Journal of Hydrology 651 (2025) 132610

Contents lists available at ScienceDirect

Journal of Hydrology

journal homepage: www.elsevier.com/locate/jhydrol

Research papers

Understanding the error patterns of multi-satellite precipitation products during the lifecycle of precipitation events for diagnostics and algorithm improvement

Runze Li [a,*], Clement Guilloteau [a], Pierre-Emmanuel Kirstetter [b,c], Efi Foufoula-Georgiou [a,d]

[a] Department of Civil and Environmental Engineering, University of California, Irvine, CA, USA
[b] Hydrometeorology and Remote Sensing Laboratory, University of Oklahoma, Norman, OK, USA
[c] NOAA/Severe Storms Laboratory, Norman, OK, USA
[d] Department of Earth System Science, University of California, Irvine, CA, USA

ARTICLE INFO

This manuscript was handled by Emmanouil Anagnostou, Editor-in-Chief, with the assistance of Yang Hong, Associate Editor

Keywords:
Precipitation
Satellite precipitation product
Precipitation event
Event stage
Storm environment
IMERG

ABSTRACT

Satellite precipitation products are not free of errors. These errors may show specific temporal patterns related to the life cycle of precipitation events. Understanding such patterns is key to uncertainty quantification, product improvements, and hydrologic applications. Here we investigate satellite error patterns during the life cycle of precipitation events over the contiguous United States (CONUS), using the Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG) as the satellite product and the Ground Validation Multi-Radar/Multi-Sensor (GV-MRMS) as reference to define "events" in a Eulerian perspective, both at the 30 min and $0.1^\circ \times 0.1^\circ$ native resolution of IMERG. We reveal significant variation in IMERG's biases (both detectability and intensity bias) before/during/after the events, with a marked temporal asymmetry with respect to the mid-point of the event duration. Overall, the miss/false proportions of precipitation occurrence peak near the event temporal boundaries, with miss proportion higher near event ends, and false detection proportion higher near event beginnings. Precipitation intensity tends to be overestimated near the boundaries as well, while it is underestimated during the early- to mid-stages of the event. Diagnostic analysis controlling for data source inhomogeneity in IMERG and intensity variations throughout events traces back this stage-dependent performance to the Passive Microwave (PMW) retrieval algorithm, possibly due to the variation in cloud physical properties during event life cycles. Further investigations over different seasons/regions/times of day reveal distinct event-stage-dependent error patterns, which are likely linked to the different convective precipitation proportions. Consequently, a conditional analysis of error patterns on storm-type-related environmental variables, i.e., Convective Available Potential Energy (CAPE) and dewpoint is performed revealing significant relationships that underscore their prognostic value for characterizing the stage-dependent error curves. This study underscores the robust dependency of satellite errors on event stages and conversely, indicates the possible accuracy improvement of satellite precipitation products by integrating event stage information, as well as comprehensively leveraging environmental variables in future algorithms.

1. Introduction

The deployment of specialized precipitation monitoring satellites, exemplified by the Tropical Rainfall Measuring Mission (TRMM) and Global Precipitation Measurement (GPM) missions, coupled with advances in retrieval and multi-sensor fusion techniques, has facilitated the development of global high-resolution precipitation products (Huffman et al., 2023b; Kubota et al., 2020; Sadeghi et al., 2021; Xie et al., 2017). Such high-resolution products are especially valuable given the inherent high spatiotemporal variability of precipitation, e.g., aiding in understanding the dynamics of short-duration intense storms and their impacts on critical regions (steep terrains, urban areas, etc.). Nevertheless, satellite products exhibit non-negligible biases and uncertainties, underscoring the need for a comprehensive understanding of their error patterns across all scales. However, existing error characterizations mostly focus on relatively coarse scales such as seasonal and

* Corresponding author.
E-mail address: runzl11@uci.edu (R. Li).

https://doi.org/10.1016/j.jhydrol.2024.132610
Received 20 July 2024; Received in revised form 11 November 2024; Accepted 16 December 2024
Available online 27 December 2024

regional, or lump all the sub-daily/hourly precipitation data together to calculate time-independent statistical performance indicators (Loew et al., 2017; Maggioni et al., 2016; Pradhan et al., 2022), such as mean detection rate/rain rate bias, implicitly assuming stationarity and homoscedasticity at fine scales and ignoring possible dependence on the stage of precipitation events (Kirstetter et al., 2020; Guilloteau and Foufoula-Georgiou, 2024). Thorough investigation into the fine-scale error patterns of satellite products thus remains insufficient. Such a gap can be attributed not only to the still limited availability of high-resolution ground reference data across the globe (Kidd et al., 2017), but also to the challenge of discerning coherent patterns at these scales, which could appear as random without in-depth analysis (Derin and Kirstetter, 2022; Derin et al., 2022; Derin et al., 2021).

In this study, we aim to further elucidate the errors of satellite precipitation products at fine time scales, i.e., 30 min, important for hydrologic, atmospheric, and other applications. We note that the occurrence of precipitation is discrete in nature, essentially comprising individual "events" with certain durations and evolution patterns in time at a given location (Li et al., 2021). Precipitation events are thus considered as the structuring element of our temporal pattern analysis. The microphysics and thermodynamic properties of precipitating clouds are known to undergo significant changes during the course of a precipitation event (Bouniol et al., 2016; Imaoka and Nakamura, 2012). The sensor signal-to-rainfall rate relationship is sensitive to these physical variations (Petkovic and Kummerow, 2017), inevitably leading to varying retrieval accuracy throughout the events' life cycle (O and Kirstetter, 2018; Yamamoto et al., 2008; Guilloteau and Foufoula-Georgiou, 2024). Analyzing this variability is expected to help understand the causes of the seemingly random fine-scale error patterns in satellite precipitation time series.

Previously, event-based evaluations of satellite precipitation products have mostly focused on evaluating event-integrated characteristics (e.g., event count, duration, total depth, timing) (Freitas et al., 2020; Li et al., 2023a; Sutton et al., 2024), without detailed scrutiny of the error evolution during the events. Alternatively, Li et al. (2021) preliminary demonstrated a dependency of satellite precipitation product performance on the event stages during the warm season in China by using gauge data at the national scale. The present study builds on this analysis to further our understanding and comprehensively investigate event-based error evolution patterns under various conditions, provide a diagnostic evaluation of their underlying causes, and examine their connection with the storm environment in the contiguous United States (CONUS). Specifically, we want to further understand: (1) What is the general pattern of satellite errors during precipitation events over CONUS and how significant and robust is the stage-dependency of the errors? (2) Is the stage-dependency of the errors intrinsic to the Passive Microwave (PMW) retrieval or a reflection of the evolving rain rate/data sources of the multi-satellite product during the events? (3) Do the stage-dependent error patterns vary across diverse conditions (i.e., regions, seasons, and times of day), and can storm-related environmental variables serve as unified prognostic indicators for these variations? The Integrated Multi-satellite Retrievals for GPM (IMERG) version 7 is evaluated against the Ground Validation-Multi-Radar/Multi-Sensor (GV-MRMS) product over CONUS (Kirstetter et al., 2014; Kirstetter et al., 2012). The physical variables characterizing the storm environment are obtained from the European Centre for Medium Range Weather Forecasts (ECMWF) Re-Analysis version 5 (ERA5).

2. Data and methods

2.1. Data

IMERG is the principal level-3 gridded global multi-satellite merged precipitation product from NASA's GPM project (Huffman et al., 2023b). It integrates all available PMW data from the GPM constellation, with interpolated PMW and Infrared (IR) data filling in any gaps to form a globally seamless product. All the data sources used in IMERG are intercalibrated by the high-accuracy instruments onboard the GPM Core Observatory (i.e., GPM Microwave Radiometer (GMI) and Dual-frequency Precipitation Radar (DPR)) to ensure maximum consistency across space and time. Additionally, IMERG is adjusted by the monthly GPCC gauge analysis, which further enhances its reliability. These features have positioned IMERG as one of the most popular high-resolution precipitation products for various atmospheric and hydrologic applications (Bai and Schumacher, 2022; Kukulies et al., 2021; Zhang et al., 2023), drawing at the same time considerable attention on its evaluation and improvement (Derin et al., 2022; Derin et al., 2021; Gebregiorgis et al., 2018; Guilloteau et al., 2021; Li et al., 2022). Here we use the IMERG V07A Final Run product which has a resolution of $0.1^\circ \times 0.1^\circ$ and 0.5 h (Huffman et al., 2023a).

GV-MRMS is a high-accuracy ground radar-gauge blended quantitative precipitation estimation (QPE), which is tailored specifically for GPM product validation and used as the ground reference data here (Kirstetter et al., 2020). It is built upon the NOAA MRMS that integrates multi-source ground-based radars, rain gauges, and forecast model data (Zhang et al., 2016), and then undergoes substantial post-processing for adaptation to satellite evaluation (Kirstetter et al., 2014; Kirstetter et al., 2012). GV-MRMS is widely acknowledged as a high-fidelity ground validation dataset over CONUS and has been frequently applied in verification tasks (e.g., Derin et al., 2022; Milani et al., 2021; Upadhyaya et al., 2020; Guilloteau et al., 2021). The spatiotemporal resolution of GV-MRMS is $0.01^\circ \times 0.01^\circ$ and 0.5 h, therefore we aggregate it spatially to IMERG's $0.1^\circ \times 0.1^\circ$ resolution. The production of GV-MRMS also generates a Radar Quality Index (RQI), which represents the quality of radar observations on a 0–100 scale (Petersen et al., 2020). To conduct quality control, only data scoring no less than 60 are retained for the analysis (Li et al., 2023a). In the following text, "GV-MRMS" is abbreviated for simplicity to "MRMS".

The atmospheric environmental variables are sourced from ERA5 (Hersbach et al., 2020). ERA5 is a cutting-edge global atmospheric reanalysis dataset with hourly 0.25° resolution and 137 vertical layers, spanning 1940 to the present. ERA5 assimilates a vast range of observational data into the ECMWF Integrated Forecasting System (IFS) model outputs, yielding a comprehensive and consistent record of global atmospheric, land, and oceanic variables, and has been widely used in weather and climate studies (Bi et al., 2023; Ham et al., 2023; Yuan et al., 2023). Here ERA5 is used to acquire environmental conditions related to precipitation events, and to examine their relationship with the error evolution during the events. It is worth noting that while regional ground-based or radiosonde observations, as well as regional reanalysis data, might offer more accurate environmental variables, our ultimate goal is to provide insights for improving global satellite precipitation retrievals, rather than establishing precise physical relationships. Therefore, we chose to use the globally available ERA5 data to explore its potential in indicating satellites' event-based errors. To match IMERG's resolution, raw ERA5 data are interpolated to a finer scale of 0.1°, half an hour prior to the onset of each precipitation event analyzed. The study period is 2018–2020.

2.2. Methods

A precipitation event at every pixel of 0.1^0x$0.1^0$is defined using MRMS data as "an uninterrupted series of half-hourly time steps with non-zero precipitation intensity ($\geq$ 0.1 mm/h)" (Li et al., 2023a). Based on this definition, we first extract all the precipitation events from MRMS over CONUS and throughout the study period. The spatial distribution of basic characteristics of the extracted precipitation events (i.e., number of events, event duration, event depth, mean event precipitation rate) is shown in Fig. S1. Events with identical durations are then grouped together, with error statistics computed at every half-hour timestep for each group.

