Chapter 19: Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement (GPM) mission (IMERG)

George J. Huffman¹, David T. Bolvin^{1,2}, Dan Braithwaite³, Kuolin Hsu³, Robert Joyce^{4,5}, Christopher Kidd^{1,6}, Eric J. Nelkin^{1,2}, Soroosh Sorooshian³, Erich F. Stocker⁷, Jackson Tan^{1,8}, David B. Wolff⁹, Pingping Xie⁵

Abstract The Integrated Multi-satellitE Retrievals for the Global Precipitation Measurement (GPM) mission (IMERG) is a U.S. GPM Science Team precipitation product. IMERG uses intercalibrated estimates from the international constellation of precipitation-relevant satellites and other data, including monthly surface precipitation gauge analyses, to compute half hour, 0.1° x 0.1° gridded datasets over 60°N-S (and partially outside of that latitude band) in three "Runs"—Early (4 h after obs time), Late (14 h after obs time), and Final (3.5 months after obs time). The concepts behind IMERG are briefly reviewed, together with major shifts related to changes in versions from the at-launch Version 03 to Version 05, and an outline of Version 06, which was released in late 2019.

19.1 Introduction

The critical role of fresh water for life on Earth, and the necessity of using satellite data to estimate its ultimate source, precipitation, is a repeated theme throughout this book. Most of the relatively

¹G. J. Huffman, NASA/GSFC Earth Science Division, Greenbelt, Maryland USA e-mail: george.j.huffman@nasa.gov Not sure how to get the repeated and multiple affiliations right: 2: Science Systems and Applications, Inc. Lanham, Maryland USA 3: University of California Irvine Irvine, California USA 4: Innovim Greenbelt, Maryland USA 5: NOAA/NWS Climate Prediction Center College Park, Maryland USA 6: University of Maryland / ESSIC College Park, Maryland USA 7: NASA/GSFC Global Change Data Center Greenbelt, Maryland USA 8: Universities Space Research Association Columbia, Maryland USA 9: NASA/GSFC Wallops Flight Facility Wallops Island, Virginia USA

accurate satellite-based precipitation estimates are provided by passive microwave (PMW) sensors flying on low Earth orbit (LEO) platforms. Fig. 19.1 shows the entire history of "modern" PMW sensors, which started in mid-1987 with the Special Sensor Microwave/Imager (SSM/I). It is also possible to estimate precipitation from infrared (IR) sensors, with geosynchronous Earth orbit (GEO) satellites providing convenient and relatively complete global coverage, except in polar latitudes, albeit with lower quality.

The proliferation of LEO-PMW satellites around the start of the millennium made it possible to create multi-satellite precipitation data sets that had uniform and increasingly finer-scale time/space grids. Typically, the focus has been on providing the best short-interval estimates (at the expense of less homogeneity in the data record), referred to as High Resolution Precipitation Products (HRPP). Examples include the Climate Prediction Center (CPC) Morphing algorithm (CMORPH; Joyce et al. 2004), the Global Satellite Mapping of Precipitation (GSMaP; Kubota et al. 2007), and the Tropical Rainfall Measuring Mission (TRMM) Multi-Satellite Precipitation Analysis (TMPA; Huffman et al. 2007, 2010).

With the launch of the Global Precipitation Measurement (GPM) mission Core Observatory (GPM-CO) in February 2014, the U.S. GPM science team instituted the Integrated Multi-satellitE Retrievals for GPM (IMERG) merged precipitation product. This algorithm is intended to intercalibrate, merge, and interpolate all available satellite PMW precipitation estimates, together with microwave-calibrated infrared (IR) satellite estimates, precipitation gauge analyses, and other precipitation estimators, to produce fine time- and space-scale estimates spanning the TRMM and GPM eras for the entire globe. IMERG is computed three times for each observation time, first providing an initial estimate and then successively better estimates as more data arrive. The final step uses monthly gauge data to create research-level products. The at-launch Version 03 and subsequent Versions 04 and (current) 05 only covered the GPM era, while Version 06 extends back through the TRMM era.

