FAQ for IMERG V07 Applications Users

Andrea Portier, Jacob Reed, Jackson Tan, George Huffman

February 24, 2025

Summary

This FAQ document provides information specifically aimed at the Integrated Multi-satellitE Retrievals for GPM (IMERG) Version 07 **for applications users**. As of August 2024, both Initial Processing (IP) and retrospective processing (RP) for the IMERG Early, Late and Final products have been completed for Version 07. IMERG V07 involves a wide range of changes to all aspects of the product, many of which are implemented in response to feedback on V06, which was the first time the IMERG record spanned the TRMM and GPM eras. IMERG V07 was reprocessed for the entire TRMM / GPM record and thus supersedes all prior IMERG versions, as well as superseding the prior TRMM-based TMPA products (3B42RT, 3B42, 3B43).

We encourage users to review the IMERG V07 Release Notes and the technical documentation at https://gpm.nasa.gov/resources/documents/imerg-v07-release-notes and https://gpm.nasa.gov/resources/documents/imerg-v07-technical-documentation. The Release Notes summarizes key changes in the IMERG V07 algorithm, together with important information about its use, while the Technical Documentation is a detailed log of all information and news relating to IMERG, including decisions made in previous versions.

For other questions about IMERG and other GPM products, please visit the GPM Data FAQ: https://gpm.nasa.gov/data/faq. Additionally, for any other questions, please do not hesitate to reach out to our GPM "Ask A Question" page at https://gpm.nasa.gov/data/ask-a-question.

Summary of Topics

IMERG V07 Product

Where can I find information (latency, resolution, dates covered), documentation, and download links for IMERG products?

Links to the IMERG files are listed on the Data Directory of the GPM website: https://gpm.nasa.gov/data/directory, which point to IMERG archives hosted at the Goddard Earth Sciences Data and Information Services Center (GES DISC): https://disc.gsfc.nasa.gov/datasets?keywords=gpm%20imerg%2007&page=1&subject=Precipitation&source=Multi-sensor%20Analysis&project=GPM, and Precipitation Processing System (PPS): https://jsimpsonhttps.pps.eosdis.nasa.gov/imerg/ (Early, Late) and https://arthurhouhttps.pps.eosdis.nasa.gov/gpmdata/YYYY/MM/DD/imerg/ (Final, where YYYY, MM, and DD are the year, month, and day numbers, respectively), as well as several interactive sites. Please visit https://gpm.nasa.gov/data/news for the latest updates.

What are the major differences between IMERG V06 and IMERG V07?

There are several changes from V06 to V07. Initial evaluation results suggest that the changes in V07 collectively improve the performance of IMERG estimates, both in terms of its precipitation detection, systematic bias, and random bias. This includes improvements in IMERG over frozen, orographic, and coastal surfaces; in the consistency of histograms (including extreme values) between the TRMM and GPM eras; and performance over regions lacking gauge data, particularly the oceans. Below is a quick list of IMERG product changes and improvements. We encourage users to review the "IMERG V07 Release Notes" document for full details: https://gpm.nasa.gov/resources/documents/imerg-v07-release-notes. The IMERG V07 Release Notes document can be read in about 15 minutes.

