Day 1 IMERG Final Run Release Notes

George J. Huffman, David T. Bolvin, Eric J. Nelkin

20 January 2015, revised 31 January 2015

The Integrated Multi-satellitE Retrievals for GPM (IMERG) Final Run, which is the post-real-time research product in the IMERG suite of products, is now available at PPS at *https://storm-pps.gsfc.nasa.gov/storm* and in the next few days from GES DISC at the DOIs

half-hourly	10.5067/GPM/IMERG/HH/3B
monthly	10.5067/GPM/IMERG/MONTH/3B

In the case of PPS, note that you will need to be a registered user to access the data. Register online at *http://registration.pps.eosdis.nasa.gov* (contact *helpdesk@pps-mail.nascom.nasa.gov* with questions). This simple, free, and automatic process satisfies NASA data system requirements.

IMERG is presently available for the period mid-March 2014 to the present (delayed by about 3 months), which is the most current data given the latency of the ECMWF input data required for the production runs. The half-hourly products have the prefix "*3B-HHR*" and the monthly products have the prefix "*3B-MO*". The complete file naming convention can be found at

http://pps.gsfc.nasa.gov/Documents/FileNamingConventionForPrecipitationProductsForGPMMissionV1.4.pdf .

The half-hourly data start on 12 March. Due to the reduced sampling, it is recommended that users ignore the March monthly IMERG estimates (i.e., the "-MO" file for March).

The version number for the initial release is Version 03D. An initial posting on 15 January 2015 was recalled when we found a minor mismatch between the metadata and "missing" values. The fields named *precipitationCal* and *precipitation* contain the "complete" IMERG precipitation estimate for the half-hour and monthly files, respectively.

1. Introduction to IMERG

IMERG is the Day-1 U.S. multi-satellite algorithm for GPM, based on components from three prior multi-satellite algorithms from NASA (TRMM Multi-Satellite Precipitation Analysis; TMPA), NOAA (CPC Morphing – Kalman Filter; CMORPH-KF), and University of California Irvine (Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks – Cloud Classification System; PERSIANN-CCS), developed with varying degrees of GPM and other funding. [The corresponding Japanese algorithm is Global Satellite Map of Precipitation (GSMaP).] The current best reference for IMERG is the IMERG Algorithm Theoretical Basis Document (ATBD; Huffman et al. 2014), accessible at

http://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V4.4.pdf .

The technical document (Huffman et al. 2015) is posted at

http://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf .

In brief, the input precipitation estimates computed from the various satellite passive microwave sensors are intercalibrated to the GPM Combined Instrument product (because it is presumed to be the best snapshot GPM estimate), then "morphed" and combined with microwave precipitation-calibrated geo-IR fields, and adjusted with monthly surface precipitation gauge analysis data (where available) to provide half-hourly and monthly precipitation estimates on a 0.1° lat./long. grid over the domain 60°N-S. Precipitation phase is diagnosed using analyses of surface temperature, humidity, and pressure. The current period of record is mid-March 2014 to the present (delayed by about 3 months).

One key fact about all GPM algorithms is that the Science Team took very seriously the notion that a new mission provided the perfect opportunity for a zero-base review of all products. Every algorithm is substantially different than its predecessor in TRMM, and consequently the "Day-1" designation is a red flag to the user that the results will be rough around the edges on this first pass. In the case of IMERG, the changes include input precipitation estimates that are Day-1 GPROF2014 for microwave imagers and for microwave sounders, and Day-1 GMI/DPR combined precipitation estimates. As well, Day-1 IMERG is the first integration of the code components from TMPA, CMORPH-KF, and PERSIANN-CCS. The intercalibration of microwave sensors is a two-step process in IMERG, namely climatologically between GPM Microwave Imager (GMI) and all partners, and then by rolling 45-day histogram matching between GPM Combined Instrument (GCI) and GMI. We have chosen to make the Day-1 GMI-partner calibration a pass-through (i.e., with no calibration), since the GPROF-imager flavors are already supposed to be consistent with GMI. On the other hand, the GPROF-sounder flavors are known to be somewhat different, so an ad hoc climatological GMI-partner calibration, based on a limited sample of several months, is computed for sounders.

