*Article*

# Validation of Satellite-Based Precipitation Products from TRMM to GPM.

**Jianxin Wang[1,2], Walter A. Petersen[3] and David B. Wolff[4]**

[1] Science System and Applications, Inc., Lanham, Maryland
[2] NASA Goddard Space Flight Center, Greenbelt, Maryland
[3] NASA Marshall Space Flight Center, Huntsville, Alabama
[4] NASA Wallops Flight Facility, Wallops Island, Virginia
\* Correspondence: Jianxin.wang@nasa.gov

**Abstract:** The Global Precipitation Measurement mission (GPM) has been in operation for seven years and continues to provide a vast quantity of global precipitation data at finer temporospatial resolutions with improved accuracy and coverage. GPM's signature algorithm, the Integrated Multi-Satellite Retrievals for GPM (IMERG) is a next-generation of precipitation product expected for wide variety of research and operational applications. This study evaluates the latest version (V06B) of IMERG and its predecessor, the Tropical Rainfall Measuring Mission (TRMM) Multi-satellite Precipitation (TMPA) 3B42 (V7) using ground-based and gauge-corrected multi-radar multi-sensor system (MRMS) precipitation products over the conterminous United States (CONUS). The spatial distributions of all products are analyzed. The error characteristics are further examined for 3B42 and IMERG in winter and summer by an error decomposition approach which partitions total bias into hit bias, biases due to missed precipitation and false precipitation. The volumetric and categorical statistical metrics are used to quantitatively evaluate the performance of the two satellite-based products. All products show a similar precipitation climatology with some regional differences. The two satellite-based products perform better in the eastern CONUS than in the mountainous Western CONUS. The evaluation demonstrates the clear improvement in IMERG precipitation product in comparison with its predecessor 3B42, especially in reducing missed-precipitation in winter and summer, and hit bias in winter, resulting in better performance in capturing lighter and heavier precipitation.

**Keywords:** GPM; TRMM; IMERG; 3B42; MRMS; GPCC; precipitation; bias

## 1. Introduction

Reliable precipitation data are critical for a wide variety of applications such as water budget studies, and prevention or mitigation of natural hazards caused by extreme precipitation events. Precise precipitation measurements are always a challenge because of its large spatiotemporal variability, as well as inherent errors of various measuring instruments. Traditional rain gauges provide direct rainwater measurements, and often serve as the reference for validation of radar- and satellite-based precipitation products[1-5]. However, gauges can only make what are essentially point-measurements at a specific site. The areal distribution of rain gauge networks is usually sparse, irregular, incomplete, and therefore insufficient for accurately describing the spatial variability of precipitation[1,3]. Ground-based weather radars estimate precipitation from reflectivity measurements over relatively large areas. The implementation and retrofitting of dual-polarization to many new weather radars leads to a more accurate precipitation estimation [6,7]. However, the radar networks are mostly deployed over continents only, not dense enough over most parts of the world over land, and provide little coverage over oceans. In addition, the radars suffer from various sources of problems, such as beam blockage in mountainous regions[1, 5]. In contrast, remote sensors from earth observation satellites have been utilized to address the problem of comprehensive precipitation coverage because of their wide and continuous coverage of regional and global precipitation[1,8].

Since the launch of the Tropical Rainfall Measuring Mission (TRMM) satellite, satellite remote sensing-based precipitation products have matured [8]. Among them, TRMM Multi-satellite Precipitation Analysis products (TMPA)[9], and its global successor - the Integrated Multi-satellitE Retrievals for Global Precipitation Measurement (GPM; IMERG)[10], have been widely used due to their relatively high accuracy and applicability [9-13]. However, satellite-based products are not immune from various sources of systematic biases and random errors [12-15]. The evaluation of satellite precipitation products is vital to identify the strength, as well as weakness, improve the accuracy, and ultimately advance their use for a variety of applications.

NASA GPM ground validation (GV) routinely conducts verification for various satellite-based products. Petersen et al. [12] provided an overview of the structure and activities associated with the NASA GV efforts, as well as examples of IMERG verification. As a spin up to full GPM GV efforts, Wang et al (2014) compared Version-7 TMPA 3B42 and nearly one-month pre-launch "test-version" IMERG rain estimates with ground radar product, and found three products were generally in good agreement [16]. Gebregiorgis et al. [17] evaluated the Version 3 (Day 1) Late Run IMERG product and highlighted the improvements of IMERG upon the real-time TMPA product with reduced missed-rain and false-rain bias, and increased hit rate. A number of these evaluations have found the overall improved performance of GPM IMERG products over the TMPA 3B42 predecessor [18-22].

Importantly, the IMERG performance has also been noted to vary significantly spatially and seasonally[20-27]. For example, Chen and Li [23] found that the IMERG performance was poor in winter, especially over North China. Murali Krishna et al. [24] showed that both IMERG and 3B42 products underestimated the observed rainfall over the Indian subcontinent. More recently, Moazami and Najafi [26] demonstrated that IMERG performed better over most parts of interior plains across Canada, but overestimated the moderate to heavy precipitation events. More comprehensive validation is still necessary to better understand regional and seasonal uncertainties in satellite-based precipitation products.

The TMPA products ended on 31 December 2019 after its successor product IMERG is retrospectively processed to the start of TRMM [28]. The time overlap for both products provides an opportunity to investigate the performance of the new IMERG products compared to the TMPA. Since the IMERG V06B were newly released recently, few studies have used it to conduct the research for the transition in multi-satellite products from TMPA to IMERG, especially at long-term and in different seasons.

This study describes one of the NASA GPM GV efforts. The final version of TRMM 3B42 (V7) and latest version of GPM IMERG (V06B) are evaluated using the ground-based and gauge-corrected multi-radar multi-sensor system (MRMS) precipitation products filtered with radar quality index (RQI) over the conterminous United States (CONUS). Because of its diverse physiographic and meteorological features, as well as relatively dense and well-maintained gauge and weather radar networks, CONUS is one of the best regions in the world for the satellite precipitation product validation.

The paper is structured as follows. Section 2 describes the study domain, and satellite-based and ground-based products. Section 3 presents the methodology and statistical metrics. In Section 4, comparisons are presented through four subsections. The discussion and conclusions are provided in Section 5.

## 2. Data

### 2.1. Study Domain

Figure 1a shows the study domain with land and ocean mask. We limited the study domain to land only because of a lack of rain gauges and radar coverages over the ocean, dwindlement of radar data utility that rapidly dwindles with range along the coastal oceans (Figure 2), as well as the significant errors in satellite retrievals due to beam filling effects of part land and part water. A land/ocean mask (Figure 1a) at 0.25$^\circ$ resolution is constructed from the GPM Microwave Imager footprint surface type data over the latitude/longitude box region enclosed by [125$^\circ$W – 65$^\circ$W, 25$^\circ$N – 50$^\circ$N]. Here we define "land" as lands, lakes, rivers and other inland waters. All others are defined as "Ocean", and excluded in the analysis.

Climatologically, precipitation features in CONUS vary with latitude and topography. Generally, it is wetter in the Southeast and along West Coast, but drier in the north and west. Based on Köppen-Geither climate

125 classification, CONUS is typically classified into five main climate regions: A
126 - Tropical Climates, B - Dry Climates, C – Moist Subtropical Mid-Latitude
127 Climates, D - Moist Continental Mid-Latitude Climates, and H – Highlands
128 (Figure 1b).

129
130

**(a)**                                                                                                         **(b)**
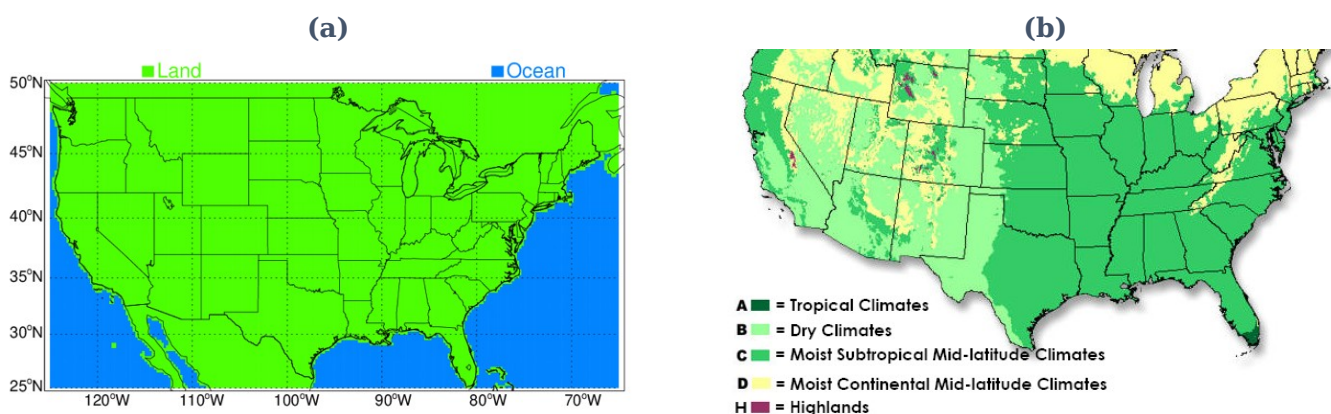
131 **Figure 1.** (a) Study domain with land/ocean mask. (b) Major Köppen-Geiger climate zones in CONUS
132 (figure from https://www.weather.gov/jetstream/climates)

133

134 *2.2. TRMM 3B42*
135       TMPA is a 3-hourly precipitation product with a spatial resolution of
136 $0.25^\circ$ in latitude and longitude, covering from $50^\circ$ S to $50^\circ$ N [9]. The product
137 record started from 1 January 1998 when TRMM data became available, and
138 for continuity, continued into the GPM era even after TRMM terminated, and
139 ended on 31 December 2019 when IMERG product was regarded as its
140 ~~substitude~~substitute [28].
141       TRMM 3B42 combines measurements from different platforms onboard
142 multiple satellites. The passive microwave precipitation estimates are inter-
143 calibrated to the TRMM Combined Instrument (TCI) product, and finally
144 adjusted with rain gauge data on monthly basis. The intercalibration was
145 changed from TCI to a climatological calibration starting with the data from

146 Oct 2014 because of the end of the TRMM satellite mission. This created at
147 least a slight inhomogeneity, primarily over the oceans[28]. The
148 inhomogeneity over land should be minimized by the gauge adjustment [29].

149 *2.3. GPM IMERG*

150 As a global successor to TMPA, the newly released Version 06B IMERG
151 is a multi-satellite precipitation product at fine spatial (0.1°) and temporal
152 (0.5-h) resolution with global coverage. The National Aeronautics and Space
153 ~~Administrtion~~Administration (NASA) Precipitation Processing System
154 (PPS) generates IMERG by combining passive microwave (PMW) and
155 infrared (IR) observations from the GPM constellation satellites, and
156 calibrated with gauge analysis of the Global Precipitation Climatology
157 Centre (GPCC) [30]. IMERG is a unified U.S. national algorithm [10] which
158 combines the strengths of several products including: (1) NASA TMPA for
159 inter-satellite calibration and gauge adjustment [9], (2) the National Oceanic
160 and Atmospheric Administration (NOAA) Climate Prediction Center (CPC)
161 morphing technique with Kalman filter (CMORPH-KF) for time interpolation
162 of PMW-based estimates [31], and (3) the University of California (Irvine)
163 Precipitation Estimation from Remotely Sensed Information using Artificial
164 Neural Networks– Cloud Classification System (PERSIANN-CCS) for
165 retrieval of microwave calibrated IR estimates [32]. One key difference
166 between 3B42 and IMERG is that CMORPH-KF is applied in IMERG to
167 minimize the use of low quality IR estimates.