- Changed and reorganized variable names such as "precipitationCal" to "precipitation" to minimize misinterpretation of the variable names;
- Included PMW estimates over frozen surfaces, resulting in near-global coverage;
- Notable improvement to the algorithm inputs to IMERG (e.g., CORRA V07 and GPROF V07), which helps reduce biases and have greater consistency over the entire record;
- Fixed spatial offset in CORRA and GPROF gridding that caused these IMERG inputs to be spatially offset by one grid box;
- Improved intercalibration of passive microwave (PMW) sensors to correct the biases such as persistent overestimation observed in IMERG V06;
- Changed source of motion vector, aiding in an appreciable improvement in orographic regions and a marginal improvement globally;
- Implemented automated infrared (IR) product quality control that automatically scans for anomalous features (e.g., cold artifacts) and attempts to mask out these artifacts;
- Upgraded the IR precipitation algorithm to the more-modern PERSIANN Dynamic Infrared–Rain Rate (PDIR-Now);
- Increased the precipitation rate cap to 200 mm / h, allowing the possibility of greater extremes in the IMERG estimates when the retrievals compute them;
- Excluded the SAPHIR instrument from the PMW inputs in this version of IMERG as SAPHIR demonstrated spatial patterns that were very different from those of other PMW sensors;
- Developed and implemented the Scheme for Histogram Adjustment with Ranked Precipitation Estimates in the Neighborhood (SHARPEN) scheme as post-processing to the Kalman filter to approximately correct distorted precipitation rate distributions;
- Combined instantaneous PMW retrievals with propagated precipitation, increasing the half-hourly Quality Index values on average and giving a smoother visual appearance to animations;

- Implemented a joint Fuchs-Legates undercatch correction to the gauge analysis;
- Applied climatological gauge adjustment to the Early and Late Runs, ensuring more consistency in the bias characteristics among all IMERG latencies: Early, Late, and Final Runs;
- Refined precipitation phase specification using reanalysis variables, result in higher probability of liquid phase estimates in V07 compared to V06; and
- Rounded precipitation rate values for improved data compression.

What are the changes to the precipitation data fields / variable names?

IMERG V06 Field Name	IMERG V07 Field Name	Description	Units
Grid/precipitationCal	Grid/precipitation	Precipitation rate in mm/hr – This is the most useful variable for most application users	mm/hr
Grid/HQprecipitation	Grid/Intermediate/MWprecipitation	Merged microwave-only precipitation estimate only includes microwave data meaning it has significant gaps.	mm/hr
Grid/HQprecipSource	Grid/Intermediate/MWprecipSource	Microwave satellite source identifier	Index values (see below this Table for key)
Grid/HQobservationTime	Grid/Intermediate/MWobservationTime	Microwave satellite observation time	min. into half hour
Grid/IRprecipitation	Grid/Intermediate/IRprecipitation	IR-only precipitation estimate	mm/hr
Grid/IRkalmanFilterWeight	Grid/Intermediate/IRinfluence	Aggregate influence of the IR precipitation to the multi-satellite precipitation	percent

The index values for *MWprecipSource* are (with 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.

IMERG V07 Data Access and Use

Any advice on what precipitation variable I should use in IMERG V07?

We recommend for general use the "precipitation" field as it represents the complete multi-satellite precipitation estimate with gauge calibration - the main IMERG field to get precipitation estimates. Using other variables is only encouraged for specific analysis such as identifying microwave satellite

source, Quality Index, or probability of liquid phase, among others. Please review Table 5 in the IMERG ATBD for more information on these variables: https://gpm.nasa.gov/resources/documents/imerg-v07-atbd.

How do I update my script to use the new IMERG V07?

At the minimum, the filenames and variable names will need to be updated. In particular, if you use the variable *precipitationCal* in previous versions, it has been renamed *precipitation* in V07.

Are there any new fields/variables, or changes to how existing fields are calculated?

The variable "IRinfluence" has a substantially different definition compared to "IRkalmanFilterWeight". Please see the IMERG V07 ATBD for more information.

Did IMERG V06 and V07 overlap?

Yes. However, IMERG V06 and V07 are sufficiently different that users are discouraged from mixing data from the two versions. Retrospective processing of IMERG V07 Final Run product covers the entire TRMM-GPM record. Please review the GPM Data News website for the latest updates and notices at https://gpm.nasa.gov/data/news.

IMERG V07 and Precipitation Extremes

How will IMERG V07 impact extremes in terms of statistics?

