



## Evaluation of IMERG Satellite Precipitation over the Land–Coast–Ocean Continuum. Part I: Detection

YAGMUR DERIN,<sup>a,b</sup> PIERRE-EMMANUEL KIRSTETTER,<sup>a,b,c,d</sup> AND JONATHAN J. GOURLEY<sup>d</sup>

<sup>a</sup> *Advanced Radar Research Center, University of Oklahoma, Norman, Oklahoma*

<sup>b</sup> *School of Civil Engineering and Environmental Sciences, University of Oklahoma, Norman, Oklahoma*

<sup>c</sup> *School of Meteorology, University of Oklahoma, Norman, Oklahoma*

<sup>d</sup> *NOAA/National Severe Storms Laboratory, Norman, Oklahoma*

(Manuscript received 20 March 2021, in final form 8 July 2021)

**ABSTRACT:** As a fundamental water flux, quantitative understanding of precipitation is important to understand and manage water systems under a changing climate, especially in transition regions such as the coastal interface between land and ocean. This work aims to assess the uncertainty in precipitation detection over the land–coast–ocean continuum in the Integrated Multisatellite Retrievals for Global Precipitation Measurement (IMERG) V06B product. It is examined over three coastal regions of the United States—the West Coast, the Gulf of Mexico, and the East Coast, all of which are characterized by different topographies and precipitation climatologies. Detection capabilities are contrasted over different surfaces (land, coast, and ocean). A novel and integrated approach traces the IMERG detection performance back to its components (passive microwave, infrared, and morphing-based estimates). The analysis is performed by using high-resolution, high-quality Ground Validation Multi-Radar/Multi-Sensor (GV-MRMS) rainfall estimates as ground reference. The best detection performances are reported with PMW estimates (hit rates in the range [25%–39%]), followed by morphing ([20%–34%]), morphing + IR ([17%–27%]) and IR ([11%–16%]) estimates. Precipitation formation mechanisms play an important role, especially in the West Coast where orographic processes challenge detection. Further, precipitation typology is shown to be a strong driver of IMERG detection. Over the ocean, IMERG detection is generally better but suffers from false alarms ([10%–53%]). Overall, IMERG displays nonhomogeneous precipitation detection capabilities tracing back to its components. Results point toward a similar behavior across various land–coast–ocean continuum regions of the CONUS, which suggests that results can be potentially transferred to other coastal regions of the world.

**KEYWORDS:** Precipitation; Remote sensing; Satellite observations; Statistical techniques; Coastal meteorology; Radars/radar observations; Uncertainty

### 1. Introduction

More than a billion people are estimated to live in coastal regions globally (Neumann et al. 2015). Climate change and anthropogenic activities modify the water fluxes and processes between land and ocean. Rising sea levels and changing precipitation patterns may enhance sea–land exchange and expose these regions to more impactful hazards (i.e., hurricanes and floods) from both ocean and land. As a fundamental water flux, precipitation is particularly important to study in transition regions that involve diverse processes such as the coastal interface between land and ocean. A detailed understanding and observation of precipitation over the land–coast–ocean continuum is still lacking, especially since observations over oceans are limited by remoteness and scale. Observations of precipitation over the ocean can be conducted by using weather radars situated on islands and coastlines and with rain gauges on ships and buoy gauge arrays (e.g., Anagnostou et al. 1999; Smith et al.

2009; Bowman et al. 2003; Serra and McPhaden 2003; Morrissey et al. 2012). The weather radars provide high-spatiotemporal-resolution precipitation measurements; however, they are limited in coverage. The buoy gauges take point measurements at high temporal resolution, but they have limited spatial representativeness and uncertainty associated with wind-induced undercatch, stability of the measuring platforms, and evaporation during daytime (Sapiano and Arkin 2009; Serra et al. 2001; Serra 2018). Hence, oceans remain among the regions that suffer from scarcity of precipitation measurements.

Satellite-based precipitation products (SPPs) provide the best coverage spatially and temporally at the global scale and especially over the land–coast–ocean continuum. The great success of the Tropical Rainfall Measuring Mission (TRMM) has led to the Global Precipitation Measurement Mission (GPM) and accelerated the development of new SPP algorithms, such as the Integrated Multisatellite Retrievals for Global Precipitation Measurement (IMERG). IMERG merges available Level-2 passive microwave (PMW) and infrared (IR) precipitation retrievals at a high spatial (0.1°) and temporal resolution (30 min). Yet, multiple uncertainties are

Corresponding author: Pierre-Emmanuel Kirstetter, pierre.kirstetter@noaa.gov

associated with Level-2 PMW and IR precipitation retrievals and Level-3 merged SPPs, especially in coastal regions. Precipitation retrievals from PMW observations highly depend on the ability to separate emission and scattering radiances of rain, ice, and clouds from Earth's surface. This is a complicated task since the ocean surface is radiometrically cold and homogeneous, while the land surface is warm and heterogeneous. Coastal regions include contrasting radiative contributions from both ocean and land surfaces, and hence are particularly challenging regions that yield among the most uncertain satellite precipitation retrievals. PMW algorithms utilize a surface mask to condition different Bayesian algorithms over ocean and land to constrain overland retrievals. GPM-era retrievals have not been evaluated over the land–coast–ocean continuum yet, and both the GPM ground validation (GV) group and IMERG algorithm developers highlight the importance of evaluation studies in “nontraditional” regions such as coasts and ocean.

Ground-based precipitation measurements provide an independent source of validation to determine the uncertainty characteristics of SPPs. The performance of GPM-era and TRMM-era products has been studied in detail over the globe (McCollum et al. 2002; Ebert et al. 2007; Anagnostou et al. 2010; Derin and Yilmaz 2014; Derin et al. 2016, 2019; Mei et al. 2014; Maggioni et al. 2016, 2017; Gebregiorgis et al. 2017, 2018; Manz et al. 2017; Ayat et al. 2021). Several studies focused on SPPs' uncertainty characteristics over the ocean. Wolff and Fisher (2009) assess the relative performance and skill of multiple PMW retrievals at the instantaneous time scale over two GV sites. Acknowledging that the surface type affects the relative performance of SPP algorithms, they conditioned the performance analyses based on ocean, land and coast surfaces. They found that TRMM PMW sensors overestimated precipitation rates over land and coast and underestimated precipitation over the ocean. Carr et al. (2015) evaluated TMI against the radar–gauge Ground Validation Multi-Radar/Multi-Sensor (GV-MRMS) over different surface types over the CONUS. They also found that over the ocean the TMI estimates were underestimating precipitation relative to GV-MRMS, although the convective fraction was overestimated. Khan and Maggioni (2019) evaluated IMERG, version 5 (IMERGV05), including its passive microwave (IMERG-PMW) and the infrared (IMERG-IR) components, against the shipboard OceanRAIN precipitation dataset (Ocean Rainfall and Ice-Phase Precipitation Measurement Network) and Level-3 DPR (3DPRD) at 0.5° resolution and at the daily time scale. They concluded that IMERGV05 correctly detected precipitation 80% of the time, but IMERG-IR showed poor linear correlation and large departures with the daily reference, and IMERG-PMW performed better than other products in terms of precipitation detection and quantification. Note that the variability of precipitation at all scales challenges the comparison of point measurements like OceanRAIN and SPP gridded estimates. In general, analyses focusing on the evaluation of the SPPs over the ocean have been performed over limited periods or regions, which limits their representativeness. Comparisons at time scales finer than daily (e.g., from 30 min to 3 h) are needed to characterize the ability of high-resolution satellite estimates to capture the precipitation variability. While some studies focused on the land–coast–ocean continuum (Carr et al. 2015; Wolff and Fisher 2009) for

TRMM-era algorithms, the performance of GPM-era algorithms remains largely unknown.

The primary objective of this study is to evaluate IMERGV06B (hereinafter IMERG) over the land–coast–ocean continuum characterized by high  $T_b$  (brightness temperature) radiometric gradients at the native resolution (0.1° spatial and 30 min temporal) of IMERG. Because of differences in satellite-based instrument operating frequencies, footprint sizes, and scan strategies, sensors in the GPM constellation have different sensitivities to the land–coast–ocean transition zone. Their associated precipitation retrievals exhibit different performances, which represents a challenge for the intercalibration of all PMW estimates that aims to mitigate the mean bias of each sensor in IMERG. Knowledge of biases in the calibrated PMW and IR estimates can help the IMERG developers refine the intercalibration process and improve precipitation in areas that are among the most populated globally. Our aim is to identify the sources of uncertainty and evaluate IMERG against the ground reference. This is performed through a novel and integrated approach using ancillary variables provided by IMERG to trace the uncertainty back to the sources of precipitation estimates. The analysis is conducted by providing statistics including two-dimensional histogram skill scores and graphical measures to complete the picture of SPP performance. To achieve this goal, IMERG Final, Late, and Early; IMERG-PMW; and IMERG-IR abilities to detect precipitation over the land–coast–ocean continuum were examined and linked to surface and precipitation characteristics.

It is critical to evaluate the transferability of regional evaluation results to different but similar areas. This important question has implications in terms of assessing SPPs performances in regions of the world devoid of ground observations. The question of transferability is addressed by examining the similarities and differences in IMERG detection performances over different coastal regions of the CONUS that are characterized by various topographies and precipitation climatologies. Three different land–coast–ocean continuum regions are considered, namely, the West Coast, Gulf of Mexico, and East Coast. The evaluation is conducted against the radar- and gauge-based GV-MRMS over a period of 12 months in 2015. The evaluation of the IMERG is conducted in two broad steps as detection and quantification analysis. This first part (Part I) of the study encompasses the detection step, and a companion paper (Part II; Y. Derin et al., unpublished manuscript) encompasses the quantification step over the same study region and period.

This paper is organized as follows. Section 2 provides the details of the study area and datasets. The evaluation method is represented in section 3, and section 4 discusses the results. Section 5 summarizes our conclusions and offers recommendations for future research directions.

## 2. Study region and dataset

### a. Study region

The CONUS has diverse climatic and topographic features. In this study we will be focusing on the land–coast–ocean continuum, and we divide this continuum into the following three regions: West Coast, Gulf of Mexico Coast, and East

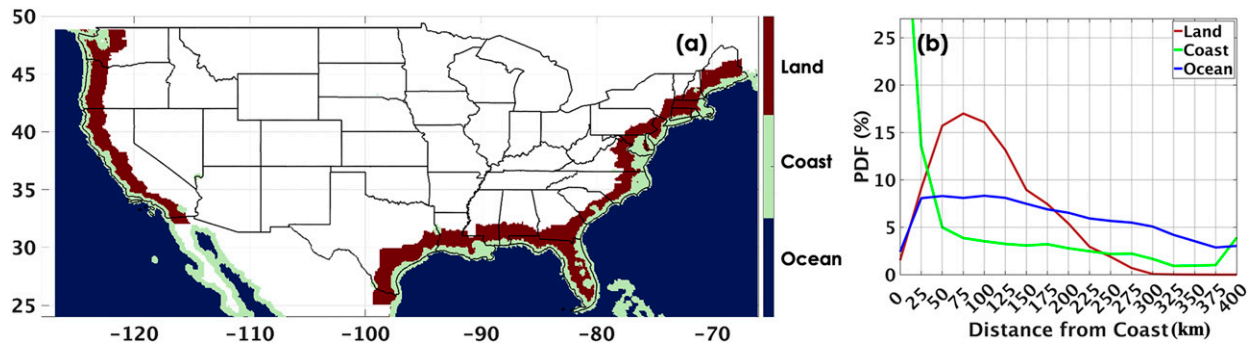


FIG. 1. The study region, with (a) a spatial map of the surface type classification and (b) the distribution of each surface type as a function of distance from the coastline.

