

Article

Investigating various products of IMERG for precipitation retrieval over surfaces with and without snow and ice cover

Alireza Arabzadeh ¹, Ali Behrangi ^{1,2,*}

¹ Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA.; arabzadeh@email.arizona.edu

² Department of Geosciences, University of Arizona, Tucson, AZ, USA; behrangi@email.arizona

* Correspondence: behrangi@email.arizona

Abstract: Various products of the Integrated Multisatellite Retrievals for GPM (IMERG) and passive microwave (PMW) sensors are assessed with respect to near-surface wet-bulb temperature (T_w), precipitation intensity, and surface type (i.e., with and without snow and ice on the surface) over the CONUS and using Stage-IV product as reference precipitation. IMERG products include precipitation estimates from infrared (IR), combined PMW, and their combination. PMW products generally have higher skills than IR over snow- and ice-free surfaces. Over snow- and ice-covered surfaces (1) PMW products (except AMSR-2) show a higher correlation coefficient than IR, (2) IR and PMW precipitation products tend to overestimate precipitation, but at colder temperatures (e.g., $T_w < -10^\circ\text{C}$) PMW products tend to underestimate and IR product continues to show large overestimations, and (3) PMW sensors show higher overall skill in detecting precipitation occurrence, but not necessarily at very cold T_w . The results suggest that the current approach of IMERG (i.e., replacing PMW with IR precipitation estimates over snow- and ice-surfaces) may need to be revised.

Keywords: IMERG, Stage IV, Infrared, Passive microwave, Snow, Ice, Precipitation, GPM, Wet-bulb temperature, AMSR-2

1. Introduction

Accurate precipitation estimation with high spatiotemporal resolution is key to many hydrologic studies. Rain gauges and ground radars have enabled high-quality observation and estimate of precipitation at a point or regional scales and satellite observations have enabled precipitation estimate with global coverage at sub-daily temporal sampling, which is important for regional and global studies.

The high temporal sampling from satellite observations comes through the use of infrared images from geostationary satellites, multiple Passive Microwave (PMW) sensors on low-earth-orbiting satellites, or a combination of both. PMW sensors often provide more information about the hydrometeors, thus tend to result in more accurate precipitation estimates than precipitation retrieval based on IR data. However, PMW-based precipitation estimation may also face large uncertainties due to several factors including errors related to the poor understanding of precipitation microphysics, difficulties in distinguishing between light rain and cloud [1-3], and challenges in determining surface emissivity, especially over snow and ice [4, 5].

The methods for blending IR- and PMW-based precipitation estimates have been different across different merged precipitation products. Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks (PERSIANN) [6] and PERSIANN Cloud Classification System (PERSIANN-CCS) [7] are mainly based on IR brightness temperature-precipitation rate relationships established using geostationary IR observations and PMW precipitation estimates. On the other hand, products such as the Integrated Multisatellite Retrievals for GPM (IMERG) [8], Climate Prediction Center's morphing technique (CMORPH) [9], and JAXA's Global Satellite Mapping of

Precipitation (GSMaP) [10] use PMW precipitation as their main input and IR precipitation might be used to fill gaps in time. For example, if the time interval from the nearest PMW observation is longer than 30 minutes, IMERG tends to use a combination of IR and PMW estimates and for time intervals of more than 90 minutes, the precipitation rate mainly comes from IR-estimate. This is because by about ± 90 minutes the IR precipitation has shown a higher correlation with reference precipitation than the propagated microwave precipitation estimates [8].

IMERG also uses IR precipitation estimate over snow and ice surfaces, regardless of the time distance from PMW observations and type of PMW sensors [8]. This is because PMW precipitation estimates are perceived as unreliable over snow and ice surfaces. However, the performance of IR precipitation over snow and ice surfaces has also not been well investigated. Furthermore, it has been found that different PMW sensors have dissimilar skills for precipitation retrieval over land and ocean [11]. It is important to assess how precipitation estimates from different PMW sensors (or a combination of them as used in IMERG) are compared with IR precipitation over surfaces with and without snow and ice. IMERG prefers precipitation estimates from PMW imagers over PMW sounders, which may not always be the best choice over land [11]. Examples of recent PMW imagers used in IMERG are the Special Sensor Microwave Imager/Sounder (SSMIS) on the Defense Meteorological Satellite Program (DMSP) platforms, the GPM Microwave Imager (GMI) on the GPM Core Observatory satellite, and the Advanced Microwave Scanning Radiometer Earth Observing System (AMSR-E) on the Aqua satellite; and its follow-on satellite (AMSR2) on board the GCOM-W1 satellite. Among the main PMW sounders used in IMERG are: the Advanced Technology Microwave Sounder (ATMS) on board Suomi National Polar-Orbiting Partnership (Suomi-NPP) and NOAA-20 satellites; the Microwave Humidity Sounder (MHS) on board NOAA-18, NOAA-19, MetOp-A, MetOp-B, and MetOp-C satellites. In IMERG, IR precipitation is obtained from PERSIANN-CCS.

Besides considering surface conditions (e.g., surfaces with and without snow and ice) in comparison of PMW and IR precipitation, it is important to perform the analysis as a function of surface temperature, as surface temperature is used to discriminate the precipitation phase. Behrangi, Tian [12] compared AMSR-E and CloudSat precipitation detection as a function of surface air temperature at 2m (T2m) and showed that AMSR-E significantly underestimates CloudSat precipitation for T2m below the freezing temperature. Zhang, Lin [13] also used T2m to evaluate high-resolution (0.1° /hourly) precipitation estimates from the Weather Research and Forecasting (WRF) Model and IMERG over the central United States. Results show that the WRF estimates exhibit higher correlations with the reference data when the temperature falls below 280 K, while IMERG estimates show higher correlation than WRF for T2m greater than 280 K. They also showed that the complementary behavior of the WRF and the IMERG products conditioned on T2m does not vary with either season or location. However, their study did not provide information on the performance of PMW and IR precipitation estimates that are used in IMERG and did not consider surface condition (i.e. surfaces with and without snow and ice) that determines the use of IR or PMW precipitation components in IMERG.

In the present study, we focus on evaluating the performance of IMERG and its precipitation components. In addition, analysis as a function of precipitation rate provides insights into the performance of the precipitation products under light, moderate, and intense precipitation events. These analyses should add insights to refine strategies for combining IR and PMW precipitation estimates in the merged products such as IMERG.

2. Materials and Methods

2.1 comparison approach and metrics

Using three years of Stage IV data (2015-2017) as a reference over the CONUS, the performance of each product is assessed within each 1-degree T_w bin, separately over surfaces with and without snow and ice cover. Precipitation thresholds of 0, 0.1, 0.3, 1, and 2mm/hr are used to separate precipitation from no-precipitation. For example, by using a

threshold of 2mm/hr, only grids with precipitation rates of greater than 2mm/hr are considered as precipitating and anything less than that is considered as non-precipitating, so the emphasis will be more on intense precipitation events. The calculation of T_w is described in [14] and uses 2-m air temperature, 2-m dewpoint temperature, and surface pressure that are obtained from reanalysis. The score metrics used in this study are: Probability of Detection (POD), False Alarm Ratio (FAR), Bias, and Heidke Skill Score (HSS) for assessment of precipitation occurrence as well as Correlation Coefficient (CC) and Volume Bias (VBias) for assessment of precipitation rate. 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 Bias and VBIAS is 1. Bias is calculated by dividing the number of estimated precipitation occurrences from each product by the corresponding value from the reference product (here Stage IV). VBIAS is similar to Bias, but is calculated by dividing the amount of estimated precipitation by the corresponding value from observation. It should be noted that Bias of 1 alone does not necessarily indicate a perfect prediction, because Bias of 1 means that the number of grids identified as precipitation is the same for the two products being compared. However, a product might miss or falsely determine precipitation occurrence as can be inferred from POD and FAR. Details for calculation of the above metrics are provided in [15].

Note that it is known that most satellite products underestimate orographic precipitation enhancement and are not able to capture that accurately [16]. Therefore, to separate this effect from our analysis, using maps of mountains (See Section 2.2), regions susceptible to orographic precipitation enhancement were removed from the analysis.

