nasa.gov/resources/faq>

.....

The **GPM data access pages** provide a one-stop shop for data and documentation at NASA:

<https://gpm.nasa.gov/data/directory>

.....

The **TRMM data access pages** are integrated into the GPM data access pages for datasets using GPM algorithms. Access pages for the previous TRMM products and documentation are available at:

<https://disc.gsfc.nasa.gov/datasets?keywords=trmm&page=1> .

.....

11. News Archive

25 October 2022 Files for the 16:00 and 17:00 UTC 25 October CPC 4 km Merged Global IR failed to be downloaded. They were later retrieved, but all 4 half hours were not available in the IMERG Early and Late Runs.

1 September 2022 SNPP ATMS went into safe mode at 09:00 UTC 28 August. Near-Real-time data were restored at 19:00 UTC 1 September.

18 August 2022 SNPP went into safe mode at 16:09 UTC 26 July. Spacecraft recovery, diagnostics, and instrument/product verification enabled near-real-time SNPP ATMS data starting with 21:30 UTC 16 August, with some expectation that embargoed data from the test period might start as early as 15:33:30 UTC 10 August for the post-real-time archive.

9 May 2022 The upgrade of GPROF and the Combined Radar-Radiometer Algorithm (CORRA) code to Version 07 as of 00 UTC on 9 May, while IMERG continues at V06 until August,

requires the Early and Late Runs to change slightly and be given the version number 06C. Specifically, the V07 GPROF and CORRA data must be used in V06 IMERG, and the GPCP climatological calibration has been recomputed to account for the change in CORRA. In addition, the GMI-other sensor calibrations will shift over the next month as V07 data are incorporated into the accumulation files. We expect the Early and Late time series to be relatively consistent across this data boundary.

IMERG was upgraded to V06C when the 00 UTC 9 May data became available for the Early Run. At that point the Late Run only had V06 input available, but the new GPCP-adjusted CORRA calibration started being applied. After about 10 hours the Late also shifted to V07 input at 00 UTC 9 May. We expect the changes to be modest, and think that users desire a continuous record of data, so the additional small inhomogeneity in the Late record is accepted. All Late files with V06C before 00 UTC 9 May are affected.

Starting with 00 UTC 9 May a gridding problem prevented the use of GPROF-AMSR2 input in IMERG Early and Late. The problem was rectified around 16 UTC and data started flowing into IMERG. As well, note that the Final Run was paused when the radar products (which contribute to CORA) were upgraded to V07, and it will next be computed when IMERG moves to V07 in August.

- 24 March 2022* A mis-set error counter on the Core Observatory provoked a safe hold at 13:31:29 UTC on 23 March 2022, shutting down DPR and GMI observations. DPR and GMI were restored at 10:54 and 14:54 UTC, respectively on 24 March 2022. No data were recoverable from the outage period, although IMERG Early and Late continued to be produced (without GMI input, of course). The same will be true for the Final Run when it is produced.
- 24 March 2022* An outage at FNMOC prevented transmission of DMSP (SSMIS) from 15:51 UTC 23 March 23 to 00:47 UTC 24 March 2022. Gaps in the data are: F16: 14:34:07-19:29:39, F17: 14:56:24-19:46:48, F18: 14:47:47-19:47:28. These affect the Early and Late Runs.
- 23 March 2022* Megha-Tropiques ceased providing useful data, signaling the end of the SAPHIR record.
- 24 February 2022* Noise in the 12:00 UTC 22 February 2022 CPC Merged 4 km IR data stalled parts of the near-real-time processing system, eliminating IR data from the Early for 12:00 UTC 22 February 2022 through 05:00 UTC 23 February 2022, and the Late for 12:00-20:00 UTC 22 February 2022. These hours were reprocessed, with Early having backward propagation included.
- 17 February 2022* The GEOS Forward Processing (GEOS FP) system, will be upgraded from GEOS-5.27.1 to GEOS-5.29.4 starting with the 06:00 UTC cycle on Tuesday, March 1, 2022. GOES16 (GOES-W) winds are specifically cited as improving tropical winds, which presumably means the impact is in the east/central Pacific and western Americas. As well, “statistically significant improvements also occur in the temperature and humidity fields”, but since the time series of vertically integrated vapor is used in morphing, the impact on Early and Late propagation vectors should be minimal.
- 9 January 2022* CPC computer systems were restored around 0300 UTC on 8 January 2022, with data recovered back to 1300 UTC on the 7th. IR data re-entered the Early Run with 1500 UTC on the 8th, and the Late Run with 0000 UTC on the 8th. Additional data recovery at CPC will fill in all IR datasets for the Final Run.