Coast. The West (Pacific) Coast of the United States is characterized by steep mountain chains. This region experiences among the highest precipitation rates in the United States brought by the combination of large water vapor fluxes impinging upon steeply sloping terrain and the basins generate the highest number of floods and the highest magnitude of flooding (Saharia et al. 2017). Over this region orographic mechanisms tend to generate low-level enhancement of precipitation with lower ice content aloft. It challenges satellite precipitation estimation since the signature of ice content in the  $T_b$  is a primary driver of satellite precipitation over land. During the warm season, the cool marine layer and coastal mountains lead to marine stratus and fog. During the winter season synoptic-scale cyclones and fronts from the Pacific Ocean undergo complex interactions with coastal mountains within 50 km of the coastline. In contrast to the West Coast, the Gulf of Mexico and tropical southeast Atlantic coasts are characterized by flatter terrain. During the warm season, sea-breeze fronts, convective storms, and tropical cyclones are observed, while during the cool season the Gulf Coast is dominated by the return flow of Gulf-modified warm, moist air and cold front passages. The Atlantic coast (East Coast) extending from Georgia to the Northeast is characterized by the Appalachian Mountains and complex coastal features consisting of barrier islands and coastal inlets. During the warm season, the Gulf Stream is significantly different from the land in terms of wind, humidity and temperature. During the cool season, offshore Gulf Stream frontal features are combined with cold air over land, producing complex mesoscale cloud systems. These mesoscale cloud systems include intense coastal cyclogenesis that often brings freezing rain, sleet, and snow on regions extending from North Carolina through New England.

### b. Surface type classification

Surface type is determined by using the GPROFV05 surface type dataset (Fig. 1a). This classification conditions the PMW retrievals of precipitation, hence it is consistent with IMERG. The surface classification identifies three surface types as land, ocean and coast for each pixel. This classification is derived from the Colorado State University surface classification scheme (NASA 2017) at  $0.25^\circ$  resolution and monthly intervals and based on surface microwave emissivity. The surface type

classification begins with ten land classes using self-similar mean emissivities from all available SSM/I observations from 1993 to 2008 under clear sky conditions (Prigent et al. 1997), to which ocean, sea ice, land-ocean, and ocean-sea ice surface types are added. Then the GPROF algorithm defines the GPM land/ocean boundary by using an  $\sim 1$ -km MODIS/SeaWiFS/ Ocean Color land mask based on the World Vector Shoreline database. As can be seen on Fig. 1a, this classification allows inland pixels to be classified as “coast.” It should be noted that surface properties like emissivity are characterized by high gradients in these transition regions. Figure 1b provides the distributions of the distances between the matched grids and the coastline taken from the NOAA medium resolution shoreline dataset (NOAA; <https://shoreline.noaa.gov/data/datasheets/medres.html>). While most of the coast pixels are within 50 km of the coastline, some of them are located as far as 300 km inland. Most land-matched pixels are located in the 25–150 km range from the coastline. Given the range of WSR-88D instruments, ocean-matched pixels are located up to 425 km from the coastline. We select ocean pixels within 100 km to the nearest radar to ensure optimal radar sampling, and we perform the analysis only with the higher-quality GV-MRMS reference.

### c. GV-MRMS

The evaluation of SPPs requires deriving high-quality reference rainfall datasets at the satellite product pixel spatial and temporal resolution. In this study as a reference dataset the high-resolution, ground-based, radar-rain gauge corrected precipitation dataset GV-MRMS (Kirstetter et al. 2012, 2018) is used. GV-MRMS builds on MRMS that uses advanced data integration techniques to create high-resolution 3D reflectivity mosaic grids and quantitative precipitation estimates at  $0.01^\circ$  spatial and 2-min temporal scales (Zhang et al. 2016). GV-MRMS blends radar and gauge data at the maximum time and space resolution providing a high-quality reference dataset, which is matched to IMERG’s native time and spatial resolution ( $0.1^\circ$  and 30 min). It provides a radar quality index (RQI) to represent the level of uncertainty associated with reflectivity changes with height and near the melting layer. A high-quality and standardized reference is obtained by filtering less trustworthy GV-MRMS estimates associated with lower RQI estimates. Additional gauge-based corrections,

quality and quantity controls, and resampling procedures are applied and described in detail in [Kirstetter et al. \(2012, 2014, 2015\)](#). In this study the radar- and gauge-based GV-MRMS is used over a period of 12 months in 2015.

#### d. IMERG algorithm

The IMERG algorithm is designed to intercalibrate, merge, and interpolate all available satellite microwave retrievals, microwave-calibrated IR satellite estimates, and rain gauge measurements. The merging of different sensors is conducted to compensate for the limited sampling of low-Earth-orbiting PMW satellites. Level-2 PMW precipitation is retrieved using the GPROF2017 (GPROFV05) Bayesian-based algorithm ([Huffman et al. 2019](#)). Basically, GPROF2017 searches an a priori database (based on DPR and GMI of the previous version of GPROF algorithm) for potential rain profiles and retrieves a weighted average of these precipitation profiles. Note that the PMW retrieval is conditioned by the surface microwave emissivity, which in turn depends on dielectric properties, the roughness, and the material of the surface. Specifically, the emissivity from ocean, land, and ice/snow-covered surfaces present large differences. The ocean surface has a low and polarized emissivity and land has a high microwave emissivity across the microwave spectrum. Snow- and ice-covered surface microwave emissivities vary significantly depending on the snow and ice properties ([Takbiri et al. 2019](#); [Kummerow 2020](#)). Over the ocean, PMW algorithms can separate the surface radiation signal from the emission signal for liquid hydrometeors. Over land, microwave emissivity is sensitive to soil moisture, particularly at lower frequencies, which obscures the liquid hydrometeor emission signals. Hence, coastal areas present high emissivity gradients that challenge GPROF2017 precipitation retrieval. These retrievals are then intercalibrated to CORRA (the GPM combined radar-radiometer precipitation estimates) using probability matching. When low-Earth-orbiting PMW observations are too sparse, geostationary IR precipitation observations are merged within the algorithm using a Kalman-filter approach. The IR retrievals are provided by Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks–Cloud Classification System (PERSIANN-CCS; [Hong et al. 2007](#)). The CPC morphing–Kalman-filter (CMORPH-KF; [Joyce and Xie 2011](#)) quasi-Lagrangian time interpolation scheme uses the PMW and IR estimates to create half-hourly estimates. This system is run twice in real time for the “early” multisatellite product (~4 h after observation time), “late” multisatellite product (~14 h after observation time), and “final” satellite–gauge product (~3.5 months after the observation month). For the “final” product, half-hourly multisatellite estimates are adjusted to the monthly satellite–gauge combination. One of the major upgrades to IMERG V06 is the calculation of displacement vectors for morphing the PMW by using global total precipitable water vapor (TQV) estimates. Note that the minimum detectable rain rates from the Ka- and Ku-band radar are 0.2 and 0.5 mm h<sup>-1</sup>, respectively ([Hou et al. 2014](#)). GPROF is calibrated by these radar-based products, so the IMERG estimate could produce rain rates that are lower than these two thresholds. Hence, in the literature mostly 0.1 mm h<sup>-1</sup> is used for the minimum rain/no-rain detection threshold.

TABLE 1. Ancillary variables and definitions provided by IMERG.

Ancillary variable name	Definition
HQprecipitation	Merged microwave-only precipitation estimate
HQprecipSource	Microwave satellite source identifier
IRprecipitation	IR-only precipitation estimate
IRkalmanFilterweight	Weighting of IR-only precipitation relative to the morphed merged microwave-only precipitation
precipitationCal	Multisatellite precipitation estimate with gauge calibration
precipitationUncal	Multisatellite precipitation estimate

In this study we use IMERG Early, Late, and Final (IM-E, IM-L, and IM-F, precipitationCal fields are used) with the aim of understanding the performance of each run, specifically the morphing and the effectiveness of rain gauge correction over the land–coast–ocean continuum. To track uncertainties from Level-2 algorithms, IMERG-PMW (PMW) and IMERG-IR (IR) precipitation estimates are examined using ancillary variables provided by IMERG. The HQprecipitation and IRprecipitation fields are used for PMW and IR, respectively. The new PMW morphing procedure (morph) is targeted by selecting instances from precipitationUnCal field where there is no PMW retrieval (HQprecipSource reports zero) and no IR contributed estimate (IRkalmanFilterWeight of 0%). IRkalmanFilterWeight between 0% and 100% corresponds to a mixture of morphed PMW and IR estimates (morph+IR) from precipitationUnCal field. The ancillary variables and their definitions provided by IMERG can be seen in [Table 1](#).

#### e. Data matchup

IMERG matchups are extracted following a temporal and spatial matching with GV-MRMS that leaves the SPP temporal and spatial resolution untouched, so that it remains free of undesirable impacts caused by statistical or dynamical resampling ([Kirstetter et al. 2012, 2015](#)). Rainfall characteristics and their spatial distributions vary significantly across the United States. [Figure 2](#) maps the frequency of rainfall detected by IMERG-Final, -PMW, and -IR over the land–coast–ocean continuum. To highlight the differences between IMERG and GV-MRMS in detecting precipitation, only cases where rainfall is detected either by IMERG or GV-MRMS are displayed. GV-MRMS conditional relative rainfall occurrence (%) decreases farther away from the coast ([Fig. 2a](#)). GV-MRMS detects less rainfall because radar sampling conditions deteriorate with distance (e.g., with radar beamwidth and height). Optimal reference conditions are targeted in this study by selecting GV-MRMS grids within 100 km of each NEXRAD location [provided in [Fig. 2](#) as a red circle representing 100-km radius around each NEXRAD (black dots)]. It provides a consistent representation of the rainfall to focus on SPP differences across surface types. IMERG-IR rainfall occurrence is generally lower than 50% and less than

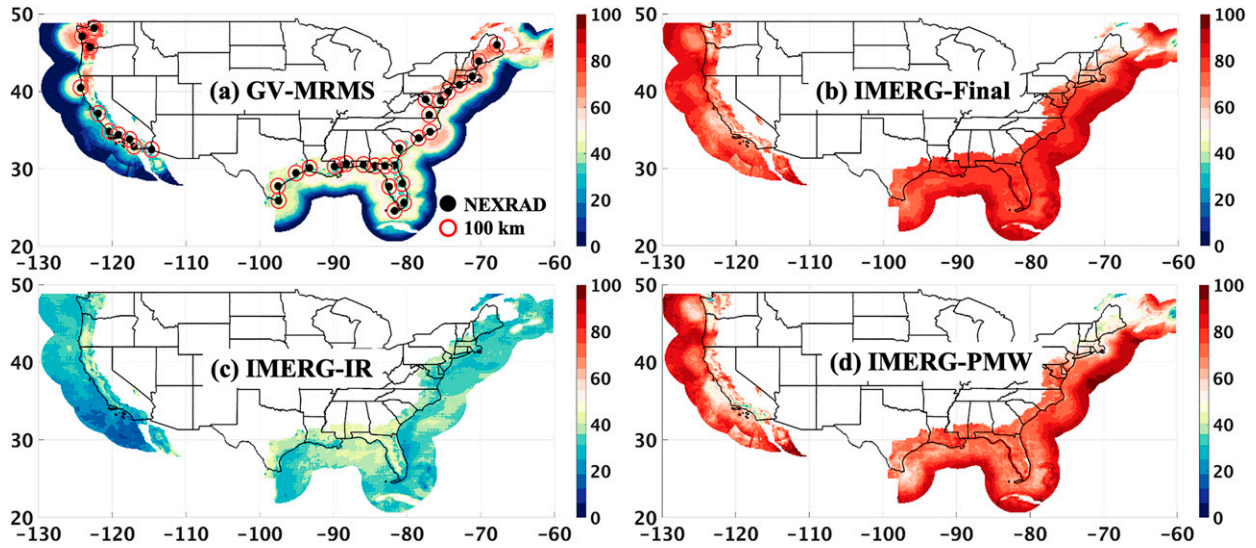


FIG. 2. Conditional relative rainfall occurrence of (a) GV-MRMS, (b) IMERG-Final, (c) IMERG-IR, and (d) IMERG-PMW. The locations of each NEXRAD radar are displayed (black filled circles) along with the corresponding 100-km radii (red open circles).

30% over the East Coast and the northwest coast (Fig. 2c). Meanwhile, rainfall occurrence is reported to be greater than 80% over the East Coast and northwest coast regions by GV-MRMS (Fig. 2a). This highlights rainfall detection challenges with the IR-based retrievals. IMERG-PMW reports higher rainfall occurrence than IMERG-IR, also higher over the ocean than coastal regions and land. IMERG-Final (Fig. 2b) displays the highest rainfall occurrence over all surface types.

As mentioned previously, the dataset is divided into three different surface types namely, land, coast, and ocean using the GPROFV05 surface classification dataset. The sample size of the matched dataset with and without quality control for each surface type is presented in Table 2.