Our focus is on calculating error evolution patterns throughout the

event life cycles, which are conducted from two aspects: detectability of precipitation occurrence (classification error) and rain rate bias (quantitative error). For the detectability, two metrics, *Miss Proportion (%)* and *False Proportion (%)*, are defined. According to our event definition, misses in IMERG occur only inside the events, while false alarms happen only outside the events. Therefore, we also incorporate a short period before/after the event into our analysis to examine false alarms potentially associated with the target event. We set this examination period to 5 h, as it would encompass most event-related misdetection causes (e.g., inappropriate precipitation detection thresholds, temporal lags between hydrometeors and surface precipitation, and spatial displacement due to satellite viewing geometry) (Li et al., 2023a; Guilloteau et al., 2018). Specifically, the *Miss Proportion* and *False Proportion* at the $\boldsymbol{t}$-th half-hour (referred to as "event stages") during/before/after the $\boldsymbol{d}$-hour MRMS events are calculated as:

$$\boldsymbol{MissProportion}(\boldsymbol{d},\boldsymbol{t}) = \frac{\boldsymbol{N}_{\boldsymbol{miss}}(\boldsymbol{d},\boldsymbol{t})}{\boldsymbol{N}_{\boldsymbol{rainy}}(\boldsymbol{d},\boldsymbol{t})} \times 100\%, \tag{1}$$

$$\boldsymbol{FalseProportion}(\boldsymbol{d},\boldsymbol{t}) = \frac{\boldsymbol{N}_{\boldsymbol{false}}(\boldsymbol{d},\boldsymbol{t})}{\boldsymbol{N}_{\boldsymbol{non-rainy}}(\boldsymbol{d},\boldsymbol{t})} \times 100\%, \tag{2}$$

where $\boldsymbol{N}_{\boldsymbol{miss}}(\boldsymbol{d},\boldsymbol{t})$ and $\boldsymbol{N}_{\boldsymbol{false}}(\boldsymbol{d},\boldsymbol{t})$ are the numbers of miss and false cases under specific $\boldsymbol{d}$ and $\boldsymbol{t}$, respectively, while $\boldsymbol{N}_{\boldsymbol{rainy}}(\boldsymbol{d},\boldsymbol{t})$ and $\boldsymbol{N}_{\boldsymbol{non-rainy}}(\boldsymbol{d},\boldsymbol{t})$ are the numbers of rainy ($\geq$0.1 mm/h) and non-rainy ($<$0.1 mm/h) cases in terms of MRMS also under specific $\boldsymbol{d}$ and $\boldsymbol{t}$. Therefore, these two metrics indicate the proportion of instances during the events that should be rainy but are mistakenly identified by IMERG as non-rainy, and the proportion of instances immediately before or after the events that should be non-rainy but are erroneously identified as rainy, which can be understood as the "False Negative Rate" and "False Positive Rate" in the general classification model, respectively. While statistics computed in the 5-hour window before/after the event of interest can certainly be affected by other nearby-in-time events on a case-by-case basis, on average, the pre- and post-event statistics show robust and interpretable temporal patterns. Excluding cases with nearby-in-time events would not only dramatically reduce our sample size, but it would also bias our analysis with the overrepresentation of isolated events.

Besides, the *Intensity Bias* is defined as the mean intensity difference (IMERG − MRMS) during their concurrent precipitation detection (i.e., "hit"). Specifically, the *Intensity Bias* at the $\boldsymbol{t}$-th half-hour during the $\boldsymbol{d}$-hour events is calculated as:

$$\boldsymbol{IntensityBias}(\boldsymbol{d},\boldsymbol{t}) = \boldsymbol{I}_{\boldsymbol{IMERG}}(\boldsymbol{d},\boldsymbol{t}) - \boldsymbol{I}_{\boldsymbol{MRMS}}(\boldsymbol{d},\boldsymbol{t}), \tag{3}$$

where $\boldsymbol{I}_{\boldsymbol{IMERG}}(\boldsymbol{d},\boldsymbol{t})$ and $\boldsymbol{I}_{\boldsymbol{MRMS}}(\boldsymbol{d},\boldsymbol{t})$ are the mean intensity of IMERG and MRMS when they "hit" under specific $\boldsymbol{d}$ and $\boldsymbol{t}$. Based on the above equations, the continuous error evolution curves before/during/after the events of different durations can be computed. Additionally, we divide the event duration in three equal intervals rounded to the nearest half hour and refer to them for brevity as the "beginning", "middle", and "end" of events.

To determine whether the event-stage dependent error pattern is an intrinsic property of the precipitation retrieval algorithm, rather than reflection of other factors, we conduct additional control experiments for diagnostic analysis. Two factors are highlighted here: the temporal inhomogeneity of IMERG data sources and the intensity variations during the events, both known to significantly affect error magnitudes (Gao and Liu, 2013; Li et al., 2018). Concerning data source inhomogeneity, we recalculate the error in equations (1)-(3) using only GMI-sourced IMERG data and compare these to the full-IMERG-derived results, assessing if the errors fundamentally stem from PMW retrievals rather than the multi-source merging in IMERG. The specific approach is to add a sensor source constraint when extracting each of the right-hand side terms of Equations (1)-(3), (e.g., $\boldsymbol{MissProportion}(\boldsymbol{d},\ \boldsymbol{t},\ \boldsymbol{s}) = \frac{\boldsymbol{N}_{\boldsymbol{miss}}(\boldsymbol{d},\boldsymbol{t},\boldsymbol{s})}{\boldsymbol{N}_{\boldsymbol{rainy}}(\boldsymbol{d},\boldsymbol{t},\boldsymbol{s})} \times 100\%$, where **s** denotes a specific data source, similar for the other two formulae). This enables us to construct complete sensor-specific error evolution curves, despite individual sensors' inability to capture the entire life cycle of precipitation events. Regarding intensity variations during the events, we compare the IMERG and MRMS products conditional on event stages to examine whether intensity alone explains the error curves, or whether errors stem from the changing radiometric-signal-to-rain-rate relationships.

Analysis of the annual average error patterns at the national scale demonstrates a robust dependence on the event stage, prompting further examination on how these patterns may vary under different conditions, i.e., seasons, regions, and times of day. Acknowledging that the potential varying error patterns across these different conditions could somewhat stem from the differences in storm types, which, in turn, could be indicated by the storm environment to some extent (Heuscher et al., 2022; Zhang and Villarini, 2019), we further explore the potential relationship between the error curves and the storm environment. The event-associated environment is represented by a number of environmental variables 30-min before the event onset to minimize the influence of possible feedback from the precipitation processes (Petkovic and Kummerow, 2017; Song et al., 2019). We collect various environmental variables (e.g., wind shear, moisture divergence, vertical velocity) from ERA5 and test the relationships between their values and the error evolution curve shapes during the events. Considering the distinctiveness of the derived patterns and relatively more direct dependence with convective precipitation, Convective Available Potential Energy (CAPE), denoting atmospheric instability, and Dewpoint, reflecting moisture availability, are selected as representatives. For a quantitative comparison, two parameters that describe the shapes and magnitudes of the error curves are introduced: the *Event-wide Average* of miss proportion/false proportion/rain rate bias across all duration events, and the *Asymmetry Index* in the error curves, defined as the difference between the average miss proportion/false proportion/rain rate bias of the latter halves and the former halves of all events.

3. Results

3.1. Nationwide annual quantification

We start by presenting national, annual mean statistics (Figs. 1 and 2). Fig. 1 first displays the half-hourly detectability of precipitation occurrence as a function of duration and time step (half-hourly) of MRMS-derived events. It is evident that the miss proportion significantly depends on both the stage and duration of events (Fig. 1a), consistent with findings from Li et al. (2021). Specifically, longer events tend to have a lower miss proportion than shorter ones, while the beginning and end of events show a higher miss proportion than their middle parts. In addition, most of the curves display pronounced asymmetry, with precipitation being more likely missed near the ends than the beginnings of events. Such asymmetry is more evident in prolonged events, as evidenced by a disparity of up to 14 % between the initial (45 %) and final (59 %) timesteps for the 10 h events. Additionally, upon observing the original intensity patterns of MRMS events (Fig. S2), we find that the variation in miss proportion is somewhat linked to the intensity evolution (i.e., the higher intensity is, the lower miss proportion is). However, the persistent significant variations in missed precipitation intensity further suggest that intensity evolution is at least not the only factor dominating the error variations during the events (otherwise, the missed precipitation would consistently remain of low intensity over time) (Fig. S3a), but there must be more fundamental causes at play, which will be further discussed later.

Similarly, the false proportions calculated within the 5-hour window before/after the events also reveal a clear correlation with event duration and proximity to the event starting/ending times (Fig. 1b). Longer events typically have higher false proportions than shorter ones, and the

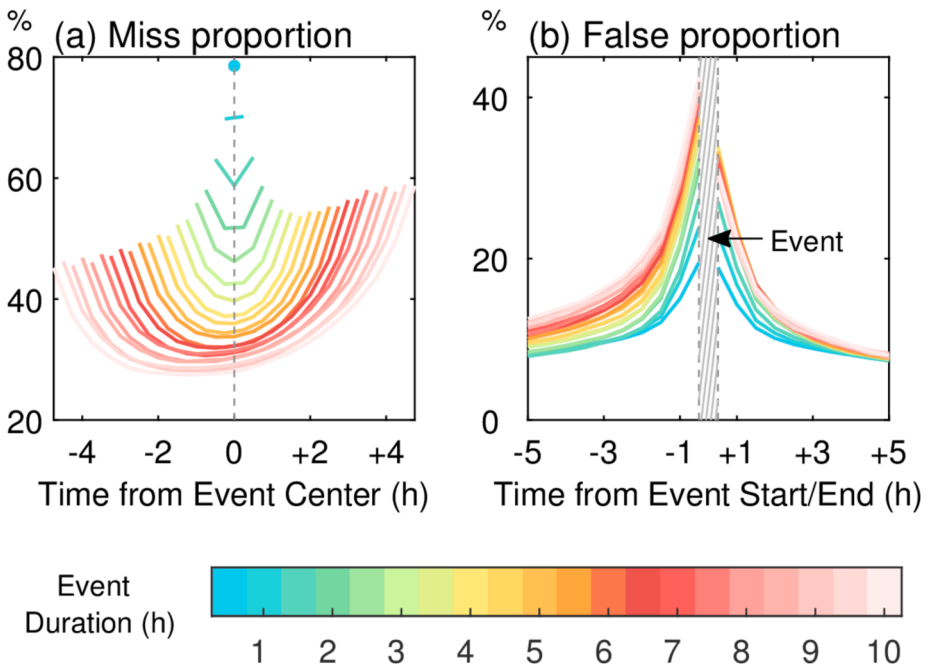

Fig. 1. (a) *Miss Proportion* of rainy hours at each half-hour during the MRMS-derived events of different durations (color-coded curves), aligned in the event centers (defined as the midpoints of the event durations). (b) *False Proportion* of rainy hours at each half-hour reported for 5-hour before the start and after the end of an event (events shown with the shaded bar). These biases are reported for events of duration up to 10 h, resulting in miss proportion curves with time axes spanning −4.75 to + 4.75 h (the centers of the half-hour first and last time increments of the event duration). The values are calculated from all the data across CONUS from 2018 to 2020.

proportions are elevated before the event starts than after it ends. These patterns are both contrary to those of miss proportion (Fig. 1a), which aligns with the expected inherent trade-off between false positives (false proportion) and false negatives (miss proportion) in classification models. For the asymmetry, the greatest disparity could reach up to 12 % between the half-hour before the starts (43 %) and after the ends (31 %) of the 10-hour events. False proportions peak around event boundaries and decrease with increasing temporal distance from the events. This pattern is intuitively expected, as the moments closer to the events are inevitably more susceptible to being misclassified as part of the events themselves. However, we observe that the false proportion still remains at moderately high level (~10 %) at times further away from the target event (~5 h). This can be mainly attributed to the impacts of the potential nearby events in both space and time. It should be noted that our statistical analysis focuses only on periods of time where an MRMS event has occurred, and thus the probability of another event occurring nearby in space and time would naturally be higher than in a randomly selected environment, leading to the higher-than-expected false proportion near the events.