2.2. Dataset

A brief description of the products used in this study is provided below:

- IMERG products

The latest version of the IMERG products (V06) used in this study. IMERG provides gridded precipitation maps with high spatiotemporal resolution (0.1x0.1 deg. every ½ hr) within 90°S-N by blending precipitation estimates from two sources: (1) IR images using PERSIANN-CCS, and (2) GPM Microwave Imager (GMI) and a constellation of GPM PMW sensors. IMERG uses the latest version of PMW precipitation products (V05) retrieval based on the 2017 version of the Goddard Profiling Algorithm (GPROF2017) [17]. Details of the IMERG algorithm are described in Huffman, Bolvin [18] and in brief includes four main steps: (1) precipitation estimates from the GPM constellation radiometers are gridded, intercalibrated to the radar-microwave combined product (2BCMB), and combined into half-hourly 0.1°x0.1° fields (variable name: "HQprecipitation"), hereafter referred to as IMERG-HQ, (2) maps of half-hourly IR precipitation rate ("IRprecipitation") are calculated using PERSIANN-CCS, hereafter referred to as IMERG-IR, (3) MW and IR estimates are used to create half-hourly estimates ("precipitationUncal") by utilizing the Climate Prediction Center Morphing-Kalman Filter (CMORPH-KF) Lagrangian time interpolation scheme, and (4) the multi-satellite half-hour estimates are adjusted so that they sum to a monthly satellite-gauge combination ("precipitationCal"). This bias-adjusted product is referred to as IMERG-Final which is available ~3.5 months after observation for accurate estimation with monthly gauge adjustments to reduce bias. IMERG also reports sources of PMW sensors "HQprecipSource", but it does not specify the platform names. Note that IMERG provides users with two other runs. The Early run is available ~4 hours after observation for real-time applications such as flood prediction and includes only forward morphing and may not benefit from all PMW sensors. The Late run (IMERG-Late) is available ~14 hours after observation, implementing forward and backward morphing. The current version of IMERG-Late (V6) does not apply climatology bias adjustment, but the future versions will.

In this study, we use IMERG-Late, IMERG-Final, IMERG-HQ (PMW-only), and IMERG-IR (IR-only) products available at half-hour at 0.1x0.1 degree resolution. As

discussed in Section 1, in IMERG, IR is used over snow and ice surfaces or when the time distance from PMW observations is large. For the period of this study, IMERG-HQ uses imagers (AMSR-2 and GMI), sounders (ATMS, and MHS), and combined imager and sounder (SSMIS) sensors. All of these sensors, regardless of their platform, are used in our analysis. When there are available data from sounders and imagers, IMERG prioritizes imagers over sounders with consideration of observation time closest to the center of the half-hour (Huffman et al. 2020).

- National Centers for Environment Prediction (NCEP) Stage IV

The hourly Stage IV product is used in this study as reference data to analyze IMERG products' performance over CONUS. Since 2002 the hourly Stage IV product provides hourly estimates using Z-R relationship from over 150 Doppler Next Generation Weather Radars (NEXRAD), and a combination of 5500 hourly rain gauge measurements to produce hourly 4×4 km resolution data [19, 20]. As of April 2017, this dataset includes Alaska and Puerto Rico stations. Stage IV benefits from manual quality control (QC) on Stage III data gathered at each River Forecast Center (RFC) unlike Stage II which does not include manual QC (<https://data.eol.ucar.edu/dataset/21.093>). Including quality control, gauge measurements give assurance of the quality of the data. This becomes more important in the case of snowfall measurements when the temperature is below the freezing point because most of the radar-based estimations under freezing conditions are controversial [21]. Cocks, Martinaitis [22] shows the value of nine winter precipitation events over the Rocky Mountains is in good agreement with gauge measurements. Altogether we found out NCEP Stage IV data is the most suitable and convenient for our study with fairly reliable accuracy [21].

- ERA5-Land

This dataset is a replay of ECMWF original land component ERA5 climate reanalysis with a finer resolution at ~9 Km grid spacing with an hourly time interval. This dataset provides surface variable data from 1981 to 2-3 months before the present. In this study to calculate wet-bulb temperature, three different variables were obtained from ERA5-Land including 2-m air temperature, 2-m dewpoint temperature, and surface pressure for 2015 to 2017 over CONUS. We used wet-bulb temperature to distinguish rainfall from snowfall following as it is a better separator than air temperature [14, 23, 24].

- NOAA Autosnow Product

The NOAA Autosnow product provides daily surface ice and snow map with global coverage. It is a gridded product with 0.04° lat/lon resolution. It uses data gathered from different sensors on various satellites. For details of sensors used in this product see Romanov [25]. IMERG uses Autosnow in two stages: (1) mask snow- and ice-covered surfaces for Kalman statistics computation and, (2) mask IMERG precipitation estimates. In this study, we use Autosnow to delineate surfaces with and without snow and ice cover.

- K3 mountain map

In this study to detach the issues related to the underestimation of orographic precipitation by satellite products from our analysis, K3 mountain mask was used. K3 provides a GIS-based, global map of mountains derived from digital elevation models (DEM) with a 250m resolution. It has been developed after K1 and K2 mountains raster with a coarser resolution [26-28]. For comparison of different K mountain maps see Roger, Charlie [29]. K3 characterizes mountains into four different groups: high and high-scattered mountains with elevation exceeding 900 m and low and low-scattered mountains with elevation ranging between 301-900 m. K3 maps are found useful to identify regions that might contain orographic precipitation (Personal communication with Dr. Paula Brow of Colorado State University). Since in this study we are not focusing on orographic precipitation and radar precipitation estimates over mountainous regions may not be accurate [30], we removed all high and high-scattered mountains (see supplementary Figure S1) from our analysis.

3. Results

The results of this study are presented under three main sections: (1) general characteristics and differences of IMERG and Stage IV products including spatial distribution and seasonality accompanied by a case study, (2) analysis of IMERG components versus Stage IV data over snow-and ice-covered and snow- and ice-free surfaces with different intensities and, (3) investigating the performance of individual PMW sensor types.

3.1 General Characterization

Figure 1 shows seasonal mean precipitation maps from Stage IV (Figs. 1a-1d), IMERG-Final (Figs. 1e-1h), IMERG-HQ (Figs. 1i-1l), IMERG-IR (Figs. 1m-1p), and IMERG-Late (Figs. 1q-1t) using three years (2015-2017) of data over CONUS. For a more detailed assessment, maps of seasonal differences in mean precipitation rate between IMERG products and Stage IV are also plotted in **Figure 2**. From these two figures, few points can be highlighted: (1) because IMERG-Final is bias-adjusted with gauges at monthly scale, it is closest to Stage IV in terms of both magnitude and pattern, (2) over mountainous regions, mainly in the west, IMERG products tend to underestimate precipitation rates during winter (DJF) where snow and ice are on the surface. This is also observed in spring (MAM) and fall (SON), but not in summer (JJA), although the underestimation is more noticeable for IMERG-HQ than IMERG-IR that is consistent with previous studies [31, 32], (3) IMERG products (except IMERG-Final that is bias-adjusted) show larger precipitation amount than Stage IV over the central and eastern parts of the CONUS, especially during DJF and MAM. It is not clear if this overestimation, especially during the DJF, is due to overestimation of rainfall, snowfall, or due to surface conditions (e.g., snow and ice on the surface). Therefore, to better understand the performance of the satellite products, a more detailed analysis as a function of surface type and precipitation phase is needed.

The observed underestimation over the mountainous west can be due to poor skill of satellite products in capturing orographic enhancement [16, 33] and the fact that most of the annual precipitation over this region occurs in wintertime [34] and a large fraction of the that is through the atmospheric rivers [35, 36]. However, it is important to note that radar beam blockage may also contribute to the lower quality of Stage IV product in mountainous regions. Most of these regions are masked out using the K3 mask applied in this study.

