Enhancing PMW Satellite Precipitation Estimation: Detecting Convective Class

VELJKO PETKOVIC
Earth System Science Interdisciplinary Center/Cooperative Institute for Satellite Earth System Studies, University of Maryland, College Park, College Park, Maryland, and Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

MARKO ORESCANIN
Department of Computer Science, Naval Postgraduate School, Monterey, California

PIERRE KIRSTETTER
School of Meteorology, and School of Civil Engineering and Environmental Science, and Advanced Radar Research Center, University of Oklahoma, and NOAA/National Severe Storms Laboratory, Norman, Oklahoma

CHRISTIAN KUMMEROW
Department of Atmospheric Science, Colorado State University, Fort Collins, Colorado

RALPH FERRARO
Earth System Science Interdisciplinary Center/Cooperative Institute for Satellite Earth System Studies, University of Maryland, College Park, and NOAA/NESDIS/STAR, College Park, Maryland

(Manuscript received 26 January 2019, in final form 27 August 2019)

ABSTRACT

A decades-long effort in observing precipitation from space has led to continuous improvements of satellite-derived passive microwave (PMW) large-scale precipitation products. However, due to a limited ability to relate observed radiometric signatures to precipitation type (convective and stratiform) and associated precipitation rate variability, PMW retrievals are prone to large systematic errors at instantaneous scales. The present study explores the use of deep learning approach in extracting the information content from PMW observation vectors to help identify precipitation types. A deep learning neural network model (DNN) is developed to retrieve the convective type in precipitating systems from PMW observations. A 12-month period of Global Precipitation Measurement mission Microwave Imager (GMI) observations is used as a dataset for model development and verification. The proposed DNN model is shown to accurately predict precipitation types for 85% of total precipitation volume. The model reduces precipitation rate bias associated with convective and stratiform precipitation in the GPM operational algorithm by a factor of 2 while preserving the correlation with reference precipitation rates, and is insensitive to surface type variability. Based on comparisons against currently used convective schemes, it is concluded that the neural network approach has the potential to address regime-specific PMW satellite precipitation biases affecting GPM operations.

1. Introduction and motivation

Variability in precipitation typology affects vertical water and energy fluxes though the associated precipitation structure, dynamics, microphysical processes, and latent heat release. The distribution of convective and stratiform precipitation impacts Earth’s radiative properties and atmospheric circulation. While the differences in microphysical processes and dynamics in convective and stratiform systems are well documented in the literature (e.g., Houze 1997), distinguishing between them remains a major challenge for remote sensing. A lack of measurements of updraft speed prevents a direct identification of precipitation types, which in turn hampers the interpretation of remotely sensed storm properties. Inaccurate understanding of hydrometeor...
distribution and evolution throughout the precipitating column ultimately affects the quality of precipitation rate estimates at the surface from both active and passive remote sensing observations. This challenge is especially pronounced in satellite observations.

Since the first spaceborne passive microwave instruments were launched in early 1970s, satellite precipitation retrievals have exploited the link between upwelling radiation and state of atmospheric column. Leveraging decades of ever-improving algorithms, coverage, and data latency, the Global Precipitation Measurement (GPM) mission (Skofronick-Jackson et al. 2018; Hou et al. 2014) represents the most advance satellite precipitation project to date. Combining both direct (gauges) and remote (radar/radiometer) measurement techniques, using ground and in-orbit observations complemented by the state-of-the-art atmosphere simulations, the GPM constellation offers full global coverage of rain and snow every 30 min at a resolution of only 0.1° and a latency of only a few hours. Freely available precipitation products are implemented across a spectrum of decision-making scientific tools, ranging from hydrology to world health. To ensure user demands for accuracy are met over broad time/space scales (e.g., nowcasting to climate), GPM precipitation products undergo continuous validation against the latest reference standards before reaching the users. In this rigorous process ground-based measurements (i.e., gauges and ground radars) typically perform well (Kirstetter et al. 2012) but their satellite counterparts face a number of challenges (Tang et al. 2014; Meyers and Ferraro 2016; Ciabatta et al. 2017). While averaged precipitation estimates from satellite precipitation products are generally quite good, their consistency across specific scales, atmospheric conditions and raining regimes is still far from the ideal (Petković and Kummerow 2017; Tan et al. 2017, 2018; Henderson et al. 2018). Currently, GPM passive microwave sensor estimates do not use precipitation type information to constrain the retrievals.

