V05 IMERG Final Run Release Notes

George J. Huffman, David T. Bolvin, Eric J. Nelkin, Erich F. Stocker, Jackson Tan

9 November 2017; Rev. 13 November 2017, 17 November 2017, 1 February 2018

Note: 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. The Early and Late Runs are labeled Version 05B for consistency with the Final, even though no V05A Early and Late were released.

The algorithm for the Integrated Multi-satellitE Retrievals for GPM (IMERG) has now been upgraded to Version 05. The transition to V05 for the IMERG Final Run began 13 November at PPS and the new data started flowing down to the GES DISC as well. However, additional testing showed that there was a problem with the metadata that required us to withdraw the V05A files and re-release the data as Version 05B. This started on 20 November. The V05B data values were unchanged from V05A; only a few metadata items are involved. Access (detailed below) and data fields continue to be the same as for Version 04, except the version number is 05B. The data record begins with 12 March 2014; note that the March 2014 monthly file is based the partial month of data starting with 12 March. Version 05B supersedes all prior IMERG versions, and users are urged to shift to the new datasets as soon as practical.

Changes from Version 04 to 05

- Use GPROF V05 to compute precipitation estimates for all microwave sensors as input.
- Provide GPROF estimates from all constellation members in the microwave-only precipitation field (HQprecipitation) over the fully global domain (90ºN - 90ºS). Note: estimates over snowy/icy surface types are not masked out.
- Include these fully global GPROF estimates in the complete precipitation fields (precipitationCal, precipitationUncal) outside the standard IR domain (60°N-S), although without morphing or IR fill-in. Note: estimates over snowy/icy surface types are masked out.
- Refine gauge error estimates to provide proper weighting when combined with satellite-only estimates.
- Add Quality Index for all 0.5-hour and monthly products. See "IMERG Quality Index" document (*https://pmm.nasa.gov/sites/default/files/document_files/IMERG_QI.pdf*) for a summary of these new fields.
- Continue the practice of thresholding GPROF (now V05) precipitation rates for all input precipitation estimates to adjust fractional coverage; all GPROF estimates are currently thresholded at 0.03 mm/hour.
- Continue the practice started with Version 04 of calibrating 2BCMB to GPCP V2.3 over ocean (at middle and high latitudes) and land (globally) to compensate for low (high) 2BCMB bias over non-tropical oceans (land).
- The Version 05 GPROF estimates for MHS and ATMS do not provide estimates for the 5 and 8 footprints (respectively) at each swath edge for the period starting at the beginning of the GPM era (12 March 2014). This is due to algorithm issues as revealed in early testing.
- The Version 05 GPROF-TMI estimates had not been computed for the GPM era when the IMERG Runs were retrospectively processed, so TMI is not presently included in the V05 IMERG datasets.

- Taken together, the restriction on MHS and ATMS swaths and the lack of TMI somewhat reduce the amount of microwave-based data contained in Version 05 IMERG, compared to previous versions.

Additional Access Information

The data archive sites have now populated the various repositories of original and value-added data with the new Version 05. See *https://pmm.nasa.gov/data-access/downloads/gpm* to find all of the various formats and their locations. Recall that access to the various systems (PPS, PPS near-real time, and GES DISC) requires separate simple, free, and automatic registrations to satisfy NASA data system requirements.

IMERG Final Run data are computed about 2.5 months after observation time, a month at a time. The native HDF5 half-hourly products have the prefix "*3B-HHR*", while the monthly products have the prefix "*3B-MO*". Note that the monthly product is only computed as a native product for the Final Run. The complete file naming convention can be found at

   http://pps.gsfc.nasa.gov/Documents/FileNamingConventionForPrecipitationProductsForGP MMissionV1.4.pdf .

The version number for the revised release is Version 05B. The field named *precipitationCal* contains the "complete" IMERG precipitation estimate.