The intensity bias variation during MRMS-derived events is subsequently analyzed when IMERG correctly identifies precipitation occurrences (i.e., 'hit'), also under nationwide and annual average conditions (Fig. 2). The mean rain rates of both MRMS and IMERG are also presented for reference (Fig. 2a and b). A notable disparity is observed in the temporal profiles of mean intensity between MRMS and IMERG, with IMERG mostly failing to capture the earlier, more intense peak of the MRMS events, yet exhibiting higher precipitation rates near the start/end of events. These differences in event development profiles naturally give rise to robust stage-dependent bias patterns, with overestimation near the start/end and underestimation in the early to middle stages (Fig. 2c). In summary, both the detectability and rain rate bias results, as illustrated in Figs. 1 and 2, underscore the event stage- and duration-dependent nature of satellite errors. This, combined with similar findings in China (Li et al., 2021), emphasizes the universality of such error patterns across different major geographic regions, laying the groundwork for the following analysis.

3.2. Diagnosis of the stage-dependent bias

As described in the Methods Section, it cannot yet be excluded that the error variations observed during the events are merely a reflection of other co-evolving factors, rather than an intrinsic property of the inversion relationship in the precipitation retrieval algorithm, which highlights the need for additional diagnostic analyses. Given the well-known temporal inhomogeneity of IMERG data sources, we first conducted diagnostics controlling for the data sources (Fig. 3), primarily focusing on the comparison between results from solely GMI and the full IMERG. While the main focus is on the comparison between results from solely GMI and the full IMERG, average results from all PMW observations and the interpolated data (called "morph" in IMERG) are also presented for reference. Overall, results from different data sources generally exhibit similar event stage-dependent and asymmetric features, suggesting that the multi-source integration does not make an essential contribution to satellites' varying performance during the events. When focusing on GMI-sourced results, they exhibit higher accuracy across all metrics (Fig. 3a1, a2, and b3), which is expected given GMI's superior accuracy among all IMERG's data sources. However, their temporal patterns mirror those of IMERG, with even more pronounced asymmetries (which is likely due to the smoothing effect of interpolation in IMERG, which tends to suppress any asymmetries (Li et al., 2023a)). Such a pattern thus indicates an inherent stage-dependent error in PMW retrievals that propagates to IMERG. Additionally, the results, in turn, suggest that using temporally continuous IMERG data to study event-related error characteristics can provide generally reliable insights for PMW retrievals (which is the core of precipitation inversion).

Another factor evolving significantly during the duration of an event is precipitation intensity, which is clearly established to have a non-negligible impact on the error (Kirstetter et al., 2014; Kirstetter et al., 2012). To demonstrate that error variations during events are not solely

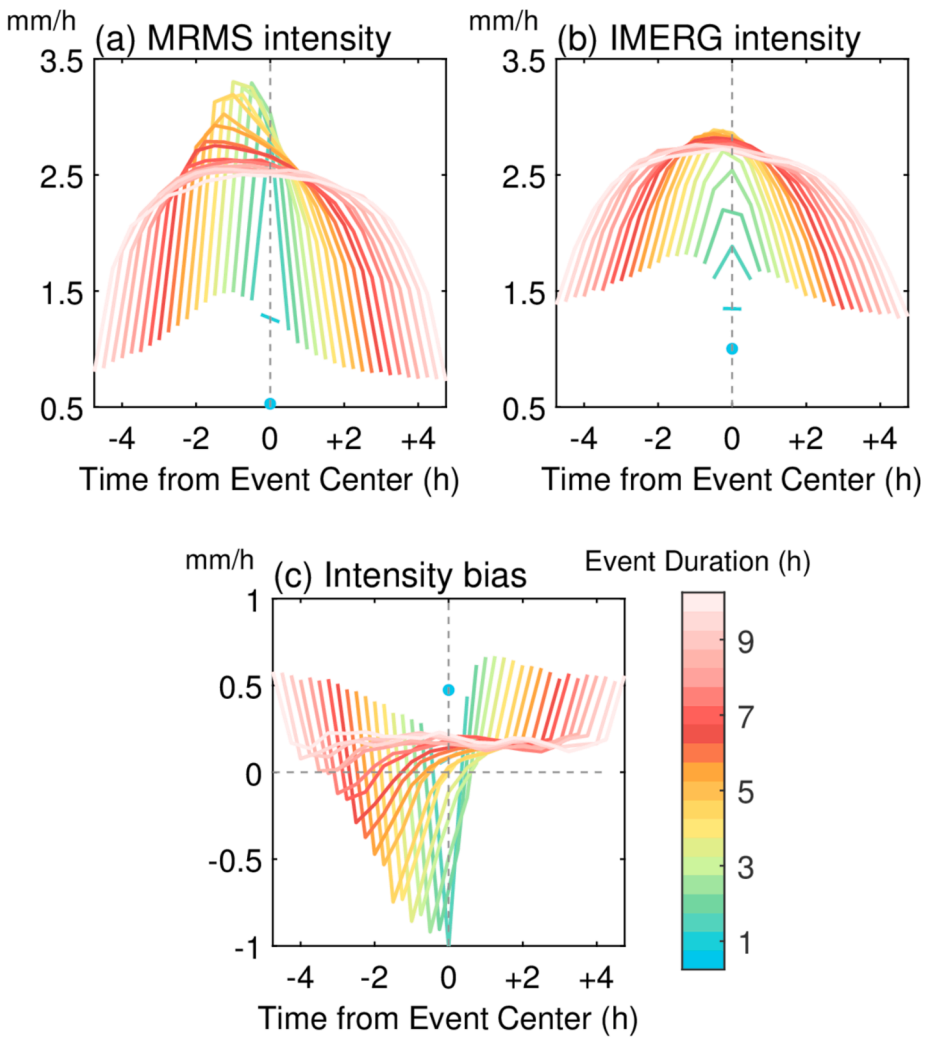

Fig. 2. Mean (a) MRMS intensity, (b) IMERG intensity, and (c) *Intensity Bias* in rainy hours (IMERG – MRMS, when both IMERG and MRMS detect precipitation occurrence, i.e., "hit") at each half-hour during the MRMS-derived events of different durations (color-coded curves), aligned in the event centers. The MRMS/IMERG intensities and their bias are reported for events of duration up to 10 h, resulting in their curves with time axes spanning −4.75 to + 4.75 h. The values are calculated from all the data across CONUS from 2018 to 2020.

driven by intensity changes, Fig. 4 illustrates the intensity correspondence between IMERG and MRMS at different event stages (specifically, the first/middle/last half-hours of the events here). The graph reveals that controlling intensity somewhat mitigates the bias asymmetry, as indicated by the similarity between the curves for the initial and final half-hours. However, there remains a significant disparity between the curves for the middle half-hour and those for the beginning and end, suggesting notable differences in the IMERG-to-MRMS intensity correspondence at various event stages, which fundamentally reflects shifts in the signal-to-rain-rate relationship during events. This pattern persists whether employing the full IMERG or solely GMI (Fig. 4a and b). Thus, the analysis further demonstrates the presence of inherent event-stage-dependency that modulates the long-known intensity-dependence, which is also supported by a similar analysis for the miss proportion (Fig. S4).

After isolating important influencing factors such as inhomogeneity in data sources and intensity, the results suggest that more fundamental factors are at play that affect the inversion relationship. This, linking with our Eulerian-based event approach, could involve the complex temporal evolution of cloud properties during the event life cycle (Bouniol et al., 2016; Samanta et al., 2021), as well as the intricate spatial structure of the storm system mapped onto the temporal dimension during the storm movement (Houze, 1997; Rickenbach et al., 2008), among other factors.

3.3. Comparative analysis across regions, seasons, and times of day

We next explore potential differences in these patterns across various conditions characterized by different precipitation regimes, i.e., regions, seasons, and times of day (Figs. 5 and 6). For conciseness, two contrasting seasons/regions/time periods of day are selectively presented here, with the complete results provided in the supplementary material (Figs. S5-S11). Overall, despite the still clear dependency of the error metrics on event stage and duration across various scenarios, the specific shape and magnitude of the error curves vary significantly (Figs. 5 and 6).

For example, notable differences are present between summer and winter in terms of miss/false proportion (Fig. 5a1-a2 and b1-b2), with a lower miss proportion but higher false proportion in summer (JJA), and vice versa in winter (DJF). This aligns with the long-recognized error patterns of satellite precipitation retrievals: summer's prevalent deep convection generally accompanies lower cloud top temperatures and denser ice crystal concentrations that facilitate easier detection (and reduce false alarms) (AghaKouchak et al., 2012; Zhang et al., 2022). This knowledge is also reflected in the generally high average intensity of MRMS events during the summer (Figs. S12a_1-a_2). In contrast, winter's

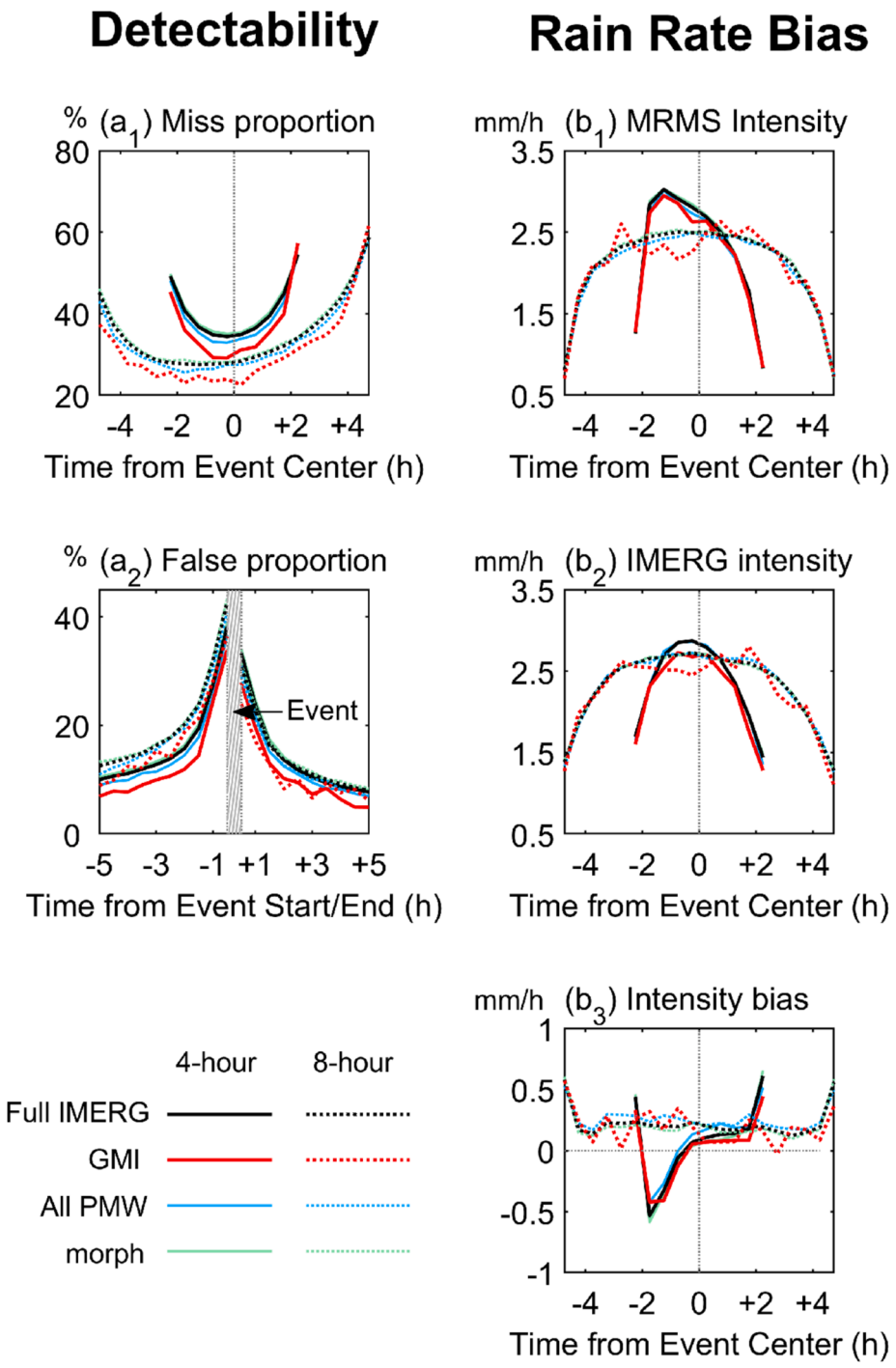

Fig. 3. Similar to Fig. 1 and Fig. 2, but showing the (a_1-a_2) detectability and (b_1-b_3) rain rate bias obtained from different data sources of IMERG, separately. Specifically, "Full IMERG" represents results calculated from full IMERG data, serving as the reference. "GMI" represents the results exclusively from the GMI sensor, while "PMW" indicates the results derived from any arbitrary PMW sensor and "morph" indicates the results from the interpolation sources in the absence of direct PMW observations. For clarity and conciseness, only results for 5-hour (solid lines) and 10-hour (dash lines) events are shown as representatives. The values are calculated from all the data across CONUS from 2018 to 2020.

snow/ice surface and shallow large-scale precipitation could contribute to the increasing likelihood of misses (Tian et al., 2009; Wang et al., 2021). Besides, the asymmetry in the miss/false proportion curves is more evident in winter compared to summer, which might be partially attributed to the longer fall time of snow relative to rainfall, leading to a temporal delay between the ground-level precipitation and the retrieved precipitation inferred from the cloud properties aloft (You et al., 2019). Such a discrepancy consequently results in the higher false proportion/lower miss proportion near the event starts and vice versa near the ends, depicted in Fig. 5a1-a2 and b1-b2.