The challenge in obtaining consistent satellite precipitation estimates across varying atmospheric conditions can be attributed to the inability of current passive microwave (PMW) retrievals to properly distinguish between distinct storm profiles and their corresponding surface precipitation rates. Due to the nonlinear response of the atmospheric content (i.e., hydrometeors) to the upwelling microwave radiation, the individual radiometric signatures of surface and atmospheric properties are obscured. Hence the inverse problem of retrieving the atmospheric contribution from the MW radiance is underconstrained by nature. In other words, solving an inverse problem using an incomplete observation vector results in a deficient estimate of the corresponding state vector.

To demonstrate this effect, instantaneous precipitation estimates from a passive and an active microwave instruments, mounted side-by-side on board the GPM core satellite, are compared for two distinct precipitating regimes—convective and stratiform, as defined by the GPM radar (details in section 2). Using a full year of globally observed (66°S–66°N) precipitation rates over land, upon removal of the overall difference, GPM PMW satellite retrieval reveals opposite mean deviations (hereafter referred to as biases) relative to the radar-based estimates under the two precipitation regimes. Illustrated in Fig. 1, the comparison of the distributions of active- and passive-estimated precipitation rates, for convective and stratiform cases, suggests a negative 29% bias for convective regime and a positive 26% bias for the stratiform cases. A closer inspection of active- and passive-estimated precipitation distributions confirms that PMW biases exist over the majority of their characteristic precipitation rates. Results over ocean, not shown here, yield the same general conclusions.

A review of the existing literature confirms the impact of precipitation type on PMW retrieval performance seen in Fig. 1 even when ground radars are used as a reference (e.g., Henderson et al. 2017). While both PMW and radar retrievals certainly contribute to the bias dipole, the passive microwave’s bias against both spaceborne and gauge-calibrated surface radar observations leads us to believe that the bulk of the bias is coming from the passive microwave retrievals themselves. The cause is typically attributed to upwelling radiation appearing similar for convective and stratiform regimes within the observation frequencies of PMW radiometers, despite their distinct rainfall profiles. With limited ability to recognize and address the exact source of subtle changes in the observed radiances, the retrieval recognizes the average value of the two as the best fit in its minimization function. Consequently, the solution is centered between the two regimes.

To provide a better insight, the distribution of total precipitation estimated by the two GPM sensors is given in Fig. 2 as a function of GPM radar-defined convective fraction. Clearly, the PMW retrieval (light blue) struggles to match the radar’s (gray) distribution of global precipitation when sampled by precipitation type. While this result may in fact be the best performance a PMW precipitation retrieval has achieved to date, the comparison raises a simple question: Is there any information the PMW retrieval could use to improve this performance? We hypothesize that accounting for the precipitation type in the retrieval has the potential to generate more
accurate precipitation estimates. Figure 2 shows a precipitation fraction distribution of the PMW retrieval when the current operational algorithm is allowed to use radar-observed convective fraction information (bright blue) as an additional ancillary parameter (see section 4a). As expected, a better match to the reference suggests that the information on convective fraction might be a key to mitigating PMW biases seen in Fig. 1. However, radar observations of precipitation from space are sparse, typically limited to research missions (e.g., TRMM and GPM) and intended to serve as a reference rather than a supplement to PMW observations. It is therefore important to assess if convective/stratiform information can be inferred from the passive microwave information itself. Yet, despite sustained, decades-long effort to identify a robust link between PMW observations and convective fraction, only a few regression methods with modest skill are available. These methods largely utilize the spatial variability of the brightness temperature (Tb) of the high-frequency channels (e.g., 30 GHz and above). Thus, the convective fraction of a radiometer field of view (FOV) is typically calculated by employing 37–89-GHz signal ratio (the liquid water emission and ice scattering), spatial variability of 85-GHz Tb depression, or the gradients and standard deviation of 85-GHz Tb adjacent pixels (Anagnostou and Kummerow 1997; Grecu and Anagnostou 2001; Kummerow et al. 2001; Olson et al. 2001; McCollum and Ferraro 2003; Dinku and Anagnostou 2006; Gopalan et al. 2010). These techniques still find their applications in operational PMW retrievals at NASA and NOAA centers (details in section 4). More recent studies have investigated the use of lightning information in complementing the PMW brightness temperatures to improve PMW algorithm’s ability to discriminate between convective and stratiform regimes (Wang et al. 2012). While the results indicate a great potential, especially considering the expansion of Geostationary Lightning Mapper...
(GLM) coverage in the near future (Goodman et al. 2013), this information is limited to the lightning-active clouds and era of GLM satellite sensors.

