Automatic Detection and Classification of Orographic Precipitation using Machine Learning

Ana Barros[1] and Malarvizhi Arulraj[1]

[1]Duke University

November 24, 2022

Abstract

Ground-clutter is a major cause of large detection and underestimation errors in satellite-based (e.g. Global Precipitation Measurement Dual Polarization Radar, GPM DPR) precipitation radar retrievals in complex terrain. Here, an Artificial Intelligence (AI) framework consisting of sequential precipitation detection and vertical structure prediction algorithms is proposed to mitigate these errors using machine learning techniques to uncover predictive associations among satellite- and ground-based measurements aided by Numerical Weather Prediction model analysis, specifically the High-Resolution Rapid Refresh (HRRR) model. The framework is implemented and tested for quantitative estimation of orographic precipitation in the Southern Appalachian Mountains (SAM). Precipitation detection relies on a Random Forest Classifier to identify rainfall based on GPM Microwave Imager (GMI) calibrated brightness temperatures (Tbs) and HRRR mixing ratios in the lower troposphere (˜ 1.5 km above ground level). The vertical structure of precipitation prediction algorithm is a Convolution Neural Network trained to learn associations among GPM DPR Ku-band reflectivity profiles, GMI Tbs, and orographic precipitation regimes in the SAM including low level light rainfall, shallow rainfall with low-level enhancement, stratiform rainfall with bright band, and deep heavy rainfall with low- and mid-level enhancement. Vertical structure classes corresponding to the distinct orographic precipitation regimes were isolated through k-means clustering of ground-based Multi-Radar/Multi-Sensor radar reflectivity profiles. The AI framework is demonstrated for automatic retrieval of warm season precipitation in the SAM over a 3-year period (2016-2019) achieving large reductions in false alarms (77%) and missed detections (82%) relative to GPM Ku-PR precipitation products, and significant rain-rate corrections (up to one order of magnitude) by using a physically-based model to capture the microphysics of low-level enhancement (i.e. seeder-feeder interactions).

Hosted file

`essoar.10502701.1.docx` available at https://authorea.com/users/543253/articles/604603-automatic-detection-and-classification-of-orographic-precipitation-using-machine-learning

Automatic Detection and Classification of Orographic Precipitation using Machine Learning

Malarvizhi Arulraj[1], Ana P. Barros[1]

1. Dept. of CEE, Duke University, Durham, NC

Corresponding Author:

Dr. Ana P. Barros

E-mail: barros@duke.edu

Phone: +1 919 660 5539

Abstract

Ground-clutter is a major cause of large detection and underestimation errors in satellite-based (e.g. Global Precipitation Measurement Dual Polarization Radar, GPM DPR) precipitation radar retrievals in complex terrain. Here, an Artificial Intelligence (AI) framework consisting of sequential precipitation detection and vertical structure prediction algorithms is proposed to mitigate these errors using machine learning techniques to uncover predictive associations among satellite- and ground-based measurements aided by Numerical Weather Prediction model analysis, specifically the High-Resolution Rapid Refresh (HRRR) model. The framework is implemented and tested for quantitative estimation of orographic precipitation in the Southern Appalachian Mountains (SAM). Precipitation detection relies on a Random Forest Classifier to identify rainfall based on GPM Microwave Imager (GMI) calibrated brightness temperatures (Tbs) and HRRR mixing ratios in the lower troposphere (~ 1.5 km above ground level). The vertical structure of precipitation prediction algorithm is a Convolution Neural Network trained to learn associations among GPM DPR Ku-band reflectivity profiles, GMI Tbs, and orographic precipitation regimes in the SAM including low level light rainfall, shallow rainfall with low-level enhancement, stratiform rainfall with bright band, and deep heavy rainfall with low- and mid-level enhancement. Vertical structure classes corresponding to the distinct orographic precipitation regimes were isolated through k-means clustering of ground-based Multi-Radar/Multi-Sensor radar reflectivity profiles. The AI framework is demonstrated for automatic retrieval of warm season precipitation in the SAM over a 3-year period (2016-2019) achieving large reductions in false alarms (77%) and missed detections (82%) relative to GPM Ku-PR precipitation products, and significant rain-rate corrections (up to one order of magnitude) by using a physically-based model to capture the microphysics of low-level enhancement (i.e. seeder-feeder interactions).

Keywords: Precipitation Detection, Orographic Precipitation, Convolution Neural Network, Global Precipitation Measurement Mission, Precipitation Radar

1. Introduction

Satellite-based Quantitative Precipitation Estimation (QPE) products are the only information on precipitation intensity and amount available over vast regions of the world where ground-based observing systems are lacking. The accuracy of tropical and subtropical precipitation products improved significantly in the two decades after the launch of the Tropical Rain Measurement Mission (TRMM, Simpson et al. 1996) satellite in 1997 with an instrument payload that included a Ku-Band (13.8 GHz) precipitation radar (PR) for the first time. Extensive error analysis of TRMM-PR precipitation products (Barros et al. 2000; Tian and Peters-Lidard, 2010; Amitai et al. 2009, 2012; Prat and Barros, 2010; Kirstetter et al. 2013; Duan et al. 2015; Maggioni et al. 2016) showed strong dependence of QPE on topography and precipitation regime, with detection errors predominating in light and low-level rainfall, and underestimation errors for heavy precipitation and cold season storm systems (Prat and Barros, 2010; Duan et al. 2015; Wilson and Barros, 2014; Duan and Barros, 2017). Further, Prat and Barros (2010) and Duan et al. (2015) assessed the TRMM-PR QPE estimates against measurements from a long-term spatially dense rain-gauge network in the Southern Appalachian Mountains (SAM, Fig. 1) and found robust multi-year patterns in the spatial and temporal organization of detection and estimation errors at diurnal and seasonal scales conditional on regional precipitation regime.

<Figure 1 here please>

The Global Precipitation Measurement Mission (GPM; Hou et al. 2014) was launched in 2014 as a TRMM follow-on to observe and quantify the three-dimensional structure of precipitation systems on a global scale (approximately between latitudes ±67°). GPM has a dual-frequency precipitation radar (DPR) that operates at Ku- (13.8 GHz; Ku-PR) and Ka-Bands (35.5

GHz; Ka-PR). The Ka-PR was included in the GPM mission to improve the detectability of light precipitation and snow. The minimum detectable precipitation-rate for TRMM-PR retrievals was 0.5 mm/hr and is approximately 0.2 mm/h for GPM Ku-PR (Hou et al. 2014; Speirs et al. 2017). In practice, however, both TRMM-PR and GPM-DPR exhibit higher sensitivity than their nominal design specifications (Hamada and Takayabu, 2016). Overall, comparative error diagnostic studies indicate improved performance in GPM DPR detection and estimation scores relative to TRMM-PR (Liu, 2016; Arulraj and Barros; 2019), albeit retaining similar spatial and temporal organization characteristics. This extends to precipitation products such as the Integrated Multisatellite Retrievals for GPM (IMERG) that combines precipitation estimates from several microwave and infrared sensors calibrated using GPM measurements to produce global maps. Rios Gaona et al. (2016) determined that IMERG underestimates precipitation by only approximately 2% in the smooth topography of the Netherlands, and Khan and Maggioni (2019) reported a detection accuracy of 80% for oceanic rain in IMERG albeit with underestimated intensity. This is in contrast with results from similar studies in mountainous regions. Barros and Arulraj (2020) assessed IMERG against ground-based radar and rain-gauge observations in the SAM, and documented persistent space-time patterns of very low probability of detection (0.3-0.4) consistent with the spatial organization of the diurnal cycle of low-level clouds and fog (see also Wilson and Barros 2014, 2015 and 2017; Duan and Barros, 2017). Speirs et al. (2017) found that GPM-DPR precipitation products underestimate precipitation by more than 50% in the winter season compared to ground-based radar QPE in the Swiss Alps and Plateau. Severe underestimation of precipitation in the SAM was reported by Arulraj and Barros (2019), especially where and when seeder-feeder interactions (SFI) play a governing role on low-level orographic precipitation enhancement. Satellite-based radar QPE in complex terrain remains therefore a critical challenge to satellite-based precipitation monitoring in the GPM era.

Previous studies including Prat and Barros (2010), Duan et al. (2015), Speirs et al. (2017) and Arulraj and Barros (2019) point to three key sources of error associated with TRMM and GPM PR measurements of shallow precipitation systems in complex terrain such as the SAM: (1) non-uniform beam filling (NUBF) artifacts tied to the horizontal resolution of the radar beam; (2) ground-clutter contamination in the near-surface reflectivity; and (3) incorrect microphysical parameterization in the radar-retrieval algorithms. NUBF artifacts are enhanced in mountainous regions due to the three-dimensional (3D) complexity of low-level circulations modulated by the terrain resulting in high spatial variability of clouds and precipitation systems at scales below the radar measurement scale. Ground clutter severely contaminates reflectivity profiles up to 3 km above ground level (AGL) depending on the radar viewing-angle (Arulraj and Barros, 2019). At present, the PR retrieval algorithm extrapolates the reflectivity factor at the lowest no-clutter level to the surface. This yields uniform reflectivity profiles in the clutter affected vertical levels that lead to underestimation (UND) and missed detection (MD) of shallow precipitation (Prat and Barros, 2007; Wilson and Barros, 2014; Porcacchia et al. 2018; Duan and Barros, 2017; Arulraj and Barros, 2019). Concurrent ground-clutter and NUBF artifacts can result in overestimation as well as spurious detection (false alarms, FA) of precipitation thus undermining the reliability of common statistically-based correction of these errors a posteriori. These errors are further compounded by retrieval uncertainty due to high spatial and temporal heterogeneity in the vertical structure of low level orographic precipitation systems that is apparent in the spatial variability of the diurnal and seasonal cycles of the vertical structure of hydrometeor size distributions (Prat and Barros, 2010; Wilson and Barros, 2014; Duan et al. 2015; Barros and Arulraj, 2020).

Arulraj and Barros (2019) demonstrated the effectiveness of a physically-based retrieval approach to improve QPE in complex terrain by applying a dynamic stochastic column model of rainfall microphysics including layered low-level clouds and fog (LLCF) and initial (hydrometeor size distributions) and boundary conditions based on GPM Ku-PR reflectivity above ground-clutter height to simulate observed low-level enhancement in the SAM. They proposed two conceptual models of the vertical structure of LLCF based on extensive observations including concurrent GPM overpasses and ground measurements during the Integrated Precipitation and Hydrology Experiment (IPHEx; Barros et al. 2014) and showed improvements in rain-rate estimates up to one order of magnitude. However, typically such comprehensive data are not available. Critical advances needed in satellite-based orographic QPE including low-level enhancement processes are two-fold: 1) precipitation detection; and 2) vertical structure diagnostics (i.e. layered LLCF configuration).

Previously, Arulraj and Barros (2017) demonstrated the potential for using multi-frequency (Ku, Ka, and W-band) satellite and, or surface-based radar to improve detection and classification of shallow precipitation systems. However, dual-frequency measurements at specific locations are only possible at present where GPM and CloudSat overpasses (EarthCare in the future) are nearly coincident, or where ground-based radars operate. To overcome this limitation, an alternative Artificial Intelligence (AI) framework is proposed here to detect and characterize the vertical structure of orographic precipitation systems leveraging coupled high-resolution Numerical Weather Prediction (NWP) models and GPM measurements, which can be used subsequently to constrain physically-based retrieval of near-surface rainfall rates (Arulraj and Barros, 2019; Fig. 2).

<Figure 2 here please>

Section 2 of this manuscript provides a brief description of the data used in this study including the reference ground-based radar reflectivity observations from Multi-Radar Multi-Sensor System (MRMS), satellite-based GPM DPR and GPM Microwave Imager (GMI) data, NWP model analysis, and the IPHEx rain-gauge network used for ground-validation. Section 3 explains the overall AI framework including precipitation detection and classification algorithms to predict low-level precipitation structure, and error analysis methodology to characterize uncertainty. Algorithm implementation and application results are presented in Sections 4 and 5, respectively. Section 6 provides a brief summary and conclusion.

2. Data Description

2.1 Multi-Radar/ Multi-Sensor System (MRMS)

The MRMS reflectivity profiles are derived from the S-Band (3 GHz) dual-polarization Weather Surveillance Radar- 1988 Doppler (WSR-88D) radars operating as a part of the Next Generation Weather Radar (NEXRAD) network across the contiguous United States (CONUS). S-Band radar reflectivity values are comparable to the Ku-Band radar for reflectivity values below 35 dBZ (Biswas and Chandrasekar, 2018). The data used in this study are the merged, quality-controlled and gridded NEXRAD 3D reflectivity profiles at spatial resolution of 0.01×0.01 degrees (~1×1 km^2), vertical height varying from 0.5 km to 19 km, range resolution between 250 m and 1 km, and temporal resolution of 2 minutes. The time-period of analysis is between November 2016 and May 2019 that is the period for which MRMS reflectivity profiles were available for this project. The quality-control process removes echoes from non-hydrometeors and random clutters due to beam blockage (Zhang et al. 2011). NEXRAD radars operate at Plan Position Indicator (PPI) scanning mode. The schematic of the PPI scanning mode is shown in Fig. 3(a). Due to geographic location of the radars, surrounding topography, and precipitation type, some of the profiles are severely attenuated near the surface. All the reflectivity profiles with reflectivity less than 10 dBZ in the lowest 2 km are removed from processing and analysis.

<Figure 3 here please>

In addition to the 3D reflectivity profiles, Level 2 MRMS precipitation rate and precipitation type products in the nearest 2-minutes of a GPM overpass are used in this study. The resolution of these products is also 0.01×0.01 degrees (~ 1×1 km^2). MRMS precipitation rates are gauge-corrected using 9000 rain-gauges across the CONUS, while the snow events are radar only estimates (Hong and Gourley, 2015). MRMS precipitation types identified based on MRMS reflectivity profiles and collocated operational NWP model temperature profiles (e.g. Rapid Refresh (RAP); see Section 2.3) are as follows: warm stratiform rain, snow, convective, hail, tropical/stratiform mix, tropical/convective mix, and cool stratiform. A detailed description of the precipitation type classification methodology is available in Hong and Gourley (2015).

2.2 Global Precipitation Measurement (GPM) DPR and GMI

The GPM DPR operates at Ku- (13.6 GHz) and Ka- (35.5 GHz) band with a spatial resolution of approximately 5×5 km^2. GPM Ku-PR Level-2A Version 06A data are used in this study. In particular, measured reflectivity profiles (Z_m), corrected reflectivity profiles (Ze), near-surface precipitation rate, no clutter bin height, melting layer height and terrain elevation are used for the analysis. The Ku-PR operates in normal scan (NS) mode with a cross-track swath width of 245 km, a sampling resolution of 125 m, and a range resolution of 250 m (Hou et al., 2014 and Iguchi et al., 2017). The viewing angle varies from nadir (0°) to ±18°.

The GMI is a multi-channel conical scanning microwave radiometer that operates at thirteen microwave channels in the frequencies ranging between 10 GHz and 183 GHz at vertical (V) and horizontal (H) polarization: 10.65 V/H, 18.70 V/H, 23.8 V, 36.64 V/H, 89 V/H, 188 V/H, 183.31±3 V and 183.31±7 V. GMI Level-1C Version 05A calibrated brightness temperatures (Tbs) are used in this study. The mean footprint (pixel) resolution depends on the operating frequencies varying between 6 and 25 km (Draper et al. 2015). To avoid excessive spatial resolution differences among GMI and Ku-PR products, the 183 and 188 GHz channels are not considered for algorithm development. The GPM DPR and GMI data are available from March 2014 to present. Here, data between March 2014 and May 2019 are considered for the analysis.

2.3 High Resolution Rapid Refresh (HRRR) Model

The RAP is a version of the Weather Research and Forecasting model developed by the NOAA Earth System Research Laboratory Global Systems Division. This is an hourly updating, cloud-resolving, convection-allowing model run operationally by the National Centers for Environmental Prediction's Environmental Modeling Center with a nominal horizontal resolution of 13 km (Benjamin et al., 2016) over CONUS. Subsequently, a high-resolution nested version of the RAP, the HRRR, was developed with 3-km horizontal resolution. HRRR is a convection-allowing model and is strongly dependent upon RAP data assimilation including radar reflectivity from the NEXRAD network. Whereas the HRRR model produces 1- to 18-hour lead-time forecasts, only the 0-hour HRRR analysis (data assimilation update) product has vertical profiles of mixing ratios. These products are downloaded from a private cloud object store developed by the Center for High Performance Computing at the University of Utah (Blaylock et al. 2017).

HRRR data are available from July 2016 onwards and the database is updated in real-time. Specifically, hourly instantaneous of surface precipitation rate [mm/hr], rain water mixing ratio [RWMR; kg/kg], snow water mixing ratio [SWMR; kg/kg], graupel mixing ratio [GRLE; kg/kg], specific humidity [SPFH; kg/kg], temperature [TMP; K] and geopotential height [HGT; gpm] at pressure levels between 50 hPa and 1000 hPa with a vertical resolution of 25 hPa were obtained for the SAM from July 2016 to May 2019.

2.4 Rain-gauge Network

The ground-based rain-gauge data used for the evaluation of HRRR model results are obtained from the IPHEx long-term spatially dense tipping bucket rain-gauge network operating in the Great Smoky Mountains National Park within the SAM (Barros et al., 2014; Barros et al. 2017). The instantaneous observations from the rain-gauges are available from June 2007 to present. The network has three different types of tipping bucket rain-gauges operating at various locations of the SAM: Hydrological Services (HS) HS-TB3 model with tipping resolution of 0.2 mm tip−1, HS-TB3/0.1 with tipping resolution of 0.1 mm tip−1, and HS-305 with tipping resolution of 1.0 mm tip−1. HS-TB3 rain-gauges were collocated with HS-305 models at selected locations for quality control purposes due to significant differences in tipping resolution. Duan et al. (2015) provide a detailed description of rain-gauge data processing and quality-control.

3. Methods

3.1 Data Evaluation

The proposed AI machine learning strategy to predict low-level precipitation structure is to implement detection and classification algorithms sequentially (see Fig. 2). The output of the precipitation detection algorithm (PDA) is binary: rain or no-rain. The PDA is driven by GPM and HRRR data. Rainfall is not occurring most of the time during GPM overpasses, and therefore it is critical to distinguish active from non-active precipitation conditions to establish a balanced PDA training database. Precipitation occurrences according to MRMS and rain-gauge measurements are used for training with cross-error characterization to quantify GPM Ku-PR and HRRR detection errors relative to the reference. Statistics of GPM Ku-PR detection discrepancies with respect to the rain-gauges are available from previous work (Arulraj and Barros, 2019; Barros and Arulraj, 2020), and thus error analysis is not repeated here.

The precipitation classification model parses GPM data and predicts the underlying precipitation vertical structure (tied to orographic precipitation regime class) by selecting one among the various classes derived from MRMS climatology through clustering analysis. Thus, MRMS is the reference data set (i.e. "ground-truth"). It is important to highlight that MRMS is limited by the number of radars and radar operations in the region of study that are strongly constrained by topography blocking. Nevertheless, the underlying assumption is that the information content in GPM measurements regarding vertical structures of precipitation in the SAM is generalizable to other geographic regions, even if it is not comprehensive to capture the full breath of orographic precipitation regimes across the world's mountains.

In regions where ground-based measurements are not available, using the vertical distribution of clouds and precipitation from NWP model output can be potentially used instead. Indeed, a HRRR-based classification of shallow precipitation systems was attempted to mitigate and complement MRMS in the inner region and along the outer regions of the SAM where blocking and ground-clutter are artifacts that affect how NEXRAD operates. For this purpose, HRRR climatologies of precipitable water, cloudiness, and precipitation rate were evaluated, and 4-years of hourly reflectivity profiles over the SAM were derived using a radar simulator based on Leinonen et al. (2015). Although HRRR shows good agreement in the timing of precipitation maxima on the eastern ridges of the SAM, which is attributed to the assimilation of NEXRAD observations, it fails to capture the spatial distribution of cloudiness and precipitation maxima over the western ridges (see Figures S1-S4 in Supplementary Data) likewise attributed to biases in NEXRAD data that result from radar operations in regions of complex terrain (e.g. overshooting to mitigate ground-clutter and blocking), and consequently HRRR and MRMS reflectivity profiles are largely incompatible. Therefore, MRMS reflectivity profiles alone are used to characterize vertical precipitation structures representative of predominant orographic precipitation regimes.

3.1.1 Error Analysis: GPM Ku-PR and MRMS

Instantaneous GPM Ku-PR QPE from individual overpasses are evaluated against the nearly (within 2 minutes) coincidental Level-2 MRMS precipitation rates. Note that the spatial resolution of the MRMS products is approximately 1 km while the GPM Ku-PR footprint resolution is approximately 5 km, and thus one GPM pixel corresponds to 25 MRMS pixels as illustrated by the schematic of the footprints in Fig. 3(b). Because of the different foot-print resolutions, the QPE comparison can be performed by either averaging the MRMS to match GPM Ku-PR's resolution and, or by comparing the GPM Ku-PR against the nearest MRMS pixel. In both these approaches, the spatial heterogeneity of the precipitation systems is not taken into consideration, resulting in biased detection metrics such as the probability of detection (POD) and false alarm ratio (FAR; Wilks, 2011). To address this challenge, error analysis was performed by thresholding the fractional area "x%" of MRMS rainy pixels within a GPM pixel (25 MRMS pixels in all): if "x%" of the MRMS pixels observe precipitation greater than 0.1 mm/h then the ground truth is an affirmative "rain" detection, else the default is "no-rain". Based on this approach, the number of correct hits (CD; when both GPM Ku-PR and MRMS detect precipitation), the number of correct misses (NN; when both GPM Ku-PR and MRMS register no rain), the number of false alarms (FA; when GPM Ku-PR detects rain and MRMS does not), and the number of missed detections (MD; when MRMS detects rain and GPM Ku-PR does not) are computed for different values of "x%" ranging from 4% (if 1 pixel out of 25 pixels observes rain, then the ground-truth is rain) to 100% (ground-truth is rain only if all the MRMS pixels observe rain). The frequency bias (FB; Wilks, 2011) for every threshold value "x%" is computed subsequently as follows:

$$FB = \frac{YY + FA}{NN + MD} \tag{1}$$

FB is the ratio of frequency of precipitation detection by the GPM-DPR to the frequency of precipitation detection by MRMS. If FB is greater than 1, then increased FA cases are observed; and if FB is less than unity, MDs dominate. The optimal value of FB is 1 signifying that the number of FA cases are equal to the number of MD cases. Thus, the "x%" value with FB close to 1 is considered the optimal threshold to classify as “rain” the aggregated MRMS pixel corresponding to the GPM pixel.

The standard POD and FAR are the detection error metrics used to evaluate GPM Ku-PR near-surface precipitation estimates. POD is the probability of precipitation detection by GPM Ku-PR given that the MRMS detects precipitation:

$$POD = \frac{YY}{YY + MD} \quad (2)$$

The desired value of POD is 1. FAR is the probability of false alarms given the GPM Ku-PR detects precipitation:

$$FAR = \frac{FA}{YY + FA} \quad (3)$$

The optimal value of FAR is 0. If both GPM and MRMS detect precipitation, the GPM-MRMS discrepancy (estimation error) is the bias (ε):

$$\varepsilon = \log\left(\frac{\sum_i R_{i,GPM}}{\sum_i R_{i,MRMS}}\right) \quad (4)$$

where $R_{i,GPM}$ is the near-surface precipitation rate estimated by GPM and $R_{i,MRMS}$ is average MRMS rain-rate within the GPM radar footprint. The optimal value of ε is 0. Negative values of ε signify underestimation, and positive values signify overestimation.