### 3. Evaluation method

We investigate the performance of IMERG-Final, IMERG-PMW, IMERG-IR, IMERG-PMW morphing, and IMERG-PMW morph and IR mix using categorical and graphical measures. The evaluation is conducted at IMERG’s native spatial and temporal resolution of  $0.1^\circ$  and 30 min. As mentioned in the previous section, radar sampling conditions deteriorate with distance (e.g., radar beamwidth and height). To conduct a fair analysis, an optimum distance of 100 km (red circles in Fig. 2a) from the closest NEXRAD radar is chosen. GV-MRMS rain/no-rain threshold is set as  $0.1 \text{ mm h}^{-1}$ , and matchup datasets that have RQI of 100 are used. Moreover, GV-MRMS provides the precipitation type for each grid cell; those that are reported as snow are disregarded from this analysis.

The detection capabilities of IMERG are analyzed with categorical skill scores. IMERG and GV-MRMS either do or do not detect rainfall and the performance of IMERG is formulated in terms of hits, false alarms, misses and correct rejections with  $2 \times 2$  contingency tables. In the literature there has been significant work to identify and quantify the properties of binary verification. Gandin and Murphy (1992) introduced the

definition of equitability, which is the property of being awarded the same expected score for all random forecasting systems including those that always predict occurrence or nonoccurrence.

The conventional approach consists of calculating values for one or more summary measures and drawing conclusions about IMERG performance as the percentage of hits ( $H$ : both GV-MRMS and IMERG detect rain), false alarms ( $F$ : IMERG detects rain while GV-MRMS does not), misses ( $M$ : IMERG does not detect rain while GV-MRMS detects rain), and correct rejections ( $C$ : both GV-MRMS and IMERG do not detect rain). For a  $2 \times 2$  contingency table, the dimensionality is three where the fourth degree of freedom is fixed by the joint probabilities summing to unity. Therefore, a full description of IMERG detection performance requires only three parameters that can be expressed in terms of probability of detection (POD), probability of false detection (POFD), and base rate  $s$ . The formulas for each score are

$$\text{POD} = \frac{H}{H + M}, \tag{1}$$

$$\text{POFD} = \frac{F}{F + C}, \quad \text{and} \tag{2}$$

$$s = \frac{H + M}{H + M + F + C}. \tag{3}$$

The POD indicates what fraction of the observed events was correctly detected by IMERG; a perfect score is 1. The POFD is the fraction of IMERG events that actually did not occur; a perfect score is 0. The  $s$  indicates sample climatological relative frequency. It should be emphasized that these scores do not measure how well the IMERG rainfall magnitudes correspond to the reference but are only measures of relative frequencies. Note also that no performance score is meaningful when considered in isolation, and all these three scores should be used in combination to make inferences on IMERG detection performance. In our case this could be cumbersome when comparing multiple regions, datasets, and surface types, not to

TABLE 2. The sample size of GV-MRMS and IMERG for each surface type. Quality control refers to disregarding the grids that reported snow, RQI of less than 100, and grids that are farther away from the closest NEXRAD radar by more than 100 km.

Surface type	No. of data pairs without quality control	No. of data pairs after quality control
Land	24 691 850	9 239 994
Coast	24 823 923	9 853 882
Ocean	63 495 762	3 263 487

mention that it is difficult to systematically discuss the properties of particular performance scores. Also, POD and POFD are inequitable as performance measures since they degenerate for rare events; that is, they will score higher for random retrievals of rare events than for random retrievals of common events.

To obtain an equitable performance score, one should consider summing over all possible contingency table scores, each of which is multiplied by the probability of it occurring by chance (Jolliffe and Stephenson 2012). One example of such a score is the Heidke skill score (Heidke 1926) for multicategory verification. HSS is a truly equitable score that awards random and constant forecasts with an expected score of zero. It quantifies the accuracy of the estimate relative to that of random chance. HSS eliminates random forecasts that would be correct due to pure random chance and measures the fraction of correct forecasts. The HSS can be calculated as

$$\text{HSS} = \frac{2(HC - FM)}{F^2 + M^2 + 2HC + (F + M)(H + C)}. \quad (4)$$

It measures the skill of an algorithm at making a discrete classification and varies from  $-1$  to  $+1$ , with  $-1$  indicating perfect negative skill, zero indicating no skill relative to chance, and  $+1$  indicating perfect positive skill. An HSS value greater than 0.3 indicates significant skill with respect to random chance.

The HSS score has been used to characterize the performance of IMERG and its components in detecting rain from no rain with respect to a reference dataset using a predefined threshold. The detection performance can be evaluated by predicting the exceedance relative to a range of rainfall thresholds. Since higher rainfall rates are less frequent, larger exceedance thresholds naturally lead to fewer hits but also to fewer false alarms. As pointed out by Conner and Petty (1998), this sensitivity to thresholds is a challenge for an objective analysis. Hence, working with a single HSS value computed with an a priori threshold can be misleading. Instead, the HSS can be computed and plotted as a continuous bivariate function of two continuously varying rainfall rate thresholds based on GV-MRMS and IMERG. The maximum IMERG skill at distinguishing rain rates that are greater than any threshold applied to the GV-MRMS can be evaluated, and the maximum skill found across all GV-MRMS thresholds is independent of any bias (linear or nonlinear) in the IMERG retrieval (Chiu and Petty 2006). Hence, the performance of IMERG at discriminating between high and low rain rates can be inferred

independent of any biases. Moreover, any (conditional) bias can be diagnosed from the relationship between the GV-MRMS and IMERG threshold magnitudes at which HSS is maximized. In short, it is possible to evaluate and objectively compare the maximum skill of IMERG at distinguishing rain rates that are greater than any threshold applied to the GV-MRMS and IMERG. The two-dimensional HSS approach (2D-HSS) is particularly useful for determining the optimum delineation of nonzero rain rates over challenging conditions including surface types, regions, and rainfall climatologies (Petty and Li 2013).

The IMERG rain-rate threshold  $R_{\text{spp}}$  that optimizes the delineation of any given nonzero GV-MRMS rain rate  $R_{\text{gv}}$  can be determined by 2D-HSS analysis. The following summarizes the useful information 2D-HSS bivariate histogram provides (Chiu and Petty 2006; Petty and Li 2013; Wolff and Fisher 2009; Munchak and Skofronick-Jackson 2013):

- (i) The IMERG rain-rate threshold that maximizes the HSS as a function of any given GV-MRMS rain-rate threshold identifies the optimal balance between hits, misses, false alarms, and correct rejections and the corresponding performance with respect to random chance. It can be compared with the 1:1 line to infer IMERG biases. If the line of the maximum HSS is above or below the 1:1 line, then the estimate is biased high or low, respectively.
- (ii) Conversely, the GV-MRMS rain-rate threshold that maximizes HSS for any given IMERG rain-rate threshold (e.g., thick white lines in Fig. 4, below) can be inferred as the effective sensitivity of IMERG for this threshold.
- (iii) The maximum HSS curve can be plotted as a function of  $R_{\text{gv}}$  (i.e., HSS values at the location of the thick black line on Fig. 4, below). This plot provides maximum skill of IMERG at delineating GV-MRMS rain rates exceeding any given threshold. In particular, the maximum skill at delineating all nonzero GV-MRMS rain rates is determined at the minimum GV-MRMS threshold (e.g., the intersection of the curve with the y axis on Fig. 5, below).
- (iv) The IMERG threshold value  $R_{\text{spp}}$  for which the maximum skill is observed at delineating all GV-MRMS positive rain rates. For example, in Fig. 4 (below),  $R_{\text{spp}}$  corresponds to the intersection of a thick black line with the y axis.

## 4. Rainfall detection results

### a. General performance of IMERG products and components

Detection performance is evaluated with categorical skill scores: percentage of hits, misses, false and correct negatives are presented in Fig. 3. The results are broken down by regions (West Coast, Gulf of Mexico, and East Coast) and surface types (land, coast, and ocean surface types) and shown for retrievals IMERG Final (IM-F), IMERG Late (IM-L), IMERG Early (IM-E), IMERG PMW (PMW), IMERG IR (IR), IMERG PMW morph (morph), and morph and IR mix (morph+IR). The detection performance is discussed in terms of algorithms, regions, and surface types.

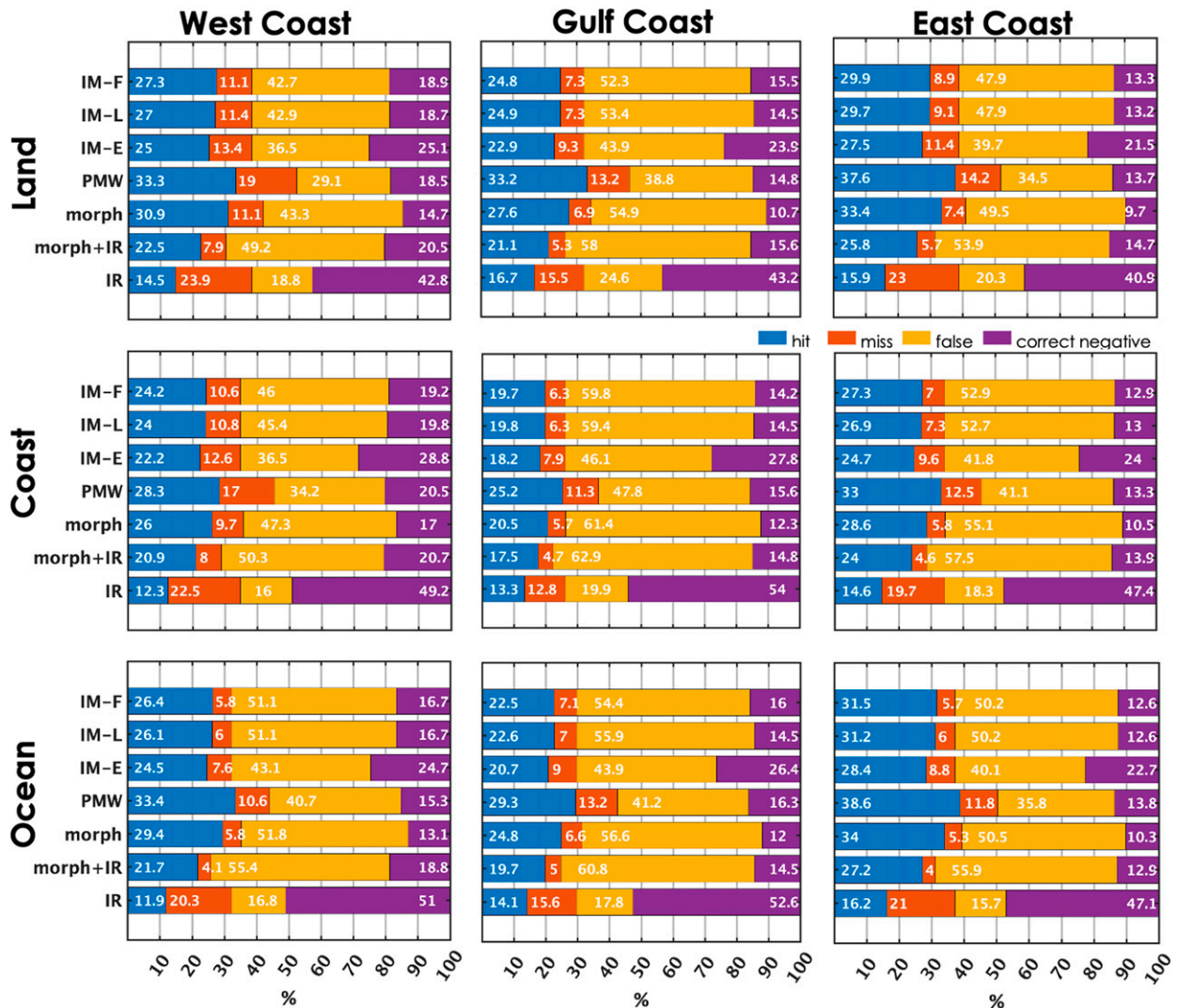


FIG. 3. Hits (blue), misses (orange), false alarms (yellow), and correct negatives (purple) for each SPP over the (left) West, (center) Gulf, and (right) East Coasts for the (top) land, (middle) coast, and (bottom) ocean surface types.

In Fig. 3, the dominant categorical statistic across the merged products IM-F, IM-L, and IM-E is the false alarm rate, which ranges from 36% to 59%. The next highest categorical statistic is hit rate ([18%–31%]), followed by correct negative rate ([12%–28%]). For most regions and surface types, the lowest categorical skill is the miss rate ([5%–13%]; Fig. 3). These indicate that the algorithms are biased toward retrieving rainfall that is not detected by GV-MRMS. The IM-F detection capability is on par with IM-L over all regions and surface types. It suggests that either the monthly gauge correction does not influence detection capabilities, or that gauges are not dense enough along coastal pixels to be used in IMERG and impact the detection capabilities of IM-F. Furthermore, the impact of morphing is more substantial, with IM-F and IM-L showing higher hit rates and lower misses relative to IM-E. Note also that there is an increase in false rates from IM-E to IM-L (+10% on average) at the expense of correct negatives.

