Quantifying and Elucidating the Physical Basis of Uncertainty in GPM-DPR

Precipitation in Mountain Regions Using Multi-Frequency Observations and Models

by

Malarvizhi Arulraj

Department of Civil and Environmental Engineering
Duke University

Date:_____
Approved:

_____
Ana P. Barros, Advisor

_____
Nathaniel Chaney

_____
Wenhong Li

_____
Robert Meneghini

_____
Simone Tanelli

Dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor
of Philosophy in the Department of
Civil and Environmental Engineering in the Graduate School
of Duke University

2019

<u>ABSTRACT</u>

Quantifying and Elucidating the Physical Basis of Uncertainty in GPM-DPR

Precipitation in Mountain Regions Using Multi-Frequency Observations and Models

by

Malarvizhi Arulraj

Department of Civil and Environmental Engineering
Duke University

Date:_____

Approved:

_____

Ana P. Barros, Advisor

_____

Nathaniel Chaney

_____

Wenhong Li

_____

Robert Meneghini

_____

Simone Tanelli

An abstract of a dissertation submitted in partial
fulfillment of the requirements for the degree
of Doctor of Philosophy in the Department of
Civil and Environmental Engineering in the Graduate School of
Duke University

2019

Abstract

Quantitative precipitation estimation (QPE) in mountainous regions remains a challenging task owing to its high spatial and temporal variability. Satellite-based radar observations at high resolution have the best potential to capture the spatial patterns of precipitation, but there is high uncertainty in the interpretation of low-level measurements due to ground clutter effects, observing geometry, and sub-grid scale vertical and horizontal heterogeneity of precipitation systems that result from interactions among orographic clouds and propagating storm systems. In the high elevation tropics and in middle mountains everywhere, the landscape is often immersed in multi-layered cloud systems that modify precipitation significantly at low levels in a complex manner depending on time of day and location that is very different from the classical understanding of orographic precipitation enhancement with elevation, and are not easily parameterized or corrected for in QPE algorithms. The overall objective of this proposal is to characterize and elucidate the physical basis of uncertainty in Global Precipitation Measurement Mission (GPM) Dual-frequency Precipitation Radar (DPR) QPE in mountainous regions and develop an improved retrieval framework for orographic precipitation. The following science objectives will be addressed specifically: i) to characterize the dependencies among the spatial and temporal variability of errors in orographic QPE and associated hydrometeorological regimes; ii) to characterize the

vertical structure of radar reflectivity associated with QPE retrieval error and establish a physics-based retrieval model; and iii) to develop an operational framework to integrate DPR observations and Numerical Weather Prediction (NWP) model toward improving the retrieval of orographic QPE. The research hypothesis are two folds: a) satellite QPE errors (false alarms, missed detections, underestimations and overestimations) exhibit a robust spatial and temporal organization that is explained by the spatial and temporal variability in the vertical microstructure of precipitation; and b) current satellite-based QPE algorithms fail because the vertical structure of precipitation cannot be detected. If the vertical structure of precipitation systems can be predicted, then the key microphysical processes can be modelled to improve QPE. The study will be conducted using observations from the Southern Appalachian Mountains (SAM) region including a high-density rain gauge network, Micro Rain Radars (MRR), and Parsivel disdrometers since the launch of GPM as well as Integrated Precipitation and Hydrology Experiment (IPHEx) 2014 data.

This research approach consists of integrating ground-based point measurements from long-term observation networks, fields campaign (IPHEx), multi-satellite data, and modeling studies to develop a physically-based retrieval framework for orographic precipitation. To characterize the dependencies among the spatial and temporal variability of errors in orographic QPE, GPM estimations were evaluated using the ground-based precipitation observations and investigated for the robust organization of

uncertainty. The physics-based retrieval of near-surface rain-rates was demonstrated through explicit modeling of rain shaft microphysical processes constrained by GPM-DPR observations. Finally, a general data-driven operational framework was developed to improve the detection and to predict the vertical structure of precipitation systems by integrating GPM observations with NWP model simulations.

Dedication

To my loving husband, sister, parents and in-laws.

Contents

List of Tables

List of Figures

xvii

xviii

Acknowledgements

1. Introduction

Mountains act as "water towers" enhancing precipitation processes at high elevations by forcing moist air upslope, the classical orographic enhancement of precipitation (e.g. Barros and Lettenmaier 1994), which subsequently is returned via recharge, streamflow and snowmelt to adjacent low-level regions. Indeed, sixty to seventy percent of the global population relies on orographic precipitation for freshwater necessary for irrigation, drinking water, and hydropower generation (Immerzeel et al. 2010; Viviroli et al. 2011). Light stratiform rainfall and fog in mountain regions are also essential for regional freshwater sustainability (Bruijnzeel et al. 2011; Barros 2013), especially during drought, heavy precipitation associated with myriad convective systems from thunderstorms to tropical storms. Thus, understanding orographic precipitation significantly improves studies related to land surface-atmosphere interactions and water cycle including hydrological feedbacks, biogeochemical cycles and regional and global climate changes (Bales et al. 2006; Barros, 2013). However, even a short-term intense precipitation in the orographic region can result in severe natural hazards such as flash floods, avalanches and land-slides impacting livelihood, infrastructure and even global economy (Caine 1980; Sharma and Ganju 2000; Geertsema et al. 2006; Tao and Barros 2013; Tarolli et al. 2013; Eghdami and Barros, 2019). Besides, spatial variations in the orographic precipitation rate are highly correlated with the erosion rate in the mountain regions suggesting that the influence of climate variability in erosion (Burbank et al. 2003; Roe et

al. 2003; Lowman and Barros 2014). These observations illustrate the need to characterize, study and obtain reliable estimates of the orographic precipitation processes to significantly improve climate-related studies, weather, and natural hazard forecasting, prediction of water-borne diseases and agricultural industry.

Theoretically, orographic precipitation refers to an increase in precipitation intensity with the elevation where moist air was adiabatically lifted upwards by the terrain to reach lifting condensation level and further forms clouds and produces rain (Roe 2005; Barros 2013). However, in reality, orographic precipitation is a complex process controlled by multiple factors such as the terrain size, microphysical time-scale of the hydrometeors, dynamics of air-flow encountering the mountains and thermodynamics of air rising over the terrain and is tied to diurnal cycle of the solar forcing at landform scale (Houze 2012; Barros 2013). In addition, the vertical structure and microphysics of the general precipitation types such as convective, frontal or tropical cyclones are significantly altered due to the terrain effects, wind and thermodynamics over the mountain regions (Barros and Letternmaier 1994; Houze 2012). The basic mechanism in which terrain affects the precipitating clouds are upslope and upstream triggering of convection due to shear stress, diurnal forcing triggering the convection, local convection due to the upwind and

lee-side triggering and seeder-feeder interactions of precipitation (Figure 1-1; Houze 2012;

Barros 2013).
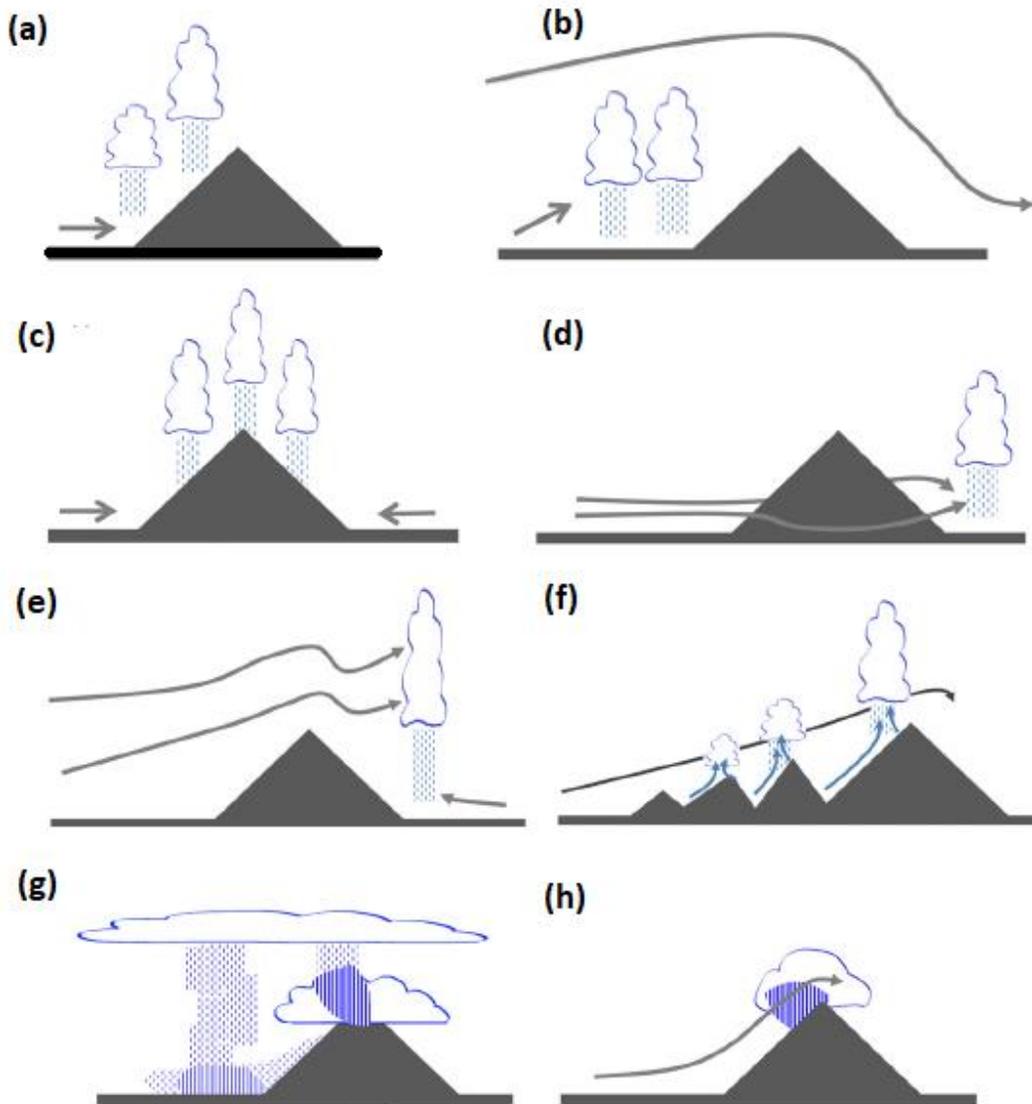
Figure 1-1 Mechanisms of orographic precipitation. (a) upslope triggering of convection; (b) upstream triggering of convection; (c) thermal triggering of convection; (d) leeside triggering of convection; (e) leeside enhancement of convection; (f) organization of upslope and leeside triggering and enhancement of convection by complex terrain; (g) seeder-feeder enhancement of rainfall; and (h) upstream condensation and stationary deep orographic cloud (from Barros 2013).

It is essential to have reliable quantitative precipitation estimates in the mountain regions. For example, Tao and Barros (2013) showed that frequent (3 hourly) satellite-based estimates, even if imperfect, could significantly improve the space-time distribution of precipitation from Numerical Weather Prediction (NWP) models in the Southern Appalachian Mountains (SAM) to increase the lead times for natural hazard forecasting and reduce false alarms. The measurement of orographic precipitation through ground-based networks is a challenging task due to the heterogeneity of the terrain requiring a highly dense observational network to be representative of the resulting spatial variability in precipitation patterns that is difficult to achieve due to remoteness and the difficulty in accessing high elevations sites. For example, the Southern Appalachian region has high heterogeneity in terms of topography and annual precipitation climatology (Figure 1-2). Ground-based point observations of the space-time structure of rainfall in the Great Smoky Mountains in the SAM over the last 10 years show strong spatial gradients with elevation and landform, as well as large temporal variability at diurnal, monthly, seasonal and inter-annual scales. Previous studies in the region (Wilson and Barros 2014, 2015 and 2017) revealed that light rainfall in the SAM contributes up to 50% of the total annual precipitation received in the region. In particular, Wilson and Barros (2014) identified low-level precipitation systems with high-frequency light rainfall between 11:00 and 16:00 LT (i.e. mid-day) in the inner regions of the SAM in all seasons.

(a)

(b)
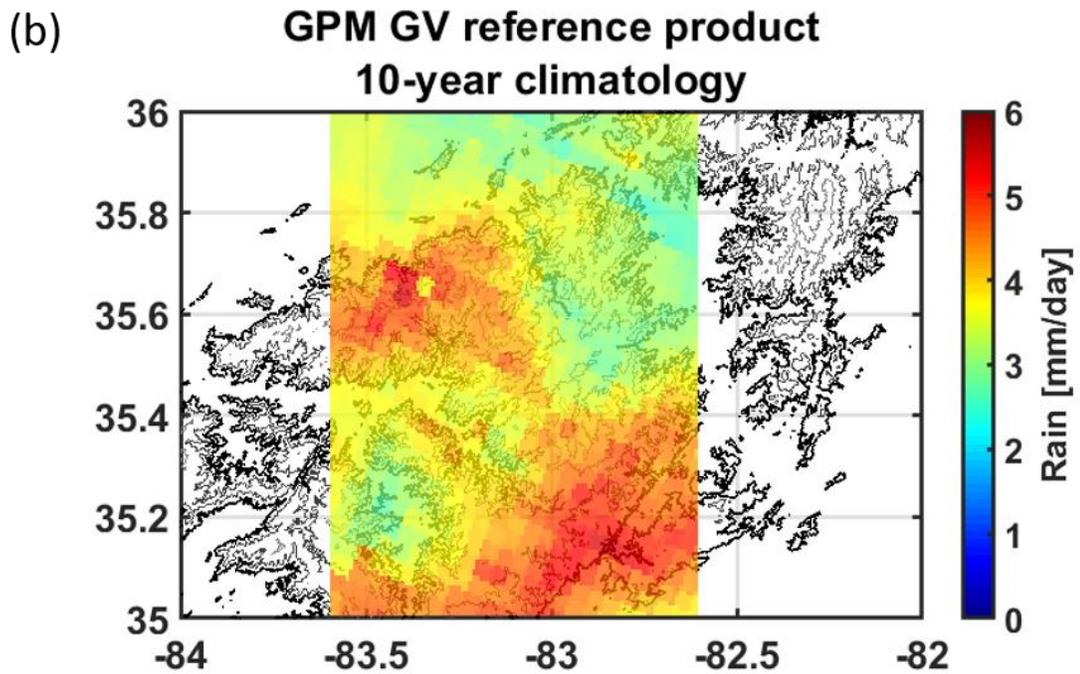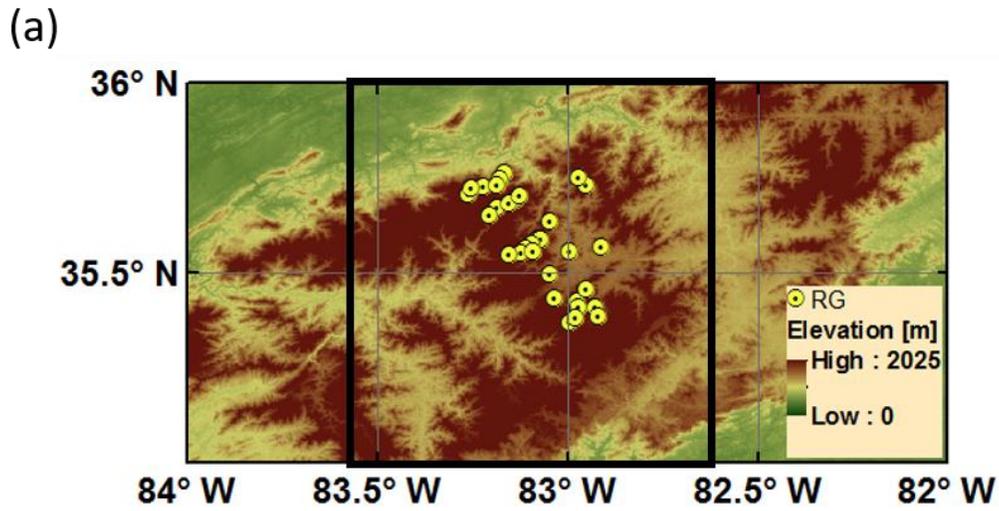
Figure 1-2: a) Topography map of the SAM. (b) Spatial distribution of the 10-year average precipitation climatology from GPM GV reference precipitation product (Liao and Barros 2019).

High heterogeneity between the valley and the ridge precipitation in the SAM is attributed to the spatial and temporal variation in the precipitation microphysics (Prat and Barros 2010; Wilson and Barros 2014, 2015 and 2017). Interestingly, as opposed to the classical orographic enhancement of precipitation with elevation on upwind slopes (e.g. Barros and Lettenmaier 1994), there are many instances in the SAM where mountain valleys receive more precipitation than ridges, up to an order of magnitude compared to the ridges due to the Seeder-Feeder Interactions (SFI; Bergeron 1960; Wilson and Barros, 2014; Duan and Barros 2017) among rainfall and multilayer cloud systems resulting in the reverse-orographic effect identified by Wilson and Barros (2015). SFI refers to the process where the raindrops produced by the high-level clouds (seeders) interact with the low-level clouds and fog (LLCF; feeders) to enhance coalescence and accelerate raindrop growth, and thus surface precipitation at low levels. The morning and mid-afternoon precipitation process are due to the interaction between stratiform and low-level orographic clouds and advection fog while nighttime and early morning SFI is governed by interactions among passing stratiform systems, convective clouds and radiation fog (Wilson and Barros 2014). A synthesis of these processes is illustrated in Figure 1-3.
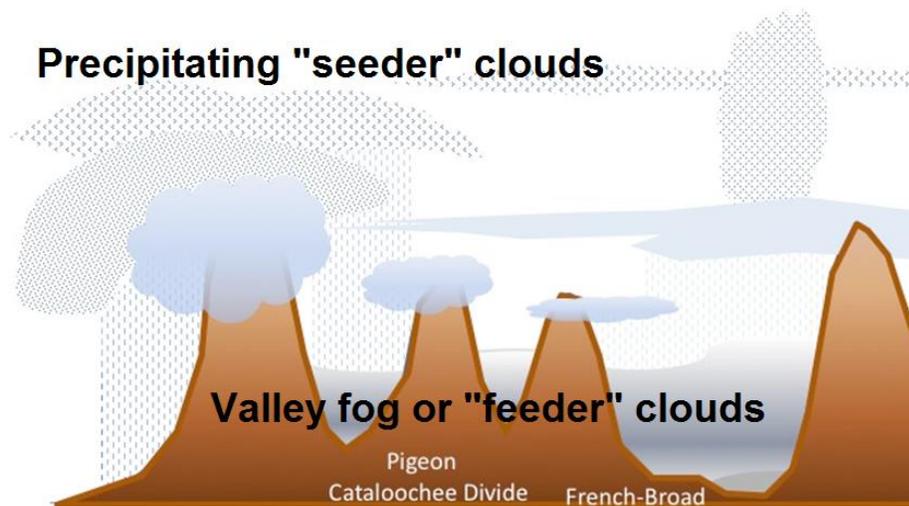
Figure 1-3: Conceptual representation of seeder-feeder interactions over complex terrain (after Wilson and Barros, 2017).

The hydrometeorological cycle of the small and middle mountains (Bergeron 1960; Storebo 1976; Hill et al. 1981; Passarelli and Boehme 1983; Robichaud and Austin 1987; Purdy et al. 2005; Wilson and Barros 2014), tropical cloud montane forests (Bruijnzeel 2001) and coastal mountain regions (Torregrosa et al. 2014) are dominated by LLCF and SFI processes. These hydrometeorological regimes are dominated by the high frequency of precipitation events and, seasonal and temporal distribution of low-level clouds and fog which is favorable to SFI with higher level of precipitating systems. Comparison of observations from fog gauges and rain-gauges in the SAM suggests that the fog plays a major role in the morning while rain dominates in the late afternoon (Wilson and Barros 2015). Long-term climatological analysis of MODIS cloud top height in the SAM illustrates that the low-level clouds are spatially, temporally and seasonally organized along the ridges and foothills of the SAM (Duan and Barros 2017). Numerical simulations of SFI

cases suggest that collision-coalescence process dominates for drops with diameter between 0.1 mm and 1 mm while break-up is the dominant process for the non-SFI precipitation events (Wilson and Barros 2014). These simulation results were further validated by observations from collocated Meteorological Particle Spectrometer, Parsivel disdrometer and Micro Rain Radar (Duan and Barros 2017).

Satellite remote sensing and the recent introduction of multi-frequency radar-based precipitation products have the potential to improve the estimation of orographic precipitation. Radars are active sensors that transmit and receive monochromatic radiation. The transmitted radiation undergoes scattering and absorption as it interacts with hydrometeors in the atmosphere. Typically, single scattering, a first-order process whereby the incident photons are scattered only once, is assumed in operational radar algorithms (Marzano et al. 2003). However, the incident photons can undergo multiple scattering based on the incoming wavelength of the photons, horizontal and vertical distribution of hydrometeors, and the optical properties of the hydrometeors. Multiple scattering (MS) becomes significant at higher frequencies, larger particle sizes, larger antenna footprints, and with asymmetry in the scattering phase function that describes volume scattering (Battaglia et al. 2005).

Satellite-based precipitation radars operate at microwave frequencies and capture the 3D structure of precipitation storms. Reflected and backscattered electromagnetic signals from the vertical distribution of hydrometeors are used to estimate the vertical profiles of

precipitation rate. When the Rayleigh approximation holds, the reflectivity factor Z at height z is related to the diameter (assuming spherical shapes) and concentration of atmospheric hydrometeors at the same height (Bringi and Chandrasekar, 2004):

$$Z = \int D^6 N(D) dD \tag{1.1}$$

where N(D) is the drop-size distribution and D is the diameter size. Rainfall rates are related to Z through a power law in some retrieval algorithms. The accuracy of satellite-based precipitation estimates in mountainous areas is affected by multiple sources of error including bright-band (BB) and dark-band (DB) effects, non-uniform beam-filling (NUBF) artifacts that are exacerbated by the complex spatial variability of rainfall systems, complex multiple-scattering processes and ground-clutter artifacts. Whereas these artifacts are not specific to mountainous regions, the heterogeneous vertical organization of hydrometeors depending on storm systems and topography, and the organization of precipitation processes on the terrain are unique to mountainous regions. A brief description of the physical basis of error sources is presented below.
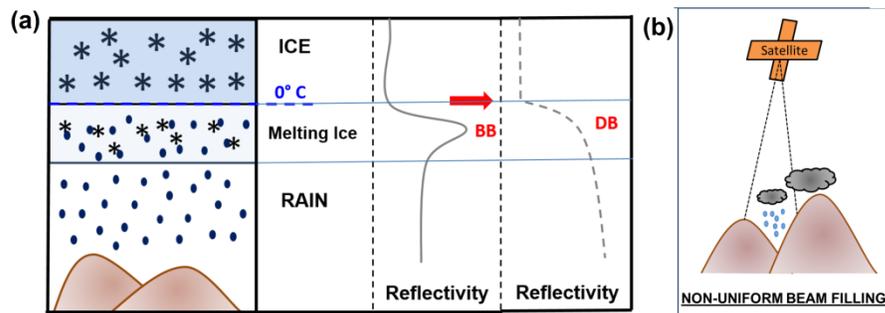
Figure 1-4: Conceptual Representation of (a) Bright-band and dark-band signatures in reflectivity profiles and (b) non-uniform beam filling effect.

The melting of ice and snow at 0-degree isotherm in the atmosphere leads to the formation of a mixed-phase layer of stratiform precipitation (Figure 1-4a). Changes in size due to coalescence of liquid and ice hydrometeors and or partially melted particles and changes in the dielectric constant due to phase changes lead to an increase in the observed reflectivity factor, the BB effect (Smyth and Illingworth, 1998). The BB is absent in convective precipitation as strong updrafts and warm temperatures keep the ice particles aloft, whereas, in the case of large frozen hydrometeors like graupel and hail, fall velocities are very high and may only melt at the ground, or not melt at all. The absence of a BB is one of the characteristics of the radar return that is used to distinguish convective from stratiform precipitation (e.g., Awaka, 1997). Underestimation and False Alarms (FA) errors in the case of stratiform precipitation are linked to a decrease in the gradient of reflectivity profiles near-surface (Duan et al. 2015). W-Band radars (95 GHz) observe a "dark-band" effect near 0-degree isotherm, where there is a sudden increase in the reflectivity profiles at the melting layer without being followed by a decrease in the reflectivity values (i.e. absence of BB), because the average attenuation in the melting layer is comparable with the attenuation in the rain layer underneath (Kollias and Albrecht, 2015; KA15, hereafter). The reflectivity factor Z is directly proportional to the sixth moment of the drop diameter (e.g. Eq. 1-1) in the Rayleigh [wavelength is larger than hydrometeor size] scattering regime and exhibits an oscillating behavior for drop-sizes larger than 1-mm in Mie [wavelength is smaller than hydrometeor size] scattering regime

(Kollias et al., 2002). For lower frequency radar, the assumption of Rayleigh scattering behavior holds well for large drop-sizes and the reflectivity shows sensitivity in the melting layer. KA15 noted that the total backscattering coefficient does not show variability for drop diameters greater than 1 mm in the W-band (~ 3.2 mm wavelength) explaining the BB absence. KA15 also points out though that in the case of very small drop sizes (less than 0.8 mm), the DB vanishes aloft, and a BB signature is present at W-band. The emphasis of this dissertation is on satellite-based Ku- and Ka-Band precipitation radar. In GPM Ku- (13.6 GHz; wavelength - 22 cm) and Ka-Band (35.5 GHz; wavelength ~ 8 cm) precipitation radar, non-Rayleigh effects are expected to be significant for Ka-band compared to Ku- band.

Non-Uniform Beam Filling (NUBF) refers to the integration of the radar signal over a volume that is not uniformly filled with hydrometeors due to the heterogeneity in the vertical and horizontal structure of precipitation systems within radar beam, including clear-sky and non-rainy clouds (Figure 1-4b). The effect of NUBF on the rain-rate retrieval depends on the methods such as the attenuation correction. Thus, NUBF usually leads to estimation errors and these errors generally result in negative bias (underestimation of rain-rate). As the distance and the antenna beam-width increases, the heterogeneity increases, and thus there can be significant underestimation in the precipitation retrievals generally (Durden et al. 1997), and low-level and small-scale local orographic precipitation systems (Duan et al. 2015).

MS effects are dominant in mesoscale convective precipitation systems leading to the overestimation of precipitation rate (Marzano et al. 2003 and Bouniol et al. 2008). Monte-Carlo simulations suggest that the CloudSat (satellite-based W-Band radar) retrievals of stratiform precipitation with intensities greater than 5 mm/h suffer from 30 to 50% overestimation (Matrosov et al. 2008). The occurrence of MS is rare and expected to be dominant in higher frequency radars (Battaglia et al. 2010).

1.1 Challenges and Motivation

Satellite-based precipitation estimates suffer from different types of detection and estimation errors: false alarm (FA) and missed detection (MD) are the detection errors while underestimation (UND) and overestimation (OVR) are the estimation errors. Previous work documenting the climatology of rainfall estimation errors from TRMM v6 and v7 precipitation estimates in the SAM reported larger errors for shallow precipitation (Prat and Barros, 2010; Duan et al., 2015). Duan et al (2015) classified the entire region into three parts: eastern, western and inner ridge to perform the error analysis with the help of data from a ground-based high-density rain gauge network. It was revealed that the precipitation uncertainties follow robust spatial and temporal (seasonality and diurnal cycle) organizations of FAs, MDs, UNDs, and OVRs. The study also suggested that the uncertainties are linked to issues in the interpretation of low-level measurements in complex terrain by satellites owing to ground-clutter effects, observing geometry, and non-uniform beam filling effects that result from interactions among fog and orographic

clouds and propagating storm systems. A similar error analysis conducted on tropical mountains suggests high uncertainty in the satellite-based precipitation estimates. For example, TRMM-v6 precipitation estimates in the central Andes highlights a slight underestimation of precipitation and increased false alarm rates (Scheel et al., 2011). The precipitation on the windward side of the Western Ghats is underestimated by about 50% due to TRMM's failure to capture short-lived convective precipitation systems, while the precipitation in the rain-shadow region is over-estimated (Nair et al. 2009).

The GPM-DPR is more sensitive compared to the TRMM-PR especially in the measurement of snow and light precipitation. The evaluation of GPM-DPR profiles obtained over Africa highlighted that the multiple scattering phenomena were observed on Ka-Band reflectivity profiles greater than 30 dBZ (high-density ice particles) (Battaglia et al., 2015). In GPM DPR, for extreme deep convective events (which is rare) such as a tornadic supercell, multiple scattering effects are observed in both Ka- and Ku- Band precipitation radars (Battaglia et al. 2016). Kirstetter et al., (2013) evaluated the TRMM-PR v6 and v7 products using the NOAA/NSSL Ground Radar-Based National Mosaic QPE (Q2 precipitation products). The Q2 precipitation products are derived from various ground-based radar networks with a high spatial resolution (0.01°; approximately 1 km). This study suggests overestimation of lighter rain rates are due to the convective Z-R relationship and underestimation of higher rain-rates is due to the NUBF correction.

Most of these problems arise due to the operating frequency, antenna orientation, and beam-width, viewing angle, surface elevation, and topography that affect the measurements near the surface. These effects can be a major issue in the retrieval of precipitation in mountainous regions and developing a correction algorithm addressing these issues is a complex task as the problem is highly stochastic (Germann et al., 2006).

1.2 Research Objectives and Hypothesis

The <u>overarching objective</u> of this proposal is to characterize and elucidate the physical basis of uncertainty in GPM-DPR QPE in mountainous regions and develop an improved retrieval framework for orographic precipitation with a focus on shallow precipitation systems (warm rain processes). The <u>research</u> <u>hypotheses</u> are of two folds: a) satellite-based QPE errors (false alarms, missed detections, underestimations and overestimations) exhibit a robust spatial and temporal organization that is explained by the spatial and temporal variability in the vertical microstructure of precipitation; and b) current satellite-based QPE algorithms fail because the vertical structure of precipitation cannot be detected. If the vertical structure of precipitation systems can be predicted, then the key microphysical processes can be modelled to improve QPE. The study will be conducted using observations from the Southern Appalachian Mountains (SAM) region including a high-density rain gauge network, fog gauges, Micro Rain Radars (MRR), Parsivel disdrometers and Meteorological Particle Spectrometer (MPS) since the launch of GPM as well as IPHEx 2014 data. This research will be transferable to other regions

around the world such as small and middle mountains, cloud montane forests and coastal mountain regions. The following <u>research objectives</u> will be addressed specifically: i) to characterize the dependencies among the spatial and temporal variability of errors in orographic QPE and associated hydrometeorological regimes; ii) to characterize the vertical structure of radar reflectivity and establish a physics-based precipitation retrieval model; and iii) to develop an operational framework to integrate DPR observations and NWP model toward improving the prediction of the vertical structure of orographic precipitation systems.

The following science questions (SQs) will be addressed specifically:

1) What is the detection and estimation uncertainty associated with the satellite-based precipitation estimates in the SAM? How can the observations from multi-frequency radar improve the detection and classification of shallow precipitation?

2) How does the surface clutter influence the near-surface reflectivity profiles for different hydrometeorological regimes in satellite-based observations?

3) How can the vertical structure of the precipitation systems be predicted?

1.3 Approach and Outline

This research addresses the physical basis of uncertainty involved in the satellite-based precipitation estimates, especially in mountain regions. The research approach consists of integrating the ground-based observations such as the long-term tipping

bucket rain-gauge network in the Great Smoky Mountains National Park, collocated MRRS and Parsivel disdrometer, the satellite-based observations of reflectivity and precipitation rate profiles from the GPM-DPR Ka- and Ku-Band precipitation radar with the Numerical Weather Prediction model results to develop a physical-based framework to address the uncertainty in QPEs.

To address the first part of SQ1, observations from ground-based instruments such as the tipping bucket rain-gauges, Parsivel disdrometers, Micro Rain Radars (MRRs) and Stage IV gridded products are used to characterize the detection and estimation errors in GPM Ku-PR estimations and to elucidate the physical basis of these errors. The error analyses of different versions of GPM Ku-PR precipitation estimations with different ground-based observations are presented in Chapters 2, 4 and 5.

The second part of SQ1 is to develop an algorithm to improve the detection of shallow precipitation systems by using observations from the collocated multi-frequency radar reflectivity profiles. The ground-observations from Maggie Valley, North Carolina obtained during IPHEx-IOP (1st May 2014 to 15th June 2014) are used to develop the shallow rain detection and classification (SRDC) algorithm. The algorithm was further validated in other ground-based locations such as Finland (Hyytiala; ARM-TMP) and Southern Great Plains (Oklahoma, ARM-SGP), model simulation of GPM-DPR and CloudSat-CPR near the SAM and the GPM-DPR and CloudSat-CPR overpass in Borneo near the Kapuas Mountains.

To address SQ-2, the reflectivity profiles from the GPM-DPR are investigated for the impact of ground-clutter in precipitation estimation over the SAM region. The ground-clutter effects are investigated with respect to different viewing angle and the impact on the estimation of drop size distributions (DSD) and rain-rate are evaluated using ground-observations such as the rain-gauges and Parsivel disdrometers. The evaluation suggests that the ground-clutter contamination leads to the underestimation of low-level enhanced events and missed detection of shallow precipitation events. Ground-clutter leads to severe retrieval errors in high elevation and during winter precipitation events. To address the lack of surface observations due to ground-clutter, a new physically-based framework was developed by coupling GPM observations with ground-observations to simulate the near-surface precipitation profiles.

SQ-3 deals with the development of a physical framework by integrating observations and model simulations to improve the detection and estimation of orographic precipitation systems. A general, data-driven deep learning based model was developed to detect and classify the precipitation systems using coupled GPM and NWP simulations. The NWP simulations assist in improving the precipitation detection while the classification algorithm assigns the class based on the precipitation vertical structure.

The structure of this dissertation is organized as follows: Chapter 2 provides a brief description highlighting the importance and challenges in the remote sensing of orographic precipitation. This chapter is currently accepted for publication as a book

chapter on Satellite Precipitation Measurement (Barros and Arulraj, 2020). Chapter 3

describes the shallow precipitation detection and classification algorithm developed using

collocated multi-frequency radars to improve the detection of shallow precipitation. This

chapter was published in the *Journal of Atmospheric and Oceanic Technology* (Arulraj and

Barros, 2017).   Chapter 4 addresses the detection and estimation errors in the GPM DPR

precipitation estimates and proposes a new physically-based framework to model the

impact of low-level clouds and fog using GPM and ground observations. This chapter was

published in *Remote Sensing of Environment* (Arulraj and Barros, 2019). Chapter 5 aims to

couple the observations from GPM and High Resolution Rapid Refresh (HRRR)

simulations to improve the detection rate and successfully estimate the vertical structure

of the precipitation systems. This chapter will be submitted to *Remote Sensing of*

Environment. Chapter 6 summarizes the major findings of the dissertation and provide

future recommendations.

2. Remote Sensing of Orographic Precipitation

Note this chapter appeared as reference (Barros and Arulraj, 2020).

2.1 Introduction

Orographic precipitation is the primary source of fresh water for more than half of the world's population and it provides more than 40% of all water used for irrigation in adjacent lowlands going up to 90% in semi-arid regions impacting up to 70% of the world's population (Garrido and Dinar 2009; Nellemann et al. 2009). Further, because of the preponderance of rain-fed agriculture in regions of complex terrain such as the Andes and the Himalayas, inter-annual food security of mountain populations is tightly linked to precipitation, and thus water resilience of downstream low-lying landscapes in the face of global change (FAO 2014 and 2016; Falkenmark and Rockström 2010). Heavy rainfall triggers flash floods and landslides that can destroy crops and erode away high-quality arable soils, causing roads and infrastructure to fail, thereby disrupting transportation tied to a myriad of socio-economic impacts. In the US, orographic precipitation (rain and snow) is the key water resource providing 70% of all water resources in the West , it is the key source of runoff for the dams operated by Tennessee Valley Authority (TVA), and essential to regional-scale groundwater recharge east of Appalachian Divide (e.g. Barros et al. 2017).

Mountain landscapes can be described as altitudinal scaffolds of topographically delineated collectors of precipitation (i.e. watersheds) interlinked through a system of

converging channel connectors (the river network). Because of erosional processes, the very spatial structure of mountain landscapes reflect the co-evolution of geological mountain building processes and regional climate, in particular precipitation patterns (e.g. Lowman and Barros 2014; Barros et al. 2006). Depending on the latitude, season, and topography, and environmental conditions, orographic precipitation can be classified as stratiform, convective, or stratiform with embedded convection (Houze 2012). Whereas light stratiform rainfall and fog in mountains environments are essential for regional freshwater sustainability (Bruijnzeel et al. 2011; Barros 2013), especially during drought, heavy precipitation associated with myriad convective systems from thunderstorms to tropical storms is not only a key source of freshwater but it is also the main driver of natural hazards, flashfloods, landslides and other natural hazards. The value of orographic precipitation measurement is strongly tied therefore to its hydrological utility at the desired (application dependent) spatial and temporal scales: for example, flood (days) or flashflood (< 6 hours) and landslides (< 1 hour) forecasting versus water budget studies (inter-annual to decadal, e.g. Lowman and Barros 2016; Lowman et al. 2018) with or without (sub-seasonal to seasonal) groundwater recharge and transboundary exchanges (e.g. Tao and Barros 2010; Tao and Barros 2013, 2014; Tao et al. 2016).

Whereas a general understanding of the classical mechanisms of orographic precipitation enhancement by which topography modifies the advection of moist air masses, modulates cloud development and impacts precipitation intensity and accumulation patterns is well

established (e.g. Barros and Lettenmaier 1994; Barros 2013), the goal in orographic precipitation measurement and prediction is to quantify precipitation everywhere at any time (when, how fast, how long, and where in the landscape) given the high spatial and temporal variability of nonlinear land–atmosphere interactions that dynamically redistribute precipitation from one watershed to another (e.g. Barros 2013). Because of remoteness and access, mountainous regions remain among the least observed regions of the planet, and even where observations are available long-term science-grade observations are rare due to the difficulties of maintaining instruments and collecting data (Barros and Lettenmaier 1994; Viviroli et al. 2011; Barros 2013).

Transformative advancements in precipitation science and precipitation measurement have been possible under the auspices of the TRMM and GPM missions over the last twenty years (Skofronick-Jackson et al. 2017). This is well illustrated in Fig. 2-1 that shows remarkable improvement in the precipitation estimation from TRMM 3B42 (gray, Huffman et al. 2007) to IMERG (blue, Huffman et al. 2015) along the rain-gauge transect maintained by Duke University in the Eastern Andes (Barros 2013), and which cannot be explained simply by improved spatial resolution in IMERG compared to 3B42. The TRMM Precipitation Radar (PR) made possible unprecedented and systematic monitoring of precipitation in Middle and Low Mountains probing into the inner regions of complex terrain to map the vertical reflectivity profiles of orographic precipitation systems.
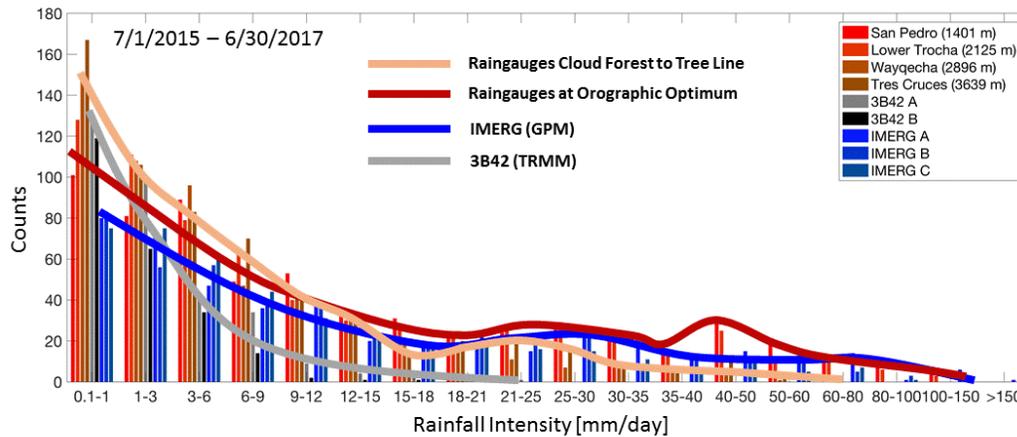
Figure 2-1: Histogram of daily rainfall observed along an altitudinal rain-gauge transect on the eastern slopes of the Peruvian Andes (Barros, 2013). The red (tropical montane forest) and orange (cloud forest) lines represent the rainfall envelope for the two main ecosystems.

The measurement of orographic precipitation through ground-based networks is a challenging task due to the heterogeneity of the terrain requiring a highly dense observational network in order to be representative of the resulting spatial variability in precipitation patterns that is difficult to achieve due to remoteness and the inaccessibility of high elevations sites. Ground-based point observations of the space-time structure of rainfall in the Himalayas (Barros et al. 2000) and in Southern Appalachian Mountains (SAM) show strong spatial gradients with elevation and landform, as well as large temporal variability at diurnal, monthly, seasonal and inter-annual scales (Wilson and Barros 2014 and 2015). In the SAM, light rainfall in the SAM contributes up to 50% of the total annual precipitation received in the region. In particular, Wilson and Barros (2014)

identified low-level precipitation systems with high-frequency light rainfall between 11:00 and 16:00 LT (i.e. mid-day) in the inner regions of the SAM in all seasons.

High heterogeneity between the valley and the ridge precipitation in the SAM is attributed to the spatial and temporal variation in the precipitation microphysics (Prat and Barros 2010a; Wilson and Barros 2014, 2015 and 2017; Duan and Barros, 2017). Interestingly, as opposed to the classical orographic enhancement of precipitation with elevation on upwind slopes (e.g. Barros and Lettenmaier 1994), there are many instances in the SAM where mountain valleys receive more precipitation than ridges, up to an order of magnitude compared to the ridges due to the Seeder-Feeder Interactions (SFI; Wilson and Barros 2014) among rainfall and multilayer cloud systems resulting in the reverse-orographic effect identified by Wilson and Barros (2015). SFI refers to the process where the raindrops produced by the high-level clouds (seeders) interact with the low-level clouds and fog (feeders) to enhance coalescence and accelerate raindrop growth, and thus surface precipitation at low-levels. The mid-afternoon precipitation process is due to the interaction between stratiform and low-level orographic clouds and advection fog while nighttime and early morning SFI is governed by interactions among passing stratiform systems, convective clouds and radiation fog (Wilson and Barros 2014). A synthesis of these processes is presented in Fig. 2-2.

Besides the spatial and temporal variability linked to landform and weather, aerosol-cloud-precipitation interactions can be another important source of variability.

Specifically, changes in cloud condensation nuclei (CCN) activation (Shrestha et al. 2010, 2012) behavior on orographic precipitation processes (Barros et al. 2018; Duan and Barros 2019) can result in significant spatial shifts in cumulative rainfall distributions impacting flashflood forecasting, water resources management, and hydropower production.
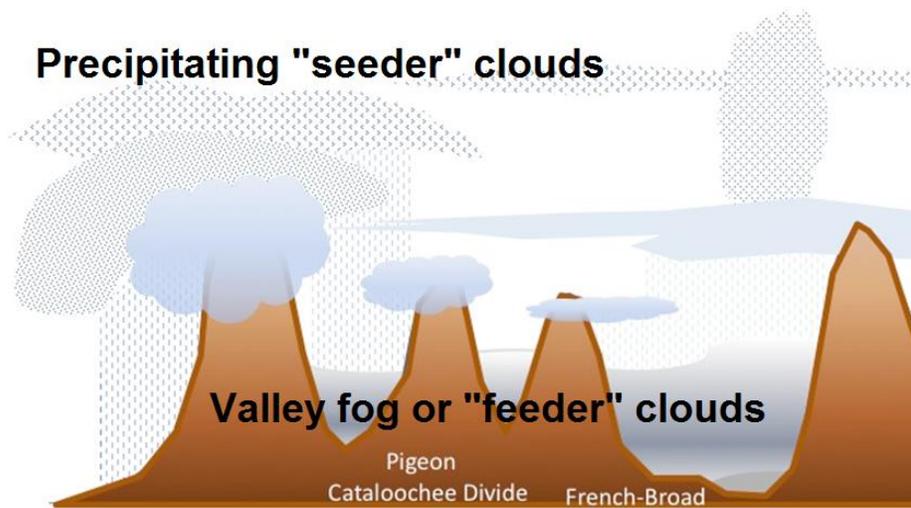
Figure 2-2: Conceptual representation of orographic precipitation mechanisms in the Southern Appalachian Mountains (After Wilson and Barros, 2017).

2.2 Orographic Precipitation Measurement

Satellite remote sensing and the recent introduction of multi-frequency radar based precipitation products have the potential to improve the estimation of orographic precipitation. Radars are active sensors that transmit and receive monochromatic radiation. The transmitted radiation undergoes scattering and absorption as it interacts with hydrometeors and other gaseous species and particles in the atmosphere. Notionally, radar retrieval of precipitation consists of estimating the size distribution

N(D) (spectrum of number of particles N as a function of equivalent diameter D) of the hydrometeors from backscatter measurements which can be used subsequently used to derive bulk rainfall properties such as rainfall volume and intensity. Because of complex microphysics, N(D) can change significantly in the vertical and in time in the atmosphere (see Testik and Barros, 2007), and thus a key challenge in radar measurements is to infer the profiles of rainfall from the instantaneous measurements of backscatter signal with coarse spatial resolution (in km scale).

Specific details on the operational radar retrieval algorithm, and how it approaches the different sources of error are provided by Iguchi et al. (2017) specifically for GPM-DPR measurements. The ground-clutter effect is illustrated in Fig. 2-3. The high reflectivity measurements (red band) in the right panel are contaminated by ground-clutter and cannot be used to estimate rainfall. This results in eliminating near-surface reflectivity measurements (low level rainfall) depending on the incidence angle of the radar unless additional constraints are introduced by using multiple sensors (Arulraj and Barros 2017). The low-level vertical structure of reflectivity from the MRR (range resolution is 100 m) in the left panel is indicative of missed rainfall when ground-clutter corrections are applied to the satellite-based measurements.
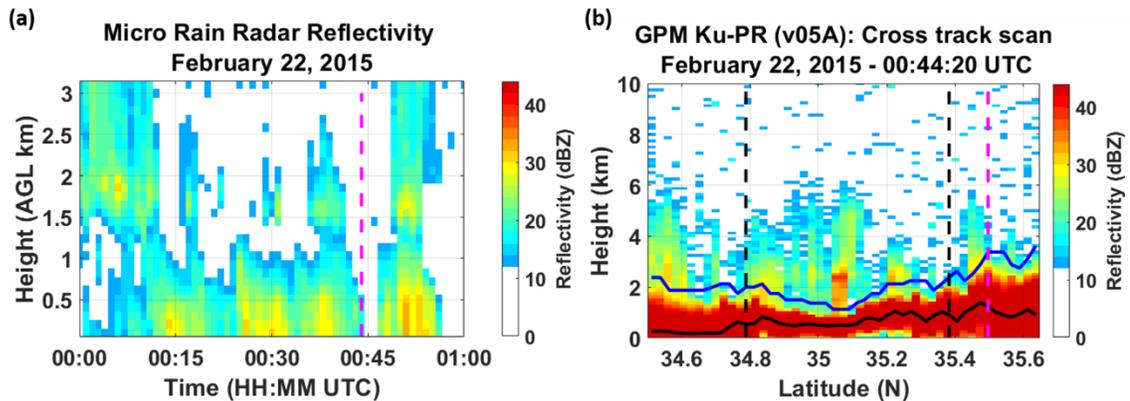
Figure 2-3: Example of ground-clutter effects on radar measurements of low level rainfall. (a) Vertical structure of MRR reflectivity above ground level (AGL) with GPM overpass marked by the pink dashed line in right panel. (b) Reflectivity cross-section over the Southern Appalachian Mountains. Black dashed lines indicate the boundaries of near-nadir scan (angle < 8.5 degrees). Pink line marks the position of the MRR where rainfall was missed by the GPM DPR algorithm. The terrain is marked in black line and the minimum no-clutter bin height is shown in blue color.

Note the intermittency of MRR fine reflectivity structure that illustrates the question of temporal representativeness of satellite measurements (instantaneous) and spatial representativeness of ground-based measurements (point scale) that must be reconciled in the evaluation of precipitation products (Prat and Barros 2010b; Duan et al. 2015). Attempts to address GC by "filling" the reflectivity profiles below the height at which no GC effects are detected often do not work either because of underestimation of SFI and cloud layering but also because they increase the detection and estimation errors. Further, it is important to note that rainfall detection and measurement sensitivity depend strongly on the geometry of the measurement proper as shown in Fig. 2-4 as a function of satellite orbit (descending and ascending overpasses) and viewing angle. Note that the number of

instances in each zenith angle class (Fig. 2-4b) is limited which led to high FARs in the at

near-nadir zenith angle.
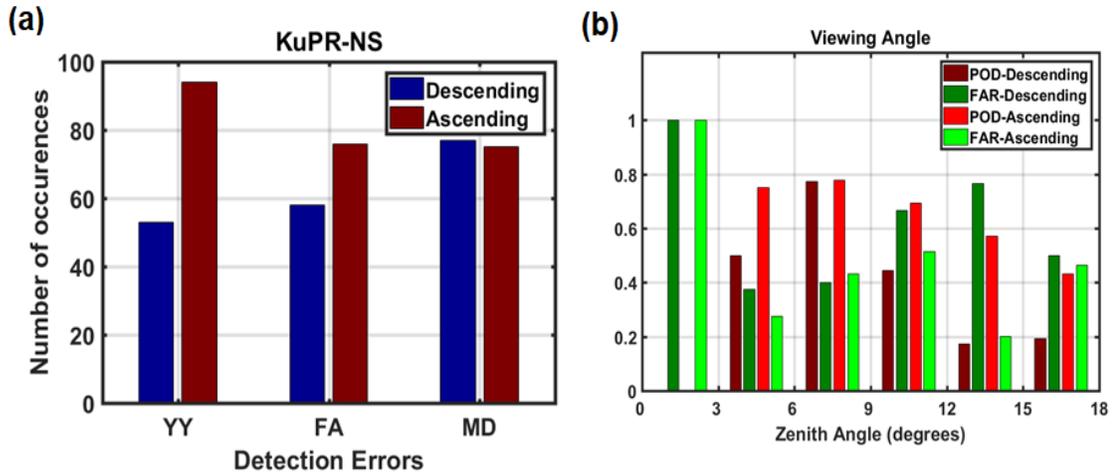
Figure 2-4: Dependence of detection errors to the satellite orientation.
Detection errors varying with **(a)** satellite geometry and **(b)** local viewing angle. Near-surface precipitation estimates of Level 2, version 05A products of GPM Ku-PR compared with rain-gauges that lie within 2.5 km of the center of DPR pixels. Time period considered for analysis: March 2014 to May 2017. Note- YY is when both GPM and RG detects precipitation. FA is false alarm when GPM detects precipitation while RG did not detect any precipitation. MD is missed detection where RG detects rain and GPM misses the detection.

This also illustrates the potential for retrieval ambiguity as complex precipitation processes appear different depending on measurement geometry, which explains why simple calibration and, or optimization of retrieval algorithms with a physical basis does not result in QPE improvements (e.g. compare Prat and Barros 2010b and Duan et al. 2015). Indeed, inspection of Fig. 2-1 shows that the IMERG product severely underestimates the frequency of light rainfall days from the cloud forest up to tree line corresponding to about 50% of annual rainfall at high elevations in the Andes, and it underestimates moderate rainfall at the orographic optimum at lower elevation during the monsoon, which combined with missed detection of light rainfall amounts to 2.5 m of precipitation that is approximately 40% of annual rainfall in the rainforest altitudinal band. Deep convection forms at the foothills of the Andes and in the Altiplano and occasionally at nighttime and early morning at the confluence of river networks (Giovanettone and Barros, 2009). Most moderate rainfall events on the Andean orographic envelope (as in the Himalayas) are stratiform with embedded shallow convection (< 5km AMSL), and thus contamination of radar measurements in the lower 2 km on steep altitudinal gradients and complex terrain geometry is one major source of error. Whereas only GC is specific to mountainous regions, the heterogeneous vertical organization of hydrometeors depending on storm systems and topography, and the organization of precipitation processes on the terrain are unique to mountainous regions (e.g. Fig. 2-2).

2.3 Ground-Validation

A ground-validation (GV) strategy relying on intense ground-based data collection of precipitation accumulations and precipitation structure during limited-duration field campaigns such as the Integrated Precipitation and Hydrology Experiment (IPHEx; Barros et al. 2014) and the Monsoon Himalaya Orographic Precipitation Experiment (MHOPrEx; Barros and Lang, 2003) in the Central Himalayas including the deployment of long-term science-grade observing systems (e.g. Barros et al. 2000; Prat and Barros, 2010a) enables scientific data analysis and discovery to inform detailed evaluation of precipitation retrievals and GV synthesis in the past.
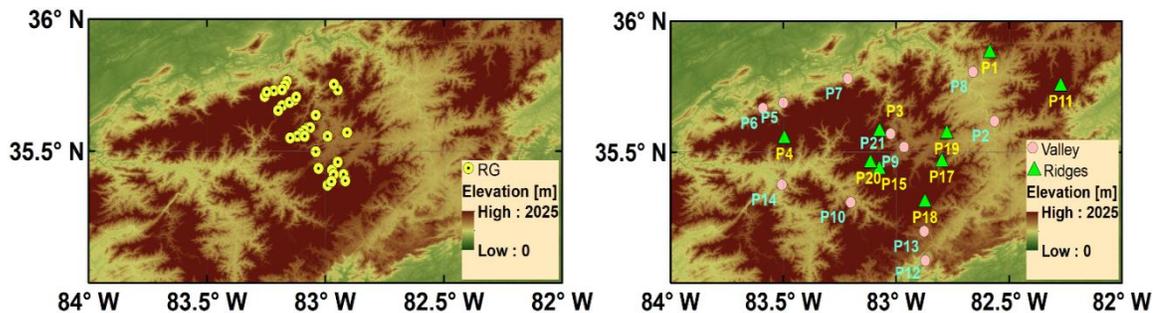
Figure 2-5: Location of IPHEx GV (left) rain-gauges and (right) Parsivel disdrometers in the SAM. Triangles denote ridge (elevation > 900 m) and circles denote valley locations.

From synthesis, a geography of nonstationary retrieval errors emerges exhibiting robust spatial modes and diurnal and seasonal cycles tied to physical processes that may vary from one region to another, and from one location to another within the same region (e.g. Barros et al. 2004; Prat and Barros, 2010b; Duan et al. 2015). The GV framework for error analysis and attribution therefore sets the stage for process studies, including

models and observations, to elucidate the physical-basis of error toward ultimately improving quantitative precipitation estimation (QPE).

Next, we rely on IPHEx data over the SAM (Fig. 2-5) to examine GPM DPR measurements and precipitation products aimed at elucidating the physical underpinnings of retrieval errors. In the Southern Appalachians (SAM), seeder-feeder interactions (SFI) between rainfall and low level clouds and fog (LLCF) modify the vertical structure of rainfall enhancing drop coalescence efficiency between drop diameter of 0.1 and 1 mm that results in increasing the number of raindrops between the diameter range (0.1 to 1 mm; Wilson and Barros, 2014). Detailed studies within sub-regions of the SAM to characterize SFI using spatially distributed observations to monitor different microclimates (Fig. 2-2) indicate that there is substantial heterogeneity in the spatial and temporal organization of LLCF. This heterogeneity in turn leads to variability in rainfall microphysics from one location to another depending on time and season as illustrated by the climatology of the regional climatology of LLCF in Fig. 2-6 (Wilson and Barros, 2017; Duan and Barros, 2017).
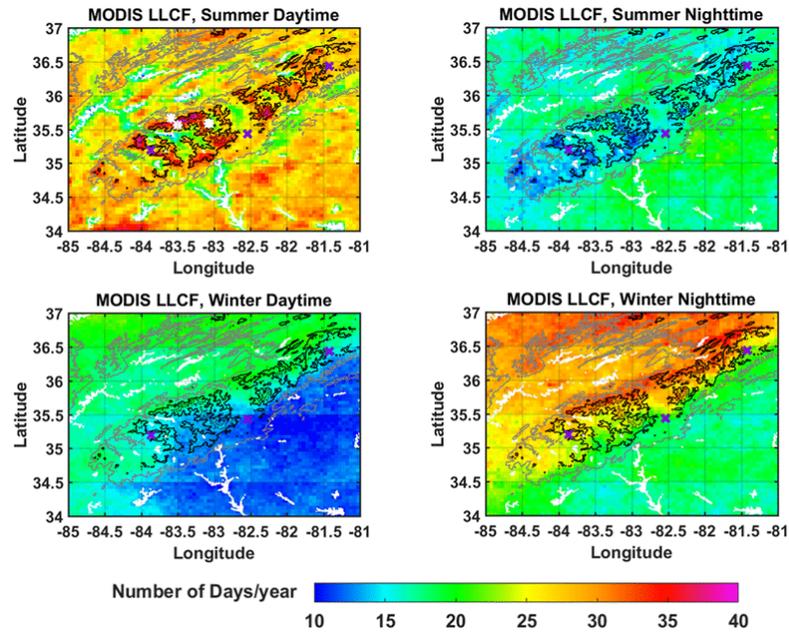
Figure 2-6: Spatial distributions of AQUA MODIS LLCF (CTH <5 MSL, confident cloudy only; 0.05°× 0.05°) during daytime and nighttime overpasses in summer and winter. (After Duan and Barros, 2017).

Fig. 2-7 summarizes analyses of rainfall microphysics from two disdrometer locations, P4 and P6 (see Fig. 5 for location). From midmorning to midnight, the number of drops is approximately the same in the ridge and in the foothills (not shown), but D_m is larger by 20% at P6 due to a deeper atmospheric column for SFI among low level layered clouds (Rain on LLC).
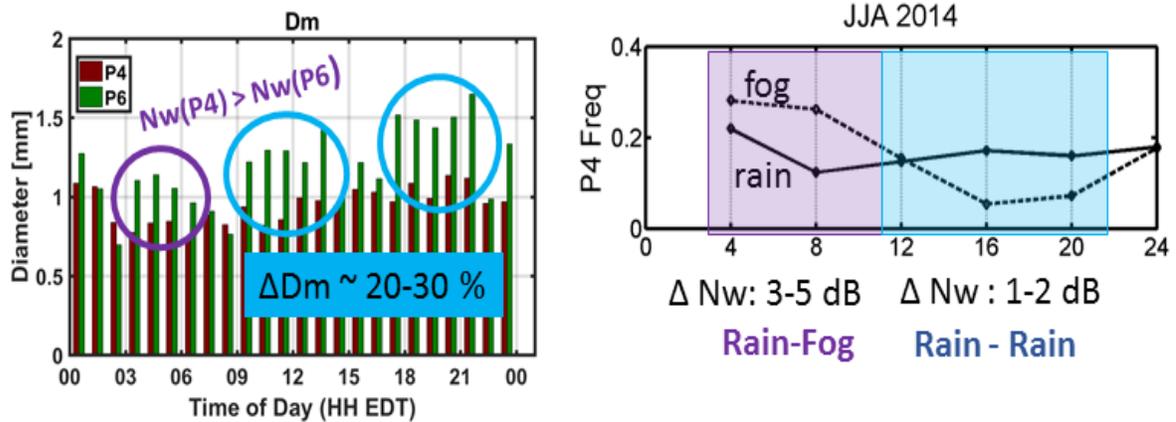
Figure 2-7: Fingerprinting SFI on DSD metrics. Left: diurnal cycle of D_m (ratio of fourth moment to the third moment of the diameter) at P4 (western ridge) and P6 (foothills). Right: diurnal cycle of rain and fog occurrences at P4. Note that D_m (P6) always ≥ D_m (P4).

The value of N_w, the normalized DSD intercept, is the same in the ridge and in the valley from mid-morning to the evening hours, but there is a 3-5 dB increase at P4 between 3 and 7 AM (purple). Between 3 and 7 AM, fog forms on the ridge resulting in a significant increase in N_w at P4 (Rain on Fog and Cap Clouds; Rain on FCC), but the fog layer is not deep enough to impact drop dynamics and significantly change Dm.

2.4 Physical-Basis of Retrieval Errors

In addition to measurement geometry linked (Fig. 2-4), overpass frequency is also an important source of variability in satellite-based remote sensing for precipitation as demonstrated in Fig. 2-8, which reconciles the diurnal cycle of revisit time frequency over the IPHEx region globally (left panel) and its spatial distribution for times when it is

raining (right panel). Note the low frequency revisit bias in the morning (left panel) and the inner mountain region (right panel), where SFI are prominent and strongly impact the diurnal cycle of rainfall.
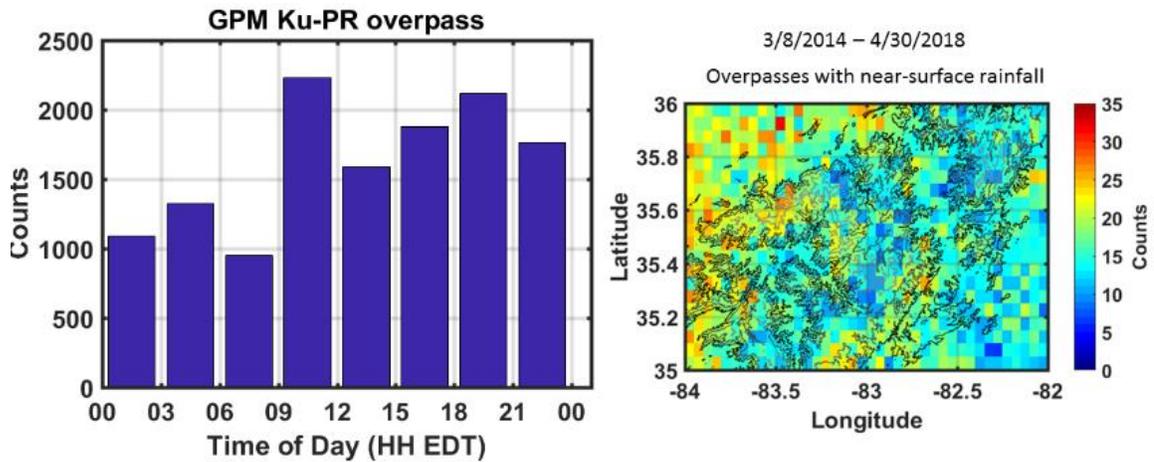
Figure 2-8: (a) Diurnal cycle of GPM revisit overpasses over the SAM. (b) Spatial distribution of rainfall detection. The blue tones in the inner region valleys and over the complex terrain at low elevations along the eastern ridges of the SAM are indicative of low frequency bias.
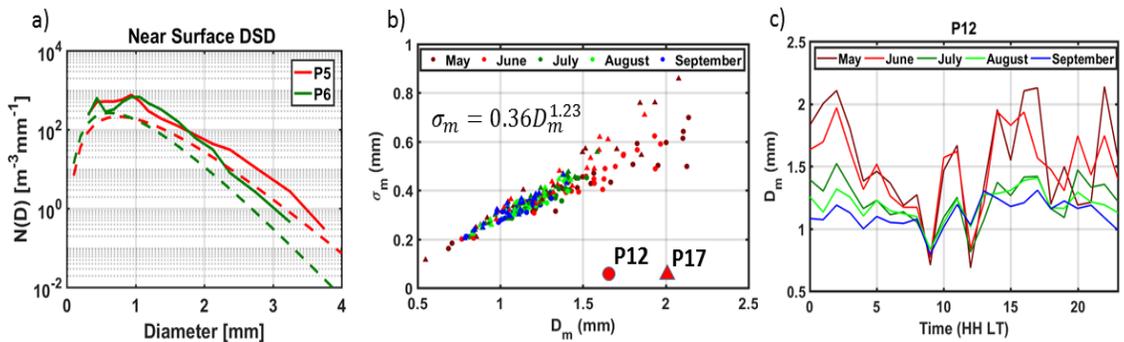
Figure 2-9: DSD non-stationarity on altitudinal gradients: a) disdrometer (continuous lines) and GPM Ku-PR estimates (dashed lines) on the western slopes of the SAM; b) and c) DSD statistics on the eastern slopes of the SAM.

The underestimation of rainfall along the western slopes of the SAM results from underestimating the number of small and intermediate sized raindrops (D < 2mm) at low elevations as shown in Fig. 2-9(a) for locations P5 and P6 (see Fig. 2-5 for locations). Note the much higher D_m for May and June compared to the remainder summer months, consistent with the regional climatology of mesoscale convective systems that produce heavy rain with larger size drops in spring. Consequently, the seasonal and diurnal cycle of detection errors reflects the relative contribution of SFI enhancement versus classical orographic enhancement and spatial modulation of stratiform versus convective precipitation processes (Fig. 2-10 and 2-11): high FARs in the winter (DJF) and in the morning (high elevation) and in the summer (JJA) at mid-day in contrast with the improved detection for convective processes generally; overestimation of light rainfall rates due to NUBF and overcorrection of GC effects, and underestimation of convective rainfall due to lack of sensitivity for high rainfall rates.
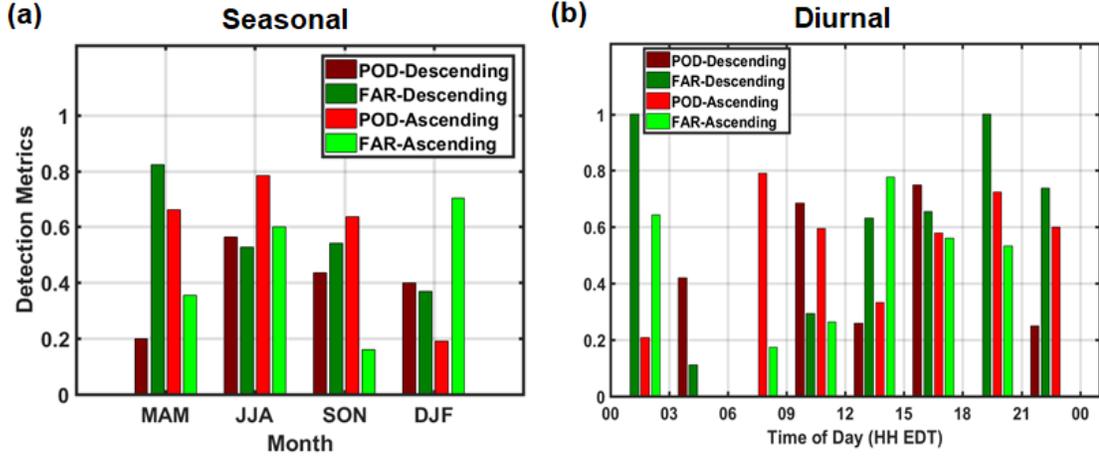
Figure 2-10: a) Seasonal and b) diurnal distribution of GPM Ku-PR detection errors in the Southern Appalachian Mountains for ascending and descending modes of the GPM.

The DSD of liquid hydrometeors in the GPM Level 2A products is assumed to follow a normalized gamma distribution function with specified shape factor $\mu = 3.0$ as follows:

$$N(D) = N_w f(\mu) \left(\frac{D}{D_m}\right)^\mu exp\left(-\frac{(4+\mu)D}{D_m}\right) \tag{2-1}$$

where $f(\mu) = \frac{6(4+\mu)^{(\mu+4)}}{4^4\Gamma(\mu+4)}$ and Γ is the gamma function.