- 6 January 2022 “Major computer problems” prevented processing/transmission of the CPC 4 km Global IR data files starting with 1600 UTC on 5 January 2022.
- 28 September 2021 The GPCC discovered an error in the monthly Monitoring Analysis in a small region in Italy for January 2021 and reposted their analysis; the January 2021 files for the affected Final Run have now been replaced by updated files.
- 15 August 2021 DMSP data delivery was interrupted from about 10:30 to 23:00 UTC on 14 August 2021; all missing data were backfilled, but were not available for some half hours of the Early and Late Runs.
- 8 August 2021 SNPP data transmission was interrupted at 02:45 UTC 8 August 2021 and resumed 07:42 UTC 8 August 2021, with data recovery following.
- 5 August 2021 SNPP went into safe mode at 12:46 UTC 3 August 2021 during star tracker updates. Data were restored by 14:17 UTC 4 August 2021, although real-time access started at 21:00 UTC 5 August 2021.
- 23 July 2021 GOES-17 (GOES-W) went into safe mode around 05:37 UTC 22 July 2021. ABI-L1B data were restored by 15:06 UTC 23 July 2021. For most half-hours, this caused small triangles of missing IR data in the Gulf of Alaska and the corresponding location in the Southern Hemisphere, as well as less reliable estimates on either side of that region due to higher zenith angles for the adjacent Himawari and GOES-Central data.
- 18 May 2021 NOAA PDA processing system conducted a failover test that resulted in a gap in ATMS data in Early and Late products for the period 18:30 UTC 17 May to 02:00 UTC 18 May.
- 15 February 2021 The CPC 4-km IR data are susceptible to episodic periods of noise in some channels due to an issue with the cooling system aboard the GOES-17 (GOES-W) satellite. Beginning with data from February 15, 2021, a screening technique was implemented in the post-real-time IMERG V06 to prevent these artifacts from affecting downstream processing. See **IR sensor error detection/correction**.
- 7 January 2021 The time field was behind by a day in the Early Run for the span 00:00 UTC 1 January 2021 to 15:30 UTC 5 January 2021 and in the Late Run from 00:00 UTC 1 Jan 2021 to 04:30 UTC 5 Jan 2021, apparently due to a stale code module. These were reprocessed on 7 and 6 January, respectively. All other information, including date/times in the files and the file names was correct.
- 27 September 2020 As of 07:43:54 UTC on 27 September 2020 the ProductVersion metadata for IMERG Early and Late was corrected from 06A to 06B, which presumably was incorrect since the shift to 06B processing at 13:33 EDT 22 May 2019. [Note: the actual product hours of the shifts will be later by the latency of the product.]
- 22 September 2020 A bug fix was installed in GPROF that eliminates negative POP values in sounder retrievals.
- 9 September 2020 The first Fall 2020 GOES-W noise season was less intrusive than previously, running up to 3 half-hours during 25-30 August 2020.
- 27 August 2020 GPM Core Observatory went into safe-hold ~1100 UTC 19 August 2020 over the Southern Indian Ocean due to a single electro-magnetic event that affected the spacecraft processor. Science data were lost until operations resumed on 26 August 2020, starting ~1400 UTC for GMI and ~2300 UTC for DPR. IMERG continued production throughout the outage, but without the GMI data or accumulating calibration information from CORRA.

