On the Propagation of Satellite Precipitation Estimation Errors: From Passive Microwave to Infrared Estimates

SHRUTI A. UPADHYAYA

Cooperative Institute for Mesoscale Meteorological Studies, Norman, Oklahoma

PIERRE-EMMANUEL KIRSTETTER

School of Meteorology, and School of Civil Engineering and Environmental Science, and Advanced Radar Research Center, University of Oklahoma, and NOAA/National Severe Storms Laboratory, Norman, Oklahoma

JONATHAN J. GOURLEY

NOAA/National Severe Storms Laboratory, Norman, Oklahoma

ROBERT J. KULIGOWSKI

NOAA/NESDIS/Center for Satellite Applications and Research, College Park, Maryland

ABSTRACT

The launch of NOAA's latest generation of geostationary satellites known as the Geostationary Operational Environmental Satellite (GOES)-R Series has opened new opportunities in quantifying precipitation rates. Recent efforts have strived to utilize these data to improve space-based precipitation retrievals. The overall objective of the present work is to carry out a detailed error budget analysis of the improved Self-Calibrating Multivariate Precipitation Retrieval (SCaMPR) algorithm for GOES-R and the passive microwave (MW) combined (MWCOMB) precipitation dataset used to calibrate it with an aim to provide insights regarding strengths and weaknesses of these products. This study systematically analyzes the errors across different climate regions and also as a function of different precipitation types over the conterminous United States. The reference precipitation dataset is Ground-Validation Multi-Radar Multi-Sensor (GV-MRMS). Overall, MWCOMB reveals smaller errors as compared to SCaMPR. However, the analysis indicated that that the major portion of error in SCaMPR is propagated from the MWCOMB calibration data. The major challenge starts with poor detection from MWCOMB, which propagates in SCaMPR. In particular, MWCOMB misses 90% of cool stratiform precipitation and the overall detection score is around 40%. The ability of the algorithms to quantify precipitation amounts for the Warm Stratiform, Cool Stratiform, and Tropical/Stratiform Mix categories is poor compared to the Convective and Tropical/Convective Mix categories with additional challenges in complex terrain regions. Further analysis showed strong similarities in systematic and random error models with both products. This suggests that the potential of high-resolution GOES-R observations remains underutilized in SCaMPR due to the errors from the calibrator MWCOMB.

1. Introduction

Precipitation is undeniably one of the important components of various environmental cycles, in particular, the water cycle. Quantifying precipitation is vital to understand and model various components of these cycles. Given the dearth of dense and uniform ground observational networks on Earth, quantitative precipitation estimations (QPEs) derived from the vantage point of space provide the best option at the global scale. QPEs with low latency and at high spatiotemporal resolutions are critical for near-real-time applications such as rapid monitoring and forecasting of high-impact societal events like flash floods, debris flows, and shallow landslides. Such resolution can be obtained primarily from satellite sensors on board geostationary Earth orbit (GEO) platforms.

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

NOAA's Advanced Baseline Imager (ABI) sensor on board the latest generation of Geostationary Operational Environmental Satellites (GOES-R Series) provides 3 times more spectral channels, 4 times the resolution, and 5 times faster scanning when compared to its predecessor imager on board previous-generation GOES (Schmit et al. (2017). This new generation of GEO sensors has opened new opportunities in quantifying precipitation rates, specifically in the western United States where the data from these GEO sensors can complement the degraded weather radar coverage of the Weather Service Radar-1988 Doppler (WSR-88D) network (Gebregiorgis et al. 2018). Along with the ABI sensor on board *GOES-16* and *GOES-17* satellites, other new generation GEO sensors include Spinning Enhanced Visible Infrared Imager (SEVIRI; *Meteosat-8*, *Meteosat-9*, *Meteosat-10*, *Meteosat-11*), Multichannel Scanning Unit-Geostationary (MSU-GS; *Electro-L N1–2*), Advanced Himawari Imager (AHI; *Himawari-8* and *Himawari-9*), Advanced Geosynchronous Radiation Imager (AGRI; *FY-4A*), Advanced Meteorological Imager (AMI; *GEO-KOMPSAT-2A*), all with more than 10 spectral channels. These next generation GEO satellites provide global coverage at very high spatiotemporal and spectral resolutions (https://www.goes-r.gov/users/abiScanModeInfo.html). With this significant progress in GEO sensor technology, the next generation of precipitation retrieval algorithms must follow suit. Recently, several attempts to utilize these datasets aimed at improving the precipitation retrievals (e.g., Kirstetter et al. 2018; Kuligowski et al. 2016; Ma et al. 2018; Meyer et al. 2016; Thies et al. 2008).

The Self-Calibrating Multivariate Precipitation Retrieval (SCaMPR; Kuligowski 2002; Kuligowski et al. 2016) algorithm is the operational precipitation retrieval algorithm for *GOES-16/GOES-17*. SCaMPR is trained with passive microwave (PMW) precipitation rate estimates from the Climate Prediction Center (CPC) combined microwave (MWCOMB) dataset (Joyce et al. 2004). While visible/infrared (VIS/IR) sensors on board GEO platforms are primarily sensitive to the cloud-top properties indirectly related to surface precipitation rates, PMW observations sense total cloud water and/or ice content (depending on the frequencies) and provide information on integrated water content that enables more accurate surface precipitation retrievals. MWCOMB provides global instantaneous estimates with a 15-h time delay and is also used as a reference for other satellite precipitation algorithms such as the CPC morphing algorithm (CMORPH; Joyce et al. 2004). Within SCaMPR, several statistical models are developed to link ABI observations and MWCOMB rates in order to detect and quantify precipitation at a spatial scale of ∼2 km at

nadir and 5-min temporal resolution across the conterminous United States (CONUS) (15 min across North and South America).

The advancement in algorithms demands the assessment of their performance. Evaluating the accuracy of satellite precipitation products has always been one of the primary objectives of the International Precipitation Working Group (IPWG: http://www.isac.cnr.it/~ipwg/). This is important not only to advance the products' utility for various applications and locations, but also to provide insights regarding strengths and weaknesses to the algorithm developers. Blended satellite precipitation products [e.g., Integrated Multisatellite Retrievals for GPM (IMERG), CMORPH] combine precipitation measurements from PMW and VIS/IR and ensure consistency by calibrating the GEO-based products at higher spatiotemporal resolution with PMW estimates. While the propagation of QPE uncertainty from PMW precipitation estimates to VIS/IR is of critical significance for precipitation estimation from space, this topic has not been extensively studied in the literature. The overall objective of the present work is to carry out an error budget analysis of the new improved SCaMPR algorithm and its calibrator data MWCOMB. A quality-controlled radar and gauge-based precipitation dataset across the CONUS from the Multi-Radar Multi-Sensor (MRMS) system is used as reference to evaluate both MWCOMB and SCaMPR algorithms (Zhang et al. 2016).

The CONUS is an important validation site for performance assessment of satellite precipitation products (e.g., Gebregiorgis et al. 2018; Khan et al. 2018; Kirstetter et al. 2012, 2014, 2018; Tan et al. 2017) mainly due to the availability of high-resolution, quality-controlled reference data and also the large diversity in precipitation types, rates, and topographic and climatic characteristics. Recently, Kirstetter et al. (2012, 2014) set up a ground validation framework across the CONUS for the Global Precipitation Measurement (GPM) mission using quality-controlled MRMS data called GV-MRMS. GV-MRMS has been used as reference for several validation, calibration, and algorithm development studies listed previously. GV-MRMS also provides a precipitation type product used in the assessment and error budget analysis of SCaMPR and MWCOMB in this paper. Additionally, Köppen–Geiger climate classifications are utilized to investigate the impact of different climatic characteristics on the spatial distribution of errors across regions (Peel et al. 2007). Overall, this error budget work is part of an ongoing effort in exploring the potential of observations from new-generation GEO sensors for improving the state-of-the-art precipitation estimation from space, with the ultimate goal of providing seamless high-resolution and low-latency precipitation estimates across the CONUS and beyond.