The accuracy ( $H + C$ ) decreases from IM-E ([46%–51%]) to IM-L ([34%–45%]), and lowest accuracy is reported by IM-F ([33%–46%]). These results indicated that forward- and backward-propagated PMW estimates (morph) used in IM-L and IM-F show some benefit in comparison with the forward-propagated PMW estimates (morph) used in IM-E. It also suggests that the morphing tends to increase the occurrence of retrieved rainfall. The influence of PMW estimates on the merged products is further confirmed by the PMW categorical statistics, which display similar features as IM-E, IM-L, and IM-F, albeit with systematic differences, that is, higher hits (in the range [25%–39%]), higher misses ([11%–19%]), and lower false alarms ([29%–49%]; Fig. 3).

These performances justify the use of PMW as the backbone to build the merged products, where preference is given to PMW estimates with respect to IR estimates in the IMERG merging procedure. As expected, because of lack of direct

observations, the morph detection performance is lower relative to PMW, with lower hits ([20%–34%]) and significantly higher false alarm rates ([43%–61%], Fig. 3). The higher false alarm rates confirm the impact of morphing toward increasing the occurrence of estimated rainfall. Moreover, out of PMW, morph, morph+IR and IR products, PMW reports the highest hit rates followed by morph, then morph+IR, and IR reports the lowest hit rates. Misses decrease from PMW to morph to morph+IR while false alarm rates increase (especially from PMW to morph), again highlighting a tendency to overestimate rainfall occurrence with increasing time lag from PMW observations. IR displays the lowest detection performance when compared with all other products, with significantly lower hit rate (–9% on average), higher miss rate (+12% on average), and the highest correct negative ([40%–54%]) in comparison with all other products. This indicates that IR tends to miss rainfall occurrence, and these detection issues may be carried over to morph+IR (morph has higher detection performance relative to morph+IR) and eventually to the IM-F.

While the detection performances of each IMERG product are mostly consistent across regions, one can note some variations that highlight the impact of terrain and climatology. For almost all products, the Gulf Coast records lower hit rates than other regions, that is, on average 3.2% and 7% less relative to the West and East Coast regions, respectively (Fig. 3), with the exception of IR recording lower hits on the West Coast. Moreover, the Gulf Coast consistently records an average of 1.8% and 0.8% higher false rates than do the West and East Coast regions, respectively. Consistently, miss rates tend to be lower in the Gulf Coast over all surface types except ocean. Clearly this indicates that products detect more rainfall than what actually occurs over Gulf Coast (Fig. 3). The accuracy ( $H + C$ ) over these regions is very similar, but the Gulf Coast records on average 6% and 4% lower accuracy relative to the West and East Coasts, respectively.

Surface types represent another conditioning factor for rainfall detection by IMERG (Fig. 3). Across all regions and products there is a tendency for lower hits over coastal surfaces. The coastal transitions remain a challenge for rainfall estimation from space. Misses are slightly more prevalent over land, reflecting the challenges in distinguishing the rainfall signal from the radiometrically warm and variable land surface. Surprisingly, false alarms tend to be higher over the ocean than over land.

#### *b. IM-F delineation of various rainfall magnitudes*

In Fig. 4, the ability of IM-F in detecting rain rate is compared with GV-MRMS with 2D-HSS plots. 2D-HSS plots quantify IMERG accuracy relative to that of random chance and determine the optimum precipitation delineation. Figure 4 shows 2D-HSS plots over each surface type and region for IM-F. GV-MRMS rain/no-rain threshold values are indicated on the  $x$  axis and the IM-F rain/no-rain threshold values are on the  $y$  axis. The colored map represents the HSS values for different thresholds applied both to GV-MRMS and IM-F. The black line is the maximum HSS of IM-F conditioned on the GV-MRMS threshold and the white line is the maximum HSS of GV-MRMS conditioned on the IM-F threshold.

Common features reveal the consistent behavior of IM-F across regions and surfaces. The 2D-HSS is maximum around the 1:1 line (with exceptions) and takes on minimum values at extreme thresholds applied on GV-MRMS or IM-F (Fig. 4). This indicates an overall agreement in between GV-MRMS and IM-F in delineating similar rainfall rates. IM-F displays skill in delineating rainfall for certain rainfall rate ranges, with maximum HSS values above 0.3 and often reaching 0.5. IM-F delineates rainfall the best along the maximum HSS lines (black lines in Fig. 4). The maximum HSS values occur in the range [0.5–1]  $\text{mm h}^{-1}$  when the maximum HSS lines tend to be closer to the 1:1 line, reflecting rainfall rates at which IM-F overall delineates rainfall the best and shows minimum bias. At rainfall rates lower than 0.5  $\text{mm h}^{-1}$ , the IM-F maximum HSS lines are systematically above the 1:1 line, reflecting the overestimation of low rainfall rates. Moreover, HSS values (colored map in Fig. 4) decreases below 0.5  $\text{mm h}^{-1}$ , confirming a general challenge in correctly delineating light rainfall. Still, at the lowest GV-MRMS reference threshold of 0.1  $\text{mm h}^{-1}$ , the HSS takes on highest values [0.38–0.45] at the IM-F rainfall rate threshold [0.4–0.6  $\text{mm h}^{-1}$ ]. This indicates that IM-F shows the best skill in delineating rainfall magnitude of  $\sim 0.5 \text{ mm h}^{-1}$ . It should also be noted that IM-F shows moderate skill [0.25–0.45] in delineating rainfall rates [0.1–2  $\text{mm h}^{-1}$ ]. Conversely, HSS values decrease to 0 and below at high rainfall rate thresholds (GV-MRMS  $> 15 \text{ mm h}^{-1}$ ; Fig. 4) indicating that rainfall rate extremes are not well delineated. The IM-F maximum HSS curve also reveals a region-dependent slight overestimation and underestimation of rainfall precipitation rates above 5  $\text{mm h}^{-1}$ . The IM-F effective sensitivity line (white line in Fig. 4) corresponds to the GV-MRMS maximum HSS conditioned on IM-F threshold and shows which GV-MRMS features are best reproduced by IM-F. It indicates that IM-F effective sensitivity for rates above 0.1  $\text{mm h}^{-1}$  is around 0.3  $\text{mm h}^{-1}$ . The effective sensitivity remains below the 1:1 line for low-rate thresholds (e.g., below 0.5  $\text{mm h}^{-1}$ ), indicating that lower precipitation rates tend to be missed. Conversely, the effective sensitivity is above the 1:1 line at high-rate thresholds (e.g., above 0.5  $\text{mm h}^{-1}$ ), indicating a possible overestimation in IM-F rates. In general, effective sensitivity (white line) and corresponding rainfall thresholds of IM-F are lower than the maximum HSS curve (black line).

In relation to varying climatologies, regions display distinguishing features in the 2D-HSS plots (Fig. 4). The East Coast and the Gulf Coast have larger areas of positive HSS values than does the West Coast, which indicates that IM-F tends to delineate rainfall over a larger range of rainfall magnitudes (e.g., GV-MRMS threshold  $> 15 \text{ mm h}^{-1}$ ; Fig. 4) in these regions. The West Coast transitions to HSS values close or below zero for GV-MRMS rainfall rate thresholds above 10  $\text{mm h}^{-1}$ . A possible explanation is orographic precipitation enhancement in complex terrain close to the coastline that is missed by IM-F and negatively impacts its ability to delineate rainfall rates above this threshold. High precipitation rates can be associated with shallow warm-rain processes that generate lower ice content aloft (e.g., Carr et al. 2017; Porcaccchia et al. 2017). As ice content is a primary driver of precipitation detection for PMW retrievals, orographic events are a challenge. The detection issues of higher rainfall magnitudes in IMERG



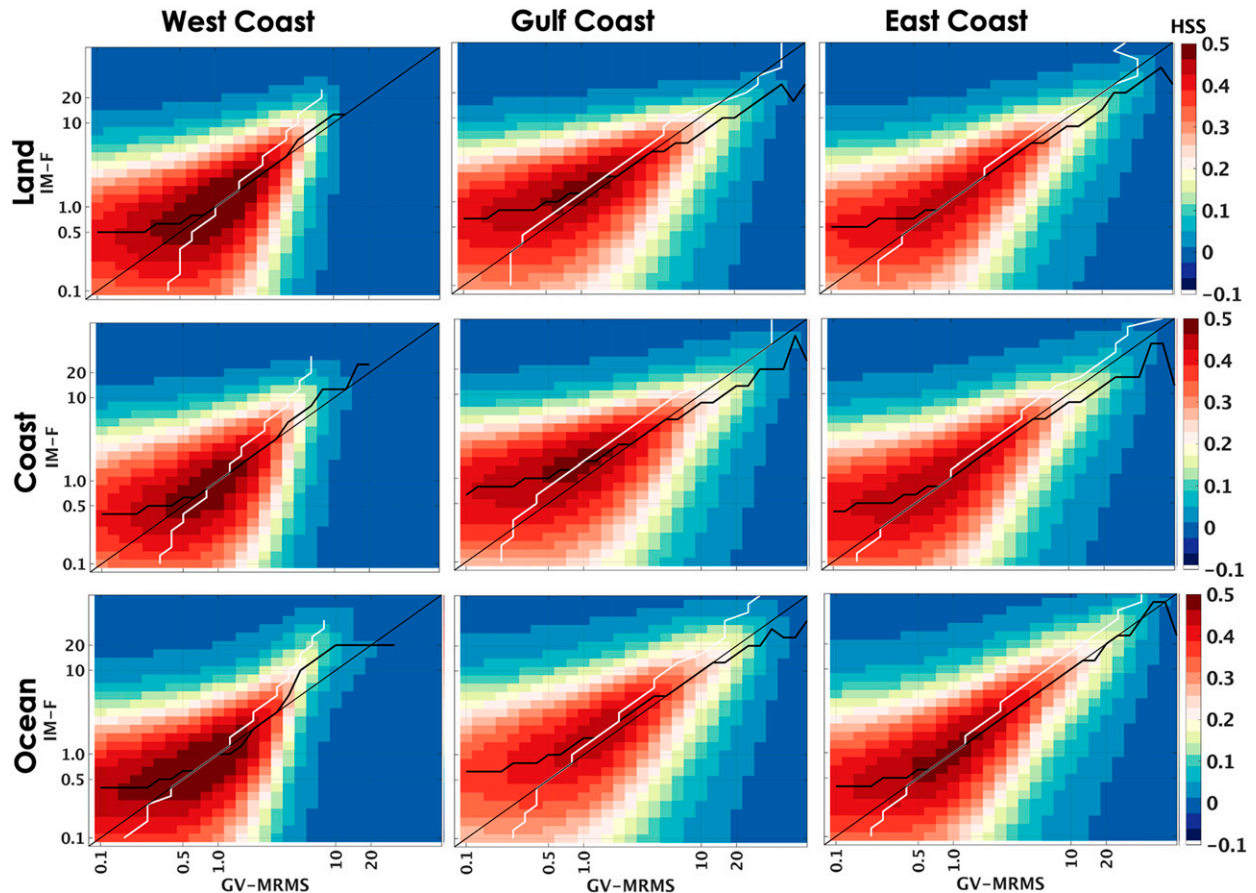


FIG. 4. 2D-HSS over each surface type and region for IM-F. The thick black line corresponds to the IM-F rain-rate threshold that maximizes the HSS as a function of GV-MRMS rain-rate threshold. The white line corresponds to the GV-MRMS rain-rate threshold that maximizes the HSS as a function of IM-F.

and PMW products are in agreement with the findings of past studies and require better understanding of the ice content in heavy precipitation events (Petković and Kummerow 2015; Derin et al. 2018). The West Coast region reports the highest HSS values across all surfaces in the range  $0.3\text{--}0.5\text{ mm h}^{-1}$ , indicating higher IM-F ability to delineate moderate rainfall in this region. A detailed performance analysis of PMW retrievals of specific West Coast precipitation events such as atmospheric rivers is warranted and will be considered in a future study. By contrast, the East Coast land and coast and the Gulf Coast ocean report lowest HSS values. Again, it could be explained by specific precipitation processes experienced in these regions that challenge the PMW retrievals inputs to IM-F. For example, precipitation in the Gulf Coast has more tropical characteristics than in the West Coast, that involve warm rain process. Over the East Coast, IM-F shows a lower conditional bias than in other regions, as the maximum HSS curves are closer to the 1:1 line across the range of rainfall rate thresholds. One significant difference between regions can be seen with the effective sensitivity of IM-F (white curve). Over the West Coast it displays a higher conditional deviation from the 1:1 line relative to the Gulf and East Coasts. The best IM-F effective sensitivity can be observed over the East Coast, followed by the Gulf Coast.