- 30 *April 2020* At about 13:15 UTC on 30 April 2020 the Ka was returned to service, including driving the “full” DPR-based Combined, which then replaced the Ku-only Combined in the CORRA-GMI calibration accumulators.
- 27 *April 2020* At about 07:50 UTC on 22 April 2020 the DPR went off-line (over Antarctica). The DPR-Ku was returned to service at about 18:17 UTC, but the DPR-Ka continued to be off-line. A Ku-only CORRA was instituted around 13:34 UTC on 26 April 2020, which started feeding the CORRA-GMI calibrator accumulation files around 11:20 UTC on 27 April 2020, meaning the NRT calibration will be about 6 days short of a full set, but very likely still stable.
- 26 *February 2020* As in August, the GOES-W cooling problems resulted in noise that contaminated the CPC 4-km merged global IR product, and so the Early and Late estimates. The date span was 15-26 February 2020. Furthermore, it appears that this problem is predictable for 4 times a year in February, April, August, and October. The IMERG team is experimenting with automated QC.
- 7 *February 2020* The METOP-A channel 2 started degrading in early December 2019 and has been judged unusable for IMERG. Following the discovery of this problem in early February 2020, use was halted in the Early and Late Runs after the 3 February 2020 orbit ending 15:33:06 UTC. Users should expect to see some noise from METOP-A starting in mid-December 2019 up to this date/time. The Final Run stops using METOP-A after 5 December 2019 based on a conservative analysis that includes the fact that METOP-B and -C are in nearly redundant orbits.
- 13 *December 2019* Anomalous SAPHIR longitude data were discovered starting in orbit 9904 and ending in 10001, which covers most of the period 12 September 2013 09 04:00 UTC to 19 September 2013 02:29 UTC. The bad data have been set to missing, although they are included in all Runs of V06B IMERG.
- 20 *November 2019* METOP-C data began to be used in Early and Late Runs with 10 November 2019 at 0740 UTC.
- 18 *November 2019* Difficulties with accessing NOAA CPC 4 km Merged Global IR data resulted in the loss of IR data in IMERG Early Run for 12 November 12:00-21:30 UTC and 14 November 10:00-15 November 10:30 UTC, and in IMERG Late Run for 12 November 12:00-20:30 UTC and 14 November 10:00-15 November 09:30 UTC. Subsequently, the IR data were retrieved for use in the Final Run.
- 25 *October 2019* Minor GOES-W noise reappeared for the 23 October 23 12:30-13:00 UTC frames and the 24 October 12:30 UTC frame.
- 8 *October 2019* The Final Run will transition to the new ERA5 analysis for determining PLPP as of September 2019 because the previously used ERA-I ends with August 2019.
- 20 *September 2019* NOAA-19 NRT 1C and GPROF are providing answers as of 20 September 2019 at 102210 UTC.
- 9 *September 2019* Beginning with 8 September 2019 and becoming progressively worse, to almost complete failure by 9 September, Channel 2 on NOAA-19 began failing.
- 1 *September 2019* The GOES-W noise appears to have been corrected as of 1 September 2019.
- 27 *August 2019* Beginning with 19 August 2019 at 13:00 UTC, noise in the form of random horizontal lines over the Eastern Pacific began appearing, with most of the problems appearing each following day in the 12:00, 12:30, and 13:00 half-hours, with at least one instance of 13:30. This was traced to noise in the GOES-W Tb’s (as the result of the faulty sensor cooling system), which made it past the filtering at CPC, where the Merged Global 4-km IR dataset is

assembled. The Early and Late datasets contain these errors, which will be dealt with in the Final computation.