3.1.2 Error Analysis: Uncertainty in HRRR rainfall

NWP simulations can exhibit 3 to 5-hour delays (phase errors) in predicting the arrival and propagation of certain types of precipitation systems (e.g. Wilson and Barros, 2015 and 2017; Erlingis and Barros, 2014 and many others). Lag-correlations between HRRR analysis and rain-gauge observations were examined to investigate timing errors. For this purpose, the instantaneous HRRR rain-rates at every hour are computed at 500 m AGL by fitting the simulated rain-water mixing ratios to the Marshall-Palmer distribution. The Marshall-Palmer drop size distribution (Marshall and Palmer, 1948) follows a negative exponential distribution of the form:

$$N(D) = N_0 \exp\left(-\Lambda_r D\right) \tag{5}$$

where N_0 is the intercept parameter with fixed value of 8x10^6 m^{-4} while the slope parameter is derived from the rain-water mixing ratios. The slope parameter is computed based on Thompson et al. (2004) as follows:

$$\Lambda_r = \left(\frac{\pi N_0 \rho_r}{\rho_{air} q_r}\right)^{\frac{1}{4}} \tag{6}$$

where N_0 is the intercept parameter, ρ_r is the density of rain (1000 kg m^{-3}), ρ_{air} is the density of air and q_r is the rain-water mixing ratio from the model simulations (kg/kg). The density of air is computed as follows:

$$\rho_{air} = \frac{P}{R_D T_v} \tag{7}$$

P is the pressure in hPa, R_D is the gas constant for air [287 Jkg^{-1}K^{-1}], and T_v is the virtual temperature [K]. The virtual temperature is computed as:

$$T_v = T \times \frac{0.622 + q_v}{0.622 \times (1 + q_v)} \tag{8}$$

Where T is the temperature [K] and q_v is the mixing ratio of water vapor [kg/kg].

Finally, the rain-rate [mm/h; R_{HRRR}] is calculated as shown below:

$$R_{HRRR} = \sum \frac{6\pi}{10^5} N_0 \exp(-\Lambda D) D^3 v(D) \Delta D \tag{9}$$

where v(D) is the fall velocity of drops [m/s] with diameter D [mm] and ΔD is the bin size of drop diameter [mm]. For the inter-comparison, rain-gauge observations [R_{RG}] correspond to the 30-minute accumulation of precipitation centered at the HRRR simulation time stamps.

The Pearson correlation coefficient (r^2) is computed as follows:

$$r^2 = \frac{\widehat{Cov}^2(R_{HRRR}, R_{RG})}{\widehat{Var}(R_{HRRR}) \times \widehat{Var}(R_{RG})} \tag{10}$$

Where *Cov* and *Var* are respectively the covariance and the variance. The value of r^2 varies between 0 and 1, the latter being the perfect score.

3.1.3 Clustering of MRMS Reflectivity Profiles

A primary objective of this study is to predict the precipitation vertical structure in the lower 2 km that is contaminated by ground clutter in GPM-DPR measurements. In order to characterize and classify the vertical structure of precipitation systems in the SAM, the reflectivity profiles from MRMS are organized into precipitation regime classes using a k-means clustering algorithm after Anderberg (1973) in a manner similar to Zhang et al. (2007), who clustered CloudSat Cloud Profiling Radar (CPR) reflectivity profiles and successfully identified 5 cloud regimes in the tropics (low cloud and cirrus, subtropical maritime stratus, anvil cirrus cloud, cumulus congestus, and deep convection).

The k-means clustering algorithm aggregates data points with 'N' different features into 'K' different clusters based on the intra-cluster and inter-cluster distance. The features considered here are the maximum reflectivity (Z_{max}), maximum reflectivity height (H_{max}), reflectivity near the surface (Z_{surf}), echo top height of precipitation systems (H_{top}), and the slope of the reflectivity profiles within 2 km near the surface as illustrated in Fig. 4. The near-surface slope is computed as follows:

$$Slope_{surf} = \frac{Z_{H1} - Z_{H2}}{H_1 - H_2} \tag{11}$$

where Z_{H1} and Z_{H2} are reflectivity values at heights H_1 and H_2 respectively. H_1 is specified at 2 km AGL since the focus here is on shallow precipitation systems, and H_2 is specified at 500 m AGL. The optimal number of clusters is determined according to the Davies Bouldin (DB) index (Davies and Bouldin, 1979) that is calculated based on the ratio of intra-cluster (minimization of variance within each cluster) and inter-cluster (maximization of variance among clusters). The optimum number of clusters is the number corresponding to the lowest DB index indicating a balance between low variance within each cluster and high variance among clusters.

<Figure 4 here please>

Finally, the reflectivity profiles in each cluster are examined in the light of precipitation type and intensity to identify the underlying precipitation regime. Each cluster is further expected to be associated with specific detection and estimation errors tied to their vertical structure. Subsequently, these clusters are used to train the classification framework and therefore to guide the configuration of layered LLCF in the physically-based rainshaft model and thus effectively fill in the near-surface structure of GPM Ku-PR reflectivity contaminated by ground-clutter.

3.2 AI Framework

3.2.1 Precipitation Detection

Whether the optimal number of clusters can be mapped to physically meaningful reflectivity morphologies depends strongly on the quality of the dataset as measured by the ability to identify unambiguous precipitation regimes. In the context of this work, this translates to improved precipitation detection by removing from training the large number of no-rain cases that introduce high frequency bias. For this purpose, GMI Tbs that are concurrent but independent of the DPR measurements and low-level HRRR mass ratios are selected to drive the precipitation detection framework. The GMI Tbs at various frequencies can be useful to discriminate between deep convection (e.g. ice scattering signal at 89 GHz) and the presence of mid-level clouds and rain (e.g. attenuation at 37 GHz). The low-level condensed water mass from HRRR provides information relevant to identify shallow clouds as shallow precipitation.

The precipitation detection model relies on a random forest classifier (RFC; Breiman, 2001) to learn from the data. RFCs are well-suited to handle high-dimensional non-linear classification problems and have been applied with great success in remote-sensing applications such as land-cover classification (Ham et al. 2005; Belgiu and Dragut, 2016; and Kulkarni and Lowe, 2016 among others). RFC input features include data from HRRR and GMI. GPM DPR metadata such as the terrain elevation, the pixel-specific ground-clutter bin height, and the melting layer height are considered as input features also. The GMI input features consist of calibrated multichannel Tbs from 9 channels (10.65 – 89 GHz V/H) within 10 km of the GPM Ku-PR pixel. Input features from HRRR include RWMR, SNWR, and GRLE values and the depth (number of non-zero vertical model layers) of the rain, snow, and graupel with 1.5 km AGL. Because the hourly HRRR analysis is available on–the-hour, the two nearest times (before and after the GPM overpass) are used. This corresponds to time-differences on the order of 30 minutes at least between the overpass and model real-time which can introduce ambiguities especially in the case of fast-evolving convective precipitation systems, though less severe in the case of stratiform and shallow precipitation in the SAM. Nonetheless, higher temporal resolution NWP data would be highly desirable in realistic operational applications.

The RFC is trained using 5-fold cross-validation (i.e. using separately 5 mutually exclusive subsamples and then training and testing the algorithm 5 times using 4 subsamples to train and the remainder for validation) to generate five decision tree ensembles each with 100 members, and thus the total number of decision trees is 500. The maximum depth of 20 (i.e. the number of splits in each decision tree), and the minimum number of features per node is set as 5. When any node reaches 5 samples, further splitting is halted. Further detail regarding the RFC algorithm is provided in Appendix A. The precipitation detection model architecture is shown in Fig. 5. Because there are substantially fewer "rain" than "no-rain" samples, class weights are defined to penalize the misclassification of precipitation events.

<Figure 5 here please>

3.2.2 Precipitation Classification

Artificial Neural Network (NN) algorithms such as back-propagation NNs have proven successful in precipitation classification and estimation over the past three decades (Heermann and Khazenie, 1992; Bruzzone and Serpico, 1997; Kuligowski and Barros, 1998). Deep learning algorithms (e.g. Fig. 6) such as convolutional neural networks (CNN) utilized in supervised classification and regression problems (Li et al. 2014; Kim and Moon, 2016; Maggiori et al. 2017; Li et al. 2017; Faridee et al. 2018; Zhang et al. 2018; Shao et al. 2019) are similar to back-propagation NNs with fully-connected multiple hidden layers for classification, and include upfront data feature extraction capabilities conceptually similar to unsupervised self-organizing maps (LeCun et al. 1998). Feature extraction refers to the systematic processing of the data in the convolution and the pooling layers. A convolution layer consists of a set of convolving filters with specific kernel size to isolate features from the input data. The pooling layer down-samples the number of features identified in the convolution layer to retain only the most important according to a specified criterion. In deep CNNs, feature extraction goes through several stages that are implemented by stacking multiple convolution-pooling layer pairs. Classification proper is carried out by a back-propagation NN. Output from feature extraction is organized into a vector (flattening step) as input to the first fully-connected layer of the classification NN.

<Figure 6 here please>

The precipitation classification model (Fig. 6) relies on a deep CNN with 2-stage feature extraction and 2-layer backpropagation NN for classification to predict the vertical structure of precipitation given GPM Ku-PR and GMI measurements in the absence of ground-truth (observations) by classifying the precipitation regime in the GPM pixel according to MRMS cluster classes. A detailed description of the CNN is given in Appendix B. The GPM Ku-PR input features are the Z_m, the melting layer height, the minimum ground-clutter free height and the local terrain elevation. GPM Ku-PR Z_m profiles are defined by reflectivity values at 176 heights above mean sea level (AMSL). To prepare the data for input, the reference height of Ku-PR Z_m profiles is adjusted to AGL and the reflectivity values are interpolated every 125 m between 0.125 km (near surface) and 15 km (top of the profile) corresponding to 120 equally spaced heights. Besides, Z_m values lower than the minimum detectability of Ku-PR (12 dBZ) were ignored in the analysis. GPM GMI input features are the calibrated Tbs from 9 channels (10.65-89 GHz V/H). For training, the “ground-truth” is the concurrent collocated MRMS cluster class (Section 3.1.3) for each Ku-PR pixel where “rain” is detected within a GPM overpass. All 25 MRMS reflectivity profiles within each Ku-PR footprint are assigned first to a cluster class, and the dominating cluster class (mode) determines the collocated MRMS class used in training.

4. Application

4.1 Data Evaluation

4.1.1 Error Analysis: GPM Ku-PR and MRMS

GPM Ku-PR profiles for each overpass are evaluated against MRMS reflectivity, precipitation rate and precipitation type based on the availability of the MRMS data. This results in the identification of 28005 GPM Ku-PR profiles. Recall that one pixel of GPM Ku-PR corresponds to 25 (5×5) MRMS pixels (Fig. 3b). The contingency matrix (not shown) is calculated first for “homogeneous” cases (in terms of precipitation detection) when all 25 MRMS pixels within the GPM Ku-PR footprint either register precipitation or not. Out of the 28005 profiles, 24919 are identified as “homogeneous” cases with more than 92% for no-precipitation conditions, and the remainder “heterogeneous” cases (3086 profiles) correspond to positive detection of non-uniform precipitation (Section 3.1.1). Table 1 shows the contingency matrix of both homogeneous and heterogeneous cases. The small number of “rain” cases (5.31% are CDs, 1.44% are MDs and 0.79% are FAs) highlights a critical challenge in data-driven modeling of precipitation, that is the need to amass long historical records to assure physical representativeness and statistical robustness.

<Table 1 here please>

Figure 7 shows the frequency bias computed by comparing GPM Ku-PR against MRMS as a function of sub-grid scale precipitation fraction “x%”: FAs dominate MDs up to 60% (15/25 MRMS pixels), and the opposite is true for higher values. Overall, more than 30% of the total precipitation detected by MRMS is missed by GPM Ku-PR, and the number of MDs exceeds the number of FAs. The spatial distributions of POD and FAR exhibit robust spatial structure with isolated low POD patterns and high FAR over the western ridges and in the inner region (Fig. 8) which is attributed in part to observational bias in MRMS due to NEXRAD operations to mitigate ground-clutter artifacts. GPM Ku-PR parallax errors for large viewing angles can result in significant mapping errors and consequently strongly affect POD and FAR statistics (see Supplementary Data, Fig. S5), but this was the case for only one overpass for the period of study.

<Figure 7 here please>

<Figure 8 here please>

The number of samples for bias analysis conditional on MRMS precipitation rate (Fig. 9a) is higher in the 0.5-2 mm/h range, which explains the lower bias of GPM Ku-PR precipitation rates with respect to the mean, maximum and nearest MRMS precipitation rates within the same PR pixel (Fig. 9b). GPM Ku-PR overestimates light precipitation (< 1 mm/h) and underestimates heavy precipitation similar to error metrics of GPM Ku-PR against rain-gauge measurements in Arulraj and Barros (2019). Figures 9(c) and 9(d) show respectively the number of GPM Ku-PR samples and the estimation bias categorized by MRMS precipitation class. Ku-PR underestimation of mean MRMS precipitation rate occurs across all precipitation types except snow, with severe underestimation of relatively rare (very small number of samples) hail, tropical/convective mix and tropical/stratiform mix precipitation regimes. Note, only GPM samples with liquid near-surface precipitation are considered for this analysis. The bias is very small for the most frequent case of stratiform precipitation, suggesting that robust relationships between the two data sets can be found. Spatial sampling (Fig. 9e) and spatial bias patterns (Fig. 9f) should be interpreted in the light of the dominant precipitation regimes. Note the lack of correct detection (white pixels) of precipitation between [35.6 and 35.7 N] and [-83.4 and -83.3W. There is large variability among dominant precipitation regimes in the SAM (see Fig. S6 in Supplementary Data). The eastern region is dominated by frontal and tropical cyclones, while fog and low-level clouds contribute the most to the precipitation observed in the inner region, and mesoscale convective systems predominate in the west. Even though GPM Ku-PR estimates are low relative to MRMS in the inner valleys of the SAM, precipitation along the eastern ridges and in the northernmost regions is overestimated. Nevertheless, a note of caution is warranted here. The NEXRAD QPE estimates contributing to MRMS are strongly affected by beam overshooting in the northern regions distant from regional radars operating at Knoxville, TN and Greenville, SC thus missing precipitation at lower levels, whereas overcorrection of ground-clutter artifacts along mountain ridges to the south result in MRMS underestimation of actual precipitation rates (see also Liao and Barros, 2019).

<Figure 9 here please >

4.1.2 MRMS Precipitation Vertical Structure Classes

A total of 56682 historical MRMS reflectivity profiles concurrent with GPM overpasses are used for clustering analysis. Profiles affected by beam blockage were removed from consideration. The Davis-Bouldin (DB) index computed for different number of clusters in the k-means clustering algorithm (Fig. S7 in Supplementary Data) shows a minimum for k=4 clusters. Because the DB sensitivity is weak, exploratory analysis (not shown) was conducted to assess the impact of choosing 3-7 clusters on independent misclassification of MRMS profiles to confirm best performance for four clusters corresponding to low-level, low-level enhanced, stratiform with bright-band and deep precipitation with mid- and low-level enhancement precipitation regimes. Contoured frequency by altitude diagrams (CFADs) of the reflectivity profiles for each of four clusters are in shown Fig. 10, and Table 2 summarizes the maximum, mean and standard deviation of the MRMS surface precipitation rate for each of the four-clusters.

<Figure 10 here please>

<Table 2 here please>

Cluster-1 reflectivity profiles represent light shallow precipitation (Fig. 10a). Note that the echo top heights for the reflectivity profiles in this cluster are concentrated within the lower 4-6 km AGL and the reflectivity values (S-Band) are within 20 dBZ. The mean precipitation of this cluster is 1.22 mm/h and the maximum is 18.89 mm/h. Cluster-2 captures shallow precipitation with low-level enhancement (Figure 10b) with slightly higher mean and maximum precipitation rate compared to Cluster-1 (Table 2). Cluster-2 reflectivity profiles exhibit an increase in reflectivity values in the near the surface which is similar to the increases in the number and size of raindrops by SFI as shown in Prat and Barros (2007), Wilson and Barros (2014), Duan and Barros (2017), Porcacchia et al. (2018), and Arulraj and Barros (2019). The profiles in Cluster-3 show peak reflectivity at ~ 4 km MSL [2 to 4 km AGL] similar to bright band morphology, and echo-top heights around 6 to 8 km AGL. The mean precipitation rate is 2.20 mm/h and the maximum precipitation rate is 74.74 mm/h. Cluster-3 is representative of the vertical structure of stratiform rainfall with bright-band. Finally, Cluster-4 corresponds to deep precipitation with mid- and low-level enhancement. The echo-top height is approximately around 8 km AGL and can extend up to 10 km, and the near surface reflectivity varies between 25 and 40 dBZ. Cluster-4 mean (9.11mm/h) and maximum (138.75 mm/h) precipitation rates are the highest among the 4 clusters, with the highest standard deviation among cluster members as well.

The synthesis of the spatial organization of higher frequency zones (hot-spots) for each cluster over the SAM topography shown in Supplementary Data Fig. S8 reveals that Cluster-1 is widespread across the ridges of the inner mountain region, whereas Cluster-3 is constrained to the two broad and deeper valleys in the region: the Broad River to the north and the Little Tennessee River to the South, and Cluster-2 and -4 are aligned with the outer western and eastern ridges. The only class without apparent low-level enhancement in the CFAD is Cluster-3. Because the inner valleys are blocked to NEXRAD radars by the outer ridges, low-level SFI processes that can significantly enhance precipitation in these valleys (e.g. Wilson and Barros, 2014 and 2017) are not captured in MRMS, and Cluster-3 morphology reflects these observational biases.

Figure 11 summarizes the statistics of MRMS reflectivity profile features used in the clustering algorithm. Figures 11(a) and (b) show the distribution of maximum reflectivity values and height at which the maximum reflectivity occurs. Cluster-4 shows higher reflectivity values within the lower 2 km AGL on average reflecting low-level enhancement processes. Class-1 shows the lowest maximum reflectivity occurring at ~ 2 km AGL, whereas in Cluster-3 the maximum reflectivity is observed at approximately 3 km AGL that is also close to the melting layer height (in AGL). Figure 11(d) shows the distribution of the slope (Eq. 11) where negative values indicate near-surface enhancement of precipitation. The near surface reflectivity slopes are negative at the 50^{th} percentile except for Cluster-3, and Cluster-2 shows negative slopes also at the 75^{th} percentile. The near-surface reflectivity (Fig. 11e) is representative of rain-rate near-surface. Class-1 shows lower near-surface reflectivity indicating light precipitation followed by Class-3, whereas the impact of low-level enhancement is apparent in Cluster 2 and 4.

<Figure 11 here please>

Table 3 examines the association between GPM Ku-PR precipitation detection errors and the cluster class of concurrent MRMS reflectivity profiles. For each pixel within an overpass of GPM Ku-PR, the most frequent cluster type (mode) within the nearest (5x5) MRMS pixels is considered as the ground-truth. GPM Ku-PR misses approximately 65% of all low-level light precipitation cases (Cluster-1), while only 0.2% of the deep precipitation events (Cluster-4) are missed. MDs in Cluster-2 and Cluster-3 amount to ~18% and ~9% respectively. Underestimation errors dominate by more than 56% in all clusters, with the smaller errors in Class-4 (deep precipitation systems) and larger discrepancies in Class-1 (light low-level precipitation systems). These results are in agreement with previous error analysis studies performed in the SAM (Prat and Barros, 2010; Duan et al. 2015; Arulraj and Barros, 2019; Barros and Arulraj, 2020), and thus provide further support to the four cluster classification to capture the principal precipitation regimes in the SAM even if low-level enhancement of stratiform system with bright-band is missed by Cluster-3. Figure S9 in Supplementary Data contrasts the statistics of GPM Ku-PR and MRMS precipitation estimates when both detect precipitation. Although Ku-PR estimates are lower than MRMS and exhibit higher variance, both products show an increase in mean precipitation with depth of the reflectivity profile except for Cluster-3. Interestingly, the variance is much higher for Ku-PR than MRMS in the case of Cluster-3 due to ground-clutter artifacts affecting Ku-PR measurements in the inner region valleys below the orographic envelope in contrast to NEXRAD radar beam overshooting above the mountain ridges for MRMS.

<Table 3 here please>

4.1.3 Error Analysis - HRRR

The spatial distribution of the number of HRRR pixels with precipitation at 500 m AGL over the 4-year period of interest can be found in Supplementary Data (Fig. S10). The figure shows high frequency of precipitation along the high elevation regions of the Pigeon River Basin. In addition, this spatial pattern agrees with the 10-year climatology map obtained from the rain-gauge merged Stage-IV (GPM GV reference product V1; Liao and Barros, 2019) shown in Supplementary Data, Fig. S11, even if amounts are underestimated especially over the western ridges.

The HRRR analysis was evaluated using rain-gauge observations from different regions in the SAM to investigate timing errors via correlation analysis at different time lags as summarized in Fig. 12. Recall that the HRRR variables are instantaneous and the rain-gauge observations represent 30-minute accumulations centered at HRRR model time stamps. The maximum correlation is observed at 0-lag indicating that HRRR does not exhibit significant timing errors (i.e. delay of precipitation arrival), and thus HRRR analysis captures the diurnal cycle. While this is expected due to the assimilation of NEXRAD data, it is important that no time corrections or adjustments need to be applied to the HRRR data relative to the GPM overpass.

<Figure 12 here please>

HRRR precipitation rates are much lower than the rain-gauge measurements especially for intense precipitation as shown in Supplementary Data, Fig. S12 (for rain-gauge locations see map in Fig. S6). This is also consistent with the previous error analysis conducted in the SAM region comparing the rain-gauge observations with different satellite products (Barros and Arulraj, 2020). Despite ambiguity in the comparison of instantaneous areal estimates (HRRR, 3 km spatial resolution) with time-average point measurements (rain-gauges), an even though HRRR underestimates precipitation by 2-3 mm/h at all times, the overall structure of the diurnal cycle is well captured in particular over the western ridges (Fig. 13). Underestimation of rain-rates in the west is accompanied by underestimation of cloudiness in the model, in particular shallow cap clouds mapped by Duan and Barros (2017) using MODIS data. In the east, the diurnal cycle simulated by HRRR differs from that of rain-gauges in the early morning hours until mid-day, while in the inner region the difference is observed only at mid-day tied to LLCF and SFI among layered clouds that are not described in the model.

<Figure 13 here please>

In summary, the HRRR analysis is a good representative of the climatological and diurnal behavior of the precipitation observed in the SAM even though rain-rates and cloudiness are underestimated. Whereas the vertical structure of water mass in HRRR is not representative of the actual vertical structure of clouds and precipitation in the region, especially at low levels, the fact that the diurnal cycle is captured well suggests that it can be used identify conditions favorable to precipitation and precipitation type. This is the basis for using HRRR RWMR, SWMR, and GRLE in the lower 1.5 km (the depth of Cluster-1 CFAD) as input to the PDA.

4.2 Precipitation Detection

The implementation of the precipitation detection algorithm (PDA) follows the methodology described in Section 3.2.1 and Appendix A (Fig. 5). First, HRRR and GMI input data are normalized between 0 and 1 prior to training the RFC. Using the 5-fold cross-validation to train the RFC yields a classification accuracy of approximately 96%. Training reveals that the GMI Tbs at 89 GHz for both VV and HH polarizations and the RWMR from HRRR are the two most important sources of information, thus suggesting that the input data can be further reduced for operational applications.

Table 4 presents the precipitation detection contingency matrix. MDs are reduced by 82%, and FAs are reduced by 77% compared to the GPM Ku-PR V06A product. This result is a dramatic reduction of detection errors. Further, all the instances when "rain" was detected can be mapped to one of the MRMS precipitation clusters, which demonstrates the robustness of the precipitation detection model in capturing the breath of regional precipitation systems. Figure 14 shows the diurnal cycle of the detection error metrics for GPM Ku-PR and for the PDA application. PDA significantly reduces FAR at all times while POD increases significantly at mid-day and during the night and early morning when low level enhancement processes are important. The slight decrease in POD between 06 and 12 h EDT is due to missed detection of 2 precipitation events.

<Table 4 here please>

<Figure 14 here please>

4.3 Precipitation Classification

The classification model is implemented according to Section 3.2.2 and Appendix B. Training is independent of the precipitation detection model by selecting only GPM data for correct "rain" detections. First, the GPM DPR and GMI data for rainy conditions are divided into training, validation and test subsets with equal representation of data samples in each class. Instead of the original reflectivity profiles, the data were submitted to dimension reduction using two different methods: Principal Component Analysis and Auto-Encoders with similar validation and test accuracy. The input features are then submitted to min-max normalization, and the normalized features are provided as input to the CNN for feature extraction. The number of convolution and pooling layers and CNN hyper-parameters were defined to avoid overfitting. In particular, the feature extraction module in the CNN consists of two convolutional layers with 16 and 8 filters respectively, and a hyperbolic tangent activation function to improve sensitivity and the ability to capture nonlinear relationships. A random dropout rate of 0.25 is enforced after each convolution layer and a maximum pooling layer of size 2 is introduced after each convolution layer. Classification proper is conducted in the second module using a NN (Fig. 6) that consists of two fully-connected layers and a fully-connected output layer with a Softmax activation function. The Adam optimizer (Kingma and Ba, 2014) is used to train the model, and the loss function is computed using categorical cross-entropy. Performance is evaluated using categorical accuracy. The categorical cross-entropy loss function is mathematically defined as,

$$L(y,\hat{y}) = y \times \log(\hat{y}) \tag{12}$$

where y is the actual MRMS Cluster class and $\hat{y}$ is the predicted Cluster class.