Given the four-decade-long effort in linking PMW observations and storm morphology, little, if any, room has been left for a potentially novel physically based approach to emerge. However, recent advances in deep learning methods with neural networks may offer perhaps not new but for the first time fully applicable models that could better exploit the information content in PMW observations. This study seeks to investigate such a possibility through the use of deep learning for both retrieving precipitation types and improving the performance of PMW precipitation retrievals.

Defined nearly 70 years ago, artificial neural networks (NNs) have been generally labeled as “black box” mathematical methods of limited application and performance. Major obstacles to their successful application resulted from their extensive need for computational power, and lack of large-enough datasets that could provide a robust, noise-resistant relationship between the predictor and predictand. The late 1980s and early 1990s brought advances in affordable computing hardware and availability of large satellite datasets. This initiated NN satellite applications, leading to a number of NN schemes present today, mainly covering image processing, classification, series prediction, and geophysical retrievals in general. A study of Tapiador et al. (2004) offers an extensive overview and description of the use of deep learning for both retrieving precipitation types and improving the performance of PMW precipitation retrievals.

The paper is organized as follows: details on the instruments and data collection, in addition to detailed description of the DNN model are given in section 2. Results on the accuracy of DNN model in predicting convective/stratiform separation are in section 3, followed by applications and conclusions in the last two sections.

2. Data collection and methods

Deep learning neural networks are data-driven models. The performance of DNNs in terms of accuracy and their ability to generalize to new inputs depends on the representativeness, quantity and quality of the training dataset. To establish a baseline model and evaluate the performance of the DNN approach we propose a relatively simple scheme and a widely available satellite dataset. Detailed descriptions of the datasets and DNN model are given below.

a. Instruments and data

This study employs 2 years, from September 2014 to August 2015 and from January to December 2017, of the GPM core satellite global observations (66°S–66°N) to explore accuracy and potential of neural network approach in retrieving precipitation type from PMW measurements. Combined observations from GPM’s Microwave Imager (GMI) and Dual-Frequency Precipitation Radar (DPR) are used to relate brightness temperatures to the convective/stratiform information.

1) GPM MICROWAVE IMAGER

Since the launch of the GPM core platform in 2014, the GMI instrument (Draper et al. 2015) serves as a
calibration standard for PMW conical-scanning radiometers in the GPM constellation. The GMI channel-dependent FOV decreases from approximately 25 to 6 km as the frequency increases from 10 to 183 GHz, leaving a gap in coverage at high-frequency channels between consecutive, 221-pixel-wide, 13-km apart, scan lines. Brightness temperatures observed at 13 microwave channels (10.65 H/V, 18.7 H/V, 23.8 V, 36.5 H/V, 89.0 H/V, 166 V/H, and 183.3 ± 3/7 V GHz) are stored in GPM level-1 standard product (GPM_BASEGPMGMI_XCAL–V05; GPM Science Team 2016) and freely available through NASA’s data exchange portals (e.g., https://storm.pps.eosdis.nasa.gov).

2) GMI PRECIPITATION RATE

The GPM Goddard profiling (GPROF) algorithm (Kummerow et al. 2015), described in section 4 and slightly modified for purposes of this study, uses the GMI Tbs to provide precipitation rate estimates for most of the GPM PMW observations. The product used in this study is GMI GPROF standard output for level-2 data (GPM_2AGPROFgpmgmi; Iguchi and Meneghini 2016; NASA 2018), freely available through the same portal as the GMI Tbs.

3) GPM COMBINED PRODUCT PRECIPITATION RATE AND TYPE

Designed to extract the best from both the passive and active microwave instruments observations, the GPM combined algorithm (hereafter DPR-combined; Grecu et al. 2016) produces the best high-resolution precipitation estimates from a spaceborne platform. As such, the DPR-combined product (GPM_2BCMB; Olson 2017) is used in this study to provide a reference regarding instantaneous storm structure, namely the surface precipitation rate and convective/stratiform class. While the precipitation rate retrieval (Olson and Masunaga 2016) is developed specifically to use information from both microwave sensors (active and passive), the precipitation type in DPR-combined product relies strictly on GPM’s DPR. The technical description of the DPR instrument and full algorithm description can be found in Iguchi et al. (2015), while a brief overview of its precipitation type classification criteria is provided here for completeness purposes.

