



Applications of Dynamic Land Surface Information for Passive Microwave Precipitation Retrieval

SARAH RINGERUD,^{a,b} CHRISTA PETERS-LIDARD,^b JOE MUNCHAK,^b AND YALEI YOU^a

^a *Earth System Science Interdisciplinary Center, University of Maryland, College Park, College Park, Maryland*

^b *NASA Goddard Space Flight Center Greenbelt, Maryland*

(Manuscript received 6 April 2020, in final form 22 October 2020)

ABSTRACT: Accurate, physically based precipitation retrieval over global land surfaces is an important goal of the NASA/JAXA Global Precipitation Measurement Mission (GPM). This is a difficult problem for the passive microwave constellation, as the signal over radiometrically warm land surfaces in the microwave frequencies means that the measurements used are indirect and typically require inferring some type of relationship between an observed scattering signal and precipitation at the surface. GPM, with collocated radiometer and dual-frequency radar, is an excellent tool for tackling this problem and improving global retrievals. In the years following the launch of the GPM *Core Observatory* satellite, physically based passive microwave retrieval of precipitation over land continues to be challenging. Validation efforts suggest that the operational GPM passive microwave algorithm, the Goddard profiling algorithm (GPROF), tends to overestimate precipitation at the low ($<5 \text{ mm h}^{-1}$) end of the distribution over land. In this work, retrieval sensitivities to dynamic surface conditions are explored through enhancement of the algorithm with dynamic, retrieved information from a GPM-derived optimal estimation scheme. The retrieved parameters describing surface and background characteristics replace current static or ancillary GPROF information including emissivity, water vapor, and snow cover. Results show that adding this information decreases probability of false detection by 50% and, most importantly, the enhancements with retrieved parameters move the retrieval away from dependence on ancillary datasets and lead to improved physical consistency.

KEYWORDS: Precipitation; Algorithms; Remote sensing

1. Introduction

Accurate, physically based precipitation retrieval over global land surfaces is an important goal of the joint NASA/JAXA Global Precipitation Measurement Mission (GPM) (Hou et al. 2014; Skofronick-Jackson et al. 2017). This is a challenging problem for the passive microwave constellation component of GPM, as the hydrometeor signal over radiometrically warm land surfaces in the microwave frequencies means that the measurements used in retrievals are more indirect than over ocean. Over ocean, the multispectral emission and scattering characteristics of hydrometeors can be used in algorithms for microwave imagers such as the 13-channel GPM Microwave Imager (GMI), whereas over land, retrievals typically reduce to looking for an ice scattering signal associated with precipitation (Levizzani et al. 2020; You et al. 2017; Petty and Krajewski 1996). Additionally, the passive microwave (PMW) brightness temperatures (Tbs) are sensitive to the column integrated precipitation content, so additional assumptions about the vertical distribution of hydrometeors are required to calculate the surface precipitation rates desired by users of the data (Smith et al. 1994). GPM, with collocated radiometer and

dual-frequency (Ku/Ka band, or 14 and 35 GHz) precipitation radar (DPR), along with a constellation of partner radiometers, is an excellent tool for exploring these issues and relationships, with application toward the improvement of global retrievals.

The operational NASA GPM retrieval for the PMW sensors is the Goddard profiling algorithm (GPROF), a parametric scheme applied for each constellation radiometer (Kummerow et al. 2015). GPROF is a Bayesian-type retrieval and will be more fully described in section 2. Early iterations of GPROF did not use the Bayesian scheme over land surfaces, instead employing more empirical relationships relating ice scattering to surface precipitation (Gopalan et al. 2010); however, products in the GPM era (2014–present) have moved away from the empirical schemes. Regardless of formulation, the scattering signal forms the bulk of the information available to the retrieval over land simply due to the highly variable nature of the surface emissivity.

Using coincident radar–radiometer observations from the predecessor Tropical Rainfall Measuring Mission (TRMM), Berg et al. (2006) showed that disagreement between rain-rate estimates from the active and passive sensors displayed regional patterns that could be correlated to regional climate conditions. Subsequent versions of GPROF addressed this by

Corresponding author: Sarah Ringerud, sarah.e.ringerud@nasa.gov

DOI: 10.1175/JTECH-D-20-0048.1

© 2021 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).

Brought to you by UNIVERSITY OF CALIFORNIA Irvine | Unauthenticated | Downloaded 05/04/21 04:29 PM UTC

constraining the TRMM (ocean only) GPROF retrievals by two environmental parameters, namely total precipitable water (TPW) and sea surface temperature (SST) (Kummerow et al. 2011). Moving forward to GPM, these same techniques were adapted to land surfaces, by replacing the SST with the 2 m air temperature commonly available from forecast and reanalysis models. In a series of papers describing and testing the Cloud Dynamics and Radiation Database for Bayesian precipitation retrieval, Casella et al. (2013) and Smith et al. (2013) showed that solution ambiguity is greatly reduced by considering additional environmental parameters including vertical velocity and convective available potential energy (CAPE), which correlate to microphysical properties (Casella et al. 2013; Smith et al. 2013). Petković and Kummerow (2017) demonstrated that constraining the retrieval over land in particular may require including other parameters that more accurately relate storm dynamics and associated scattering signals, with results suggesting the best constraint to be CAPE over land surfaces. The authors found that stratifying by CAPE, which describes the potential for vertical motion and convection, decreased the variability in the relationship between surface precipitation and ice scattering aloft. While this gives some good physical insight into the processes that control the scattering–rain rate relationships, straight implementation of this in the algorithm requires more reliance on ancillary model data. Relying heavily on coarse model data decouples the algorithm from the observations and may potentially lead to problems near strong gradients, such as cold fronts, where the retrieval may be assuming conditions associated with different boundary layer and freezing level heights, with different scattering signal–rain rate relationships as a result.

While the Bayesian technique is adopted for operational NASA GPM retrievals, NOAA has implemented a 1D variational (1DVAR) technique—the Microwave Integrated Retrieval System (MiRS)—which does an iterative inversion of the radiative transfer in the same way over all surfaces (Boukabara et al. 2011; Meng et al. 2017). This method has the benefit of being fully radiometrically consistent and having no reliance on ancillary model data beyond initial offline development of first-guess parameters from climatology. The downside is that in retrieving all relevant parameters simultaneously in this way, retrievals of this type are relatively underconstrained.

False alarms are often an issue in PMW retrievals, particularly over land where dynamic surface characteristics can sometimes be mistaken for a precipitation signal. For this reason, many retrievals have historically employed techniques for “screening” of nonraining pixels (Ferraro et al. 1998). In the GPROF algorithm, the desire to be fully parametric discourages any empirically based screening. Version 5 of GPROF does this instead by running the retrieval, which returns a probability of precipitation, twice. Following the first pass, areas below a probability threshold are assumed nonraining and set to zero (Kummerow et al. 2017). MiRS also runs twice. In the first pass, the atmosphere is assumed to not contain any hydrometeors. Areas that fail to converge under emission-only conditions are identified and the retrieval is run a second time in these areas with scattering enabled

(Boukabara et al. 2011). Even with a screening system in place, the GPM GPROF algorithm tends to greatly overestimate light precipitation as a result of false alarms. Kidd et al. (2018) show relatively high false alarms from all GPROF products as compared to surface radar and gauges, for both the continental United States and Europe.

In this work, we explore a hybrid method, using output from an optimal estimation (OE) retrieval similar in basis to the MiRS 1DVAR first pass, as a first step to screen for precipitation, and a Bayesian precipitation retrieval in areas where the OE fails to converge. The Bayesian retrieval is constrained using dynamic retrieved parameters describing the surface and atmospheric conditions retrieved from the first step. In this way the algorithm relies on dynamic and radiometrically consistent information about the surface rather than solely ancillary model data. The goal of this work is to explore possible information enhancements to the retrieval for improving light precipitation estimates where the surface signal is important, while keeping generally within the operational GPROF framework.

2. Algorithms and data

a. GPROF

GPROF, version 5, builds upon a long heritage and evolution of the PMW Bayesian retrieval (Kummerow and Giglio 1994; Olson et al. 1996; Kummerow et al. 2011, 2015), employing the same algorithm over all surfaces. Broadly, the Bayesian inversion scheme retrieves the most likely precipitation state using the full vector of observed Tb. The retrieval space is defined by an a priori database constructed using the GPM Combined Radar-Radiometer (2BCMB) product, which retrieves precipitation profiles that best match all active and passive observations from the GPM *Core Observatory* satellite (Greco et al. 2016). The 2BCMB retrievals are currently (through version 6) carried out only in the presence of an active radar signal. The database is constructed from 1 year of 2BCMB profiles and is organized by surface type (including ocean, sea ice, coast, five land surface classes of varying vegetation level, and four classes of increasingly snow-covered land), 2 m temperature, and TPW derived from ancillary model data, which serve to constrain the retrievals. (An earlier iteration utilized an OE retrieval for the TPW field over ocean only, but this has been discontinued.) It is important to note that in this formulation, the 2BCMB retrievals constitute the mechanically correct answer from an algorithm perspective—GPROF cannot retrieve anything that does not exist in the database, which means no retrievals of light precipitation below the sensitivity of the active radar. This is a separate issue from the spurious widespread light precipitation common in retrieval schemes of this type. Such values are simply due to Bayesian weighting and the probabilistic nature of the algorithm (Kummerow et al. 2015). Tb used for the retrievals are taken from the level 1C-R product, which delivers calibrated Tb values in a format where each channel field of view has been collocated to the same center point (NASA PPS and X-Cal Working Group 2017).