A few users were given very early test versions of IMERG files for the month of June 2012, which used legacy input data from the TRMM era and contained fields and metadata that were in use at the time. Since then, several modifications have been made to the HDF5 files. We dropped the notion of averaging multiple overpasses by microwave sensors in a half-hour interval. Accordingly, the fields HQprecipSource2 and HQobservationTime2 were dropped and the "1" was dropped from HQprecipSource1 and HQobservationTime1. As well, fields and metadata values were made no shorter than two bytes. The metadata were upgraded to be more compatible with standard application packages, including providing vectors of latitude and longitude, with names "lat" and "lon". These changes won't solve all the format problems that some applications encounter – users can get a sense of alternative formats that should gradually become available after IMERG is released by browsing through the datasets listed in

http://pmm.nasa.gov/data-access/downloads/gpm .

As well, the beta test for IMERG revealed a problem with the date/times that was corrected. The complete current file specification for the IMERG products can be found at

ftp://gpmweb2.pps.eosdis.nasa.gov/pub/GPMfilespec/filespec.GPM.V1.pdf .

Note that PPS provides subsetting by parameter on the *storm* access system, so it is possible for users to reduce the volume of data that they download. On the other hand, users should consider whether fields in the IMERG files that they initially consider unnecessary might, in fact, subsequently prove to be interesting, such as the diagnostic precipitation phase.

2. Performance and Known Issues

The preliminary analysis during beta testing may be summarized as follows:

a) The half-hourly maps are appealing and show reasonable continuity. As an example, a 3-day movie is posted at *ftp://meso.gsfc.nasa.gov/agnes/huffman/3IMERGHH_20140601-03_3vid.qt* . In large part, this performance reflects the forward-backward morphing scheme, as well as the use of a consistent algorithm for all input data sets. On the other hand, we do see "flashing" as the different sensors give different views of the same precipitation systems, particularly at high latitudes. What remains to be demonstrated at this fine scale is the fidelity of the resulting statistics, including fraction of the time precipitating and the precipitation rate histogram. We believe that continued algorithm development for GPROF2014, as well as the future use in IMERG of proper climatological calibrations of partner imagers and sounders to GMI, will improve the agreement between successive satellite overpasses and provide more coherent precipitation estimates.

b) At the monthly scale (Fig. 1), IMERG (top) has a close resemblance to the monthly TRMM Multi-satellite Precipitation Analysis, product 3B43 (bottom), for June 2014. [IMERG is averaged up to the 0.25° lat./long. native grid for 3B43.] Over land the agreement is not surprising; the monthly pattern is strongly influenced by the precipitation gauge analysis, the same analysis is used in both (from the Global Precipitation Climatology Centre), and the same satellite-gauge combination procedure is used, although at the different resolutions of the final products.

c) On the other hand, IMERG is smoother than 3B43 over ocean, mostly at higher latitudes. This was a goal for IMERG, which provides estimates every half hour, versus the 3-hourly interval for the satellite data contributing to 3B43.

d) In both maps, land and coastal regions show blockiness in some areas, such as the southern coast of Jamaica, the eastern coast of Brazil, and the southwestern coast of Australia. This appears to happen when there are mismatches between gauge and satellite mean values, and points to the need for a more careful treatment of the satellite-gauge combination.

e) The monthly high-latitude IMERG precipitation over oceans displays some minor blockiness and some latitudinal noise; this is perhaps a result of mismatches in the spatial scales of the various calibrations currently applied. For the most part this is at latitudes outside the 3B43 domain of 50°N-S.

f) Both products display "picture frames" of low values in some near-coastal oceanic areas, including the eastern coast of Africa and the southern coast of Madagascar. This largely reflects the performance of GPROF2014 and its processor version for imagers.

g) Both products occasionally show anomalously high values over some inland water bodies, again due to GPROF performance.

h) The monthly IMERG precipitation field has values in some areas above latitude 60°N (masked out in Fig. 1). These values are strictly precipitation gauge analysis, lacking satellite data. For this release we haven't restricted the latitude range of the output to 60°N-S, but these values should not be used. On the other hand, the fields related to the microwave data, including HQprecipitation, HQprecipSource, and HQobservationTime, are meaningful over the entire global domain.