The CNN validation accuracy is ~ 70%, which is close to the training accuracy (~ 71%) and indicates that the model is not overfit. There is large epistemic uncertainty implicit in the model because the empirical ground-truth that the MRMS profiles represent suffers from large bias with robust spatial patterns due to the location (range limitations) and operations of NEXRAD radars to mitigate ground-clutter and blocking. Other sources of ambiguity in the assignment of precipitation class stem from the difference in the resolution of MRMS and GPM Ku-PR and NUBF effects that hinder the representativeness of the Ku-PR reflectivity profiles.

Table 5 presents the classification contingency matrix. There is significant improvement in elucidating the vertical structure of precipitation associated with dominant orographic precipitation regimes that are not explicitly captured in the stratiform/convective classification in the Ku-PR V06A product, albeit with leftover ambiguity especially between Class-2 and Class-3, and between Class 3 and Class 4. The latter is attributed to the fact that the correct reflectivity profiles from GPM Ku-PR that are used in training the model are not as deep as the MRM reflectivity profiles, which creates ambiguity unless there is a strong bright-band.

<Table 5 here please>

Higher misclassification rates are observed in the attribution of Cluster-2 and Cluster-3 classes. The spatial distribution of the MRMS cluster classes and the algorithm predictions were examined to determine whether the ambiguity in the MRMS-based cluster classification in the SAM could be interpreted in light of algorithm externalities such as the configuration of the observing system (i.e. NEXRAD) and regional MRMS precipitation climatology, that is to say the systematic handicap in detecting shallow precipitation systems in the inner mountain region as discussed in Section 3.1.3. The maps are available in the Supplementary Data, Figs. S13-S16. Indeed, the data show that most of the misclassified cases are in the inner region where NEXRAD radars have limited view on account of blocking, and thus overshooting effectively reduces the depth of the precipitation system over which reflectivity measurements are obtained.

Classification ambiguity due to the spatial heterogeneity of precipitation (e.g. NUBF) is evaluated at the scale of the GPM Ku-PR pixel with a focus on NUBF using the metric SSH defined as follows,

$$SSH = \frac{Number\ of\ MRMS\ pixels\ with\ predicted\ class\ mode\ X}{Number\ of\ MRMS\ \mathrm{rain}\ pixels} \tag{13}$$

SSH is a linear index of sub-grid scale heterogeneity. Table 6 shows the number of misclassification pixels and number of pixels with corresponding SSH values. Class-2 shows high spatial heterogeneity with approximately 50% instances for SSH < than 0.5, and thus the relatively high misclassification rate of Class-2 is attributed to NUBF effects.

<Table 6 here please>

5. Discussion

The performance of the precipitation detection and classification algorithm is illustrated first for a precipitation event on October 11, 2019 at 06:59 EDT. The precipitation rate and precipitation type according to MRMS are shown in Figs. 15(a-b). This is a stratiform storm system (blue color) over the mountains with a tropical mix stratiform sector to the east and western foothills (red color). GPM Ku-PR overpasses along the eastern region of the SAM capture the tropical mix stratiform system (Fig. 15c). Overall, the spatial structure of the precipitation event is well detected by the GPM, albeit underestimating precipitation especially at higher viewing angles (Fig. S5 in Supplementary Data).

<Figure 15 here please>

Figure 16(a) shows the spatial distribution of precipitation regimes for the same event. Class-0 corresponds to “no-rain”. Classes 1-4 correspond to MRMS Clusters 1-4. Figure 16(b) shows the corresponding map for GPM predicted classes. The model detects precipitation and predicted the precipitation class accurately although some of the Class-2 events along the edge of the overpass are misclassified as Class-3 and some Class-3 events are misclassified as class-4 by the model.

<Figure 16 here please>

Consider the location marked in Fig. 15(c) (black circle) in the section of the GPM overpass where viewing angles are the largest, and thus estimation errors are expected to be large. Indeed, the Ku-PR near-surface precipitation rate estimate is 1.76 mm/h. By contrast, the average MRMS precipitation rate within the 5 km field of view of GPM Ku-PR is 9.1 mm/h with maximum and minimum precipitation-rate within the GPM overpass of 4.86 mm/h and 13.46 mm/h respectively. The MRMS pixel nearest to the center of the Ku-PR pixel registered a precipitation rate of 10.02 mm/h. The classification algorithm predicts shallow precipitation with low level enhancement (Class-2), and therefore this is a case that is suitable for physically-based retrieval using the near-surface LLCF configuration following Arulraj and Barros (2019).

The top boundary (TBC; 2 km AGL) and initial conditions of the rain microphysics model are derived from reflectivity profiles from GPM Ku-PR. The reflectivity profile of the nearest MRMS pixel, and Z_m and Z_e of GPM Ku-PR overpass are shown in Fig. 17(a). The difference in the storm top height of Z_e and MRMS reflectivity profiles is due to the detectability of GPM Ku-PR and the attenuation correction on Z_m. First, negative exponential DSDs are fit to the GPM Ku-PR Z_e at the TBC height and throughout the column to initialize the rainshaft model. Second, the LLCF drop size distribution (DSD) is specified based on the mean of the diurnal climatology derived from a ground-based spectrometer at the same time of day (see Arulraj and Barros, 2019). The model runs for 20 min with the same TBC to reach equilibrium, and the LLCF layer is introduced then for 30 min. Three different LLCF depths (e.g. 300 m, 400 m and 500 m) are used for sensitivity analysis. The surface rain-rate predicted by the physically-based model varies between 12.5 and 15 mm/h (Fig. 17b). The result for the 300m LLCF case (12.5 mm/hr) that is more consistent with climatology in this region of the SAM than the deeper layers that are typical of the inner region (e.g. Duan and Barros, 2017) is slightly higher than the average MRMS estimate. Nevertheless, NEXRAD based precipitation products also tends to underestimate precipitation rate near-surface over the eastern ridges (Liao and Barros, 2019). To quantify uncertainty in the physical-retrieval estimates, 1000 additional simulations were conducted for the 300m LLCF configuration by perturbing the microphysics based on DSD variance statistics from the spectrometer observations (Fig. 17c). The ensemble rain-rates in the 25^{th} to 75^{th} percentile intervals are within the range of uncertainty as described by variance of MRMS estimates within the Ku-PR pixel, that is the Ku-PR subpixel scale spatial variability at the MRMS spatial resolution. This gives rise to the interesting challenge that is to determine the spatial scale, or spatial scale range depending on precipitation regime, beyond which microphysical processes prevail over dynamics to govern the scaling behavior of precipitation.

<Figure 17 here please>

The AI framework (precipitation detection and classification) was also independently applied in fully predictive mode for the (2014-2016) and in the second half of 2019, in which case confirmed "rain" conditions for error analysis are based on rain-gauge observations alone. Recall that MRMS reflectivity profiles that are ground-truth for precipitation type are only available between November 2016 and May 2019, and HRRR simulations are available only after July 2016. Figure 18 shows the distribution of GPM Ku-PR underestimation and overestimation errors for the predicted four precipitation regimes over the 2014-2019 period. Underestimation errors are more frequent in all precipitation regimes (Fig.18a). The full climatology of precipitation rate at the rain-gauges during the same period, thus including both Ku-PR and CD and MD cases, is shown in Fig. 18(b) and in Fig. 18(c) considering only the Ku-PR CD events the climatology of which is presented in Fig. 18(d). Overall, Ku-PR overestimation errors are very small, and the underestimation errors are larger than the overestimation errors in all classes as it can be seen from comparing the skew of the precipitation rate distributions for Ku-PR correct detections Fig. 18(d) and for the same events at the rain-gauges Fig.18(c). On average the Ku-PR CDs underestimate the rain-gauges by ~ 25% for C-4 and ~50% for the other three classes. Whereas a paired inspection of Figs. 18(b) and 18(c) indicates that MDs occur mostly for light rainfall events (classes C-1 and C-2), the distribution for C-3 is more right-skewed when all events are accounted for at the rain-gauges than when only CDs are accounted for. This suggests that the MD events in the case of stratiform rainfall with bright-band are heavy rainfall events suggesting low-level enhancement below the ground-clutter height in the inner mountain region.

<Figure 18 here please>

Finally, the specific underestimation events identified as cases where low-level enhancement could be unambiguously attributed to SFI by Arulraj and Barros (2019) are revisited here in Table 7. A location map can be found in Supplementary Data, Fig. S17. Except for two events (Case 1 and Case 4), all the events are classified as Stratiform type by the GPM algorithm and all fall into Class-4 and Class-3 according to the AI framework, suggesting deep precipitation systems with mid and low level enhancement (C-4) and stratiform precipitation systems with bright-band (C-3). For all of these cases, the physically-based correction algorithm yields precipitation rate estimates close to the rain-gauge and disdrometer observations. However, it is difficult to distinguish between layered SFI (Case 3) and low-level SFI (all others). This suggests that more information on vertical precipitation structure needs to be assessed, specifically longer records of MRMS data with known layered SFI cases. Even if the vertical structure of C-3 according to MRMS does not exhibit low-level enhancement, all C-3 cases are identified by Arulraj and Barros (2019) as low-level enhancement cases, which further supports earlier discussion with regard to observational biases in MRMS that reflect the constraints of NEXRAD operations. Both cases classified as Convective by GPM are placed into C-4 class by the AI framework: the first (Case 1) is an NUBF case, and the second (Case 4) is a case of low-level enhancement with significant improvement relative to Ku-PR (one order of magnitude), but still an underestimation of the rain-gauge measurements. The latter can be addressed by lifting the TBC in the rainshaft model to level of highest Z_e. One possible route to addressing NUBF requires introducing metrics of sub-grid scale heterogeneity such as SSH (Eq.13, Table 6) explicitly into the classification model. Note that similar metrics can be inferred from high-resolution infrared observations from geostationary satellites such as the

Geostationary Operational Environmental Satellites – R series (GOES-R) even when MRMS data are not available.

<Table 7 here please>

6. Conclusion

A data-driven AI framework was developed to improve orographic precipitation detection and classification of low-level precipitation structure by integrating passive (GMI) and active (Ku-PR) GPM measurements and NWP analysis, specifically the HRRR model that is available over CONUS, and it was demonstrated in the Southern Appalachian Mountains with dramatic improvement in detection skill as compares to the GPM products. The new AI framework can be used to identify cases where low level enhancement of precipitation is present, and thus to guide correction of Ku-PR reflectivity profiles contaminated by ground-clutter according to precipitation regime informed by MRMS precipitation structure classes over CONUS. The major findings of this manuscript are as follows:

1. Comparison of GPM Ku-PR precipitation estimates with MRMS precipitation rate confirm that missed detections and false alarms are aligned along the western ridges of the SAM, which is the region where precipitation is higher and orographic enhancement effects stronger. In addition, GPM Ku-PR predominantly underestimates precipitation rate in this region for most of the precipitation types.
2. NWP analysis (e.g. HRRR) profiles of condensed water mixing ratios in the lower troposphere provide reliable and useful information regarding the likelihood of precipitation activity that can be used in precipitation detection.

3. Clustering analysis of reflectivity profiles from ground-based radars (e.g. MRMS) allowed identification of four classes of vertical structure of precipitation that correspond to distinctive orographic precipitation regimes with well-defined spatial patterns linked to topography, and strong enhancement of near surface precipitation rates due to seeder-feeder interactions among low and mid-level layered clouds: Cluster-1 represents shallow light precipitation; Cluster-2 captures shallow precipitation with low-level enhancement; Cluster-3 consists of stratiform precipitation with bright-band; and Cluster-4 captures deep precipitation systems with mid and low-level enhancement. Error diagnostics indicate that GPM Ku-PR missed detection errors are tightly associated with shallow light precipitation (Cluster-1) while the highest number of correct detection cases is for deep precipitation systems (Cluster-4).
4. A precipitation detection model was developed using a random forest classifier. The inputs of the precipitation detection algorithm include GPM GMI multichannel brightness temperatures, DPR Ku-band reflectivity profiles, and HRRR water mixing ratios. The most important features used in the random forest classifier to accurately detect precipitation are the calibrated brightness temperatures at 89 GHz both vertical and horizontal polarizations, followed by the average HRRR rainwater mixing ratio in the lowest 1.5 km AGL, and the depth of low level clouds as measured by the number of HRRR model layers with non-zero rainwater in the nearest 1.5 km AGL.
5. The precipitation detection model significantly improves the probability of detection and reduces the number of false alarms in GPM KU-PR retrievals. The number of

missed detections is reduced by 82% while the number of false alarms is reduced by 77% over a three-year period (2016-2019).

6. A convolution neural network algorithm (CNN) was implemented to identify the vertical structure of detected precipitation. GPM observations from DPR and GMI are used to train a supervised classification algorithm that maps GPM observations to different MRMS precipitation classes. The algorithm is general and can be applied in regions with precipitation climatology similar to the SAM in the absence of ground-observations.

The AI framework is composed of sequential detection and classification models. The data used in the detection model are not specific to the Southern Appalachian Mountain and are generally available, thus the detection model can be applied globally. The classification model is trained to identify orographic precipitation regimes with low-level enhancement, and thus it is applicable in regions with similar climatology and precipitation physics, which can be identified by clustering TRMM and, or GPM observations. To identify additional precipitation regimes, an adaptive model would be necessary to identify precipitation classes in different climatic regions with more input from NWP to compensate for the lack of ground-based radars and support from ground-based measurements.

CRediT author statement

M. Arulraj: Methodology, Data curation, Writing- First draft preparation, Formal analysis, Software, Investigation. A. P. Barros: Conceptualization, Methodology, Analysis, Investigation, Writing- Reviewing and Editing, Supervision, Project administration, Funding acquisition.

Acknowledgments

Access to the MRMS data was made possible by the GPM Ground Validation Program. The authors are grateful to Dr. Robert (Bob) Meneghini, Dr. Walter Petersen, and Dr. Simone Tanelli for their advice, comments and suggestions regarding this work. The research was funded in part by NASA Earth System Science Fellowship Grant NNX16AO10H with the first author, and by NASA Grant 80NSSC19K0685 with the second author.

References

Amitai, E., C. L. Unkrich, D. C. Goodrich, E. Habib, and B. Thill. 2012. "Assessing Satellite-Based Rainfall Estimates in Semi-arid Watersheds Using the USDA-ARS Walnut Gulch Gauge Network and TRMM PR." *J. Hydrometeorol.* 13: 1579-1588. doi:10.1175/jhm-d-12-016.1,.

Amitai, E., X. Llort, and D. Sempere-Torres. 2009. "Comparison of TRMM Radar Rainfall Estimates with NOAA Next-Generation QPE." *J. Meteorol. Soc. Jpn.* 87A: 109-118. doi:10.2151/jmsj.87A.109.

Anderberg. 1973. *Cluster Analysis for Applications.* New York: Academic Press. doi:10.1016/C2013-0-06161-0.

Arulraj, M., and A. P. Barros. 2019. "Improving quantitative precipitation estimates in mountainous regions by modelling low-level seeder-feeder interactions constrained by Global Precipitation Measurement Dual-frequency Precipitation Radar measurements." *Remote Sens. Environ.* 231: 111213. doi:10.1016/j.rse.2019.111213.

Arulraj, M., and A. P. Barros. 2017. "Shallow Precipitation detection and classification using multifrequency radar observations and model simulations." *J. Atmos. Oceanic Technol.* 34: 1963-1983. doi:10.1175/JTECH-D-17-0060.1.

Barros, A. P., and M. Arulraj. 2020. "Remote sensing of orographic precipitation." In *Satellite Precipitation Measurement*, edited by V. Levizzani, C. Kidd, D. Kirschbaum, C. Kummerow, K. Nakamura and F. J. Turk, 483. Springer International Publishing. doi:10.1007/978-3-030-24568-9.

Barros, A. P., M. Joshi, J. Putkonen, and D. W. Burbank. 2000. "A study of the 1999 monsoon rainfall in a mountainous region in central Nepal using TRMM products and rain gauge observations." *Geophys. Res. Lett.* 27: 3683-3686. doi:10.1029/2000gl011827.

Barros, A. P., W. Petersen, M. Schwaller, R. Cifelli, K. Mahoney, C. Peters-Liddard, and D. Starr. 2014. *NASA GPM-Ground Validation: Integrated Precipitation and Hydrology Experiment 2014 Science Plan.* Tech. Rep., Durham, U.S.A.: Duke University.

Barros, A.P., D. Miller, A.M. Wilson, G. Cutrell, M. Arulraj, P. Super, and W.A. Petersen. 2017. "IPHEx-Southern Appalachian Mountains -- Rainfall Data 2008-2014." Duke Digital Repository. doi:https://doi.org/10.7924/G8CJ8BJK.

Belgiu, M., and L. Dragut. 2016. "Random forest in remote sensing: A review of applications and future directions." *ISPRS Journal of Photogrammetry and Remote Sensing* 114: 24-31. doi:10.1016/j.isprsjprs.2016.01.011.

Benjamin, S. G., and Co-authors. 2016. "A North American Hourly Assimilation and Model Forecast Cycle: The Rapid Refresh." *Mon. Wea. Rev.* 144: 1669-1694. doi:10.1175/MWR-D-15-0242.1.

Biswas, S. K., and V. Chandrasekar. 2018. "Cross-Validation of Observations between the GPM Dual-Frequency Precipitation Radar and Ground Based Dual-Polarization Radars." *Remote Sens.* 10 (11): 1173. doi:10.3390/rs10111773.

Blaylock, B. K., J. D. Horel, and S. T. Liston. 2017. "Cloud archiving and data mining of High-Resolution Rapid Refresh forecast model output." *Computers & Geosciences* 109: 43-50. doi:10.1016/j.cageo.2017.08.005.

Breiman. 2001. "Random Forests." *Machine Learning* 45 (1): 5-32. doi:10.1023/A:1010933404324.

Breiman, L., L. Friedman, R. Olshen, and C. Stone. 1984. *Classification and regression trees.* Belmont, California: Wadsworth Inc.

Bruzzone, L., and S. B. Serpico. 1997. "An iterative technique for the detection of land-cover transitions in multitemporal remote-sensing images." *IEEE Transactions on Geoscience and Remote Sensing* 35 (4). doi:10.1109/36.602528.

Davies, D. L., and D. W. Bouldin. 1979. "A Cluster Separation Measure." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 2: 224-227. doi:10.1109/TPAMI.1979.4766909.

Draper, D. W., D. A. Newell, F. J. Wentz, S. Krimchansky, and G. M. Skofronick-Jackson. 2015. "The Global Precipitation Measurement (GPM) Microwave Imager (GMI): Instrument Overview and Early On-Orbit Performance." *IEEE J. Sel. Top. Geosci. Remote Sens.* 8 (7): 3452-3462. doi:10.1109/JSTARS.2015.2403303.

Duan, Y., A. M. Wilson, and A. P. Barros. 2015. "Scoping a field experiment: error diagnostics of TRMM precipitation radar estimates in complex terrain as a basis for IPHEx2014." *Hydrol. Earth Syst. Sci.* 19: 1501-1520. doi:10.5194/hess-19-1501-2015.

Duan, Y., and A. P. Barros. 2017. "Understanding How Low-Level Clouds and Fog Modify the Diurnal Cycle of Orographic Precipitation Using In Situ and Satellite Observations." *Remote Sens.* 9: 920. doi:10.3390/rs9090920.

Erlingis, J. M., and A. P. Barros. 2014. "A Study of the Role of Daytime Land–Atmosphere Interactions on Nocturnal Convective Activity in the Southern Great Plains during CLASIC." *J. Hydrometeor.* 15: 1932-1953. doi:10.1175/JHM-D-14-0016.1.

Faridee, A. Z. M., S. R. Ramamurthy, H. M. S. Hossain, and N. Roy. 2018. "HappyFeet: Recognizing and Assessing Dance on the Floor." *19th International Workshop on Mobile*

Computing Systems & Applications. Tempe, Arizona: ACM. 49-54. doi:10.1145/3177102.3177116.

Ham, J., Y. Chen, M. M. Crawford, and J. Ghosh. 2005. "Investigation of the random forest framework for classification of hyperspectral data." *IEEE Transactions on Geoscience and Remote Sensing* 43 (3): 492-501. doi:10.1109/TGRS.2004.842481.

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. Ocean. Technol.* 33: 653-667. doi:10.1175/JTECH-D-15-0097.1.

Heermann, P. D., and N. Khazenie. 1992. "Classification of multispectral remote sensing data using a back-propagation neural network." *IEEE Transactions on Geoscience and Remote Sensing* 30 (1): 81-88. doi:10.1109/36.124218.

Hong, Y., and J. J. Gourley. 2015. *Radar hydrology: principles, models, and applications.* Boca raton: CRC Press. doi:10.1201/b17921.

Hou, A. Y., R. K. Kakar, N. Steven, A. A. Azarbarzin, C. D. Kummerow, K. Kojima, R. Oki, K. Nakamura, and T. Iguchi. 2014. "The Global Precipitation Measurement Mission." *Bull. Amer. Meteor. Soc.* 95: 701-722. doi:10.1175/BAMS-D-13-00164.1.

Iguchi, T., S. Seto, R. Meneghini, N. Yoshida, J. Awaka, M. Le, V. Chandrasekar, and T. Kubota. 2017. *GPM/DPR Level-2 algorithm theoretical basis document.* NASA/JAXA Tech. Rep., 81 pp. https://pmm.nasa.gov/sites/default/files/document_files/ATBD_DPR_201708_whole_1.pdf.

Khan, S., and V. Maggioni. 2019. "Assessment of Level-3 Gridded Global Precipitation Mission (GPM) Products Over Oceans." *Remote Sens.* 11 (3): 255. doi:10.3390/rs11030255.

Kim, Y., and T. Moon. 2016. "Human Detection and Activity Classification Based on Micro-Doppler Signatures Using Deep Convolutional Neural Networks." *IEEE Geoscience and Remote Sensing Letters* 13 (1): 8-12. doi:10.1109/LGRS.2015.2491329.

Kingma, D. P., and J. Ba. 2014. "Adam: A Method for Stochastic Optimization." *arXiv preprint* 1412.6980.

Kirstetter, P.-E., Y. Hong, J. J. Gourley, M. Schwaller, W. Petersen, and J. Zhang. 2013. "Comparison of TRMM 2A25 Products, Version 6 and Version 7, with NOAA/NSSL Ground Radar-Based National Mosaic QPE." *J. Hydrometeorol.* 14: 661–669. doi:10.1175/jhm-d-12-030.1, 2013.

Kohai, R. 1995. "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection." *International Joint Conference on Artificial Intelligence.* Montreal, Quebec, Canada. 1137-1143. doi:10.5555/1643031.1643047.

Kuligowski, R. J., and A. P. Barros. 1998. "Using Artificial Neural Networks to estimate missing rainfall data." *J. Am. Water Resour. Assoc.* 34: 1437-1447. doi:10.1111/j.1752-1688.1998.tb05443.x.

Kulkarni, A. D., and B. Lowe. 2016. "Random Forest Algorithm for Land Cover Classification." *International Journal on Recent and Innovation Trends in Computing and Communication (IJRITCC)* 58-63.

Lecun, Y., L. Bottou, Y. Bengio, and P. Haffner. 1998. "Gradient-based learning applied to document recognition." *IEEE.* 2278-2324. doi:10.1109/5.726791.

Leinonen, J., M. D. Lebsock, S. Tanelli, K. Suzuki, H. Yashiro, and Y. Miyamoto. 2015. "Performance assessment of a triple-frequency spaceborne cloud–precipitation radar concept using a global cloud-resolving model." *Atmos. Meas. Tech.* 8: 3493–3517. doi:10.5194/amt-8-3493-2015.

Li, Q., W. Cai, X. Wang, Y. Zhou, D. D. Feng, and M. Chen. 2014. "Medical Image Classification with Convolutional Neural Network." *13th International Conference on Control, Automation, Robotics & Vision.* Marina Bay Sands, Singapore: IEEE. 844-848.

Li, Y., H. Zhang, and Q. Shen. 2017. "Spectral–Spatial Classification of Hyperspectral Imagery with 3D Convolutional Neural Network." *Remote Sens.* 9 (1): 67. doi:10.3390/rs9010067.