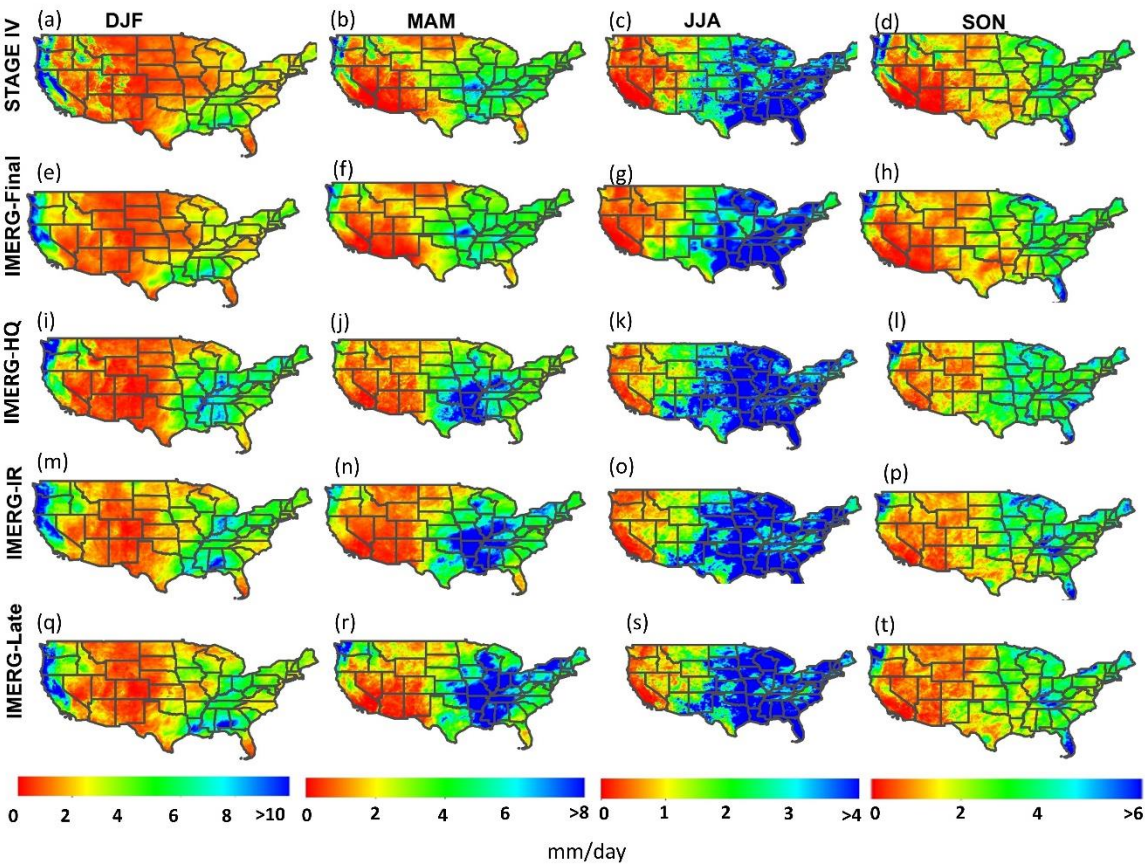


Figure 1: Spatial distribution of mean seasonal precipitation rate (mm/day) over the CONUS using STAGE IV and various IMERG products, each presented in a row. Average intensity is calculated from three years (2015 to 2017) of data.

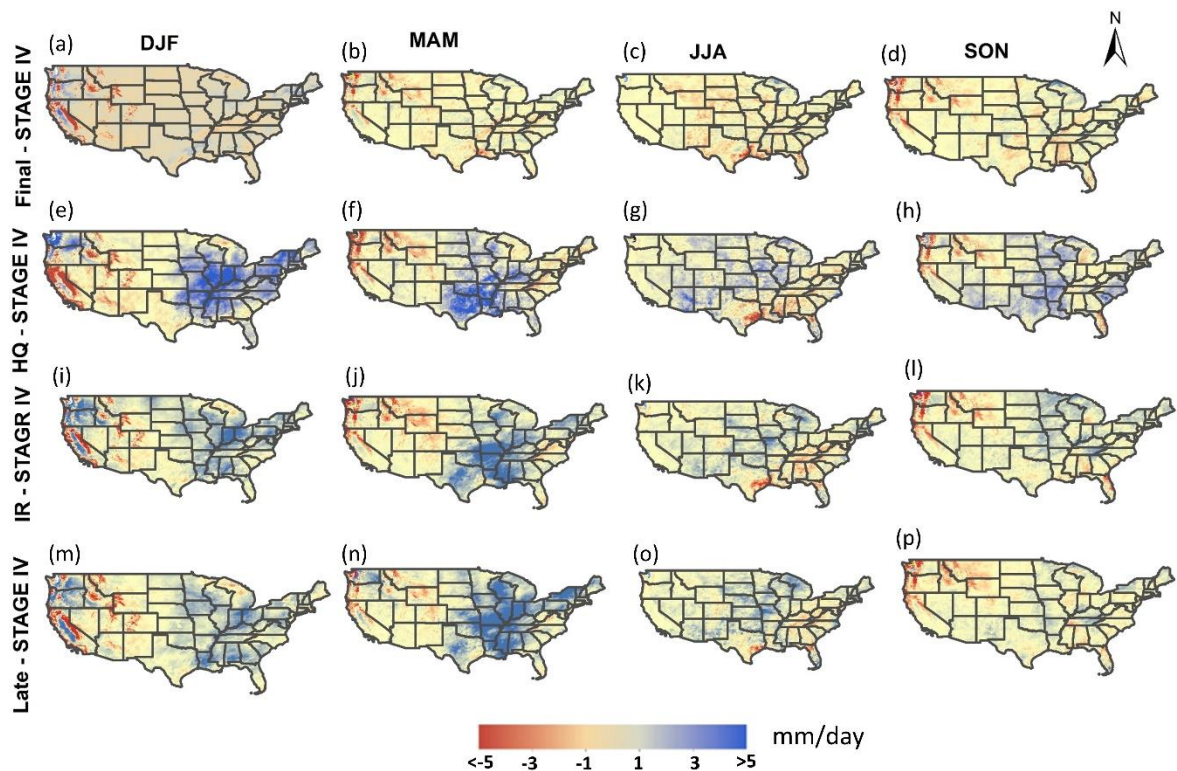


Figure 2: Seasonal spatial distribution of the difference in mean daily precipitation between IMERG products and STAGE IV for three years (2015-2017). Average intensity includes zero precipitation rate. Each row is labeled with IMERG product minus Stage IV data

Figure 3 shows a precipitation event on December 29, 2017, over the Washington and Oregon states (**Figure 3a**). This event contains both rain and snowfall, rainfall mainly occurs over the western and southern parts and snowfall occurs over the north-eastern part of the study area as can be seen from the liquid probability (**Figure 3h**) and Wet-bulb temperature (**Figure 3g**) maps. The snow- and ice-covered surfaces are also shown in **Figure 3i** using the Autosnow product. Stage IV precipitation map (**Figure 3b**) is used as a comparison reference. The white areas in **Figure 3b** represent regions with no data either because it is over the ocean located in the western part of the region (Stage IV has no coverage over the ocean) or missing radar data inland (**Figure 3a**). As discussed in Section 2, IMERG-Final and IMERG-Late are produced by combining the IR-based and PMW-based precipitation estimates. IMERG-Final bias-adjusts IMERG-Late using rain-gauges at a monthly scale. While the bias adjustment is often effective at a monthly scale, it does not necessarily improve the product at a sub-monthly scale, which could be a reason for the considerable difference between the IMERG-Final and STAGE IV products. IMERG-Late uses IR-based precipitation over snow and ice surfaces and when PMW overpasses are far in time from each other. Therefore, over the northeast, IR-precipitation is directly used in IMERG-Late products as can be inferred from **Figure 3d-3j**. Over the rainfall area, IMERG-Late is produced by combining the IMERG-IR and IMERG-HQ products as can be inferred from **Figures 3e** and **3f**.

As discussed in the introduction section, there is a need for a more detailed analysis of the performance of IR (represented by IMERG-IR) and PMW (represented by IMERG-HQ) precipitation estimates over snow and ice surfaces that is investigated in the next section.

