Geophysical Research Letters

RESEARCH LETTER
10.1029/2023GL105343

A New Event-Based Error Decomposition Scheme for Satellite Precipitation Products

Runze Li[1], Clement Guilloteau[1], Pierre-Emmanuel Kirstetter[2,3], and Efi Foufoula-Georgiou[1,4]

[1]Department of Civil and Environmental Engineering, University of California, Irvine, Irvine, CA, USA, [2]Hydrometeorology and Remote Sensing Laboratory, University of Oklahoma, Norman, OK, USA, [3]NOAA/Severe Storms Laboratory, Norman, OK, USA, [4]Department of Earth System Science, University of California, Irvine, Irvine, CA, USA

Key Points:

- A new event-based error decomposition scheme for satellite precipitation products is proposed, dividing the total bias into 10 components
- Inaccuracies in event occurrence, timing, and intensity contribute on average to about 30%, 20%, and 50% of the total amount bias
- A large fraction of errors is associated with events starting/ending too early in the satellite product

Supporting Information:

Supporting Information may be found in the online version of this article.

Correspondence to:

R. Li,
runzl11@uci.edu

Citation:

Li, R., Guilloteau, C., Kirstetter, P.-E., & Foufoula-Georgiou, E. (2023). A new event-based error decomposition scheme for satellite precipitation products. *Geophysical Research Letters*, *50*, e2023GL105343. https://doi.org/10.1029/2023GL105343

Received 5 JUL 2023
Accepted 4 NOV 2023

Abstract Understanding the nature and origin of errors in satellite precipitation products is important for applications and product improvement. Here we propose a new error decomposition scheme incorporating precipitation event (continuous rainy periods) information to characterize satellite errors. Under this framework, the errors are attributed to the inaccuracies in event occurrence, timing (event start/end time), and intensity. The Integrated MultisatellitE Retrieval for Global Precipitation Measurement (IMERG) is used as our test product to apply the method over CONUS. The above-listed factors contribute approximately 30%, 20%, and 50% to the total bias, respectively. Significant asymmetry exists in the temporal distribution of biases throughout events: early event endings cause threefold more precipitation amount bias than late event beginnings, while early event beginnings cause fourfold more bias than late event endings. Dominant contributors vary across seasons and regions. The proposed error decomposition provides insight into sources of error for improved retrievals.

Plain Language Summary Satellite remote sensing offers unique capabilities to map global precipitation at daily to sub-daily scales, important for hydrologic applications and decision-making. However, inherent uncertainties and errors in satellite precipitation products necessitate a comprehensive understanding of their characteristics and sources for effective utilization and enhancement. Here, we propose a new event-based error decomposition scheme to characterize satellite errors. This approach is based on the understanding that precipitation occurs as individual events (continuous rainy periods); hence, any quantitative inaccuracy in a satellite product can be attributed to the imperfect delineation of diverse event facets: (a) occurrence (completely missed/falsely detected events), (b) timing (wrong start/end times of the detected events), and (c) intensity (inaccurate precipitation rates during the events). We apply the method to a popular high-resolution satellite product, the Integrated MultisatellitE Retrieval for Global Precipitation Measurement (IMERG) over CONUS. Results show that, nationwide, the above three error types contribute on average about 30%, 20%, and 50% to the total bias, respectively. A large fraction of errors is associated with events starting/ending too early in the satellite product. Dominant error types are season- and region-dependent. The event-based error breakdown offers potential to diagnose error sources and guide algorithm improvement for satellite precipitation products.

1. Introduction

Recent decades have witnessed the rapid development of remote sensing of precipitation, which is nowadays providing indispensable regional and global data to advance research and applications in the hydrologic, atmospheric, and climate sciences (AghaKouchak et al., 2015; Fassoni-Andrade et al., 2021; Foufoula-Georgiou et al., 2020). Nevertheless, due to its indirect measurement nature, satellite-derived precipitation is inevitably subject to both random and systematic errors (Smith et al., 2006), prompting significant research toward understanding the error sources to guide algorithm improvements (Guilloteau et al., 2022; Kirstetter, Karbalaee, et al., 2018; Tian et al., 2009; Wright et al., 2017).

The discontinuous nature of precipitation, as contrasted to other variables such as temperature, humidity, etc., makes it necessary to decompose the estimation errors into, first, errors arising from misclassification of precipitation occurrence (i.e., *Miss*/*False Alarm*), and subsequently, errors arising from inaccurately quantifying precipitation rates when precipitation occurrence is properly detected (i.e., *Hit*). Based on this categorization, one of the most popular error characterization methods for satellite precipitation products involves decomposing satellites' *Total Bias* over a given time period into three components: "*Miss Bias*" (negative bias introduced by "*Misses*"),

"*False Bias*" (positive bias introduced by "*False Alarms*"), and "*Hit Bias*" (negative/positive bias introduced by inaccurate precipitation rates of "*Hits*") (Chen et al., 2021; Su et al., 2018; Tian et al., 2009). This decomposition approach extracts additional information as compared to only assessing satellites' mean errors, providing valuable insight into the nature and origins of errors. Recent studies have taken one step further by subdividing *Hit Bias* into *Hit Positive Bias* (when the satellite product overestimates rain rates) and *Hit Negative Bias* (when the satellite product underestimates rain rates) (Chaudhary & Dhanya, 2020, 2021; Zhang et al., 2021).

However, the information gleaned from the three/four component error categories based solely on detection and intensity is still limited. Here we propose a more comprehensive error decomposition framework that incorporates a crucial, yet overlooked, aspect of precipitation—it essentially occurs as individual "events" with specific timing, duration, and intensity (Dunkerley, 2008; Hanel & Maca, 2014). Such an event-based error characterization approach can furnish additional useful information on satellite errors, especially at high temporal resolutions. Specifically, the *Miss/False Bias* can be further sub-categorized depending on their occurrence throughout a precipitation event: at the beginning (satellite-derived event starting too early/too late), or at the end (satellite-derived event ending too early/too late); or be attributed to the satellite's entirely overlooking a "real" event or inventing a nonexistent one. Furthermore, the event-based error information can be associated with more intrinsic factors, such as satellite performance fluctuations induced by changing cloud microphysics during precipitation processes (Bouniol et al., 2016), or spatial-temporal mismatches between satellite and ground observations (Li et al., 2023). Therefore, the proposed approach offers opportunities for deeper exploration and attribution of error sources.

In what follows we describe the new event-based error decomposition scheme for satellite precipitation products by factoring errors due to inaccuracies in event occurrence, timing, and intensity. We apply this scheme to evaluate the high-resolution Integrated MultisatellitE Retrieval for Global Precipitation Measurement (GPM) (IMERG) at the half-hourly scale over the Continental United States (CONUS) from 2018 to 2020.

2. Data

IMERG is the primary multi-satellite merged precipitation product developed by the U.S. GPM Science Team with a spatial-temporal resolution of 0.1°, 0.5 hr and a global coverage (Huffman et al., 2019a). The version we use is IMERG V06B Final Run Half Hourly product (Huffman, Stocker, et al., 2019). This product integrates data from multiple passive microwave (PMW) and infrared (IR) sensors, ensuring consistency and accuracy through intercalibration with state-of-the-art precipitation measurement instruments onboard the GPM Core Observatory, and is ultimately adjusted by the monthly gauge analysis. IMERG is recognized as one of the most accurate high-resolution satellite precipitation data sets available (Guilloteau et al., 2021; Pradhan et al., 2022; Tang et al., 2020), and has been widely employed in various applications (Nie & Sun, 2022; Orland et al., 2022; Zhang et al., 2023).