The mass-weighted diameter D_m is defined as the ratio of fourth moment to the third moment of diameter with the drop diameter varying between D_{min} and D_{max} with interval of dD:

$$D_m = \frac{\sum_{D_{min}}^{D_{max}} N(D)D^4 dD}{\sum_{D_{min}}^{D_{max}} N(D)D^3 dD} \tag{2-2}$$
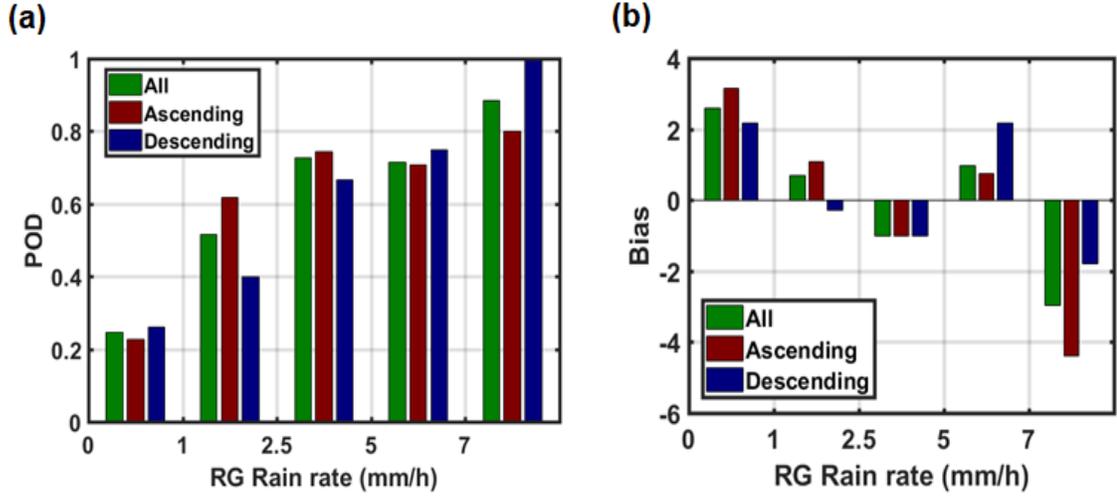
Figure 2-11: a) Detection and b) estimation errors varying with rain-gauge rain-rate observations for ascending and descending modes of the GPM Ku-PR.

The normalized number concentration or intercept factor, N_w (m^{-3} mm^{-1}) is calculated from the liquid water content q (g m^{-3}) [where $q = \frac{\pi}{6}\rho_w \sum_{D_{min}}^{D_{max}} N(D)D^3 dD$, ρ_w is the density of water (10^{3} g mm^{-3})] and D_m:

$$N_w = \frac{4^4}{\pi\rho_w}\left(\frac{q}{D_m^4}\right) \tag{2-3}$$

One approach to elucidate the physical basis of retrieval error is to investigate whether the conceptual microphysics model underlying radar retrieval is representative of observed microphysics. GPM Ku-PR v05A product is analyzed in this Chapter. Figure 2-12 shows the spatial structure of the average D_m and N_w retrieved since GPM launch 2014. One salient feature in the time-average spatial distribution of Ku-PR D_m and N_w is that they vary in opposite directions within their respective ranges. This results in higher N_w values over the ridges than in the valleys, and larger D_m values in the foothills on the

36

eastern and western slopes of the SAM as well as at the confluences of wide open

tributaries in the inner region (Little Tennessee and the French-Broad). This is consistent

with disdrometer observations at some, but not at all the times of day in parts of the

western sector due to widespread mid-day and afternoon LLCF in the inner region (not

shown, Duan and Barros 2017; Wilson and Barros 2014). Even though the higher (lower)

values of N_w (D_m) in the sub-regions of low frequency bias in Fig. 2-8 is physically-true,

bias in the DSD parameters are observed due to the lack of constraints in single-frequency
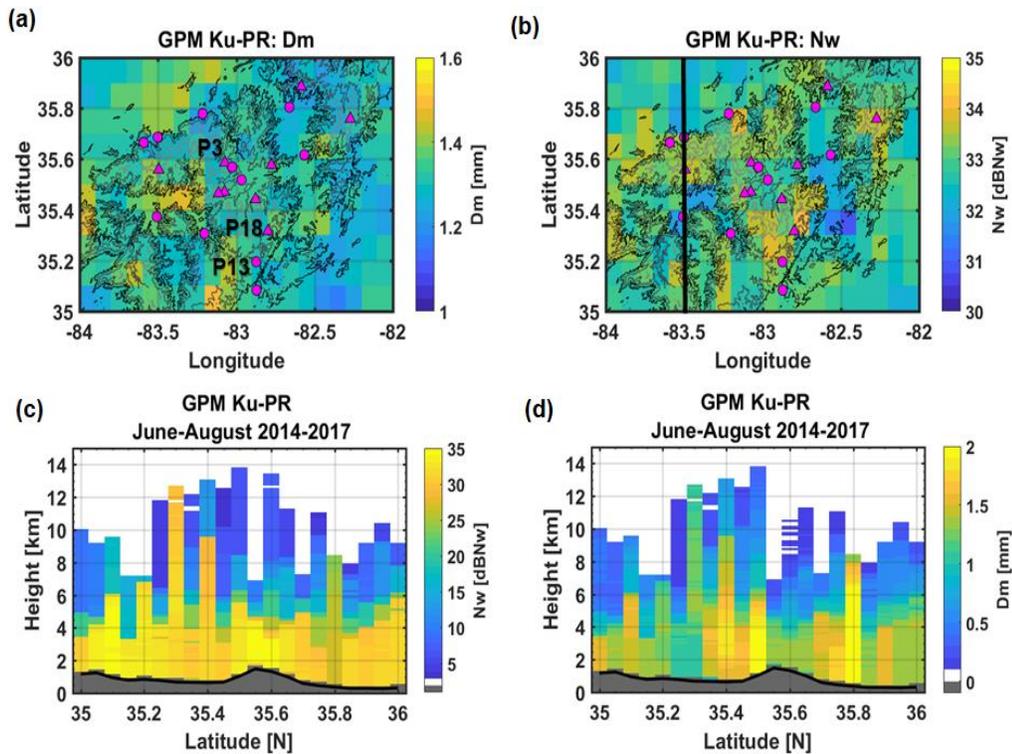
Figure 2-12: Average GPM Ku-PR Version 05A DSD parameters in the SAM. Spatial distribution: a) D_m; b) N_w. Vertical structure along transect at 83.5W: c) N_w; d) D_m. The black line in denotes the transect. The pink circles and triangles mark respectively valley and ridge disdrometer locations.

retrieval algorithm and the fixed shape factor assumption in the gamma distribution in GPM-DPR products.
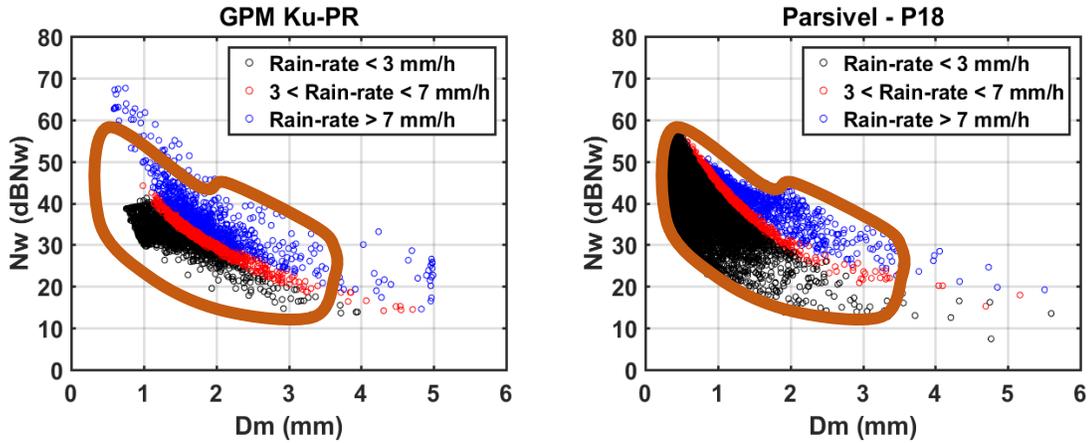
Figure 2-13: Phase-space maps of (D_m, N_w) from GPM Ku-PR Version 05A product and from GV Parsivel disdrometer P18 during IPHEx.

Figure 2-13 contrasts the GPM Ku-PR single frequency DSD estimates (top right panel) with the disdrometer observations at P18 (see Fig. 2-5 for location) which further supports this hypothesis. The DSD parameters at P18 are calculated from the disdrometer DSD distributions for each time using Eqs. (2) and (3). The Ku- PR parameters exhibit steep trade-offs in (D_m, N_w) space over a relatively narrow region governed by the inverse relationship also apparent in the spatial distributions in Fig. 2-12 (see also Prat and Barros, 2009). This behavior is not consistent with the disdrometer at either low (black) or moderate to high rainfall intensity (blue), and only to a much lesser degree for intermediate rain rates (red). The closed brown contour line delineates a region of (D_m,

N_w) space filled by the multiplicity of microphysical states (variable μ) observed within the disdrometer measurement capabilities at high temporal resolution [1 minute], in line with the near-instantaneous duration of the satellite overpass. In addition, increases in disdrometer D_m values do not imply substantial increases or decreases N_w indicating convective rainfall (see P12 in Fig. 2-9) or SFI on layered LLC as previously discussed at P6 (Fig. 2-7). Figure 2-14 shows simulation results of Rain on Fog simulations at P6 (see Fig. 2-7) using an existing model (Prat and Barros, 2007; Prat et al. 2012) that describes stochastic raindrop dynamics in the atmospheric column that has been modified to simulate SFI through layered low level clouds, orographic cap clouds and, or fog (Wilson and Barros, 2014; Duan and Barros, 2017). The left panel shows the sensitivity of simulated rainfall intensity to the DSD spectra of fog, and thus the specific fog microphysical regime matters in order to capture the observed rainfall at the collocated rain-gauge. Detailed description of simulations is provided in Duan and Barros (2017). The right panel shows that the trajectory of an individual precipitation event occupies a large area of phase-space from start to end of the Rain on Fog event. Thus, the space-filling behavior exhibited by the disdrometer data in Fig. 2-13 reflects change in rainfall intensity that emerge as the microphysics change during a storm. This nonlinear behavior translates into changes in the shape parameter μ that characterize the evolution raindrop population collision dynamics (coalescence versus breakup) at difference stages of the event (Testik and Barros, 2007; Prat and Barros, 2010a; Testik et al. 2011; Prat et al. 2012).
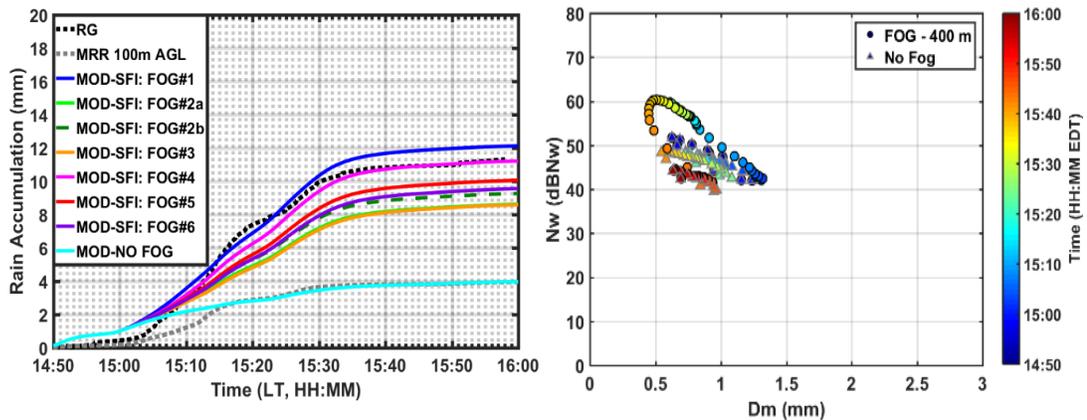
Figure 2-14: Sensitivity analysis of surface rainfall accumulation to the DSD spectra of fog in the simulation of RoF SFI at P6 in the western foothills of the SAM. RG–dashed black lines are local rain-gauge observations. After Duan and Barros (2017). Right: simulated microphysical trajectory for stratiform rainfall with and without RoF SFI at P6 corresponding to Fog#4 DSD and fog depth of 400m. Circles describe microphysical state at 1-min intervals from start to end of the event as per the legend.

The space-time time organization of errors by topography and microphysical regime propagates from GPM Level 2A to higher level products such as IMERG QPE (Huffman et al. 2017). This is demonstrated in Fig. 2-15 for the SAM by comparing IMERG-v04A to the Stage IV combined radar-rain-gauge product between March 2014 and February 2017. The spatial resolution of IMERG-v04A precipitation product is approximately 11 km while Stage IV has a spatial resolution of approximately 4 km. The 1-hour accumulation of Stage IV precipitation data can be obtained from http://www.emc.ncep.noaa.gov/mmb/ylin/pcpanl/stage4/.

The spatial maps show the detection error metrics such as the probability of detection (POD) and false alarm rate (FAR). The probability of detection is computed as:

$$POD = \frac{YY}{YY + MD}$$

Where, YY is the number of correct detections of precipitation and MD is the number of missed detections. FAR is mathematically computed as:

$$FAR = \frac{FA}{YY + FA}$$

Where FA is the number of false alarms. The optimal value of POD is 1 while the best value of FAR is 0. Figure 2-15 shows that the missed detections dominate in the eastern and western ridges of the SAM. However, FAs dominate in the east of the Appalachian mountain ranges.
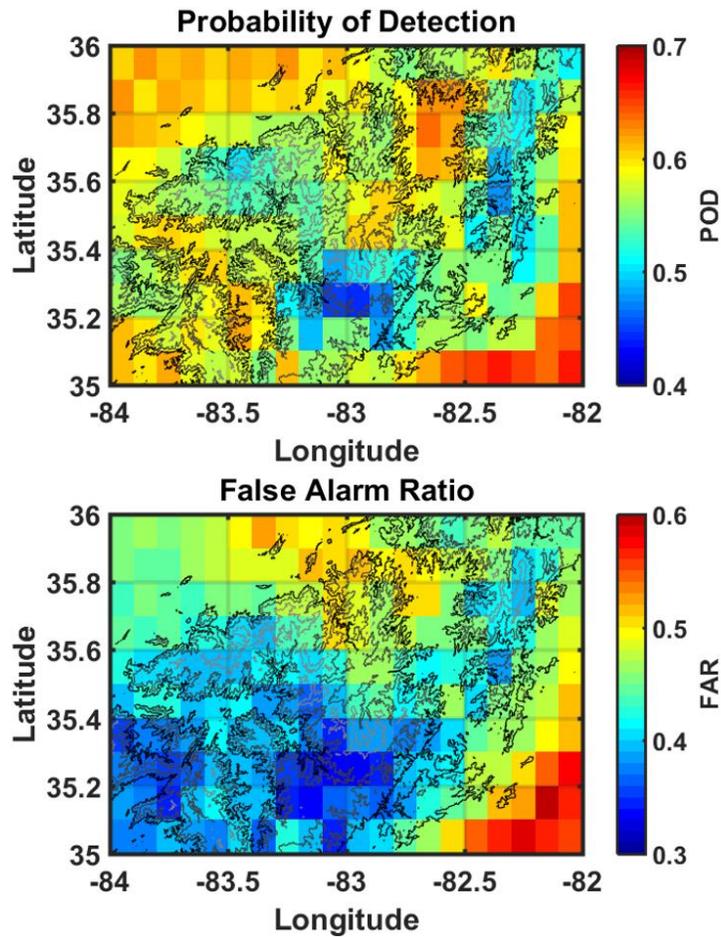
Figure 2-15 (a) Probability of Detection and (b) False Alarm Ratio from error diagnostics of IMERG-v04A when compared with STAGE IV precipitation data. Time period of analysis- March 2014 to February 2017.

The spatial maps show the diurnal cycle climatology of POD are shown in Fig. 2-16. Note the concurrence between the spatial patterns in Fig. 2-16 and Fig. 2-8 and the temporal patterns in Fig. 2-4. The time-series graphs on the right show the DSD climatologies at three Parsivel disdrometer locations (P3, P18, and P13) marked in the spatial maps. The lowest skill of the IMERG product is at mid-day (12-17 LT) in the inner

ridge and eastern slopes of the SAM, which are times during which there is significant rainfall with small D_m (< 1 mm) at very high concentrations and large spatial variability. This outcome, while illustrative of the challenges in orographic QPE, clearly makes the case for the importance of capturing the underlying physical processes in retrieval algorithms which cannot be statistically retrieved or corrected.
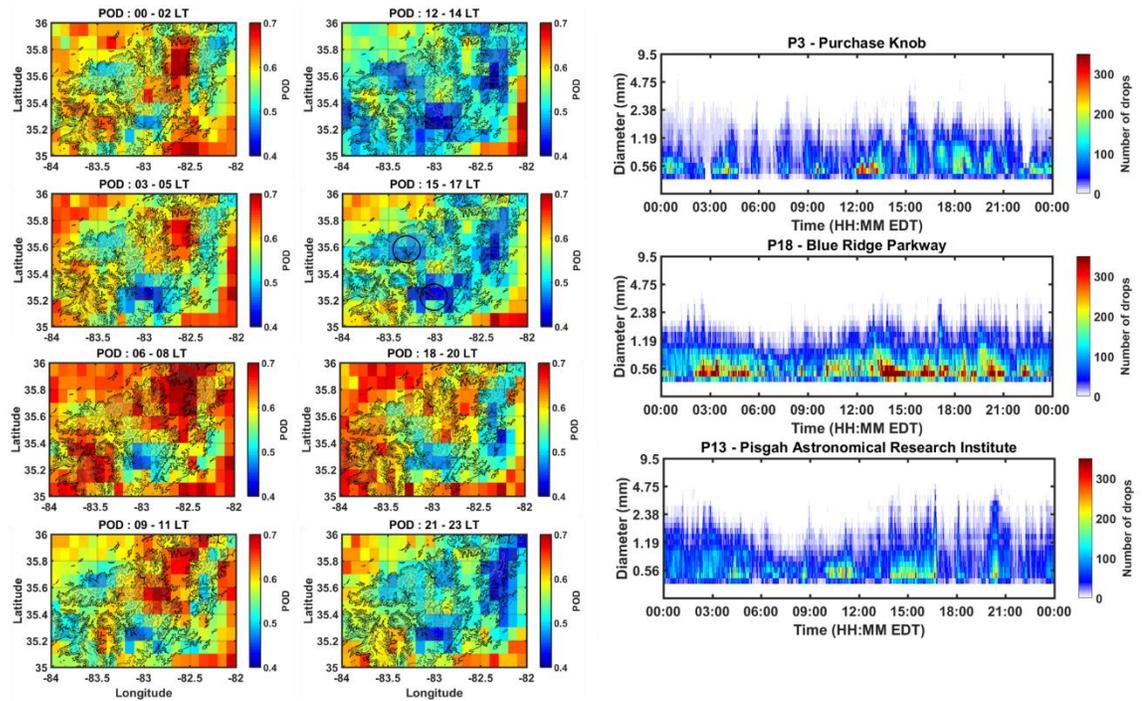
Figure 2-16 (a) Diurnal cycle of the spatial distribution of POD of IMERG-v04A precipitation compared with STAGE IV and (b) Parsivel disdrometer DSD measurements .Time period of analysis - March 2014 to February 2017.

2.5 Summary

The grand challenge of precipitation measurement in mountainous regions is therefore to measure rainfall extremes, both light, and heavy rainfall rates, at the relevant hydrometeorological spatial and temporal scales. The major findings of this chapter are as follows:

- Previous studies suggest the presence of robust spatial, temporal and seasonal patterns of low-level clouds and fog in the SAM. The impact of the low-level clouds and fog are expected to be missed by the GPM DPR estimations due to ground-clutter contamination of near-surface profiles especially at off-nadir pixels.

- The ground-clutter correction methodology followed in the retrieval algorithm is expected to increase the detection and estimation errors of low-level light precipitation and low-level enhanced collision-coalescence precipitation events.

- Error analysis of GPM DPR with ground-based rain-gauge observations suggests that the errors are dependent on the GPM DPR viewing geometry and viewing angle.

- Analysis of ground-based Parsivel disdrometer data suggests high spatial and variability in the DSD parameters. However, the spatial variability highlighted in the GPM Ku-PR estimates is limited.

- Comparison of Stage-IV estimates with IMERG precipitation estimates provides a spatial and diurnal patterns of detection errors in the SAM. MDs dominate along the ridges and the inner region while FAs in the northern part of the SAM and eastern sides of the valleys. The disdrometer observations in the MD dominated region suggest that the DSDs are dominated by high concentrations of small drop sizes suggesting the presence of low-level clouds and fog.

3. Shallow Precipitation Detection and Classification using Multifrequency Radar Observations and Model Simulations

Note this chapter appeared as reference (Arulraj and Barros, 2017).

3.1 Introduction

The space–time variability of hydrometeors in convective precipitation shows high variability in vertical and horizontal structures typically with distinct deep cores of heavy rainfall and broad drop size distributions (DSD) in contrast with narrow DSDs for stratiform rainfall (Houze 1993; Zafar and Chandrasekar 2004). Independently of precipitation regimes, DSDs can change significantly with height over time because of changes in the surrounding environment and drop–drop interactions, including coalescence and break-up dynamics (e.g., Prat et al. 2008). In regions where spatially persistent multilayer cloud systems and fog banks form, seeder–feeder interactions modify rainfall at low levels, such as in the case of reverse orographic enhancement in the Southern Appalachians (Wilson and Barros 2014, 2015, 2017). Seeder–feeder interactions refer to the accelerated growth of raindrops produced by high-level clouds (seeders) as they fall through multilayer cloud systems, coalesce with low-level cloud and fog drops (feeders), resulting in a significant increase in raindrop sizes near the ground and, consequently, significant increases in surface precipitation. The measured precipitation radar reflectivity factor $Z_{l,m}$ at wavelength l and range R is a nonlinear function of the equivalent reflectivity factor $Z_{l,e}$ and the specific attenuation k_l:

$$Z_{\lambda,m}(R) = Z_{\lambda,e}(R)exp\left(-0.2\ln(10)\int_0^R k_\lambda(s)ds\right) \qquad (3\text{-}1)$$

$$Z_{\lambda,e}(R) = \frac{\lambda^4}{\pi^5|K_w|^2}\int_{D_{min}}^{D_{max}} N(D,R)\sigma_{\lambda,b}(D,\lambda)dD \qquad (3\text{-}1a)$$

$$k_\lambda(R) = \int_{D_{min}}^{D_{max}} N(D,R)\sigma_{\lambda,e}(D,\lambda)dD \qquad (3\text{-}1b)$$

where $\sigma_{\lambda,b}$ and $\sigma_{\lambda,e}$ are respectively the backscattering and extinction cross sections of hydrometeors of diameter size D, N(D,R) is the DSD of rainfall at range R, and $|K_w|^2 =$ 0.93 for water. The received power $P_\lambda(R)$ can be calculated using the generalized radar equation $[P_\lambda(R) \approx C\, Z_{\lambda,m}(R)\,/\,R^2]$, where the radar constant C varies directly with range resolution ΔR and inversely with λ^2. The reflectivity factors $Z_{\lambda,e}$ and k_λ vary non-linearly with wavelength, hydrometeor size distribution, and rainfall rate (Liao et al. 2014; Marzano et al. 2003).

Reflectivity profiles of radars operating at short wavelengths differ from long wavelength radars as a result of differences in attenuation and scattering behavior as the concentration of non-Rayleigh scattering particles changes with wavelength (Kollias et al. 2007). Strong attenuation and large non-Rayleigh scattering artifacts for moderate and heavy intensity precipitation, typically associated with a high number of large raindrop sizeslimit the use of millimeter-wave radars (e.g., W-band) to clouds and light rainfall (Kollias et al. 2007; Mead et al. 1996). For example, in the lower 6 km of the troposphere, attenuation of the CloudSat Cloud Profiling Radar (CPR; 94GHz, W band) signal is much stronger than for the Tropical Rainfall Measurement Mission Precipitation Radar (TRMM PR; 13.8GHz, Ku band); conversely, the TRMM PR misses the anvil part of the cloud

system (Sindhu and Bhat 2013). Because differences in the structure of reflectivity profiles at distinct frequencies result from specific precipitation structure features, combining and contrasting multifrequency radar observations should help to overcome their independent disadvantages (Figure 3-1). This approach could be particularly useful to leverage Global Precipitation Measurement Dual-Frequency Precipitation Radar (GPM DPR, centimeter wavelength) and CloudSat CPR (millimeter wavelength) observations, including upcoming satellite missions [e.g., the European Space Agency's Earth Clouds, Aerosol and Radiation Explorer (EarthCARE)].

In the absence of physics-based models, operational precipitation retrievals from satellite-based observations such as the TRMMPR and the GPM DPR rely on DSD parameterizations that vary according to rainfall regime (e.g., convective or stratiform; Iguchi et al. 2000; Short and Nakamura 2000). Rainfall regime detection and classification are therefore a key step in the retrieval work flow prior to rainfall estimation proper (see Appendix B for a review of existing rainfall classification strategies developed for TRMM and GPM precipitation retrieval algorithms).Orographic processes and/or precipitation regimes in regions of complex topographic transitions are not explicitly addressed in operational algorithms because of the complex microphysics of orographic precipitation coupled to high spatial and temporal variability in a precipitation structure modulated by
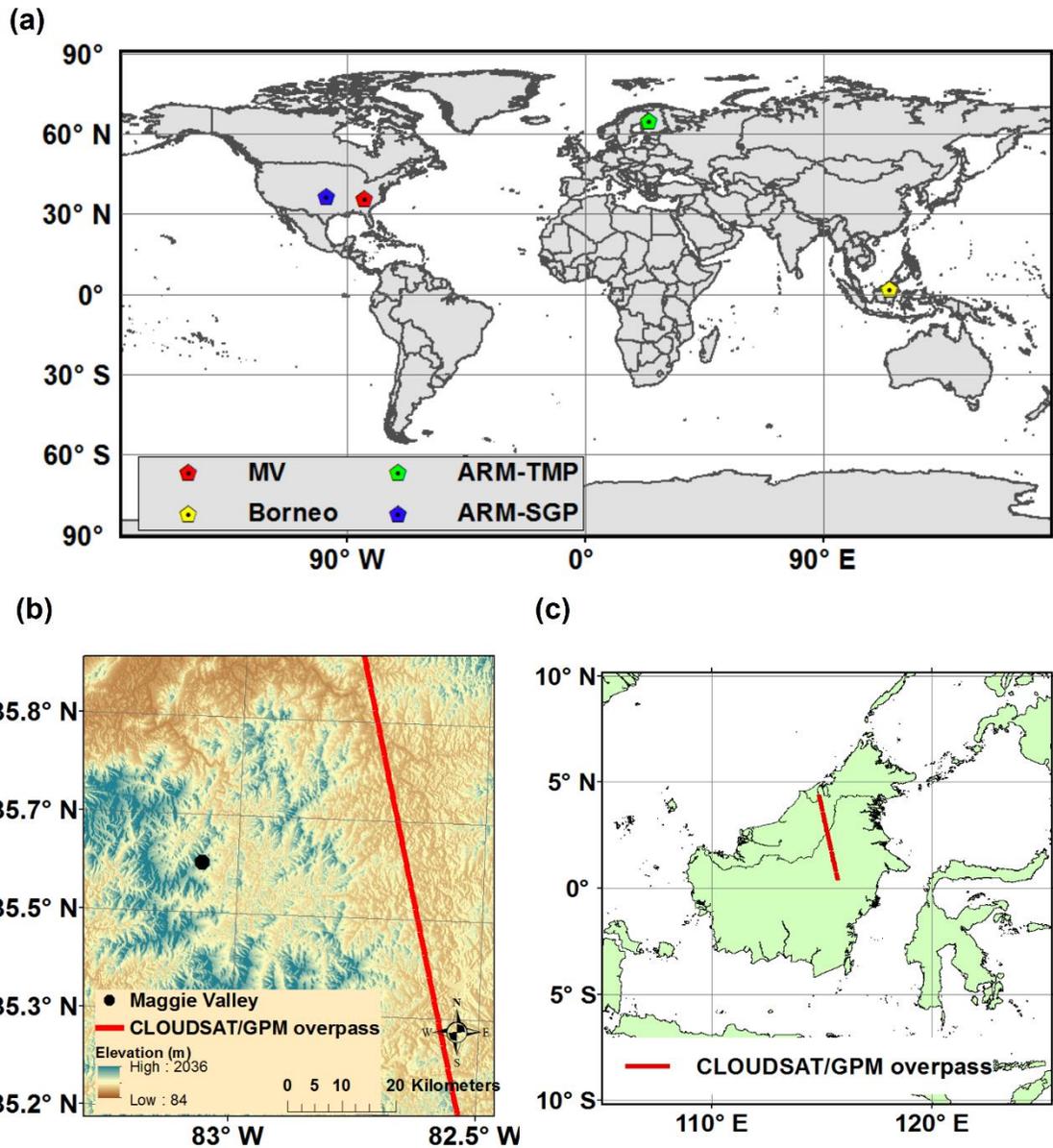
Figure 3-1 (a) The locations were SRDC algorithm was tested. Concurrent GPM-DPR and CloudSat-CPR overpass for (b) May 15, 2014 and (c) June 03, 2014 event.

topography (Barros et al. 2000; Houze 2012). Precipitation retrievals in mountainous

regions consequently exhibit large detection errors (~50%) and severe underestimation of

precipitation rates (e.g., Barros 2013; Prat and Barros 2010). Further, Duan et al. (2015)

showed that, independent of rainfall intensity, large errors and uncertainty in TRMM PR rainfall estimates in the Southern Appalachian Mountains (SAM) exhibited diurnal and seasonal patterns linked to spatial and temporal patterns of light and shallow orographic rainfall. This behavior is expected in complex terrains generally (e.g., Barros et al. 2000).

In this chapter, a new algorithm to detect and identify shallow precipitation using collocated radars operating at two distinct wavelengths (i.e., centimeter vs millimeter) is presented. The algorithm, referred to as the shallow rain detection and classification algorithm (SRDC), relies on an integrated measure of the space–time correlation between the reflectivity profiles, and is demonstrated and evaluated using specifically W- and Ka-band ground-based radars from monitoring sites in different climatic regions with distinct terrain complexity (Figure 3-1a). Note that the algorithm formulation presented here should be applicable to any pair of frequencies with significant contrast in attenuation and scattering behavior (e.g., millimeter vs centimeter wavelengths), such as, for example, W and Ku, X, or C bands; the focus here is on W and Ka bands because of ground validation data availability.

The algorithm was implemented first using observations at Maggie Valley (MV), North Carolina, during the intensive observing period of the Integrated Precipitation and Hydrology Experiment (IPHEx IOP) from 1 May 2014 to 15 June 2014 in SAM (Barros et al. 2014). The second demonstration is for Hyytiala, Finland (ARM-TMP), a relatively flat region of terrain crisscrossed with many land–water boundaries, during the Biogenic

50

Aerosols–Effects on Clouds and Climate (BAECC) field campaign from February through September (Petäjä 2013). An independent evaluation was conducted using ground-based radar data from the Atmospheric Radiation Measurement Southern Great Plains (ARM SGP) site in Oklahoma, a region without salient topographic or coastline contrasts. For proof of concept and to examine the potential generalization and application of the algorithm to satellite-based observations, the SRDC algorithm is applied to observing system simulations (OSS) of nearly concurrent GPM and CloudSat overpasses on 15 May 2014 during IPHEx (Figure 3-1b), and to GPM DPR and CloudSat CPR observations along a complex terrain transect in Borneo (Figure 3-1c).

The organization of the chapter is as follows. Section 3.2 describes the data and instruments employed in locations such as MV and ARM-TMP. Section 3.3 presents the algorithm formulation followed by results and statistical analysis of the accuracy of the SRDC algorithm in section 3.4, and a summary and conclusions are given in section 3.5.

3.2 Data Description

3.2.1 Maggie Valley, North Carolina (MV)

The IPHEx IOP was conducted between 1 May 2014 and 15 June 2014 over the southeastern United States with collocated instruments installed at MV (35.5198°N, 83.0947°W), including a scanning W-band dual-polarization radar, a Micro Rain Radar (MRR), an all-sky imager (ASI), a ceilometer, and an OTT Parsivel disdrometer (Barros et

al. 2014). The Aerosol, Cloud, Humidity, Interactions Exploring and Validating Enterprise (ACHIEVE) W-band radar operates at 95GHz in a vertically pointing mode with a 0.25° beam-width and a minimum detectable threshold of -55 dBZ at 1 km. The range resolution is 25 m, where the vertical range extends from 475m to 8.4km. The data acquisition rate ranges from 1 to 2 s; however, data utilized in the analysis are aggregated to 1-min temporal resolution. The MRR operates at 24GHz (K band) with minimum detectable reflectivity of -5dBZ and 1.5° beam-width in vertically pointing mode to measure the profiles of Doppler power spectra. It was set up with a vertical resolution of 50m with range extending to a maximum height of 1.5 km, and the measurements were aggregated to 1-min temporal resolution. The equivalent reflectivity factor obtained from the MRR is used for the analysis with negative values reset to zero. The OTT Parsivel[2] is a laser-based optical disdrometer that simultaneously measures the particle size and velocity of the raindrops with a temporal resolution of 10 s, and thus records the number of raindrops and rain intensity. The formulation and implementation of the SRDC algorithm is informed by past research, including comprehensive analysis of collocated MRR and Parsivel observations in the SAM (e.g., Prat and Barros 2010; Wilson and Barros 2014).

3.2.2 ARM-TMP

The ARM Mobile Facility (AMF2) was installed in a Scots pine forest at TMP (Figure 3-1a) from February 2014 until mid-September 2014 during the BAECC field campaign. The objective was to measure biogenic aerosols from forests and to study

aerosol–cloud–precipitation interactions and climate impacts (Petäjä 2013). Various instruments were collocated at the ARM TMP site, including the marine W-band ARM cloud radar (MWACR), the Ka zenith-pointing ARM radar (Ka-ZR), a ceilometer and a two-dimensional video disdrometer (2DVD). The MWACR operates at 95GHz in zenith-pointing mode, and the data were aggregated to 1-min temporal resolution. The radar was deployed at an elevation of 160m above mean sea level (MSL), and the vertical range of the reflectivity profiles extended from 210 m to 18.2 km. The Ka-ZR is a 35-GHz zenith-pointing Doppler radar with a data acquisition rate of approximately 1 s, deployed at an altitude of 180m MSL. The vertical range of data extended from 500mto 18 km. Data are available from the Ka-ZR in general and moderate sensitivity modes. The general mode detects full range but is less sensitive; the moderate sensitivity mode is a compromise between sensitivity and range, and thus the moderate sensitivity data are used here for analysis (Feng et al. 2014). The range resolution of the MWACR and Ka-ZR is approximately 30 m. The 2DVD comprises two orthogonally aligned video cameras to record the hydrometeor size distribution, fall velocity, geometry, and precipitation properties. The 2DVDdata have a temporal resolution of 1min, and the 2DVD rain rate was used for evaluating the algorithm.

3.2.3 ARM-SGP

The ARM Program has been operating a long-term observational network in the Southern Great Plains (SGP) for the past few decades. This site is characterized by

homogeneous topography, and large intra- and inter-seasonal variability of radiative energy fluxes. The central facility of ARM-SGP site is located at Lamont in north Central Oklahoma (Figure 3-1a), and it consists of various collocated ground-based instrumentation supporting research related to surface meteorology, aerosols, cloud properties, and atmospheric profilers. In this study, observations from the collocated Ka-band Zenith Radar (Ka-ZR), W-band Scanning ARM Cloud Radar (W-SACR), and optical rain gauge are used.

The ARM-SGP Ka-ZR is placed at 316 m AMSL and it operates at 35 GHz with range resolution of 30 m approximately, with a dual polarization transmitter and an antenna with 3-dB beam width of 0.2° Widener et al. (2012a). It operates in two modes, general and moderate sensitivity mode. The general sensitivity mode is used here considered for analysis due to the long-term availability of data. The W-SACR radar operates at 94 GHz and is placed at 318 m AMSL. It is equipped with a horizontal linear transmitter operating in vertically pointing mode, and the antenna beam width is around 0.33° (Widener et al. 2012b).

Various collocated instruments are available at the ARM-SGP site including a tipping bucket rain gauge, a disdrometer, and an optical rain gauge to measure surface precipitation. Due to the lack of continuity of the quality-controlled radar records, various periods April 20, 2011 and June 2014 were selected for analysis depending on data

availability, and for these periods, the optical rain gauge provided the most consistent and continuous records of precipitation rate measurement at the same elevation as the radars.

3.2.4 GPM DPR and CloudSat CPR Goddard SDSU

The Goddard Satellite Data Simulator Unit (G-SDSU; Matsui et al. 2013) simulates satellite observations based on atmospheric conditions generated by a numerical weather prediction (NWP) model [e.g., Weather Research and Forecasting Model (WRF)]. Here, G-SDSU GPMDPR and CloudSat CPR simulations of the 15 May 2014 event during IPHEx previously studied by Wilson and Barros (2015, 2017) are examined in detail. Note that the spatial resolution of both G-SDSU GPM DPR and CloudSat CPR simulations is 1.25 km, which is different from the along-track sampling resolution of the CPR (~1 km) and DPR (~5 km) products. Further, the CloudSat CPR G-SDSU simulation output is on a two-dimensional grid; the GPM DPR G-SDSU simulation output is on a three-dimensional grid with the vertical profiles aligned along the satellite overpass (single swath). The grid points with the closest spatial coordinates between the CloudSat CPR and the GPM DPR G-SDSU swaths were identified for analysis in section 3.4. Previous studies of CloudSat CPR G-SDSU simulations reported that it overestimates reflectivity profiles for heights above 4 km but that the overall macrostructure of the observations is captured well (Tao et al. 2009).

3.2.5 GPM DPR and CloudSat CPR Observations

The GPM DPR is a space-borne Earth-pointing precipitation radar on board the GPM satellite and operates in the Ku band (13.6GHz) and Ka band (35.5GHz). The Ka-band is tuned for high sensitivity to detect light precipitation and snow events up to 65° latitude. The nominal sensitivity of Ku- band is approximately 12 dBZ while the sensitivity of matched Ka-band radar is 17-18 dBZ. The sensitivity of interleaved Ka- band is approximately 15 dBZ (Iguchi et al. 2015; Hou et al. 2014). Ka-band radar reflectivity profiles from the DPR Level 2A product were used for analysis. The CloudSat CPR operates at 94 GHz to detect and distinguish weakly scattered radiation from clouds at the global scale (Stephens et al. 2008). Reflectivity data from the CloudSat geometric profile product (2B-GEOPROF) are used here.

The CloudSat–GPM coincidence product (2BCSATGPM; including CloudSat CPR (W-band), GPM DPR's Ku-band (13.6 GHz) and Ka-band (35.5 GHz) radar profiles, and the passive microwave brightness temperatures from the 13-channel (10–183 GHz) GMI radiometer for coincident CloudSat and GPM overpasses within 15-min of each other are utilized to extract nearly concurrent observations (Turk 2016). The minimum detectable radar reflectivity of CloudSat CPR is -30dBZ. The 2BCSATGPM products extract the nearest GPM DPR pixels that coincide with the CloudSat CPR track. Precipitation estimates for the concurrent CloudSat and GPM tracks extracted from the DPR Level 2A product [2ADPR, version 04A (2ADPR V04A)] include the precipitation rate from three

swaths—normal scans (NS), matched scans (MS), and high-sensitivity scans (HS). The HS detect light precipitation in the near-nadir region (within 8.5° angle; and with a swath width of 125 km) with a minimum detectable reflectivity factor of 15 dBZ. The MS by the Ka-band radar have minimum detectable reflectivity of 18 dBZ (0.5mm/h rain rate) with a swath width of 125 km. The NS of the Ku-band precipitation radar have a swath width of 245 km and a nominal minimum detectable reflectivity similar to that of MS, though Toyoshima et al. (2015) suggest that in practice the actual Ku-band radar NS detectability is 12 dBZ. Because of the differences in viewing geometry and resolution (Fig. 1 in Turk 2016), only the GPM DPR Ka-band HS, specifically the nadir-looking 2ADPR V04A product with all the CloudSat pixels aligned with the GPM DPR grid, are presented here, consistent with the ground-based cases.

3.3 Shallow Rain Detection and Classification Algorithm

3.3.1 Vertical Correlation Structure Calculations

The SRDC algorithm detects the presence of precipitation and classifies precipitation features into shallow and deep structures. In this study, deep structure events are defined as 1-min precipitation rates associated with reflectivity profiles exceeding reflectivity criteria above 5 km above ground level (AGL), whereas shallow precipitation events have reflectivity profiles exceeding reflectivity criteria below 5 km. Note that this reference height (i.e., 5 km) is expected to vary depending on regional

precipitation physics and topography. As discussed earlier, the algorithm takes advantage of contrasts between radar measurements in the centimeter (e.g. Ka-band) and millimeter (e.g. W-band) wavelengths to extract information masking their respective operational disadvantages as illustrated in Figure 3-2 (Kollias et al. 2007; Sindhu and Bhat 2013).
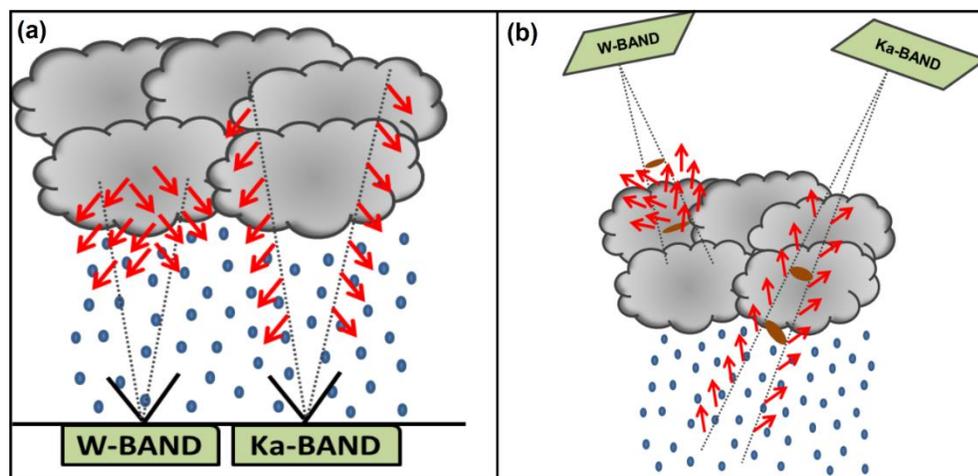
Figure 3-2: Physical Representation of operating principle of (a) Space-Borne and (b) Ground-based radars operating at Ka-band and W-band wavelength

Here, concurrent and collocated W-band (mm) and Ka-band (cm) radars operating in vertically pointing mode are the basic inputs into the algorithm. The time resolution and period considered for analysis should be identical because the algorithm involves the calculation of the space–time correlation between the two bands. The data are averaged to a common temporal resolution of 1 min. To leverage the dual-frequency observations, a two-pronged strategy is adopted: method 1 compares Ka-band profiles with W-band profiles over the entire vertical range of observations to identify the presence or absence

(clear sky) of deep convective structures; method 2 captures all precipitation events, and it is used in conjunction with method 1 to identify shallow precipitation events. The work flow of the SRDC algorithm is illustrated in Figure 3-3, and the formulation is detailed next.
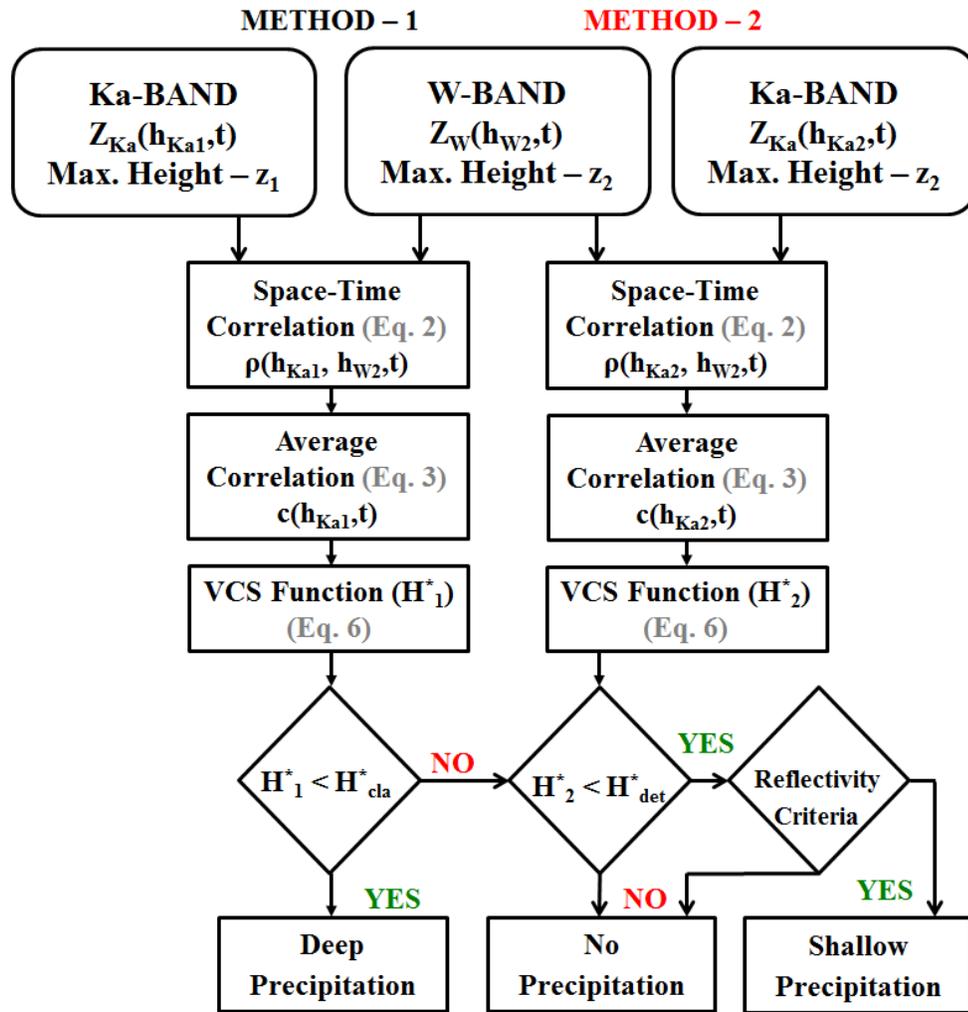
Figure 3-3: Schematic representation of the SRDC algorithm. The algorithm input is the equivalent radar reflectivity from collocated W-band and Ka-band radars. H^*_{cla} and H^*_{det} are the VCS thresholds for classification and detection, respectively.

Take the time-dependent reflectivity profiles $Z_{Ka}(h_{Ka1}, t)$ and $Z_W(h_{W2}, t)$ of Ka- and W- band radars, where h_{Ka1} and h_{W2} are vertical range bins of the Ka and W bands from the surface to maximum heights z_1 and z_2, respectively. First, the space–time correlation analysis between Ka- and W-band radar reflectivity factors is performed to obtain a three-dimensional (3D) space–time correlation $\varrho(h_{Ka1}, h_{W2}, t)$ as follows:

$$\rho(h_{Ka1}, h_{W2}, t) = \frac{\sum\limits_{k=0}^{N_t-1}[Z_{Ka}(h_{Ka1}, t-k) - \overline{Z_{Ka}(h_{Ka1}, t)}] \times [Z_W(h_{W2}, t-k) - \overline{Z_W(h_{W2}, t)}]}{SD_{Ka1} \times SD_{W2}} \quad (3\text{-}2a)$$

where N_t is the length of the moving time window (the appropriate length of the time window is examined in section 3.4), and

$$\overline{Z_{Ka}(h_{Ka1}, t)} = \frac{\sum\limits_{k=0}^{N_t-1} Z_{Ka}(h_{Ka1}, t-k)}{N_t} \quad (3\text{-}2b)$$

$$\overline{Z_W(h_{W2}, t)} = \frac{\sum\limits_{k=0}^{N_t-1} Z_W(h_{W2}, t-k)}{N_t} \quad (3\text{-}2c)$$

$$SD_{Ka1} = \sqrt{\sum_{k=0}^{N_t-1}(Z_{Ka}(h_{Ka1}, t-k) - \overline{Z_{Ka}(h_{Ka1}, t)})^2} \quad (3\text{-}2d)$$

$$SD_{W2} = \sqrt{\sum_{k=0}^{N_t-1}(Z_W(h_{W2}, t-k) - \overline{Z_W(h_{W2}, t)})^2} \quad (3\text{-}2e)$$

The 3D space-time correlation $\rho(h_{Ka1}, h_{W2}, t)$ is then averaged over the W-band vertical range converting the 3D data into the 2D matrix $c(h_{Ka1}, t)$:

$$c(h_{Ka1}, t) = \frac{\sum\limits_{j=1}^{N_{W2}} \rho(h_{Ka1}, h_{W2}(j), t)}{N_{W2}} \quad (3\text{-}3)$$

Where N_{Ka1} and N_{W2} are the number of height bins in the Ka- and W- band radar reflectivity profiles, respectively, within heights z_1 and z_2. At a generic time t, the average dual-frequency correlation (DFC) $c(h_{Ka1},t)$ is a one-dimensional vertical function defined at the height of each bin, that is a spatial series varying with height.

Changes in precipitation microphysics are captured differently in the reflectivity profiles at the two frequencies in the precipitation column, and therefore the DFC captures variations in the vertical structure of precipitation as a function of time. The height z_c at which $c(h_{Ka1}, t)$ first becomes negative (the first zero-crossing) is a robust estimate of the active depth of precipitation jointly detected by both radars. Note that the roles of Ka- and W-band reflectivity profiles are predicated on data availability over a vertical range suitable to quantify nonlinear attenuation differences between the two bands for the same vertical structure of precipitation. If the vertical range of the Ka-band radar is lower than that for W-band, then their roles in Eqs. 2(a-e) and Eq. 3 can be exchanged, albeit only for rainfall conditions that result in small to moderate attenuation at W-band (light rainfall). This will be further discussed in Section 3.4.

Next, the variations observed in the structure of the correlation function are described quantitatively using a measure of the information content in each column of the DFC (Balestrino et al. 2007; Denis and Cremoux 2002). In particular, Denis and Cremoux (2002) successfully used the mean of the absolute value of the time derivative of a time series X of length t to identify non-stationarity as follows:

$$H(\tau) = m_{|X'|} = \sum_{0}^{\tau} \frac{|X(t+\Delta t) - X(t)|}{\tau} \qquad (3\text{-}4)$$

Here, we are interested in spatial variability as a function of height z defined along the reflectivity profile rather than a time series. Assuming the process is stationary, the function defined in Eq. (3.4) is applied instead to the DFC c(h$_{Ka1}$, t$_0$) obtained from Eq. (3.3) [i.e., X is replaced by c(h$_{Ka1}$, t$_0$), where t$_0$ is the time representing the temporal window of width N$_t$ (i.e., τ = N$_t$) used to estimate the space–time correlation in Eq. (3.2)]. The DFC varies from -1 to 1, the integration length of the spatial series is the height d of the Ka-band reflectivity profile (km) [i.e., d = N$_{Ka1}$ x Δz, with radar bin resolution Δz], and thus the corresponding spatial metric Hs is calculated as follows:

$$H_s(t_0) = \frac{L(t_0)}{d}, \qquad (3\text{-}5a)$$

$$L(t_0) = \sum_{i=1}^{N_{Kal}-1} |c\,[h_{Kal}(i+1), t_0] - c[h_{Kal}(i), t_0]| \qquad (3\text{-}5b)$$

where $L(t_0)$ is the cumulative result of the absolute average correlation differences along the Ka-band profile. Given the small range of variability of the DFC during a storm event, a logarithmic transformation is applied to Eq. (4.5a) aiming to enhance the sensitivity of H_s by "stretching" $L(t_0)$ to detect small changes that can be attributed to changes in the vertical structure of precipitation processes. For a uniform spatial series, H_s = 0 and the logarithmic of a zero function is not defined. To avoid this problem, a shift was introduced to Hs so that when the DFC is constant, the modified metric is 1. Finally, the form of the measure of the variability of the average vertical correlation structure (VCS) function H^* for a generic time t is

$$H^*(t) = 1 + log_2 \left[\frac{L(t)}{d}\right] = log_2 \left[\frac{2L(t)}{d}\right] \; for \; L \neq 0 \tag{3-6a}$$

$$H^*(t) = 1 \; for \; L = 0 \tag{3-6b}$$

Note that $H^*(t)$ is tied to the specific length N_t of the window used for temporal averaging in Eq. (3-2), and different window lengths should yield different VCS values.

3.3.2 Precipitation Detection and Classification

The reflectivity profiles are processed following the two work flow paths depicted in Figure 3-3: method 1(column VCS; H_1^*) applies to deep reflectivity profiles extending to cloud-top height (CTH; e.g., z_1 = 8km at MV); method 2 (shallow VCS; H_2^*) applies to low-level radar reflectivity profiles up to a height z_2 to capture changes in surface precipitation due to low-level processes (e.g., z_2 = 1.5km at MV). A reflectivity constraint is included for no-precipitation event cases to reduce noise artifacts. Specifically, minimum reflectivity thresholds of 12 and 0 dBZ were used to detect precipitation for the Ka- and W- band radars, respectively, used in this study. Methods 1 and 2 follow the same steps except that the Ka-band (W band) reflectivity extends vertically to either CTH or z_1 (z_2). Note that z_1, z_2, and CTH vary from site to site depending on how the radars operate and thus data availability, and depending on local hydrometeorology (clouds and rainfall regimes). Operationally, a sliding (moving and overlapping) time window is used to detect and analyze short-duration precipitation. The ''optimal'' width of the sliding time window is extended to depend on the characteristic time scales of regional hydrometeorological processes.

The radar VCS functions using method 1 (H_1^*) and/or method 2 (H_2^*) show significant

sensitivity to rainy versus nonrainy conditions, and for rainy conditions with precipitation

structure, which can be verified against independent ground-based observations.

The detection threshold H_{det}^* is determined with a Bayesian classifier aiming to minimize

the overlap of the VCS probability density functions (PDFs) for method 1 (no

precipitation) and method 2 (precipitation). Similarly, the classification threshold H_{cla}^* is

determined to minimize the overlap of method 1 PDFs of H_1^* for deep and shallow

precipitation. These two thresholds are used subsequently to classify events based on

reflectivity measurements alone: 1) deep precipitation if (H_1^* , H_{cla}^*), else 2)

nonprecipitation conditions correspond to (H_2^* , H_{det}^*) and shallow precipitation

otherwise. The application and evaluation of the algorithm for different

hydrometeorological regimes are presented next.

3.4 Results and Discussion

3.3.1 Maggie Valley

The SRDC algorithm was developed initially for the IPHEx-IOP MV site as

described earlier. Figure 3-4 shows W-band radar and MRR reflectivity profiles on 11 June

2014. Intermittent precipitation activity was present throughout the day (from 10:00 until

22:00 LT) with shallow precipitation ahead of the convective rainfall after 20:30 LT. The

reflectivity profiles of the W-band radar and MRR distinctly show five events (S1, S2, S3,

S4, and D1), where the first four events are shallow precipitation (S1: 10:15–10:30 LT, S2: 14:30–14:45 LT, S3: 17:30–17:45 LT, S4: 18:30–18:45 LT) and the last event corresponds to a deep structure (D1: 21:00–21:15 LT). The DFC is computed between the W-band radar (z_1 5.8 km) and MRR (z_2 =1.5 km) reflectivity, and it is subsequently averaged over the MRR vertical profile range as shown in Figure 3-5(a). The method 1 average correlation shows high fluctuations around zero at higher levels after it first becomes negative at an altitude
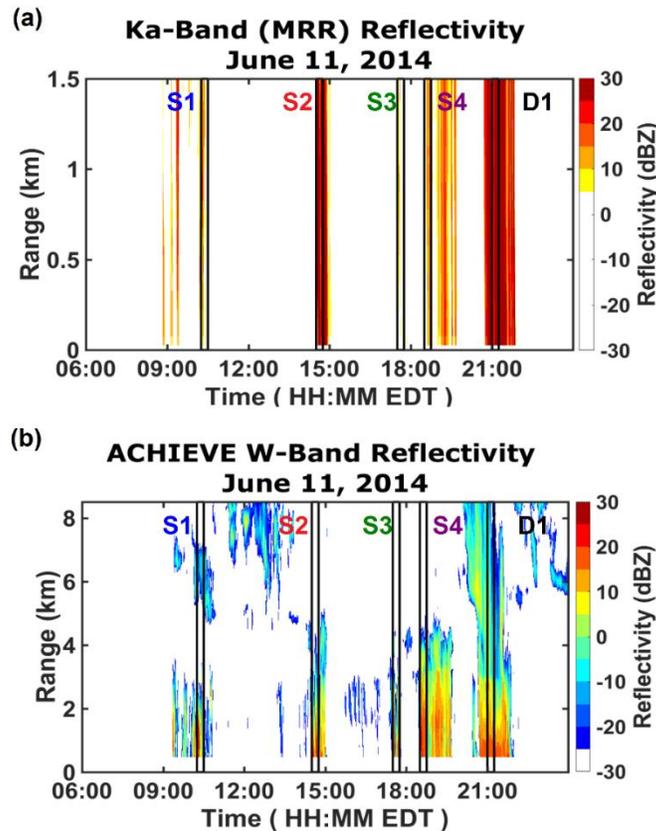
Figure 3-4: Equivalent reflectivity factor of (a) MRR and (b) ACHIEVE W-band radar observed at MV on 11 Jun 2014 during IPHEx IOP. Note that the MRR maximum level (1.5 km) is lower than the ACHIEVE W-band maximum level (8 km) as a result of the specific radar operational configuration. Five events (S1–S4, and D1; black boxes) considered for analysis.

below 5 km in the case of shallow events (S1–S4); however, the method 1 average

correlation structure is comparatively smooth for the deep structure event, D1, until it first

becomes negative at an altitude above 8 km. The MRR's operational setup at MV prevents
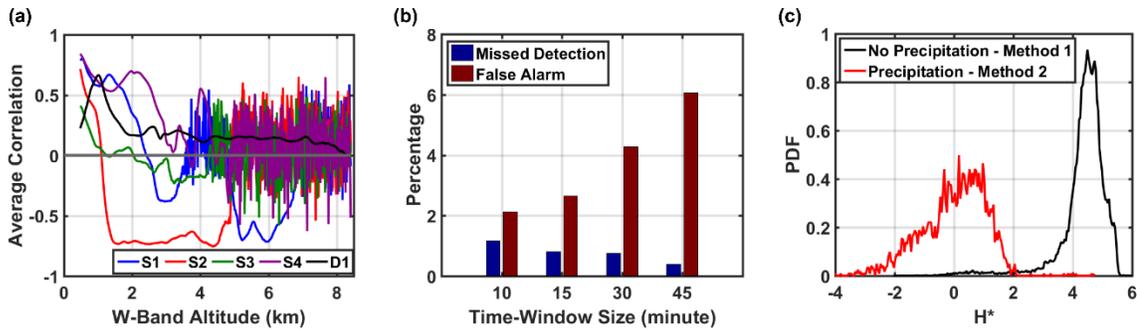
Figure 3-5: Space-time correlation, optimal time-window and VCS metrics for Maggie Valley, NC. (a) Average of the method 1 space–time correlation computed in step 2 of the SRDC algorithm for the five events highlighted in Figure 3-4. Changes in the pattern of the curve exhibit the evolution of D1 from S1–S4. (b) Trade-off between the MD and FA errors for varying time-window width. (c) PDF of the optimum (15-min) moving time window for the computation of VCS by the SRDC algorithm.

from using the Ka band for column VCS calculations; instead, W band is used here though

it suffers from severe attenuation for deep structure cases. This approach was adopted

after careful inspection of the adequacy of W-band observations, which revealed

consistently deep and well-defined W-band reflectivity profiles as a result of the high

frequency of light and shallow precipitation in this region. Thus, interchanging the Ka-

and W-band profiles is deemed acceptable at MV during IPHEx, and indeed generally if

attenuation is small enough that there is information content through the active depth of

precipitation as discussed in section 3.3.

The SRDC algorithm was applied to the IPHEx IOP data from 1 May 2014 and 15 June

2015, comprehending 2738 rain events (1-min time scale) and 35,693 min without rainfall.

To determine the width for the moving time window, a sensitivity study of the PDF of

precipitation occurrence using methods 1 and 2 for different time intervals of
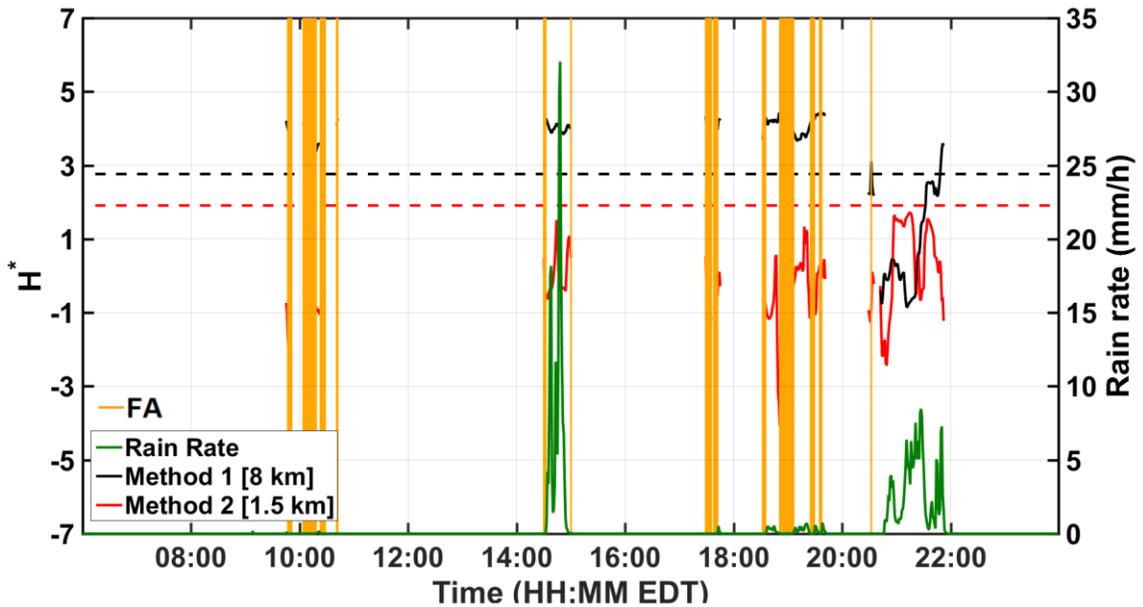
Figure 3-6: VCS functions computed from SRDC methods 1 and 2 for the 11 June 2014 event compared with the rain-rate recorded by a Parsivel disdrometer collocated with the ACHIEVE W-band radar and MRR at MV. Threshold for precipitation detection (red dotted lines) and classification (black dotted lines). SRDC algorithm detects precipitation when H_2^*(red solid line) is less than the detection threshold (red dotted horizontal line). FA occurrences (orange lines). SRDC algorithm classifies the event as deep when H_1^* (black solid line) is less than the classification threshold (black dotted horizontal line).

length N_t was performed, with the time-averaged values posted at the end of the

averaging window interval (i.e., t_0). The optimal window size (here, 15 min) produces the

best trade-off between the missed detection and false alarm (FA) errors as shown in Figure

3-5b. Note the increase and decrease in the false alarm rate (FAR) and missed detection

rate (MDR) as the width of the window increases. Figure 5c shows the PDF for the 15-min

temporal width to highlight the missed detection (MD) and false alarm statistics. The
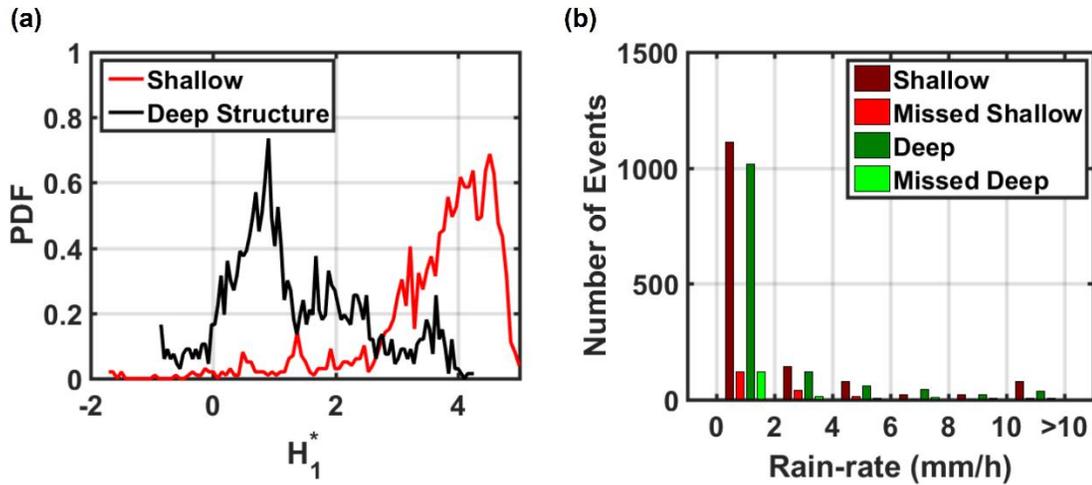
Figure 3-7: Probability Distribution Function of VCS metrics at MV. (a) PDF of 15-min method 1 VCS for classification of precipitation events at 1-min time scale. (b) Histogram of detection skill for different deep and shallow precipitation classes as a function of precipitation intensity during IPHEx IOP (1 May–15 Jun 2014) at MV.

percentage of MDR and FAR errors for a 15-min window (N_t = 15) width are 0.8% and

2.64%, respectively—the best compromise at MV. Method 1 performs well for the no-

precipitation cases as expected, whereas method 2 is skillful at detecting precipitation,

therefore enabling, in combination with method 1, the separation of deep and shallow

events.

The performance of the algorithm with the 15-min moving window for 11 June 2014 is

examined in detail next. The time series of VCS functions computed from method 1 and

method 2, and the corresponding time series of the rainfall intensity recorded by the

collocated Parsivel disdrometer are shown in Figure 3-6. Green and pink dotted lines identify the precipitation detection and classification thresholds, respectively. Most of the precipitation events are detected by tracking the VCS function values, after applying the reflectivity threshold criteria. Column VCS H_1^* and shallow VCS H_2^* are higher than the precipitation detection threshold for noprecipitation conditions, whereas shallow VCS H_2^* is lower than the detection threshold for precipitation events. For example, the algorithm detects the start of the D1 event 2 min in advance because of averaging artifacts, resulting in two false alarm counts. Similarly, method 1 VCS values above and below the classification threshold are identified as shallow and deep, respectively. Inspection of the PDFs of precipitation and no-precipitation cases detected by the algorithm (Figure 3-5c) indicates that $H_{det}^* = 1.91$, the point of intersection between the two curves, is appropriate for rainfall detection.

