Evaluation of GPROF V05 Precipitation Retrievals under Different Cloud Regimes

JACKSON TAN, a,b NAYEONG CHO, a,b LAZAROS OREOPOULOS, b AND PIERRE KIRSTETTER c,d,e

a University of Maryland, Baltimore County, Baltimore, Maryland
b Earth Sciences Division, NASA Goddard Space Flight Center, Greenbelt, Maryland
c School of Meteorology, University of Oklahoma, Norman, Oklahoma
d School of Civil Engineering and Environmental Science, University of Oklahoma, Norman, Oklahoma
e NOAA/Severe Storms Laboratory, Norman, Oklahoma

ABSTRACT: Precipitation retrievals from passive microwave satellite observations form the basis of many widely used precipitation products, but the performance of the retrievals depends on numerous factors such as surface type and precipitation variability. Previous evaluation efforts have identified bias dependence on precipitation regime, which may reflect the influence on retrievals of recurring factors. In this study, the concept of a regime-based evaluation of precipitation from the Goddard profiling (GPROF) algorithm is extended to cloud regimes. Specifically, GPROF V05 precipitation retrievals under four different cloud regimes are evaluated against ground radars over the United States. GPROF is generally able to accurately retrieve the precipitation associated with both organized convection and less organized storms, which collectively produce a substantial fraction of global precipitation. However, precipitation from stratocumulus systems is underestimated over land and overestimated over water. Similarly, precipitation associated with trade cumulus environments is inaccurately retrieved due to retrieval limitations. These findings highlight the important role of atmospheric states on the accuracy of satellite precipitation and the potential of cloud regimes for categorizing the global atmosphere.

SIGNIFICANCE STATEMENT: To understand how the accuracy of satellite precipitation depends on weather conditions, we compare the satellite estimates of precipitation against ground radars in the United States, using cloud regimes as a proxy for different recurring atmospheric systems. Consistent with previous studies, we found that errors in the satellite precipitation vary under different regimes. Satellite precipitation is, reassuringly, more accurate for storm systems that produce intense precipitation. However, in systems that produce weak or isolated precipitation, the errors are larger due to retrieval limitations. These findings highlight the important role of atmospheric states on the accuracy of satellite precipitation and the potential of cloud regimes for categorizing the global atmosphere.

KEYWORDS: Clouds; Precipitation; Rainfall; Climate classification/ regimes; Cloud retrieval; Microwave observations; Radars/Radar observations; Remote sensing; Satellite observations; Clustering

1. Introduction

The intimate connection between clouds and precipitation reflects their integral role in the global water and energy cycle. To quantify these relationships, satellite observations are an invaluable resource for understanding these processes at a global scale. In particular, gridded satellite products are useful due to their completeness in coverage and consistency in processing. Gridded satellite precipitation products are often composed of precipitation retrievals from passive microwave (PMW) observations and use other sources only when PMW observations are not available. PMW precipitation retrievals are prioritized first because the microwave signal interacts with precipitation-sized hydrometeors and not cloud-sized hydrometeors, so—compared to other sources such as infrared observations—PMW provides a more direct observation of the precipitation. Nevertheless, the PMW retrieval process is challenging due to the underconstrained nature of the problem (Stephens and Kummerow 2007), especially over land where, due to the high and varying surface emissivity, the retrieval process tends to rely on the scattering of upwelling energy emitted by the surface, cloud water, and liquid hydrometeors at higher frequencies. Over ocean, the surface generally has a lower and more polarized emissivity, which permits the use of the emission signal at lower frequencies. Due to the complexity of the retrievals, successful PMW algorithms often have to rely on parametric approaches. One such approach is the Goddard profiling (GPROF) algorithm (Kummerow et al. 2001, 2011, 2015; Randel et al. 2020), which...
provides PMW precipitation estimates that are ingested into products such as the Integrated Multi-satellite Retrievals for GPM (IMERG; Huffman et al. 2019, 2020). GPROF relies on Bayesian techniques to estimate the surface precipitation using databases that relate observed brightness temperatures to hydrometeor profiles and are conditioned on surface class, surface temperature, and total column water vapor.