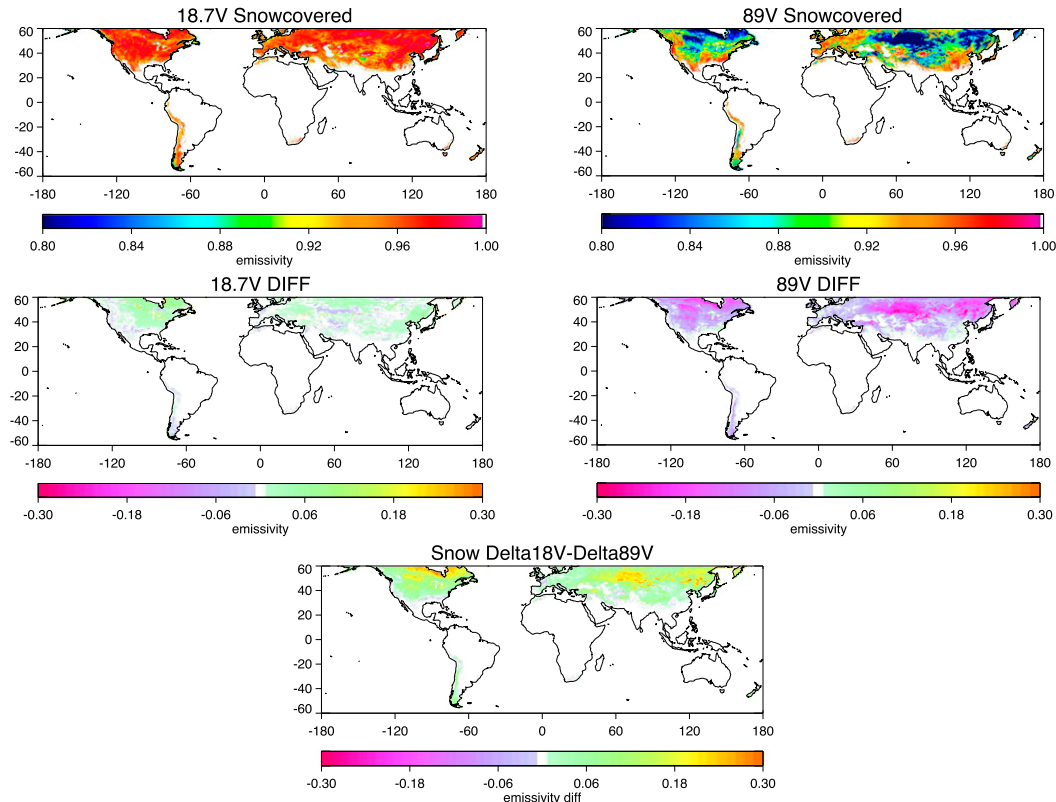


FIG. 1. (top) September 2014–August 2015 annual mean emissivity for snow-covered GMI overpasses for the 18.7 and 89 GHz channels used for development of the snow-cover retrieval. (middle) Differences between snow-covered and non-snow-covered emissivity. (bottom) Frequency differences for snow cover used in the final formulation.

b. The OE emissivity retrieval

Munchak et al. (2020) describe an optimal estimation-based emissivity retrieval for the GPM Combined algorithm, anticipated to be included in version 7 of the operational product. The retrieval uses GMI Tb as input and retrieves surface emissivity at the GMI frequencies along with the water vapor profile. It should be noted that the emissivity retrieval uses the Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2; Gelaro et al. 2017), dataset as a first-guess value, meaning that this method is not completely free of ancillary model data. Importantly, the emissivity retrieval output includes an error parameter, Φ_N , which describes the ability of the OE forward model to converge and match the observed Tb. The retrieval (like the first step of MiRS) assumes a nonprecipitating, nonscattering atmosphere. Therefore, areas where this parameter indicates a failure to converge likely are associated with characteristics not accounted for under the nonscattering assumption, most likely hydrometeors. This parameter then indicates likely areas of precipitation.

c. The hybrid retrieval and dynamic constraints

The retrieval itself involves several steps. The presence or absence of snow cover is first determined using Tb observations.

The OE parameters are then used to screen for areas of possible precipitation, where a Bayesian scheme is applied for the actual precipitation retrieval and constrained using retrieved information from the OE, in contrast to the GPROF constraints using ancillary model data. Each step is discussed individually in detail below.

1) STEP 1: SNOW COVER

Because of the large effect of the emissivity change due to snow cover on observed Tb, the first step in the retrieval is to determine whether snow on the ground is present. Within the GMI frequencies, Tbs are sensitive to the presence, depth, and character of snow, making quantitative measurement in this framework difficult (Mätzler 1994; Chang et al. 2003; Kongoli et al. 2007). The operational GPROF retrieval uses ancillary, daily model data from the NOAA Autosnow product (Romanov et al. 2000) to identify snow-covered areas (Kummerow et al. 2017). Snow cover can be a dynamic variable, with both its presence and properties changing at time scales less than a day and has a large effect on the surface emissivity and associated Tb. In this application, the presence of snow cover is diagnosed within the retrieval process using the Munchak et al. (2020) OE-retrieved emissivities. Remote sensing of snow cover has a long history, and PMW observations at lower frequencies are

particularly valuable due to sensitivity and the ability to sense the surface in the presence of clouds at lower frequencies, with greatest sensitivity to snow volume and shallow snow in the vertical polarization, but better differentiation from dry soil in the horizontal polarization (Dietz et al. 2012). At the lower GMI frequencies, snowpack and its effect on the dielectric properties of the surface leads to a positive emission signal, whereas at the higher frequencies scattering tends to decrease emissivity (Shahroudi and Rossow 2014). Munchak et al. (2020) demonstrate a high information content in the retrieved emissivities up to 89 GHz for TPW values under about 50 mm. To lessen the effects of scattering due to vegetation, deviations from the local mean non-snow-covered emissivity are used, rather than absolute emissivity following the technique of Shahroudi and Rossow. The 19 and 89 GHz channels are chosen due to their presence on all GPM constellation imagers.

For development of the snow-cover detection scheme, a 1-yr $0.25^\circ \times 0.25^\circ$ climatology of emissivity, in the absence of precipitation (meaning none detected by the active radar), at each frequency is computed and separated into snow-covered and non-snow-covered, as determined by the GPROF ancillary data (NOAA Autosnow) for calibration. Light precipitation below the sensitivity of the radar or below the ground clutter cutoff height has the potential to contaminate results here, but in such cases the impact on the dominant surface emission signal is likely small and should have little effect on the results. An exception would be in cases where the Tb is affected by liquid water emission. In cases where liquid clouds are present over snow cover, for example, this technique may fail. It should also be noted here that while the 18.7 and 89 GHz channel footprints share the same scan line and have been collocated via the 1C-R product, the actual resolution of each channel differs, so this is a possible source of error in that the signal from the 19 GHz corresponds to a larger area than the 89. The September 2014–August 2015 mean emissivities and (snow – nonsnow) emissivity differences are shown in Fig. 1 for the 18.7 and 89 GHz vertical polarization (V-pol) channels. The left side of the figure demonstrates the emissivity effects of snow cover at the 18.7 GHz V-pol channel. The top panel illustrates the high values (near 1) for snow-covered pixels, and the middle left, which plots the difference between these values and the non-snow-covered mean, shows the general net positive, emission-based signal from the snowpack at this frequency. The right side plots the same indicators for the 89.0 GHz V-pol channel, illustrating the net decrease in emissivity associated with scattering from snow on the ground. The bottom panel shows a nearly uniformly positive value for the difference between the net emission signal at the lower frequency, and net scattering signal at the higher frequency. The relationship is positive everywhere with the exception of a few of the most equatorward points.

The snow-cover detection scheme is then tested using data from September 2015 to May 2016 (the following year) for all pixels. Several threshold values are tested for optimization of the scheme. The statistical metrics probability of false detection (POFD—the number of false alarms divided by the

TABLE 1. Boreal winter 2015/16 global probability of false detection, probability of false alarm, probability of detection, and HSS for all tested values of snow emissivity cutoff values as described by Eq. (1).