The next step is to classify the precipitation events based on vertical structure. The PDFs of the VCS functions of shallow and deep structure events are shown in Figure 3-7a, where $H_{cla}^* = 2.76$ is the point of intersection. At MV, the ambiguity in the classification of deep structure precipitation is 14.0%, whereas for shallow or low-level rainfall the value is close to 12.0%. The H_{det}^* and H_{cla}^* thresholds were used to differentiate between no-precipitation, deep, and shallow precipitation events over the duration of the IPHEX-IOP,

and the corresponding contingency matrix is shown in Table 3-1. It is evident from Figure

3-7(b) that many shallow events observed at MV during the IPHEx IOP had very high

precipitation intensity compared to that of deep structure events. This is an indicator of

Table 3-1: Contingency table for precipitation detection and classification by SRDC methods 1 and 2 for profiles observed during IPHEX IOP at MV.

		Classified as		
		Deep (%)	Shallow (%)	No precipitation(%)
Actual	Deep	85.3	14.4	0.3
	Shallow	12.3	79.0	8.7
	No Precipitation	4.0	2.6	93.4

the robustness of the algorithm that captures the reverse orographic enhancement effects

associated with seeder–feeder interactions among stratiform rainfall, low-level clouds,

and deep fog in the inner mountain region (Wilson and Barros 2014, 2015, 2017). The

algorithm misses shallow precipitation events with rain rates less than 5mm h^{-1}, though

most missed events are for rain rates less than 2 mmh^{-1} based on the disdrometer

observations often not recorded by the collocated rain gauges, and thus approaching

instrumental sensitivity. In addition, some misclassification errors result in part from

temporal offsets at the beginning and ending of rainy periods, which can be addressed

moving the windowed VCS values from the end (t_0 = 15th minute) to the center of the

window (t_0 = 8th minute). In the case of deep structure precipitation, events with rain rates

greater than 8mm h^{-1} are missed for the IPHEx-IOP due to the use of W-band for column

calculations constrained by the field setup as discussed earlier, and the statistical sample

is also small because of the lack of convective activity during the IPHEx IOP.

(a)

(b)
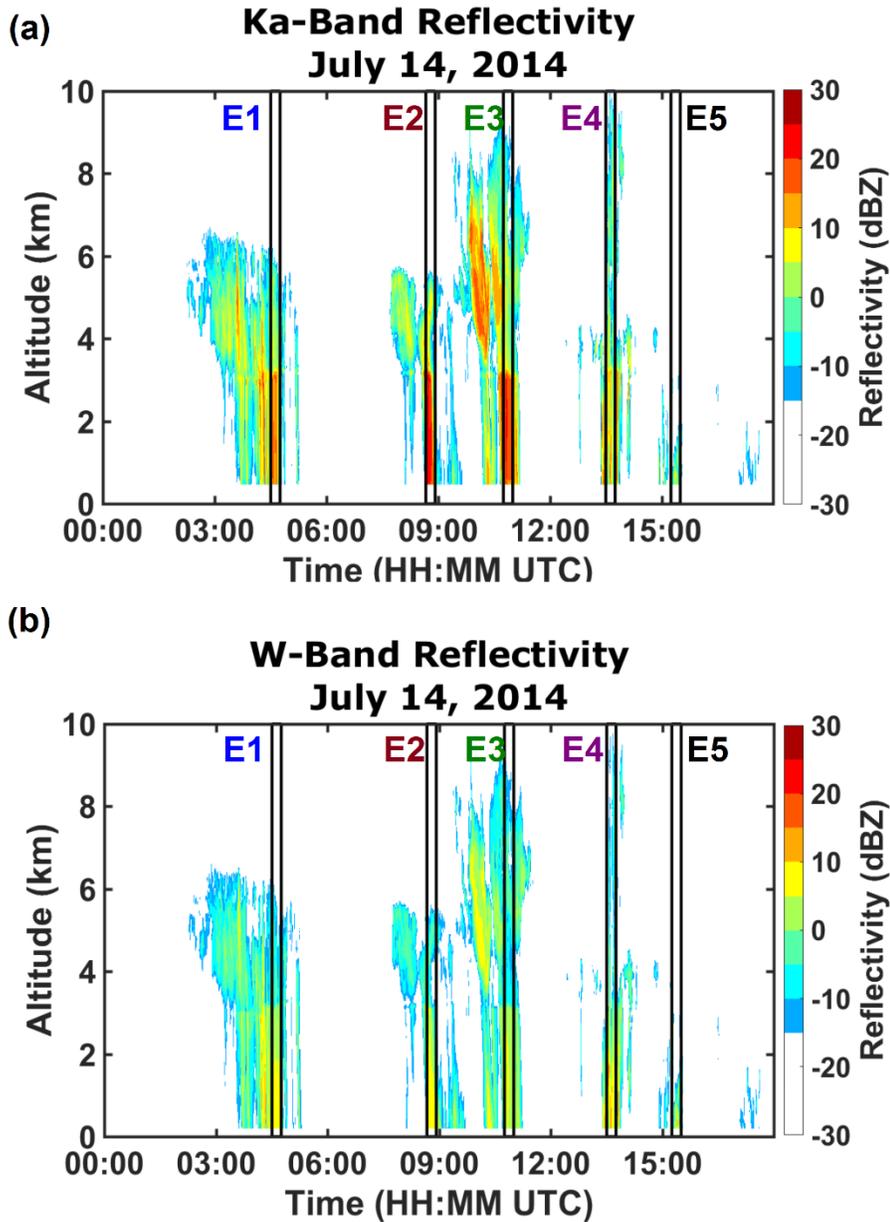
Figure 3-8: Equivalent reflectivity factor of (a) Ka-ZR and (b)MW ACR observed at ARM TMP on 14 Jul 2014. Events considered for analysis (black boxes).

3.2.2 ARM-TMP

The observations at ARM TMP exhibit frequent ''deep'' non convective

precipitation structures that are quite different compared to MV. Indeed, ARM TMP

reflectivity profiles for deep structure events generally show enhanced reflectivity below
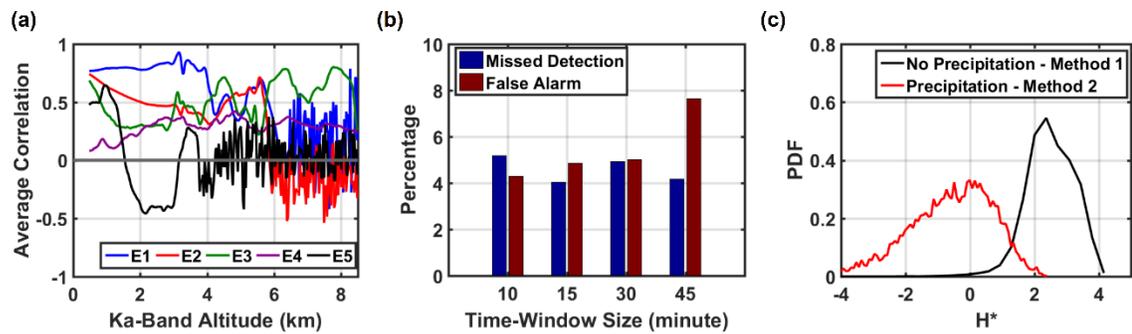
Figure 3-9: Space-time correlation, optimal time-window and VCS metrics for ARM-TMP. (a) Method 1 average space–time correlation structure computed in step 2 of the algorithm for the five events highlighted in Figure 3-8 (E1–04:30 to 04:45 UTC; E2–08:40 to 08:55 UTC; E3–10:45 to 11:00 UTC; E4–13:30 to 13:45 UTC; E5–15:15 to 15:30 UTC) in Finland on 14 Jul 2014. (b) Comparison of MDR and FAR errors (expressed as a percentage of the total number of events) as a function of the time window width. (c) PDF of the optimum (15 min) moving time window for the computation of VCS by the SRDC algorithm.

the melting layer as a result of downward transport of ice from aloft, the so called dark-

band effect (see, e.g., Fig. 9 in Kollias and Albrecht 2005). For this reason, the altitude

threshold to distinguish shallow and deep structure events is lower (z_c = 4 km). To account

for the significant differences in precipitation physics, method 1 is applied here to the Ka-

band reflectivity profiles up to z_1 = 8.5 km and method 2 is computed using z_2 = 1.5 km

profiles for both Ka- and W- bands. The maximum elevation for method 1 is selected based

72

on the regional climatology of typical precipitation structure and clouds. Studies based on models and observations suggest that the maximum CTH closely depends on the height of the tropopause, and cloud tops in the high latitudes do not exceed altitudes of 8–9 km (Pan et al. 2010; Pan and Munchak 2011).
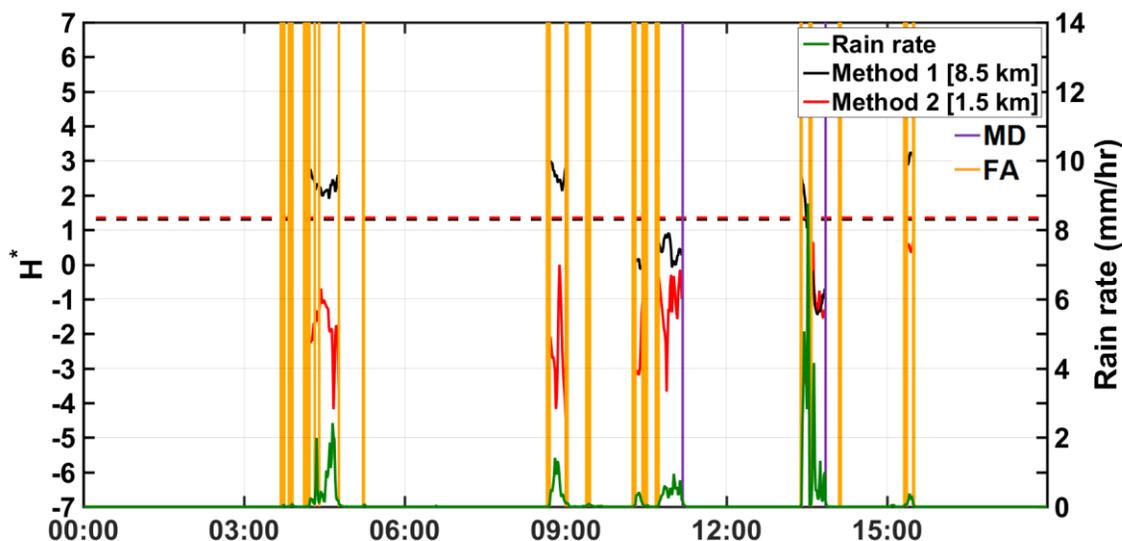
Figure 3-10: Rain rate observed by 2DVD compared with VCS computed by methods 1 and 2 of the algorithm at ARM TMP on 14 Jul 2014. Threshold for precipitation detection (red dotted lines) and classification (black dotted lines). FA occurrences (orange lines) and MD occurrences (purples lines).

The 14 July 2014 event selected for detailed examination exhibits intermittent shallow and deep structure precipitation from 0300 to 1600 UTC (Figure 3-8). Out of the five precipitation events (E1: 0430–0445 UTC, E2: 0840– 0855 UTC, E3: 1045–1100 UTC, E4: 1330–1345 UTC, and E5: 1515–1530 UTC) detected, E2 and E5 are classified as shallow events. Space–time correlations for the individual events are displayed in Figure 3-9(a). Note how the average correlation pattern is quite similar to that obtained at MV (Figure

3-5a), with smooth average correlation patterns for deep events below z_c with a significant

increase in reflectivity above. For the shallow events, there are large noisy oscillations in

the correlation value above z_c. The full ARM TMP dataset collected from 6 May 2014 until

11 September 2014 was used in the analysis corresponding to 10,391 min of precipitation

and 82,352 min of non-rainy conditions. The next step was to find the appropriate time

width for the moving window to optimize precipitation detection. Figure 3-9b

summarizes the FAR and MDR errors, and Figure 3-9(c) shows the PDFs of precipitating

occurrence for optimal window sizes. Again, as in MV, a 15-min interval achieves the best

error trade-off. The algorithm's application for the 14 July 2014 event with the 15-min

window, and the variations of the VCS patterns obtained from method 1 and method 2

are shown in Figure 3-10. Note that the FA and MD occur at either the beginning or the

end of the event. The VCS values for method 1 show the expected variations between

clear-sky and deep structure precipitation events, while method 2 captures both shallow

and deep precipitation as in MV.
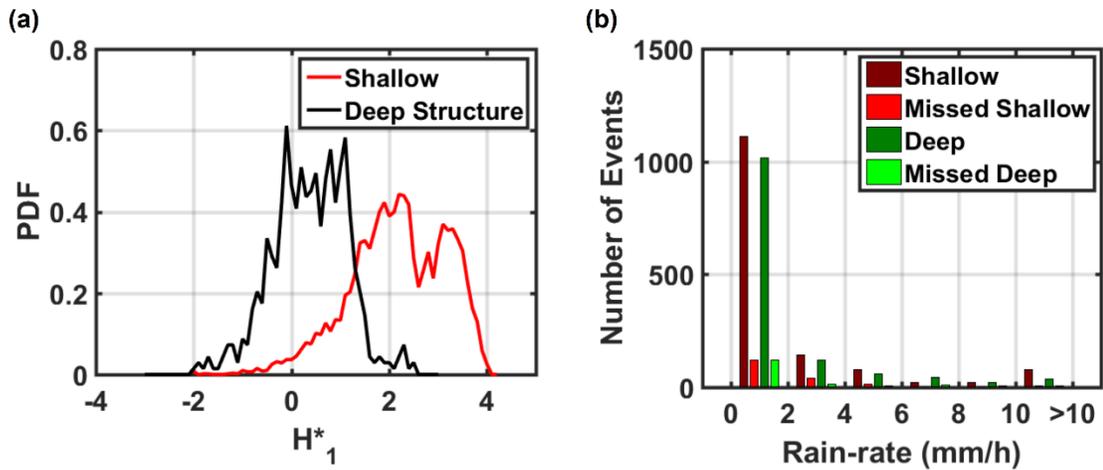
Figure 3-11: Probability Distribution Function of VCS metrics at ARM-TMP. (a) PDF of 15-min method 1 VCS values for classification of precipitation events at 1-min time scale. (b) Histogram of detection skill for different deep and shallow precipitation classes as a function of precipitation intensity during BAECC (May 2014–September 2014) at ARM TMP.

When contrasting the precipitation detection PDFs obtained for ARM TMP (Figure 3-9c) and MV (Figure 3-5c), it is apparent that method 1 VCS for clear-sky conditions is very different for the two locations. By contrast, the PDF of method 2 VCS at the ARM TMP site (Figure 3-9c, red curve) closely resembles the PDF obtained for MV (Figure 3-5c, red curve). The differences in H_1^* PDFs of noprecipitation in MV and ARM TMP explain the different H_{det}^* values for each location. This is attributed to the different operational setup of the Ka-band radars and differences in the vertical structure of light rainfall between the two hydrometeorological regimes as well.

Table 3-2: Contingency table for precipitation detection and classification by SRDC methods 1 and 2 for profiles observed at ARM TMP

		Classified as		
		Deep (%)	Shallow (%)	No precipitation (%)
Actual	Deep	91.7	8.2	0.1
	Shallow	17.0	78.4	4.6
	No Precipitation	4.3	4.8	90.9

The detection threshold for ARM TMP is $H_{det}^* = 1.36$ with MDR and FAR of 3.9% and 4.5%, respectively (Figure 3-9c). A similar approach was followed for the precipitation classification, and the threshold for precipitation classification can be obtained from the PDFs in Figure 3-11a ($H_{cla}^* = 1.3$). These thresholds are used to compute the contingency matrix shown in Table 3-2. Shallow precipitation events are associated with higher ambiguity corresponding to a misclassification error of 17%, whereas deep structure precipitation is misclassified 8% of the time.

A comparison among the histograms of precipitation intensity of actual and detected shallow and deep structure precipitation events is presented in Figure 3-11(b). Overall, the rain-rate PDF observed at ARM TMP is similar to MV with most of the observed precipitation being light rainfall with intensities below 2 mmh^{-1}. Note that, if dark band cases were classified as shallow (not shown), then the VCS threshold value for

classification is H_{cla}^* = 1.5, yielding higher misclassification errors of 7.3% and 16.1% for deep and shallow systems, respectively.

3.4.3 ARM-SGP

Observations from the ARM-SGP Central facility in Lamont, Oklahoma were used to test the SRCD algorithm under a broader range of storm regimes. The W-band Scanning ARM Cloud Radar (W-SACR) placed at SGP has been operating intermittently for the past 5 years. A total of thirty days of collocated W-SACR and Ka-ZR data (8 days between April 20, 2011 and June 12, 2011 and 22 days from June 4, 2012 to June 26, 2012) are considered for this analysis corresponding to 149 minutes of precipitation events. Most of the precipitation events considered during this period of analysis were short-lived (duration less than 2 hours). Three cases of isolated deep structure events showed strong reflectivity values extending up to 10 km. The shallow events were generally restricted to z_c = 5 km. The SGP observations exhibit entirely different hydrometeorological properties and precipitation structure compared to ARM-TMP and MV. For this case, the vertical range for the Ka-Band profile was z_1 = 10 km and z_2 = 2 km for the W-band profile in the column VCS calculations. These vertical ranges are consistent with Clothiaux et al. (2000) who developed climatology of cloud heights and radar reflectivities at the ARM SGP site and found that most of the hydrometeors are observed between 1 and 10 km.

(a)

Ka-Band Reflectivity
July 08, 2012

(b)

W-Band Reflectivity
July 08, 2012
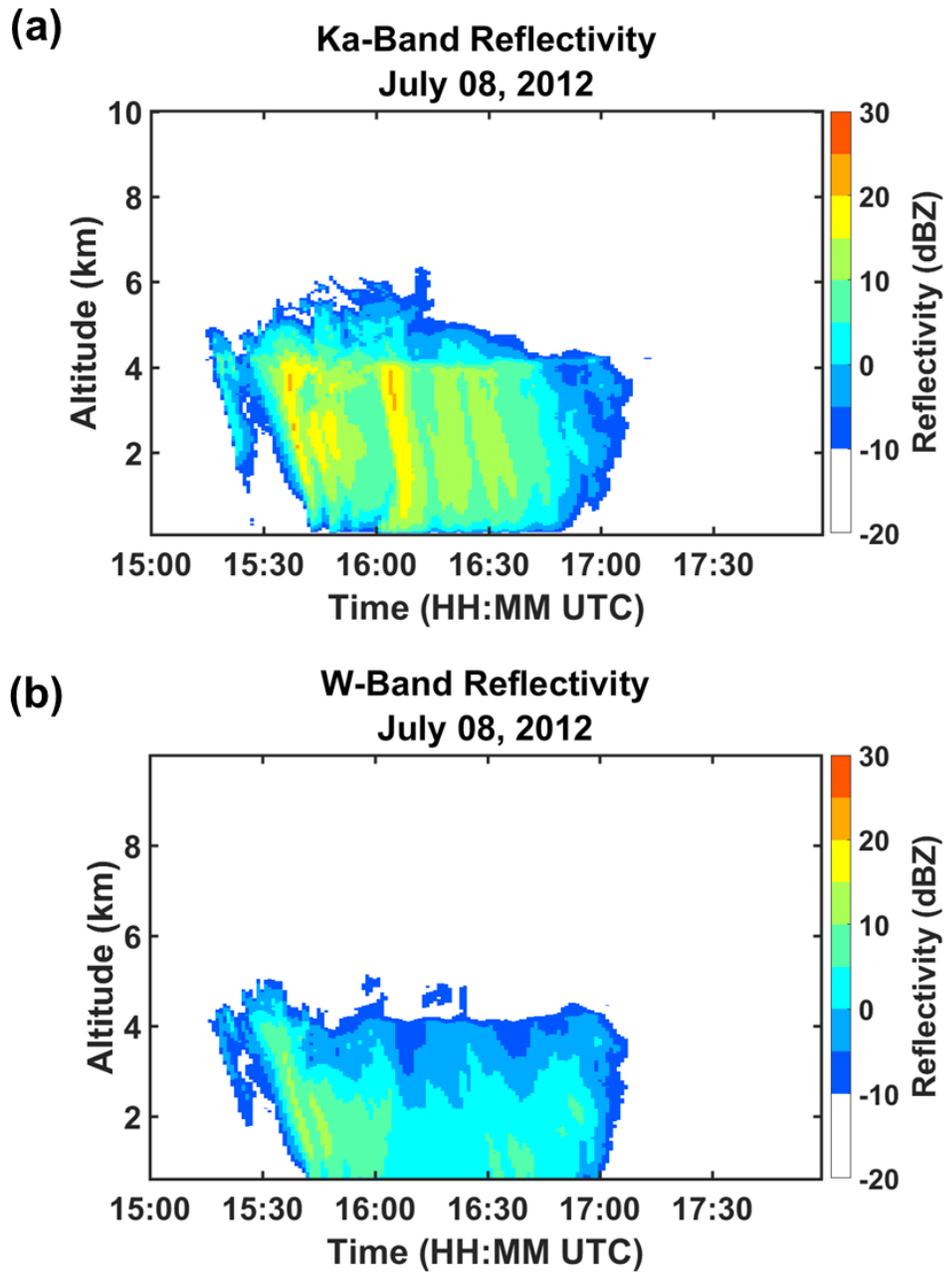
Figure 3-12 Equivalent Reflectivity factor of (a) Ka-Band Zenith Radar and (b) W-Band Scanning ARM Cloud Radar observed at ARM-SGP in Oklahoma, USA, on July 08, 2012
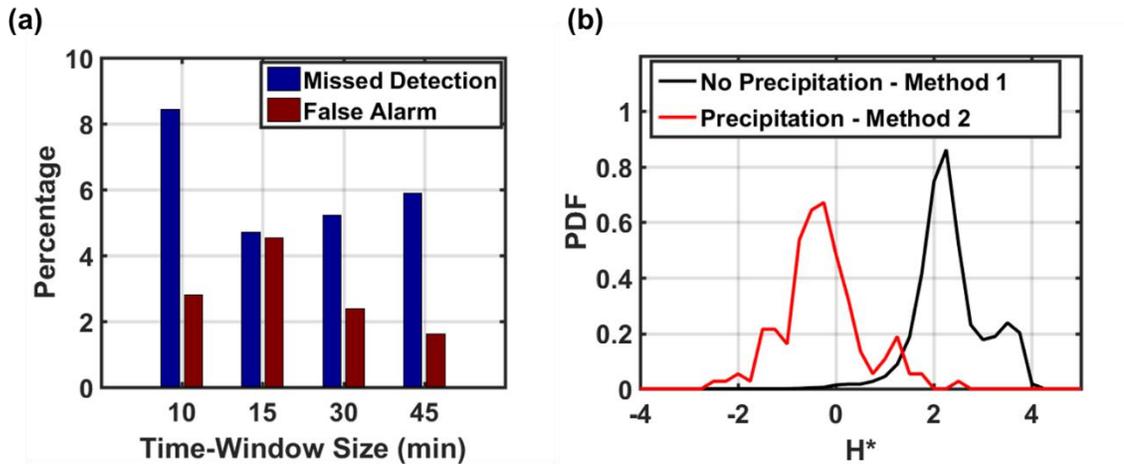
Figure 3-13 (a) Performance of the SRDC precipitation detection algorithm with varying window-size. (b) PDF of precipitation detection for the ARM-SGP in Oklahoma, USA.

Figure 3-12 shows the W- and Ka-band radar reflectivity profiles for a shallow precipitation event observed at the SGP on July 8, 2012. Detailed analysis similar to that presented for MV and ARM-TMP was repeated for this case. The optimal time-window computation showed similar results with 15-minute duration providing the best trade-off between missed detection and false alarms (Figure 3-13a). The PDFs of precipitation detection for the 15-minute moving window are shown in Figure 3-13(b). The threshold value for the precipitation detection from modified entropies calculated following Method-1 and -2 is $H^*_{det}=1.3$. Precipitation detection fails for 7 minutes only out of the 149 mins tested here, corresponding to 4.7% error; the false alarms amount to 4.54% error.
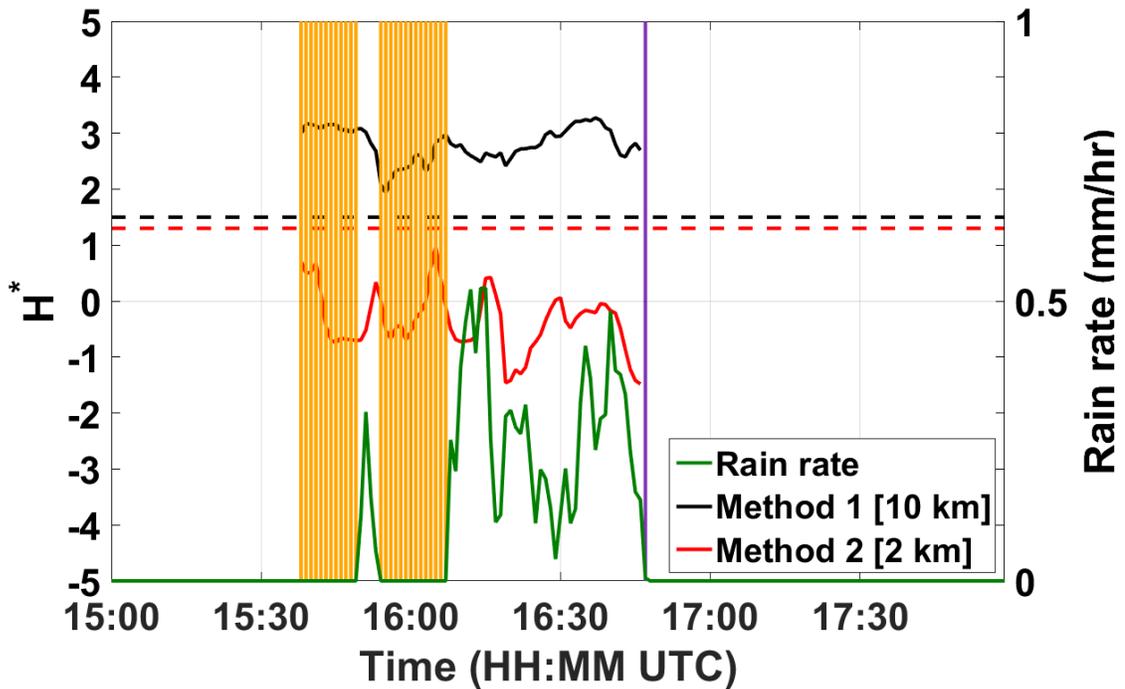
Figure 3-14: Rain-rate compared with VCS computed by SRDC Method-1 and Method 2 of the algorithm at the ARM-SGP in Oklahoma, USA, on July 8, 2012. The red and black dotted lines denote the threshold for precipitation detection and classification, respectively. The orange and purple lines mark false alarm (FA) and missed detection (MD) occurrences.

Figure 3-14 shows the VCS time-series for the July 8, 2012 event. VCS computed according to Method-1 fails to detect rainfall, while the opposite is true for Method-2. The precipitation detection algorithm accurately detects 42 deep-structured and 100 shallow events (events with vertical extent below z_c = 5 km). The column VCS (Method-1) is used for the classification of precipitation events, and Figure 3-15(a) shows the PDF of the precipitation classification. The precipitation classification algorithm misses 7 deep-structured precipitation events and no shallow event using H^*_{cla} = 1.5 as the threshold value. Table 3-3 shows the contingency matrix derived using the detection and

classification thresholds. The precipitation rate distributions of shallow and deep-structure events are shown in Figure 3-15(b). Around 75% of the shallow and deep events considered for this study are light rainfall with precipitation rate less than 2 mm/h, while 15% of the light, deep-structured precipitation events were missed by the algorithm. The classification skill at ARM-SGP for the deep structured events was better than for shallow events for MV and ARM-TMP. We attribute the contrasting behavior at ARM-SGP to the small sample size of quality controlled data available at the time this study was conducted and the specific period of study.
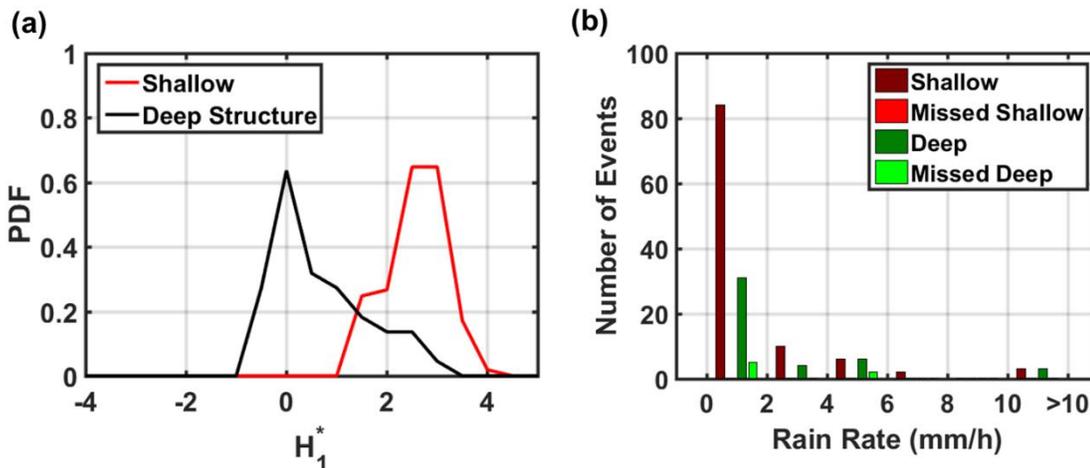
Figure 3-15 Probability Distribution Function of VCS metrics at ARM-TMP. (a) PDF of 15-min Method-1 VCS values for classification of precipitation events at 1-min time-scale; (b) Histogram of detection skill for different deep and shallow precipitation classes as a function of precipitation intensity at the ARM-SGP site in Oklahoma, US.

3.4.4 GPM DPR and CloudSat CPR G-SDSU Simulations

The G-SDSU was used to simulate the radar equivalent reflectivity profiles observed by GPM DPR and CloudSat CPR at 1855 UTC 15 May 2014 in the SAM (Wilson and Barros

2015, 2017). Simulated reflectivity profiles along the coinciding satellite overpasses for the GPM DPR and CloudSat CPR are shown in Figure 3-16. The maximum spatial distance between the GPM and CloudSat overpass pixels is 0.19268, and over 94% of the pixels are collocated at the model resolution of 1.25 km. The simulated GPM DPR (Ka-band) and CloudSat CPR (W-band) reflectivity profiles are the inputs into the SRDC algorithm toward evaluating the performance for satellite-based (nadir looking) radar at the same resolution. A terrain elevation mask was applied to the G-SDSU-simulated GPM-DPR and CloudSat-CPR profiles, and the analysis is conducted only for reflectivity profiles above the terrain.

Table 3-3: Contingency table for precipitation detection and classification by SRDC Method 1 and 2 for profiles observed at ARM-SGP

		Classified as		
		Deep	Shallow	No Precipitation
Actual	Deep	79.6 %	13.6 %	6.8 %
	Shallow	0 %	95.2 %	4.8 %
	No Precipitation	7.2 %	0.2 %	92.6%

The simulated reflectivity profiles for GPM DPR and CloudSat CPR exhibit deep and shallow precipitation structures with heavy and light rainfall, respectively. Note the difference between the downward-looking perspective from satellite-based sensors and the upward-looking perspective in the case of ground-based sensors (MV, TMP, and SGP)

is well illustrated by the severe attenuation of simulated CloudSat CPR reflectivity in the

lower troposphere for the deep convective event with heavy rainfall between 32.5° and

33°N.
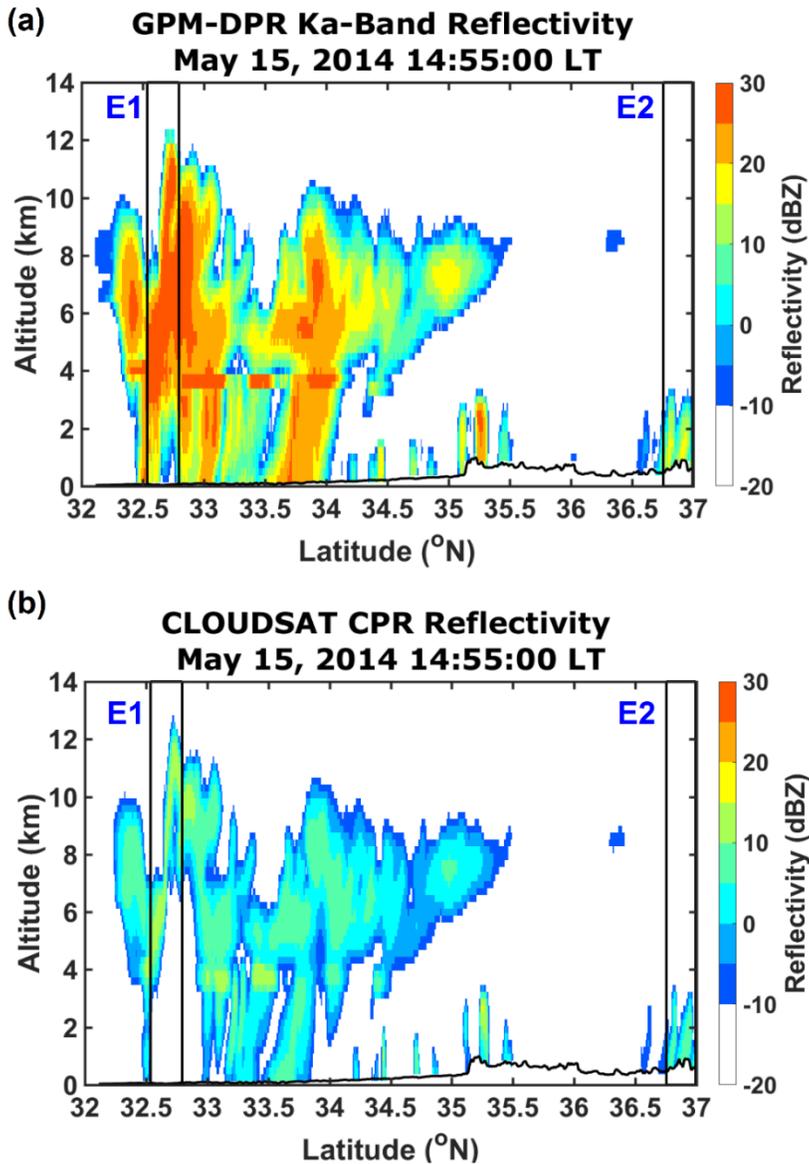
Figure 3-16: Equivalent reflectivity factor simulated by the G-SDSU for a concurrent overpass of (a) GPM DPR and (b) CloudSat CPR for 15 May 2014 event during IPHEx. Events E1 and E2 considered for analysis (black boxes).

Attenuation artifacts are still present in the case of the convective rainfall between 32.5°

and 33° N (E1), but not so for the light rainfall event between 36.5° and 37.5° N (E2).

The simulated reflectivity profiles were input into the SRDC algorithm using a spatial

moving average window width of five consecutive grid points (spatial length of 6 km) to

calculate the correlation between the two frequencies along the CloudSat daytime

overpass trajectory—that is the algorithm applied here by trading local time for space

along the satellite overpass, with the pixel number along the overpass tract used as the

spatial coordinate. Figure 3-17 shows the comparison of low-level (method 2) and column

VCS (method 1). The G-SDSU does not simulate ground clutter artifacts that are present

in the satellite-based observational datasets. Thus, the low-level height for method 2 is

taken as 1 km AGL, which is the active depth of low level seeder–feeder processes in the

region (Wilson and Barros 2014, 2015, 2017; Duan et al. 2015); to capture the attenuation

effects at high frequencies (e.g. CPR), the depth of the column for method 1 is taken as 8

km AGL. Indeed, the impact of CPR attenuation is apparent in the large negative values

of low-level VCS for the cases of heavy rainfall (from left to right until latitude 34 N).

Beyond 34N latitude, shallow rainfall is captured by the method 2 VCS (note the feature

between 36.5 and 37.5 N). The detection and classification thresholds were not computed

for this limited case study (an extensive OSSE study would be required for statistical

robustness). Nevertheless, applying the zero-crossing criterion to the method 1
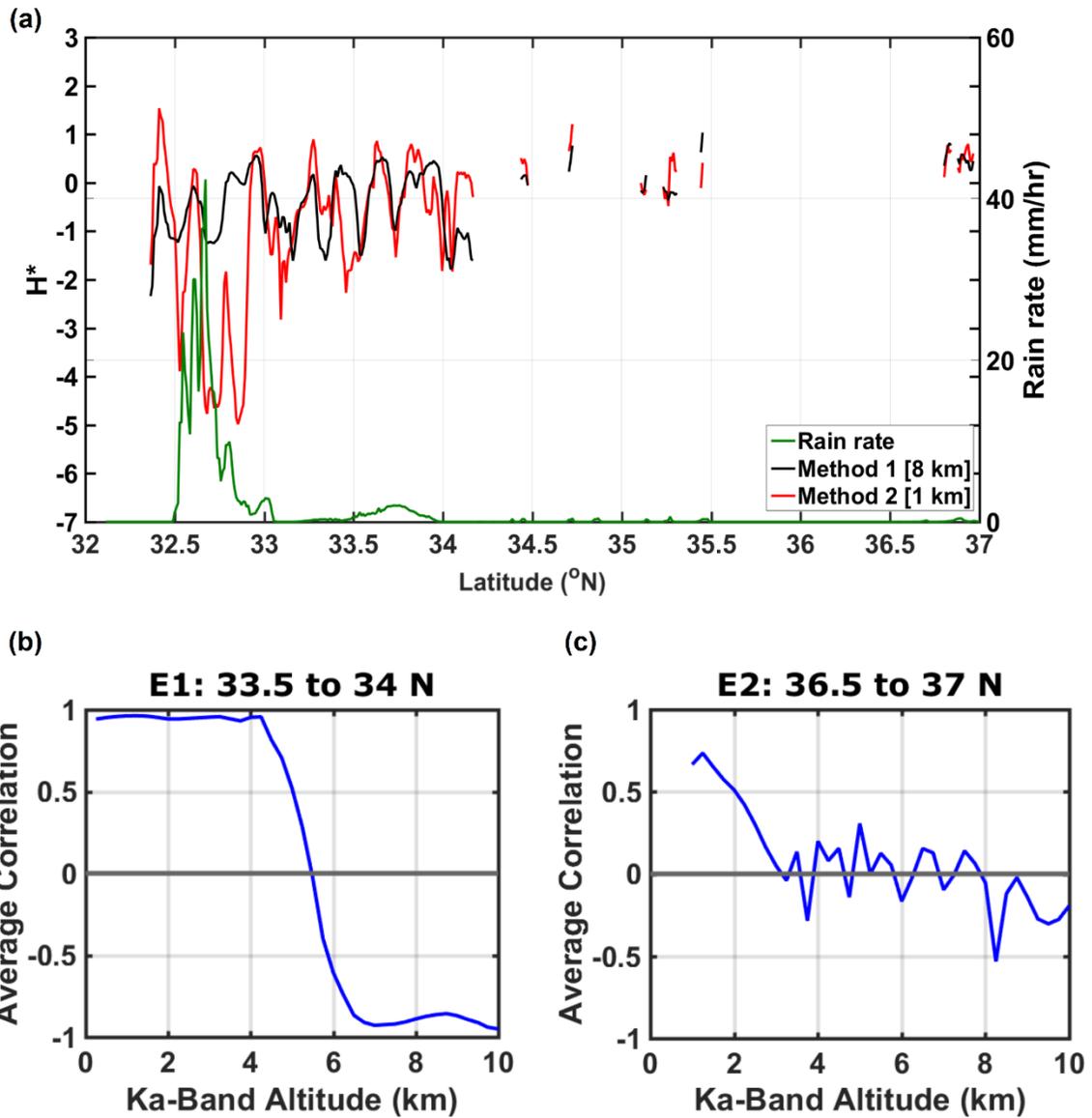
Figure 3-17: Rain rate simulated by WRF compared with 15-min VCS computed by SRDC methods 1 and 2 for the 15 May 2014 event. Method 1 average correlation profile computed from the space–time correlation DFC, where (b) E1: 32.5–33N and (c) E2: 36.5–37N. E1 is a deep precipitation event, while E2 is a shallow light rainfall event with precipitation depth below 4 km. Average correlation curves for these events show different behavior and zero-crossing altitudes consistent with different precipitation structures.
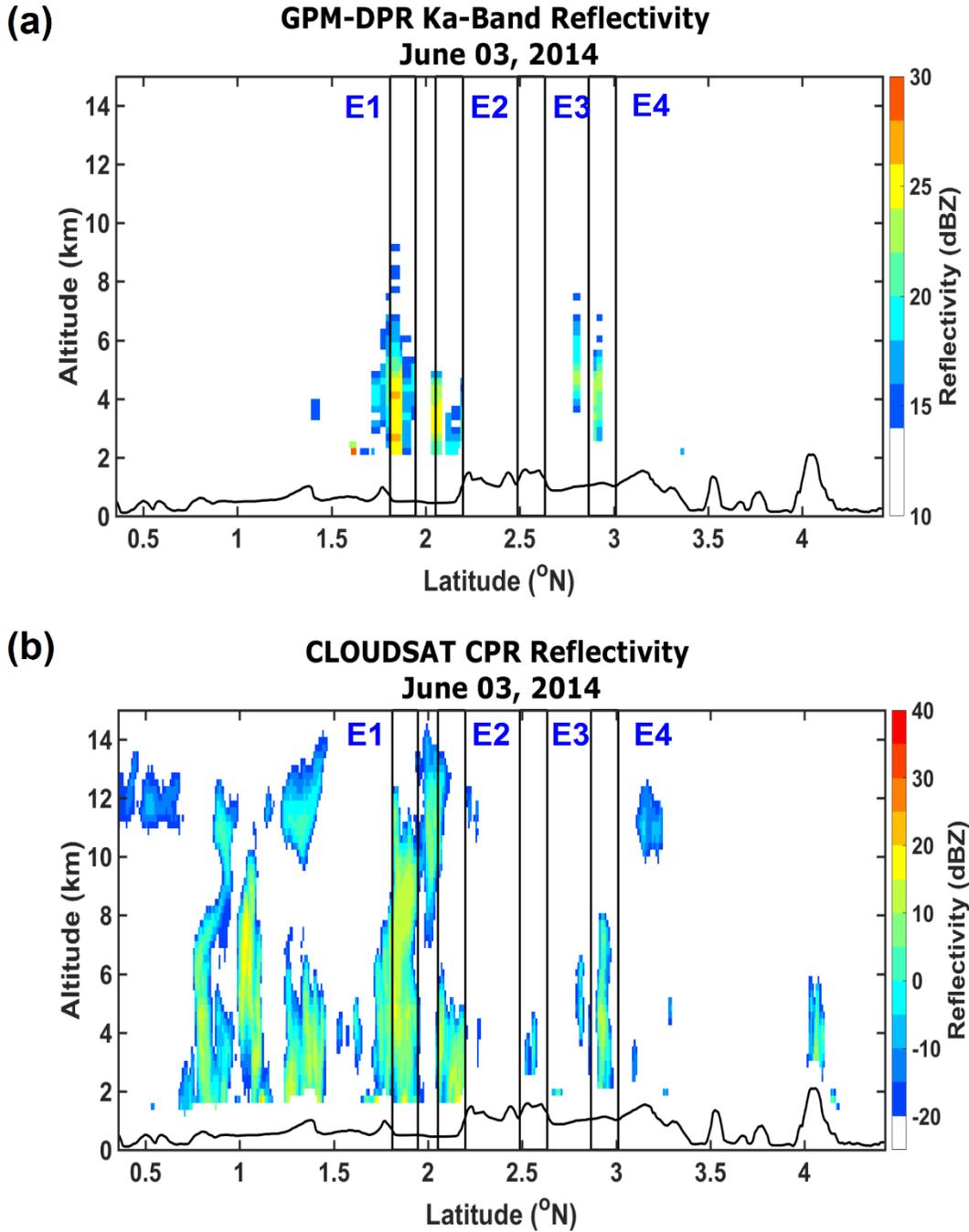
Figure 3-18: Equivalent reflectivity factor observed by concurrent overpasses of (a) GPM DPR Ka-band radar and (b) CloudSat CPR radar from the combined 2BCSATDPR product on 3 Jun 2014 over Borneo (1158E). Events considered for the analysis (black boxes).

average correlation to identify the active depth of rainfall yields results that are consistent with the ground-based observations, as shown in Figure 3-17(b) and Figure 3-18(c) for E1 and E2, respectively. This case study is indicative of the potential applicability of the SRDC algorithm at least for nadir- or near-nadir-looking satellite-borne radar systems. It also highlights the need to further characterize the impact of attenuation at high frequencies, as well as the need for large datasets to establish operational metrics, such as the VCS thresholds estimated for the ground-based applications. The performance of the algorithm using actual GPM DPR and CloudSat CPR observations is illustrated next.

3.4.5 GPM DPR and CloudSat CPR Observations

Reflectivity profiles observed by CloudSat CPR and GPM DPR Ka-band on 3 June 2014 over complex terrain in Borneo (115°E) from the Kapuas Mountains to the South China Sea (see Figure 3-1 for map) were extracted to demonstrate the application of the SRDC algorithm to observations from space-based platforms (Figure 3-18). Because the reflectivity profiles are affected strongly by ground clutter, the digital elevation model (DEM) mask included in the CloudSat CPR products was used to remove the radar bins in the lowest 2 km AGL. The GPM DPR Ka-band radar overpass coincides with the CloudSat CPR between 1.5° and 3.5°N, and all the pixels are within the resolution of GPM DPR products (5 km). The time difference between the two overpasses is approximately 7 min. Missing values above 9 km AGL in the 2ADPR V04A HS product and low-level values affected by ground clutter were replaced with a Gaussian random variable of mean

0 and variance 0.01. The ground clutter values in the CloudSat CPR reflectivity product were replaced by a Gaussian random variable of mean 0 and variance 0.01. The mean values for the missing value correction of GPM DPR and CloudSat CPR were chosen based on the minimum detectable threshold of the instruments. Low-level VCS (method 2) values are computed for 4 km profiles of the two satellite products. Column VCS (method 1) values are calculated using maximum heights of 12 and 4 km AGL for GPM DPR and CloudSat CPR respectively. The spatial window chosen for the analysis is approximately 10 km, corresponding to 10 CloudSat-CPR and two GPM-DPR co-aligned pixels. The joint temporal evolution of the VCS and the precipitation rate estimates derived from Ka-band radar (2ADPR V04A HS product) is examined next.

Table 3-4: Detection and classification VCS thresholds used in the SRDC algorithm for MV, ARM TMP, and ARM SGP.

Location	H_{det}^*	H_{det}^*
MV	1.91	2.76
ARM TMP	1.36	1.3 (1.5)
ARM SGP	1.3	1.5
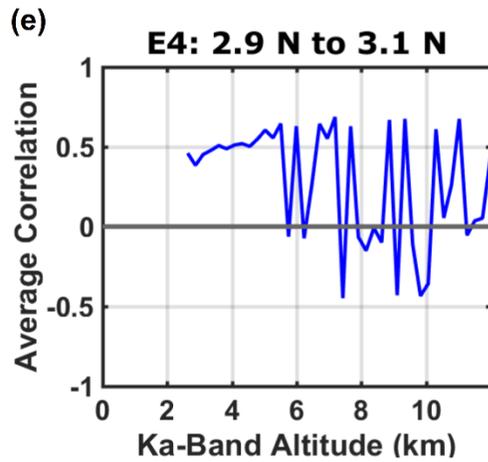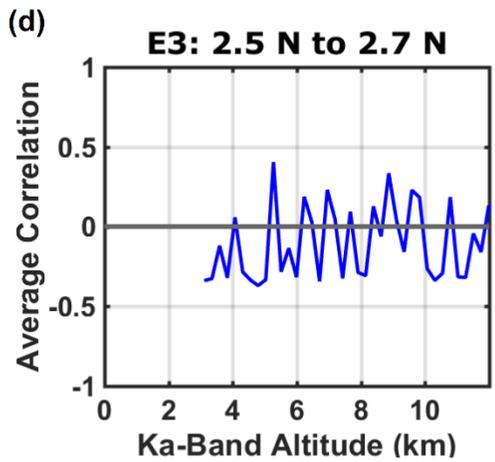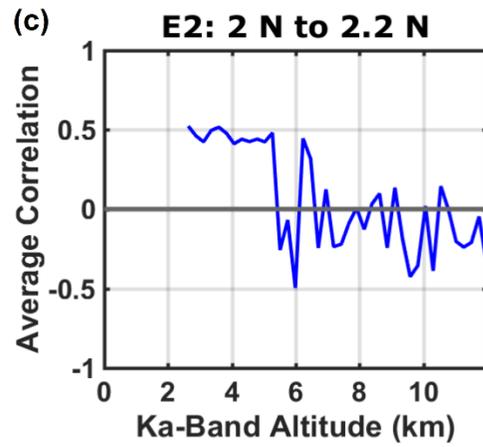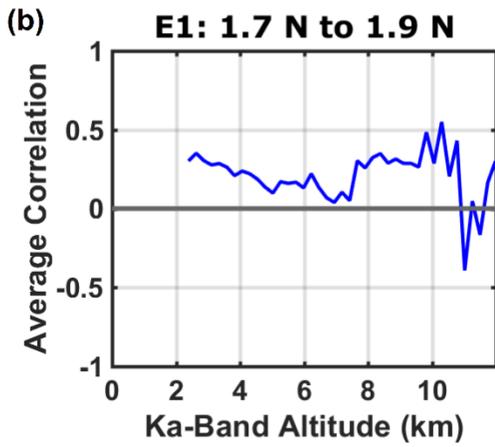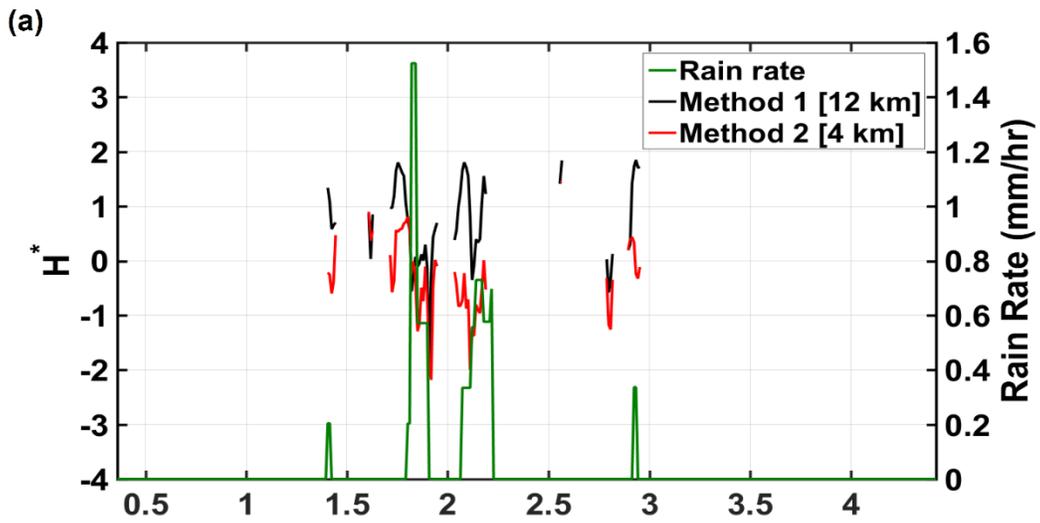
Figure 3-19: VCS functions computed by SRDC methods 1 and 2 for the reflectivity profiles and rain rates from the GPMDPR 2A product. Method 1 average correlation profile computed from the space–time correlation computed in step 2 of the algorithm, where (b) E1: 1.7°–2°N, (c) E2: 2°–2.2°N, (d) E3: 2.5°–2.7°N, and (e) E4: 2.9°–3.1°N. E1 and E4 are deep structure events, E2 is a shallow event, and E3 is a no precipitation event. Note the change in the shape and altitude of first zero-crossing points of the correlation profiles based on the structure of precipitation.

The deep structure events observed between 1.7° and 1.9°N (E1), and 2.9° and 3.1°N (E4) are captured by both method 1 and method 2, while the shallow structure between 2 and 2.2N (E2) is captured only by method 2 as expected (Figure 3-19a). As in the previous case study, the detection and classification thresholds are not computed because of the lack of enough samples for robust statistics. Note how the CloudSat CPR observes a shallow cloud system between 2.5 and 2.7N (E3) that is missed by the GPM DPR Ka-band. The method 1 (column) average correlation patterns for the four highlighted events (E1–E4) are presented in Figure 3-19(b–e). Like for ground-based observations at MV and ARM-TMP, the first zero crossing of the average correlation is a robust criterion to determine the depth of precipitation structures from space. Indeed, the SRDC detects most of the precipitation features observed where and when concurrent observations exist, despite differences in instrument sensitivity and observing geometry. Exploratory applications of the SRDC using Ku-band NS instead of Ka-band HS (Appendix C). The results are consistent for Ka- and Ku, although the zero-crossing heights are different as expected with Ku-band being higher by less than 1 km. This indicates that because attenuation is lower for Ku-band then the potential to retrieve vertical structure information from Ku-W is even better than for Ka-W.

3.5 Conclusions

This study presents a dual-frequency radar algorithm for shallow rainfall detection and classification (SRDC). The algorithm aims to provide an efficient way to combine multiple radar observations, taking advantage of different operating frequencies to characterize the vertical structure of precipitation using an integrated VCS metric to describe the evolution of the space–time correlation structure of radar reflectivity profiles. The SRDC formulation is applicable to any pair of frequencies with contrasts in attenuation and scattering behavior, such as, for example, replacing Ka band with either Ku, or X, or C band. Indeed, performance should be expected to improve using frequencies lower than Ka band. The focus here is on W and Ka bands because of data availability constraints only.

The SRDC was tested using collocated ground-based W- and Ka-band radar reflectivity observations with a focus on warm rainfall events at the surface only. Snowfall at the ground was not included because the vertical structure of reflectivity (and the underlying microphysics) for snow events above and below 0°C is out of the scope of this work. Method 1 and method 2 VCS are used to detect and classify precipitation events. However, the detection and classification thresholds can be estimated with confidence only when large collocated datasets at different frequencies are available. The behavior of the first zero crossover illustrates that method 1 and method 2 work as intended and are immune to the exact z_c location and the correlation variability. Further, over a given

geographic region, when the algorithm is first applied, the z_c can be used to aid in the classification until robust VCS statistics are determined. The algorithm was applied first to radar profiles obtained from the IPHEx IOP ground validation campaign in MV with successful detection of precipitation 99% of the time. Further, the algorithm performed well in distinguishing between shallow and deep precipitation in the complex terrain of the SAM, even with column VCS calculated using W-band radar instead of Ka-band radar, which exhibits less attenuation at low levels because of setup constraints.

An independent evaluation was conducted at the ARM TMP in Finland. Furthermore, a third independent application of the SRDC was carried out at the ARM SGP in the central United States. The SGP is representative of the mid-latitude continental plains, while the TMP site was located close to and in between land–water margins and coastlines. Consequently, the maximum vertical range considered for the computation of column VCS (method 1) is quite different, reflecting climatic differences in precipitation climatology. Nonetheless, the MDR and FAR obtained from the precipitation detection algorithm for both TMP and SGP agree with each other and are approximately 4.5% and 0.5%, respectively. The classification algorithm provides better classification skill for deep structures compared to shallow events for MV and ARM TMP. However, the opposite is true in the SGP case, although only three cases (42 min of rainfall) were available at the time when the study was conducted. The detection and classification threshold values using W-band as support for the column correlation calculations (method 1) are very close

in both TMP and SGP (Table 3-4), when both sensors were operating with a similar set up

and in contrast to MV, which supports the general applicability of the algorithm.
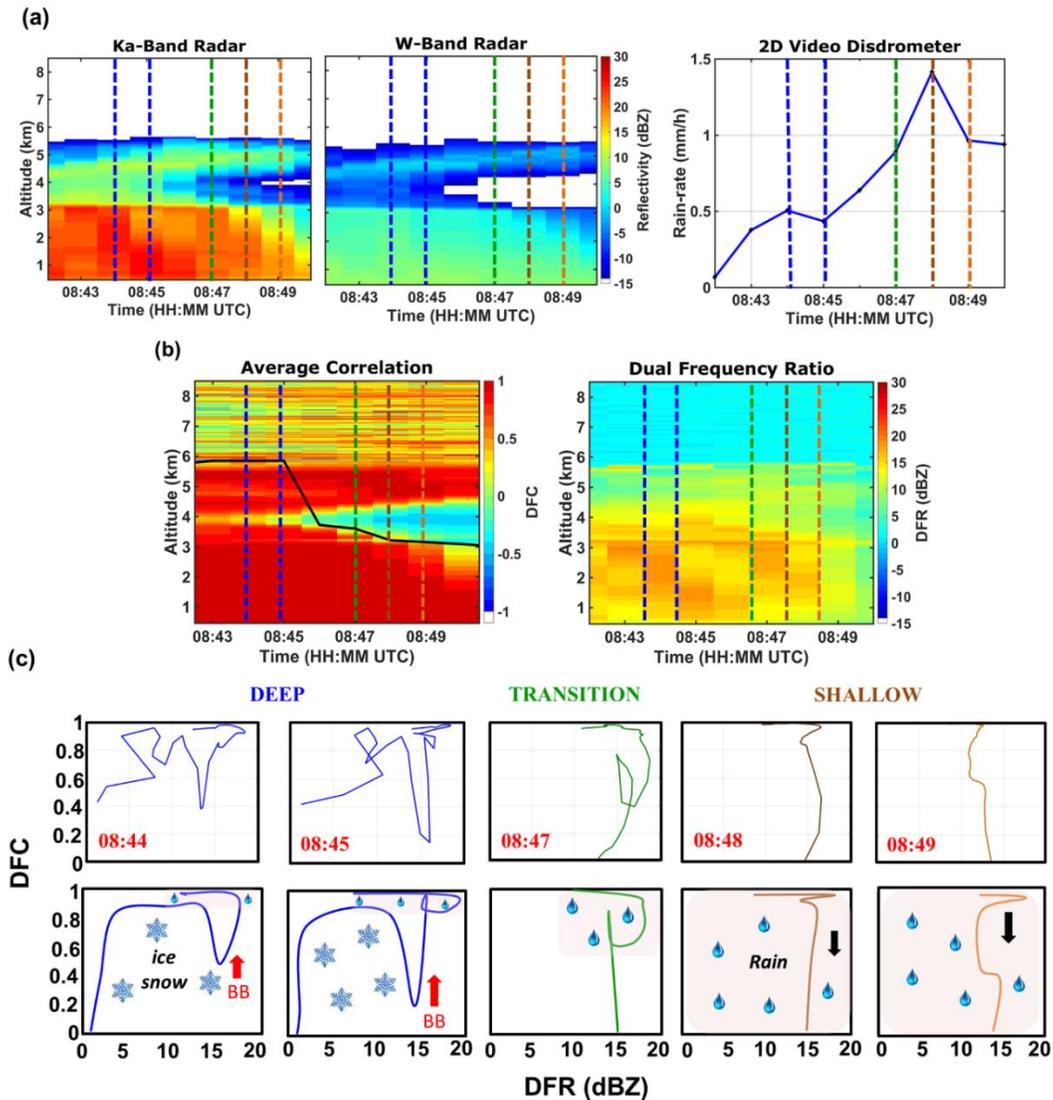
Figure 3-20: Physical Basis of SRDC using dual-frequency correlation and dual frequency ratio: (a) (left to right) Observed Ka- and W-band reflectivity, and surface rainfall rate at ARM TMP on 14 Jul 2014, including a transition from deep structure to shallow structure rainfall (see also Figure 3-8). (b) (left to right) DFC and DFR profiles computed from the observations. (c) (top) Temporal evolution of DFC–DFR trajectories at the times marked by the dashed lines of the same color in (a) and (b). (bottom) Schematic generalization of evolution of the relationship between DFR and

DFC at select times from (two leftmost panels) the deep structure stage of the event as the bright-band effect increases and the drop size distribution broadens near the surface, (middle) through the transition phase without bright-band signature and heavy surface rainfall, and (two rightmost panels) into the shallow structure stage of the event as it weakens and vertical depletion of a large drop becomes apparent. In (a), profiles in the deep structure (blue line), transition (green line), and shallow (brown line) structure stages of the events.

The SRDC physical basis can be further examined in light of the phase-space trajectories of concurrent dual-frequency correlation (DFC) and dual-frequency ratio [DFR ;z_{Ka} (dBZ)– Z_W (dBZ)] below the first zero crossing height. This is illustrated in Figure 3-20 using 8 min of observations during the 14 July 2014 event at the ARM TMP site (see Figure 3-8). Ka- and W-band reflectivity profiles exhibit the signature of deep structure precipitation, including mixed-phase processes in the first 3 min (08:43–08:45 UTC) followed by a transition to shallow structure warm rainfall in the next 2 min (Figure 3-20). The DFC and DFR profiles are shown in Figure 3-20 followed by an analysis of DFC–DFR relationships in Figure 3-20(c). In Figure 3-20(c), left panels, note the strengthening of the bright band (red arrows) and its corresponding signature in the DFC–DFR phase space (low DFC and high DFR excursion in the bottom-right-hand corner), both of which are accompanied by intensification of low-level rainfall captured by the DFR protrusion at high DFC values (top-right corner of the phase space). As the event evolves into shallow structure precipitation, the bright band disappears in the middle panels, the low-level protrusion deepens initially as a result of large melted hydrometeors from above, and eventually recedes (black arrows) as the column is selectively depleted of large raindrops

94

first. The right panels show how the DFR progressively decreases (and DSD narrows) from the surface upward (from high to low DFC). The DFC–DFR analysis supports therefore the notion of VCS thresholds (H_{det}^* and H_{cla}^*) that should be precipitation regime dependent, not location dependent. Nevertheless, fingerprints of regional weather require adaptive VCS thresholds to correct for artifacts that may introduce ambiguity errors. For example, the presence of dark-band effects in the TMP observations prompted the lowering of the classification VCS threshold with significant improvement in the SRDC skill statistics. Other relevant examples include the attenuation at high frequencies for heavy rainfall in the tropics along with bright-band artifacts.

4. Improving Quantitative Precipitation Estimates in Mountainous Regions by Modeling Low-Level Seeder-Feeder Interactions Constrained by Global Precipitation Measurement Dual-Frequency Precipitation Radar Measurements

4.1 Introduction

Reliable estimates of orographic precipitation are critical to understand the water cycle including hydrological feedbacks, bio-geochemical cycles, and regional and global climatic variations (e.g. Bales et al. 2006; Barros 2013). The classical notion of orographic precipitation refers to the increase of precipitation with elevation (orographic enhancement) as approaching moist air is forced upward, reaches its lifting condensation level, forms clouds and produces rain, thus maintaining strong spatial precipitation gradients modulated by elevation. In reality, orographic precipitation is controlled by multiple factors such as regional weather patterns affecting moisture convergence and stability conditions, and multiscale circulations tied to the diurnal cycle of solar forcing at landform scale (Barros and Lettenmaier 1994; Houze 2012; Wilson and Barros 2017). Bergeron (1960a, b) noted that rain-gauges at low elevations adjacent to small hills systematically registered higher precipitation during the passage of storm systems, and proposed seeder-feeder interactions (SFI) to explain the observed enhancement. SFI is the process by which clouds aloft (seeder clouds) precipitate over low-level clouds, orographic cap clouds and fog banks (feeder clouds) thereby enhancing drop coalescence efficiency, and increasing the precipitation-rate near the surface. Subsequently, many

96

studies confirmed the role of SFI in small-mid range mountains around the world through modelling and ground-based measurements (Hill et al. 1981; Storebo 1976; Passarelli and Boehme 1983; Robichaud and Austin 1987; Purdy et al. 2005). In the Southern Appalachian Mountains (SAM), low-level clouds and fog (LLCF) SFI can enhance the surface rainfall intensity by a factor of 5-10 leading to high spatial and temporal variability due to persistent low-level moisture convergence in the presence of both weak and strong synoptic forcing (Wilson and Barros 2014, 2015 and 2017; Duan and Barros 2017). Generally, SFI should play a similar role on the climatology of precipitation albeit of different magnitude depending on regional conditions wherever LLCF are a prominent feature of the diurnal cycle of moist processes (e.g. middle-and coastal mountains).

Spatial and temporal organization of LLCF in the SAM leads to high spatial heterogeneity in the low-level microstructure of rainfall and consequently in near-surface precipitation-rates (Wilson and Barros 2017; Duan and Barros 2017). Satellite-based Quantitative Precipitation Estimates (QPE) typically misses or underestimates low-level, layered and bottom-heavy precipitation systems due to ground-clutter contamination of reflectivity profiles in the lower 1-2 km, as well as non-uniform beam filling (NUBF) artifacts when the spatial scale of precipitation systems is smaller than the radar or sensor resolution (relatively coarse). This explains the robust space-time organization of retrieval errors in Tropical Rain Measurement Mission (TRMM) – Precipitation Radar (PR; spatial resolution ~ 5 km) products tied to the topography, and the diurnal and seasonal cycles

of precipitation systems in the SAM (Prat and Barros 2010a; Duan et al. 2015). In addition to QPE errors due to sensor specifications and measurement artifacts, radar-QPE relies on simplified drop size distribution (DSD) models that cannot be well constrained from reflectivity observations and thus are a major source of retrieval error and uncertainty (Liao et al. 2014).

Numerical modelling studies were conducted previously to characterize orographic precipitation microphysics and to quantify the measurement uncertainty associated with disdrometer resolution and detection limits (Prat et al. 2008; Prat and Barros 2010b; Wilson and Barros 2014). Wilson and Barros (2014) used the Duke rain microphysics column model (DRCM; Prat et al. 2012) to examine the impact of large numbers of small drops (50 μm and 0.1 mm) from fog and/or cap clouds on DSD evolution that results in significant decreases in the mass-weighted mean diameter D_m and mass spectrum standard deviation σ_m (Appendix D, Eqns. D2 and D3 respectively). Limitations in the detection and counting of very small drops lead to high bias in the estimation of rain-rates as well as large errors in radar-based precipitation and DSD parameter retrievals, thus highlighting the need to study the full DSD spectra from μm-mm scale (Barros et al. 2008; Prat et al. 2008; Prat and Barros 2010b; Wilson and Barros 2014; Williams et al. 2014; Angulo-Martinez and Barros 2015; Thurai et al. 2017; Duan and Barros 2017). Duan and Barros (2017) and Thurai et al. (2017) utilized the Meteorological Particle Spectrometer (MPS; Baumgardner et al. 2002) to study the precipitation processes

using the full DSD spectra observations from μm-mm scale. Thurai et al. (2017) explored the effect of considering submillimeter scale drop-diameter observations on DSD parameters using collocated MPS and 2D-video disdrometer (2DVD; Schönhuber et al. 2008) measurements. They showed that the normalized intercept parameter N_w (Appendix D, Eqn. D6) and the ratio σ_m/D_m are highly underestimated by the 2DVD since the small drops are neglected or underestimated consistent with Prat et al. (2008) and Wilson and Barros (2014). The MPS used by Duan and Barros (2017) was built to meet finer resolution specifications than the standard instrument (e.g. 25 μm vs 50 μm), and they used the MPS DSD spectra prior to rainfall onset to specify LLCF microphysics (the feeder cloud) in the DRCM whereas seeder microphysics aloft were derived from collocated Micro Rain Radar (MRR; Peters et al. 2002) reflectivity profiles. The model successfully estimated surface rain-rates in the western foothills of the SAM, and highlighted the importance of SFI on the temporal evolution of the vertical DSD profiles especially for drop diameters between 0.7 and 2 mm

The focus of the present study is on quantifying the impact of low-level orographic precipitation processes on the vertical structure of DSD inferred from satellite-based PR measurements, specifically the Ku-PR on GPM. The specific objectives are: 1) to characterize the impact of SFI on DSD parameters and to examine the signature of these changes in Ku-PR corrected reflectivity profiles; and 2) to use the findings from (1) toward demonstrating a physics-based framework to improve satellite-based orographic QPE

and microphysical estimates affected by SFI and describe their associated uncertainty. The application region is the SAM due to the availability of long-term spatially dense ground validation (GV) networks of rain-gauges and Parsivel disdrometers (Figure 3-1 a-b) that were installed to support Integrated Precipitation and Hydrology Experiment (IPHEx; Barros et al. 2014). Nevertheless, the methods are strictly physics-based and thus of general applicability for similar precipitation regimes.

