Evaluation of Numerical Weather Model–Based Satellite Precipitation Adjustment in Tropical Mountainous Regions

XINXUAN ZHANG AND EMMANOUIL N. ANAGNOSTOU

Department of Civil and Environmental Engineering, University of Connecticut, Storrs, Connecticut

(Manuscript received 17 January 2018, in final form 22 January 2019)

ABSTRACT

The study evaluated a numerical weather model (WRF)-based satellite precipitation adjustment technique with 81 heavy precipitation events that occurred in three tropical mountainous regions (Colombia, Peru, and Taiwan). The technique was applied on two widely used near-real-time global satellite precipitation products—the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center morphing technique (CMORPH) and the Global Satellite Mapping of Precipitation project (GSMaP)—for each precipitation event. The WRF-adjusted satellite products along with the near-real-time and gauge-adjusted satellite products as well as the WRF simulation were evaluated by independent gauge networks at daily scale and event total scale. Results show that the near-real-time precipitation products exhibited severe underestimation relative to the gauge observations over the three tropical mountainous regions. The underestimation tended to be larger for higher rainfall accumulations. The WRF-based satellite adjustment provided considerable improvements to the near-real-time CMORPH and GSMaP products. Moreover, error metrics show that WRF-adjusted satellite products outperformed the gauge-adjusted counterparts for most of the events. The effectiveness of WRF-based satellite adjustment varied with events of different physical processes. Thus, the technique applied on satellite precipitation estimates of these events may exhibit inconsistencies in the bias correction.

1. Introduction

Satellite remote sensing plays an irreplaceable role in precipitation measurement because it is the only mean of gathering data with uninterrupted, quasi-global coverage. Precipitation-related satellite observations are of four main types: longwave infrared (IR), visible spectrum (VIS), passive microwave (PMW), and active microwave retrievals. The satellite IR and VIS sensors measure the cloud-top brightness temperature or reflectivity that researchers use to derive precipitation rates by certain retrieval algorithms (Ebert and Manton 1998). These estimates represent an indirect measurement of precipitation, and their accuracy is largely affected by different cloud types, rain systems, and hydroclimatic regimes. The PMW measurements observe the microwave energy emitted by rain droplets or scattered by precipitating ice particles. While the IR/VIS and PMW techniques can only capture horizontal precipitation

patterns and intensities, the precipitation radar (PR) can provide three-dimensional storm structure.

Nowadays, mainstream high-resolution satellite precipitation products are usually generated by combining IR/VIS, PMW, and PR measurements, a conjunction that takes advantage of the different techniques. Examples of these products include the National Oceanic and Atmospheric Administration (NOAA) Climate Prediction Center (CPC) morphing technique (CMORPH; Joyce et al. 2004); the Global Satellite Mapping of Precipitation project (GSMaP; Kubota et al. 2007; Mega et al. 2014); the Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN; Sorooshian et al. 2000); and the Tropical Rainfall Measuring Mission (TRMM) Multisatellite Precipitation Analysis (TMPA; Huffman et al. 2007, 2010, 2015). In addition, the Global Precipitation Measurement (GPM) mission, which was launched in 2014, has provided a new-generation satellite product, the Integrated Multi-satellite Retrievals for GPM (IMERG; Huffman et al. 2015).

Although the past two decades have brought considerable progress in satellite precipitation retrieval

Corresponding author: Emmanouil Anagnostou, manos@engr.uconn.edu

techniques and algorithms, producing reliable satellite products over mountainous areas remains a big challenge. Many studies have been devoted to satellite precipitation evaluation over complex terrain. In South America, [Dinku et al. \(2010\)](#) found severe overestimation by PERSIANN and significant underestimation by GSMaP over Colombia. In Africa, [Hirpa et al. \(2010\)](#) showed significant underestimation of precipitation by PERSIANN over the Ethiopian Plateau, and [Milewski et al. \(2015\)](#) reported underestimation of precipitation by the TMPA product over high-elevation areas of Morocco. In Europe, [Stampoulis and Anagnostou \(2012\)](#) found that CMORPH and TMPA underestimated rainfall over the Italian Alps region. In Asia, [Chen et al. \(2013\)](#) showed significant underestimation of the 2009 extreme Typhoon Morakot by the CMORPH, PERSIANN, and TMPA precipitation products, while [Tong et al. \(2014\)](#) found that TMPA underestimated precipitation over the Tibetan Plateau. Finally, a recent comprehensive error analysis of nine satellite precipitation products over nine mountainous regions showed that all tended to underestimate the high rain rates significantly ([Derin et al. 2016](#)).

Typically, correction methods for satellite precipitation systematic error (bias) rely on comparisons of near-real-time satellite precipitation products with ground observations over large spatial and temporal scales (1° – 5° and monthly; [Xie and Arkin 1997](#); [Mega et al. 2014](#)). To be efficient, this bias estimation requires data from dense in situ gauge networks, which are rarely available over mountainous areas, especially in some tropical regions. Some studies have even shown that the use of in situ gauge observations in data-sparse regions associated with significant spatial precipitation gradients could lead to increased errors in the gauge-adjusted satellite precipitation estimates ([Ghajarnia et al. 2015](#); [Derin et al. 2016](#)).

To overcome this barrier in complex terrain, [Zhang et al. \(2013\)](#) developed a bias-correction technique based solely on high-resolution numerical weather prediction (NWP) simulations. This technique is designed to reduce satellite precipitation underestimation, which is typically due to the low-level orographic enhancement processes in mountainous areas. The technique has been tested for CMORPH using a few case studies in the Alpine region of northern Italy, the Massif Central Mountains in France ([Zhang et al. 2013](#)), the southern Appalachian Mountains in North America ([Zhang et al. 2016](#)), and the Rocky Mountains in Colorado in the western United States ([Nikolopoulos et al. 2015](#)). Results based on these studies have shown that the NWP-based adjustments can reduce the CMORPH underestimation of high rain rates and moderate the magnitude-dependent bias. Authors have argued that

although the NWP-based adjustment is independent of any ground observation, the improvements are comparable to or even better than those from the postprocessed, gauge-adjusted CMORPH precipitation product.