i) Considering a somewhat arbitrarily chosen day from the middle of the month (June 2014) shown in Fig. 1, Fig. 2 compares the daily accumulation of IMERG and 3B42 (the 3-hourly counterpart to 3B43) over the coterminous U.S. (CONUS) to the NOAA Multi-Radar Multi-Sensor (MRMS) analysis. IMERG and MRMS are averaged up to the 0.25° lat./long. native grid for 3B42. The rain band stretching from Wisconsin to Kansas is better depicted by IMERG, although it overdoes the feature in northern Minnesota. In some cases the two satellite schemes look more like each other than the MRMS, including Montana and Maine. As well, both satellite maps display the major advantage of such data sets: the precipitation system over the Atlantic doesn't end at the edge of the collection of radar umbrellas used in MRMS.

j) Fig. 3 compares the three data sets as PDFs and CDFs of 3-hourly, 0.25°-gridbox averages for all available data in CONUS in June 2014. As expected, IMERG is considerably closer to MRMS for precipitation occurrence, and more modestly better for precipitation volumes.

k) This improvement at the fine spatial scale carries over to improved performance by IMERG in 3-hourly, CONUS-wide averages. Fig. 4 shows the scatter plots for all three data sets over the 4-month period April-July 2014. Most statistics improve, notably the RMSE and correlation, but IMERG shows the same tendency as 3B42 to have a high bias for high precipitation rates. This is consistent with the fact that both satellite products have higher occurrences of high-end rain rates than MRMS in Fig. 3. It remains for detailed validation to demonstrate the degree to which these improvements arise from improvements in the input data set retrievals, compared to improvements in the multi-satellite merger procedures.

A few additional comments are in order based on IMERG's structure:

- Calibrations tend to be monthly (or longer) to ensure stability, but fast-changing weather patterns might introduce variations that such calibrations cannot well represent.
- The morphing scheme used to time-interpolate between microwave sensor overpasses essentially uses linear interpolation. Particularly as the time between overpasses becomes more than about two hours, the morphed and actual evolution of precipitation can diverge significantly, and the use of IR data in these cases introduces lower-quality data.
- The IR-based displacement vectors used in morphing are not guaranteed to reproduce the actual motions of precipitation systems. This is certainly true at scales below about 2.5° of latitude/longitude and when the motions of IR-sensed cloud tops are not well-coupled to the motions of lower-altitude precipitation features.
- Morphing will propagate the error in a microwave retrieval along with the "good" precipitation signal, creating correlated error at nearby times.

Other cautions arise from the input data used:

- Current-generation microwave algorithms work best over tropical ocean and progressively less well over mid-latitude ocean, tropical land, mid-latitude land, regions of complex topography, and regions with frozen surface.
- Over land the liquid-sensing capabilities of the emission signal in the lower-frequency channels cannot be used in current-generation algorithms, reducing the potential skill of algorithms.
- The retrievals make estimates over many snowy or icy surfaces, but are necessarily lower in skill in those regions.

- In regions of complex terrain a variable amount of precipitation enhancement and suppression take place in the liquid phase. Such variations cannot be detected by the high-frequency channels, so skill will be lower in these areas.

The error estimates contained in the IMERG Day-1 datasets are a first approximation and should be used with caution. Development work being done in the GPM science team and elsewhere holds promise for upgrades to error estimation in later releases. In common with all fine-scale precipitation datasets, the validation statistics for IMERG are relatively modest at full resolution. Averaging to progressively larger time/space data cubes results in improved error characteristics. We choose to release relatively uncertain estimates at these fine scales to allow the users to craft averages appropriate to their particular studies. This can include, for example, streamflow in sufficiently large basins, which implicitly average the data.

3. IMERG Timeline

- The first IMERG Final Run data sets (for the GPM era, March 2014-present, delayed by 3 months) are released as of 15 January 2015.
- The first IMERG Early and Late Run data sets (again, for the GPM era, but with 4- and 12-hour delay) are planned to be released in late February 2015. Note that all IMERG runs have the same file structure and format, so switching from one run to another is merely a matter of specifying the appropriate file name.
- The first retrospectively processed TRMM/GPM-era IMERG data sets (March 2000-present) are planned to be released in early 2016. The goal is to start at the beginning of 1998, but at the present the appropriate geo-IR data are not available before mid-February 2000. This issue affects all runs, including the Final, and it's being worked.

4. Request for Feedback

We anticipate that in using the data you will find "features", errors, and messy details that we won't necessarily have a chance to find right away, and we look forward to hearing what you find.

