Assessment of Satellite Precipitation Products in Relation with Orographic Enhancement Over the Western United States

Abishek Adhikari and Ali Behrangi

Department of Hydrology and Atmospheric sciences, University of Arizona, Tucson

Submitted to: *Journal of Earth and Space Science*

Key words: Orographic precipitation, satellite remote sensing, passive microwave precipitation retrieval, precipitation products over the western United States

Corresponding author address: Abishek Adhikari, Department of Hydrology and Atmospheric sciences, University of Arizona, Tucson

Email: abi.adk@gmail.com

Abstract:

Different precipitation products are assessed for their skill in capturing orographic precipitation over the western United States using two popular methods used to delineate orographic precipitation events. The first method defines orographic indices using orographic enhancement and moisture content that represents the amount of moisture advected over sloping terrain, whereas the second method classifies precipitation events into orographic and non-orographic events based on the orographic enhancement and moisture flux convergence. The precipitation products studied here include Dual-frequency Precipitation Radar (DPR), combined radar and radiometer (COMBINE), The Global Precipitation Measurement (GPM) microwave Imager (GMI), Microwave Humidity sounder (MHS), infrared precipitation of the Integrated Multi-satellite Retrievals for GPM (IMERG-IR), and a reanalysis product (ERA5). NCEP Stage-IV product used as a reference. All of the evaluated products show more significant errors for the orographic than the non-orographic events. DPR, COMBINE, MHS, and GMI severely underestimates the precipitation rates, especially for heavy precipitation ($>$ 4 mm/day), whereas IMERG-IR, and ERA5 show relatively better estimation. Satellite products tend to show lower fraction of precipitation occurrence and amount for orographic than no-orographic classes. This is more sever for radar only product and less sever for IR estimates. It was found that rate BIAS varies with seasons, so in cold seasons satellite precipitation products tend to underestimate while in warm season they (except DPR) tend to overestimate precipitation amount. A fairly clear relationship between the BIAS and the surface temperatures and total precipitable water (TPW) is observed. Most of the satellite products severely underestimate precipitation volume at relatively colder surfaces ($<$ 10 ^{0}C) and lower TPW ($<$15mm), but ERA5 shows little rate BIAS in such cases. The underestimation tends to be larger for orographic than non-orographic events. In contrast, ERA5 shows relatively large underestimation at warmer temperatures ($>$20 ^{0}C), where satellite products tend to overestimate precipitation amount. Overall, none of the

methods used to identify orographic events is found robust, suggesting that there is a clear need for determining a more effective method to better delineate orographic from non-orographic precipitation events.

1. Introduction

Satellite-based precipitation estimates have been widely used in many research and application areas, including short-term weather and long-term climate prediction (Arkin and Ardanuy 1989; Huffman et al., 1995). Precipitation estimates from satellite measurements are unique due to their global coverage, including oceanic and high terrain mountainous regions where ground-based observations are not always possible (Kidd et al., 2017). It is difficult to get broad spatial coverage in mountainous areas using rain gauges, and ground-based radar observation has limitations (Sapiano and Arkin 2009). Despite the great coverage of satellite products, their precipitation retrieval accuracy in mountain regions is still a challenge (Shige et al. 2013; Mei et al., 2014). The primary types of sensors used to estimate precipitation from satellite measurements are radar, passive microwave imager/sounder, and Infrared. The space-borne radar onboard satellites such as the Tropical Rainfall Measuring Mission (TRMM; Kummerow et al., 1998) and the Global Precipitation Measurement (GPM; Hou et al., 2014; Skofronick-Jackson et al., 2018) have shown success in measuring precipitations globally, but they still suffer from uncertainties especially in mountainous regions (Adhikari et al., 2019). One major source of uncertainty associated with space-borne radars is the contamination of near-surface reflectivity profiles due to ground-clutter (Arulraj and Barros 2019, Liao et al., 2014). The Passive microwave sensors are popular and have been widely used in estimating precipitation for decades (Wilheit 1986; Prigent 2010). The benefit of using PMW sensors compared to radar is their higher sampling rate, although they can misclassify rainfall over mountainous regions (Yamamoto and Shige, 2015). For example, snow and ice-covered surfaces over the mountains could be classified wrongly as precipitating clouds resulting in an overestimation of precipitation. The PMW precipitation estimates mainly rely on the cumulative signal from vertical column of cloud, hydrometeors, and ice particles and their emission or scattering properties. The heavy rainfall associated with mountainous regions can be generated from clouds with no or little ice aloft, resulting in underestimation in PMW estimates (Dinku et al. 2010; Houze 2012; Shige et al. 2013). Studies have shown that satellite-based precipitation products underestimate precipitation rate over several mountainous regions of the globe such as Japan (Kubota et al. 2009; Shige et al. 2013), Africa and South America (Dinku et al. 2010), Taiwan (Chen et al. 2013), and Vietnam (Ngo-Duc et al. 2013). They stated that satellite-based passive microwave retrieval underestimates heavy precipitation that is associated with shallow orographic systems because the algorithm assumes that only the deep systems, with large scattering signals, are responsible for heavy precipitation. However, shallow orographic systems can produce heavy rainfall due to orographic enhancement with no major scattering signals from ice particles (Takeda and Takase 1980, Sakakibara 1981).

Another sensor used in precipitation estimation is Infrared (IR), which is used to estimate precipitation from cloud top temperatures (Arkin and Meisner 1987). Because intense orographically enhanced precipitation over mountains may be produced by shallow clouds with little ice aloft, these systems may appear relatively warm in IR images. Therefore, IR-based algorithms may miss or significantly underestimate precipitation events over these regions (Hong et al. 2007, Bitew and Gebremichael 2010). Negri and Alder (1993) demonstrated that the IR retrieval methods are generally inadequate over the mountainous regions, especially those associated with shallow orographic rain systems. Therefore, topography can have a massive impact on sensor retrievals, including radar, PMW, and IR.

Several previous studies made efforts to improve precipitation retrievals over land, especially over mountainous regions. The correction techniques are necessary because the ground-based radar or rain gauge measurements are limited and not always possible over mountainous areas. Vicente et al. (2002) and Scofield and Kuligowski, (2003) have developed an IR based correction technique, whereas Kwon et al. (2008) proposed topographic correction factors based on the terrain slope, low-level wind, and moisture, which later implemented in the Goddard Profiling (GPROF) algorithm (Kummerow et al., 2001, Wang et al., 2009, Kummerow et al., 2015). Shige et al. (2013) classified orographic and non-orographic precipitation areas using terrain slope, low-level wind and moisture and developed a correction technique using the dynamic selection of lookup table (LUTS). The LUTS are used to correct orographic precipitation retrievals in the Global Satellite Mapping of Precipitation products (GSMaP; Ushio et al., 2009) over the Kii Peninsula of Japan (Shige et al., 2013), Taiwan (Taniguchi et al., 2013), Indian Subcontinent (Shige et al., 2014), and North America (Yamamoto and Shige, 2014). The method was further improved by Yamamoto et al. (2017). Although these correction techniques are shown helpful to enhance orographic precipitation over various regions mentioned above, they have been less successful over the United States and Mexico (Yamamoto and Shige, 2014).

This study examines six different precipitation products and evaluates their performance in orographic events. The products are precipitation retrieval from: GPM Dual-Frequency Precipitation Radar (GPM-DPR), GPM Microwave Imager (GMI), Microwave Humidity Sounder (MHS), Infrared (IR), GPM combined radar and microwave imager (GPM-2BCMB), and a reanalysis product (ERA5). Using Stage-IV precipitation rate as reference over the western US, Fig 1 shows that while MHS and GMI indicate skill in capturing light precipitation (<1 mm/hr) compared to Stage-IV there is severe underestimation for the heavier precipitation rates (e.g., > 3 mm/hr) (Fig 1b and 1c). DPR (Fig 1a) perform better than MHS and GMI, but still significantly underestimate intense precipitation rates compared to Stage IV. IMERG-IR show a relatively better estimation than the other products, as indicated by the rate BIAS score close to 1 and intense precipitation rates captured by the two products (Fig 1d). The severe underestimation of heavy precipitation rates by most of the prod-

ucts might be associated with the orographic precipitation events. Note that all the products in Fig 1 are averaged in 0.25^0x0.25^0 longitudes and latitudes grid boxes, so the skill scores represent grid to grid comparison rather than a footprint to grid.

This study aims to evaluate precipitation detection and estimation skill of various precipitation products by focusing on orographic precipitation. The data and method used in this study are summarized in Sections 2 and 3, followed by results in Section 4 and conclusion in Section 5.

2. Data

Two years (2018-2019) of precipitation data sets over the western United States (30^0 N-50^0 N & 105^0 W-135^0 W) are evaluated with respect to orographic indices in this study. The reference data in this study is the NCEP Stage-IV precipitation product to evaluate variety of sensors such as radar (e.g., DPR), passive microwave imager (e.g., GMI), passive microwave sounder (e.g., MHS), combined PMW and radar (i.e., 2BCMB), IR, and reanalysis (i.e., ERA5) products (Table 1). Each data set is briefly introduced as follows:

2.1 Stage-IV:

The National Center for Environmental Prediction (NCEP) Stage-IV is an hourly accumulated data with 4 km spatial resolution in the polar stereographic grid (Lin, 2011). It assimilates both ground-based radar and gauge observations and is quality controlled by 12 River Forecast Centers. This product agrees well with the gauge measurements and also performs relatively well for snowfall (Zhang, 2018; Nelson et al. 2016), so this study utilizes Stage-IV as a reference product.

2.2 GPM DPR/2BCMB

DPR is on the GPM core satellite, which has a non-sun-synchronous orbit. DPR operates at Ku (13.5 GHz) and Ka (35.5 GHz) frequencies. The DPR precipitation product incorporates both Ku and Ka-bands and provides precipitation rate at each footprint with a spatial resolution of about 5 km x 5km. With the dual-frequency radar, DPR should enable estimating light rain and snowfall at minimum rate of 0.2 mm/hr (Hou et al., 2014; Skofronick-Jackson et al., 2017; Adhikari et al., 2018). The algorithm theoretical basis is discussed in Iguchi et al. (2018). The GPM also offers a combined radar-radiometer product (hereafter, "COMBINE") which has a similar resolution to DPR product, but in addition to the dual-frequency measurement, it incorporates microwave measurement (Grecu et al. 2016). It is expected that the COMBINE provides better precipitation estimation than the DPR, although studies have shown that it can miss large fraction of snowfall and precipitation in high latitudes (Skofronick-Jackson et al., 2019; Behrangi et al. 2020). The GPM Combined Radar-Radiometer Precipitation Algorithm Theoretical Basis Document is discussed in Olson et al. (2018).