Surfaces are consistently driving IM-F performances. Over the ocean, IM-F displays better detection skills relative to other surfaces, for example, with improved delineation of light precipitation and high rates of HSS, with the exception of the Gulf Coast, where IM-F shows the lowest HSS values relative to all other regions and surface types. The effective sensitivity of IM-F tends to be higher over the ocean (i.e., closer to 1:1 line;  $[0.15\text{--}0.2]\text{ mm h}^{-1}$ ) than other surfaces ( $[0.2\text{--}0.4]\text{ mm h}^{-1}$ ), highlighting a better delineation of light precipitation ( $<0.3\text{ mm h}^{-1}$ ). This is expected since GPROF can take advantage of the low-frequency channels over ocean, which are more directly sensitive to light and shallow (predominately liquid) precipitation (Kidd et al. 2017). Over coast and land the detection skills tend to be similar. The GV-MRMS maximum HSS (white curve) has distinct differences especially for rainfall magnitudes  $0.1\text{--}0.4\text{ mm h}^{-1}$  over the West Coast land and coast surface types.

#### c. Maximum skill of IMERG products and components at delineating various GV-MRMS rainfall magnitudes

The gradients of HSS magnitudes in Fig. 4 provide subtle differences in terms of IM-F detection performance over different regions and surface types. All IMERG products' detection performance are reported in Fig. 5 with the maximum

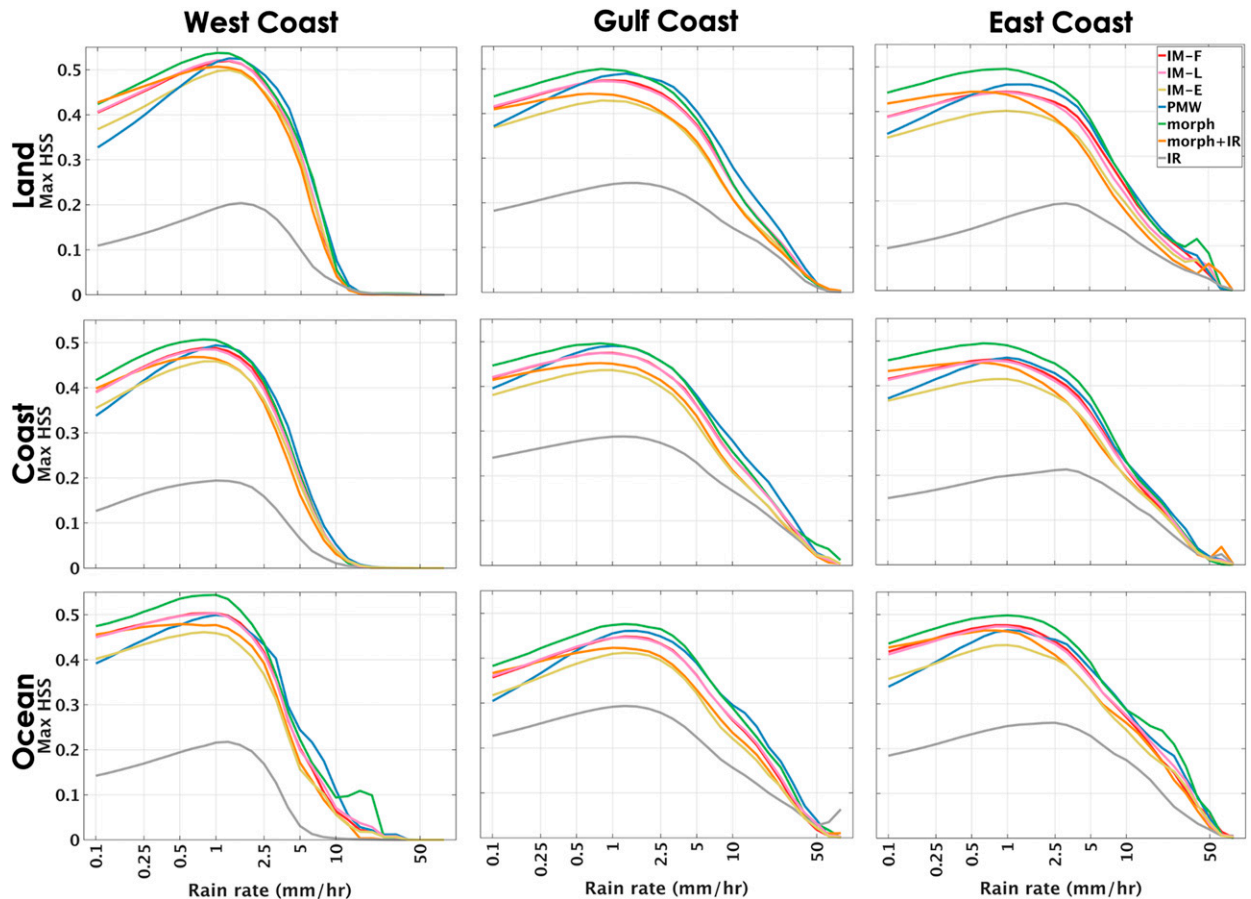


FIG. 5. Maximum skill of IMERG at discriminating GV-MRMS rain-rate thresholds over each region and surface type.

skill in delineating GV-MRMS rates above a given threshold (e.g., black curves in Fig. 4). It represents the intrinsic (bias independent) discrimination ability of the satellite algorithms with respect to GV-MRMS.

In general, all IMERG products (except IR) display significant skill ( $HSS > 0.43$ ) in discriminating rainfall occurrence with respect to GV-MRMS up to  $\sim 2 \text{ mm h}^{-1}$ , above which HSS decreases significantly especially with higher rainfall thresholds. Surprisingly morph has the highest discrimination ability with respect to GV-MRMS up to  $\sim 2 \text{ mm h}^{-1}$  threshold, then above this threshold PMW shows the highest discriminating skill.

IM-F and IM-L exhibit similar performances in delineating all nonzero ( $>0.1 \text{ mm h}^{-1}$ ) GV-MRMS rain rates (HSS in the range 0.36–0.45) meanwhile IM-E performance is lower (HSS in the range 0.3–0.4). As noted earlier, IM-L applies forward and backward morphing and IM-E applies forward morphing only, which can explain the significant detection skill difference between the two products. Morph exhibits the highest performance in delineating all nonzero GV-MRMS (HSS in the range 0.39–0.48). PMW skill in delineating GV-MRMS thresholds in the range  $[0.1\text{--}0.5] \text{ mm h}^{-1}$  is lower relative to morph and morph+IR products but is better than morph+IR above  $0.5 \text{ mm h}^{-1}$  and better than morph above  $2 \text{ mm h}^{-1}$ .

The detection performance of the PMW, morph, and morph+IR products should be taken into consideration when merging them into IM-F, IM-L, and IM-E. Note that both morph and PMW display better skill at delineating GV-MRMS rainfall at thresholds higher than  $1 \text{ mm h}^{-1}$  relative to IM-F, IM-L, and IM-E.

Over all regions, the products show varying skill at delineating GV-MRMS rain-rate thresholds, and specifically higher GV-MRMS rain-rate thresholds. Products have the highest HSS values over the West Coast up to approximately  $2 \text{ mm h}^{-1}$ . Above this value, HSS sharply decreases to 0 for thresholds higher than approximately  $10 \text{ mm h}^{-1}$ . Over the Gulf Coast and East Coast the HSS curves display smoother features and take on values above 0 over a wider range of GV-MRMS rainfall thresholds. HSS values over the Gulf and East Coast are similar except over the Gulf Coast ocean where they are lower as noted earlier.

As expected, HSS values tend to be higher for all GV-MRMS rain-rate thresholds over the ocean surface type when compared with other surface types (Gulf Coast ocean being an exception), as GPROF uses the full range of PMW channels to delineate precipitation. The PMW product displays significant differences in delineating GV-MRMS rain rates over the ocean, coast, and land surfaces especially over the West Coast. IMERG products have lower skill delineating GV-MRMS

TABLE 3. The  $R_{spp}$  (mm h<sup>-1</sup>) optimum SPP rain-rate threshold and corresponding HSS values for delineating all nonzero rain rates of GV-MRMS.

	Surface type	IM-F		IM-L		IM-E		PMW		Morph		Morph+IR		IR	
		$R_{spp}$	Skill	$R_{spp}$	Skill	$R_{spp}$	Skill	$R_{spp}$	Skill	$R_{spp}$	Skill	$R_{spp}$	Skill	$R_{spp}$	Skill
West Coast	Land	0.5	0.41	0.5	0.41	0.5	0.37	0.5	0.33	0.4	0.42	0.4	0.43	0.6	0.11
	Coast	0.4	0.39	0.4	0.39	0.4	0.35	0.4	0.34	0.4	0.41	0.4	0.39	0.1	0.13
	Ocean	0.4	0.45	0.4	0.45	0.4	0.4	0.5	0.39	0.4	0.47	0.4	0.45	0.13	0.14
Gulf Coast	Land	0.6	0.41	0.8	0.42	0.8	0.37	0.8	0.38	0.8	0.44	0.8	0.41	0.4	0.18
	Coast	0.6	0.42	0.8	0.42	1.0	0.38	0.8	0.39	0.8	0.45	0.8	0.41	0.2	0.24
	Ocean	0.6	0.36	0.6	0.36	0.6	0.32	0.6	0.3	0.5	0.38	0.6	0.37	0.1	0.23
East Coast	Land	0.5	0.39	0.5	0.39	0.5	0.34	0.6	0.35	0.5	0.44	0.5	0.42	0.3	0.09
	Coast	0.4	0.42	0.5	0.41	0.5	0.37	0.5	0.37	0.4	0.46	0.4	0.4	0.1	0.15
	Ocean	0.4	0.42	0.4	0.41	0.4	0.36	0.4	0.34	0.4	0.44	0.4	0.43	0.1	0.18

rain-rate thresholds higher than ~10 mm h<sup>-1</sup>. This lack of skill is exacerbated over the land and coast surface type over the West Coast. Again, this indicates the direct effect of complex terrain proximity on the performance of IMERG products.

As stated in section 3, another set of information we can discern from the 2D-HSS plots is the rain-rate threshold at which the SPP estimates maximize HSS ( $R_{spp}$ ); that is,  $R_{spp}$  is the rainfall magnitude at which SPP estimates show the best agreement in delineating all nonzero GV-MRMS rain rates (>0.1 mm h<sup>-1</sup>). This value is determined from the intersection of the black thick line in Fig. 4 with the y axis. The  $R_{spp}$  values for each SPP, region and surface type are provided in Table 3.

In general, all products (except IR) display significant skill in identifying rain versus no rain with respect to random chance, i.e., HSS values are higher than 0.3. Overall IM-F and IM-L HSS values are similar to each other ([0.36–0.45]), but the corresponding  $R_{spp}$  values differ slightly with IM-F  $R_{spp}$  being in the range [0.4–0.6] mm h<sup>-1</sup> and IM-L  $R_{spp}$  being in the range [0.4–0.8] mm h<sup>-1</sup>. On the other hand, IM-E has lower HSS ([0.32–0.4]) and higher corresponding  $R_{spp}$  ([0.4–1.0 mm h<sup>-1</sup>]) magnitudes. One interesting result is that morph ( $R_{spp}$  [0.4–0.8 mm h<sup>-1</sup>] and HSS [0.38–0.47]) has higher HSS in discriminating between zero and nonzero GV-MRMS rain rates relative to PMW ( $R_{spp}$  [0.4–0.8 mm h<sup>-1</sup>] and HSS [0.30–0.39]) with similar  $R_{spp}$  values. In comparison with morph, morph+IR displays similar  $R_{spp}$  ([0.4–0.8 mm h<sup>-1</sup>]) but with lower HSS ([0.37–0.45]), and IR has the lowest skill ([0.11–0.24]). Overall, these results show that a 0.1 mm h<sup>-1</sup> rain/no-rain threshold (commonly used threshold in the literature) is not the optimum rain/no-rain threshold for IMERG and each subproduct has different threshold and performance in discriminating between GV-MRMS rain and no rain. These differences have implications in estimating the spatial extent of precipitation fields and should be considered accordingly.