5. References

Huffman, G.J., D.T. Bolvin, D. Braithwaite, K. Hsu, R. Joyce, P. Xie, 2014: Algorithm Theoretical Basis Document (ATBD) Version 4.4 for the NASA Global Precipitation Measurement (GPM) Integrated Multi-satellitE Retrievals for GPM (I-MERG). GPM Project, Greenbelt, MD, 30 pp. *http://pmm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V4.4.pdf* .

Huffman, G.J., D.T. Bolvin, E.J. Nelkin, 2015: Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. NASA/GSFC Code 612, 47 pp. *http://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf* .

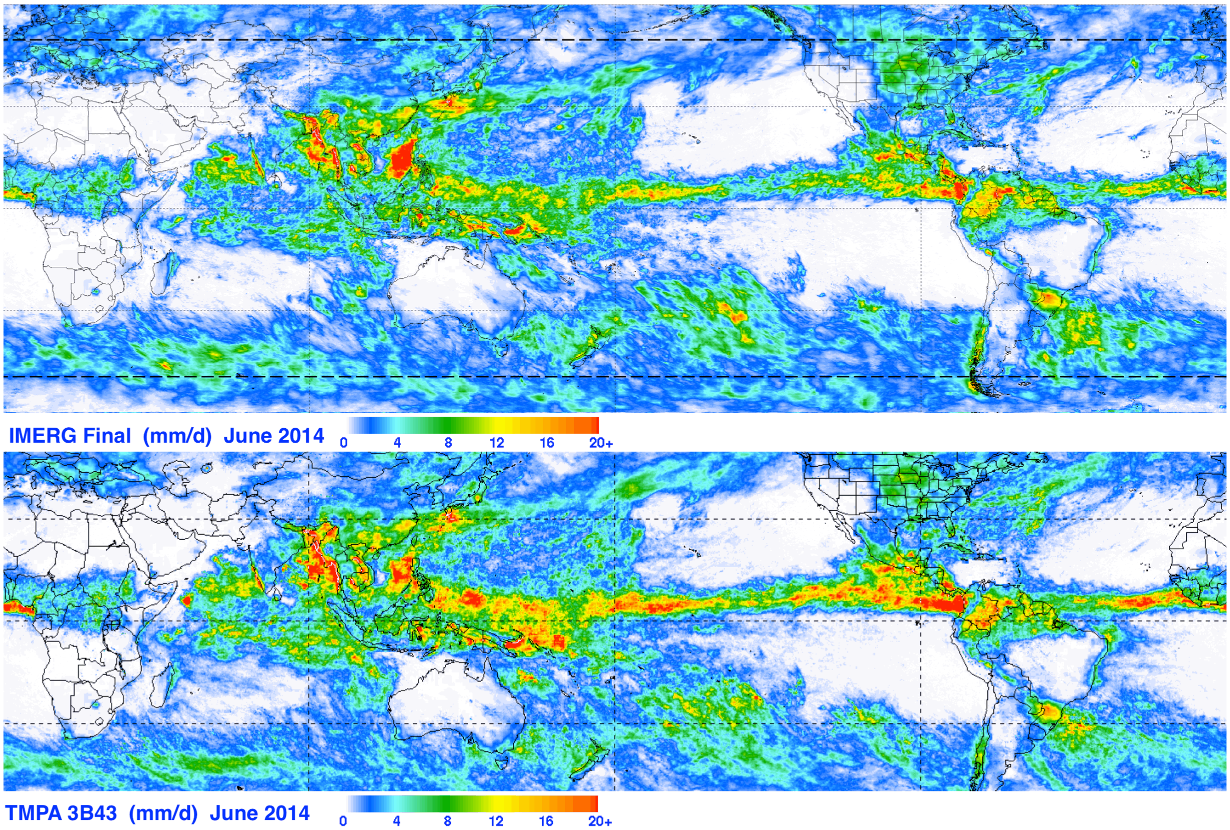

Fig. 1 Version 03D IMERG Final Run (top) and Version 7 3B43 (bottom) monthly precipitation fields (in mm/d) for June 2014. The long-dash black lines on the IMERG field show the latitude boundaries for the 3B43 field (at 50°N and 50°S).