The Ground Validation-Multi-Radar/Multi-Sensor (GV-MRMS) data product is a high-quality ground-based radar-gauge merged quantitative precipitation estimation (QPE), which serves as the ground reference data here (Kirstetter, Petersen, & Gourley, 2018). It combines data from 180 ground-based radar and about 7,000 rain gauges, along with model analyses (Zhang et al., 2016). Designed for GPM ground validation, GV-MRMS is further quality-controlled, adjusted, and integrated to the IMERG's resolution (0.1°, 0.5 h) (Kirstetter et al., 2012, 2014), and has been extensively used in GPM validation studies (e.g., Derin & Kirstetter, 2022; Guilloteau et al., 2021; Tan et al., 2022). Additionally, the Radar Quality Index (ranging from 0 to 100 (best)) generated concurrently with GV-MRMS QPE, is employed here to filter out low-quality estimates (Petersen et al., 2020). This index characterizes the uncertainty in radar QPE by considering factors such as distance from the radar, beam blockage, and altitude of the freezing level. In this study, gridboxes with an average RQI < 60 during the study period were removed from the analysis.

3. Methods

In this study, a precipitation event is defined as an uninterrupted sequence of half-hourly time steps with non-zero precipitation ($\geq$0.1 mm/hr) at any given pixel of IMERG or GV-MRMS (Li et al., 2023). By this definition, all precipitation within a certain period could be perceived as an assembly of individual events. Therefore, any half-hourly bias throughout this period can be attributed to the flawed representations of distinct event aspects,

which can be summarized as inaccuracies in event (a) occurrence (i.e., completely missed/falsely detected events), (b) timing (i.e., wrong start/end times of the detected events), and (c) intensity (i.e., incorrect precipitation rates during the events). Based on these event-related inaccuracies, the event-based error decomposition is proposed (Figure 1).

Figure 1a depicts two hypothetical precipitation time series from GV-MRMS and IMERG, respectively (Figure 1a, upper panel). The traditional and event-based error components are sequentially displayed, based on these two series (Figure 1a, middle and bottom panels). Depending on the detection scenarios, the IMERG-derived precipitation at each time step (here half-hour) may be a "*Miss*" (GV-MRMS: rainy, IMERG: dry), a "*False Alarm*" (GV-MRMS: dry, IMERG: rainy), or a "*Hit*" (GV-MRMS: rainy, IMERG: rainy) (Figure 1a, upper panel), which result in a "*Miss Bias*", a "*False Bias*," or a "*Hit Bias*" (when the rain rate is inaccurate), respectively (Figure 1a, middle panel). Referring to the original time series (Figure 1a, upper panel), despite the varying causes and timings of errors, they are all grouped into these three categories, thus providing only limited diagnostic information.

In contrast, by integrating event occurrence, timing, and intensity, satellite errors are further subdivided into 10 categories, which are schematically represented within the time series (Figure 1a, lower panel). The corresponding explanation of each error type is articulated more specifically in Figure 1b. From the perspective of total error breakdown, Figure 1c further elucidates the relationships among the total bias, the traditional error components, and the proposed event-based error components. As shown, the four "*Miss-*," four "*False-*," and two "*Hit-*" components in the event-based scheme provide more detailed diagnostics and dissect the "*Miss Bias*," "*False Bias*" and "*Hit Bias*" aggregate errors of the traditional scheme, correspondingly. Further, the four "*Miss-*" components in combination with the "*Hit-Negative*" component comprise the total underestimation of IMERG, whereas the four "*False-*" components along with "*Hit-Negative*" bias make up the total overestimation. Ultimately, the total underestimation and overestimation collectively form the total bias.

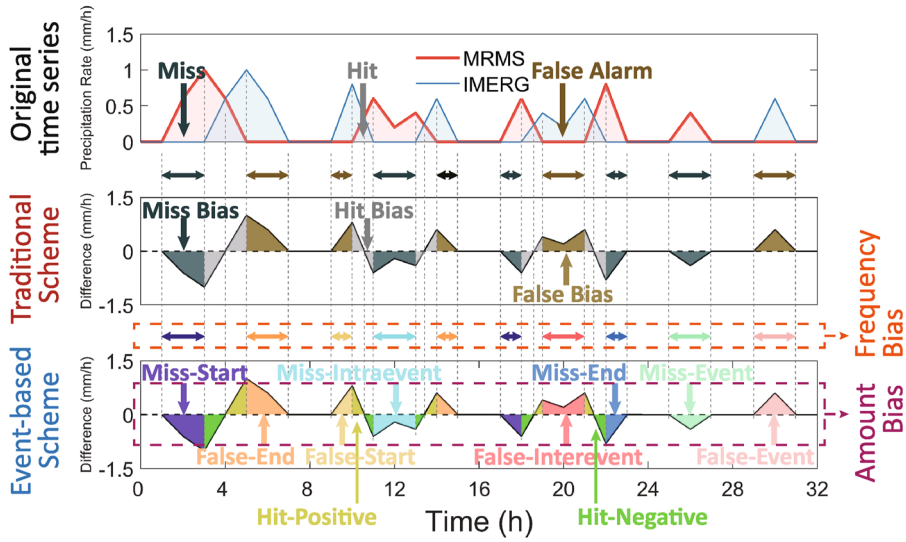
Here, we perform this error analysis over CONUS in different seasons to analyze the regional and seasonal error components and their respective contributions to the total errors in terms of both precipitation amount and frequency (i.e., bias in rainy hours). The intensity of each error component is also shown, calculated as the mean absolute half-hourly intensity difference (IMERG - MRMS) during the occurrence of each specific error type.

4. Results

Figure 2 shows the spatial patterns of total precipitation amount bias per year (the original values of IMERG and GV-MRMS are provided in Figure S1 in Supporting Information S1), as well as each error component from both the traditional and event-based error decomposition schemes. Compared to *Total Bias* (Figure 2a), all three error components from the traditional scheme exhibit distinct spatial patterns (Figures 2b–2d), which demonstrates the additional error information provided by the decomposition. Consequently, the extended event-based scheme contributes to a deeper knowledge of error characteristics (Figures 2b1–2b4, 2c1–2c2, and 2d1–2d4). In particular, among the four subcomponents of *Miss Bias* (Figures 2b1–b4), the main contributors are the misses due to IMERG events ending too early (*Miss-End*) and the completely ignored events (*Miss-Event*) (Figures 2b2 and 2b4). The *Miss-End* and *Miss-Event* error components are both noticeable in the northeast, which is probably owing to the less accurate retrieval of snow and coastal precipitation in these areas (Derin et al., 2021; You et al., 2017). *Miss-Event* is additionally notable near the Rocky Mountains (Figure 2b4), potentially attributable to satellites' oversight of certain orographic warm rain processes entirely (Figure S2b$_2$ in Supporting Information S1), an effect induced by the weak signatures on brightness temperatures for liquid hydrometeors aloft (Derin & Kirstetter, 2022). However, this pattern is not reflected in *Miss Bias* due to the compensatory effects from the smaller *Miss-End* error in this region (Figures 2b and b2), which primarily results from the lower precipitation intensity in this semi-arid area (Figure S3b$_2$ in Supporting Information S1), highlighting the additional diagnostics that the proposed error decomposition provides.