Focusing on the current Version 05, Sect. 19.2 describes the input data sets, while Sect. 19.3 describes the IMERG processing and Sect. 19.4 provides the status of the data sets. Sect. 19.5 describes some tests of the algorithm performance with examples of IMERG applied to typical applications. Sect. 19.6 presents the status for Version 06 and concluding remarks.

19.2 Input Data Sets

The precipitation-relevant satellites mentioned above all enter the IMERG computation as singlesensor precipitation estimates. Subsequently, inter-calibration is performed to ensure that these are as consistent as possible. The main groupings are:

- 1. GPM-CO: The GPM Microwave Imager (GMI) is a well-calibrated conical-scan multi-channel, dual-polarization PMW sensor, and is considered the standard for the other PMW sensors. Its precipitation retrievals (Kummerow et al. 2015 are used in IMERG as direct estimates, as well as contributing to the inter-calibration process. The GPM-CO also carries the scanning Dual-frequency Precipitation Radar (DPR) which is key to the inter-calibration process as part of the Combined Radar-Radiometer Algorithm (Olson et al. 2011) estimates.
- 2. PMW constellation: The rest of the GPM era constellation (cf. Fig. 19.1) is composed of satellites of opportunity. The U.S. Defense Meteorological Satellite Program (DMSP) F16, F17, F18, and F19 Special Sensor Microwave Imager/Sounder (SSMIS); the Japanese Global Change Observation Mission-Water (GCOM-W1) Advanced Microwave Scanning Radiometer (AMSR2); and TRMM Microwave Imager (TMI) are conically scanning imagers, while the Joint Polar Satellite System (JPSS-1) and Suomi National Polar-orbiting Partnership (SNPP) Advanced Technology Microwave Sounder (ATMS); and European Meteorological operational

satellites (METOP-A and ,METOP-B), and the U.S. National Oceanic and Atmospheric Administration (NOAA-18 and NOAA-19) Microwave Humidity Sounder (MHS) are cross-track scan sounders. Observations from all of these sensors are processed using the Goddard Profiling (GPROF; Kummerow et al. 2015) algorithm to provide precipitation estimates.

- 3. GEO-IR constellation: IR data from the U.S. Geostationary Operational Environmental Satellite (GOES), the Japanese Himawari, and the European Meteosat satellite series cover the central Atlantic through to the central Pacific, from there to the Indian Ocean, and from there to the central Atlantic, respectively. The data are processed into precipitation retrievals using Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks Cloud Classification System (PERSIANN-CCS; Hong et al. 2004).
- 4. Work with previous datasets has shown that incorporating a uniform precipitation gauge analysis, even at the monthly scale, is important for controlling the biases that satellite precipitation estimates typically contain. We use the Deutscher Wetterdienst (DWD) Global Precipitation Climatology Centre (GPCC) V5 Monitoring Product (Becker et al. 2013, Schneider et al. 2014, Schneider et al. 2017) through March 2018, and the V6 thereafter.
- 5. Ancillary products: The IMERG algorithm also accesses surface type (the standard static map of percent water coverage from the Precipitation Processing System [PPS]); snow/ice surface extent (AUTOSNOW from NOAA); and surface temperature, relative humidity, and surface pressure (Japan Meteorological Agency forecasts [for Early and Late] of meteorological data; and the European Centre for Medium-range Weather Forecasts [ECMWF] analysis [for Final, for consistency with the GPROF "climatological" run]).