Snow cutoff value	POFD	POFA	POD	HSS
SON 2015				
0.0	0.51	0.65	0.84	0.22
0.01	0.25	0.50	0.76	0.43
0.02	0.12	0.35	0.68	0.55
0.03	0.067	0.25	0.62	0.59
0.04	0.043	0.19	0.56	0.58
0.05	0.031	0.16	0.51	0.55
DJF 2015–16				
0.0	0.51	0.19	0.93	0.46
0.01	0.32	0.13	0.90	0.60
0.02	0.19	0.086	0.87	0.66
0.03	0.12	0.058	0.84	0.67
0.04	0.083	0.042	0.80	0.65
0.05	0.06	0.032	0.77	0.62
MAM 2016				
0.0	0.44	0.51	0.85	0.34
0.01	0.22	0.37	0.79	0.53
0.02	0.12	0.25	0.74	0.62
0.03	0.073	0.18	0.69	0.64
0.04	0.050	0.14	0.64	0.63
0.05	0.037	0.11	0.59	0.61

number of times the event did not happen) and probability of false alarm (POFA—the number of false alarms divided by the total number detected) are computed as defined by the contingency table in Barnes et al. (2009), along with the POD (probability of detection, which gives the ratio of correct detections to all snow-cover observations). The Heidke skill score (HSS) is also computed, which gives a measure of retrieval skill as the proportion of correct retrievals scaled with those expected from random chance (Heidke 1926). Results, presented in Table 1, suggest an inflection point at a threshold of 0.03, which maximizes HSS for all three seasons. Skill is lowest in the fall season, suggesting a particular difficulty in this type of early season snow. The POD is maximized at a slightly lower threshold value, but the lower values are also associated with more false alarms. The final relationship, balancing all factors, is determined as

$$\Delta 18.7V\epsilon - \Delta 89.0V\epsilon > 0.03, \quad (1)$$

where the delta (Δ) indicates a difference between observed emissivity at and 1 yr mean snow-free emissivity at each frequency (i.e., $\Delta 18.7V\epsilon$ is the difference in emissivity (ϵ) from the snow-free mean at 18.7 GHz V-pol). This relationship is designed to detect seasonal snow cover, leading to a deviation from snow-free emissivity, and adjustments would be required for application to the polar regions, for example, where snow cover is the default state. It should be noted that the Autosnow data, while used for calibration, are not to be viewed as pure validation data as they are a daily product, whereas the scheme here is intended to identify instantaneous snow cover. As this designation is used for determining retrieval constraints, the choice is made to err on the side of detection. Figure 2 shows

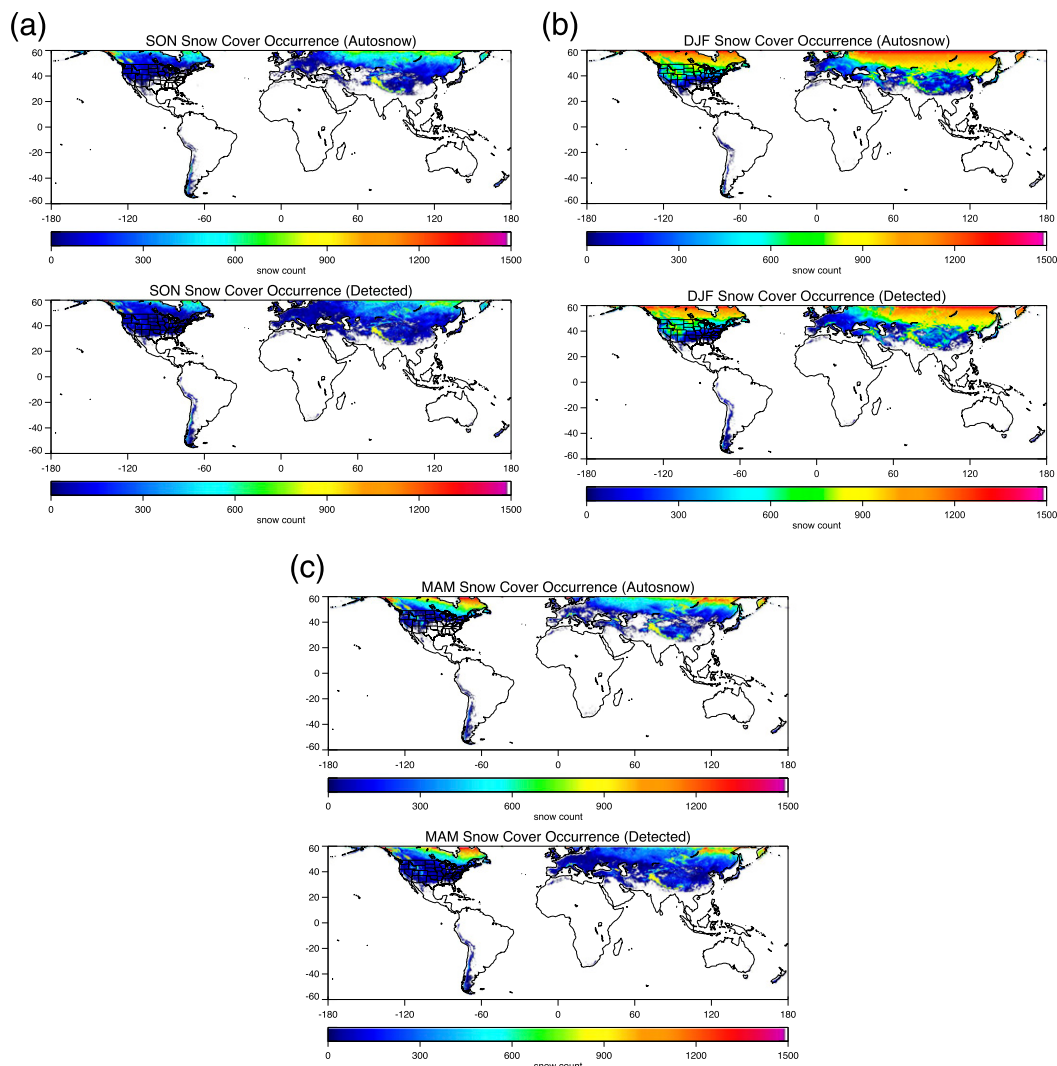


FIG. 2. Snow-cover detection occurrence for GMI overpasses from (top) the Autosnow product and (bottom) the detection scheme described here over the period of boreal winter 2015/16 for the seasonal periods (a) September–November (SON), (b) December–February (DJF), and (c) March–May (MAM).

the snow-cover detection results for the test period, September 2015–May 2016, broken down by season. As indicated by the statistics in Table 1, the colder season snow (DJF) is most consistently detected. Notable differences are located in mountainous areas, particularly eastern Europe and the U.S. Pacific Northwest region, where snow cover is underdetected as compared to Autosnow. In the fall and spring season, the southern extent of snow-cover areas shows widespread but very small in number false detections as compared to Autosnow, though since the comparison is of an instantaneous versus daily product this comparison is a bit ambiguous. While the purpose of this simple scheme is purely to screen for the likely presence of snow cover, there is certainly information about the character of the snow within the multispectral emissivity data (see Munchak et al. 2020 for a deeper discussion of emissivity and snow characterization), and this will be a

topic of future work. The differences observed in areas with more complex orography may indicate problems with the scheme for particular snow types, and due to the orbital constraints of the GPM *Core Observatory* satellite there has been no testing closer to the poles. Based upon the results in Table 1, the detection performs reasonably well for the current application, comparing nicely to the operational Autosnow product, and will be used here to define the presence of snow cover without the need for coincident ancillary data.