For the various "*False-*" components (Figures 2d1–2d4), the bias from completely invented events also plays a significant role (*False-Event*) (Figure 2d4). However, the other major contributor is the satellites' false alarms around the event beginnings, that is, *False-Start* (Figure 2d2). Combined with the pronounced *Miss-End* bias, it is easy to infer an error pattern: when IMERG detects an event, it tends to both start earlier (resulting in false alarms) and end earlier (resulting in misses), which is in accordance with Li et al. (2023). The distinct spatial patterns of *False-Start* and *False-Event*, evidenced by a spatial correlation coefficient of only 0.41, validate the

(a) Diagrammatic Error Component Definitions



(b) Descriptive Error Component Definitions

Components	Causes
Miss-Start	IMERG event starts too late
Miss-End	IMERG event ends too early
Miss-Intraevent	IMERG falsely generates an intermittency within the event (erroneously splits the event)
Miss-Event	IMERG misses the whole event
Hit-Negative	IMERG underestimates the intensity during the event
Hit-Positive	IMERG overestimates the intensity during the event
False-Start	IMERG event starts too early
False-End	IMERG event ends too late
False-Interevent	IMERG misses the intermittency between two events (erroneously merges the events)
False-Event	IMERG falsely detects a whole event

(c) Error Decomposition Scheme

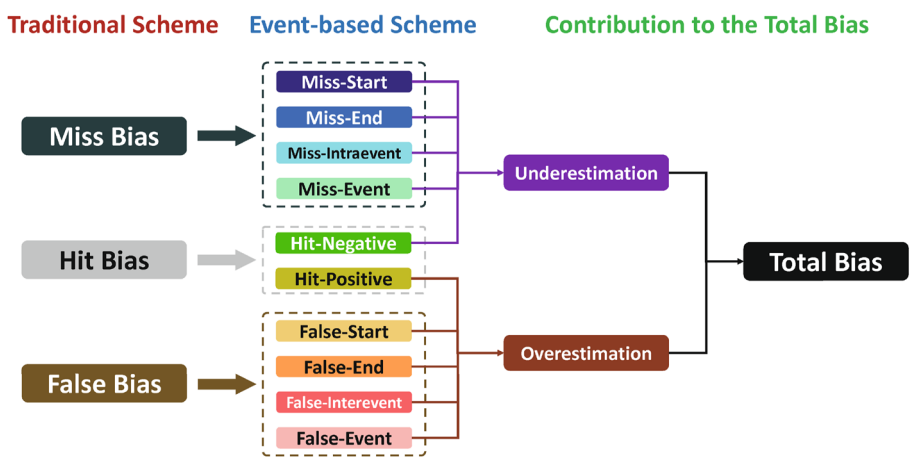


Figure 1. (a) A pair of hypothetical precipitation half-hourly time series from Integrated MultisatellitE Retrieval for Global Precipitation Measurement and GV-MRMS (considered as the reference product) (upper panel), as well as the traditional (middle panel) and event-based (lower panel) error components defined based on the differences of the two products. (b) A detailed descriptive explanation of each event-based error component. (c) The proposed event-based error decomposition scheme composed of 10 diagnostic error components and its relationship with the traditional scheme (left) and the total bias (right).

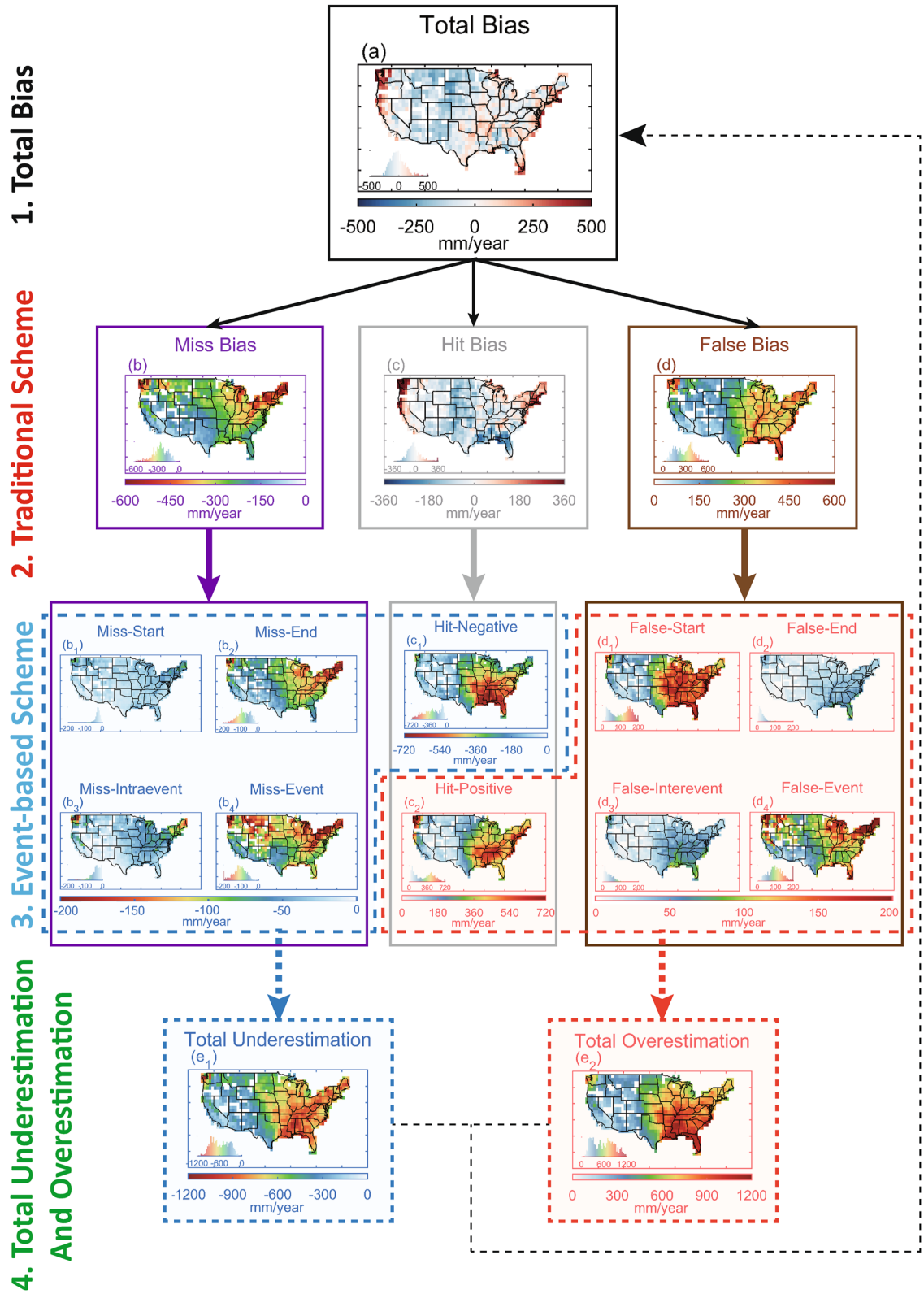


Figure 2. Spatial patterns of (a) total precipitation amount bias per year; (b)–(c) three bias components from the traditional error decomposition scheme; (b_1–b_4, c_1–c_2, and d_1–d_4) 10 bias components from the proposed event-based error decomposition scheme; (e_1–e_2) total underestimation and overestimation, with the components contributing to each, consistent with Figure 1c. The values are computed from all the half-hourly data during the study period of 2018–2020 and displayed in 1° × 1° gridboxes.

successful extraction of greater information content and diagnostics via the proposed decomposition (Figures 2d1 and 2d4). In detail, *False-Start* is significantly high across the eastern U.S. (Figure 2d1), predominantly attributed to a higher frequency in the southeast and increased intensity in the Great Plains, respectively (Figures S2d$_1$ and S3d$_1$ in Supporting Information S1). However, *False-Event* hotspots are observed in the northwest, northeast, and southeast corners of the U.S., which are primarily coastal regions with inherent greater bias (Figure 2d4) (Derin et al., 2021). The spatial pattern of the less significant *False-End* and *False-Interevent* bias also differs from the above two components, exhibiting higher values throughout the southeast (Figures 2d2–2d3). These results suggest that the overall *False Bias* has significant variations in its sources among different areas (Figure 2d).