Section 2 details the dataset and methodology. Section 3 presents results from the error budget analysis, and section 4 concludes the article and describes future steps.

2. Dataset characteristics and methods

a. Multi-Radar Multi-Sensor

The MRMS system integrates approximately 180 operational radar observations and ingests approximately 7000 hourly rain gauge observations, atmospheric model analyses, and satellite data to generate a suite of severe weather and QPE products at 0.01° (~1 km) and 2-min resolution across the CONUS and southern Canada (Zhang et al. 2016). The MRMS v11 surface precipitation rate product is obtained by applying Z–R relationships at each grid cell based on an automated precipitation type classification to mosaicked reflectivity fields. The categories of precipitation types include 1) Warm Stratiform Rain, 2) Cool Stratiform Rain, 3) Convective Rain, 4) Tropical–Stratiform Rain Mix, 5) Tropical–Convective Rain Mix, 6) Hail, and 7) Snow. Kirstetter et al. (2012, 2014) set up a standardized reference for GPM ground validation based on MRMS (GV-MRMS; Kirstetter et al. 2018). Highly quality-controlled and gauge-corrected GV-MRMS precipitation rate and type products are used here as the reference for the summer season of 2018. Only the most trustworthy data identified with maximum radar quality index (RQI) are used in the analysis.

b. SCaMPR and MWCOMB

The newly improved SCaMPR (Kuligowski et al. 2016) is the operational QPE algorithm for the GOES-R/ABI sensor. SCaMPR derives various indices from five ABI spectral channel observations to detect and quantify precipitation using discriminant and multiple linear regression techniques, respectively. To account for spatiotemporal variations in the relationship between cloud-top properties and surface precipitation, SCaMPR builds models for each cloud type that are adapted geographically and regularly updated. To mitigate false precipitation detections due to subcloud evaporation, a relative humidity (RH) correction is applied using numerical weather prediction model outputs. MWCOMB precipitation estimates are used to train the statistical models linking cloud-top properties and surface precipitation. The MWCOMB product (CPC combined microwave dataset; Joyce et al. 2004) intercalibrates satellite-based PMW QPEs, and combines them to generate a gridded precipitation product at a resolution of 0.08° × 0.08° and 30 min. During the

TABLE 1. Validation sample sizes for rain/no-rain classes for different dataset used in this study. "All rain" is the number of pixels where all three products have precipitation rates above 0.01 mm h^{-1}. Rain = precipitation > 0.01 mm h^{-1}.

Product	Total size	Rain	No rain
MRMS	18 288 370	2 708 153 (14.81%)	15 580 217 (85.19%)
MWCOMB	18 288 370	1 618 210 (8.85%)	16 670 160 (91.15%)
SCaMPR	18 288 370	1 754 528 (9.59%)	16 533 842 (90.41%)
All rain	711 473	711 473	

period of study, MWCOMB used only the Advanced Microwave Sounding Unit (AMSU) or Microwave Humidity Sounder (MHS) retrievals from the POES series and the EUMETSAT MetOp series. Being a remote sensing, space-based QPE product, it is expected that limitations of MWCOMB, such as underestimation of precipitation rates generated by warm clouds, propagate in the SCaMPR estimates (Kuligowski et al. 2016). Note that SCaMPR is available at the native resolution of the ABI sensor, i.e., ~2 km at nadir and 5 min across the CONUS. For uniformity in the following intercomparison, both GV-MRMS and SCaMPR precipitation estimates are resampled to match the spatiotemporal scale of the coarser MWCOMB grids.

Table 1 shows the sample sizes with proportion of rain and no-rain in MWCOMB and aggregated GV-MRMS and SCaMPR precipitation products. Note that in the present article the terms "rain" and "precipitation" are used synonymously given the period of study during the warm season. A rainy pixel is defined when the precipitation rate is greater than 0.01 mm h^{-1}. "All rain" indicates the sample of pixels where all three products are defined as rainy. For satellite products, approximately 90% of the matched datasets in all three products is no-rain while the remaining ~10% is rain.

Table 2 shows the distribution of GV-MRMS validation sample sizes including "all rain" data across the eight MRMS precipitation types. A pixel (0.08°) is classified as a particular precipitation type when at least 80% of the MRMS precipitation types belong to the same precipitation type in the matched pixel; otherwise it is assigned to the Combined (Comb) category. Most data points fall under the No-Precipitation type, followed by the Comb type of precipitation. Note that the Comb precipitation type is not analyzed in the present study. Considering broad precipitation types, Stratiform (consisting of Warm Stratiform, Tropical Stratiform/Mix, and Cool Stratiform) dominates over Convective (consisting of Convective, Tropical Convective/Mix, and Hail) and Snow occurrence across the CONUS. Within the stratiform category, the highest proportion is observed for Warm Stratiform, followed by Tropical Stratiform/Mix

TABLE 2. Validation samples sizes as a function of MRMS precipitation types. A data point is classified in a particular precipitation type when at least 80% of the time in space (0.08° grid) and time (30 min) the same precipitation type is observed in all the MRMS pixels; otherwise, it is classified as Comb category. "All rain" means a pixel has a precipitation rate > 0.01 mm h^{-1} in all three products (MRMS, SCaMPR, and MWCOMB).

	Precipitation type from MRMS	Precipitation type (acronym used in study)	MRMS (count)	MRMS (%)	All rain (count)	All rain (%)
1	Convective	Convec	3203	0.02	2927	0.41
2	Cool Stratiform	Cool_Strat	4981	0.03	311	0.04
3	Hail	Hail	228	0.00	224	0.03
4	Combined Type Data	Comb	1 309 046	7.16	362 035	50.89
5	No Precipitation	NoPrecip	16 368 756	89.50	67 093	9.43
6	Snow	Snow	153	0.00	7	0.00
7	Tropical/Convective Mix	Trp_ConvMix	1325	0.01	1268	0.18
8	Tropical/Stratiform Mix	Trp_StratMix	32 976	0.18	27 395	3.85
9	Warm Stratiform	WarmStrat	567 702	3.10	250 213	35.17

and Cool Stratiform. Within the convective category, Convective has higher occurrences, followed by Tropical Convective/Mix and Hail. Note that since it is the warm season, Snow occurrences are least frequent followed by Hail when compared to other types. Table 2 gives more details of the sample sizes associated with the different precipitation types.

The error budget analysis has been carried out considering the dependence on precipitation types and climate regions. Climate regions are obtained across the CONUS from Köppen–Geiger climate classifications matched with other datasets (Peel et al. 2007). There are more than 20 Köppen–Geiger climate classes across the CONUS, which are grouped into six land classes and one ocean class (Khan et al. 2018). Figure 1 shows the climate regions and Table 3 indicates the corresponding sample sizes. Note that regions R1, R2, and R3 (western CONUS) are complex topography regions and receive less frequent precipitation compared to other regions in the eastern part of the CONUS. Table 4 provides the sample sizes broken down according to both precipitation types and climate regions. While precipitation types are generally distributed across all regions, most Snow and Cool Stratiform occurrences are found in climate region R5 (northern and high-altitude regions) and parts

of region R2. Across R3 (a complex topography region) and R1 (an arid desert region) occurrences of Warm Stratiform precipitation type are found.