2) STEP 2: ERROR PARAMETER SCREENING

Due to the added complexities of retrieving precipitation over snow-covered surfaces, this work will focus on non-snow-covered surfaces only, with the snow-cover retrieval described in Step 1 used for classification; snow-covered surfaces are relegated to future work. After screening for snow cover, the

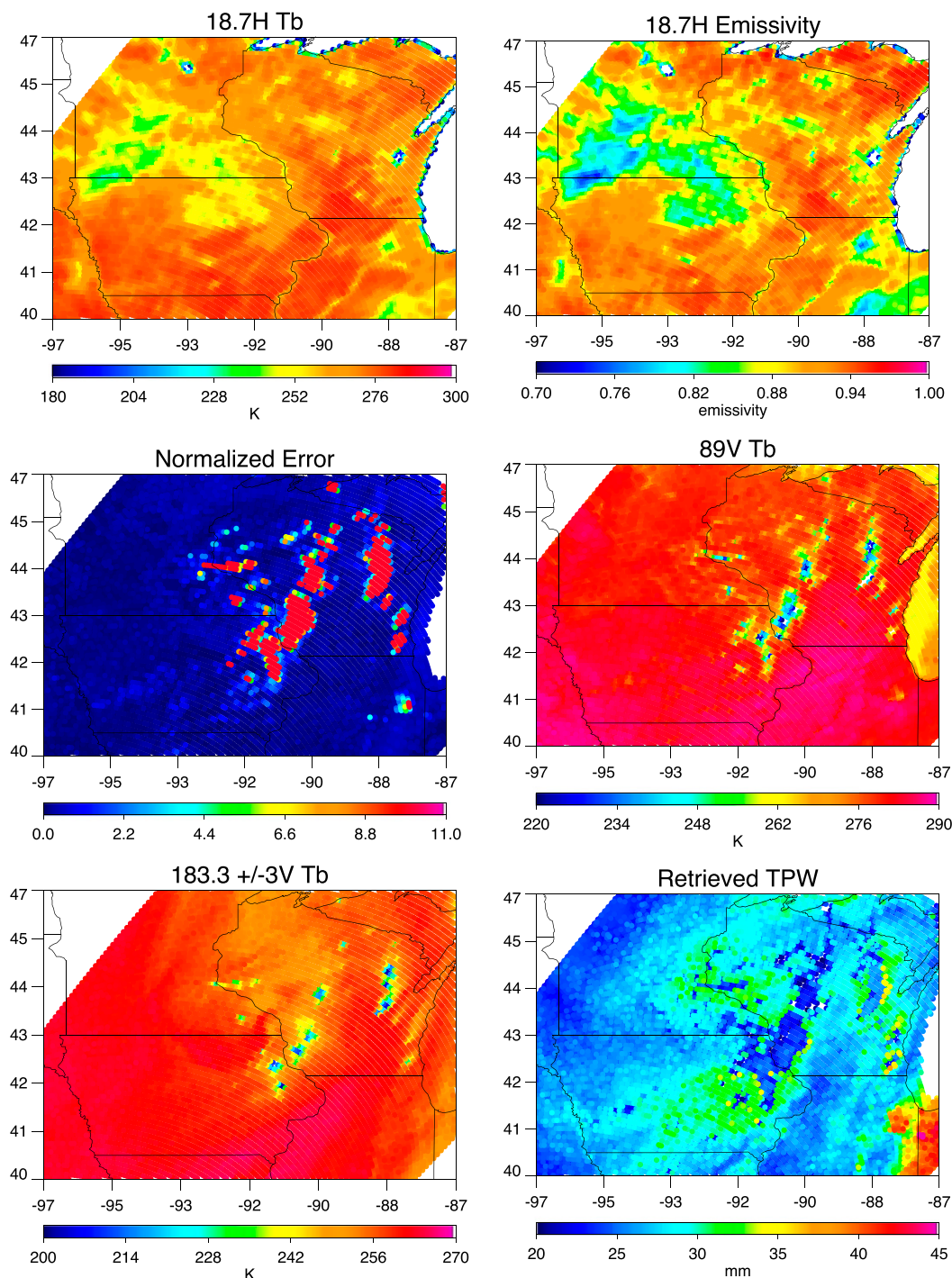


FIG. 3. Observed and retrieved parameters from a GMI overpass on 4 Jun 2016. (top left) Observed Tb and (top right) retrieved emissivity at 18.7 GHz. (middle left) The normalized error parameter from the OE retrieval and (middle right) the observed 89V Tb. (bottom left) Observed Tb at 183 ± 3 GHz alongside (bottom right) retrieved TPW.

OE parameters are obtained. Figure 3 shows emissivity at the 18.7 GHz horizontal polarization (H-pol) channel (upper right) along with associated observed GMI Tb (upper left) for an overpass of the GPM satellite on 4 June 2016 over the upper

Midwest region of the United States. Unusually low emissivities are observed in southwestern Minnesota, associated with prior precipitation. The observed Tb from the 89 GHz V-pol channel are shown in the middle-right panel, and the 183.3 ± 3 GHz

V-pol in the bottom left. These observations suggest intense convective precipitation associated with the cold scattering cloud tops running diagonally over Wisconsin associated with increased water vapor values. Retrieved TPW from the OE is plotted in the lower right panel, and the normalized error in the middle-left (for plotting purposes, the Φ_N values are cutoff at a value of 100) and are collocated with the clearly erroneous TPW values in the areas associated with the scattering signal in the Tb fields. Observed precipitation from the NOAA Multi-Radar Multi-Sensor (MRMS) system (Zhang et al. 2011) surface radar product at the overpass time is shown in Fig. 4, and confirms that most of the radar observed precipitation is associated with this scattering signal and in the areas of elevated Φ_N . These areas then are candidates for the Bayesian precipitation retrieval, and the TPW and emissivity values here are not to be used as the OE has difficulty converging.

3) STEP 3: DEFINE OBSERVATIONALLY BASED CONSTRAINTS FOR BAYESIAN RETRIEVAL AND DATABASE

As described previously, the operational GMI algorithm constrains retrievals using ancillary parameters: static surface types along with TPW and 2 m temperature from model reanalysis products. In an effort to make the constraints more observationally based and fully utilize the information content in the Tbs, the hybrid retrieval constrains are based upon emissivity in the 18.7 GHz H-pol channel and TPW, both associated with the OE output. Constraining the database by emissivity at only a single frequency will not fully describe variability in the surface state but adding further dimensions to the database leads to the danger of overconstraint and a sparsely populated database. The 18.7 GHz is chosen as it represents the lowest, and therefore most surface-sensitive, frequency common to all of the GPM constellation imagers. The normalized error parameter is used to screen out footprints without reliable values. The threshold is fixed and may include precipitating or nonprecipitating conditions. Over these areas, TPW is interpolated and a recent (previous 1 year) climatology of retrieved seasonal emissivity is used. To allow for dynamic variability associated with the presence of precipitation, an expansion of the Bayesian constraints is allowed in these areas and will be explained further in the next section. This methodology differs from the operational GPROF algorithm in that the constraints (snow cover, TPW, and emissivity) are each associated with the OE retrieval and thereby the observed Tb, rather than ancillary model input.

4) STEP 4: BAYESIAN PRECIPITATION RETRIEVAL

Following the OE procedure and association of each footprint identified by the normalized error parameter as possible for precipitation, Bayesian retrieval is performed on only these identified footprints following the GPROF methodology described in section 2. Each database profile is associated with a TPW and emissivity value computed offline using the OE. The database is searched within a region defined by the TPW interpolated from the retrieved values, over a range up to 2 mm above the interpolated value, to allow for increases associated

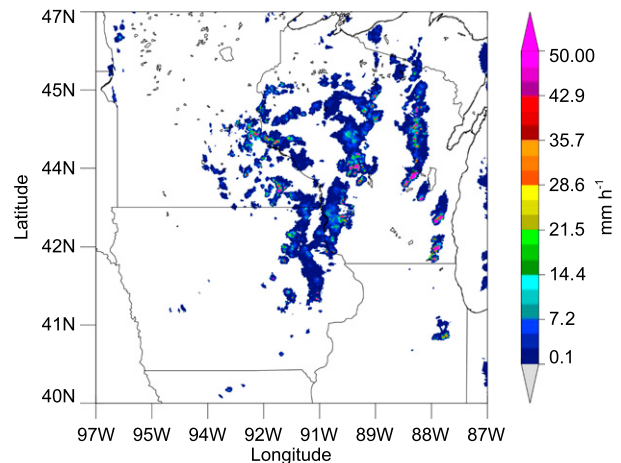


FIG. 4. MRMS surface radar rain rates for the 4 Jun 2016 case shown in Fig. 3.

with the precipitating area, and 18.7 GHz H-pol emissivity from the previous year's climatology, plus a decreasing range of 0.05 to allow for decreased emissivity due to recent precipitation (Munchak et al. 2020). Figure 5 shows retrieved precipitation for the overpass shown in Fig. 3, for six different values of Φ_N cutoff from zero (Bayesian retrieval everywhere) to 0.5. A clear pattern emerges as the cutoff value increases, with widespread light precipitation areas decreasing with Φ_N . Values starting with 0.3 start to look quite similar to the surface radar observations shown in Fig. 4. Note that values are truncated at 0.1 mm h^{-1} to compare with the MRMS product, which has a 0.1 mm h^{-1} minimum in order to be consistent with DPR sensitivity.

The hybrid technique demonstrated here is run as described for the 1-yr period September 2015–August 2016 globally for non-snow-covered land surfaces. Results are discussed in section 3.