Regarding *Hit Bias*, although both the U.S. central plains and southeast exhibit comparable magnitudes of negative errors (Figure 2c), further subdivision indicates the substantially more severe *Hit-Negative* and *Hit-Positive* bias in the southeast (Figures 2c1–2c2). This highlights again the role of error segmentation in uncovering additional insights. As a constituent of *Total Underestimation* (Figure 2e1), *Hit-Negative* bias exhibits a nearly inverse pattern in the eastern U.S. compared to the other four constituents (i.e., the four "*Miss-*" components) (Figures 2b1–2b4, and 2c1), partially due to a certain degree of opposite patterns between the missed precipitation and hit precipitation in these regions (i.e., more misses in the snowy northeastern mountain area may naturally lead to fewer hits, while more hits in the relatively warm southern plains could result in fewer misses). Consequently, the primary cause of the pronounced *Total Underestimation* in the eastern U.S. is the substantial misses in the northeast and the underestimated intensity of detected precipitation in the southeast, respectively (Figure 2e1). *Total Overestimation* exhibits a similar "higher-east, lower-west" pattern compared to *Total Underestimation*, as their sum corresponds to the relatively low *Total Bias* after gauge correction (Figure 2a) (Huffman et al., 2019a). However, the pronounced overestimation in the southeast is a collective outcome of the mistiming (*False-Start*, *False-End*, and *False-Interevent*) and overestimated intensity (*Hit-Positive*), (Figures 2d1–2d3, and 2c2), whereas it is primarily engendered by the too-early event beginnings (*False-Start*) and entirely false events (*False-Event*) in the southwest (Figures 2d1 and 2d4).

Figure 3 further visualizes the relative importance of each error component by quantifying their fractional contributions to the total underestimation/overestimation of precipitation amount and frequency, which could inform prioritizing algorithm improvements. The mean intensity of each error component is also presented. As shown, the total precipitation amount underestimation is roughly equally split between the underestimated intensity of the detected precipitation (*Hit-Negative*) and various types of misses (Figure 3a1). Of the latter's 50% contribution, about 30% comes from the wrong start/end time of the detected events (*Miss-Start*, *Miss-End*, and *Miss-Intraevent*), while the remaining 20% results from completely missed events (*Miss-Event*). A similar partition goes for the precipitation amount overestimation (Figure 3a2). Besides, the fractional contributions also clearly illustrate the asymmetric satellite errors in capturing precipitation events, with events ending-too-early (*Miss-End*) contributing to the total amount underestimation threefold more than events starting-too-late (*Miss-Start*), but events staring too early (*False-End*) contributing to the total amount overestimation fourfold than events ending-too-early (*False-Start*).

For the composition of precipitation frequency bias (Figures 3b1–3b2), the 0% contribution of the "*Hit-*" components causes a natural expansion in other components' contributions. Consequently, satellites' disregarding partial detected events (*Miss-Start*, *Miss-End*, and *Miss-Intraevent*) causes about 50% frequency underestimation, while their ignoring the entire precipitation processes (*Miss-Event*) accounts for the other half. The asymmetry is also evident in the frequency graph, with *Miss-End* contributing about 2.5 times higher error than *Miss-Start*. In conjunction with the intensity (Figure 3c1), it is understandable that precipitation in the middle of events (*Miss-Intraevent*) with higher intensity has a lower chance of being missed, while the completely ignored events (*Miss-Event*) are those with relatively low mean intensity, as this pattern aligns with satellite detection rationale (Hou et al., 2018). However, the concurrently higher frequency and intensity of missed precipitation near the event ending times (*Miss-End*), as compared to the beginnings (*Miss-Start*), unexpectedly deviates from this "lower-intensity, higher-frequency" pattern of missed precipitation. This suggests that the primary mechanism for the pronounced *Miss-End* may not be the neglect of abundant weak precipitation stemming from reduced ice content in clouds during the event dissipation stage (Braga & Vila, 2014), but rather attributable to alternative factors. As suggested by Li et al. (2023), one possibility is that satellite-derived events are systematically earlier than the "actual" events, resulting in infrequent misses at the event beginnings and overlooking even intense precipitation at the ends. The delay between the detection of hydrometeor signals in clouds and the actual arrival of precipitation on the ground, together with the spatial mismatch stemming from the inclined angle of satellite