19.3 IMERG Processing

The input data from the various satellite PMW sensors are assembled, received at PPS as Level 1 brightness temperatures (Tb) from the relevant providers, intercalibrated to GMI as GPM Level 1C brightness temperatures (Tc), then converted to Level 2 precipitation estimates using the V05 GPROF scheme. All estimates are gridded, inter-calibrated to the Ku swath Combined Radar-Radiometer (CORRA) product on a rolling 45-day basis using probability matching, and climatologically calibrated to the Global Precipitation Climatology Project (GPCP) monthly Satellite-Gauge estimates with a simple ratio in high latitude oceans (where GPM-CO Version 05 products are deficient in precipitation) and over all land areas (where the CORRA tends to be high). These "high quality" (HQ) data are combined into half-hourly fields, masked for surface snow and ice (due to uncertain quality in the GPM-CO products), and provided to both the recalibration of PERSIANN-CCS infrared estimates and to the semi-Lagrangian time interpolation scheme adapted from CMORPH-Kalman Filter (CMORPH-KF; Joyce et al. 2011). In parallel, CPC assembles the zenith-angle-corrected, inter-calibrated "even-odd" and merged geo-IR fields and forwards them to the PPS for use in the CMORPH-KF semi-Lagrangian time interpolation scheme and the PERSIANN-CCS computation routines, respectively. The PERSIANN-CCS estimates are computed (supported by an asynchronous 30-day re-calibration cycle) and sent to the CMORPH-KF weighting scheme. The CMORPH-KF weighting scheme (supported by an asynchronous KF weights 3-monthly updating cycle) uses the PMW and IR estimates to create half-hourly estimates. Note that various intermediate fields are carried through the processing as necessary to populate the fields in the output file (Table 19.1). Precipitation phase is computed in the PMW merger step as a diagnostic using surface type, surface pressure, surface temperature, and surface humidity (after Sims and Liu 2015). Finally, user requests for a simple Quality Index (QI) led to two distinctly different QI definitions for the half-hourly and monthly data sets in Version 05 (https://pmm.nasa.gov/sites/default/files/document files/IMERG QI.pdf, last accessed 17 Nov. 2018).

IMERG is processed twice in near-real time:

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

For the Final product the half-hour multi-satellite estimates are adjusted to equal the monthly satellite-gauge combination computed in a monthly IMERG estimate (following the TMPA approach). In all Runs, the output contains multiple fields that provide information on the input data, selected intermediate fields, and estimation quality (Table 19.1).

To ensure a consistent archive of data for all users, all three runs of IMERG are retrospectively processed, including for Version 05. In practice, IMERG is first retrospectively processed as the Final Run, then again with calls that mimic the processing for Early and Late, but using the data available to the Final. Specifically, the selection of input data available to the retrospective Early Run is approximated by limiting the forward time span of data to the typical latency time (~3 h) before the Early run time (currently 4 h after observation time). The Late uses both backward- and forward-morphing, but neither the Early nor Late Runs are given gauge data. These choices cause the Early and Late Runs to be reprocessed with a superset of input data covering the original runs, and the input data from a particular sensor are produced by the climatological GPROF estimates (computed with more-carefully prepared reanalysis data). Retrospective processing for both the Early and Late Runs is carried out after retrospective processing for the Final Run.

All output data files have multiple data fields with project-mandated metadata and are written in HDF5, which is compatible with NetCDF4. All fields are produced for all Runs, as listed in Table 19.1. Since the PPS provides interactive data subsetting by time, region, and parameter, users are not required to download the entire file. Furthermore, both the PPS and the Goddard Earth Science Data and Information Services Center (GES DISC) create value-added products that give the user additional formats and/or periods of accumulation (*https://pmm.nasa.gov/data-access/downloads/gpm*, last accessed 17 Nov. 2018, provides a current summary).

19.4 IMERG Data Set Status

The at-launch IMERG V03 was processed using pre-launch calibrations, as were all the other GPM products such as GPROF. The cross-track sounder PMW precipitation retrievals computed from the MHS were processed using an alternative scheme (Kidd et al. 2016). Version 04 used initial GPM-based calibrations, as well as V04 GPROF for all PMW data, including the cross-track sounder retrievals. Among other changes, in this version we started dynamically calibrating PERSIANN-CCS to the HQ merged PMW data, and HQ was extended to cover the entire globe.

IMERG was upgraded to V05 in November 2017, with retrospective processing occurring over the following 2.5 months. In this version the fully global GPROF estimates were inserted into the complete precipitation fields (precipitationCal, precipitationUncal) outside the standard IR domain (60°N-S), although without morphing or IR fill-in. Recall that GPROF estimates over snowy/icy surface types are set to missing, so data coverage outside 60°N-S is quite limited. Version 05 also marked the introduction of the Quality Index. Because the Version 05 GPROF-TMI estimates had not been computed for the GPM era when the IMERG Runs were retrospectively processed, TMI retrievals are not included in the V05 IMERG datasets.