Additional Notes

Recall that the Final Run uses calibrations based on accumulations of match-ups that include the entire current month and data from previous months, while Early and Late Runs necessarily use calibrations based on trailing accumulations of match-ups, since these cannot be computed into the future. In addition, the Early Run only has forward propagation of the microwave data (unlike both the Late and Final Runs), and neither has calibration to the monthly gauge data as in the Final Run. Both the Early and Late Runs used "seed" calibration files for the Kalman coefficients and the 2BCMB-GMI and HQ-IR calibrations from the respective last runs of V04A on 30 November 2017 to begin Initial Processing on 1 December 2017. Accordingly, users should expect the estimates in December 2017 to be less accurate than following months of data that will have calibrations fully populated with Version 05. By contrast, the Final Run does not require a seed file, since enough data exists in March 2014 to approximately fill the (centered) accumulation files. Validation results will be posted as they are developed.

Complete records of all Runs are posted (honoring the necessary latency), labeled as Version 05B for consistency with the Final, even though no V05A Early and Late were released. It is planned that all three Runs will be retrospectively processed back to the start of the Tropical Rainfall Measuring Mission (TRMM), 1998 (or 2000, depending on some data issues), in Spring 2018.

The Quality Index (QI) document (see below) describes the new QI variables for both half-hourly and monthly data. This includes advice about what might be considered "good" ranges of QI. We encourage users to report their experience with QI (and any other variable) to aid in developing the next versions of IMERG.

Key IMERG Documents

Algorithm Theoretical Basis Document
https://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V4.7.pdf

Quality Index
https://pmm.nasa.gov/sites/default/files/document_files/IMERG_QI.pdf

Technical Documentation
https://pmm.nasa.gov/sites/default/files/document_files/IMERG_technical_doc_11_09_17.pdf

Early Validation Results

Initial Comparison to MRMS-Based Surface Reference (Jackson Tan)

One key validation is to compare IMERG precipitation estimates to the Multi-Radar Multi-Sensor (MRMS) suite of precipitation observations processed in support of the GPM mission for ground validation. MRMS is produced by NOAA/NSSL, based primarily on the WSR-88D network over conterminous US with bias correction using gauge data from the Hydrometeorological Automated Data System and regional rain gauge networks. Only the MRMS pixels with the best quality are selected, requiring

1) a perfect radar quality index, which filters out sampling and estimation uncertainty such as beam issues due to orography and bright band;
2) gauge correction factors within the range [0.5, 2], which removes pixels in which the radar and gauge estimates differ by more than a factor of 2; and
3) only liquid-phase precipitation.

The first evaluation here is conducted with the IMERG Final Run "complete" precipitation field (precipitationCal) for both Version 04 and 05 at the native IMERG resolution (0.1°, 30 min) from January to September 2015. The results are computed for the "best quality" MRMS pixels across all of CONUS, which favors data relatively near radars in flat terrain not involving snowy conditions. The rain/no-rain threshold is chosen to be 0.2 mm/h.
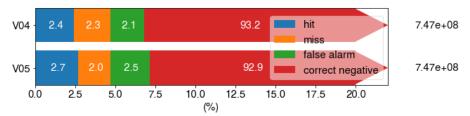

Fig. 1. Hits, misses, false alarms, and correct negatives. The numbers on the right indicate the sample sizes. V05 shows an improvement in the percentage of hits and misses, but with an increase in false alarms and slight decrease in correct negatives.
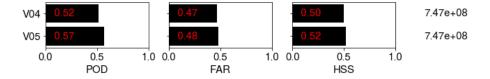
Fig. 2. Probability of detection (POD), false alarm ratio (FAR), and Heidke Skill Score (HSS). The numbers on the right indicate the sample sizes. There is a clear improvement from V04 to V05 in POD and HSS, though with a marginal increase in FAR.
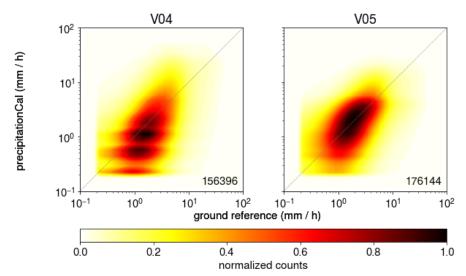