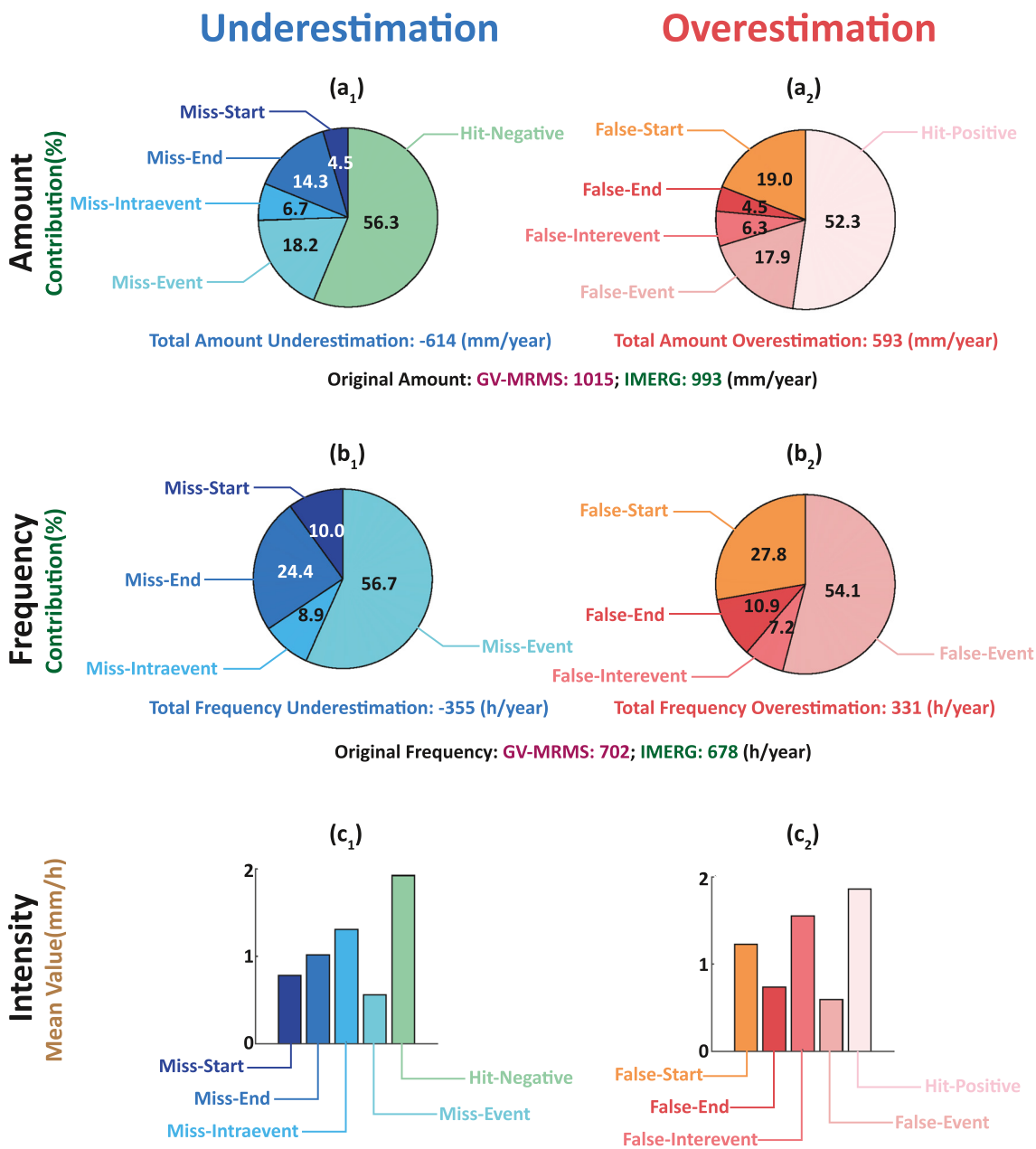


Figure 3. The contribution of each event-based error component to the total (a_1–a_2) amount and (b_1–b_2) frequency underestimation/overestimation, as well as (c_1–c_2) the mean intensity of each component. The values are calculated from all the data during the study period of 2018–2020 over CONUS. The annual mean total underestimation/overestimation and the original amount/frequency of GV-MRMS and Integrated MultisatellitE Retrieval for Global Precipitation Measurement are also shown for reference. It is noted that the total underestimation amount almost cancels the total overestimation amount, hiding error diagnostics and highlighting the importance of splitting the error to further components.

observations (i.e., parallax shift) that translates into a temporal difference, might be the possible underlying causes (Guilloteau et al., 2018). Undoubtedly, further investigation is warranted to substantiate this conjecture or identify more fundamental causes. Moreover, a similar pattern is manifested in the cases of precipitation overestimation (Figures 3b2 and 3c2).

We further analyze the error composition structure by different regions and seasons to understand the relative importance of different error components under varying conditions. Two regions and seasons with significant contrasts are selected for illustration (Figure 4), with more comprehensive information in Supporting Information S1 (Figures S4 and S5). Specifically, the predominant factor for SE is the inaccurate intensity

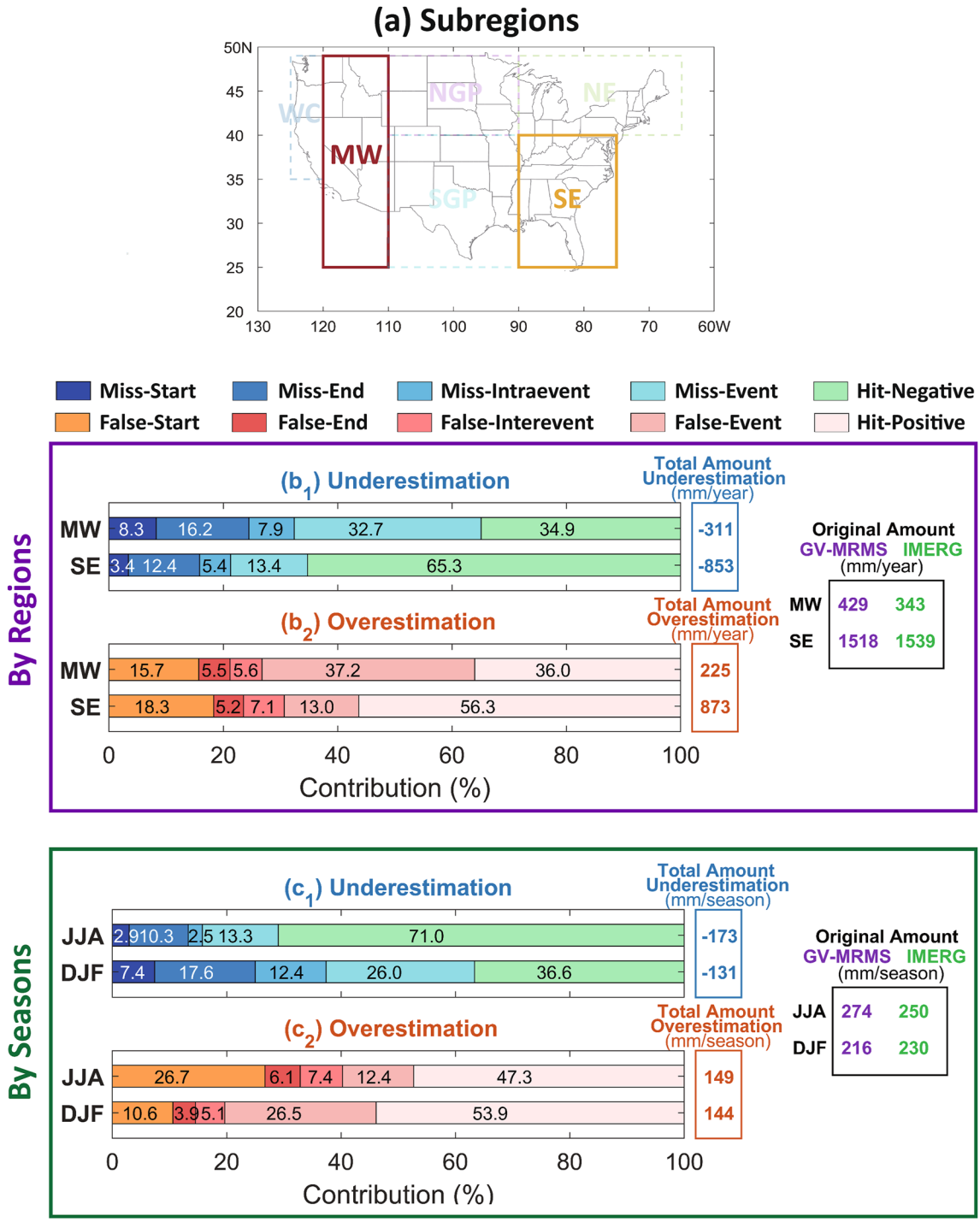


Figure 4. (a) The two primary focus subregions: **MW**: Mountain West and **SE**: Southeast (highlighted), alongside secondary subregions (semi-transparent, full results in Figure S4 in Supporting Information S1), following Cui et al. (2017). The contribution of each error component to the total amount underestimation/overestimation across the (b_1–b_2) two primary subregions (for the full year) (**MW** and **SE**) and (c_1–c_2) two primary seasons (for the full CONUS) (**JJA** and **DJF**). The corresponding annual/seasonal mean total underestimation/overestimation, and original amount of GV-MRMS and Integrated MultisatellitE Retrieval for Global Precipitation Measurement are displayed for reference.

(*Hit-Negative/Hit-Positive*) for both total underestimation and overestimation, while that for MW is all forms of "*Miss-*"/"*False-*" components (Figures 4b1–4b2). The primary cause for the latter is a significantly higher number of completely missed/false events in MW, which is approximately 2.5 times greater than those in SE.

This demonstrates significant disparities in error composition between low-altitude, plain, humid regions with minimal snowfall, and high-altitude, mountainous, semi-arid areas.