2.3 GPM GMI

GMI is a passive microwave imager on the GPM core observatory. It is a multifrequency imager that operates at 13 different frequencies ranging from 10 GHz to 183 GHz (Draper et al. 2015). The microwave radiances at lower frequency channels (37 GHz or less) are primarily influenced by the emission signatures from the liquid droplets, whereas the higher frequency channels (89 GHz or higher) are more sensitive to the scattering signatures. The surface precipitation rate is estimated by the GPROF 2017 algorithm (Kummerow et al., 2015; Randel et al. 2020), which uses a Bayesian approach. The brightness temperatures are used to estimates surface precipitation rates from emission or scattering signatures of hydrometeors. Over land, which is the primary focus of this study, GPROF relies mainly on the ice- scattering signatures. The GMI level 2 orbital surface precipitation data is used in this study with approximately 13 km of spatial resolution.

2.4 NOAA 19 MHS

MHS on NOAA-19 satellite is a microwave humidity sounder that operates at five high-frequency microwave channels ranging from 89 GHz to 190 GHz. The spatial resolution of MHS is approximately 16 km near nadir and becomes larger with increasing the scan angle (Mo, 2010). The GPROF 2017 algorithm is used to estimate surface precipitation which mostly relies on the scattering signature of ice particles. The depression in the brightness temperatures due to the ice particles is linked to the precipitation rate. Because of the high-frequency channels, MHS has shown skill in retrieving snowfall (Bennartz & Bauer, 2003, Noh & Liu, 2004, Adhikari et al. 2020). MHS is available on several NOAA and MetOp satellites.

2.5 IMERG-IR

IR precipitation rate was obtained from the Integrated Multi-Satellite Retrievals for GPM (IMERG; Huffman et al. 2019) precipitation product at 0.1^{o}x0.1^{o} spatial and 30 minutes temporal resolutions. The IMERG product combines precipitation estimates from geostationary IR and low-Earth-orbit PMW sensors. IMERG-IR precipitation estimates are based on the Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN) Cloud Classification System (PERSIANN-CCS) (Hong et al. 2004) and is available as part of the IMERG output field (hereafter we refer to it as IMERG-IR).

2.6 ERA5

ERA5 product is developed by the European Centre for Medium-Range Weather Forecasts (ECMWF) using 4D-Var data assimilation techniques. ERA5 is a reanalysis product that provides various meteorological parameters at the surface and the vertical levels in each hour at 0.25^{o}x0.25^{o} longitude and latitude grid box (Hersbach et al. 2020). This study utilizes precipitation rate, 2 m (T2M) temperatures, total precipitable water (TPW), and Specific humidity at surface level, and U and V component of the wind at vertical pressure levels. The wind profile data is used to estimate "near-surface" wind components.

3. Methods:

This study utilizes 2 years (2018-2019) of precipitation data, from all of the products introduced in Section 2, over the western US. The reason to choose this time duration is that during these two years hourly Stage-IV data has little or no gap, but for other years large missing values over the western U.S. exist. All the data sets were mapped onto a common 0.25^{o}x0.25^{o} resolution to match the resolution of ERA5 that is used to calculate orographic indices. Due to the hourly resolution of Stage IV product, all the data sets are matched within an hour. Orographic rainfall indices are calculated based on two major methods described in the literature as follows:

3.1 Method1:

In this method, orographic rainfall index (ORI) is defined by the same approach used by Nieman et al., (2002, 2009) and Bikos et al. (2014). The ORI is estimated by utilizing vertically upward motion and total moisture in the given vertical column, as shown in equation 1 below.

$$ORI = TPW * V \bullet H \quad \ldots\ldots\ldots\ldots \quad (1)$$

Where TPW is the ERA5 total precipitable water (kg m^{-2}), V is the ERA5 near-surface wind (ms^{-1}), and H is the terrain elevation (m). The first expression in equation 1 (i.e., TPW) ensures that enough moisture content is in the column and the second expression provides enough upward motion. So, the ORI represents the amount of moisture advected over the sloping terrain. The ORI has a unit of kg s^{-1} $m^{-1,}$ and its values roughly range from -10 to +10. The positive ORI values indicate the upslope terrain, the negative values indicate downslope, and the values close to zero indicate relatively plain surfaces. The interest of this study is on positive values of ORI, which might be associated with orographic rain. Note that ORI does not predict precipitation; rather, it is an index that shows the amount of moisture advected over the sloping terrain (Bikos et al., 2014). The ORI is further subdivided into five categories based on its value as ORI1: ORI<-5, ORI2: ORI -5 to -0.25, ORI3: ORI -0.25 to 0.25, ORI4: ORI 0.25 to 5, and ORI5: ORI $>$ 5.

Method 2:

In this method, precipitation events are classified into orographic and non-orographic events based on conditions used in previous studies such as Shige et al. 2013, Shige et al. 2015, and Taniguchi et al. 2013. Any precipitation event is identified as an orographic if there is abundant orographically forced upward motion and moisture flux convergence. The orographically forced upward motion (w) and moisture flux convergence are estimated as shown in Equations 2 and 3.

$$w = Dh/Dt = V_H \cdot \nabla h \quad \ldots\ldots \quad (2)$$

$$Q = -\nabla \cdot V_H * q \quad \ldots\ldots \quad (3)$$

where, V_H is the horizontal surface wind, q is the water vapor mixing ratio derived from ERA5 data sets, and h is the terrain elevation. Previous studies have chosen different threshold values of w and Q for different geographic locations (See Yamamoto and Shige, 2014 for details). In this study, the threshold values of w and Q are chosen to be 0.1 ms^{-1} and $0.1x10^{-6}$ (s^{-1}), respectively. If any precipitation event satisfies both the criteria, the event is identified as orographic events otherwise non-orographic. The justification of these threshold values will be discussed in the result section.

4. Result:

The orographic precipitation events from both methods are determined, then all the products are compared with Stage-IV precipitation to explore the detection and estimation capability of the studied products over the western US.

4.1: Evaluation of precipitation products based on method1:

In this section, the detection and estimation skill of the precipitation products for orographic events are compared with Stage-IV. Note that method1 defines orographic precipitation similar to Bikos et al. (2014) method where moisture content and upward vertical motion is used to determine orographic index, as shown in equation 1. The orographic index is further subdivided into five categories, as explained in section 3.1. First, the detection statistics of each product are calculated in pixel level that are: False Alarm Rate (FAR), Hidke Skill Score (HSS), Probability of Detection (POD), and occurrence BIAS (here BIASo). HSS provides a more generalized skill score for assessing the accuracy of the predictions relative to the random chance. In other words, HSS shows the fraction of correct predictions by excluding correct predictions due to random chance. The range of the HSS is $-\infty$ to 1. Negative values indicate that the forecast by chance is better, 0 means no skill, and 1 means a perfect forecast. The ideal score for BIASo is 1. BIASo is calculated by dividing the number of estimated precipitation occurrence from each product by the corresponding value from the reference product (here Stage IV). Recall that orographic categories ORI1 and ORI5 represent extreme orographic index values for downward and upward terrain slope, respectively, and ORI3 represents a relatively plain terrain. The main interest of this study is precipitation events that are associated with upward terrain slope, which might be represented by ORI4 and ORI5. MHS shows high FAR for all ORI classes and relatively less for extreme ORIs (ORI1 and ORI5). The HSS and POD scores for extreme ORIs are relatively higher than the ORI3 (Table 2). This indicates that MHS can detect precipitation better when events are associated with orography, perhaps because they are more intense. However, the BIASo score suggests that MHS overestimates (BIASo = 1.25) Stage-IV for extreme ORI class. Similar statistics are seen for GMI, except BIASo scores are relatively low, meaning GMI underestimates precipitation when compared to Stage-IV for all ORI classes, and it underestimates more for extreme ORI (BIASo = 0.90 for ORI5). In terms of BIASo, microwave sounder (MHS) overestimates Stage-IV precipitation events, whereas microwave imager (GMI) underestimates it. However, both MHS and GMI have

comparable overall skill based on HSS.

Radar products (i.e., DPR and COMBINE) show similar statistics to the PMW products (e.g., in terms of FAR, HSS, and POD). However, DPR severely underestimates Stage-IV for all ORI classes, and the underestimation is more for extreme ORIs (BIASo = 0.69). The COMBINE product also shows underestimation for all ORI classes but relatively less in magnitude for extreme ORI class (BIASo = 0.75). While spaceborne radar may outperform PMW products in capturing orographic precipitation (Shige et al. 2013), signal clutter issues near the surface might result in missing shallow precipitation events. The IMERG-IR has generally low skill in detecting precipitation regardless of the ORI class, as can be seen by its significantly lower HSS and higher FAR compared to the other products. On the other hand, its BIASo score shows underestimation for extreme ORI class and overestimation for the other classes, suggesting that IMERG-IR is relatively less skillful than other products in detecting precipitation events in the mountainous regions. ERA5 overestimates Stage-IV precipitation for all ORI classes as indicated by the higher BIAS value (BIASo = 1.35). However, it has higher POD, lower FAR, and better overall skill (based on HSS) than other products for detecting precipitation occurrence.

The precipitation rates from all the products are also evaluated in each ORI class with Stage-IV. The Correlation coefficient (CC-value), RMSE, and rate BIAS scores (here, BIASr) are demonstrated in table 3 for all the products. CC values are fairly comparable across different OIR classes for most products, whereas RMSE scores are much larger for extreme ORI classes that for relatively plain classes (Table 3). MHS and GMI BIASr scores are close to 1 and slightly lower values (0.93 and 0.86) for extreme ORI classes, respectively. This means that MHS and GMI retrieved precipitation volume is consistent with Stage-IV with slight underestimation for orographic events. The DPR severely underestimates precipitation volume for all the ORI classes , especially for extreme ORI classes (BIASr = 0.59 for ORI1 and BIASr = 0.56 for ORI5). This indicates that GPM DPR precipitation is heavily affected by orography, and a special attention is needed to overcome the underestimation over the mountainous regions. Similar to the DPR, the GPM COMBINE product also sensitive to the orography and underestimates the total precipitation for all the ORI classes, but the underestimation is less when compared to the DPR (BIASr =0.76 for ORI5). Since COMBINE incorporates both DPR and GMI products in its retrieval algorithm, it is likely that the GMI would aid COMBINE product to prevent severe underestimation as seen by the DPR. IMERG-IR product overestimate Stage-IV for relatively plain surfaces (e.g., ORI=3), whereas in the extreme ORI class, the total precipitation from both products is almost close to the reference. While IMERG-IR shows relatively good BIASr, it has the lowest CC-value among all other products, suggesting that the relationship between cloud-top brightness temperature information and precipitation rate is poor across all ORIs. ERA5 overestimates Stage-IV in each ORI class, but shows slightly higher CC-values and lower RMSE compared to most of the products studied here (Table 3). However, RMSE of ERA5 is still higher for extreme ORI classes than relatively

plain surfaces.