Several evaluation efforts on GPROF have uncovered biases under different retrieval conditions, which include precipitation intensity, channel availability, footprint size, surface type, convective or stratiform nature of the precipitation, variability of precipitation within the satellite pixel, and the ice water content (Berg et al. 2006; Wolff and Fisher 2008, 2009; Carr et al. 2015; Kirstetter et al. 2014; Henderson et al. 2017; Petković and Kummerow 2017; Kidd et al. 2018; Tan et al. 2018; You et al. 2020). It has been shown that the errors from some of these conditions such as precipitation types propagate downstream and impact IMERG (Kirstetter et al. 2020). Of particular relevance here are Henderson et al. (2017) and Petković and Kummerow (2017), both of which classified the evaluation of GPROF by different precipitation regimes defined using the Precipitation Radar onboard the Tropical Rainfall Measuring Mission (TRMM) satellite (Elsaesser et al. 2010). Both studies found that retrievals within organized convective systems led to overestimations compared to the Precipitation Radar (Petković and Kummerow 2017) and a ground radar over the ocean (Henderson et al. 2017), which the latter further traced to widespread stratiform precipitation within the system. Henderson et al. (2017) also identified an underestimation of precipitation in convective events within both deep isolated regimes and organized convection. Furthermore, Petković and Kummerow (2017) found that GPROF underestimates the precipitation of shallow systems. Such a regime approach identifies recurring conditions under which retrieval errors manifest that would otherwise be masked in a broadscale evaluation.

In this study, we build upon these prior efforts to classify the evaluation of GPROF V05 retrievals against ground radars in the United States under different regimes. Specifically, we use Moderate Resolution Imaging Spectroradiometer (MODIS) cloud regimes (Cho et al. 2021), which categorize the global atmosphere into a discrete number of recurrent modes, such as organized convection and stratocumulus environments. Not only do MODIS cloud regimes provide a larger coverage than the TRMM Precipitation Radar regimes, but they also allow an extension of the evaluation to other sensors as well as more suppressed environments.

2. Data and methods

a. GPROF

In this study, we focus on precipitation retrievals from PMW sensors in the GPM constellation based on the GPROF V05 algorithm from 2015 to 2020 (Kummerow et al. 2001, 2011, 2015; Randel et al. 2020). These estimates are Level 2 products in the official NASA GPM suite, containing derived geophysical parameters at the resolution of individual sensor footprints (Kummerow 2017). GPROF is a physically based parametric retrieval algorithm that estimates the surface precipitation rate from the brightness temperature observed by passive microwave sensors (Kummerow et al. 2015). GPROF performs a conditional Bayesian retrieval of precipitation by comparing the observed brightness temperature to a priori databases according to the surface class, surface temperature, and total column water vapor. Surface classes are defined using self-similar microwave emissivity at a monthly time scale from Special Sensor Microwave Imager observations (Aires et al. 2011), supplemented with daily snow maps from NOAA’s AutoSnow product (Romanov et al. 2000). These a priori databases match brightness temperatures with associated surface precipitation rates and hydrometeor profiles from the Ku-band Dual-frequency Precipitation Radar (DPR) and the Combined Radar–Radiometer product (except over snow-covered surfaces where ground radar data are used), with adjustments to the threshold for cloud liquid water retrievals to match the probability of precipitation observed by the Cloud Profiling Radar on board CloudSat to account for drizzle.

The GPROF algorithm is applied to most of the PMW sensors in the GPM constellation of satellites. These sensors can be classified as conical-scanning or cross-track-scanning sensors, with the former having imaging and sometimes sounding channels and the latter having only sounding channels. For the purposes of precipitation retrievals, imaging channels (or imagers) are of low frequencies—suitable for detecting the emission signal of precipitation—while sounding channels (or sounders) are of high frequencies—suitable for detecting the scattering signal of precipitation. PMW precipitation retrievals over land can use the emission signal of the precipitation because the low emissivity of the ocean surface allows the emission signal from the precipitation to be more easily discriminated from the background. PMW precipitation retrievals over land generally rely on the scattering of the background microwave emission by the precipitation because the higher and more variable emissivity of the land surface makes the detection of the emission signal from the precipitation harder. It should be pointed out that GPROF generally uses the same channel information over both land and ocean, relying on the Bayesian process to pick out the precipitation information.

The imagers examined in this study are the GPM Microwave Imager (GMI) on board the GPM Core Observatory; the Advanced Microwave Scanning Radiometer 2 (AMSR2) on board the GCOM-W1 satellite; and the Special Sensor Microwave Imager/Sounder (SSMIS) on board the DMSP F16, F17, F18, and F19 satellites. The sounders examined in this study are the Microwave Humidity Sounder (MHS) on board the NOAA-18, NOAA-19, MetOp-A, and MetOp-B satellites; and the Advanced Technology Microwave Sounder (ATMS) on board the Suomi-NPP satellite. All these sensors except for GMI are on sun-synchronous orbits with equator-crossing times that are fixed (if fuel is expended to maintain its orbit) or slowly varying (if allowed to drift). While data from the TRMM Microwave Imager on board the TRMM satellite are available, they are excluded from this analysis.
due to the shortness of the record in the study period. Similarly, data from the Sondeur Atmosphérique du Profil d’Humidité Intertropicaire par Radiométrie on board the Megha-Tropiques satellite are excluded from this analysis because its low latitude orbit barely overlaps with the ground reference (see next section).