The differences in error curves between the Northeastern (NE) and Southeastern (SE) U.S. can also be partly explained by the seasonal differences discussed above (Fig. 5a3-a4 and b3-b4), although the complex terrain in the Northeast characterized by mountainous snow-caps, high spatial variability in precipitation, and warm cloud precipitation could also contribute to the higher miss proportion/lower false proportion in this region (Derin and Kirstetter, 2022). In comparison, the aforementioned factors (i.e., rain versus snow, convective versus

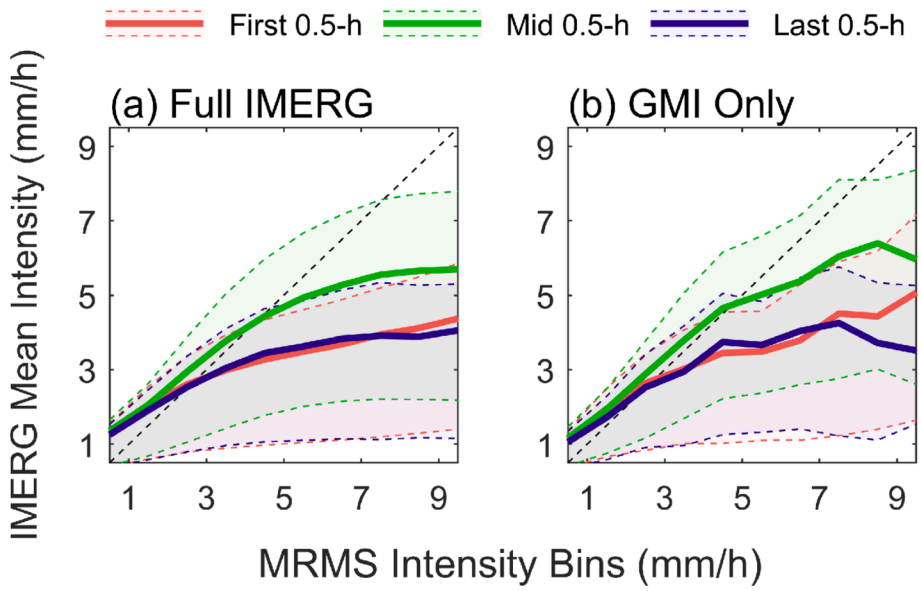

Fig. 4. Mean intensity correspondence curves (solid lines) between MRMS and IMERG during the initial (red), middle (green), and final (blue) half-hours of MRMS events. These curves are derived by calculating the mean IMERG intensity when the corresponding MRMS precipitation falls within each 1 mm/h bin (i.e., (0,1], (1,2], …, (9,10] mm/h). Curves representing the 25th and 75th percentile of IMERG intensities (dash lines) within each bin, along with the distribution range (shading areas) formed between these two lines, are also presented for reference. To obtain the middle timestep, results are derived from events with odd-numbered durations (i.e., 2.5, 3.5, …, 18.5, 19.5 h, 0.5-/1.5-hour events are not selected due to the blurred boundaries across different stages). The midpoint of each bin is set as the x-value (i.e., 0.5, 1.5, …, 9.5). The values are calculated from all the data across CONUS from 2018 to 2020.

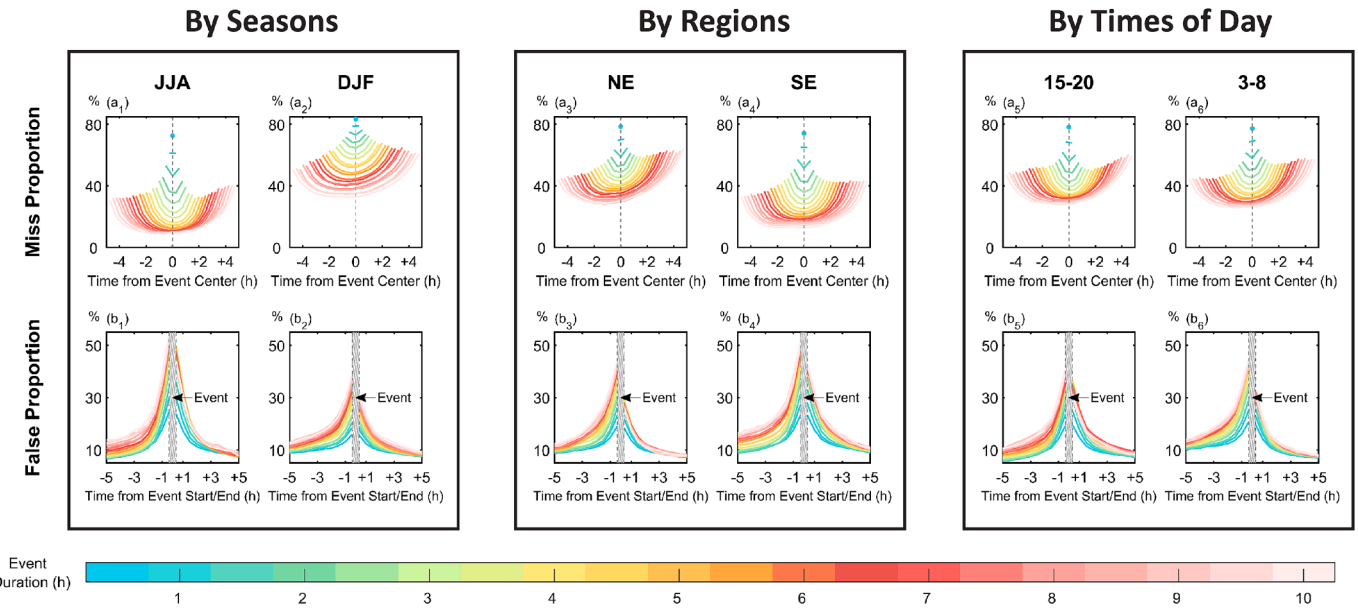

Fig. 5. Similar to Fig. 1, but illustrating the contrasts between two typical (a_1-a_2 and b_1-b_2) seasons (JJA and DJF), (a_3-a_4 and b_3-b_4) regions (NE (90°-65°W, 40°-49°N) and SE (90°-75°W, 25°-40°N)), and (a_5-a_6 and b_5-b_6) time periods of day (15–20 and 3–8 Local Solar Time (LST)). The diurnal timing of events is determined based on their peak times. Detailed analysis for other regions, seasons, and times of day is provided in the supplementary material (Figs. S6-S8).

large-scale precipitation, etc.) do not introduce a significant difference between late afternoon (15–20 LST) and early morning (3–8 LST) precipitation, with the detection discrepancy in the two diurnal periods less pronounced than in the previous two groups (Fig. 5a5-a6 and b5-b6). From the perspective of storm types, large-scale synoptic dynamics-driven precipitation may not have an evident response to the diurnal cycle of radiative forcing (Wang et al., 2022), while for convective precipitation, local convective precipitation induced by surface heating in the afternoon, as well as the Mesoscale Convective Systems (MCSs) typically peaking at night over CONUS (Feng et al., 2019), can both develop deep enough to allow satellites a comparable detectability for them. Notably, a slightly stronger asymmetry is observed in the morning curves, for which one potential reason could be the natural asymmetric spatial structures of MCSs that map onto the time dimension when considering a Eulerian scheme of reference (Houze, 2004).

From the "hit" and the overall intensity of MRMS events (Fig. 6a1-a6 and S11a_1-a_6), the summer, low latitudes, and late afternoon exhibit earlier peaks and higher average and peak intensities, compared to the winter, high latitudes, and early morning, respectively, which aligns with our knowledge of stronger convective activity in the first three

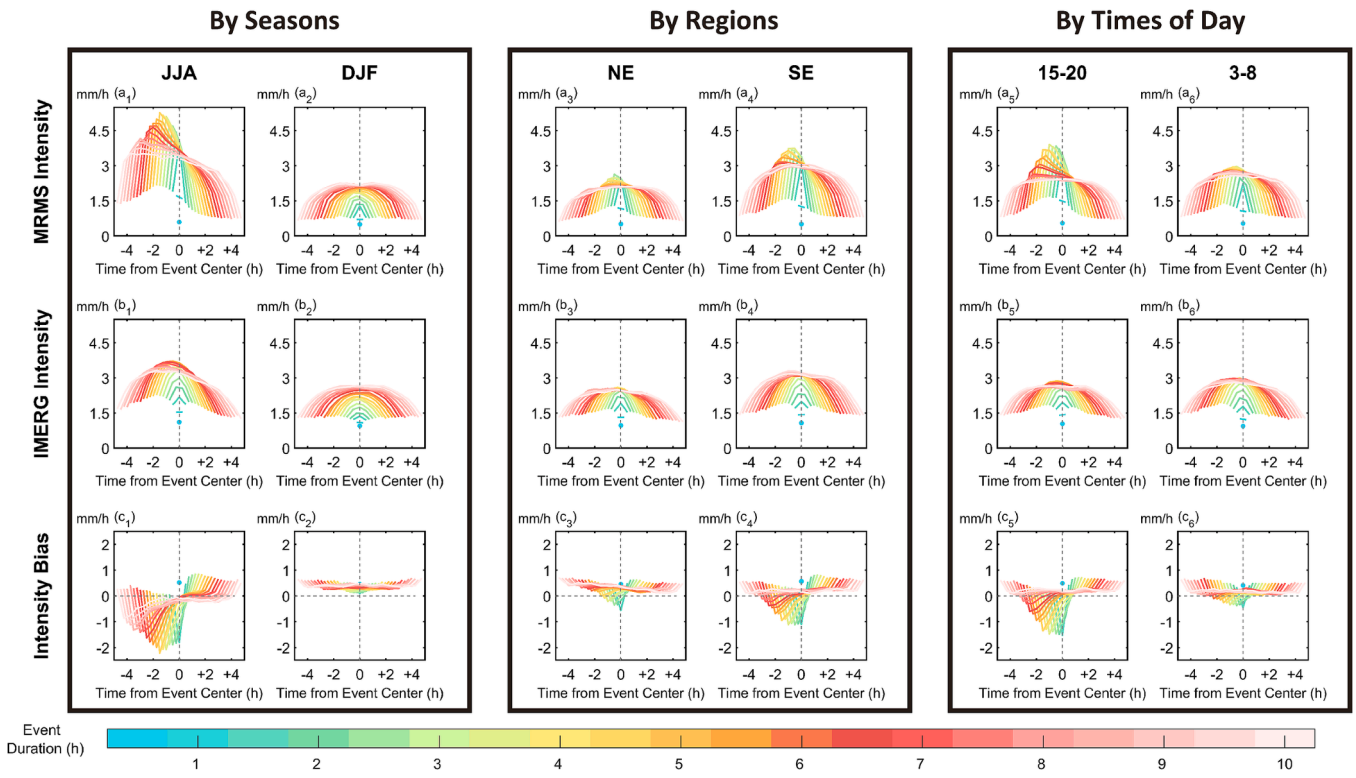

Fig. 6. Similar to Fig. 2, but illustrating the contrasts of error variation during events across between two typical (a_1-a_2, b_1-b_2, and c_1-c_2) seasons, (JJA and DJF), (a_3-a_4, b_3-b_4, and c_3-c_4) regions (NE and SE), and (a_5-a_6, b_5-b_6, and c_5-c_6) time periods of day (15–20 and 3–8 LST). The diurnal timing of events is determined based on their peak times. Detailed analysis for other regions, seasons, and times of day is provided in the supplementary material (Figs. S9-S11).