Figure 2 shows mean precipitation rates for all the products as a function of ORI values. Note that each product is separately matched up with Stage-IV, so each product is evaluated based on its respective matched Stage-IV (Fig 2). For example, the green dashed and solid lines in figure 2 represent MHS and Stage-IV that is matched with MHS precipitation rate, respectively. Fig 2a shows mean precipitation rates for the cold season (October to April) only. It is seen that all the sensor products except IMERG-IR and ERA5 underestimate precipitation in each ORI index. The underestimation in PMW and radar estimation might be associated with the shallow clouds and snow systems that are common in the western US winter months (Creamean et al., 2013). The IMERG-IR product seems to have higher skills than the PWM and radar products in the cold seasons, especially in orographic conditions. During the warm season (May to September) (Fig 2b), the difference between the sensor product and the reference is smaller than cold season. IMERG-IR overestimates the Stage-IV, whereas DPR underestimates it during the warm season. It is also seen from the figure that the mean precipitation rates are higher for extreme ORIs and lower if the ORIs are close to 0. To understand ORI dependencies of sensor products, BIASr is calculated as a function of ORI indices for cold and warm seasons separately and shown in Fig 3a and 3b, respectively. Note that BIASr is estimated by dividing total precipitation from the sensor to the matched Stage-IV in each ORI bin. During the cold season, BIASr scores of IMERG-IR and ERA5 are close to 1 or higher (e.g., BIASr scores tend to higher than 1 for ORIs between -5 and 2.5 and less than 1 for extreme ORIs). The rest of the other products clearly underestimates Stage-IV and show smaller ORI dependency (Fig 3a). During the warm season, all the sensor products except the DPR and ERA5 overestimates Stage-IV. IMERG-IR has the highest values of BIASr ($> =1.2$) for all the ORIs, whereas DPR shows 0.8 or lower BIASr scores. Overall, it can be seen that the dependency of the BIASr to ORIs is fairly week, but generally a smaller fraction of precipitation is captured at positive ORIs (upslope) compared to the central and negative ORIs (downslope).

4.1: Evaluation of precipitation products based on method 2:

The orographic precipitation index using method 1 was useful to compare precipitation products, but no strong ORI dependency was observed. Here method 2 is used to define orographic events, where any precipitation event is defined as the orographic precipitation (or "ORO") if it has enough moisture flux convergence (> 0.1 ms^{-1}) and orographic enhancement ($Q>0.1 \times 10^{-6}$ s^{-1}) otherwise non-orographic (or "NO-ORO") as shown in the equation 2 and 3. Choosing the correct threshold values for moisture flux convergence and orographic enhancement is not straightforward and has been defined differently in past studies (Yamamoto and Shige, 2014). The reason to choose the above-mentioned threshold values is demonstrated in Figure 4 that shows the BIASr as a function of w and Q for all the precipitation products. BIASr is defined as the ratio of total precipitation from sensor product to Stage-IV. Fig 4a shows DPR's BIASr as

a function of w and Q and black solid cross lines are the threshold lines that have been chosen in this study. We can see that above the threshold values, DPR significantly underestimates the Stage-IV (BIASr values up to 0.1). The BIASr score is close to 1 in the region where there is not enough uplifting and moisture flux convergence. These thresholds are fairly consistent with the previous studies (e.g., Shige et al., 2013; Taniguchi et al., 2013; Shige et al., 2014; Yamamoto and Shige, 2014), thus here we use them for separating orographic and non-orographic precipitation events. Although BIASes are observed in the other quadrant, our region of interest is the top-right quadrant ("orographic quadrant" or "orographic region"), where both orographic uplifting and abundant moistures are present (Fig 4). Similar patterns are observed in the other products in the orographic region. The MHS and GMI show significant underestimation (BIASr < 0.5) in the entire orographic quadrant, whereas slightly less underestimation is seen in IMERG-IR and ERA5 products.

Fig 5 shows precipitation occurrence BIAS (BIASo) in all of the studied products as a function of Q and w. DPR and COMBINE heavily underestimate the precipitation events almost everywhere, but they show slightly better detection for higher values of moisture flux convergence and orographic enhancement (fig 5a and b). MHS shows BIASo values close to 1, almost everywhere including the orographic quadrant. Note that the higher BIASo of MHS comes with higher FAR as can be seen in Table 2, suggesting that the improved precipitation detection is at the expense of larger false detection. GMI and IMERG-IR tend to underestimate in all quadrants, but show BIASo values near 1 in regions where both w and Q are close to 0 (Figs 5d and 5e). ERA5 slightly overestimates in all quadrants, including the orographic quadrant. However, the overestimation is more along the region where the orographic enhancement term (Q) is close to 0. Overall, despite that observed for precipitation amount (Fig 4), the second method does not show a clear difference between orographic and no-orographic quadrants for precipitation occurrence BIAS.

Because method 2 does the binary classifications of the precipitation events i.e., either the precipitation event is an orographic ("ORO") or a non-orographic (NO-ORO"), all the precipitation products are classified into these two categories and compared with their respective Stage-IV match up. The primary focus of the evaluation is given to the "ORO" class, which is the primary interest of this study. Table 4 lists the detection statistics of the sensors, which includes FAR, HSS, POD, and BIASo for both ORI classes. The FAR scores of the ORO class are slightly better (less) than the NO-ORO class for all the sensor products. A similar pattern is also observed with HSS and POD scores, where all products have better detection statistics for the ORO class than the N0-ORO. Although the "ORO" class has better HSS and POD scores, there are severe underestimations observed in few products. MHS, GMI, DPR, and COMBINE show underestimation in both classes, and the underestimation is more in orographic events. For example, MHS and GMI have BIASo scores of 0.67 and 0.73, which are less when compared to N0-ORO class 0.80 and 0.83, respectively. DPR and COMBINE also show severe underestimation in the ORO class when

compared with their respective N0-ORO classes (Table 4). IMERG-IR has a relatively better BIASo scores compared to other satellite products, but still underestimates Stage-IV for ORO class as indicated by lower BIASo score 0.75. A different pattern in the BIASo scores is seen with the ERA5 product, where ERA5 overestimates both ORO and N0-ORO classes with the BIASo values of 1.19 and 1.25, respectively (Table 4).

The sensor products are further evaluated by plotting precipitation occurrences in both N0-ORO and ORO classes and compared with the respective Stage-IV match-up. Fig 6 shows histogram of occurrences that is calculated by dividing precipitation samples by total samples (normalized), so that each curve adds up to 1. For ORI0 class, both DPR and COMBINE precipitation occurrences are close to Stage-IV with slight underestimation for higher precipitation rates ($>$ 1 mm/hr). On the other hand, DPR and COMBINE ORO class shows severe underestimation when precipitation rates are greater than 1 mm/hr. Interestingly, DPR and COMBINE overestimate Stage-IV for both NO-ORO and ORO classes when the precipitation rate is 0.3 to 1 mm/hr (Fig 6 a and b). A similar pattern is observed in MHS and GMI, where the underestimation is observed in higher precipitation rates (>0.8 mm/hr), and overestimation is seen for the lower end ($<$ 0.3 m/hr). The occurrence difference is more for ORO class when compared to NO-ORO for MHS and GMI products (Fig. 6c and d). IMERG-IR precipitation occurrences for both N0-ORO and ORO are close to their respective Stage-IV match-up and tend to have higher precipitation occurrence in the ORO class for extreme precipitation rates ($>$ 9 mm/hr) (Fig 6e). Overall, the histograms of ORO and NO-ORO are more distinct for Stage-IV than the studied sensors, suggesting that Stage-IV better captures orographic precipitation enhancement than sensors. In other words, there is considerably higher occurrence of intense precipitation for the ORO class than NO-ORO class for Stage IV compared to that observed for different sensors. The histograms of precipitation from GMI (Fig 6d) shows relatively smaller difference between ORO and NO-ORO compared to the other sensors and ERA5 (Fig 6f), suggesting that GMI precipitation product has little skill in capturing the orographic precipitation enhancement.

The precipitation products are further assessed by comparing precipitation rates in both NO-ORO and ORO classes. Table 5 summarizes the precipitation estimation statistics where CC-value, RMSE, and BIASr scores of both N0-ORO and ORO classes are presented. All of the satellite products have relatively low CC-value (~0.3) and show only slightly higher CC for ORI than NO-ORI classes. The IMERG-IR correlates the least (CC = 0.20) with Stage-IV for the orographic events, whereas ERA5 correlates the most (CC = 0.54). In terms of BIASr, MHS, GMI, DPR, and COMBINE underestimate in both ORI classes. The BIASr score for an orographic event is much less than the non-orographic. For example, MHS and GMI BIASr for the orographic class are 0.63 and 0.79, whereas the same for non-orographic events are 0.86 and 0.93, respectively (Table 5). The DPR product significantly underestimates orographic precipitation rates, as indicated by the low value of bias (BIASr=0.49). IMERG-IR and ERA5

show BIASr>1 for N0-ORO, but for ORO, ERA5 has BIASr=1 and IMEG-IR slightly underestimates Stage-IV (BIASr = 0.86) for orographic events (Table 5).

Figure 7 shows precipitation amount contribution at each intensity range for each precipitation product and its corresponding Stage-IV match up. Note that for each product the contribution is calculated by multiplying samples counts in each bin by its intensity and dividing precipitation volume in each precipitation rate bin by the total precipitation volume (normalization). Fig 7a shows DPR and Stage-IV precipitation contribution for N0-OROand ORO classes. DPR shows double peaks for orographic events near 0.5 mm/hr and 8 mm/hr bins, whereas Stage-IV has a single peak near the 3 mm/hr bin. For N0-ORO events, DPR peaks near 0.5 mm/hr bin whereas Stage-IV peaks at or near 1 mm/hr, and both contribute approximately 3% to the total precipitation. COMBINE shows a similar pattern to DPR (Fig 7b). The contribution pattern of the two PMW products (MHS and GMI) for orographic events are also similar to each other (but different from DPR and COMBINE products). GMI and MHS contribute more at lower precipitation rates (< 1 mm/hr) and Stage-IV contributes more at higher precipitation rates (>1 mm/hr) (Fig 7c, d). In contrast, IMERG-IR shows higher contribution of intense precipitation than Stage-IV for both orographic and non-orographic classes (Fig 7e). Overall, based on Figure 7, Stage–IV shows more contribution from intense precipitation in ORO class than N0-ORO, but this is not necessarily the case in the satellite products. This suggests that satellite products generally have difficulties in capturing the orographic enhancement of precipitation rate.