168 Three types of IMERG products are provided for different user
169 requirements and applications: near-real time "Early" run (IMERG-E), "Late"
170 run (IMERG-L), and post-real time "Final" run (IMERG-F) with approximate
171 latency of 4h, 12h and 4 months, respectively. The monthly gauge
172 adjustment is used in IMERG-F to reduce bias. IMERG-F is a research
173 quality product and generally considered to be more accurate than IMERG-E
174 and L. Here the latest version (V06B) IMERG-F is evaluated. The details of
175 the algorithm are described by Huffman et al. [10], and the ~~lastest~~latest
176 changes to the morphing algorithm in V06B are further provided by Tan et
177 al. [33]. IMERG data files can be accessed from
178 https://gpm.nasa.gov/data/directory.

179 *2.3. Ground MRMS*

180 To evaluate the quality of IMERG and 3B42 products, the "Level 3"
181 MRMS is selected as a ground-based reference. Level 3 denotes a gridded
182 product in GPM project. The NOAA's National Severe Storms Laboratory
183 (NSSL) integrates information from about 180 ground-based operational
184 radars and generates a 3D radar mosaic across the CONUS and southern
185 Canada (130°-60°W, 20°-55°N) [34].The Level 3 MRMS products used in this
186 study are at half-hourly and 0.01° resolutions, generated through significant
187 MRMS data postprocessing which involves gauge-based bias adjustments,
188 resampling, and quality controls, specifically for support of GPM GV [35]. A

Radar Quality Index (RQI) is also produced along with precipitation estimates in the MRMS suite. RQI ranging from 0 (worst) to 100 (best) indicates sampling and estimation uncertainty, depending on the height of the lowest radar elevation angle relative to the bright band and percentage of beam blockage by topography [36]. Figure 2 shows mean RQI maps averaged from half-hourly data for three winters and five summers, respectively, during the 55-month period (June 2014 to December 2015, June 2016 to July 2017, and March 2018 - December 2019). Data from Jan-May 2016, Aug 2017 – Feb 2018 are excluded, which will be discussed in Section 4.1. Figure 2 clearly indicates that the MRMS data quality in summer (JJA) is better than in winter (DJF). In addition, the gaps of radar coverages do exist in western CONUS even in summer. Further, compared to eastern CONUS, the MRMS quality is evidently lower in the Intermountain West because of severe blockages caused by the complex terrain. RQIs in coastal oceanic areas are generally very low due to large distances between neighboring radars as well sharp topographic transition. The coastal oceanic areas are masked out from the study (Figure 1).
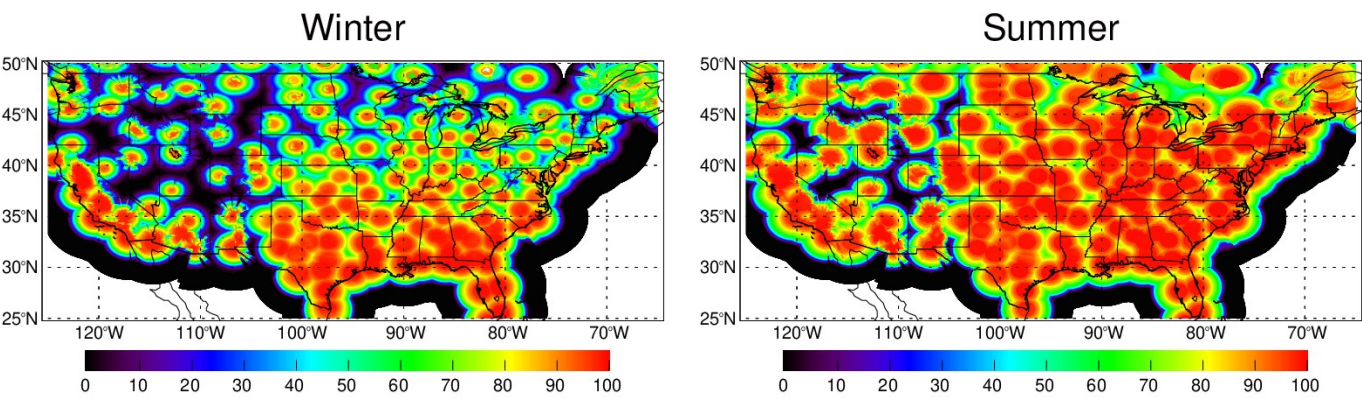


**Figure 2.** Maps of mean RQIs from three winters (DJF) and five summers (JJA) during the 55- month period (June 2014 to December 2015, June 2016 to July 2017, and March 2018 - December 2019).

*2.5. Montly GPCC*

The Global Precipitation Climatology Centre (GPCC) provides gridded precipitation analyses derived from quality-controlled gauges worldwide via the Global Telecommunication System of the World Meteorological Organization[30]. One of the GPCC products is Full Data Monthly Analysis Version 2020, which is developed from up to 54,000 gauges using the new GPCC Precipitation Climatology Version 2020 as analysis background. The GPCC monthly gauge analysis is one of the contributing data sources for TMPA [9] and IMERG [10]. The Full Data Monthly Analysis Version 2020 at 0.25° spatial resolution is used in this study to confirm the consistency of satellite-based precipitation products and check the MRMS reliability at monthly scale.

## 3. Methodology and Evaluation Metrics

Due to the difference in spatial and temporal resolutions among the above-described products, the validation is conducted using an approach that matches and temporally resamples MRMS, 3B42 and IMERG data to a common spatiotemporal resolution of 0.25° and 3-hour. Note that, for each day, the 3-hour cutoff times for 3B42 files are 0130, 0430, ..., 2230UTC. The first 3-hour of a day has a starting time of 2230UTC on the day before, and the last 3-hour has the ending time of 2230UTC on the data day. Accordingly, we accumulate 30-minute data for MRMS and IMERG to the same 3-hour interval as 3B42.

Probability density function (PDF) for precipitation occurrence and volume, scatter density plots, time series of daily and instantaneous 3-hourly precipitation maps are routinely produced [https://wallops-prf.gsfc.nasa.gov/NMQ/IMERG_3B42_MRMS_30m/index.php]. The statistical metrics such as relative bias (Bias), Pearson correlation coefficient (Corr), and normalized mean absolute error (NMAE), normalized root-mean-square error (NRMSE) are utilized to quantify the performance of satellite products.

Relative bias shows an overestimation (positive) or underestimation (negative) of total precipitation. Corr ranging from -1 to 1, is an indicator of an a priori assumed linear relationship between two products. NRMSE is normalized root-mean-square-error (RMSE), denoting the relative mean error magnitude. By using NRMSE, the issue that RMSE values increase with the precipitation intensities can be mitigated. The equations of these metrics are not listed in this manuscript. Readers are referred to statistics books, e.g.[37] and peer-reviewed papers, e.g. [15].

In order to further examine precipitation detection capability of satellite-based products (IMERG and 3B42) against ground-based reference product (MRMS), a contingency table is constructed for each product. The contingency table, also called two-way frequency table, consistes of occurrences of hit (h), miss (m), false alarm (f), and correct negative (c). For the satellite product to be evaluated, "hit" means that precipitation is correctly detected when it occurs, and "miss" means that precipitation is not

detected when it occurs. "False alarm" means that precipitation is mistakenly detected, and "correct negative" means that zero-precipitation is correctly flagged as such.

From the hits, misses, false alarms, and correct negatives, three widely applied categorical statistical skill metrices: Probability of Detection (POD), False Alarm Ratio (FAR) and Heidke Skill Score (HSS) can be calculated. POD is the fraction of actual precipitation occurrences that are correctly detected; a perfect score is 1 and worst is 0. The FAR measures the fraction of falsely detected precipitation occurrences, ranging from 0 (the perfect) to 1 (all detected precipitation rates are fake). HSS, a generalized skill score ranging from -∞ to 1, quantifies whether the estimate is equal to (0), worse (negative value) or better (positive value) than random chance relative to the reference; the perfect value is 1.

$$POD = \frac{h}{h+m} \tag{1}$$

$$FAR = \frac{f}{h+f} \tag{2}$$

$$HSS = \frac{2(hc - fm)}{(h+m)(m+c) + (h+f)(f+c)} \tag{3}$$

The categorical statistical skill metrics provide categorical measures of performance based on precipitation occurrences, but do not provide any information as to what fraction of the precipitation volume is detected. Therefore, a volumetric error decomposition approach developed by Tian et al.[38] and Habib et al. [39] is adapted to completely decompose the total error (E) of precipitation volume into three independent components: 1) hit bias (H) - precipitation is detected, but the precipitation volume is biased either positively or negatively; 2) bias due to missed precipitation (-M) - precipitation is not detected, which obviously leads to a negative bias; and 3) bias due to false precipitation detection (F) - precipitation is mistakenly detected while there is no precipitation, which apparently causes a positive bias. These three components are related by Equation (4)

$$E = H - M + F. \tag{4}$$

It is possible that H, -M or F is larger than E simply because they could cancel one another, which results in a smaller E. Therefore, this approach can reveal much more information on error characteristics by looking at individual components, instead of merely the total error.

In the calculation of the above categorical metrics, as well as in the error decomposition approach, the precipitation and no-precipitation threshold is assumed to be 0.1 mm h$^{-1}$. This ~~threshould~~threshold has been applied in many previous studies (e.g. [13]). If the precipitation rate is below 0.1 mm h$^{-1}$, it is considered not precipitating. The selection of this threshold is rather subjective, but the results are relatively robust, which is demonstrated when other thresholds (0.0, 0.2, 0.5, 1.0 mm h$^{-1}$) are also

295  selected for the test. The trace precipitation below the threshold of 0.1 mm
296  h$^{-1}$ is not very important in meteorological and hydrological applications. It
297  only accounts for 1.5%, 0.3%, 3.1% of total precipitation for MRMS, 3B42
298  and IMERG, respectively, which will be discussed in Section 4.4.   The
299  threshold of 0.1 mm h$^{-1}$ is below the minimum detectable precipitation rates
300  of remote sensors on ground radars and satellites. It is common to consider
301  a threshold to eliminate spurious light precipitation caused by different
302  instruments and retrieval algorithms [12,15].

303  **4. Results and Analysis**

304  *4.1. Time series of spatially averaged monthly precipitation*
305  Figure 3 is the time series of monthly precipitation spatially averaged
306  over the study domain for GPCC, MRMS, 3B42 and IMERG, respectively,
307  from June 2014 to December 2019. MRMS data are missing from January to
308  May 2016.  Precipitation from MRMS during July - November 2017 is
309  evidently abnormally lower than all other products. Daily and 3-hourly
310  precipitation                                                                    maps
311  (https://wallops-prf.gsfc.nasa.gov/NMQ/IMERG_3B42_MRMS_30m/index.php
312  )  reveal that MRMS failed to capture major rain events due to some
313  unknown issues during this period. After the period, there were large
314  amounts of missing data in MRMS from December 2017 to February 2018.
315  Hence, we only use the data in the common time span of 55 months from
316  June 2014 to December 2015, June 2016 to July 2017, and March 2018 -
317  December 2019 in this study. The data from January to May 2016 and from
318  August 2017 to February 2018 are excluded. From the commonly available
319  55-month datasets for all products, there are three winters (December 2014
320  - February 2015,  December 2016- February 2017, and December 2018-
321  February 2019) and five summers (June-August in 2014, 2015, 2016, 2018,
322  and 2019), which will be studied in a winter/summer contrast.
323  Figure 3 illustrates that the satellite-based monthly precipitation from
324  TRMM 3B42 and GPM IMERG have a pattern similar to the observations
325  from ground-based MRMS and GPCC products. All products are able to
326  reasonably reproduce the precipitation seasonal variation. IMERG and 3B42
327  appear to overestimate the winter precipitation. The overestimation is likely
328  because IR precipitation estimation in satellite products is indirectly inferred
329  from cloud top temperature which limits the accuracy for the cold
330  precipitation      detection      and      quantification      [17].      Product
331  ~~charicteristics~~characteristics will be further analyzed in next section.