scenarios. Notably, a closer examination indicates more intense long-duration events at night but more intense short-duration events during the day (Fig. 6a5-a6 and S11a_5-a_6), echoing the presence of prolonged nocturnal MCSs and brief diurnal isolated convection (Yu et al., 2007), as discussed earlier. Regarding the hit intensity bias, IMERG tends to have a greater underestimation for MRMS events with higher average and peak intensities (Fig. 6c1-c6 and S11a_1-a_6), in line with previous research (Adhikari et al., 2019; Kirstetter et al., 2020; Petkovic and Kummerow, 2017). This relationship, combined with the earlier analysis (Fig. 5), results in the obvious pattern across various scenarios where the lower the miss proportion, the higher the false proportion is, and the greater the intensity bias.

From the perspective of asymmetry, the miss/false proportion with larger asymmetry typically corresponds to the intensity bias with smaller asymmetry (Figs. 6 and 7). Combined with Fig. S12, a possible explanation is that, for events with relatively symmetric intensity temporal profiles (e.g., Fig. 6 a2 and S11a_2), the miss/false proportion is inherently asymmetric (typically with higher false proportion/lower miss proportion near the event beginnings) due to factors like the advanced timing of IMERG events (e.g., Fig. 5a2 and b2) (Li et al., 2023b). In contrast, when event intensity profiles are asymmetric (typically with greater intensity in the first half) (e.g., Fig. 6a1 and S12a_1), it results in decreases (increases) of the miss (false) proportion at the initial stages, counterbalancing the effects of premature event initiation and resulting in the symmetric detectability (e.g., Fig. 5a1 and b1). Additionally, the similarity in error curve shapes (for the overall intensity, hit intensity, and missed intensity) of MRMS events across scenarios indicates that IMERG misses the precipitation of various intensities in a somewhat proportional manner (Fig. 5, S1, and S12). This again suggests that intensity is not the sole dominant driver behind varying detectability during the events, pointing instead to more fundamental underlying causes (e.g., shift in the inversion relationship).

3.4. *Indicative role of the event-associated environment*

The results of our analysis and the physical insights gained by comparing the error patterns in different precipitation systems, leads us to the hypothesis that these patterns might relate to cloud microphysical processes, which differ significantly between convective (rapid, asymmetric) and large-scale (stable, uniform) precipitation over their life cycles (Morrison et al., 2020). To test this hypothesis, we explore possible relationships between the error patterns and environmental variables indicative of storm type (particularly convective precipitation).

We first examine the statistical distribution of pre-event CAPE and dewpoint values from ERA5 across seasons, regions, and time periods of day described in Figs. 5 and 6 (Fig. 7). It is observed that the atmospheric instability (CAPE) and moisture availability (dewpoint) display large variability but are higher for summer, the southeast, and late afternoon, compared to winter, the northeast, and early morning, correspondingly. Although the magnitude of these variables does not absolutely characterize event types, it is generally true that strong convective weather, compared to large-scale-lifting-induced precipitation, tends to occur under conditions of high atmospheric instability and humidity level, which jointly provide essential energy and moisture support necessary for the initiation and development of deep moist convection (Dong et al., 2019; Yang et al., 2019). Hence, the results obtained from Fig. 7 are consistent with the commonly held understanding that convective precipitation proportions are greater for summer, the southeast, and late afternoon. This echoes well our conjecture in the analysis of Figs. 5 and 6 that "the similar error patterns across these three scenarios may correspond to stronger convective precipitation". Therefore, it can be concluded that the shape and magnitude of the error curves under various scenarios can indeed be linked to environmental conditions, with the proportion of convective precipitation serving as the key factor connecting them.

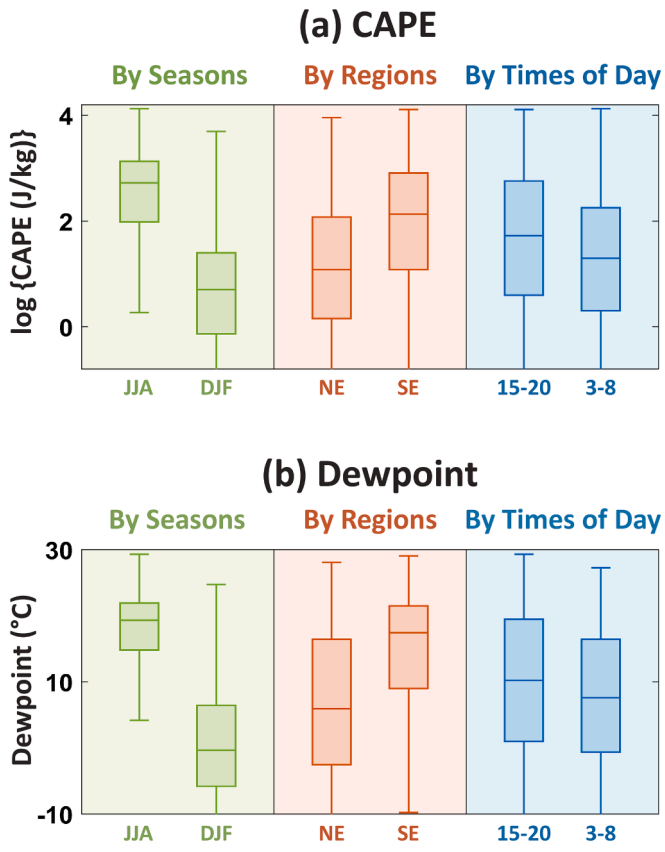

Fig. 7. Boxplots of the environmental conditions, specifically (a) CAPE and (b) Dewpoint of events across the two seasons, regions, and time periods of days depicted in Figs. 5 and 6. The environmental conditions of events are defined based on values recorded half an hour prior to events' onsets. The 25th (Q1), 50th (Q2), and 75th (Q3) percentiles, along with the whiskers (Q1 − 1.5IQR and Q3 + 1.5IQR; IQR (Interquartile range) = Q3 − Q1) are displayed in the boxplots.

Building on the above analysis, we directly investigate the relationship between the magnitude of the environmental variables and the shape of the error curves to explore the potential explanatory power of the former on the latter (Fig. 8). After extracting the CAPE and dewpoint for all events, we isolate events with their CAPE/dewpoint values falling within the top 30 % and bottom 30 % of the overall distributions of these values, respectively, and calculate their detectability and rain rate bias, as in Figs. 1 and 2. It is readily apparent from Fig. 8 that the error curves of events corresponding to the top and bottom 30 % of CAPE/dewpoint values exhibit significant disparities. Specifically, for larger CAPE/dewpoint we see a lower average miss proportion, higher false proportion, and greater intensity bias, as well as more symmetric false/miss curves and more asymmetric intensity bias curves. These results clearly demonstrate that these two environmental variables can play an indicative role in the error curves during the events, with their manifestations corresponding to the analytical results depicted in Figs. 5, 6, and 7.

The contrast in the error temporal patterns between the top 30 % and the bottom 30 % scenarios (Fig. 8) is even more pronounced than that between summer and winter, the northeast and the southeast, or afternoon and morning (Figs. 5 and 6), underscoring that environmental variables have larger explanatory power and are potentially better predictors of error patterns than season, region, and time of day. Moreover, environmental variables offer quantifiable, continuous, and practicable measures for delineating error curve shapes during the events. This is further evidenced by the progressively varying error curves calculated within successive percentiles (i.e., 0–10 %, 10–20 %, …, 90 %-100 %) of environmental variables (Figs. S14-S19). Additionally, the intensity correspondence curves at different event stages (similar to Fig. 4) under diverse environmental conditions (i.e., bottom/top 30 % of CAPE/dewpoint values) depicted in Fig. S20, not only again highlight the marked differences in the MRMS-to-IMERG intensity relationship across various stages but confirm that the extent of these differences varies with environmental conditions as well, offering the potential of using these variables for improving retrieval algorithms.

We further use two quantitative descriptors of the error curve shapes (whose spatial distributions are shown in Fig. S21 for reference), specifically the event-wide average and the asymmetry index as outlined in the Methods Section, to more distinctly quantify the variations in the error curve shape across each percentile interval of the CAPE/dewpoint values (Fig. 9). For the event-wide average values of miss proportion/false proportion/intensity bias (Fig. 9a1-c1), both variables exhibit a near-monotonic relationship with them. Consistent with the earlier analysis, higher CAPE/dewpoint values basically correspond to the lower miss proportion, higher false proportion, and lower intensity bias. Regarding the asymmetry index of the three error metrics (Fig. 9a2-c2), although the values across ten percentile intervals exhibit non-monotonic patterns, they still maintain a discernible and robust relationship with the environmental variables. For the miss proportion (Fig. 9a2), the predominance of positive asymmetry index values indicates a tendency for the miss curves to skew towards their latter half in most instances. Specifically, a shift from increasing to decreasing asymmetry index is observed at around the median CAPE/dewpoint value, with a transition asymmetry index at a high CAPE/dewpoint value (around 80 % quantile). The false proportion basically exhibits the opposite trend (Fig. 9b2), while the intensity bias follows a similar trend but with less variability within the first 50 % range, where its error curves primarily display a symmetric pattern (Fig. 9ac). A latter-stage skewness emerges only when exceeding 60 %, which corresponds to the maximum underestimation in the first halves of the intensity bias curves (Figs. S18 a7-a10 and S19 a7-a10). These results quantitatively demonstrate the complex interplay between precipitation and the storm environment, which manifests itself in non-linear relationships between the error characteristics and the environmental variables. However, the relationships between the examined environmental variables and the error characteristics are still well-defined, demonstrating the potential of such variables to improve retrieval algorithms, despite the need for further efforts to understand and parameterize them.

4. Discussion

In this study, our goal was to understand and quantify the temporal error patterns of satellite precipitation products at fine time scales and gain quantitative insight into the relation of these patterns to physical attributes of the storm environment. We followed a Eulerian approach, in which the temporal evolution of "precipitation events" (defined as uninterrupted half-hour intervals of > 0.01 mm/hr precipitation) was tracked at fixed locations using MRMS as the reference. Tracking "real" precipitation systems under a Lagrangian framework (i.e., follow the storm in space and time) using ground observations is quite challenging due to limited radar coverage especially over complex terrain (e.g., the Rocky Mountains) (Petersen et al., 2020; Zhang et al., 2016), the absence of radar observations over oceans where many storms originate (e.g., atmospheric rivers, hurricanes, Nor'easters) (Elsner, 2003; Rutz et al., 2019), and the complex splitting and merging behavior of precipitation systems that create ambiguities in determining their life cycle stages (Müller et al., 2022; Prein et al., 2024). These challenges also explain why very few studies have tracked precipitation systems with radar networks and used them to evaluate satellite products (Ayat et al., 2021). In contrast, the Eulerian event analysis is much simpler and easier to implement, which can be conducted wherever ground-based observations (whether from radar or gauges) are available.