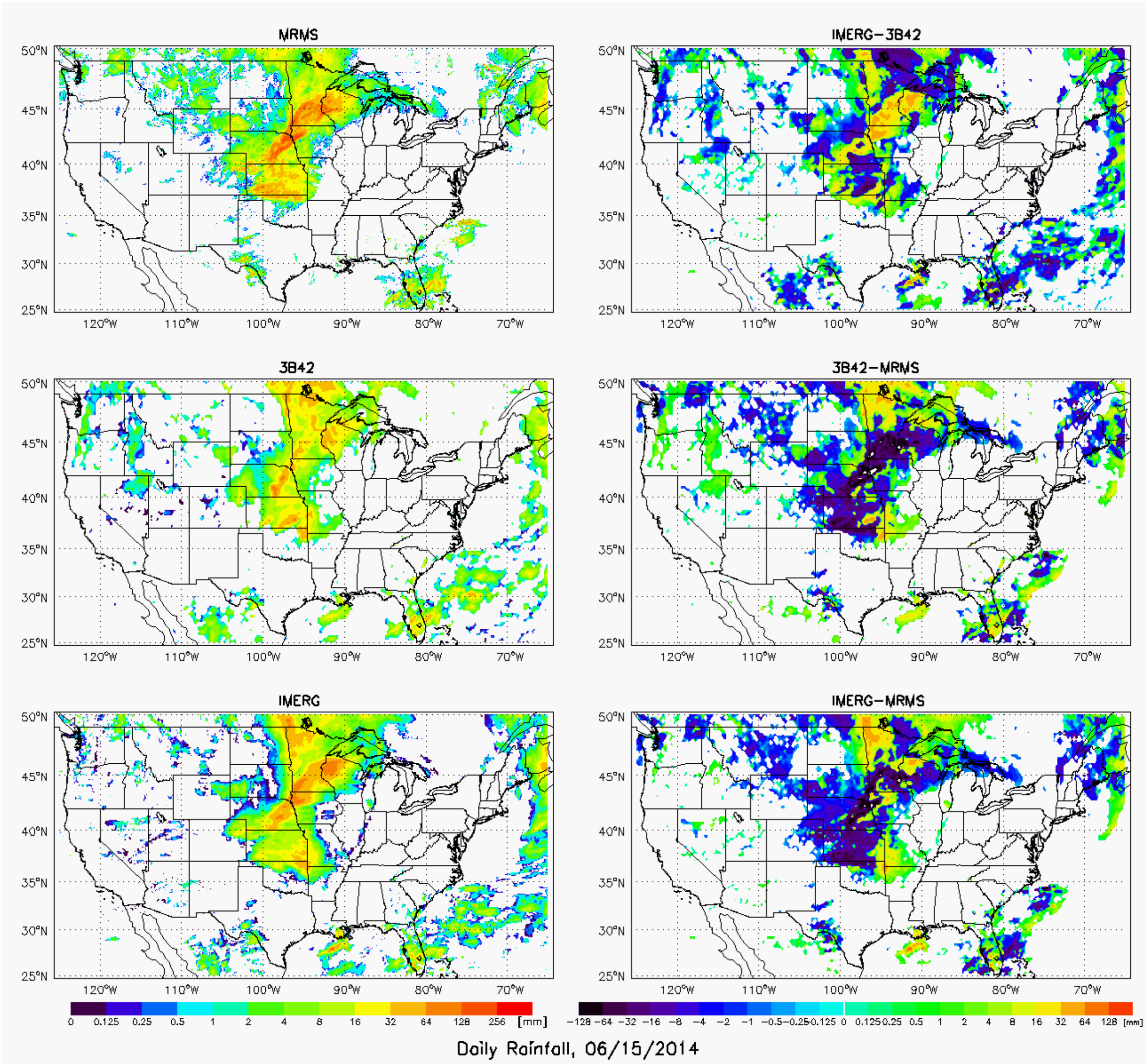

Fig. 2 Maps comparing NOAA MRMS, V.7 TRMM 3B42, and V.03D GPM IMERG for the UTC day 15 June 2014 (with the 90-minute 3B42 offset) over CONUS. All data are averaged to 0.25° lat./long. in units of mm/d. [Courtesy J. Wang (SSAI; NASA/GSFC 612).]