- 20 August 2019* Retrospective processing for the V06B Early and Late Runs are finished, providing a record from June 2000-delayed present. This sets the TMPA and TMPA-RT products on a planned end date of December 31, 2019.
- 3 July 2019* Retrospective processing for the V06B Final Run finished, providing a record from June 2000-March 2019.
- 19 June 2019* Data networking problems with NOAA/CPC prevented access to an extended period of CPC 4 km Global Merged IR. Initial Processing for both Early and Late was suspended as of about 14:00 UTC on 19 June, and the existing files starting with 22:00 UTC 16 June are being reprocessed and replaced, up until 03:30 UTC 19 June for Early and 23:30 UTC 18 June for Late. Data in this time range that was previously pulled should be discarded and replaced with the new files. The Early and Late returned to their nominal latencies with the 19:30 and 12:30 UTC 19 June datasets, respectively.
- 10 June 2019* Effective around 12:00 UTC 10 June 2019 the MHS and ATMS input to near-real-time IMERG suddenly became unavailable due to an unexpected change in NOAA protocols. While PPS scrambles to remedy the issue, the data content of the Early and Late Runs may be reduced for the next several days, meaning lower Quality Index due to longer morphing of microwave estimates and greater contributions by IR.
- 5 June 2019* The new, corrected Kalman filter stats were implemented for the Early and Late as of 1337 UTC on 5 June 2019, later than announced on 22 May 2019 as planned for 1 June.
- 4 June 2019* All V06A files are being deleted and a new retrospective processing for the Final Run as V06B is starting on 5 June 2019. The completion date will depend on balancing this new processing and the other PPS scheduled activities.
- 22 May 2019* after 22 May 13:33 EDT all HDF files will have V06B on them in NRT. This reflects that the new surfrac files have been installed to replace the one in V06A used in the Kalman filter that is shifted several grid boxes east and south. However, the effect of this change will not be seen in the output until the next computation of Kalman filter stats on 1 June.
- 2 May 2019* Early and Late Runs were upgraded to V06 as of 00 UTC 1 May 2019.
- 16 April 2019* SSMIS data transmission went out for the following UTC times on 16 April 2019:
F16: 05:54:00 - 14:40:32
F17: 06:00:09 - 14:40:12
F18: 04:38:52 - 15:06:30
- As well, ATMS data for NOAA20 failed to arrive in sufficient time to appear in the Early and Late from 00:00:00 04 April 2019 (or earlier) to 03:00:26 16 April 2019; and ATMS for SNPP from 00:00:00 04 April 2019 (or earlier) to 08:39:20 10 April 2019, followed by partial data to 10:30:16, and late data to 14:00:21.
- 20 March 2019* Communication issues led to an outage of AMSR2 data from 14:50:55 UTC 18 March 2019 to 05:42:23 UTC 19 March 2019.
- 15 March 2019* An error has been discovered in processing the initial batches of V06 IMERG Final Run months. A design choice in the code ended up retaining microwave precipitation estimates in the latitude band 60°N-S when there is snow/ice on the surface, rather than masking out the estimates due to low performance in such cases. [Masking for microwave retrievals over surface snow and ice outside that band is correct.] All V06A Monthly and Half-hourly IMERG products are being retracted, as shown in this table:

| Data Type | Version | Start of Data | End of Data | # Granules | # Empty | Total Size |
|-----------|---------|------------------------|------------------------|------------|---------|--------------|
| 3IMERGHH | V06A | 2014-06-01 00:00:00 | 2015-12-31 23:59:59 | 27792 | 0 | 273.23 GB |
| 3IMERGM | V06A | 2014-06-01 00:00:00 | 2015-12-31 23:59:59 | 19 | 0 | 0.93 GB |

Replacements are anticipated to be released the week of March 25, and will continue to be named V06A since the retracted files are considered faulty.

- 13 *March 2019* IMERG Version 06 retrospective processing has begun. The Final Run will be computed for the GPM era, followed by retrospective Early and Late Runs. Thereafter, these Runs will be computed for the TRMM era, starting with April (Final) and June (Early, Late) 2000, and eventually extending back to 1998 when all data become available. The complete dataset is expected in late Summer. Several months later, the legacy TMPA dataset will be retired.
- 22 *October 2018* NOAA18 started showing degraded channel data at 18:00 UTC 20 October 2018-10-21 due to irregular function of the stepper motor.
- 4 *October 2018* 2BCMB was promoted to V06 starting at 11:41:28 UTC 4 October 2018. This means that the Early and Late calibrations will shift to V06 over the next 45 days; the differences are small enough that these products will be continued with no adjustments. As well, this means that July 2018 will be the last month of Final data that can be run in V05.
- 25 *August 2018* SSMIS data were unavailable from 0530-1900 UTC 24 August 2018. They were subsequently backfilled, but Early estimates likely lack for roughly 0530-1500 UTC.
- 23 *August 2018* For the period 03 UTC 21 August to 16 UTC 22 August 2018, the CPC 4-km IR data were unavailable. Despite the outage, IMERG Early and Late products continued to be produced, but users should be aware that the quality of these products during this period may be suspect and should be used judiciously. Furthermore, the products for the first few hours after the outage are less than optimal due to IMERG restart but should be acceptable for most users. Details:
1. The precipitation propagation vectors for the period slowly degraded after 03 UTC 21 August due to the missing CPC 4-km even-odd IR files. Mechanisms are in place to extrapolate motion vectors into the future in case of brief IR outages but will degrade quickly if the outage is beyond a few hours.
 2. As IR precipitation estimates were unavailable for the outage period, microwave estimates propagated beyond ± 90 minutes were used "as is". Normally, IR precipitation is weighted in the final precipitation products outside the ± 90 minute window (*precipitationCal* and *precipitationUncal*).
- 25 *March 2018* Because of processing errors that occurred in the 11:00 UTC 24 March 2018 cycle, both Early and Late Run IMERG halted dataset production with that date/time. As of 19:00 UTC 25 March 2018 we have successfully restarted both datasets. However, due to staff travel, only the Early Run IMERG will be caught up to its typical (4 hr) latency on 25 March. The Late Run will be caught up to its typical (14 hr) latency some time during the morning of 26 March Central European Time.
- 27 *January 2018* Effective 27 January 2018 all the IMERG Early products are retrospectively processed to V05B for the entire GPM mission (starting in March 2014). V04 products are available, but retired, while V03 products are no longer available. IMERG Late products are retrospectively processed to V05B through January 2017, with a complete set available in another week. The GIS TIFF products should all be at V05B in approximately 3 weeks.