3. Retrieval results and discussion

In heavily precipitating areas, the component of the Tb signal from the surface becomes less important than the precipitation signal. It is assumed therefore, that use of the OE emissivity retrieval in a “hybrid” configuration defining areas for a GPROF-like Bayesian precipitation retrieval will have the most impact on light precipitation cases and false alarms. To explore these differences in a consistent manner, the retrieval is also performed for the same year using the GPROF constraints—a climatological surface type (or Class) defined by GPROF developers, model TPW, and model 2 m temperature. This is procedurally identical to the GPROF product, with the following exceptions: Database entries are compared to observations individually, rather than using clustered values as GPROF does (Kummerow et al. 2017), and the first “screening” pass is not performed. The purpose of using this, which will be labeled Class, for comparison, rather than the operational GPROF product, is to examine differences in retrieval results using the exact same database profiles but

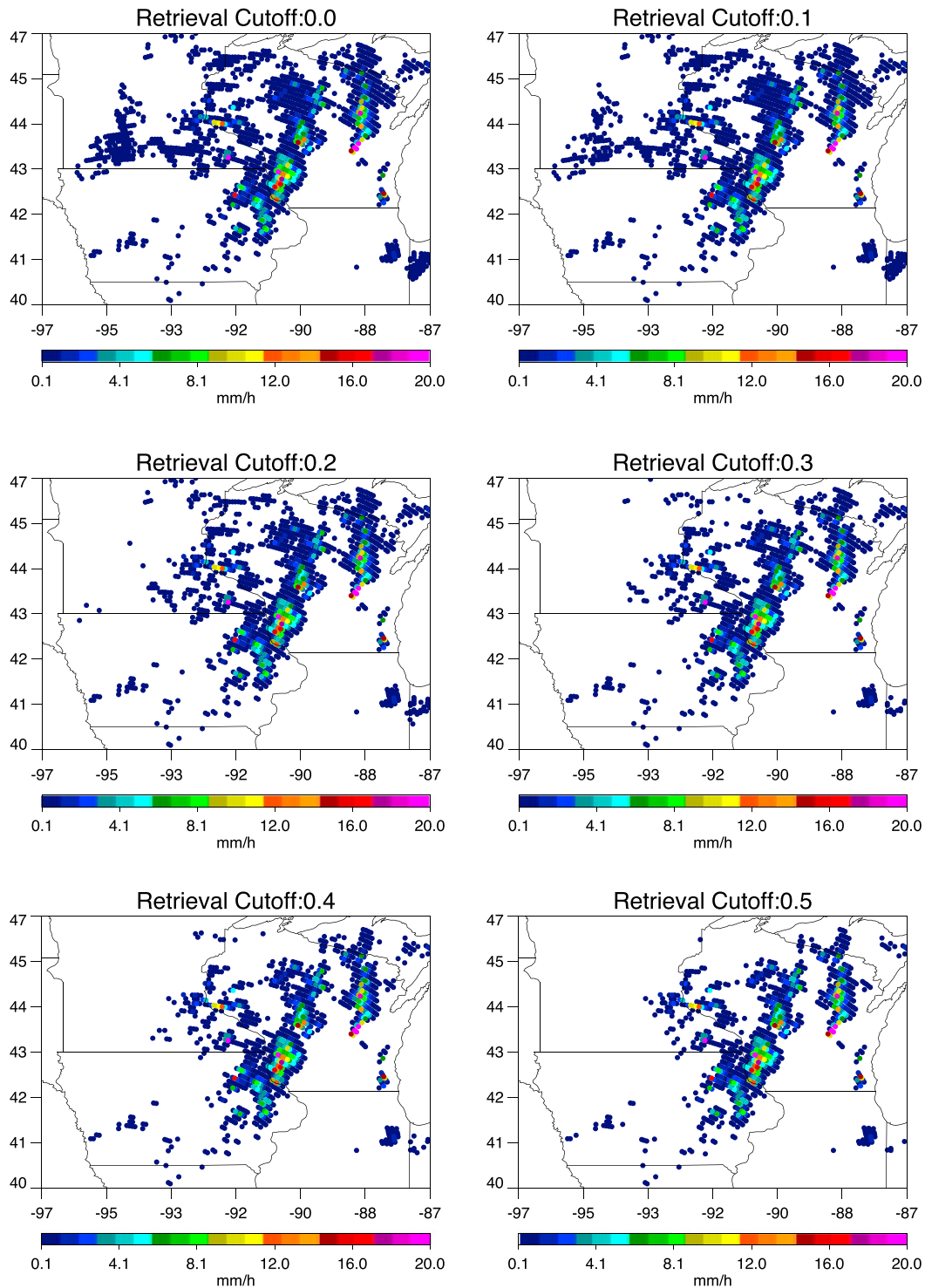


FIG. 5. Retrieved precipitation for the 4 Jun 2016 case shown in Figs. 3 and 4. Each panel shows the retrieved rates using different values of the normalized error parameter as a cutoff to define areas where the Bayesian retrieval is performed.

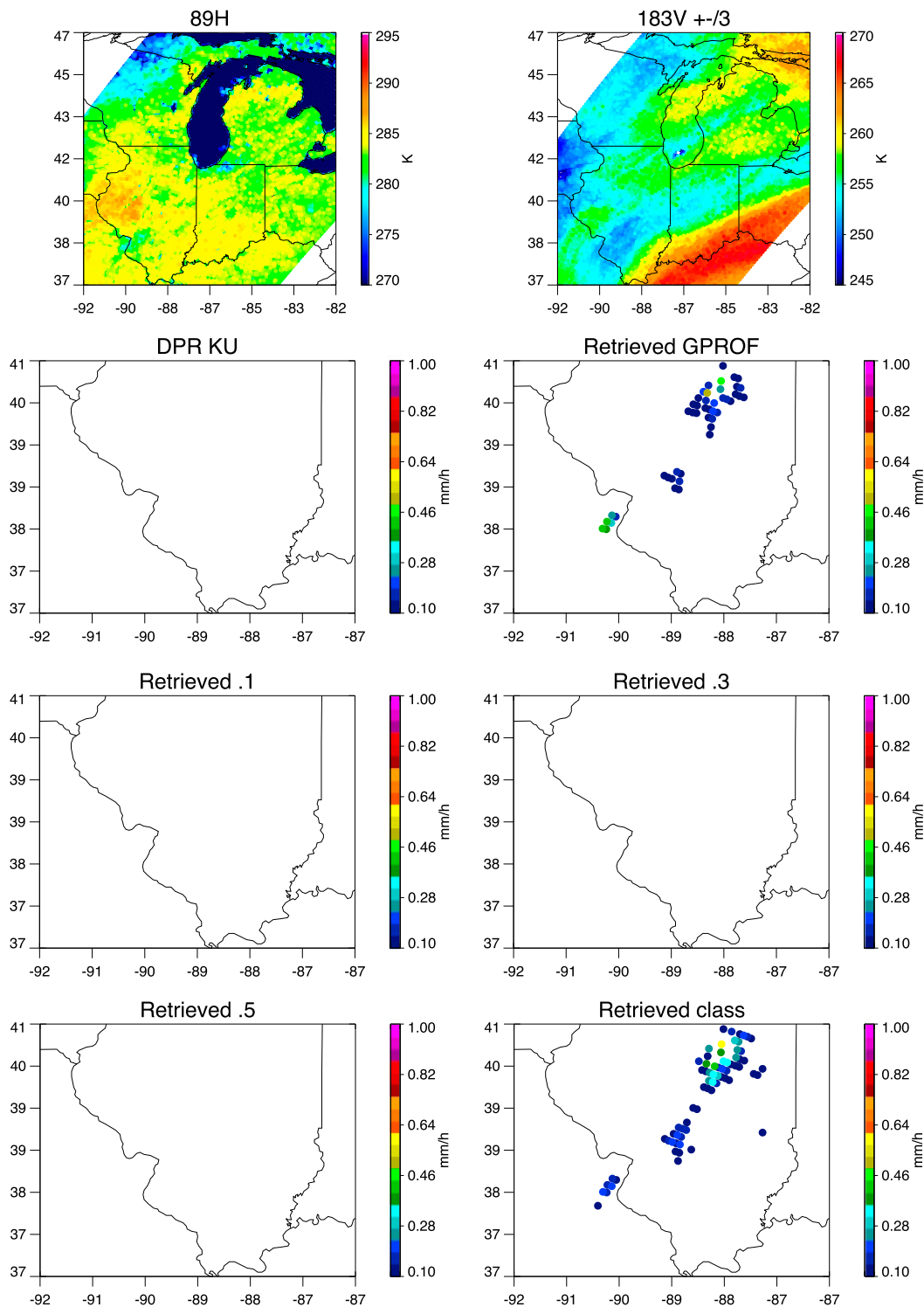


FIG. 6. Retrieved precipitation rates from a GMI overpass on 8 Sep 2015. (top left) Observed GMI Tb at the 89 and (top right) 183 ± 3 GHz channels. (second row) GPM retrievals from (left) the active Ku radar and (right) the operational GPROF retrieval are shown over Illinois. The additional rows show the hybrid retrievals using several values of normalized error parameter as a cutoff for the Bayesian retrieval areas, and (bottom right) retrieval using the GPROF classes.