Improvements to the algorithm that are anticipated to affect the estimation of precipitation extremes are listed below. However, extremes have large uncertainties, so improvements may not be obvious in certain cases. See "Appendix 2: Initial Evaluation of IMERG V07" in the IMERG V07 Release Notes for more information: https://gpm.nasa.gov/resources/documents/imerg-v07-release-notes.

a) ***Increased the precipitation rate cap to 200 mm / h*:** This was intended to be implemented in V06 but remnant internal checks capped the maximum rates to 50 mm / h in the IR module and 120 mm / h in the PMW module. These remnant caps were eliminated in V07, allowing the possibility of greater extremes in the IMERG estimates when the retrievals compute them.

b) ***Implemented the SHARPEN scheme*.** The averaging in the Kalman filter distorts the distribution of precipitation rates by reducing extremes and increasing precipitation area. In response, SHARPEN was developed and implemented in V07 (Tan et al. 2021). Effectively, this means that the SHARPEN-adjusted V07 Kalman filter precipitation rates have a distribution and thus extreme fraction comparable with the aggregate of the propagated PMW and the IR inputs, weighted by their Kalman filter contributions.

c) ***Skill in the GPROF V07 retrievals varies by surface type.*** The GPROF V07 algorithm, used to retrieve precipitation from all passive microwave (PMW) inputs for IMERG, has made progress in handling difficult surface types that tend to yield lesser quality results, including frozen surfaces, orographic areas, and coastal zones. It remains to be demonstrated how well the resulting retrievals work in each of these areas, and users should seek confirmation of the values in the corresponding IMERG estimates.

   - *Specifically, IMERG estimates over frozen surfaces should be examined critically.* The IMERG team believes that GPROF retrievals over frozen surfaces are sufficiently advanced that IMERG V07 includes such retrievals (for the first time), effectively providing fully global estimates, but they are flagged with low Quality Index values. Retrievals over Antarctica are particularly problematic due to the very low (and thus hard-to-detect) rates of snowfall. Furthermore, it has been determined that the CORRA

values used to calibrate IMERG are systematically low when the precipitation type is snow.

What is the cap on precipitation rates? Why should users use caution for precipitation rates from IMERG close to 200 mm/ h?

IMERG V07 increased the precipitation rate cap to 200 mm / h.

Extreme precipitation rates are often associated with challenging conditions (heavy attenuation, large ice particles, etc.) that make it difficult to retrieve accurate precipitation rates, resulting in large uncertainties. In these areas it is likely that extreme weather is occurring, but the specific value is not precisely captured by the algorithm and should be taken with a grain of salt. Bias estimates should be treated carefully. Additionally, there are potential spatial errors and lower sample sizes in a priori databases for extremes that can lead to uncertainties. Rain gauges can also have trouble with accurately measuring heavy precipitation rates, leading to further inaccuracy.

Do you anticipate any of the IMERG V07 changes with respect to changing values to the extreme rainfall rates to vary by season or region in any systematic way?

Not specifically with regards to season or region beyond the general improvements described above.

What do users need to consider and know when evaluating precipitation extremes over time?

The changing constellation of microwave radiometers used in a High-Resolution Precipitation Product (HRPP), of which IMERG is one, may introduce artifacts due to different sensors having different performances, and users should be careful when using to study trends and interannual variability.

IMERG V07 and Reliability

Using IMERG V07 data over snow/ frozen surfaces and for snowfall:

Overall, data may be unreliable over frozen surfaces due to challenges in remote sensing retrievals of precipitation, as sensors can misinterpret ice and snow-covered surfaces as precipitation.

Prior to V07, IMERG dropped PMW retrievals for grid boxes with estimated surface type of snow/ice-covered land and sea ice. As such, the IMERG estimates over frozen surfaces were either purely IR-based (in the latitude band 60°N-S) or recorded as missing data (outside the IR coverage). An assessment of the GPROF V07 PMW estimates over frozen surfaces suggested that they may be useful, so they are now included in IMERG V07. The result is that, except for a handful of grid boxes mostly at the poles, IMERG has complete global coverage. Nevertheless, *users should be aware of the diminished performance of estimates at the polar latitudes, exercise appropriate caution in their interpretation of the results, and account for the reduced accuracy of these estimates in their application*s. Accordingly, the half-hourly Quality Index (QI) values for precipitation estimates over frozen surfaces are reduced based on You et al. (2023) to reflect the lower confidence. The monthly QI was not changed because gauges are also less certain in polar regions, but we don't currently have a good basis for changing the gauge coefficients in the monthly QI. Furthermore, it has been determined that the CORRA values used to calibrate IMERG are systematically low when the precipitation type is snow.