b. GV-MRMS

As a ground reference, we use a specially processed version of Multi-Radar Multi-Sensor (MRMS) product for GPM ground validation (henceforth referred to as GV-MRMS) (Kirstetter et al. 2012, 2014, 2015). The MRMS system is a gridded product by NOAA/NSSL based primarily on the WSR-88D radar network over the conterminous United States (Zhang et al. 2016). Reflectivity data are mosaicked onto a 3D grid with quality control for beam blockages and bright band. From the reflectivity structures and environmental fields, precipitation types (e.g., snow, warm stratiform rain, convective rain) are determined using physically based heuristic rules, and a corresponding reflectivity–precipitation relationship is applied to estimate the surface precipitation rate. These precipitation rates are bias-corrected at the hourly time scale using gauge data from the Hydrometeorological Automated Data System and regional rain gauge networks. MRMS includes a radar quality index alongside each precipitation estimate (Zhang et al. 2011), providing a numerical value that reflects sampling and estimation uncertainty, such as beam interacting with orography and bright bands. Evaluation of MRMS shows better performances with the gauge correction and quality control based on the radar quality index (Chen et al. 2013; Kirstetter et al. 2015).

GV-MRMS applies further gauge-based corrections and quality control measures on MRMS so that its estimates can reliably be used to evaluate satellite retrievals. GV-MRMS covers the conterminous United States with a high spatial resolution of 0.01° and is available from June 2014 onward except for the periods of 1 January–31 May 2016 and 14 October 2020–31 December 2020 due to missing data. We select pixels with the best radar quality index, thus keeping only the highest quality GV-MRMS pixels in computing the areal averages and ensuring the highest possible quality reference (Gebregiorgis et al. 2017), though this introduces gaps in coverage especially over the mountainous western United States. The radar quality index filter also minimizes situations in which GV-MRMS misses shallow precipitation due to radar beam heights above the precipitation. Furthermore, we also mask pixels in which the gauge correction factor is outside the range of 0.5–2, thus removing pixels in which the radar and gauge estimates disagree greatly, which may indicate potential discrepancies in the ground measurements.

c. MODIS cloud regimes

In this study, we separate the evaluation of GPROF against GV-MRMS by MODIS cloud regimes (CR), which are also sometimes referred to as “weather states.” First developed using data from the International Satellite Cloud Climatology Project (Jakob and Tselioudis 2003; Rossow et al. 2005), CRs classify the globe into a number of mesoscale atmospheric systems at a 110-km grid every 3 h during daylight hours based on its cloud distributions (Tan and Oreopoulos 2019). Specifically, passive observations of cloud top pressure (CTP) and cloud optical thickness (COT) from the MODIS instruments are aggregated into joint-histograms on which a k-means clustering algorithm is applied. Two MODIS instruments are currently in operation, both on low-Earth-orbit satellites: Terra, with a nominal equator crossing time of 1030 local time, and Aqua, with a nominal equator crossing time of 1330 local time.

While previous MODIS CRs are derived using daily gridded cloud properties at 1° from the MOD08_D3 and MYD08_D3 products (Oreopoulos et al. 2016), we instead use the CR product defined in a 110-km equal-area grid at 3-h windows (Cho et al. 2021; Oreopoulos 2021). The equal-area grid has a reduced distortion in grid area at higher latitudes, which may lead to a more consistent representation of global cloud fields in the clustering process. More important for this study is the finer time window, which improves the collocation between MODIS, GPROF, and GV-MRMS observations, though sampling biases arise due to the nature of the orbits (see section 2e). This CR product also has stricter quality control measures, such as screening out joint histograms with fewer than 120 pixels, which are deemed insufficient sampling of the grid box. There are 11 global CRs plus a clear sky CR, available for the period January 2003–December 2020, and their occurrences on the equal-area grid are interpolated to a 1° equal-angle grid using the nearest neighbor technique. For this study we use only a subset of the CRs, and then only when they occur within the geographical and temporal parameters of the GV-MRMS and GPROF datasets.