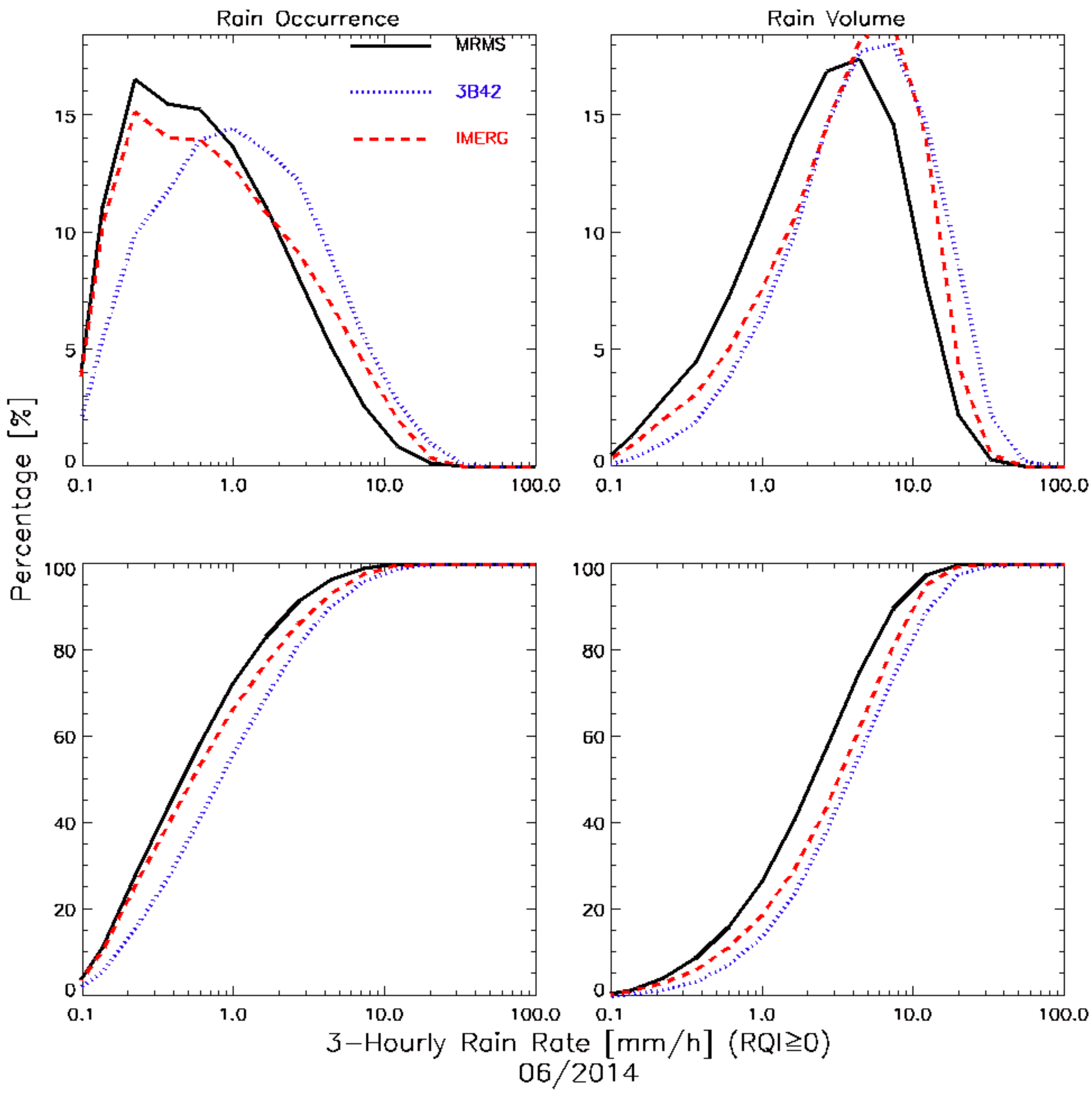

Fig. 3 PDF's and CDF's comparing NOAA MRMS, V.7 TRMM 3B42, and V.03D GPM IMERG for the UTC day 15 June 2014 over CONUS. All data are averaged to 0.25° lat./long. for each day (using the 3B42 definition), the X axis is in mm/h (logarithmic scale), and the Y axis has units of percent. [Courtesy J. Wang (SSAI; NASA/GSFC 612).]