These previous studies were focused on subtropical or temperate zone climates. However, approximately two-thirds of global precipitation occurs in tropical areas, which highlights a need to test the satellite precipitation adjustment technique in those regions. In this study, we provide a comprehensive evaluation of the NWP-based adjustment technique by applying the technique on two high-resolution satellite precipitation products for 81 heavy precipitation events occurring in three tropical mountainous regions.

The next section describes the study regions, rain gauge and satellite precipitation data, and numerical model setups. [Section 3](#) introduces the methodologies of the NWP-based adjustment technique and error analyses. Results and discussions are presented in [section 4](#), and the conclusions are summarized in [section 5](#).

2. Study regions and datasets

a. Study regions

Three tropical mountainous regions are included in this study, two in the Andes Mountains and one in Taiwan. All regions have independent dense gauge network datasets for data evaluation. Although the gauge data time periods vary across the regions, all the selected heavy precipitation events took place during their common period from 2004 to 2010. The criteria of event selection are described in [section 3a](#).

The Andes Mountains are located in South America, running from north ($\sim 10^{\circ}\text{N}$) to south ($\sim 53^{\circ}\text{S}$) along the western coast of the continent. The Colombia domain is a portion of the Northern Andes ([Fig. 1a](#)), where the climate is typically wet and warm. The Colombian Andes can be divided from east to west into three mountain ranges. There are 113 rain gauges used for data evaluation in this region. Most of them are located at the eastern mountain range. The spatial distribution of the gauge network is biased toward high-elevation area ([Fig. 1d](#)). More than 90% of gauges are located at areas with elevation higher than 1000 m. [Figure 1d](#) shows the scatterplot of gauge elevation and total precipitation of all events from this study. The precipitation magnitude is positively correlated to elevation for gauges located below 1000 m, while the correlation is negative for gauges located above 1000-m elevation.

The second study region is the Peru domain in Central Andes where the summer season (from December to February) contributes over 60% of the annual

FIG. 1. Gauge information. (top) Maps of terrain elevation and gauge locations for each study region. The elevation data are obtained from the ASTER Global Digital Elevation Model (<https://doi.org/10.5067/ASTER/ASTGTM.002>). (bottom) Gauge elevation vs total precipitation of all events in each region.

precipitation (Garreaud and Aceituno 2001). Ground observations of Peru region are from 124 rain gauges distributed throughout the mountains (Fig. 1b). Fifty of the gauges are located below 1000-m elevation and 39 of them are located above 1000 m. The rest of the gauges lacked elevation information. There is not clear trend between precipitation and elevation for this region (Fig. 1e).

The third study region is in southeastern Taiwan (Fig. 1c). Complex terrain and an average annual precipitation of more than 2500 mm characterize the island of Taiwan. The eastern part consists mostly of rugged mountains and the western part of the Chianan Plains. Taiwan's climate is influenced by the East Asian monsoon. The monsoon is especially significant in the southeastern region, where approximately 90% of the annual precipitation occurs during the wet season (from May to October; Yu et al. 2006). The ground observations for this study area came from 40 rain gauges in the Tsengwen River basin, where the elevations vary from near sea level to 2540 m. There are 40% of the gauges located below 100-m elevation and 22% of them located above 1000 m. The rest of the gauges are located between 100- and 1000-m elevations. Figure 1f exhibits a

linear positive correlation between the precipitation magnitude and gauge elevations.

In addition to elevation, we also looked at the gauge topographic locations. Table 1 summarizes the percentage of gauges located in valley bottoms, hilltops, along the slopes or flat terrain for the three study regions. This information was derived from DEM data. Specifically, we calculated average DEM values within 1-km radius neighborhood and then subtract the average values from the original DEM to obtain the Topographic Position Index (TPI). Finally, we extracted TPI values of all gauge locations. The strongly negative TPI values indicate valley bottoms, near-zero values indicate flat terrains or slopes, and strongly positive values indicate hilltops. As shown in Table 1, most of the gauges (63%–70%) in the three study regions are located either on mountain slopes or flat terrain. The Colombia region has 21% of the gauges located in valley bottoms and 16% located on hilltops. Peru follows the same trend with more gauges in valley bottoms (22%) than on hilltops (8%). The Taiwan region has more gauges on hilltops (25%) than in valley bottoms (7%). The gauges in valley bottoms may underestimate area-average precipitation that could affect the error statistics presented in this

TABLE 1. Percentage of gauges at different topographic positions.

Study region	Total No. of gauges	Percent of gauges in valley bottoms	Percent of gauges on hilltops	Percent of gauges at slopes or flat terrain
Colombia	113	21	16	63
Peru	124	22	8	70
Taiwan	40	7	25	68

study. Fortunately, the portion of these gauges is relatively low, so, we proceed with the study considering their effect on results to be low.

b. Satellite precipitation products

We applied the NWP-based adjustment technique to two passive microwave-based high-resolution satellite precipitation products, CMORPH and GSMaP. Both apply gauge-based corrections to the near-real-time satellite precipitation estimates.

1) CMORPH

CMORPH is a satellite rainfall retrieval algorithm that uses motion vectors derived from half-hour-interval, geostationary satellite IR imagery to propagate rainfall estimates obtained from Earth-orbiting satellite-based PMW sensors (Joyce et al. 2004). This study used the CMORPH V1.0 near-real-time and gauge-adjusted products with 0.073°/30-min resolution. The gauge-adjusted product is corrected by two widely used long-term datasets, the CPC unified gauge analysis over land and the pentad Global Precipitation Climatology Project (GPCP) over the ocean. CMORPH has a newer version named V0.x, which employed more advanced algorithms. However, CMORPH V0.x does not provide a gauge-adjusted product, and it is not available for the time period of storms in this study.