d. GPM DPR

To characterize global precipitation properties of CRs and interrogate the GPROF database, we use the DPR on board the GPM Core Observatory (Iguchi 2020). Just like GPROF, DPR estimates are a Level 2 dataset in the official NASA GPM products. Specifically, we will use the V06 Ku-band radar retrievals, a cross-track scanning active instrument at 13.6 GHz with a swath width of 245 km. The instrument has a surface horizontal resolution of about 5 km at nadir, with vertical range resolution of 250 m sampled every 125 m. With the ability to determine the three-dimensional structure of precipitation, the DPR retrievals are expected to provide precipitation estimates—specifically for this study the estimated surface precipitation rate—that are more reliable than GPROF, as well identifying the convective/stratiform classification of the precipitating pixel and whether the precipitation is of a shallow nature. Nevertheless, limitations in the DPR algorithm, such as the inability to detect precipitation beneath a certain height (about 1–2 km from the surface) due to ground clutter and potential inaccuracies in the DPR convective/stratiform classification technique, may lead to some inaccuracies.

e. Methods

To collocate the data, we coarsened the GV-MRMS data and grid the Level 2 GPROF pixels to the 1°, 3-h grid of...
MODIS CRs. For GV-MRMS, we simply averaged the 0.01° pixels to the 1° grids, under the condition that at least half of the pixels were available. Figure 1 shows the geographical distribution of the 1° estimates GV-MRMS used in this study. For GPROF, we gridded the pixels to the 1° grid over the contiguous United States within 3-h windows based on the geolocation of the pixel footprint. While this may introduce some error on the ice-scattering signature aloft for footprints near the edge of the grid due to the nonzero footprint size and the parallax offset of the slanted viewing angle, the effect should be minimal at the 1° scale. This is done separately for each sensor, though this gridding does not distinguish between the same sensors on different satellite platforms (e.g., the sensor SSMIS on the satellites F16, F17, F18, and F19). The pixels within each 1° grid box at each 3-h window are then averaged if there are a minimum of 20 pixels to ensure spatial representativeness. Note that, with the exception of the GPM Core Observatory, these platforms are on a sun-synchronous orbit similar to the MODIS-bearing Aqua and Terra satellites, which may result in uneven sampling between sensors. This is most prominent example in SSMIS, which, despite being on board four different satellites during the analysis period, has the lowest sample size of all five sensors considered here. On the other hand, both AMSR2 on board the GCOM-W1 satellite and the MODIS instrument on board the Aqua satellite are in the A-Train formation, so the coincident observations resulted in a large sample size. With four years of data, the CRs generally have a sufficient sample size even for SSMIS to produce a representative result. It should also be noted that the accuracy of precipitation retrievals is highly dependent on resolution, so we expect relatively high scores at the 1° scale.

With the GPROF and GV-MRMS data on the same grid, we can compare the values and categorize the pairs into different CRs. In addition, as PMW retrievals of precipitation generally rely on different signals between surface types, the comparisons are further segregated by surface type into land and water. Here, water refers to large inland bodies of water and near-coast oceans; the range of ground radars limits the extent of GV-MRMS over oceans. However, with five different sensors categorized by 12 CRs and two different surface types, the analysis can become unwieldy. Therefore, to simplify the results, we will focus on four CRs: CR1 (organized convection), CR3 (tropical and midlatitude storms), CR7–9 (stratocumulus), and CR10 (fair weather or trade cumulus). For the purposes to this study, we consider CR7–9 as a single CR, since their cloud and precipitation characteristics, as well as the evaluation results when separated (not shown), are similar. Results from the excluded CRs either do not provide additional insights beyond those from these four CRs or suffer from low sample sizes that prevents a reliable and robust analysis.

To visualize and quantify the comparisons between GPROF and GV-MRMS, we will utilize scatter density diagrams, relative bias (bias divided by GV-MRMS mean), and the Pearson correlation coefficient (or simply, correlation). The bias metrics and the correlation are computed using all values, including zeroes. These will characterize the systematic and random errors in the GPROF retrievals. Here, we do not compute any detection metrics such as probability of detection or false alarm ratio because the contrast in spatial scales between the evaluation (1° ≈ 100 km) and the satellite footprints (~10 km) limit our ability to draw direct inference on detection skill. For example, a 1° grid box in which half the pixels are misses and half the pixels are false alarms may average out to a “successful detection.” Therefore, we restrict our analysis to retrievals of the precipitation rates themselves.