The error budget analysis has been carried out by focusing on the detection and quantification of satellite precipitation products. Detection is assessed using various categorical statistics [Eqs. (1)–(4)] computed for each climate region and precipitation type separately:

$$\text{probability of detection (POD)} = \frac{h}{h + m}, \quad (1)$$

$$\text{false alarm ratio (FAR)} = \frac{f}{h + f}, \quad (2)$$

$$\text{bias} = \frac{h + f}{h + m}, \quad (3)$$

Heidke skill score (HSS)

$$= \frac{2(hc - fm)}{(h + f)(f + c) + (h + m)(m + c)}, \quad (4)$$

where h, m, f, and c stand for hits, misses, false detections, and correct rejections by occurrence, respectively. It is complemented with decomposing the error by volume as hits, miss, and false rain:

$$\text{volumetric hit index (VHI)} = \frac{\sum_{i=1}^{n} [(\text{SRE}_i > t \ \& \ \text{REF}_i > t)]}{\sum_{i=1}^{n} [(\text{SRE}_i > t \ \& \ \text{REF}_i > t)] + \sum_{i=1}^{n} [(\text{SRE}_i \leq t \ \& \ \text{REF}_i > t)]}, \quad (5)$$

$$\text{volumetric miss index (VMI)} = \frac{\sum_{i=1}^{n} [(\text{SRE}_i \leq t \ \& \ \text{REF}_i > t)]}{\sum_{i=1}^{n} [(\text{SRE}_i > t \ \& \ \text{REF}_i > t)] + \sum_{i=1}^{n} [(\text{SRE}_i \leq t \ \& \ \text{REF}_i > t)]}, \quad (6)$$
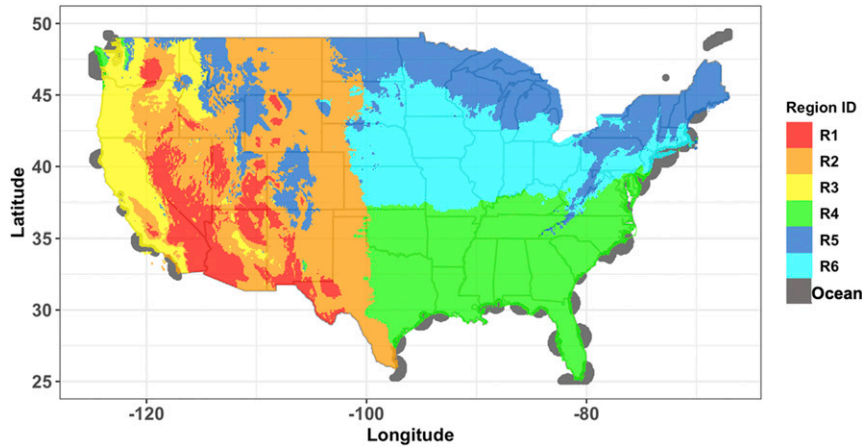
FIG. 1. Simplified Köppen–Geiger climate map and locations of points near shore (represented as ocean here).

$$\text{volumetric false alarm ratio (VFAR)} = \frac{\sum_{i=1}^{n}[(\text{SRE}_i > t \,\&\, \text{REF}_i \leq t)]}{\sum_{i=1}^{n}[(\text{SRE}_i > t \,\&\, \text{REF}_i > t)] + \sum_{i=1}^{n}[(\text{SRE}_i > t \,\&\, \text{REF}_i \leq t)]}, \tag{7}$$

where SRE is the satellite-based precipitation product (i.e., MWCOMB or SCaMPR) and REF is the reference precipitation, n is the validation sample size, and t is the precipitation threshold above which VHI and VFAR are computed.

For quantification, the following statistics are used:

$$\text{correlation coefficient (CC)} = \frac{\sum_{i=1}^{n}(\text{REF}_i - \overline{\text{REF}})(\text{SRE}_i - \overline{\text{SRE}})}{\sqrt{\sum_{i=1}^{n}(\text{REF}_i - \overline{\text{REF}})^2} \times \sqrt{\sum_{i=1}^{n}(\text{SRE}_i - \overline{\text{SRE}})^2}}, \tag{8}$$

$$\text{root-mean-square error (RMSE)} = \sqrt{\frac{1}{n}(\text{SRE}_i - \text{REF}_i)^2}, \tag{9}$$

$$\text{additive error} = \frac{1}{n}\sum_{i=1}^{n}\text{SRE}_i - \text{REF}_i, \tag{10}$$

TABLE 3. Validation sample size as a function of climate regions; "rain" and "no-rain" is according to MRMS and "all rain" is defined as in Table 2.

	ID	Class description	Total (count)	Rain (count)	No rain (count)	All rain (count)	Proportion of rain in region (%)
1	Ocean	Not land	658 529	102 690	555 839	29 386	4.13
2	R1	Arid desert	317 662	31 685	285 977	5142	0.72
3	R2	Arid steppe	2 349 051	257 173	2 091 878	61 237	8.61
4	R3	Temperate, Mediterranean, continental with warm summers	433 302	25 620	407 682	2509	0.35
5	R4	Temperate, oceanic, subtropical	4 183 360	716 997	3 466 363	229 237	32.22
6	R5	Warm summer continental, continental	6 199 358	890 547	5 308 811	178 309	25.06
7	R6	Hot summer continental	4 147 108	683 441	3 463 667	205 653	28.91

TABLE 4. Validation sample size as a function of both precipitation types and climate regions.

Precipitation type from MRMS	Simplified Köppen–Geiger climate regions						
	R1	R2	R3	R4	R5	R6	Ocean
Convec	13	331	0	760	716	1333	50
Cool_Strat	0	1005	2	0	3780	191	3
Hail	0	33	0	60	19	116	0
Comb	14 592	125 101	11 208	369 740	403 540	333 054	51 811
NoPrecip	298 472	2 170 739	418 151	3 691 179	5 562 769	3 642 966	584 480
Snow	0	96	0	0	56	1	0
Trp_ConvMix	0	40	0	801	44	267	173
Trp_StratMix	76	2034	1	16 058	3705	9159	1943
WarmStrat	4509	49 672	3940	104 762	224 729	160 021	20 069

$$\text{relative error } (\%) = \frac{\sum_{i=1}^{n} \text{SRE}_i - \text{REF}_i}{\sum_{i=1}^{n} \text{REF}_i} \times 100. \quad (11)$$

In addition to these statistics, the systematic and random components of additive error are extracted. Following Kirstetter et al. (2013), the systematic error is defined as the median of the error while the random error is defined as the interquantile range (10%–90%) of error.

3. Error budget analysis

a. Precipitation detection assessment

The categorical and volumetric statistics [Eqs. (1)–(7)] for both SCaMPR and MWCOMB are given in Table 5 at a threshold of 0.01 mm h^{-1} and for various thresholds in Fig. 2. MWCOMB and SCaMPR have similar scores and dependencies on thresholds, but overall MWCOMB has better scores compared to SCaMPR. This is expected since PMW-based retrievals are generally more accurate than IR-based ones. For example, the POD for MWCOMB is higher than SCaMPR and FAR is higher with SCaMPR. Both products detect less than 50% of precipitation occurrence, and out of the detected precipitation, 34% (44%) is falsely detected by MWCOMB (SCaMPR). When considering the volume of precipitation with VHI (Table 5), MWCOMB (SCaMPR) detects around 77% (60%) of precipitation volume, and volumetric FAR is only 23% (33%). These observations indicate that most pixels with low precipitation rates are misclassified in

terms of rain versus no rain by both products. As the rain rate threshold increases both products' performances decline (Fig. 2). At thresholds higher than 1 mm h^{-1} the SCaMPR FAR statistic follows more closely with MWCOMB. This may be due to SCaMPR using a detection threshold of 1 mm h^{-1} for separating raining from nonraining pixels in the calibration process. The major performance difference between SCaMPR and MWCOMB is with POD. SCaMPR has weak performances in detecting precipitation, which affects HSS. It can be noted from Fig. 2 that POD and HSS increase slightly as the threshold increases up to around 1 mm h^{-1} then deteriorate for higher thresholds. FAR monotonically increases with the threshold, with a higher rate of increase after 1 mm h^{-1}.

The impact of P_{rain} (percent coverage of precipitating pixels from GV-MRMS within one grid cell of MWCOMB) on the detection performance of MWCOMB and SCaMPR is shown in Table 6 with POD for different P_{rain} classes. As P_{rain} increases, POD also increases, which indicates that the satellite products tend to miss pixels that are partially covered with precipitation compared to pixels completely covered by precipitation. Figure 3 summarizes the distributions of P_{rain} values for hits, misses, false alarms, and correct rejections with box-and-whisker plots. As expected, P_{rain} is almost 0 for correct rejections c and false alarms f for both satellite products, whereas more than half of the pixels correctly detected are completely filled with precipitation (median of P_{rain} for hits h is almost 1). For missed m precipitation, P_{rain} exhibits large variability for both satellite products. It confirms that extreme cases ($P_{\text{rain}} = 1$ and $P_{\text{rain}} = 0$) are well classified, but intermediate situations (P_{rain} in the

TABLE 5. Categorical and volumetric statistics for SCaMPR and MWCOMB at threshold of 0.01 mm h^{-1}.