Fig. 3. Joint distributions of the rain rates between IMERG and the MRMS-based ground reference, limited to cases where both meet the 0.2 mm/h threshold. The number in the lower-right corner of each plot shows the normalization factor for the color scale. There is a distinct improvement from V04 to V05, with (i) the elimination of horizontal bands, (ii) a reduction in the underestimation of low rain rates ($\leq$ 1 mm / h), and (iii) a slight reduction in the overestimation of high rain rates ($\geq$ 3 mm / h).
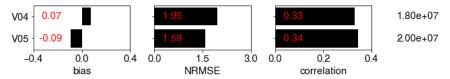

Fig. 4 Bias, normalized root mean square error (NRMSE), and Pearson correlation coefficient for cases where both IMERG and the MRMS-based ground reference meet the 0.2 mm/h threshold. The numbers on the right indicate the sample sizes (number of hits). Between V04 and V05, there is a slight increase in the magnitude of the bias, though with an opposite sign, a distinct reduction in NRMSE, and a marginal increase in the correlation. There is also a higher number of hits (for this threshold) in V05.
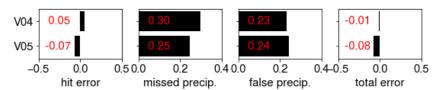

Fig. 5. Hit error, missed precipitation, false precipitation, and total error, normalized against the total rain amount in the MRMS-based ground reference. There is a slight increase in the magnitude of hit error (but with opposite sign) and false precipitation in V05, but a distinct

improvement in the reduction of missed precipitation. Overall, however, the magnitude of the total error increased.

Next, we repeat the evaluation for the same conditions, namely at the native IMERG resolution (0.1°, 30 min) from January to September 2015 using the "best quality" MRMS pixels across all of CONUS, with a rain/no-rain threshold set to 0.2 mm/h, but with the IMERG Final Run "complete" precipitation field (precipitationCal) for both Version 04 and 05 categorized by the sensor or approximation that is the source for each particular value (inferred using the HQprecipitation and IRkalmanFilterWeight data fields). Each of these figures shows that in V05, TMI is absent and the cross-track scanners (MHS and ATMS) have reduced sampling, and most of these cases transfer to the IR+morph category.
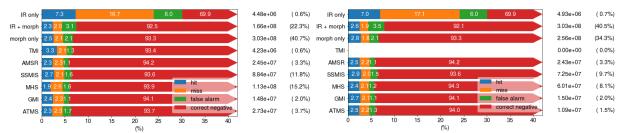

Fig. 6. Hits, misses, false alarms, and correct negatives by source for V04 (left) and V05 (right). The differences are sufficiently small between V04 and V05 for most sources that it is easier to discuss them in the following figures.
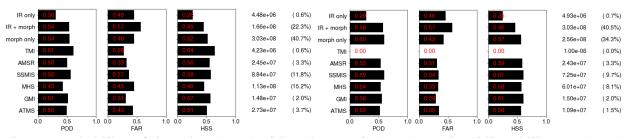

Fig. 7. Probability of detection (POD), false alarm ratio (FAR), and Heidke Skill Score (HSS) by source for V04 (left) and V05 (right). There is generally improvement for most of the sources, with SSMIS being more noticeable.
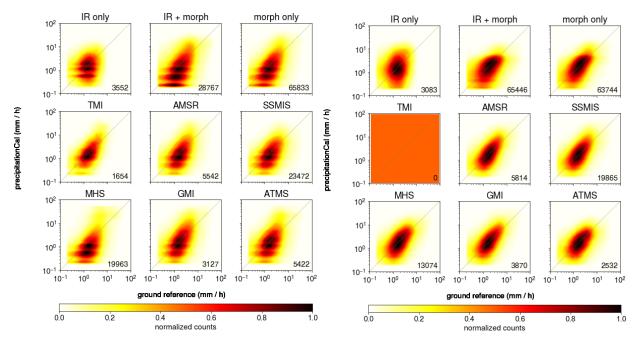
Fig. 8. Joint distributions of the rain rates between IMERG and the MRMS-based ground reference, limited to cases where both meet the 0.2 mm/h threshold by source for V04 (left) and V05 (right). The number in the lower-right corner of each plot shows the normalization factor for the color scale. The overall improvement from V04 to V05 described in Fig. 3 largely applies to each source, although the IR has a longer "tail" of low values.
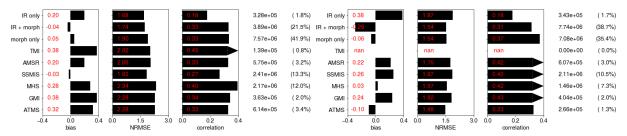