3. Results

a. Climatology of the cloud regimes

Figure 2 shows the mean joint histograms and geographical distributions of the four CRs (one of which being actually a combination of three original CRs) we selected for our analysis. These CRs are identified based on the dominant cloud field in the 1° grid box, though other cloud types may be present in smaller amounts as well. The mean joint histogram of CR1 indicates the presence of substantial deep convective clouds and widespread stratiform anvils, which is typically associated with organized tropical convection. While its global geographical distribution is concentrated in the tropics (Cho et al. 2021), its relative frequency of occurrence (RFO) in contiguous United States is not negligible, especially in the region adjacent to the Gulf of Mexico (Fig. 2). CR3 has a mean joint histogram with an abundance of thick stratiform anvil clouds but with a lower cloud top height, as well as an appreciably lower population of deep convective clouds. Representing storms less organized than CR1, it occurs both in the tropics and in the midlatitudes, so its presence over the domain of our analysis is higher. A notable feature of the mean joint histograms of CR1 and CR3 is the substantial presence of ice content aloft, which will give a strong microwave scattering signature that is typically associated with high surface precipitation rates in GPROF. The stratocumulus environments of CR7–9 have a mean joint histogram that shows a large occurrence of clouds with low cloud top heights over a range of optical thickness. Globally, it is dominant in
the marine stratocumulus regions; over the GV-MRMS domain, it has a high occurrence off the Pacific coast as well as in the northeast. In contrast, CR10 represents grid boxes with lower cloud fractions and optical thicknesses, likely corresponding to shallow cumulus fields. While it is mostly an oceanic regime, it does have some presence over land in the southeast of the domain. In contrast to CR1 and CR3, the general absence of clouds above the freezing level means that
CR7–9 and CR10 will likely have a weaker microwave scattering signature and therefore pose a greater challenge for GPROF.

To get a sense of the precipitation character associated with these four CRs, Fig. 3 shows the composites of precipitation properties—surface precipitation rate, convective/stratiform classification, and shallow/nonshallow classification—from DPR from 2015 to 2020. To ensure representativeness, these composites are drawn from grid boxes that have at least 100 DPR pixels. We should point out that the convective/stratiform and shallow/nonshallow classifications are defined only for precipitating pixels and thus do not reflect the absolute population of such precipitation across CRs. As expected, the organized convection represented by CR1 is associated with high precipitation rates and a large fraction of nonzero precipitation within the grid. Of the precipitating pixels in CR1, most are stratiform, consistent with previous observations on organized convective systems (e.g., Jakob and Schumacher 2008). Similarly, an overwhelming fraction of precipitation pixels is nonshallow. For the tropical and midlatitude storms represented by CR3, the grid-mean precipitation rates are lower; while both the mean and median are lower than those of CR1 by several factors, the 25th percentile is lower by more than an order of magnitude. CR3 has precipitating grid fractions that are slightly lower, but with a higher stratiform and nonshallow fraction at the same time. In contrast, CR7–9 are suppressed environments generally not associated with precipitation, though DPR struggles to detect the small particles associated with stratocumulus precipitation (Hayden and Liu 2018). Even so, precipitation, when present, appears to be generally of more convective nature and possesses a more shallow character. This suggests the presence of isolated shallow convection in CR7–9 aside from the stratiform drizzle common in stratocumulus clouds, which serves as a reminder that CRs are classified based on dominant cloud populations and thus may contain a substantial contribution from other cloud types and environments. Finally, CR10, similar to CR7–9, is a suppressed environment, though with a marginally higher grid-mean precipitation rate and precipitating fraction. However, its precipitating pixels make up the highest convective fraction of all four CRs, with approximately equal amounts of convective and stratiform pixels. Similarly, the precipitating pixels of CR10 form the highest shallow fraction of all four CRs. These DPR-based precipitation characteristics of the CRs may be different over land and water (supplementary Figs. 1 and 2 in the online supplemental material).
Specifically, except for CR7–9 due to the abovementioned reason, the fraction of nonzero precipitation within the grid box for all CRs is higher over water than over land, which in turn leads to a higher grid-mean precipitation rate. This suggests that our interpretation of the nature of the CRs based solely on cloud patterns may have different precipitation properties; it is possible that further including precipitation information in the definition of the regimes (Jin et al. 2021) can reduce such cloud–precipitation variability, though this is beyond the scope of this study. Furthermore, Leinonen et al. (2016) have shown that the CRs have substantial regional variability, which is evident when we limit the DPR precipitation characteristics over land to over the northeast United States and the southeast United States separately (supplementary Figs. 3 and 4). This regional variability likely reflects the diversity in the relationship between the cloud fields that define the CRs and other meteorological variables.