332

333
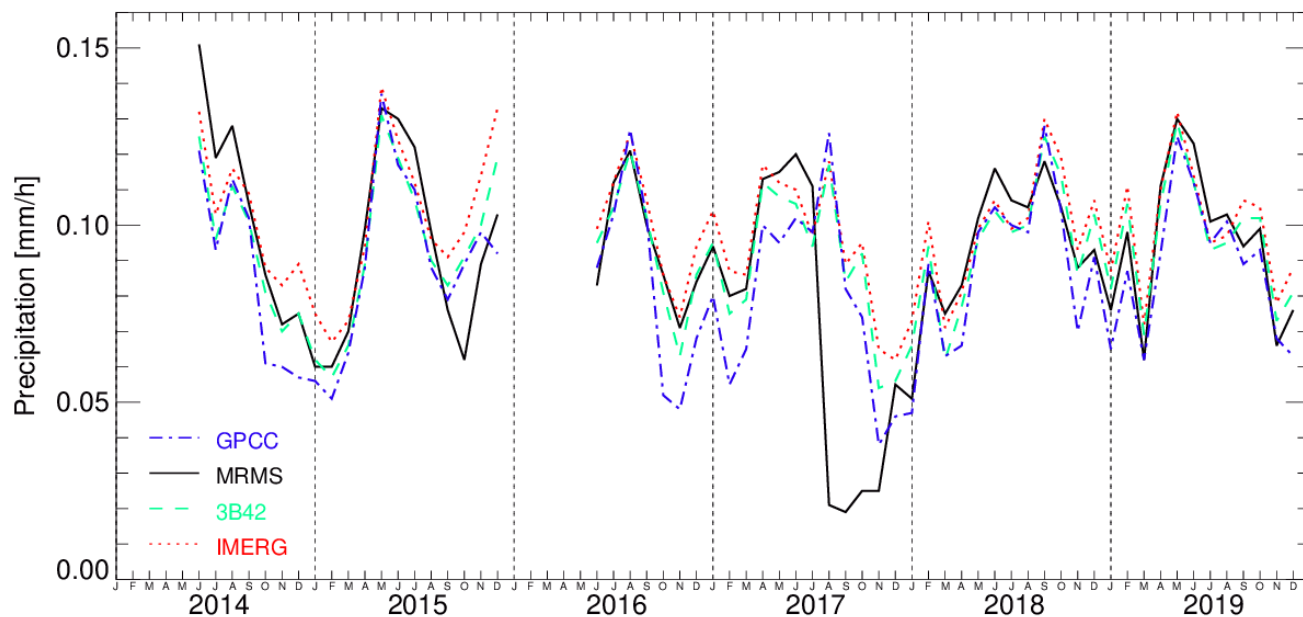
334 **Figure 3.** Time series of monthly precipitation spatially averaged over the study
335 domain for GPCC, MRMS, 3B42 and IMERG, respectively, from June 2014 to
336 December 2019.

337    Figure 4 is the time series of ratios and correlations for IMERG (3B42)
338 relative to MRMS, calculated from 3-hourly precipitation at 0.25° resolution
339 for each month. GPCC is not avaliable for the comparison because it is a
340 monthly product. Only pixels associated with the perfect MRMS RQI value
341 (100) and precipitation at least 0.1 mm h$^{-1}$ for both IMERG (3B42) and
342 MRMS are used in the calculation. This qualitatively filtering precedure can
343 effectively remove noise from MRMS data [40]. Figure 4 clearly
344 demostrates that IMERG constantly outperforms 3B42 in terms of ratio and
345 correlation. Relative to MRMS, both products overestimate precipitation in
346 winter whereas understimate precipitation in summer.  The correlation
347 between IMERG (3B42) and MRMS is higher in winter than in summer.

348
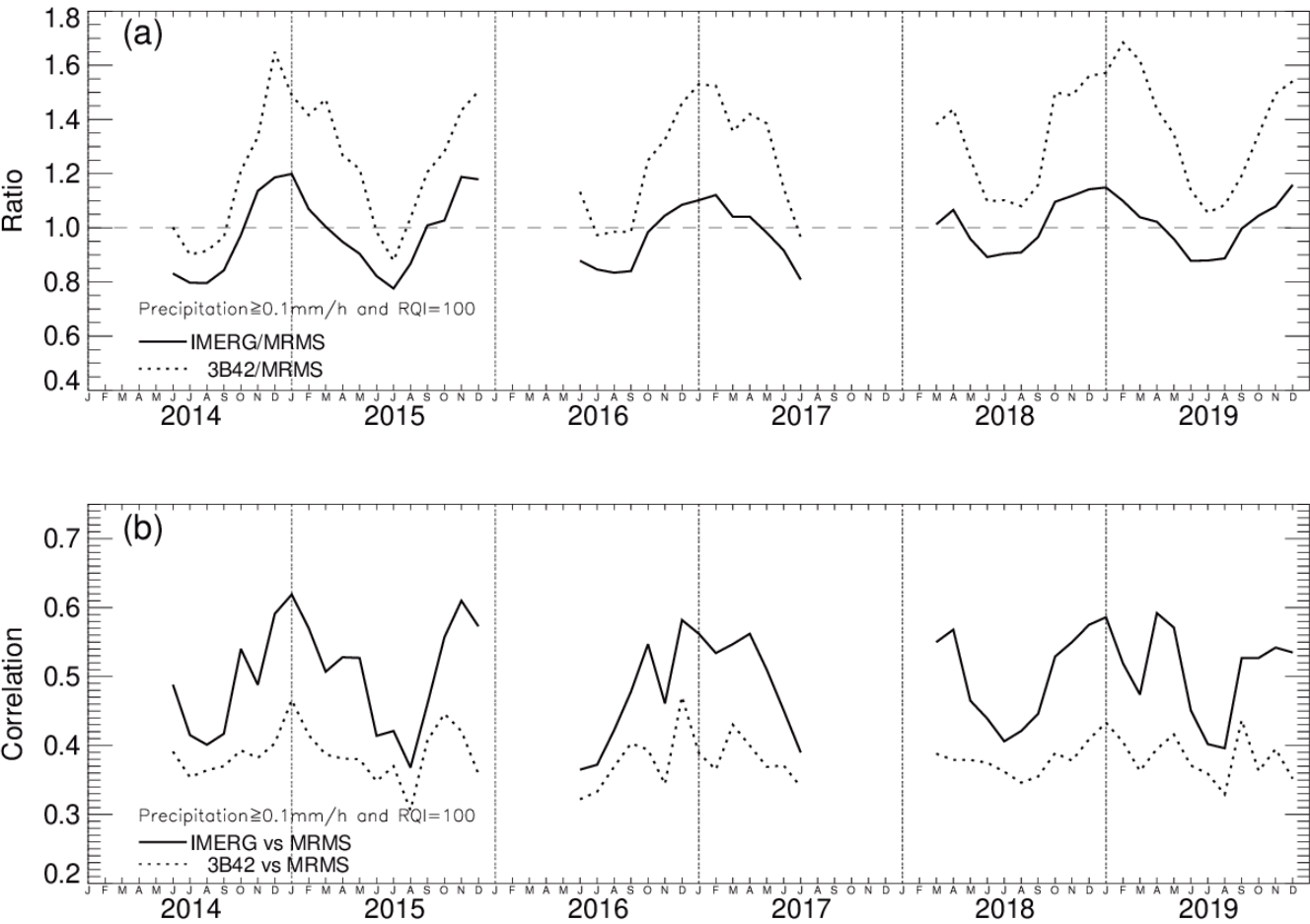
349  **Figure 4.** Time series of monthly precipitation (a) ratios and (b) correlations for
350  IMERG and 3B42 relative to MRMS, respectively calculated from 3-hourly
351  precipitation at 0.25° resolution. Only pixels associated with MRMS RQI value of 100
352  and precipitation at least 0.1 mm h$^{-1}$ for both IMERG (3B42) and MRMS are used.

353        Figure 4 is just one of the plots for the time series of monthly ratios and
354  correlations from all combinations of RQIs ($\geq$ 0, 25, 50, 75, 100) and
355  precipitation thresholds ($\geq$ 0.0, 0.1, 0.2, 0.5 mm/h) for 3B42 and IMERG.
356  The complete distribution of the ratios and correlations can be dispalyed on
357  boxplots with emphasis on the directions of bias (i.e. positive or negative)

and outliers. In Figures 5 and 6, one boxplot conresponds to the 55 months from each of the 80 time series plots. For example, the data in the boxplot (last column in Figure 6f) for the correlation between IMERG and MRMS with RQI of 100 and precipitation threshold of 0.1mm $h^{-1}$ is from the solid line in Figure 4b.

From Figures 5b-d, the large positive bias in 3B42 is evident, with the interquartile range and mean all well above 1. Large outliers contribute a considerable overestimation of precipitation rates.

All boxplots, except for Figures 5a and 6a, share a common feature that the ratio decreases towards 1 with increasing RQI. It is worth noting that the ratio for 3B42 relative to MRMS with threshold of 0.0 mm $h^{-1}$ (Figures 5a) is lower than all other ratios with larger precipitation thresholds (Figures 5b-d and 6b-d). This implies that 3B42 missed a considerable amount of total rain volume due to light precipitation. This will be shown in probability distribution of 3-hourly precipitation rates (Figure 11). On the other end, 3B42 overestimates precipitation with intensities greater than 0.1mm $h^{-1}$ (Figures 5b-d). This issue is largely mitigated in Figures 6b-d for IMERG.

The correlation coefficient is relatively stable with respect to the change of RQI. This is expected because the correlation quantifies the co-variability between two products. The best correlation is for the cases with a threshold of 0.0 mm $h^{-1}$ (Figures 5e and 6e) because of large concurrences of zero-precipitation rates included.