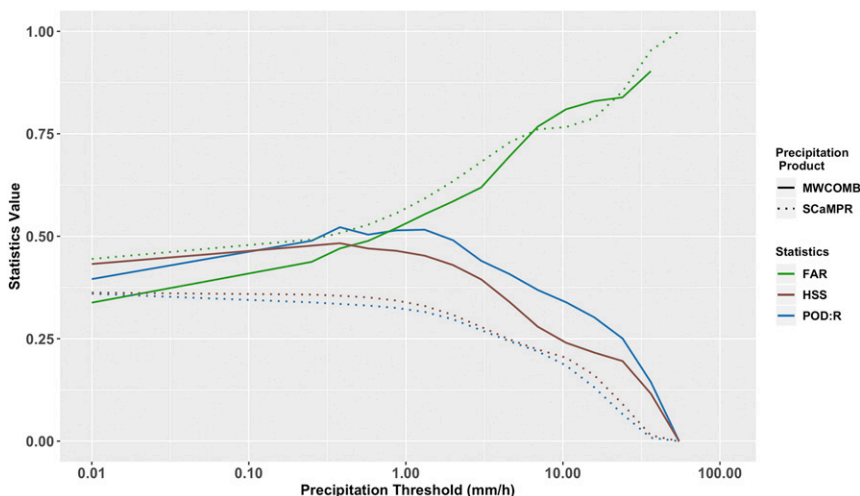|	h (%)	m (%)	f (%)	c (%)	POD	FAR	Bias	HSS	VHI	VMI	VFAR
MWCOMB	5.86	8.95	2.99	82.2	0.40	0.34	0.60	0.43	0.77	0.23	0.23
SCaMPR	5.33	9.48	4.27	80.93	0.36	0.44	0.65	0.36	0.60	0.40	0.33

FIG. 2. Categorical statistics for SCaMPR and MWCOMB as a function of precipitation threshold.

range [0–1]) are challenging to classify. This is attributed to the interplay between coarse satellite sampling resolution and edges of rain areas [nonuniform beam filling (NUBF)] as highlighted in Kirstetter et al. (2012, 2013). Note that the median of P_{rain} for missed m precipitation in slightly closer to 1 for SCaMPR compared to MWCOMB, which indicates that SCaMPR tends to miss precipitation with partially filled pixels to relatively larger extent compared to MWCOMB. This result is expected as MWCOMB is based on observations more directly related to surface precipitation than SCaMPR.

Table 7 shows the distribution of hits and POD of both satellite products as functions of precipitation type from GV-MRMS. Note that GV-MRMS precipitation types imply that GV-MRMS detects precipitation, thus only detection hits are displayed here. Precipitation types generally associated with high precipitation intensity, such as Convective, Hail, and Tropical/Convective Mix types are well detected (hit percent > 90) by both the products, with SCaMPR detecting 100% of the raining pixels classified as Hail. However, the Tropical/Stratiform Mix type is better detected by MWCOMB (94.0%) than by SCaMPR (85.9%). In MRMS, this precipitation type tends to be associated collision–coalescence processes with limited vertical extent and ice content, which makes it more challenging to detect with IR and water vapor (WV) channels than with microwave channels (Cecil and Zipser 2002). Clouds producing moderate precipitation rates and with limited ice content and vertical extent are not as well detected by satellite sensors. More than 50% of the Warm Stratiform precipitation type is detected by both products with SCaMPR detection performance (54%) lower than MWCOMB (64%).

Cold environments are more challenging. The Cool Stratiform type is observed to be poorly detected by both products, with SCaMPR showing surprisingly better detection rate (21%) than its calibrator MWCOMB (10%). Similar observations are made with the Snow type, with overall better detection with SCaMPR (23%) than MWCOMB (9%). A possible explanation is that surface emissivity affects microwave observations in the range [10–37] GHz more significantly than infrared observations. Surface emissivity variability associated with surface snow is particularly challenging for microwave observations (e.g., Takbiri et al. 2019; Gebregiorgis et al. 2017). The Cool Stratiform category may be associated with snow at the surface as suggested by (Chen et al. 2016). Note that the limited observations of Snow type and the influence of surface emissivity associated with cold surface conditions (Zhang et al. 2016; Takbiri et al. 2019) merits further investigation. Finally, Comb types (i.e., data points where no precipitation type is greater than 80%) and No-Precipitation (i.e., 80% of MWCOMB grid is zero precipitation but the mean GV-MRMS rain rate > 0.01 mm h^{-1}) are poorly detected by both products (hits < 50%).

Figure 4 reports the detection statistics for both products across different climate regions as defined in

TABLE 6. POD as a function of P_{rain}.

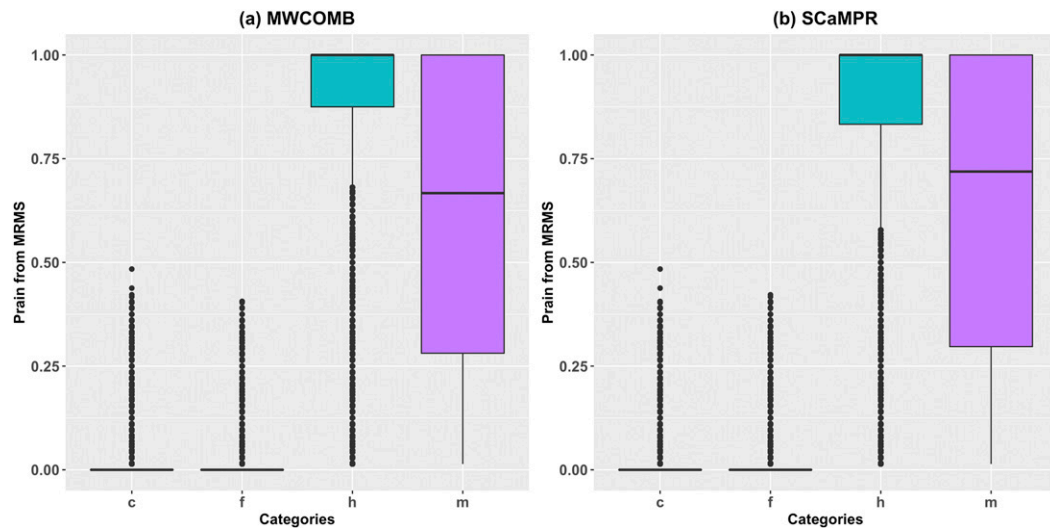|	POD	
P_{rain}	MWCOMB	SCaMPR
0.0–0.2	0.15	0.16
0.2–0.4	0.19	0.19
0.4–0.6	0.23	0.21
0.6–0.8	0.27	0.25
0.8–1.0	0.32	0.3

FIG. 3. Distribution of P_{rain} of validation samples for hits h, misses m, false alarms f, and correct rejections c in both tested products.

section 2. MWCOMB shows better detection performance than SCaMPR across all climate regions, which impacts the detected volumes of precipitation. Across the mountainous region R3, both products show poor performance. The detection of precipitation is low in occurrence and volume for both the products with high missed precipitation errors. This can be attributed to orographic enhancement of rainfall as well as climatology of the region where the Warm Stratiform precipitation type is prominent and is poorly detected by both products (Table 7). Challenges in detecting precipitation in complex terrain from space have been well documented (Derin et al. 2016; Golian et al. 2015; Upadhyaya and Ramsankaran 2018). Both products also show lower precipitation detection over R5 (northern and high-altitude regions). This can be attributed to the climatology as R5 has high proportions of Snow and Cool Stratiform precipitation types, which are detected relatively poorly (Table 7). Note that SCaMPR POD

(0.32) is slightly better than MWCOMB (0.29), which can again be attributed to the fact that Snow and Cool Stratiform precipitation types are detected better by SCaMPR than MWCOMB (Table 7). However, FAR is relatively higher for SCaMPR (0.43) than MWCOMB (0.27) which affects HSS.