b. Evaluation of GPROF under different cloud regimes

CR1 is the regime associated with the highest precipitation rates, with a disproportionate contribution to the total precipitation despite its relatively low frequency of occurrence. Thus, it is important for retrieval algorithms to accurately capture the precipitation associated with CR1. In this respect, GPROF performs well at the 1° spatial resolution, both over land (Fig. 4) and over water (Fig. 5). The retrieved precipitation rates have high correlations (0.724–0.765 over land and 0.669–0.816 over water) indicating a low random error relative to other CRs in the retrieval process, though these numbers will likely be lower at finer resolutions. Such a result is expected because the intense precipitation associated with organized convection is easier to quantify than the weak, scattered precipitation of other regimes, consistent with the high ice content, and thus PMW scattering signal, implied by its mean joint histogram (Fig. 2). We note that the precipitation rates generally have greater spread in their scatter density diagrams and lower correlations over land than over water. The relative biases are small (0.21–0.31) compared to other CRs (e.g., 1.31–2.31 in CR7–9 over water), largely because they are normalized to the high precipitation rates associated with this system (see section 2e). However, it is worth noting that, despite a low relative bias, SSMIS underestimates the precipitation associated with CR1 over water compared to the other sensors, a possible consequence of its large footprint size, especially for the low frequencies used for retrievals over the water.

CR3 has a cloud pattern broadly similar to organized convection but with lower cloud tops; with an occurrence extending into the midlatitudes, CR3 may represent storms that are less organized or weaker compared to CR1. While it is associated with less precipitation, it has a higher frequency of occurrence and thus still a large impact on total precipitation. For CR3, GPROF retrievals are excellent both over land (Fig. 6) and over water (Fig. 7) at the 1° scale. Compared to the organized convection as represented by CR1, the retrievals generally have a slightly lower correlation (0.674–0.718 over land, 0.632–0.690 over water), consistent with wider spread at lower intensities in the density scatter diagrams. However, unlike CR1, the performance over land is not inferior to over water, even though both CRs broadly have substantial presence of ice content (Fig. 2). On the other hand, the biases of the sensors for this less organized storm regime over both land (0.02–0.14) and water (0.07–0.30) are similar to, or, in some cases, smaller than those of CR1.
CR7–9 represents regimes dominated by stratocumulus at various stages of development. Such environments have a low contribution to global precipitation (Tan et al. 2013; Tan and Oreopoulos 2019), but light precipitation mostly in the form of drizzle often occurs (Wood 2012). Over land, there is an underestimation of precipitation by most sensors, with SSMIS being the exception due to its tendency to overestimate at higher intensities (Fig. 8). The relative biases range falls between $-0.43$ and $-0.65$ except for SSMIS, which has a relative bias of 0.04. This underestimation by GPROF is likely due to the absence of ice hydrometeors and the weak scattering signature in the shallow precipitation associated with stratocumulus clouds, resulting in lower quantification of precipitation and hence an overall underestimation. It is unclear why SSMIS is unaffected by this issue, though we note that SSMIS has the lowest sample size of all five sensors.

On the other hand, over water, precipitation is substantially overestimated across all sensors, with relative biases of between 1.31 and 2.31; i.e., GPROF overestimates by a factor of more than two. While this overestimation is observed only
from a small sample size (Fig. 9), the differences in the GPROF and GV-MRMS means are statistically different with a $p$ value of less than 0.01 based on the Mann–Whitney $U$ test. The fact that PMW retrievals over land do not suffer the same bias as over water implies the presence of biases in the GPROF database, but with a sign contrary to expectations. Hayden and Liu (2018) found that DPR, used to construct the database, struggles to detect the small drop sizes associated with stratocumulus precipitation, which suggests that GPROF should have underestimated the precipitation of this stratocumulus environment. Therefore, there are likely other more dominant factors driving the bias of the PMW retrievals. One possible explanation is that the Bayesian nature of GPROF yields overestimates of low values and underestimates of high values (see, e.g., Petković et al. 2018 for an explanation). Indeed, this interpretation is consistent with GV-MRMS precipitation rates for CR7–9 being lower over water than over land (supplementary Fig. 5), although other factors likely also contribute to this overestimation. The correlations at the $1^\circ$ scale are relatively low both over land...
(0.446–0.680) and over water (0.341–0.565), possibly reflecting the challenges of retrieving PMW precipitation that is light, shallow, and often isolated. This is most severe in MHS (0.382) and ATMS (0.341) over water, likely exacerbated by their lack of low-frequency channels that are primarily used for emission-based retrievals.