- 17 November 2017* A problem with the metadata in the original Version 05A IMERG Final files required us to withdraw the V05A files and re-release the data as Version 05B, starting 20 November. When the Early and Late Runs are released, starting on 1 December, they will be labeled Version 05B for consistency with the Final, even though no V05A Early and Late were released.
- 9 November 2017* IMERG has now been upgraded to Version 05. The transition to V05 for the IMERG Final Run is planned for 13 November at PPS and the new data should start flowing down to the GES DISC as well. The version number is 05A. Initially, the data record begins with 12 March 2014 and additional months of data will be added over the next few weeks. It is planned to commence Version 05 Early and Late Run in December.
- 25 October 2017* At about 4:57 UTC on 9 October 2017 the NOAA-19 MHS precipitation estimates started displaying artifacts, which was eventually traced to the instrument going into safe mode without shutting down data delivery. Since this happened over the long Columbus Day weekend, it took until 01:47 UTC on 10 October 2017 to shut down the data stream, so the Early and Late IMERG have these (very obvious) artifacts for almost 24 hours. No reprocessing is planned. The sensor resumed operations at 17:31:08 UTC on 16 October 2017, but because the basis for the safe mode was unknown, GPM chose to monitor the data stream at first; GPROF-MHS estimates from NOAA-19 were restarted for the orbit segment beginning 10:32:49 UTC on 25 October 2017.
- 30 September 2017* A malfunction in the AMSR2 data recorder aboard the GCOM-W1 satellite resulted in the loss of data from 16:34 UTC on 27 September 2017 to 15:39 on 28 September 2017.
- 30 May 2017* Starting at about 15:00 UTC on 30 May the Version 05 GPROF-ATMS and GPROF-MHS Level 2 precipitation estimates were given quality flags of 3 in the outer swath region, suggesting that users (in particular IMERG) should not use these footprints due to unrealistic areas of widespread light rain in moist tropical oceanic cases. However, the IMERG Early and Late results from the start of Version 05 processing for core products (9 May) have erroneous data in these footprints. It is not obvious that these footprints yield unrealistic values outside of moist tropical ocean areas, but at present not enough is known to do a more sophisticated screening.
- 16 May 2017* At about 11:30 UTC on 12 May the Version 05 GPROF-ATMS Level 2 precipitation estimates were turned off because the scan positions were misnavigated, which caused frequent unreasonable values in the last 6 footprints. A corrected version was restarted about 23:00 UTC on 15 May.
- 9 May 2017* At 02:20 UTC Version 04 processing was shut down for GPM core products and subsequently started as Version 05. At this point IMERG continued to run in V04, with backward compatibility adjustments to V05 inputs.
- 22 March 2017* IMERG Final Run was upgraded to Version 04 beginning on 22 March with a release of 12 March 2014 through May 2015. Retrospective processing up to the nominal latency (about 3.5 months) should take a few weeks.
- 14 March 2017* IMERG Early and Late Runs were upgraded to Version 04 beginning at 15 UTC on 13 March and both were back to their target latency by 18 UTC on 14 March. The initial record starts on 5 February, with retrospective processing to the start of GPM (12 March 2014) after the Final Run is processed.
- 23 December 2016* Anomalies in the DMSP F-16 input data for the period 09 UTC 18 December 2016 to 18 UTC 19 December 2016 caused the orbits to be deleted.