Liao, M., and A. P. Barros. 2019. *GPM Ground Validation Reference Precipitation IPHEx.* Huntsville, Alabama, U.S.A.: NASA Global Hydrology Resource Center DAAC. doi:10.5067/GPMGV/IPHEX/GAUGES/DATA401.

Liu, Z. 2016. "Comparison of Integrated Multisatellite Retrievals for GPM (IMERG) and TRMM Multisatellite Precipitation Analysis (TMPA) Monthly Precipitation Products: Initial Results." *J. Hydrometeor.* 17: 777-790. doi:10.1175/JHM-D-15-0068.1.

Maggioni, V., P. C. Meyers, and M. D. Robinson. 2016. "A Review of Merged High-Resolution Satellite Precipitation Product Accuracy during the Tropical Rainfall Measuring Mission Era." *J. Hydrometeor.* 17: 1101-1117. doi:10.1175/JHM-D-15-0190.1.

Maggiori, E., and Y. Tarabalka. 2017. "Convolutional Neural Networks for Large-Scale Remote-Sensing Image Classification." *IEEE Transactions on Geoscience and Remote Sensing* 55 (2): 645-657. doi:10.1109/TGRS.2016.2612821.

Marshall, J. S., and W. McK. Palmer. 1948. "The Distrinution of Raindrops with Size." *J. Meteor.* 5: 165-166. doi:10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2.

Porcacchia, L., P-E. Kirstetter, V. Maggioni, and S. Tanelli. 2019. "Investigating the GPM Dual-frequency Precipitation Radar signatures of low-level precipitation enhancement." *Q J R Meteorol. Soc.* 1-15. doi:10.1002/qj.3611.

Prat, O. P., and A. P. Barros. 2007. "A Robust Numerical Solution of the Stochastic Collection–Breakup Equation for Warm Rain." *J. Appl. Meteor. Climatol.* 46: 1480-1497. doi:10.1175/JAM2544.1.

Prat, O. P., and A. P. Barros. 2010. "Assessing satellite-based precipitation estimates in the Southern Appalachian mountains using rain gauges and TRMM PR." *Adv. Geosci.* 25: 143-153. doi:10.5194/adgeo-25-143-2010.

Qian, N. 1999. "On the momentum term in gradient descent learning algorithms." *Neural Networks* 12 (1): 145-151. doi:10.1016/S0893-6080(98)00116-6.

Rios Gaona, M. F., A. Overeem, H. Leijnse, and R. Uijlenhoet. 2016. "First-Year Evaluation of GPM Rainfall over the Netherlands: IMERG Day 1 Final Run (V03D)." *J. Hydrometeor.* 17: 2799-2814. doi:10.1175/JHM-D-16-0087.1.

Safavian, S. R., and D. Landgrebe. 1991. "A survey of decision tree classifier methodology." *IEEE Transactions on Systems, Man, and Cybernetics* 21 (3): 660-674. doi:10.1109/21.97458.

Shao, Z., Y. Pan, C. Diao, and J. Cai. 2019. "Cloud Detection in Remote Sensing Images Based on Multiscale Features-Convolutional Neural Network." *IEEE Transactions on Geoscience and Remote Sensing* 57 (6): 4062-4076. doi:10.1109/TGRS.2018.2889677.

Simpson, J., C. Kummerow, W.-K. Tao, and R. F. Adler. 1996. "On the Tropical Rainfall Measuring Mission (TRMM)." *Meteorol. Atmos. Phys.* 60 (1-3): 19-36. doi:10.1007/BF01029783.

Speirs, P., M. Gabella, and A. Berne. 2017. "A Comparison between the GPM Dual-Frequency Precipitation Radar and Ground-Based Radar Precipitation Rate Estimates in the Swiss Alps and Plateau." *J. Hydrometeor.* 18: 1247-1268. doi:10.1175/JHM-D-16-0085.1.

Srivastava, N., G, Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. 2014. "Dropout: A Simple Way to Prevent Neural Networks from Overfitting." *J. Mach. Learn. Res.* 15: 1929-1958.

Thompson, G., R. M. Rasmussen, and K. Manning. 2004. "Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis." *Mon. Wea. Rev.* 132: 519-542. doi:10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2.

Tian, Y., and C. D. Peters-Lidard. 2010. "A global map of uncertainties in satellite−based precipitation measurements." *Geophys. Res. Lett.* 37: L24407. doi:10.1029/2010GL046008.

Wilks, D. 2011. *Statistical Methods in the Atmospheric Sciences.* (3rd ed.). Vol. 100. Academic Press. doi:10.1016/B978-0-12-385022-5.00008-7.

Wilson, A. M., and A. P. Barros. 2014. "An Investigation of Warm Rainfall Microphysics in the Southern Appalachians: Orographic Enhancement via Low-Level Seeder–Feeder Interactions." *J. Atmos. Sci.* 71: 1783-1805. doi:10.1175/JAS-D-13-0228.1.

Wilson, A. M., and A. P. Barros. 2015. "Landform controls on low level moisture convergence and the diurnal cycle of warm season orographic rainfall in the southern appalachians." *J. Hydrol.* 531: 475-493. doi:10.1016/j.jhydrol.2015.10.068.

Wilson, A. M., and A. P. Barros. 2017. "Orographic Land–Atmosphere Interactions and the Diurnal Cycle of Low-Level Clouds and Fog." *J. Hydrometeor.* 18: 1513-1533. doi:10.1175/JHM-D-16-0186.1.

Zhang, J., and Co-authors. 2011. "MULTI-RADAR MULTI-SENSOR (MRMS) Quantitative Precipitation Estimation." *Bull. Amer. Meteor. Soc.* 97: 621-638. doi:10.1175/BAMS-D-14-00174.1.

Zhang, J., P. Liu, F. Zhang, and Q. Song. 2018. "CloudNet: Ground-Based Cloud ClassificationWith Deep Convolutional Neural Network." *Geophysical Research Letters* 45: 8665-8672. doi:10.1029/2018GL077787.

Zhang, Y., S. Klein, G. G. Mace, and J. Boyle. 2007. "Cluster analysis of tropical clouds using CloudSat data." *Geophys. Res. Lett.* 34: L12813. doi:10.1029/2007GL029336.

Appendix A

Random Forest Classifier **-** A random forest classifier (RFC) is a supervised classification method based on decision trees that relies on ensemble statistics to predict individual classes. An individual decision tree is sensitive to the input data and a complex tree structure can easily tend to overfit the test data (Safavian and Landgrebe, 1991). To address these issues, a random forest classifier was proposed by Breiman (2001) that consists of an ensemble of multiple decision trees. The diversity or variability within the trees in a forest can be increased by a two-stage randomization procedure as follows: 1) bootstrap the training data, and 2) randomize the features. In the first stage of randomization, the data samples for each tree are chosen randomly with replacement (bootstrapping). Next, random subsets are extracted for each tree separately to achieve low correlation among the trees in the forest. The goal is to generalize the model to avoid overfitting.

Every "parent" node of the decision tree is split into individual "offspring" nodes according to impurity measures such as the Gini index (Breiman et al. 1984). The Gini index at node 't' is defined as follows:

$$G(t)=\sum_{i\neq j} p(i/t)\,p(j/t) \qquad \text{(A-1)}$$

Where p(i/t) is the probability that the random variable belongs to class 'i' given at node 't'. The split at the parent node is performed for the minimal change in the impurity measure between the "parent" and the "offspring" nodes. These impurity measures implicitly act as a feature selection method and provide the most important features in the training data for the prediction of classes.

Finally, the output class of a random forest is predicted by applying the most frequent criterion to the pool of individual predictions among all the trees. Since the output is based on a collective decision of all the trees, the variance of the final output decreases by providing a better prediction. The RFC algorithm's ability to highlight important features is key to extract the physical interpretability of the model. Performance is evaluated based on the accuracy of prediction on data that are not used in the training. A k-fold cross-validation (Kohavi 1995) approach is used to further generalize and reduce the bias of the model. In 'k'-fold cross-validation, the data are divided into 'k' mutually exclusive subsets. The model is trained and tested k times where 'k'-1 subsets are used as train data and the remaining subset is the test data used to determine the accuracy of the method. The schematic flowchart of a random forest classifier is shown in Supplementary Data (Fig. S18).

Appendix B

***Convolutional Neural Network* -** Assume '**X**' is the input data with 'N' samples and 'm' features ($\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N]$; $\mathbf{x}_1=[x_{11}, x_{12}, \ldots, x_{1m}]$) and '**Y**' is the final output vector with class labels [Number of classes – 'c']. '**X**' is passed as input to the 1st convolutional layer consisting of 'K' filters with kernel size 'k_1'. The output from a filter is

$$z_i = b_i + \sum_{j=1}^{m} Convolution\left(w_{i,j} * x_j\right) \qquad \text{(B-1)}$$

Where $\mathbf{w}_{i,j}$ is the weight vector between i^{th} and j^{th} feature, $\mathbf{b}_i$ is the bias of the i^{th} feature and $\mathbf{z}_i$ is the output of i^{th} filter. Further, $\mathbf{z}_i$ is transformed to $\mathbf{s}_i$ based on the activation function 'f'. At the end of the 1st convolutional layer with 'K' filter, the output will have the dimension of '$(N-k_1+1) \times K$'. Next, pooling is applied for the output from the convolutional layer ($s = [s_1, s_2, \ldots, s_K]$)]. The pooling layer will down-sample the output of the convolution layer by choosing the maximum or the average within the kernel size to extract local features. For a pooling of length ‘P’, the pooling layer output will have a dimension of $\frac{N-k_1+1}{P} \times K$.

The Dropout regularization technique (Srivastava et al. 2014) by which randomly selected features (neurons) are temporarily removed during training and not passed to the next layer, and the weights of the dropped neurons are not updated in the backward pass is applied to avoid overfitting. Finally, the outputs from the last layer are stacked in a 1-D vector (extracted features; **F**) in the flatten layer and sent to the fully connected layer (FC) or multi-layer perceptron layer for the classification. The output at the FC is computed as follows:

$$\hat{y}_{f1}=f\,¿ \tag{B-2}$$

Where $\mathbf{w}_{1,i}$ is the weight of the i^{th} neuron, **F** are the features extracted, b_{f1} is the bias, c_1 is the number of neurons and f is the activation function of the FC layer. The final class prediction will be performed by the FC output layer where the input will be from the previous FC layer (or from the flatten layer in the absence of multiple FC).

In this study, the hyperbolic tangent (tanh) function is used as the activation function for all the convolutional and FC layers except the FC output layer. The tanh is a monotonic function which is similar to the logistic sigmoid with range between -1 and 1. Here, the negative inputs are mapped to strong negative values while the zeros inputs are mapped close to zero. In the FC output layer, a Softmax function is chosen as the activation function since it is a generalized logistics function used for a multiclass classification. In the pooling layer, maximum pooling of size 2 is preferred. Maximum pooling considers the largest element within the kernel size.

The weights of the filters are calculated iteratively by minimizing the loss function using an optimizer. Some of the common optimizers are the gradient descent and the Root Mean Square prop (RMSprop) optimizer. In a gradient descent optimizer, at each iteration, a gradient of the loss function is computed to update the weights and biases to get the global minima. The gradient descent with momentum (Qian 1999) has faster converging rate that the traditional gradient descent algorithm since the exponential weighted averages are computed from the gradients to update the weights. The weights in a gradient descent with momentum are computed as follows:

$$v_{dW} = \beta v_{dW} + (1-\beta) dW \qquad \text{(B-3a)}$$

$$v_{dW} = \beta v_{dW} + (1-\beta) dW \qquad \text{(B-3b)}$$

$$W = W - \alpha v_{dW} \qquad \text{(B-3c)}$$

$$B = B - \alpha v_{dB} \qquad \text{(B-3d)}$$

Where W is the weight vector, B is the bias, v_{dW} is gradient update of the weights at t^{th} iteration, v_{dB} is gradient update of the bias at t^{th} iteration, α is the learning rate and β is the momentum parameter.

The RMSprop optimizer are similar to that of the gradient descent with momentum but it normalizes the gradient using moving average. The weights in RMSprop are calculated as follows:

$$s_{dW} = \beta s_{dW} + (1-\beta) dW^2 \qquad \text{(B-4a)}$$

$$s_{dW} = \beta s_{dW} + (1-\beta) dW^2 \qquad \text{(B-4b)}$$

$$W = W - \frac{\alpha \, dW}{\sqrt{s_{dW}}} \tag{B-4c}$$

$$B = B - \frac{\alpha \, dB}{\sqrt{s_{dB}}} \tag{B-4d}$$

The Adaptive Momentum Estimation optimizer (ADAM; Kingma and Ba 2014) is the combination of gradient descent with momentum and RMSprop optimizer where a decreasing learning rate is adapted as the global minima approaches.

List of Tables

Table 1: Contingency matrix comparing GPM Ku-PR and MRMS rainfall occurrences. The values in the parenthesis are the corresponding percentages.

	MRMS = 0	MRMS ≠0
GPM KU-PR = 0	24438 (86.9%) NN	824 (3.3%) MD
GPM KU-PR ≠ 0	779 (2.8%) FA	1964 (7.0%) CD

Table 2: Mean, maximum and standard deviation of the MRMS precipitation rates corresponding to each of the 4 vertical structure classes after k-means clustering of MRMS reflectivity profiles.

	MAXIMUM PRECIPITATION [mm/hr]	MEAN PRECIPITATION [mm/hr]	STANDARD DEVIATION [mm/h]
CLUSTER 1	18.89	1.22	0.98
CLUSTER 2	48.76	2.45	2.14
CLUSTER 3	74.74	2.20	2.77
CLUSTER 4	138.75	9.11	10.21

Table 3: Distribution of GPM Ku-PR precipitation errors relative to MRMS.

	CLUSTER 1	CLUSTER 2	CLUSTER 3	CLUSTER 4
MISSED DETECTION GPM = 0 AND MRMS ≠ 0	649 (65.1 %)	112 (17.7%)	62 (9.3%)	1 (0.2%)
CORRECT DETECTION GPM ≠ 0 AND MRMS ≠ 0	348 (34.9%)	522 (82.3%)	604 (90.7%)	490 (99.8%)
UNDERESTIMATION GPM < MRMS	245 (70% OF CD)	331 (63% OF CD)	381 (63% OF CD)	275 (56% OF CD)
OVERESTIMATION GPM > MRMS	103 (30% OF CD)	191 (37% OF CD)	223 (37% OF CD)	215 (44% OF CD)

Table 4: Contingency matrix of predicted GPM detection and MRMS rainfall occurrences. The values in the parenthesis are the corresponding percentages.

	MRMS = 0	MRMS ≠0
GPM PREDICTION = 0	25040 (89.4%)	148 (0.5%)
GPM PREDICTION ≠ 0	177 (0.6%)	2640 (9.5%)

Table 5: Contingency matrix of the predicted GPM classification against MRMS.

MRMS	GPM PREDICTED			
	CLUSTER-1	CLUSTER-2	CLUSTER-3	CLUSTER-4
CLUSTER-1	**874**	61	61	1
CLUSTER-2	151	**359**	106	18
CLUSTER-3	41	80	**466**	79
CLUSTER-4	2	13	71	**405**

Table 6: Distribution of predicted classification errors conditional on subgrid scale heterogeneity as measured by SSH (Eq.13) at the GPM Ku-PR pixel scale.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Number of Misclassifications	123	275	200	86
SSH < 0.5	80 (65.0%)	138 (50.2%)	117 (58.5%)	52 (60.5%)
0.5 ≤ SSH < 0.75	24 (19.5%)	60 (21.8%)	50 (25.0%)	22 (25.6%)
SSH ≥ 0.75	19 (15.5%)	77 (28.0%)	33 (16.5%)	12 (13.9%)

Table 7: Vertical structure classes (MRMS cluster index) for GPM Ku-PR underestimation cases when rain-gauge (RG) measurements are available. AB19 - QPE estimates using the rainshaft microphysics model reported by Arulraj and Barros (2019). Large disagreement among AB19 estimates and rain-gauge measurements for Case 1, Class 4 is highlighted in italics and superscript *. This case is associated with high NUBF by AB19.

Case ID	**Date – Time EDT**						
		GPM Ku-PR	**RG**	**AB19 Without LLCF**	**AB19 With LLCF**	**Predicted GPM Class**	**GPM Precipitation Type**
1*	May 31, 2016 19:20	*5.9*	*13.9*	*27.9*	*31.0–33.0*	*4**	*Convective*
2	May 21, 2017 11:30	2.2	9.6	8.4	11.3 – 13.4	3	*Stratiform*
3	June 17, 2018 16:49	0.5	6.4	0.8	4.5 – 7.2	4	*Stratiform*
4	August 11, 2014 19:52	2.2	20.0	9.1	12.7 – 15.2	4	*Convective*
5	August 17, 2016 20:27	2.0	8.3	5.9	10.4 – 13.5	3	*Stratiform*
6	August 17, 2016 20:27	2.0	10.8	5.3	9.8 – 13.0	4	*Stratiform*
P5	August 8, 2014 20:55	4.67	11.56	14.5	18.1 – 22.0	4	*Stratiform*
P6	September 2, 2014 22:59	2.51	10.41	5.24	10.2 – 16.4	3	*Stratiform*
RG110	June 17, 2018 16:49 EDT	1.60	37.4	4.5	30.5 – 48.5	4	*Stratiform*

List of Figures

enhancement; Cluster-3: stratiform precipitation with bright-band; and Cluster-4: deep precipitation with mid- and low-level enhancement.

Figure 11: Mean (marker) with the standard deviation (error bars) of the features used to cluster MRMS reflectivity profiles: (a) maximum reflectivity; (b) height of maximum reflectivity [AGL]; (c) storm top height [AGL]; (d) near surface slope; and (e) near surface reflectivity.

Figure 12: Pearson correlation coefficient computed between the instantaneous hourly HRRR rain-rate at 500 m AGL and concurrent rain-gauge records.

Figure 13: Diurnal cycle of precipitation rate observed by the HRRR compared with that of the rain-gauge observations. The blue boxes highlight the time of day when the diurnal cycle patterns differ the most.

Figure 14: Diurnal cycle of (a) POD and (b) FA. GPM Ku-PR – GPM Algorithm Level 2A Version 6A. PDA - Precipitation Detection Algorithm (this work).

Figure 15: (a) MRMS precipitation rate, (b) MRMS precipitation type, and (c) GPM Ku-PR estimated near-surface precipitation rate for the overpass on October 11, 2018. The black circle in (c) marks the location selected for illustrating physically-based retrieval in Fig. 17.

Figure 16: (a) Cluster classes to GPM pixels by matching Ku-PR pixels with the MRMS mode. (b) Model predicted precipitation classes for the GPM overpass. Class 0: represents no-precipitation; Class-1: shallow light precipitation systems; Class-2: shallow surface enhanced precipitation; Class-3: stratiform precipitation; Class-4: deep intense precipitation systems with mid and low-level enhancement.

Figure 17: (a) Concurrent and nearest MRMS reflectivity and GPM Ku-PR Z_m and Z_e for the October 11, 2018 GPM overpass on the eastern ridges of the SAM (black circle in Fig. 15c). (b) Surface rain-rate simulated by the column microphysics model simulation for a shallow SFI scenario for three LLCF depths (300, 400 and 500 m) compared with MRMS mean precipitation rate within the Ku-PR pixel (dashed pink line) and KU-PR estimates (black dotted line). The pink shaded band delimits ± MRMS variance within the Ku-PR pixel. (c) Ensemble statistics (25^{th}, 50^{th} and 75^{th} percentiles) for 1000 simulations of near-surface rain-rate by the column microphysics model to capture uncertainty in LLCF microphysics and LLCF depth of 300m.

Figure 18: (a) Underestimation (blue) and overestimation (orange) errors for correct detection (CD) GPM Ku-PR cases organized by predicted precipitation class (i.e. MRMS cluster) and rain-rate relative to collocated rain-gauge measurements. Statistics of precipitation rate [25^{th}, 50^{th} (black circle) and 75^{th} percentiles]: (b) rain-gauge (RG) measurements; (c) RG measurements only for Ku-PR CD; and (d) CD for Ku-PR

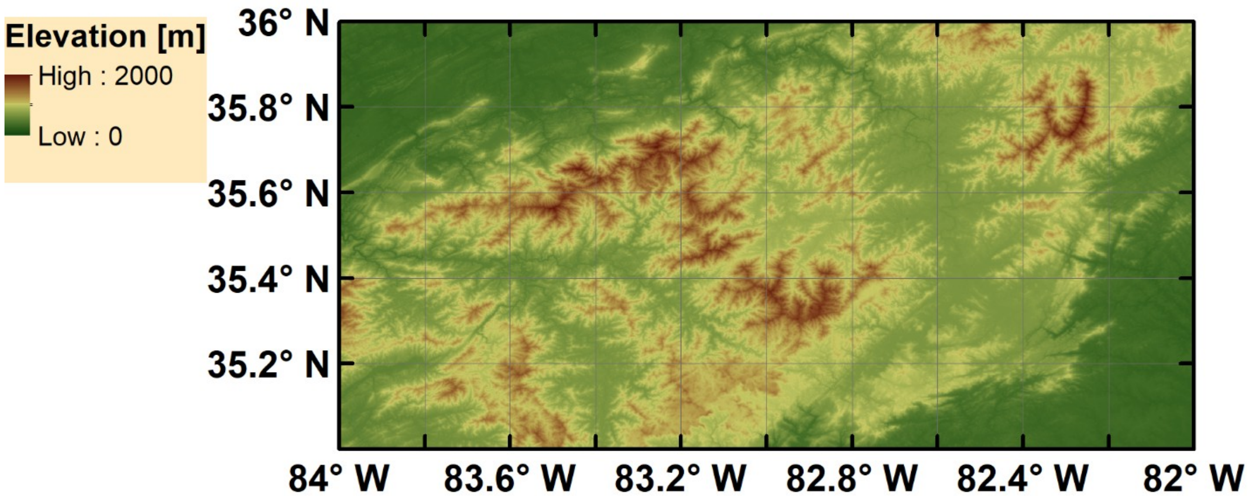

Figure 1: Digital Elevation Model (DEM) map of the Southern Appalachian Mountains (SAM).

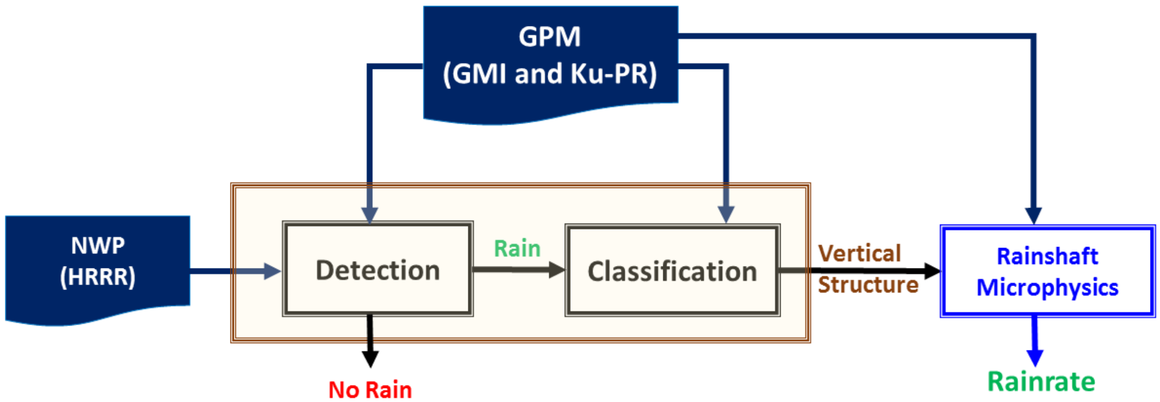

Figure 2: Overall architecture for physically-based precipitation retrieval. Dark blue - Input data. Brown and black- AI framework including detection and classification models (this work); Bright Blue – Physically-based model for quantitative precipitation estimation after Arulraj and Barros (2019).