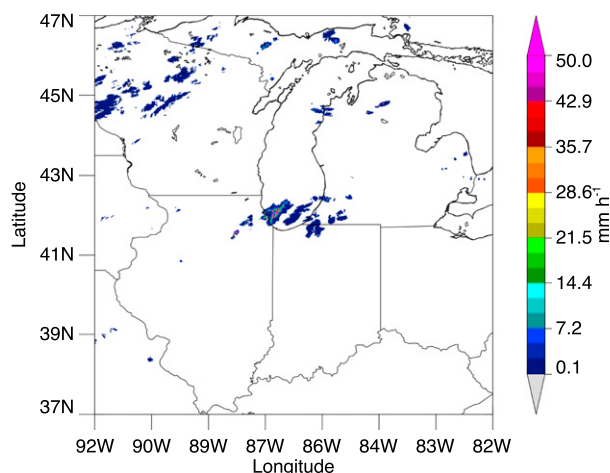


FIG. 7. MRMS surface radar rain rates for the 8 Sep 2015 case shown in Fig. 6.

organized into different database constraining schemes. An example of successful light precipitation screening is shown in Fig. 6 for a GMI overpass on 8 September 2015. The top panels show observed GMI Tb at the 89 and 183 ± 3 GHz frequencies over the Great Lakes region of the United States. Cold Tb at 89 GHz suggest high cold clouds in central Wisconsin as well as south of Lake Michigan and a weak scattering signal over southwestern Illinois. A strong gradient is apparent in the 183 ± 3 GHz water vapor channel, suggesting a frontal boundary in the area. Focusing in on the area of Illinois south of Lake Michigan, the DPR Ku-band (DPR Ku) retrieved values as well as the operational GPROF product are plotted in the second row, with the hybrid retrievals at several Φ_N cutoff

values below. Each of these panels shows an absence of any retrieved precipitation, regardless of Φ_N . The bottom-right panel, labeled “Retrieved Class” shows the retrieval using the same database organized into the GPROF classification. Results are similar to the operational GPROF results. Here the weak scattering signal over the center of the state appears to have been identified by the Class formulation as precipitation, whereas in the hybrid retrieval no precipitation is retrieved, regardless of Φ_N cutoff value. The MRMS validation data, shown in Fig. 7 over the full Great Lakes region for a sense of the larger scale situation, does not indicate precipitation in this area. For this particular case, the normalized error parameter is around 1.0 over the area, so the reorganization of the database entries appears to have led to the differences in this case. A midlatitude cyclone is present here, associated with a strong water vapor gradient as illustrated by the plot of observed 183 ± 3 GHz V-pol Tb shown in the top row. As there are no retrievals across the threshold values, the constraint of TPW from retrieved values versus model values has clearly made a large difference here. The importance of the TPW bin is further demonstrated in Fig. 8, which plots the TPW values from the GPROF model-based classification (left) and the TPW calculated from the OE procedure (right). The 1DVAR is run over coastal and land surfaces only, resulting in the absence of data over open water. The model-derived TPW values are smoother and lower by up to 10 mm in the area of false alarms, which notably coincides with a tight gradient. Therefore, each method, though using the same profiles in the database, is constraining the retrieval with significantly different TPW values in the same area, yielding different results.

For a global comparison, the retrievals are run as described for a 1-yr period. To compare with DPR active radar estimates,

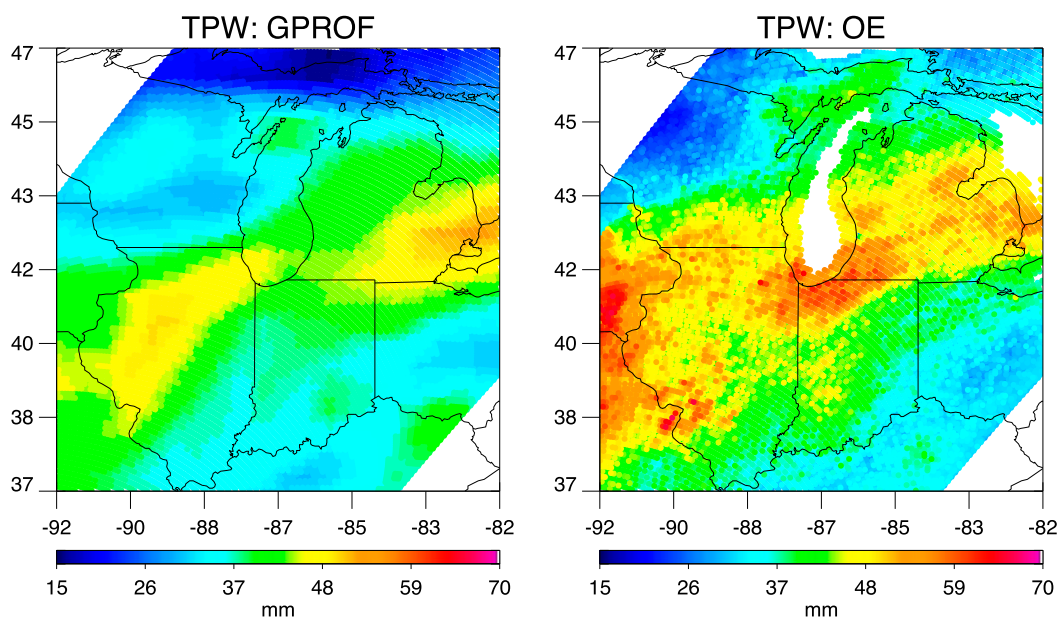


FIG. 8. TPW values from the (left) GPROF model values and (right) OE retrieved values over the Great Lakes region for the 8 Sep 2015 case shown in Fig. 6.

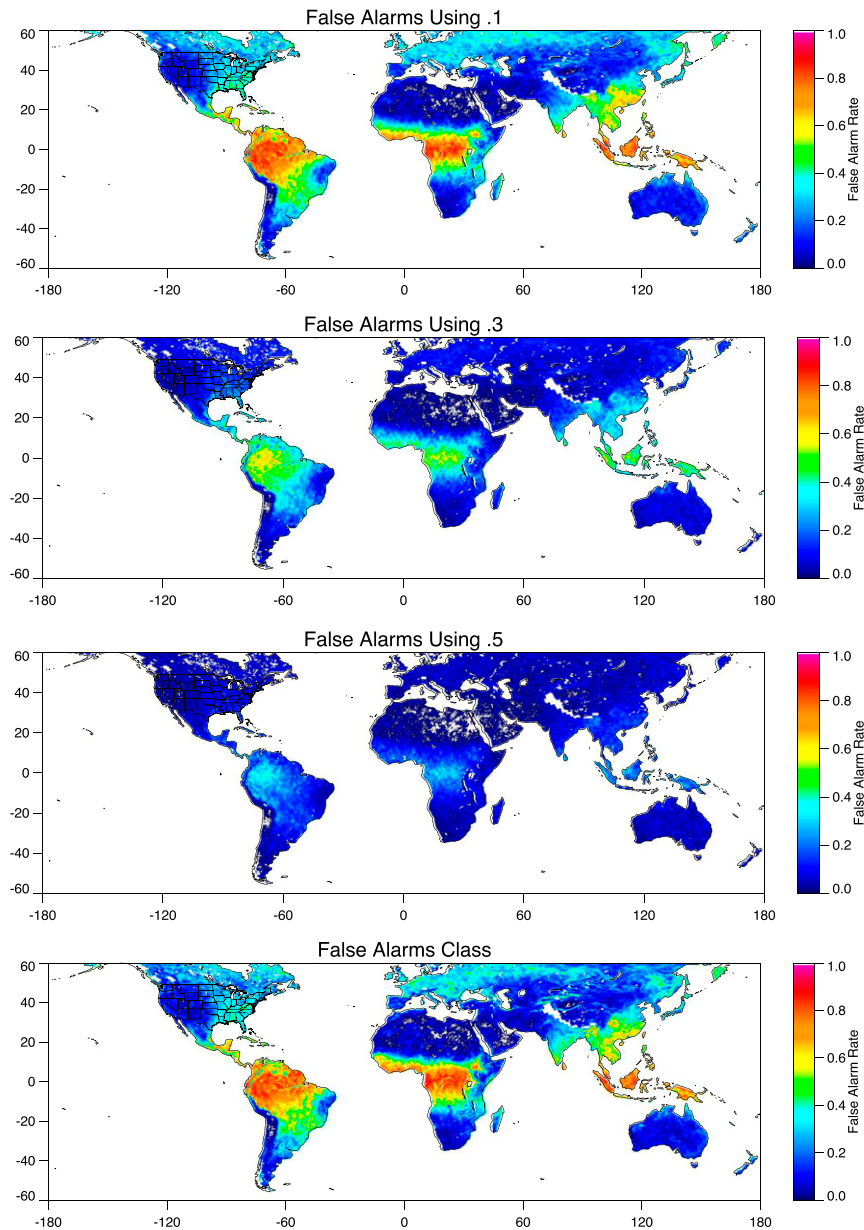


FIG. 9. Ratio of retrieval false alarms (not detected by the active radar) to all observations for the 1-yr period September 2015–August 2016. (bottom) Results using the GPROF classification scheme; (top three rows) results using the hybrid retrieval with several values of normalized error parameter cutoff defining locations for the Bayesian retrieval.