381
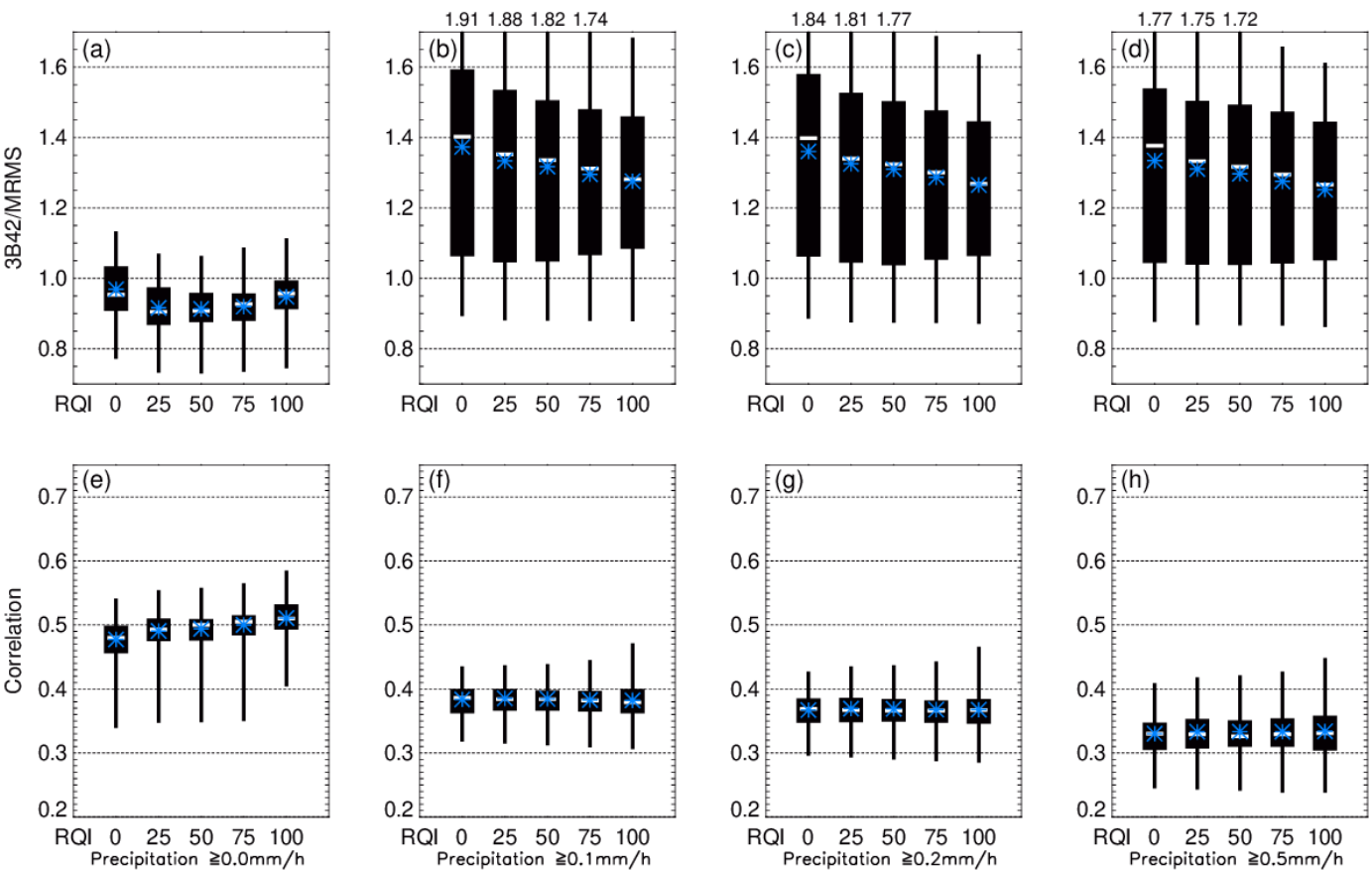
382        **Figure 5.** Boxplots of ratios and correlations for 3B42 relative MRMS with respect to
383        different RQI and precipitation thresholds. The RQI thresholds are 0, 25, 50, 75 and
384        100 listed at the bottom of each panel. The precipitation thresholds are 0.0 mm h⁻
385        ¹(Figures 5a and 5e), 0.1 mm h⁻¹(Figures 5b and 5f), 0.2 mm h⁻¹(Figures 5c and 5g)
386        and 0.5 mm h⁻¹(Figures 5d and 5h). The black box is the interquartile range
387        extending from the lower quartile (25th) to the upper quartile (75th). The white bar
388        inside the black box is the median, and the blue asterisk is the mean.  The vertical
389        line represents whiskers extending from the minimum to maximum. The upper
390        whisker is truncated and listed at the upper border of each boxplot if the maximum is
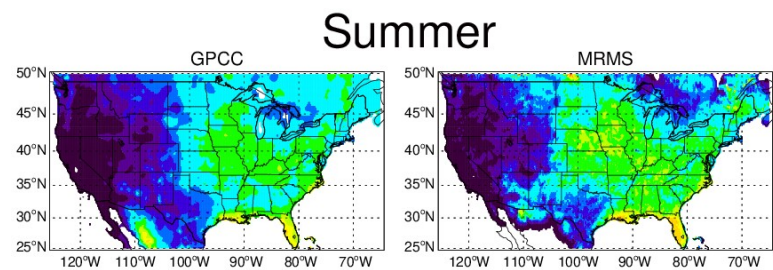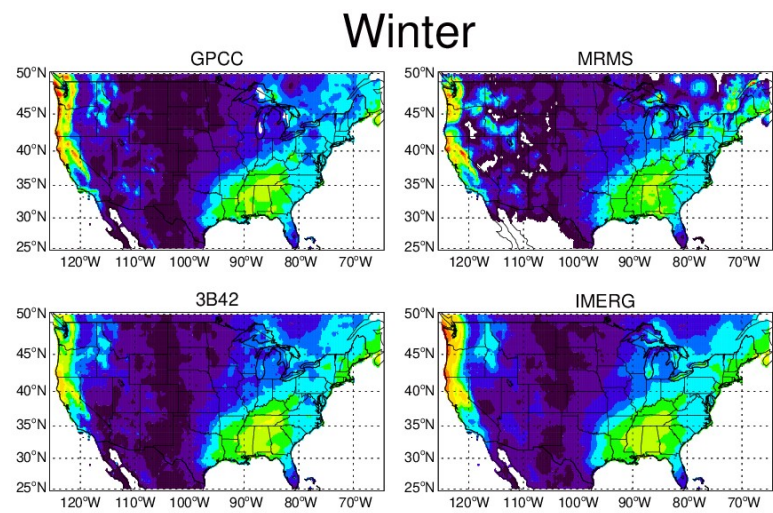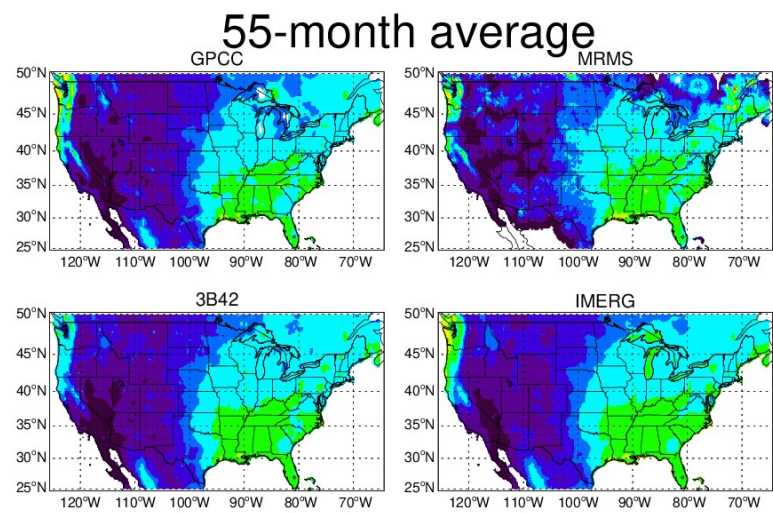391        out of the plot area.

392
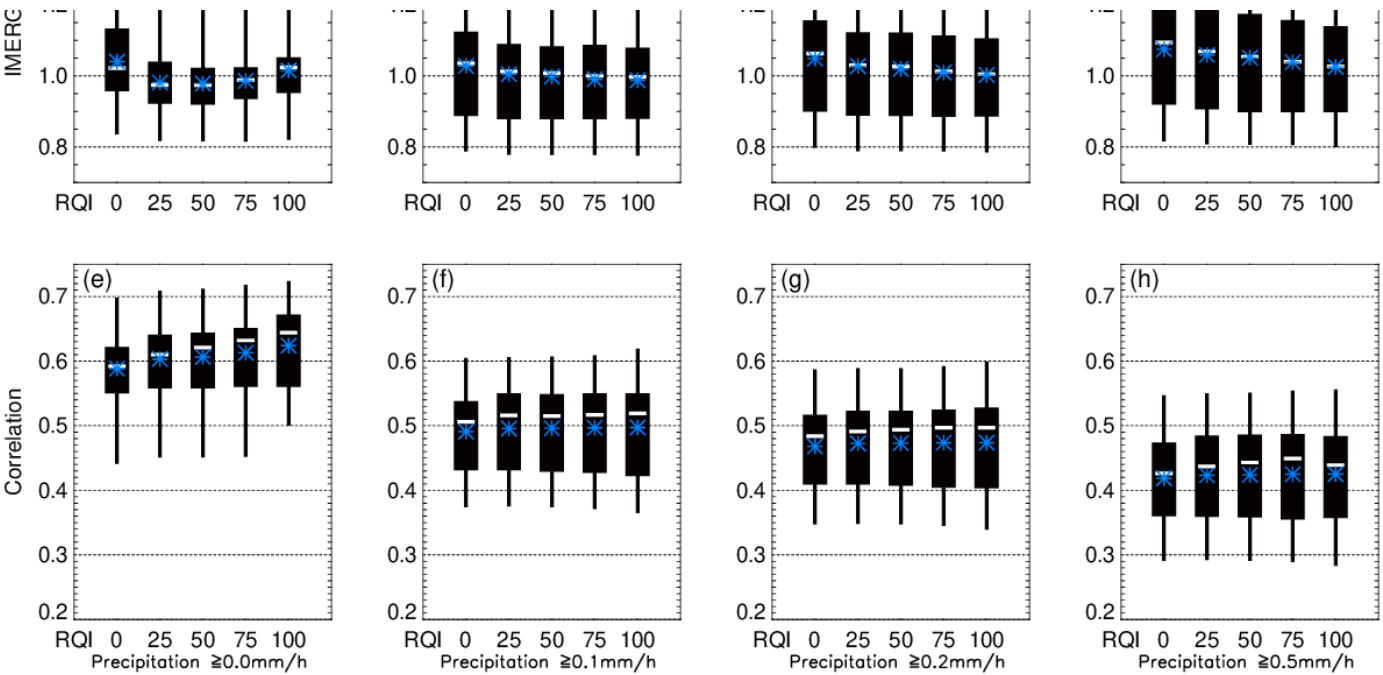
**Figure 6.** Same as Figure 5 but for IMERG.

393

*4.2. Spatial Distribution of Mean Monthly Precipitation and Error Decomposition Analysis*

The spatial distributions of mean monthly ~~preicicipitation~~precipitation averaged from 55 months, three winters, and five summers for ground (GPCC and MRMS ) and satellite (3B42 and IMERG) derived products are shown in Figure 7. For the 55-month ave~~rarr~~rage, all products display a high degree of consistency in spatial distribution patterns with copious precipitation in the Southeast and Pacific Northwest, with relatively less precipitation in the large area of western CONUS. However, differences are still visually discernable among the four products. There are missing data over the Great Lakes in GPCC because of the lack of floating rain gauges. A gradual precipitation gradient descent from south and southeast to north and northwest in IMERG, 3B42 and GPCC is spatially smoother than that in MRMS. That is the smoothing effect in GPCC gauge interpolation[30] reflected in 3B42 and IMERG through the gauge adjustment. IMERG best

409    reproduces the spatial variation of GPCC which appears to be benefit from
410    the IMERG's monthly calibration against GPCC gauges. IMERG has the best
411    pattern correlation (0.963) with GPCC, followed by 3B42 (0.947) and MRMS
412    (0.81).
413          In winter, precipitation along the northwest coastal area is considerably
414    underestimated by 3B42, mainly because of the difficulty in detecting warm,
415    coastal

## 55-month average

GPCC

MRMS

3B42

IMERG

## Winter

GPCC

MRMS

3B42

IMERG

## Summer

GPCC

MRMS

**Figure 7.** Mean monthly precipitation maps from GPCC, MRMS, 3B42 and IMERG over the period of 55 months, three winters and five summers, respectively. The precipitation unit is mm h$^{-1}$.

rainfall. This is also reported by Tang et al. [41]. The underestimation is well corrected in IMERG.

Both IMERG and 3B42 show different levels of overestimation near the Great Lakes area, especially in summer, due to the lack of gauge calibration. The spatial characteristics of the error components for 3B42 and IMERG in winter and summer seasons can be further analyzed using the error decomposition approach [38, 39].