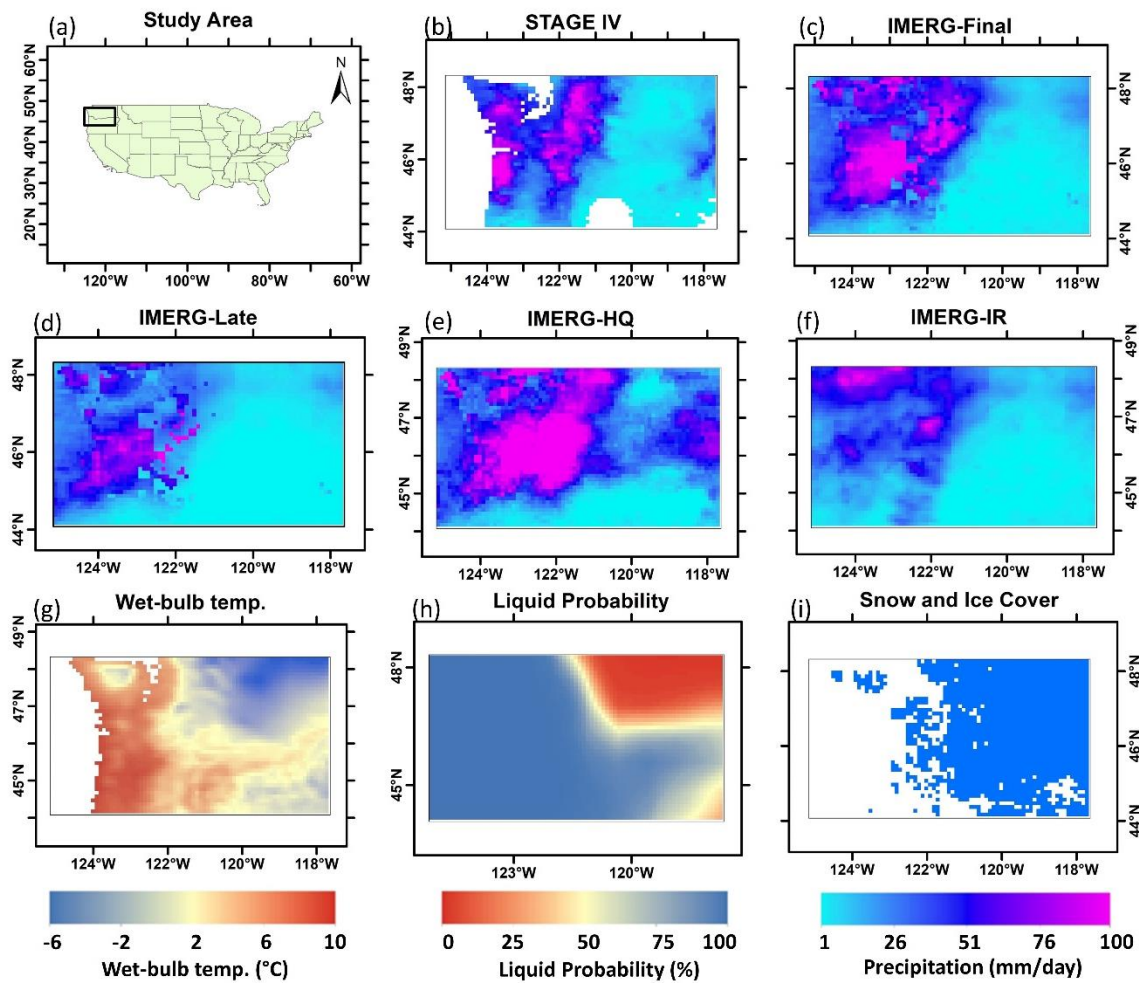


Figure 3: A precipitation event on Dec. 29 2017 for a selected study area shown in (a). Precipitation intensity from STAGE IV and IMERG products are shown in Panels b-f. Wet-bulb temperature, liquid probability, and snow- and ice-covered surfaces are shown in panels (g), (h), and (i), respectively.

3.2 Assessment of the IMERG products as a function of precipitation rate, surface, and environmental conditions

This section compares different IMERG products (e.g., IMERG-Final, IMERG-HQ, IMERG-IR and, IMERG-Late) with the Stage IV precipitation estimates by considering precipitation intensity, temperature effects (T_w ; that is also related to the precipitation phase), and surface type (i.e., surface with or without snow and ice cover). To investigate how the performance of the precipitation products varies with precipitation rate, five different thresholds (e.g., 0, 0.1, 0.3, 1 and, 2 mm/hr) were used and statistical scores were calculated for precipitation events identified by these thresholds (**Figure 4 and 5**). In other words, precipitation rates smaller than the thresholds were set to non-precipitating events and precipitation rates equal or greater than the threshold were considered as precipitation events.

Figure 4 shows the performance of different IMERG products over snow- and ice-covered surfaces as a function of T_w , plotted in the X-axis. Five different intensity thresholds are used to assess how products perform under events with higher precipitation rates. The top row shows the number of samples used in the analysis (**Figure 4a-4d**), followed by CC, POD, FAR, Bias, Volume Bias (VBias), and HSS in the lower rows. Note that the ideal score for Bias and VBias is one as discussed in Section 2. In general, IMERG-Late and IMERG-Final are similar to IMERG-IR, because by design IMERG uses IMERG-IR over snow and ice surfaces. IMERG-Final can be slightly different from IMERG-Late due

to the monthly bias adjustment utilized in IMERG-Final. Comparison of IMERG-IR with IMERG-HQ, as a function of T_w and precipitation intensity, enables us to assess whether the use of IR instead of PMW precipitation in IMERG is effective or not, especially at cold temperatures and over snow- and ice-covered regions. **Figure 4** shows that as T_w decreases CCs tend to decrease for all the products. While IMERG-HQ is more sensitive than IMERG-IR to changes in T_w , IMERG-HQ has higher CC than IMERG-IR at all T_w bins (**Figure 4e-4h**). This is the case for all precipitation intensities, but note that CC is generally lower once we focus on higher precipitation rates.

IMERG-HQ has a higher POD than IMERG-IR at T_w greater than $\sim -10^\circ\text{C}$, but at colder temperatures POD of IMERG-HQ is generally lower than IMERG-IR (**Figure 4i-4l**). Overall, both IMERG-IR and IMERG-HQ show reduction in POD at colder temperatures, although this is not necessarily the case for precipitation rates greater than 1 mm/hr, in which both IMERG-IR and IMERG-HQ show a slight increase in POD as T_w decreases. Note that precipitation is well detected (i.e., POD is about 1) by both products for precipitation intensities greater than 1 mm/hr. **Figure 4m-4p** show that FAR tends to be higher at colder T_w and lower at higher precipitation rates for both IMERG-IR and IMERG-HQ. IMERG-HQ, however, has a slightly lower FAR than IMERG-IR overall. This results in IMERG-HQ having generally higher HSS than IMERG-IR, almost at all T_w ranges, except at very cold temperatures (e.g., $T_w < -15^\circ\text{C}$) where the two products are comparable (**Figure 4z-4aa**). With respect to Bias (**Figure 4q-4t**), both IMERG-IR and IMERG-HQ show values around 1 for precipitation intensities greater than 0.1 mm/hr, but at colder temperatures (e.g., T_w less than -10°C) IMERG-HQ tends to under detect (Bias<1) and IMERG-IR tends to over detect (Bias>1) precipitation occurrences. When precipitation intensities lower than 0.1 mm/hr are included in the analysis, both IMERG-IR and IMERG-HQ show larger Bias values, suggesting that the products tend to have large false detection of light precipitation as can also be noticed from the FAR plots. This is not necessarily the case for IMERG-HQ for $T_w < -10^\circ\text{C}$ where the products tend to under detect precipitation incidences. VBias plots (**Figure 4u-4x**) show that over snow- and ice-covered surfaces both IMERG-HQ and IMERG-IR tend to overestimate precipitation amount, except for IMERG-HQ for light precipitation at cold temperatures (e.g., T_w less than -12°C). This is consistent with previous studies showing that PMW tends to underestimate light precipitation over cold surfaces [13]. Clearly, the impact of PMW underestimation is larger in high latitude regions, where light precipitation is dominant [37, 38]. Note that for precipitation rates higher than 1 mm/hr, there are large overestimations by both IMERG-HQ and IMERG-IR at colder temperatures that might be related to the confusion of the retrieval methods over snow- and ice-covered surfaces. As can be seen, the bias adjustment employed in IMERG-Final can only slightly improve the IMERG-Late product, likely because the adjustment is performed at a monthly scale, while statistical scores shown in **Figure 4** are based on the instantaneous matchups. Overall, results show that IMERG-HQ may outperform IMERG-IR over snow and ice surfaces, at least over the CONUS, suggesting that unconditional use of IR as a replacement for IMERG-HQ over snow and ice surfaces needs to be revisited.