Based on Iguchi et al. (2009), the DPR algorithm classifies all precipitating FOVs into three major categories—convective, stratiform, and other—using a detection of a bright band (BB) as a criterion in so-called vertical and horizontal methods (Awaka et al. 2016). In this process, if a BB is detected, the pixel is considered to be stratiform, unless the attenuation-corrected reflectivity below the BB exceeds a threshold of 39 dBZ, in which case pixel is flagged as convective. If no BB is detected, the pixel is classified as convective if corrected reflectivity anywhere in the profile exceeds a 39-dBZ threshold. In addition, if the maximum reflectivity for a given pixel stands out against the background of the surrounding pixels, or exceeds a predetermined threshold (i.e., 40 dBZ), the pixel and its first neighbors are regarded as convective. In any other scenario the pixel is classified as other (this class typically accounts for less than 10% of all precipitating pixels and is ignored in this study).

Using both DPR-combined precipitation rate and type, this study assigns a precipitation type to a GMI FOV through the volumetric convective fraction (i.e., a fraction of convective relative to the total precipitation rate within a GMI FOV). When more than 50% of the FOV’s total precipitation volume is convective the FOV is labeled as convective; otherwise the FOV is labeled as stratiform. No attempt toward improvement of the DPR classification scheme is made, albeit the known challenges (Iguchi et al. 2015; Kirstetter et al. 2014).

4) GPM GROUND VALIDATION MULTI-RADAR/ MULTI-SENSOR PRECIPITATION RATE AND TYPE

The GPM Ground Validation Multi-Radar/Multi-Sensor (GV-MRMS) dataset is used in this study as an independent reference for qualitative assessment of the satellite DPR-combined precipitation type product (see section 4b). MRMS quantitative precipitation estimates are based on ground polarimetric WSR-88D radar and automated rain gauge networks (Zhang et al. 2011, 2016) offering products at 0.01° spatial and 2-min temporal resolution over the conterminous United States and Canada. Extensive gauge correction and quantity filters are applied on MRMS data to generate GV-MRMS products specifically adapted to satellite purposes and needs for the highest data quality, as described in Kirstetter et al. (2012, 2014). To provide fair comparisons to DPR-combined product, GV-MRMS precipitation type flags are grouped into two categories using the MRMS-provided flags (Zhang et al. 2016): convective (including hail, convective and tropical convective) and stratiform (including tropical stratiform and stratiform), and the same volume threshold (50%) as in the DPR-combined case.

b. Machine learning: Deep neural network, training, and validation

A common approach to a classification problem would be to engineer input features for a classification model (e.g., random forest, support vector machines, k-nearest neighbor). Such an approach requires a high level of expert knowledge to hand-craft the features
(e.g., spatial properties of the input fields; cross-channel correlations; derived class metrics) while bounding the space the classification model considers in the search for a solution. To minimize constraints and biases posed by hand-crafted features and subjective choices we opted to employ neural networks, which in contrast can learn representations of data and use them for classification. Studies by Meyer et al. (2016) and Tang et al. (2018) provide an overview of four neural network employed schemes used in recent years for satellite rainfall estimates. In this study, the problem of classifying precipitating satellite FOVs is posed as a supervised learning problem given the availability of the data and the ability to prepare and label a large dataset for the training of the classification algorithm. Specifically, deep learning methods are used to establish a link between the raw GMI channels and two precipitation classes (i.e., convective and stratiform precipitation type) using the DPR’s convective/stratiform flag, given by GPM_2ADPR product (Iguchi and Meneghini 2017), as a relative reference. In this application of deep learning, a feedforward neural network is used with fully connected architecture (LeCun et al. 2015) relying on TensorFlow API, an open source software library for machine learning intelligence (Abadi et al. 2015).

Using one year of observations (September 2014–August 2015) the described DNN model is trained to predict a precipitation type class. Randomly choosing 24-h intervals of labeled predictors, this 12-month period is split into training, validation, and test subsets using a 70/20/10 ratio. The input data are chosen to be $3 \times 5$ FOV fields of GMI brightness temperatures at all 13 channels, centered on the retrieving FOV. These approximate $20 \times 20$ km patches of Tbs provide temperature gradient information that is known to be well correlated with precipitation type change, while 13 GMI frequencies capture hydrometeor and surface type variability. The resulting 195 predictors are trained against FOV class determined by the ratio of total to convective DPR-combined rain rate within the GMI FOV. Training is performed separately for land and ocean surface.

### THE ARCHITECTURE

Figure 3 summarizes the architecture of the neural network: 195 input features, two fully connected hidden layers, where the first hidden layer has 195 neurons, the second hidden layer has 96 neurons, and the last layer (i.e., the output) has 2 neurons as the number of classes. A hyperbolic tangent (tanh) function is used as the activation function for the hidden layers, while a softmax activation function is used for the output layer. Additionally, batch norm is applied to the input layer in order to normalize batch statistics (Ioffe and Szegedy 2015) and speed up the training. Cross-entropy loss is used as loss function and coupled with the Adam optimization method (Kingma and Ba 2014).