2) GSMaP

The second satellite product examined in this study was GSMaP (Kubota et al. 2007; Tian et al. 2010). The GSMaP—abbreviated in full as GSMaP_MVK (version 5)—employs a morphing algorithm similar to that used by CMORPH to derive the cloud motion vectors. Unlike CMORPH, however, the GSMaP applies a Kalman filter to update the rain rates derived from the IR brightness temperature (Ushio et al. 2009). The spatial and temporal resolutions of the GSMaP product are 0.1° and hourly, respectively. The gauge-adjusted GSMaP product is available at the same resolutions (Mega et al. 2014). We acknowledge that GSMaP has two newer versions, v6 and v7, in which the algorithm was updated regarding to orographic rainfall retrievals (Shige et al. 2013; Yamamoto and Shige 2015; Yamamoto et al. 2017). However, the GSMaP v6 and v7 are not available before 2014, meaning

that the recent algorithm improvement does not cover storms in this study.

c. Numerical weather simulations

To simulate storm events in the different study areas, we used the numerical Weather Research and Forecasting (WRF) Model, version 3.7.1 (Skamarock et al. 2008). The periods of our WRF storm simulations ranged from 1 to 5 days, with a 12-h spinup prior to each. We initialized and constrained the simulations at the model boundaries by NCEP Global Forecast System (GFS) final analysis fields of 0.5° or 1° (<http://nomads.ncdc.noaa.gov/data/gfsanl>), depending on the availability of GFS data. The WRF Model uses a two-way interactive mode and a three-domain spatial configuration (18, 6, and 2 km). Figure 2 presents the WRF domain setting for each study region. The 2-km inner domains cover the entire essential target areas (as shown in Fig. 1) and output hourly simulated dataset. The Grell 3D cumulus scheme was enabled on the 18- and 6-km domains, but not on the 2-km inner domain.

Before using the WRF simulations in the satellite error adjustment technique, we tested them using various parameterizations against in situ gauge observations to verify the model's ability to reproduce quantitatively the structure of those heavy precipitation storms and their interactions with topography. The main items of the final parameterizations are summarized in Table 2.

3. Methodology

a. Selection of precipitation events

The storms used in this study varied from one-day to multiday events. We based our selection of the events on their severity, represented by daily precipitation derived from gauge observations. We set two thresholds for the area-average rainfall accumulation over each study region. The first, $R_{\max}$, constrained the storm maximum rainfall intensity; the second, $R_{\min}$, constrained the storm length. Storm events are considered finished once the precipitation is lower than $R_{\min}$. For example, an n -day storm event had to satisfy two conditions: 1) $\max(R_i) \geq R_{\max}$ and 2) $R_i \geq R_{\min}$, where R is the region-average gauge daily rainfall intensity and $i \in [1, n]$ represents the event days.

FIG. 2. WRF domain settings for each study region. The map extent is the WRF parent domain (18 km), the blue box is the first nested domain (6 km), and the red box is the second nested domain (2 km).

The threshold values were empirical and unique to each study region. The Colombia and Peru regions had moderate threshold values, while Taiwan had much higher thresholds because of the frequent typhoons in the region. Table 3 summarizes the threshold values and number of events for the three study regions.

b. NWP-based adjustment technique

Before applying the adjustment technique, we spatially averaged the WRF-simulated hourly precipitation data, available at 2-km resolution, to match the coarser spatial resolutions of satellite products (8 km for CMORPH and 10 km for GSMaP). We then applied the adjustment procedure separately for each rainfall event and each satellite product. Since the adjustment focused on land areas only, we ignored all the precipitation values over ocean background surfaces.

First, we adjusted the near-real-time satellite hourly precipitation rates by a power-law function [Eq. (1)] derived from WRF and near-real-time satellite precipitation quantile values:

$$Y = a \times X^b, \quad (1)$$

where X and Y are corresponded to hourly precipitation quantile values of satellite and WRF datasets, respectively. The quantiles are derived from 19 cumulative probability

values (5%, 10%, 15%, ..., 95%) based on nonzero precipitation rates only. The quantile–quantile plots between near-real-time satellite product and WRF are shown in Fig. 3. The CMORPH and GSMaP products have lower quantiles than the WRF simulations for most of the events, which confirms that it is feasible to increase the satellite precipitation by WRF. Parameters a and b were determined based on the least squares method by fitting the X and Y datasets for each rainfall event. Figure 4 presents the values of parameters a and b for each event. The parameters of Colombia are grouped together in narrow ranges. Peru has a slightly wider range for parameter a , but the parameter b values are in the similar range as Colombia. The parameter ranges of Taiwan are much wider. Some of the Taiwan events have high values for either a or b . This could be because the satellite severely underestimated the precipitation.

Once the power-law parameters were obtained from the quantile–quantile datasets, we applied Eq. (1) again on the near-real-time satellite precipitation product to derive the NWP-adjusted precipitation product. This time, X represented each hourly precipitation rate of the near-real-time satellite dataset, and Y represented the NWP-adjusted satellite hourly precipitation rate. The adjustment procedures were applied on each event separately.

TABLE 2. WRF v3.7.1 model parameterizations.

WRF parameter	Scheme
Microphysics	WRF double-moment 6-class scheme
Longwave radiation	Rapid Radiative Transfer Model (RRTM) scheme
Shortwave radiation	Goddard shortwave: Two-stream multiband scheme with ozone from climatology and cloud effects
Surface layer	MM5 similarity: Based on Monin–Obukhov with Carlson–Boland viscous sublayer and standard similarity functions from lookup tables
Land surface	Unified Noah land surface model
Planetary boundary layer	Yonsei University scheme
Cumulus parameterization	Grell 3D: An improved version of the Grell–Devenyi scheme that may also be used on high resolution

TABLE 3. Summary of event selection criteria and number of events.

Study region	$R_{\max}$ (mm day ⁻¹)	$R_{\min}$ (mm day ⁻¹)	No. of events
Colombia	13.5	10	28
Peru	12	10	29
Taiwan	100	50	24

c. Error metrics

The primary task of the error analysis was to evaluate the improvement coming from the WRF-based adjustment relative to the postanalysis gauge-adjusted satellite product and the near-real-time (nonadjusted) counterpart. We compared the WRF-based adjusted and gauge-adjusted satellite products to find out which worked best in each study region. The gauge networks used in data evaluation are independent from the ones producing gauge-adjusted satellite products.