The MWCOMB performance across arid regions R1 and R2 (Arid Desert and Arid Steppe, respectively) is marked by more frequent missed precipitation than false precipitation. However, when considering the volume of rainfall, false errors dominate. In the case of SCaMPR, errors due to missed precipitation (both by occurrence and by volume) dominate the false error. The performance of both products for the climate regions R4 and R6 (Subtropical and Continental, respectively) is relatively better than other regions and reveal similar miss and false alarm errors. Overall, volumetric statistics show that both products have larger errors across the western CONUS (R1, R2, and R3)

TABLE 7. MWCOMB and SCaMPR error decomposition of hits as functions of MRMS precipitation types.

Precipitation type	MWCOMB		SCaMPR	
	Hits pixels (count)	POD (%)	Hits pixels (count)	POD (%)
Convection	3154	98.59	2960	92.53
Cool Stratiform	474	9.63	1038	21.1
Hail	224	98.25	228	100
Comb	536 438	42.15	494 540	38.86
No-Precip	134 440	16.23	140 940	17.02
Snow	9	5.88	35	22.88
Tropical Convection Mix	1304	98.42	1285	96.98
Tropical Stratiform Mix	30 961	94.02	28 292	85.92
Warm Stratiform	364 186	64.51	305 131	54.05

	h	*m*	*f*	*c*	*POD*	*FAR*	*Bias*	*HSS*	*VHI*	*VMI*	*VFAR*
MWCOMB											
Ocean	6.61	8.98	1.92	82.5	0.42	0.22	0.55	0.49	0.73	0.27	0.16
R1	4.01	5.96	3.58	86.5	0.4	0.47	0.76	0.41	0.8	0.2	0.37
R2	4.6	6.35	3.15	85.9	0.42	0.41	0.71	0.44	0.83	0.17	0.28
R3	1.41	4.5	1.15	92.9	0.24	0.45	0.43	0.31	0.6	0.4	0.36
R4	8.08	9.06	4.88	78	0.47	0.38	0.76	0.46	0.79	0.21	0.24
R5	4.22	10.1	1.56	84.1	0.29	0.27	0.4	0.37	0.67	0.33	0.21
R6	7.25	9.23	3.46	80.1	0.44	0.32	0.65	0.46	0.81	0.19	0.23
SCaMPR											
Ocean	6.11	9.48	3.83	80.6	0.39	0.39	0.64	0.41	0.63	0.37	0.24
R1	2.03	7.95	2.52	87.5	0.2	0.55	0.46	0.23	0.36	0.64	0.49
R2	3.32	7.63	3.39	85.7	0.3	0.51	0.61	0.32	0.57	0.43	0.4
R3	0.97	4.94	1.31	92.8	0.16	0.57	0.38	0.21	0.37	0.63	0.45
R4	6.94	10.2	5.36	77.5	0.41	0.44	0.72	0.38	0.64	0.36	0.31
R5	4.61	9.76	3.41	82.2	0.32	0.43	0.56	0.34	0.55	0.45	0.31
R6	6.5	9.98	5.45	78.1	0.39	0.46	0.72	0.37	0.63	0.37	0.36

FIG. 4. Categorical and volumetric statistics of MWCOMB and SCaMPR as a function of climate regions. Note: For ease of interpreting table, the statistics scores are represented as bars.

compared to the eastern CONUS. There is a more significant degradation in performance in SCaMPR going from east to west. Along with orographically enhanced rainfall, other possible reasons for this could be the related to the relatively shallow viewing angle of the ABI over the western United States relative to the eastern United States. This causes the satellite footprints to become larger (coarser spatial resolution) and also increases the optical pathlength of the absorbing atmosphere between the satellite and Earth as viewing zenith angle increases, which results in lower brightness temperatures

known as limb cooling (Elmer et al. 2016). It is important to highlight that the SCaMPR algorithm is trained at a coarser spatial resolution of MWCOMB and applied at a finer resolution of ABI/GOES-R, which could also aggravate the error in SCaMPR compared to MWCOMB.

b. Precipitation quantification assessment

All the statistics reported in this section are calculated based on the validation sample with "all rain" (where all three products GV-MRMS, MWCOMB, and SCaMPR have precipitation rates $> 0.01 \, mm \, h^{-1}$). Figure 5 shows
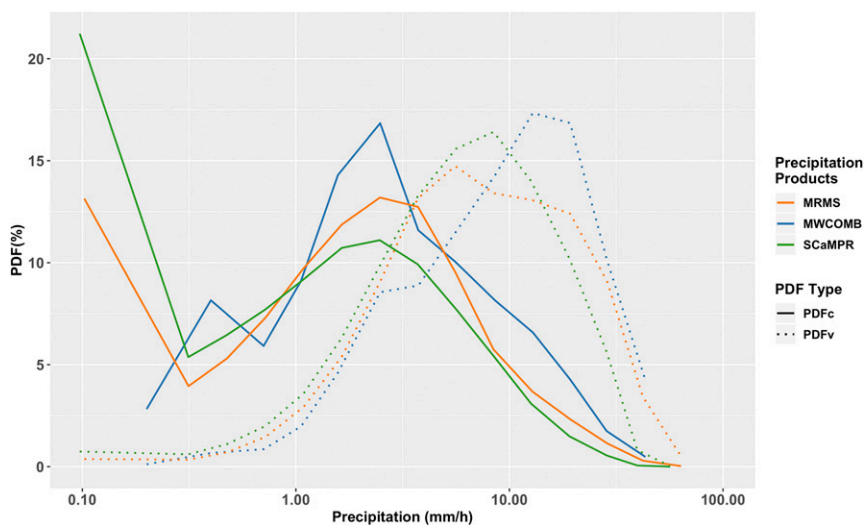

FIG. 5. Precipitation distribution by occurrence (PDFc) and by volume (PDFv) for all the three products.
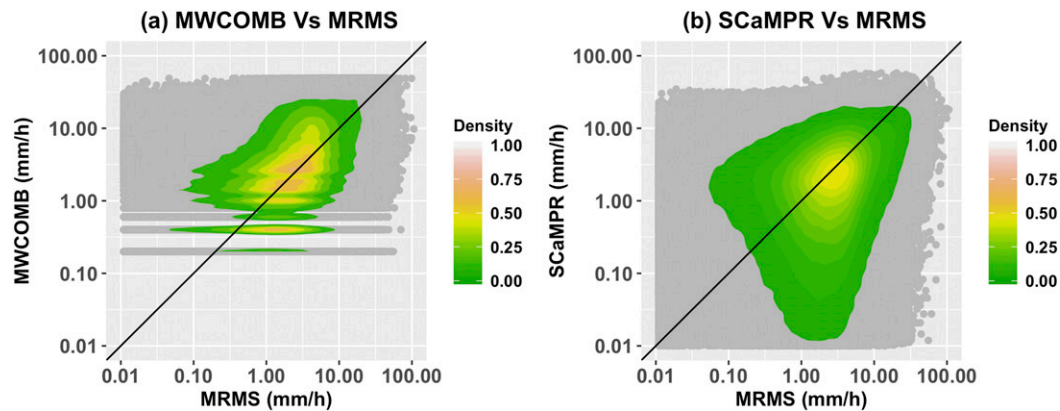
FIG. 6. Density colored scatterplot of (a) MWCOMB and (b) SCaMPR against reference GV-MRMS precipitation rates.

the precipitation density distribution by occurrence and volume for all three products following Kirstetter et al. (2012). Note that the lowest precipitation rate reported by MWCOMB is $0.2\,\mathrm{mm\,h^{-1}}$. However, GV-MRMS and especially SCaMPR report significant proportions of precipitation rates $< 0.2\,\mathrm{mm\,h^{-1}}$, and SCaMPR largely over samples low precipitation rates compared with GV-MRMS. These discrepancies can be traced back to the design and the training of SCaMPR. While SCaMPR is trained for precipitation detection with MWCOMB rates greater than $1\,\mathrm{mm\,h^{-1}}$, it retrieves rates lower than this threshold especially lower than the MWCOMB threshold ($0.2\,\mathrm{mm\,h^{-1}}$). This extrapolation outside of its training domain likely affects the accuracy of the retrievals and mixes detection and quantification of very light precipitation values.