Additionally, while users should take caution of precipitation estimates over ice and snow-covered surfaces, IMERG estimates *for snowfall should be examined critically*. As described above, the IMERG team believes that GPROF retrievals over frozen surfaces are sufficiently advanced that IMERG V07 includes such retrievals (for the first time), effectively providing fully global estimates, but they are flagged with low Quality Index values. Retrievals over Antarctica are particularly problematic due to the very low (and so hard-to-detect) rates of snowfall. Recent analysis shows low precipitation rates for snowfall over land due to deficiencies in the detection and rate retrieval by CORRA, which is used to intercalibrate the PMW estimates, and consequently suppresses the apparent skill in the GPROF retrievals.

Using IMERG V07 data over orographic regions:

Greater uncertainty near orographic regions may lead to underestimated precipitation rates. The switch to GPROF V07 to provide retrievals from the various passive microwave sensors takes into account orographic enhancement of precipitation, but how well it does so remains to be verified. Propagation (morphing) and IR estimates do not directly account for orographic enhancement yet.

The GPROF algorithm has been upgraded to Version 07. Among its many improvements, GPROF V07 introduces new surface classes for mountains (with further subclasses for different orographic enhancement conditions), which are expected to improve retrievals over orography. For more information please refer to the GPROF V07 ATBD: https://gpm.nasa.gov/resources/documents/gpm-gprof-algorithm-theoretical-basis-document-atbd

How do you recommend that users verify the quality of data in each of these scenarios?

Compare against reliable ground observations under similar conditions, such as rain gauge data from a country's national weather service. The Quality Index gives an estimate of the reliability of the data. Ground validation data links can be found at https://www.eorc.jaxa.jp/IPWG/calval-links.html (this is not an exhaustive list). Additionally, an example of validation using the multi-radar multi-sensor system (MRMS) precipitation products over the conterminous United States can be found in Wang et al. (2021).

IMERG Quality Index and Handling Errors, Uncertainty and Bias

What is the Quality Index (QI)? Is there a space/time range beyond which the quality index is no longer relevant (e.g., for groups working with long term anomalies or aggregating data beyond specific hours/days)?

QI is an indicator of the expected reliability of the data and is included in the native half-hourly and monthly files under the variable *precipitationQualityIndex*. This can help the user to decide whether they would like to use the data or not for their given application. However, "low-quality data vs. no data" is a decision that only the user can make, and depends on the importance of completeness.

Moreover, remarking on the "so what?" question associated with QI; namely, if a precipitation estimate is classified as "red" (see the table below), what estimate should the user use instead? This is a question that only the user can answer depending on their application. If the application can tolerate some missing information or have alternative sources of precipitation estimates, then the user can choose to discard precipitation estimates of high uncertainty, though they should be aware that this may introduce diurnal sampling bias (more samples at one or a few times of day than at others) given the Sun-

synchronous orbits of most of the satellites. If the application requires complete sampling, then perhaps having low-quality data is better than no data at all. We provide the QI as information on the reliability of the precipitation estimate, and it is up to the user to use this information in a way that is relevant to their application.

There are two types of QI: half-hourly and monthly. Higher half-hourly QI means better performance on average, but this average contains huge variability in performance. The documentation further recommends three discrete categories based on QI ("red", "yellow", "green"):

Half-Hourly QI

0-0.5	"red" questionable quality
0.5-0.75	"yellow" mid-range quality
0.75-1	"green" good-to-excellent quality

Monthly

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
2-4	"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
4+	"green" these are developed areas with good-to-excellent gauge networks

Unfortunately, we cannot provide clear guidance currently on how to aggregate QI to coarser scales.