The cloud pattern of CR10, which occurs predominantly over oceans of all latitudes, suggests fair weather or trade cumulus conditions, but perhaps also associated with stratocumulus clouds breaking up. Based on composites with DPR, CR10 is likely associated with isolated convection, both deep and shallow. Over land, all sensors except SSMIS underestimate light precipitation, resulting in an overall negative bias with relative biases ranging from −0.34 to −0.50 (Fig. 10), which can be attributed to the lack of ice-scattering signal for the shallow convective precipitation just as with CR7–9. This is further supported by the fact that the underestimations are generally confined to the lighter precipitation rates that isolated shallow precipitation would produce at the 1° scale. Though higher than in CR7–9, correlations are generally low.
which is possibly due to the isolated nature of the precipitation. Over water, there is a distinct difference in bias between imagers and sounders (Fig. 11). Specifically, while imagers (GMI, AMSR2, and SSMIS) have a slight overestimation in precipitation (relative biases of 0.34–1.07), sounders (MHS and ATMS) underestimate the precipitation rate (relative biases from −0.12 to −0.29). Furthermore, a discernibly larger degree of scatter can be seen for the sounders, which again may be attributed to the absence of low frequency channels.

c. Evaluation of GPM DPR under different cloud regimes

Since the GPROF relies primarily on DPR to construct its a priori database, we briefly investigate how DPR retrievals compare with GV-MRMS under the different CRs. Due to the limited sample size, we focus largely on the retrievals over land. Figure 12 shows that DPR almost always has higher correlations than GPROF; individual sensors may have correlations that exceed DPR, but such occasions are rare. In particular, DPR has higher correlations than GPROF GMI for all four CRs, suggesting that the apparently superior performance of some sensors over DPR may be attributed to variability due to differences in sampling. This is an entirely expected result since, in principle, GPROF can only be as good as its database, so its performance is capped by the quality of DPR. A more interesting picture emerges in the biases over land, where the biases in GPROF broadly reflect that of DPR: low magnitudes for CR1, even closer to zero for CR3, and appreciably negative for both CR7–9 and CR10. This correspondence suggests that, at least over land, GPROF inherits its biases from DPR. Over water, the small sample size makes an assessment more challenging, with the biases in DPR largely corresponding to those in GPROF but with greater variability (supplementary Fig. 6). A striking observation is the appreciably positive bias in DPR for CR7–9 over water, which may be a factor in the large positive biases of GPROF (Fig. 9), though the scatter diagrams suggest that the overestimation in GPROF is more severe. Therefore, it appears that DPR limitations propagate into GPROF, but other factors also affect GPROF retrievals such as in the case of stratocumulus environment over water.

4. Discussion and conclusions

This study evaluated GPROF against GV-MRMS at 1° and 3-h scales under different environmental conditions determined by four MODIS CRs and further segregated by surface type. Generally, GPROF is skillful in capturing the precipitation associated with both organized convection and weaker or less organized storms—with low relative biases and high correlations—which is reassuring given the substantial fraction of global precipitation coming from such systems. In contrast, precipitation from stratocumulus systems, in general a suppressed environment, is underestimated over land by most sensors—likely due to quantification issues in the absence of ice hydrometeors—and is overestimated over water by all sensors—possibly because of Bayesian averaging effects and/or a bias in the GPROF database. As for fair weather and trade cumulus conditions, the underestimation of light precipitation over land suggests again quantification issues of shallow rain, while over water imagers overestimate and sounders slightly underestimate, possibly due to the lack of low frequency channels in the latter. Additionally, despite the suppressed environments in both cases, retrievals under trade cumulus conditions, the underestimated light precipitation over land suggests again quantification issues of shallow rain, while over water imagers overestimate and sounders slightly underestimate, possibly due to the lack of low frequency channels in the latter. Additionally, despite the suppressed environments in both cases, retrievals under trade cumulus conditions generally suffer from slightly lower random error compared to retrievals under stratocumulus environments. A further investigation into DPR retrievals revealed that the biases in GPROF generally reflect those in

Fig. 11. As in Fig. 4, but for CR10 over water.
DPR, suggesting that limitations in DPR propagate into GPROF through the a priori database.