The results of our study demonstrated a distinct time-dependence of

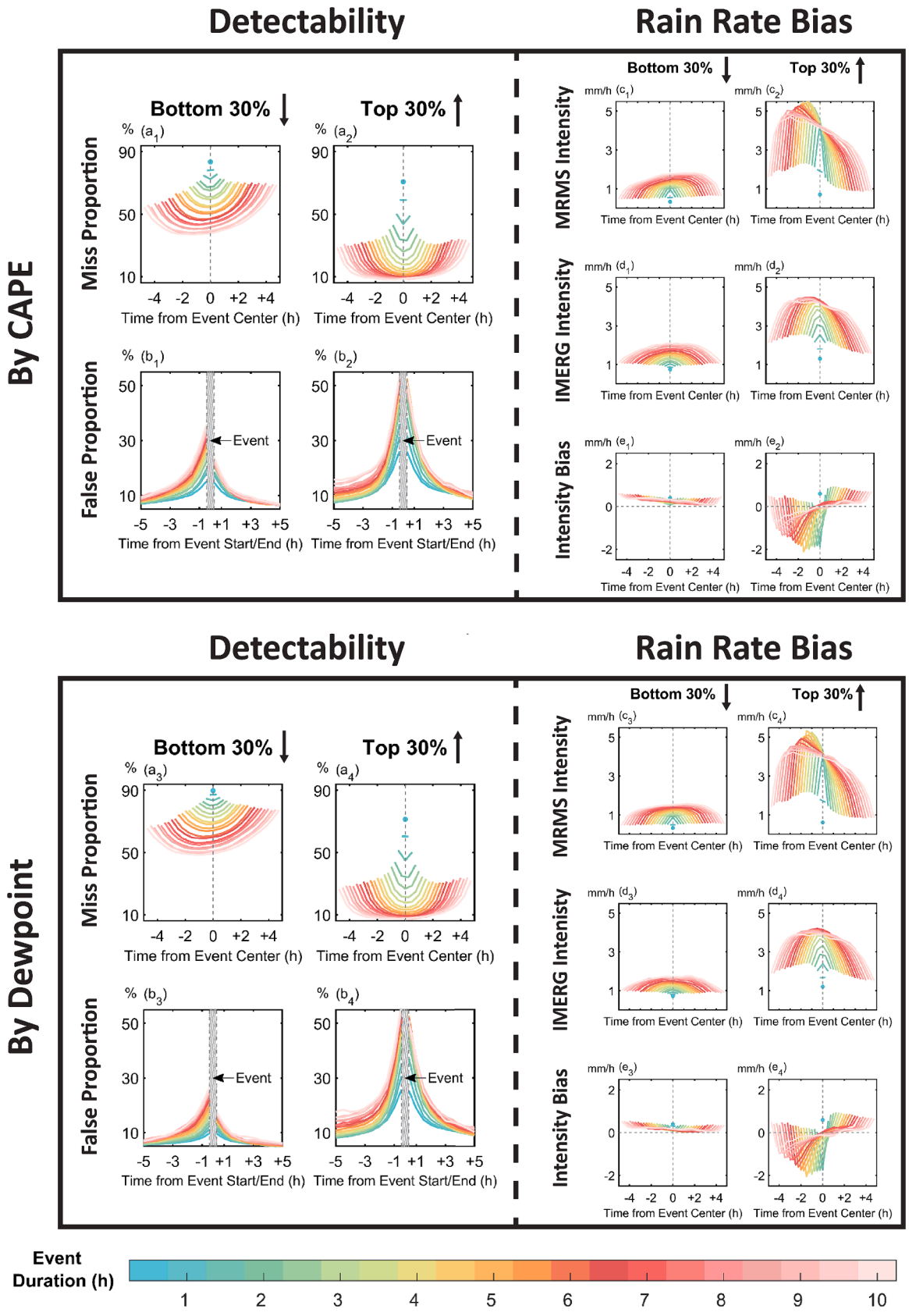

Fig. 8. Evolution of half-hourly (a_1-a_4 and b_1-b_4) miss/false proportion (similar to Fig. 1) and (c_1-c_4, d_1-d_4, and e_1-e_4) intensity bias (similar to Fig. 2) before/during/after the events under different environmental conditions (specifically CAPE (the upper orange panel) and Dewpoint (the lower purple panel) here). The environmental conditions of events are defined based on values recorded half an hour prior to the events' onset. Results computed from events with their environmental condition values falling within the bottom and top 30% of the overall distribution of these values are displayed for comparison. The values are calculated from all the data across CONUS from 2018 to 2020.

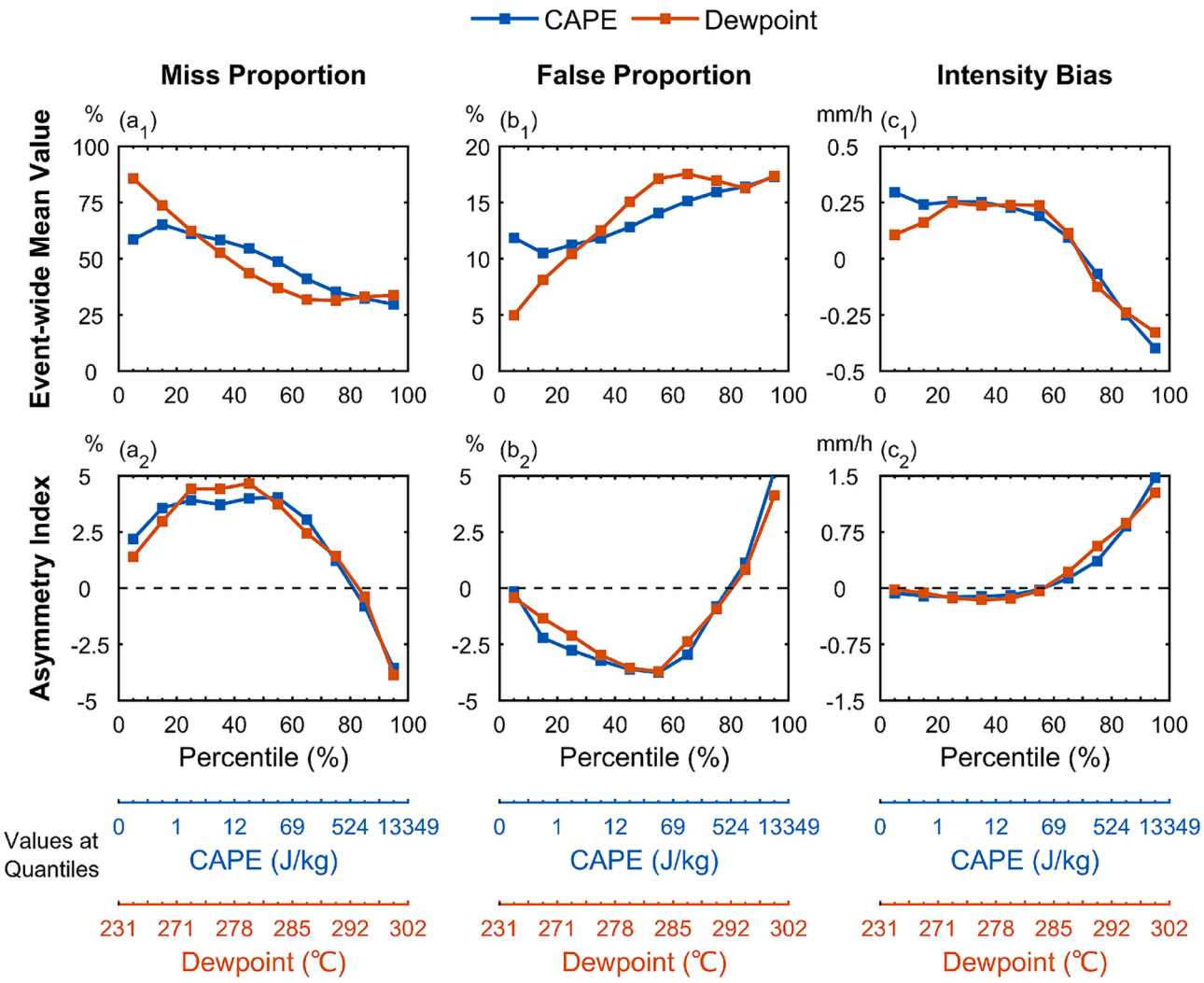

Fig. 9. (a_1-c_1) Event-wide averages and (a_2-c_c) the asymmetry indices of Miss Proportion/False Proportion/Intensity Bias curves under different pre-event CAPE (blue line) and Dewpoint (red line) conditions, which are classified based on the percentile intervals (i.e., 0%-10%, 10%-20%, …, 90%-100%) of the distribution of their values extracted from the entire event set. Data points are plotted at the midpoints of these intervals (i.e., 5%, 15%, …, 95%). Corresponding CAPE and dewpoint values for each percentile are also depicted for reference. The values are calculated from all the data across CONUS from 2018 to 2020.

the error curves during the event evolution and explored the potential of event-associated environmental variables (exemplified by CAPE and dewpoint) to provide unified quantitative indicators for the shapes of the error curves, towards their possible incorporation in the retrieval algorithms to reduce systematic biases. Considering that the operational Goddard Profiling Algorithm (GPROF) PMW retrieval already incorporates some environmental variables (two-meter surface Temperature (T2m), Total Column Water Vapor (TCWV), etc.) at the time of inversion (Passive Microwave Algorithm Team Facility, 2022), our results additionally demonstrate the predictive role of environmental variables on the overall satellite error evolution patterns during the events, emphasizing the importance of further exploration and utilization of a broader array of environmental variables to improve retrieval algorithms (Upadhyaya et al., 2022). Considering the complex relationship between precipitation and the environment (Berg et al., 2006; Liu et al., 2020), extensively integrating various model output environmental variables into traditional PMW retrieval algorithms for improvement is challenging and has thus been seldomly addressed in previous research. This approach, however, shows unprecedented promise within increasingly advanced deep learning frameworks (for example, Google's pure AI-driven weather forecasting model, GraphCast, successfully utilizes current and 6-hour earlier precipitation, along with basic environmental variables (temperature, wind, humidity, pressure, etc.) to autoregressively predict future precipitation up to 10 days ahead (Lam et al., 2023)), and is expected to play a critical role in future retrieval algorithms.

Based on our analysis, the proposed idealized guidance for satellite algorithm improvement is to use the pre-event environmental conditions to indicate the overall shape and magnitude of the error evolution curves during the events, and then, through the event-stage-dependent nature of errors, to leverage the stage information to further pinpoint the specific satellite error magnitudes. Yet, a practical challenge emerges as the actual precipitation event stage is not readily available in the absence of a ground-truth; it is thus necessary to seek alternative proxies of the "real" event stages. A potentially viable approach is to utilize not only the pre-event environmental conditions as we did here, but also the continuous temporal evolution patterns of environmental and cloud microphysics variables by identifying specific patterns related to the occurrence and distinct development stages of precipitation events. (Bouniol et al., 2016). This concept is theoretically grounded in cloud physics' natural evolution during precipitation processes (e.g., changes in the ice water content) (Braga and Vila, 2014), the effect of the environment on precipitation development and maintenance (e.g., the persistence of moisture transport) (Lee et al., 2016), the feedback of

precipitation itself on the environment (e.g., downburst-induced wind field alteration) (Romanic et al., 2022), and so on. Another implementable approach might be using the satellite-derived events' stage information as the proxy of the "real" stages, feeding it back into the products for iterative improvement. Fig. S22 performs such an analysis, similar to Figs. 1 and 2 but using IMERG-extracted events as reference. As demonstrated, the errors still show a clear dependency on the stages of the IMERG events, suggesting the potential of this method. Additionally, the cloud life cycle, easily obtained from IR sensors on geostationary satellites, could also be tried as an approximation for the precipitation life cycle, which has been explored in some studies (Guilloteau and Foufoula-Georgiou, 2024; Tadesse and Anagnostou, 2009).