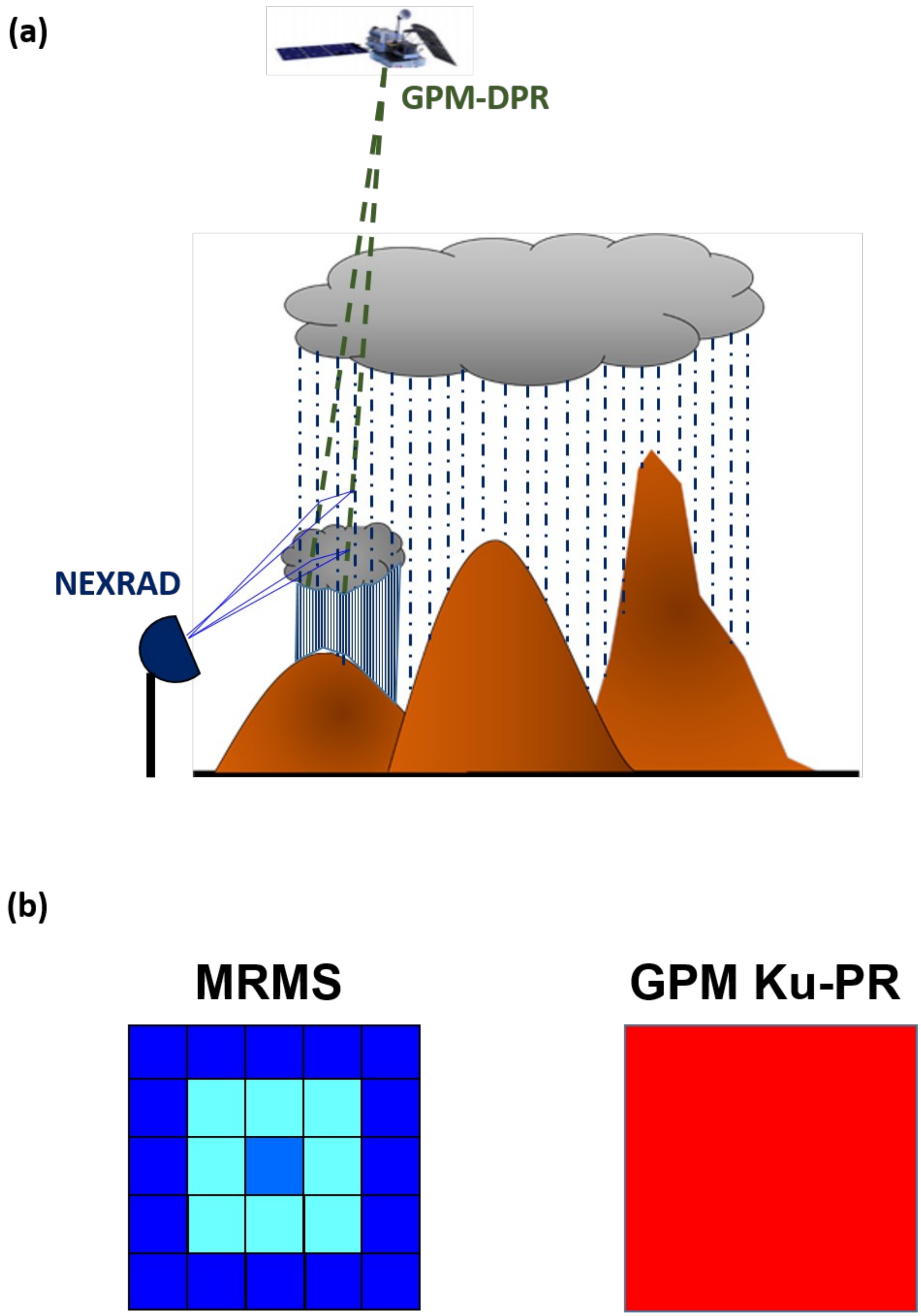

Figure 3: (a) Schematic showing the PPI (Plan Position Indicator) scanning pattern of the ground-based radar network (NEXRAD) used in MRMS and the vertical scanning pattern of the GPM DPR. (b) Contrast in spatial resolution between MRMS and GPM DPR with 25 (5×5) MRMS pixels (1×1 km^2) for each GPM DPR pixel.

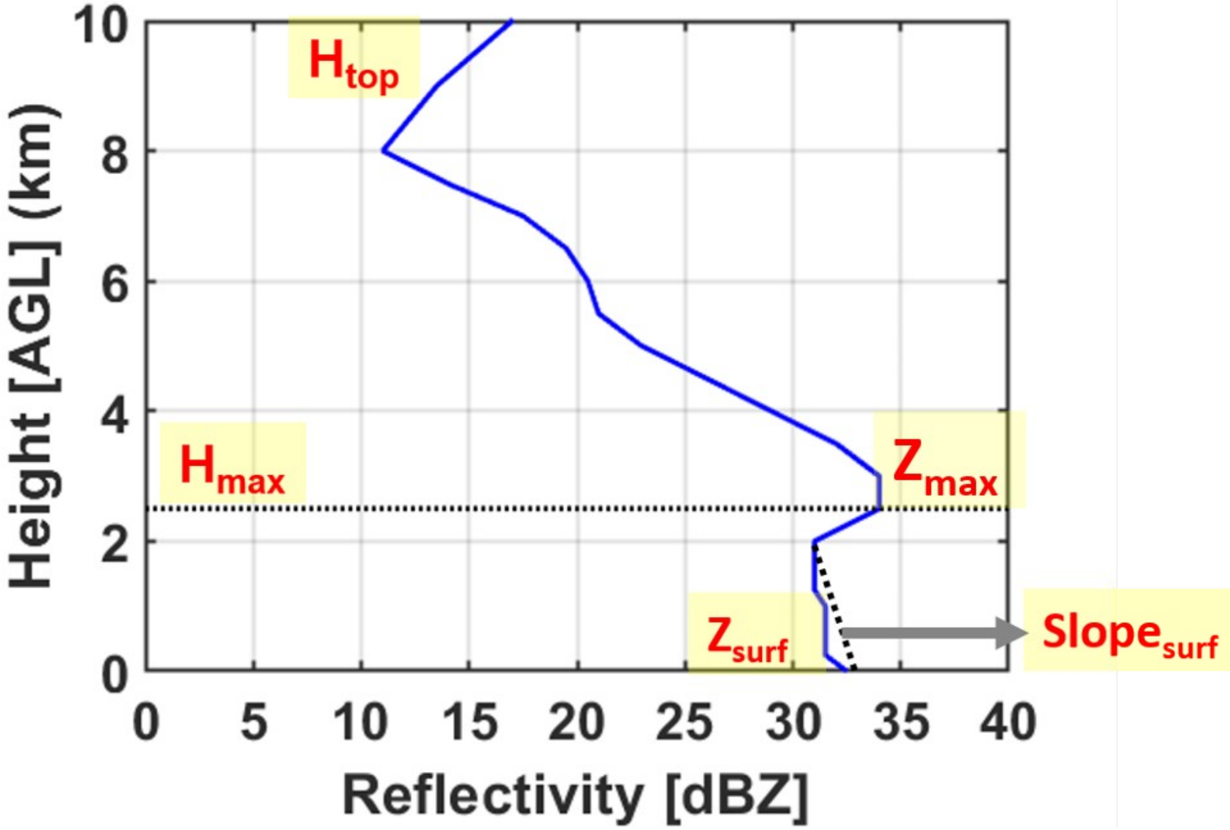

Figure 4: Features extracted from the MRMS reflectivity profiles for clustering algorithm.

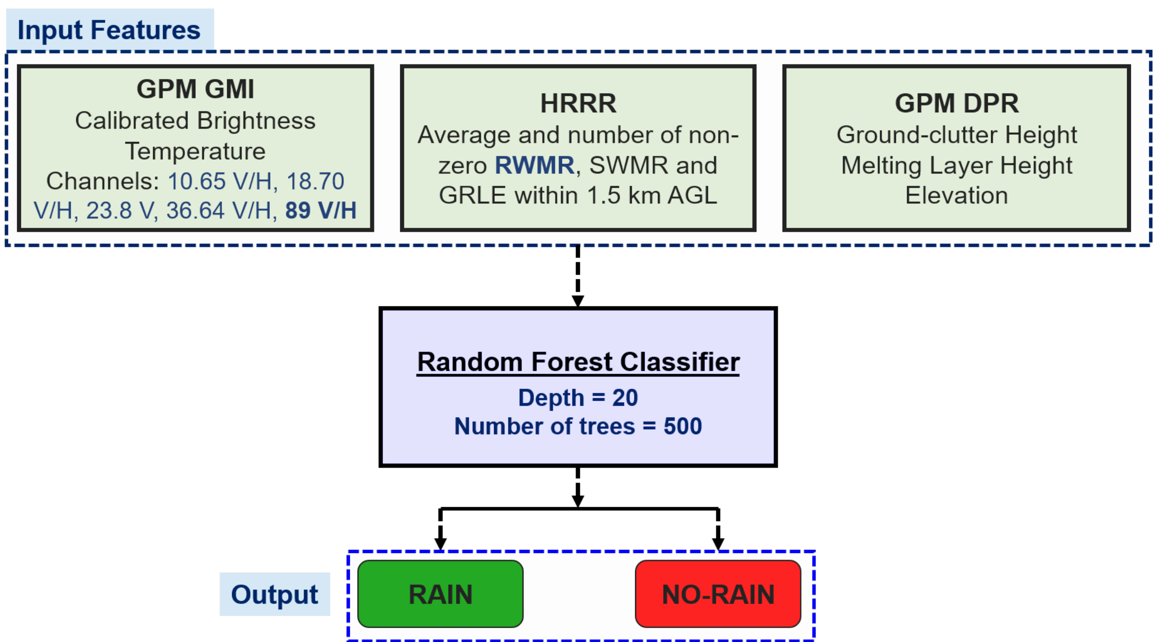

Figure 5: Schematic of the precipitation detection model (Appendix A, Section 3.2.1). The inputs of this framework are obtained from GPM GMI calibrated brightness temperatures and selected features from GPM DPR reflectivity profiles, and water mixing ratios (MR) from HRRR analysis (Rainwater RW; Snow SW; and Graupel GRLE).

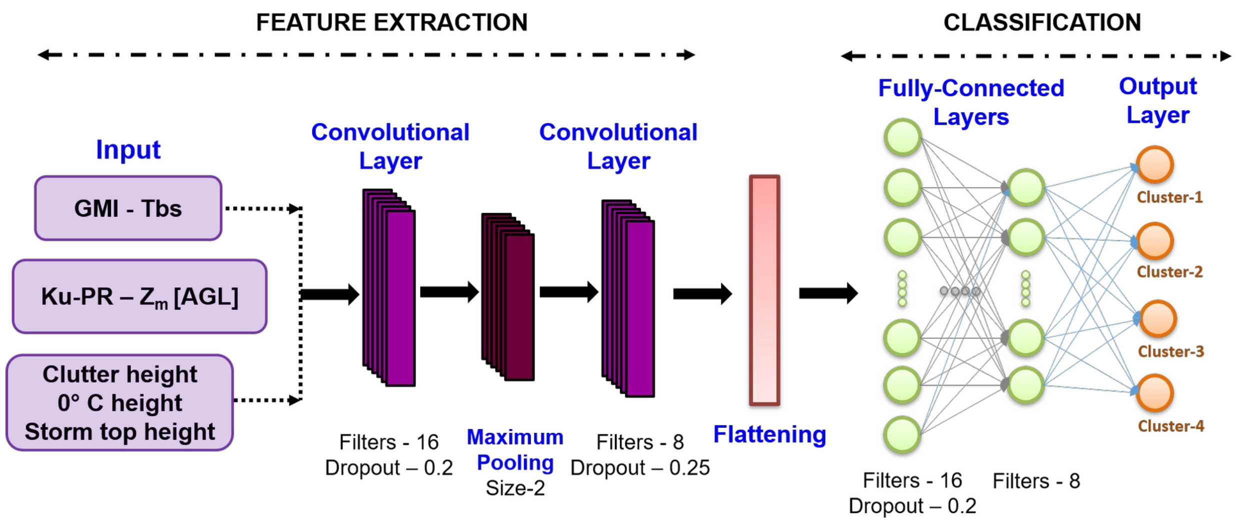

Figure 6: Schematic of the precipitation classification model (Appendix B, Section 3.2.2). The convolutional neural network (CNN) maps individual GPM Ku-PR pixels where "rain" is detected to MRMS vertical structure clusters (i.e. precipitation regimes). The CNN consists of a feature extraction sub-model that identifies the main associations among the input data, and a classification neural network sub-model that links the features to precipitation class (i.e. MRMS vertical structure clusters).

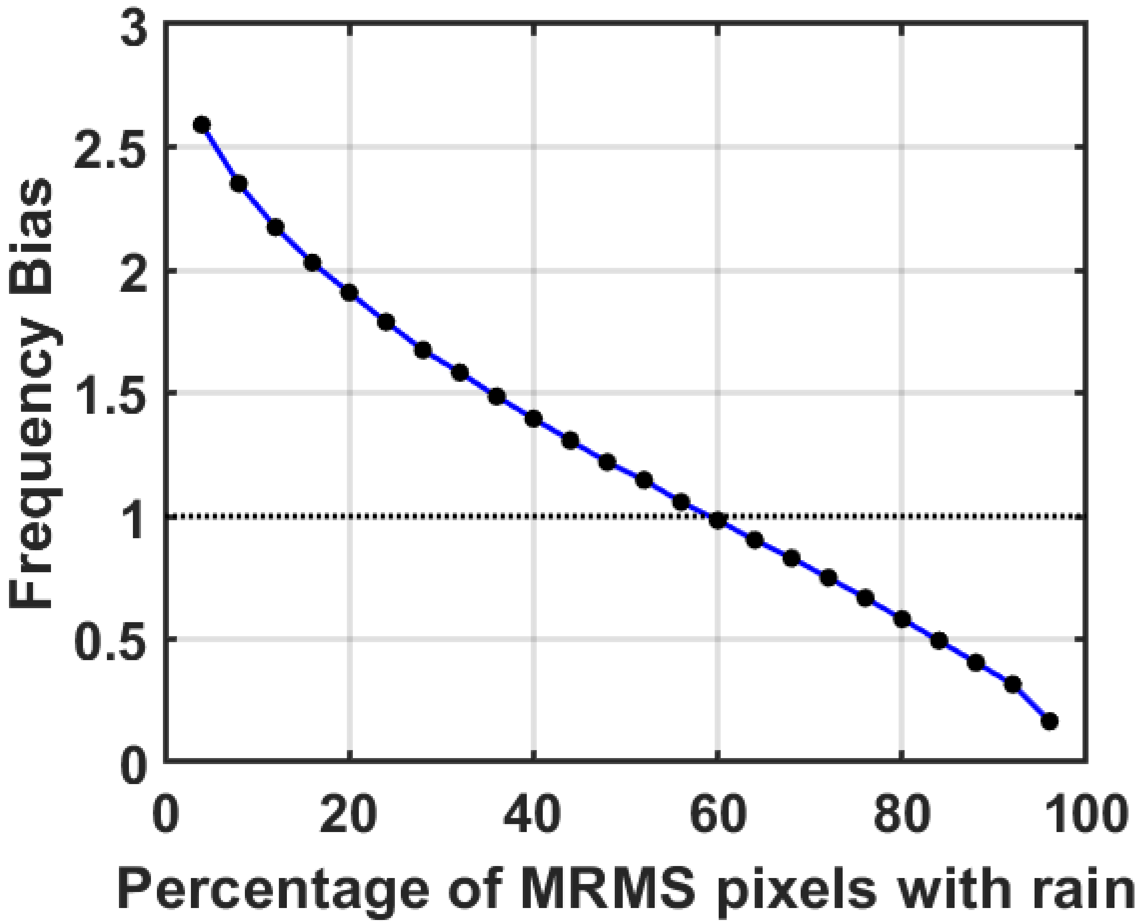

Figure 7: Frequency bias as a function of the fraction of non-uniformly precipitating MRMS pixels within the GPM Ku-PR footprint.

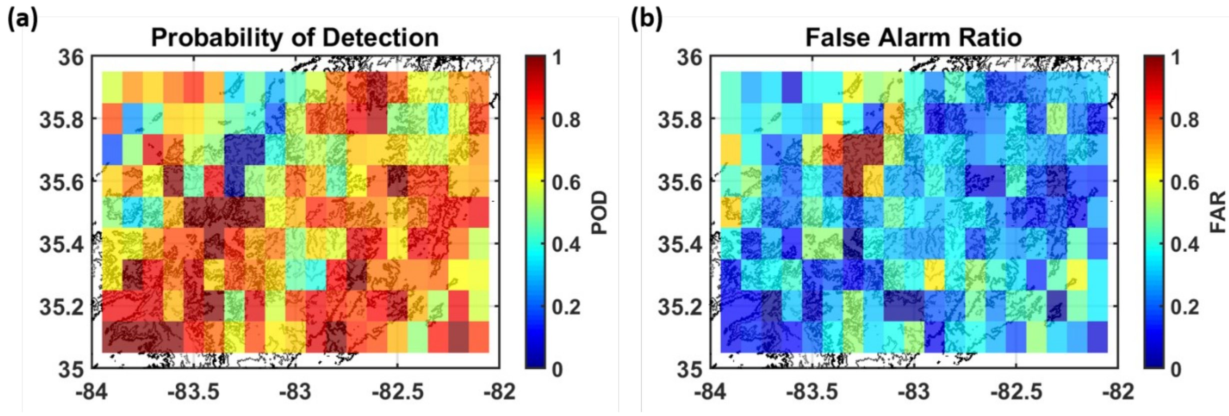

Figure 8: Spatial distribution of the detection metrics of GPM Ku-PR near-surface precipitation estimates relative to Level 2 MRMS precipitation. a) Probability of Detection - POD; b) False Alarm Ratio - FAR. The underlying black and gray contours represent elevation contours (Fig.1).

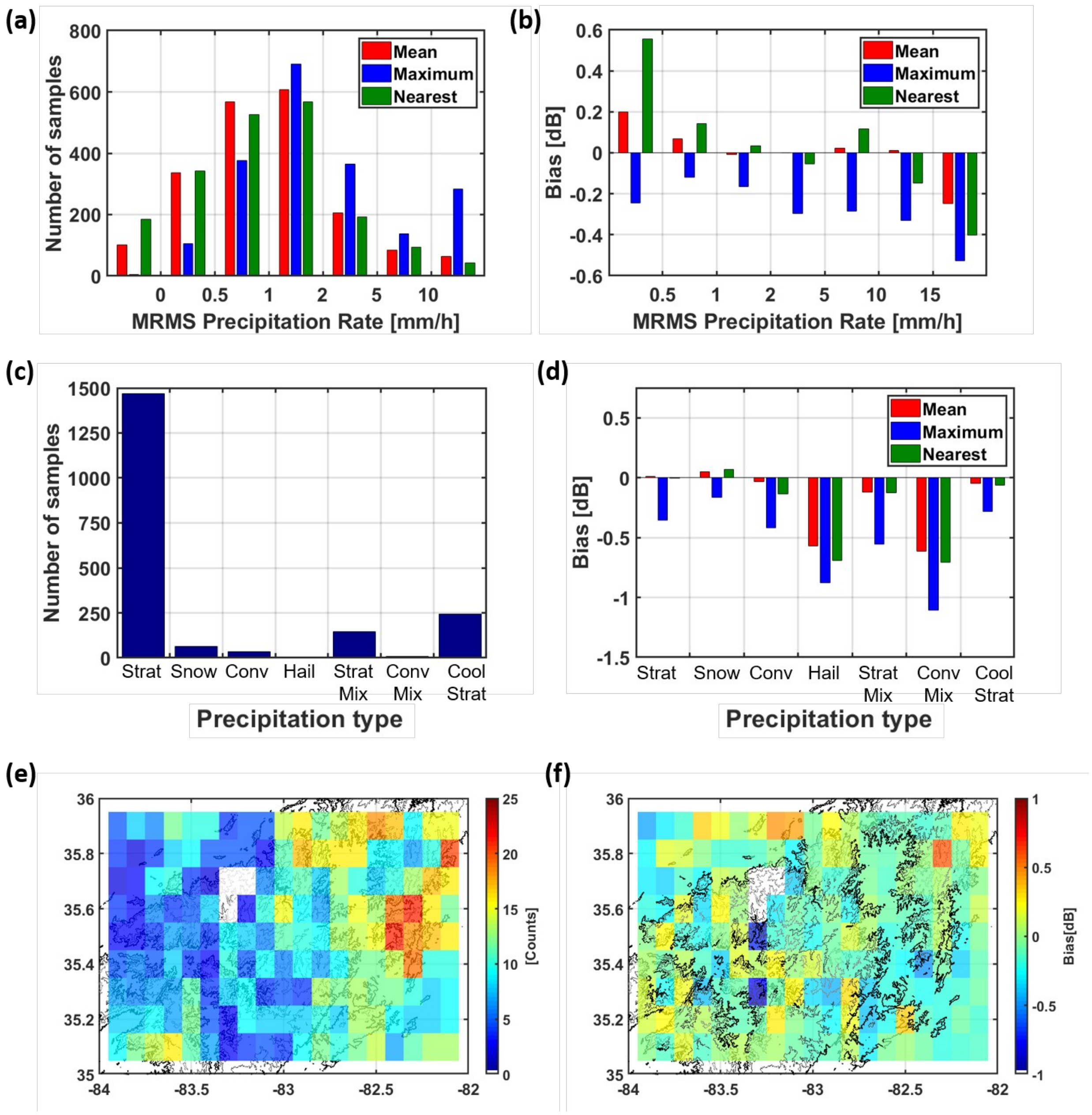

Figure 9: Histograms of number of samples (a, c) and the corresponding bias (b, d) in GPM Ku-PR relative to MRMS precipitation rate and type. Spatial distribution of number of samples (e) and bias (f). Precipitation type: Strat - warm stratiform; Conv - convective; Strat Mix - tropical/stratiform mix; Conv Mix - tropical/convective mix; Cool Strat - cool stratiform.

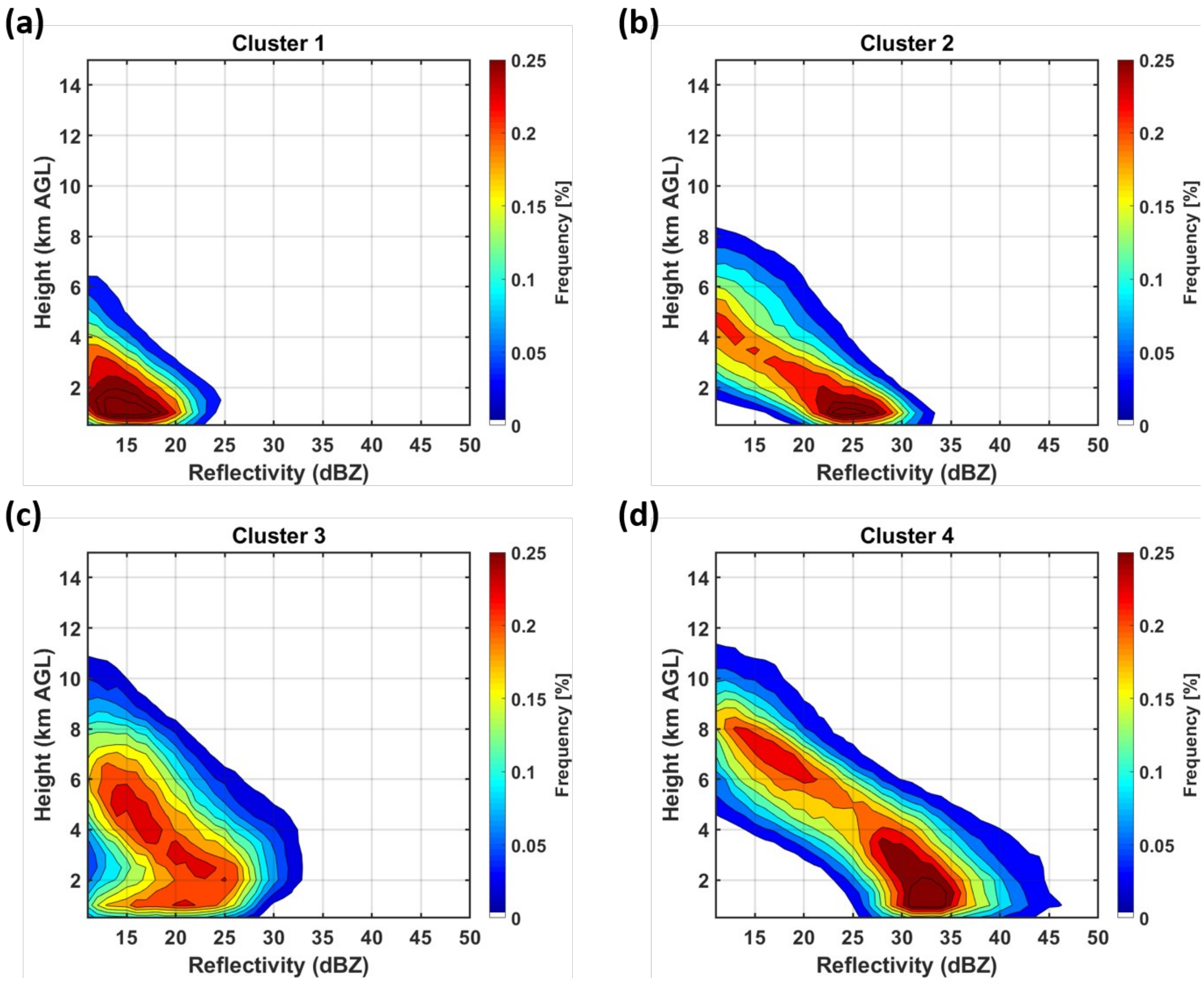

Figure 10: Contoured frequency altitude diagram (CFADs) of the four distinct clusters of vertical structure of MRMS reflectivity profiles identified by the k-means algorithm in the SAM domain. Cluster-1: light low-level precipitation; Cluster-2 Shallow precipitation with low level enhancement; Cluster-3: stratiform precipitation with bright-band; and Cluster-4: deep precipitation with mid- and low-level enhancement.