As described in Section 3, we ~~decomose~~decompose the total bias into hit, missed-precipitation and false-precipitation biases for each product during each month at each 0.25° grid in the study domain (Figure 1), and then accumulate monthly total biases and bias components as percentages of total MRMS precipitation into seasonal scales, ~~particually~~particularly for winter and summer seasons. In this analysis, 3-winter-average and 5-summer-average are presented (Figure 8). The spatial characteristics of the error components for the 55-month average is not shown here as it basically represents the summation of winter and summer in Figures 8a-d, the spatial distributions of the total bias and three error components with RQI filtering. Both 3B42 and IMERG share considerable similarity in their spatial bias patterns, but with different levels of biases. One obvious feature in total bias is that the superior performance of satellite products in the central and southern Plains. This may be resultant of the gauge calibration in 3B42 and IMERG algorithms. Both IMERG and 3B42 appear to underestimate the orographically-forced precipitation over the large areas of mountainous western CONUS in summer and winter, whereas they overestimate the precipitation along the Rockies, Sierra Nevada, Mexican border and Great Lakes areas where MRMS quality is questionable (Figure 2). The overestimation is mostly dominated by the enhanced false precipitation detection in these areas. The positive hit bias also contributes to the overestimation over some of these areas. However, after the RQI filter (RQI=100) is applied, these biases are largely removed (Figures 8 b, d).

For winter season, the missed-precipitation bias has the largest amplitude (Figure 9a) among all three components, especially for 3B42 at higher latitudes and mountainous areas (Figures 8a, b). Missed precipitation might be mostly due to the inability of multi-satellite PMW sensors for snowfall measurements or rainfall observations over icy land surface [41, 42]. Substantial improvements can be seen in IMERG product where the magnitude of missed precipitation error is dramatically decreased. This can be the result of the use of morphing and PERSIANN-CSS in IMERG. Compared to 3B42, the hit and missed-precipitation biases in IMERG are greatly reduced, whereas the false-precipitation bias is slightly reduced, which is more clearly summarized in Figure 9. The missed-precipitation bias has the largest amplitude among all three components.

These three components cancel one another, which results in a smaller total bias. IMERG displays obvious improvement to 3B42 in missed-precipitation and hit bias reduction in winter. However, the hit bias in IMERG is over-reduced in summer, resulting in large area of negative hit bias (Figures 8 c, d).

For summer season (Figures 8c, d), both products are noticeably similar with slight underestimation in most of CONUS except for along the Mexican border and the Great Lakes region. The missed-precipitation bias is largely reduced compared to its winter counterpart. The total bias pattern appears similar to that of the hit bias (Figures 8c, d). The overall false precipitation and missed-precipitation biases are not as pronounced as in winter (Figures 8 and 9). This can likely be traced to the more effective gauge correction in 3B42 and IMERG algorithms, which benefits from the fact that the gauge measurement is more reliable for summer rainwater than for winter solid precipitation.

Note that the selection of RQI=100 is a very conservative approach which only retains the MRMS data with the best quality; as a result, it significantly limits the data availability especially in the Intermountain West where the ~~the~~ quality of radar measurements is lower than other CONUS areas because of radar beam blockage by terrain. Hence, MRMS is not a good GV reference over mountainous regions as it might not correctly represent the weather systems in the regions [43].

We also select other RQI thresholds for comparisons (Figure 9). Figure 9 shows the averaged error components, as well as the total errors, as percentages of the total MRMS precipitation for all 3-hourly precipitation over 0.25° grids with different RQI filtering for winter and summer, respectively.

For winter, missed precipitation is the major source of error, accounting for about 50% of the total precipitation for 3B42 and 20-30% for IMERG, much larger than the total bias. The huge missed-precipitation is countered by the sum of positive hit bias and positive false bias, resulting in a very small total bias.

For summer, both missed and false precipitation biases are smaller compared to those in winter. Compared to winter, the positive hit bias in summer is greatly reduced for 3B42, but it turns to be negative for IMERG.

One easily noticeable feature in Figure 9a is that the missed- and false-precipitation biases for both IMERG and 3B42 obviously decrease with the increasing RQI in winter season. In summer, this feature is not very salient. RQI filtered less area in summer than in winter because RQI is better in the warm season than in the cold season (Figure 2) due to the presence of stronger low-level temperature gradients during winter periods.

39


(a) Winter (RQI≧0)


(b) Winter (RQI=100)


(c) Summer (RQI≧0)


(d) Summer (RQI=100)
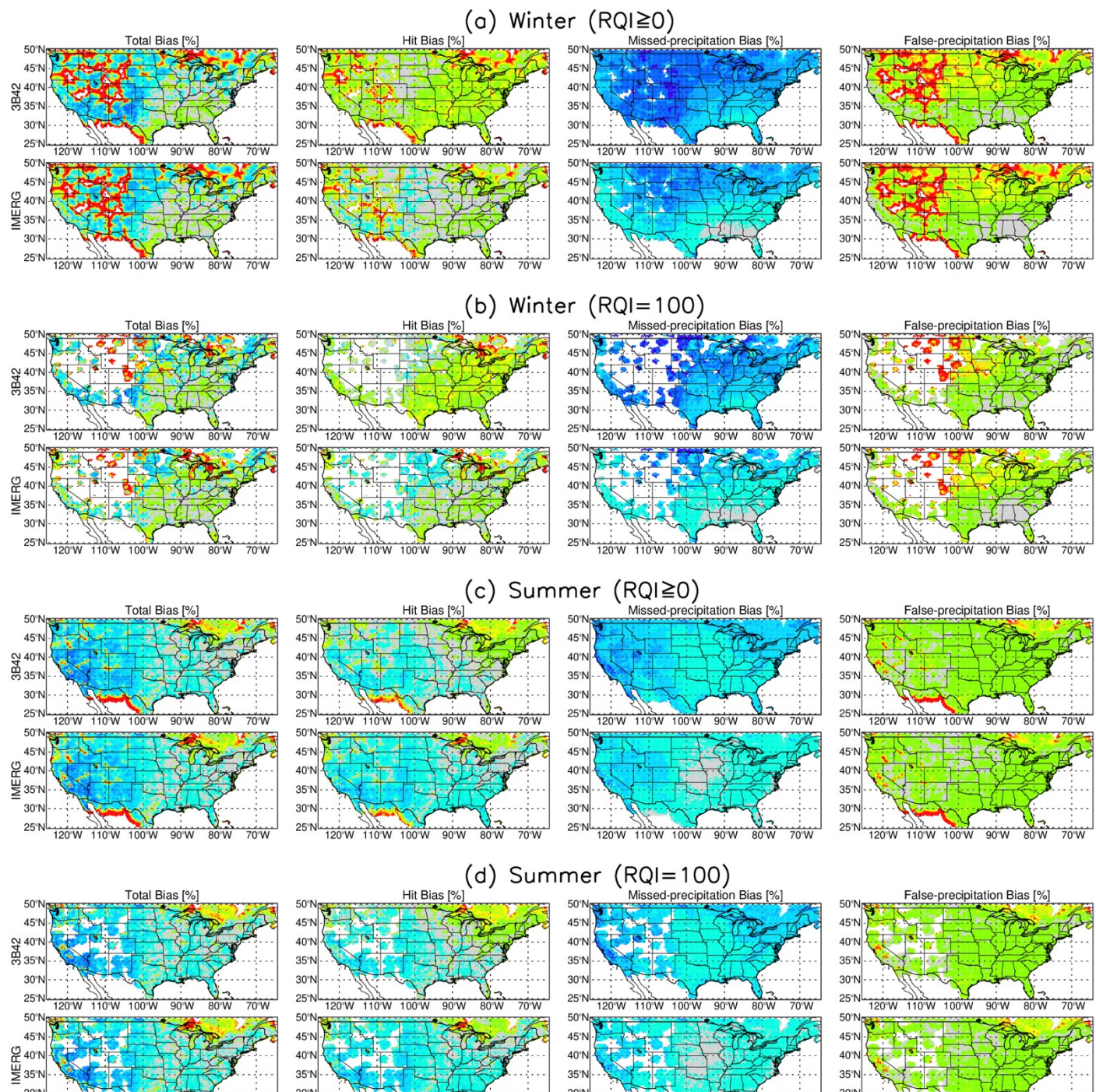
504 **Figure 8.** The spatial distributions of error components for 3B42 and IMERG as
505 percentages of total MRMS precipitation: total (E), hit (H), missed-precipitation (M),
506 and false-precipitation (F) biases for three winters and five summers with RQI≥0 and
507 RQI=100. The components are related by E = H - M + F.



508
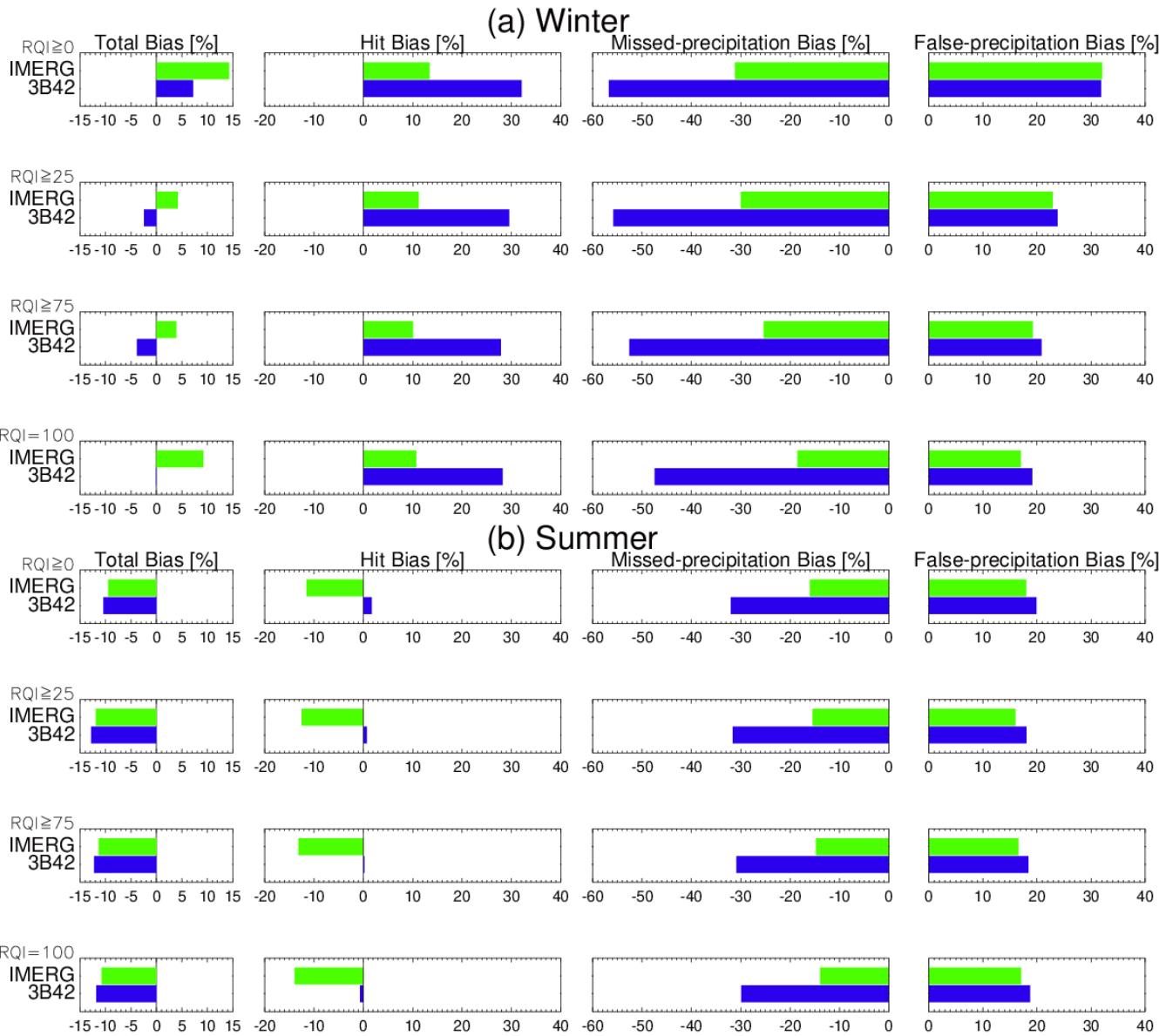
**Figure 9.** Averaged error components for IMERG and 3B42 relative to MRMS at 3-hour and 0.25° resolution for (a) winter and (b) summer with with RQI thresholds from 0 to 100.