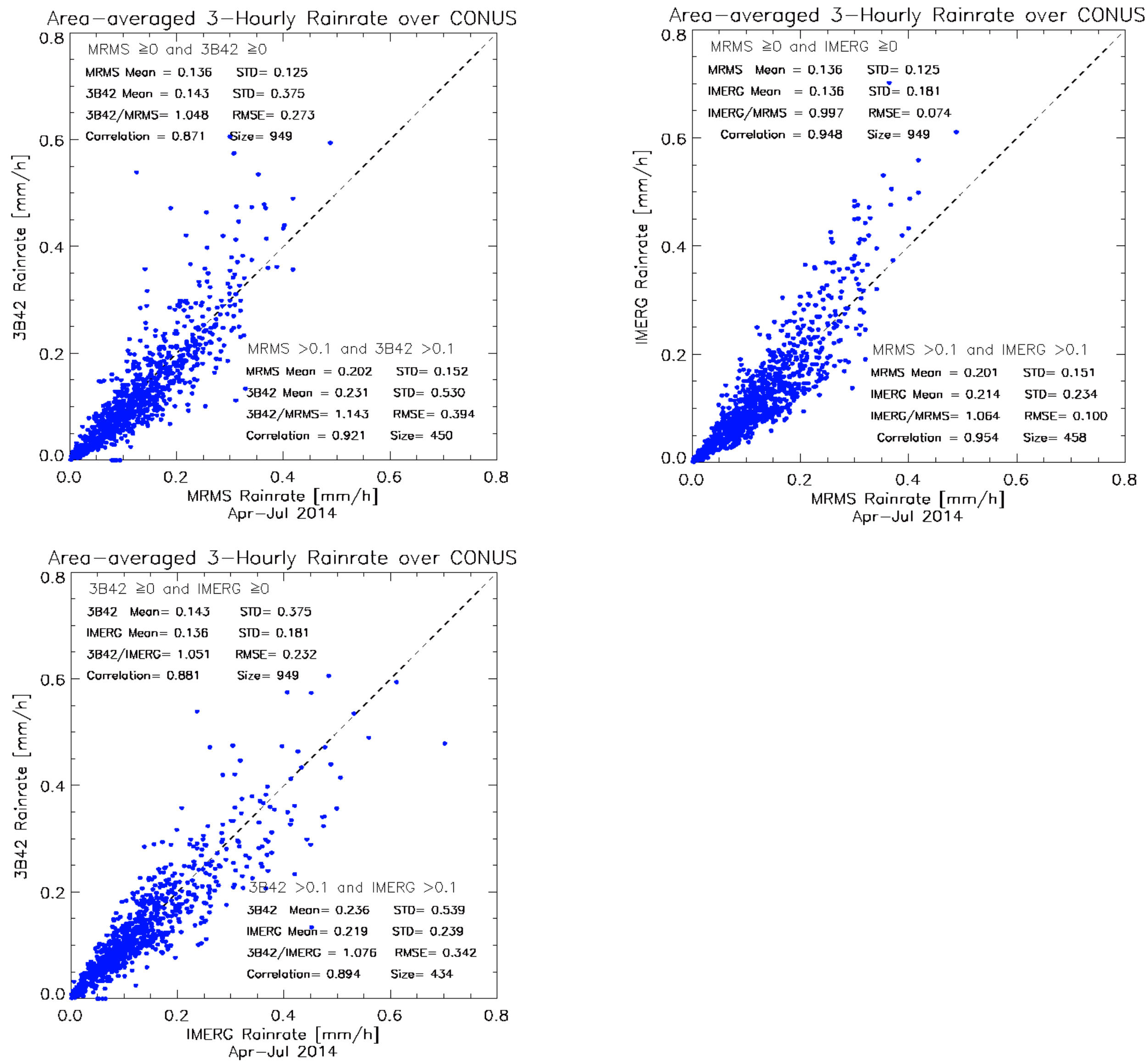

Fig. 4 Scatterplots comparing 3-hourly V.7 TRMM 3B42 to NOAA MRMS (top left), V.03D GPM IMERG to NOAA MRMS (top right), and V.7 TRMM 3B42 to V.03D GPM IMERG (bottom left) for all available periods in the months April through July 2014. All data are averaged over CONUS, and the IMERG and MRMS data are approximately averaged to the 3B42 3-hourly intervals. All rainrate units are mm/h. [Courtesy J. Wang (SSAI; NASA/GSFC 612).]