Over the regions SPPs have different performances. Over the West Coast and East Coast similar  $R_{spp}$  ([0.1–0.5 mm h<sup>-1</sup>]) and HSS ([0.09–0.47]) values are reported. Over the Gulf Coast,  $R_{spp}$  ([0.1–1.0 mm h<sup>-1</sup>]) values are significantly higher than over other regions with comparable HSS values ([0.23–0.45]), suggesting that SPPs products tend to miss light precipitation in this area. As expected, the ocean surface type

shows the highest HSS values ([0.13–0.47]) with lowest  $R_{spp}$  ([0.1–0.6 mm h<sup>-1</sup>]) magnitudes. It indicates that SPPs show skill in capturing the rain/no-rain limit at lower rates over ocean than over other surfaces. Meanwhile land and coast surface types have similar HSS ([0.11–0.44]) and ([0.13–0.46]), respectively), whereas  $R_{spp}$  has higher range over the coast ([0.1–1.0 mm h<sup>-1</sup>]) than over land ([0.3–0.8 mm h<sup>-1</sup>]) surface types.

d. Precipitation typology

As typology is a key characteristic of precipitation (Kirstetter et al. 2020), IMERG detection performance is investigated as a function of precipitation types, that is, GV-MRMS stratiform and convective rainfall. In this section, a SPP–reference matched pair is included in the statistics if the reference is nonzero (>0.1 mm h<sup>-1</sup>). Any pixel with positive convective volume contribution is considered as a convective case. Figure 6 provides hits and misses for each product over each surface type and region for stratiform (bold colors) and convective (pale colors) precipitation types.

Overall, the various SPPs display similar rainfall detection performances for both convective and stratiform rainfall, with hits generally above 70%. IR stands out with significantly lower performance than other estimates (around 50% detection hits), and a markedly higher (+11% on average) detection of convective rainfall occurrence than stratiform rainfall occurrence. Lower IR detection capability of stratiform rainfall has been reported in the literature (e.g., Stampoulis et al. 2013; Kirstetter et al. 2020). Convective rainfall is associated with higher/colder cloud tops that are more easily detectable in IR observations than stratiform rainfall. Morph and morph+IR tend to have higher hit rates than other products for both rainfall types over all regions and surface types. Apart from IR, PMW estimates tend to have lower detection scores relative to other estimates, which highlights the challenges in detecting rainfall from Level-2 observations. PMW relative performances in detecting precipitation type occurrence also vary across surface types. This lower detection ability seems to affect IM-E as it shows similar performance as PMW across all regions and surfaces. The morphing has a slight but positive impact on precipitation detection and benefits IM-L and IM-F.

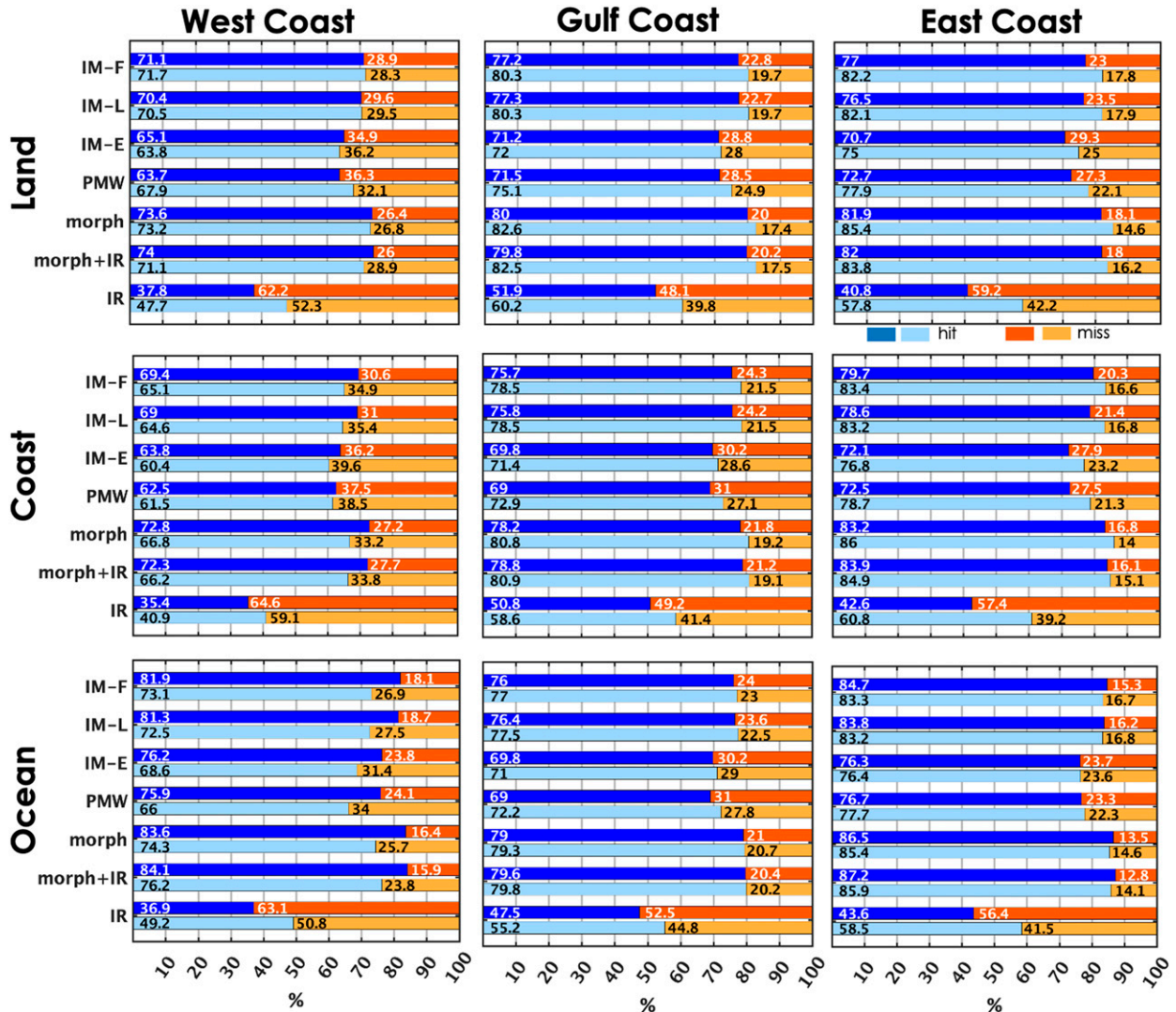


FIG. 6. Hits (blue) and misses (orange) for each SPP over the (left) West, (center) Gulf, and (right) East Coasts for the (top) land, (middle) coast, and (bottom) ocean surface types for stratiform (bold color hues) and convective (pale color hues) precipitation type. The GV-MRMS rain/no-rain threshold is set as  $0.1 \text{ mm h}^{-1}$ .

Detection of stratiform and convective rainfall occurrence varies significantly across regions, which highlights the impact of precipitation regimes. Detection of precipitation types varies also slightly depending on surfaces. Over the land and coastal surfaces of the East Coast, all products detect convective rainfall occurrence slightly better than stratiform rainfall occurrence. It tends to be the other way around on the West Coast, where slightly higher hit rates are associated with stratiform rainfall. For example, the West Coast IM-F, IM-L, and IM-E stratiform hit rates are slightly higher (~6%) over coastal and ocean surface types relative to the convective hit rate. Again, it probably reflects the difference in precipitation generation mechanisms across both regions, and the challenges in detecting orographic precipitation from space on the West Coast. Over the ocean, most products display similar or slightly better detection of stratiform rainfall (except IR). On the East

Coast and West Coast, the detection of stratiform rainfall is higher over ocean than over land and coastal surfaces (+11% hits with IM-F on the West Coast), as expected. The detection of convective rainfall occurrence follows the same trend, albeit with less difference across surfaces. It is likely that the transition of surfaces and environments (e.g., surface emissivity gradients) impacts more stratiform than convective satellite estimates. On the Gulf Coast, the detection performance of both convective and stratiform precipitation remains about the same across surfaces for both precipitation types.

Figure 7 provides IMERG products maximum skill in delineating stratiform and convective GV-MRMS rates above a given threshold (e.g., black curves in Fig. 4). Because IM-F, IM-L, and IM-E performances are very similar, only IM-F is reported. In general, IMERG products and their components display significant skill in delineating stratiform and convective

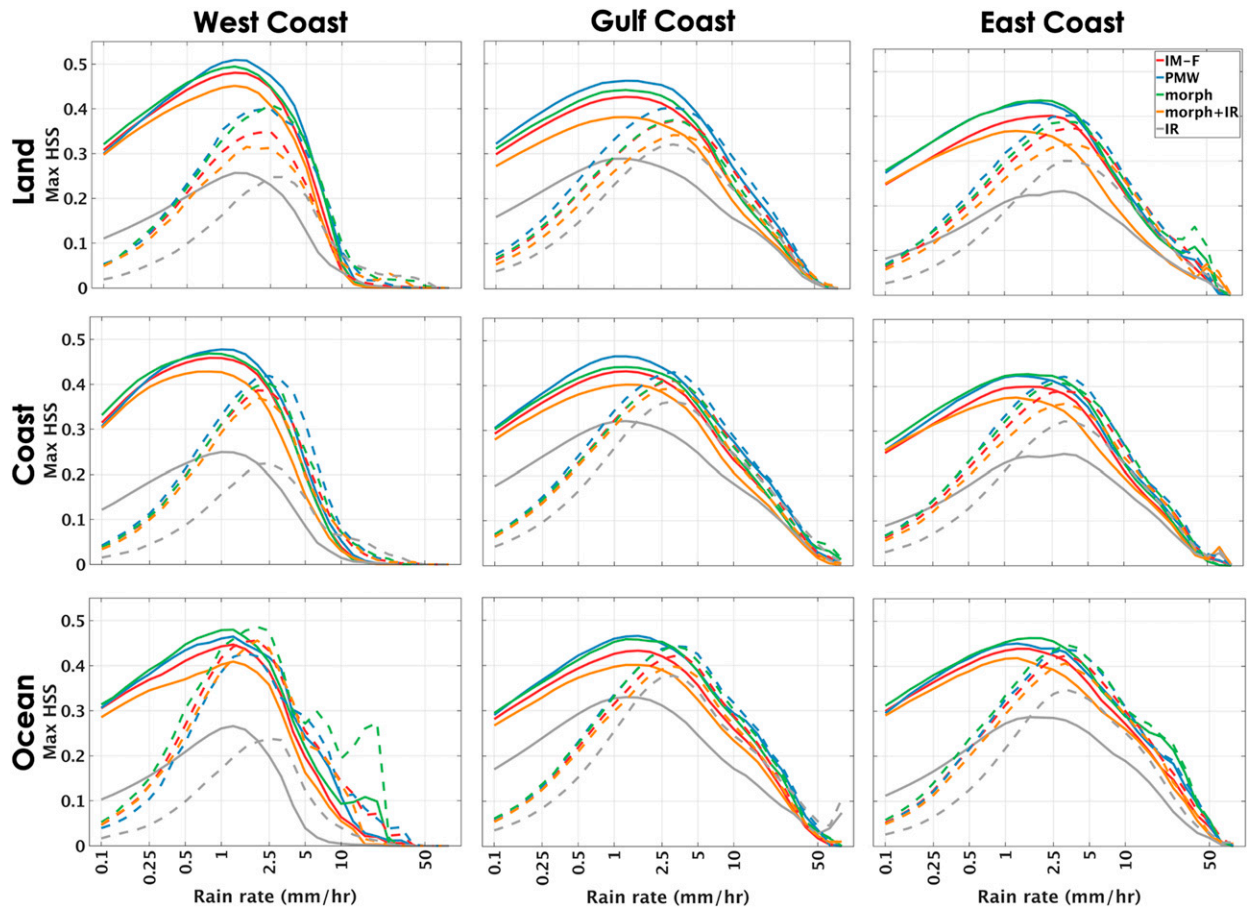


FIG. 7. Maximum skill at discriminating SPP stratiform (solid lines) and convective (dashed lines) rain rates exceeding the indicated GV-MRMS threshold values over each region and surface type.

rainfall occurrence as all curves (except IR's) display maximum HSS values greater than 0.3. However, the ability to delineate rainfall displays a strong dependence with rainfall magnitude and type. Rainfall occurrence delineation skills are higher with stratiform rainfall at low to moderate rates ( $<5 \text{ mm h}^{-1}$ ) and higher with convective rainfall at high rates ( $>5 \text{ mm h}^{-1}$ ). At threshold  $0.1 \text{ mm h}^{-1}$ , all products except IR show significant skill with stratiform rainfall (HSS value around 0.3) while the delineation of convective rainfall occurrence shows performance close to random chance. The stratiform curves reach their maximum around  $1.5 \text{ mm h}^{-1}$  with HSS in the range [0.35–0.52], except IR that shows significantly lower skills (HSS in the range [0.23–0.33]). The convective curves consistently reach their maximum at higher rainfall rates around  $2.5 \text{ mm h}^{-1}$  with generally lower HSS ([0.3–0.52], except IR that again shows significantly lower skill (HSS in the range [0.22–0.35])). Often, the IR maximum convective HSS is higher than the maximum stratiform HSS, illustrating the dependence of the IR retrieval to cloud top information. The differing detection performances of the PMW, morph and morph+IR products for different precipitation types provide us with further insights on where the detection challenges lie. First, precipitation typology is a strong

characteristic that is currently not accounted for in IMERG while these results call for ingesting the GPROF precipitation type information into IMERG. Second, the poor delineation of convective precipitation at low-rate thresholds may reveal the impact of subpixel rainfall nonuniformity on IMERG.