The Early, Late, and Final Runs all cover the practically useful data record from GPM, starting in March 2014. The Early and Late Runs occasionally suffer data dropouts and faults in the input data that are transient, but are not corrected until after processing. PPS consults with the IMERG team in

such cases, usually allowing the deficiency to stand, but occasionally reprocessing the datasets. Such cases are recorded in the IMERG Technical Documentation (see the link to this document at *https://pmm.nasa.gov/data-access/downloads/gpm*, last accessed 17 Nov. 2018, for the latest listing).

One easy way to get a quick qualitative review of recent IMERG performance is by viewing the "last week of IMERG" visualization provided by the Goddard Science Visualization Studio (<u>https://svs.gsfc.nasa.gov/cgi-bin/details.cgi?aid=4285</u>, last accessed 17 Nov. 2018).

19.5 IMERG Performance and Examples

A number of studies are available that document IMERG performance, including the "V05 IMERG Final Run Release Notes" (https://pmm.nasa.gov/sites/default/files/document_files/IMERG_FinalRun_V05_release_notesrev3.pdf, last accessed 17 Nov. 2018), various hydrological assessments (e.g. Sharif et al. 2017, Yuan et al. 2018), and more-general statistical analyses (e.g., Beck et al. 2018, Kim et al. 2017). Additional studies are listed in the GPM citation list (https://pmm.nasa.gov/resources/gpmpublications, last accessed 17 Nov. 2018).

One interesting example of IMERG (Late) performance occurred during Hurricane Harvey, which deposited significant, and even all-time record rainfall in the Houston, Texas area in August 2017. The initial analysis showed that IMERG Late underestimated precipitation in the Houston area (just west of, and extending through, Area 1 on Fig. 19.2). However, as the figure shows, Area 2, inland to the north, showed IMERG overestimates. Furthermore, the time series of area-averages in Fig. 19.3 demonstrates that these under- and over-estimates occurred simultaneously. Clearly, the details of the meteorological setting in the two regions are driving systematic and different retrieval errors by GPROF. The IR-based precipitation (Fig. 19.3) is systematically low, a fact we ascribe to its trailing calibration to microwave data: if rain events similar to those in hurricanes have not occurred in the previous 45 days, the calibrations will very likely lead to weaker estimates than are needed to match the hurricane's rainfall.

19.6 Status for Version 06 and Concluding Remarks

The upgrade to Version 06 for IMERG computed in Fall 2019, after this manuscript went to press. The first major change in this version is carrying the retrospective processing back into the TRMM era. For dates before the launch of GPM, the TRMM satellite products will serve as the calibration standard, although some adjustments are necessary to ensure compatibility with GPM-CO products. For example, TRMM's orbital inclination is 35°, so calibration to CORRA must be approximated in the latitudinal band from that point to 65° (GPM's inclination) in both hemispheres. The initial start date is June 2000 to accommodate data availability, but the plan is to extend back to January 1998 when the necessary data are made available.

Another significant change in Version 06 is shifting the source of the motion vectors from IR to vertically integrated vapor (TQV) as depicted in the Modern-Era Retrospective analysis for Research and Applications, Version 2 (MERRA-2) for the Final Run, and in the Global Earth Observing System Forward Processing (GEOS FP) for the Early and Late Runs. This change was necessitated by data availability issues and our intent to extend the morphing technique beyond the 60° latitude extent of the GEO-IR. Testing indicates that the TQV-based vectors have skill that equals or slightly exceeds the skill of the IR-based vectors.

The new TQV-based vectors are computed over the entire globe, so V06 includes morphed PMW estimates wherever the PMW GPROF estimates are considered reliable. Specifically, in the time available, the IMERG team was not able to develop the adjustments necessary to tune GPROF over ice and snow surface types, so areas with ice or snow cover continue to default to IR precipitation (as in previous versions) in the latitude band 60N-S, and are set to missing at higher latitudes. Nonetheless, providing morphed PMW estimates at higher latitudes will provide a welcome increase in coverage, in summertime land areas, and year-round for open ocean water.