As described in section 2a, seven gridded precipitation datasets were available for each event. The data came from the WRF simulation, the near-real-time CMORPH, the gauge-adjusted CMORPH, the WRF-based adjusted CMORPH, the near-real-time GSMaP, the gauge-adjusted GSMaP, and the WRF-based adjusted GSMaP. The data evaluations were done by gauge locations at two

temporal scales: daily precipitation rate and event accumulated precipitation.

Evaluation of daily precipitation rate was performed by two statistical error metrics: bias ratio score (BS) and Heidke skill score (HSS; Heidke 1926):

$$BS = \frac{A + B}{A + C}, \quad (2)$$

$$HSS = \frac{2(A \times D - B \times C)}{(A + C)(C + D) + (A + B)(B + D)}, \quad (3)$$

where A , B , C , and D were the numbers of occurrences for any specific precipitation threshold:

A was counted when estimator > threshold and gauge observation > threshold;

B was counted when estimator > threshold and gauge observation < threshold;

C was counted when estimator < threshold and gauge observation > threshold;

D was counted when estimator < threshold and gauge observation < threshold.

The value of “estimator” was extracted from the gridded precipitation datasets by a simple nearest-neighbor method, according to the gauge location.

FIG. 3. Quantile–quantile plot of the hourly precipitation rate for each event. (top) Near-real-time CMORPH vs WRF. (bottom) Near-real-time GSMaP vs WRF.

FIG. 4. Scatterplot of parameter a and b values for each event. (left) WRF-based adjustment for CMORPH product. (right) WRF-based adjustment for GSMaP product.

We calculated the BS and HSS at three daily precipitation thresholds for each event. The thresholds were selected by fixed quantiles (10%, 30%, and 50%) of daily gauge precipitation in each study region. Thus, the threshold values differed for each region. For example, Taiwan has much higher thresholds than Colombia and Peru because most of the events in Taiwan are more intense and typhoon related. A BS of 1 is considered as an unbiased estimation, while above or below 1 represents overestimation or underestimation, respectively. The HSS is defined as the number of correct estimated occurrences minus the number of correct estimated occurrences by chance, which is then divided by the total number of estimated occurrences minus the number of correct estimated occurrences by chance. The HSS expression can be found in Zhang et al. (2013). The HSS values range from $-\infty$ to 1, where 1 indicates a perfect estimation and less than or equal to zero indicates a random estimation. Spatial mismatches in the rainfall patterns of different products would affect HSS but not BS. In other words, BS evaluates the overall rainfall occurrences in study region no matter of rainfall location, while HSS would be lower for inaccurate location estimations.

We compared storm-length accumulated precipitation using scatterplots of region-average precipitation and three quantitative statistics: correlation R^2 , normalized root-mean-square error (NRMSE), and mean relative error (MRE). The R^2 is simply the square of Pearson correlation coefficient. The equations of NRMSE and MRE are shown below:

$$\text{NRMSE} = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (E_i - G_i)^2}}{\frac{1}{n} \sum_{i=1}^n G_i}, \quad (4)$$

$$\text{MRE} = \frac{\frac{1}{n} \sum_{i=1}^n (E_i - G_i)}{\frac{1}{n} \sum_{i=1}^n G_i}, \quad (5)$$

where n is the number of events in each study region, and E and G are the region-average accumulated precipitation for each event from the estimator and gauge, respectively. Both NRMSE and MRE are scale independent, which made cross-region comparison more convenient.

4. Results and discussion

Results are discussed below for each study region. The daily precipitation error metrics (BS and HSS) are rendered as boxplots, and the comparisons of event total precipitation are shown in scatterplots with bulk statistics summarized in Table 4.

a. Colombia domain

The Colombia domain is part of the tropical Andes area. Most of the storms analyzed in this study received precipitation from deep convective systems that developed over the two mountain ranges. Figure 5a presents the BS at three daily rainfall accumulation thresholds (2, 6, and 11 mm day⁻¹) for the Colombia domain. The near-real-time CMORPH product exhibits large underestimation for 6 and 11 mm day⁻¹ thresholds. The gauge-adjusted CMORPH product was not able to moderate the underestimation, while the WRF-adjusted CMORPH product effectively reduced the underestimation. The performance of the GSMaP group is quite different from CMORPH. First of all, the near-real-time

TABLE 4. Statistics of event total precipitation for each region.

Study region	Statistics	WRF (upscaled to CMORPH grid)	CMORPH	Gauge-adjusted CMORPH	WRF-adjusted CMORPH	WRF (upscaled to GSMaP grid)	GSMaP	Gauge-adjusted GSMaP	WRF-adjusted GSMaP
Colombia	R^2	0.75	0.62	0.59	0.59	0.72	0.65	0.48	0.63
	NRMSE	0.31	0.49	0.47	0.47	0.34	0.38	0.38	0.61
	MRE	0.12	-0.37	-0.34	0.13	0.14	-0.23	-0.07	0.32
Peru	R^2	0.64	0.56	0.76	0.64	0.63	0.49	0.45	0.67
	NRMSE	0.45	0.66	0.58	0.47	0.45	0.60	0.60	0.50
	MRE	-0.15	-0.53	-0.50	-0.06	-0.15	-0.44	-0.43	0.04
Taiwan	R^2	0.60	0.49	0.56	0.76	0.61	0.65	0.70	0.82
	NRMSE	0.62	0.93	0.78	0.41	0.64	0.82	0.75	0.36
	MRE	-0.39	-0.69	-0.54	-0.03	-0.43	-0.60	-0.53	-0.10

GSMaP product has less underestimation than CMORPH. The gauge-based and WRF-based GSMaP adjustments are comparable to each other. However, the WRF-adjusted GSMaP has a narrower BS value range, which makes it a better product than gauge-adjusted GSMaP.