These results are largely consistent with those of Henderson et al. (2017) and Petković and Kummerow (2017). For example, we identified a positive bias in most of the sensors under organized convective systems (represented by CR1 in this study), both over land and over water. Similarly, consistent with Petković and Kummerow (2017), we found that GPROF underestimates shallow precipitation over land associated with stratocumulus environments (CR7–9) and trade cumulus environments (CR10), though we also found additionally that there is an overestimation over water under these environments. However, our finding that the less organized convection (CR3) has lower biases despite possessing a greater stratiform fraction contradicts the results of Henderson et al. (2017) who identified a positive bias for stratiform precipitation. A possible explanation for this discrepancy may be the difference in environment [the less organized convection of CR3 compared to the organized convection in Henderson et al. (2017)] and thus the contribution of other factors such as precipitation intensity. Another explanation for the discrepancy may be the difference in precipitation processes in the tropical region examined by Henderson et al. (2017) and the midlatitude region of this study. Our study also identified different biases between land and water for the trade cumulus environment, suggesting the importance of surface type in determining PMW retrieval performance for such isolated convective conditions.

One of the challenges in attributing the results to the retrieval process is the coarseness of the comparison, which is a constraint imposed by the 1° resolution of the CRs. This means that any inference about the performance of the GPROF retrieval at the pixel level is indirect. Since CRs are defined using cloud fields at mesoscales as opposed to individual cloud pixels, options of increasing the spatial resolution are limited. The alternative approach of evaluating GPROF at its native resolution but categorized by the CR environment based on the 1° grid box that contains the pixel seems precarious. For example, it is unclear how representative the described conditions associated with the CR are throughout the entire grid box; similarly, one may question whether the conditions associated with the CR are representative of a pixel at the edge of the grid box. On the other hand, an evaluation of GPROF at the pixel level would allow us to better interpret the comparisons, especially with regard to detection skill. For this study, we prioritized the representativeness of the CR over the more detailed spatial matching of GPROF with GV-MRMS. Another limitation in this analysis is the sun-synchronous orbits of the MODIS-bearing satellites, which, when coupled with the fact that the CRs are only defined during daylight hours, means that all the results are based on retrievals at about 1030 and 1330 local time. Therefore, this analysis does not capture the full diurnal range of precipitation, which can vary substantially over land (Tan et al. 2019); in this respect, cloud regimes defined from geosynchronous satellites (Tselioudis et al. 2021) have the advantage of increased sampling and a better representation of the diurnal cycle. However, this study does not intend to evaluate the full diurnal cycle of precipitation, but to interrogate the PMW retrieval algorithm, an objective that should be resilient to the time of day. A third limitation of using CRs instead of precipitation regimes is the possibility of a disconnect between the observations that define the regime and the precipitation retrievals, one of the consequences of CRs being only
imperfect proxies of their precipitation environment. In other words, the same CR occurring in different regions may have different precipitation characteristics and thus precipitation retrieval conditions (section 3a). Any interpretation of the results has to contend with the variability in the cloud–precipitation relationship (e.g., Jin et al. 2018; Tan and Oreopoulos 2019), which may be regionally dependent (Leinonen et al. 2016). Still, while these regional variations may blur the lines between two convectively active regimes (e.g., organized convection versus tropical and midlatitude storms), it is unlikely to confuse radically different cloud regimes (e.g., organized convection versus stratocumulus environment).

The approach of this study—separating the evaluation into different regimes—largely follows a similar concept of using the TRMM Precipitation Radar regimes (Henderson et al. 2017; Petković and Kummerow 2017). Indeed, these studies were able to provide a more precise characterization of the environment in clearly precipitating scenes. However, these precipitation regimes were only defined when precipitation was detected by the active instrument, and then categorized into only three regimes. On the other hand, CRs can distinguish between different convective and suppressed environments, which are associated with occasional, isolated, or scattered precipitation that is generally more challenging for PMW retrievals. Therefore, CRs provide a complementary perspective to the precipitation regimes. Such a regime-based approach to evaluation adds value by allowing us to establish the context in which the retrievals perform well or poorly, thereby identifying avenues of potential improvements.

Acknowledgments. JT, NC, and LO acknowledge support from NASA’s MEaSUREs and PMM programs. PK acknowledges support through the NASA’s Precipitation Measurement Missions Award 80NSSC19K0681 and Award NNX16AL23G.

Data availability statement. GPROF products can be downloaded from https://gpm.nasa.gov/data/directory. GV-MRMS can be obtained from http://dx.doi.org/10.5067/GPMGV/MRMS/DATA101. MODIS can be obtained from https://gpm.nasa.gov/sites/default/files/document_files/IMERG_ATBD_V06.pdf. GPM Data availability statement. GV-MRMS/DATA101. MODIS can be obtained from https://doi.org/10.5067/MEASURES/MODISCR/EQANG3H/DATA301.

References


Kiststetter, P.-E., and Coauthors, 2012: Toward a framework for systematic error modeling of spaceborne precipitation radar with NOAA/NSSL ground radar-based national mosaic