5. Conclusions

Understanding the error patterns of satellite precipitation products at fine time scales (i.e., half-hourly) is crucial for their hydrologic, atmospheric, and other applications. Acknowledging that precipitation occurs discretely over time as "precipitation events" with specific start/end times and durations, this study explores the error evolution characteristics during the life cycle of the events and their relation to the storm environment. Using GV-MRMS as the reference product, IMERG is evaluated at its native resolution of $0.1^\circ \times 0.1^\circ$ and 30mins over CONUS. The main conclusions are as follows:

(1) The nationwide annual average results reveal significant variations in both detectability and rain rate bias of IMERG before/during/after the events. Specifically, both miss and false proportions peak near event margins. An asymmetry is observed in detectability, with higher miss proportion near event ends than the beginnings, and vice versa for the false proportion. IMERG and MRMS show diverse mean intensity profiles during the events, leading to an intensity bias pattern where IMERG overestimates near event beginnings/ends but underestimates in the early- to mid-stages. These findings underscore a pronounced dependency of IMERG errors on the stages of precipitation events.

(2) Error diagnoses performed using solely GMI-sourced data reveal similar error patterns during the events as observed with full IMERG data, indicating the inherent event-stage-dependent error patterns in PMW retrievals. Further investigation reveals evident variation in the MRMS-to-IMERG intensity correspondence curves across different event stages, which implies substantial shifts in the satellite-signal-to-rain-rate relationship throughout the events and suggests that the stage-dependent error pattern is not a simple temporal projection of intensity dependence. These diagnostic results indicate the inherent nature of satellite errors' dependency on event stages, whose more essential causes might be the cloud microphysics variations during the event life cycles.

(3) The analysis over different seasons, regions, and times of day reveals distinct event-stage-dependent error patterns, with significant differences between winter/the northeast/early morning and summer/the southeast/late afternoon. Specifically, the latter exhibits a larger more symmetric miss proportion, smaller more symmetric false proportion, and greater more asymmetric negative intensity, most likely associated with the increased presence of convective precipitation. Observed higher pre-event CAPE/dewpoint values under these situations additionally corroborate this point. Further comparison of error curves, calculated from events ranked within the top and bottom 30 % of the environmental conditions, demonstrates the indicative role of environmental variables in the shape error curves during the events. Two metrics that describe the shape and magnitude of the error curves further quantify the clear error dependence of the error patterns on the CAPE/dewpoint quantile.

In summary, this paper provides guidance and insights into the fine-scale error patterns of satellite precipitation products through the analysis of the error evolution during precipitation events. It highlights the limitations of current retrieval algorithms in fully resolving the inversion relationship which changes throughout the precipitation event life cycles due to changes in cloud microphysics, and conversely emphasizes the potential benefits of incorporating event stage information into future algorithms for enhanced accuracy. Moreover, the research demonstrates that storm environmental variables could serve as indicators not only of the retrievals at specific times (partially used in the current operational retrieval algorithms), but also for the error patterns throughout the entire event sequence. This highlights the necessity of more effectively utilizing various environmental variables (including CAPE/dewpoint and beyond) in future algorithms, although specific implementation strategies require further investigation. Future efforts should aim to integrate the insights from this study to practically explore and test the enhancement in the satellite precipitation product accuracy under deep learning frameworks.

CRediT authorship contribution statement

Runze Li: Writing – original draft, Visualization, Investigation, Formal analysis, Conceptualization. **Clement Guilloteau:** Writing – review & editing, Methodology, Investigation, Formal analysis. **Pierre-Emmanuel Kirstetter:** Writing – review & editing, Resources, Funding acquisition, Data curation. **Efi Foufoula-Georgiou:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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) Division of Information and Intelligent Systems (Expand AI2ES, Grant IIS 2324008).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jhydrol.2024.132610.

Data availability

The GV-MRMS data are available at NASA's Global Hydrometeorology Resource Center (GHRC) Distributed Active Archive Center (DAAC) (https://doi.org/10.5067/GPMGV/MRMS/DATA101). IMERG data are archived at NASA's Goddard Earth Sciences (GES) Data and Information Services Center (DISC) (https://10.5067/GPM/IMERG/3B-HH/07). ERA5 data are sourced from ECMWF's Copernicus Climate Change Service (C3S) Climate Data Store (CDS) (https://doi.org/10.24381/cds.adbb2d47).

References

Adhikari, A., Liu, C.T., Hayden, L., 2019. Uncertainties of GPM Microwave Imager Precipitation Estimates Related to Precipitation System Size and Intensity. J. Hydrometeorol. 20 (9), 1907–1923. https://doi.org/10.1175/Jhm-D-19-0038.1.

AghaKouchak, A., Mehran, A., Norouzi, H., Behrangi, A., 2012. Systematic and random error components in satellite precipitation data sets. Geophys. Res. Lett. 39. https://doi.org/10.1029/2012gl051592.
Ayat, H., Evans, J.P., Sherwood, S., Behrangi, A., 2021. Are Storm Characteristics the Same When Viewed Using Merged Surface Radars or a Merged Satellite Product? J. Hydrometeorol. 22 (1), 43–62. https://doi.org/10.1175/Jhm-D-20-0187.1.
Bai, H.D.Q., Schumacher, C., 2022. Topographic Influences on Diurnally Driven MJO Rainfall Over the Maritime Continent. J Geophys Res-Atmos 127 (6). https://doi.org/10.1029/2021JD035905.
Berg, W., L'Ecuyer, T., Kummerow, C., 2006. Rainfall climate regimes: The relationship of regional TRMM rainfall biases to the environment. J Appl Meteorol Clim 45 (3), 434–454. https://doi.org/10.1175/Jam2331.1.
Bi, K., et al., 2023. Accurate medium-range global weather forecasting with 3D neural networks. Nature 619 (7970), 533–538. https://doi.org/10.1038/s41586-023-06185-3.
Bouniol, D., Roca, R., Fiolleau, T., Poan, E., 2016. Macrophysical, Microphysical, and Radiative Properties of Tropical Mesoscale Convective Systems over Their Life Cycle. J. Clim. 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. J. Hydrometeorol. 15 (4), 1486–1497. https://doi.org/10.1175/Jhm-D-13-0206.1.
Derin, Y., Kirstetter, P.E., 2022. Evaluation of IMERG Over CONUS Complex Terrain Using Environmental Variables. Geophys. Res. Lett. 49 (19). 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. J. Hydrometeorol. 22 (11), 2843–2859. https://doi.org/10.1175/Jhm-D-21-0058.1.
Derin, Y., Kirstetter, P.E., Brauer, N., Gourley, J.J., Wang, J.X., 2022. Evaluation of IMERG Satellite Precipitation over the Land-Coast-Ocean Continuum. Part II: Quantification. J. Hydrometeorol. 23 (8), 1297–1314. https://doi.org/10.1175/Jhm-D-21-0234.1.
Dong, W.H., et al., 2019. Precipitable water and CAPE dependence of rainfall intensities in China. Clim. Dyn. 52 (5–6), 3357–3368. https://doi.org/10.1007/s00382-018-4327-8.
Elsner, J.B., 2003. Tracking hurricanes. Bull. Am. Meteorol. Soc. 84 (3), 353–356. https://doi.org/10.1175/Bams-84-3-353.
Feng, Z., et al., 2019. Spatiotemporal Characteristics and Large-Scale Environments of Mesoscale Convective Systems East of the Rocky Mountains. J. Clim. 32 (21), 7303–7328. https://doi.org/10.1175/Jcli-D-19-0137.1.
Freitas, E.D., et al., 2020. The performance of the IMERG satellite-based product in identifying sub-daily rainfall events and their properties. J. Hydrol. 589, 125128. https://doi.org/10.1016/j.jhydrol.2020.125128.
Gao, Y.C., Liu, M.F., 2013. Evaluation of high-resolution satellite precipitation products using rain gauge observations over the Tibetan Plateau. Hydrol. Earth Syst. Sci. 17 (2), 837–849. https://doi.org/10.5194/hess-17-837-2013.
Gebregiorgis, A.S., et al., 2018. To What Extent is the Day 1 GPM IMERG Satellite Precipitation Estimate Improved as Compared to TRMM TMPA-RT? J. Geophys. Res.: Atmos. 123 (3), 1694–1707. https://doi.org/10.1002/2017jd027606.
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. J. Hydrometeorol. 22 (11), 2805–2823. https://doi.org/10.1175/jhm-d-21-0075.1.
Guilloteau, C., Foufoula-georgiou, E., 2024. Life Cycle of Precipitating Cloud Systems from Synergistic Satellite Observations: Evolution of Macrophysical Properties and Precipitation Statistics from Geostationary Cloud Tracking and GPM Active and Passive Microwave Measurements. J. Hydrometeorol. 25 (5), 789–805. https://doi.org/10.1175/Jhm-D-23-0185.1.
Ham, Y.G., et al., 2023. Anthropogenic fingerprints in daily precipitation revealed by deep learning. Nature 622 (7982), 301–307. https://doi.org/10.1038/s41586-023-06474-x.
Hersbach, H., et al., 2020. The ERA5 global reanalysis. Q. J. r. Meteorolog. Soc. 146 (730), 1999–2049. https://doi.org/10.1002/qj.3803.
Heuscher, L., Liu, C.T., Gatlin, P., Petersen, W.A., 2022. Relationship Between Lightning, Precipitation, and Environmental Characteristics at Mid-/High Latitudes From a GLM and GPM Perspective. J Geophys Res-Atmos 127 (13). https://doi.org/10.1029/2022JD036894.
Houze, R.A., 1997. Stratiform Precipitation in Regions of Convection: A Meteorological Paradox? Bull. Am. Meteorol. Soc. 78 (10), 2179–2196. https://doi.org/10.1175/1520-0477(1997)078<2179:Spiroc>2.0.Co;2.
Houze, R.A., 2004. Mesoscale convective systems. Rev. Geophys. 42 (4), RG4003. https://doi.org/10.1029/2004rg000150.
Huffman, G.J. et al., 2023a. V07 IMERG Release Notes. 6 pp. [Available online at https://gpm.nasa.gov/sites/default/files/2023-07/IMERG_V07_ReleaseNotes_final_230713.pdf.].
Huffman, G.J. et al., 2023b. Integrated Multi-satellitE Retrievals for GPM (IMERG) Technical Documentation. 5 pp. [Available online at https://gpm.nasa.gov/sites/default/files/2023-07/IMERG_TechnicalDocumentation_final_230713.pdf.].
Imaoka, K., Nakamura, K., 2012. Statistical Analysis of the Life Cycle of Isolated Tropical Cold Cloud Systems Using MTSAT-1R and TRMM Data. Mon. Weather Rev. 140 (11), 3552–3572. https://doi.org/10.1175/Mwr-D-11-00364.1.
Kidd, C., et al., 2017. So, how much of the Earth's surface is covered by rain gauges? Bull. Am. Meteorol. Soc. 98 (1), 69–78. https://doi.org/10.1175/BAMS-D-14-00283.1.
Kirstetter, P.E., et al., 2012. Toward a Framework for Systematic Error Modeling of Spaceborne Precipitation Radar with NOAA/NSSL Ground Radar Based National Mosaic QPE. J. Hydrometeorol. 13 (4), 1285–1300. https://doi.org/10.1175/Jhm-D-11-0139.1.
Kirstetter, P.E., et al., 2014. Research Framework to Bridge from the Global Precipitation Measurement Mission Core Satellite to the Constellation Sensors Using Ground-Radar-Based National Mosaic QPE, Remote Sensing of the Terrestrial Water Cycle. Geophysical Monograph Series 61–79. https://doi.org/10.1002/9781118872086.ch4.
Kirstetter, P.E., Petersen, W.A., Kummerow, C.D., Wolff, D.B., 2020. Integrated Multi-satellite Evaluation for the Global Precipitation Measurement: Impact of Precipitation Types on Spaceborne Precipitation Estimation. In: Levizzani, V. (Ed.), Satellite Precipitation Measurement. Advances in Global Change Research. Springer International Publishing, Cham, pp. 583–608. https://doi.org/10.1007/978-3-030-35798-6_7.
O, S., Kirstetter, P.E., 2018. Evaluation of diurnal variation of GPM IMERG-derived summer precipitation over the contiguous US using MRMS data. Q. J. R. Meteorolog. Soc., 144: 270-281. https://doi.org/10.1002/qj.3218.
Kubota, T., et al., 2020. Global Satellite Mapping of Precipitation (GSMaP) Products in the GPM Era. In: Levizzani, V. (Ed.), Satellite Precipitation Measurement. Advances in Global Change Research. Springer International Publishing, Cham, pp. 355–373. https://doi.org/10.1007/978-3-030-24568-9_20.
Kukulies, J., Chen, D.L., Curio, J., 2021. The Role of Mesoscale Convective Systems in Precipitation in the Tibetan Plateau Region. J Geophys Res-Atmos 126 (23). https://doi.org/10.1029/2021JD035279.
Lam, R., et al., 2023. Learning skillful medium-range global weather forecasting. Science 382 (6677), 1416–1421. https://doi.org/10.1126/science.adi2336.
Lee, K.O., et al., 2016. Convective initiation and maintenance processes of two back-building mesoscale convective systems leading to heavy precipitation events in Southern Italy during HyMeX IOP 13. Q. J. r. Meteorolog. Soc. 142 (700), 2623–2635. https://doi.org/10.1002/qj.2851.
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. J Geophys Res-Atmos 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. J Geophys Res-Atmos 126 (1). https://doi.org/10.1029/2020JD033692.
Li, R.Z., Qi, D., Zhang, Y., Wang, K.C., 2022. A new pixel-to-object method for evaluating the capability of the GPM IMERG product to quantify precipitation systems. J. Hydrol. 613, 128476. https://doi.org/10.1016/j.jhydrol.2022.128476.
Li, R.Z., Guilloteau, C., Kirstetter, P.E., Foufoula-Georgiou, E., 2023a. How well does the IMERG satellite precipitation product capture the timing of precipitation events? J. Hydrol. 620, 129563. https://doi.org/10.1016/j.jhydrol.2023.129563.
Li, R.Z., Guilloteau, C., Kirstetter, P.E., Foufoula-Georgiou, E., 2023b. A New Event-Based Error Decomposition Scheme for Satellite Precipitation Products. Geophys. Res. Lett. 50 (22). https://doi.org/10.1029/2023GL105343.
Liu, N.N., Liu, C.T., Chen, B.H., Zipser, E., 2020. What Are the Favorable Large-Scale Environments for the Highest-Flash-Rate Thunderstorms on Earth? J. Atmos. Sci. 77 (5), 1583–1612. https://doi.org/10.1175/jas-d-19-0235.1.
Loew, A., et al., 2017. Validation practices for satellite-based Earth observation data across communities. Rev. Geophys. 55 (3), 779–817. https://doi.org/10.1002/2017rg000562.
Maggioni, V., Meyers, P.C., Robinson, M.D., 2016. A Review of Merged High-Resolution Satellite Precipitation Product Accuracy during the Tropical Rainfall Measuring Mission (TRMM) Era. J. Hydrometeorol. 17 (4), 1101–1117. https://doi.org/10.1175/Jhm-D-15-0190.1.
Milani, L., et al., 2021. Extreme Lake-Effect Snow from a GPM Microwave Imager Perspective: Observational Analysis and Precipitation Retrieval Evaluation. Journal of Atmospheric and Oceanic Technology 38 (2), 293–311. https://doi.org/10.1175/jtech-d-20-0064.1.
Morrison, H. et al., 2020. Confronting the Challenge of Modeling Cloud and Precipitation Microphysics. J. Adv. Model. Earth Syst., 12(8). https://doi.org/ARTN e2019MS001689 10.1029/2019MS001689.
Müller, S.K., et al., 2022. Evaluation of Alpine-Mediterranean precipitation events in convection-permitting regional climate models using a set of tracking algorithms. Clim. Dyn. 61 (1–2), 939–957. https://doi.org/10.1007/s00382-022-06555-z.
Passive Microwave Algorithm Team Facility, 2022. GLOBAL PRECIPITATION MEASUREMENT (GPM) MISSION Algorithm Theoretical Basis Document GPROF2021 Version 1 (used in GPM V7 processing). 20 pp. [Available online at https://gpm.nasa.gov/sites/default/files/2022-06/ATBD_GPM_V7_GPROF.pdf.].
Petersen, W.A., Kirstetter, P.E., Wang, J.X., Wolff, D.B., Tokay, A., 2020. The GPM Ground Validation Program. In: Levizzani, V. (Ed.), Satellite Precipitation Measurement. Advances in Global Change Research. Springer International Publishing, Cham, pp. 471–502. 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. J Appl Meteorol Clim 56 (3), 597–614. https://doi.org/10.1175/Jamc-D-16-0174.1.
Pradhan, R.K., et al., 2022. Review of GPM IMERG performance: A global perspective. Remote Sens. Environ. 268, 112754. https://doi.org/10.1016/j.rse.2021.112754.
Prein, A.F., et al., 2024. Km-Scale Simulations of Mesoscale Convective Systems Over South America-A Feature Tracker Intercomparison. J Geophys Res-Atmos 129 (8). https://doi.org/10.1029/2023JD040254.
Rickenbach, T., et al., 2008. The Relationship between Anvil Clouds and Convective Cells: A Case Study in South Florida during CRYSTAL-FACE. Mon. Weather Rev. 136 (10), 3917–3932. https://doi.org/10.1175/2008mwr2441.1.