*4.3. Statistical metrics and Categorical skill score*

Figure 10 shows statistical metrics for IMERG and 3B42 performances relative to MRMS with RQI filtering (zero or 100%) for winter and summer, respectively. IMERG clearly outperforms 3B42 in both winter and summer in terms of all metrics except for a slightly worse false rate, FAR, and relative bias in summer. The morphing algorithm improves the precipitation detection with better hit and miss rates, as well as increased POD and HSS scores in Figures 10a-d, but at the price of slightly increased false rate and FAR. The increased FAR might be an issue associated with PMW retrieval which sometimes may erroneously treat the dynamic surface characteristics over land as precipitation signals [44].

Looking at Figure 10a, a striking feature is that the relative bias in winter season is drastically reduced from 74.7% for 3B42 to 19.1% for IMERG without RQI filtering, and the bias is reduced from 54.1% to 12.8% when RQI=100 is applied. Another salient distinction between TRMM 3B42 and GPM IMERG is that there is a large increase in POD, and a marginal increase in FAR; as a result, a perceived increase in HSS. As discussed in Su et al. [14], the improved POD is mainly attributed to the enhanced sensitivity of sensors used in producing the precipitation products and increased sampling frequency of the GPM mission.

Considered over a continental scale, Figures 10c and d indicate that IMERG precipitation estimates are lower relative to MRMS during summer. The similar result is also found by comparing IMERG with radar-based Stage IV product [45]. This suggests that IMERG may underestimate precipitation from convective systems. In summer, strong convective precipitation events occur more frequently. Due to their short duration, localization and spatial complexity, these events are generally more difficult to be captured by satellite remote sensors.

The hit bias in Figure 9 and relative bias in Figure 10 have the same direction (positive or negative) but different amplitudes. This behavior occurs because the relative bias represents a difference between satellite and MRMS precipitation as the percentage of MRMS precipitation, conditioned on the presence of a satellite and MRMS precipitation rate of at least 0.1mm h$^{-1}$, whereas the hit bias is the difference as the percentage of MRMS precipitation greater than 0.1 mm h$^{-1}$.

To further measure the agreement between satellite-based and ground-based products, the Kling-Gupta efficiency (KGE) [46,47], a recently more widely used criterion for the predictive skill assessment of hydrological models, is also introduced as another statistic in this study. By taking into account correlation, bias and variability in a more balanced way, KGE provides an additional interesting insight into the performance of satellite-based products. KGE ranges from -∞ (worst) to 1 (best). Table 1 lists the

KGE values calculated from 3-hourly precipitation rates at 0.25° resolution for 3B42 and IMERG using MRMS as a reference during the periods of the three winters, five summers and 55 months, respectively. Table 1 further demonstrates the obvious performance improvement from 3B42 to IMERG in winter.

**Table 1.** KGE calculated from 3-hourly precipitation rates at 0.25° resolution for 3B42 and IMERG using MRMS as a reference during three-winter, five-summer and all 55-month periods, respectively.

| | **Winter** | | **Summer** | | **55 months** | |
|---|---|---|---|---|---|---|
| | 3B42 | IMERG | 3B42 | IMERG | 3B42 | IMERG |
| **RQI≥0** | -0.193 | 0.372 | 0.359 | 0.362 | 0.286 | 0.465 |
| **RQI=100** | 0.026 | 0.466 | 0.338 | 0.335 | 0.330 | 0.449 |

A threshold of 0.1 mm h$^{-1}$ is used in the calculation of these statistical metrics. If other thresholds are used, the numerical values for these statistics would be different. However, as indicated in Figures 5 and 6, the general conclusion should be robust.
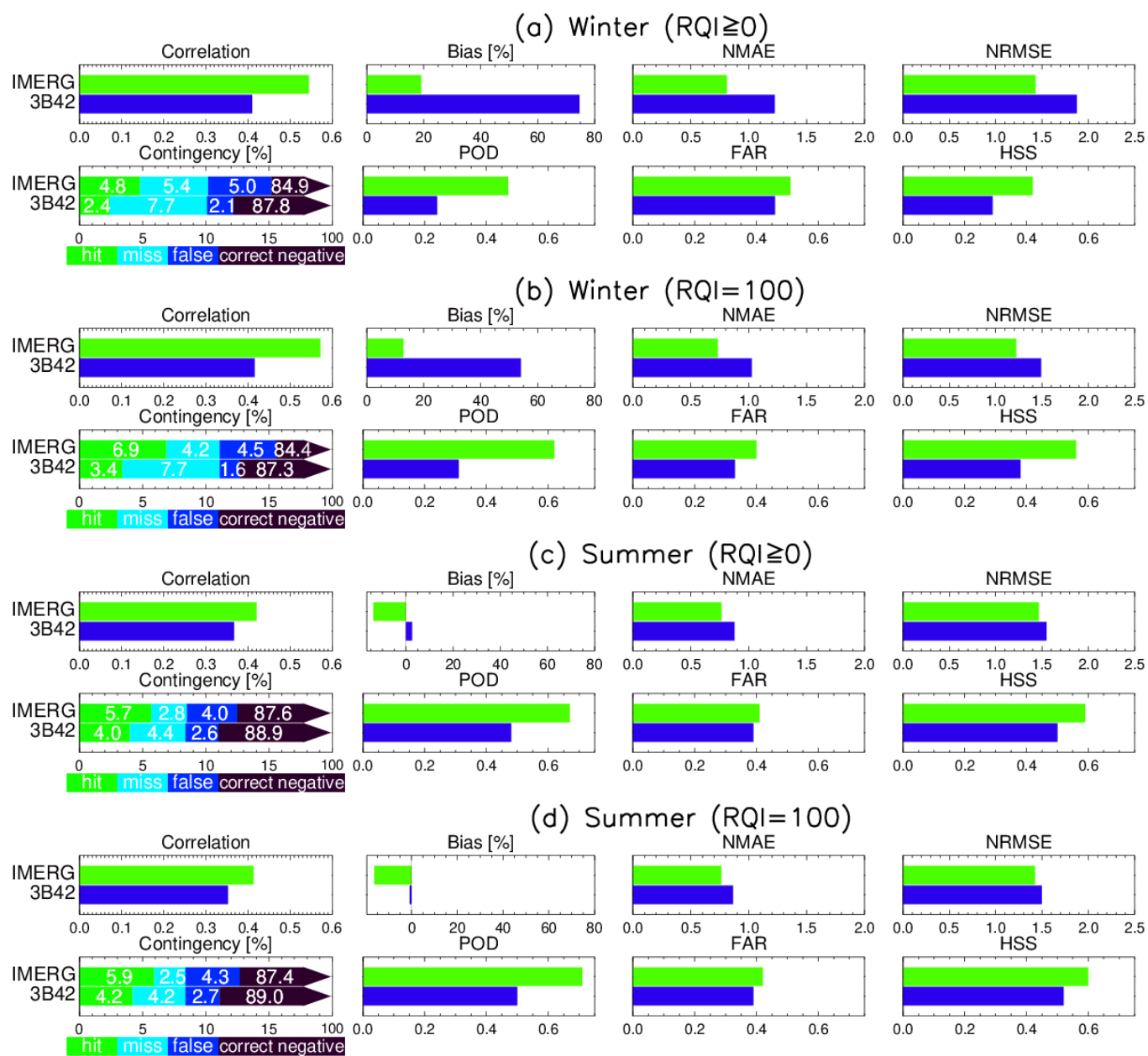
**Figure 10.** Statistical metrics for IMERG and 3B42 performances relative to MRMS for (a,b) winter and (c,d) summer with (a,c) RQI≥0 and (b,d) RQI=100. Top panels: Pearson correlation coefficient (Correlation), relative bias (Bias), normalized mean absolute error (NMAE), and normalized root mean square error (NRMSE). Bottom:

574  contingency table with hit (green), miss (aqua), false alarm (blue), and correct
575  negative (black). Probability of detection (POD), False alarm ratio (FAR) and Heike
576  skill score (HSS) are also displayed.

577  *4.4. Probability distribution analysis*

578  To further investigate the distribution of precipitation intensities,
579  probability distribution functions are constructed as shown in Figure 11 for
580  3-hourly precipitation occurrences and volumes at 0.25° resolution without
581  RQI filtering over the study domain during the 55-month period. The insets
582  are for the data in the range between 0 and 0.1 mm h$^{-1}$, representing either
583  no precipitation or drizzle. In this range, 3B42 accounts for 96.1% of total
584  data points but only 0.3% of precipitation volume, whereas the numbers are
585  93.0% and 3.1% for IMERG, 93.5% and 1.5% for MRMS. This also indicates
586  the improved capability of IMERG in detecting no precipitation or drizzle
587  occurrences compared to 3B42, which could be attributed to the GPM DPR
588  with its higher sensitivity relative to TRMM PR [48, 49]. For precipitation
589  under 0.1 mm h$^{-1}$, IMERG very well reproduces the precipitation occurrences
590  but overestimates the volumes.

591  The distribution for the precipitation above 0.1 mm h$^{-1}$ is separately
592  constructed for each product. 3B42 product shows an obvious
593  overestimation in occurrences for the precipitation with the intensity
594  between 1 and 30 mm h$^{-1}$, and underestimation for precipitation less than 1
595  mm h$^{-1}$. IMERG shows an improved agreement with MRMS both in
596  occurrence and volume distributions.

597  Similar features can be also seen in the joint distributions of 3-hourly
598  precipitation rates at 0.25° resolution for 3B42 vs MRMS, and IMERG vs
599  MRMS, respectively, conditional on both precipitation rates at least 0.1 mm
600  h$^{-1}$(Figure 12). More importantly, the random error for IMERG is very low
601  compared to 3B42, as seen by the tight scatter in density plot. The
602  distribution is more concentrated along 1:1 line for IMERG, consistent with
603  NMSE and NMAE results in Figure 10. Figure 12 ~~demostrates~~demonstrates
604  that both higher and lower precipitation rates in IMERG are better
605  represented than in 3B42, though IMERG slightly overestimates the higher
606  precipitation rates and underestimates the lower precipitation rates. The
607  improvement can be mainly attributed to the improved morphing scheme
608  and enhanced sensitivity of sensors used in IMERG.