For seasonal analysis, the underestimation in summer is primarily attributed to the intensity underestimation (*Hit-Negative*), whereas that in winter is predominantly due to the "*Miss-*" components (Figure 4c1). The former might be attributed to satellites' inherent under-representativeness of intense precipitation prevalent in summer due to the "regression-to-mean" tendency inherent in the Bayesian inversion, coupled with the "smoothing effect" of IMERG's "morphing" interpolation algorithm (Rajagopal et al., 2021; Tan et al., 2021). In contrast, the latter is related to the large uncertainties over ice/snow covers in winter, where less accurate IR retrievals are used instead of PMW observations (Huffman et al., 2019a; Passive Microwave Algorithm Team Facility, 2017). For the overestimation (Figure 4c2), the severity of the overestimated intensity (*Hit-Positive*) remains similar in both seasons. However, the summer exhibits significant errors due to too early event start times (*False-Start)*, while the bias in winter is more prominently influenced by the totally invented events (*False-Event)*. The former might be related to the enhanced lengthening effect of interpolation procedures in the merging algorithm for the more intense, early peaking summer events (Li et al., 2018, 2021), while the latter still relates to the ice/snow surfaces.

Besides, a similar attribution based on time of day is displayed but exhibiting less pronounced diurnal variations (Figure S6 in Supporting Information S1). The composition predicated on the sensor sources of IMERG reveals more information (Figure S7 in Supporting Information S1). For instance, Special Sensor Microwave Imager/Sounder (SSMIS) displays a more noticeable intensity underestimation, while for Sounder for Probing Vertical Profiles of Humidity (SAPHIR), *Miss-Start* surpasses *Miss-End*, diverging from all other sources (though SAPHIR's representation within IMERG is minimal). Additionally, the smoothing effect in generating "*morph*"/"*moprh + IR*" sources leads to a substantially greater intensity underestimation compared to PMW sources (Huffman et al., 2019a). Excluding interpolated sources, PMW retrievals mostly exhibit a more pronounced asymmetry in the temporal structure of their errors (i.e., the contrast between the "*-Start*" and the "*-End*" components) than that observed in IMERG overall. This occurs as the interpolation generally exerts a symmetric "smoothing effect" near both the start and end of events, thus reducing any asymmetry. Furthermore, the utilization of *IR-only* sources over snow/ice surfaces proportionally magnifies various "*Miss-*" bias, but uniquely triggers numerous *False-Event* among all "*False-*" components, making it a major contributor to the total underestimation (85.7%). The above observations could explain the error composition pattern in winter (Figures 4c1–4c2).

5. Conclusions and Discussion

In this study, we introduce a novel event-based error decomposition scheme for satellite precipitation products, which segments the total bias within a certain time period into 10 independent components. This scheme is founded on the understanding that viewing precipitation as event-based constructs, all fine scale (half-hourly here) satellite biases can be attributed to the inaccuracies in event occurrence (completely missed/falsely detected events), timing (wrong start/end times of the detected events), and intensity (inaccurate precipitation rates during the events). The scheme is applied over CONUS, with IMERG as the sample product for analysis and GV-MRMS as the reference. The attribution results show that about 20%, 30%, and 50% of the total national amount bias comes from the above three causes, respectively. The error composition exhibits significant asymmetry concerning precipitation events. Specifically, missed precipitation from premature event end times in the satellite product is about threefold than that from delayed event start times, whereas false precipitation from premature event start times is about fourfold than that from delayed event end times. The dominant error components vary significantly among regions, seasons, and sensor sources.

Compared to our previous work (i.e., Li et al. (2023)) that evaluated various characteristics of precipitation events to assess the suitability and limitations of satellite precipitation products for hydrologic applications, this study instead approaches the problem from the perspective of error diagnosis and algorithm improvement, aiming at quantifying how the revealed inaccuracies in capturing precipitation event features are mapped onto satellite fine-scale (e.g., half-hourly/hourly) errors (as illustrated in Figures 1a and 1b), and whether we can gain a new understanding of satellite errors through the lens of precipitation events. By our event-based error decomposition, error components with distinct patterns are dissected. Each component might be linked to diverse underlying causes, thus providing additional insights into satellite error reduction and performance enhancement for tailored improvement strategies. For example, the decomposition allows for the extraction of completely missed and

false events, whose patterns suggest that they might be linked to the faint spectral signature of shallow nimbus clouds with low ice content (warm rain), and to the difficulty of distinguishing between the surface emission and the atmospheric component in satellited-measured radiances over frozen or snow-covered surfaces (Derin & Kirstetter, 2022; You et al., 2017), respectively. Such biases, from completely ignored or fabricated events, which account for about 20% of the total bias, might be particularly challenging to correct and require focused attention.

Regarding the errors stemming from incorrect event start/end times, the aforementioned analysis suggests that the underlying reason might be a prevalent early timing of satellite-derived events compared to the "real" events, as identified by Li et al. (2023). One possible explanation is the spatial-temporal mismatches arising from differences in observation methods, observation geometry and sampling between satellites and ground-based radars (e.g., parallax shift, or delay caused by hydrometeors' fall time) (Guilloteau et al., 2018). These findings corroborate our previous assertion made in Li et al. (2023) that "numerous seemingly minor event time shifts may contribute significantly to the total bias," as "premature event onset. alone accounts for about 20%/30% of total amount/frequency bias, respectively. Consequently, minor post-processing adjustments in event timing could potentially mitigate a substantial portion of the overall bias. Moreover, the asymmetric satellite performance at the beginning and end of events may still be related to the varying accuracy in satellite responses to the changing thermodynamic and microphysical properties of clouds during the precipitation life cycle (Anagnostou et al., 2013; Bouniol et al., 2016). For example, for mesoscale convective systems, robust convective precipitation at initial stages is more readily detected by satellites (corresponding to the low *Miss-Start* and high *False-Start* error components), whereas the dissipation stages primarily feature stratiform precipitation with reduced detectability (corresponding to the high *Miss-End* and low *False-End* error components) (Anagnostou et al., 2013). This suggests the potential to enhance precipitation retrieval accuracy by incorporating additional ancillary information indicative of event stages (e.g., cloud evolution, lightning activity, cloud liquid/ice water content, atmospheric instability) (Petkovic & Kummerow, 2017; Tadesse & Anagnostou, 2009).

The central focus of this research lies in introducing a universal method for characterizing and diagnosing errors in high spatio-temporal resolution satellite precipitation products. It must be noted that the outcomes, such as the configuration of error components, may vary due to multiple factors like the spatial and temporal resolution of the products, the study period, the selection of precipitation occurrence thresholds, and the specific definition of "precipitation events" (e.g., whether short intermittent periods within events are allowed, as e.g., could be the case if events were defined on the basis of precipitation attributed to distinct meteorological systems (e.g., Restrepo-Posada & Eagleson, 1982)). Therefore, a more comprehensive experimental investigation is needed for product developers to accurately pinpoint the error components, thus facilitating product improvements.

In future work, it will also be necessary to disentangle the diverse potential factors affecting satellite errors within the context of events, such as interpolation procedures, sensor inhomogeneity, spatial-temporal displacement, and cloud property evolution, to identify the root causes of biases. In addition, a similar decomposition in the spatial domain can be pursued. For example, if knowing that estimates derived from passive satellite measurements tend to overestimate the frequency of occurrence of low-intensity precipitation (Tian et al., 2018), it is interesting to know whether it is primarily due to satellites' tendency to invent non-existent precipitation systems, or due to the overestimation of precipitating area associated with properly detected precipitating systems. These error diagnostics are valuable for developing uncertainty prediction schemes, which is a longstanding challenge for the satellite precipitation community. Finally, the proposed method is applicable to various satellite precipitation products, as well as numerical weather and climate model outputs, offering additional insights into their respective errors.

Acknowledgments
This research was supported by NASA through the Global Precipitation Measurement Mission (Grant 80NSSC22K0597), Weather and Atmospheric Dynamics (Grant 80NSSC23K1304) and the Ground Validation Program (Grant 80NSSC21K2045). The research was also partially supported by the National Science Foundation (NSF) under Grant IIS 2324008. The authors would like to thank George Huffman and Jackson Tan from the NASA IMERG development team for their invaluable insights and comments regarding the interpretation of the results, which enhanced the depth of the manuscript.