The chapter is organized as follows. GV and remote sensing data used in the study are described in Section 2. The methodology for error analysis, a brief description of the stochastic rain microphysics model, and the proposed physics-based modelling framework to correct for the low-level SFI enhancement in satellite-based radar retrievals of orographic precipitation are described in Section 3. Section 4 presents the results from error analysis of GPM Ku-PR retrievals relative to GV data to elucidate the physical basis of detection and estimation errors. The application of the physically-based correction framework is detailed in Section 5 including step-by-step demonstration, and systematic application to Ku-PR underestimation cases in the SAM since GPM launch in 2014. Potential caveats and limitations as well as potential for generalized application of the proposed methodology are discussed in Section 6, followed by conclusions in Section 7.
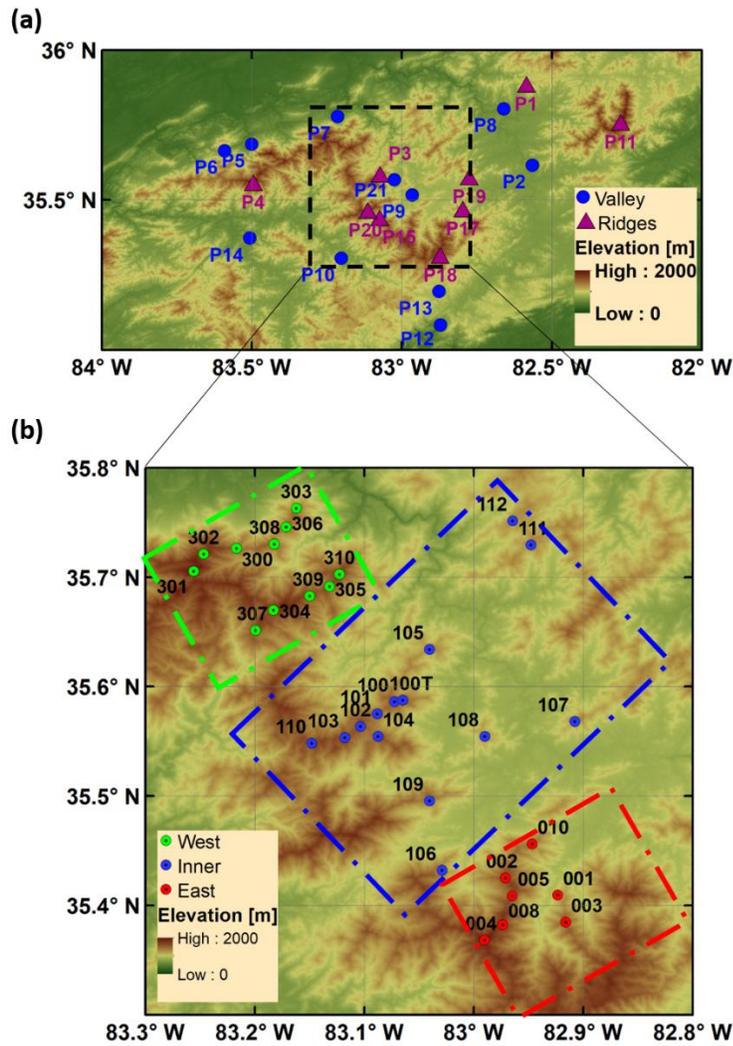
Figure 4-1 Digital Elevation Model of the Southern Appalachian Mountains (SAM) with locations of (a) Parsivel disdrometers and (b) long-term rain-gauge network. The green, blue and red boxes denote western, inner and eastern region.

4.2 Data

4.2.1 Ground Valid Observations

A long-term spatially dense tipping bucket rain-gauge network has been operating since 2007 in the Great Smoky Mountains National Park within the SAM (Prat and Barros 2010; Barros et al. 2014 and 2017). The geographical locations of the rain-gauges are shown in Figure 4-1(b) and the details of the locations and data considered for analysis are shown in the Appendix E (Table E1). Different types of tipping bucket rain gauges were operated in 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 the differences in tipping resolution. The study region is categorized into three different regions according to hydrometeorology and precipitation regimes: Eastern, Inner and Western regions. The most frequent precipitation systems in the eastern ridges are tropical storms and south-easterly convective storms, with SFI among passing systems (convective and/or frontal) and LLCF dominating the inner region especially at mid-day (Wilson and Barros 2014, 2015). Westerly mesoscale convective systems are predominantly observed in the western ridges of the SAM, along with Southeasterly atmospheric rivers fueled by Gulf of Mexico (Miller et al. 2019). In addition, the water-cycle in the study area is primarily driven by LLCF and light precipitation (< 3 mm/h). Previously, the physical basis of detection and

estimation error analysis in TRMM precipitation estimates was examined and the main challenges in TRMM-PR QPE were attributed to ground-clutter, NUBF artifacts and failure to detect light precipitation (Prat and Barros 2010a; Duan et al. 2015; Arulraj and Barros 2017).

During IPHEx (Barros et al. 2014) Intensive Observation and Extended Observation Periods (respectively, IOP and EOP), surface observations of DSDs were collected using 20 OTT Parsivel-P1 and P2 disdrometers (Loffler-Mang and Joss 2000) between May-1 and December-1, 2014. The disdrometer locations are marked in Figure 4-1(a), and the coordinates are provided in the Appendix E (Table E2). The OTT-Parsivel is a laser-based optical disdrometer that simultaneously measures and classifies the drops into 32 diameter-size and 32 velocity bins. The mid-value of the diameter bins ranges from 0.062 mm to 24.5 mm, and the bin interval varies between 0.125 mm for smaller drops to 3 mm for larger drops. The bins with drop-diameter < 0.25 mm and greater than 9 mm were cleared due to the sensitivity of the instrument. In addition, solid precipitation data are not considered for analysis in this study. The terminal velocity is used to compute the DSD to eliminate instrument measured fall velocity artifacts following Tokay et al. (2014). The drop velocity bins range between 0 and 22.4 ms^{-1} with bin interval ranging between 0.1 ms^{-1} and 3.2 ms^{-1}. The measurement sampling interval is 10s and the data are averaged to 1-minute time interval. In addition to the collocated Parsivel disdrometers and rain-gauges, some sites were also equipped with MRRs. The MRR manufactured by Metek,

operates at 24 GHz in a vertically pointing mode resolving 30 vertical gates. Antialiasing techniques are not applied on the MRR data used in this study. After IPHEx, between September 25 and November 6, 2015, a fog-trailer equipped with collocated rain-gauge, MPS, MRR and OTT Parsivel disdrometers was deployed in the western foothills (Elkmont; P6 in Figure 4-1a) to study precipitation microphysics. MPS is also a laser-based sensor that was custom-designed for the fog-trailer to measure the DSD in the 25 μm to 1.5 mm range with drop-size bin resolution of 25 μm and 1 Hz sampling frequency. Comparison (not shown) of the fog-trailer MPS and Parsivel disdrometer drop counts at 1-min time-scale shows underestimation by the MPS for bins with diameter > 0.3 mm and suggest underestimation by the Parsivel disdrometer for diameters < 0.3 mm as predicted by Prat et al. (2008).

4.2.2 Satellite Observations

The GPM satellite has onboard a DPR operating at Ku- (13.6 GHz) and Ka- (35.5 GHz) bands. GPM covers the spatial region approximately between latitudes ± 67º and operates at three different modes (Iguchi et al. 2017). The Ku-PR operates in normal scan (NS) mode with a cross-track swath width of 245 km and sampling resolution of 125 m, range resolution of 250 m and spatial resolution of approximately 5 km (Hou et al. 2014 and Iguchi et al. 2017). The viewing angle varies from nadir (0º) to ±18º (Iguchi et al. 2017). Hou et al (2014) states that the minimum detectable precipitation rate of Ku-PR is 0.5 mm/h; however, in practice, the Ku-PR has detectability similar to Ka-PR (~0.2 mm/h;

Speirs et al. 2017). The Ka-PR operates at matched scans (MS) and high sensitivity (HS) modes. The MS mode captures the near-nadir pixels with 25 beams matching with the Ku-PR scans and the spatial and vertical resolution is same as the Ku-PR NS. In HS mode, the Ka-PR is sensitive to light precipitation and snowfall with minimum detectable precipitation rate of 0.2 mm/h and its beams are interlaced with the MS beams. The HS mode sampling interval is 250 m and the range resolution is 500 m, which is coarser than the NS and MS modes, and the horizontal resolution is 5 km. Note that the GPM DPR and TRMM PR have been proven to exhibit higher sensitivity to light precipitation than the official detectability of 0.2 mm/h and 0.5 mm/h respectively (Hamada and Takayabu, 2016).

The Level-2 GPM products consist of DSD parameters estimates such as the D_m and N_w. In a single frequency retrieval algorithm, the apparent reflectivity factor (Z_m; mm^6mm^{-1}) is corrected for attenuation to derive the effective radar reflectivity (Z_e, mm^6mm^{-1}):

$$Z_m(r) = A(r)Z_e(r) \tag{4-1}$$

where A(r) [unitless] is the attenuation factor due to precipitating and non-precipitating particles at height r (Iguchi et al. 2017). The attenuation corrected reflectivity factor Z_e is dependent on D_m and N_w:

$$Z_e = N_w F(D_m) \tag{4-2a}$$

$F(D_m)$ [mm^7] is defined as follows:

$$F(D_m) = \frac{\lambda^4}{\pi^5} \left| \frac{n_1^2+2}{n_1^2-1} \right|^2 \int_{D=0}^{\infty} \sigma_b(D) \, f(\mu) \left(\frac{D}{D_m} \right)^{\mu} exp \left(-\frac{(4+\mu)D}{D_m} \right) dD \qquad \text{(4-2b)}$$

where σ_b is the backscattering cross section [mm^2], λ is the wavelength of the sensor [mm] and n_1 is the refractive index of water in liquid phase at temperature T. First, D_m will be estimated using Z_e at two wavelengths. Finally, N_w is estimated using D_m and Z_e using Eqn. 2(a) (Seto et al. 2013). For liquid hydrometeors, the normalized gamma distribution function (Appendix D, Eqn. D4) is assumed with fixed shape factor μ= 3.0 (Iguchi et al. 2017; Seto et al. 2013; Liao et al. 2014). The GPM Ku-PR version 05A data are used in this study since Ku-PR covers a larger swath compared to Ka-PR. Due to the off-nadir viewing angle, the Ku-PR no-clutter height is very high, and thus the utility of near-surface measurements is not reliable for angles within the outer swath.

4.3 Methods

4.3.1 Error Metrics

Previous research established robust evidence from ground- and satellite-based sensors linking the diurnal cycle of QPE errors to the diurnal cycle of LLCF, where LLCF is tied to near-surface enhancement of precipitation (Prat and Barros, 2010b; Wilson and Barros, 2014, 2015 and 2017; Duan and Barros, 2017). GPM Ku-PR near-surface precipitation estimates are compared against observations from GV tipping bucket rain-gauges following the protocol in Duan et al. (2015) and Prat and Barros (2010a). GPM Ku-PR has a spatial resolution of 5 km and any rain-gauge located within a 2.5 km radius from the center of the radar beam are used in the comparison for the corresponding GPM

Ku-PR pixel. If more than one rain-gauge lies within a radar beam, then each of the rain-gauges is considered as an independent sample for error analysis. Errors statistics were also calculated using the average rain-gauge observations within the radar beam. Orographic precipitation exhibits high spatial variability within the spatial resolution of GPM Ku-PR and by assuming each rain-gauge is an independent sample, the spatial heterogeneity of rainfall is represented to some degree. To capture light rainfall cases consistent with collocated disdrometer observations, the rain-gauge tip data are first converted to rain-rate [mmh^{-1}] at every time-stamp (e.g. for each tip) by calculating the rainfall accumulation within an hour period centered at the time-stamp. In this context, the minimum rain-rate for rain-gauges with 0.1-, 0.2- and 1.0-mm tipping resolution is 0.1, 0.2 and 1.0 mmh^{-1} respectively. Based on Prat and Barros (2010a) and Duan et al. (2015), the intercomparison is conducted for the rain-rates obtained by averaging hourly rainfall accumulation tips within the time-window lengths depending on the rain-gauge measurement resolution: 10-minute for rain-gauges with 0.1- and 0.2-mm tipping resolution, and 30-minute integration for 1.0 mm tipping resolution. The minimum detectable rain-rates will not differ since averaging is performed with respect to the number of tips within the time-window. The averaging time-window lengths are determined for different tipping-resolution rain-gauges based on the best detection metrics and bias to account for the difference in comparing time-averaged observations with instantaneous GPM Ku-PR estimations (Prat and Barros, 2010a; Duan et al. 2015).

Nevertheless, it is important to highlight that uncertainty inherent to gaps in spatial (point vs area) and temporal (instantaneous vs time-averaged) scales between satellite-based measurements and rain-gauge measurements remains. Studies suggest that the direct comparison between point-scale gauge observations and areal radar estimates penalizes the radar (Raupach and Berne, 2016), but it is difficult to establish conclusive performance metrics over large areas because of the space-time variability of radar operations, spatial intermittency due to LLCF, orographic effects and storm dynamics.

Two types of detection errors were evaluated – False Alarms (FA) and Missed Detections (MD). FA are the cases when GPM Ku-PR detects precipitation and rain-gauges did not record any rain. MD are the cases where GPM Ku-PR missed the precipitation recorded by the rain-gauges. Standard detection metrics such as the Probability of Detection (POD; Wilks 2011), False Alarm Ratio (FAR; Wilks 2011) and Critical Success Index (CSI; Schaefer, 1990) are used to evaluate the performance of GPM Ku-PR precipitation estimations in the SAM. POD is the probability of precipitation detection by satellite-based estimates provided the ground-based observations detect precipitation. POD varies between 0 and 1 while the prefect score is 1. POD is calculated as,

$$POD = \frac{YY}{YY+NY} \tag{4-3}$$

where YY is the number of overpasses when both GPM Ku-PR and collocated rain-gauges detect precipitation. NY is the number of missed detection overpasses. FAR is the

probability of false alarm occurrences given the GPM Ku-PR detects precipitation. FAR also varies between 0 and 1, however, the perfect score is 0:

$$FAR = \frac{YN}{YY+YN} \tag{4-4}$$

where YN is the number of false alarm overpasses. CSI is also known as Threat Score. CSI measures the ratio of concurrent correct detections of precipitation by both GPM Ku-PR and rain-gauges to the total number of overpasses for which either GPM Ku-PR or ground-based gauges detect precipitation. The perfect score for CSI is 1, and it varies between 0 and 1:

$$CSI = \frac{YY}{YY+YN+NY} \tag{4-5}$$

Estimation errors are the bias in the estimation of the rain-rate between GPM Ku-PR and rain-gauges. The relative estimation bias (ε) is computed when both the rain-gauge and GPM Ku-PR detect precipitation and is calculated as:

$$\varepsilon = \frac{RR_{GPM}-RR_{RG}}{RR_{RG}} \tag{4-6}$$

where RR_{GPM} is the instantaneous near-surface rain-rate (mm/h) estimated by GPM Ku-PR and RR_{RG} is the rain-rate (mm/h) recorded by the rain-gauge. The overpasses are classified as overestimation (OVR), underestimation (UND) and correct detection (CD) based on the estimation bias. If the relative bias value lies between ±0.5, the overpass is identified as a *correct* detection as proposed by Prat and Barros (2010a) and subsequently confirmed by Duan et al. (2015). Thus, in the context of this dissertation, a correct detection is not necessarily an accurate detection, but rather a detection that may have a

moderate discrepancy (< 50% error) between GV point measurements and satellite-based areal estimates. If the relative bias is less than -0.5, the satellite estimate is classified as UND. If the relative bias is greater than 0.5, the satellite estimate is classified as OVR. The generous CD criterion aims to recognize the space-time scale gap between rain-gauges and the DPR providing for a representative distribution of errors among the three classes to focus on large errors as discussed by Prat and Barros (2010a).

4.3.2 Stochastic Rainshaft Microphysics Model

This study uses the rain microphysics column model that can simulate the evolution of rain-drop dynamics in liquid phase due to drop-drop interactions from the cloud base to the surface (Prat and Barros, 2007a and 2007b, and Prat et al. 2012). This model explicitly solves the continuous stochastic collection and stochastic break-up equation (SCE-SBE) using a grid independent number and mass conservative scheme. The continuous SCE-SBE equation is discretized using fixed pivot techniques and an Eulerian-Lagrangian strategy to resolve advection. The discrete formulation of the SCE-SBE equation in the ith bin is written as

$$\frac{dn_i(z,t)}{dt} + V_i \frac{dn_i(z,t)}{dz} = S_i(z,t) = \sum_{\substack{j,k \\ x_{i-1} \leq (x_j+x_k) \leq x_{i+1}}}^{j \geq k} \left(1 - \frac{1}{2}\delta_{j,k}\right)\eta C_{j,k} n_j(z,t)n_k(z,t) -$$

$$n_i(z,t)\sum_{k=1}^{nbin} C_{i,k}\, n_k(z,t) + \frac{1}{2}\sum_{j=1}^{nbin}\sum_{k=1}^{nbin} n_j(z,t)n_k(z,t)B_{i,k}\kappa_{i,j,k} -$$

$$n_i(z,t)\sum_{k=1}^{nbin} B_{i,k}n_k(z,t) \tag{4-7}$$

Where $n_i(z,t)$ [m^{-3}] is the total number density of drops in the ith diameter bin, at height z and time t, V_i [ms^{-1}] is the effective fall drop velocity at i^{th} bin, nbin is the number

of diameter bins, $\delta_{j,k}$ is the Dirac function, η [unitless] is the coalescence contribution parameter and $\kappa_{i,j,k}$ [unitless] is the breakup contribution parameter for drops at i^{th} bin resulting from the collisional breakup of two drops at j^{th} and k^{th} bin. The total number drop density $n_i(z,t)$ is mathematically related to the DSD concentration [$N_i(z,t)$; $m^{-3}mm^{-1}$]

$$N_i(z,t) = \frac{n_i(z,t)}{dD} \tag{4-8}$$

where dD is the diameter bin resolution. The first term from the left represents the time evolution of drop number concentration at height z [m] and time t [s]. The second term represents the effective fall drop velocity $V_i = V_{t,i} - w$, where $V_{t,i}$ [ms^{-1}] is the drop terminal velocity for diameter-bin i and w [ms^{-1}] is the vertical wind velocity. The S [$m^{-3}sec^{-1}$] term represents the source term which denotes the coalescence-breakup dynamics. C [m^3sec^{-1}] is the coalescence kernel and B [m^3sec^{-1}] is the breakup kernel. The first term on the right denotes the rate of formation of drops from the coalescence of small drops. The second term denotes the rate of removal of drops due to the coalescence of other drops. The third term denotes the creation of drops due to breakup and the final term represents the removal of drops due to coalescence with other drops. The model also includes different breakup processes such as disk, sheet and filament based on We-p (We stands for Weber number and p stands for raindrop diameter) regime (Prat et al. 2012). This model was modified to include LLCF as per Wilson and Barros (2014) and Duan and Barros (2017) to study the rain microphysics of orographic precipitation and bottom-heavy events such as the SFI cases.

The column microphysics model was run with a time-step of 1s and vertical resolution of 10 m. The drop diameter bins considered in the microphysics model include 60 irregularly spaced diameter bins ranging from 10 μm to 6 mm. The DSD corresponding to the first 13 bins with drop diameter less than 0.05 mm are used for the LLCF-DSD. The model is driven by the DSD at the top of the rainshaft (i.e. top boundary condition; TBC), and by steady injection of LLCF DSD spectra at low-levels over a predetermined depth. The initial condition profiles for the column are derived from the ground-observations or GPM estimates. Because ice microphysics and lateral advection effects are not considered in the model at the current spatial resolution of 4 km and the short-time scales of measurement, the TBC height should be strictly placed below the 0°C isotherm. More details about the model description and the numerical schemes can be obtained from Prat and Barros (2007a) and Prat et al. (2012). Further details on the model implementation for layered cloud systems, model initialization and specification of boundary conditions are available from Duan and Barros (2017) and Wilson and Barros (2014).

4.3.3 Post Retrieval Physical-Correction Framework

GPM Ku-PR corrected reflectivity profiles are obtained by eliminating the ground-clutter contaminated reflectivity values and replacing them with the reflectivity from the nearest no-clutter affected bin, thus imposing a uniform reflectivity profile in the lower troposphere. This strategy to correct ground-clutter leads to the underestimation of bottom-heavy precipitation events (SFI enhancement) and missed detection of low-level

precipitation. The alternative strategy proposed here is to apply a physical-correction to GPM Ku-PR ground-clutter artifacts using a rainshaft microphysical model with initial and boundary conditions at the top of the rainshaft specified using either GPM Ku-PR retrieved DSD parameters assuming a normalized gamma distribution (Method-1) or corrected reflectivity factors to derive DSD parameters assuming a negative exponential distribution (Method-2). A graphical schematic representation of the proposed correction framework is shown in Figure 4-2(a). Note that the methodology is not frequency dependent, and it can be applied to Ka-PR products as well.
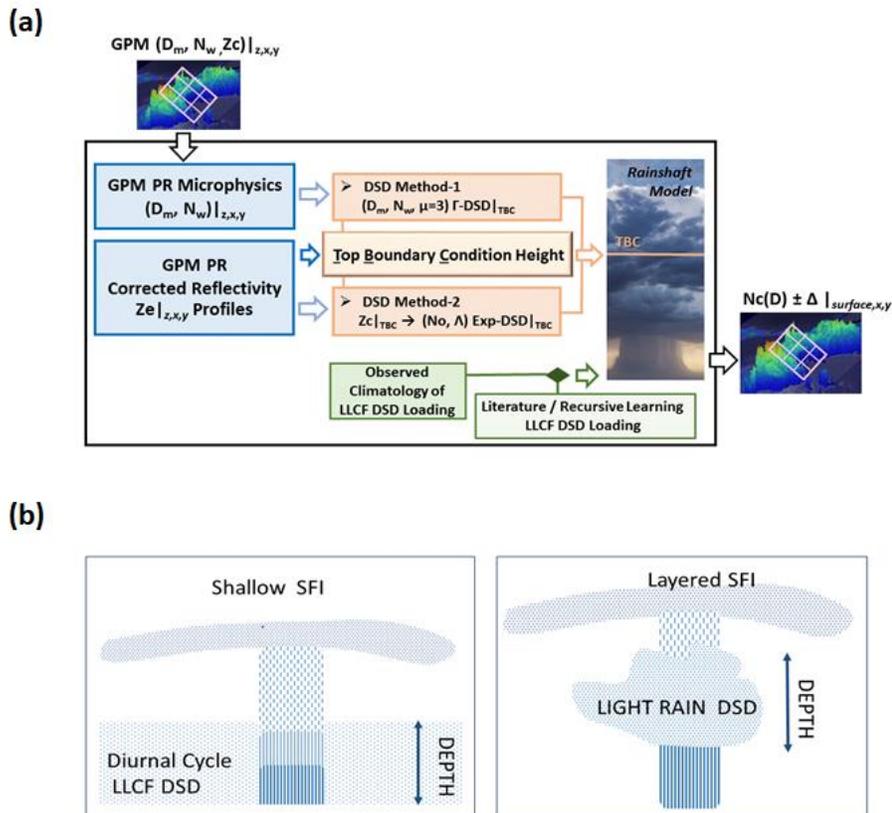
Figure 4-2 a) Schematic representation of physically-based correction framework; b) Illustration of the conceptual SFI models in identified in (a).

Method-1 aims to elucidate the impact of GPM Ku-PR DSD estimates on the rain-rate retrievals, specifically the propagation of retrieval errors to precipitation estimates. DSD spectra profiles from GPM Ku-PR overpasses are used as the TBC and initial conditions for the rain microphysics model simulation. Method-2 relies on the corrected GPM Ku-PR reflectivity factors. Assuming Rayleigh scattering, the reflectivity factor is related to the sixth moment of the DSD spectra:

$$Z = \sum_{D_{min}}^{D_{max}} N(D)D^6 dD \qquad (4\text{-}9)$$

The initial and TBCs for Method-2 rely on DSDs derived from the GPM Ku-PR corrected reflectivity factor assuming it is a negative exponential function of the form:

$$N(D) = N_0 \exp(-\Lambda D) \qquad (4\text{-}10)$$

where, N_0 [m^{-3}mm^{-1}] is the intercept parameter and Λ [mm^{-1}] is the slope of the distribution. First, based on the reflectivity values, the intercept parameter is constrained between 8000 (Marshall-Palmer distribution) and 80000 m^{-3}mm^{-1} based on a look-up-table generated by the DRCM assuming Rayleigh scattering (Eqn. 4-9) following Prat and Barros (2010b) sensitivity analysis. Subsequently, the intercept parameter is fixed, and the slope parameter is estimated iteratively to match the reflectivity factor. Because of the virtually instantaneous Ku-PR measurements corresponding to a specific GPM overpass, here the TBC is fixed for the duration of the simulations. The difference between Method-1 and Method-2 is in the DSD at the initial and TBC.

In general, two types of SFI regimes are observed: Shallow-SFI and Layered-SFI. Conceptual depictions of the two regimes are illustrated in Figure 4-2(b). In shallow-SFI, near-surface rain-rate is enhanced due to the presence of low-level sub-millimeter and μm-scale hydrometeors such as fog, haze and very small cloud droplets. In layered-SFI, light stratiform precipitation is enhanced locally by low-level precipitating clouds. To capture the shallow-SFI effects, it is proposed that climatological observations of LLCF microphysics either from regional observations or from the literature be used to specify LLCF microphysics loading. For layered-SFI, a cloud layer is introduced in the rainshaft using light rain DSDs (DSD with d<0.3 mm; narrow spectra) for the LLCF microphysics loading. For both SFI regimes, LLCF effective depth is also an important model parameter in addition to the microphysics. Here, simulations with and without low-level microphysical forcing are conducted for both SFI regimes using both Method-1 and Method-2. It is envisioned that for operational implementation, a detection and classification algorithms such as Arulraj and Barros (2017) and other information from GPM measurements should be used to identify the prevailing SFI regime. LLCF forcing (i.e. LLCF-DSD derived from MPS climatology for shallow-SFI or light rain-DSD for layered-SFI) is introduced in the simulation after the initial rain-rate reaches steady state and the simulation continues until the enhanced rain-rate reaches steady state. After the low-level forcing is removed, the simulations continue until the rain-rate drops to the initial steady state value.

The DSD parameters used for the layered-SFI microphysics are derived from the initial DSD conditions by removing all bins ≥ 0.3 mm. The DSD parameters used for the shallow-SFI microphysics are derived from the climatology of LLCF parameters obtained from MPS observations available in the SAM. In particular, the MPS was placed along with a fog-trailer at P6 in the western foothills of SAM between September 25, 2015 and November 6, 2015 (Section 4.2.1). The LLCF-DSD spectra is assumed also to follow a negative exponential distribution as in Eqn. (10) (Wilson and Barros, 2014). The LLCF-DSD parameters such as the intercept (N_0) and the slope (Λ) are calculated by fitting a negative exponential distribution to the MPS observed DSD spectra from drop-diameter between 25 and 50 μm. Then, the climatology of the LLCF characteristics is obtained by averaging the DSD parameters at different time-scales. Figure 4-3 shows the diurnal cycle of the LLCF parameters averaged at 10-min, 30-min and 60-min time-scales to illustrate the impact of temporal integration and parameter variability on the final parameter estimates. The high values of the intercepts reflect morning and mid-day fog that are prominent in the western foothills of the SAM. The DSD parameters used as the LLCF loading for the model simulations are extracted from the diurnal cycle climatology at the same time of day as the overpass. Also, the DSD for microphysics loading is for diameters ranging from 10 to 50 μm. This range is considered based on Wilson and Barros (2014) and Prat and Barros (2017). For all simulations presented in this chapter the TBC was placed at 2 km AGL (Above Ground Level), and only warm rain conditions were considered.

(a)

(b)
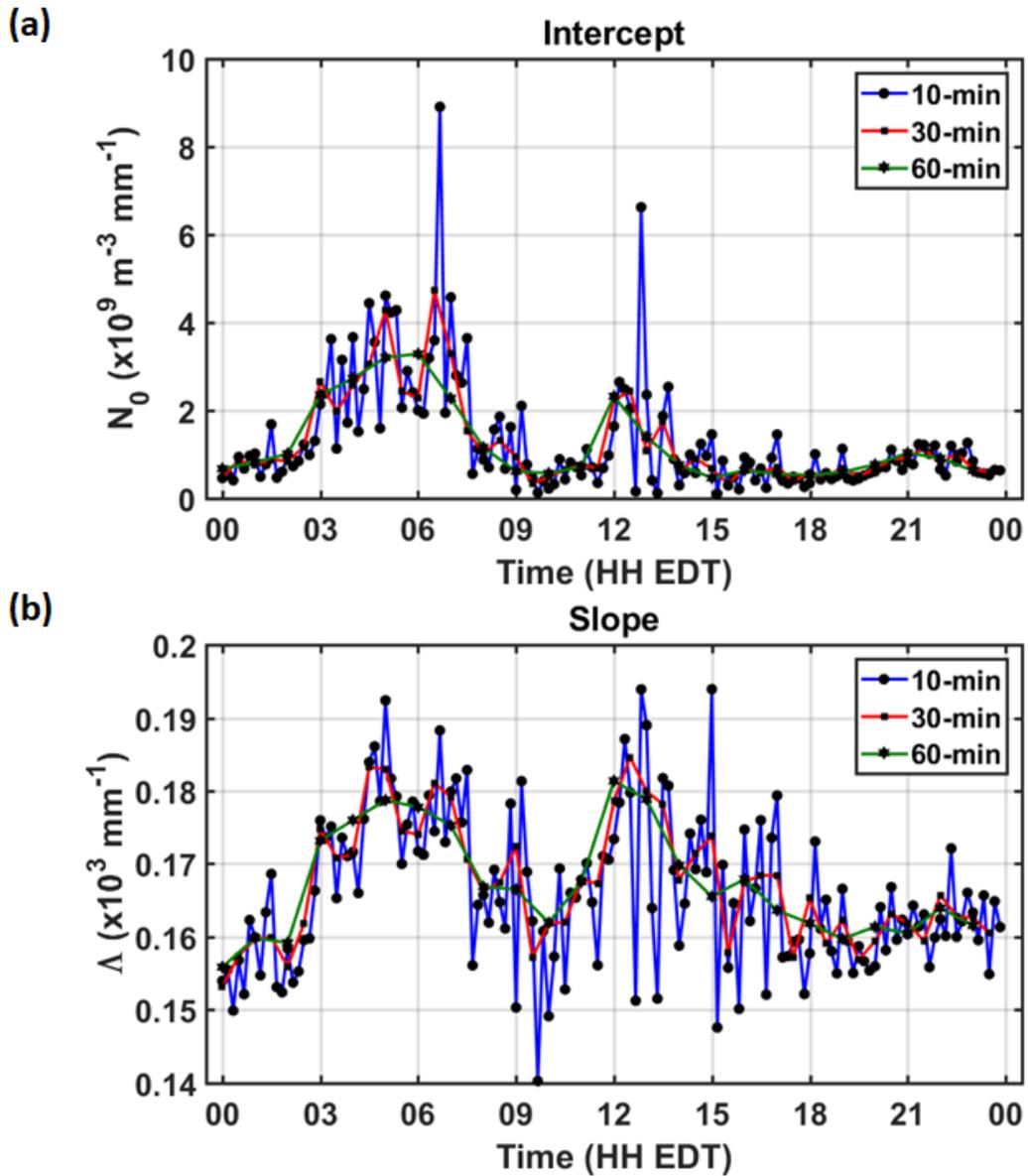
Figure 4-3 Diurnal cycle of LLCF microphysical parameters measured by the MPS at Elkmont (P6) between September 25, 2015 and November 6, 2015 at 10-, 30-, and 60-minute time-scales: (a) Intercept and (b) Slope.

4.4 Error Analysis Results

4.4.1 Statistics of Retrieval Error

The error analysis contingency matrix is shown in Table 4-1. Following discussion in Section 3.1, if more than one rain-gauge lies within a radar beam, then each rain-gauge is considered as an independent sample. For reference, the values of the contingency matrix calculated by averaging all rain-gauge observations within the GPM Ku-PR radar beam at the time of overpass are also reported in Table 4-1. This results in an increase of 0.3% and 0.2% respectively in MDs and correct hits (YY), and a decrease of 0.2% in FAs. Most of the GPM overpasses occurred during no-rain conditions. In the case of TRMM products, the number of FA cases was the dominant type of error compared to the number of MD cases (Prat and Barros, 2010a; Duan et al. 2015). However, in GPM Ku-PR estimates, the number of MD cases is higher than the number of FA and YY cases.

Table 4-1 Contingency matrix comparing rainfall detection by the GPM Ku-PR radar vis-à-vis the rain-gauges expressed in number of occurrences and as percentage of all measurements considering each rain-gauge measurement as an independent sample. Numbers between ()

	RG=0	RG≠0
GPM Ku-PR = 0	10226 – 94.3% (7118 – 93.5%)	NY/MD 211 – 2.0% (177 – 2.3%)
GPM Ku-PR ≠ 0	YN/FA 208 – 1.9% (132 – 1.7%)	YY 197 – 1.8% (150 – 2.0%)

Further, analysis of the diurnal cycle of detection and estimation errors considering each rain-gauge observation as an independent sample is presented in Figure 4-4(a-d). Similar analysis performed using the average rain-gauge observations within the radar beam as reference is provided in Appendix E, Figure E 1 E1. Figure 4-4(a) shows the distribution of the detection error metrics with respect to the time of the day. Note, higher FARs denote high FAs and lower PODs denote high MDs. High FARs and low PODs lead to low CSIs between 00-03h and 12-15h EDT. In addition, higher FARs and higher PODs suggest FAs dominate MDs after 15h EDT. Figure 4-4(b) shows the corresponding seasonal distribution of the detection errors. Lower PODs in the fall and winter seasons suggest higher frequency of MDs, and high values of FAR in the summer season denote higher frequency of FAs. Spring season shows higher CSI comparatively. The intra-annual variability in POD suggests that MDs are associated with light precipitation cases and low-level processes systematically not captured by reflectivity measurements (Prat and Barros, 2010a; Duan et al. 2015), whereas FAs are associated with stratiform precipitation as classified by GPM Ku-PR retrieval algorithm (82% of 208 cases) and retrieved rain-rates < 1 mm/h (67% of 208 cases; the histogram of GPM Ku-PR rain-rates during FAs is shown in Appendix E, Appendix E1). This could be the result of orographic rainfall's spatial heterogeneity and NUBF artifacts. Figure 4-4(c) and (d) show the diurnal cycle and seasonal cycle of the estimation errors. A total of 48 overpasses were identified as OVR cases, while 61 cases were classified as UND during the period of study. The diurnal cycle

119

of the estimation errors suggests that the frequency of UND is higher than that of OVR in the mornings reflecting regional LLCF climatology (Wilson and Barros, 2014 and 2015). Nevertheless, a note of caution is warranted given that the total number of overpasses for rainy conditions is still small and insufficient to establish statistical conclusions with high confidence. In the SAM, the UND number is higher than OVR in all seasons except spring, and in particular in the warm season (June, July and August) consistent with the intra-seasonal climatology of LLCF in the region (Duan and Barros, 2017).
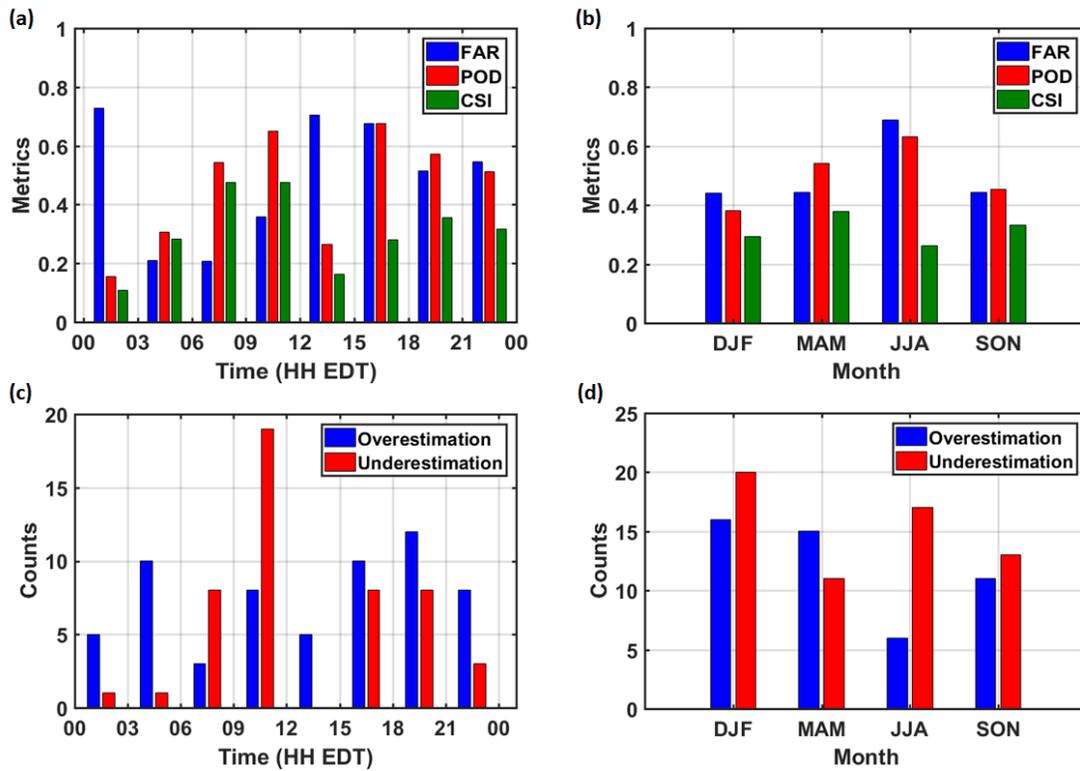
Figure 4-4 Diurnal and seasonal cycle of detection and estimation errors in the SAM. (a) Diurnal and (b) Seasonal cycle of detection error metrics by comparing rain-gauges and GPM Ku-PR near-surface precipitation estimates in the SAM. (c) Diurnal and (d) Seasonal cycle of number of cases with estimation errors in the SAM.

Error classes OVR and UND in Figure 4-4(c-d) are based on the relative bias ε (Eqn. 6) using a threshold of ± 0.5. If no threshold is applied, that is $\varepsilon=0$, all CD occurrences are redistributed to OVR (80) and UND (117). The diurnal cycles of OVR and UND errors estimated according to the standard meaning of under- and overestimation are shown in Appendix E, Fig. E3. The diurnal cycle in Figure E3 is similar to Figure 4-4(c) except in the morning between 09-12 EDT when the number of OVR cases doubles consistent with high POD during the same period. UND increase in MAM, but the seasonal cycle does not change.

Figure 4-5(a) shows the bias of the GPM Ku-PR estimated precipitation-rates when compared with the rain-gauge observations, and the diurnal cycle of the estimation bias is shown in Figure 4-5(b). OVR occurs when the rain-rate detected by the rain-gauges is less than 1 mm/h, and UND is dominant when GV rain-rates are greater than 7 mm/h. Most of the overpasses with GV rain-rates greater than 7 mm/h between 16-23 h EDT were underestimated by GPM Ku-PR. Similarly, most of the overpasses with GPM Ku-PR rain-rates greater than 10 mm/h between 06-12h EDT were overestimated by GPM. The diurnal cycle of bias errors follows therefore the diurnal cycle of rainfall and hydrometeorological regime (Wilson and Barros, 2014).
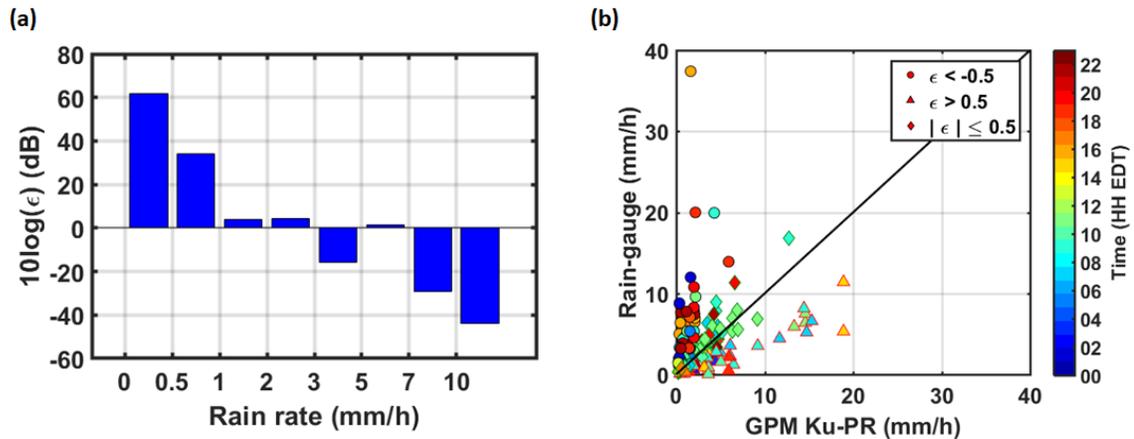
Figure 4-5 Estimation errors in GPM Ku-PR near-surface precipitation estimates. (a) Bias between GPM Ku-PR near-surface rain-rate estimates and rain-gauge observations. (b) Diurnal cycle of GPM Ku-PR near-surface rain-rates compared with the rain-gauge observations.

4.4.2 Physical Basis of Retrieval Error

The specified minimum detectable precipitation-rate for TRMM-PR is 0.5 mm/h. Even though the Ku-PR minimum detectability is 0.5 mm/h as stated by Hou et al. (2014), in practice, as mentioned earlier, GPM Ku-PR has detectability close to that of Ka-PR HS (0.2 mm/h; Speirs et al. 2017). The improvements in minimum rainfall detectability in GPM Ku-PR compared to TRMM-PR explain the lower relative number of mid-day MDs where light precipitation is present (e.g. results in Duan et al. (2015) vis-à-vis the present chapter). However, this also increases the FAs during mid-day and late-afternoon. High positive bias was also observed when the rain-gauges detect rain-rates less than 1.0 mm/h.

The ground-clutter signature in the reflectivity profiles of GPM Ku-PR extends from 0.5 to 2.5 km AGL depending on the PR viewing angle, which results in the severe underestimation of bottom-heavy or layered precipitation cases (Figure 4-6). In the SAM,

however, low-level enhancement of precipitation is frequently observed as illustrated in Figure 4-7, where the enhancement of precipitation profiles near surface is measured by the difference between 5-minute averaged GV MRR reflectivity (24 GHz with vertically pointing mode) at 0.2 km and 2 km AGL during the warm season (June, July and August) from valley and ridge sites in the inner mountain region. The diurnal cycle of this reflectivity difference in both valleys and ridges peaks at approximately 3-4 dBZ in the late afternoon (Figure 4-7) suggesting low-level enhancement of afternoon and evening rainfall (e.g. mid-day low-level clouds, and radiation fog in the evening as shown by Wilson and Barros, 2014 and Duan and Barros, 2017).
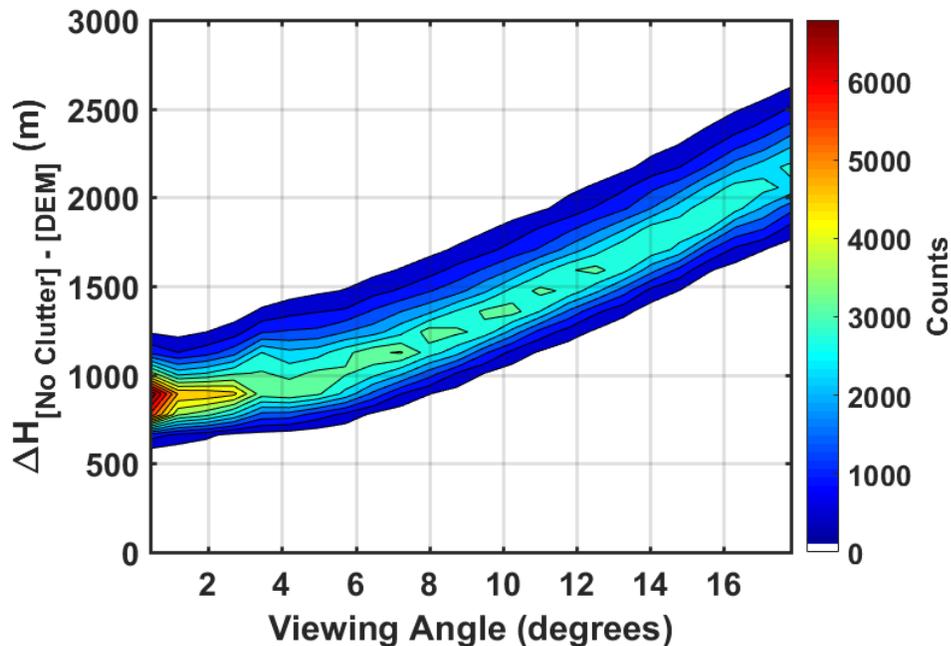
Figure 4-6 Histogram showing the height (AGL) affected by ground-clutter depending on the viewing angle of the GPM Ku-PR in the SAM
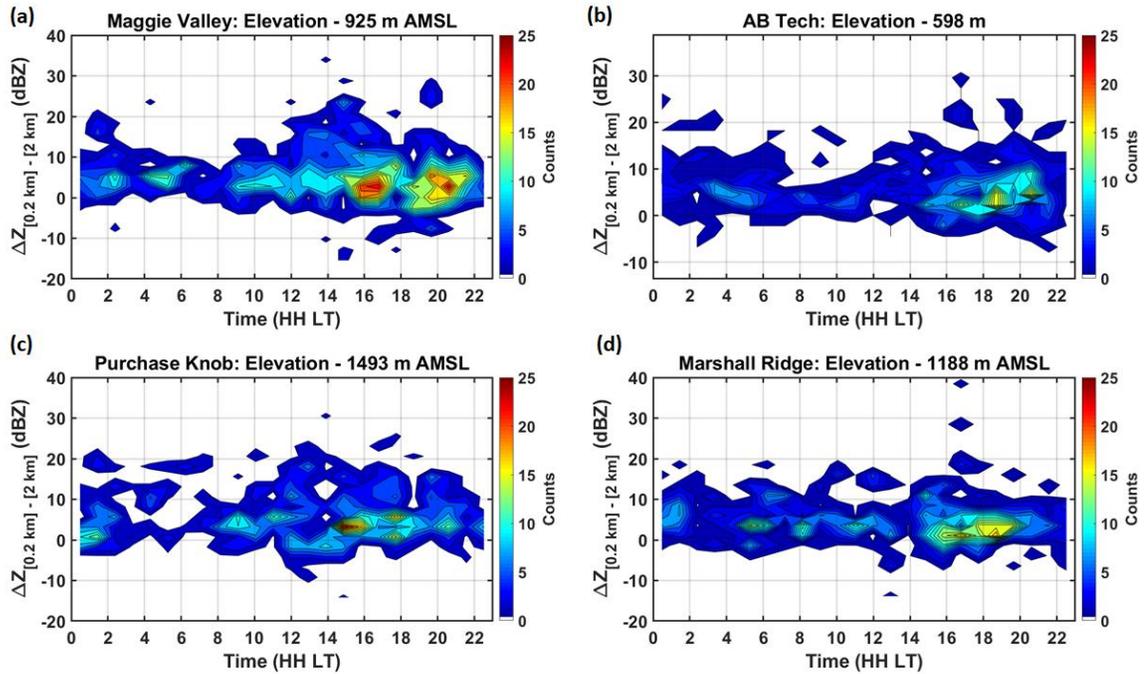
Figure 4-7 Histogram of difference between the MRR reflectivity at 0.2 km and reflectivity at 2 km in the warm season (JJA) at (a) Maggie Valley, (b) AB Tech (c) Purchase Knob and (d) Marshall Ridge. Purchase Knob and Marshall Ridge are located on ridges while AB Tech and Maggie Valley are located at low elevations in the inner region.

A consequence of failing to detect low-level enhancement processes is high uncertainty in the estimates of microphysical characteristics of bottom-heavy precipitation from GPM PR measurements. Figure 4-8 shows the D_m-N_w relationship from the Parsivel disdrometer at Elkmont (P6 in Figure 4-1a) and GPM Ku-PR estimated DSD parameters according to precipitation-rate. The orange contour outlines the region of D_m-N_w phase-space occupied by the climatology of disdrometer microphysics. There is good agreement between disdrometer microphysical parameters and GPM Ku-PR estimates for intermediate rain-rates between 3 and 7 mm/h. This is in keeping with the lower bias for

intermediate rain-rates shown in Figure 4-5(a). However, there are large differences for

rain-rates < 3 mm/h (light precipitation) and > 7mm/h (heavy rainfall).
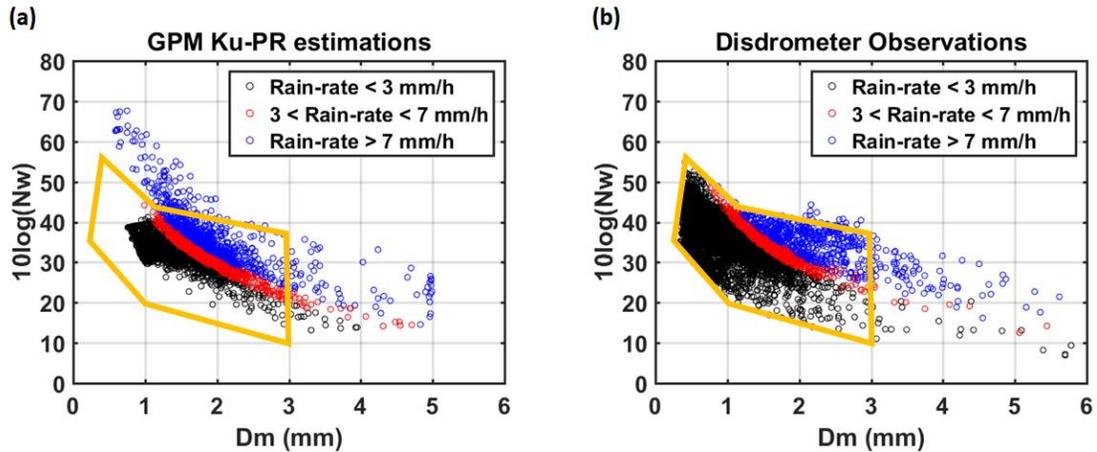
Figure 4-8 Climatology of D_m-N_w estimated by (a) GPM Ku-PR (Time period – March 2018 – April 2018) and (b) Parsivel Disdrometer (Time period – May to October, 2014). Orange lines delineate the region of Dm-Nw space filled by the Parsivel disdrometer.

Duan and Barros (2017) simulated the evolution of the vertical structure of a mid-

day rainfall event observed on October 1, 2015 at P6 (Parsivel disdrometer site in Figure

4-1a). These simulation results are used here to perform proof of concept analysis of GPM

Ku-PR DSD retrievals against GV observations. The model simulations are conducted for

70 minutes during the event between 14:50 and 16:00 Local Time (LT), and the low-level

enhancement in the near-surface precipitation is observed between 15:00 and 15:30 LT.

Contemporary reflectivity profiles for this SFI case are shown in the Appendix E (Figure

E2), and details of the simulation are also summarized in Appendix E (Table E 3) following

Duan and Barros (2017), where 'LLCF – 400 m' is the simulation of SFI involving a 400m

deep LLCF layer with uniform DSD and the DSD TBC derived from MRR reflectivity is

updated every 1 minute. 'No LLCF' refers to the simulation of stochastic collision-coalescence and collision-break up processes in the rainshaft without SFI. The latter severely underestimates the surface precipitation accumulation at collocated rain-gauges by approximately 7 mm on average (see also Figure 4-6 in Duan and Barros, 2017).

The evolution of N_w and D_m relationship for with-LLCF (circles) and without-LLCF (triangles) simulations for the duration of the event is depicted in Figure 4-9. For reference, the squares denote the D_m-N_w relationship obtained from the collocated MPS observations that were used to specify the LLCF spectra. The MPS N_w and D_m are computed for 1-minute averages for drop diameter bins between 12.5 μm and 1.5 mm. The right panel shows the model simulated rainfall rate compared with the rain-gauge records. For the simulation with the LLCF-DSD forced in the lower 400 m, the N_w gradually increases while the D_m decreases. As LLCF retreats and precipitation-rate diminishes after 15:30 EDT, the D_m-N_w trajectory (circles) moves in a wide loop and retreats to the D_m-N_w values close to the initial conditions. By contrast, the excursion of D_m-N_w phase-space trajectory for the no-LLCF case (triangles) is much narrower, reflecting the dominance of break-up processes as D_m decreases and N_w increases between 15:10 and 15:35 EDT when rainfall intensity is the highest. Thus, the microphysical signature of SFI can be measured in terms of a decrease in D_m and an increase in N_w for the simulated event. The red and black dotted contours in Figure 4-9 show the histograms of one-minute averaged D_m-N_w from the existing GV Parsivel disdrometer at Elkmont (P6

in Figure 4-1a) on June-October 2014, and from GPM Ku-PR observations of liquid phase precipitation respectively. GPM Ku-PR DSD estimates considered for the analysis are near-surface DSD parameters between March 2014 and April 2018 at the SAM. Note the most probable value of the D_m-N_w for GPM Ku-PR estimates shifted towards higher D_m and lower N_w compared to the Parsivel disdrometer. In addition, the MPS observations show lower D_m and high values of N_w confirms the large numbers of small drops present during the event, consistent with LLCF conditions favorable to SFI. Note that the MPS measurements are not representative of the full DSD (D < 1.5 mm by design see Section 2.1). The truncation of the DSD spectra affects the D_m and N_w computation, and therefore the MPS spectra cannot be used to estimate rain-rates. Nevertheless, apparent coordination among dominant microphysical processes and DSD parameters estimated from different sensor measurements suggests that for a given region there is potential to use the climatology of the microphysical parameters (e.g. diurnal cycle, seasonality, and topographic organization) tied to rainfall regime (e.g., see Figure 4-1 in Wilson and Barros, 2017) in order to develop general SFI corrections to GPM retrievals. It is important to note that very small drops will add little to the Ku- and Ka- band reflectivity and increase the attenuation especially at Ka-band.

(a)

Observations vs Model Simulations

(b)
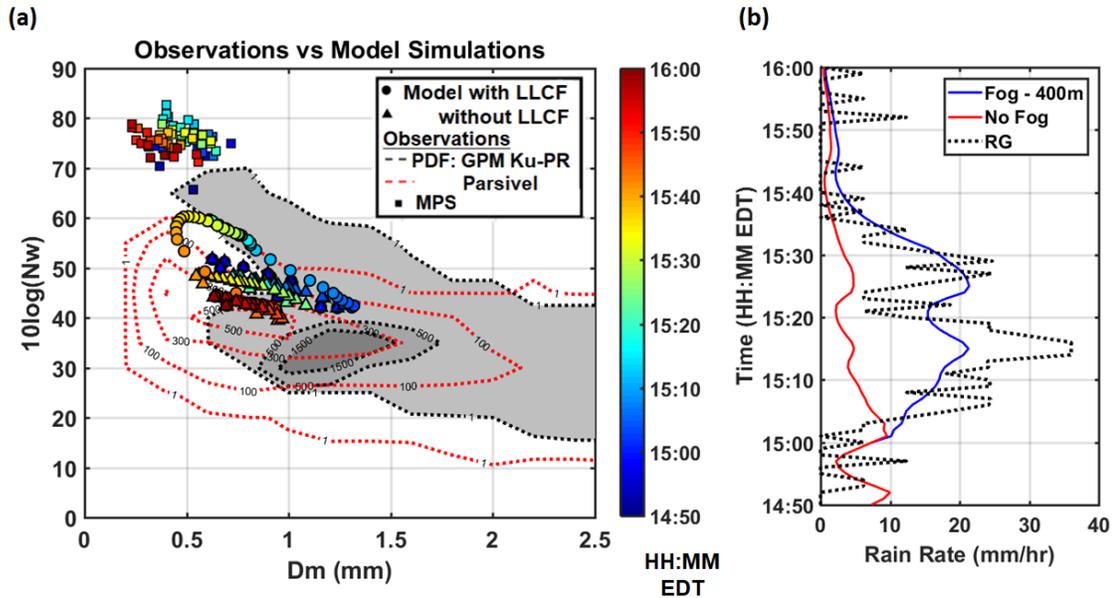
Figure 4-9 D_m-N_w relationship for model simulations compared with observations. The circles and triangles denote the DSD for model simulations with and without LLCF respectively for October 1, 2015 SFI case. LLCF of 400 m depth is forced between 15:00 and 15:30 EDT. The corresponding rain-rate for the simulations and rain-gauge (RG) observations are shown in the right panel. For further details please see Duan and Barros (2017). The squares denote the MPS observations for the October 1, 2015 case. The black dotted lines denote the PDF of GPM Ku-PR liquid phase precipitation between March 2014 and April 2018. The red dotted lines show the PDF of Parsivel observations between June and October, 2014.

4.4.3 Case Study

Error analysis of TRMM-PR (Barros et al. 2000; Prat and Barros, 2010a; Islam et al. 2012; Duan et al. 2015), GPM Ku-PR Level-2 (Speirs et al. 2017) products using ground validation rain-gauge networks and discussions on possible physical basis of errors in satellite-based precipitation radar retrievals suggest three possible categories of error sources: 1) contamination of near-surface reflectivity profiles due to ground-clutter; 2) high spatial heterogeneity of precipitation systems interacting with complex landform

that magnify NUBF artifacts; 3) DSD parameter estimation such as D_m and N_w by assuming a constant shape parameter $\mu=3$ for the normalized gamma distribution. The physical basis of detection and estimation errors are further investigated on a warm precipitation case in the SAM with ground-based disdrometer networks and GPM Ku-PR overpass. The warm precipitation overpass occurred at 22:59 EDT on September 2, 2014 when GPM Ku-PR was in descending orbit. The spatial distribution of near-surface estimates of corrected reflectivity factor, rain-rate and DSD parameters (D_m and N_w) are shown in Appendix E (Figure E 3) with a summary of results in Table 4-2.

Table 4-2: Comparison of GPM Ku-PR estimates with the GV Parsivel disdrometer observations on September 03, 2014 at 22:59 EDT

ID	Site	Rain-rate (mm/h)		D_m (mm)		$10\log(N_w)$ (dBNw)	
		GV	Ku-PR	GV	Ku-PR	GV	Ku-PR
MD-1	P2 - valley	1.38	-	1.22	-	34.76	-
FA-1	P7 - valley	-	0.52	-	1.13	-	32.24
FA-2	P19 - ridge	-	0.66	-	1.18	-	32.43
UND-1	P6 – valley	10.41	2.51	1.48	1.45	39.54	34.01
CD-1	P5 – valley	11.24	10.20	1.87	1.41	35.23	40.78
CD-2	P8 - valley	1.65	1.73	1.31	1.36	34.39	33.59

Observations from GV disdrometers are used next to evaluate the DSD parameters and rain-rates estimated by GPM Ku-PR. Valley locations (elevation < 850 m) and ridges are denoted in Figure 4-1(a) by circles and triangles respectively. The missed detection at P2 is due to spatial heterogeneity and NUBF effects due off-nadir viewing geometry, even as surrounding pixels detect precipitation. The vertical profile of measured reflectivity

129

factor at P2 is shown in Appendix E (Figure E 4). P5 and P6 are located in the foothills of the western ridges and heavy precipitation above 10 mm/h was recorded at both locations. Rain-rate at P5 is estimated correctly (hereafter, CD-1), and the rain-rate at P6 is underestimated (hereafter, UND-1). Note, CD and UND are classified based on the bias estimation criteria highlighted in Section 3.1. The GPM Ku-PR estimated D_m for UND-1 is close to the observations, but N_w is underestimated by 5 dBNw which leads to the underestimation of rain-rate. In the case of CD-1, the GPM Ku-PR underestimates D_m by 0.4 mm and overestimates N_w by 5 dBNw. The errors in the overestimation of N_w and underestimation of D_m are compensated in CD-1 leading to correct estimation of rainfall amounts, even if not the DSD spectrum. Analysis of the GPM retrievals indicates that practically a trade-off between N_w and D_m takes place. This leads to the alternate sign of error in the DSD parameter estimation. FA cases at P7 and P19 are overpasses with light precipitation (<1 mm/h). However, the collocated disdrometers did not register precipitation within 1-minute of overpass. This could be explained in part by space-time intermittency in precipitation.
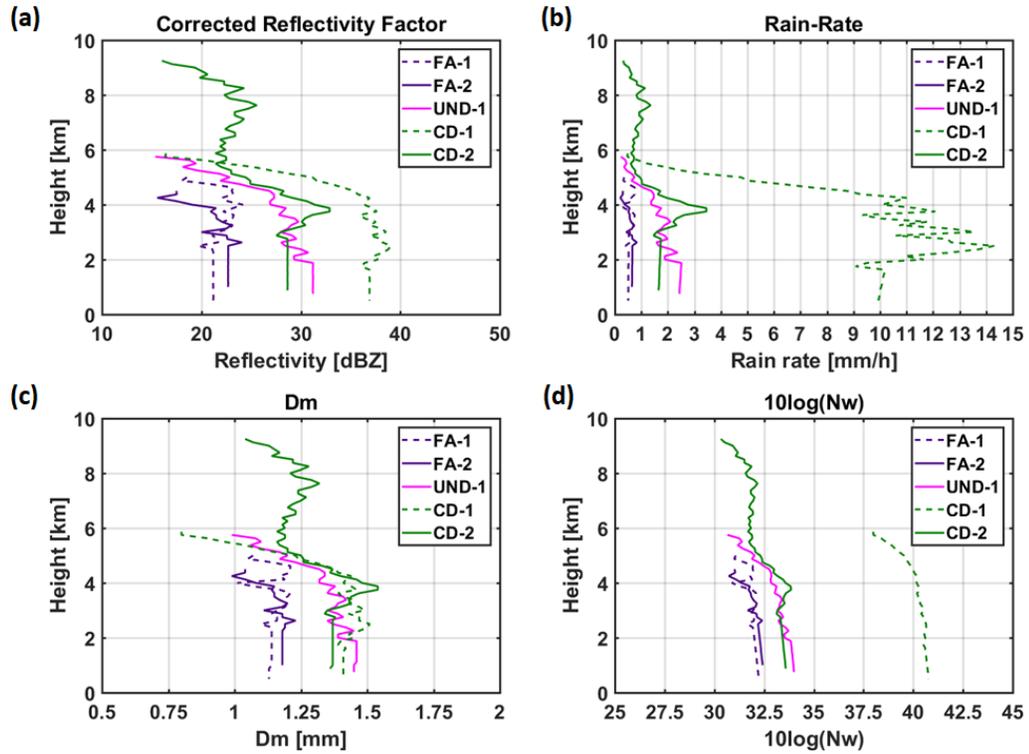
Figure 4-10 GPM Ku-PR estimated profiles of (a) Reflectivity factor, (b) rain-rate, (c) Dm and (d) Nw for the cases observed on September 2, 2014 at 22:59 EDT.

Ku-PR profiles of the two FA cases, two CD cases and the UND case for this overpass are shown in Figure 4-10. Note that the reflectivity values, rain-rate and D_m are approximately constant near-surface according to the ground-clutter correction. Rain-rate at CD-1 decreases by approximately 0.2 mm h^{-1} near-surface. However, N_w shows increase of 0.1-0.5 dBNw in the near-surface bins in most of the cases. Z_m and D_m are constant near surface. This is due to the sensitivity of the attenuation corrected reflectivity factor Z_e to physical temperature in the near-surface bins, specifically due to the dependence of the refractive index of liquid water on temperature (Eqn. 2a). Consequently, as $F(D_m, n_l)$ decreases, the near-surface N_w values increase for the same value of Z_e. The shapes of the

131

reflectivity profiles for CD-1 and UND-1 are similar, but the UND-1 near-surface reflectivity is 7 dBZ below CD-1 while the UND-1 D_m is 0.04 mm higher than CD-1. The N_w profile for CD-1 shows an increase of nearly 6 dBNw compared to other profiles, illustrating the impact of number concentration (hydrometeor mass) on reflectivity independently of size.
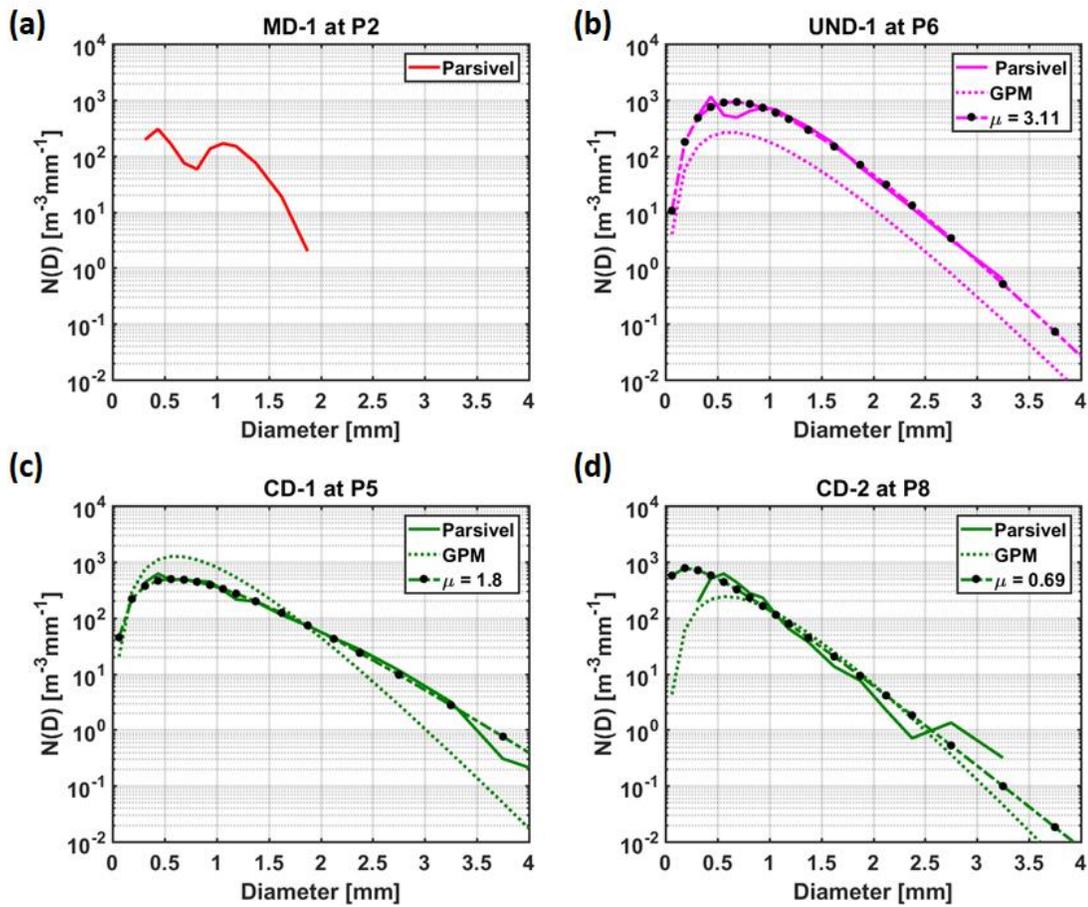
Figure 4-11 DSD from Ku-PR and Parsivel disdrometer for the (a) MD-1, (b) UND-1, (c) CD-1, and (d) CD-2 cases observed on September 2, 2014 at 22:59 EDT. The shape factor (μ) is computed by fitting the gamma distribution to the Parsivel observations. GPM Ku-PR estimated DSD profiles corresponds to μ=3.