609
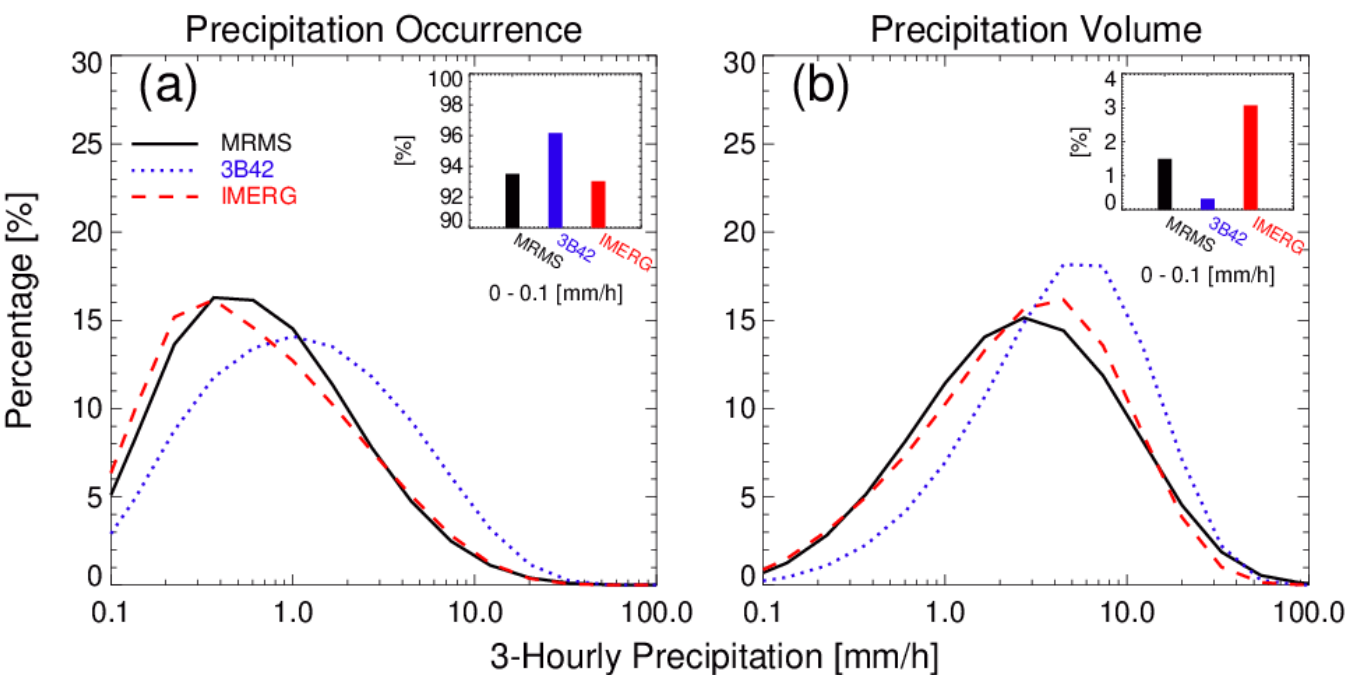
**Figure 11.** Probability distributions of 3-hourly precipitation (a) occurrences and (b) volumes for MRMS , 3B42 and IMERG at 0.25° resolution without RQI filtering over the study domain. The insets are percentages of occurrences and volumes for precipitation less than 0.1 mm h$^{-1}$, including zeros.
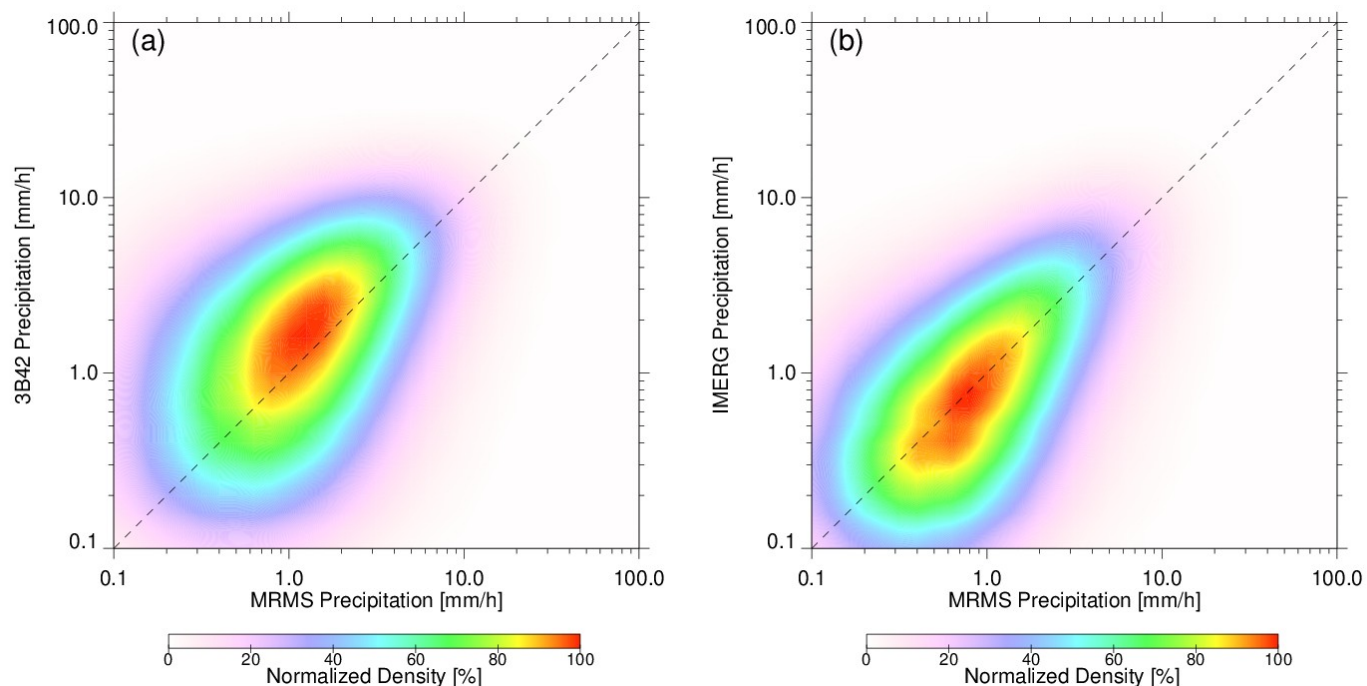
**Figure 12.** Joint distributions of 3-hourly precipitation rates at 0.25° resolution for (a) 3B42 and (b) IMERG against MRMS, conditional on both precipitation rates at least 0.1 mm h$^{-1}$.

## 5. Discussion

From the launch of the TRMM satellite in November 1997 to the currently active GPM mission, satellite-based remotely sensed precipitation measurement has been advanced steadily. Built upon the success of TRMM, GPM is developed to provide not only measurement continuity, but also an improvement on the TRMM instruments, algorithms and precipitation products. GPM IMERG, as a successor of TRMM 3B42, is a unified satellite algorithm developed to provide multi-satellite precipitation products over the globe. All IMERG products are retrospectively processed back to the start of TRMM era (January 1998), and continue for the entire life of the GPM mission, currently expected to last to the mid-2030's or beyond. The evaluation of a new version of IMERG is particularly important for the algorithm development and application.

It should be noted that the evaluation of GPM IMERG and TRMM 3B42 precipitation products in this study is carried out over the land area only (Figure 1). The results drawn from this land area would be different for oceanic sites because of challenging problem for the PMW retrievals over

635
636
complex land surfaces. A validation study over ocean (e.g., Kwajalein Atoll) is underway.

637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
In addition, any ground-based products themselves are certainly not perfect, such as MRMS, especially in winter and over mountainous regions, although they have been often used as references in the satellite-based product validation. The validation results are likely unreliable if questionable products are used as references [50]. Satellite-based products in western CONUS appear to exhibit reduced performance in comparison to the eastern CONUS where RQI is high. This does not necessarily mean that the satellite products are the only problem, as there are also problems in the MRMS product as indicated in Figure 2. It is our conclusion that the MRMS product herein might not be suitable as a GV reference over the mountainous west CONUS regions.  On the contrary, one may reasonably argue that the satellite products can be utilized to fill the gaps or even substitute the ground product where the ground observations are sparse, rare, poor or completely missing, as considering that the satellite products perform well over many areas with adequate radar or gauge coverage.

653
## 6. Conclusions

654
655
656
657
658
659
660
661
662
663
664
This study evaluates two satellite-based precipitation products TRMM 3B42 V7 and GPM IMERG V06B using MRMS as the reference over CONUS for the 55-month period.  GPCC is also used in the comparison with MRMS and satellite products. The spatial distributions of precipitation climatology from four products are analyzed. The error characteristics are further examined for 3B42 and IMERG in winter and summer by an error decomposition approach which partitions total bias into hit bias, biases due to missed precipitation, and those due to false precipitation. The continuous and categorical statistical metrics are used to quantitatively evaluate the performance of the two satellite-based products. The main findings are summarized as follows:

665
666
667
(1) All products display a high degree of consistency in spatial distribution patterns though some differences are visually discernable.

668
669
(2) The IMERG shows substantial improvements in terms of nearly all statistical metrics, compared to its predecessor 3B42.

670
671
672
673
(3) For winter, the improvement in IMERG is primarily from significantly reduced missed-precipitation bias, and from largely reduced positive hit bias. For summer, the improvement is mainly from notably reduced missed-precipitation bias and marginally

reduced false-precipitation bias but at the expense of worse hit bias.

(4) The precipitation intensity distribution shows a significant improvement of IMERG algorithm in comparison with 3B42 which obviously overestimates heavy precipitation but underestimates light precipitation.

(5) Missed-precipitation bias over mountainous regions, especially over frozen surfaces in winter, is still a challenging problem in satellite-based precipitation retrieval algorithms. The bias correction is of particular importance in mountainous regions such as Serra Nevada Mountains in California and Rocky Mountains in Colorado.

(6) All the statistical metrics, as well as the error decomposition approach, work together, are effective in evaluation of the performances for the satellite-based precipitation products.