The precipitation estimation BIAS (BIASr) is further evaluated based on the ERA5 total precipitable water ("TPW") and surface 2 m temperatures ("T2M"). Fig 8 shows BIASr in each T2M and TPW bins for both orographic and non-orographic precipitation events. Note that the BIASr is estimated by dividing sensor precipitation volume by Stage-IV precipitation volume, separately for each bin. For the orographic class, DPR underestimates Stage-IV for most of the T2M and TPW bins. There is severe underestimation (BIASr < 0.3) for the relatively colder surface (< 10 ^{0}C) and drier environment (TPW < 15 mm) (Fig 8a) that might be related to shallow cloud systems. For the non-orographic class, DPR shows a similar pattern of underestimation for colder surfaces but for the warmer surface (T2M > 20 ^{0}C), DPR overestimates Stage-IV (Fig 8b). COMBINE shows a similar BIASr pattern to DPR (Fig. 8c and 8d). GMI, MHS, and IMERG-IR also underestimate Stage-IV precipitation when the surface is colder than 10 ^{0}C and overestimates above it. The underestimation tends to be larger for orographic than non-orographic events. When the surface is warmer, it is expected to have relatively deeper cloud systems with large vertical ice column (Rabin et al., 1990). In such cases almost all the sensor products tend to overestimate Stage-IV precipitation for both orographic and non-orographic events (Fig. 8e-8j). However, ERA5 stands out among the others with a different BIASr pattern. For relatively colder surfaces (<20 ^{0}C), ERA5 precipitation is often consistent with Stage-IV as indicated by a BIASr score close to 1. This

indicates that the ERA5 is best among the others, especially for colder surfaces. However, for warmer surfaces (e.g., T2M>25 ^{0}C) with TPW> 10mm, ERA5 severely underestimates Stage-IV for both orographic and non-orographic events. This is an opposite pattern than all other sensor products, which calls for further investigations into the differences between satellite and reanalysis products.

5. Conclusion

Two years (2018-2019) of various precipitation products over the western US are evaluated using Stage-IV as a reference and through separating orographic and non-orographic precipitation events. The precipitation products include variety of sources such as radar, passive microwave imager, passive microwave sounder, infrared sensors, and reanalysis products. The precipitation products are categorized and evaluated based on two popular types of orographic precipitation classification schemes. The first method utilizes moisture and orographic enhancement to estimate orographic indexes, and all the products are assessed based on the orographic class. In the second method, precipitation products are classified into orographic and non-orographic based on the orographic enhancement and moisture flux convergences and by applying thresholds to them. Both precipitation occurrence and estimation statistics are evaluated for the two methods, and major findings are summarized as follows:

Method 1 suggests that the IMERG-IR and ERA5 overestimates, and GMI, DPR, COMBINE, and MHS underestimate Stage-IV precipitation rates in almost all ORI classes, but for the extreme orographic classes the ratio of precipitation amount from products over Stage-IV is smaller than that from classes associated with no or moderate orographic precipitation events. All the products show more skill in capturing precipitation occurrence (but with higher RMSE) for extreme orographic classes than no or moderate orographic classes. With respect to season, all of the products (except IMERG-IR and ERA5) show large underestimation in cold season. In contrast, for warm season, the products (except ERA5 and DPR) tend to overestimate precipitation rate, DPR shows larger underestimation than ERA5. Although the fraction of precipitation captured by products tend to be smaller at positive ORIs (upslope) than negative ORIs (downslope), BIASr patterns are not found very much sensitive to the orographic indexes.

Method 2 shows that all sensor products except ERA5 tend to underestimate the Stage-IV precipitation occurrence. The underestimation is more on orographic events than the non-orographic. Among the products, DPR and COMBINE severely underestimate orographic precipitation occurrence as indicated by the relatively low BIASo values of 0.57 and 0.61, respectively. The orographic precipitation events tend to have higher precipitation rates than the non-orographic events. MHS and GMI orographic-precipitation-occurrence histograms are different from the Stage-IV, where MHS and GMI precipitation events occur less frequently at higher intensities (>1 mm/hr) and more frequently at lower rates (< 1mm/hr). DPR and COMBINE orographic precipitation occurrences are somewhat consistent with Stage-IV. In contrast, IMERG-IR shows higher fre-

quency of occurrence than Stage-IV at the very intense tail of the histogram. Overall, the histograms of orographic and non-orographic precipitation occurrence and volume are more distinct for Stage-IV than the studied sensors, suggesting that Stage-IV is likely more skillful in capturing orographic precipitation enhancement than sensor products.

The joint analysis of moisture flux convergence and orographic enhancement (Fig. 8) shows that satellite products present severe underestimation when the environment is dry (i.e., TPW is low) and surface temperatures is colder than about 10 ^{0}C. The underestimation tends to be larger for orographic than non-orographic events. For warmer surfaces and higher TPW, satellite products tend to overestimate Stage-IV for both orographic and non-orographic events. However, ERA5 shows an opposite pattern compared to the other product for both orographic and non-orographic events. ERA5 heavily underestimates Stage-IV precipitation when the environment is relatively humid (e.g., TPW>10mm) and warm (e.g., T2M greater than 25 ^{0}C but less than about 40 ^{0}C).

In sum, all precipitation products show difficulty in detecting and estimating orographic precipitation events. From the two methods of defining orographic precipitation the first method showed relatively week dependency between ORI and orographic precipitation enhancement. The second method showed clearer signals. Besides improvements in satellite precipitation products for capturing orographic precipitation enhancement, there seems to be a need for developing more effective methods to delineate orographic precipitation events. This is critical not only for evaluation of precipitation products, but also for improving retrieval algorithms.

Acknowledgment:

The NCEP-STAGE-IV product was obtained from Earth Observing Laboratory (EOL) website at https://data.eol.ucar.edu/. NOAA-19 MHS, GPM, and IMERG products were obtained from the Goddard Earth Sciences Data and Information Services Center (GES DISC, https://disc.gsfc.nasa.gov/). ERA5 data is downloaded from the Climate Data Store (https://cds.climate.copernicus.eu/). Financial support was made available from NASA MEaSUREs (NNH17ZDA001N-MEASURES) and NASA Weather and Atmospheric Dynamics (NNH19ZDA001N-ATDM) grants.

References:

Adhikari, A., Liu, C., & Hayden, L. (2019). Uncertainties of GPM Microwave Imager Precipitation Estimates Related to Precipitation System Size and Intensity. *Journal of Hydrometeorology*, *20*(9), 1907-1923.

Adhikari, A., Liu, C., & Kulie, M. S. (2018). Global distribution of snow precipitation features and their properties from 3 years of GPM observations. *Journal of Climate*, *31*(10), 3731-3754.

Adhikari, A., Ehsani, M. R., Song, Y., & Behrangi, A. (2020). Comparative assessment of snowfall retrieval from Microwave Humidity Sounders using machine learning methods. *Earth and Space Science*, *7*(11), e2020EA001357.

Arkin, P. A., & Meisner, B. N. (1987). The relationship between large-scale convective rainfall and cold cloud over the western hemisphere during 1982-84. *Monthly Weather Review*, *115*(1), 51-74.

Arkin, P. A., & Ardanuy, P. E. (1989). Estimating climatic-scale precipitation from space: A review. *Journal of Climate*, *2*(11), 1229-1238.

Arulraj, M., & Barros, A. P. (2019). Improving quantitative precipitation estimates in mountainous regions by modelling low-level seeder-feeder interactions constrained by Global Precipitation Measurement Dual-frequency Precipitation Radar measurements. *Remote Sensing of Environment*, *231*, 111213.

Behrangi, A., and Y. Song (2020), A new estimate for oceanic precipitation amount and distribution using complementary precipitation observations from space and comparison with GPCP, Environmental Research Letters, 15(12), 124042, doi: 10.1088/1748-9326/abc6d1.

Behrangi, A., Lebsock, M., Wong, S., & Lambrigtsen, B. (2012). On the quantification of oceanic rainfall using spaceborne sensors. *Journal of Geophysical Research: Atmospheres*, *117*(D20).

Bennartz, R., & Bauer, P. (2003). Sensitivity of microwave radiances at 85–183 GHz to precipitating ice particles. *Radio Science*, *38*(4).

Bitew, M. M., & Gebremichael, M. (2010). Evaluation through independent measurements: Complex terrain and humid tropical region in Ethiopia. In *Satellite rainfall applications for surface hydrology* (pp. 205-214). Springer, Dordrecht.

Chen, S., Hong, Y., Cao, Q., Kirstetter, P. E., Gourley, J. J., Qi, Y., ... & Wang, J. (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?. *Journal of hydrology*, *506*, 4-13.

Conner, M. D., & Petty, G. W. (1998). Validation and intercomparison of SSM/I rain-rate retrieval methods over the continental United States. *Journal of Applied Meteorology*, *37*(7), 679-700.

Creamean, J. M., Suski, K. J., Rosenfeld, D., Cazorla, A., DeMott, P. J., Sullivan, R. C., ... & Prather, K. A. (2013). Dust and biological aerosols from the Sahara and Asia influence precipitation in the western US. *science*, *339*(6127), 1572-1578.

Dinku, T., Connor, S. J., & Ceccato, P. (2010). Comparison of CMORPH and TRMM-3B42 over mountainous regions of Africa and South America. In *Satellite rainfall applications for surface hydrology* (pp. 193-204). Springer, Dordrecht.

Draper, D. W., Newell, D. A., Wentz, F. J., Krimchansky, S., & Skofronick-Jackson, G. M. (2015). The global precipitation measurement (GPM) microwave imager (GMI): Instrument overview and early on-orbit performance. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *8*(7), 3452-3462.

Grecu, M., Olson, W. S., Munchak, S. J., Ringerud, S., Liao, L., Haddad, Z., ... & McLaughlin, S. F. (2016). The GPM combined algorithm. *Journal of Atmospheric and Oceanic Technology*, *33*(10), 2225-2245.

Hersbach, H., Bell, B., Berrisford, P., Hirahara, S., Horányi, A., Muñoz-Sabater, J., ... & Thépaut, J. N. (2020). The ERA5 global reanalysis. *Quarterly Journal of the Royal Meteorological Society*, *146*(730), 1999-2049.

Hong, Y., Adler, R. F., Negri, A., & Huffman, G. J. (2007). Flood and landslide applications of near real-time satellite rainfall products. *Natural Hazards*, *43*(2), 285-294.

Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., ... & Iguchi, T. (2014). The global precipitation measurement mission. *Bulletin of the American Meteorological Society*, *95*(5), 701-722.

Iguchi, T. et al. (2018), GPM/DPR Level-2Algorithm Theoretical Basis Document, available online at: https://arthurhou.pps.eosdis.nasa.gov/Documents/ATBD_DPR_201811_with_Ap

Houze Jr, R. A. (2012). Orographic effects on precipitating clouds. *Reviews of Geophysics*, *50*(1).

Huffman G J, Adler R F, Bolvin D T, Hsu K, Kidd C, Nelkin E J, Tan J and Xie P 2019 Algorithm Theoretical Basis Document (ATBD) Version 06, NASA Global Precipitation Measurement (GPM)Integrated Multi-satellitE Retrievals for GPM (IMERG) (available at: https://gpm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_ V06.pdf)

Huffman, G. J., Adler, R. F., Rudolf, B., Schneider, U., & Keehn, P. R. (1995). Global precipitation estimates based on a technique for combining satellite-based estimates, rain gauge analysis, and NWP model precipitation information. *Journal of Climate*, *8*(5), 1284-1295.