Google’s Cloud Machine Learning Engine (https://cloud.google.com/ml-engine/), ml-engine, is utilized for the training of the model with both training and validation data accessible by the framework through the Google Cloud storage service. The ml-engine hardware configuration was set to a single worker GPU and a single parameter server, both of Google’s internal “standard_type.” Hyperparameter constants include learning rate $\lambda$ and a batch size. A search for learning rate spans $\lambda \in [5 \times 10^{-5}, 3 \times 10^{-3}]$ and the batch size $\in [256, 1024]$. The ml-engine built-in hyperparameter search method (Golovin et al. 2017) and the accuracy on the validation data to determine these constants are utilized.

The highest accuracy in training on validation data was found for $\lambda = 9.2 \times 10^{-5}$ and batch size of 432. The overfitting is monitored during training via validation.
dataset. Figure 4 shows validation dataset accuracy and loss, confirming the model does not overfit in training. The next subsection reports the model’s performance on the test dataset.

3. Results

Using a predicted class probability of 50% as a threshold and testing the model on an independent 12-month period (January–December 2017) dataset, the model achieved the 87% of overall accuracy in precipitation type classification (by count) when DPR retrieval is used as a reference. Summarized results, given in Table 1, indicate an imbalance between the two classes with accuracy of 97% for stratiform and approximately 40% for convective class. This result is shown to be insensitive when changes are made to convective-to-stratiform count ratios in the training dataset. When assessed using the total precipitation amount by each of the classes, the results suggest the correctly classified convective scenes account for almost 70% of total convective precipitation volume. This implies that most of the classification errors occur for relatively light precipitation classified as convective. The method correctly classifies 98% of stratiform precipitation volume, with the relatively high Heidke skill score (HSS; Wilks 2011) of 0.47 and the two subgroups of accurately assigned classes accounting for nearly 84% of total precipitation.

To provide more insight to the DNN model’s performance, Fig. 5 shows the probability of correctly predicted class as a function of convective fraction (left) and precipitation rate (right). With a mean probability (red diamonds) above 75% and with 75% of FOVs (the bottom of blue box) being above the 65% probability value, the model shows a robust performance in respect with both precipitation regime and intensity.

The stability and high percentage of accurately classified precipitation volumes are strong indicators of model’s potential to mitigate PMW precipitation type biases presented in Fig. 1. The following section evaluates the extent of this potential.

4. Applications

This section aims at testing the potential benefit of the above classification scheme on the passive microwave retrieval algorithm used in GPM.

a. PMW rainfall retrieval

Developed at NASA’s Goddard Space Flight Center in the mid-1990s (Kummerow and Giglio 1994), the GPROF algorithm is used operationally for PMW precipitation retrieval at NASA and NOAA. NASA’s Precipitation Processing System (PPS) runs the most recent GPROF Version 5 (V05) algorithm to process PMW observations from a constellation of cross-track (e.g., Kidd et al. 2016) and conical scanning sensors, including GMI, SSMI/S (Kunkee et al. 2008), AMSR2 (Shimoda 2005), ATMS (Muth et al. 2005), MHS (Edwards and Pawlak 2000), and others (see the appendix for a full list of acronyms and abbreviations). NOAA’s operations rely on an older GPROF version (2010v2) to retrieve AMSR2 precipitation rates (Gopalan et al. 2010; Meyers et al. 2015). Only the most recent version of the algorithm is detailed below, while the evolution of the retrieval and differences

<table>
<thead>
<tr>
<th>Observed convective</th>
<th>Observed stratiform</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Retrieved convective</td>
<td>29152 (39%)</td>
<td>44705 (61%)</td>
</tr>
<tr>
<td>Retrieved stratiform</td>
<td>5104 (2%)</td>
<td>299859 (98%)</td>
</tr>
<tr>
<td>Total</td>
<td>34256 (9%)</td>
<td>344564 (91%)</td>
</tr>
</tbody>
</table>
between the two versions are documented in work of Kummerow et al. (2015).