Romanic, D., Taszarek, M., Brooks, H., 2022. Convective environments leading to microburst, macroburst and downburst events across the United States. Weather Clim. Extremes 37. https://doi.org/10.1016/j.wace.2022.100474.
Rutz, J.J., et al., 2019. The Atmospheric River Tracking Method Intercomparison Project (ARTMIP): Quantifying Uncertainties in Atmospheric River Climatology. J Geophys Res-Atmos 124 (24), 13777–13802. https://doi.org/10.1029/2019jd030936.
Sadeghi, M., et al., 2021. PERSIANN-CCS-CDR, a 3-hourly 0.04 degrees global precipitation climate data record for heavy precipitation studies. Sci. Data 8 (1), 157. https://doi.org/10.1038/s41597-021-00940-9.
Samanta, S., et al., 2021. The Life Cycle of a Stationary Cloud Cluster during the Indian Summer Monsoon: A Microphysical Investigation Using Polarimetric C-Band Radar. Mon. Weather Rev. 149 (11), 3761–3780. https://doi.org/10.1175/Mwr-D-20-0274.1.
Song, F.F., et al., 2019. Contrasting Spring and Summer Large-Scale Environments Associated with Mesoscale Convective Systems over the US Great Plains. J. Clim. 32 (20), 6749–6767. https://doi.org/10.1175/Jcli-D-18-0839.1.
Sutton, J.R.P., Kirschbaum, D., Stanley, T., Orland, E., 2024. Evaluating Precipitation Events Using GPM IMERG 30-Minute Near-Real-Time Precipitation Estimates. J. Hydrometeorol. 25 (7), 991–1006. https://doi.org/10.1175/jhm-d-23-0141.1.
Tadesse, A., Anagnostou, E.N., 2009. The Effect of Storm Life Cycle on Satellite Rainfall Estimation Error. J. Atmos. Oceanic Technol. 26 (4), 769–777. https://doi.org/10.1175/2008jtecha1129.1.
Tian, Y.D., et al., 2009. Component analysis of errors in satellite-based precipitation estimates. J Geophys Res-Atmos 114, D24101. https://doi.org/10.1029/2009jd011949.
Upadhyaya, S.A., Kirstetter, P.E., Gourley, J.J., Kuligowski, R.J., 2020. On the Propagation of Satellite Precipitation Estimation Errors: From Passive Microwave to Infrared Estimates. J. Hydrometeorol. 21 (6), 1367–1381. https://doi.org/10.1175/jhm-d-19-0293.1.
Upadhyaya, S.A., Kirstetter, P.E., Kuligowski, R.J., Searls, M., 2022. Towards improved precipitation estimation with the GOES-16 advanced baseline imager: Algorithm and evaluation. Q. J. r. Meteorolog. Soc. 148 (748), 3406–3427. https://doi.org/10.1002/qj.4368.
Wang, D., et al., 2022. Linking Synoptic Patterns to Cloud Properties and Local Circulations Over Southeastern Texas. J Geophys Res-Atmos 127 (5). https://doi.org/10.1029/2021JD035920.
Wang, J.X., Petersen, W.A., Wolff, D.B., 2021. Validation of Satellite-Based Precipitation Products from TRMM to GPM. Remote Sens. 13 (9). https://doi.org/10.3390/rs13091745.
Xie, P.P., et al., 2017. Reprocessed, Bias-Corrected CMORPH Global High-Resolution Precipitation Estimates from 1998. J. Hydrometeorol. 18 (6), 1617–1641. https://doi.org/10.1175/Jhm-D-16-0168.1.
Yamamoto, M.K., Furuzawa, F.A., Higuchi, A., Nakamura, K., 2008. Comparison of diurnal variations in precipitation systems observed by TRMM PR, TMI, and VIRS. J. Clim. 21 (16), 4011–4028. https://doi.org/10.1175/2007jcli2079.1.
Yang, Y., Gan, T.Y., Tan, X.Z., 2019. Spatiotemporal Changes in Precipitation Extremes over Canada and Their Teleconnections to Large-Scale Climate Patterns. J. Hydrometeorol. 20 (2), 275–296. https://doi.org/10.1175/Jhm-D-18-0004.1.
You, Y.L., Meng, H., Dong, J., Rudlosky, S., 2019. Time-Lag Correlation Between Passive Microwave Measurements and Surface Precipitation and Its Impact on Precipitation Retrieval Evaluation. Geophys. Res. Lett. 46 (14), 8415–8423. https://doi.org/10.1029/2019gl083426.
Yu, R.C., Xu, Y.P., Zhou, T.J., Li, J., 2007. Relation between rainfall duration and diurnal variation in the warm season precipitation over central eastern China. Geophys. Res. Lett. 34 (13). https://doi.org/10.1029/2007gl030315.
Yuan, X., et al., 2023. A global transition to flash droughts under climate change. Science 380 (6641), 187–191. https://doi.org/10.1126/science.abn6301.
Zhang, J., et al., 2016. MULTI-RADAR MULTI-SENSOR (MRMS) QUANTITATIVE PRECIPITATION ESTIMATION Initial Operating Capabilities. Bull. Am. Meteorol. Soc. 97 (4), 621–637. https://doi.org/10.1175/Bams-D-14-00174.1.
Zhang, L.L., Chen, X.H., Lai, R.Y., Zhu, Z.H., 2022. Performance of satellite-based and reanalysis precipitation products under multi-temporal scales and extreme weather in mainland China. J. Hydrol. 605. https://doi.org/10.1016/j.jhydrol.2021.127389.
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. J. Hydrol. 620. https://doi.org/10.1016/j.jhydrol.2023.129386.
Zhang, W., Villarini, G., 2019. On the weather types that shape the precipitation patterns across the US Midwest. Clim. Dyn. 53 (7–8), 4217–4232. https://doi.org/10.1007/s00382-019-04783-4.