The Parsivel disdrometer DSD spectra and GPM Ku-PR estimates are compared in Figure 4-11 for the same pixel locations. The observed DSD spectrum for MD-1 exhibits bimodality with peaks at 0.5 mm and 1 mm and the largest diameter does not exceed 2mm (Figure 4-11a). Figure 4-11(b) shows the DSD spectra of the underestimation case at P6. The solid line and dotted lines denote the disdrometer observations and GPM Ku-PR estimates respectively. Dashed lines represent the best-fit normalized gamma distribution (Appendix D, Eqn. D4) with D_m and N_w values obtained from the disdrometer observations and fitted with shape parameter (μ). The optimal μ for UND-1 is close to the GPM retrieval algorithm specifications (μ=3), and the estimated DSD spectrum is clearly shifted downwards compared to the observations. This shift is due to the underestimation of N_w by the GPM Ku-PR, and underestimation of rain-rate due to the underestimation of liquid water content for this case with bottom heavy corrected reflectivity. Note the bimodal shape for D < 1mm consistent with the enhanced coalescence efficiency by SFI as discussed by Duan and Barros (2017).

The CD cases are presented in Figure 4-11(c) and (d): the GPM Ku-PR derived DSD spectrum overestimates the number of small drops and underestimates the number of large drops for CD-1 showing a much steeper slope; for CD-2 the number of large drops (>2.5 mm) and small drops (0.3-0.75 mm) is slightly underestimated and the number of mid-size drops is overestimated by GPM Ku-PR, although the DSD slope is in good agreement with the disdrometer observations. CD-2 is a light rainfall case with intensity

less than 2 mm/h, while CD-1 is a heavy precipitation case. This illustrates the challenge in capturing the right-hand side tail of the spectrum with substantial differences for diameters larger than 2.5 mm. Note also the differences in the shape parameter estimates between the two Parsivel disdrometer spectra (μ=1.8 for CD-1, μ=0.69 for CD-2), and between the two disdrometer spectra and the GPM algorithm (μ=3). Even though the GPM Ku-PR estimates similar D_m for both UND-1 and CD-1, the observed DSD for UND-1 suggests the influence of SFI and thus presence of local LLCF, while the precipitation for CD-1 is determined by convection aloft.

In the retrieval algorithm, D_m is the first parameter to be estimated and, therefore, any errors in the estimation of D_m will be propagated to the estimation of N_w and further to the rain-rate retrieval. In addition, the error in the fixed shape parameter (μ) assumption will lead to unrealistic DSD shapes and rain-rate retrieval errors. For example, the two CD cases correspond to very different rainfall regimes, which further suggests that the use of the μ=3 for both cases in the algorithm is likely a source of error, and instead the shape parameter should be estimated depending on rainfall regime.

4.5 Post-Retrieval Physically-Based Correction Results

Three underestimation cases are considered to demonstrate the proposed framework. Two are shallow-SFI cases for GPM Ku-PR measurements collocated with GV Parsivel disdrometers observations at P5 and P6 (Figure 4-1a) in the western foothills of

134

the SAM. The third case is a layered-SFI case in the inner mountain ridge collocated with

rain-gauge RG110 (Table E 1; Figure 4-1b).

4.5.1 Shallow-SFI
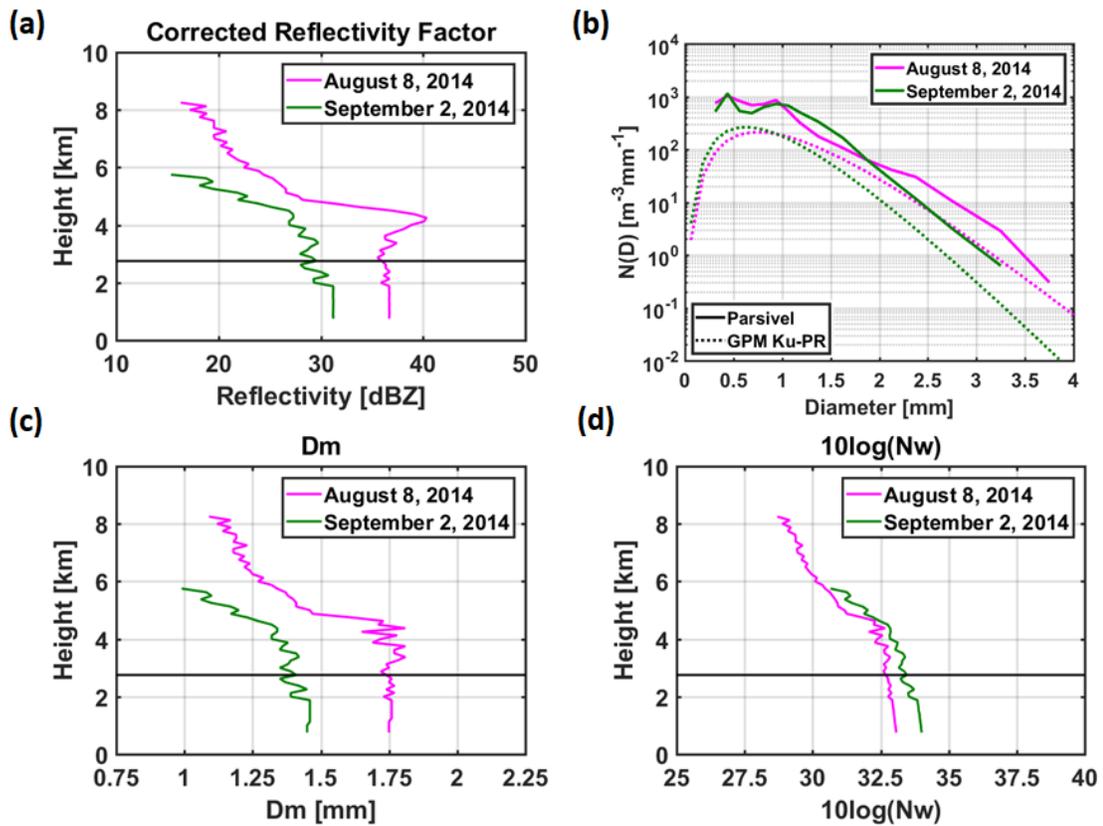
Figure 4-12 GPM Ku-PR estimated (a) reflectivity factor, (b) DSD spectra, (c) D_m and (d) N_w for the two underestimation cases (August 8, 2014 and September 2, 2014). Comparison of DSD spectra from Parsivel disdrometer observations and GPM Ku-PR.

The P5 case was detected on August 8, 2014 at 20:55 EDT, while the P6 case was

detected on September 2, 2014 at 22:59 EDT. Figure 4-12(a) shows the GPM Ku-PR

corrected reflectivity factor profiles for the two cases: P5 (pink) is a stratiform precipitation

case with well-defined bright-band in the reflectivity profiles at 4 km AMSL, whereas P6 (green) suggests shallow warm precipitation with light to moderate intensity. Figure 4-12(b) shows the observed disdrometer spectra (solid lines) and the corresponding GPM Ku-PR estimates (dashed lines). There is approximately 60-80% difference in the number concentration between the disdrometer and GPM spectra for drop diameters in the range of 0 to 2 mm for both the cases. Figure 4-12(c) and (d) show the vertical structure of GPM Ku-PR DSD parameter estimates: D_m profiles closely mimic the reflectivity profiles. Theoretically, N_w is expected to be positively correlated with the reflectivity values. However, the reflectivity factor of the P6 profile is higher than that of the P5 profile, and GPM Ku-PR estimated N_w values for P5are higher than for P6. Surface rain-rates and DSD parameter estimates are summarized in Table 4-4. The D_m values estimated by the GPM Ku-PR algorithm and from surface disdrometer observations are approximately the same, while the N_w values are underestimated by approximately 4-5 dBNw, which explains the underestimation of rain-rate.

Table 4-3: GV Surface observations (Parsivel disdrometers or rain-gauges) and GPM Ku-PR products for three underestimation cases in the western foothills of SAM.

Date	Site	Rain-rate (mm/h)		D_m (mm)		$10\log(N_w)$ (dBNw)	
		GV	Ku-PR	GV	Ku-PR	GV	Ku-PR
August 8, 2014	P5 – valley	11.56	4.67	1.72	1.75	36.96	33.07
September 2, 2014	P6 – valley	10.41	2.51	1.48	1.45	39.54	34.01
June 17, 2018	RG110 - ridge	37.4	1.60	-	1.43	-	32.38

For Method-1 simulations, DSD parameters below 2 km AGL (marked by black solid line in Figure 4-12) estimated by Level-2 GPM Ku-PR are used to obtain DSD spectra assuming the DSD follows a gamma distribution with shape factor μ=3.0 as in the retrieval algorithm. These retrieved DSD spectra profiles are used to initialize the model and specify the TBCs for Method-1 rain microphysics model simulations. The simulations are conducted for a total time of 80 minutes with the same TBC as the initial condition to simulate the GPM Ku-PR instantaneous rain-rate. In addition, varying depths of low-level forcing with LLCF-DSDs from 10 μm to 50 μm introduced for 30 minutes between the 20- and 50-minute marks after the start of the simulations. The LLCF-DSD parameters at the

time of overpass are extracted from the diurnal cycles shown in Figure 4-3 and imposed homogeneously through the depth of the LLCF layer (300 m, 400 m and 500 m AGL) following previous fog studies in the SAM (Wilson and Barros, 2014; Duan and Barros, 2017). The 10-minute averaged diurnal cycle for the intercept and slope shows high variations compared to the 60-min and 30-min averages. Consequently, the simulations are repeated using the nearest 10- and 60-min diurnal cycle of the LLCF parameters to produce the feasible range of rain-rate observations and uncertainty associated with temporal integration and variability.
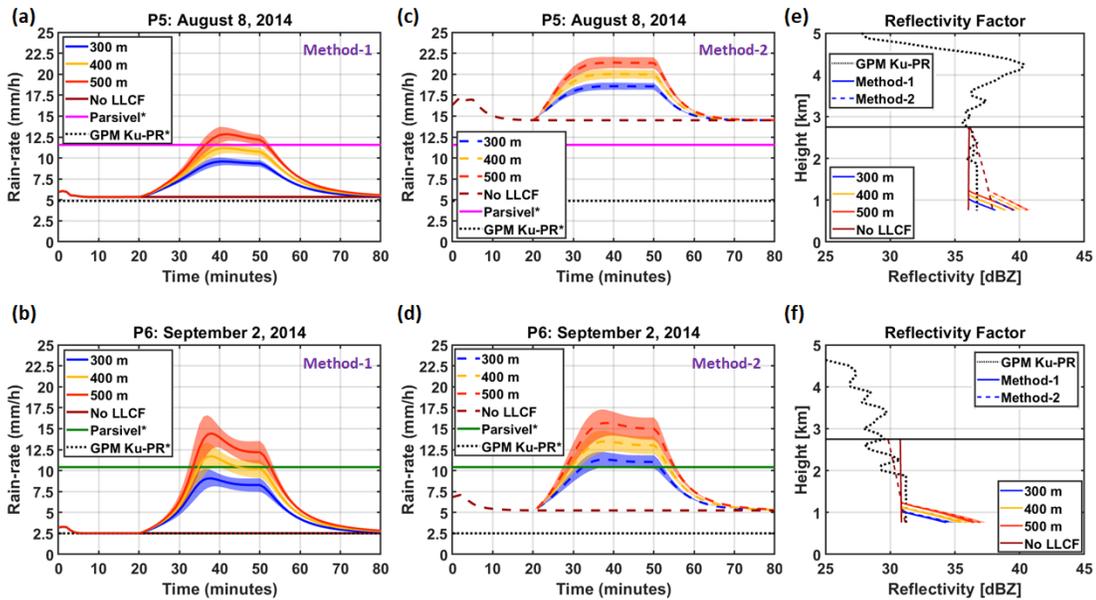
Figure 4-13 Simulated surface rain-rate by Method-1 with and without shallow-SFI for the underestimation cases on (a) P5-August 8, 2014 and (b) P6-September 2, 2014 compared against GPM Ku-PR near-surface precipitation estimates (dotted black lines) and 1-min time-averaged Parsivel disdrometer measurements centered at the GPM overpass time (pink and green lines for P5 and P6, respectively). The asterisk identifies non-model QPE. (c-d) Same as (a-b) but for Method-2 simulations for the two cases. (e-f) Reflectivity profiles from GPM Ku-PR,

and for the results from Method-1 and Method-2 simulations for the two cases. The solid black line denotes the TBC height chosen for the model simulations.

Figure 4-13(a–b) show the surface rain-rates obtained from Method-1 for August 8, 2014 (P5) and September 2, 2014 (P6), respectively. Method-1 simulations without SFI produce surface rain-rates close to the GPM Ku-PR estimates that are 60-75% lower than the disdrometer for the stratiform and shallow precipitation cases. For both cases, Method-1 simulations with shallow-SFI in the lower 400 m produce surface rain-rates close (within 5-10%) to the disdrometer observations. The range of rain-rates obtained using the nearest 10- and 60-min diurnal cycle of LLCF parameters for low-level forcing are shown as the shaded region for each simulation scenario. Figure 4-13(c-d) show simulation results for Method-2. In Method-2, the corrected reflectivity factor profiles are used to estimate the initial and TBC DSD parameters assuming the DSDs follow negative exponential distributions (Eqn. 4-10). For the August 8, 2014 case at P5, the model rain-rate estimates without SFI are already close to the disdrometer observations (~ 20% difference), and indeed including shallow-SFI does not improve the results. This is in contrast with the results for September 2, 2014 at P6 where no-LLCF simulations yield rain-rates of 5mm/h, which is double the GPM Ku-PR estimates, but still severely underestimating the observed rainfall. In the stratiform case at P5, Method-2 no-LLCF precipitation-rate exceeds the observations and thus the surface microphysics results strictly from stochastic rainfall dynamics below the bright-band including differential advection, drop collisions and coalescence and breakup mechanisms. By contrast, to

predict the observed 10 mm/h rain-rates in the shallow precipitation case at P6, the introduction of low-level forcing and simulation of shallow-SFI is necessary in Method-2 as it was in Method-1.

Overall there is 5-15% uncertainty in rain-rate that can be attributed to shallow-SFI impacts on DSD intermittency with the lower values for the stratiform case and the higher values for the shallow precipitation. The hypothesis is that LLCF microphysical loading more strongly impacts the shape of the DSD for light rainfall with smaller D_m. Likewise, the uncertainty linked to the LLCF depth (300 m – 500 m) is about 7.5 mm/h (75% of the observed rain-rate; Figure 4-13b and d) for the shallow precipitation case at P6, and 5.0 mm/h (50% of the observed rain-rate; Figure 4-13a and c) for the stratiform case at P5. The LLCF depth is a scaling factor to describe the hydrometeor mass that can be scavenged through SFI. The LLCF depth is smaller than the depth of the reflectivity profiles impacted by ground-clutter. This highlights the challenges in detecting and estimating shallow precipitation in regions of complex terrain addressed previously by Arulraj and Barros (2017). The changes in reflectivity profiles in the lower 2-km AGL with and without LLCF for both Method-1 and -2 simulations are presented in Figure 4-13 (e-f). The simulated profiles display a reflectivity enhancement of ~ 2 dBZ for Method-2 simulations of the stratiform case without LLCF and negligible changes for Method-1. The simulations with LLCF for the shallow precipitation yield ~5dBZ near-surface enhancement for both methods corresponding to 200% increase in linear terms. The discontinuity in the Method-

1 reflectivity profiles in Figure 4-13(f) which is due to the error in the GPM retrieved DSD

parameters.
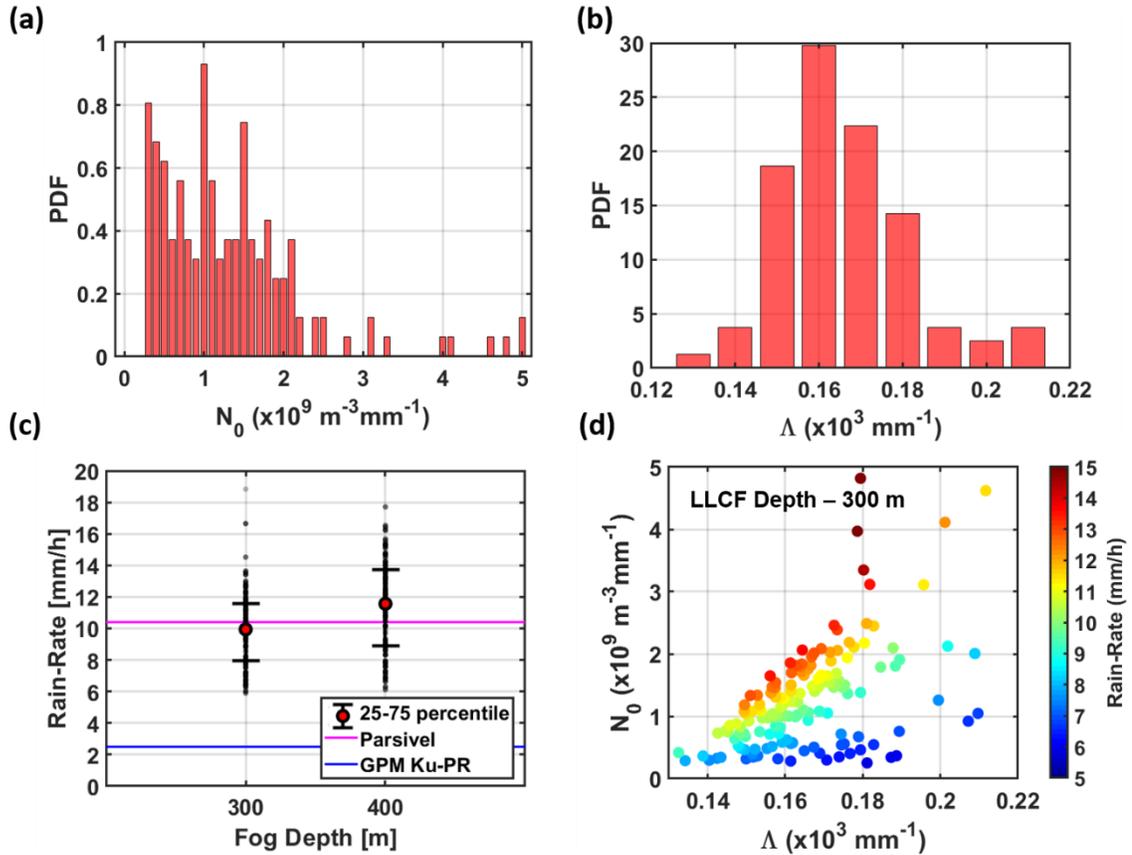
Figure 4-14 Probability distribution Function (PDF) of LLCF microphysical parameters to fit an exponential distribution to MPS DSDs: (a) Intercept parameter (N0) and (b) Slope (Λ). (c) 25th and 75th percentile and mean of near-surface rain-rate obtained from the ensemble of 1000 simulations obtained by bootstrapping with replacement to extract pairs of (N$_0$, Λ) from (a) and (b) and for two different LLCF depths; (d) Comparison of rain-rate for different slope and intercept parameters for the 300 m LLCF ensemble of the September 2, 2014 case at P6.

The LLCF microphysical parameters and depth are an important source of uncertainty

in the application of the correction algorithm as they too are inferred from point

measurements. We rely on average values that neglect the large sub-seasonal, seasonal

and inter-annual variability of the diurnal cycle precipitation systems. Recall that Figure 4-3 shows the averages at 10-, 30- and 60-min time-scales of the diurnal cycle of MPS DSD parameters. Whereas a comprehensive uncertainty analysis for all times-of-day and at all scales is out of the scope of this study, we present here results from ensemble model simulations for September 2, 2014 case at P6 using the empirical MPS Probability Distribution Functions (PDFs) of N_0 and Λ at 22 h EDT to represent the true distributions as shown in Figure 4-14(a-b), respectively. To specify the LLCF microphysics, 1000 pairs of (N_0, Λ) parameters were sampled from the MPS PDFs using bootstrapping with replacement. Method-2 was applied to generate two ensembles of 1000 DRCM simulations each for 300 m and 400 m LLCF depth. The simulated near-surface rain-rates for 300 m and 400 m LLCF depth along with the 25th and 75th percentile intervals are summarized in Figure 4-14(c). The mean rain-rates for 300 m and 400 m are approximately 10 mm/h and 11.5 mm/h respectively, whereas the Parsivel disdrometer observe 10.4 mm/h, and both the 5th and 25th percentiles are higher than the GPM Ku-PR estimate. Although the uncertainty from LLCF microphysics is in line with standard uncertainty in rainfall measurements, it is interesting to examine the variation of rain-rate with respect to N_0-Λ values of LLCF in Figure 4-14 that is mapping uncertainty in the LLCF DSD space. The range of Λ decreases as N_0 increases, and for constant values of N_0, rain-rate shows a decrease of ~ 3 mm/h with the slope as larger drops break. Note the apparent convergence and reduction of uncertainty to $N_0 < 1$. Despite the exploratory nature of these results,

because N_0 ultimately depends on Cloud Condensation Nuclei (CCN) concentration, this map is illustrative of potential uncertainty tied to regional aerosol concentrations (e.g. Duan et al. 2019).

Finally, Method-1 and Method-2 DSD spectra at the 50th minute of the simulation when rain-rates are at steady-state are compared with GPM Ku-PR DSD estimates and disdrometer observations in Figure 4-15. The gray shaded bars denote the expected uncertainties in the simulated DSDs due to uncertainty in the LLCF spectra and underestimation of small drop counts by the Parsivel disdrometer (Angulo-Martinez and Barros, 2014). Method-1 spectra, both without and with LLCF, underestimate number concentrations for drop diameters between 0.75 and 1.5 mm for both cases at P5 and P6 even though the rain-rates produced for scenarios with SFI at P6 (green line) and at P5 are close to the observations. This behavior suggests that there is memory of errors due to the use of a fixed shape parameter (μ=3) to specify the TBC and initial conditions. Figure 4-15(c-d) show Method-2 simulated DSDs. The agreement between observed and simulated spectra with shallow-SFI is excellent for the shallow precipitation regime at P6. For the stratiform regime at P5, there is overestimation of number concentrations by Method-2 simulations for diameters between 1 and 2.5 mm with good agreement only for the larger drop sizes to correct the overestimation of rain-rates with and without LLCF included in the model.
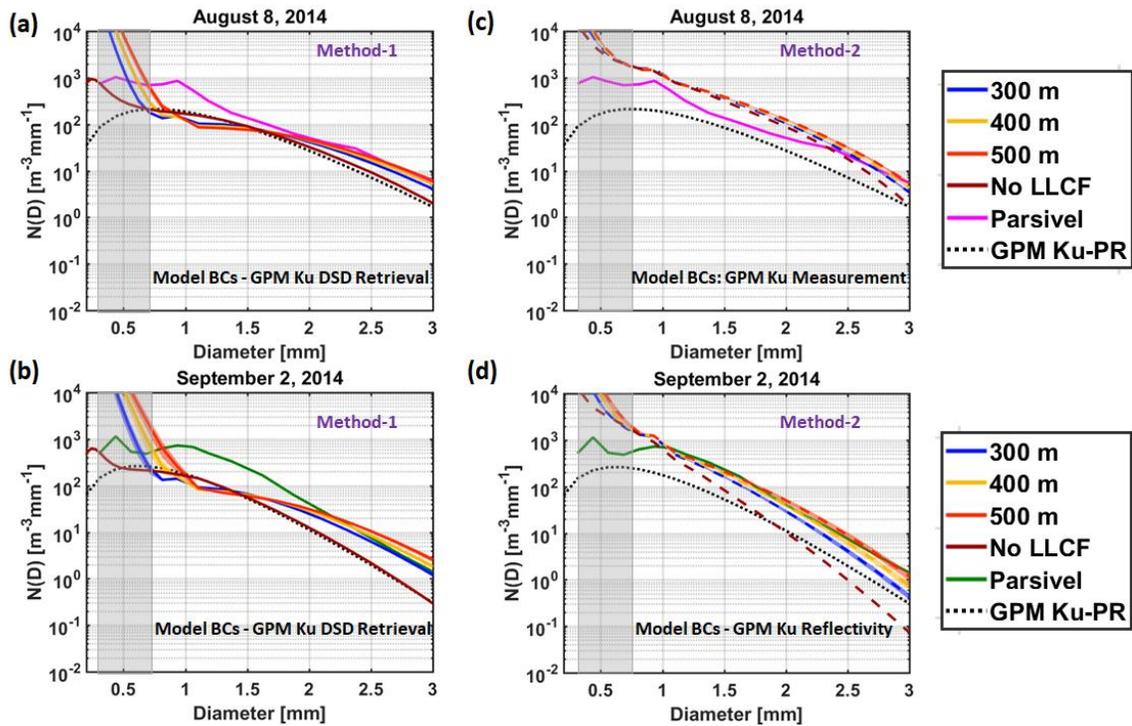
Figure 4-15 DSD profiles from GPM Ku-PR, Parsivel disdrometer and simulated by Method-1 and Method-2 with and without shallow-SFI for the underestimation cases on August 8, 2014 and September 2, 2014. The gray shaded bars denote the expected uncertainties in the simulated DSDs due to uncertainty in the LLCF spectra and underestimation of small drop counts by the Parsivel disdrometer (Angulo-Martinez and Barros, 2014).

4.5.2 Layered-SFI

The layered-SFI case was detected on June 17, 2018 at 16:49 EDT in the inner ridges of the SAM collocated with RG110. RG110 is a tipping bucket rain-gauge with tipping resolution of 0.1 mm/tip. In this case, the rain-gauge recorded rain-rate of ~37 mm/h while the GPM Ku-PR estimated less than 2 mm/h. The surface rain-rate from both Method-1 (not shown here) and Method-2 simulations of shallow-SFI for this case severely underestimate the observations by ~90% (Figure 4-16a). GPM Ku-PR estimated DSD

parameters are shown in Table 4-3, and the back dotted line in Figure 4-16(b) shows the corrected reflectivity factor measured by the Ku-PR. Surface measurements of DSD parameters for this case are not available due to the lack of a collocated disdrometer. The near-surface corrected reflectivity profiles show approximately 2-5 dBZ difference with and without shallow-SFI simulations (Figure 4-16b).
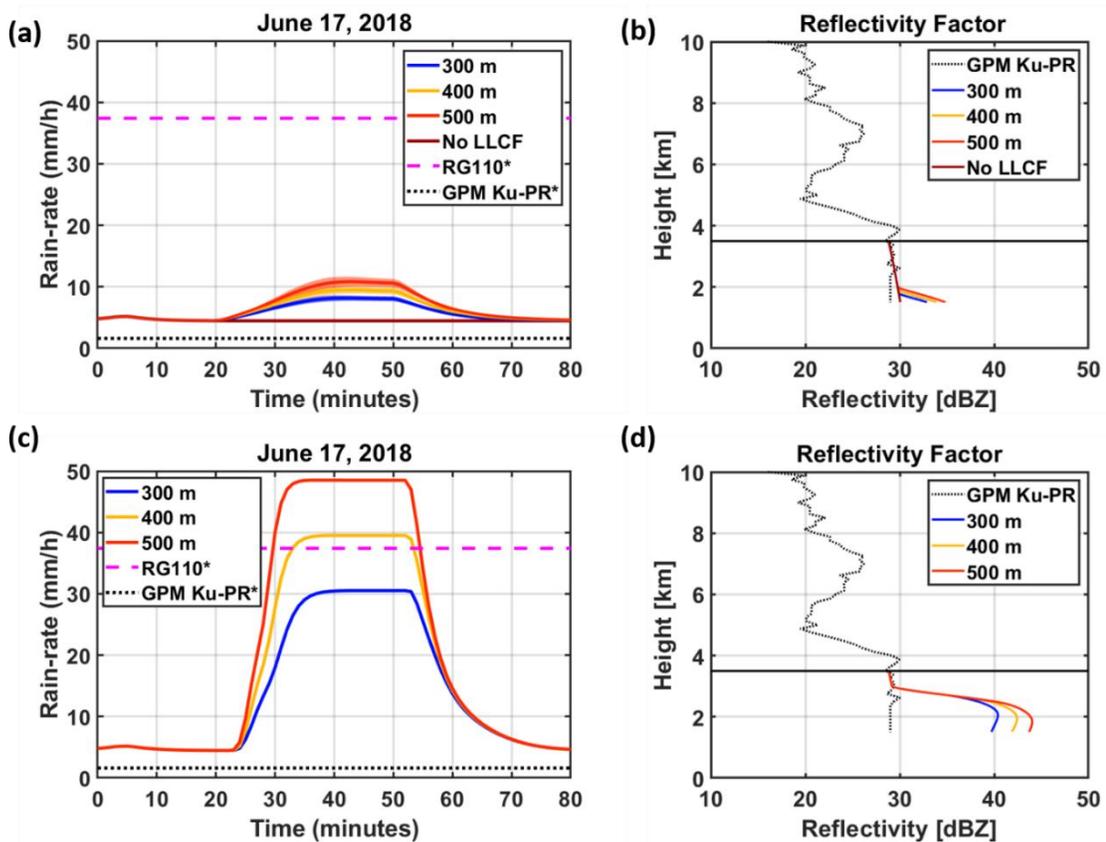
Figure 4-16 Physically-based framework tested on layered-SFI cases : (a) Surface rain-rate by Method-2 with and without LLCF for the case observed on June 17, 2018. (b) Corrected reflectivity profiles from GPM Ku-PR and Method-2 shallow-SFI. (c) Surface rain-rate by Method-2 layered-SFI model simulations. (d) Corrected reflectivity profiles from GPM Ku-PR and Method-2 layered-SFI

Note that the freezing level is at 4.5 km AMSL and ground-clutter impacts the profiles below 0.5 km AGL. The reflectivity profile clearly highlights the presence of ice clouds at ~ 7 km AMSL aloft of a shallow precipitation system with cloud top height around 4 km AMSL. This is an example of how multiple scattering effects might play a key role due to the ice-clouds aloft overlaying warm clouds at low levels. Systematic rain-gauge and disdrometer observations suggest the presence of local precipitating clouds at low-levels (similar to the conceptual model for layered-SFI shown in Figure 4-2b). For this case, Method-2 layered-SFI simulations were conducted by introducing a cloud layer with cloud-top height at 1.5 km AGL and varying depth, and the first 25 bins (with drop-diameter < 0.3 mm; light-rain DSD) from the TBC spectra were retained to specify the layered-SFI microphysics. The rain-rates for the layered-SFI simulations are shown in Figure 4-16(c) with predicted values within 20% of the observations. The modified reflectivity profiles show that the reflectivity below 1.5 km AGL increases by ~10-15 dBZ due to the presence of the cap-clouds (Figure 4-16d), a substantial change from the nominal corrected reflectivity. Whereas this example is very promising because of the improvement in rainfall rates, detection and classification of layered-SFI cases is challenging. In the next section, the methodology is systematically applied for various warm season shallow-SFI underestimation scenarios as a preliminary evaluation.

4.5.3 Systematic Implementation of Physical-Correction

The physically-based framework is systematically applied to six additional shallow-SFI underestimation cases of warm precipitation identified in the SAM since the launch of GPM. The six overpasses were identified from the error diagnostics shown in Section 4.1 by comparing rain-gauge observations with GPM Ku-PR observations. Overpasses with rain-gauge observations greater than 5 mm/h are considered for analysis. Table 4-4 lists the six cases along with rain-gauge locations, rain-gauge observations and GPM Ku-PR estimates between May and August. Despite different LLCF microphysics in the western foothills of the SAM compared to the western, inner and eastern ridges, due to limited data availability, the framework is tested with the LLCF diurnal cycle derived from the MPS observations at P6 in the western foothills of the SAM. The simulations for these independent cases were also conducted with and without shallow-LLCF forcing of varying depths for both Method-1 and Method-2 for 80 minutes and low-level forcing between 20 and 50 minutes. The simulations with and without LLCF reach steady rain-rate before 50 minutes for all cases, and the model simulated rain-rates and DSD spectra at 50th minute were considered for the analysis.

Figure 4-17(a) shows the intercomparison of rain-gauge observations (solid horizontal lines) and estimates from the GPM Ku-PR retrieval (triangles) and model simulations (other symbols with LLCF error bars). Overall, there is a substantial improvement for model simulations. GPM Ku-PR estimates and no-LLCF Method-1

simulations have the largest bias as indicated by the distance away from the rain-gauge observations, while the various model configurations with LLCF produce rain-rates closer to the rain-gauge measurements. In particular, note how the diamond symbols (Method-2) are closer to the rain-gauges than all other model simulations except for Case-1 and Case-5. In Case-5, the mean rain-rates for Method-1 and Method-2 simulations with LLCF are within 2 mm/h of each other with results from Method-1 LLCF simulations closer to the observations. Ground-clutter effects dominate up to 4 km AMSL in Case-1 where the reflectivity factor is approximately 40 dBZ. The lack of reliable reflectivity values below 4 km AMSL prevents detecting the bright-band, thus resulting in severe Method-2 model overestimation due to large drop sizes in the initial conditions, TBCs and rainfall rates (Table 4-4, dark blue values). The reflectivity and evolution of the model simulated rain-rate for individual cases are shown in the Appendix E (Figure E 5-10). In addition to ambiguity in detecting the bright-band, there is also very large variability of rainfall structure among surrounding GPM pixels (Figure E 5) suggesting potential NUBF effects in Case-1.
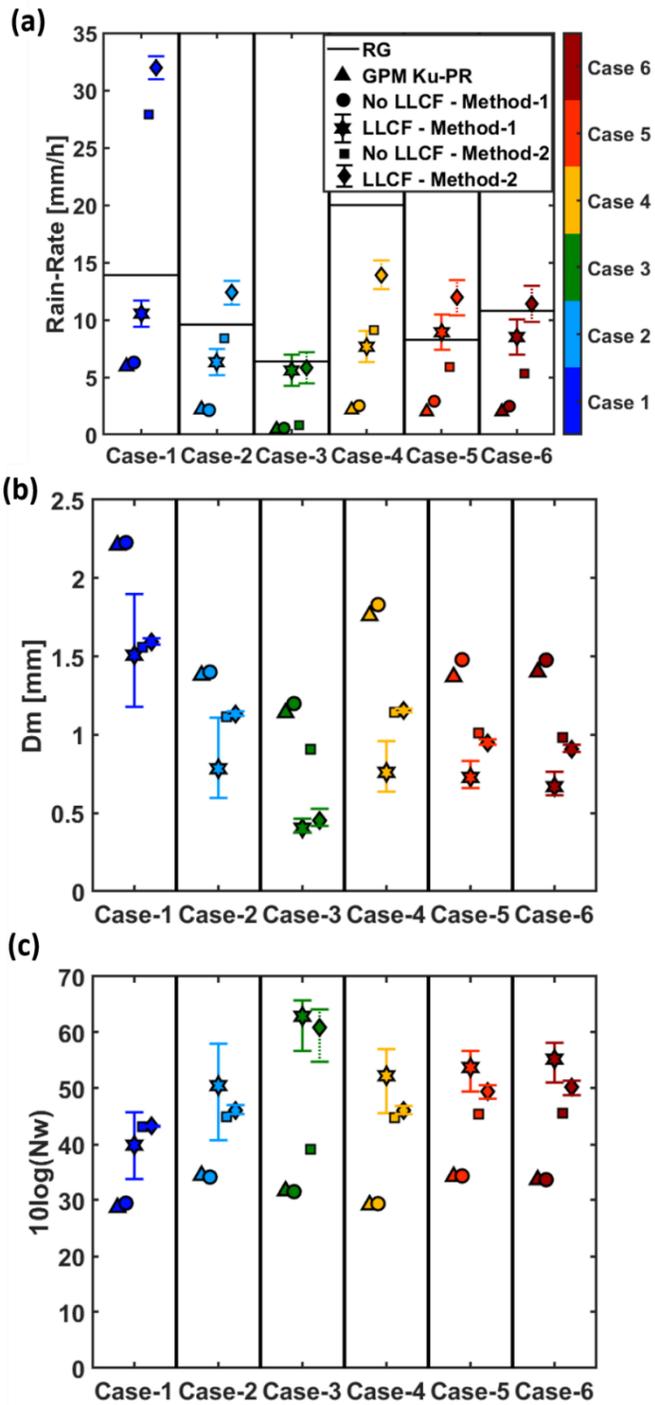
Figure 4-17 Systematic analysis of microphysical parameters in UND cases (a) Comparison of rain-gauge observations with rain-rate estimates from Method-1 and Method-2 with and without shallow LLCF and GPM Ku-PR estimates. Variations in

(b) Dm and (c)Nw for simulations following Method-1 and Method-2 with and without shallow LLLCF and GPM Ku-PR estimates. The error bars denote the mean, maximum and minimum (a) rain-rate, (b) D_m and (c) N_w obtained from 6 simulations corresponding to combinations of three fog depths (300, 400 and 500 m) with fog microphysical parameters estimated at 10- and 60-min timescales

Table 4-4: List of warm season underestimation cases with rain-rate observed by GPM Ku-PR, rain-gauges and Method-2 corrected precipitation estimates. Case description is strictly based on GPM Ku-PR products.

Case	Case Description	Date – Time EDT	Site	Rain-Rate (mm/h)			
				GPM Ku-PR	RG	Method-2 without LLCF	Method-2 with LLCF
1	*Stratiform, deep ground-clutter, spatial heterogeneity*	May 31, 2016 19:20	RG103	5.9	**13.9**	27.9	**31.0– 33.0**
2	*Bright-band between 3.175 and 4.25 AMSL*	May 21, 2017 11:30	RG112	2.2	**9.6**	8.4	**11.3 – 13.4**
3	*Stratiform case. Peak reflectivity at 4 km AMSL*	June 17, 2018 16:49	RG308	0.5	**6.4**	0.8	**4.5 – 7.2**
4	*Ground-clutter at 3.5 km AMSL. 0º at 5 km AMSL.*	August 11, 2014 19:52	RG106	2.2	**20.0**	9.1	**12.7 – 15.2**
5	*Stratiform case. 0º at 4.75 km AMSL.*	August 17, 2016 20:27	RG303 S	2.0	**8.3**	5.9	**10.4 – 13.5**
6	*Stratiform case. 0º at 4.75 km AMSL.*	August 17, 2016 20:27	RG310	2.0	**10.8**	5.3	**9.8 – 13.0**

Application of the physical-correction framework to Case-2 through Case-6 is summarized in Table 4-4 to show improvements in the precipitation estimates and their associated uncertainty in model simulations considering the presence and absence of shallow-SFI. Figure 4-17(b) and (c) show the D_m and N_w values for the GPM Ku-PR estimates and the model results for the six underestimation cases. Similar to rain-rate, GPM Ku-PR D_m and N_w estimates are approximately the same as for Method-1 without LLCF. LLCF microphysics is introduced in Method-1 (circles and stars) simulations resulting in D_m decreases and N_w increases, thus demonstrating the influence of LLCF on raindrop collision and break-up statistics. Method-2 simulations show increase in N_w and decrease in D_m values (squares and diamonds) compared to GPM Ku-PR estimates. Overall, D_m and N_w values for simulations with LLCF are consistently lower (D_m) and higher (N_w) than the GPM Ku-PR estimated DSD parameters. Method-2 without LLCF simulations also shows lower D_m and higher N_w than the Ku-PR estimates. Thus, using the exponential DSD to specify the TBC and initial conditions leads to more realistic DSDs near the surface independently of SFI.

The application of the proposed framework improves the rain-rate estimates and demonstrates the effect of LLCF in the estimation of DSD parameters for shallow and stratiform precipitation regimes. Although the differences between Method-1 and Method-2 with LLCF are relatively small, the differences for simulations without LLCF are large (circles vs squares in Figure 4-17) suggesting that Method-2 using GPM Ku-PR

corrected reflectivity to impose initial and top boundary conditions for the model is preferable to using retrieved DSD parameters in Method-1.

4.6 Discussion

In order to address the retrieval errors due to low-level processes in satellite-based QPE, a physically-based framework was demonstrated to estimate precipitation-rate based on the satellite-based DSD retrievals (D_m and N_w; Method-1) and proposed DSD retrieval method from measured reflectivity assuming a negative exponential distribution (Method-2) using a column rainshaft model. Direct applications of this approach to support interpretive process studies have been reported previously in the literature (Prat et al. 2008; Prat et al. 2010b; Wilson and Barros, 2014; Duan and Barros, 2017). Here, the proposed framework is applied to multiple case studies since GPM launch with substantial improvement in the estimation of rain-rate and realistic DSD spectra.

D'Adderio et al. (2015) reported that in natural rain at high rain-rates (> 5 mm/h), when break-up is dominant and the equilibrium DSD is not reached, the actual DSDs exhibit the same type of concave shape of the DSD at intermediate diameters as that for P5. This behavior raises the question of whether the Method-2 TBC DSD may be overestimating the number concentration of very small drops and underestimating the large drop sizes, thus overwhelmingly favoring collision-coalescence without forming large enough drops for break-up to play an important role. To address this question, an exploratory experiment was conducted by applying a Marshall-Palmer fit (MP, less steep DSD

spectrum; Marshall and Palmer, 1948) to define the TBC and initialize the model. In the MP distribution, the intercept factor is assumed to be a constant value of 8000 $m^{-3}mm^{-1}$ irrespective of reflectivity value leading to less steep DSD spectra compared to the methodology described in Section 3.3. The MP slope parameter depends on the rain-rate. The difference between the simulations with TBC specified using MP or the exponential fit is in the number of large drops. Indeed, the model results fit better the surface DSD in the intermediate size-range due to the increase in collision-breakup efficiency among larger drop sizes. Although the model simulation with the Marshall-Palmer fit underestimates the rain-rate unless LLCF forcing is applied (shown as Appendix E, Fig. E11), the results of the experiment do suggest there is potential gain in modifying the approach used by Wilson and Barros (2014) to estimate the TBC DSD for shallow precipitation cases toward a wider DSD spectrum in the case of stratiform rainfall with bright-band due to melting snow. Nevertheless, note that the ambiguities in the DSD spectra from Method-1 simulations are rectified in Method-2 simulations for this case.
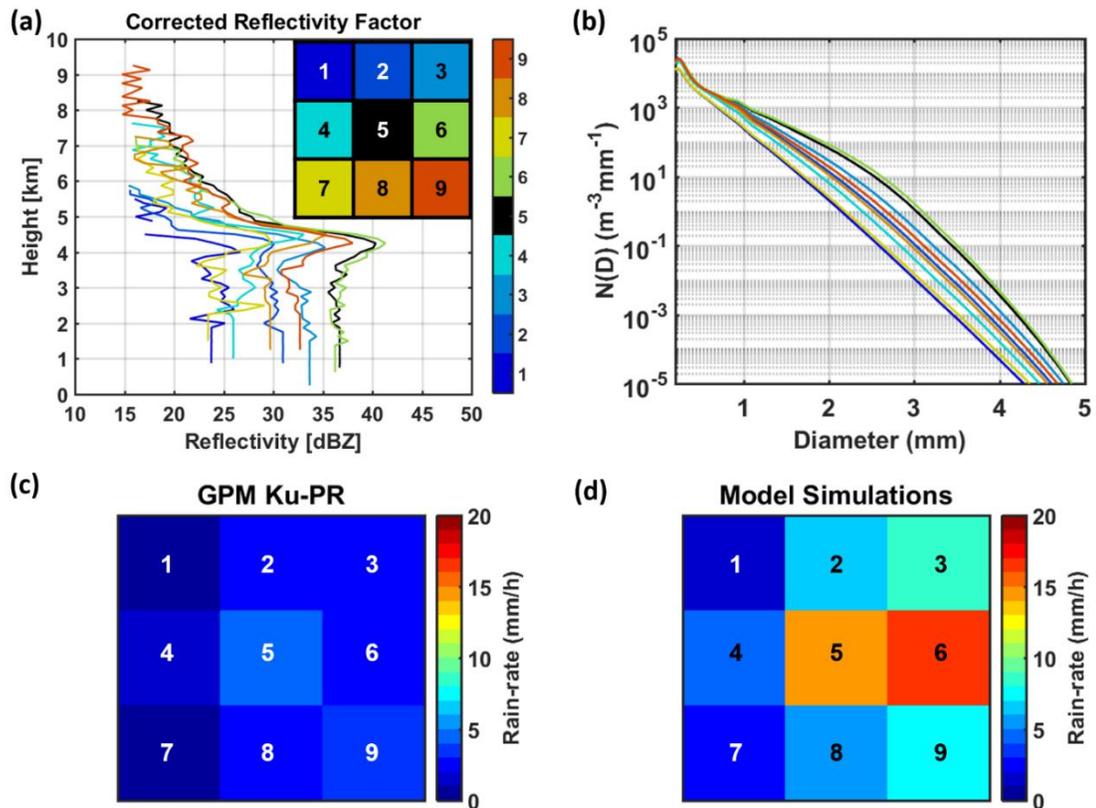
Figure 4-18 Spatial variability of precipitation microphysics (a) GPM Ku-PR corrected reflectivity factor for the August 8, 2014 case at P5 and 8 adjacent pixels. (b) Surface DSD estimated by Method-2 without LLCF forcing for August 8, 2014 case at P5 and 8 surrounding pixels. Note the two orders of magnitude difference in N(D) between pixels 1 and 5 for Diameter > 2 mm, which translate into one order of magnitude differences in rain-rate. The spatial distribution of rain-rates within the 8 pixel neighborhood centered at P5 is shown in the bottom row: (c) GPM Ku-PR and (d) Method-2 model simulations.

Table 4-5 Spatial variability of precipitation characteristics around P5 on August 8, 2014

Parameters	Method 2 Spatial Variability		GPM Ku-PR Spatial Variability	
	Minimum	Maximum	Minimum	Maximum
Rain-rate [mm/h]	1.95	16.97	0.74	4.88
D_m [mm]	0.85	1.30	1.20	2.02
$10\log(N_w)$	43.76	45.55	27.86	34.73
μ	1.02	6.67	3	3

Artifacts in GPM Ku-PR measurements due to NUBF and spatial heterogeneity can be severe for deep convective storms and in steep-terrain gradients that are not factored in the rainshaft model. One approach to addrss this uncertainty is to examine the spatial variability in microphysics and surface precipitation within neighborhoods of pixels around a location of interest. This is demonstrated here using the neighborhood of 8 pixels surrounding P5 to generate an ensemble of 9 possible solutions for the August 8, 2014 case (one solution at the pixel of interest P5, and 8 at surrounding pixels) using Method-2 without LLCF. The simulated ensemble shows that change in the magnitude and shape of the reflectivity profiles (Figure 4-18a) strongly impacts the model predicted surface DSD spectra parameters including the shape factor (Figure 4-18b, Table 4-5) with differences of up to two orders of magnitude in number concentrations for diameters > 1

mm. Inspection of Figure 4-18(c-d) shows that GPM Ku-PR estimates over the same neighborhood do not exceed 5 mm/h, whereas model rain-rates are in the range 1-20 mm/h and their spatial organization clearly delineates the rainfall gradient between active and weak areas within the storm system. The sensitivity of surface rain-rate and DSD parameters, and associated uncertainty in terms of spatial variability is summarized in Table 4-5. Note the spatial variability between model and GPM Ku-PR results in pixel-to-pixel differences in D_m of ~ 1mm and in N_w of ~ 10 dBNw. The shape parameter μ varies between 1 and 6 for the model predicted DSD depending on the reflectivity profile and amplitude and shape of the bright-band in contrast with fixed $\mu=3$ in the GPM Ku-PR algorithm.

4.6 Conclusion

This chapter presents a physically-based framework to improve QPE retrievals for bottom-heavy and layered orographic precipitation systems (low-level enhancement of precipitation). The framework was applied in the SAM to leverage the wealth of IPHEx GV data, and to maximize the concurrence of satellite-based and ground-based observations over the period of study. However, the methodology is not region or measurement dependent. The major findings of this study are:

1) MDs in GPM Ku-PR are attributed to light and low-level precipitation. More than 82% of the FAs are for stratiform precipitation, with GPM Ku-PR rain-rates below 1.0 mm/h for 67% of the cases.

2) Error analysis indicates that GPM Ku-PR detection and estimation errors are linked to shallow precipitation systems and low-level enhancement of precipitation by SFI, which are frequently observed regionally in the summer.

3) Based on the error diagnostics and physical interpretation, the retrieval errors are attributed to ground-clutter contamination of near-surface reflectivity values, errors in estimation of DSD parameters, and subgrid-scale spatial heterogeneity of precipitation systems leading to NUBF artifacts.

4) Comparison of DSD parameters from Parsivel disdrometers with GPM Ku-PR estimates suggest that the D_m is overestimated by approx. 0.03- mm, N_w is underestimated by 3-5 dBNw, and rain-rate from shallow precipitation systems is underestimated by as much as one order of magnitude in the presence of SFI.

5) A stochastic rainshaft model with prescribed LLCF microphysics (Duke Rain Microphysics Column Model, DRCM) is shown to predict realistic near-surface precipitation-rate and DSD from reflectivity measurements below the bright-band and above the ground-clutter contamination height with and without SFI depending on time-of-day and location.

6) Model top boundary and initial conditions were specified using two Methods: (1) directly from GPM DSD estimates, and (2) by fitting a distribution to GPM reflectivity factor measurements. Model simulations using both approaches result in similar improvements in near-surface QPE. However, Method-2 simulations

produce realistic surface DSDs close to the Parsivel observations for most of the case-studies.

7) Precipitation estimation errors due to ground-clutter artifacts can be explained by missing physics, specifically transient raindrop dynamics in the absence of LLCF, and low-level enhancement of rainfall due to SFI when LLCF is present. The physics-based correction framework addresses principally this type of errors.

8) Preliminary analysis suggests that spatial uncertainty in regions of strong spatial gradients in precipitation physics such as along the boundaries of storm systems can be as high as uncertainty due to missed microphysics.

Synergies between remote sensing instruments can be exploited to determine whether and when LLCF effects must be included at any location following for example Arulraj and Barros (2017) using multi-frequency radar to detect light rain and low–level enhancement, and Duan and Barros (2017) using spectrometer and lidar observations to detect persistent LLCF features. Mesoscale numerical weather prediction guidance and observations from geostationary satellites can be used between overpass times as needed either to fill-in data gaps, or in the context of data assimilation. In the applications presented here, LLCF microphysical parameters were obtained from the diurnal cycle climatology derived from ground-based observations in the SAM. Whereas preliminary results suggest that the uncertainty associated with LLCF microphysical parameters is substantially lower than retrieval errors due to missed processes, it is necessary to develop

strategies for parameter estimation that are not location specific, that is anywhere anytime. Finally, robust statistical analysis of the physical-correction framework is constrained in this study by the number of overpasses over the GV sites during rainy conditions since GPM launch, and thus further systematic evaluation is required.

5. Coupling GPM Observations and Numerical Weather Prediction Simulations to Predict Vertical Microstructure of Precipitation.

5.1 Introduction

Satellite-based Quantitative Precipitation Estimates (QPE) provide the most realistic representation of precipitation intensity on a global scale since it is impossible to establish spatially dense long-term in-situ measurement networks. The accuracy of tropical and subtropical precipitation products has improved significantly in the past two decades after the launch of the Tropical Rain Measurement Mission satellite in 1997 (TRMM; Simpson et al. 1996). TRMM's payload included a Ku-Band (13.8 GHz) precipitation radar (TRMM PR). Extensive error analysis of TRMM-PR precipitation products in different regions around the world (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) shows strong dependence on topography and precipitation regime, detection errors predominating in light and low-level precipitation, and severe underestimation of heavy and winter precipitation events (Prat and Barros, 2010; Duan et al. 2015; Wilson and Barros, 2014; Duan and Barros, 2017). In particular, Prat and Barros (2010) and Duan et al. (2015) compared the TRMM-PR estimates with measurements from a long-term spatially dense rain-gauge network in the Southern Appalachian Mountains (SAM) and found robust patterns in the spatial and temporal organization of the detection and estimation errors at diurnal and seasonal scales.

The Global Precipitation Measurement Mission (GPM; Hou et al. 2014) was launched in 2014 as a successor to the TRMM 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- (35.5 GHz; Ka-PR) Band. 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 was approximately 0.5 mm/h while GPM Ku-PR is 0.2 mm/h (Hou et al. 2014; Speirs et al. 2017). However, in practice, both TRMM-PR and GPM-DPR exhibit higher sensitivity than their nominal detectability (Hamada and Takayabu, 2016).

Speirs et al. (2017) evaluated GPM-DPR precipitation estimates using radar observations in the Swiss Alps and Plateau and highlighted that the GPM-DPR underestimates precipitation by more than 50% in the winter season. Arulraj and Barros (2019) also found severe underestimation of precipitation events in the SAM especially in the presence of multi-layer cloud systems and seeder-feeder interactions. Overall, error diagnostics studies of GPM-DPR indicate improved performance in detection and estimation scores compared to TRMM-PR (Liu, 2016; Arulraj and Barros; 2019), but retaining the spatial and temporal organization characteristics. The latter extends to the Integrated Multisatellite Retrievals for GPM (IMERG) that integrates precipitation estimates from several microwave and infrared sensors calibrated using GPM

measurements to produce global maps. Rios Gaona et al. (2016) found that IMERG underestimates precipitation by approximately 2% in the smooth topography of the Netherlands. Likewise, Khan and Maggioni (2019) determined that the IMERG detects rain events with an accuracy of 80% but underestimates the intensity over the ocean. By contrast, Barros and Arulraj (2020) evaluated the IMERG precipitation product with Stage-IV precipitation products in the SAM and observed spatial and temporal patterns of detection errors consistent with the climatology of low-level clouds and fog (Wilson and Barros 2014, 2015 and 2017; Duan and Barros, 2017).

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 satellite-based radar measurements of shallow precipitation systems in complex terrain such as the SAM (Fig. 5.1): (1) non-uniform beam filling (NUBF) artifacts tied to the resolution of the radar beam; (2) ground-clutter contamination in the near-surface reflectivity; and (3) incorrect microphysical parameters. NUBF artifacts are enhanced in mountainous regions due to the 3-D complexity of low-level circulations modulated by the terrain resulting in high spatial variability 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). Presently, the PR retrieval algorithm extrapolates the reflectivity factor at the lowest no-clutter level to the surface yielding uniform reflectivity profiles in the lower 2 km 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). However, concurrent ground clutter and NUBF can result in overestimation as well as increased spurious detection (false alarms, FA) thus undermining the reliability of common posteriori statistically-based correction of these errors. These errors are further compounded in complex terrain by high spatial and temporal heterogeneity in the vertical structure of low level precipitation that is apparent in the spatial distribution 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).
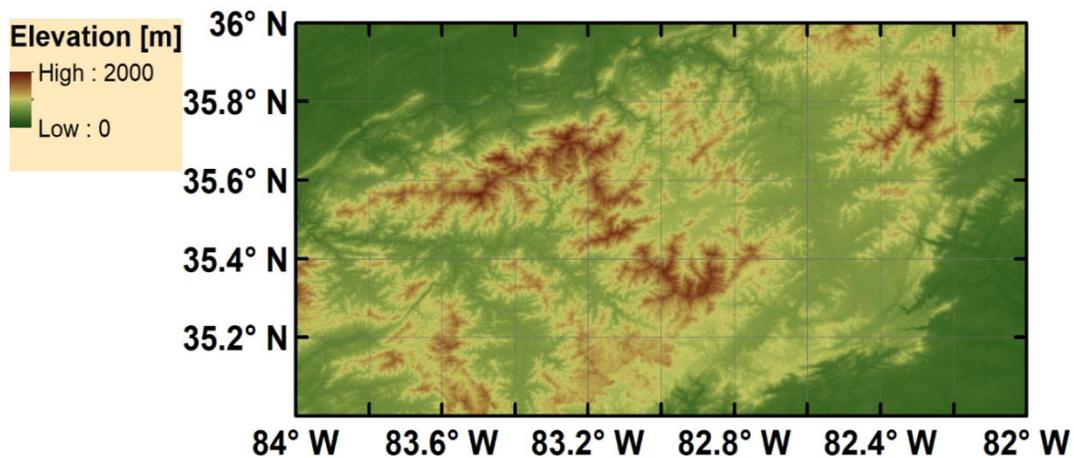
Figure 5-1: Digital Elevation Model (DEM) map of the Southern Appalachian Mountains (SAM).

Based on long-term observations in the SAM, Arulraj and Barros (2019, see also Chapter 4) demonstrated the effectiveness of using a physically-based approach to

quantitative precipitation estimation (QPE) from shallow precipitation systems in complex terrain (e.g. in the SAM) using two alternative conceptual models of the vertical structure of layered and low level clouds and fog (LLCF), and a dynamic stochastic model of rainshaft microphysics with boundary conditions based on Ku-PR reflectivity above ground-clutter height and initial conditions (hydrometeor size distributions) based on the reflectivity profile. An impediment to the systematic (and global) application of their approach is the lack of criteria to guide the configuration of the LLCF vertical structure where ground-based observations are not available. 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 at point locations where ground radars operate. To overcome this limitation, it is hypothesized that vertical information on the distribution of liquid and solid phase water in Numerical Weather Prediction (NWP) model simulations or reanalysis depending on desired latency that runs globally can provide adequate constraints for this purpose even if the model themselves lack the required fidelity to simulate and, or predict clouds and precipitation. The idea is to infer metrics to detect and classify shallow precipitation systems from coupled high-resolution and long-term NWP information and GPM-DPR measurements.

To test this hypothesis and investigate the operability of the proposed approach, a new data-driven framework is proposed that aims to improve the convective/stratiform classification algorithm. Early on, a strategy consisting of developing a look-up table classification of shallow precipitation systems from NWP-based simulations was pursued. The NWP model of choice is the High-Resolution Rapid Refresh (HRRR; Benjamin et al. 2016) that runs operationally at 3 km resolution over CONUS. To evaluate the HRRR reliability for this purpose, 4 years of hourly simulations were analyzed over the SAM domain (Fig. 5-1) with a focus on simulating model reflectivity profiles following Leinonen et al. (2015). Evaluation of HRRR derived reflectivity profiles was conducted against the database of 3-D Multi-Radar Multi-Sensor System (MRMS) reflectivity fields concurrent with GPM overpass over the SAM. In addition, a climatology of precipitable water, cloudiness, and precipitation were also derived for evaluation against ground-based and satellite observations. The HRRR climatology shows lack of timing errors (e.g. 3-5 hours in the SAM for the Weather and Research Forecasting model over the SAM, Wilson and Barros, 2015) and it captures well precipitation maxima on the eastern ridges of the SAM, but it fails to capture the cloudiness and precipitation maxima over the western ridges of the SAM (Figures in Appendix F). Based on this result, it was decided to change strategy and use the MRMS reflectivity profiles to identify shallow precipitation systems directly from GPM measurements. Because of the overwhelming frequency of no-rain observations vis-à-vis rain observations, the data-driven model consists of two

separate sub-models for detection and classification. The detection sub-model relies on GPM and HRRR low level (< 1.5 km AGL) water mass distribution (for example rain-water mixing ratio, snow-water mixing ratio, and graupel mixing ratio). The classification sub-model classifies GPM-DPR profiles based on LLCF structure classes inferred from MRMS climatology. To accomplish this goal, the following specific tasks are required: 1) to quantify the detection estimation errors in GPM-DPR using rainfall estimations from ground-based MRMS; 2) to analyze and quantify the HRRR results with respect to the ground-based rain-gauge observations and the MRMS precipitation product; 3) to cluster the MRMS reflectivity profiles based on vertical structure with special attention to LLCF; 4) to develop a precipitation detection algorithm and deep learning-based classification algorithm to identify error types and vertical structure during the absence of ground-based observations.

Section 5.2 provides a brief description of the GPM DPR and GMI data, the ground-based radar reflectivity observations from MRMS, the HRRR model and the ground-based rain-gauge network. Section 5.3 explains the methodology of error analysis, clustering of ground-based reflectivity observations, detection and classification framework to predict the error types and the near-surface vertical structure of reflectivity. The results are presented in Section 5.4 followed by the discussion in Section 5.5. Section 5.6 provides a summary and conclusion.

5.2 Data Description

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

The 3-D reflectivity profiles from the MRMS network will be utilized in this study. The 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
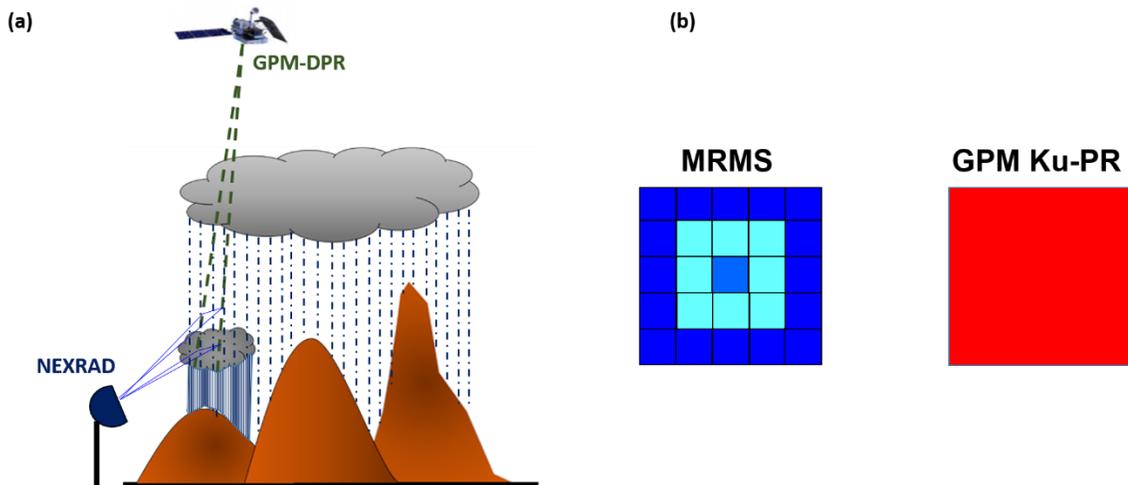
Figure 5-2: (a) Schematic showing the PPI scanning pattern of the NEXRAD and the vertical scanning pattern of the GPM DPR. (b) Footprint of the MRMS and GPM DPR.

(CONUS). The data used in this study are the merged, quality-controlled and gridded 3D reflectivity profiles with a spatial resolution of 0.01 degree (approximately 1 km) and the vertical height ranging from 0.5 km until 19 km with a temporal resolution of 2 minutes. The range resolution varies between 250 m and 1 km. The quality-control process removes echoes from non-hydrometeors and random clutters due to beam blockage (Zhang et al.

167

2011). The NEXRAD radars operate at Plan Position Indicator (PPI) scanning mode. The schematic of the PPI scanning mode is shown in Fig. 5-2(a).

In addition to the 3-D reflectivity profiles, Level 2 precipitation rate and precipitation type products are also used in this study. The spatial resolution of these products is 0.01 degrees (~ 1 km) and the temporal resolution is 2 minutes. The precipitation rates are gauge corrected using 9000 collocated rain-gauges during rain events while the snow events are radar only estimations (Yang and Gourley 2018). The precipitation type product classifies the data into seven different categories: warm stratiform rain, snow, convective, hail, tropical/stratiform mix, tropical/convective mix and cool stratiform. The classification is performed based on the reflectivity values from the radar and 3-D temperature data from the Rapid Refresh model (Yang and Gourley, 2018). The detailed description of the precipitation type classification methodology is available in Yang and Gourley (2018). MRMS data in the nearest 2-minutes of GPM overpass are considered for this study.

5.2.2 Global Precipitation Measurement Mission

The GPM-DPR operates at Ku- (13.6 GHz) and Ka- (35.5 GHz) band with a spatial resolution of approximately 5 km. The Ku-PR operates in normal scan (NS) mode with a cross-track swath width of 245 km and 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 of the Ku-PR varies from nadir (0°) to ±18°. Ka-PR operates in two modes: matched scan (MS) with

sampling and range resolution similar to that of Ku-PR NS; high sensitivity scan (HS) with a sampling resolution of 250 m and range resolution of 500 m. Level-2 GPM Ku-PR and version 06A data are used in this study.

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. The mean footprint resolution depends on the operating frequencies varying between 6 and 25 km (Draper et al. 2015). The calibrated brightness temperatures from GMI Level 1C Version 05A data will be used in this study.

5.2.3 High Resolution Rapid Refresh (HRRR) Model

The Rapid Refresh (RAP) is a version of the Weather Research and Forecasting (WRF) model developed by the NOAA Earth System Research Laboratory (ESRL) Global Systems Division (GSD). This is an hourly updating, cloud-resolving, convection-allowing model run operationally by the National Centers for Environmental Prediction's Environmental Modeling Center (EMC) with a nominal resolution of 13 km (Benjamin et al., 2016). The model domain covers the entire North America with up to 21 forecast lead times. In the spring of 2016, a high-resolution nested version of the RAP called the High Resolution Rapid Refresh (HRRR) was developed with 3-km horizontal grid and one hour update. The HRRR covers the CONUS with 0 to 18 h forecast durations. HRRR is a convection-allowing model and is strongly dependent upon RAP data assimilation

including radar reflectivity. Hourly HRRR simulation (data assimilation and forecast modeling systems) data used here are obtained from a private cloud object store developed by the Center for High Performance Computing (CHPC) at the University of Utah (Baylock et al. 2017).

In this study, variables at every pressure level from the HRRR model at 0 h forecasts are extracted every hour for the study region. The HRRR model is available from July 2016 and the database is updated in real-time. The time-period considered for this study is between July 2016 and May 2019. Hourly instantaneous variables such as the 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] are downloaded at the pressure levels between 50 mb and 1000 mb with a resolution of 25 mb.

5.2.4 Ground-Based Observations

The ground-based rain-gauge data used for the evaluation of HRRR model results are obtained from the long-term spatially dense tipping bucket rain-gauge network operating in the Great Smoky Mountains National Park within the SAM (Barros et al., 2017). The geographical locations of the rain-gauges are shown in Fig. 4-1(b) and the details of the geographical coordinates and elevation of the rain-gauge locations are shown in the Appendix E (Table E1). The network has three different type 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.