Similar to findings in section 4a (Fig. 4), morph and PMW have the highest discrimination ability for both stratiform and convective rainfall and for all rainfall thresholds, consistently followed by morph+IR and finally IR that shows the lowest skill. The skills of the combined IM-F lay between morph+IR and morph. PMW exhibits the highest performance in delineating all nonzero stratiform rainfall (HSS in the range 0.3–0.35 for GV-MRMS  $> 0.1 \text{ mm h}^{-1}$ ).

The three regions are characterized by different precipitation regimes and generation mechanisms. Accordingly, all products show regionally varying skill at delineating stratiform and convective GV-MRMS precipitation. Again, West Coast stands out with the highest detection performance difference between stratiform (higher) and convective (lower) rainfall types over land and coastal surfaces. West Coast overland detection of convective rainfall is the lowest across regions and surfaces. It confirms that the challenges in this region are associated with orographic generation of convective precipitation.

Detection skills are generally more uniform across stratiform than convective precipitation over the ocean.

## 5. Conclusions

The ability to detect rainfall over the land–coast–ocean continuum is examined over three coastal regions of the United States, that is, the West Coast, Gulf of Mexico, and East Coast, each of which are characterized by different topographies and precipitation climatologies. Detection capabilities are contrasted over different surfaces (land, coast, and ocean) to understand the performances over the land–coast–ocean transition. Moreover, the analysis is performed by examining the detection of stratiform and convective rainfall occurrence. An integrated and novel approach is developed to trace the detection performance of IMERG V06B (IM-F, IM-L, IM-E) back to their components (PMW, morph, morph+IR, and IR) and sources of rainfall estimates. It is performed by using high-resolution GV-MRMS rainfall estimates as ground reference at the IMERG native spatial and temporal resolution (0.1° and 30 min).

The main results for precipitation detection characteristics are summarized as follows:

- Decreasing detection performances are consistently noted from retrievals that rely on observations more indirectly related to surface precipitation. The best performances (in terms of hit rate) are reported with PMW estimates followed by morph, morph+IR, and finally IR estimates. IR displays the lowest detection performance when compared with all other products, with significantly lower hit rate (−9% on average), higher miss rate (+12% on average), and the highest correct negative ([40%–54%]) relative to all other products. Overall, IMERG displays nonhomogeneous precipitation detection properties that vary according to which component is used, and that translates into nonhomogeneous accuracy in space and time.
- The positive impact of the morphing is substantial on precipitation detection. IM-F and IM-L (with forward- and backward-propagated PMW estimates) show higher hit rates and lower misses than IM-E (forward-propagated only PMW estimates). Surprisingly, morph displays the highest discrimination ability with respect to GV-MRMS up to  $\sim 2 \text{ mm h}^{-1}$  threshold and above this threshold PMW shows the highest discrimination skill. Moreover, morph and PMW display better skill at delineating the reference rainfall at thresholds higher than  $1 \text{ mm h}^{-1}$  than do IM-F, IM-L, and IM-E.
- IM-F detection capability is on par with IM-L over all regions and surface types, which suggests that the monthly gauge correction does not influence the detection capabilities.

Varying rain/no-rain detection thresholds for IMERG and each subproduct have implications for estimating the spatial extend of precipitation fields. It is recommended to account for the detection performance of the PMW, morph and morph+IR retrievals when merging them into IM-F, IM-L, and IM-E.

Regional differences in terms of detection performance have significant implications for the generalization of the results to other coastal regions of the world:

- The detection performances of each IMERG product or its component are mostly consistent across regions, yet one can note some variations that highlight the impact of terrain and climatology.
- The Gulf Coast has lower hit rates (3.2%, i.e., 7% average less relative to other regions) and higher false rates (1.8%, i.e., 0.8% average higher).
- From the 2DHSS analysis, IMERG tends to better delineate rainfall over a larger range of rainfall magnitudes in the East and Gulf Coast regions than in the West Coast. Orographic precipitation enhancement associated with the proximity of complex terrain to the coastline challenges IMERG ability to detect the occurrence of a large range of rainfall magnitudes, particularly at the high end.

Surface types have a significant influence that is attributed to impact retrievals through emissivity gradients and contrasts in land/ocean precipitation characteristics:

- Coastal transitions remain a challenge for rainfall estimation. Misses are slightly more prevalent over land, reflecting the challenges in distinguishing the rainfall signal from the radiometrically warm and variable land surface. Moreover, across all regions and products there is a tendency for lower hit rates and higher false alarms over coastal surfaces.
- Over the ocean, IMERG displays better detection skill with improved delineation of light precipitation and high rates relative to other surfaces. This is expected since GPROF takes advantage of more brightness temperature channels over the ocean.
- Over the ocean, IMERG detection generally suffers more from false alarms in proportions ranging from 10% to 53% depending on the satellite product or component.

These results point to similar behavior across the coastal transitions regions of the United States, which suggests that the results of this study can be generalized to other coastal regions of the world to some extent. For example, terrain and climatological patterns similar to the West Coast can be found in the Andes, Spain, and so on. Regional differences (especially the West Coast) probably result from precipitation processes and generation mechanisms that are not captured by IMERG. This is highlighted with the impact of precipitation typology:

- Overall IMERG products and components display similar rainfall detection performances for both convective and stratiform rainfall, with hits generally above 70%. IR stands out with significantly lower performance ( $\sim 50\%$  hits), and a markedly higher detection of convective rainfall occurrence (+11% on average) than stratiform rainfall occurrence. Convective rainfall is associated with higher/colder cloud tops that are more easily detectable in IR observations than stratiform rainfall.
- Detection of stratiform and convective rainfall occurrence varies significantly across regions, which highlights the impact of precipitation regimes.
- The ability to delineate rainfall displays a strong dependence with rainfall magnitude and type. Rainfall occurrence delineation skills are higher with stratiform rainfall at low to moderate rates ( $< 5 \text{ mm h}^{-1}$ ) and higher with convective

rainfall at high rates ( $>5 \text{ mm h}^{-1}$ ). Poor delineation of convective precipitation at low rate thresholds may reveal the impact of subpixel rainfall nonuniformity within IMERG pixels.

- Morph and PMW have the highest discrimination ability for both stratiform and convective rainfall and for all rainfall thresholds, consistently followed by morph+IR and finally IR showing the lowest skill. These varying detection performances provide us with further insight on where the detection challenges lie.

Because precipitation typology is a strong characteristic that is currently not accounted for in IMERG, these results call for ingesting the GPROF precipitation type information into IMERG. It would potentially increase IMERG accuracy and its consistency across regions. A complementary study (Part II) will focus on the quantification performance of IMERG V06B as a second step of the evaluation over the land–coast–ocean continuum.

*Acknowledgments.* The authors acknowledge the efforts made by the NASA science team for making IMERG precipitation data accessible. Support from the NASA Global Precipitation Measurement Ground Validation program under Grant NNX16AL23G and the Precipitation Measurement Missions program under Grant 80NSSC19K0681 is acknowledged. The authors are thankful for suggestions made by Drs. David Wolff and Joseph Munchak. The authors thank three anonymous reviewers whose comments helped to improve the paper.

*Data availability statement.* The GV-MRMS precipitation data are available at the NASA Global Hydrology Resource Center (GHRC) (<http://dx.doi.org/10.5067/GPMGV/MRMS/DATA101>). The IMERG precipitation products are available online (<https://arthurhouhttps.pps.eosdis.nasa.gov>).

## REFERENCES

- Anagnostou, E. N., W. F. Krajewski, and J. Smith, 1999: Uncertainty quantification of mean-areal radar-rainfall estimates. *J. Atmos. Oceanic Technol.*, **16**, 206–215, [https://doi.org/10.1175/1520-0426\(1999\)016<0206:UQOMAR>2.0.CO;2](https://doi.org/10.1175/1520-0426(1999)016<0206:UQOMAR>2.0.CO;2).
- , V. Maggioni, E. Nikolopoulos, T. Meskele, F. Hossain, and A. Papadopoulos, 2010: Benchmarking high-resolution global satellite rainfall products to radar and rain-gauge rainfall estimates. *IEEE Trans. Geosci. Remote Sens.*, **48**, 1667–1683, <https://doi.org/10.1109/TGRS.2009.2034736>.
- Ayat, H., J. P. Evans, and A. Behrangi, 2021: How do different sensors impact IMERG precipitation estimates during hurricane days? *Remote Sens. Environ.*, **259**, 112417, <https://doi.org/10.1016/j.rse.2021.112417>.
- Bowman, K. P., A. B. Phillips, and G. R. North, 2003: Comparison of TRMM rainfall retrievals with rain gauge data from the TAO/TRITON buoy array. *Geophys. Res. Lett.*, **30**, 1757, <https://doi.org/10.1029/2003GL017552>.
- Carr, N., and Coauthors, 2015: The influence of surface and precipitation characteristics on TRMM microwave imager rainfall retrieval uncertainty. *J. Hydrometeorol.*, **16**, 1596–1614, <https://doi.org/10.1175/JHM-D-14-0194.1>.
- , P. E. Kirstetter, J. J. Gourley, and Y. Hong, 2017: Polarimetric signatures of midlatitude warm-rain precipitation events. *J. Appl. Meteor. Climatol.*, **56**, 697–711, <https://doi.org/10.1175/JAMC-D-16-0164.1>.
- Chiu, J. C., and G. W. Petty, 2006: Bayesian retrieval of complete posterior PDFs of oceanic rain rate from microwave observations. *J. Appl. Meteor. Climatol.*, **45**, 1073–1095, <https://doi.org/10.1175/JAM2392.1>.
- Conner, M. D., and G. W. Petty, 1998: Validation and intercomparison of SSM/I rain-rate retrieval methods over the continental United States. *J. Hydrometeorol.*, **37**, 679–700, [https://doi.org/10.1175/1520-0450\(1998\)037<0679:VAIOSI>2.0.CO;2](https://doi.org/10.1175/1520-0450(1998)037<0679:VAIOSI>2.0.CO;2).
- Derin, Y., and K. K. Yilmaz, 2014: Evaluation of multiple satellite-based precipitation products over complex topography. *J. Hydrometeorol.*, **15**, 1498–1516, <https://doi.org/10.1175/JHM-D-13-0191.1>.
- , and Coauthors, 2016: Multi-regional satellite precipitation products evaluation over complex terrain. *J. Hydrometeorol.*, **17**, 1817–1836, <https://doi.org/10.1175/JHM-D-15-0197.1>.
- , E. Anagnostou, M. N. Anagnostou, J. Kalogiros, D. Casella, A. C. Marra, G. Panegrossi, and P. Sano, 2018: Passive microwave rainfall error analysis using high-resolution X-band dual-polarization radar observations in complex terrain. *IEEE Trans. Geosci. Remote Sens.*, **56**, 2565–2586, <https://doi.org/10.1109/TGRS.2017.2763622>.
- , and Coauthors, 2019: Evaluation of GPM-era global satellite precipitation products over multiple complex terrain regions. *Remote Sens.*, **11**, 2936, <https://doi.org/10.3390/rs11242936>.
- Ebert, E. E., J. E. Janowiak, and C. Kidd, 2007: Comparison of near-real-time precipitation estimates from satellite observations and numerical models. *Bull. Amer. Meteor. Soc.*, **88**, 47–64, <https://doi.org/10.1175/BAMS-88-1-47>.
- Gandin, L. S., and A. H. Murphy, 1992: Equitable scores for categorical forecasts. *Mon. Wea. Rev.*, **120**, 361–370, [https://doi.org/10.1175/1520-0493\(1992\)120<0361:ESSFCF>2.0.CO;2](https://doi.org/10.1175/1520-0493(1992)120<0361:ESSFCF>2.0.CO;2).
- Gebregiorgis, A., P. E. Kirstetter, Y. Hong, N. Carr, J. J. Gourley, and Y. Zheng, 2017: Understanding overland multisensor satellite precipitation error in TRMM-RT products. *J. Hydrometeorol.*, **18**, 285–306, <https://doi.org/10.1175/JHM-D-15-0207.1>.
- , —, —, J. J. Gourley, G. J. Huffman, W. A. Petersen, X. Xue, and M. R. Schwaller, 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**, 1694–1707, <https://doi.org/10.1002/2017JD027606>.
- Heidke, P., 1926: Berechnung der Erfolge und der Gute der Windstarkevorhersagen im Sturmwarnungsdienst (Calculation of the success and goodness of wind strength forecasts in the Storm Warning Service). *Geogr. Ann.*, **8**, 301–349, <https://doi.org/10.1080/20014422.1926.11881138>.
- Hong, Y., D. Gochis, J. Cheng, K. Hsu, and S. Sorooshian, 2007: Evaluation of PERSIANN-CCS rainfall measurement using the NAME event rain gauge network. *J. Hydrometeorol.*, **8**, 469–482, <https://doi.org/10.1175/JHM574.1>.
- Hou, A. Y., and Coauthors, 2014: The global precipitation measurement mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, <https://doi.org/10.1175/BAMS-D-13-00164.1>.
- Huffman, G. J., and Coauthors, 2019: NASA Global Precipitation Measurement (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG). NASA Algorithm Theoretical Basis Doc., version 6, 34 pp., [https://gpm.nasa.gov/sites/default/files/document\\_files/IMERG\\_ATBD\\_V06.pdf](https://gpm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf).
- Jolliffe, I. T., and D. B. Stephenson, 2012: *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. 2nd ed. John Wiley and Sons, 296 pp.