Figure 5 is similar to **Figure 4**, but it is over snow- and ice-free surfaces. Here IMERG-Late and IMERG-Final mainly follow IMERG-HQ scores rather than IMERG-IR. However, the impact of IMERG-IR on IMERG-Late (thus IMERG-Final) can be seen (at least in the CC plot), as IR is still used in IMERG-Late when the time-distance between successive PMW overpasses is large. It can be seen that both IR and PMW products tend to have better scores over snow- and ice-free than over snow-and ice-covered surfaces and both show worse skill scores as T_w decreases. Furthermore, IMERG-HQ tends to outperform IMERG-IR almost regardless of the T_w and intensity ranges. Both products also show higher skill scores when light precipitation is removed from the analysis, except for CC and VBias. Both IMERG-IR and IMERG-HQ show large VBias when precipitation intensities greater than 1 mm/hr are assessed at lower temperatures, but VBias is clearly larger for IR at T_w less than 5°C . At T_w less than 0°C , IMERG-HQ shows slight underestimation for events that include light precipitation. Overall, it appears that the use of IMERG-IR

performs worse than IMERG-HQ over snow- and ice-free surfaces. Note that IMERG-HQ does not include morphed PMW estimates, so the outcomes of this analysis do not apply to the morphed PMW estimates that may show less skill at a longer time distance from the time of PMW observations [39].

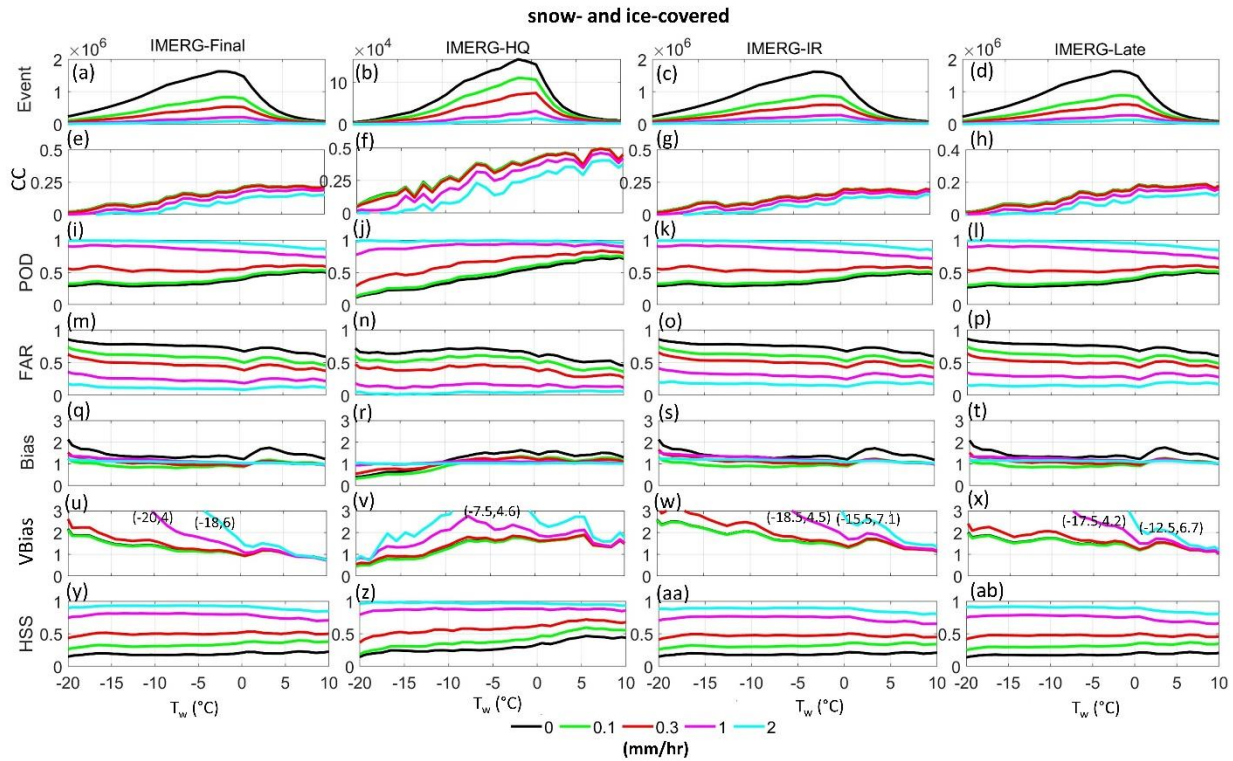


Figure 4: Comparison of different IMERG products (from left to right column: Final, HQ, IR and, Late) with Stage IV data for three years (2015-2017) over CONUS over snow- and ice-covered surfaces. Orographic effects are excluded using mountains mask. The X-axis indicates wet-bulb temperature 1-degree bins. Each color indicates the threshold for rainfall intensity from 0 mm/hr (in black) to 2 mm/hr (in cyan). The number of events for each threshold is shown in the first row. The second row to the seventh row shows statistical indices for correlation coefficient (CC), probability of distribution (POD), false alarm ratio (FAR), bias, volume bias (VBias) and, Heidke skill score (HSS).

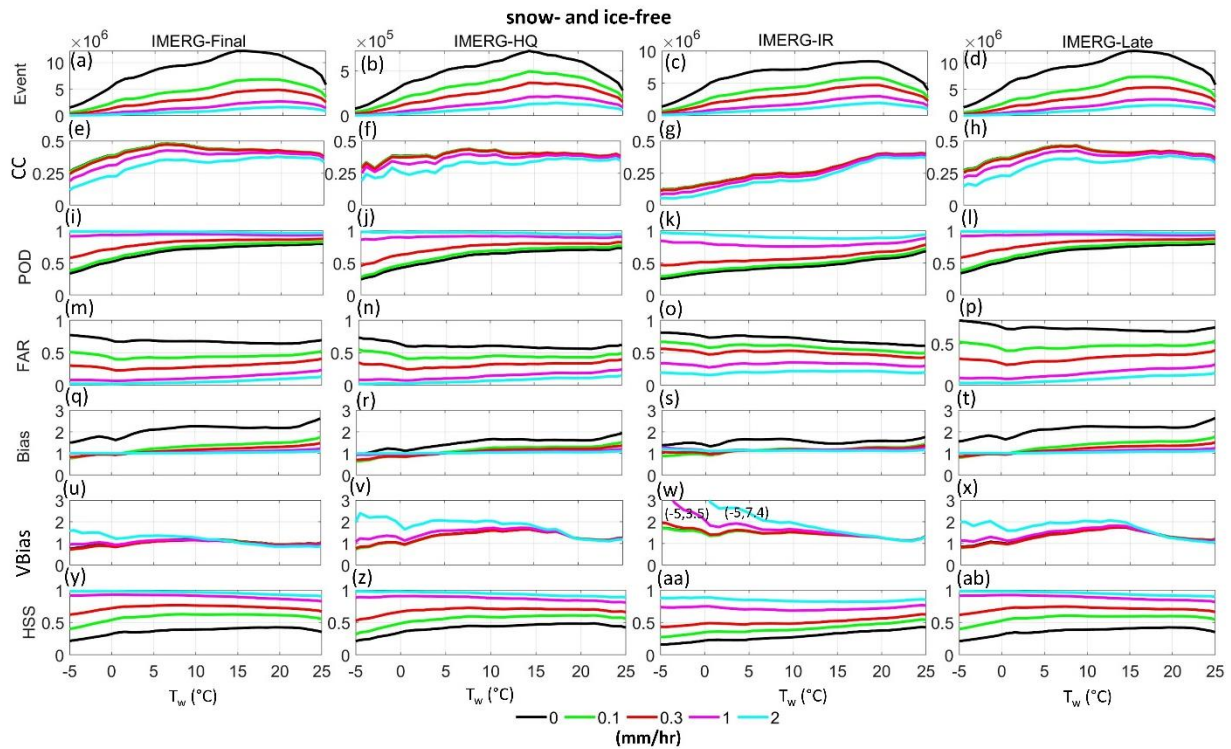


Figure 5: Similar to figure 4, but for snow- and ice-free surfaces.