5.3. Methods

5.3.1 Error Analysis

5.3.1.1 GPM Ku-PR and MRMS

Error analysis is first conducted by comparing instantaneous GPM Ku-PR QPE with 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. 5-2(b). Because of the different foot-print resolutions, the 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 pixels. In both these approaches, the spatial heterogeneity of the precipitation systems is not taken into consideration which might lead to a bias in the estimation of detection error metrics such as the probability of detection (POD; Wilks, 2011) and false alarm ratio (FAR; Wilks, 2011). Consequently, a different strategy was developed to perform the error analysis based on the fractional area

of MRMS rainy pixels "x%" within a GPM pixel: if "x%" of the MRMS pixels observe precipitation greater than 0.1 mm/h, then the ground truth is considered as rain, or else the ground-truth is assumed to be no-rain. Based on this approach, the number of correct hits (YY; 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 (MR; 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 value of "x" is computed subsequently as follows:

$$FB = \frac{YY+FA}{NN+MD} \tag{5-1}$$

FB is the ratio of frequency of precipitation detection by the GPM-DPR to the frequency of precipitation detection by the ground-truth (MRMS). If FB is greater than 1, then increased FA cases are observed; and if FB is less than unity, then number of MDs dominates the number of FAs. 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 percentage of precipitation pixels to classify the MRMS aggregated pixel corresponding to the GPM pixel as rain.

172

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 and is calculated as follows:

$$POD = \frac{YY}{YY+MD} \tag{5-2}$$

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

$$FAR = \frac{FA}{YY+FA} \tag{5-3}$$

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

$$\varepsilon = log\left(\frac{\sum_i R_{i,GPM}}{\sum_i R_{i,MRMS}}\right) \tag{5-4}$$

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 underestimations of GPM, and overestimations are indicated by the positive values of ε.

5.3.1.2 Evaluation of HRRR results using rain-gauges

Generally, NWP models exhibit 3 to 5-hour delay in detecting certain type of precipitation systems (e.g. Wilson and Barros, 2015 and 2017; Erlingis and Barros, 2014

and many others). Thus, the first step of the error analysis is to conduct correlation

analysis between the HRRR results and the rain-gauge observations at different time lags.

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 follows a negative exponential distribution of

the form:

$$N(D) = N_0 \exp(-\Lambda D) \tag{5-5}$$

where N_0 is the intercept parameter with fixed value of 8×10^6 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:

$$\Lambda_r = \left(\frac{\pi N_0 \rho_r}{\rho_{air} q_r}\right)^{\frac{1}{4}} \tag{5-6}$$

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:

$$\rho_{air} = \frac{P}{R_D T_v} \tag{5-7}$$

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:

$$T_v = T \times \frac{0.622 + q_v}{0.622 \times (1 + q_v)} \tag{5-8}$$

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

Finally, the rain-rate [mm/h] is calculated as shown below:

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

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]. The rain-gauge observations 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:

$$r^2 = \frac{\widehat{Cov}^2(R_{HRRR}, R_{RG})}{\widehat{Var}(R_{HRRR}) \times \widehat{Var}(R_{RG})} \qquad (5\text{-}10)$$

The value of r^2 varies between 0 and 1, the latter being the perfect score.

5.3.2 Clustering of Reflectivity Profiles

Previously, Zhang et al. (2007) clustered CloudSat Cloud Profiling Radar (CPR) reflectivity profiles in the tropics identified 5 clusters corresponding to low cloud and cirrus, subtropical maritime stratus, anvil cirrus cloud, cumulus congestus, and deep convection and heavy precipitation. In order to characterize the structure of precipitation systems in the SAM, the reflectivity profiles from MRMS are organized into classes of precipitation regime using a clustering algorithm in a manner similar to Zhang et al. (2007). Specifically, the MRMS reflectivity profiles are analyzed using a k-means algorithm (Anderberg, 1973) to identify multiple clusters based on the vertical structure of the reflectivity values. The k-means clustering algorithm clusters the data points with

'N' different features into 'K' different clusters based on the intra-cluster and inter-cluster distance.
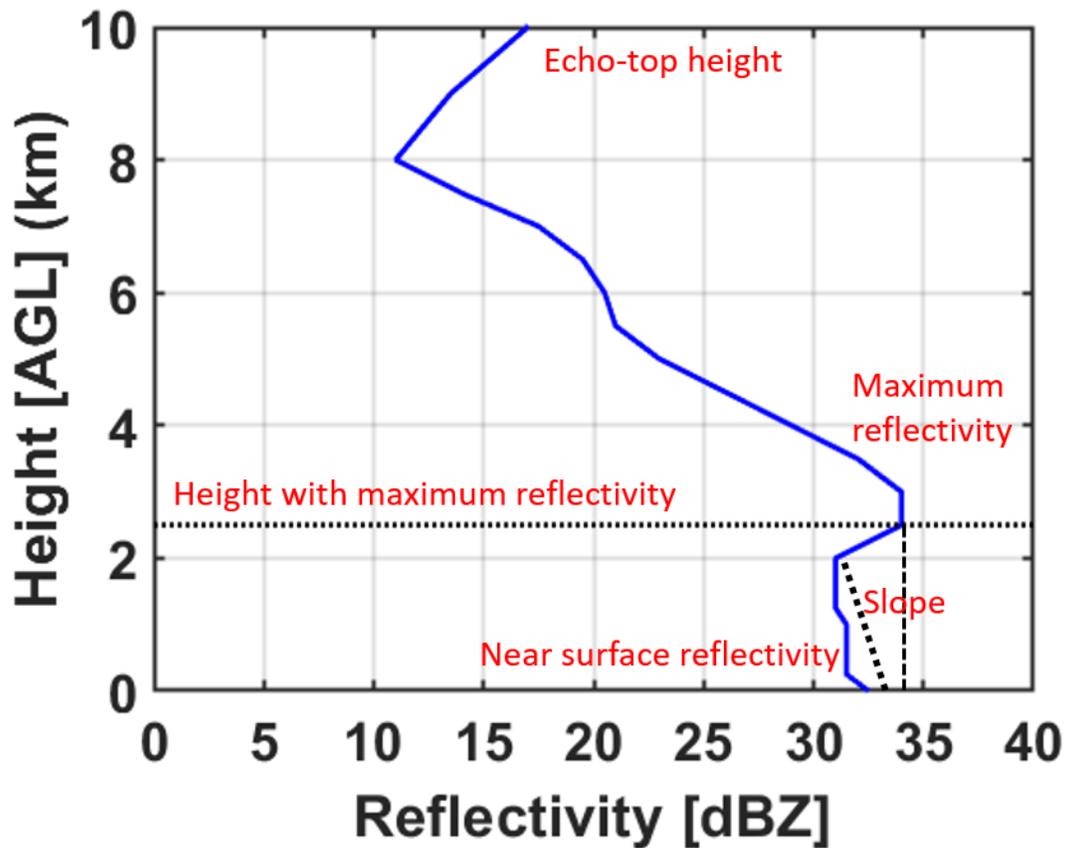
Figure 5-3: Features extracted from the MRMS reflectivity profiles for clustering algorithm.

In this study, the primary objective is to predict the precipitation vertical structure, especially in the reflectivity values at the lower 2-3 km near-surface that are generally missed by the GPM-DPR due to ground-clutter. Thus, the MRMS clustered into different categories based on the specific features to address the objective of this study. The features considered here for the k-means clustering algorithm are the maximum reflectivity,

maximum reflectivity height, reflectivity near the surface, echo top height of precipitation

systems, and the slope of the reflectivity profiles within 2 km near the surface. Figure 5-3

shows an example of a reflectivity profile with the features marked. The near-surface

slope is computed as follows:

$$Slope = \frac{Z_{H1} - Z_{H2}}{H1 - H2} \tag{5-11}$$

Where Z_{H1} and Z_{H2} are respectively the reflectivities at heights H1 and H2. H1 is

specified at 2 km AGL that is the GC height in GPM DPR since the focus here is on shallow

precipitation systems, and H2 is the surface. The optimal number of clusters is decided

using the Davies Bouldin (DB) index (Davies and Bouldin, 1979) that is calculated based

on the ratio of intra-cluster and inter-cluster distances. The best number of clusters is the

number corresponding to the lowest DB index indicating low variance within each cluster

and high variance among clusters.

Finally, the reflectivity profiles in each cluster are classified according to the

underlying precipitation regime conditional on precipitation type and intensity. Further,

each cluster is expected to be associated with certain type of detection and estimation

errors also. Subsequently, these clusters are used to train the error diagnostics framework

and to select an appropriate LLCF model to fill in the near-surface structure of GPM Ku-

PR reflectivity that is contaminated by ground-clutter.

5.3.3 Precipitation Detection Framework

Whether the optimal number of clusters can be mapped to physically meaningful reflectivity morphologies depends strongly on the quality of the data set as measured by the ability to identify unambiguous precipitation regimes. In the context of this work, this translates to improved precipitation detection in order to remove from training the large number of no-rain cases that introduce high frequency bias. For this purpose, all GPM GMI calibrated brightness temperatures (Tbs) that are concurrent but independent of the DPR measurements and low-level water mass from HRRR are the data selected to drive the precipitation detection framework. The GMI Tbs at various frequencies are expected to provide guidance with regard to the presence of 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), where the low-level water mass from HRRR is indicative of the likely presence of shallow precipitation. The precipitation detection algorithm was developed based on a random forest classifier (Breiman, 2001). A random forest classifier has been successfully used in various applications including land-cover classification of remote sensing observations (Ham et al. 2005; Belgiu and Dragut, 2016; Kulkarni and Lowe, 2016).

A random forest classifier 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 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 subset 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 son nodes based on impurity measures such as the Gini index (Breiman et al. 1984). The Gini index at node 't' is defined as follows:

$$G(t) = \sum_{i \neq j} p(i/t)p(j/t) \tag{5-12}$$

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 "son" nodes. These impurity measures implicitly act as a feature selection method and provide the most important features in the training data in 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 algorithm's ability to provide important features helps to extract the physical interpretability of the model. The performance of the model 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 is 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 Appendix F (Fig. F5).

In the study, a random forest classifier is used for precipitation detection since the classifier is robust in handling non-linear datasets with a large size of features. Here, the input features include data from HRRR and GMI. Other metadata from DPR such as the terrain elevation, the pixel-specific ground-clutter bin height, and the melting layer height are also considered as input features. The GMI input features considered are the calibrated multichannel Tbs from the 9 channels (10.65 – 89 GHz V/H) within 10 km of the GPM Ku-PR radar beam pixel. The input features from HRRR data include the RWMR, SNWR, and GRLE values and the depth (number of non-zero values) of the rain, snow, and graupel with 1.5 km AGL. Because the HRRR simulations are available on an hourly basis, the two nearest times (before and after the overpass) are used in the analysis. It is recognized that 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 precipitation systems. Thus, higher temporal resolution NWP data would be highly desirable in practical applications.

The 5-fold cross-validation approach is used to train a random forest classifier. The random forest classifier has a maximum depth of 20 and the total number of decision tree ensembles is 500. The minimum number of samples per node is set as 5. When any node reaches 5 samples, further splitting of the corresponding node will not be performed on the data. The schematic of the framework is shown in Figure 5-4. Another major issue with the input data is the imbalance in the representation of different classes. This is a common issue in many fields such as fraud detection, disease detection, and precipitation detection. Here, the number of samples with precipitation is usually lesser than the no-precipitation class which will cause a bias in the metrics evaluating the performance of the classification algorithm. Thus, class weights are defined to penalize the misclassification of precipitating events. Finally, all the samples that are classified as rain will be further classified into the type of precipitation profiles and the methodology is explained in the next section.
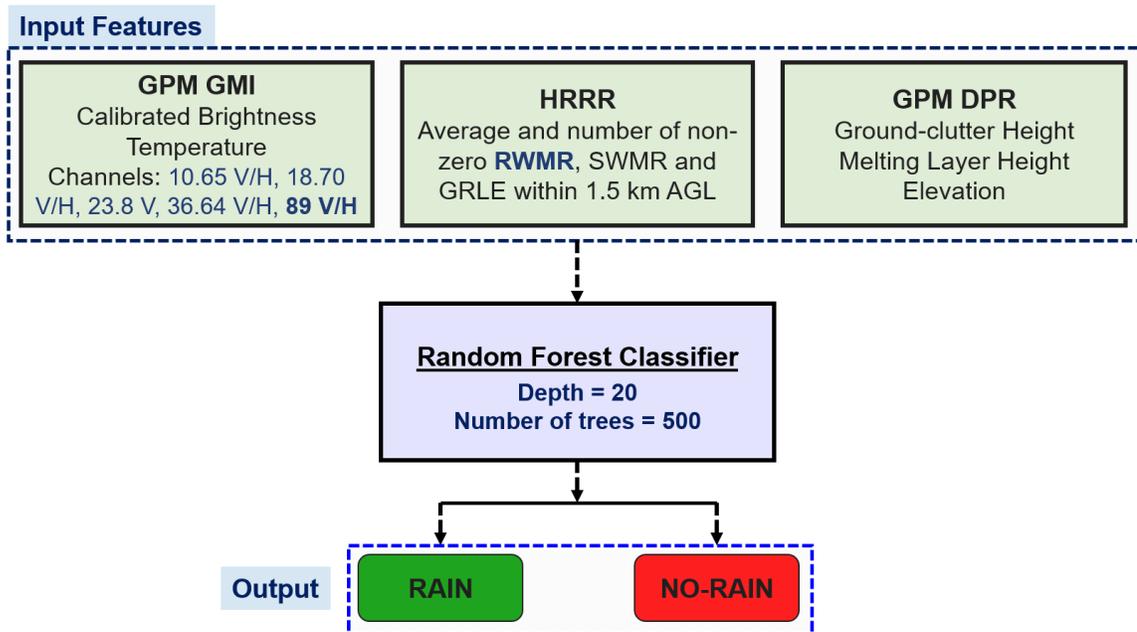
Figure 5-4 Schematic of the Random forest classifier precipitation detection framework. The inputs of this framework are obtained from GPM observations and HRRR model.

5.3.4 Precipitation Classification Framework

Artificial neural network algorithms such as the back propagation has proven to provide better results in classification and estimation problem in the past two decades (Heermann and Khazenie, 1992; Bruzzone and Serpico, 1997; Kulogowski and Barros, 1998). The increase in computation power and large dimensional data in the past decade has led to the growth of deep learning algorithms such as convolutional neural networks (CNN). Numerous studies in various field has already highlighted the potential of CNN 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).

CNN are similar to regular back-propagation neural networks with neurons in the input layer that receive inputs which are connected to multiple hidden layers and finally the output layer provides the outputs. The weights in all the layers are calculated iteratively based on the loss function. The architecture of the CNN can be categorized into two parts: feature extraction and classification (LeCun et al. 1998). Feature extraction part has two layers namely the convolution layer and the pooling layer. A convolution layer consists of a set of convolving filters with specific kernel size that extracts meaningful features from the input data. In convolution layer, a dot product is performed between weights of the convolving filters and the different parts of the input data followed by the activation function (f).

Assume 'X' is the input data with 'N' samples and 'm' features ($X = [x_1, x_2,…, x_N]$; $x_1=[x_{11},x_{12},…,x_{1m}]$) and '$Y$' is the final output vector with class labels [Number of classes – 'c']. 'X' is passed as input to the 1st convolutional layer with 'K' filters with kernel size 'k_1', then the output from a filter will be,

$$z_i = b_i + \sum_{j=1}^{m} Convolution(w_{i,j} * x_j) \tag{5-13}$$

Where $w_{i,j}$ is the weight vector between i^{th} and j^{th} feature, b_i is the bias of the ith feature and z_i is the output of ith filter. Further, z_i is transformed to 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 '$(N-k_1+1)xK$'. Next, pooling is applied for the output from the convolutional layer ($s = [s_1, s_2,…, s_K]$)]. 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 $\frac{N-k_1+1}{P} \times K$.

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:

$$\hat{\boldsymbol{y}}_{f1} = f(b_{f1} + \sum_{i=1}^{c_1}(\boldsymbol{w}_{1,i} \times \boldsymbol{F})) \tag{5-14}$$

Where $\boldsymbol{w}_{1,i}$ is the weight of the ith 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 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:

$$v_{dW} = \beta v_{dW} + (1 - \beta)dW \tag{5-15a}$$

$$v_{dB} = \beta v_{dB} + (1 - \beta)dB \tag{5-15b}$$

$$W = W - \alpha v_{dW} \tag{5-15c}$$

$$B = B - \alpha v_{dB} \tag{5-15d}$$

Where W is the weights, 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:

$$s_{dW} = \beta s_{dW} + (1 - \beta)dW^2 \tag{5-16a}$$

$$s_{dB} = \beta s_{dB} + (1 - \beta)dB^2 \tag{5-16b}$$

$$W = W - \frac{\alpha \mathrm{dW}}{\sqrt{s_{dW}}} \qquad\qquad (5\text{-}16c)$$

$$B = B - \frac{\alpha \mathrm{dB}}{\sqrt{s_{dB}}} \qquad\qquad (5\text{-}16d)$$

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.

Here, categorical cross-entropy is used as the loss function to train the model since the problem is a multiclass classification problem. The categorical cross-entropy loss function is mathematically defined as,

$$L(y, \hat{y}) = y \times \log(\hat{y}) \qquad\qquad (5\text{-}17)$$

where, y is the actual class label and $\hat{y}$ is the model predicted class label.

This CNN-based framework proposed in this study is a supervised classification problem that can predict the shape of the near-surface GPM Ku-PR reflectivity profiles given the inputs from GPR Ku-PR and GMI. The ground-truth is the class label from the MRMS reflectivity clusters and the objective is to predict the reflectivity shape using the trained model even in the absence of ground-truth (observations). The measured reflectivity factor profile, melting layer height, minimum height with no ground-clutter and elevation are the inputs from GPM Ku-PR. The calibrated brightness temperatures for 9 channels of the GMI are the input features from GMI. Only 9 channels are considered for analysis due to the spatial resolution. The GPM Ku-PR overpass are assigned a class name based on cluster labels (from section 5.3.2) provided to the MRMS reflectivity

profiles. Note, for every GPM overpass case, all the 25 MRMS reflectivity profiles within the radar footprint are assigned a cluster. The dominating cluster number (mode) within the PR's footprint is assigned as the class label.

Firstly, preprocessing is performed on the features extracted from the GPM estimations. The Ku-PR measured reflectivity profile has a dimension of 176x1 for each beam that is reflectivity values varying with height above mean sea level. In order to reduce the dimension of the reflectivity profiles, auto-encoders are used to retrieve 10 features from each reflectivity profiles. Auto-encoder is a dimension reduction technique where the dominant patterns are extracted and are commonly used in compression of dataset. Thus, the input to the framework comprises of 21 features, 10 features from Ku-PR measured reflectivity factor, 9 features from GMI estimated brightness temperatures and the 3 features from Ku-PR estimations including the melting layer height, no-clutter height and elevation.

In this study, a deep CNN is used where several stages of convolution layers, pooling layers and fully-connected layers are stacked to train the precipitation detection and classification framework. The deep learning models are prone to overfitting due their complicated structure and large number of features compared to traditional neural networks. Overfitting can be avoided using the regularization techniques. Dropout (Srivastava et al. 2014) is used in this study as a regularization technique. Dropout randomly ignores selected neurons during training where their contribution to the

forward layer are temporary removed and the weight update in the backward pass are

not performed on the dropped neurons. The schematic of the deep CNN architecture

proposed for the error diagnostics of GPM is shown in Figure 5-5.
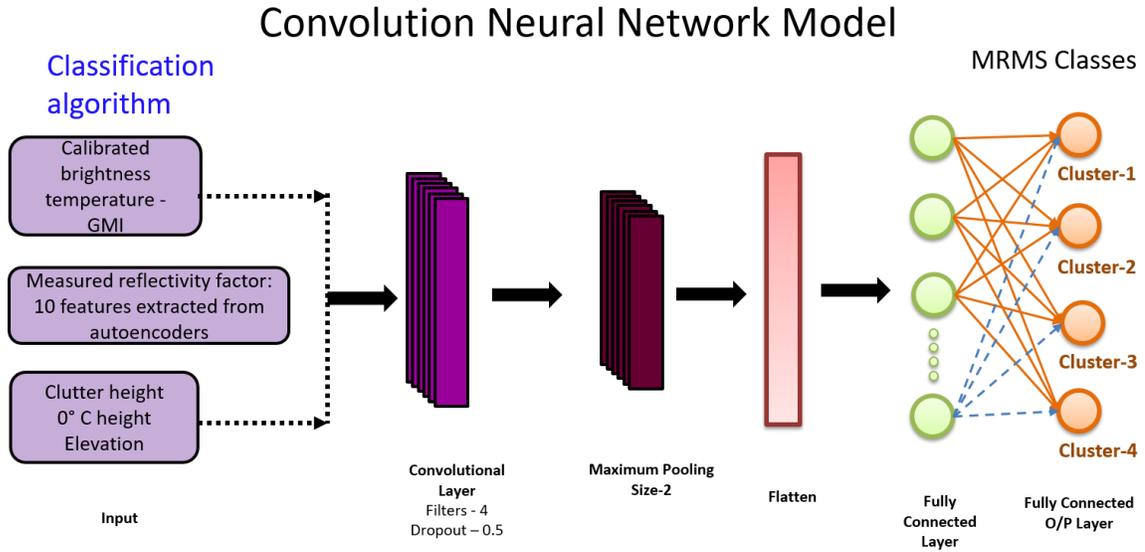
Figure 5-5 Schematic of the convolutional neural network framework used to classify rain profiles with respect to MRMS precipitation structure clusters.

5.4 Results

5.4.1 Training Data Model - MRMS Precipitation Structure Classes

5.4.1.1 Error Analysis of GPM Ku-PR Using MRMS

Table 5-1: Contingency matrix comparing the rainfall from the GPM Ku-PR with the MRMS cases. The values in the parenthesis are the percentage.

	MRMS = 0	MRMS ≠0

GPM Ku-PR = 0	24438 (86.9%)	917 (3.3%)
GPM Ku-PR ≠ 0	779 (2.8%)	1972 (7.0%)
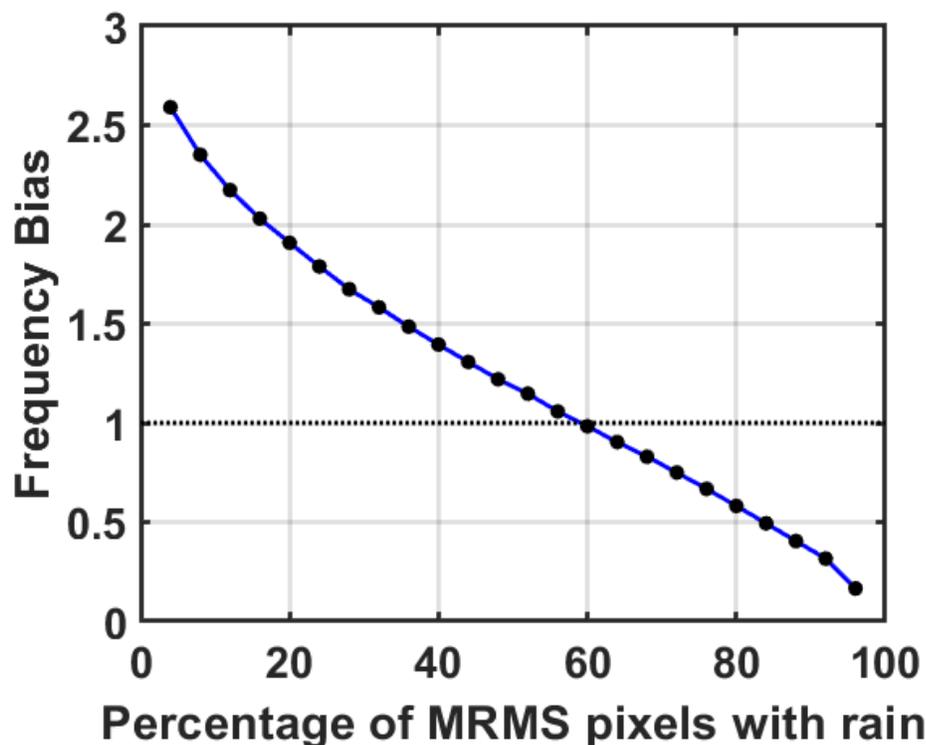
Figure 5-6: Frequency bias for non-uniformly precipitating MRMS within the GPM Ku-PR footprint

GPM Ku-PR overpasses are compared with the MRMS reflectivity, precipitation rate and precipitation types. Based on the availability of MRMS data, 28106 GPM Ku-PR radar beams are selected for the analysis. Also, note that one pixel of GPM Ku-PR corresponds to 5x5 pixels of MRMS as shown in Fig. 5-2(b). First, the contingency matrix is calculated for the "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 28106 profiles, 24930 are identified as "homogeneous" cases. Note that more than 92% of the "homogeneous" cases are no-precipitation, 5.31% of the cases are correct detection of precipitation, 1.47% are MDs and 0.79% are FAs. This illustrates the frequency bias issue raised earlier. For the "heterogeneous" cases (3176 profiles) with partially precipitating pixels, the methodology described in Section 5.3.1.1 is followed. Figure 5-6 shows the frequency bias computed by comparing GPM Ku-PR with the MRMS for various values of the percentage of precipitating pixels (x). Note, for smaller values of x, the FAs dominate the MDs while the MDs dominate FA for larger values of x. The optimal FB (FB=1) is reached when x is 60%, that is, the ground truth is considered as raining if more than 15 pixels of the 25 MRMS pixels observe precipitation. The contingency matrix is shown in Table 1. More than 30% of the total precipitation detected by MRMS is missed by GPM Ku-PR. In addition, the number of MD cases exceed the number of FAs. Figure 5-7 shows the spatial distribution of POD and FAR. POD and FAR exhibit robust spatial structure with low POD and high FAR over the western ridges and inner regions. Precipitation in the western ridges is dominated by mesoscale convection precipitation systems while LLCF predominates in the inner SAM (Wilson and Barros, 2014, 2015 and 2017; Duan and Barros, 2017).
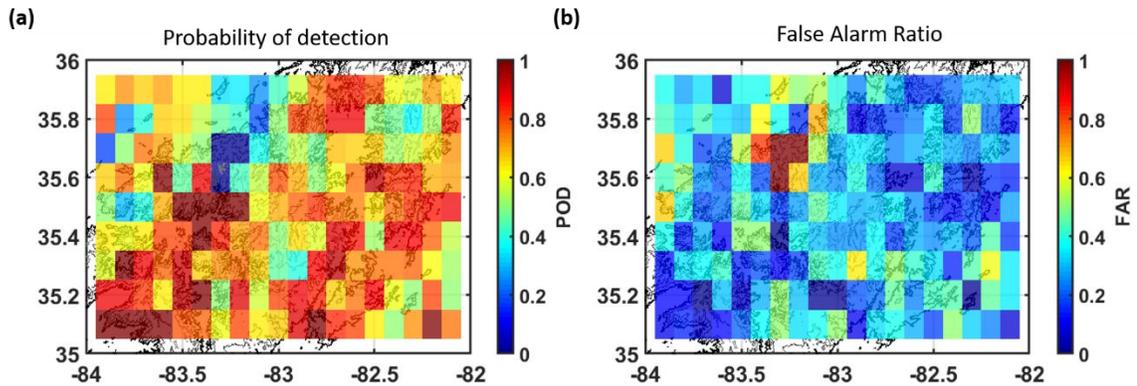
Figure 5-7 Spatial distribution of the detection metrics by comparing the GPM Ku-PR near-surface precipitation estimations with Level 2 MRMS precipitation rates. The black and gray contours denote DEM of the SAM.

The estimation bias is computed assuming the average and maximum MRMS precipitation rate as the ground-truth. Figure 5-8(b) shows the bias of GPM Ku-PR precipitation-rate with respect to the mean, maximum and nearest MRMS precipitation rate with the number of samples considered for the bias shown in Fig. 5-8(a). GPM Ku-PR overestimates light precipitation (less than 1 mm/h) while heavy precipitation is underestimated. These results are consistent with the results shown in Arulraj and Barros (2019) where GPM Ku-PR is compared against rain-gauge observations. Figure 5-8(d) shows the estimation bias categorized with respect to the precipitation type as classified by MRMS, and corresponding number of samples are shown in Fig. 5-8(c). Note that most of the events considered in the analysis are classified as stratiform precipitation by the MRMS. Underestimation of mean MRMS precipitation rate occurs across all the

191

precipitation types except that of snow, and the underestimation is severe in hail, tropical/convective mix and tropical/stratiform mix regions. The number of samples considered for analysis and the corresponding bias in spatial distribution is shown in Fig. 5-8 (e-f). There were no instances of correct detection of precipitation between [35.6 and 35.7 N] and [-83.4 and -83.3W], thus the bias discontinuity in this region. Even though GPM Ku-PR precipitation underestimates precipitation relative to MRMS in most of the SAM, precipitation in the southern and northernmost regions is overestimated. A note of caution is warranted here as the NEXRAD radar measurements contributing to MRMS along the eastern ridges are strongly affected by ground clutter, and therefore it is possible that MRMS is overcorrecting in this region. Given the lack of ground-truth over the National Forest, it is not possible to fully elucidate the sign of the errors in this case.
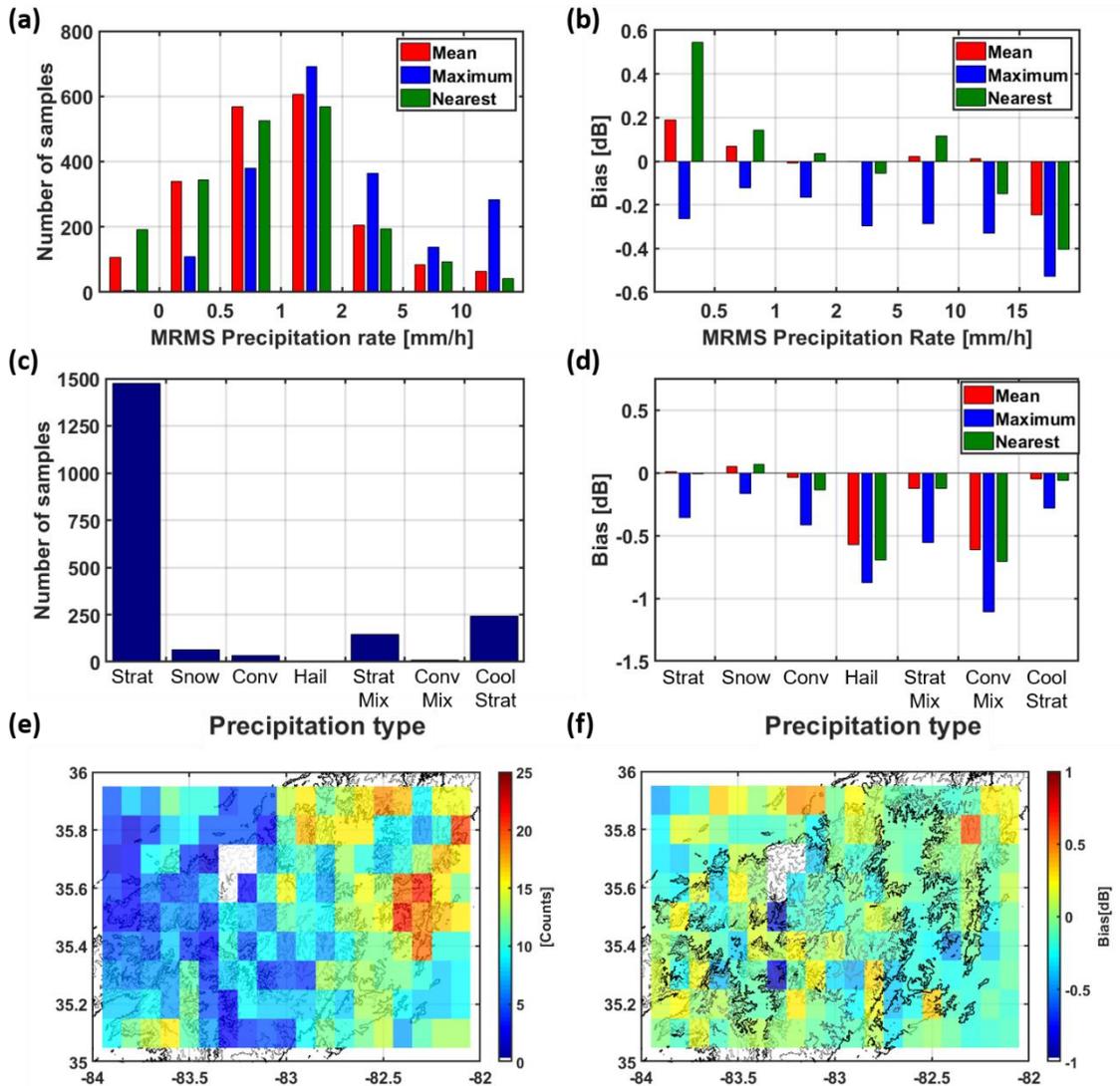
Figure 5-8 Histogram of (a, c and e) number of samples and (b, d and f) the corresponding bias in GPM Ku-PR with MRMS precipitation rate, type and spatial distribution. Note: 'Strat' denotes warm stratiform, 'Conv' denotes convective, 'Strat Mix' denotes tropical/stratiform mix, 'Conv Mix' denotes tropical/convective mix and 'Cool Strat' denotes cool stratiform type.

5.4.1.2 Clustering of MRMS Reflectivity profiles

The MRMS reflectivity profiles within the GPM Ku-PR overpass are used for this clustering analysis. Thus, for each GPM Ku-PR beam, (5x5) pixels of MRMS reflectivity profiles with precipitation are considered for this analysis which amounts to a total of 74724 reflectivity profiles. Profiles affected by beam blockage were removed from consideration. Figure 5-3 shows the first five features extracted from the MRMS reflectivity profiles. Second, the DB index is computed for different number of clusters in the k-means clustering algorithm. Figure 5-9 shows the variation of DB index for different number of clusters in k-means algorithm. The minimum value of DB index is obtained for four clusters, although there is some ambiguity in DB for values between 3 and 7 (representation of cluster-3 and 5 are shown in Appendix F). The contoured frequency by altitude diagrams (CFADs) of the reflectivity profiles for each of four clusters are in shown Figure 5-10. Table 5-2 shows the maximum, mean and standard deviation of the surface precipitation rate for the clusters.
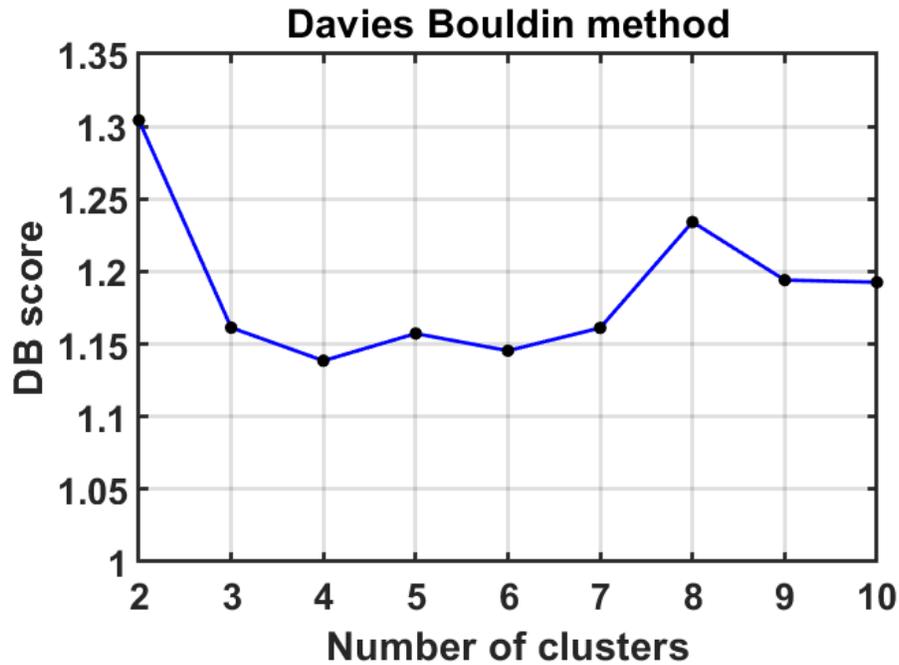
Figure 5-9 Davies-Bouldin (DB) index computed for different number of clusters of MRMS reflectivity features in the k-means algorithm.

Table 5-2: Mean, maximum and standard deviation of the MRMS precipitation rates corresponds to different clusters of reflectivity profiles.

	Maximum precipitation [mm/h]	Mean precipitation [mm/h]	Standard deviation [mm/h]
CLUSTER 1	18.89	0.86	0.77
CLUSTER 2	48.76	1.88	1.64
CLUSTER 3	74.74	1.62	2.27
CLUSTER 4	138.75	7.26	8.98

Cluster-1 reflectivity profiles represent light low-level precipitation systems (Figure 5-10a). Note that the echo top heights for the reflectivity profiles in this cluster are concentrated within the lower 4 km AGL and the reflectivity values are within 20 dBZ. The mean precipitation of this cluster is less than 1 mm/h while the maximum precipitation rate is 18.89 mm/h. Cluster-2 represents the shallow low-level enhanced precipitation systems (Figure 5-10b). These reflectivity profiles exhibit an increase in reflectivity values in the near the surface which is similar to enhanced seeder-feeder precipitation events as shown in Prat and Barros (2007), Wilson and Barros (2014), Duan and Barros (2017), Porcacchia et al. (2018), and Arulraj and Barros (2019). This cluster shows slightly higher mean and maximum precipitation rate compared to the cluster-1. The profiles in Cluster-3 show peak reflectivity around 2 to 4 km AGL with reflectivity decreases underneath the peak similar to bright band effects. The echo-top heights of these reflectivity profiles are around 6 to 8 km AGL. The mean precipitation rate is 1.62 mm/h and the maximum precipitation rate is 74.74 mm/h. This cluster has higher standard deviation compared to clusters 1 and 2. This cluster represents stratiform rainfall with shallow embedded convection at times. Cluster-4 corresponds to deep and heavy precipitation (convective) reflectivity profiles. The echo-top height is approximately around 8 km AGL and can extend up to 10 km. The near surface reflectivity value also varies between 25 and 40 dBZ consistent with heavy precipitation. This is also highlighted in the mean and maximum precipitation rate that is the highest among the 4 clusters.

196

Finally, the detection errors associated with each cluster are summarized in Table 5-3. For each pixel within an overpass of GPM Ku-PR, the maximum occurrence of cluster type (mode) within the nearest (5x5) MRMS pixels is considered as the ground-truth. GPM Ku-PR misses around 78% of all low-level light precipitation cases (cluster-1), while only 1.1% of the deep precipitation events (cluster-4) are missed. MD in cluster-2 and cluster-3 amounts to ~26% and ~15% respectively. 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 credence that the four clusters capture the principal precipitation regimes in the SAM. In almost all the clusters, underestimations dominate the overestimations by more than 57%. The least percentage of underestimation cases are seen in Cluster 4 (deep cases) while the highest underestimations are seen in Cluster-1 and Cluster-2. The next step is to develop the precipitation detection and classification algorithm relying solely on GPM and HRRR observations. This encompasses two steps: 1) Precipitation Detection; and 2) Precipitation System identification.
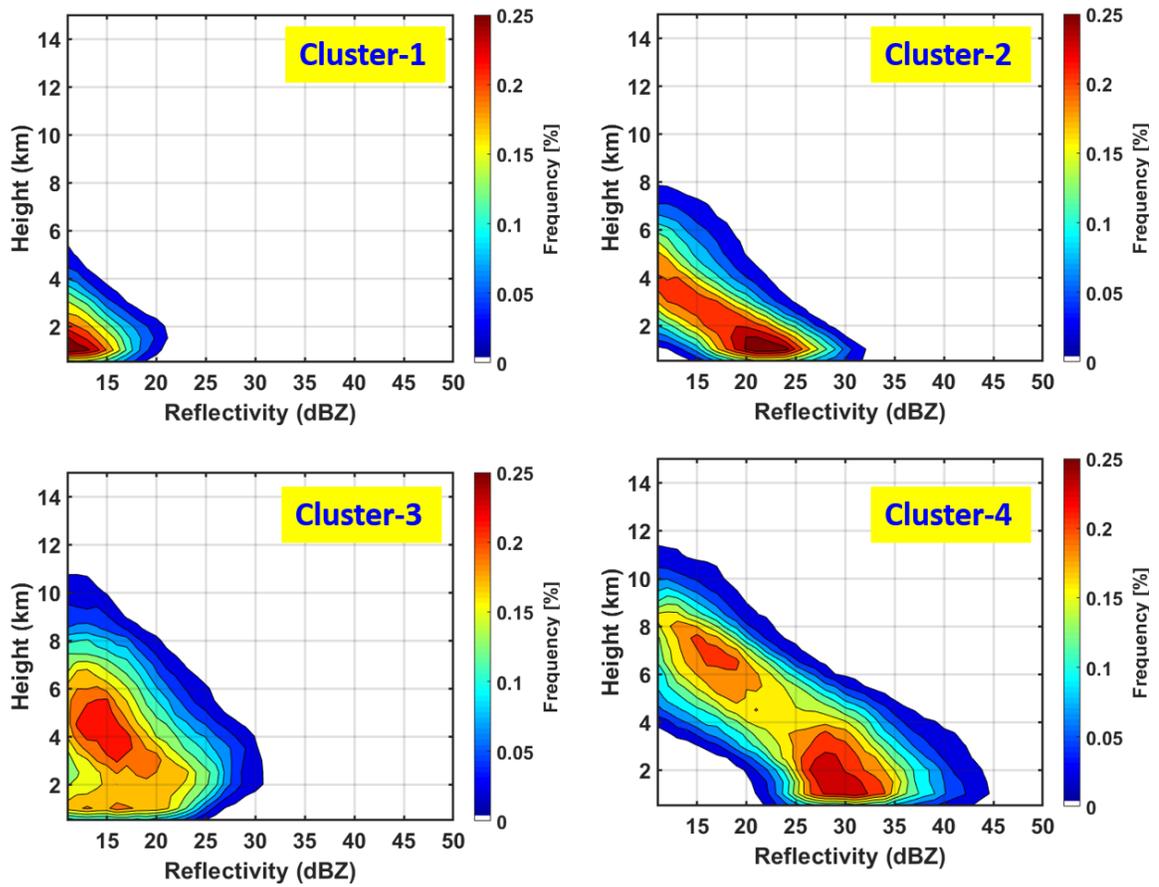
Figure 5-10 Contoured frequency altitude diagram (CFADs) of the MRMS reflectivity profiles clustered into different clusters using k-means clustering algorithm.

198

Table 5-3: Distribution of missed detection and correct detection of the reflectivity profiles in different clusters while comparing with GPM Ku-PR.

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Missed detections (MD) **GPM = 0 and MRMS ≠ 0**	621 (78.0 %)	184 (26.2%)	104 (15.8%)	8 (1.1%)
Correct detections (CD) **GPM ≠ 0 and MRMS ≠ 0**	175 (22.0%)	518 (73.8%)	554 (84.2%)	725 (98.9%)
Underestimations **GPM < MRMS**	121 (69% of CD)	354 (68% of CD)	348 (63% of CD)	414 (57% of CD)
Overestimations **GPM > MRMS**	54 (31% of CD)	164 (32% of CD)	206 (37% of CD)	311 (43% of CD)

5.4.2 Precipitation Detection

5.4.2.1 Statistical Analysis of HRRR

Figure 5-9 shows the spatial distribution of the number of HRRR pixels that observed precipitation at 500 m AGL over the 4-year period of interest. 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 Figure 1-2, even if amounts are largely underpredicted especially over the western ridges.
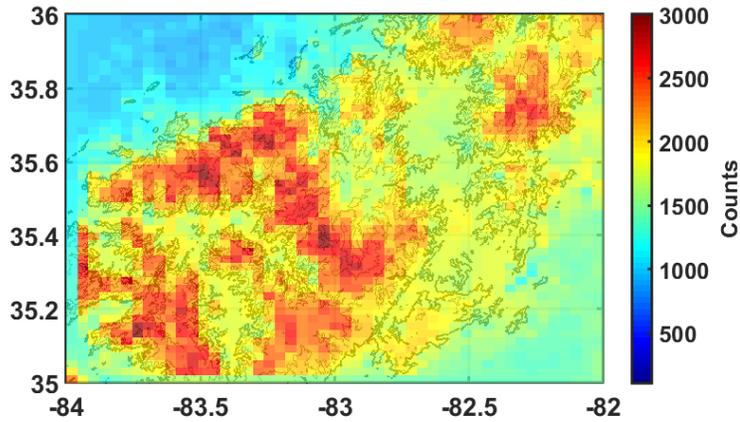
Figure 5-11 Spatial distribution map showing the frequency of precipitation instances observed in the HRRR model data.

Next, the HRRR model were 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 Figure 5-10. The maximum correlation is observed at 0-lag indicating that the HRRR does not exhibit significant timing errors (i.e. delay of precipitation onset), and thus HRRR captures well the diurnal cycle of rainfall. This is an important result in that no time corrections or adjustments need to be applied to the HRRR data relative to the GPM overpass.

Correlation coefficient vs Lag in RG rain-rate
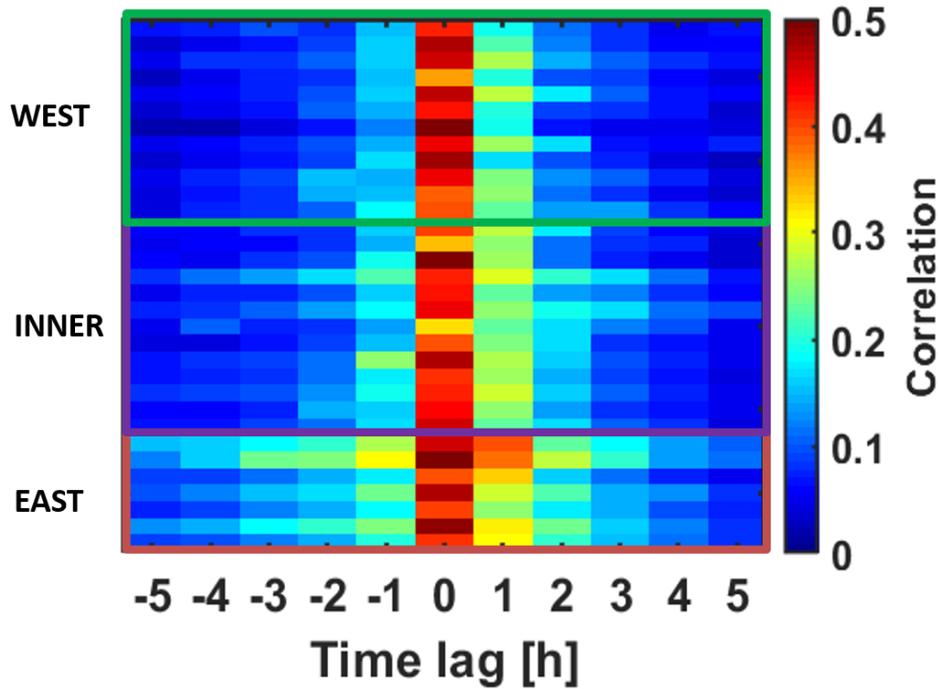
Figure 5-12 Pearson Correlation coefficient computed between the HRRR rain-rate at 500 m AGL and rain-gauge accumulations for lag in the time-scale of the rain-gauges

Next, the estimation bias observed in the HRRR is quantified against rain-gauge observations. Recall that the HRRR are instantaneous observations and the rain-gauge observations represent 30-minute accumulations centered at HRRR model time stamps. Figure 5-13 shows the density plot comparing the HRRR simulations against the rain-gauges. HRRR precipitation rates are much lower than the raingauge measurements especially for intense precipitation. This is also consistent with the previous error analysis conducted in the SAM region comparing the rain-gauge observations with different satellite products. Nevertheless, it is important to note that the uncertainties due to the

comparison of instantaneous areal estimates (HRRR) with time-average point measurements (rain-gauges) are not addressed fully.
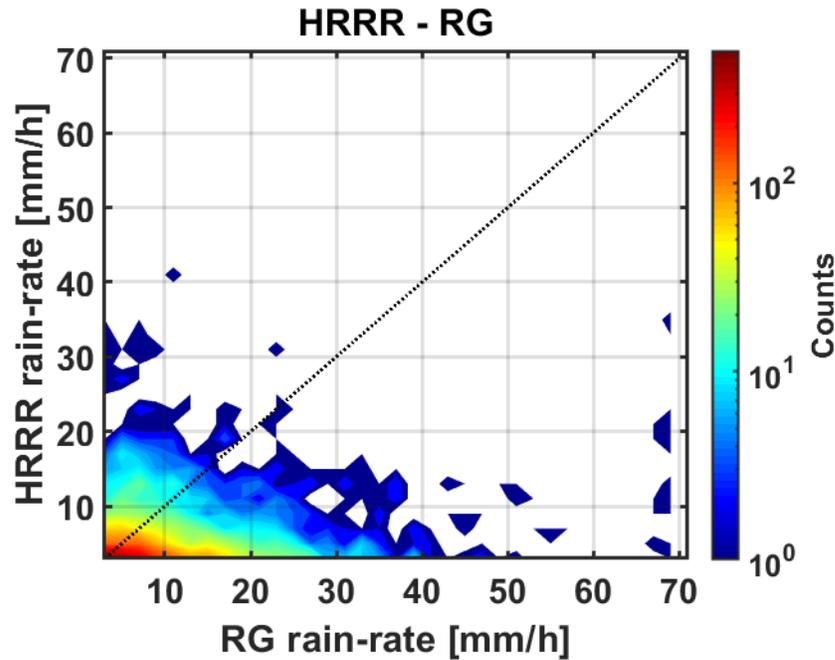
Figure 5-13 Density plot comparing the rain-rate derived from the rain-gauges with the HRRR simulated rain-rates at 500-m AGL.

Next, the diurnal climatology of HRRR precipitation rates is compared with rain-gauge observations in Fig. 5-12. 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 HRRR underestimates precipitation by 2-3 mm/h at all times, the overall structure of the diurnal cycle is well captured by the model in particular over the western ridges. This underestimation of rain-rates in the west is accompanied by underestimation of cloudiness, in particular shallow cap clouds such as those 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 seeder-feeder interactions among layered clouds.

These analyses suggest that the HRRR are a good representative of the climatological and diurnal behavior of the precipitation observed in the SAM even though rain-rates and cloudiness are underestimated. Thus, 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 water mass can be used as a robust indicator of rainfall activity. This is the basis for using HRRR RWMR, SWMR, and GRLE in the lower 1.5km (the depth of cluster-1 CFAD) as input to the precipitation detection model.
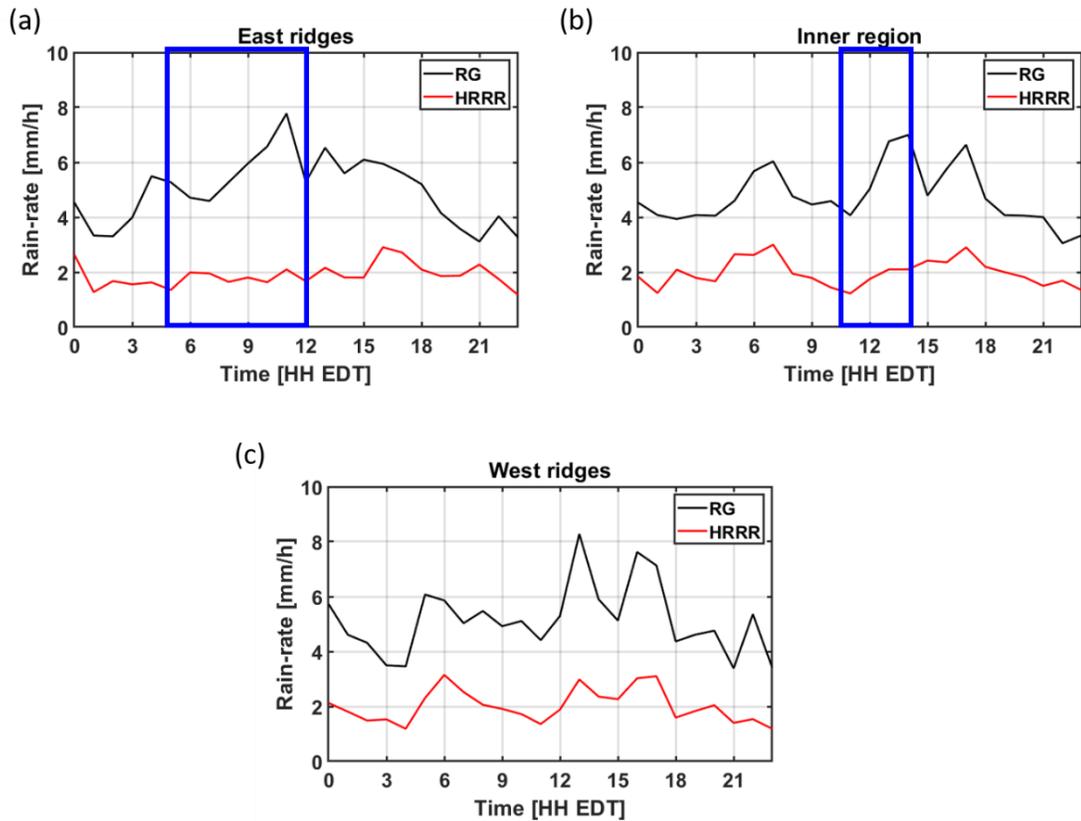
Figure 5-14 Diurnal cycle of precipitation rate observed by the HRRR compared with that of the rain-gauge observations. The blue boxes highlight the time when the diurnal cycle pattern of the HRRR and the rain-gauges differ.

5.4.2.2 Precipitation Detection Model

The precipitation detection model follows the methodology described in Section 5.3.3. The HRRR and GMI input data are normalized between 0 and 1 first. Then, the random forest classifier is applied to the normalized data sets for training. In order to avoid overfitting, a 5-fold cross validation is performed on the entire dataset. In a 5-fold cross validation, the dataset is divided into 5 subsets and one part is considered as

validation dataset while the remainder 4 parts are used to train the model in each iteration yielding a classification accuracy of approximately 0.96%.

The final model is then test on the entire dataset and the resulting contingency matrix is shown in Table 5-4. MDs are reduced by 84%, while FAs are reduced by 75%. This result is significant in that by itself reveals dramatic reduction of GPM Ku-PR detection errors. All the instances when precipitation was predicted can be mapped to one of the precipitation cluster classes derived from MRMS profiles. This further highlight the robustness of the precipitation detection algorithm that ties regional precipitation physics. The most important source of information value from GMI is the 89GHz channel for both VV and HH polarizations.

Table 5-4: Contingency matrix derived by comparing MRMS is compared with predictions from the framework

	MRMS = 0	MRMS ≠0
MODEL = 0	25030 (89.1%)	151 (0.5%)
MODEL ≠ 0	187 (0.7%)	2738 (9.7%)

5.4.3 Precipitation Classification Framework

Next, the classification of precipitation profiles is performed using a convolutional neural network (CNN) approach. Figure 5-5 depicts the CNN architecture used in this study is depicted in Figure 5-5. The input data consists of GPM Ku-PR and GMI observations and the target is to map GPM features to MRMS reflectivity cluster classes.

First, the model is developed (i.e. trained) using only GPM data for confirmed MRMS rainfall occurrences. After the model was developed, then it was implemented in series with the precipitation detection model for application in fully predictive model.

The GPM DPR and GMI data for the precipitating instances are divided into training, validation and test subsets with equal representation of data samples in each class. The input features are then submitted to min-max normalization. The information dimension of the GPM Ku-PR reflectivity profiles is compressed using auto-encoders with 10 features extracted. These features are provided as input to the first module of the CNN architecture. Here, a simple 1-layer network was chosen to avoid overfitting due to the small sample size of the training data. The number of filters in the convolution layer is four and a hyperbolic tangent activation function to improve sensitivity and the ability to capture nonlinear relationships. A random dropout of 50% of the information is enforced after each convolution layer with a maximum pooling layer of size 2 introduced after the first layer. The classification module consists of one fully-connected layer with 4 neurons 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.

Table 5-5: Contingency matrix to assess the performance of the data-driven framework to classify GPM estimations to clusters based on the shape of the MRMS reflectivity profiles.

GROND TRUTH	MODEL PREDICTION			
	Cluster-1	**Cluster-2**	**Cluster-3**	**Cluster-4**
Cluster-1	**572**	82	66	2
Cluster-2	169	**346**	140	25
Cluster-3	35	82	**406**	107
Cluster-4	2	40	112	**552**

The model is iteratively trained for 200 epochs and the training accuracy is 0.69. The validation accuracy is 0.68 which is close to the training accuracy and indicates that this CNN model is not overfit. Table 5-5 shows the contingency matrix of the classification algorithm using all the data except no-rain cases. Despite significant improvement in the classification, there is still significant ambiguity between Clusters 1 and 2 on the one hand, and Clusters-3 and 4 on the other corresponding to about 20-30% of rainfall instances.

5.5 Discussion

The performance of the precipitation detection and classification algorithm is illustrated here through a case-study. The precipitation event occurred on February 12, 2019 at 02:26 EDT. The precipitation rate and type classification by the MRMS are shown

in Figure 5-15. Note that this is largely a warm stratiform storm (blue color) with a cool

sector to the east (red color).
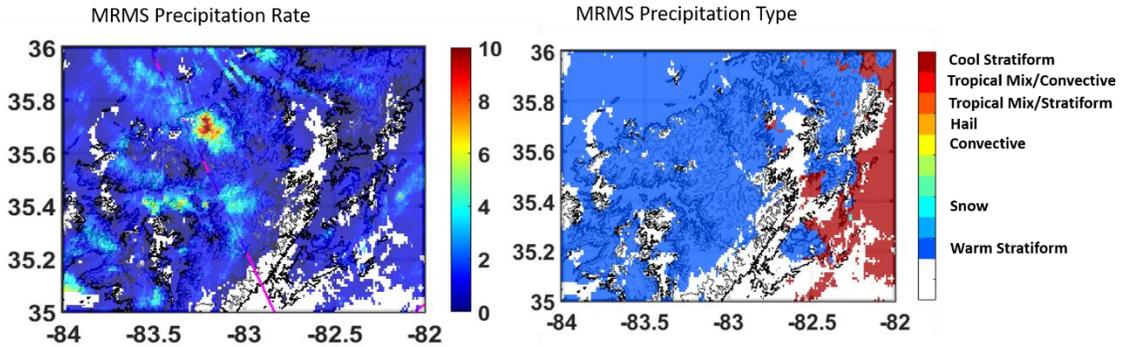
Figure 5-15 Precipitation rate and Precipitation type as estimated by MRMS precipitation product.

The GPM Ku-PR estimated near surface precipitation rate at the time of overpass

is shown in Figure 5-16. GPM detects the overall spatial structure of the stratiform

precipitation but misses the cool stratiform pixels and generally the edges of individual

precipitation systems.
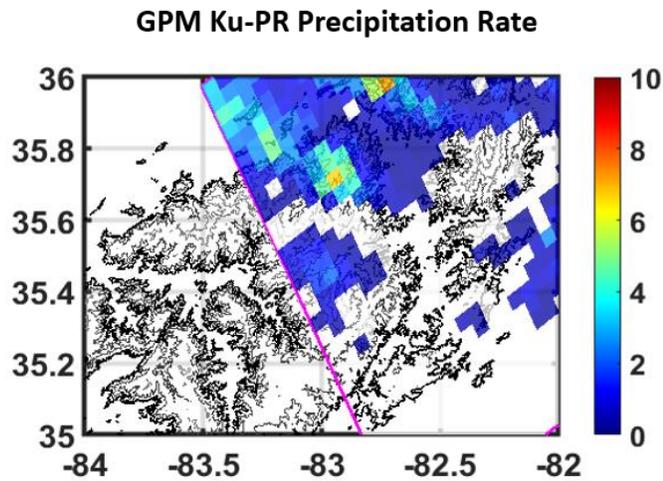
Figure 5-16 GPM Ku-PR estimated near-surface precipitation rate of February 12, 2019 event.

Figure 5-17(a) shows the spatial distribution of precipitation regimes according the MRMS cluster classes as the ground-truth. Class-0 corresponds to no-precipitation events. Classes 1-4 represent clusters 1-4 of the MRMS precipitation types. Figure 5-17(b) shows the corresponding map for the model predicted classification labels. Note that the edge of the precipitation systems in Fig. 5-17 (a) that classified as low- and light precipitation events are the missed detection events in Fig. 5-16. The model predictions capture the missed cool stratiform event and generally represent the overall spatial structure of the ground-truth more accurately. False alarm instances in this case-study are mostly classified as Cluster-1 precipitation systems (low-level light precipitation systems).
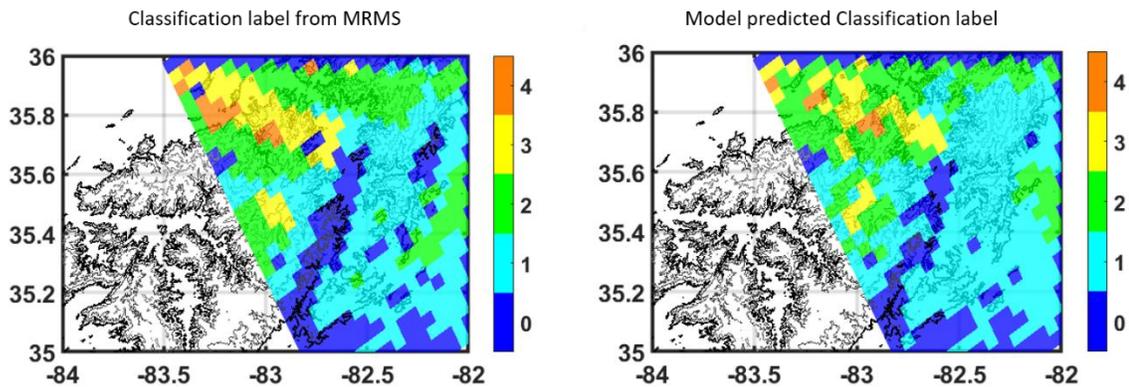
Figure 5-17 (a) Classification labels assigned to the GPM pixels by comparing Ku-PR estimations with MRMS pixels. (b) Model predicted classification labels for the GPM. Note, Class 0 represents no-precipitation, Class-1 corresponds to low-level precipitation systems, Class-2 are surface enhanced profiles, Class-3 are bright-band cases and Class 4 belongs to deep intense precipitation systems.

The classification framework identifies the near-surface structure of the reflectivity profile. Cluster 1 corresponds to light and shallow precipitation while Cluster-2 identifies

the collision-coalescence precipitation processes or the SFI cases. To determine the enhancement of the precipitation systems, tied to low-level cloud and fog microphysics can be determined using the physically-based framework proposed in Chapter 4. In order to test this hypothesis, the physically-based model is tested on an underestimation case from cluster-2. The case was observed on November 24, 2018 at 07:40 UTC (03:40 EDT) in the SAM. The GPM Ku-PR estimates a near-surface precipitation rate of 1.42 mm/h. However, the closest MRMS pixel to the GPM overpass bin observed a precipitation rate of 6.4 mm/h. The average MRMS precipitation rate with the 5 km field of view of GPM Ku-PR is 6.02 mm/h with a standard deviation of 3.24 mm/h. The reflectivity profiles of the nearest MRMS pixel and GPM Ku-PR overpass is shown in Fig. 5-18.
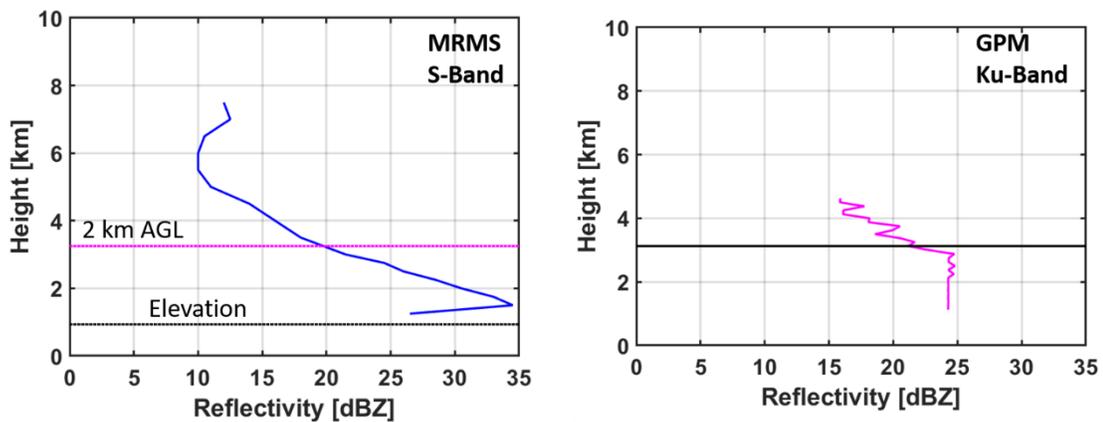
Figure 5-18 Reflectivity profiles of (a) MRMS and (b) GPM Ku-PR for November 24, 2018 precipitation case.

The top boundary condition (TBC; 2 km AGL) and initial condition of the rain microphysics model are derived from reflectivity profiles from GPM Ku-PR. A negative exponential distribution of DSD is derived from the GPM Ku-PR reflectivity values. The

microphysics of low-level cloud and fog is obtained from the diurnal climatology derived

from the MPS (Figure 4-3). The model is run for 80 minutes with the same top boundary

condition in demonstrate instantaneous overpass of GPM. The shallow SFI is introduced

between 20 and 50 minutes and is repeated for 300 m, 400 and 500 m fog depth. The

surface rain-rate from the physically-based framework are shown in Fig. 5-19. The surface

precipitation rate with S-SFI for different LLCF depth varies between 9 and 15 mm/h

which is higher than the average MRMS estimations. However, note that MRMS also

tends to severely underestimate precipitation rate near-surface in the SAM due to clutter

and beam blockage issues.
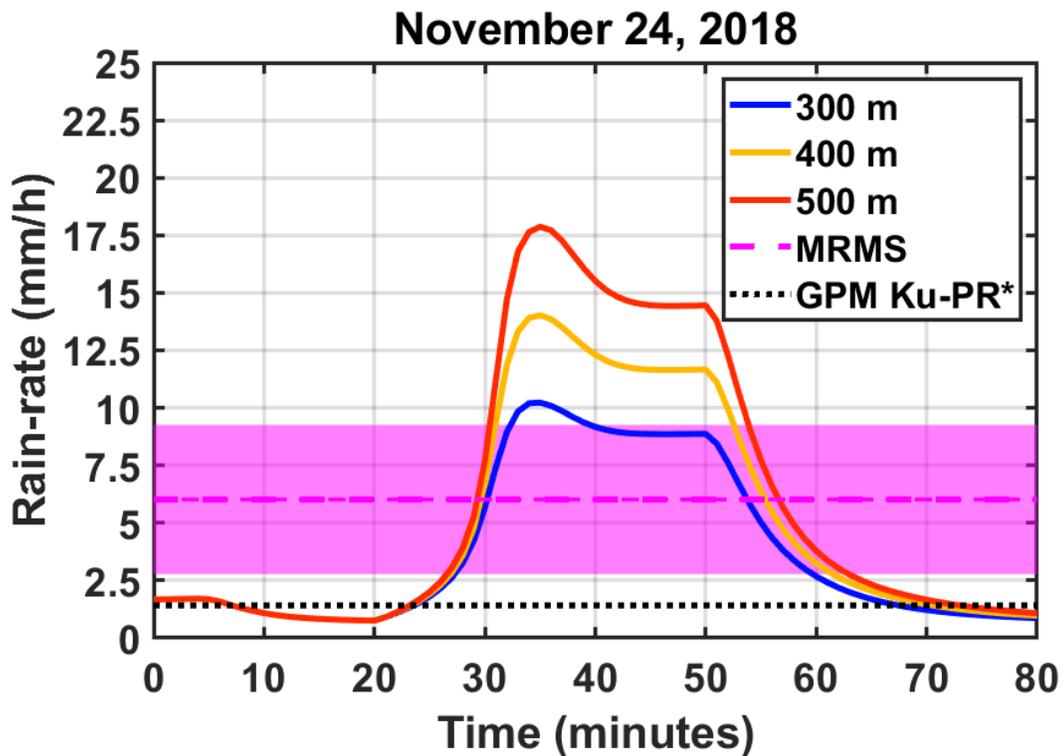
Figure 5-19 Surface rain-rate from the rain-microphysics model simulation for a shallow SFI scenario compared with MRMS precipitation rate.