We computed the HSS metric at the same daily precipitation thresholds as the BS. The variation in the HSS boxplots (Fig. 5b) among the different satellite products was not as significant as for the BS boxplots. The WRF-adjusted CMORPH product had the highest HSS values in CMORPH retrievals and for 6 and 11 mm day⁻¹ thresholds. We noted that the performance of the gauge-adjusted CMORPH product was similar to the near-real-time CMORPH product, while in the GSMaP group the HSS decreased with the gauge adjustment. This demonstrated that the gauge adjustment could worsen the performance of the product by introducing random error (Ghajarnia et al. 2015; Derin et al. 2016). The WRF-adjusted GSMaP product performed better than gauge-adjusted GSMaP and similarly to the near-real-time GSMaP product.

While the WRF-simulated precipitation performed better than all the satellite products in terms of the BS, it does not benefit the HSS. This is because high-resolution WRF simulations can resolve the low-level orographic enhancement that results in more accurate rainfall magnitudes over a relatively large area, but the model has difficulty in locating the orographic rainfall in space and time, which can result in significant errors in hydrological applications (Baldwin et al. 2001; Ducrocq et al. 2002; Zhang et al. 2013). Further discussion about spatial distribution of orographic rainfall is provided in section 4d.

To illustrate our investigation of the error pattern on a larger temporal scale, we display in Figs. 5c and 5d the event total precipitation scatterplots of CMORPH and GSMaP products, respectively. The figure clearly shows the near-real-time CMORPH and GSMaP products had the most significant underestimation in each satellite product, confirming the finding in the BS boxplots. The two adjustment methods effectively moderated this underestimation. In fact, the WRF-based adjustment tended to overestimate a bit, while the gauge-adjusted product continued to underestimate. Table 4 reports the quantitative comparisons for each product. WRF simulation had better performance than any of the satellite-related products. The WRF-adjusted CMORPH was better than the near-real-time and gauge-adjusted CMORPH products. The gauge-adjusted GSMaP product performed better than the near-real-time and WRF-adjusted GSMaP products.

Overall, the WRF-based adjustment did not perform well in Colombia, especially for GSMaP dataset. Some possible reasons include the following:

FIG. 5. Data evaluation of events in the Colombia region. (a) BS for daily precipitation. (b) HSS for daily precipitation. (c) Event total precipitation of WRF and CMORPH products. (d) Event total precipitation of WRF and GSMaP products.

- 1) More than 90% of the Colombia gauges are located above 1000-m elevation, thus data evaluation may not be representative enough due to the limited range of gauge elevations.
- 2) There are two events having severe overestimation on total precipitation (Figs. 5c,d). Error metrics (not shown here) without those two events exhibited that the WRF-adjusted products are consistently better than the near-real-time and gauge-adjusted products.
- 3) The Colombia domain is an inland region that is more likely to have deep convective rain systems.

WRF, as a mesoscale numerical model, generally does well on such simulations and consequently showed the best error metrics. It is noted that although WRF performed well in Colombia in terms of Table 4, it exhibited the lowest HSS score in Fig. 5b, while WRF-adjusted satellite products tend to be higher.

b. Peru domain

This study domain is located in northern Peru, where mountainous convective systems dominate the formation of storms. We again calculated the BS values (Fig. 6a) for three daily rain rate thresholds (2, 6, and 12 mm day⁻¹). As in the Colombia region, the precipitation in the Peru region has been considerably underestimated by near-real-time CMORPH and GSMaP

products, especially for the higher rain rates. In the CMORPH group, we noted no significant difference between the near-real-time and gauge-adjusted products, which may have been because of the limited in situ gauge data. In contrast, the WRF-adjusted product revealed noticeable improvement. It brought the BS median value from 0.3 to 0.95 at the 6 mm day⁻¹ threshold and from 0.5 to 0.9 at the 12 mm day⁻¹ threshold. The performance of GSMaP was consistent with that of CMORPH. Gauge-adjusted GSMaP showed similar results to the near-real-time version, while the WRF-adjusted GSMaP product had median values very close to 1. In both CMORPH and GSMaP retrievals, the BS value ranges of the WRF-adjusted products at the 9 mm day⁻¹ threshold were a little wider than those of the near-real-time products, representing a slight increase in uncertainty. Moreover, the Peru BS boxplot demonstrates that the uncertainty of satellite products increased with rainfall magnitude. The underestimations from near-real-time satellite products tended to be more significant in higher precipitation thresholds, and the corrections from WRF and gauge adjustments of higher rainfall thresholds were shown to be more effective than their corrections of low rainfall thresholds.

Results of the Peru HSS metrics are illustrated in Fig. 6b. WRF simulations showed the lowest score values for 6 and 9 mm day⁻¹ thresholds. This could be due to

FIG. 6. As in Fig. 5, but for the Peru region.

inaccuracies of WRF capturing the spatiotemporal distribution of precipitation. The WRF-adjusted CMORPH product showed the highest values in the CMORPH group, while the gauge-adjusted CMORPH has even lower HSS values than the near-real-time product. In the GSMaP group, neither the gauge-based adjustment nor the WRF-based adjustment provided any improvement at 2 and 6 mm day⁻¹ thresholds. However, the WRF-based adjustment slightly increased HSS values at the 9 mm day⁻¹ threshold.

The event total rainfall comparisons are captured in Figs. 6c and 6d. The WRF adjustment significantly improved the near-real-time satellite estimation, especially for the heavier rainfall (>20 mm) events, while the gauge-adjusted products were almost identical to the near-real-time products for most events. In the quantitative comparison shown in Table 4, the WRF-adjusted satellite products outperformed all other products because of their higher correlation values and lower MRE values. The Peru domain provided a very convincing example of the value of using high-resolution numerical weather simulations to evaluate bias adjustment over data-sparse complex terrain regions.

c. Taiwan domain

The results of the Taiwan domain are captured in Fig. 7. Unlike the other two study domains, heavy precipitation events in Taiwan are mostly caused by typhoons. As a

result, events in this region have very high amount of precipitation. In fact, the average precipitation of all Taiwan events was more than 400 mm, while the corresponding rainfall values in Colombia and Peru were 28 and 21 mm, respectively. So the BS and HSS were calculated at much higher rain rate thresholds in Taiwan: 10, 60, and 110 mm day⁻¹.