Data Availability Statement

The GV-MRMS data can be accessed at NASA GHRC (Kirstetter, Petersen, & Gourley, 2018). The IMERG data are available at NASA GES DISC (Huffman et al., 2019b).

References

AghaKouchak, A., Farahmand, A., Melton, F. S., Teixeira, J., Anderson, M. C., Wardlow, B. D., & Hain, C. R. (2015). Remote sensing of drought: Progress, challenges and opportunities. *Reviews of Geophysics*, *53*(2), 452–480. https://doi.org/10.1002/2014rg000456

Anagnostou, E. N., Pathak, C. S., & Morales, C. A. (2013). Use of storm life cycle information and lightning data in radar-rainfall estimation. *Journal of Hydrologic Engineering*, *18*(2), 168–174. https://doi.org/10.1061/(Asce)He.1943-5584.0000557

Bouniol, D., Roca, R., Fiolleau, T., & Poan, E. (2016). Macrophysical, microphysical, and radiative properties of tropical mesoscale convective systems over their life cycle. *Journal of Climate*, *29*(9), 3353–3371. https://doi.org/10.1175/Jcli-D-15-0551.1

Braga, R. C., & Vila, D. A. (2014). Investigating the ice water path in convective cloud life cycles to improve passive microwave rainfall retrievals. *Journal of Hydrometeorology*, *15*(4), 1486–1497. https://doi.org/10.1175/Jhm-D-13-0206.1

Chaudhary, S., & Dhanya, C. T. (2020). Expanding contingency table for intensity and frequency based "true" detection of rainy events in precipitation datasets. *Atmospheric Research*, *244*, 105119. https://doi.org/10.1016/j.atmosres.2020.105119

Chaudhary, S., & Dhanya, C. T. (2021). An improved error decomposition scheme for satellite-based precipitation products. *Journal of Hydrology*, *598*, 126434. https://doi.org/10.1016/j.jhydrol.2021.126434

Chen, H. Q., Yong, B., Kirstetter, P. E., Wang, L. Y., & Hong, Y. (2021). Global component analysis of errors in three satellite-only global precipitation estimates. *Hydrology and Earth System Sciences*, *25*(6), 3087–3104. https://doi.org/10.5194/hess-25-3087-2021

Cui, W. J., Dong, X. Q., Xi, B. K., & Kennedy, A. (2017). Evaluation of reanalyzed precipitation variability and trends using the gridded gauge-based analysis over the CONUS. *Journal of Hydrometeorology*, *18*(8), 2227–2248. https://doi.org/10.1175/Jhm-D-17-0029.1

Derin, Y., & Kirstetter, P. E. (2022). Evaluation of IMERG over CONUS complex terrain using environmental variables. *Geophysical Research Letters*, *49*(19), e2022GL100186. https://doi.org/10.1029/2022GL100186

Derin, Y., Kirstetter, P. E., & Gourley, J. J. (2021). Evaluation of IMERG satellite precipitation over the land-coast-ocean continuum. Part I: Detection. *Journal of Hydrometeorology*, *22*(11), 2843–2859. https://doi.org/10.1175/Jhm-D-21-0058.1

Dunkerley, D. (2008). Identifying individual rain events from pluviograph records: A review with analysis of data from an Australian dryland site. *Hydrological Processes*, *22*(26), 5024–5036. https://doi.org/10.1002/hyp.7122

Fassoni-Andrade, A. C., Fleischmann, A. S., Papa, F., Paiva, R. C. D. D., Wongchuig, S., Melack, J. M., et al. (2021). Amazon hydrology from space: Scientific advances and future challenges. *Reviews of Geophysics*, *59*(4). https://doi.org/10.1029/2020RG000728

Foufoula-Georgiou, E., Guilloteau, C., Nguyen, P., Aghakouchak, A., Hsu, K. L., Busalacchi, A., et al. (2020). Advancing precipitation estimation, prediction, and impact studies. *Bulletin of the American Meteorological Society*, *101*(9), E1584–E1592. https://doi.org/10.1175/bams-d-20-0014.1

Guilloteau, C., Foufoula-Georgiou, E., Kirstetter, P., Tan, J., & Huffman, G. J. (2021). How well do multi-satellite products capture the space-time dynamics of precipitation? Part I: Five products assessed via a wavenumber-frequency decomposition. *Journal of Hydrometeorology*, *22*(11), 2805–2823. https://doi.org/10.1175/jhm-d-21-0075.1

Guilloteau, C., Foufoula-Georgiou, E., Kirstetter, P., Tan, J. C., & Huffman, G. J. (2022). How well do multisatellite products capture the space-time dynamics of precipitation? Part II: Building an error model through spectral system identification. *Journal of Hydrometeorology*, *23*(9), 1383–1399. https://doi.org/10.1175/Jhm-D-22-0041.1

Guilloteau, C., Foufoula-Georgiou, E., Kummerow, C. D., & Petkovic, V. (2018). Resolving surface rain from GMI high-frequency channels: Limits imposed by the three-dimensional structure of precipitation. *Journal of Atmospheric and Oceanic Technology*, *35*(9), 1835–1847. https://doi.org/10.1175/JTECH-D-18-0011.1

Hanel, M., & Maca, P. (2014). Spatial variability and interdependence of rain event characteristics in the Czech Republic. *Hydrological Processes*, *28*(6), 2929–2944. https://doi.org/10.1002/hyp.9845

Hou, S., Tian, F., Yang, L., Hu, H., & Hou, A. (2018). How does the evaluation of the GPM IMERG rainfall product depend on gauge density and rainfall intensity? *Journal of Hydrometeorology*, *19*(2), 339–349. https://doi.org/10.1175/jhm-d-17-0161.1

Huffman, G. J., Bolvin, D. T., Braithwaite, D., Hsu, k., Joyce, R., Kidd, C., et al. (2019a). NASA global precipitation measurement (GPM) integrated multi-satellitE retrievals for GPM (IMERG) algorithm theoretical basis document (ATBD) version 6.0 (p. 14). Retrieved from https://pps.gsfc.nasa.gov/Documents/IMERG_ATBD_V15.11.pdf.]

Huffman, G. J., Stocker, E. F., Bolvin, D. T., Nelkin, E. J., & Tan, J. (2019b). GPM IMERG final precipitation L3 half hourly 0.1 degree × 0.1 degree V06 [Dataset]. NASA Goddard Earth Sciences Data and Information Services Center (GES DISC). https://doi.org/10.5067/GPM/IMERG/3B-HH/06

Kirstetter, P. E., Hong, Y., Gourley, J. J., Cao, Q., Schwaller, M., & Petersen, W. A. (2014). Research framework to bridge from the global precipitation measurement mission Core satellite to the constellation sensors using ground-radar-based National mosaic QPE. In *Remote sensing of the terrestrial water cycle* (pp. 61–79). https://doi.org/10.1175/Jhm-D-11-0139.1

Kirstetter, P. E., Hong, Y., Gourley, J. J., Chen, S., Flamig, Z., Zhang, J., et al. (2012). Toward a framework for systematic error modeling of spaceborne precipitation radar with NOAA/NSSL ground radar based National mosaic QPE. *Journal of Hydrometeorology*, *13*(4), 1285–1300. https://doi.org/10.1175/Jhm-D-11-0139.1