- Joyce, R. J., and P. Xie, 2011: Kalman filter based CMORPH. *J. Hydrometeorol.*, **12**, 1547–1563, <https://doi.org/10.1175/JHM-D-11-022.1>.
- Khan, S., and V. Maggioni, 2019: Assessment of level-3 gridded Global Precipitation Mission (GPM) products over oceans. *Remote Sens.*, **11**, 255, <https://doi.org/10.3390/rs11030255>.
- Kidd, C., J. Tan, P.-E. Kirstetter, and W. A. Petersen, 2017: Validation of the version 05 level 2 precipitation products from the GPM Core Observatory and constellation satellite sensors. *Quart. J. Roy. Meteor. Soc.*, **144**, 313–328, <https://doi.org/10.1002/qj.3175>.
- Kirstetter, P. E., and Coauthors, 2012: Toward a framework for systematic error modeling of spaceborne precipitation radar with NOAA/NSSL ground radar-based national mosaic QPE. *J. Hydrometeorol.*, **13**, 1285–1300, <https://doi.org/10.1175/JHM-D-11-0139.1>.
- , Y. Hong, J. J. Gourley, Q. Cao, M. Schwaller, and W. Petersen, 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, Geophys. Monogr.*, Vol. 206, Amer. Geophys. Union, 61–79, <https://doi.org/10.1002/9781118872086.ch4>.
- , —, —, M. Schwaller, W. Petersen, and Q. Cao, 2015: Impact of sub-pixel rainfall variability on spaceborne precipitation estimation: Evaluating the TRMM 2A25 product. *Quart. J. Roy. Meteor. Soc.*, **141**, 953–966, <https://doi.org/10.1002/qj.2416>.
- , N. Karbalaee, K. Hsu, and Y. Hong, 2018: Probabilistic precipitation rate estimates with space-based infrared sensors. *Quart. J. Roy. Meteor. Soc.*, **144**, 191–205, <https://doi.org/10.1002/qj.3243>.
- , W. A. Petersen, C. D. Kummerow, and D. B. Wolff, 2020: Integrated multi-satellite evaluation for the global precipitation measurement: Impact of precipitation types on spaceborne precipitation estimation. *Satellite Precipitation Measurement*, V. Levizzani et al., Eds., Advances in Global Change Research, Vol. 69, Springer, 583–608.
- Kummerow, C. D., 2020: Introduction to passive microwave retrieval methods. *Satellite Precipitation Measurement*, V. Levizzani et al., Eds., Advances in Global Change Research, Vol. 67, Springer, 123–140.
- Maggioni, V., P. C. Meyers, and M. D. Robinson, 2016: A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era. *J. Hydrometeorol.*, **17**, 1101–1117, <https://doi.org/10.1175/JHM-D-15-0190.1>.
- , E. I. Nikolopoulos, E. N. Anagnostou, and M. Borga, 2017: Modeling satellite precipitation errors over mountainous terrain: The influence of gauge density, seasonality, and temporal resolution. *IEEE Trans. Geosci. Remote Sens.*, **55**, 4130–4140, <https://doi.org/10.1109/TGRS.2017.2688998>.
- Manz, B., S. Paez-Bimos, N. Horna, W. Buytaert, B. Ochoa-Tocachi, W. Lavado-Casimiro, and B. Willems, 2017: Comparative ground validation of IMERG and TMPA at variable spatio-temporal scales in the tropical Andes. *J. Hydrometeorol.*, **18**, 2469–2489, <https://doi.org/10.1175/JHM-D-16-0277.1>.
- McCollum, J. R., W. F. Krajewski, R. R. Ferraro, and M. B. Ba, 2002: Evaluation of biases of satellite rainfall estimation algorithms over the continental United States. *J. Appl. Meteor.*, **41**, 1065–1080, [https://doi.org/10.1175/1520-0450\(2002\)041<1065:EOBOSR>2.0.CO;2](https://doi.org/10.1175/1520-0450(2002)041<1065:EOBOSR>2.0.CO;2).
- Mei, Y., E. N. Anagnostou, E. I. Nikolopoulos, and M. Borga, 2014: Error analysis of satellite precipitation products in mountainous basins. *J. Hydrometeorol.*, **15**, 1778–1793, <https://doi.org/10.1175/JHM-D-13-0194.1>.
- Morrissey, M. L., H. J. Diamond, M. J. McPhaden, H. P. Freitag, and J. S. Greene, 2012: An investigation of the consistency of TAO-TRITON buoy-mounted capacitance rain gauges. *J. Atmos. Oceanic Technol.*, **29**, 834–845, <https://doi.org/10.1175/JTECH-D-11-00171.1>.
- Munchak, S. J., and G. Skofronick-Jackson, 2013: Evaluation of precipitation detection over various surfaces from passive microwave imagers and sounders. *Atmos. Res.*, **131**, 81–94, <https://doi.org/10.1016/j.atmosres.2012.10.011>.
- NASA, 2017: GPROF2017 Version 1 (used in GPM V5 processing). NASA Algorithm Theoretical Basis Doc., 63 pp., [http://rain.atmos.colostate.edu/ATBD/ATBD\\_GPM\\_June1\\_2017.pdf](http://rain.atmos.colostate.edu/ATBD/ATBD_GPM_June1_2017.pdf).
- Neumann, B., A. T. Vafeidis, J. Zimmermann, and R. J. Nicholls, 2015: Future coastal population growth and exposure to sea-level rise and coastal flooding – A global assessment. *PLOS ONE*, **10**, e0118571, <https://doi.org/10.1371/journal.pone.0118571>.
- Petković, V., and C. D. Kummerow, 2015: Performance of the GPM passive microwave retrieval in the Balkan flood even of 2014. *J. Hydrometeorol.*, **16**, 2501–2518, <https://doi.org/10.1175/JHM-D-15-0018.1>.
- Petty, G. W., and K. Li, 2013: Improved passive microwave retrievals of rain rate over land and ocean. Part II: Validation and intercomparison. *J. Atmos. Oceanic Technol.*, **30**, 2509–2526, <https://doi.org/10.1175/JTECH-D-12-00184.1>.
- Porcaccia, L., P. E. Kirstetter, J. J. Gourley, V. Maggioni, B. L. Cheong, and M. N. Anagnostou, 2017: Toward a polarimetric radar classification scheme for coalescence-dominant precipitation: Application to complex terrain. *J. Hydrometeorol.*, **18**, 3199–3215, <https://doi.org/10.1175/JHM-D-17-0016.1>.
- Prigent, C., W. B. Rossow, and E. Matthews, 1997: Microwave land surface emissivities estimated from SSM/I observations. *J. Geophys. Res.*, **102**, 21 867–21 890, <https://doi.org/10.1029/97JD01360>.
- Saharia, M., P. E. Kirstetter, H. Vergara, J. J. Gourley, and Y. Hong, 2017: Characterization of floods in the United States. *J. Hydrol.*, **548**, 524–535, <https://doi.org/10.1016/j.jhydrol.2017.03.010>.
- Sapiano, M. R. P., and P. A. Arkin, 2009: An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *J. Hydrometeorol.*, **10**, 149–166, <https://doi.org/10.1175/2008JHM1052.1>.
- Serra, Y. L., 2018: Precipitation measurements from the Tropical Moored Array: A review and look ahead. *Quart. J. Roy. Meteor. Soc.*, **144**, 221–234, <https://doi.org/10.1002/qj.3287>.
- , and M. J. McPhaden, 2003: Multiple time- and space-scale comparisons of ATLAS buoy rain gauge measurements with TRMM satellite precipitation measurements\*. *J. Appl. Meteor.*, **42**, 1045–1059, [https://doi.org/10.1175/1520-0450\(2003\)042<1045:MTASCO>2.0.CO;2](https://doi.org/10.1175/1520-0450(2003)042<1045:MTASCO>2.0.CO;2).
- , P. A. A'Hearn, H. P. Freitag, and M. J. McPhaden, 2001: ATLAS self-siphoning rain gauge error estimates. *J. Atmos. Oceanic Technol.*, **18**, 1989–2002, [https://doi.org/10.1175/1520-0426\(2001\)018<1989:ASSRGE>2.0.CO;2](https://doi.org/10.1175/1520-0426(2001)018<1989:ASSRGE>2.0.CO;2).
- Smith, S. R., and Coauthors, 2009: The data management system for the Shipboard Automated Meteorological and Oceanographic Systems (SAMOS) initiative. *Proceedings of the OceanObs'09: Sustained Ocean Observations and Information for Society Conf.*, Vol. 2, J. Hall, D. E. Harrison,



- and D. Stammer, Eds., ESA Publication WPP-306, European Space Agency, 10 pp., <https://doi.org/10.5270/OceanObs09.cwp.83>.
- Stampoulis, D., E. N. Anagnostou, and E. I. Nikolopoulos, 2013: Assessment of high-resolution satellite-based rainfall estimates over the Mediterranean during heavy precipitation events. *J. Hydrometeor.*, **14**, 1500–1514, <https://doi.org/10.1175/JHM-D-12-0167.1>.
- Takbiri, Z., A. Ebtehaj, E. Foufoula-Georgiou, P. Kirstetter, and F. J. Turk, 2019: A prognostic nested k-nearest approach for microwave precipitation phase detection over snow cover. *J. Hydrometeor.*, **20**, 251–274, <https://doi.org/10.1175/JHM-D-18-0021.1>.
- Wolff, D. B., and B. L. Fisher, 2009: Comparisons of instantaneous TRMM ground validation and satellite rain-rate estimates at different spatial scales. *J. Appl. Meteor. Climatol.*, **47**, 2215–2237, <https://doi.org/10.1175/2008JAMC1875.1>.
- Zhang, J., and Coauthors, 2016: Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97**, 621–638, <https://doi.org/10.1175/BAMS-D-14-00174.1>.