The Taiwan BS boxplots (Fig. 7a) show the same trend as in the Colombia and Peru regions: the near-real-time satellite products underestimated rainfall, and the adjusted products significantly moderated this underestimation. Unlike in Colombia and Peru, however, the WRF simulation did not show “close to 1” values in Taiwan. In contrast, WRF exhibited large underestimation especially for the higher rain rates. After combining the WRF and satellite products, we found the WRF-adjusted satellite products performed the best (much closer to 1 than other products). This indicates that although WRF has underestimation, it can still correct near-real-time satellite products as long as the WRF-simulated precipitation is higher than the near-real-time satellite products. The power-law parameters a and b were obtained by the precipitation quantiles and applied on each satellite hourly precipitation value to produce the WRF-adjusted satellite product. Higher near-real-time satellite precipitation rates would benefit more from the exponential adjustment.

The HSS results (Fig. 7b) for Taiwan revealed different trends than in Colombia and Peru for the CMORPH

FIG. 7. As in Fig. 5, but for the Taiwan region.

and GSMaP products. Overall, the HSS metric in Taiwan has lower values, and the HSS value ranges were much wider than in other two regions. All products performed similarly at the 10 mm day⁻¹ threshold. In the CMORPH group, the gauge-adjusted and WRF-adjusted products have a wider 25th–75th percentile range at 60 and 110 mm day⁻¹ thresholds, which reveals that some of the events obtained improvements from the adjustment while some of them were led to the wrong direction by the adjustment. The GSMaP group has the same performance as the CMORPH group.

Event total precipitation plots (Figs. 7c,d) of Taiwan indicated clear benefit from the WRF-adjusted CMORPH and GSMaP products. The improvements from gauge-based adjustment were limited, and the improvements from WRF-based adjustment were significant. The advantage of WRF-based adjustment is even more noticeable in Table 4. The WRF-adjusted CMORPH and WRF-adjusted GSMaP products were better than all other products, including the WRF simulation, for all presented statistics: correlation, NRMSE, and MRE.

d. Discussion

The accuracy of WRF-based satellite adjustment depends on the spatial distribution of near-real-time satellite precipitation products and the precipitation magnitude simulated by WRF. Given that the power-law parameters of the WRF-based adjustment are derived

from the quantile comparisons between WRF and satellite hourly rainfall maps, it is important to note that for this method to work, WRF does not need to provide accurate spatial distributions of rainfall, but it needs to be relatively accurate on the magnitude of rainfall rates within the mountainous domain. Figure 8 shows total precipitation maps for three events. It is noted the spatial distributions of WRF are different than the corresponding satellite precipitation maps. Specifically, WRF-based adjustment enhanced precipitation for CMORPH, while keeping its spatial distribution as the raw near-real-time product. For example, the WRF and CMORPH maps of the Peru event in Fig. 8 show very different rainfall patterns. Specifically, the WRF exhibits intense precipitation between longitudes 78° and 76°W, while the near-real-time CMORPH exhibited very light precipitation in that area. Thus, the effect of WRF-based adjustment was limited in this region. On the other hand, there was relatively high precipitation in the near-real-time CMORPH between longitudes 80° and 78°W, and consequently the WRF-adjustment increased considerably CMORPH precipitation in this area. Similar observation for the Peru event can be found in the maps of GSMaP (Fig. 9). Arguably, there is no indication as to which dataset is more accurate until it is compared to an independent reference dataset (in our case this is the dense rain gauge network). Table 5 shows the NRMSE and MRE values of daily precipitation evaluated based on the rain gauge rainfall

FIG. 8. Event total precipitation of WRF and CMORPH products. Black dots are gauge locations.

values. WRF has higher NRMSE and MRE values than the WRF-adjusted satellite products, which highlights the value of the adjustment to the satellite precipitation. In fact, the evaluation in Table 5 shows that WRF-adjusted CMORPH is the best product for all three regions. The performance of WRF-adjusted GSMaP was not as consistent as the CMORPH product.

Considering the WRF-based satellite adjustment is based on power-law function, which contains an exponential parameter, it should be noted that the adjustment is sensitive to large precipitation magnitudes in near-real-time satellite product. It may introduce high uncertainties for intense precipitation adjustment. For example, the Colombia event total precipitation scatterplots (Figs. 5c,d) exhibited severe overestimation for some events, a few of which were even twice as much as the corresponding near-real-time product. On the other hand, the exponential parameter also allows us to correctly adjust events when the WRF and near-real-time satellite product both exhibit underestimation, but WRF values are much higher than near-real-time satellite product. For example, in the Taiwan event total

precipitation plots (Figs. 7c,d), the event with gauge data around 1400 mm showed a fairly accurate WRF-based satellite precipitation adjustment, while both WRF and near-real-time satellite products had significant underestimation.

5. Conclusions

This study evaluated a model-based satellite precipitation adjustment technique for two high-resolution satellite products over three tropical mountainous regions based on 81 heavy precipitation events. We compared the model-adjusted satellite products to the model simulations, near-real-time satellite products, and gauge-adjusted satellite products.

In general, although the general rainfall patterns are similar across the different dataset, it is clear that more detail and variability exists in the model simulations, particularly over the complex terrain domain, while the satellite products tend to smooth out the terrain effect on precipitation. In addition, the magnitude of satellite rainfall is much lower than that in the model simulations

FIG. 9. Event total precipitation of WRF and GSMaP products. Black dots are gauge locations.

over the mountainous domain. Although, both model and satellite precipitation datasets are uncertain over complex terrain, the orographic enhancement is more likely to be captured by the model than the satellite observations. On the other hand, the model simulations can dislocate the intense precipitation areas, while the satellite products are more reliable on the rainfall locations. Therefore, the combination of model simulations and satellite precipitation through the adjustment method presented in this paper exhibits considerable improvements on precipitation quantification.

Specifically, both CMORPH and GSMaP near-real-time precipitation products exhibited strong underestimation over the tropical mountainous regions. The GSMaP product exhibited less underestimation than the CMORPH product. The satellite underestimations tended to be severer for higher rainfall accumulations. Overall, the gauge-adjusted satellite precipitation products moderated the underestimation of the corresponding near-real-time products. However, for some storm events the gauge adjustment worsens the near-real-time product's accuracy.