Kubota, T. Ushio, S. Shige, S. Kida, M. Kachi, and K. Okamoto, 2009: Verification of high resolution satellite-based rainfall estimates around Japan using gauge-calibrated ground radar dataset. J. Meteor. Soc. Japan, 87A, 203–222

Kuligowski, R. J., 2002: A self-calibrating real-time GOES rainfall algorithm for short-term rainfall estimates. J. Hydrometeor., 3, 112–130.

Kummerow, C. D., Randel, D. L., Kulie, M., Wang, N. Y., Ferraro, R., Joseph Munchak, S., & Petkovic, V. (2015). The evolution of the Goddard profiling algorithm to a fully parametric scheme. *Journal of atmospheric and oceanic technology*, *32*(12), 2265-2280.

Kummerow, C., Barnes, W., Kozu, T., Shiue, J., & Simpson, J. (1998). The tropical rainfall measuring mission (TRMM) sensor package. *Journal of atmospheric and oceanic technology*, *15*(3), 809-817.

Kummerow, C., et al. (2001), The evolution of the Goddard profiling algorithm (GPROF) for rainfall estimation from passive microwave sensors, J. Appl. Meteorol., 40, 1801–1820.

Kwon, E. H., Sohn, B. J., Chang, D. E., Ahn, M. H., & Yang, S. (2008). Use of numerical forecasts for improving TMI rain retrievals over the mountainous area in Korea. *Journal of applied meteorology and climatology*, *47*(7), 1995-2007.

Liao, L., Meneghini, R., & Tokay, A. (2014). Uncertainties of GPM DPR rain estimates caused by DSD parameterizations. *Journal of Applied Meteorology and Climatology*, *53*(11), 2524-2537.

Lin, X., & Hou, A. Y. (2008). Evaluation of coincident passive microwave rainfall estimates using TRMM PR and ground measurements as references. *Journal of Applied Meteorology and Climatology*, *47*(12), 3170-3187.

Lin, Y. (2011). GCIP/EOP Surface: Precipitation NCEP/EMC 4KM Gridded Data (GRIB) Stage IV Data, Version 1.0. UCAR/NCAR-Earth

Maggioni, V., Meyers, P. C., & Robinson, M. D. (2016). A review of merged high-resolution satellite precipitation product accuracy during the Tropical Rainfall Measuring Mission (TRMM) era. *Journal of Hydrometeorology*, *17*(4), 1101-1117.

Mei, Y., Anagnostou, E. N., Nikolopoulos, E. I., & Borga, M. (2014). Error analysis of satellite precipitation products in mountainous basins. *Journal of Hydrometeorology*, *15*(5), 1778-1793.

Mo, T. (2010). Postlaunch calibration of the NOAA-19 advanced microwave sounding unit-A. *Journal of Geophysical Research: Atmospheres*, *115*(D8).

Negri, A. J., and R. F. Adler, 1993: An intercomparison of three satellite infrared rainfall techniques over Japan and surrounding waters. J. Appl. Meteor., 32, 357–373.

Neiman, P. J., Ralph, F. M., White, A. B., Kingsmill, D. E., & Persson, P. O. G. (2002). The statistical relationship between upslope flow and rainfall in California's coastal mountains: Observations during CALJET. *Monthly Weather Review*, *130*(6), 1468-1492.

Neiman, P. J., White, A. B., Ralph, F. M., Gottas, D. J., & Gutman, S. I. (2009, April). A water vapour flux tool for precipitation forecasting. In *Proceedings of the Institution of Civil Engineers-Water Management* (Vol. 162, No. 2, pp. 83-94). Thomas Telford Ltd.

Nelson, B. R., O. P. Prat, D.-J. Seo, and E. Habib (2016), Assessment and Implications of NCEP Stage IV Quantitative Precipitation Estimates for Product

Intercomparisons, Weather and Forecasting, 31(2), 371-394, doi: 10.1175/waf-d-14-00112.1.

Ngo-Duc, T., Matsumoto, J., Kamimera, H., & Bui, H. H. (2013). Monthly adjustment of Global Satellite Mapping of Precipitation (GSMaP) data over the VuGia–ThuBon River Basin in Central Vietnam using an artificial neural network. *Hydrological Research Letters*, *7*(4), 85-90.

Njoku, E. G. (1982). Passive microwave remote sensing of the earth from space—A review. *Proceedings of the IEEE*, *70*(7), 728-750.

Noh, Y. J., & Liu, G. (2004). Satellite and aircraft observations of snowfall signature at microwave frequencies. *Riv. Ital. Telerilevamento*, *30*, 101-118.

Observing Laboratory. Retrieved from http://data.eol.ucar.edu/dataset/21.093

Olson, W. S. et al. (2018), GPM Combined Radar-Radiometer Precipitation Algorithm Theoretical Basis Document, available at: https://arthurhou.pps.eosdis.nasa.gov/Documents/Combine

Prigent, C. (2010). Precipitation retrieval from space: An overview. *Comptes Rendus Geoscience*, *342*(4-5), 380-389.

Rabin, R. M., Stadler, S., Wetzel, P. J., Stensrud, D. J., & Gregory, M. (1990). Observed effects of landscape variability on convective clouds. *Bulletin of the American Meteorological Society*, *71*(3), 272-280.

Randel, D. L., Kummerow, C. D., & Ringerud, S. (2020). The Goddard Profiling (GPROF) Precipitation Retrieval Algorithm. In *Satellite precipitation measurement* (pp. 141-152). Springer, Cham.

Sakakibara, H. (1981), Heavy rainfall from very shallow convective clouds, J. Meteor. Soc. Japan, 59, 387–3

Sapiano, M. R. P., & Arkin, P. A. (2009). An intercomparison and validation of high-resolution satellite precipitation estimates with 3-hourly gauge data. *Journal of Hydrometeorology*, *10*(1), 149-166.

Scofield, R. A., & Kuligowski, R. J. (2003). Status and outlook of operational satellite precipitation algorithms for extreme-precipitation events. *Weather and Forecasting*, *18*(6), 1037-1051.

Shige, S., Kida, S., Ashiwake, H., Kubota, T., & Aonashi, K. (2013). Improvement of TMI rain retrievals in mountainous areas. *Journal of Applied Meteorology and Climatology*, *52*(1), 242-254.

Shige, S., Yamamoto, M. K., & Taniguchi, A. (2014). Improvement of TMI rain retrieval over the Indian subcontinent. *Remote Sensing of the Terrestrial Water Cycle, Geophys. Monogr*, *206*, 27-42.

Skofronick-Jackson, G., Kirschbaum, D., Petersen, W., Huffman, G., Kidd, C., Stocker, E., & Kakar, R. (2018). The Global Precipitation Measurement (GPM) mission's scientific achievements and societal contributions: reviewing four years

of advanced rain and snow observations. *Quarterly Journal of the Royal Meteorological Society*, *144*, 27-48.

Takeda, T., and K. Takase (1980), Radar observation of rainfall system modified by orographic effects, J. Meteor. Soc. Japan, 58, 500–516.

Tang, L., Tian, Y., & Lin, X. (2014). Validation of precipitation retrievals over land from satellite-based passive microwave sensors. *Journal of Geophysical Research: Atmospheres*, *119*(8), 4546-4567.

Taniguchi, A. S. Shige, M. K. Yamamoto, T. Mega, S. Kida, T. Kubota, M. Kachi, T. Ushio, and K. Aonashi (2013), Improvement of high-resolution satellite rainfall product for Typhoon Morakot (2009) over Taiwan, J. Hydrometeor, 14, 1859–1871.

Ushio, T., et al. (2009), A Kalman filter approach to the Global Satellite Mapping of Precipitation (GSMaP) from combined passive microwave and infrared radiometric data. J. Meteor. Soc. Japan, 87A, 137–151.

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.

Wang, N. Y., C. Liu, R. Ferraro, D. Wolff, E. Zipser, and C. Kummerow (2009), TRMM 2A12 land precipitation product-status and future plans, J. Meteor. Soc. Japan, 87A, 237–253.

Wang, N. Y., Liu, C., Ferraro, R., Wolff, D., Zipser, E., & Kummerow, C. (2009). TRMM 2A12 land precipitation product-status and future plans. *Journal of the Meteorological Society of Japan. Ser. II*, *87*, 237-253.

Wilheit, T. T. (1986). Some comments on passive microwave measurement of rain. *Bulletin of the American Meteorological Society*, *67*(10), 1226-1232.

Yamamoto, M. K., & Shige, S. (2014). Implementation of an orographic/nonorographic rainfall classification scheme in the GSMaP algorithm for microwave radiometers. *Atmospheric Research*, *163*, 36-47.

Yamamoto, M. K., Shige, S., Yu, C. K., & Cheng, L. W. (2017). Further improvement of the heavy orographic rainfall retrievals in the GSMaP algorithm for microwave radiometers. *Journal of Applied Meteorology and Climatology*, *56*(9), 2607-2619.

You, Y., Petkovic, V., Tan, J., Kroodsma, R., Berg, W., Kidd, C., & Peters-Lidard, C. (2020). Evaluation of V05 precipitation estimates from GPM constellation radiometers using KuPR as the reference. *Journal of Hydrometeorology*, *21*(4), 705-728.

Zhang, J., Howard, K., Langston, C., Kaney, B., Qi, Y., Tang, L., ... & Kitzmiller, D. (2016). Multi-Radar Multi-Sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bulletin of the American Meteorological Society*, *97*(4), 621-638.

Table 01: List of the data sets used in this study. Two years (2018-2019) of the dataset over the western US is used in this study. All the data sets are compared in ERA5 resolution (0.1°* 0.1 °).

Satellite/platform	Sensor/Algorithm	Variables Used
NCEP Sage-IV	Ground-based multi-sensors and gauges measurements	Surface precipitation
NOAA-19	MHS V05	Surface precipitation
GPM	GMI V05	Surface Precipitation
GPM	DPR V06	Near-surface precipitation
GPM	COMB V06	Near-surface precipitation
Geostationary	IR/ PERSIANN-CCS through IMERG-IR	IR precipitation rates
ERA5	Reanalysis data	T2M, TPW, "near-surface" win
USGS DEM		Elevation

Table 02: Comparison of the products skill scores for precipitation occurrence as a function of ORI groups based on Method 1. ORI indices are presented in 5 categories as follows: ORI1: ORI < -5, ORI2: -0.25 > ORI > = -5, ORI3: 0.25 > ORI >= -0.25, ORI4: 5 > ORI >= 0.25, ORI5: ORI >5.