the retrievals are compared over the center footprints corresponding to the radar swath only. As the interest here is in improving estimates of light precipitation, the DPR Ku-band radar is used as a global comparison for calculation of false alarms—cases where the passive retrieval has retrieved precipitation (above a value of 0.01 mm h^{-1}) where the active radar gives a value of zero. The radar has a minimum detectable rain rate of 0.30 mm h^{-1} (Hamada and Takayabu 2016), and the 0.01 mm h^{-1} cutoff conservatively ensures that even

averaged over the 18.7 GHz footprint precipitation is retrieved and not simply a mathematical artifact of the probabilistic Bayesian scheme. Recall that the retrieval database is constructed using the radar precipitation, so this is mechanically the desired “right answer,” which may differ from “truth.” Results are plotted and shown in Fig. 9 as a ratio of false alarms to total observations. The global distribution is similar in each case and largely follows the distribution of precipitation. High values are concentrated over the Amazon and Congo rain

forests. Given the relative uniformity of emissivity over the tropical rain forest, which tends to be less dynamic due to year-round vegetation and relative insensitivity to surface soil moisture changes under the heavy canopy, this presents further evidence that using a TPW value retrieved using the coincident observations versus ancillary data has the potential to improve the GPM PMW precipitation algorithm. The false alarms using the GPROF Class constraints are very similar to a Φ_N cutoff value of 0.1, and these results form a baseline from which the number of false alarms decreases as the cutoff value increases. For a deeper look at the effect of the cutoff on retrieval results, error statistics for the full year are given in Table 2 for precipitation values less than 2.0 mm h^{-1} . Correlations and root-mean-square error (RMSE) tend to be associated with precipitation at higher rain rates, and these are unchanged for each retrieval formulation. Bias and probability of false detection decrease with Φ_N value. When the Φ_N cutoff value is increased from 0.1 to 0.3, probability of false detection decreases by half. A value of 0.5 decreases POFD by an order of magnitude. Heidke skill score values increase with the more stringent cutoff values. The probability of detection statistic suggests caution however, as this value is also decreasing. The statistics here indicate a possible optimal value of about 0.3 as a middle ground for relative maximization of detection and minimization of false detection as both are decreasing. The HSS is not at a maximum here, however, and this is not necessarily a final answer. Future implementation of this technique should explore the possibility that the cutoff need not be static and may vary by a yet to be determined location or regime.

4. Conclusions

In this work we demonstrate enhancements to the GMI PMW precipitation algorithm with more dynamic observationally based constraints supplied by a concurrently run OE retrieval to maintain radiometric consistency and decrease reliance on ancillary model data as constraint. Use of observed emissivity for screening of snow cover eliminates potential problems from dynamic changes to the emissivity due to fresh snow or melting below the daily time scale of the ancillary products used operationally. Constraining of the database using the OE retrieval parameters rather than model values has an impact on detection as illustrated by the case shown in Fig. 6. Global statistics indicate the potential of such a technique to reduce false alarms and spurious widespread light precipitation often associated with pure Bayesian retrievals. The idea of “false alarms” is of course a bit tricky for a probabilistic retrieval, as an absolute zero value is highly unlikely in such a scheme. As indicated in the algorithm description, the latest version of GPROF is run twice, and zero values are assigned below a probability threshold. Spurious light precipitation over land is still common, however, and many users instead use self-defined cutoff values. The hybrid technique eliminates the need for rate or probability-based cutoffs by instead examining convergence of the OE, similar to a MiRS-style variational retrieval. In the nonconverging areas, a highly constrained Bayesian precipitation retrieval

TABLE 2. Error statistics for 1 year of retrieved precipitation over global snow-free surfaces. Each row gives statistics for the 1-yr period using a different value of normalized error parameter as cutoff for defining the area of Bayesian precipitation retrieval, with the bottom column showing statistics using the GPROF classes.

	Correlation	Bias	RMSE	POFD	POD	HSS
$\Phi_N > 0.1$	0.47	0.044	1.04	0.29	0.91	0.20
$\Phi_N > 0.2$	0.47	0.043	1.04	0.22	0.89	0.26
$\Phi_N > 0.3$	0.47	0.038	1.04	0.15	0.84	0.33
$\Phi_N > 0.4$	0.46	0.037	1.04	0.11	0.79	0.40
$\Phi_N > 0.5$	0.47	0.035	1.04	0.08	0.75	0.46
Class	0.47	0.045	1.07	0.29	0.90	0.20

is inserted, in contrast to the uniform MiRS or GPROF formulations.

As a two-step process, with individual OE emissivity retrievals required for each constellation sensor, the hybrid technique is not proposed as an operational scheme but is presented here for use as an enhanced and more observationally based, radiometrically consistent retrieval system for research. In particular, the inclusion of surface emissivity and concurrent retrieved water vapor has many applications for use in the study of land-atmosphere interaction. While the current work has been demonstrated here only for the GMI sensor, the inclusion of the OE retrieval has the potential to improve the database Tb calculations for use with the constellation sensors as well. This is relatively straightforward for similar imaging instruments, but the GPM constellation also includes several cross-track scanning platforms designed for atmospheric sounding with frequencies focused on the higher end of those described here (Hou et al. 2014). These high-frequency sounders (many of which have a 23.8 GHz channel near the water vapor absorption line but do not have the 18 GHz window channel) are a bit trickier, but the surface is much less important for sounder retrievals, and surface-related constraints may not be necessary in those cases. Dynamically retrieved water vapor, however, may have significant potential for improving these retrievals, and this will be a direction of upcoming work. For future use, the OE retrieval can be easily adapted to each sensor and would remain physically consistent for that platform. This consistency with observed Tb is particularly important for applications such as data assimilation and use of the GPM data to understand precipitation processes and relation to the geophysical environment and is the key enhancement presented in this work. The inclusion of the OE and dynamic surface information along with water vapor retrieved from the observations create a more robust, radiometrically consistent retrieval for process studies and hydrologic applications.

Acknowledgments. This research was supported by a NASA PMM work package. The support by NSF (Grant EAR-1928724) and NASA (Grant 80NSSC19K0726) to organize the 12th International Precipitation Conference (IPC12), Irvine, California, June 2019, and produce the IPC12 special collection of papers is gratefully acknowledged.

Data availability statement. The data used are available online at <https://doi.org/10.5067/GPM/GMI/GPM/GPROF/2A/05>, <https://doi.org/10.5067/GPM/DPRGMI/CMB/2B/05>, and <https://doi.org/10.5067/GPM/GMI/R/1C/05>.