This oversampling tendency of low precipitation rates in SCaMPR (Fig. 5) comes along with high occurrence of false alarms observed in SCaMPR in comparison to false precipitation by volume. It is important to highlight here that the primary application of short-fuse satellite rainfall products such as SCaMPR is for heavy rain/flash flood situations and not for climate applications where light rain and drizzle become much more important. Precipitation density by occurrence consistently peaks around $3–5\,\mathrm{mm\,h^{-1}}$ for all three products. However, maximum contributions by volume occur at higher precipitation rates ($>5\,\mathrm{mm\,h^{-1}}$). The MWCOMB precipitation density distribution by volume is shifted toward higher rates compared to the reference. The SCaMPR precipitation density distribution is shifted toward lower rates compared to GV-MRMS, indicating overestimation by MWCOMB and underestimation by SCaMPR.

Figure 6 shows the density scatterplots and Table 8 reports the quantitative metrics for MWCOMB and

SCaMPR compared to the GV-MRMS reference. There are discontinuities in the MWCOMB scatter density plot due to a minimum data resolution of $0.2\,\mathrm{mm\,h^{-1}}$. As expected, MWCOMB displays a higher correlation than SCaMPR. This is attributed to the more direct relation between microwave measurements and surface precipitation than with IR or water vapor channels. However, in terms of overall error, MWCOMB overestimates by 36% and SCaMPR underestimates by 22% (Table 8). More details can be observed from Fig. 6 such as the tendency for MWCOMB to overestimate the higher rainfall rates while SCaMPR underestimates the lightest rainfall rates, corroborating early results in Fig. 5. It is interesting to note that the RMSE for SCaMPR is slightly improved compared to MWCOMB. It may be due to the postprocessing RH adjustment in SCaMPR. Further error decomposition into systematic and random components and studying the relative error as a function of the reference rainfall rates should reveal more details about quantitative error characteristics of both the products.

Figure 7 shows the density scatterplot of additive error as a function of GV-MRMS precipitation rates. The corresponding systematic (median) and random components (90%–10% quantile) of the additive error are extracted. Table 9 reports the relative error (%) for both the products as a function of precipitation rates. Both products tend to overestimate lower precipitation rates

TABLE 8. Quantitative statistics for MWCOMB and SCaMPR.

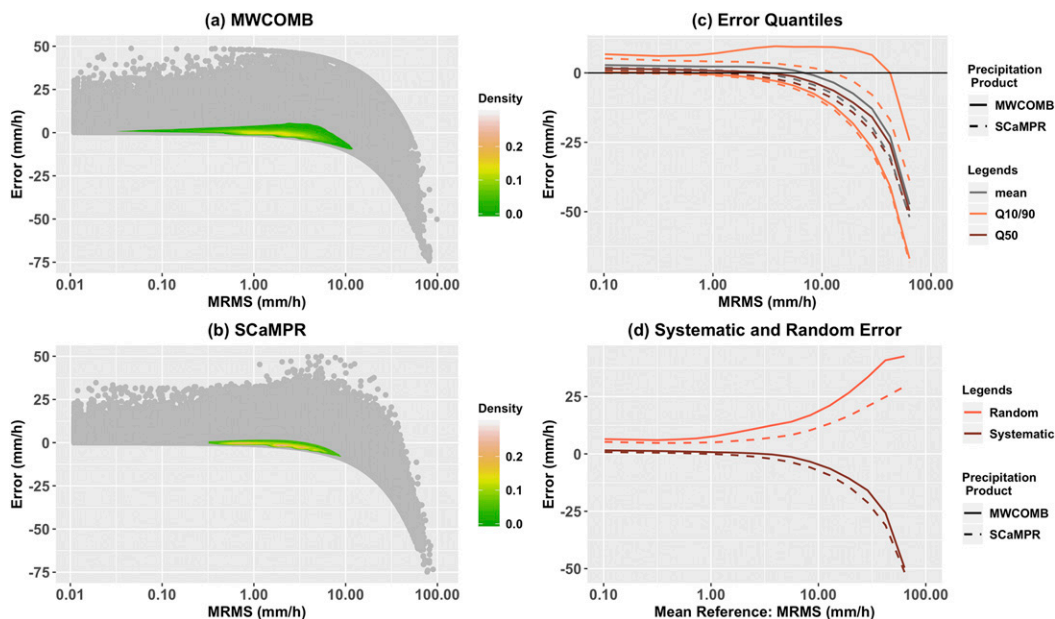|	MWCOMB	SCaMPR
CC	0.38	0.32
RMSE	6.58	5.63
Additive error	1.27	−0.82
Relative error (%)	35.31	−22.83

FIG. 7. Additive error scatterplot of (a) MWCOMB and (b) SCaMPR vs GV-MRMS, (c) error quantiles for MWCOMB and SCaMPR, and (d) random and systematic additive error for MWCOMB and SCaMPR.

and underestimate higher precipitation rates. MWCOMB underestimates precipitation rates greater than $7.5\,\mathrm{mm\,h^{-1}}$ whereas SCaMPR starts to underestimate at lower precipitation rates ($5\,\mathrm{mm\,h^{-1}}$). As discussed earlier, the overestimation of low precipitation intensity ($<0.2\,\mathrm{mm\,h^{-1}}$) is attributed to the lower threshold on MWCOMB rates used for SCaMPR training ($1\,\mathrm{mm\,h^{-1}}$) that propagates into SCaMPR. Conversely, a $50\,\mathrm{mm\,h^{-1}}$ cap is applied in MWCOMB and causes both products to underestimate precipitation rates greater than $50\,\mathrm{mm\,h^{-1}}$. Kuligowski et al. (2016) identified such quantification error characteristics in the earlier version of SCaMPR and attributed them to the algorithm structure (mean-square-error minimization of linear regression models). In the newer SCaMPR version for GOES-R a histogram matching method is adopted to overcome this limitation. However, our error budget analysis indicates that the error trend remains. It is attributed therefore to the training with MWCOMB.

It is also interesting to note that the relative error in SCaMPR is significantly less than MWCOMB for precipitation rates less than $5\,\mathrm{mm\,h^{-1}}$ (Table 9 and Figs. 7a,b). Also, SCaMPR displays a slightly higher systematic error compared to MWCOMB which is the result of underestimation (from Figs. 7c and 7d). MWCOMB shows higher random error than SCaMPR. This surprising reduction in the random error and less relative error for low intensity precipitation rates in SCaMPR compared to its own training product MWCOMB could possibly be due to post processing,

i.e., relative humidity correction in SCaMPR algorithm. Overall, the error similarities of both products highlight that a major portion of SCaMPR error originates from its training reference MWCOMB.

Table 10 shows the quantitative statistics as a function of precipitation type. Overall, the trend in both products is very similar for different precipitation types with some exceptions, again highlighting the propagation of MWCOMB errors in SCaMPR retrievals. Both products show the lowest correlation for Hail precipitation type whereas for other high-intensity precipitation types (i.e., Convective and Tropical/Convective Mix) the correlation coefficient (CC) of SCaMPR is slightly higher than that of MWCOMB. This supports the fact that IR measurements have a stronger relationship with surface precipitation from convective precipitation systems (Ba and Gruber 2001), whereas stratiform precipitation types such as cool stratiform,

TABLE 9. Relative error (%) for MWCOMB and SCaMPR as a function of GV-MRMS precipitation rates.

MRMS (mm h^{-1})	MWCOMB	SCaMPR
0.01–0.25	2779.07	1835.65
0.25–5	109.6	19.21
5–7.5	9.93	−43.03
7.5–10	−11.97	−52.13
10–25	−32.6	−61.96
25–50	−47.24	−68.55
50–100	−71.28	−79.58

TABLE 10. Quantitative statistics of MWCOMB (MW) and SCaMPR (SC) as a function of precipitation types.