GPROF utilizes a Bayesian approach that employs a priori information on the relationship between hydrometeor profiles and corresponding radiances. Using the DPR-combined algorithm as a primary source of precipitation profiles, coupled with radiative transfer models, GPROF computes Tbs for any sensor that forms part of the GPM constellation (Kummerow et al. 2011). The algorithm first groups the entire a priori database by using ancillary information (TPW, surface type, and 2-m temperature) to subset the database by the observed large-scale conditions. As part of this process, surface types are defined using SSM/I observed emissivity climatology (Aires et al. 2011) updated daily by NOAA’s AutoSnow product (Romanov et al. 2000), while TPW and 2-m temperature come from reanalysis datasets such as ECMWF (Dee et al. 2011) and JMA’s global analysis (GANAL; JMA 2000). The database elements are averaged through the Bayesian scheme where DPR-combined precipitation rates are assigned a weight proportional to their respective probability given by Eq. (1):

$$r_r = \frac{\sum r_i w_i}{\sum w_i}, \quad \text{where} \quad w_i = \exp\left[-0.5\{T_b - T_{b,f}(r_i)\}^T S^{-1} \{T_b - T_{b,f}(r_i)\}\right],$$

(1)

here, $i$ is an element of the a priori database, $S$ is the $T_b$ error covariance (to account for both forward model and instrument errors), $T_b$ is the FOV observed brightness temperature, and $T_{b,f}(r_i)$ is $r_i$-associated brightness temperature while $r_r$ is the retrieved precipitation rate.

This approach is relatively easy to apply to any PMW sensor and ensures preservation of the global precipitation rate distribution given by the a priori reference (e.g., DPR-combined product). However, when a sensor is incapable of distinguishing between radiometrically similar scenes of various precipitation rates, the Bayesian technique gives equal weight to database elements of mismatching hydrometeor profiles. Consequently, the error introduced to the instantaneous precipitation rates becomes a function of precipitation type (as seen in Fig. 1). Thus, if the averaging is limited only to a priori database elements of similar precipitation types, it should be possible to decrease, if not fully eliminate, the biases highlighted in Fig. 1. Before testing for potential of the DNN model to contribute toward this goal, a qualitative evaluation of the model’s ability to predict the convective/stratiform flag is presented.

b. Retrieving convective fraction from PMW observations

Using an example of the squall line observed over the Midwest U.S. on 13 July 2015 (Fig. 6), a performance of the DNN model is compared to that of ground and satellite retrievals. Figure 6 offers a side-by-side comparison of the system’s precipitation type partitioning observed by GPM core satellite and ground radar network at 0720 UTC. The same 50% convective volume threshold is applied to delineate two precipitation classes [note: in GPROF case, Eq. (1) defines both total and
FIG. 6. A squall line over the Midwest U.S. on 13 Jul 2015. Precipitation rate by (top left) DPR-combined and (top right) Multi Radar Multi Sensor network. Precipitation type by (middle left) DPR-combined and (middle right) Multi Radar Multi Sensor network. Precipitation type by (bottom left) GPROF algorithm and (bottom right) DNN. Black/white lines denote the edges of the DPR swath for the GPM core-satellite overpass (orbit number 7789) at 0720 UTC.
convective precipitation, using total and convective DPR rate, respectively]. Qualitative analysis of precipitation rate and type retrieved by GV-MRMS, DPR-combined, GPROF, and DNN suggest the following: 1) DPR-combined and GV-MRMS precipitation rate and type show good overall agreement but some disparity is present; 2) while GPROF algorithm captures the general structure of the squall line it largely overestimates the convective portion of the system; 3) DNN model outperforms the PMW retrieval showing only minor inconsistencies with both ground and satellite references.

Upon closer inspection of Fig. 6, it can be noted that the DNN model does not assign a convective flag to the area on the left side of the convective bow (46°N, 90.5°W)—the only area over which ground and satellite radars show significant disagreement. It is likely that this region is erroneously labeled as convective by the DPR, and does not have the typical spectral signal of other DPR-defined convective regions. While investigating this behavior (of the DNN perhaps outperforming the training data in this particular case) is beyond the scope of this study, the more robust result is that GPROF without DNN largely overestimates the convective portion of the storm, which directly links to Fig. 1 and the precipitation type bias. Given the nature of convection, one can typically expect higher precipitation rates over convective pixels relative to the stratiform ones. Therefore, the erroneous assignment of convective type over stratiform region by the GPROF retrieval (Fig. 6 bottom left) should, in general, lead to an overestimation of stratiform scenes. On the other hand, as long as the Bayesian averaging allows stratiform database elements to be included in the retrieval of convective pixels, one should expect an underestimation of precipitation rates. The current GPROF algorithm does not have an effective mechanism to separate the two, resulting in the precipitation type bias dipole, as confirmed by Fig. 1. To test if the performance of the DNN model seen in Fig. 6 can mitigate these biases, the model is implemented directly to the GPROF retrieval and tested on a longer time period of observations.