Figure 6 facilitates comparing IMERG-IR with IMERG-HQ as a function of T_w over snow- and ice-covered (left-column) and snow- and ice-free (right-column) surfaces. Here, a fixed threshold of 0.3 mm/hr is used for precipitation delineation. It can be seen that IMERG-HQ outperforms IMERG-IR over snow- and ice-free surfaces across all temperatures, although it shows slightly worse Bias and VBias than IMERG-IR for $T_w \sim 5^\circ\text{C}$ (**Figure 6h**) and T_w between 10°C and 20°C (**Figure 6j**), respectively. Over snow- and ice-covered surfaces IMERG-HQ outperforms IMERG-IR in terms of CC (**Figure 6a**) and FAR (**Figure 6e**), regardless of T_w . However, IMERG-IR shows higher POD and HSS than IMERG-HQ for $T_w < \sim 12^\circ\text{C}$ (**Figure 6c** and **6k**), but this is along with higher FAR and Bias (**Figure 6e** and **6g**). **Figure 6i** shows a significantly large VBias for IMERG-IR, especially for $T_w < -8^\circ\text{C}$.

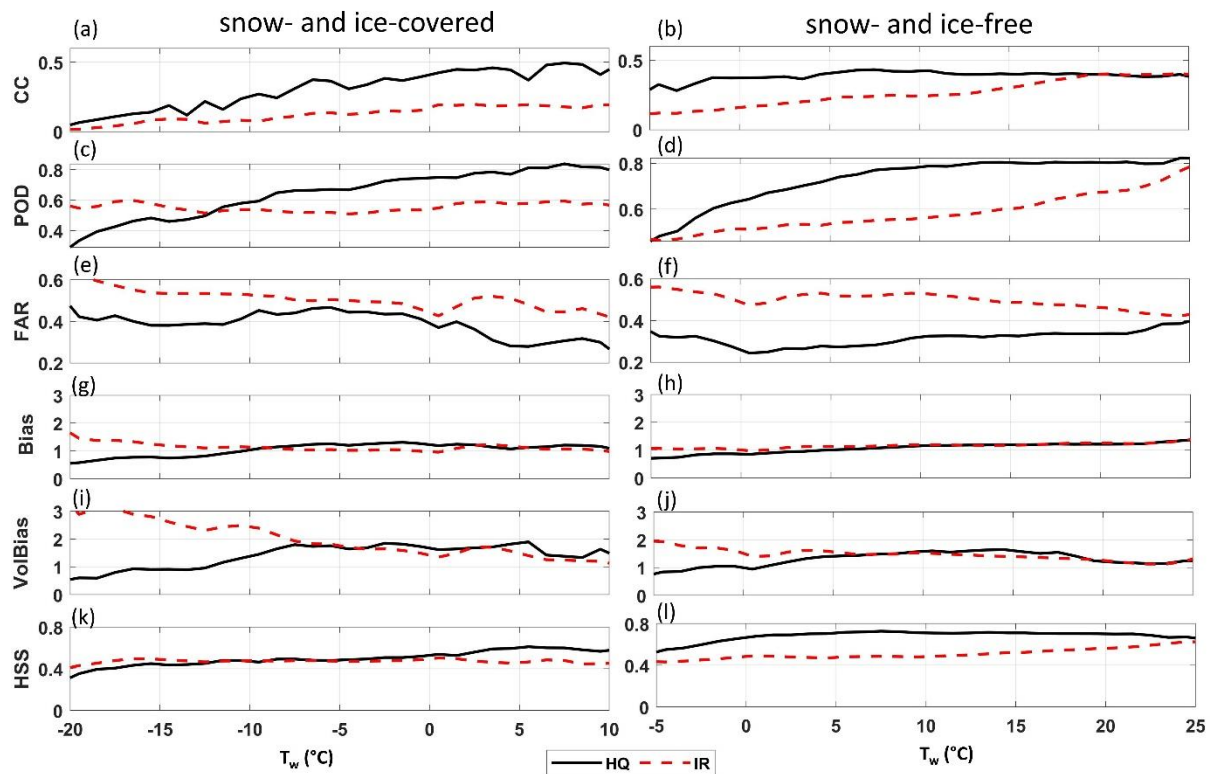


Figure 6: Comparison of IMERG-HQ and IMERG-IR over snow- and ice-covered and snow- and ice-free surfaces using three years (2015-2017) of data over CONUS. Stage IV is used as a reference. Dashed and solid lines represent IMERG-IR and IMERG-HQ, respectively. Precipitation is delineated from no precipitation using a threshold of 0.3 mm/hr. Orographic effects are excluded using mountains mask.

3.3 Performance of individual PMW precipitation estimates

In the previous section precipitation estimates from a combination of PMW sensors (i.e., through IMERG-HQ) were compared with precipitation estimates from geostationary IR observations. However, IMERG-HQ is composed of precipitation estimates from several PMW sensors that are identified in the IMERG output fields. Here the performance of individual PMWs is compared with IMERG-IR as a function of T_w and using Stage IV as reference. Similar to the previous section, analysis is conducted separately on snow- and ice-covered and snow- and ice-free surfaces. Timespan, study area, and procedure are the same as in the last section and a threshold of 0.3 mm/hr is used to delineate precipitation from no-precipitation. Orographic effects are excluded using mountains mask. Note that the analysis is based on sensor type, not satellite, so if a sensor is available on more than one platform a combination of the sensors is used. Furthermore, IMERG prefers PMW imagers over PMW sounders, so if they coincide, precipitation estimates from PMW imagers are used in IMERG-HQ (Huffman et al. 2020). Based on **Figure 7**, and over snow- and ice-covered surfaces, the following observations are highlighted: (1) all PMW sensors, except AMSR-2, have better CC than IMERG-IR regardless of T_w (**Figure 7c**), and AMSR-2 shows the lowest CC among all the studied products for $T_w < -5^\circ\text{C}$, (2) SSMIS shows the best and AMSR-2 shows the worst POD among all the PMW sensors, IMERG-IR has better POD than SSMIS for $T_w < -10^\circ\text{C}$, better POD than GMI, MHS and ATMS for $T_w < -3^\circ\text{C}$, and better POD than AMSR-2 for $T_w < 5^\circ\text{C}$ (**Figure 7e**).

POD of AMSR-2 is almost zero for $T_w < -10^\circ\text{C}$, (3) for FAR (**Figure 7g**), IMERG-IR has higher FAR than all PMW sensors except AMSR-2 that shows higher FAR than IMERG-IR for $T_w < 2^\circ\text{C}$ or generally snowfall, (4) for Bias, PMW and IMERG-IR show good skill (i.e., Bias near 1) for $T_w > 0^\circ\text{C}$, but for $T_w < 0^\circ\text{C}$, IMERG-IR tends to show Bias values greater

than one and PMW sensors tend to show Bias values smaller than 1. The Biases get worse as T_w decreases. AMSR-2 has Bias values greater than 1 for $T_w > -10^\circ\text{C}$, but Bias values less than one for $T_w < -10^\circ\text{C}$, (5) in terms of volume Bias, individual sensors tend to perform differently, but all of them overestimate Stage IV precipitation amount for $T_w > -15^\circ\text{C}$, except GMI and AMSR-2 that tend to underestimate for $T_w < -5^\circ\text{C}$. The overestimation of IMERG-IR is more significant than all PMW sensors at $T_w < -8^\circ\text{C}$ at which AMSR-2's VBias reaches almost zero. The increase of AMSR-2 VBias at $T_w < -15$ could be due to unstable sampling (note that AMSR-2 has the lowest sampling among all other sensors; **Figure 7a**), (6) as an overall score for precipitation detection, HSS suggests that IR is superior to the individual PMW sensors for $T_w < -10^\circ\text{C}$ (2°C for AMSR-2), although SSMIS shows lower HSS than IMERG-IR only for $T_w < -15^\circ\text{C}$, AMSR-2 show almost no skill at $T_w < -15^\circ\text{C}$. One reason for the relatively poor performance of AMSR-2 compared to other PMW products could be the lack of high frequency (e.g., greater than 90 GHz) and also sounding channels in AMSR-2 compared to other PMW sensors. It has been shown that high-frequency MW and sounding channels are valuable in retrieving snowfall and can be used to mask out scattering signals from snow- and ice-covered surfaces that might otherwise interfere with scattering signals from ice particles in clouds (e.g., Skofronick-Jackson, Kulie [40]).