@ >p(- 10) * >p(- 10) * >p(- 10) * >p(- 10) * >p(- 10) * >p(- 10) * @ Sensors & ORIs & FAR & HSS & POD & BIASo

MHS & & & & &
& ORI1 & 0.55 & 0.33 & 0.34 & 1.2
& ORI2 & 0.62 & 0.26 & 0.26 & 0.87
& ORI3 & 0.66 & 0.25 & 0.25 & 0.92
& ORI4 & 0.63 & 0.28 & 0.28 & 1.06
& ORI5 & 0.54 & 0.33 & 0.36 & 1.25

GMI & & & & &
& ORI1 & 0.55 & 0.32 & 0.31 & 0.81
& ORI2 & 0.63 & 0.28 & 0.26 & 0.79
& ORI3 & 0.70 & 0.26 & 0.25 & 0.90
& ORI4 & 0.65 & 0.28 & 0.27 & 0.89
& ORI5 & 0.55 & 0.32 & 0.33 & 0.88

DPR & & & & &
& ORI1 & 0.51 & 0.33 & 0.30 & 0.69
& ORI2 & 0.63 & 0.28 & 0.26 & 0.74
& ORI3 & 0.69 & 0.25 & 0.23 & 0.89
& ORI4 & 0.65 & 0.27 & 0.25 & 0.79
& ORI5 & 0.53 & 0.31 & 0.29 & 0.69

COMB & & & & &
& ORI1 & 0.52 & 0.33 & 0.31 & 0.72
& ORI2 & 0.64 & 0.28 & 0.26 & 0.79
& ORI3 & 0.71 & 0.24 & 0.24 & 0.99
& ORI4 & 0.66 & 0.27 & 0.26 & 0.84
& ORI5 & 0.53 & 0.32 & 0.30 & 0.75
IMERG- IR & & & & &
& ORI1 & 0.73 & 0.19 & 0.22 & 0.88
& ORI2 & 0.82 & 0.16 & 0.18 & 0.99
& ORI3 & 0.86 & 0.14 & 0.18 & 1.24
& ORI4 & 0.82 & 0.16 & 0.19 & 1.05
& ORI5 & 0.71 & 0.20 & 0.23 & 0.91
ERA5 & & & & &
& ORI1 & 0.59 & 0.42 & 0.52 & 1.27
& ORI2 & 0.66 & 0.36 & 0.44 & 1.29
& ORI3 & 0.70 & 0.33 & 0.41 & 1.35
& ORI4 & 0.67 & 0.36 & 0.45 & 1.35
& ORI5 & 0.58 & 0.44 & 0.56 & 1.35

Table 03: Precipitation estimation statistics (using Method 1) of all sensors/products when compared to Stage-IV.

@ >p(- 8) * >p(- 8) * >p(- 8) * >p(- 8) * >p(- 8) * @ Sensors & ORIs & CC-Value & RMSE (mm/day) & BIASr

MHS & & & &
& ORI1 & 0.35 & 0.67 & 0.97
& ORI2 & 0.32 & 0.32 & 0.97
& ORI3 & 0.32 & 0.22 & 0.99
& ORI4 & 0.33 & 0.32 & 0.89
& ORI5 & 0.35 & 0.70 & 0.93

GMI & & & &
& ORI1 & 0.34 & 0.65 & 0.86
& ORI2 & 0.30 & 0.36 & 0.94
& ORI3 & 0.28 & 0.28 & 1.09
& ORI4 & 0.30 & 0.36 & 0.86
& ORI5 & 0.33 & 0.65 & 0.78

DPR & & & &
& ORI1 & 0.33 & 0.62 & 0.59
& ORI2 & 0.27 & 0.41 & 0.66
& ORI3 & 0.26 & 0.25 & 0.75
& ORI4 & 0.27 & 0.35 & 0.65

& ORI5 & 0.34 & 0.62 & 0.56

COMB & & & &
& ORI1 & 0.30 & 0.73 & 0.78
& ORI2 & 0.26 & 0.42 & 0.85
& ORI3 & 0.24 & 0.30 & 0.97
& ORI4 & 0.25 & 0.41 & 0.80
& ORI5 & 0.32 & 0.69 & 0.76
IMERG-IR & & & &
& ORI1 & 0.19 & 0.87 & 0.99
& ORI2 & 0.16 & 0.51 & 1.14
& ORI3 & 0.15 & 0.38 & 1.33
& ORI4 & 0.16 & 0.50 & 1.08
& ORI5 & 0.20 & 0.91 & 0.95

ERA5 & & & &
& ORI1 & 0.35 & 0.55 & 1.25
& ORI2 & 0.31 & 0.36 & 1.29
& ORI3 & 0.27 & 0.29 & 1.31
& ORI4 & 0.29 & 0.39 & 1.27
& ORI5 & 0.33 & 0.52 & 1.26

Table 04: Precipitation occurrence statistics of all sensors/products using method 2. In method 2, precipitation is classified as an orographic if orographically forced upward motion and moisture flux convergence are above certain threshold values as indicated in equations 2 and 3.

Sensors	ORIs	FAR	HSS	POD	BIASo
MHS					
	ORO	0.43	0.28	0.26	0.67
	N0-ORO	0.54	0.25	0.23	0.80
GMI					
	ORO	0.51	0.32	0.30	0.73
	N0-ORO	0.62	0.29	0.28	0.83
DPR					
	ORO	0.48	0.32	0.27	0.57
	N0-ORO	0.61	0.30	0.27	0.72
COMB					
	ORO	0.50	0.32	0.27	0.61
	N0-ORO	0.61	0.30	0.28	0.78
IMERG- IR					
	ORO	0.71	0.19	0.20	0.75
	N0-ORO	0.81	0.16	0.19	0.97
ERA5					

Sensors	ORIs	FAR	HSS	POD	BIASo
	ORO	0.54	0.47	0.56	1.19
	N0-ORO	0.65	0.38	0.45	1.25

Table 05: Precipitation estimation statistics (using Method 2) of all sensors/products when compared to Stage-IV.

@ >p(- 8) * >p(- 8) * >p(- 8) * >p(- 8) * >p(- 8) * @ Sensors & ORIs & CC-Value & RMSE & BIASr
MHS & & & &
& ORO & 0.31 & 0.52 & 0.63
& N0-ORO & 0.30 & 0.26 & 0.86
GMI & & & &
& ORO & 0.36 & 0.57 & 0.79
& N0-ORO & 0.33 & 0.34 & 0.93
DPR & & & &
& ORO &

0.36

& 0.54 & 0.49
& N0-ORO &

0.31

& 0.34 & 0.65
COMB & & & &
& ORO & 0.34 & 0.60 & 0.66
& N0-ORO & 0.29 & 0.39 & 0.85
IMERG- IR & & & &
& ORO & 0.20 & 0.80 & 0.86
& N0-ORO & 0.17 & 0.49 & 1.13
ERA5 & & & &
& ORO & 0.54 & 0.45 & 1.00
& N0-ORO & 0.46 & 0.28 & 1.19

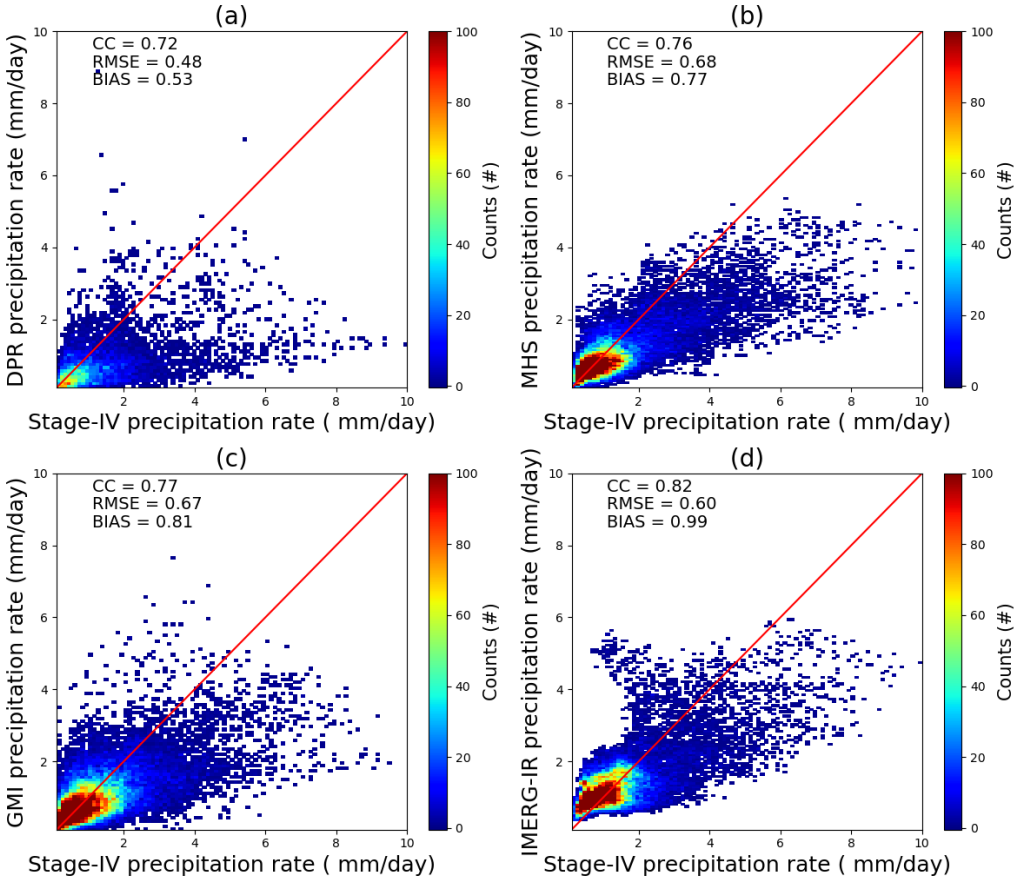

Fig 01: Scatter density plot of Stage-IV vs. (a) DPR, (b) MHS, (c) GMI, (d) IMERG-IR over the western USA. The red line in the plot is 1:1 line. Two years (2018-2019) of data is used to create the density plot where each product is averaged in 0.25^0x0.25^0 longitudes and latitudes.