The WRF-based adjustment technique is able to improve satellite precipitation estimates. Specifically, the WRF-adjusted satellite products outperformed the gauge-adjusted counterparts for most of events. As mentioned in the discussion, the accuracy of WRF-based satellite adjustment depends on the spatial distribution of near-real-time satellite precipitation products and the precipitation magnitude simulated by WRF. Given the power-law parameters of the WRF-based adjustment, the method is more effective for higher rain rates, which is important because the high rain rates are potentially causing hydrological hazards. However, the WRF-based adjustment may also cause overestimations at locations where satellite precipitation magnitudes are unbiased or exhibiting positive bias. It should be noted that the WRF-based adjustment technique could not correct for the missed rainfall in satellite precipitation products. It is also worth mentioning that the effectiveness of WRF-based satellite adjustment varies with events of different physical processes. Similar biases in the satellite products may have inconsistencies in the bias correction. Overall, WRF-based adjustment performed very well

TABLE 5. Statistics of daily precipitation for single event in each region.

Study region	Statistics	WRF (upscaled to CMORPH grid)	CMORPH	Gauge-adjusted CMORPH	WRF-adjusted CMORPH	WRF (upscaled to GSMaP grid)	GSMaP	Gauge-adjusted GSMaP	WRF-adjusted GSMaP
Colombia event (2 May 2010)	NRMSE	1.84	1.26	1.28	1.23	1.82	1.09	1.21	1.15
	MRE	0.52	-0.64	-0.55	-0.56	0.64	0.01	0.52	0.15
Peru event (30 Mar 2010)	NRMSE	1.65	1.04	1.04	1.05	1.61	1.08	1.00	1.27
	MRE	0.47	-0.50	-0.50	0.04	0.44	-0.41	-0.35	0.23
Taiwan event (24 Aug 2004)	NRMSE	1.67	1.25	1.18	0.93	1.56	1.24	1.07	0.93
	MRE	1.22	-0.87	-0.78	-0.44	1.03	-0.83	-0.62	-0.51

in Peru and Taiwan, but greatly overestimated a few events in Colombia.

The current study used postanalysis weather simulations to evaluate corrections of satellite precipitation products. Future research should investigate the feasibility of using real-time weather forecasts to correct near-real-time high-resolution satellite precipitation products (e.g., CMORPH, GSMaP, PERSIANN-CCS, and IMERG) for heavy precipitation events over complex terrain areas and evaluate hydrologic impacts in terms of flood forecasts. Furthermore, a future study should focus on demonstrating the technique on recently released versions of CMORPH and GSMaP and the GPM-era IMERG product based on recent flood-inducing storms.

Acknowledgments. This work was supported by a NASA Precipitation Measurement Mission award (NNX07AE31G). We acknowledge and appreciate valuable input from three anonymous reviewers.

REFERENCES

- Baldwin, M. E., S. Lakshmiarahan, and J. S. Kain, 2001: Verification of mesoscale features in NWP models. Preprints, *Ninth Conf. on Mesoscale Processes*, Fort Lauderdale, FL, Amer. Meteor. Soc., 255–258.
- Chen, S., and Coauthors, 2013: Performance evaluation of radar and satellite rainfalls for Typhoon Morakot over Taiwan: Are remote-sensing products ready for gauge denial scenario of extreme events? *J. Hydrol.*, **506**, 4–13, <https://doi.org/10.1016/j.jhydrol.2012.12.026>.
- Derin, Y., and Coauthors, 2016: Multiregional satellite precipitation products evaluation over complex terrain. *J. Hydrometeorol.*, **17**, 1817–1836, <https://doi.org/10.1175/JHM-D-15-0197.1>.
- Dinku, T., F. Ruiz, S. J. Connor, and P. Ceccato, 2010: Validation and intercomparison of satellite rainfall estimates over Colombia. *J. Appl. Meteor. Climatol.*, **49**, 1004–1014, <https://doi.org/10.1175/2009JAMC2260.1>.
- Ducrocq, V., D. Ricard, J.-P. Lafore, and F. Orain, 2002: Storm-scale numerical rainfall prediction for five precipitating events over France: On the importance of the initial humidity field. *Wea. Forecasting*, **17**, 1236–1256, [https://doi.org/10.1175/1520-0434\(2002\)017<1236:SSNRPF>2.0.CO;2](https://doi.org/10.1175/1520-0434(2002)017<1236:SSNRPF>2.0.CO;2).
- Ebert, E. E., and M. J. Manton, 1998: Performance of satellite rainfall estimation algorithms during TOGA COARE. *J. Atmos. Sci.*, **55**, 1537–1557, [https://doi.org/10.1175/1520-0469\(1998\)055<1537:POSREA>2.0.CO;2](https://doi.org/10.1175/1520-0469(1998)055<1537:POSREA>2.0.CO;2).
- Garreaud, R. D., and P. Aceituno, 2001: Interannual rainfall variability over the South American Altiplano. *J. Climate*, **14**, 2779–2789, [https://doi.org/10.1175/1520-0442\(2001\)014<2779:IRVOTS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2001)014<2779:IRVOTS>2.0.CO;2).
- Ghajarnia, N., A. Liaghat, and P. D. Arasteh, 2015: Comparison and evaluation of high resolution precipitation estimation products in Urmia Basin-Iran. *Atmos. Res.*, **158**, 50–65, <https://doi.org/10.1016/j.atmosres.2015.02.010>.
- Heidke, P., 1926: Berechnung des Erfolges und der Güte der Windstärkevorhersagen im Sturmwarnungsdienst. *Geogr. Ann.*, **8**, 301–349.