Kirstetter, P. E., Karbalaee, N., Hsu, K. L., & Hong, Y. (2018). Probabilistic precipitation rate estimates with space-based infrared sensors. *Quarterly Journal of the Royal Meteorological Society*, *144*(S1), 191–205. https://doi.org/10.1002/qj.3243

Kirstetter, P. E., Petersen, W. A., & Gourley, J. J. (2018). GPM ground validation multi-radar/multi-sensor (MRMS) precipitation reanalysis for satellite validation product [Dataset]. NASA Global Hydrology Resource Center (GHRC). https://doi.org/10.5067/GPMGV/MRMS/DATA101

Li, R. Z., Guilloteau, C., Kirstetter, P. E., & Foufoula-Georgiou, E. (2023). How well does the IMERG satellite precipitation product capture the timing of precipitation events? *Journal of Hydrology*, *620*, 129563. https://doi.org/10.1016/j.jhydrol.2023.129563

Li, R. Z., Wang, K. C., & Qi, D. (2018). Validating the integrated multisatellite retrievals for global precipitation measurement in terms of diurnal variability with hourly gauge observations collected at 50,000 stations in China. *Journal of Geophysical Research: Atmospheres*, *123*(18), 10423–10442. https://doi.org/10.1029/2018jd028991

Li, R. Z., Wang, K. C., & Qi, D. (2021). Event-based evaluation of the GPM multisatellite merged precipitation product from 2014 to 2018 over China: Methods and results. *Journal of Geophysical Research: Atmospheres*, *126*(1), e2020JD033692. https://doi.org/10.1029/2020JD033692

Nie, Y. B., & Sun, J. Q. (2022). Moisture sources and transport for extreme precipitation over henan in July 2021. *Geophysical Research Letters*, *49*(4). https://doi.org/10.1029/2021GL097446

Orland, E., Kirschbaum, D., & Stanley, T. (2022). A scalable framework for post fire debris flow hazard assessment using satellite precipitation data. *Geophysical Research Letters*, *49*(18), e2022GL099850. https://doi.org/10.1029/2022GL099850

Passive Microwave Algorithm Team Facility. (2017). Global precipitation measurement (GPM) mission algorithm theoretical basis document GPROF2017 version 1 and version 2 (used in GPM V5 processing) (p. 43). Retrieved from https://pps.gsfc.nasa.gov/Documents/ATBD_GPM_V45B_April15_2018.pdf.]

Petersen, W. A., Kirstetter, P.-E., Wang, J., Wolff, D. B., & Tokay, A. (2020). The GPM ground validation Program. In *Satellite precipitation measurement* (pp. 471–502). Springer International Publishing. https://doi.org/10.1007/978-3-030-35798-6_2

Petkovic, V., & Kummerow, C. D. (2017). Understanding the sources of satellite passive microwave rainfall retrieval systematic errors over land. *Journal of Applied Meteorology and Climatology*, *56*(3), 597–614. https://doi.org/10.1175/Jamc-D-16-0174.1

Pradhan, R. K., Markonis, Y., Vargas Godoy, M. R., Villalba-Pradas, A., Andreadis, K. M., Nikolopoulos, E. I., et al. (2022). Review of GPM IMERG performance: A global perspective. *Remote Sensing of Environment*, *268*, 112754. https://doi.org/10.1016/j.rse.2021.112754

Rajagopal, M., Zipser, E., Huffman, G., Russell, J., & Tan, J. (2021). Comparisons of IMERG version 06 precipitation at and between passive microwave overpasses in the tropics. *Journal of Hydrometeorology*, *22*(8), 2117–2130. https://doi.org/10.1016/0022-1694(82)90136-6

Restrepo-Posada, P. J., & Eagleson, P. S. (1982). Identification of independent rainstorms. *Journal of Hydrology*, *55*(1–4), 303–319. https://doi.org/10.1016/0022-1694(82)90136-6

Smith, T. M., Arkin, P. A., Bates, J. J., & Huffman, G. J. (2006). Estimating bias of satellite-based precipitation estimates. *Journal of Hydrometeorology*, *7*(5), 841–856. https://doi.org/10.1175/Jhm524.1

Su, J. B., Lu, H., Zhu, Y. H., Wang, X. Y., & Wei, G. H. (2018). Component analysis of errors in four GPM-based precipitation estimations over mainland China. *Remote Sensing*, *10*(9), 1420. https://doi.org/10.3390/rs10091420

Tadesse, A., & Anagnostou, E. N. (2009). The effect of storm life cycle on satellite rainfall estimation error. *Journal of Atmospheric and Oceanic Technology*, *26*(4), 769–777. https://doi.org/10.1175/2008jtecha1129.1

Tan, J., Huffman, G. J., Bolvin, D. T., Nelkin, E. J., & Rajagopal, M. (2021). SHARPEN: A scheme to restore the distribution of averaged precipitation fields. *Journal of Hydrometeorology*, *22*(8), 2105–2116. https://doi.org/10.1175/Jhm-D-20-0225.1

Tan, J. C., Cho, N. Y., Oreopoulos, L., & Kirstetter, P. (2022). Evaluation of GPROF V05 precipitation retrievals under different cloud regimes. *Journal of Hydrometeorology*, *23*(3), 389–402. https://doi.org/10.1175/Jhm-D-21-0154.1

Tang, G. Q., Clark, M. P., Papalexiou, S. M., Ma, Z. Q., & Hong, Y. (2020). Have satellite precipitation products improved over last two decades? A comprehensive comparison of GPM IMERG with nine satellite and reanalysis datasets. *Remote Sensing of Environment*, *240*, 111697. https://doi.org/10.1016/j.rse.2020.111697

Tian, F. Q., Hou, S. Y., Yang, L., Hu, H. C., & Hou, A. Z. (2018). How does the evaluation of the GPM IMERG rainfall product depend on gauge density and rainfall intensity? *Journal of Hydrometeorology*, *19*(2), 339–349. https://doi.org/10.1175/Jhm-D-17-0161.1

Tian, Y. D., Peters-Lidard, C. D., Eylander, J. B., Joyce, R. J., Huffman, G. J., Adler, R. F., et al. (2009). Component analysis of errors in satellite-based precipitation estimates. *Journal of Geophysical Research*, *114*(D24), D24101. https://doi.org/10.1029/2009jd011949

Wright, D. B., Kirschbaum, D. B., & Yatheendradas, S. (2017). Satellite precipitation characterization, error modeling, and error correction using censored shifted gamma distributions. *Journal of Hydrometeorology*, *18*(10), 2801–2815. https://doi.org/10.1175/Jhm-D-17-0060.1

You, Y. L., Wang, N. Y., Ferraro, R., & Rudlosky, S. (2017). Quantifying the snowfall detection performance of the GPM microwave imager channels over land. *Journal of Hydrometeorology*, *18*(3), 729–751. https://doi.org/10.1175/Jhm-D-16-0190.1

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., et al. (2016). Multi-radar multi-sensor (MRMS) quantitative precipitation estimation Initial Operating Capabilities. *Bulletin of the American Meteorological Society*, *97*(4), 621–637. https://doi.org/10.1175/Bams-D-14-00174.1

Zhang, Y., Li, R. Z., & Wang, K. C. (2023). Climatology and changes in internal intensity distributions of global precipitation systems over 2001–2020 based on IMERG. *Journal of Hydrology*, *620*, 129386. https://doi.org/10.1016/j.jhydrol.2023.129386

Zhang, Y. H., Ye, A. Z., Nguyen, P., Analui, B., Sorooshian, S., & Hsu, K. L. (2021). New insights into error decomposition for precipitation products. *Geophysical Research Letters*, *48*(17). https://doi.org/10.1029/2021GL094092