IMERG V06 processing was pursued during summer 2019, and all runs were processed for the complete long-record dataset in fall 2019. That milestone started the countdown to end the computation of the TMPA datasets, which had continued to be produced, even after the demise of TRMM using climatological calibrations, to support users who require a long, relatively homogeneous record. Production of the TMPA datasets was carried to the end of 2019 and ended.

Table 19.1. Lists of data field variable names and definitions to be included in each of the output datasets. Primary fields for users are in italics.

Half-hourly data file (Early, Late, Final)	
precipitationCal	Multi-satellite precipitation estimate with gauge calibration (recommended for general use)
precipitationUncal	Multi-satellite precipitation estimate
randomError	Random error for gauge-calibrated multi-satellite precipitation
HQprecipitation	Merged microwave-only precipitation estimate
HQprecipSource	Microwave satellite source identifier
HQobservationTime	Microwave satellite observation time
IRprecipitation	IR-only precipitation estimate
IRkalmanFilterWeight	Weighting of IR-only precipitation relative to the morphed merged microwave-only precipitation
probabilityLiquidPrecipitation	Probability of liquid precipitation phase
PrecipitationQualityIndex	Quality Index for precipitationCal field
Monthly data file (Final)	
precipitation	Merged satellite-gauge precipitation estimate (recommended for general use)
randomError	Random error for merged satellite-gauge precipitation
gaugeRelativeWeight	Weighting of gauge precipitation relative to the multi-satellite precipitation
probabilityLiquidPrecipitation	Accumulation-weighted probability of liquid precipitation phase
PrecipitationQualityIndex	Quality Index for precipitationCal field



Figure 19.1. PMW sensor Equator-crossing times for 12-24 Local Time (LT; 00-12 LT is the same) for the modern PMW sensor era. These are all ascending passes, except F08 is descending. Shading indicates that the precessing TRMM, Megha-Tropiques, and GPM cover all times of day with changes that are too rapid to depict at this scale. [Image by Eric Nelkin (SSAI; GSFC), 12 July 2018; *https://pmm.nasa.gov/sites/default/files/imce/times_allsat.jpg*, last accessed 1 Apr. 2019, holds the current version.]



Figure 19.2. Rainfall accumulations for the week of 25-31 August 2017 over the U.S. Gulf Coast for NOAA Multi-Radar Multi-Sensor (MRMS) data (left) and IMERG V05 Late estimates (right). Houston, Texas is just west of Area 1.



Figure 19.3. Time series of area-average rainfall for the week of 25-31 August 2017 over the U.S. Gulf Coast for the near-coastal Area 1 (top) and the more inland Area 2 (bottom). Houston, Texas is just west of Area 1. The IMERG Late averages are labeled precipitationCal, and the IR-based precipitation time series is IRprecipitation.

References

- Beck, H. E., Pan, M., Roy, T., Weedon, G. P., Pappenberger, F., van Dijk, A. I. J. M., Huffman, G. J., Adler, R. F., & Wood, E. F. (2019). Daily evaluation of 26 precipitation datasets using stage-IV gauge-radar data for the CONUS. *Hydrology and Earth System Sciences*, 23, 207–224. https://doi.org/10.5194/hess-23-207-2019.
- Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., & Ziese, M. (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*, 5, 71–99. https://doi.org/10.5194/essd-5-71-2013.
- Hong, Y., Hsu, K.-L., Sorooshian, S., & Gao, X. (2004). Precipitation estimation from remotely sensed imagery using an artificial neural network cloud classification system. *Journal of Applied Meteorology*, 43, 1834–1852. https://doi.org/10.1175/JAM2173.1.
- Huffman, G. J., Bolvin, D. T., Nelkin, E. J., Wolff, D. B., Adler, R. F., Gu, G., Hong, Y., Bowman, K. P., & Stocker, E. F. (2007). The TRMM Multi-satellite Precipitation Analysis: Quasi- global, multi-year, combined-sensor precipitation estimates at fine scale. *Journal of Hydrometeorology*, 8, 38–55. https://doi.org/10.1175/JHM560.1.