5.6 Conclusion

A data-driven framework was developed to improve precipitation detection and classification of low-level precipitation structure by integrating passive (GMI) and active (Ku-PR) GPM observations with output of NWP simulations. The objective is to develop a simple operational framework to improve precipitation retrieval. The detection model is of general applicability globally. The classification model depends on the identification of regions of similar climatology and precipitation physics, which can be identified by clustering TRMM and, or GPM observations. Outside of CONUS, an adaptive model supported by ground validation would be necessary to identify precipitation classes in different climatic regions.

In the context of operational Ku-PR retrieval, the new classification framework can be used to improve detection skill and to replace of the uniform extrapolation of reflectivity above GC by a representative reflectivity profile based on MRMS precipitation structure classes over CONUS. In the context, of physically-based retrieval as presented Chapter 4 (Arulraj and Barros, 2019), Clusters 1-2 classification classes provide the necessary information to constrain the LLCF configuration. The major findings of this chapter 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 ridge 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. The HRRR does not show issues with the timing of precipitation onset as seen in other numerical model simulations. Nevertheless, HRRR precipitation rates are severely underestimated when compared with ground-based rain-gauges, and more so to the west. Given that the western ridges are in the direct path of moisture fluxes from the Gulf of Mexico and the Southern Great Plains, systematic underestimation raises the question of whether rainfall data used in the HRRR is introducing unrealistic constraints and indeed error. Because the spatial patterns and the diurnal cycle agree well with observations even if the magnitude is incorrect, the HRRR provides reliable information regarding precipitation activity that can be used in precipitation detection.

3. The reflectivity profiles from MRMS are clustered into 4 different classes based on the vertical structure. Cluster-1 represents shallow light precipitation, cluster-2 captures shallow low-level enhanced precipitation (e.g. LLCF), cluster-3 consists of profiles that exhibit bright-band effects, and cluster-4 captures deep convective (and heavy) precipitation systems. Error diagnostics indicate that GPM Ku-PR missed detection errors are tightly associated with

cluster-1 (shallow light precipitation) while the cluster-4 (deep convection) has the highest number of correct detection cases.

4. A precipitation detection algorithm was developed using a random forest classifier. The inputs of the precipitation detection algorithm include GPM GMI multichannel brightness temperatures, DPR reflectivity profiles, and HRRR. The most important features used in the random forest classifier to accurately detect precipitation are the calibrated brightness temperatures at 89 GHz for both vertical and horizontal polarization, followed by the average RWMR in the lowest 1.5 km AGL and the number of levels with non-zero RWMR in the nearest 1.5 km AGL.

5. The precipitation detection model improves the probability of detection and reduced the false alarm ratios. The number of missed detections is reduced by 84% while the number of false alarms is reduced by 74%.

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 type clusters. The cluster information derived can be used to estimate and classify precipitation events. This algorithm can also be applied in regions with precipitation climatology similar to the SAM in the absence of ground-observations.

This data-driven framework developed in this study shows great potential to improve the detection of light precipitation globally with potential to be generalized to different geographical locations.

6. Conclusions

This section will provide a summary of the dissertation (Section 6.1), outline major research findings (Sections 6.2), and discuss limitations in this work and suggestions for future research (Section 6.3).

6.1 Research Summary

The <u>overarching objective</u> of this dissertation is two-fold: 1) to characterize and elucidate the physical basis of uncertainty in Global Precipitation Measurement (GPM) – Dual-frequency Precipitation Radar (DPR) Quantitative Precipitation Estimates (QPE) in middle mountains; and 2) to develop an improved retrieval framework for orographic precipitation using GPM measurements with a focus on shallow precipitation systems (warm rain processes). The research is demonstrated in Southern Appalachian Mountains (SAM) due to the availability of a comprehensive data-base of long-term, spatially dense ground-based observations and modelling simulations developed in the context of IPHEx (Integrated Precipitation and Hydrology Experiment) the first ground-validation (GV) field-campaign after GPM launch. The research findings are representative of shallow precipitation processes mediated by layered low level clouds and fog (LLCF) generally such as the foothills and cloud forests of the world's largest cordilleras in the tropics and extra-tropics (e.g Western cordillera of North America, the Andes, the Western Ghats, and the Himalayas among others). This dissertation builds on previous studies of warm season precipitation processes in the SAM that illuminated the importance of the spatial

and temporal variability of seeder-feeder interactions (SFI) to capture low-level enhancement of precipitation (Prat and Barros 2010a, 2010b; Wilson and Barros 2014, 2015, and 2017; Duan and Barros 2017, and 2019; and Duan et al. 2015).

Duan and Barros (2017) showed that the classical features of upwind precipitation enhancement and lee wind rain shadow are seasonal in the SAM and are only present in the winter season. By contrast, warm season clouds and precipitation patterns exhibit strong diurnal cycles that reflect distinct moisture convergence pathways modulated by landform: from the east in the inner region ridge-valley system at low levels, and from the S-SW and W-SW at higher elevations (Wilson and Barros, 2017). More than 50% of the water-cycle in the SAM region is dominated by low-level light precipitation systems with intensity less than 3 mm/h (Wilson and Barros, 2014) with intermittent enhancement (~10x) explained by seeder-feeder interactions among layered clouds, cap clouds and fog.

Duan and Barros (2017) observed that the low-level clouds and fog are organized systematically with respect to the terrain, season and time of the day using long-term satellite-based observations. This spatial, temporal and seasonal organization of low-level clouds and fog systems is consistent with the space-time structure of TRMM PR precipitation retrieval errors, and in particular detection errors, documented by Duan et al. (2015) which remain for GPM DPR as shown in Chapter 2 and 4 of this dissertation (Barros and Arulraj, 2020; Arulraj and Barros, 2019).

A principal goal of this dissertation's research was to improve the detection and estimation of shallow orographic precipitation systems. This was accomplished by following two different strategies: a) improving detection and classification of orographic precipitation systems using remote-sensing observations; and b) improving estimation through explicit modeling of microphysical processes ("physics-based retrieval"). To improve detection and classification of shallow precipitation systems, a novel radar-only dual-frequency methodology was developed and demonstrated (Chapter 3; Arulraj and Barros, 2017). The feasibility and superior performance of physics-based retrieval of near-surface rain-rates associated with shallow precipitation systems enhanced by SFI was demonstrated through explicit modeling of rainshaft microphysical processes initialized and constrained by space-based radar observations and configured to describe the vertical structure of LLCF (Chapter 4, Arulraj and Barros, 2019). Finally, a data driven model relying on NWP and GPM measurements from both GMI and DPR was developed to predict the vertical structure of shallow precipitation systems tied to LLCF, which also dramatically improved detection skill (Arulraj and Barros, 2020, in preparation).

6.2 Major Research Findings

The science questions (SQs) are revisited below along with the major findings in this dissertation.

SQ-1: What is the detection and estimation uncertainty associated with the satellite-based precipitation estimates in the SAM? How can the observations from multi-frequency radar improve the detection and classification of shallow precipitation?

- The error diagnostics performed on the GPM Ku-PR suggests that the missed detections are attributed to light and low-level precipitation while more than 82% of the false alarms are light stratiform precipitation.

- A robust space-time structure of detection errors that are linked to the low-level clouds and fog is observed in the SAM. These regions also observed low D_m and high N_w suggesting the presence of SFI.

- Based on the error diagnostics and physical interpretation, the retrieval errors are attributed to ground-clutter contamination of near-surface reflectivity values, errors in the estimation of DSD parameters, and subgrid-scale spatial heterogeneity of precipitation systems leading to NUBF artifacts.

- A shallow precipitation detection and classification algorithm was developed based on the dual-frequency space-time correlation of collocated W- and Ka- Band. Proof of concept analyses were performed on ground-based observations, model simulations and satellite-based observations (GPM Ka-PR and CloudSat CPR). The results from this algorithm suggest increased detection in shallow precipitation.

SQ-2: How does the surface clutter influence the near-surface reflectivity profiles for different hydrometeorological regimes in satellite-based observations and Numerical Weather Prediction Models??

- Error diagnostics of satellite-based observations suggest that the ground-clutter contamination leads to a severe underestimation of low-level enhanced coalescence-collision precipitation systems. In cases where the ground-clutter height is close to the melting layer, satellite-based retrievals tend to overestimate precipitation rate near the surface due to the effect of bright-band. In addition, most of the low-level precipitation systems are missed due to contamination of near-surface clutter. In GPM Ku-PR, ground-clutter contaminates reflectivity profiles from 500 m up to 3500 m above the ground level.

- The contamination of near-surface reflectivity profiles leads to high detection and estimation errors in the SAM where the water cycle is dominated by shallow light precipitation system and low-level enhanced collision coalescence precipitation events.

- A physically-based framework was developed using a stochastic rainshaft model with prescribed low-level cloud and fog microphysics ground-observations to predict the near-surface precipitation and DSD from GPM DPR measurements. This framework successfully simulated the precipitation-rate and DSD profiles in the SFI regime.

220

SQ-3: How does the integration of observations and simulations influence the accuracy of QPE?

- A general data-driven error diagnostics and classification framework was developed to improve precipitation detection and estimate the structure of near-surface reflectivity profiles using GPM observations and HRRR.

- This framework reduces the number of missed detection by 84% and false alarm cases by 75% overall improving the detection of the spatial structure of the precipitation events.

- The detected precipitation events are further classified based on the expected vertical structure of ground-based MRMS profiles. This information can assist in the existing operational precipitation detection and classification algorithm.

6.3 Limitations and Suggestions for Future Research

The limitations of this study stem largely from the relatively small number of precipitation measurements in the GPM era. The number of GPM overpasses during active precipitation is biased with respect to the time of day and that is apparent from the error analysis shown in Chapter 2 and Chapter 4. Further, the data-driven framework described in Chapter-5 is trained based on a record consisting of 3000 reflectivity profiles from 69 rainy overpasses among over 700 since 2014. Although 10-years of IPHEx related observations including several months of distributed microphysical observations of

rainfall provide confidence with regard to the physics-based and data-driven models, there is significant interannual variability that cannot be accounted for in this research.

Previous work (Wilson and Barros, 2014) relied on the literature to specify LLCF spectra (i.e. DSDs for drop sizes < 100 μ) for SFI simulations. In this dissertation, data collected at Cosby in the western valley region of the SAM are used. However, these measurements conducted over a 2-month period after the IPHEx field campaign do not reflect seasonal and interannual variability in weather systems and aerosol sources, which in turn may impact the diurnal cycle of LLCF DSDs that are critical to model SFI at local scales.

The DSD analysis performed in Chapter 4 rely on the results from GPM Ku-PR single-frequency (SF) algorithm. The SF algorithm was preferred because of the large sample size compared to the dual-frequency (DF) algorithm that was limited to near-nadir pixels only at the time. However, the DSDs from DF algorithm are more reliable because of the additional information that can be extracted from the relationship between the dual-frequency ratio (DFR) and D_m. Figure E14 (Appendix E) shows the relationship between the D_m and N_w estimated by SF and DF algorithms. The DF D_m estimates are bounded at 3 mm and slightly overestimated the observations. N_w estimated by DF and SF are not correlated and the range of DF N_w estimates is wider consistent with observations. This is further confirmed in the D_m-N_w phase diagram shown in Fig. E15. Note that the small values of D_m (high values of N_w) are still missing in the GPM DF estimations. The distinct

222

line seen in D_m-N_w phase diagram for rain-rate between 3 and 7 mm/h was evaluated also using the DF-based DSD estimations and is shown in Fig. E16 in Appendix E. Interestingly, the agreement between the D_m-N_w maps for SF and Parsivel observations is excellent for 3-7mm/hr rain rates, whereas the map is substantially different. This is an interesting and intriguing finding that will be explored via model simulations. This research can also be further extended to analyze the effect of Path Integrated Attenuation (PIA) estimates and attenuation correction in the estimation of DSD parameters.

While the spatial resolution of HRRR simulations (~3 km) and GPM Ku-PR (~ 5 km) are relatively close, the HRRR data are available at hourly intervals only. Thus, the life-cycle of developing and fast evolving precipitation systems cannot be captured. Nevertheless, it is important to point out that questions regarding severe bias of clouds and precipitation in the HRRR likely result from bias in the assimilation of biased observations in the SAM, a problem that cannot be resolved by higher temporal resolution.

Critical next steps for this research include testing and evaluation of the detection and evaluation models, and in particular the data-driven framework proposed in Chapter 5, in different geographical regions. In addition, extending the data-base of precipitation structure classes to CONUS and elsewhere where radar profiles are available would be valuable in itself, as such the data analysis and the data-driven framework are not dependent on precipitation regime and therefore their applicability is general beyond

223

mountainous regions. Furthermore, the precipitation type information derived from the

classification algorithm can be further extended to estimate precipitation rate and DSD

parameters using Bayesian or deep learning algorithms and to quantify uncertainty.

Consequently, the use of HRRR was limited to low-level water mixing ratios and

cloud depth as inputs to the precipitation detection algorithm. However, given reliable

NWP climatology, hydrometeor mixing ratio data can be easily used to derive reflectivity

profiles using existing simulators such as the fast radar model by Hogan and Battaglia

(2008) with optical parameters from Leinonen et al. (2015) and thus generate a data-base

of precipitation structures where radar measurements such as MRMS are not available

outside of CONUS. A global data-base generated from high-resolution reanalysis from

ERA5, albeit imperfect, could make a significant contribution to help reduce detection

errors and improving GPM Ku-PR rain-rate estimates by leveraging global GV sites for

various precipitation regimes.

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.

Angulo-Martinez, M., and A. P. Barros. 2015. "Measurement uncertainty in rainfall kinetic energy and intensity relationships for soil erosion studies: an evaluation using ." *Geomorphology* 28-40.

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.

Awaka, J., T. Iguchi, H. Kumagai, and K. Okamoto. 1997. "Rain type classification algorithm for TRMM precipitation radar." *IGARSS 1997.* IEEE. 317-319. doi:10.1109/IGARSS.1997.608993.

Bales, R. C., N. P. Molotch, T. H. Painter, M. D. Dettinger, and R. Dozier, J. Rice. 2006. "Mountain hydrology of the western United States." *Water Resour. Res.* 42: W08432. doi:10.1029/2005WR004387.

Balestrino, A., A. Caiti, and E. Crisostomi. 2007. "A classification of nonlinear systems: An entropy based approach." *Chem. Eng. Trans.* 11: 119-124.

Barros, A. P. 2013. "Orographic precipitation, freshwater resources, and climate vulnerabilities in mountainous." In *Climate Vulnerability: Understanding and*

Addressing Threats to Essential Resources, 57-78. Amsterdam, The Netherlands: Elsevier.

Barros, A. P., and D. P. Lettenmaier. 1994. "Dynamic Modeling of orographically induced precipitation." *Rev. Geophys.* 32 (3): 265-284. doi:10.1029/94RG00625 .

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. Springer International Publishing.

Barros, A. P., and T. J. Lang. 2003. "Monitoring the Monsoon in the Himalayas: Observations in Central Nepal, June 2001." *Mon. Wea. Rev.* 131: 1408-1427. doi:10.1175/1520-0493(2003)131<1408:MTMITH>2.0.CO;2.

Barros, A. P., G. Kim, E. Williams, and S. W. Nesbitt. 2004. "Probing orographic controls in the Himalayas during the monsoon using satellite imagery." *Nat. Hazards Earth Syst. Sci.* 4: 29-51. doi:10.5194/nhess-4-29-2004.

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., O. P. Prat, P. Shrestha, F. Y. Testik, and L. F. Bliven. 2008. "Revisiting Low and List (1982): Evaluation of Raindrop Collision Parameterizations using Laboratory Observations and Modeling." *J. Atmos. Sci.* 2983-2993. doi:10.1175/2008JAS2630.1.

Barros, A. P., P. Shrestha, S. Chavez, and Y. Duan. 2018. "Modeling aerosol-cloud-precipitation interactions in mountain regions - Challenges in the representation of indirect microphysical effects with impacts at sub-regional scales." In *Atmospheric Chemistry*. IntechBookOpen. doi:10.5772/intechopen.80025.

Barros, A. P., W. Petersen, M. Schwaller, R. Cifelli, K. Mahoney, C. Peters-Liddard, M. Shepherd, et al. 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.

Battaglia, A., M. O. Ajewole, and C. Simmer. 2005. "Multiple scattering effects due to hydrometeors on precipitation radar systems." *Geophys. Res. Letters.*

Battaglia, A., S. Tanelli, G. M. Heymsfield, and L. Tian. 2014. "The Dual Wavelength Patio Knee: A Signature of Multiple Scattering in Airborne Ku-Ka Observations." *J. Appl. Meteor. Climatol.* 1790-1808.

Battaglia, A., S. Tanelli, K. Mroz, and F. Tridon. 2015. "Multiple scattering in observations of the GPM dual-frequency precipitation radar: Evidence and impact on retirevals." *J. Geophys. Res.*

Battaglia, A., S. Tanelli, S. Kobayashi, D. Zrnic, R. Hogan, and C. Simmer. 2010. "Multiple-scattering in radar systems: A review." *J. Quant. Spectrosc. Radiat. Transfer* 111: 917-947. doi:10.1016/j.jqsrt.2009.11.024.

Baumgardner, D., G. Kok, W. Dawson, D. O'Connor, and R. Newton. 2002. "A new groundbased precipitation spectrometer: The meteorological particle sensor (MPS)." *11th Conference on Cloud Physics.* Ogden, UT: Amer. Meteor. Soc. 3-7. https://ams.confex.com/ams/pdfpapers/41834.pdf.

Belgium, 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.

Berg, W., T. L'Ecuyer, and C. Kummerow. 2006. "Rainfall climate regimes: The relationship of regional TRMM rainfall biases to the environment." *J. Appl. Meteor. Climatol.* 45 (3): 434–454. doi:10.1175/JAM2331.1.

Bergeron, T. 1960. "Operation and results of "Project Pluvius"." Edited by H. Weichmann. *Physics of Precipitation, Geophys. Monogr. Ser.* Washington D. C: AGU. 152-157.

—. 1960. "Problems and methods of rainfall investigation, address of honorary chairman of the conference." Edited by H. Weickmann. *Physics of Precipitation, Geophys. Monogr. Ser.* Washington D. C.: AGU. 5-30.

Biggerstaff, M. I., and S. A. Listemaa. 2000. "An improved scheme for convective/stratiform echo classification using radar reflectivity." *J. Appl. Meteor.* 39 (12): 2129–2150. doi:10.1175/1520-0450(2001)040h2129:AISFCSi2.0.CO;2.

Bindlish, R., and A. P. Barros. 2000. "Disaggregation of rainfall for one-way coupling of atmospheric and hydrological models in regions of complex terrain." *Global Planet. Change* 25: 111-132. doi:10.1016/S0921-8181(00)00024-2.

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.

Bohren, C. F., and D. R. Huffman. 1983. *Absorption and Scattering of Light by Small Particles.* Wiley. doi:10.1002/9783527618156.

Bouniol, D., A. Protat, A. Plana-Fattori, M. Giraud, J. Vinson, and N. Grand. 25. "Comparison of Airborne and Spaceborne 95-GHz Radar Reflectivities and Evaluation of Multiple Scattering Effects in Spaceborne Measurements." *J. Atmos. Oceanic Technol.* 1983–1995. doi:10.1175/2008JTECHA1011.1.

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

Breiman, L., J. Friedman, R. Olshen, and C Stone. 1984. *Classification and Regression Trees.* New York: Chapman and Hall, Wadsworth.

Breiman, L., J. H. Friedman, R. A. Olshen, and C., I. Stone. 1984. *Classification and Regression trees.* Belmont, California: Wadsworth.

Bringi, V., and V. Chandrasekar. 2001. *Polarimetric Doppler Weather Radar: Principles and Applications.* Cambridge: Cambridge University Press. doi:10.1017/CBO9780511541094.

Bruijnzeel, L. A. 2001. "Hydrology of tropical montane cloud forests: a reassessment." *Land Use and Water Resources Research* 1: 1.1-1.18. doi:10.22004/ag.econ.47849.

Bruijnzeel, L. A., M. Mulligan, and F. N. Scatena. 2011. "Hydrometeorology of tropical montane cloud forests: emerging patterns." *Hydrol. Process.* 25: 465-498. doi:10.1002/hyp.7974.

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.

Burbank, D. W., A. E. Blythe, J. Putkonen, B. Pratt-Sitaula, E. Gabet, M. Oskin, A. P. Barros, and T. P. Ojha. 2003. "Decoupling of erosion and precipitation in the Himalayas." *Nature* 426: 652-655. doi:10.1038/nature02187.

Caine, N. 1980. *The Rainfall Intensity: Duration Control of Shallow Landslides and Debris Flows.* Vol. 62, in *Geografiska Annaler*, 23-27. doi:10.2307/520449.

Clothiaux, E. E., T. P. Ackerman, G. G. Mace, K. P. Moran, R. T. Marchand, M. A. Miller, and B. E. Martner. 2000. "Objective Determination of Cloud Heights and Radar Reflectivities Using a Combination of Active Remote Sensors at the ARM CART sites." *J. Appl. Meteor.* 39 (5): 645-665. doi:10.1175/1520-0450(2000)039<0645:ODOCHA>2.0.CO;2.

D'Adderio, L.P., F. Porcu, and A. Tokay. 2015. "Identification and analysis of collisionalbreakup in natural rain." *J. Atmos. Sci.* 3404–3416. doi:10.1175/JAS-D-14-0304.1.

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.

Denis, A., and F. Cremoux. 2002. "Using the entropy of curves to segment a time or spatial series." *Math. Geol.* 899–914. doi:10.1023/A:1021302922108.

Dolan, B., B. Fuchs, S.A. Rutledge, E.A. Barnes, and E.J. Thompson. 2018. "Primary modes ofglobal drop size distributions." *J. Atmos. Sci.* 75: 1453–1476. doi:10.1175/JAS-D-17-0242.1.

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.

Duan, Y., M. D. Petters, and A. P. Barros. 2019. "Understanding aerosol-cloud interactions through modelling the development of orographic cumulus congestus during IPHEx." *Atmos. Chem. Phys.* 19: 1-25. doi:10.5194/acp-19-1-2019.

Durden, S.L., Z. S. Haddad, A. Kitiyakara, and F. K. Li. 1998. "Effects of Nonuniform Beam Filling on Rainfall Retrieval for the TRMM Precipitation Radar." *J. Atmos. Oceanic Technol.* 635–646. doi:10.1175/1520-0426(1998)015<0635:EONBFO>2.0.CO;2.

Eghdami, M., and A. P. Barros. 2019. "Extreme Orographic Rainfall in the Eastern Andes Tied to Cold Air Intrusions." *Front. Environ. Sci.* 7: 101. doi:10.3389/fenvs.2019.00101.

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.

Falkenmark, M., and J. Rockström. 2010. "Building Water Resilience in the Face of Global Change: From a Blue-Only to a Green-Blue Water Approach to Land-Water Management." *Journal of Water Resources Planning and Management* 136: 606-610. doi:10.1061/(ASCE)WR.1943-5452.0000118.

FAO. 2016. *Climate Change and food security: risks and responses.* Rome: Food and Agriculture Organization of the United Nations. www.fao.org/3/a-i5188e.pdf.

FAO. 2014. *The State of Food and Agriculture. Innovation in family farming.* Rome: Food and Agriculture Organization of the United Nations. www.fao.org/3/a-i4040e.pdf.

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.

Feng, Z., S. A. McFarlane, C. Schumacher, S. Ellis, J. Comstock, and N. Bharadwaj. 2014. "Constructing amerged cloud–precipitation radar dataset for tropical convective clouds during the DYNAMO/AMIE experiment at AdduAtoll." *J.Atmos.Oceanic Technol.* 31: 1021–1042. doi:10.1175/JTECH-D-13-00132.1.

Garrido, A., and A. Dinar. 2009. "Overcoming the constraints for a more integrated and adaptive water management: a new agenda for Upland Waters." In *Managing Water Resources in a Time of Global Change: Mountains, Valleys and Flood Plains*, 253-263. New York: Taylor & Francis Group.

Geertsema, M., O. Hungr, J. W. Schwab, and S. G. Evans. 2006. "A large rockslide–debris avalanche in cohesive soil at Pink Mountain, northeastern British Columbia, Canada." *Engineering Geology* 83 (1-3): 64-75. doi:10.1016/j.enggeo.2005.06.025.

Germann, U., G. Galli, M. Boscacci, and M. Bolliger. 2006. "Radar precipitation measurement in a mountainous region." *Q.J.R. Meteorol. Soc.* 132: 1669-1692. doi:10.1256/qj.05.190.

Giovannettone, J. P., and A. P. Barros. 2009. "Probing Regional Orographic Controls of Precipitation and Cloudiness in the Central Andes Using Satellite Data." *J. Hydrometeorol.* 10 (1): 167-182. doi:10.1175/2008JHM973.1.

Gourley, J. J., and Y. Hong. 2014. *Radar hydrology: principles, models, and applications.* CRC Press.

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.

Hill, F., K. A. Browning, and M. J. Bader. 1981. "Radar and raingauge observations of orographic rain over south Wales." *Quart. J. Roy. Meteor. Soc.* 107: 643-670. doi:10.1002/qj.49710745312.

Hong, Y., C. D. Kummerow, and W. S. Olson. 1999. "Separation of convective and stratiform precipitation using microwave brightness temperature." *J. Appl. Meteor.* 38 (8): 1195–1213. doi:10.1175/1520-0450(1999)038h1195:SOCASPi2.0.CO;2.

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.

Houze, R. A. 1993. *Climate Dynamics.* Vol. 53. Academic Press.

Houze, R. A., Jr. 2012. "Orographic effects of precipitating clouds." *Rev. Geophys.* 50: RG1001. doi:10.1029/2011RG000365.

Huffman, G. J., D. T. Bolvin, and E. J. Nelkin. 2015. *Integrated Multi-satellite retrievals for GPM (IMERG) technical documentation.* NASA/GSFC, 47.

Huffman, G. J., D. T. Bolvin, E. J. Nelkin, D. B. Wolff, R. F. Adler, G. Gu, Y. Hong, K. P. Bowman, and E. F. Stocker. 2007. "The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales." *J. Hydrometeorol.* 8: 38-55. doi:10.1175/JHM560.1.

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.

Iguchi, T., T. Kozu, R. Meneghini, J. Awaka, and K. Okamoto. 2000. "Rain-Profiling Algorithm for the TRMM Precipitation Radar." *J. Appl. Meteor.* 39: 2038-2052. doi:10.1175/1520-0450(2001)040<2038:RPAFTT>2.0.CO;2.

Illingworth, A. J., and T. M. Blackman. 2002. "The Need to Represent Raindrop Size Spectra as Normalized Gamma Distributions for the Interpretation of Polarization Radar Observations." *J. Appl. Meteor.* 41: 286-297. doi:10.1175/1520-0450(2002)041<0286:TNTRRS>2.0.CO;2.

Immerzeel, W. W., L. P. H. van Beek, and M. F. P. Bierkens. 2010. "Climate Change Will Affect the Asian Water Towers." *Science* 328 (5984): 1382-1385. doi:10.1126/science.1183188.

Islam, T., M.A. Rico-Ramirez, D. Han, P.K. Srivastava, and A.M. Ishak. 2012. "Performanceevaluation of the TRMM precipitation estimation using ground-based radars from theGPM validation network." *J. Atmos. Sol. Terr. Phys.* 77: 194-208. doi:10.1016/j.jastp.2012.01.001.

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." *Proceedings of the 3rd International Conference on Learning Representations (ICLR).*

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.

Kohavi, R. 1995. "A study of cross-validation and bootstrap for accuracy estimation and model selection." *IJCAI.* 1137-1145.

Kollias, P., and B. Albrecht. 2005. "Why the melting layer radar reflectivity is not bright at 94 GHz." *Geophys. Res. Lett.* 32: L24818. doi:10.1029/2005GL024074.

Kollias, P., E. E. Clothiaux, M. A. Miller, B. A. Albrecht, G. L. Stephens, and T. P. Ackerman. 2007. "Millimeter-wavelength radars:Newfrontier in atmospheric cloud and precipitation research." *Bull. Amer. Meteor. Soc.* 88: 1608-1624. doi:10.1175/BAMS-88-10-1608.

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.

Le, M., and V. Chandrasekar. 2013. "Precipitation type classification method for dual-frequency precipitation radar (DPR) on board the GPM." *IEEE Trans. Geosci. Remote Sens.* 51 (3): 1784–1790. doi:10.1109/TGRS.2012.2205698.

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, L., R. Meneghini, and A. Tokay. 2014. "Uncertainties of GPM DPR Rain Estimates Caused by DSD Parameterizations." *J. Appl. Meteor. Climatol.* 53: 2524-2537. doi:10.1175/JAMC-D-14-0003.1.

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.

Loffler-Mang, M., and J. Joss. 2000. "An optical disdrometer for measuring size and velocity of hydrometeors." *J. Atmos. Ocean. Technol.* 17 (2): 130-139. doi:10.1175/1520-0426(2000)017h0130:AODFMSi2.0.CO;2.

Lowman, L. E. L., and A. P. Barros. 2016. "Interplay of Drought and Tropical Cyclone Activity in SE US Gross Primary Productivity." *J. Geophys. Res. Biogeosci.* 120: 1540–1567. doi:10.1002/2015JG003279.

Lowman, L. E. L., and A. P. Barros. 2014. "Investigating links between climate and orography in the central Andes: Coupling erosion and precipitation using a physical-statistical model." *J. Geophys. Res. Earth Surf.* 119: 1322– 1353. doi:10.1002/2013JF002940.

Lowman, L. E. L., T. M. Wei, and A. P. Barros. 2018. "Rainfall Variability, Wetland Persistence, and Water-Carbon Cycle Coupling in the Upper Zambezi River Basin in Southern Africa." *Remote Sens.* 10 (5): 692. doi:10.3390/rs10050692.

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.M. Palmer. 1948. "The distribution of raindrops with size." *J. Meteorol.* 5: 165-166. doi:10.1175/1520-0469(1948)005<0165:TDORWS>2.0.CO;2.

Marzano, F. S., L. Roberti, S. D. Michele, A. Mugnai, and A. Tassa. 2003. "Modeling of apparent radar reflectivity due to convective clouds at attenuating wavelengths." *Radio Sci.* 38 (1): 1002. doi:10.1029/2002RS002613.

Matrosov, S. Y., A. J. Heymsfield, and Z. Wang. 2005. "Dual-frequency radar ratio of nonspherical atmospheric hydrometeors." *Geophys. Res. Lett.* 32 (13). doi:10.1029/2005GL023210.

Matsui, T., T. Iguchi, X. Li, M. Han, W. Tao, W. Petersen, T. L'Ecuyer, et al. 2013. "GPM Satellite Simulator over Ground Validation Sites." *Bull. Amer. Meteor. Soc.* 94: 1653–1660. doi:10.1175/BAMS-D-12-00160.1.

Mead, J.B., A. L. Pazmany, S. M. Sekelsky, and R. E. McIntosh. 1996. "Millimeter-wave radars for remotely sensing clouds and precipitation." *IGARSS '96: Remote Sensing for a Sustainable Future.* IEEE. 1553-1555. doi:10.1109/IGARSS.1996.516728.

Miller, D.K., C.F. Miniat, R.M. Wooten, and A.P. Barros. 2019. "An expanded investigation of atmospheric rivers in the southern Appalachian Mountains and their connection to landslides." *Atmosphere* 10: 71. doi:10.3390/atmos10020071.

Nair, S., G. Srinivasan, and R. Nemani. 2009. "Evaluation of Multi-Satellite TRMM Derived Rainfall Estimates over a Western State of India." *Journal of the Meteorological Society of Japan* 87 (6): 927-939. doi:10.2151/jmsj.87.927.

Nellemann, C., M. MacDevette, T. Manders, B. Eickhout, B. Svihus, A. Prins, and B. Kaltenborn. 2009. *The Environmental Food Crisis. The environment's role in averting future food crises. A UNEP rapid response assessment.* Arendal: UNDP.

Pan, L. L., and L. A. Munchak. 2011. "Relationship of cloud top to the tropopause and jet structure from CALIPSO data." *J. Geophys. Res.* 116. doi:10.1029/2010JD015462.

Pan, L.L., K. P. Bowman, E. L. Atlas, S. C. Wofsy, F. Zhang, J. F. Bresch, B. A. Ridley, et al. 2010. "The Stratosphere-Troposphere Analyses of Regional Transport 2008 Experiment." *Bull. Amer. Meteor. Soc.* 91: 327-342. doi:10.1175/2009BAMS2865.1.

Passarelli, R. E., Jr., and H. Boehme. 1983. "The Orographic Modulation of Pre-Warm-Fron Precipitation in Southern New England." *Mon. Wea. Rev.* 111: 1062-1070. doi:10.1175/1520-0493(1983)111<1062:TOMOPW>2.0.CO;2.

Petaja, T. 2014. *Science Plan Biogenic Aerosols – Effects on Clouds and Climate (BAECC).* US Department of Energy, Office of Science.

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., A. P. Barros, and F. Y. Testik. 2012. "On the Influence of Raindrop Collision Outcomes on Equilibrium Drop Size Distributions." *J. Atmos. Sci.* 69: 1534-1546. doi:10.1175/JAS-D-11-0192.1.

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.

Prat, O. P., and A. P. Barros. 2007. "Exploring the use of a column model for the characterization of microphysical processes in warm rain: results from a homogeneous rainshaft model." *Adv. Geosci.* 10: 145-152. doi:10.5194/adgeo-10-145-2007.

Prat, O. P., and A. P. Barros. 2010. "Ground observations to characterize the spatial gradients and vertical structure of orographic precipitation – Experiments in the inner region of the Great Smoky Mountains." *J. Hydrol.* 391: 141-156. doi:10.1016/j.jhydrol.2010.07.013.

Prat., O. P., A. P. Barros, and C. R. Williams. 2008. "An intercomparison of model simulations and VPR estimates of the vertical structure of warm stratiform rainfall during TWP-ICE." *J. Appl. Meteor. Climatol.* 11: 2797-2815. doi:10.1175/2008JAMC1801.1.

Purdy, J. C., G. L. Austin, A. W. Seed, and I. D. Cluckie. 2005. "Radar evidence of orographic enhancement due to the seeder feeder mechanism." *Meteorol. Appl.* 12: 199-206. doi:10.1017/S1350482705001672.

Qian, N. 1999. "On the momentum term in gradient descent learning algorithms." *Neural Networks* 12: 145-151.

Raupach, T.M., and A. Berne. 2016. "Small-scale variability of the raindrop size distribution and its effect on areal rainfall retrieval." *J. Hydrometeorol.* 17: 199-206. doi:10.1175/JHM-D-15-0214.1.

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.

Robichaud, A. J., and G. L. Austin. 1988. "On the modelling of warm orographic rain by the seeder-feeder mechanism." *Quart. J. Roy. Meteor. Soc.* 114: 967-988. doi:10.1002/qj.49711448207 .

Roe, G. H. n.d. "Orographic Precipitation." *Annual Reviews* 33: 645-671. doi:10.1146/annurev.earth.33.092203.122541.

Roe, G. H., D. R. Montgomery, and B. Hallet. n.d. "Orographic precipitation and the relief of mountain ranges." *J. Geophys. Res.* 108: 2315. doi:10.1029/2001JB001521.

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.

Scheel, M. L. M., m. Rohrer, c. Huggel, D. S. Villar, E. Silvestre, and G. J. Huffman. 2011. "Evaluation of TRMM Multi-satellite Precipitation Analysis (TMPA) performance in the Central Andes region and its dependency on spatial and temporal resolution." *Hydrology and Earth System Sciences* 2649-2663.

Schönhuber, M., G. Lammer, and W. L. Randeu. 2008. "The 2D-video-disdrometer." In *Precipitation: Advanced in Measurement , Estimation and Prediction*, edited by S.C. Michaelides, 3-31. Springer. doi:10.1007/978-3-540-77655-0_1.

Schumacher, C., and R. A. Houze. n.d. "The TRMM precipitation radar's view of shallow, isolated rain." *J. Appl. Meteor.* 42 (10): 1519–1524. doi:10.1175/1520-0450(2003)042h1519:TTPRVOi2.0.CO;2.

Segelstein, D. 1981. *The Complex Refractive Index of Water.* Kansas City: MS thesis, University of Missouri. https://mospace.umsystem.edu/xmlui/handle/10355/11599.

Seto, S., T. Iguchi, and T. Oki. 2013. "The Basic Performance of a Precipitation Retrieval Algorithm for the Global Precipitation Measurement Mission's Single/Dual-Frequency Radar Measurements." *IEEE Trans. Geosci. Remote Sens.* 51 (12): 5239-5251. doi:10.1109/TGRS.2012.2231686.

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.

Sharma, S. S., and A. Ganju. 2000. "Complexities of avalanche forecasting in western Himalayas – An overview." *Cold Region Sci. Technol.* 31: 95-102. doi:10.1016/S0165-232X(99)00034-8.

Short, D. A., and K Nakamura. 2000. "TRMM radar observations of shallow precipitation over the tropical oceans." *J. Climate* 4107-4124.

Shrestha, P., A. P. Barros, and A. Khlystov. 2013. "CCN estimates from bulk hygroscopic growth factors of ambient aerosols during the pre-monsoon season over Central Nepal." *Atmos. Environ.* 67: 120-129. doi:10.1016/j.atmosenv.2012.10.042.

Shrestha, P., A. P. Barros, and A. Khlystov. 2010. "Chemical composition and aerosol size distribution of the middle mountain range in the Nepal Himalayas during the 2009 pre-monsoon season." *Atmos. Chem. Phys.* 10: 11605–11621. doi:10.5194/acp-10-11605-2010.

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.

Sindhu, K. D., and G. S. Bhat. 2013. "Comparison of CloudSat and TRMM radar reflectivities." *J. Earth Syst. Sci.* 947-956.

Skofronick-Jackson, G., and Co-authors. 2017. "The Global Precipitation Measurement (GPM) Mission for Science and Society." *Bull. Amer. Meteor. Soc.* 98: 1679–1695. doi:10.1175/BAMS-D-15-00306.1.

Smyth, T. J., and A. J. Illingworth. 1998. "Correction for attenuation of radar reflectivity using polarization data." *Q.J.R. Meteorol. Soc.* 124: 2393-2415. doi:10.1002/qj.49712455111.

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.

Srivastava, N., G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov. 2014. "Dropout: a simple way to prevent neural networks from overfitting." *Journal of machine learning research* 15 (1): 1929-1958.

Steiner, M., R. A. Houze Jr., and S. E. Yuter. 1995. "Climatological characterization of three-dimensional storm structure from operational radar and rain gauge data." *J. Appl. Meteor.* 34 (9): 1978–2007. doi:10.1175/1520-0450(1995)034<1978:CCOTDS>2.0.CO;2.

Stepherns, G. L., and Coauthors. 2008. "CloudSat mission: Performance and early science after the first year of operation." *J. Geophys. Res.* 113: D00A18. doi:10.1029/2008JD009982.

Storebo, P. B. 1976. "Small scale topographical influences on precipitation." *Tellus* 28:1: 45-58. doi:10.3402/tellusa.v28i1.10193.

Tanelli, S., G. F. Sacco, S. L. Durden, and Z. S. Haddad. 2012. "Impact of non-uniform beam filling on spaceborne cloud and precipitation radar retrieval algorithms." *SPIE Asia-Pacific Remote Sensing.* International Society for Optics and Photonics.

—. 2012. "Impact of non-uniform beam filling on spaceborne cloud and precipitation radar retrieval algorithms." *SPIE 8523, Remote Sensing of the Atmosphere, Clouds, and Precipitation IV.* 852308. doi:10.1117/12.977375.

Tao, J, and A. P. Barros. 2014. "Coupled prediction of flood response and debris flow initiation during warm- and cold-season events in the Southern Appalachians, USA." *Hydrol. Earth Syst. Sci.* 18: 367-388. doi:10.5194/hess-18-367-2014.

Tao, J, D. Wu, J. Gourley, S. Q. Zhang, W. Crow, C. Peters-Lidard, and A. P. Barros. 2016. "Operational hydrological forecasting during the IPHEx-IOP campaign – Meet the challenge." *J. Hydrol.* 541: 434-456. doi:10.1016/j.jhydrol.2016.02.019.

Tao, J., and A. P. Barros. 2013. "Prospects for flash flood forecasting in mountainous regions - An investigation of Tropical Storm Fay in the Southern Appalachians." *J. Hydrol.* 506: 69-89. doi:10.1016/j.jhydrol.2013.02.052.

Tao, K., and A. P. Barros. 2010. "Using Fractal Downscaling of Satellite Precipitation Products for Hydrometeorological Applications." *J. Atmos. Oceanic Technol.* 27: 409-427. doi:10.1175/2009JTECHA1219.1.

Tao, W.-K., and Coauthors. 2009. "The Goddard multi-scale modeling system with unified physics." *Ann. Geophys.* 27: 3055–3064. doi:10.5194/angeo-27-3055-2009.

Tapiador, F.J., A. Berne, T. Raupach, A. Navarro, G. Lee, and Z.S. Haddad. 2018. "Objective characterization of rain microphysics: validating a scheme suitable for weather and climate models." *J. Hydrol.* 19 (6): 929-946. doi:10.1175/JHM-D-17-0154.1.

Tarolli, M., M. Borga, D. Zoccatelli, C. Bernhofer, N. Jatho, and F. A. Janabi. 2013. "Rainfall Space-Time Organization and Orographic Control on Flash Flood Response: The Weisseritz Event of August 13, 2002." *Journal of Hydrologic Engineering* 18 (2). doi:10.1061/28ASCE29HE.1943-5584.0000569.

Testik, F. Y., and A. P. Barros. 2007. "Toward Elucidating the Microstructure of warm rainfall: A survey." *Rev. Geophys.* 45: RG2003. doi:10.1029/2005RG000182.

Testik, F.Y., A. P. Barros, and L. F. Bliven. 2011. "Toward a Physical Characterization of Raindrop Collision Outcome Regimes." *J. Atmos. Sci.* 68: 1097–1113. doi:10.1175/2010JAS3706.1.

Testud, J., Bouar, E. L., E. Obligis, and M. Ali-Mehenni. 2000. "The Rain Profiling Algorithm Applied to Polarimetric Weather Radar." *J. Atmos. Oceanic Technol.* 17: 332-356. doi:10.1175/1520-0426(2000)017<0332:TRPAAT>2.0.CO;2.

Testud, J., S. Oury, R. A. Black, P. Amayenc, and X. Dou. 2001. "The Concept of ''Normalized'' Distribution to Describe Raindrop Spectra: A Tool for Cloud Physics and Cloud Remote Sensing." *J. Appl. Meteor.* 40: 1118-1140. doi:10.1175/1520-0450(2001)040<1118:TCONDT>2.0.CO;2.

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.

Thurai, M., and V. N. Bringi. 2018. "Application of the Generalized Gamma Model to Represent the Full Rain Drop Size Distribution Spectra." *J. Appl. Meteor. Climatol.* 57: 1197-1210. doi:10.1175/JAMC-D-17-0235.1.

Thurai, M., P. Gatlin, V. N. Bringi, W. Petersen, P. Kennedy, B. Notaros, and L. Carey. 2017. "Toward Completing the Raindrop Size Spectrum: Case Studies Involving 2D-Video Disdrometer, Droplet Spectrometer, and Polarimetric Radar Measurements." *J. Appl. Meteor. Climatol.* 56: 877-896. doi:10.1175/JAMC-D-16-0304.1.

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.

Tokay, A., D. B. Wolff, and W. A. Petersen. 2014. "Evaluation of the new version of the laser-optical disdrometer, OTT Parsivel2." *J. Atmos. Ocean. Technol.* 31: 1276-1288. doi:10.1175/JTECH-D-13-00174.1.

Torregrosa, A., T. A. O'Brien, and I. C. Faloona. 2014. "Coastal Fog, Climate Change, and the Environment." *Eos Trans. AGU* 95 (50): 473. doi:10.1002/2014EO500001.

Toyoshima, K., H. Masunaga, and F. A. Furuzawa. 2015. "Early evaluation of Ku- and Ka-band sensitivities for the Global Precipitation Measurement (GPM) Dual-Frequency Precipitation Radar (DPR)." *SOLA* 11: 14-17. doi:10.2151/sola.2015-004.

Turk, J. 18. *CloudSat-GPM coincidence dataset, version 1C.* Algorithm Theoretical Basis Doc., NASA. https://pmm.nasa.gov/sites/default/files/document_files/CSATGPM_COIN_ATBD.pdf.

Ulbrich, C. W. 1983. "Natural variations in the analytical form of the raindrop size distribution." *J. ClimateAppl. Meteor.* 22: 1764-1775. doi:10.1175/1520-0450(1983)022,1764:NVITAF.2.0.CO;2.

Viviroli, D., D. R. Archer, W. Buytaert, H. J. Fowler, Gregory B. Greenwood, Alain F. Hamlet, Yan Huang, et al. 2011. "Climate Change and Mountain Water Resources: Overview and Recommendations for Research, Management and Policy." *Hydrology and Earth System Science* 471-504.

Wang, J., M. Cheng, Y. Gao, Y. Xiong, and S. Zhu. 2008. "Classification of precipitation radar reflectivity echo with Back-Propagation ANN." *PACIIA, 2008.* IEEE. 630-633. doi:10.1109/PACIIA.2008.351.

Widener, K., N. Bharadwaj, and K. Johnson. 2012. *Ka-Band ARM Cloud Radar (KAZR).* DOE/SC-ARM/TR-106, US Department of Energy, Office of Science.

Widener, K., N. Bharadwaj, and K. Johnson. 2012. *Scanning ARM Cloud Radar (X/Ka/W-SACR).* DOE/SC-ARM/TR-113, US Department of Energy, Office of Science.

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.

Williams, C. R. 2016. "Reflectivity and Liquid Water Content Vertical Decomposition Diagrams to Diagnose Vertical Evolution of Raindrop Size Distributions." *J. Atmos. Oceanic Technol.* 33: 579-595. doi:10.1175/JTECH-D-15-0208.1.

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.

Zafar, B J, and V Chandrasekar. 2004. "Classificatication of precipitation type from space borne precipitation radar data and 2D wavelet analysis." *Geoscience and Remote Sensing Symposium* 3570-3573.

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

Glossary of Abbreviations and Symbols

CNN – Convolutional Neural Network

CSI – Critical Score Index

D – drop diameter [mm]

D_m – Mass weighted mean diameter

DPR – Dual-frequency Precipitation Radar

DSD – Drop Size Distribution

FA – False Alarms

FAR – False Alarm Ratio

FB – Frequency Bias

GPM – Global Precipitation Measurement Mission

GV – Ground Validation

H* - Vertical Correlation Structure function

HRRR – High Resolution Rapid Refresh

IPHEx – Integrated Precipitation and Hydrology Experiment

Ku-PR – Ku-Band Precipitation Radar

Ka-PR – Ka-Band Precipitation Radar

LLCF – Low-level Clouds and Fog

LWC – Liquid Water Content

MD – Missed Detections

MPS – Meteorological Particle Spectrometer

MRMS – Multi-Radar/ Multi-Sensor

MRR – Micro Rain Radar

N(D) – Drop number concentration [$m^{-3}mm^{-1}$] at drop diameter D

NN – Number of correct detections of no-precipitation events

N_w – Normalized intercept parameter of drop size distribution following normalized

gamma distributions

NWP – Numerical Weather Prediction

OVR – Overestimations

POD – Probability of Detections

PR – Precipitation Radar

QPE – Quantitative Precipitation Estimates

SAM – Southern Appalachian Mountains

SFI – Seeder Feeder Interactions

TBC – Top Boundary Condition

TRMM – Tropical Rainfall Measurement Mission

VCS – Vertical Correlation Structure

UND – Underestimations

YY – number of correct detections of precipitation events

Z – radar reflectivity factor with Rayleigh Approximation

Z_m – Measured reflectivity factor

Z_e – Equivalent reflectivity factor

z_c – Zero-crossing height of VCS function

ε – Estimation Bias

μ – shape factor of drop size distribution following normalized gamma distribution

σ_b – Backscattering cross-section

σ_e – Extinction cross-section

ϱ_w – density of water – 1000 kgm^{-3}

Appendix B

Review of Rainfall Classification Algorithm - The rainfall regime classification strategy used in the TRMM-PR (13.8 GHz; Ku-Band) retrieval algorithm uses two methods, one focusing on the horizontal structure, and the other on the vertical structure of PR reflectivity and the presence of a bright band (BB), that is the very high reflectivity layer associated with the partial melting of ice hydrometeors in the case of stratiform conditions (Duan et al. 2015; Berg et al. 2006). A final TRMM classification product, TRMM-2A23, was derived by combining the two methods (Awaka et al. 1997). Steiner et al. (1995) suggests that the precipitation classification based on BB detection (vertical method) is not operationally feasible since the presence of BB is not clear during the initial stages of stratiform precipitation or when rainfall intensity is high, thus underestimating rainfall. They developed a classification algorithm to overcome BB ambiguity by focusing instead on identifying the convective center of the radar echo profiles based on the intensity and sharpness of the peak. The premise is to locate the convective center of the radar echoes (and precipitation events) without distinguishing first between convective and stratiform conditions toward reducing the underestimation of stratiform events. Biggerstaff and Listemaa (2000) extended the Steiner et al. (1995) algorithm for heavy stratiform and light convective events with a classification based: vertical lapse rate of reflectivity, BB fraction at every grid point, and the magnitude of the horizontal gradient of the reflectivity factor. The modified method was tested with ground-based Weather

Surveillance Radar data in Houston, Texas and showed improvement in classification accuracy compared to Steiner et al. (1995), but further testing for different hydrometeorological conditions and using satellite-based radar products was not attempted.

Schumacher and Houze (2003) developed a new classification algorithm by merging the algorithm developed by Awaka et al. (1997) (Vertical Method) and Steiner et al. (1995) (Horizontal Method) to identify shallow isolated echo rain in TRMM V6 (version 6) that was a major source of error in the earlier TRMM V5 (version-5) products but BB detection errors in the Vertical Method were not eliminated in TRMM-V6, and the BB reflectivity for low freezing heights is often misclassified as maximum reflectivity in the Horizontal Method. Zafar and Chandrasekar (2004) proposed an algorithm to overcome the issues in the Vertical and Horizontal Methods by decomposing and reconstructing the profiles to Approximate, Vertical, Horizontal and Diagonal Coefficients using wavelet transforms. They reported agreement with TRMM 2A23 about 80% of the time in five case-studies with a 15% increase in BB detection, and 11% increase in the classification of stratiform events that were otherwise classified as convective in TRMM 2A23, though with small differences (approximately 5%) in the overall precipitation classification.

Other approaches to developing precipitation classification algorithms have used passive microwave observations from space (e.g., Spatial Sensor Microwave/Imager, SSM/I), and ground-based radars and disdrometers. Hong et al. (1999) classified

convective and stratiform precipitation over the tropical ocean using the brightness temperatures at 19, 37 and 85 GHz for simulated profiles of TRMM and SSM/I. Thurai et al. (2016) utilized X- band radar profiles to identify the bright band and drop size distributions measured by 2-D video disdrometers to classify precipitation events. The method developed by Thurai et al. (2016) was tested using data from Ontario, CA and Huntsville, Alabama, USA, and the authors suggest that the method would not work efficiently for frequent shallow rainfall over the tropical oceans. Wang et al. (2008) implemented a back-propagation and artificial neural network algorithm to classify the precipitation events with CAPPI (Constant Altitude Plan Position Indicator) reflectivity profiles from ground-based S-band radar. The algorithm was trained with reflectivity profiles between 5 km and 10 km altitudes only and tested at one location in China, and thus further validation is needed for generalized application.

The TRMM follow-on Global Precipitation Mission (GPM) core satellite (Hou et al. 2014) is equipped with a dual-frequency Ku-Ka Precipitation Radar (DPR). Le and Chandrasekar (2013) proposed a classification algorithm for the GPM-DPR based on the ratio of the normalized difference Dual Frequency Ratio (DFR) index calculated from the maximum and minimum of the measured DFR profiles and the mean slope of the DFR profile well above the ground-surface. The DFR is a measure of the spectral width of the drop size distributions detected using two distinct radar frequencies, and it is expected to increase as the particle size increases (Matrosov et al. 2005). The algorithm was

demonstrated for oceanic precipitation using aircraft-based Dual Frequency Ratio (DFR) profiles over the Atlantic Ocean and over Wakasa Bay, Japan. Threshold values of the classification index for convective and stratiform precipitation cases were defined for the 90% of the empirical cumulative density function, thus implying a minimum acceptable 10% misclassification error.

Early assessments of TRMM-PR rainfall products in complex terrain reported large estimation errors. (e.g., Prat and Barros 2010). Duan et al. (2015) specifically highlighted uncertainty associated with "stratiform" and "probably stratiform" classes in TRMM-V6 and V7 products that correspond to shallow rainfall systems and reverse orographic enhancement effects (Wilson and Barros 2014, 2015, 2017). Despite significant improvements in radar rainfall retrievals in the TRMM era (Iguchi et al. 2000), ongoing studies (not shown) similar to Duan et al. (2015) suggest that error statistics have not changed significantly between TRMM-PR to GPM-DPR over complex terrain in the Andes and in the Southern Appalachians.

Appendix C

Supplementary Data for Chapter 3
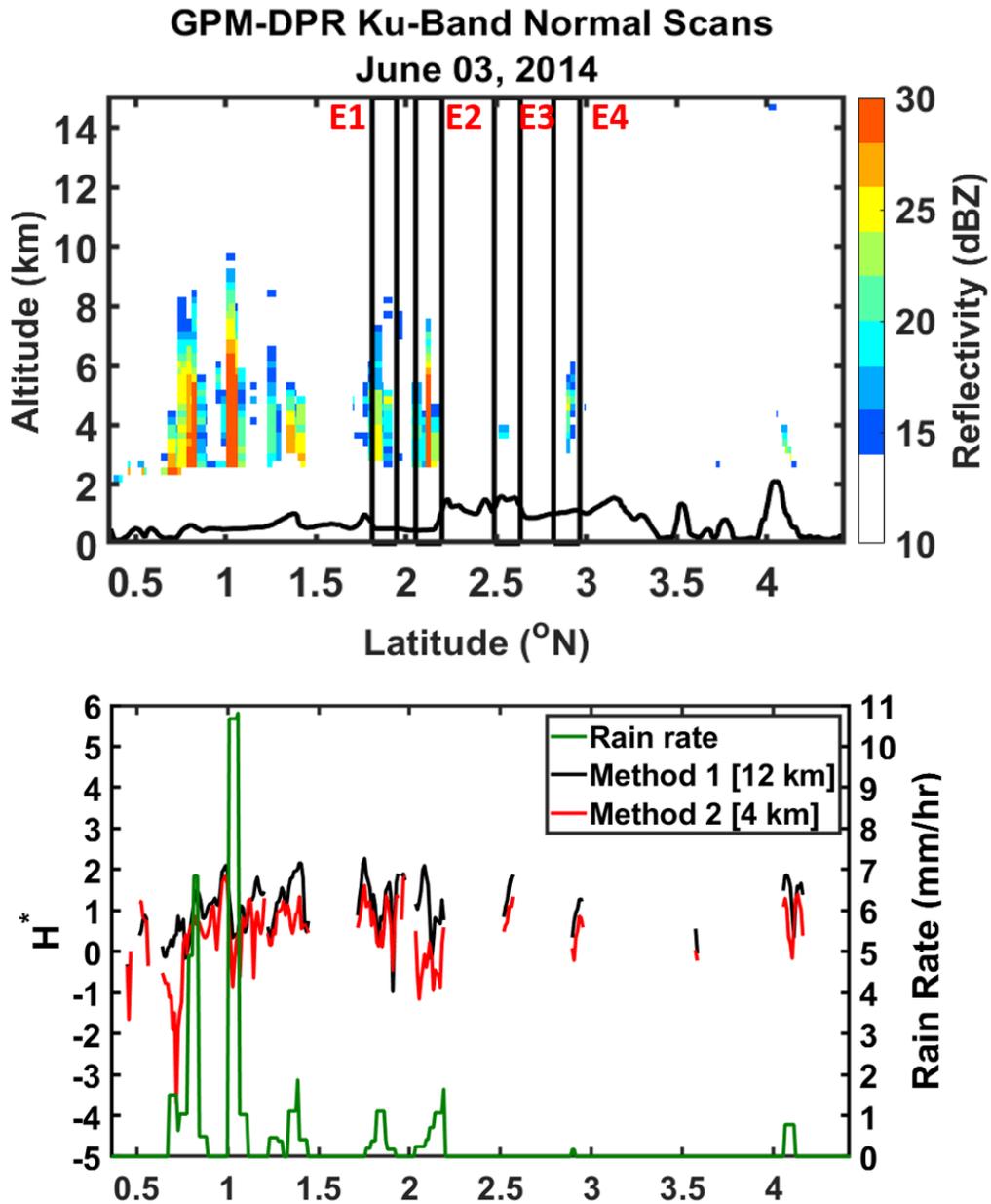
Figure C1: (a) Equivalent reflectivity factor observed by GPM Ku-band radar for Case shown in Fig. 3-18. (b) VCS function computed by method-1 and 2 for GPM Ku-band and CloudSat CPR reflectivity profiles. Black boxes denote the events marked in Fig. 3-18.
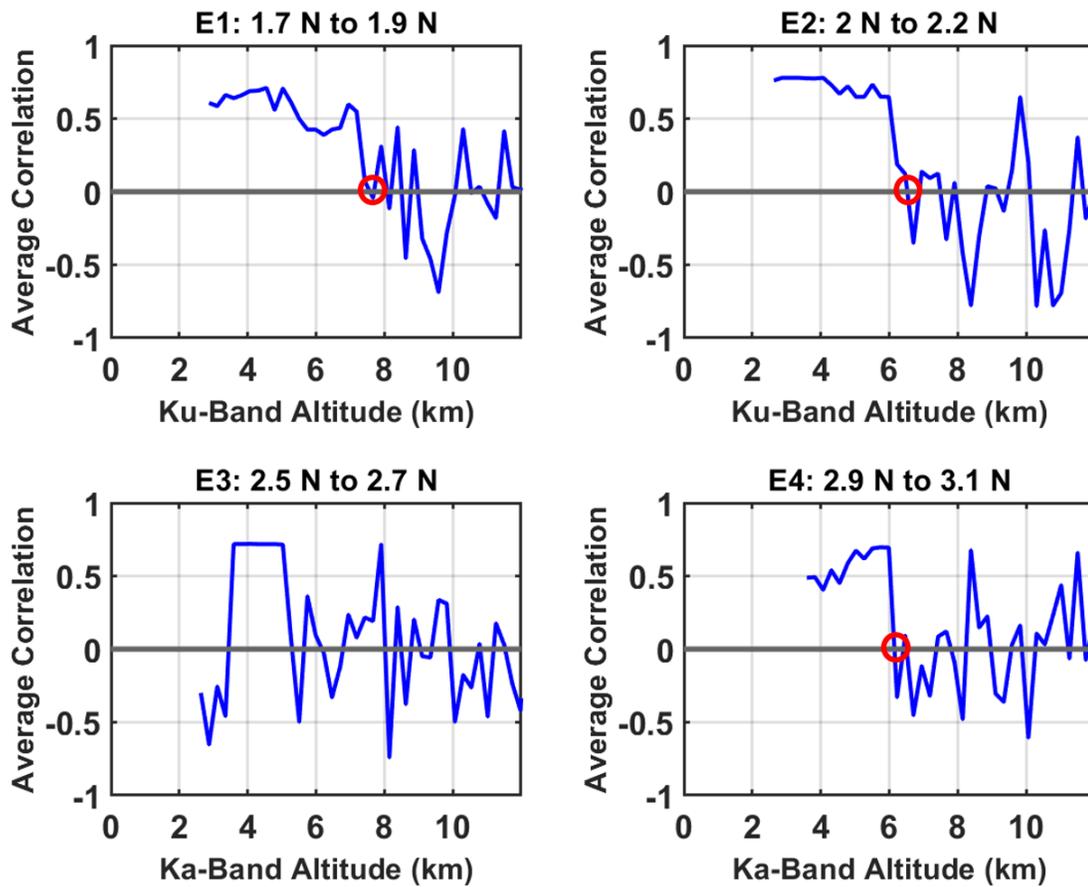
Figure C2 Method 1 average correlation profile computed from the space–time correlation computed in step 2 of the algorithm, where (b) E1: 1.7°–2°N, (c) E2: 2°–2.2°N, (d) E3: 2.5°–2.7°N, and (e) E4: 2.9°–3.1°N. E1 and E4 are deep structure events, E2 is a shallow event, and E3 is a no precipitation event. Note the change in the shape and altitude of first zero-crossing points of the correlation profiles based on the structure of precipitation.

Similar zero-crossing values are observed for GPM Ka-band HS reflectivity profiles (Fig.

3-19) which confirms that zero-crossing can be used as an indicator for storm height.

Appendix D

Drop Size Distribution Formulation - Ulbrich (1983) proposed a simple Gamma

distribution with three unknown parameters (N_0, Λ and μ) to represent the raindrop DSD

as follows:

$$N(D) = N_0 D^\mu \exp(-\Lambda D) \tag{D1}$$

where, N(D) is the raindrop number concentration [$m^{-3}mm^{-1}$], $\Lambda = {(4 + \mu)}/{D_m}$ [mm^{-1}], N_0

is the intercept parameter and a measure of drop concentration [$mm^{-1- \mu}m^{-3}$], D_m is the

mass-weighted mean diameter [mm], D is the drop diameter [mm] and μ [unitless] is the

factor that governs the shape of the DSD. The mass-weighted mean diameter, D_m is

defined as the ratio of fourth moment to the third moment of the DSD. D_m can be

expressed in terms of N(D):

$$D_m = \frac{\sum_{D_{min}}^{D_{max}} N(D)D^4 dD}{\sum_{D_{min}}^{D_{max}} N(D)D^3 dD} \tag{D2}$$

The mass spectrum standard deviation (σ_m) is expressed as:

$$\sigma_m = \left[\frac{\sum_{D_{min}}^{D_{max}}(D-D_m)^2 N(D)D^3 dD}{\sum_{D_{min}}^{D_{max}} N(D)D^3 dD}\right]^{1/2} \tag{D3}$$

The drop diameter ranges between D_{min} [mm] and D_{max} [mm] with interval dD [mm].

However, the three DSD parameters are highly correlated, and the dimension of N_0 is

dependent on shape parameter μ (Testud et al. 2000; Illingworth and Blackman, 2002; and

Tapiador et al. 2018), which introduces large ambiguity in parameter estimation. To

address these issues, the DSD can be rewritten as a normalized Gamma-shaped

distribution with normalized intercept parameter N_w [mm⁻¹mm⁻³] (Testud et al. 2000; Testud et al. 2001; Illingworth and Blackman, 2002; Bringi and Chandrasekar, 2001).

$$N(D) = N_w f(\mu) \left(\frac{D}{D_m}\right)^\mu exp\left(-\frac{(4+\mu)D}{D_m}\right) \tag{D4}$$

where,

$$f(\mu) = \frac{6(4+\mu)^{(\mu+4)}}{4^4 \Gamma(\mu+4)} \tag{D5}$$

and Γ denotes the gamma function. Similar to Eq. (A1), the normalized gamma distribution also has three unknown parameters: D_m, μ and N_w. The normalized intercept factor, N_w [m⁻³ mm⁻¹] is calculated as

$$N_w = \frac{4^4}{\pi \rho_w} \left(\frac{q}{D_m^4}\right) \tag{D6}$$

where ρ_w is the density of water [10⁻³ g mm⁻³] and q is the liquid water content [g m⁻³] derived from the drop size distribution N(D):

$$q = \frac{\pi}{6} \rho_w \sum_{D_{min}}^{D_{max}} N(D) D^3 dD \tag{D7}$$

In radar retrieval algorithms such as the TRMM-PR and the Global Precipitation Measurement (GPM) – Dual-frequency Precipitation Radar (DPR), a fixed shape parameter ($\mu = 3$) is used in the normalized gamma distribution (Eq. A4) for warm (liquid-phase) precipitation (Iguchi et al. 2000; Seto et al. 2013; Liao et al. 2014; Iguchi et al. 2017). Even though the global mean of the shape factor (μ) varies between 3 and 4, recent work shows that the μ exhibits high variation based on the geographical location and type of precipitation (Dolan et al. 2018).