## References

1.  Kidd, C.; Huffman, G. Global precipitation measurement. *Meteor. Appl.,* **2011,** *18,* 334–353. https://doi.org/10.1002/met.284.
2.  Wang, J.; Wolff, D.B. Evaluation of TRMM ground-validation radar rain errors using rain gauge measurements. *J. Appl. Meteor. Climatol.,* **2010,** *49,* 310–324.
3.  Wang, J.; Wolff, D.B. Evaluation of TRMM rain estimates using ground measurements over Central Florida. *J. Appl. Meteor.,* **2012,** *51,* 926-940.
4.  Tan, J.; Petersen, W. A; Kirchengast, G.; Goodrich, D. C.; Wolff, D. B. Evaluation of Global Precipitation Measurement Rainfall Estimates against Three Dense Gauge Networks. *J. Hydrol.* **2018,** *19,* 517-532.
5.  Wolff, D. B.; Marks, D. A.; Amitai, E.; Silberstein, D. S.; Fisher, B. L.; Tokay, A.; Wang, J.; Pippitt, J. L. Ground validation for the Tropical Rainfall Measuring Mission (TRMM). *J. Atmos. Oceanic Technol.,* **2005,** *22,* 365–380.
6.  Bringi, V. N.; Chandrasekar V. *Polarimetric Doppler Weather Radar. Principles and Applications;* Cambridge University Press: Cambridge, United Kingdom, 2001; 636 pp. https://doi.org/10.1017/CBO9780511541094
7.  Cifelli, R.; Chandrasekar, V.,;Lim, S.; Kennedy, P. C.; Wang, Y.; Rutledge, S. A.  A new dual-polarization radar rainfall algorithm: Application in Colorado precipitation events. *J. Atmos. Ocean. Tech.,* **2011,** *28,* 352–364,  https://doi.org/10.1175/2010JTECHA1488.1.
8.  Kidd, C.; Levizzani, V. Status of satellite precipitation retrievals. *Hydrol. Earth Syst. Sci.* **2010,** *15,* 1109–1116.
9.  Huffman, G. J.; Adler, R.; Bolvin, D.; Gu, G.; Nelkin, E.; Bowman, K.; Hong, Y.; Stocker, E.; Wolff, D. (2007). The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-Global, Multiyear, Combined-Sensor Precipitation Estimates at Fine Scales. Journal of Hydrometeorology. *J. Hydrometeorol*. **2007,** *8,* 38-55.
10. Huffman, G. J. and Coauthors.  NASA Global Precipitation Measurement (GPM) Integrated Multi-Satellite Retrievals for GPM (IMERG). NASA Algorithm Theoretical Basis Doc., version 06, 39 pp., https://gpm.nasa.gov/sites/default/files/2020-05/IMERG_ATBD_V06.3.pdf.
11. Gebremichael, M.; Hossain, F. *Satellite Rainfall Applications for Surface Hydrology.* Springer, Dordrecht, Netherlands, 2010; pp. 1–327.
12. Petersen, W. A.; Kirstetter P.E.; Wang J.; Wolff, D. B.; Tokay A. The GPM Ground Validation Program.  In *Satellite Precipitation Measurement,* Chapter 26, V. Levizzani, C. Kidd, D. B. Kirschbaum, C. Kummerow, K. Nakamura, F. J. Turk, Eds, 2020, Springer-Nature.
13. Guo, H.; Chen, S.; Bao, A.; Behrangi, A.; Hong, Y.; Ndayisaba, F.; Hu, J.; Stepanian, P.M.  Early assessment of integrated multi-satellite retrievals for global precipitation measurement over China. *Atmos. Res.* **2016,** *176–177*, 121–133. https://doi.org/10.1016/ j.atmosres. 2016.02.020.
14. Su, J.; Lü, H; Ryu, D.; Zhu, Y. The assessment and comparison of TMPA and IMERG products over the major basins of Mainland China. *Earth Space Sci.,* **2019,** *6*, 2461–2479, https://doi.org/10.1029/2019EA000977.
15. Tan, J.; Petersen, W. A.; Tokay, A. A novel approach to identify sources of errors in IMERG for GPM ground validation. *J. Hydrol.,* **2016,** *17,* 2477-2491.
16. Wang, J.; Petersen, W. A.; Wolff, D. B.; Bolvin, D.T.; Kirstetter, P.-E. Validation of satellite precipitation products using NMQ over the CONUS.  8th European Conference on Radar in Meteorology and Hydrology Garmisch-Partenkirchen, Germany, 2014.
17. Gebregiorgis, A. S.; Kirstetter, P.-E.; Hong, Y. E.; Gourley, J. J.; Huffman, G. J.; Petersen,W. A.; Xue, X.; Schwaller,M. R. To what extent is the day 1 GPM IMERG satellite precipitation estimate improved as compared to TRMM TMPART? *J. Geophysical Research: Atmospheres,* **2018,** *123,* 1694–1707. https://doi.org/10.1002/2017JD027606
18. Guo, H.; Chen, S.; Bao, A.; Behrangi, A.; Hong, Y.; Ndayisaba, F.; Hu, J.; Stepanian, P.M.  Early assessment of integrated multi-satellite retrievals for global precipitation measurement over China. *Atmos. Res.,* **2016,** *176–177,* 121–133. https://dx.doi.org/10.1016/j.atmosres.2016.02.020.

19. Tang, G.; Ma, Y.; Long, D.; Zhong, L.; Hong, Y. Evaluation of GPM Day-1 IMERG and TMPA version-7 legacy products over mainland China at multiple spatiotemporal scales. *J. of Hydrol.,* **2016,** *533,* 152–167. https:// doi.org/10.1016/j.jhydrol.2015.12.008

20. Xu, R.; Tian F.; Yang L.; Hu H.; Lu H.; Hou A. Ground validation of GPM IMERG and TRMM 3B42V7 rainfall products over southern Tibetan Plateau based on a high-density rain gauge network, *J. Geophys. Res. Atmos.,* **2017,** *122,* doi:10.1002/2016JD025418.

21. Wei, G.; Lü, H.; T. Crow, W.; Zhu, Y.; Wang, J.; Su, J. Evaluation of Satellite-Based Precipitation Products from IMERG V04A and V03D, CMORPH and TMPA with Gauged Rainfall in Three Climatologic Zones in China. *Remote Sens.* **2018,** *10,* 30. https://doi.org/10.3390/rs10010030

22. Ma, Q.; Li, Y., Feng, H.; Yu, Q., Zou, Y.; Liu, F.; Pulatov B. Performance evaluation and correction of precipitation data using the 20- year IMERG and TMPA precipitation products in diverse subregions of China*, Atmos. Res.,* **2021,** *249,* 105304.

23. Chen, F.; Li, X. Evaluation of IMERG and TRMM 3B43 monthly precipitation products over Mainland China. Remote Sens. **2016,** 8, 472; doi:10.3390/rs8060472

24. Murali Krishna, U.V.; Das, S.K.; Deshpande, S.M.; Doiphode, S.L.; Pandithurai, G. The assessment of Global Precipitation Measurement estimates over the Indian subcontinent. *Earth Space Sci.* **2017,** 4, 540–553. doi:10.1002/2017EA000285.

25. Islam, M.A. Statistical comparison of satellite-retrieved precipitation products with rain gauge observations over Bangladesh. *Int. J. Remote Sens.* **2018,** 39, 2906–2936, DOI: 10.1080/01431161.2018.1433890

26. Moazam, S.; Najafi M.R. A comprehensive evaluation of GPM-IMERG V06 and MRMS with hourly ground-based precipitation observations across Canada*. J. Hydrol.* **2020**, 549, 125929

27. Anjum, M.N. et al. Assessment of IMERG-V06 precipitation product over different hydro-climatic regimes in the Tianshan Mountains, North-Western China. *Remote Sens.* **2019,** 11, 2314; doi:10.3390/rs11192314

28. Huffman, G. J. The Transition in Multi-Satellite Products from TRMM to GPM (TMPA to IMERG), https://gpm.nasa.gov/sites/default/files/2020-10/TMPA-to-IMERG_transition_201002.pdf

29. Bolvin, D. T.; Huffman, G. J. Transition of 3B42/3B43 research product from monthly to climatological calibration/adjustment**. 2015** NASA TRMM Doc., 11 pp. https://pmm.nasa.gov/sites/default/files/document_files/3B42_3B43_TMPA_restart.pdf

30. Schneider, U.; Becker, A.; Finger, P.; Rustemeier, E.; Ziese, M. GPCC Full Data Monthly Product Version 2020 at 0.25°: Monthly Land-Surface Precipitation from Rain-Gauges built on GTS-based and Historical Data. DOI: 10.5676/DWD_GPCC/FD_M_V2020_025

31. Joyce, R. J.; Xie, P. Kalman filter–based CMORPH. *J. Hydrometeor.,* **2011**, *12,* 1547–1563, https://doi.org/10.1175/JHM-D-11-022.1.

32. Hong, Y.; Hsu K.-L.; Sorooshian, S.; Gao,X. Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System. *J. Appl. Meteor.,* **2004,** *43,* 1834–1853, https://doi.org/10.1175/JAM2173.1.

33. Tan, J.; Huffman, G.J.; Bolvin, D.T.; Nelkin, E.J. IMERG V06: Changes to the morphing algorithm. *J. Atmos. Ocean. Technol.* **2019***, 36,* 2471–2482.

34. Zhang, J.; Howard, K.; Langston, C.; Kaney, B.; Qi, Y.; Tang, L.;…, Kitzmiller, D. Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.,* **2016,** *97,* 621–637. https://doi.org/10.1175/BAMS-D-14-00174.1

35. Kirstetter, P.-E., and Coauthors. Toward a framework for systematic error modeling of spaceborne precipitation radar with NOAA/NSSL ground radar–based National Mosaic QPE. *J. Hydrometeor.,* **2012,** *13,* 1285–1300, https://doi.org/10.1175/JHM-D-11-0139.1.

36. Zhang, J.; Qi, Y.; Langston, C.; Kaney, B. Radar quality index (RQI) – A combined measure for beam blockage and VPR effects in a national network. Proceedings of International Symposium on Weather Radar and Hydrology, Exeter, UK, Apr **2011**. IAHS Publ., *351,* 388–393.

37. Wilks, D. S. *Statistical Methods in the Atmospheric Sciences.3rd ed. International Geophysics Series,* Vol. 100, Academic Press, **2011;** 704 pp.

38. Tian, Y.; Peters-Lidard, C.D.; Eylander, J.B.; Joyce, R.J.; Huffman, G.J.; Adler, R.F.; Hsu, K.; Turk, F.J.; Garcia, M.; Zeng, J. Component analysis of errors in satellite-based precipitation estimates. *J. Geophys. Res.* **2009,** *114 (D24),* 1–15. http://doi.org/10.1029/2009JD011949.

39. Habib, E.; Larson, B.F.; Graschel, J. Validation of NEXRADmultisensor precipitation estimates using an experimental dense rain gauge network in south Louisiana. *J. Hydrol.* **2009,** *373 (3–4),* 463–478. http://doi.org/10.1016/j.jhydrol.2009.05.010.

40. Kirstetter, P.-E.; Gourley, J.; Hong,Y.; Zhang,J; Moazamigoodarzi, S.; Langston, C.; Arthur, A. Probabilistic precipitation rate estimates with ground-based radar networks, *Water Resour. Res.,* **2015,** *51,* 1422–1442, https://doi.org/10.1002/2014WR015672.

41. Tang, L.; Tian, Y.; Yan, F.; Habib, E. An improved procedure for the validation of satellite-based precipitation estimates. *Atmos. Res.* **2015,** *163,* 61-73.

42. Shige, S.; Kida, S.; Ashiwake, H.; Kubota, T.; Aonashi, K. Improvement of TMI rain retrievals in mountainous areas. *J. Appl. Meteorol. Climatol.* **2013,** *52,* 242–254.

43. Kummerow, C.; Brown, P. GPROF: Progress towards an improved orographic precipitation product. PMM science team meeting 2020.

819    44. Ringerud, S.; Peters-Lidard, C; Munchak, J.; You, Y.  Applications of dynamic land surface information for
820        passive    microwave    precipitation    retrieval. *J. Atmos. Oceanic. Technol.,* **2021,** *38,* 167-180,
821        https://doi.org/10.1175/JTECH-D-20-0048.1

822    45. Cui, W. et al. Can the GPM IMERG Final Product Accurately Represent MCSs' Precipitation
823        Characteristics over the Central and Eastern United States? *J. Hydrometeorol*. **2020,** *21,* 39-57. DOI:
824        10.1175/JHM-D-19-0123.1

825    46. Gupta, H. V.; Kling, H.; Yilmaz, K. K.; Martinez, G. F. Decomposition of the mean squared error and NSE
826        performance criteria: Implications for improving hydrological modelling.  *J. Hydrol*., **2009,** *377(1-2),* 80-
827        91. doi:10.1016/j.jhydrol.2009.08.003.

828    47. Kling, H.; Fuchs, M.; Paulin, M. Runoff conditions in the upper Danube basin under an ensemble of
829        climate change scenarios. *J. Hydrol.,* **2012,** 424, 264-277, doi:10.1016/j.jhydrol.2012.01.011

830    48. Hou, A. Y., and Coauthors. The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.,*
831        **2014,** *95,* 701–722, https://doi.org/10.1175/BAMS-D-13-00164.1.

832    49. Skofronick-Jackson G.; Kirschbaum D.; Petersen W.; et al. The Global Precipitation Measurement (GPM)
833        mission's scientific achievements and societal contributions: reviewing four years of advanced rain and
834        snow observations. *Q. J. R. Meteorol. Soc.* **2018,** *144 (Suppl. 1)*, 27–48. https://doi.org/10.1002/qj.3313

835    50. Chen, H.; Yong, B.; Qi, W.; Wu, H.; Ren., Li; Hong, Y. Investigating the evaluation uncertainty for satellite
836        precipitation   estimates   based   on   two   different   ground   precipitation   observation   products. *J.*
837        *Hydrometeorol.,* **2020,** *21,* 2595-2606. https://doi.org/10.1175/JHM-D-20-0103.1.

838