- Huffman, G. J., Adler, R. F., Bolvin, D. T., & Nelkin, E. J. (2010). The TRMM Multi-satellite Precipitation Analysis (TMPA). In F. Hossain & M. Gebremichael (Eds.), Satellite rainfall applications for surface hydrology. New York: Springer, ISBN: 978-90-481-2914-0, 3-22.
- Joyce, R. J., Janowiak, J. E., Arkin, P. A., & Xie, P. (2004). CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *Journal of Hydrometeorology*, 5, 487–503. https://doi.org/10.1175/1525-7541(2004)005<0487%3ACAMTPG>2.0.CO%3B2.
- Joyce, R. J., Xie, P., & Janowiak, J. E. (2011). Kalman filter based CMORPH. Journal of Hydrometeorology, 12, 1547–1563. https://doi.org/10.1175/JHM-D-11-022.1.
- Kidd, C., Matsui, T., Chern, J., Mohr, K., Kummerow, C., & Randel, D. (2016). Global precipitation estimates from cross-track passive microwave observations using a physically based retrieval scheme. *Journal of Hydrometeorology*, 17, 383–400. https://doi.org/10.1175/JHMD-15-0051.1.
- Kim, K., Park, J., Baik, J., & Choi, M. (2017). Evaluation of topographical and seasonal feature using GPM IMERG and TRMM 3B42 over far-East Asia. *Atmospheric Research*, 187, 95–105. https://doi.org/10.1016/j.atmosres.2016.12.007.
- Kubota, T., Shige, S., Hashizume, H., Aonashi, K., Takahashi, N., Seto, S., Hirose, M., Takayabu, Y. N., Nakagawa, K., Iwanami, K., Ushio, T., Kachi, M., & Okamoto, K. (2007). Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation. *IEEE Transactions on Geoscience and Remote Sensing*, 45, 2259–2275. https://doi.org/10.1109/TGRS.2007.895337.
- Kummerow, C. D., Randel, D. L., Kulie, M., Wang, N.-Y., Ferraro, R., Munchak, S. J., & Petkovic, V. (2015). The evolution of the Goddard PROFiling algorithm to a fully parametric scheme. *Journal of Atmospheric and Oceanic Technology*, 32, 2265–2280. https://doi.org/10.1175/JTECH-D-15-0039.1.
- Olson, W. S., Masunaga, H., & the GPM Combined Radar-Radiometer Algorithm Team. (2011). GPM combined radar-radiometer precipitation. Algorithm theoretical basis document (Version 2). PPS, NASA/GSFC, 58 pp. Available at http://pps.gsfc.nasa.gov/Documents/GPM2011CombinedL2ATBD.pdf, last accessed 17 Nov 2018.
- Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., & Rudolf, B. (2014). GPCC's new land-surface precipitation climatology based on quality-controlled in-situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, 115, 15–40. https://doi.org/10.1007/s00704-013-0860-x.
- Schneider, U., Finger, P., Meyer-Christoffer, A., Rustemeier, E., Ziese, M., & Becker, A. (2017). Evaluating the hydrological cycle over land using the newly-corrected precipitation climatology from the Global Precipitation Climatology Centre (GPCC). *Atmosphere*, 8, 52. https://doi.org/10.3390/atmos8030052.
- Sharif, H. O., Al-Zahrani, M., & El Hassan, A. (2017). Physically, fully-distributed hydrologic simulations driven by GPM satellite rainfall over an urbanizing arid catchment in Saudi Arabia. *Water*, 9. https://doi.org/10.3390/w9030163.
- Sims, E. M., & Liu, G. (2015). A parameterization of the probability of snow-rain transition. Journal of Hydrometeorology, 16, 1466–1477. https://doi.org/10.1175/JHM-D-14-0211.1.
- Yuan, F., Wang, B., Shi, C., Cui, W., Zhao, C., Liu, Y., Rena, L., Zhang, L., Zhu, Y., Chen, T., Jiang, S., & Yang, X. (2018). Evaluation of hydrological utility of IMERG final run V05 and TMPA 3B42V7 satellite precipitation products in the Yellow River source region, China. *Journal of Hydrology*, 567, 696–711. https://doi.org/10.1016/j.jhydrol.2018.06.045.