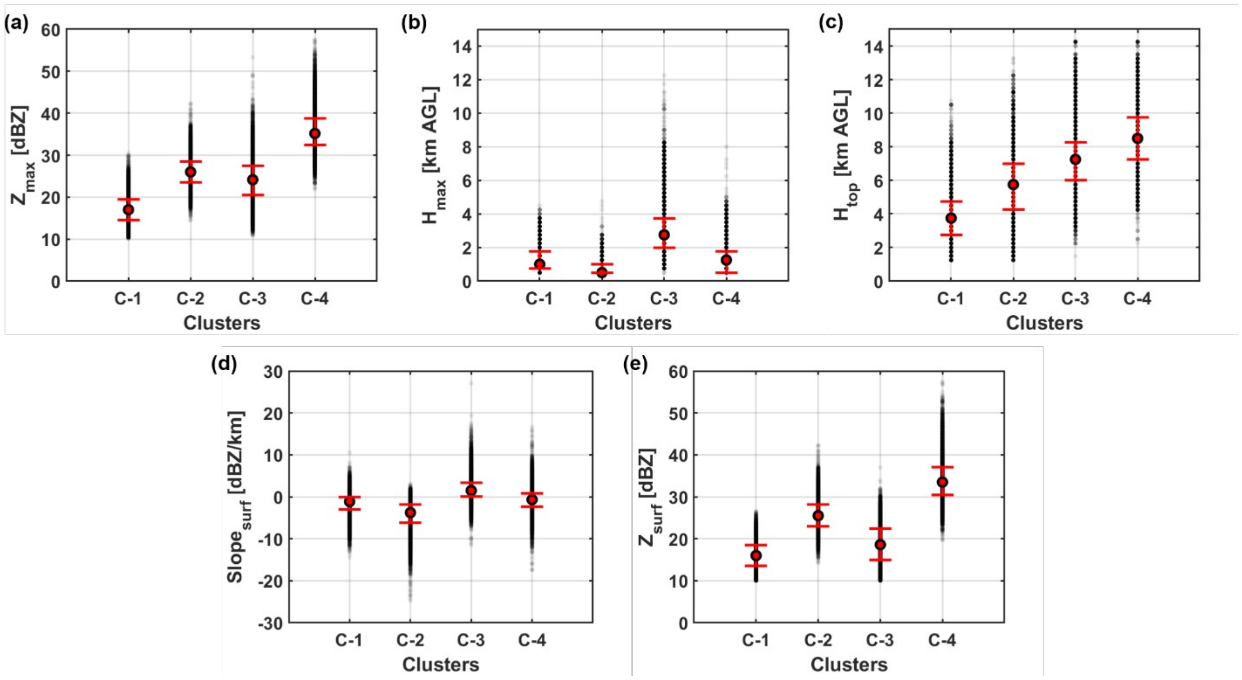

Figure 11: Mean (marker) with the standard deviation (error bars) of the features used to cluster MRMS reflectivity profiles: (a) maximum reflectivity; (b) height of maximum reflectivity [AGL]; (c) storm top height [AGL]; (d) near surface slope; and (e) near surface reflectivity.

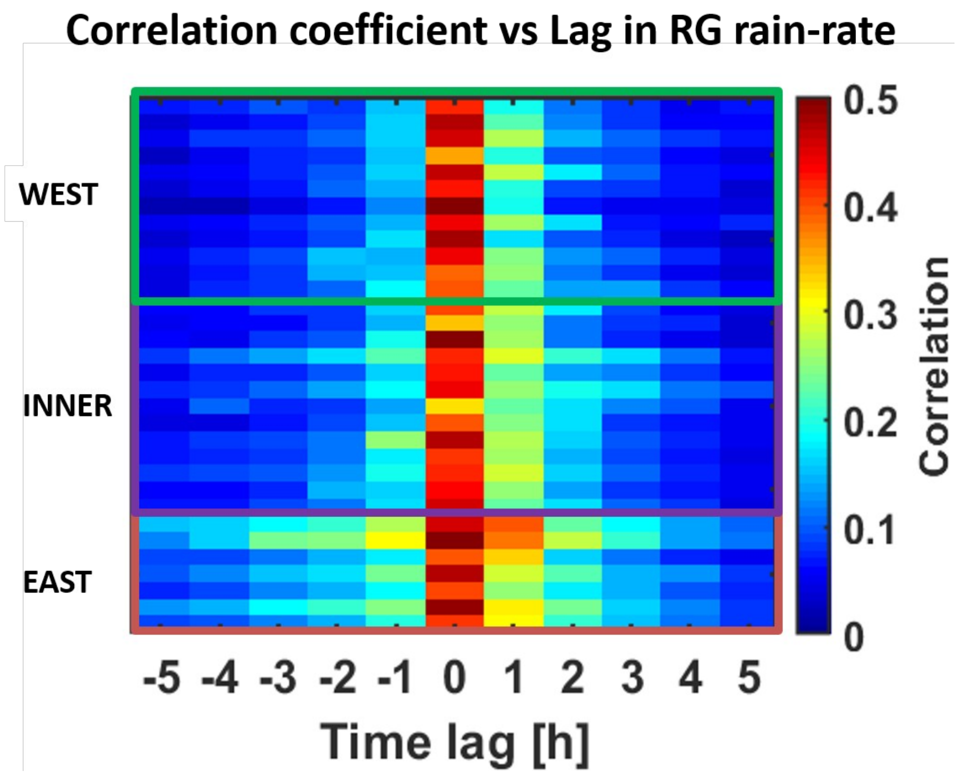

Figure 12: Pearson correlation coefficient computed between the instantaneous hourly HRRR rain-rate at 500 m AGL and concurrent rain-gauge records.

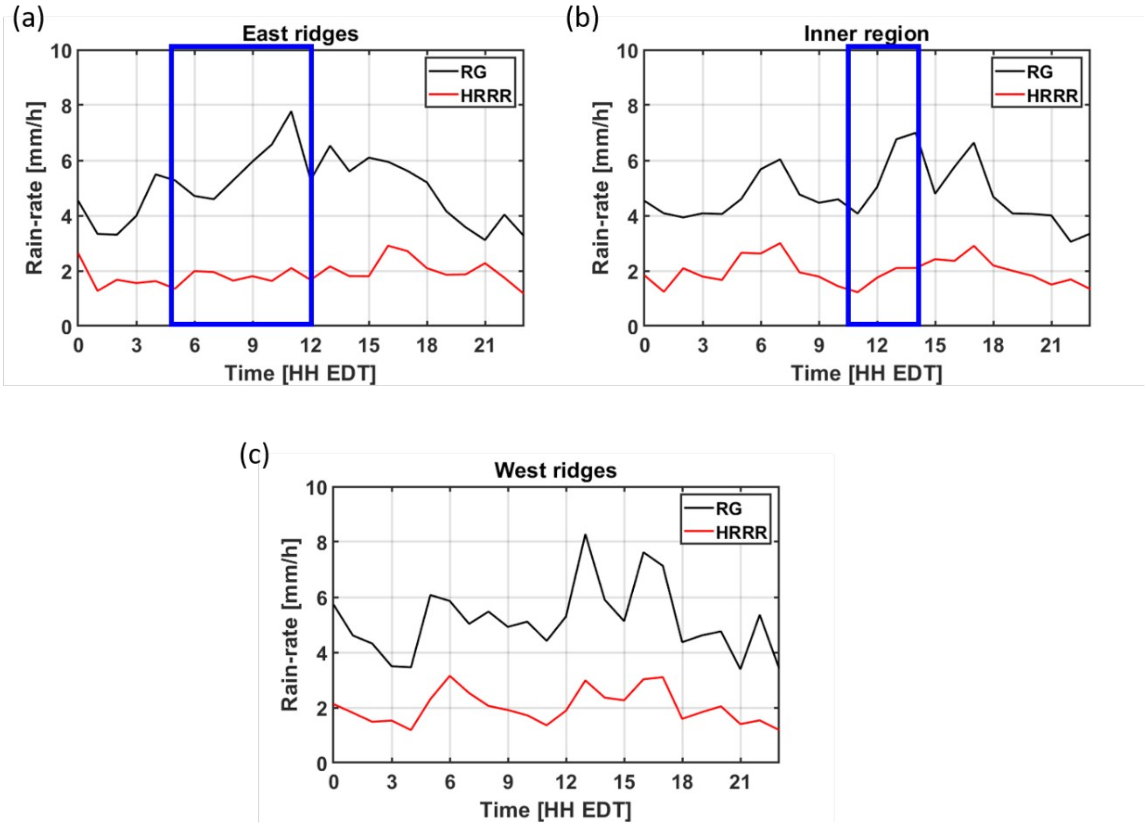

Figure 13: Diurnal cycle of precipitation rate observed by the HRRR compared with that of the rain-gauge observations. The blue boxes highlight the time of day when the diurnal cycle patterns differ the most.

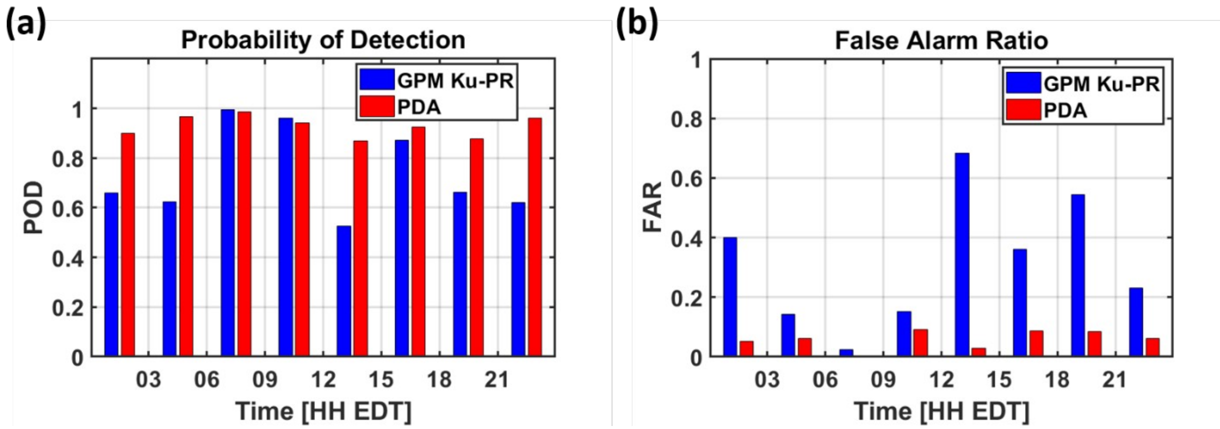

Figure 14: Diurnal cycle of (a) POD and (b) FA. GPM Ku-PR – GPM Algorithm Level 2A Version 6A. PDA - Precipitation Detection Algorithm (this work).

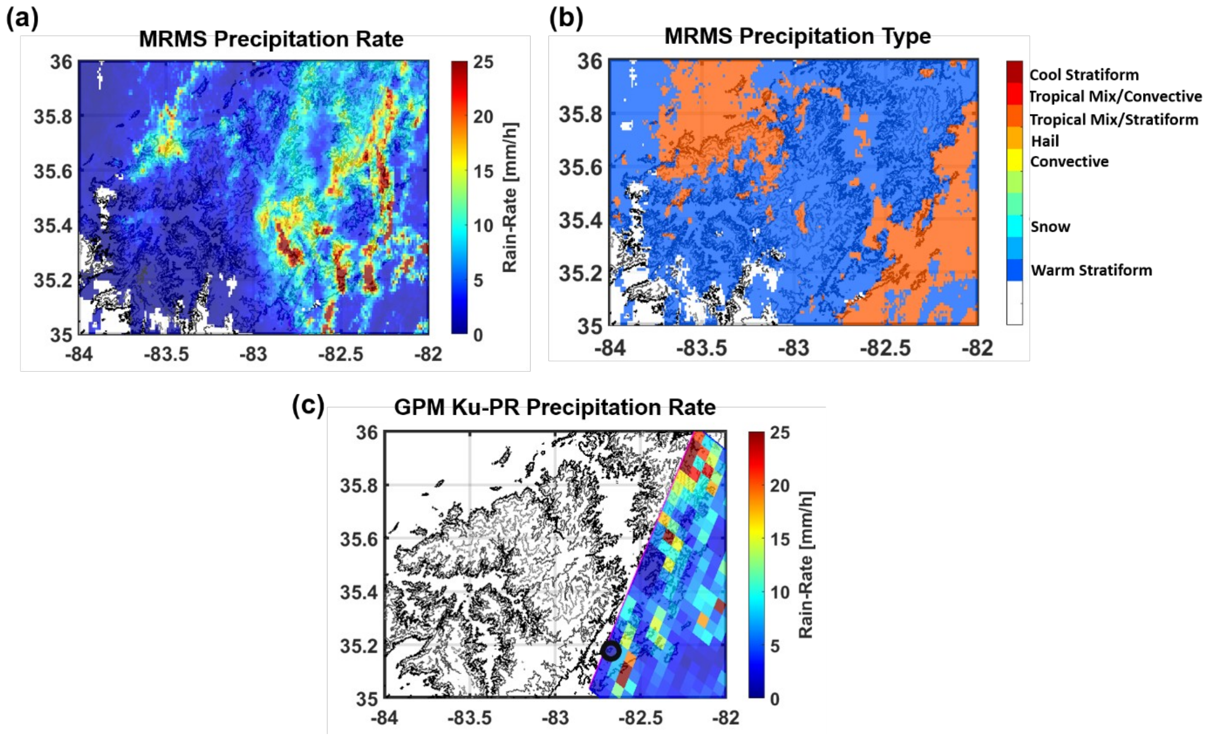

Figure 15: (a) MRMS precipitation rate, (b) MRMS precipitation type, and (c) GPM Ku-PR estimated near-surface precipitation rate for the overpass on October 11, 2018. The black circle in (c) marks the location selected for illustrating physically-based retrieval in Fig. 17.

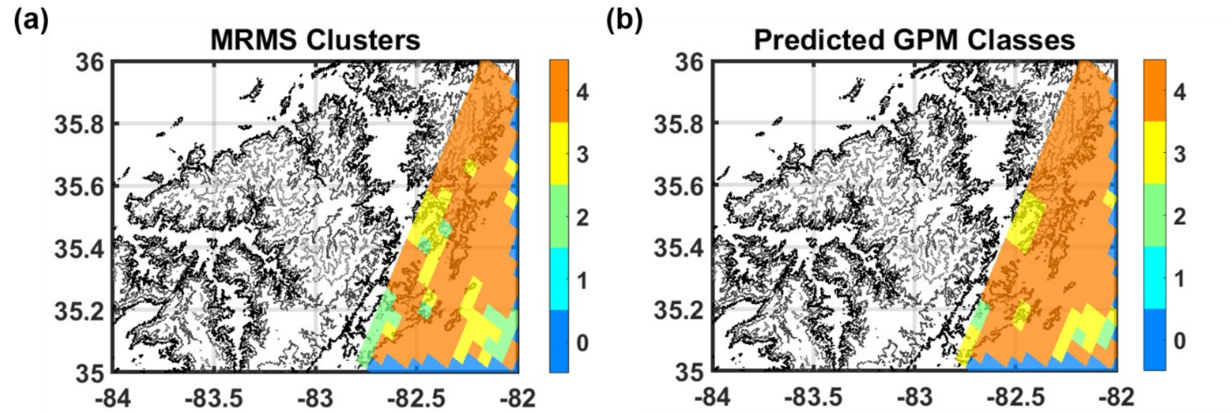

Figure 16: (a) Cluster classes to GPM pixels by matching Ku-PR pixels with the MRMS mode. (b) Model predicted precipitation classes for the GPM overpass. Class 0: represents no-precipitation; Class-1: shallow light precipitation systems; Class-2: shallow surface enhanced precipitation; Class-3: stratiform precipitation; Class-4: deep intense precipitation systems with mid and low-level enhancement.

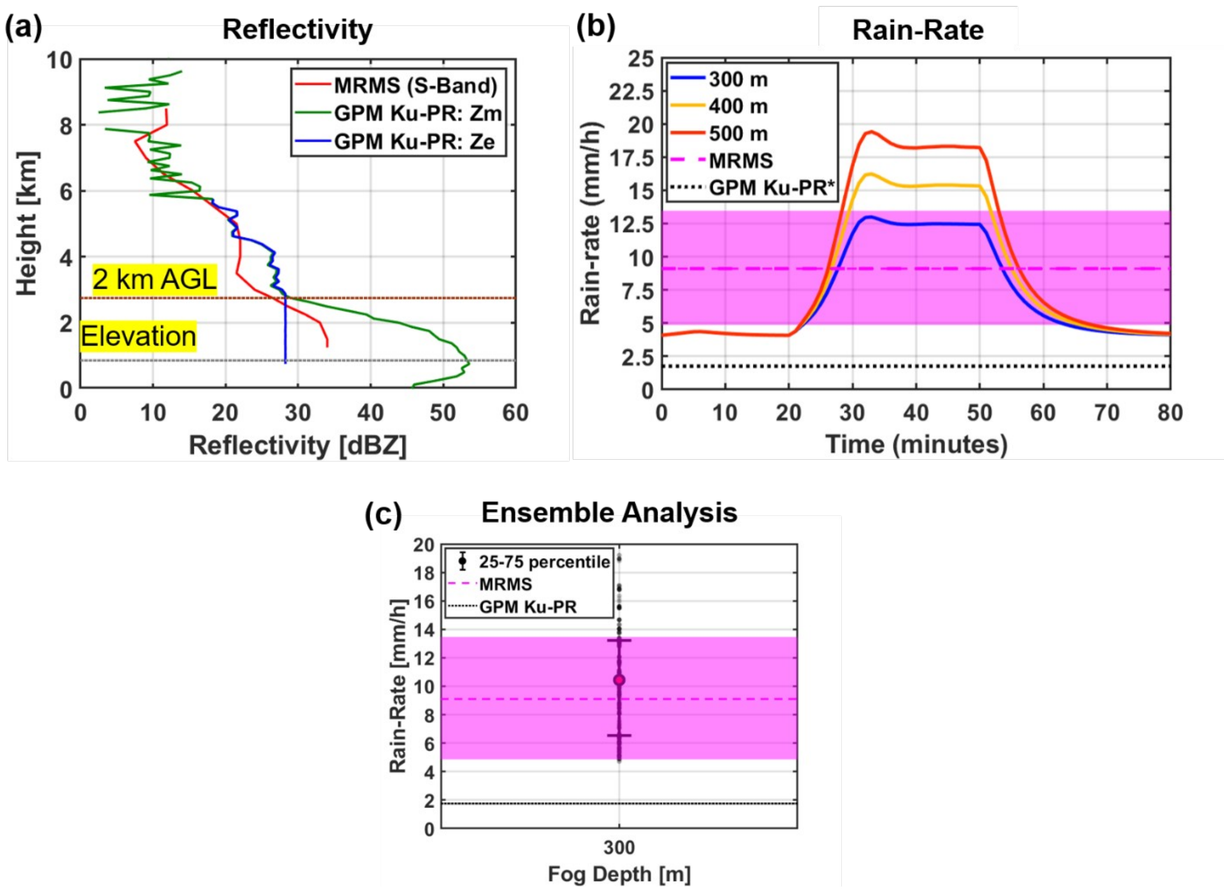

Figure 17: (a) Concurrent and nearest MRMS reflectivity and GPM Ku-PR Z_m and Z_e for the October 11, 2018 GPM overpass on the eastern ridges of the SAM (black circle in Fig. 15c). (b) Surface rain-rate simulated by the column microphysics model simulation for a shallow SFI scenario for three LLCF depths (300, 400 and 500 m) compared with MRMS mean precipitation rate within the Ku-PR pixel (dashed pink line) and KU-PR estimates (black dotted line). The pink shaded band delimits ± MRMS variance within the Ku-PR pixel. (c) Ensemble statistics (25^{th}, 50^{th} and 75^{th} percentiles) for 1000 simulations of near-surface rain-rate by the column microphysics model to capture uncertainty in LLCF microphysics and LLCF depth of 300m.

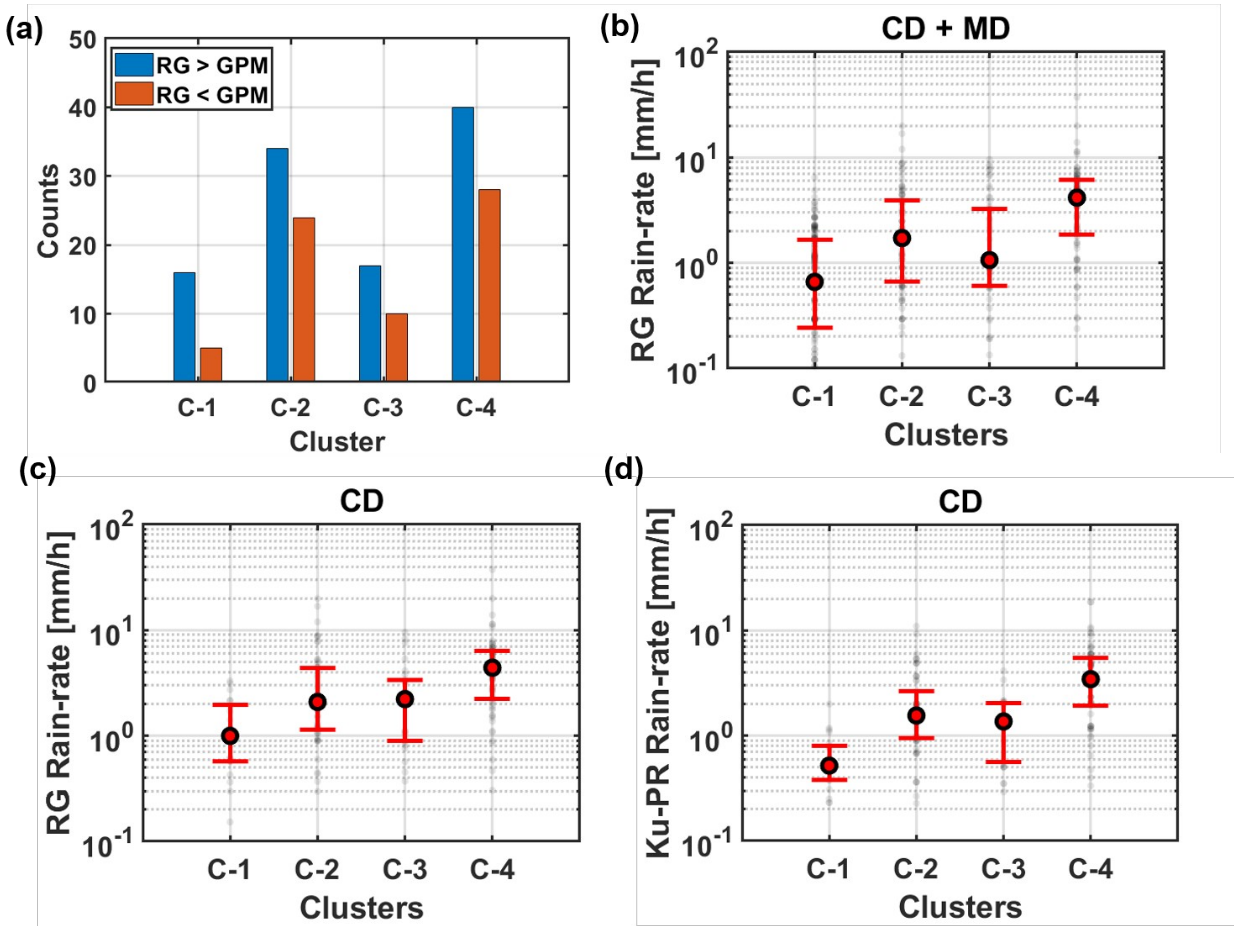

Figure 18: (a) Underestimation (blue) and overestimation (orange) errors for correct detection (CD) GPM Ku-PR cases organized by predicted precipitation class (i.e. MRMS cluster) and rain-rate relative to collocated rain-gauge measurements. Statistics of precipitation rate [25th, 50th (black circle) and 75th percentiles]: (b) rain-gauge (RG) measurements; (c) RG measurements only for Ku-PR CD; and (d) CD for Ku-PR.