Appendix E

Supplementary Data for Chapter-4
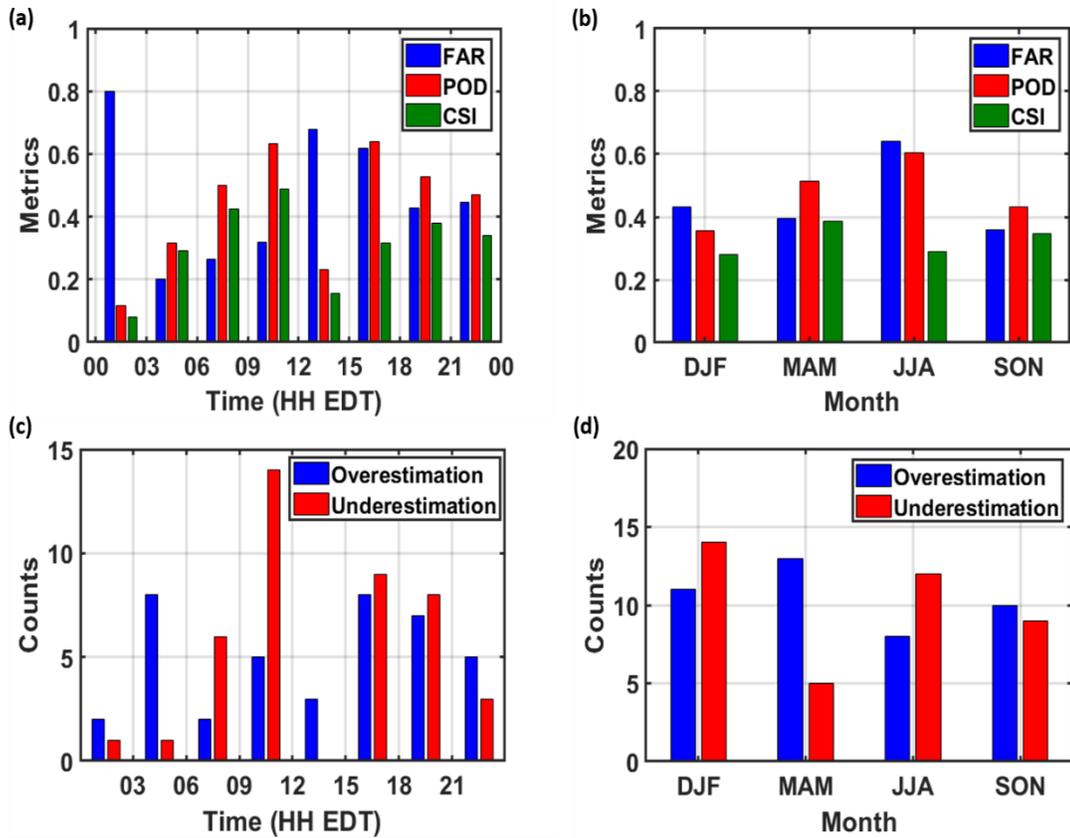
Figure E 1: Same as Figure 4-4 but using average rain-gauge measurements within the radar beam as reference [values between () in Table 4-1]. The diurnal cycle pattern of detection errors here and in Figure 4-4 is the same with a change of only 1 occurrence from OVR to UND between 15-18 EDT and a decrease in OVR of 5 occurrences between 18-21 EDT reflecting the effect of averaging rain-gauge observations among different convective cells at this time-of-day in MAM.
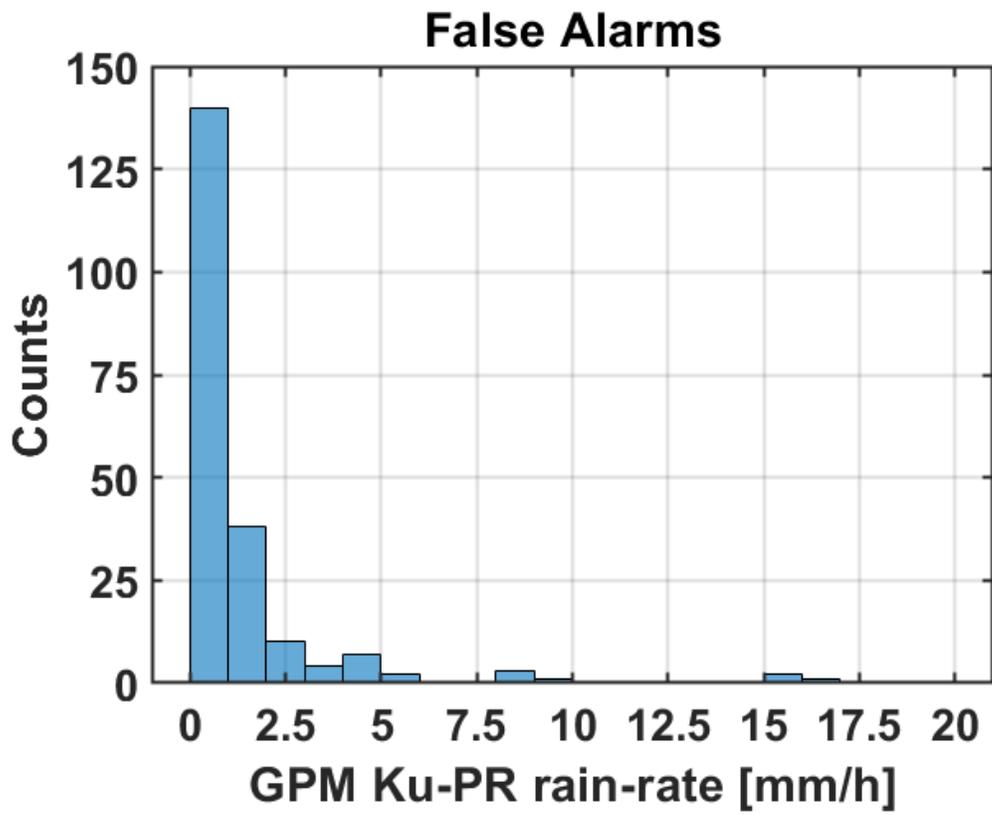
Figure E 2: Histogram of rain-rate estimated by GPM Ku-PR in False Alarm cases.
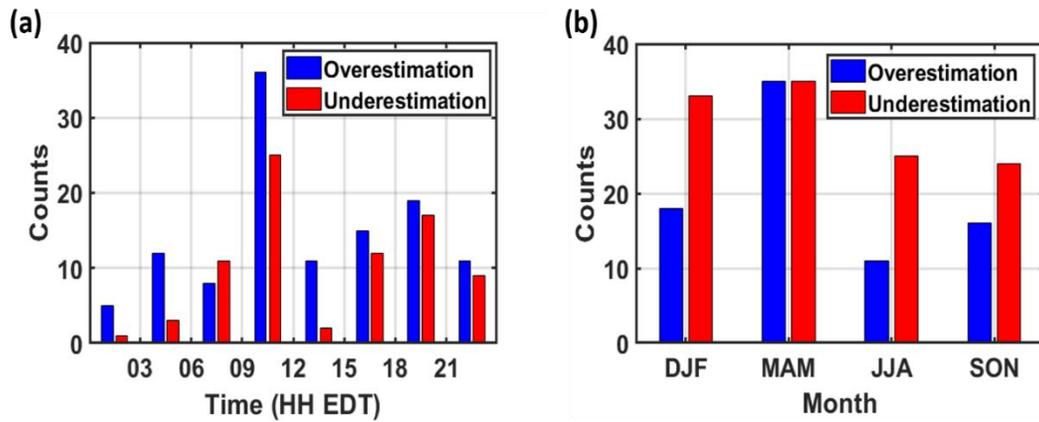
Figure E 3: Same as Figure 4-4(c-d) but OVR and UND are determined by standard metrics (UND – RR$_{RG}$ > RR$_{GPM}$; OVR – RR$_{RG}$ < RR$_{GPM}$).
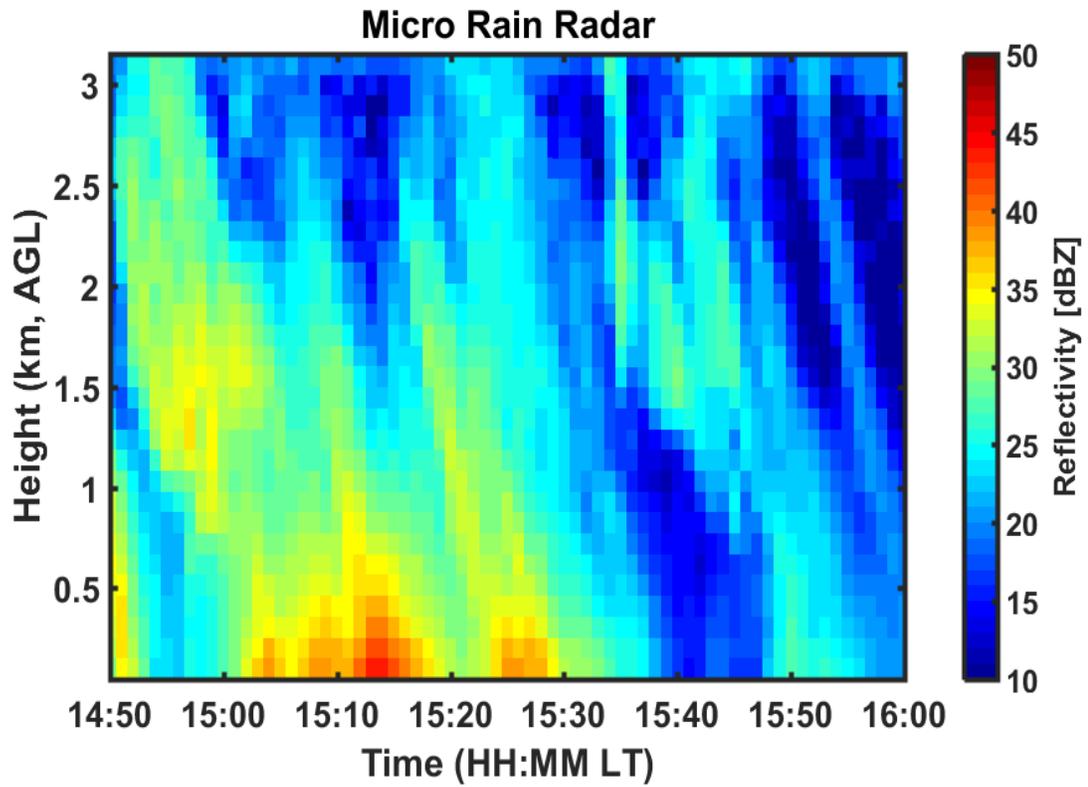
Figure E 4: Micro Rain Radar (MRR) Reflectivity profiles of SFI event observed on October 1, 2015 at Elkmont (P6) in the western foothills of the SAM. Note the low-level enhancement of precipitation between 15:00 and 15:30 LT.
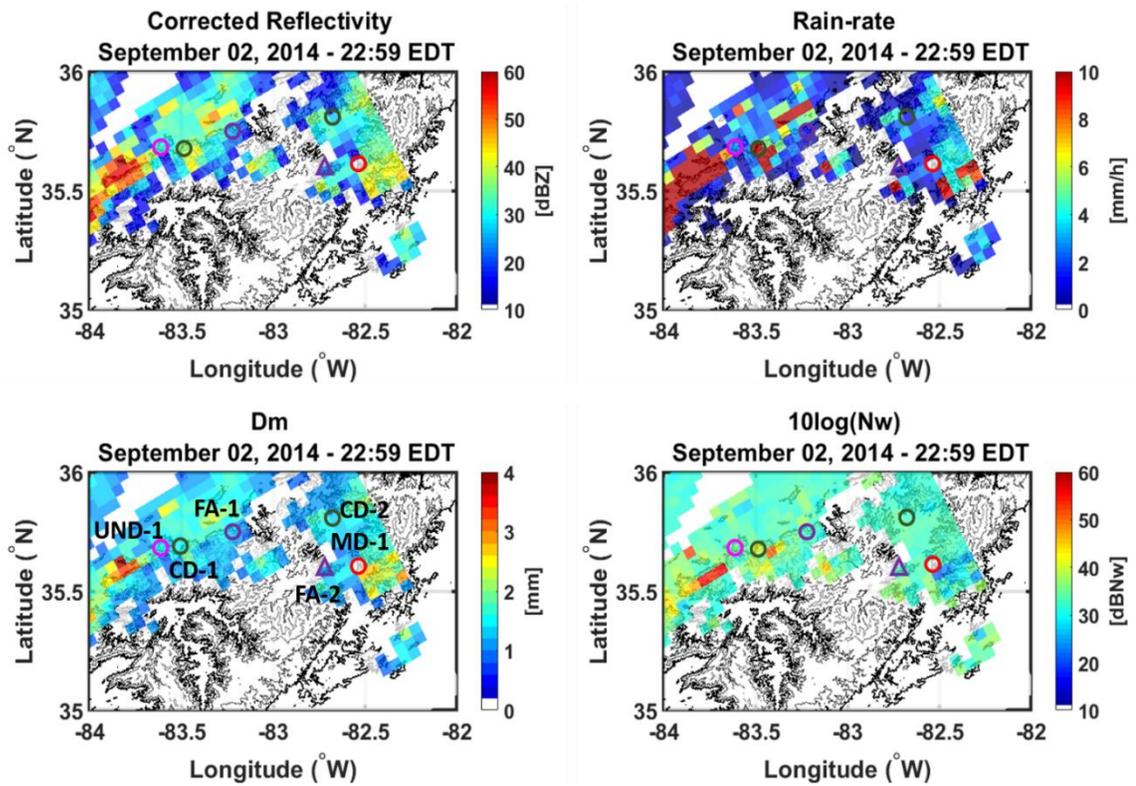
Figure E 5: Spatial distribution of (a) Corrected Reflectivity factor, (b) rain-rate, (c) D_m and (d) N_w estimated by GPM Ku-PR on September 2, 2014 at 22:59 EDT. The contours denote terrain elevations in the SAM: 750 m (black), 1000 m (dark gray) and 1500 m (light gray). Valley locations (elevation < 850 m) and ridges are denoted by circles and triangles respectively.
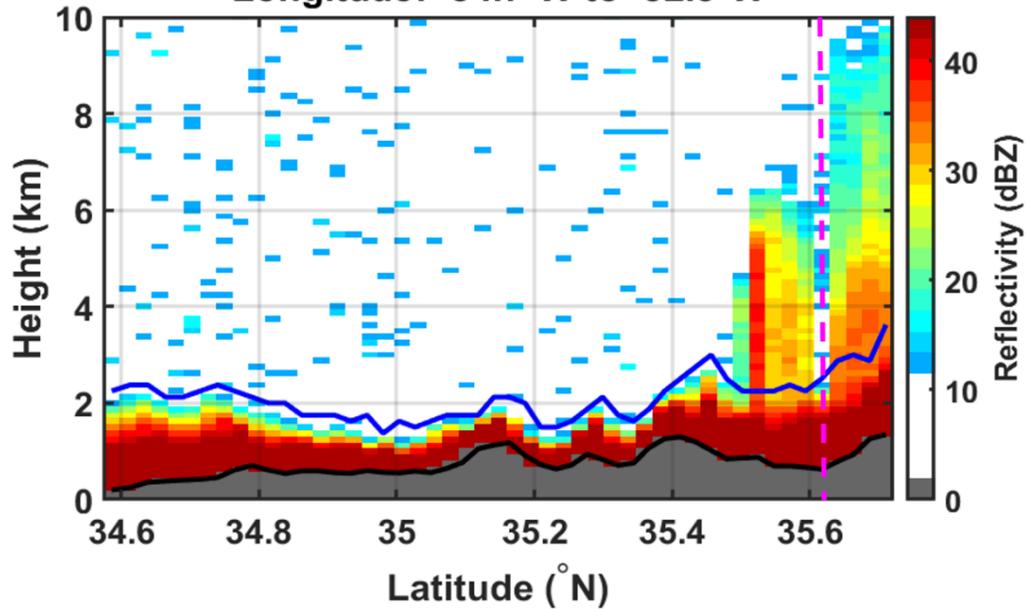
Figure E 6: Measured reflectivity factor by GPM Ku-PR for the warm precipitation case observed on September 2, 2014 at 22:59 LT. The magenta line denotes the location of P2. The gray shaded region and black line denotes the topography. The blue line is the shows the minimum height with no-clutter.
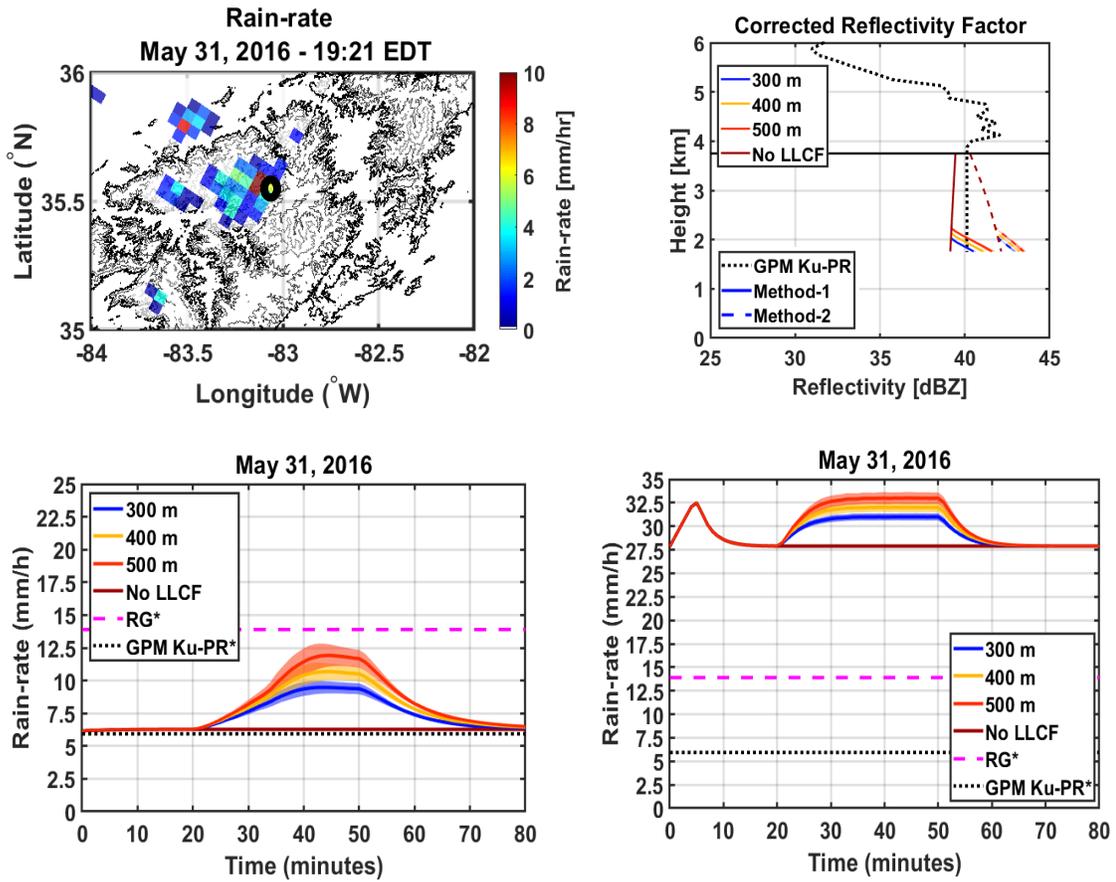
Figure E 7: For case-1 (Table 4-4): Top row- Left: Spatial plot of the GPM Ku-PR estimated near-surface rain-rate. Black circle denotes the rain-gauge location. Right: Corrected reflectivity factor from GPM Ku-PR (black dotted line) and the model simulated profiles (solid- Method-1 and dashed – Method-2). Bottom row – Simulated rain-rate for varying LLCF width compared with rain-gauge observations and GPM Ku-PR for Method-1 and Method-2 simulations.
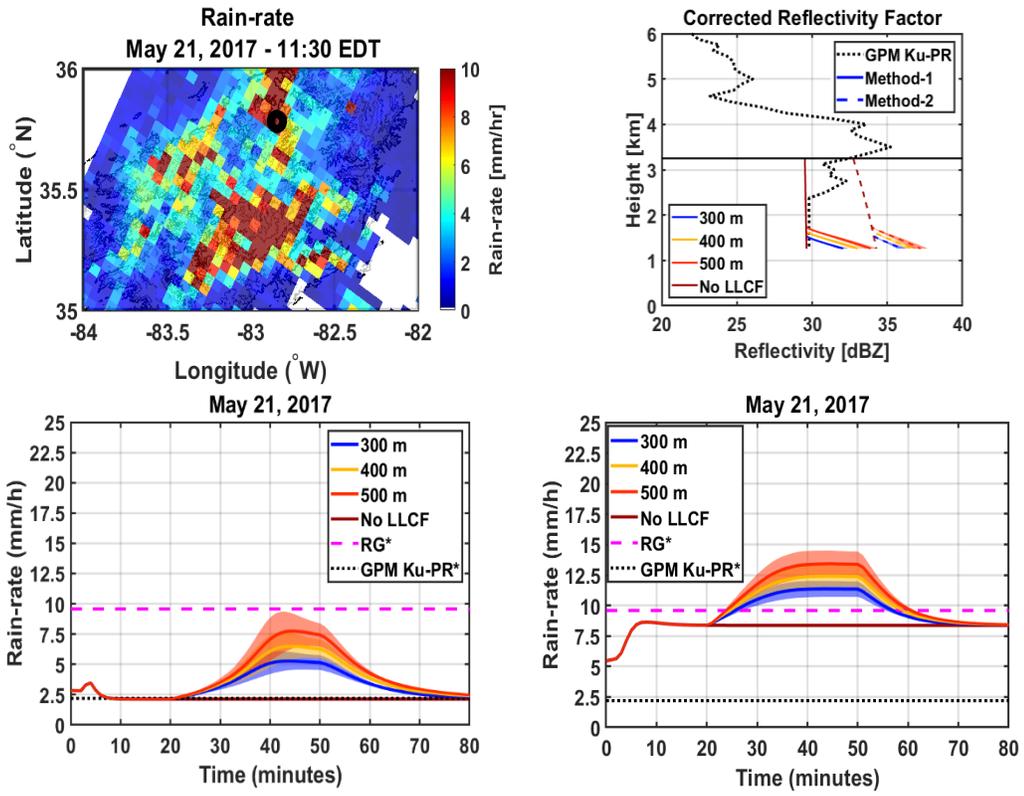
Figure E 8: Same as Fig. E7 but for Case-2 in Table 4-4
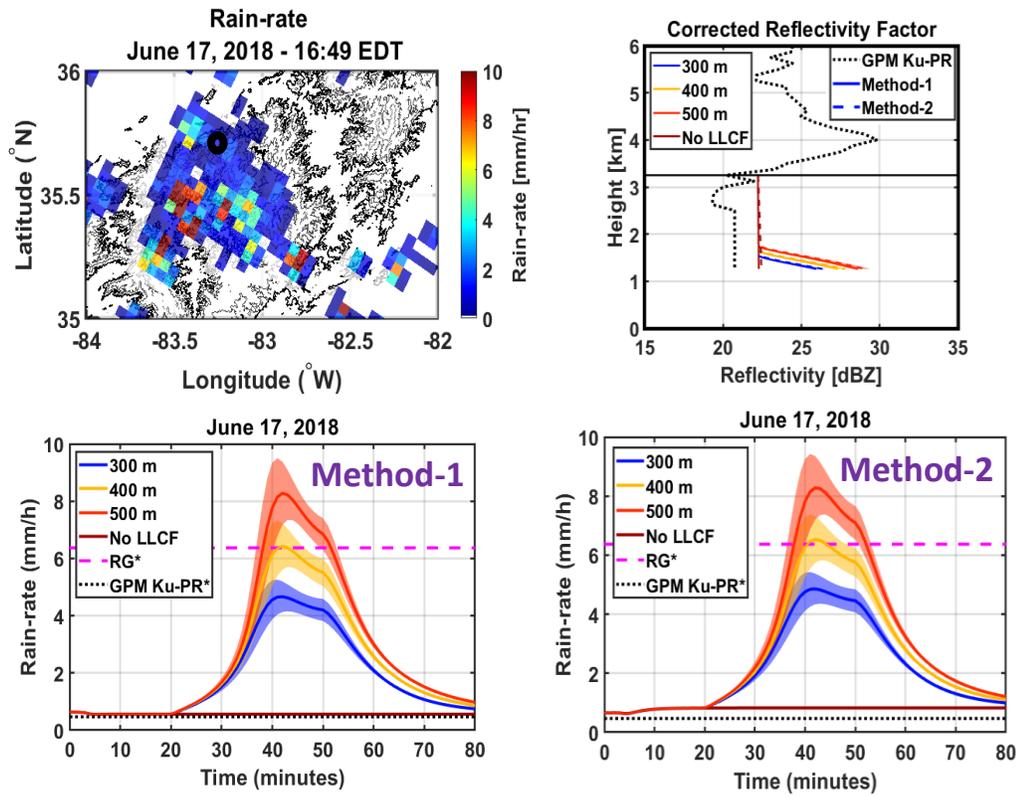
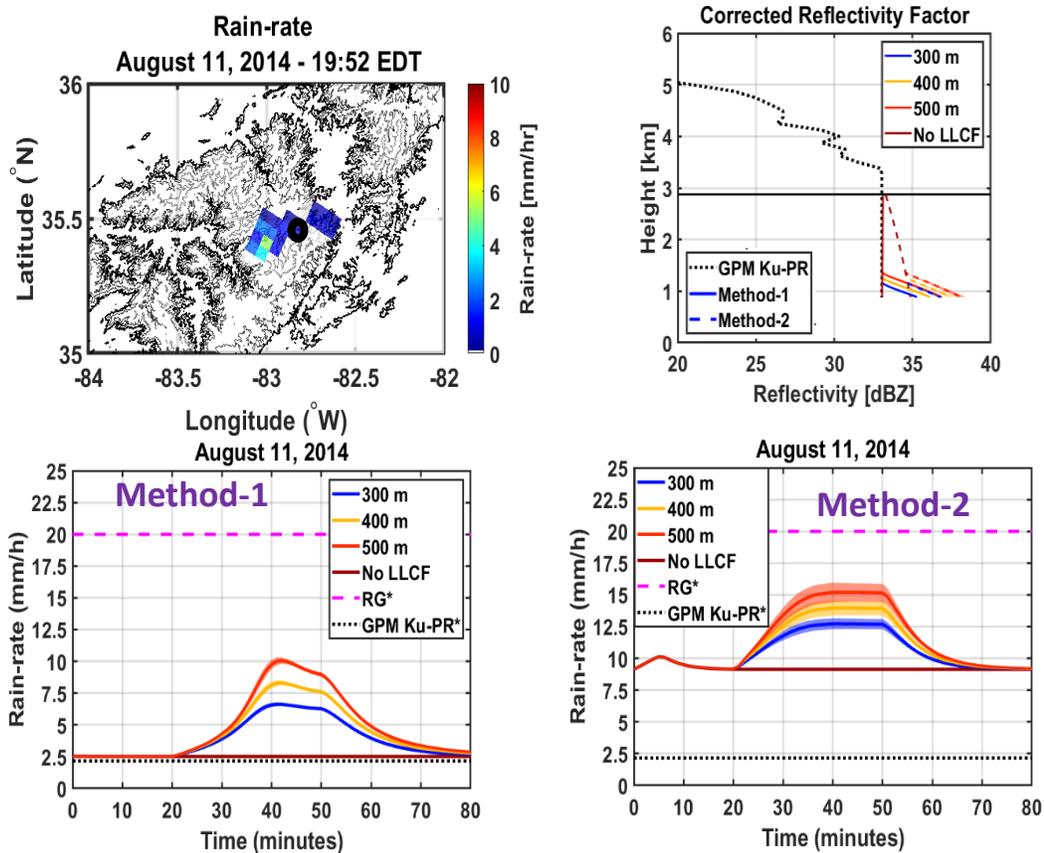**Figure E 9: Same as Fig. E7 but for Case-3 in Table 4-4.**

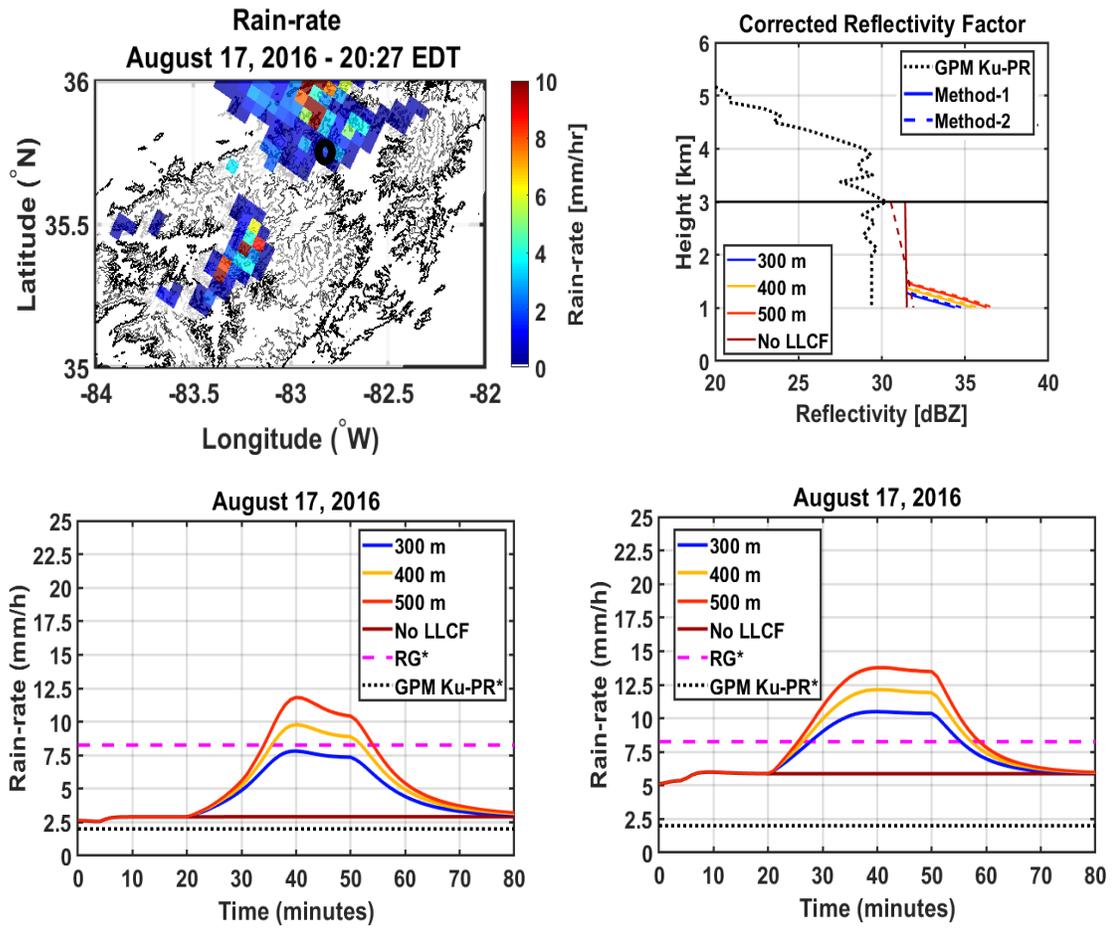**Figure E 10: Same as Fig. E7 but for Case-4 in Table 4-4.**

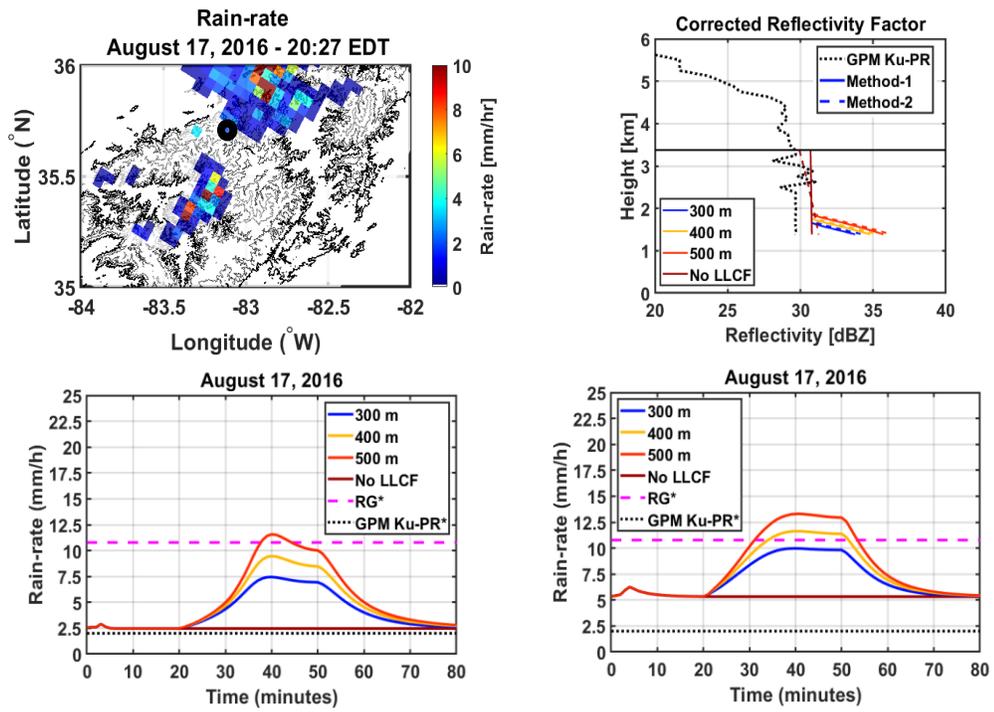**Figure E 11: Same as Fig. E7 but for Case-5 in Table 4-4.**

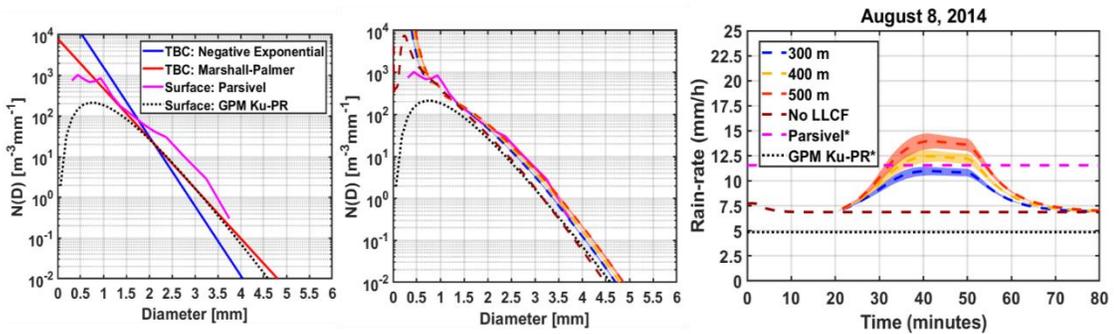**Figure E 12: Same as Fig. E7 but for Case-6 in Table 4-4.**

Figure E 13: Left panel: exponential fit vis-à-vis Marshall-Palmer for the top boundary and initial conditions at P5 (August 8, 2014 case). Mid-panel: model predicted surface DSD with and without LLCF forced by the Marshall-Palmer fit DSD at the TBC and initial conditions. Right panel: model predicted surface rain-rates corresponding to mid-panel DSD. Legend and color scheme in the mid- and right panels are the same.
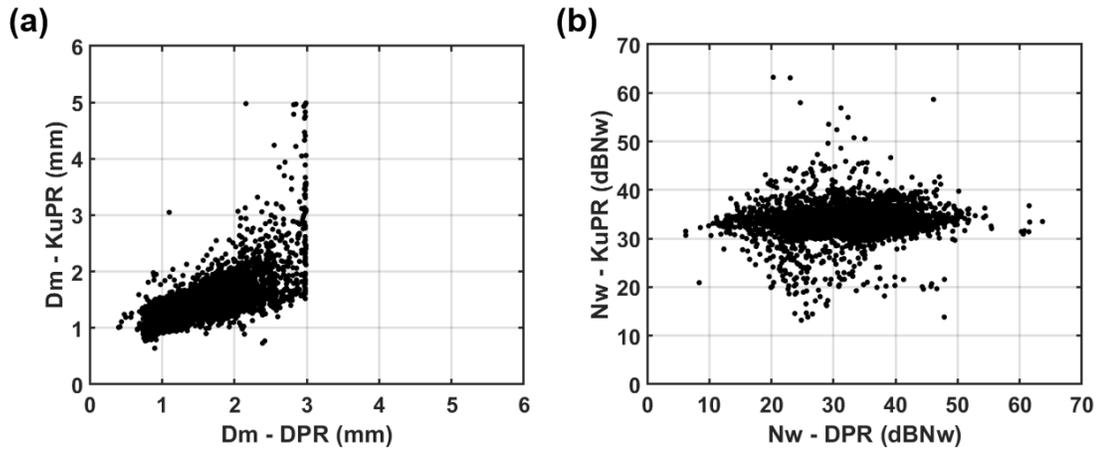
Figure E 14: Comparison of DSD parameters (a) D_m and (b) N_w by the single-frequency algorithm (Ku-PR) and dual-frequency algorithm (DPR) for near-nadir pixels.
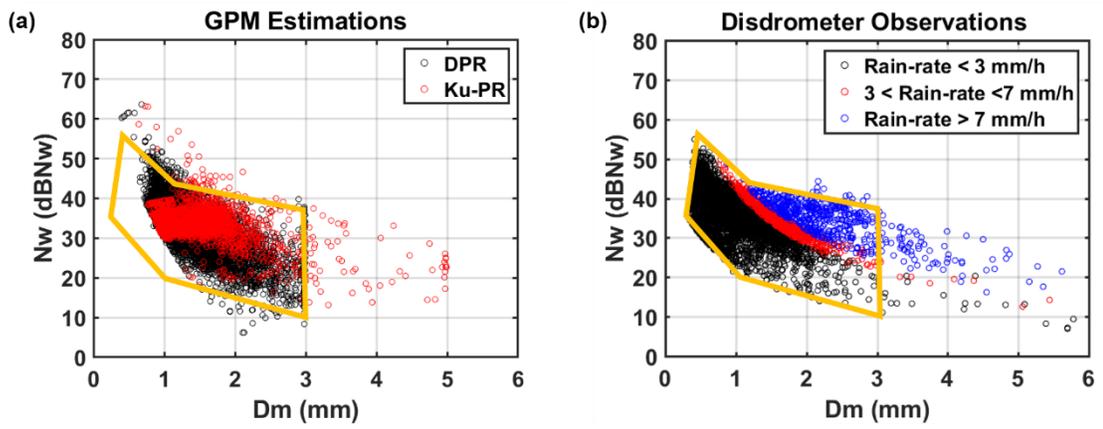
Figure E 15: Dm-Nw relationship estimated by (a) GPM Ku-PR and DPR algorithm and (b) Parsivel disdrometer. Orange lines delineate the region of Dm-Nw space filled by the Parsivel disdrometer.

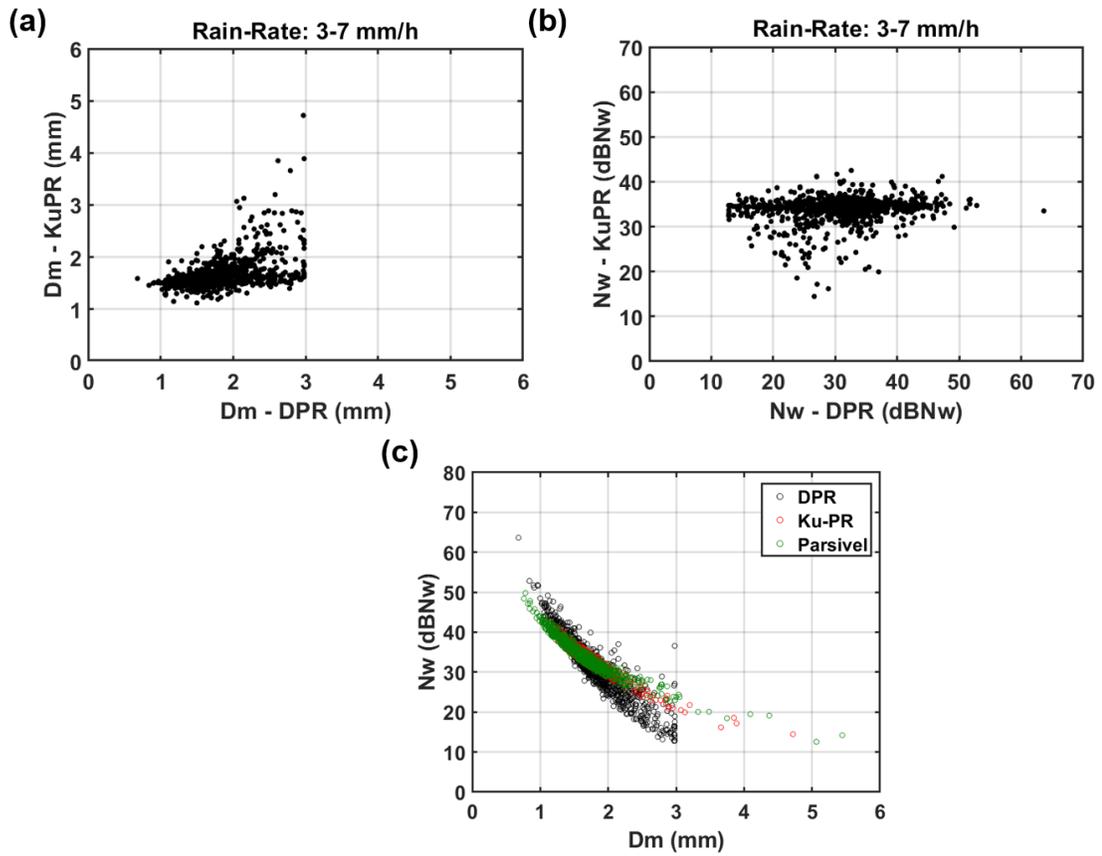**(a)** Rain-Rate: 3-7 mm/h

(b) Rain-Rate: 3-7 mm/h

(c)

Figure E 16: Comparison of DSD parameters estimated by Ku-PR and DPR algorithms: (a) D_m and (b) N_w for rain-rate between 3 and 7 mm/h. (c) D_m-N_w relationship for rain-rate between 3 and 7 mm/h.

Table E 1: List of rain-gauges located in the Great Smoky Mountains National Park at the Southern Appalachian Mountains considered for the analysis.

ID	Resolution [mm/tip]	Latitude [N]	Longitude [W]	Elevation [m]	Start time [mm/dd/yy]	End time [mm/dd/yy]
			EAST			
RG001	0.2	35.41	-82.92	1156.4	03/08/14	04/24/17
RG002	0.2	35.42	-82.97	1731.3	03/08/14	06/30/18
RG003	0.2	35.38	-82.92	1609.3	03/08/14	06/30/18
RG004	0.2	35.37	-82.99	1921.8	03/08/14	06/30/18
RG005	0.2	35.41	-82.96	1519.7	03/08/14	06/30/18
RG008	0.2	35.38	-82.97	1737.4	03/08/14	06/30/18
RG010	0.2	35.46	-82.95	1478.0	03/08/14	06/30/18
RG011	0.1	35.40	-82.91	1244.0	05/25/17	06/30/18
			INNER			
RG100	0.1	35.59	-83.07	1495.0	03/08/14	05/09/16
RG303PK	1.0	35.59	-83.07	1495.0	03/08/14	05/09/16
RG100T	0.1	35.59	-83.06	1485.0	03/08/14	06/30/18
RG101	0.1	35.58	-83.09	1519.0	03/08/14	06/30/18
RG102	0.1	35.56	-83.10	1635.3	03/08/14	06/30/18
RG103	0.1	35.55	-83.12	1688.3	03/08/14	06/30/18
RG104	0.1	35.55	-83.09	1587.4	03/08/14	06/30/18
RG105	0.1	35.63	-83.04	1344.8	03/08/14	06/30/18

RG106	0.1	35.43	-83.03	1209.8	03/08/14	06/30/18
RG107	0.1	35.57	-82.91	1359.1	03/08/14	06/30/18
RG108	0.1	35.55	-82.99	1276.5	03/08/14	06/30/18
RG109	0.1	35.50	-83.04	1500.2	03/08/14	06/30/18
RG110	0.1	35.55	-83.15	1563.0	03/08/14	06/30/18
RG111	0.1	35.73	-82.95	1393.9	03/08/14	06/30/18
RG112	0.1	35.75	-82.96	1183.8	03/08/14	06/30/18
RG113	0.1	35.89	-82.58	1193.0	03/08/14	05/10/16
RG115	0.1	35.62	-82.56	600.0	03/08/14	05/10/16
RG403	0.1	35.52	-83.10	925.0	03/08/14	05/10/16
WEST						
RG300	1.0	35.73	-83.22	1557.5	03/08/14	06/30/18
RG301	1.0	35.71	-83.26	2002.5	03/08/14	06/30/18
RG302	1.0	35.72	-83.25	1860.5	03/08/14	06/30/18
RG402	0.1	35.72	-83.25	1860.5	03/08/14	05/11/16
RG303S	0.1	35.76	-83.16	1489.6	03/08/14	06/30/18
RG304	1.0	35.67	-83.18	1820.0	03/08/14	06/30/18
RG305	1.0	35.69	-83.13	1630.4	03/08/14	06/30/18
RG306	1.0	35.75	-83.17	1535.9	03/08/14	06/30/18
RG307	1.0	35.65	-83.20	1623.7	03/08/14	06/30/18
RG401	0.1	35.65	-83.20	1623.7	03/08/14	06/30/18
RG308	1.0	35.73	-83.18	1471.0	03/08/14	06/30/18

RG309	1.0	35.68	-83.15	1603.9	03/08/14	06/30/18
RG310	1.0	35.70	-83.12	1756.0	03/08/14	06/30/18
RG400	0.1	35.70	-83.12	1756.0	03/08/14	10/24/15
RG311	1.0	35.76	-83.14	1035.7	03/08/14	06/30/18

Table E 2: Locations of Parsivel disdrometers in the study region.

ID	Latitude [N]	Longitude [W]	Elevation [m]
P1	35.88	-82.58	1188
P2	35.62	-82.57	646
P3	35.59	-83.07	1493
P4	35.56	-83.50	1956
P5	35.69	-83.50	595
P6	35.66	-83.59	634
P7	35.78	-83.21	528
P8	35.80	-82.66	598
P9	35.52	-82.96	794
P10	35.31	-83.20	690
P11	35.76	-82.27	1897
P12	35.08	-82.87	575
P13	35.20	-82.87	863
P14	35.37	-83.51	589
P15	35.44	-83.07	992
P17	35.47	-82.80	1380
P18	35.32	-82.87	1720
P19	35.58	-82.78	954
P20	35.46	-83.11	1860
P21	35.57	-83.02	788

Table E 3: Configuration of rain microphysics model integrated with MPS observations to represent LLCF (Low Level Cloud and Fog)

Simulation	Depth of Fog	Top boundary condition update	Total number of drop-size bins	LLCF Spectra (Number of bins from MPS observations)
Fog – 400 m	400 m	Every 60 s	60	13
No Fog	--------	Every 60 s	60	-----

Appendix F

Four-year climatology of HRRR precipitation

4-year climatology: 2016-2019

Liquid water content
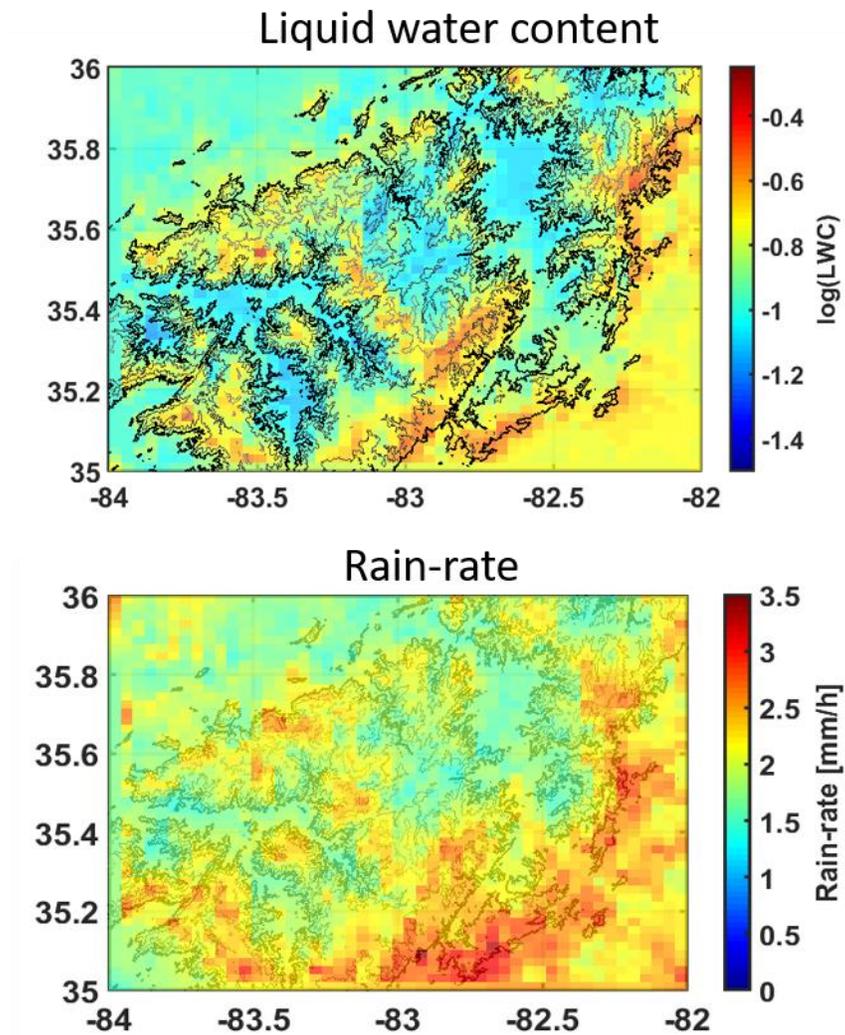
Rain-rate

Figure F1 Four-year climatology of (a) Liquid water content and (b) Rain-rate at 500 m AGL estimated by HRRR in the summer (May-October).
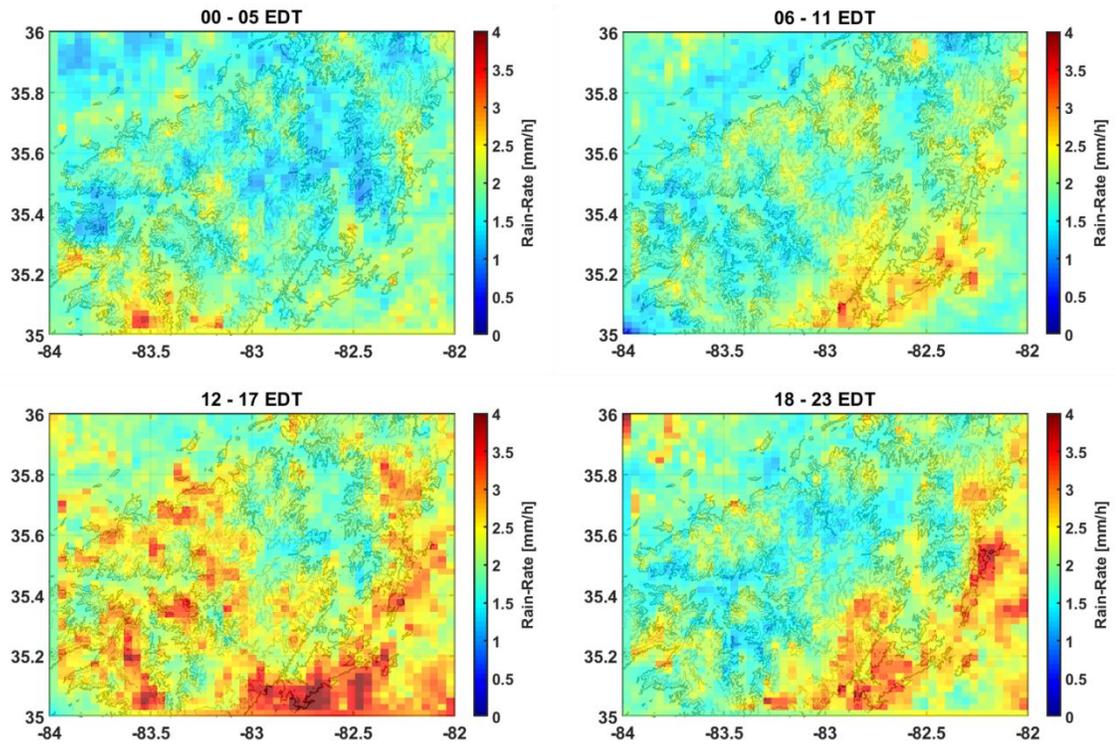
Figure F2 Diurnal cycle of the HRRR rain-rate climatology at 500 AGL in summer (May-October).
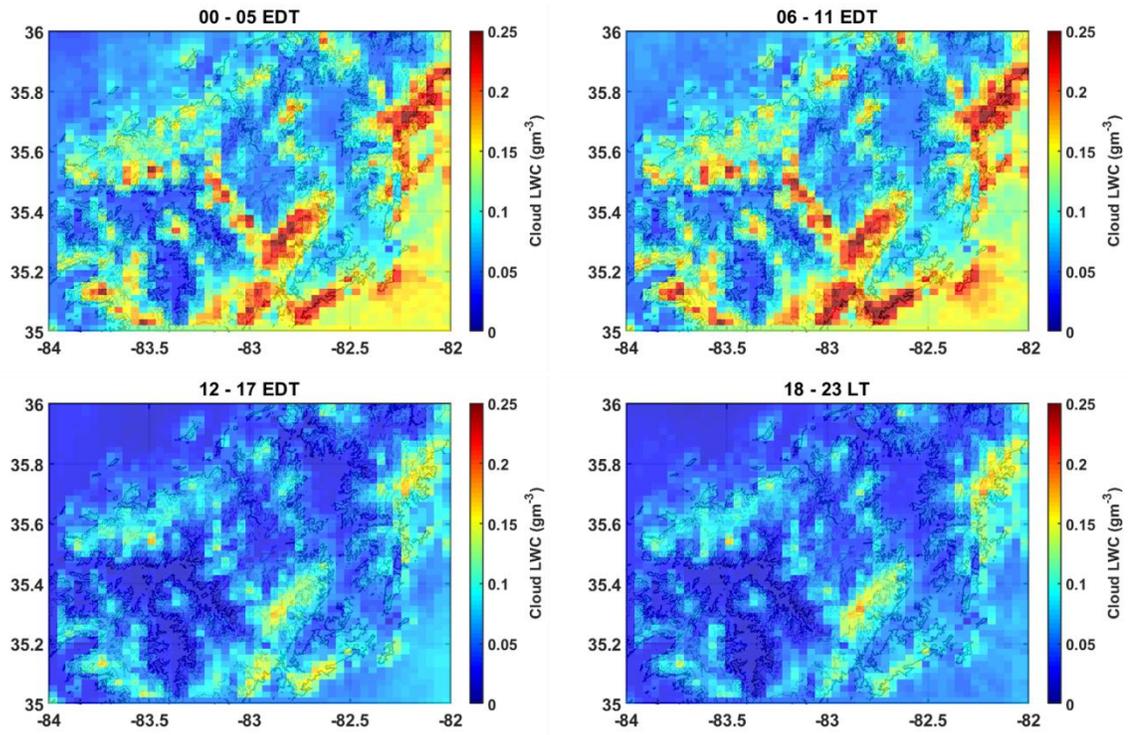
Figure F3 Diurnal cycle 4-year climatology of HRRR estimated cloud-water mixing ratio at 500 m AGL in summer months (May-October).
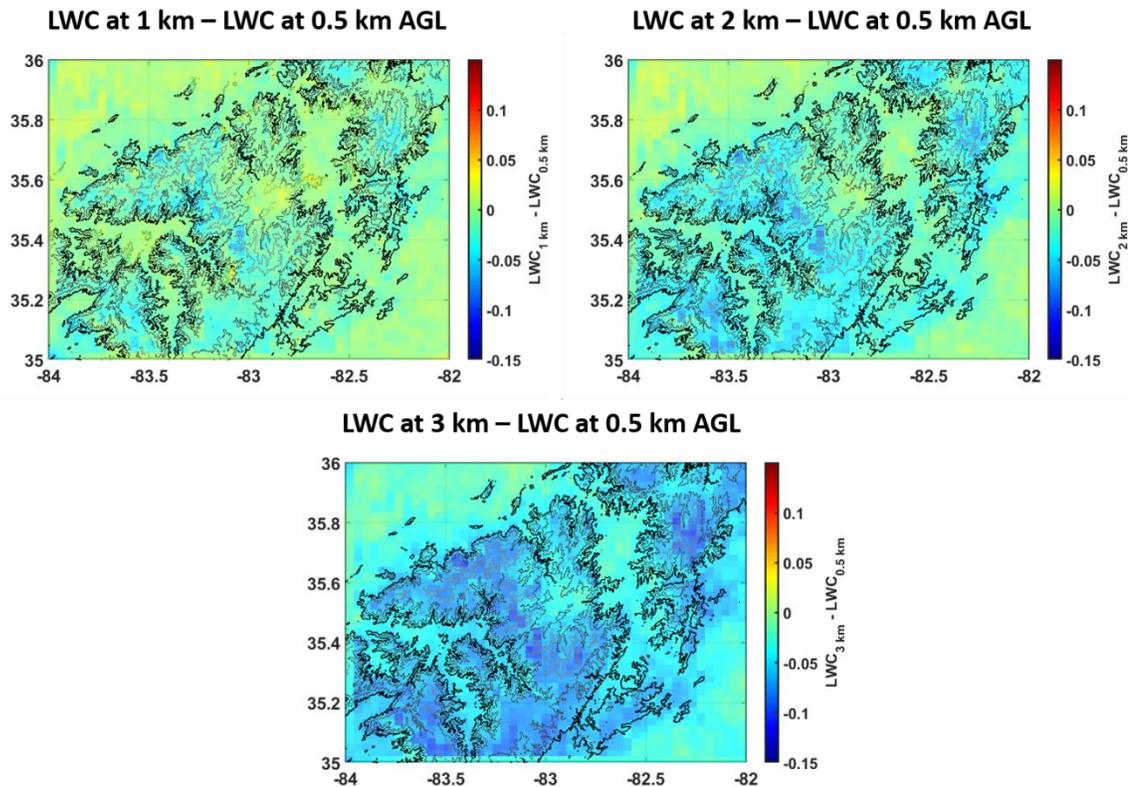
Figure F4 Climatology of showing the spatial distribution of surface enhancement of rain liquid water content. Difference between rain liquid water content at (a) 1 km and 0.5 km; (b) 2 km and 0.5 km and (c) 3 km and 0.5 km AGL. High negative value signifies near-surface enhancement of precipitation that leads to detection and estimation error in GPM DPR.
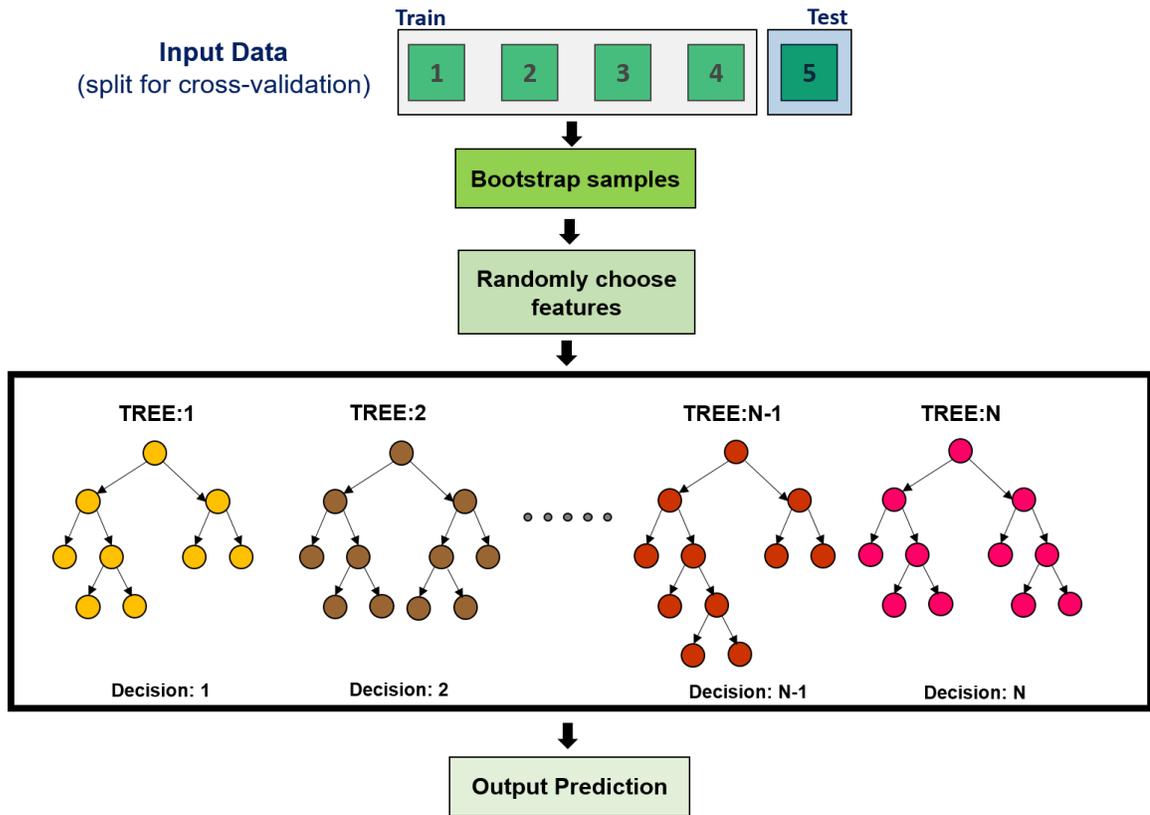
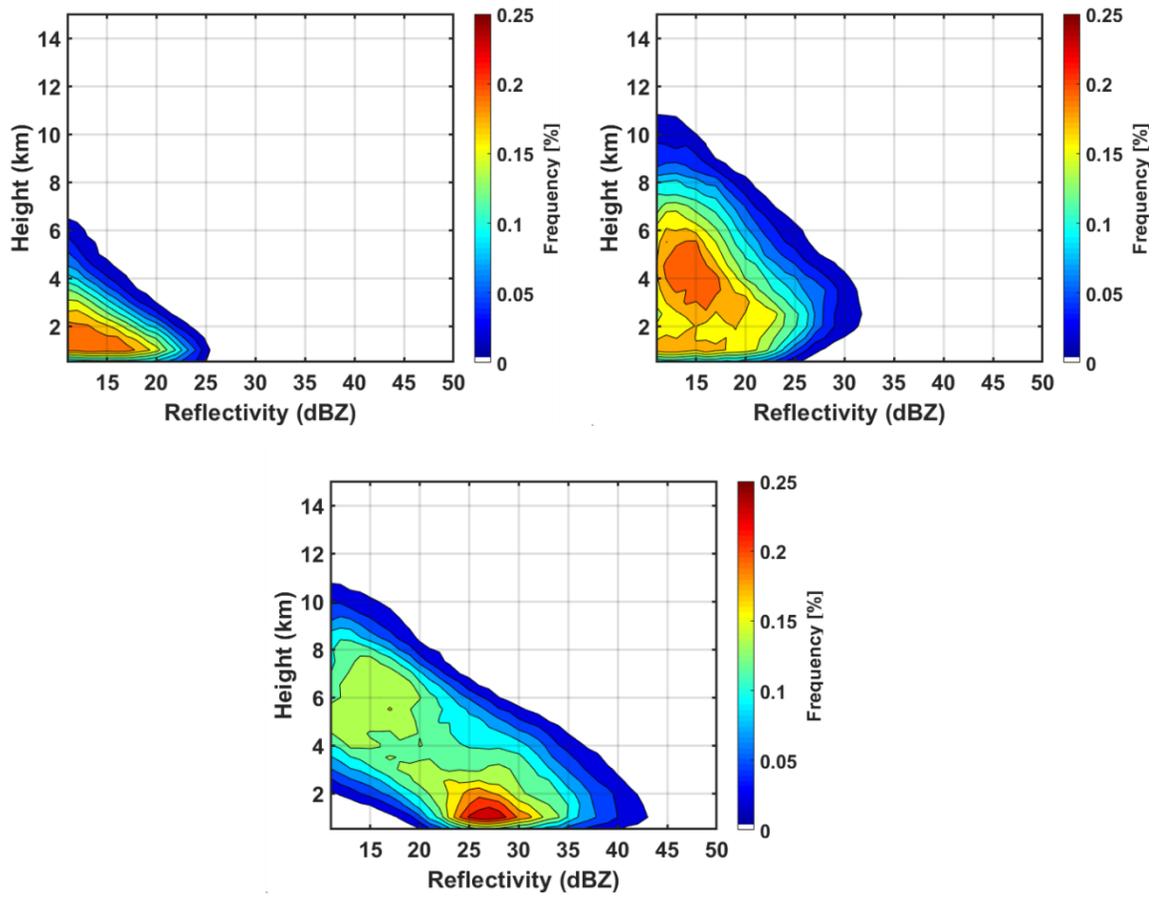**Figure F5: Schematic representation of a random forest classifier used for precipitation detection.**

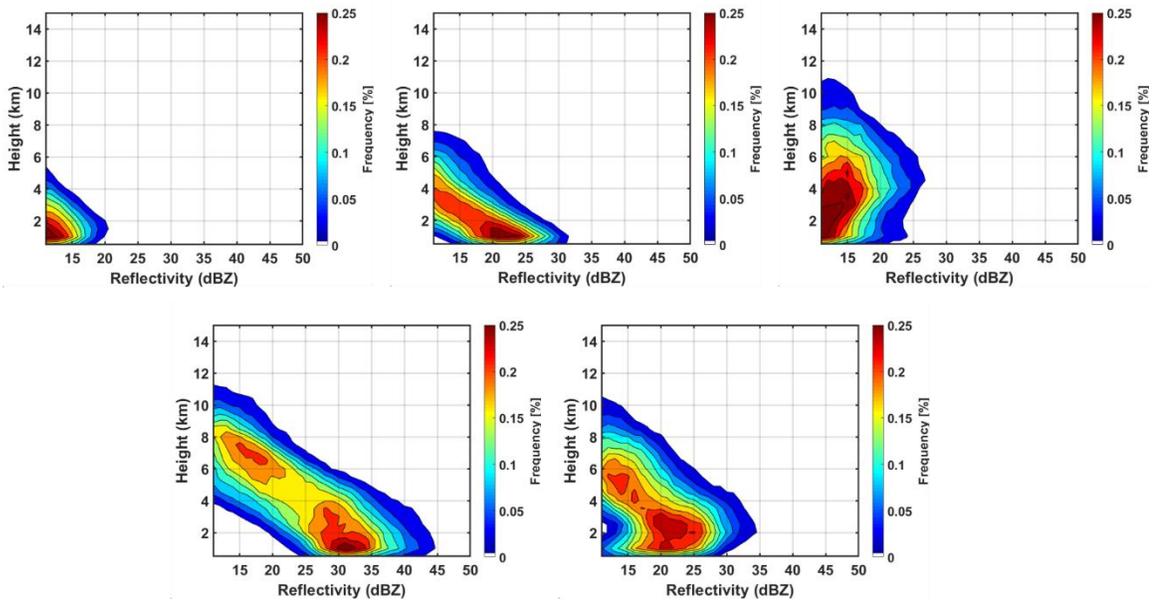**Figure F6: Same as Fig 5-9 but for 3-clusters scenario.**

Figure F7: Same as Fig 5-9 but for 5-clusters scenario.

Biography

Malarvizhi Arulraj received her B.Tech in Electronics and Communication Engineering at Amrita Vishwa Vidyapeetham at Coimbatore, India in 2012. She earned her M. Tech in Climate Science from Indian Institute of Science at Bengaluru in 2014 with her thesis focusing on temporal down-scaling of in-situ Ganges-Brahmaputra river discharge data. She completed her Ph. D. degree in Civil and Environmental Engineering at Duke University in 2019. Her research interests include Hydrology, Radar Meteorology, Precipitation microphysics and Data Science.

Malarvizhi has authored and co-authored three peer-reviewed publications and in the areas of precipitation detection, non-stationarity analysis of baseflow, remote sensing and orographic precipitation microphysics in the Journal of Atmospheric and Oceanic Technology, Environmental Research Letters and Remote Sensing of Environment. She also co-authored a book chapter in Springer which is currently in press. An additional publication from her Ph.D. research is in preparation.

Malarvizhi received a NASA Earth System Science Fellowship to pursue her Ph.D. She also received the departmental Professor Senol Utku Award in 2018 for the best pre-Ph.D. peer-reviewed journal paper for the article published in the Journal of the Atmospheric and Oceanic Technology.