c. Mitigating PMW precipitation bias via retrieved convective fraction

To test whether the DNN model’s precipitation type retrieval can offer complementary information to the GPROF algorithm, the model is implemented directly into the retrieval’s scheme. In this two-step process, using Tb vectors as an input, the DNN model first assigns a convective/stratiform flag to each element of GPROF’s a priori database. Then, in the retrieving process, the same DNN model employs observed brightness temperatures to assess each FOV precipitation type. Classified either as convective or stratiform, the FOVs are assigned Bayesian average precipitation rates through the Eq. (1) but now considering only database elements of the same precipitation type. In other words, DNN output is used to additionally constrain the a priori database and eliminate nonrelevant inputs in Eq. (1). The results are summarized in Figs. 7 and 8 and Table 2.

Figure 7 presents the same analysis shown in Fig. 1, where global over land GMI observations for the 12-month period are separated into convective and stratiform types, but in addition to the DPR-combined (gray; reference) and GPROF-GMI algorithm (light colors), it displays the DNN-adjusted result (bright colors).
Clearly, the adjusted product is a better match to the reference. The majority of the precipitation rate spectrum has the rain fraction of the adjusted product much closer to that of DPR-combined. Exceptions are convective rates below 1 mm h\(^{-1}\) and stratiform rates above 5 mm h\(^{-1}\). However, these account for less than a tenth of the overall bias and in fact do not reflect DNN-induced errors. Their origin lays in the Bayesian scheme itself, where the solution of Eq. (1) always leans toward the mean value of averaging elements. Since these precipitation rates are common for both precipitation types, the link between the precipitation type and Tb is rather weak and too ambiguous for DNN to extract from Tb vector alone. Consequently, the bias centers around the peak of precipitation fraction.

Following the approach in section 1, the DNN effect on the PMW retrieval is tested through comparison of total precipitation fraction distributions. The analysis presented in Fig. 2 are repeated and shown in Fig. 8, only this time the optimal PMW retrieval is replaced by that of the DNN-adjusted retrieval. Showing consistent improvement and reducing the total precipitation misclassification, the comparison clearly confirms the potential DNN-model has in reducing precipitation type bias of PMW retrievals.

To test the robustness of this result, the PMW retrieval is extended to a full domain of the GPM core satellite (66°S–66°N). Table 2 summaries the effect of the DNN model implementation on the bias and correlation coefficient for the two precipitation types over both land and ocean for the aforementioned 12-month period. On average, the precipitation rate bias is reduced by a factor of 2 (i.e., 45%–50%) while the correlation coefficients remain stable.

Interestingly, when combined, Figs. 7 and 8 suggest the DNN model may in fact be close to the optimal. The model’s impact to GPROF performance is poor only for precipitation rates common to both convective and stratiform regimes, where the link between DPR precipitation type and Tb is rather weak and ambiguous. While this requires further investigation, a comparison of the accuracy of the retrieved precipitation type by the DNN model and current state-of-the-art PMW algorithm is considered next. Table 3 summarizes assessments of convective/statiform separation given by NOAA’s (GPROF-2010) and NASA’s (GPROF V05) operational PMW precipitation retrievals, using once again the DPR-combined product as a reference. To maximize each scheme’s performance, all three methods consider only the domain they have been trained for and use the same 0.5 threshold of convective to total precipitation ratio as a delineator between convective and stratiform FOVs. Following Gopalan et al. (2010), GPROF-2010 is assessed over tropical land (20°S–20°N), while the other two consider land coverage of the GPM core satellite. Percentage of both volume and counts relative to the number of predicted elements per category is given in addition to the HSS using the same approach as in section 3. The results suggest that, when the DPR is used as a reference, the DNN model outperforms the other two in all aspects except in the

![Fig. 8. Total precipitation fraction for year 2017 as a function of convective fraction as retrieved by DPR-combined (reference; gray), original (light blue), and adjusted (bright blue) PMW retrieval.](image-url)

<table>
<thead>
<tr>
<th>Table 2. Bias and correlation coefficients of the original (V05) and DNN-adjusted GPROF precipitation rates retrieved globally (66°S–66°N) during 2017.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Bias (%)</strong></td>
</tr>
<tr>
<td>Convective</td>
</tr>
<tr>
<td><strong>Land</strong></td>
</tr>
<tr>
<td>GPROF V5</td>
</tr>
<tr>
<td>GPROF w/DNN</td>
</tr>
<tr>
<td><strong>Ocean</strong></td>
</tr>
<tr>
<td>GPROF V5</td>
</tr>
<tr>
<td>GPROF w/DNN</td>
</tr>
</tbody>
</table>
accuracy of the convective class. However, given the HSSs and percent of total precipitation volume under the convective class, the DNN model has the best potential for addressing precipitation-type related biases of the PMW retrievals.