Precipitation type	CC		RMSE (mm h^{-1})		Additive error (mm h^{-1})		Relative error (%)	
	MW	SC	MW	SC	MW	SC	MW	SC
Convec	0.28	0.31	12.17	14.21	−2.66	−9.86	−16.27	−60.34
Cool_Strat	0.25	0.22	2.16	2.35	−0.6	−1.18	−23.63	−46.74
Hail	−0.03	−0.01	17.39	24.12	−5.34	−21.05	−14.13	−55.71
Comb	0.42	0.36	7.31	6.55	0.82	−1.46	17.67	−31.5
NoPrecip	0.04	0.05	4.87	3.59	2.79	1.91	1445.57	987.28
Trp_ConvMix	0.12	0.29	26.12	28.33	−19.07	−23.49	−62.36	−76.79
Trp_StratMix	0.03	0.13	5.56	4.99	−0.36	−2.38	−6.73	−44.25
WarmStrat	0.18	0.11	5.54	3.86	1.87	−0.23	73.25	−9.06

warm stratiform, tropical stratiform/mix have CCs less than 0.2. Generally, RMSEs for both products are similar and highest for heavy precipitation types such as Hail, Convective, and Tropical/Convective Mix, which can be explained by the higher spatial and temporal variability associated with these types. Note that Snow type is not reported due to small sample size (Table 2).

In terms of error characteristics, MWCOMB overestimates Warm Stratiform and Comb precipitation types and underestimates other precipitation types, whereas SCaMPR underestimates all precipitation types. The general overestimation observed with MWCOMB can be attributed to retrievals in Warm Stratiform and Comb precipitation types, since the majority of data points fall in these two categories. The underestimation by SCaMPR for precipitation types associated with high-intensity precipitation such as Convective, Hail, Tropical/Convective Mix is significantly larger than MWCOMB. Again, this could be due to MWCOMB underestimation that propagates into SCaMPR and is aggravated as a result of the SCaMPR algorithm structure. It is interesting to note that both the products detect Hail with POD > 80%, however the quantification statistics are poor (note: the sample is small and needs further investigation). Also, in comparison to MWCOMB, SCaMPR exhibits better scores from most quantitative and detection statistics for Tropical/Stratiform Mix and Tropical/Convective Mix, suggesting that SCaMPR performs better in quantifying tropical precipitation types than MWCOMB.

Figure 8 reports the quantitative statistics across different climate regions. With few exceptions, the trends in the region-based statistics are similar to the overall statistics; i.e., the CC of MWCOMB is slightly higher than SCaMPR, RMSE of SCaMPR is slightly lower than MWCOMB, and overestimation in MWCOMB and underestimation in SCaMPR. There is no significant performance difference in quantitative statistics across the regions except for the complex terrain region R3 where both products have poor CC and SCaMPR overestimates

precipitation rates. As described earlier, the complex terrain region has issues with both detection and quantification from space-based passive sensors and several attempts have been made to improve space-based precipitation retrieval over complex topography (Taniguchi et al. 2013; Upadhyaya and Ramsankaran 2016; Vicente et al. 2002). It is interesting to note that detection statistics indicate significant differences across climate regions whereas the quantitative statistics reveal much less sensitivity.

4. Summary and conclusions

Considering the recent progression in the spatial, temporal and spectral resolution of new generation geostationary satellites, the present study provides an error budget analysis of the recently improved SCaMPR precipitation rates from *GOES-16* (Kuligowski et al. 2016) and of MWCOMB (Joyce et al. 2004), which is used to calibrate it. The MWCOMB product is based on passive microwave data and is used as a calibrator for several GEO-based products including SCaMPR. This error budget work is part of an ongoing effort to explore the potential of new generation GEO observations to improve state-of-the-art precipitation estimation from space, and ultimately to provide seamless, high-resolution

	CC		RMSE		Additive Error		Relative Error (%)	
	MW	SC	MW	SC	MW	SC	MW	SC
Ocean	0.33	0.34	6.13	5.74	0.15	-1.15	3.73	-29.26
R1	0.28	0.30	6.70	4.74	2.27	-1.08	79.94	-38.10
R2	0.39	0.30	7.12	5.52	2.35	-0.83	68.62	-24.35
R3	0.21	0.10	4.20	4.35	1.18	0.26	62.53	13.86
R4	0.36	0.33	6.98	6.33	0.72	-1.14	17.73	-27.95
R5	0.41	0.33	5.57	4.62	1.25	-0.43	41.77	-14.21
R6	0.43	0.33	6.84	5.66	1.73	-0.78	46.95	-21.14

FIG. 8. Quantitative statistics of MWCOMB (MW) and SCaMPR (SC) as a function of climate regions. For ease of interpreting table, the statistics scores are represented as bars.

and low-latency precipitation estimates across the CONUS and beyond. In this study, the GV-MRMS data provided by Kirstetter et al. (2012, 2014) are used as a reference. The investigation is carried out across the CONUS for the summer season (June–September) of 2018. To understand the impact of different climatic characteristics on the spatial distribution of errors across the region, the error budgets of SCaMPR and MWCOMB are analyzed for different precipitation types from GV-MRMS and across Köppen–Geiger climate regions. Overall MWCOMB shows better evaluation scores than SCaMPR. The specific conclusions from the study are summarized as follows.

a. Precipitation detection

- The training data used for SCaMPR (MWCOMB) detect only 40% of MRMS precipitating data points, which propagates to SCaMPR and limits its precipitation detection to only 36%.
- The satellite products tend to miss pixels partially covered with precipitation compared to pixels completely covered by precipitation, which is attributed to the interplay between coarse satellite sampling resolution and edges of rain areas (nonuniform beam filling effect) as highlighted in Kirstetter et al. (2012, 2013)
- Precipitation types generally associated with high precipitation intensity, such as Convective, Hail, and Tropical/Convective Mix types are well detected by both products compared to other precipitation types such as Warm Stratiform, Cool Stratiform, Tropical/Stratiform Mix, and Snow. In particular, 90% of Cool Stratiform is missed by MWCOMB, which again propagates into the accuracy of SCaMPR precipitation detection.
- Detection scores based on volume show that both products have larger errors across the western CONUS compared to the eastern CONUS, and particularly in SCaMPR there is significant reduction in performance going from east to west. This can be attributed mainly to complex topography of the region and orographic enhancement of rainfall. For SCaMPR, the coarser resolution of the ABI sensor from *GOES-16* across the western CONUS might have played a role in increasing the errors, along with the effects of longer optical pathlengths at shallower viewing angles and associated impacts on cloud-top temperatures.

b. Precipitation quantification

- As expected, MWCOMB exhibits a higher correlation with the reference rainfall product than SCaMPR. This is attributed to the more direct relation between microwave measurements and surface precipitation than with IR/WV satellite observations.

- MWCOMB has a general tendency to overestimate the precipitation rates whereas SCaMPR underestimates. However, both products tend to overestimate lower precipitation rates and underestimate higher precipitation rates, which is propagated to SCaMPR from MWCOMB.
- Underestimation by SCaMPR for precipitation types associated with high-intensity precipitation such as Convective, Hail, Tropical/Convective Mix is significantly larger than MWCOMB. This could mainly be a consequence of initial underestimation in MWCOMB that propagates to SCaMPR and is aggravated as a result of the structure of SCaMPR.
- There are strong similarities in the error models of both the products with SCaMPR displaying higher systematic error than MWCOMB and MWCOMB displaying higher random error than SCaMPR. This surprising reduction in the random error in SCaMPR compared to its own training product MWCOMB could possibly be due to postprocessing; i.e., RH adjustment of rain rates for subcloud evaporation in the SCaMPR algorithm.
- Analyzing the quantification accuracy by precipitation type also shows similar performance by both products. The quantification of precipitation in Hail and other stratiform types such as Warm Stratiform, Cool Stratiform, and Tropical/Stratiform Mix is poor compared to Convective and Tropical/Convective Mix.
- As opposed to the detection scores, analyzing quantification accuracy across climate regions showed no significant performance difference across the regions except for the complex terrain region R3 where both the products display poor CC.