PMW sensors perform generally better over snow- and ice-free than over snow- and ice-covered surfaces. PMW sensors have higher CC than IMERG-IR for $T_w < 15^\circ\text{C}$, but at warmer T_w they are relatively comparable (**Figure 7d**). Over snow- and ice-free surfaces, PMW sensors tend to have higher POD than IMERG IR (except for $T_w < 0^\circ\text{C}$ at which PMW sensors' POD decreases with reduction in T_w ; **Figure 7f**), IMERG-IR shows higher FAR than PMW sensors for all temperatures (**Figure 7h**), and both Bias and VBias of PMW sensors and IMERG-IR are fairly comparable for $T_w > 0^\circ\text{C}$. At $T_w < 0^\circ\text{C}$, IMERG-IR tends to show a slight overestimation in Bias (**Figure 7j**) and large overestimation in VBias (**Figure 7l**), but PMW sensors tend to underestimate both precipitation occurrence (Bias) and amount (VBias) at colder temperatures. As an overall metric, HSS suggests that IMERG-IR tends to outperform PMW sensors at $T_w < 0^\circ\text{C}$ for AMSR-2 and $T_w < -4^\circ\text{C}$ for other sensors over snow- and ice-free surfaces.

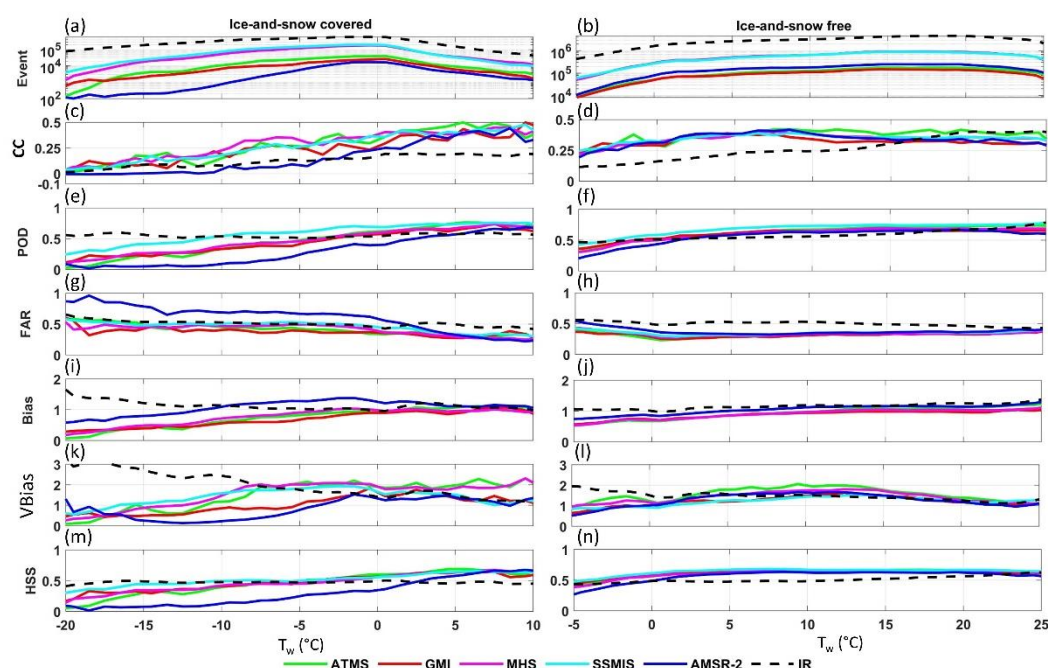


Figure 7: Comparison of precipitation estimates from different passive microwave sensors (used in IMERG-HQ) with IMERG-IR using STAGE IV data as a reference. Three years (2015-2017) of data over CONUS are used. Orographic effects are excluded using mountains mask.

4. Discussion

Although, performance of PMW sensors over snow- and ice-covered surfaces has been heavily studied, IR-based precipitation measurements have not been fully discovered. Previous studies show poor performance of PMW-based precipitation estimation as the temperature decreases especially in the presence of the snow or ice on the surface. This study confirms this issue related to PMW however, it also indicates majority of PMW sensors could outperform IR-based precipitation in cold temperatures over snow- and ice-covered surfaces. It should be mentioned that assessment of PMW- and IR-based precipitation in higher latitudes where environment and type of precipitation could be very different (light precipitation is dominant over high latitudes) from CONUS might be different from our results and could be further investigated.

5. Conclusions

Accurate estimation of precipitation is important for water cycle studies and various hydrologic applications. A long-standing challenge for remote sensing products has been an estimation of precipitation in cold regions, especially over snow and ice surfaces. Here using three years of Stage IV data (2015-2017) over the CONUS, the performance of various products of IMERG is investigated as a function of near-surface wet-bulb temperature (used for precipitation phase detection), precipitation intensity, and surface type (i.e., with and without snow and ice on the surface). The IMERG products include precipitation estimations from infrared (IR), combined passive microwave (PMW) sensors, and a combination of the precipitation estimate from IR and PMW sensors, that are either bias adjusted using in situ data (IMERG-Final) or not (IMERG-Late). In the analysis, steep mountainous regions were eliminated to reduce complexities that might be due to the orographic enhancement of precipitation.

Results show that precipitation estimates from PMW products generally have better statistics than IR over snow- and ice-free surfaces. Over snow- and ice-covered surfaces, PMW products (except AMSR-2) show a higher correlation coefficient (with Stage IV data) than IR. Both IR and PMW precipitation products tend to overestimate precipitation over snow and ice surfaces, but at colder temperatures (e.g., $T_w < -10^\circ\text{C}$) PMW products tend to underestimate while IR product continues to show large overestimations. With respect to precipitation occurrence, both PMW and IR products show considerably higher skill in capturing intense precipitation than light precipitation rates. PMW sensors outperform IR over snow- and ice-free surfaces and also show higher overall skill in detection precipitation occurrence over snow- and ice-covered regions, but not necessarily at T_w colder than -5°C . Generally, AMSR-2 performs worst and SMMIS performs best among the studied PMW sensors that also include GMI, MHS, and ATMS. The results suggest that the current approach of IMERG, replacing PMW with IR precipitation estimates over snow- and ice-covered surfaces, needs further investigations and has room for revisions.

Supplementary Materials: The following are available online at www.mdpi.com/xxx/s1, Figure S1: Mask of high and high-scatter mountains based on K3 mountain maps.

Funding: Financial support was made available from NASA MEaSUREs (NNH17ZDA001N-MEASURES) and NASA [Weather and Atmospheric Dynamics](#) (NNH19ZDA001N-ATDM) grants.

Data Availability Statement: The NCEP-STAGE-IV product was obtained from Earth Observing Laboratory (EOL) website at <https://data.eol.ucar.edu/>. IMERG products are available 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/#!/home>) and K3 mountain maps obtained from USGS website at <https://rmgsc.cr.usgs.gov/outgoing/ecosystems/Global/>.