Fig. 9 Bias, normalized root mean square error (NRMSE), and Pearson correlation coefficient for cases where both IMERG and the MRMS-based ground reference meet the 0.2 mm/h threshold by source for V04 (left) and V05 (right). The bias is smaller for most sources (changing sign for some), but the IR+morph and morph-only have such a large proportion of cases that their larger sizes drives the overall bias shown in Fig. 4. The overall reduction in NRMSE is true for nearly every source, and correlation increases across most sources.
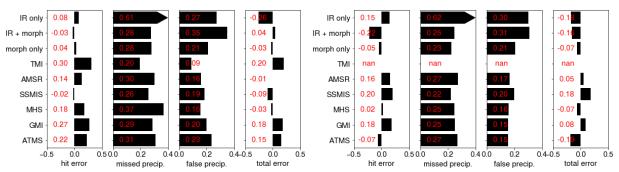
Fig. 10. Hit error, missed precipitation, false precipitation, and total error, normalized against the total rain amount in the MRMS-based ground reference by source for V04 (left) and V05 (right). As with bias in Fig. 9, the change in IR+morph and morph-only hit error and total error dominate the corresponding values averaged for all of IMERG shown in Fig. 5.

To get a sense of the large-area performance, we compare IMERG Final Run against the MRMS-based ground reference over a region in the southeastern US where MRMS is highly reliable due to excellent radar coverage, high density of gauges, and an absence of significant orography. Here, we compute the daily-averaged precipitation rates over the area bounded by 30–40°N, 93–83°W and examine the time series for V04 and V05, as well as for the gauge-calibrated and uncalibrated V05 estimates.
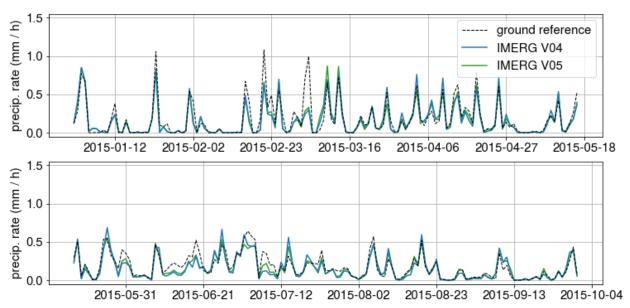

Fig. 11: Time series of the ground reference, IMERG Final Run V04 precipitationCal, and V05 precipitationCal over the southeastern US. V04 and V05 tend to be closer to each other than to the observations when significant differences occur, although it appears that V05 is generally an improvement over V04.
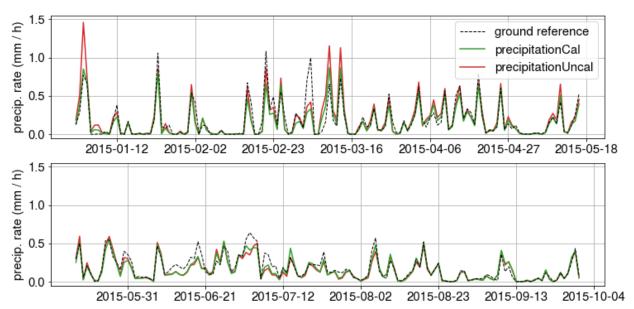
Fig. 12: Time-series of the ground reference, IMERG Final Run V05 precipitationCal, and V05 precipitationUncal over the southeastern US. It is clear that the gauge adjustment, which is applied at the monthly time scale, generally improves agreement with the ground reference.

Additional V05 validation results are being computed, and will be posted here as they are developed.