In summary, the error characteristics of SCaMPR and MWCOMB are very similar although some exceptions exist. Note that the error scores observed in this study for both the products are similar to the accuracy of other satellite precipitation products (e.g., Kuligowski et al. 2016; Kirstetter et al. 2018). The challenge starts at the initial stage where poor detection using PMW estimates limits the detection capability of IR estimates that are calibrated using them. Furthermore, in the quantification stage the limitations of PMW estimates are propagated to SCaMPR, with challenges over complex terrain regions and cold environments. Thus, the major portion of error in the SCaMPR IR estimates is propagated from the MWCOMB reference product. Furthermore, the SCaMPR algorithm is calibrated at coarser resolution than MWCOMB and applied at the higher resolution of the ABI IR data. These results suggest that the potential of high-resolution ABI data used in SCaMPR remain underutilized due to consideration of coarser-scale data

as reference. It is also important to highlight that the classification of precipitation types in SCaMPR is not directly based on retrieved cloud properties but on observed changes in the relationship between the IR brightness temperature and the PMW rain rate as a function of different GOES IR parameters (Kuligowski et al. 2016). In the future, we propose to utilize high-resolution, quality-controlled GV-MRMS precipitation rates and precipitation type products to guide the precipitation classification and quantification of an improved SCaMPR algorithm over the CONUS and surrounding regions.

REFERENCES

Ba, M. B., and A. Gruber, 2001: GOES Multispectral Rainfall Algorithm (GMSRA). *J. Appl. Meteor.*, **40**, 1500–1514, https://doi.org/10.1175/1520-0450(2001)040<1500:GMRAG>2.0.CO;2.

Cecil, D. J., and E. J. Zipser, 2002: Reflectivity, ice scattering, and lightning characteristics of hurricane eyewalls and rainbands. Part II: Intercomparison of observations. *Mon. Wea. Rev.*, **130**, 785–801, https://doi.org/10.1175/1520-0493(2002)130<0785:RISALC>2.0.CO;2.

Chen, S., J. J. Gourley, Y. Hong, Q. Cao, N. Carr, P.-E. Kirstetter, J. Zhang, and Z. Flamig, 2016: Using citizen science reports to evaluate estimates of surface precipitation type. *Bull. Amer. Meteor. Soc.*, **97**, 187–193, https://doi.org/10.1175/BAMS-D-13-00247.1.

Derin, Y., and Coauthors, 2016: Multiregional satellite precipitation products evaluation over complex terrain. *J. Hydrometeor.*, **17**, 1817–1836, https://doi.org/10.1175/JHM-D-15-0197.1.

Elmer, N. J., E. Berndt, and G. J. Jedlovec, 2016: Limb correction of MODIS and VIIRS infrared channels for the improved interpretation of RGB composites. *J. Atmos. Oceanic Technol.*, **33**, 1073–1087, https://doi.org/10.1175/JTECH-D-15-0245.1.

Gebregiorgis, A. S., P.-E. Kirstetter, Y. E. Hong, N. J. Carr, J. J. Gourley, W. Petersen, and Y. Zheng, 2017: Understanding overland multisensor satellite precipitation error in TMPA-RT products. *J. Hydrometeor.*, **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.

Golian, S., S. Moazami, P.-E. Kirstetter, and Y. Hong, 2015: Evaluating the performance of merged multi-satellite precipitation products over a complex terrain. *Water Resour. Manage.*, **29**, 4885–4901, https://doi.org/10.1007/s11269-015-1096-6.

Joyce, R. J., J. E. Janowiak, P. A. Arkin, and P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeor.*, **5**, 487–503, https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2.

Khan, S., V. Maggioni, and P.-E. Kirstetter, 2018: Investigating the potential of using satellite-based precipitation radars as reference for evaluating multisatellite merged products. *J. Geophys. Res. Atmos.*, **123**, 8646–8660, https://doi.org/10.1029/2018JD028584.

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. Hydrometeor.*, **13**, 1285–1300, https://doi.org/10.1175/JHM-D-11-0139.1.

——, Y. Hong, J. J. Gourley, M. Schwaller, W. Petersen, and J. Zhang, 2013: Comparison of TRMM 2A25 products, version 6 and version 7, with NOAA/NSSL ground radar–based National Mosaic QPE. *J. Hydrometeor.*, **14**, 661–669, https://doi.org/10.1175/JHM-D-12-030.1.

——, ——, ——, 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.

——, 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.

Kuligowski, R. J., 2002: A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates. *J. Hydrometeor.*, **3**, 112–130, https://doi.org/10.1175/1525-7541(2002)003<0112:ASCRTG>2.0.CO;2.

——, Y. Li, Y. Hao, and Y. Zhang, 2016: Improvements to the goes-r rainfall rate algorithm. *J. Hydrometeor.*, **17**, 1693–1704, https://doi.org/10.1175/JHM-D-15-0186.1.

Ma, L., G. Zhang, and E. Lu, 2018: Using the gradient boosting decision tree to improve the delineation of hourly rain areas during the summer from advanced Himawari imager data. *J. Hydrometeor.*, **19**, 761–776, https://doi.org/10.1175/JHM-D-17-0109.1.

Meyer, H., M. Kühnlein, T. Appelhans, and T. Nauss, 2016: Comparison of four machine learning algorithms for their applicability in satellite-based optical rainfall retrievals. *Atmos. Res.*, **169**, 424–433, https://doi.org/10.1016/j.atmosres.2015.09.021.

Peel, M. C., B. L. Finlayson, and T. A. McMahon, 2007: Updated world map of the Köppen-Geiger climate classification. *Hydrol. Earth Syst. Sci.*, **11**, 1633–1644, https://doi.org/10.5194/hess-11-1633-2007.

Schmit, T. J., P. Griffith, M. M. Gunshor, J. M. Daniels, S. J. Goodman, and W. J. Lebair, 2017: A closer look at the ABI on the GOES-R series. *Bull. Amer. Meteor. Soc.*, **98**, 681–698, https://doi.org/10.1175/BAMS-D-15-00230.1.

Takbiri, Z., A. Ebtehaj, E. Foufoula-Georgiou, P.-E. 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.

Tan, J., W. A. Petersen, P.-E. Kirstetter, and Y. Tian, 2017: Performance of IMERG as a function of spatiotemporal scale. *J. Hydrometeor.*, **18**, 307–319, https://doi.org/10.1175/JHM-D-16-0174.1.

Taniguchi, A., and Coauthors, 2013: Improvement of high-resolution satellite rainfall product for Typhoon Morakot (2009) over Taiwan. *J. Hydrometeor.*, **14**, 1859–1871, https://doi.org/10.1175/JHM-D-13-047.1.

Thies, B., T. Nauss, and J. Bendix, 2008: First results on a process-oriented rain area classification technique using Meteosat Second Generation SEVIRI nighttime data. *Adv. Geosci.*, **16**, 63–72, https://doi.org/10.5194/adgeo-16-63-2008.

Upadhyaya, S., and R. A. A. J. Ramsankaran, 2016: Modified-INSAT Multi-Spectral Rainfall Algorithm (M-IMSRA) at climate region scale: Development and validation. *Remote Sens. Environ.*, **187**, 186–201, https://doi.org/10.1016/j.rse.2016.10.013.

——, and ——, 2018: Comprehensive inter-comparison of INSAT multispectral rainfall algorithm estimates and TMPA 3B42-RT V7 estimates across different climate regions of India during southwest monsoon period. *Environ. Monit. Assess.*, **190**, 45, https://doi.org/10.1007/s10661-017-6411-7.

Vicente, G. A., J. C. Davenport, and R. A. Scofield, 2002: The role of orographic and parallax corrections on real time high resolution satellite rainfall rate distribution. *Int. J. Remote Sens.*, **23**, 221–230, https://doi.org/10.1080/01431160010006935.

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.