For more information about QI and use of the QI, see the Appendix in the IMERG V07 ATBD.

How are IMERG V07 uncertainty estimates quantified?

IMERG has the variable "randomError" that quantifies uncertainties, but improving uncertainty estimation on the whole remains an active work in progress. The end goal is to have a quantification of uncertainties (e.g., providing a probability distribution), beyond just a qualitative index (QI). The IMERG QI helps give guidance to users when to most trust the IMERG product.

How does an end user generate error estimates on aggregated time/space scales? Is this needed? If so, why?

This is still a very active area of research. See Guilloteau (2022) for one foray into this challenging subject.

Is there a link to where a user can go and look at characterization of errors as well understand what is known about precision and bias for IMERG estimates?

Some basic information is provided in the IMERG V07 Release Notes. This is an active area of research in the literature, with numerous studies variously based on comparisons to other satellite estimates, surface precipitation data, and comparison to hydrological data. Because IMERG (and satellite retrievals in general) is sensitive to underlying surface, weather regime, and algorithm versions, users should be cautious about drawing conclusions from results that differ in these respects from their particular situation. The IMERG development team is happy to assist users in assessing the applicability of such studies to the particular application at hand. Feel free to reach out to the IMERG development team via the GPM "Ask A Question" page at https://gpm.nasa.gov/data/ask-a-question.

Do application users need to account for error from gauges?

The precipitation gauge analysis that is used in IMERG is produced by the Global Precipitation Climatology Centre (GPCC), located in the Deutscher Wetterdienst, Offenbach am Main, Germany (Becker et al. 2013, Schneider et al. 2014, 2018). Error from the GPCC gauge adjustment gauges may lead to biases in Final. Our use of the GPCC analysis for gauge adjustment is a broad bias correction. Users should consider a "fine-tune" bias adjustment with regional gauges if possible. Users can also consider whether the Climate Hazards center IMErg with Stations (CHIMES, https://www.chc.ucsb.edu/data/chimes) is a product suitable for their purpose. CHIMES combines the IMERG Late Run with low-latency gauges to provide precipitation estimates at a 2-day latency.

If a user finds that IMERG data is overestimating or underestimating, what do you suggest to them to take for next steps?

If users find that IMERG is consistently overestimating or underestimating the precipitation over their region of interest, they are encouraged to perform their own bias adjustment.

Climatological Analysis

What needs to be considered when using IMERGv7 for evaluating trends?

Use of IMERG for trend analysis should be conducted with caution. As mentioned above, IMERG is composed of a changing constellation of sensors and may thus suffer artifacts when old sensors drop out, new sensors come online, and even when satellites drift in their orbits. One of the ways these artifacts manifest is a sudden change in the distribution—especially in quantities beyond mean precipitation rate—coinciding with the removal or introduction of a sensor into the constellation.

In general, any analysis on the long-term changes should use Climate Data Records if possible. For example, the Global Precipitation Climatology Project (GPCP) provides global precipitation estimates with a design priority of consistency across the entire record (for the current version of monthly and daily data, see https://disc.gsfc.nasa.gov/datasets/GPCPMON_3.2/summary and https://disc.gsfc.nasa.gov/datasets/GPCPDAY_3.2/summary, respectively). If HRPPs such as IMERG are used, users should pay particular attention to possible abrupt changes that coincide with changes in the constellation (e.g., as depicted in https://gpm.nasa.gov/resources/images/gpm-constellation-overpass-times), as well as compare multiple products to identify deviations of a single product from the ensemble. Shifts in histograms have been noted at the TRMM orbit boost (August 2001), a switch in the electronics for the TRMM PR (May 2009) the change-over from TRMM to GPM calibration (June 2014), and the GPM orbit boost (November 2023). In general the impact of these data boundaries is more obvious at the daily or sub-daily scale, while at the monthly scale they are less critical.