Supplementary Data

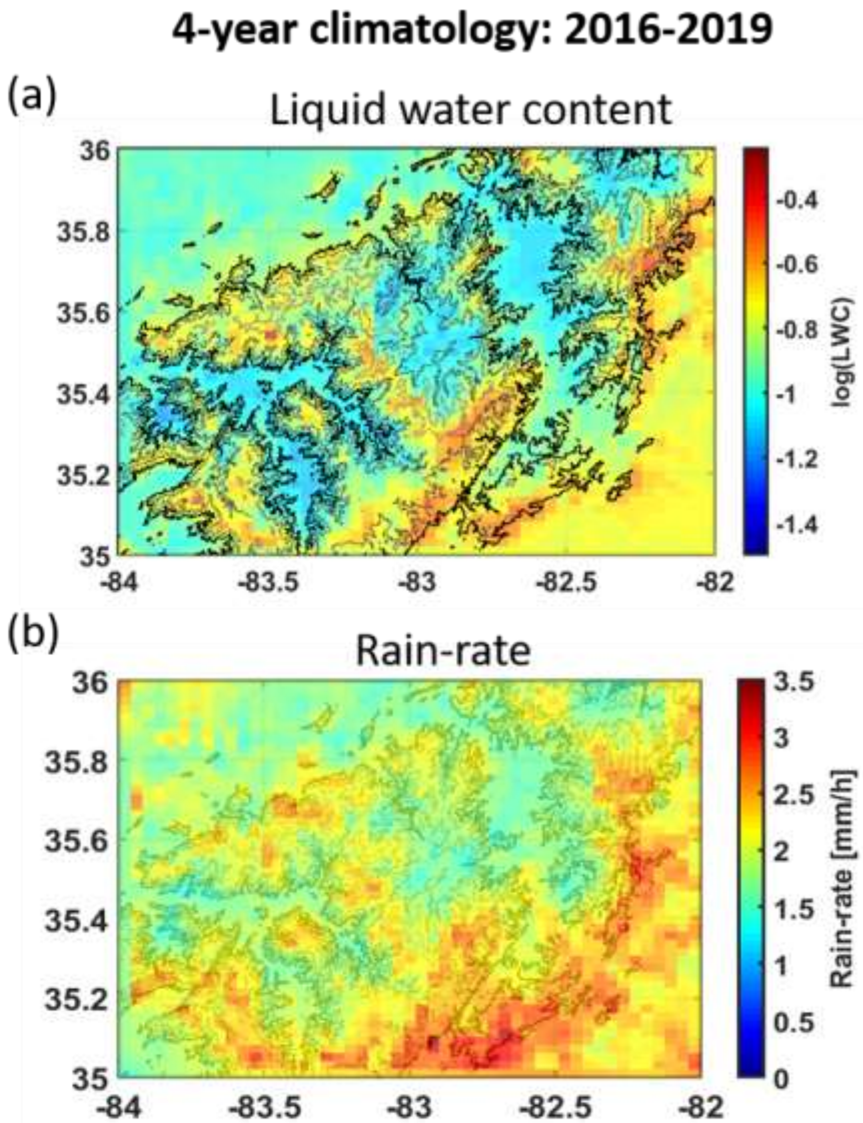

Figure S1. Four-year climatology of (a) Liquid water content and (b) Rain-rate at 500 m AGL estimated by HRRR in the summer (May-October).

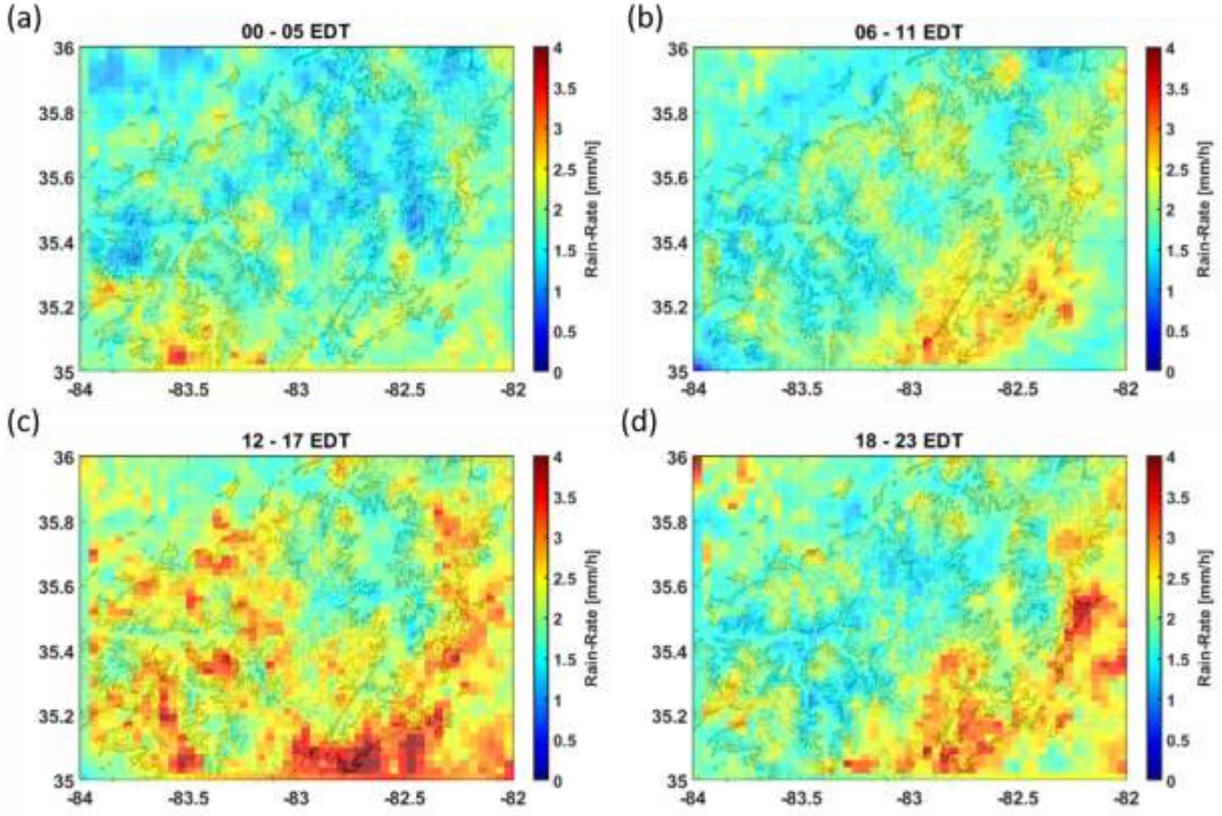

Figure S2. Diurnal cycle of the HRRR rain-rate climatology at 500 AGL in summer (May-October).

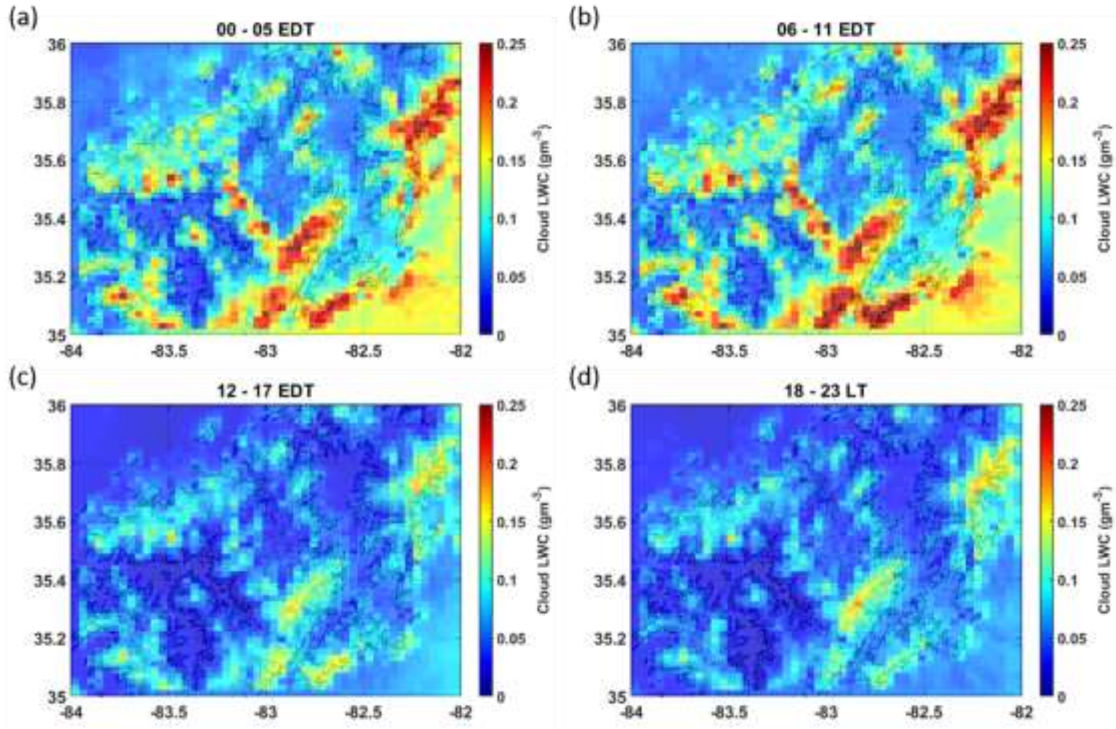

Figure S3. Diurnal cycle 4-year climatology of HRRR estimated cloud-water mixing ratio at 500 m AGL in summer months (May-October).

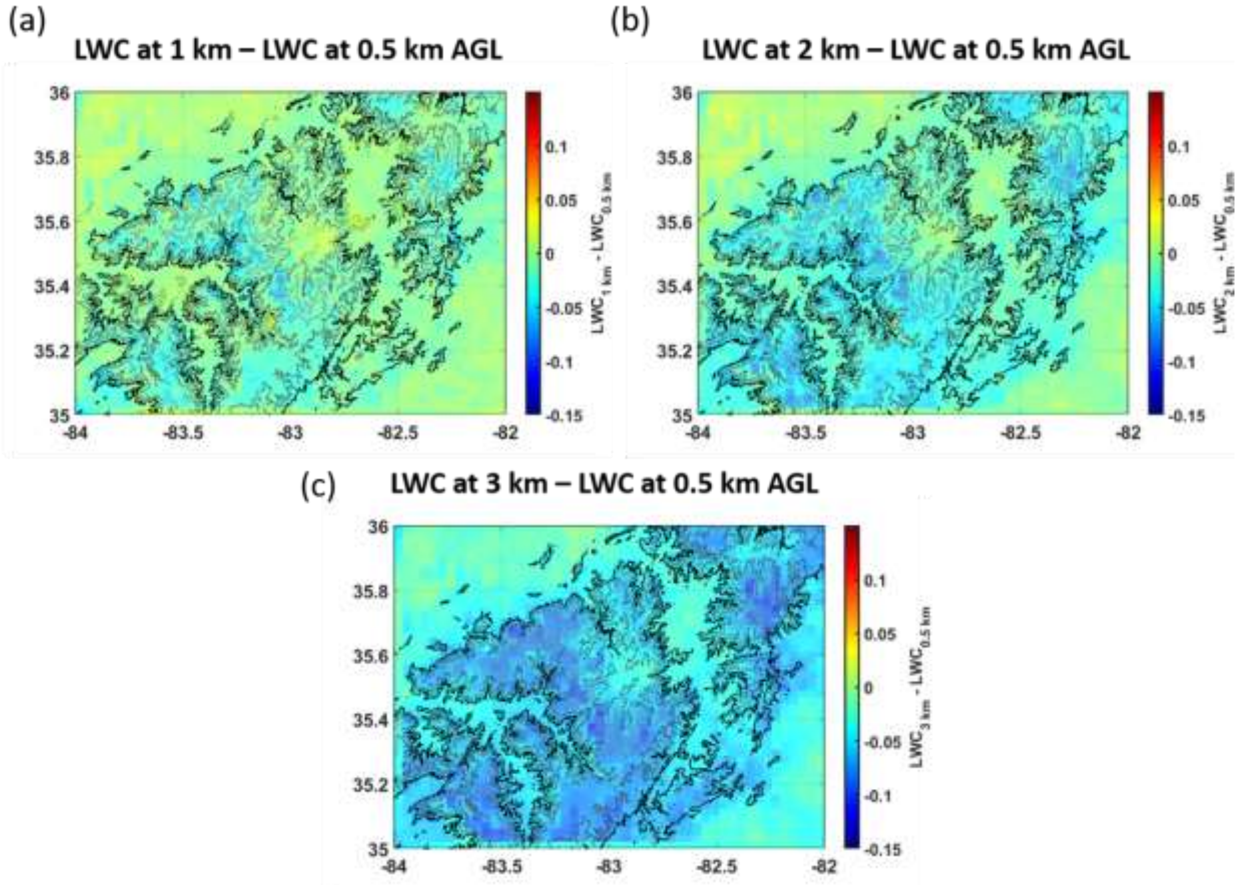

Figure S4. Spatial distribution of surface enhancement of HRRR rain liquid water content. Difference between mean rain liquid water content at (a) 1 km and 0.5 km; (b) 2 km and 0.5 km and (c) 3 km and 0.5 km AGL. High negative values indicate near-surface enhancement of precipitation tied to detection and estimation errors in GPM Ku-PR precipitation products.

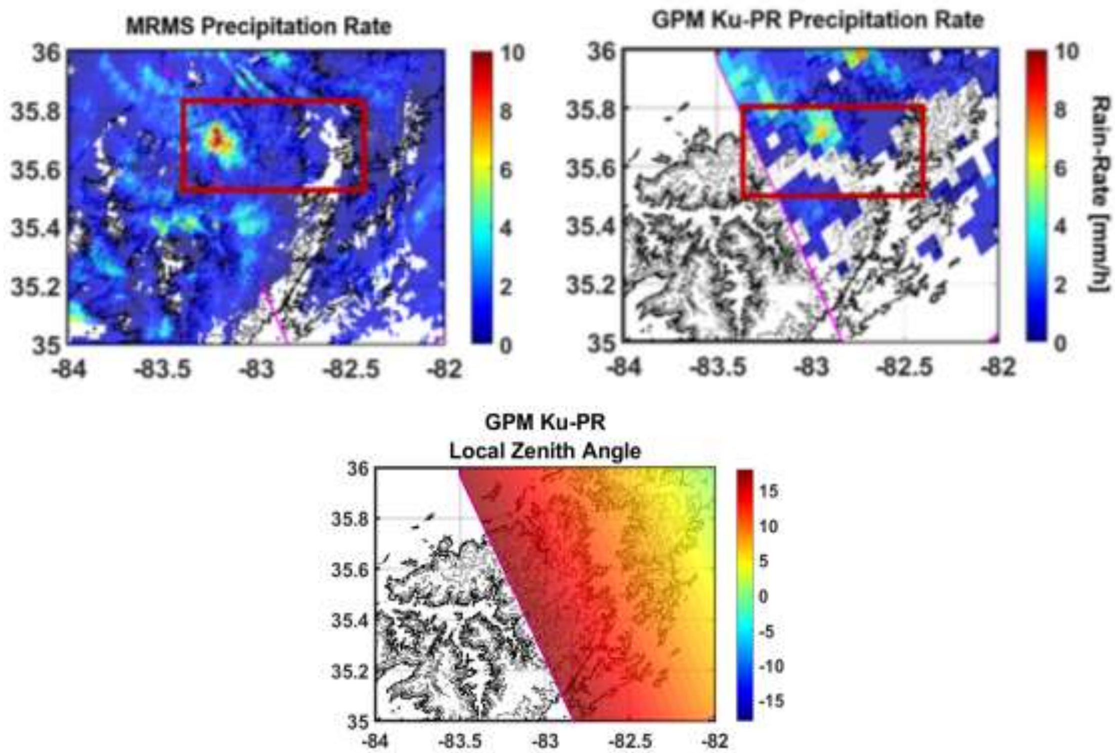

Figure S5. Example of GPM overpass on February 18, 2019 with large parallax errors in Ku-PR rainfall estimates compared to MRMS due to the large viewing angle.

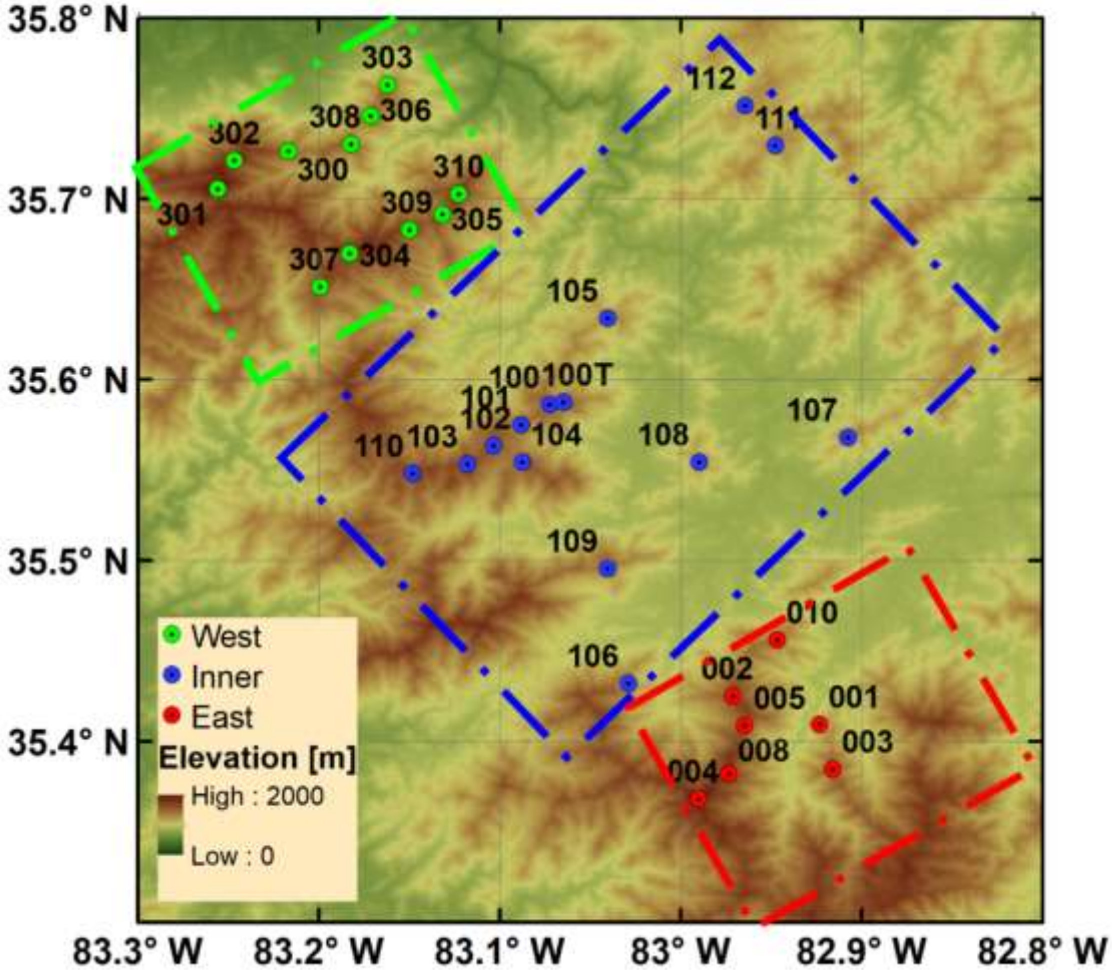

Figure S6. Map showing the region of study at the Southern Appalachian Mountains (SAM) with locations of long-term rain-gauge network. The green, blue and red boxes denote western, inner and eastern region [from Arulraj and Barros, 2019].

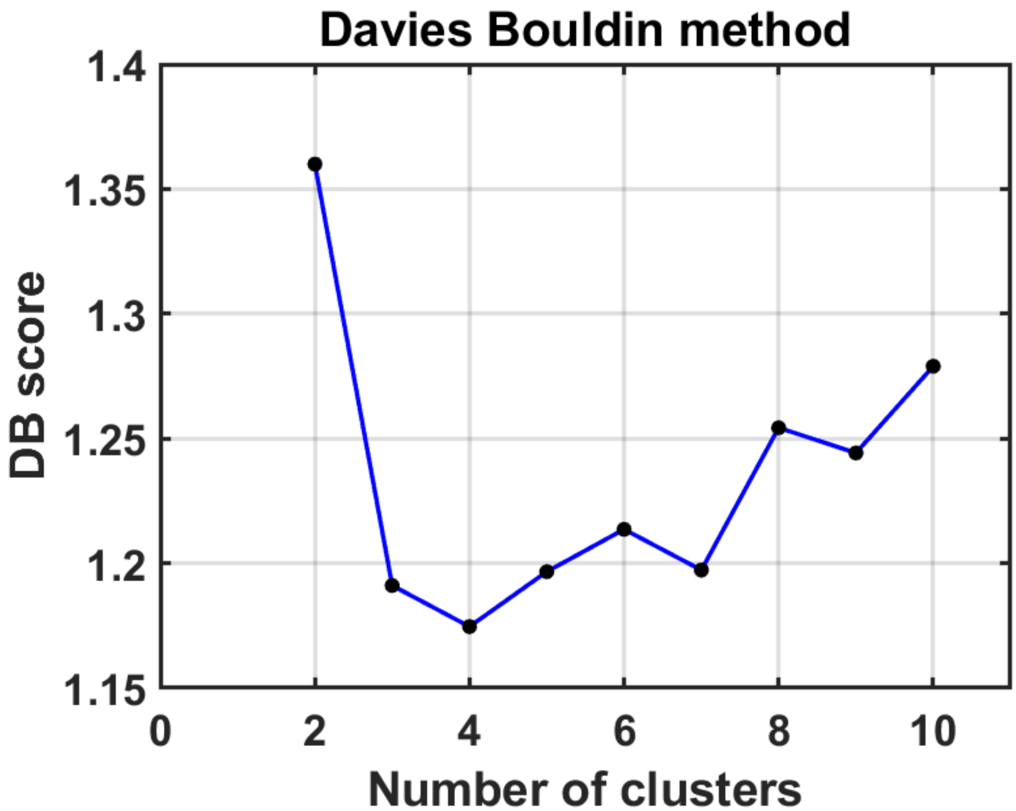

Figure S7. Davies-Bouldin (DB) index computed for different number of clusters of MRMS reflectivity features in the k-means algorithm.

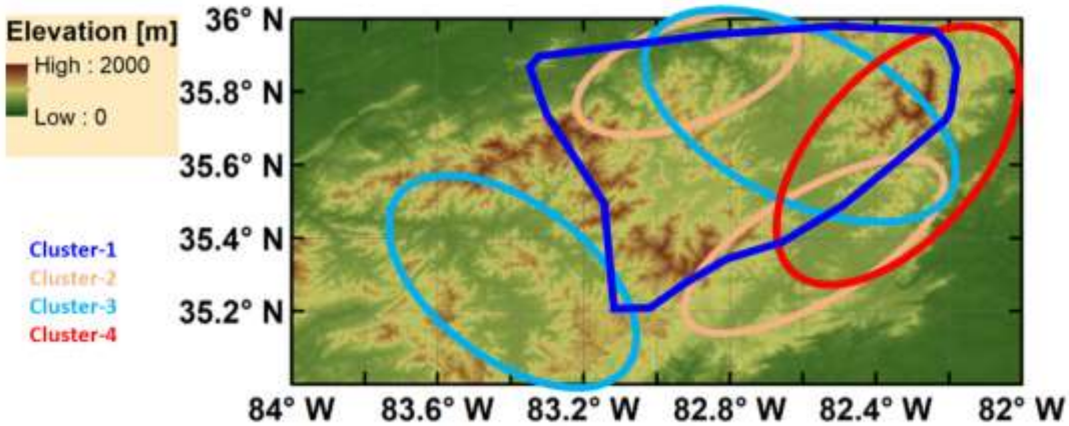

Figure S8. Synthesis of the spatial patterns of most frequent occurrences (hot-spots) of the four MRMS cluster classes.