REFERENCES

- Barnes, L. R., D. M. Schultz, E. C. Grunfest, M. H. Hayden, and C. C. Benight, 2009: Corrigendum: False alarm rate or false alarm ratio? *Wea. Forecasting*, **24**, 1452–1454, <https://doi.org/10.1175/2009WAF2222300.1>.
- Berg, W., T. L'Ecuier, and C. Kummerow, 2006: Rainfall climate regimes: The relationship of regional TRMM rainfall biases to the environment. *J. Appl. Meteor. Climatol.*, **45**, 434–454, <https://doi.org/10.1175/JAM2331.1>.
- Boukabara, S. A., and Coauthors, 2011: MiRS: An all-weather 1DVAR satellite data assimilation and retrieval system. *IEEE Trans. Geosci. Remote Sens.*, **49**, 3249–3272, <https://doi.org/10.1109/TGRS.2011.2158438>.
- Casella, D., and Coauthors, 2013: Transitioning from CRD to CDRD in Bayesian retrieval of rainfall from satellite passive microwave measurements: Part II. Overcoming database profile selection ambiguity by consideration of meteorological control on microphysics. *IEEE Trans. Geosci. Remote Sens.*, **51**, 4650–4671, <https://doi.org/10.1109/TGRS.2013.2258161>.
- Chang, A. T. C., R. E. J. Kelly, J. L. Foster, and D. K. Hall, 2003: Global SWE monitoring using AMSR-E data. *2003 IEEE Int. Geoscience and Remote Sensing Symp. Proc.*, Vol. 1, Toulouse, France, IEEE, 680–682, <https://doi.org/10.1109/IGARSS.2003.1293880>.
- Dietz, A. J., C. Kuenzer, U. Gessner, and S. Dech, 2012: Remote sensing of snow—A review of available methods. *Int. J. Remote Sens.*, **33**, 4094–4134, <https://doi.org/10.1080/01431161.2011.640964>.
- Ferraro, R. R., E. A. Smith, W. Berg, and G. J. Huffman, 1998: A screening methodology for passive microwave precipitation retrieval algorithms. *J. Atmos. Sci.*, **55**, 1583–1600, [https://doi.org/10.1175/1520-0469\(1998\)055<1583:ASMFPM>2.0.CO;2](https://doi.org/10.1175/1520-0469(1998)055<1583:ASMFPM>2.0.CO;2).
- Gelaro, R., and Coauthors, 2017: The Modern-Era Retrospective Analysis for Research and Applications, version 2 (MERRA-2). *J. Climate*, **30**, 5419–5454, <https://doi.org/10.1175/JCLI-D-16-0758.1>.
- Gopalan, K., N. Wang, R. Ferraro, and C. Liu, 2010: Status of the TRMM 2A12 land precipitation algorithm. *J. Atmos. Oceanic Technol.*, **27**, 1343–1354, <https://doi.org/10.1175/2010JTECHA1454.1>.
- Greco, M., W. S. Olson, S. J. Munchak, S. Ringerud, L. Liao, Z. Haddad, B. L. Kelley, and S. F. McLaughlin, 2016: The GPM combined algorithm. *J. Atmos. Oceanic Technol.*, **33**, 2225–2245, <https://doi.org/10.1175/JTECH-D-16-0019.1>.
- Hamada, A., and Y. N. Takayabu, 2016: Improvements in detection of light precipitation with the Global Precipitation Measurement Dual-Frequency Precipitation Radar (GPM DPR). *J. Atmos. Oceanic Technol.*, **33**, 653–667, <https://doi.org/10.1175/JTECH-D-15-0097.1>.
- Heidke, P., 1926: Berechnung des Erfolges und der Gute der Windstarkevorhersagen im Sturmwarnungsdienst (Measures of success and goodness of wind force forecasts by the gale-warning service). *Geogr. Ann.*, **8**, 301–349.
- Hou, A. Y., and Coauthors, 2014: The Global Precipitation Measurement Mission. *Bull. Amer. Meteor. Soc.*, **95**, 701–722, <https://doi.org/10.1175/BAMS-D-13-00164.1>.
- Kidd, C., J. Tan, P.-E. Kirstetter, and W. A. Petersen, 2018: Validation of the Version 05 Level 2 precipitation products from the GPM Core Observatory and constellation satellite sensors. *Quart. J. Roy. Meteor. Soc.*, **144**, 313–328, <https://doi.org/10.1002/qj.3175>.
- Kongoli, C., C. A. Dean, S. R. Helfrich, and R. R. Ferraro, 2007: Evaluating the potential of a blended passive microwave-interactive multi-sensor product for improved mapping of snow cover and estimations of snow water equivalent. *Hydrol. Processes*, **21**, 1597–1607, <https://doi.org/10.1002/hyp.6722>.
- Kummerow, C., and L. Giglio, 1994: A passive microwave technique for estimating rainfall and vertical structure information from space. Part I: Algorithm description. *J. Appl. Meteor.*, **33**, 3–18, [https://doi.org/10.1175/1520-0450\(1994\)033<0003:APMTFE>2.0.CO;2](https://doi.org/10.1175/1520-0450(1994)033<0003:APMTFE>2.0.CO;2).
- , S. Ringerud, J. Crook, D. Randel, and W. Berg, 2011: An observationally generated a priori database for microwave rainfall retrievals. *J. Atmos. Oceanic Technol.*, **28**, 113–130, <https://doi.org/10.1175/2010JTECHA1468.1>.
- , D. L. Randel, M. Kulie, N. Wang, R. Ferraro, S. Joseph Munchak, and V. Petkovic, 2015: The evolution of the Goddard profiling algorithm to a fully parametric scheme. *J. Atmos. Oceanic Technol.*, **32**, 2265–2280, <https://doi.org/10.1175/JTECH-D-15-0039.1>.
- , and Coauthors, 2017: Global Precipitation Measurement (GPM) Mission. NASA Algorithm Theoretical Basis Doc., 63 pp., https://pmm.nasa.gov/sites/default/files/document_files/ATBD_GPM_GPROF_June1_2017.pdf.
- Levizzani, V., C. Kidd, D. Kirschbaum, C. Kummerow, K. Nakamura, and F. Turk, Eds., 2020: *Satellite Precipitation Measurement*. Advances in Global Change Research, Vol. 67, Springer, 450 pp.
- Mätzler, C., 1994: Passive microwave signatures of landscapes in winter. *Meteor. Atmos. Phys.*, **54**, 241–260, <https://doi.org/10.1007/BF01030063>.
- Meng, H., J. Dong, R. Ferraro, B. Yan, L. Zhao, C. Kongoli, N.-Y. Wang, and B. Zavadsky, 2017: A 1DVAR-based snowfall rate retrieval algorithm for passive microwave radiometers. *J. Geophys. Res. Atmos.*, **122**, 6520–6540, <https://doi.org/10.1002/2016JD026325>.
- Munchak, S. J., S. Ringerud, L. Brucker, Y. You I. de Gelis, and C. Prigent, 2020: An active-passive microwave land surface database from GPM. *IEEE Trans. Geosci. Remote Sens.*, **58**, 6224–6242, <https://doi.org/10.1109/TGRS.2020.2975477>.
- NASA PPS and X-Cal Working Group, 2017: NASA Global Precipitation Measurement level GPM 1C algorithms, version 1.8. NASA Algorithm Theoretical Basis Doc., 81 pp., https://pps.gsfc.nasa.gov/Documents/L1C_ATBD.pdf.
- Olson, W. S., C. D. Kummerow, G. M. Heymsfield, and L. Giglio, 1996: A method for combined passive-active microwave retrievals of cloud and precipitation profiles. *J. Appl. Meteor.*, **35**, 1763–1789, [https://doi.org/10.1175/1520-0450\(1996\)035<1763:AMFCPM>2.0.CO;2](https://doi.org/10.1175/1520-0450(1996)035<1763:AMFCPM>2.0.CO;2).
- Petković, V., and C. D. Kummerow, 2017: Understanding the sources of satellite passive microwave rainfall retrieval systematic errors over land. *J. Appl. Meteor. Climatol.*, **56**, 597–614, <https://doi.org/10.1175/JAMC-D-16-0174.1>.
- Petty, G. W., and W. F. Krajewski, 1996: Satellite estimation of precipitation over land. *Hydrol. Sci. J.*, **41**, 433–451, <https://doi.org/10.1080/02626669609491519>.
- Romanov, P., G. Gutman, and I. Csiszar, 2000: Automated monitoring of snow cover over North America with multispectral satellite data. *J. Appl. Meteor.*, **39**, 1866–1880, [https://doi.org/10.1175/1520-0450\(2000\)039<1866:AMOSCO>2.0.CO;2](https://doi.org/10.1175/1520-0450(2000)039<1866:AMOSCO>2.0.CO;2).

- Shahroudi, N., and W. Rossow, 2014: Using land surface microwave emissivities to isolate the signature of snow on different surface types. *Remote Sens. Environ.*, **152**, 638–653, <https://doi.org/10.1016/j.rse.2014.07.008>.
- Skofronick-Jackson, G., and Coauthors, 2017: The Global Precipitation Measurement (GPM) mission for science and society. *Bull. Amer. Meteor. Soc.*, **98**, 1679–1695, <https://doi.org/10.1175/BAMS-D-15-00306.1>.
- Smith, E. A., C. Kummerow, and A. Mugnai, 1994: The emergence of inversion-type profile algorithms for estimation of precipitation from satellite passive microwave measurements. *Remote Sens. Rev.*, **11**, 211–242, <https://doi.org/10.1080/02757259409532266>.
- , and Coauthors, 2013: Transitioning from CRD to CDRD in Bayesian retrieval of rainfall from satellite passive microwave measurements: Part III—Identification of optimal meteorological tags. *Nat. Hazards Earth Syst. Sci.*, **13**, 1185–1208, <https://doi.org/10.5194/nhess-13-1185-2013>.
- You, Y., C. Peters-Lidard, J. Turk, S. Ringerud, and S. Yang, 2017: Improving overland precipitation retrieval with brightness temperature temporal variation. *J. Hydrometeor.*, **18**, 2355–2383, <https://doi.org/10.1175/JHM-D-17-0050.1>.
- Zhang, J., and Coauthors, 2011: National Mosaic and Multi-Sensor QPE (NMQ) system: Description, results, and future plans. *Bull. Amer. Meteor. Soc.*, **92**, 1321–1338, <https://doi.org/10.1175/2011BAMS-D-11-00047.1>.