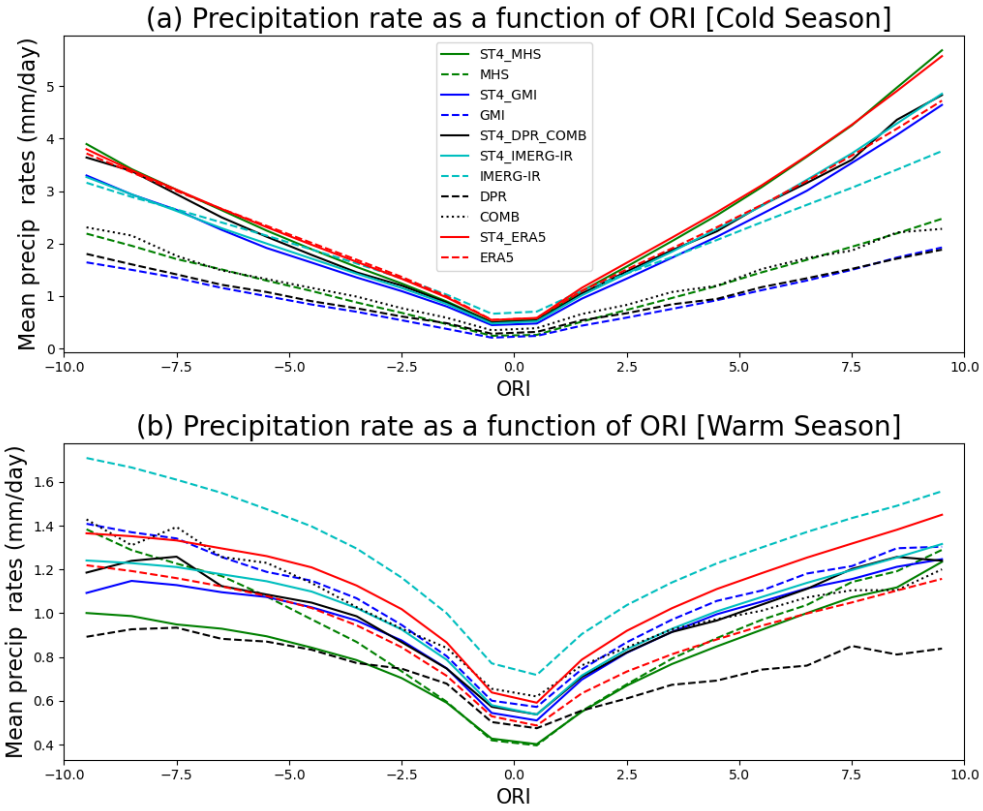

Fig 02: Two years (2018-2019) of average precipitation rates as a function of ORI indexes (based on method 1) for (a) cold season and (b) warm season. Dashed lines represent sensors precipitation rates, whereas solid lines represent Stage-IV precipitation rates that are matched with respective sensors as indicated in the legend.

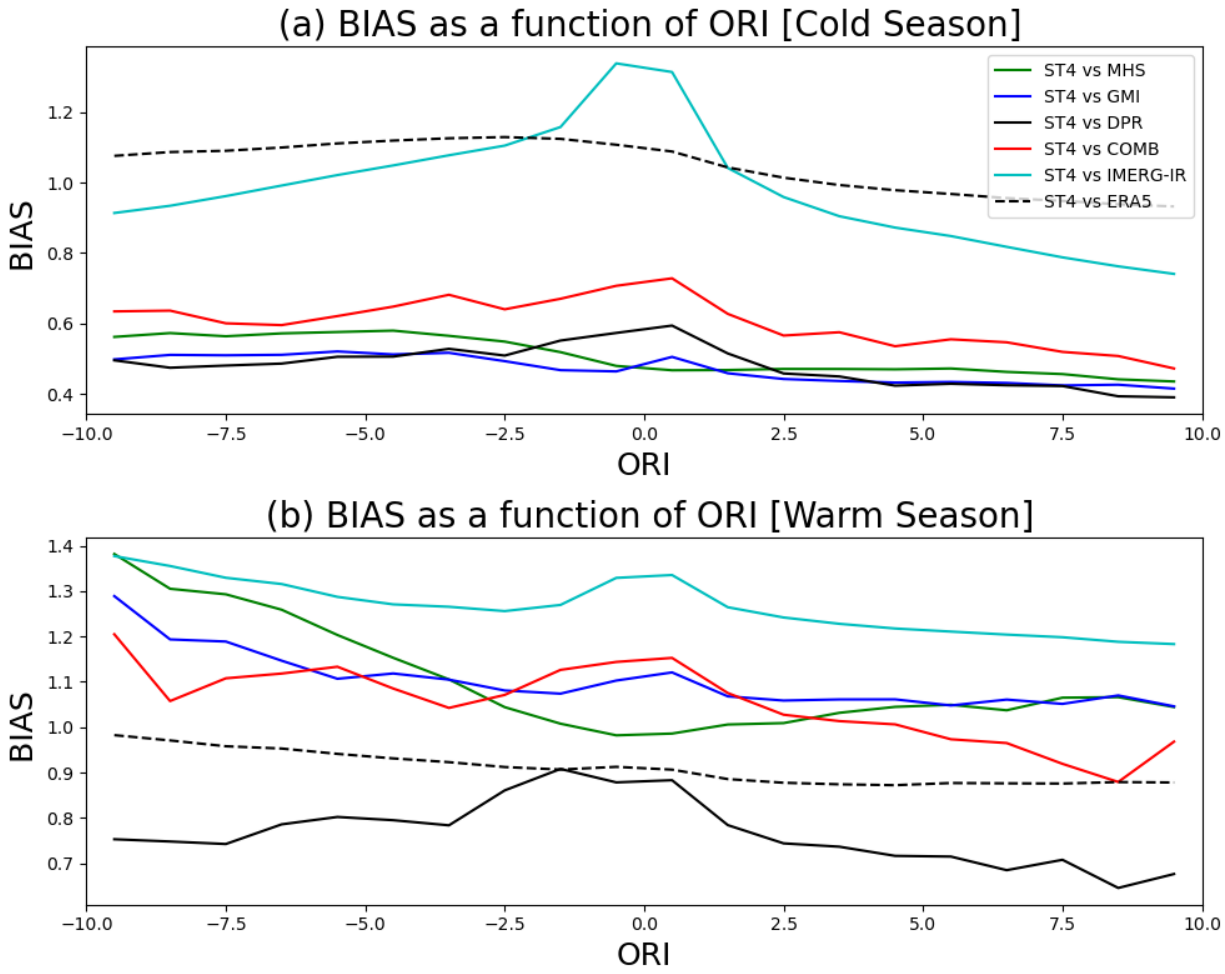

Fig 03: BIASr as a function of ORI indices (based on method 1) for (a) cold season and (b) warm season. The BIASr is defined as the ratio of sensors' total precipitation to that of Stage-IV in each ORI indices.

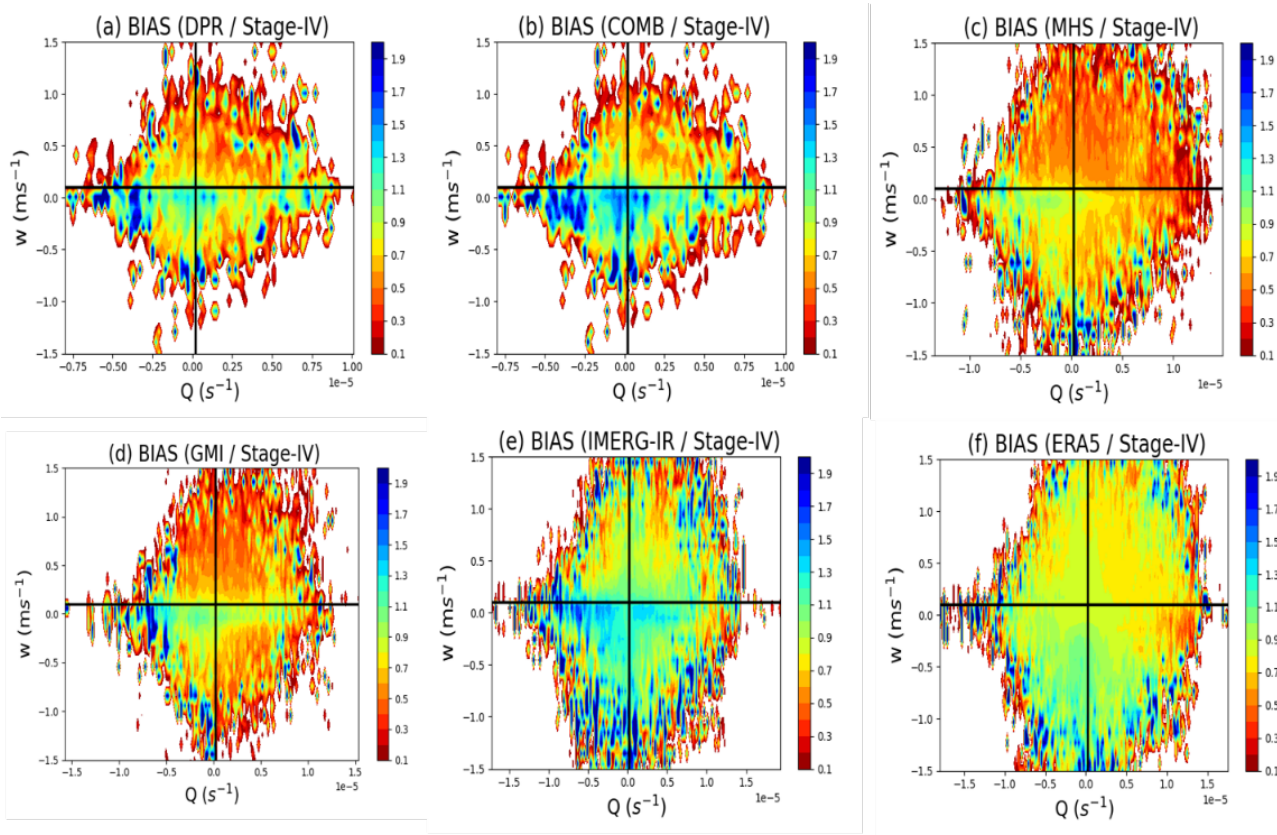

Fig 04: BIASr as a function of moisture flux convergence (w) and orographic enhancement (Q) for all the products. Here, BIASr is the ratio of sensors total precipitation to the Stage-IV total precipitation in each w and Q bin. The horizontal and vertical black lines represent the w and Q threshold values, respectively. If precipitation occurs above these threshold values, then precipitation is classified as orographic precipitation (Method 2).

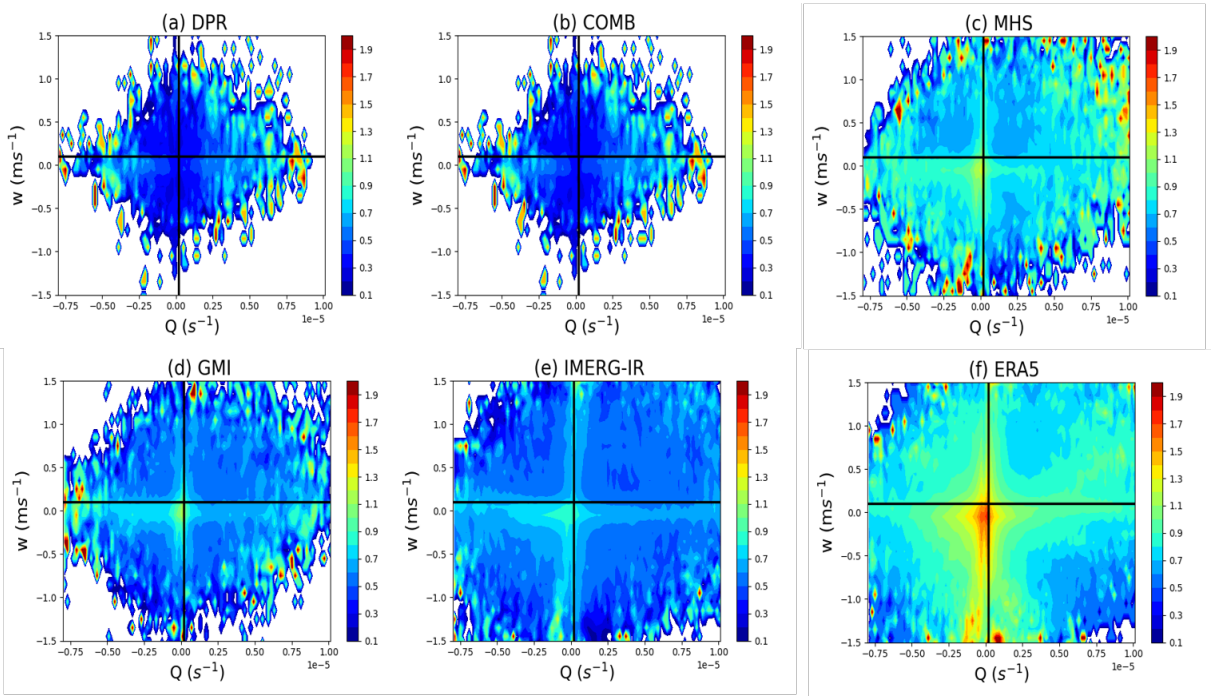

Fig 05: Occurrence BIAS (BIASo) as a function of moisture flux convergence (w) and orographic enhancement (Q) for all the products. The horizontal and vertical black lines represent the w and Q threshold values, respectively. If precipitation occurs above these threshold values, then precipitation is classified as orographic precipitation (Method 2).