References

1. Berg, W., T. L'Ecuyer, and C. Kummerow, *Rainfall Climate Regimes: The Relationship of Regional TRMM Rainfall Biases to the Environment*. Journal of Applied Meteorology and Climatology, 2006. **45**(3): p. 434-454.
2. Lebsock, M.D. and T.S. L'Ecuyer, *The retrieval of warm rain from CloudSat*. Journal of Geophysical Research: Atmospheres, 2011. **116**(D20).
3. Adhikari, A., et al., *Comparative Assessment of Snowfall Retrieval From Microwave Humidity Sounders Using Machine Learning Methods*. Earth and Space Science, 2020. **7**(11): p. e2020EA001357.
4. Ferraro, R.R., et al., *An Evaluation of Microwave Land Surface Emissivities Over the Continental United States to Benefit GPM-Era Precipitation Algorithms*. IEEE Transactions on Geoscience and Remote Sensing, 2013. **51**(1): p. 378-398.
5. Ehsani, M.R., et al., *Assessment of the Advanced Very High-Resolution Radiometer (AVHRR) for Snowfall Retrieval in High Latitudes Using CloudSat and Machine Learning*. Journal of Hydrometeorology, 2021.
6. Hsu, K.-I., et al., *Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks*. Journal of Applied Meteorology, 1997. **36**(9): p. 1176-1190.
7. Hong, Y., et al., *Precipitation Estimation from Remotely Sensed Imagery Using an Artificial Neural Network Cloud Classification System*. Journal of Applied Meteorology, 2004. **43**(12): p. 1834-1853.
8. Huffman, G.J., et al., *Integrated Multi-satellite Retrievals for the Global Precipitation Measurement (GPM) Mission (IMERG)*, in *Satellite Precipitation Measurement: Volume 1*, V. Levizzani, et al., Editors. 2020, Springer International Publishing: Cham. p. 343-353.
9. Xie, P., et al., *Reprocessed, Bias-Corrected CMORPH Global High-Resolution Precipitation Estimates from 1998*. Journal of Hydrometeorology, 2017. **18**(6): p. 1617-1641.
10. Kubota, T., et al., *Global Precipitation Map Using Satellite-Borne Microwave Radiometers by the GSMaP Project: Production and Validation*. IEEE Transactions on Geoscience and Remote Sensing, 2007. **45**(7): p. 2259-2275.
11. You, Y., et al., *Evaluation of V05 Precipitation Estimates from GPM Constellation Radiometers Using KuPR as the Reference*. Journal of Hydrometeorology, 2020. **21**(4): p. 705-728.
12. Behrangi, A., et al., *What does CloudSat reveal about global land precipitation detection by other spaceborne sensors?* Water Resources Research, 2014. **50**(6): p. 4893-4905.
13. Zhang, J., L.-F. Lin, and R.L. Bras, *Evaluation of the Quality of Precipitation Products: A Case Study Using WRF and IMERG Data over the Central United States*. Journal of Hydrometeorology, 2018. **19**(12): p. 2007-2020.
14. Behrangi, A., et al., *On distinguishing snowfall from rainfall using near-surface atmospheric information: Comparative analysis, uncertainties and hydrologic importance*. Quarterly Journal of the Royal Meteorological Society, 2018. **144**(S1): p. 89-102.
15. Behrangi, A., et al., *Evaluating the Utility of Multispectral Information in Delineating the Areal Extent of Precipitation*. Journal of Hydrometeorology, 2009. **10**(3): p. 684-700.
16. Shige, S., et al., *Improvement of TMI Rain Retrievals in Mountainous Areas*. Journal of Applied Meteorology and Climatology, 2013. **52**(1): p. 242-254.
17. Randel, D.L., C.D. Kummerow, and S. Ringerud, *The Goddard Profiling (GPROF) Precipitation Retrieval Algorithm*, in *Satellite Precipitation Measurement: Volume 1*, V. Levizzani, et al., Editors. 2020, Springer International Publishing: Cham. p. 141-152.
18. Huffman, G.J., et al., *NASA Global Precipitation Measurements (GPM) Integrated Multi-satellite Retrievals for GPM (IMERG), Algorithm Theoretical Basis Document (ATBD) Version 06* 2019.
19. Cui, W., et al., *Can the GPM IMERG Final Product Accurately Represent MCSsB—> Precipitation Characteristics over the Central and Eastern United States?* Journal of Hydrometeorology, 2020. **21**(1): p. 39-57.

20. Lin, Y. and K.E. Mitchel, *The NCEP Stage II/IV hourly precipitation analysis*, in *19th Conf. on Hydrology*. 2005: San Diego, CA, Amer. Meteor. Soc.
21. Nelson, B.R., et al., *Assessment and Implications of NCEP Stage IV Quantitative Precipitation Estimates for Product Intercomparisons*. Weather and Forecasting, 2016. **31**(2): p. 371-394.
22. Cocks, S.B., et al., *MRMS QPE Performance during the 2013/14 Cool Season*. Journal of Hydrometeorology, 2016. **17**(3): p. 791-810.
23. Wang, Y.-H., et al., *A Wet-Bulb Temperature-Based Rain-Snow Partitioning Scheme Improves Snowpack Prediction Over the Drier Western United States*. Geophysical Research Letters, 2019. **46**(23): p. 13825-13835.
24. Sims, E.M. and G. Liu, *A Parameterization of the Probability of SnowB—[Rain Transition]*. Journal of Hydrometeorology, 2015. **16**(4): p. 1466-1477.
25. Romanov, P., *GLOBAL 4KM MULTISENSOR AUTOMATED SNOW/ICE MAP (GMAI) ALGORITHM THEORETICAL BASIS DOCUMENT*. 2016.
26. Kapos, V., et al., *Developing a map of the world's mountain forests*. 2000, CABI Publishing: Wallingford. p. 4-19.
27. Körner, C., J. Paulsen, and E.M. Spehn, *A definition of mountains and their bioclimatic belts for global comparisons of biodiversity data*. Alpine Botany, 2011. **121**(2): p. 73.
28. Karagulle, D., et al., *Modeling global Hammond landform regions from 250-m elevation data*. Transactions in GIS, 2017. **21**(5): p. 1040-1060.
29. Roger, S., et al., *A New High-Resolution Map of World Mountains and an Online Tool for Visualizing and Comparing Characterizations of Global Mountain Distributions*. Mountain Research and Development, 2018. **38**(3): p. 240-249.
30. Wen, Y., et al., *Evaluation of MRMS Snowfall Products over the Western United States*. Journal of Hydrometeorology, 2017. **18**(6): p. 1707-1713.
31. Behrangi, A., K.J. Bormann, and T.H. Painter, *Using the Airborne Snow Observatory to Assess Remotely Sensed Snowfall Products in the California Sierra Nevada*. Water Resources Research, 2018. **54**(10): p. 7331-7346.
32. Wen, Y., et al., *How well were the early 2017 California Atmospheric River precipitation events captured by satellite products and ground-based radars?* Quarterly Journal of the Royal Meteorological Society, 2018. **144**(S1): p. 344-359.
33. Mei, Y., et al., *Error Analysis of Satellite Precipitation Products in Mountainous Basins*. Journal of Hydrometeorology, 2014. **15**(5): p. 1778-1793.
34. Arabzadeh, A., et al., *Global Intercomparison of Atmospheric Rivers Precipitation in Remote Sensing and Reanalysis Products*. Journal of Geophysical Research: Atmospheres, 2020. **125**(21): p. e2020JD033021.
35. Behrangi, A., et al., *On the Quantification of Atmospheric Rivers Precipitation from Space: Composite Assessments and Case Studies over the Eastern North Pacific Ocean and the Western United States*. Journal of Hydrometeorology, 2016. **17**(1): p. 369-382.
36. Guan, B., D.E. Waliser, and F.M. Ralph, *An Intercomparison between Reanalysis and Dropsonde Observations of the Total Water Vapor Transport in Individual Atmospheric Rivers*. Journal of Hydrometeorology, 2018. **19**(2): p. 321-337.
37. Song, Y., A. Behrangi, and E. Blanchard-Wrigglesworth, *Assessment of Satellite and Reanalysis Cold Season Snowfall Estimates Over Arctic Sea Ice*. Geophysical Research Letters, 2020. **47**(16): p. e2020GL088970.
38. Behrangi, A., A.S. Gardner, and D.N. Wiese, *Comparative analysis of snowfall accumulation over Antarctica in light of Ice discharge and gravity observations from space*. Environmental Research Letters, 2020. **15**(10): p. 104010.
39. Ayat, H., et al., *Are Storm Characteristics the Same When Viewed Using Merged Surface Radars or a Merged Satellite Product?* Journal of Hydrometeorology, 2021. **22**(1): p. 43-62.
40. Skofronick-Jackson, G., et al., *Satellite Estimation of Falling Snow: A Global Precipitation Measurement (GPM) Core Observatory Perspective*. Journal of Applied Meteorology and Climatology, 2019. **58**(7): p. 1429-1448.