5. Summary and conclusions

The ability of deep learning approaches in retrieving precipitation type from passive microwave observations is tested using GPM mission radiometer (GMI) and radar (DPR) observations. A fully connected DNN model suggests significant improvement in detection skill of precipitation type over the existing operational PMW schemes (i.e., GPROF) when trained on collocated DPR-combined precipitation types and GMI’s 13-channel Tbs vectors. Tasked to reproduce DPR-based classification of precipitating FOVs (into convective and stratiform scenes), the model correctly assigns a class to 84% of total precipitation volume, with overall accuracy on the scenes’ type of 87%. This is achieved with the median probability of class prediction value exceeding 75% mark at all precipitation rates for all convective fractions. Compared to some state-of-the-art satellite passive microwave retrievals, this result shows a 10%–50% improvement in both accuracy and precipitation volume within each of the two categories. The model reduces precipitation rate bias associated with convective and stratiform precipitation in the GPM operational PMW algorithm by a factor of 2 while preserving the correlation with reference precipitation rates, and is insensitive to surface type variability.

The objective of this study was to simply demonstrate the capability of DNN models, rather than optimize a specific model for applications to the GPM radiometer algorithm. The above results allow us to conclude that the deep learning neural network approach for detection of convective precipitation class from satellite passive microwave measurements has an outstanding application potential. Further improvements could be sought through more sophisticated neural network models, such as convolutional approach that could potentially allow for better extraction of spatial information content from the Tbs fields. Adding input features, such as large-scale environment (e.g., Petković and Kummerow 2018) and precipitating system features (e.g., Liu et al. 2008) would likely further strengthen representation of the complex link between DPR precipitation type and PMW Tbs relying on information content beyond the observed FOV.

Besides considering more and better predictors to detect convection, another improvement would address current assumptions in the DNN scheme and its application in the GPROF Bayesian framework: accounting for inherent uncertainties in the DPR precipitation classification. Currently, a convective precipitation flag is assigned to any GMI FOV when more than 50% of the FOV’s DPR total precipitation is flagged as convective. The explicit consideration of uncertainties and mixtures of convective and stratiform precipitation within each FOV would be more consistent with the primary precipitation type information and the Bayesian formalism. It would seamlessly address the challenge of ambiguous link between the precipitation types and Tbs and remove potential biases arising from a binary classification. Ultimately, these contributions will pave the way toward more accurate precipitation estimation at local to regional scales by mitigating the PMW precipitation retrieval error dependence to precipitation types. These benefits are expected to propagate to larger scales, for example, over ENSO cycles. It will enable more reliable studies of the climatology of global convection and its evolving impact on Earth’s water and energy cycles under a changing climate.

Acknowledgments. Authors would like to acknowledge Dr. Wesley Berg (Colorado State University, Fort Collins, CO) for the valuable feedback and help with data analyses. This study was supported by the following grants: NNX16AQ66G, NA19NES4320002, 80NSSC19K0681.

APPENDIX

List of Acronyms

AMSR2 Advanced Microwave Scanning Radiometer 2
ATMS Advanced Technology Microwave Sounder
API Application programming interface
BB
DNN
DPR
ECMWF
FOV
GANAL
GLM
GMI
GPM
GP PROF
GV-MRMS
HSS
IR
JMA
MHS
ML
NASA
NN
NOAA
PERSIANN
PMW
SSMI/S
Tb
TMPA
TPW

Bright band
Deep neural network
Dual-frequency precipitation radar
European Centre for Medium-Range Weather Forecasts
Field of view
Global analysis
Geostationary Lightning Mapper
GPM Microwave Imager
Global Precipitation Measurement
Goddard profiling algorithm
Ground Validation–Multi Radar/Multi Sensor
Heidke skill score
Infrared
Japan Meteorological Agency
Microwave Humidity Sounder
Machine learning
National Aeronautics and Space Administration
Neural network
National Oceanic and Atmospheric Administration
Precipitation Estimation from Remotely Sensed Information using Artificial Neural Networks
Passive microwave
Special Sensor Microwave Imager/Sounder
Brightness temperature
TRMM Multisatellite Precipitation Analysis
Total precipitable water

REFERENCES