Note: BIAS = (Hits + False Alarms) / (Hits + Misses)

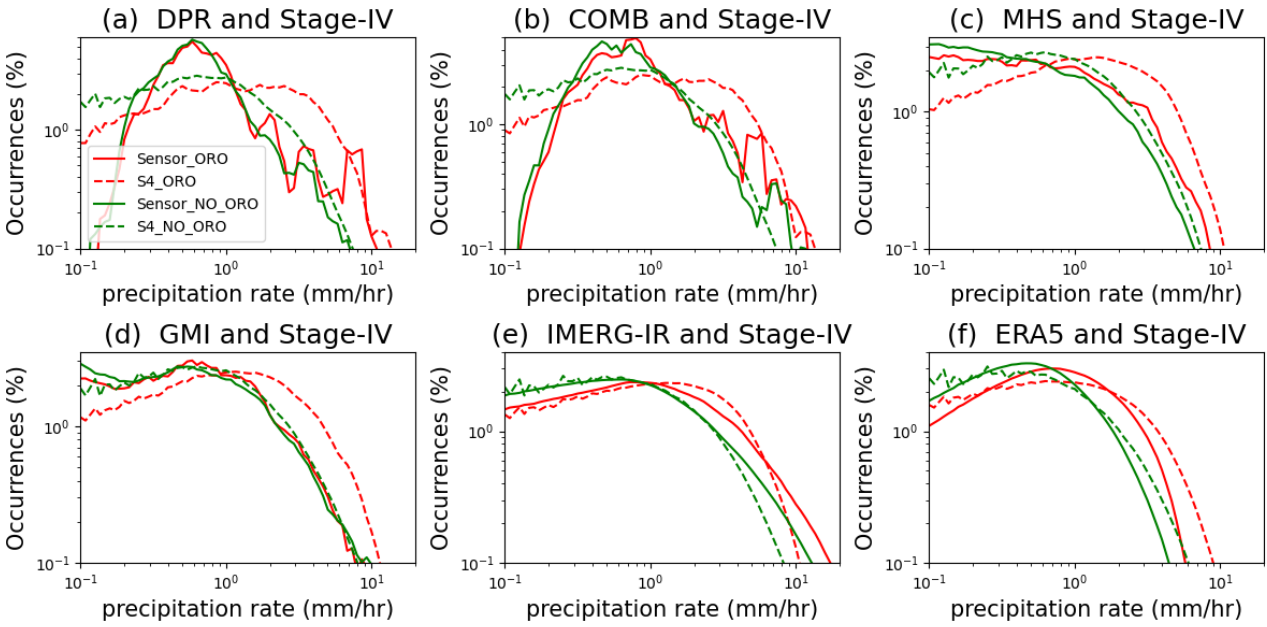

Fig 06: Occurrence of orographic (red curve) and non-orographic (green curve) precipitation as defined by the second method. The solid and dashed lines represent sensors' and Stage-IV precipitation occurrences that are matched with each other, respectively. Each curve in the figure is normalized so that each adds up to 100.

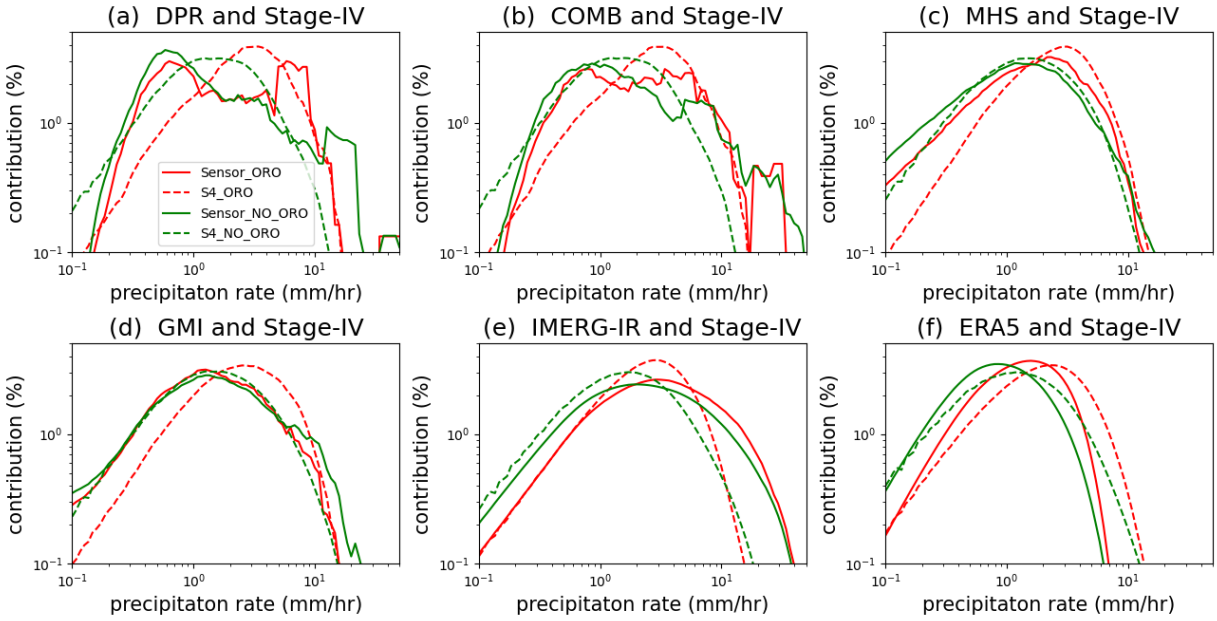

Fig 07: Precipitation amount contribution from each orographic (red curve) and non-orographic (green curve) precipitation bins as defined by the second method. The solid and dashed lines represent sensors' and Stage-IV precipitation contributions that are matched with each other, respectively. Each curve in the figure is normalized so that the area below each adds up to 100.

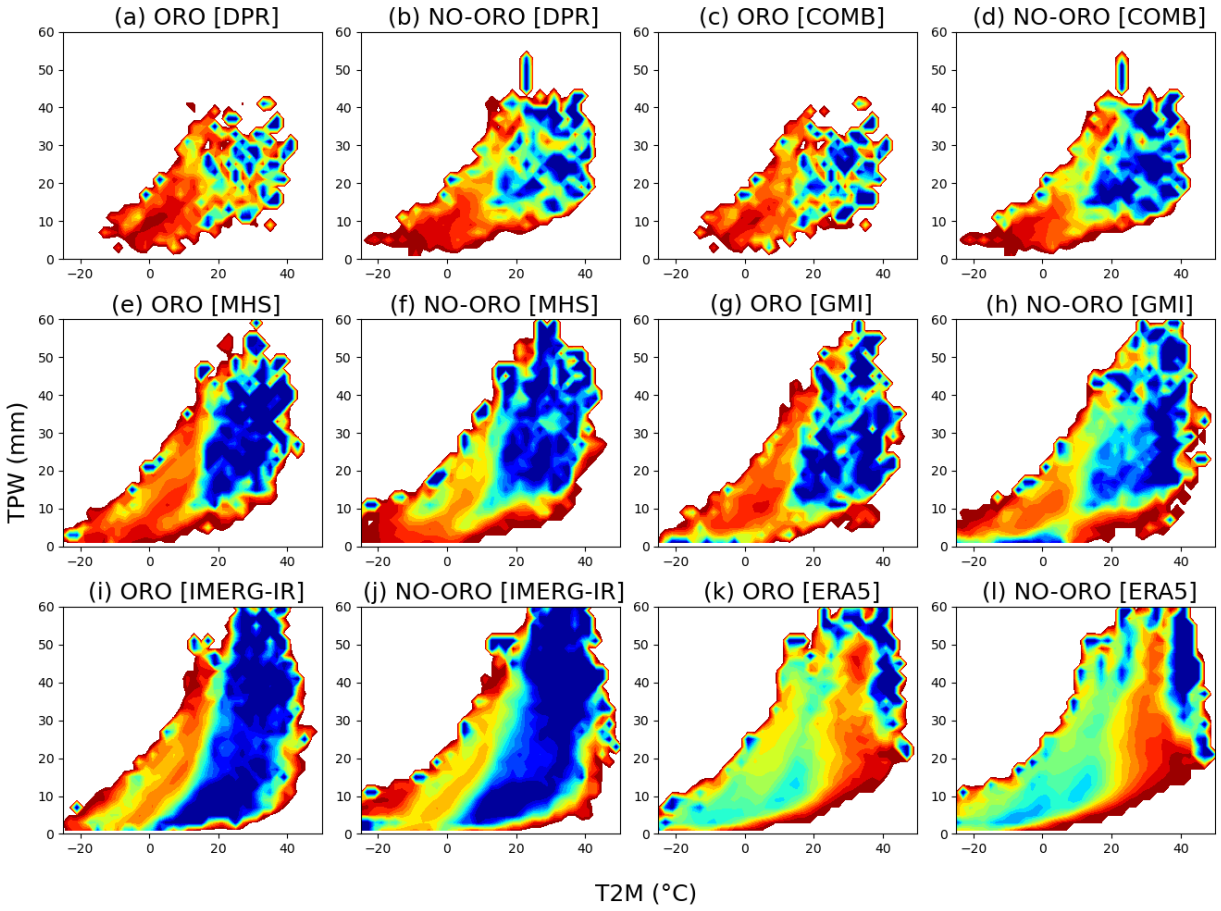

Fig 08: BIASr as a function of T2M and TPW for each sensor for both orographic and non-orographic events based on 2nd method. BIASr is calculated by dividing sensor precipitation amount by Stage-IV precipitation amount between each TPW and T2M bins by respective sensors.