- 27 *September 2016* A user discovered that the January 2016 IMERG Final Run lacked gauge data, which turned out to stem from a failed change of code to accommodate a change in format for the GPCC gauge analysis. The month was re-run.
- 9 *September 2016* SSMIS data stopped flowing starting with data timestamps around 00 UTC on 8 September 2016. By late on 8 September FNMOC commenced backfilling the data.
- 8 *July 2016* FNMOC informed Wes Berg that “spike detection” was turned off on 18 May, at which point the F17 anomaly ceased. The 37V channel was again used, and in future reprocessings we expect to set the 37V channel to missing for the period 5 April – 18 May 2016.
- 26 *May 2016* Version 4 AMSR2 1Base, 1C, and GPROF were installed 06:31:36 UTC on 26 May 2016, followed by the same for SSMIS at 08:15:25 UTC.
- 10 *May 2016* The data flow from the Himawari-8 GEO satellite was interrupted for 57 hours, 7 May, 05 UTC – 9 May, 14 UTC and during that time there is a continuous zone of missing values in the *IRprecipitation* data in the center of the Himawari-8 sector (over Japan), where data from the adjoining satellites are unable to fill the gap. The IR-based displacement vectors were computed using the standard fallback of spatial interpolation in the missings.
- 20 *April 2016* Beginning with 5 April 0823 UTC the DMSP F17 37V channel experienced intermittent noisy values. Several days of bad data were used in V03 Early and Late processing (detectable by notable swaths of high precipitation over land). Beginning 26 April 1727 UTC, V03 GPROF-F17 retrievals were run without using the 37V channel. The quality is somewhat less, but judged to be better than the alternatives (morphed or IR).
- 4 *March 2016* Version 04 GPROF-GMI began, effectively ending the Final Run with January 2016, and forcing back-calibration of V04 near-real-time data to V03 to enable the Early and Late Runs to continue until IMERG switches to V04.
- 3 *March 2016* The GPM NRT system began using V04 algorithms for: L1B and L1C GMI, L1 and L2 radar, and L2 Combined. GMI is extremely well calibrated and the brightness temperatures in L1B and L1C are an important improvement over those in V03. V04 also made improved retrievals in the Ku, Ka, DPR, and Combined. The GPROF algorithm used for radiometer retrievals will remain at V03 until approximately April 15 due to late bug fixes to Combined that impacted the GPROF database. In parallel with GPROF, IMERG is being held at V03. In order to maintain consistent V03 calibration, the GCI-GMI calibration in IMERG was frozen effective 0230 UTC on 3 March 2016. We expect the impact to be modest, if noticeable at all.
- 11 *February 2016* DMSP F19 stopped responding to command and control transmissions late in the day. Telemetry confirms that the satellite is functional, but cannot transmit data.
- 7 *July 2015* The Japanese geo-satellite transitioned from MTSat-2 to Himawari 8 effective 0200 UTC on 7 July 2015.
- 23 *June 2015* The two Early Run files with start times 0700 through 0730 are missing due to missing input. Late and Final Run files are available for these times.
- 11 *June 2015* The four Early Run files with start times 2100 through 2230 are missing due to missing input. Late and Final Run files are available for these times.
- 15 *May 2015* The four Early Run files with start times 1300 through 1430 are missing due to missing input. Late and Final Run files are available for these times.
- 28 *April 2015* NOAA reprocessed the global IR data for 10 UTC 26 April through 14 UTC 27 April due to dropped images; however there is no straightforward way to re-do the Early and

Late Runs, so that period should display somewhat lower quality in the regions that lack IR input.

17 April 2015 IMERG Early and Late Runs ceased when CPC Global 4 km Merged IR data dropped out, starting 17 UTC 14 April due to processing issues at NOAA, and were caught up after the data returned around 11 UTC 17 April.

8 April 2015 End of TMI data

1 April 2015 Initial release of IMERG Late Run (V03D); beta release of Day-1 IMERG Early Run. [1 August 2016: In fact, the Late and Early Runs were released as V03E, while the Final Run was left at V03D; see “data set file names”]

13 March 2015 Beta release of Day-1 IMERG Late Run.

20 January 2015 Revised release of IMERG Final Run (V03D).

16 January 2015 Initial release retracted due to minor inconsistencies in “missing” values.

15 January 2015 Initial release of IMERG Final Run (V03D).

4 December 2014 Beta release of Day-1 IMERG Final Run for Early Adopters.