- Hirpa, F. A., M. Gebremichael, and T. Hopson, 2010: Evaluation of high-resolution satellite precipitation products over very complex terrain in Ethiopia. *J. Appl. Meteor. Climatol.*, **49**, 1044–1051, <https://doi.org/10.1175/2009JAMC2298.1>.
- Huffman, G. J., and Coauthors, 2007: The TRMM Multisatellite Precipitation Analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales. *J. Hydrometeorol.*, **8**, 38–55, <https://doi.org/10.1175/JHM560.1>.
- , R. F. Adler, D. T. Bolvin, and E. J. Nelkin, 2010: The TRMM Multi-Satellite Precipitation Analysis (TMPA). *Satellite Rainfall Applications for Surface Hydrology*, M. Gebremichael and F. Hossain, Eds., Springer, 3–22, https://doi.org/10.1007/978-90-481-2915-7_1.
- , D. T. Bolvin, and E. J. Nelkin, 2015: Integrated Multi-satellite Retrievals for GPM (IMERG) technical documentation. NASA/GSFC Code 612 Tech. Doc., 48 pp., http://pmm.nasa.gov/sites/default/files/document_files/IMERG_doc.pdf.
- 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. Hydrometeorol.*, **5**, 487–503, [https://doi.org/10.1175/1525-7541\(2004\)005<0487:CAMTPG>2.0.CO;2](https://doi.org/10.1175/1525-7541(2004)005<0487:CAMTPG>2.0.CO;2).
- Kubota, T., and Coauthors, 2007: Global precipitation map using satellite-borne microwave radiometers by the GSMaP project: Production and validation. *IEEE Trans. Geosci. Remote Sens.*, **45**, 2259–2275, <https://doi.org/10.1109/TGRS.2007.895337>.
- Mega, T., T. Ushio, T. Kubota, M. Kachi, K. Aonashi, and S. Shige, 2014: Gauge adjusted global satellite mapping of precipitation (GSMaP_Gauge). *XXXIth URSI General Assembly and Scientific Symp.*, Beijing, China, IEEE, 1–4, <https://doi.org/10.1109/URSIGASS.2014.6929683>.
- Milewski, A., R. Elkadiri, and M. Durham, 2015: Assessment and comparison of TMPA satellite precipitation products in varying climatic and topographic regimes in Morocco. *Remote Sens.*, **7**, 5697–5717, <https://doi.org/10.3390/rs70505697>.
- Nikolopoulos, E. I., N. S. Bartsotas, E. N. Anagnostou, and G. Kallos, 2015: Using high-resolution numerical weather forecasts to improve remotely sensed rainfall estimates: The case of the 2013 Colorado flash flood. *J. Hydrometeorol.*, **16**, 1742–1751, <https://doi.org/10.1175/JHM-D-14-0207.1>.
- Shige, S., S. Kida, H. Ashiwake, T. Kubota, and K. Aonashi, 2013: Improvement of TMI rain retrievals in mountainous areas. *J. Appl. Meteor. Climatol.*, **52**, 242–254, <https://doi.org/10.1175/JAMC-D-12-074.1>.
- Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., <https://doi.org/10.5065/D68S4MVH>.
- Sorooshian, S., K.-L. Hsu, X. Gao, H. V. Gupta, B. Imam, and D. Braithwaite, 2000: Evaluation of PERSIANN system satellite-based estimates of tropical rainfall. *Bull. Amer. Meteor. Soc.*, **81**, 2035–2046, [https://doi.org/10.1175/1520-0477\(2000\)081<2035:EOPSS>2.3.CO;2](https://doi.org/10.1175/1520-0477(2000)081<2035:EOPSS>2.3.CO;2).
- Stampoulis, D., and E. N. Anagnostou, 2012: Evaluation of global satellite rainfall products over continental Europe. *J. Hydrometeorol.*, **13**, 588–603, <https://doi.org/10.1175/JHM-D-11-086.1>.
- Tian, Y., C. D. Peters-Lidard, R. F. Adler, T. Kubota, and T. Ushio, 2010: Evaluation of GSMaP precipitation estimates over the contiguous United States. *J. Hydrometeorol.*, **11**, 566–574, <https://doi.org/10.1175/2009JHM1190.1>.
- Tong, K., F. Su, D. Yang, L. Zhang, and Z. Hao, 2014: Tibetan Plateau precipitation as depicted by gauge observations, re-analyses and satellite retrievals. *Int. J. Climatol.*, **34**, 265–285, <https://doi.org/10.1002/joc.3682>.
- Ushio, T., and Coauthors, 2009: A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. *J. Meteor. Soc. Jpn. Ser. II*, **87A**, 137–151, <https://doi.org/10.2151/jmsj.87A.137>.
- Xie, P., and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Meteor. Soc.*, **78**, 2539–2558, [https://doi.org/10.1175/1520-0477\(1997\)078<2539:GPAYMA>2.0.CO;2](https://doi.org/10.1175/1520-0477(1997)078<2539:GPAYMA>2.0.CO;2).
- Yamamoto, M. K., and S. Shige, 2015: Implementation of an orographic/nonorographic rainfall classification scheme in the GSMaP algorithm for microwave radiometers. *Atmos. Res.*, **163**, 36–47, <https://doi.org/10.1016/j.atmosres.2014.07.024>.
- , —, C.-K. Yu, and L.-W. Cheng, 2017: Further improvement of the heavy orographic rainfall retrievals in the GSMaP algorithm for microwave radiometers. *J. Appl. Meteor. Climatol.*, **56**, 2607–2619, <https://doi.org/10.1175/JAMC-D-16-0332.1>.
- Yu, P.-S., T.-C. Yang, and C.-C. Kuo, 2006: Evaluating long-term trends in annual and seasonal precipitation in Taiwan. *Water Resour. Manage.*, **20**, 1007–1023, <https://doi.org/10.1007/s11269-006-9020-8>.
- Zhang, X., E. N. Anagnostou, M. Frediani, S. Solomos, and G. Kallos, 2013: Using NWP simulations in satellite rainfall estimation of heavy precipitation events over mountainous areas. *J. Hydrometeorol.*, **14**, 1844–1858, <https://doi.org/10.1175/JHM-D-12-0174.1>.
- , —, and H. Vergara, 2016: Hydrologic evaluation of NWP-adjusted CMORPH estimates of hurricane-induced precipitation in the southern Appalachians. *J. Hydrometeorol.*, **17**, 1087–1099, <https://doi.org/10.1175/JHM-D-15-0088.1>.