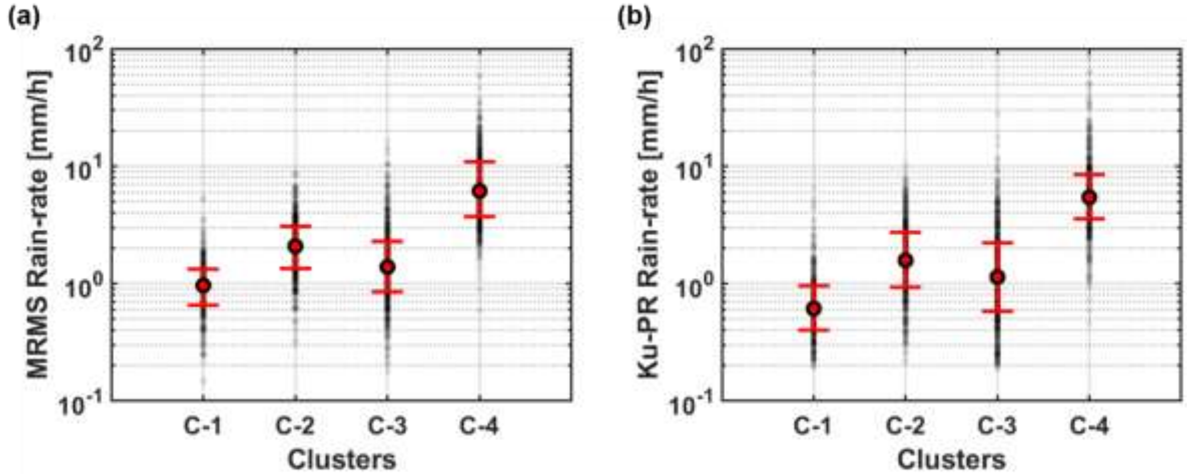

Figure S9. Concurrent mean precipitation rate (marker) with the standard deviation (error bars) for each of the MRMS clusters: (a) MRMS and (b) Ku-PR.

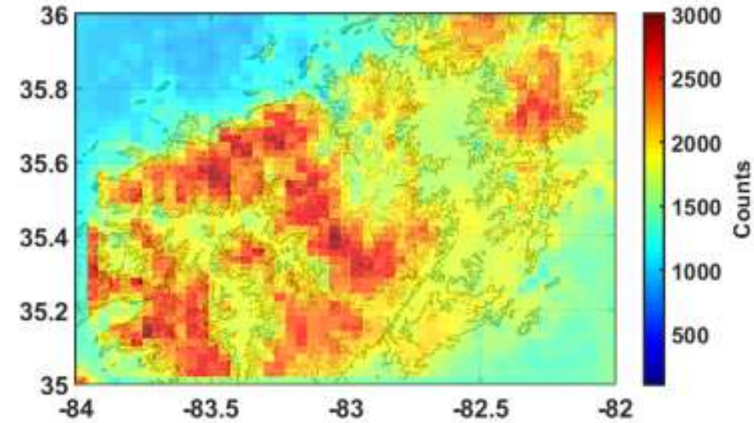

Figure S10. Spatial distribution of HRRR precipitation frequency.

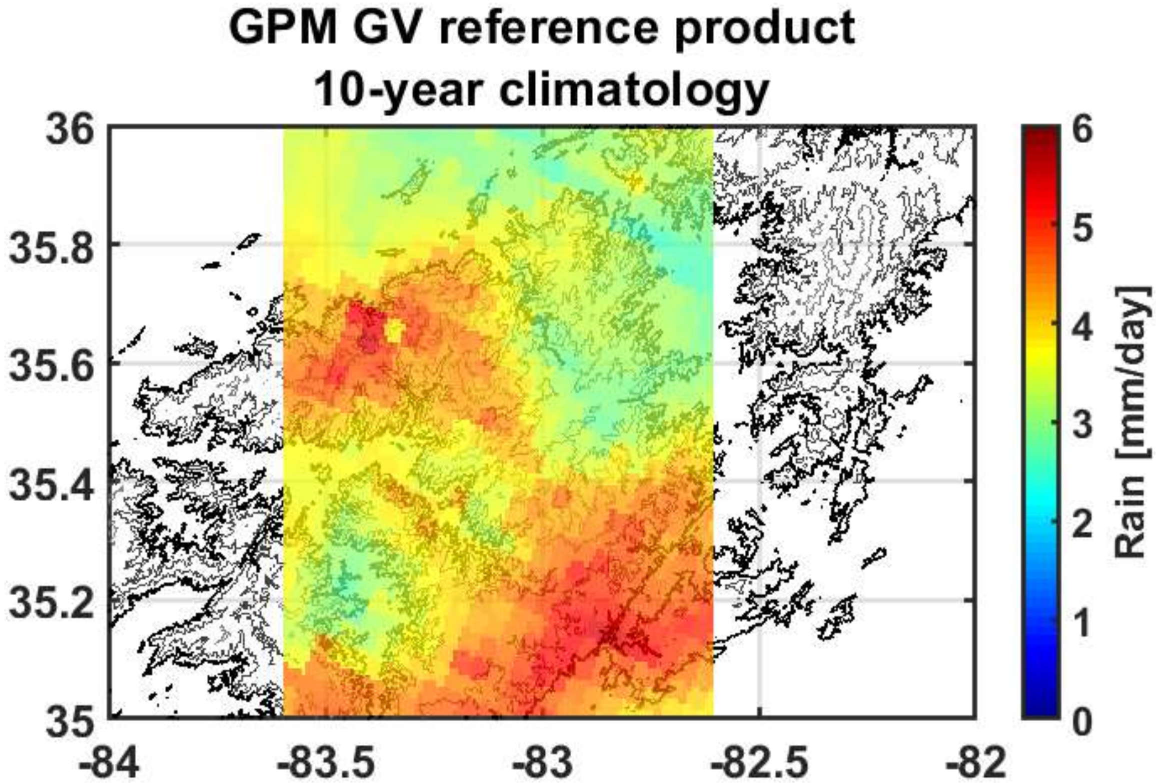

Figure S11. Spatial distribution of the 10-year average precipitation climatology from GPM GV reference precipitation product (Liao and Barros 2019).

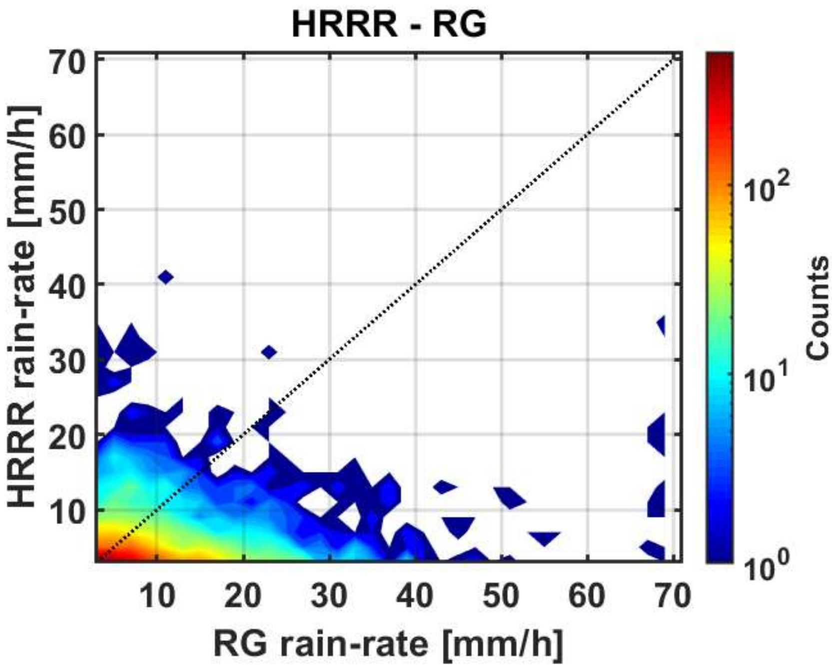

Figure S12. Density plot comparing the rain-rate derived from the rain-gauges with the HRRR simulated rain-rates at 500-m AGL.

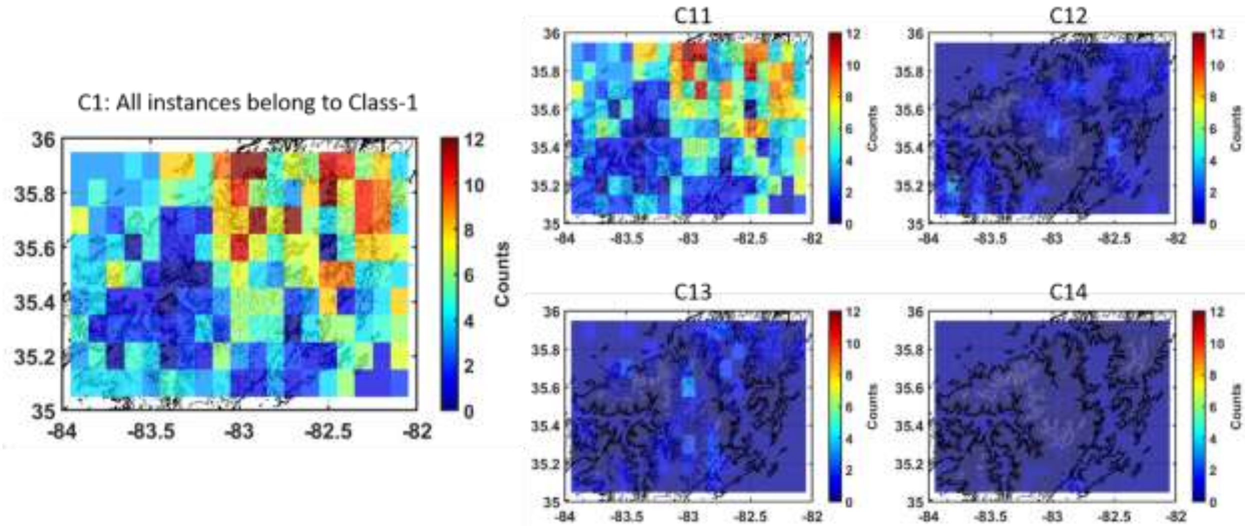

Figure S13. Left panel - Spatial distribution of MRMS Cluster-1 frequency of occurrence. Right panel – Spatial distribution of Cluster-1 GPM Ku-PR reflectivity profiles predicted by the precipitation classification algorithm. CXY denotes the output of the classification algorithm where X is the true class and Y is the model output.

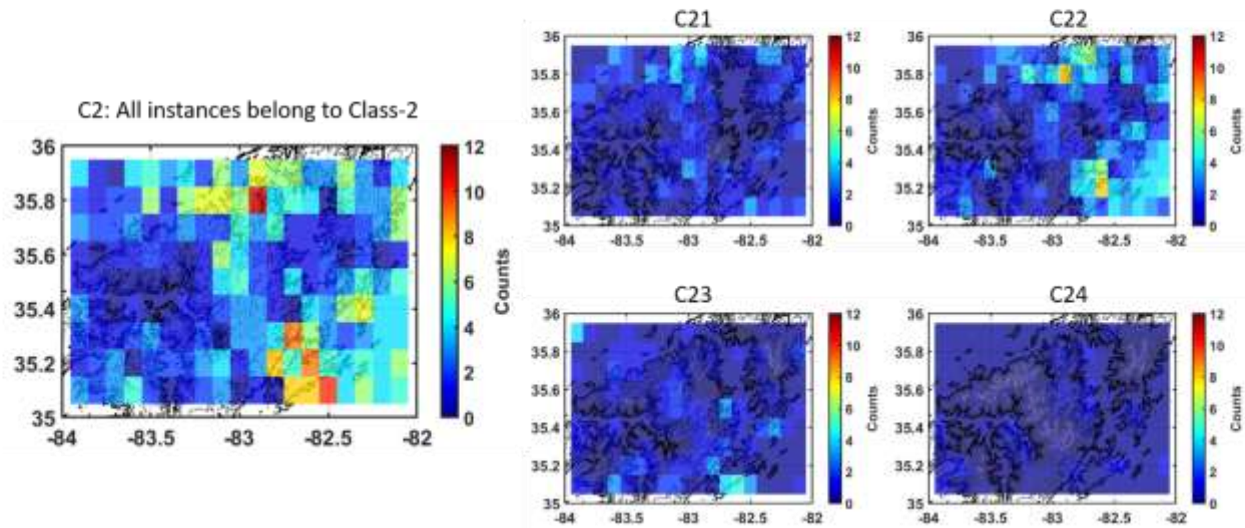

Figure S14. Same as Fig. S13 but for Class-2.

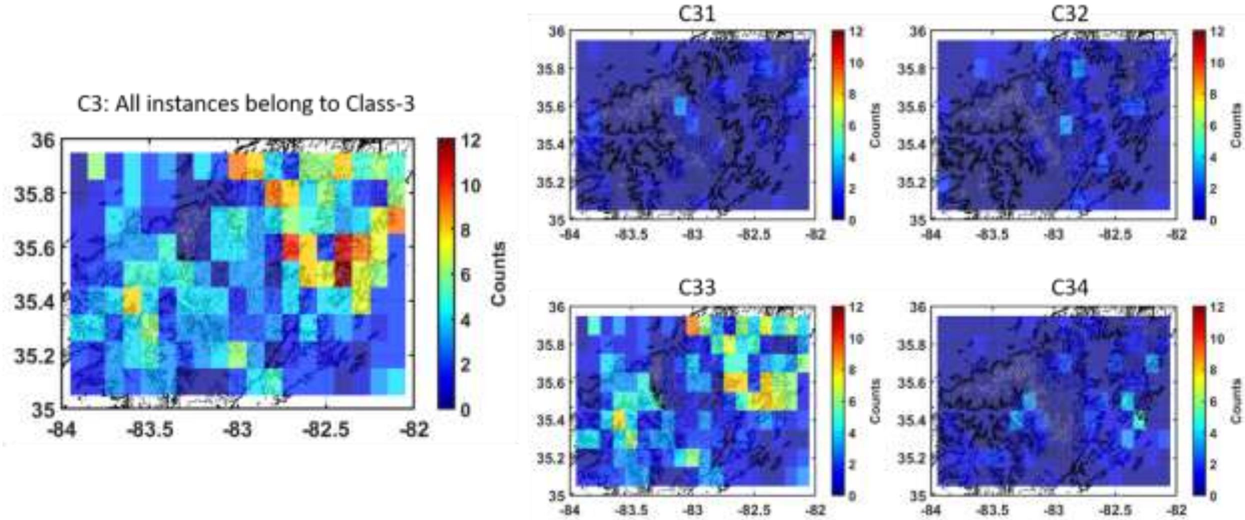

Figure S15. Same as Fig. S13 but for Class-3.

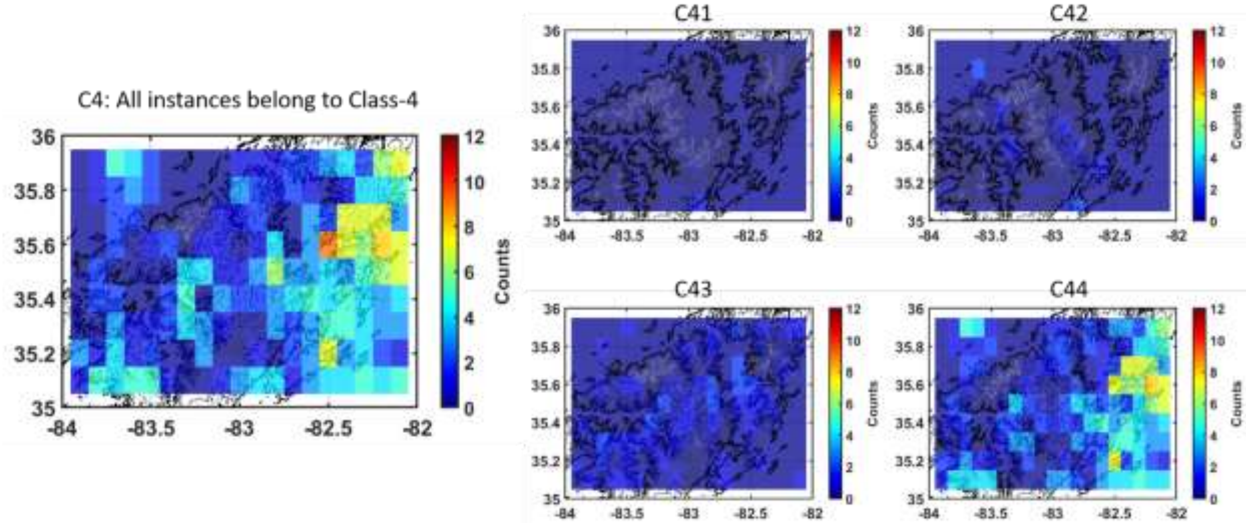

Figure S16. Same as Fig. S13 but for Class-4.

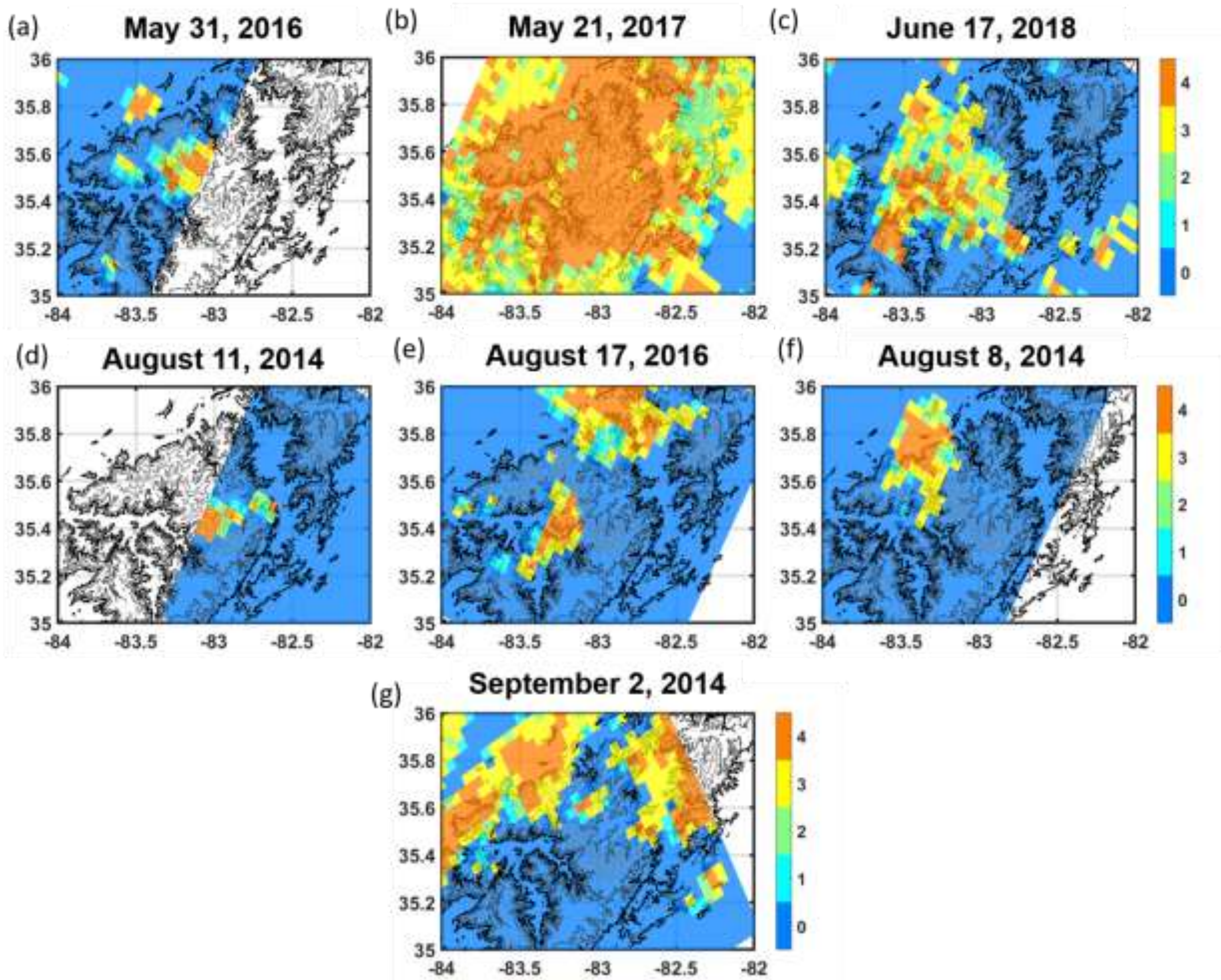

Figure S17. Spatial distribution of predicted classification labels for GPM overpasses listed in Table 7. Legend: 0 - No rain ; X – Rain in Cluster-X, X=1,2,3,4.

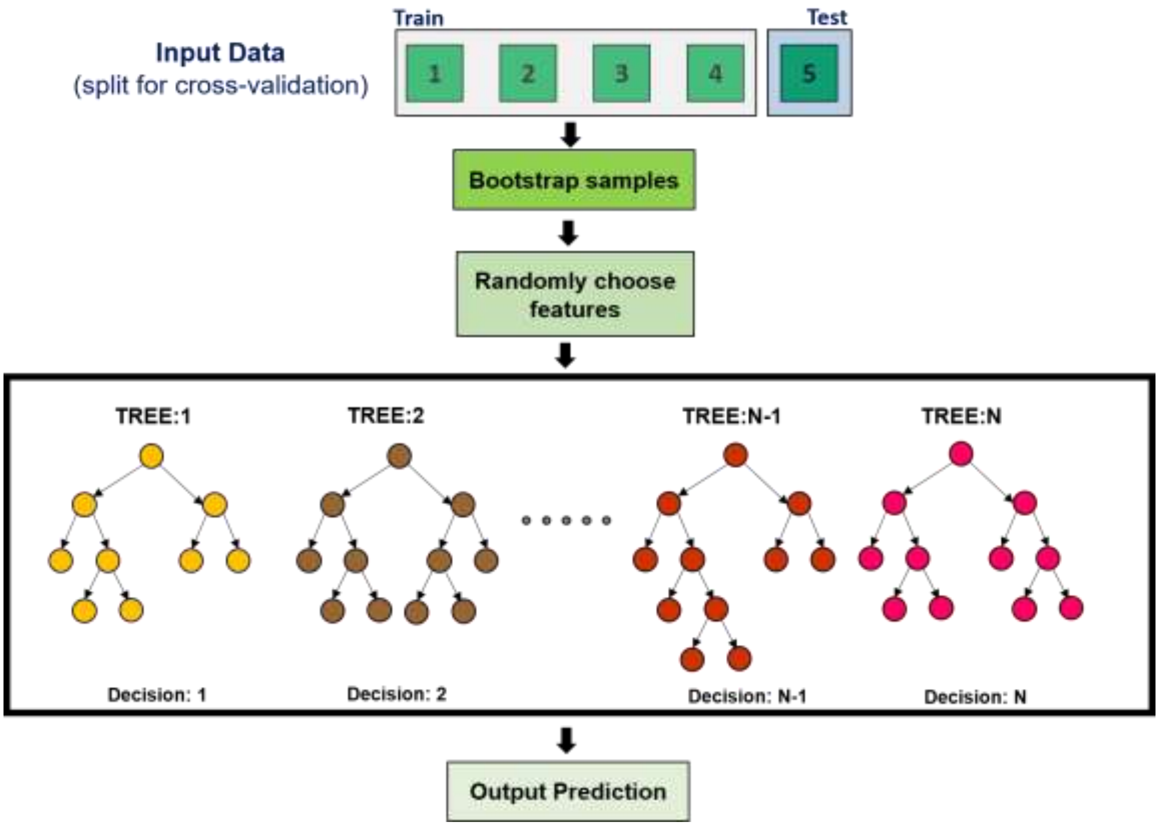

Figure S18. Schematic representation of the Random Forest Classifier (RFC) used for precipitation detection.