PARAMETERIZING PSD ASSUMPTIONS FOR REMOTE SENSING ALGORITHMS

BY

KIRSTIN JOY HARNOS

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Atmospheric Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

Associate Professor Stephen Nesbitt, Chair
Professor Robert Rauber
Professor Emeritus John Walsh
Adjunct Associate Professor David Kristovich
Abstract

The Global Precipitation Measurement (GPM) mission Core satellite is the next generation of space-based precipitation monitoring. Upgrading the ability to measure light precipitation and snowfall from a nearly global perspective. With these new capabilities comes new levels of uncertainty within the retrieval algorithms, specifically the assumptions associated with snowfall particle size distributions (PSD). In support of the GPM mission, the Ground Validation (GPM-GV) program sponsored field campaigns to collect a comprehensive precipitation dataset utilizing airborne, ground-based, and simulated data for validation and improvement to the GPM retrieval algorithms. As part of GPM-GV, two field campaigns collected data focusing in the high-latitudes (poleward of 45°N): the Light Precipitation Validation Experiment (LPVEx) in Southern Finland from September to December 2010 and the GPM Cold-Season Precipitation Experiment (GCPEx) in Southern Ontario from January to March 2012.

GCPEx utilized aircraft and ground instrumentation to sample snowfall characteristics in Ontario, Canada from January to March 2012. In-situ measurements from the University of North Dakota Citation aircraft and 2-D video disdrometers (2DVD) represent a large dataset of particle size distributions (PSD) from which statistically independent relationships between PSD parameters can be determined utilizing a new framework. This framework, introduced in Williams et al. (2014), determines relationships from the mass spectrum of the PSD and has been shown previously to reduce normalized bias in retrieved rainfall rates. Once PSD relationships are determined for snowfall from GCPEx, the variability is examined using measured environmental parameters of temperature, liquid and ice water content, and
relative humidity. While temperature and water content show organization within the data distributions, application of the environmental influence on the relationship is unlikely to be useful within the GPM algorithm.

Case studies of the 21 September and 20 October 2010 IOPs from LPVEx are performed using in-situ aircraft measurements, ground-based 2D video disdrometers (2DVD), and high-resolution simulations using the Weather Research and Forecasting (WRF) model. WRF simulations for each case use two different microphysical parameterizations: the Goddard 6-class scheme and the WRF single moment 6-class scheme. Simulations of the case studies includes construction of vertical columns using WRF output for comparison to aircraft spirals. Comparisons between observed and WRF simulated data within the vertical columns shows a WRF environment similar to what was sampled by aircraft in terms of temperature, relative humidity, and hydrometeor water content. Particle size distribution assumptions within the WRF microphysical schemes are compared to exponential size distributions from both the aircraft and surface distributed 2DVD measurements. Results shows large differences, some exceeding an order of magnitude, between assumed and measured particle size distribution characteristics, and particle fall speeds.

This project addresses three objectives in support of the GPM satellite retrieval algorithms. A framework is adapted to characterize ice phase PSDs from GCPEX in a statistically independent manner. Variability is then explored within the PSDs using environmental measurements. Finally, case studies from LPVEx focus on whether or not current microphysical assumptions within cloud-resolving model simulations are representative of high-latitude light precipitation.
Acknowledgements

I would like to express my deep appreciation and gratitude to my advisor, Dr. Stephen Nesbitt, for offering me this opportunity, providing much needed support and advise, and taking me on all of those field campaigns. My committee members for taking the time to help shape this project with their feedback. Definitely cannot forget to acknowledge those who took the raw data measurements and created all the data I used, specifically Aaron Bansemer, Ali Tokay, and Kimberly Reed. Also special acknowledgement to Christopher Williams for allowing me to adapt his methods and for answering all of my questions. Funding for this work was provided by NASA PMM Science Team grant NNX13AF86G and GPM Project grant NNX13AJ55G.

On the personal side, I would like to acknowledge my analysis and writing companions, Chase and Oscar. Sometimes it takes kitty snuggles to get keep the motivation going! My parents for continuing unconditional love and advise. My many friends who made me realize that I was never alone during my struggles. Last but not least, thank you to my husband Daniel. You have been a constant source of sympathy, comfort, and support for my entire graduate school tenure. I’m not sure I would have finished without you. Can’t wait to see where our next adventures take us!
# Table of Contents

Chapter 1  Overview and Motivations .................................. 1  
1.1  GPM Retrieval Algorithms .................................. 2  
1.2  GPM Ground Validation .................................. 3  
1.3  Science objectives .................................. 4  

Chapter 2  PSD relationships for mid-latitude ice phase precipitation .. 6  
2.1  Framework .................................. 10  
2.2  GCPEx .................................. 15  
2.3  Results .................................. 17  
2.4  Discussion .................................. 22  
2.5  Figures and Tables .................................. 23  

Chapter 3  PSD parameters and environmental influences ................. 30  
3.1  Results .................................. 32  
3.2  Discussion .................................. 40  
3.3  Figures and Tables .................................. 42  

Chapter 4  Evaluation of WRF microphysics schemes using microphysical measurements during LPVEx .................................. 59  
4.1  Data Sources and Methodology .................................. 62  
4.2  Results .................................. 70  
4.3  Summary and Discussion .................................. 75  
4.4  Figures and Tables .................................. 77  

Chapter 5  Project Conclusions, Implications, and Future Directions .... 88  
5.1  General Summary .................................. 88  
5.2  Implications .................................. 90  
5.3  Future Investigations .................................. 90  

References .................................. 92
Understanding and monitoring precipitation processes introduce a variety of implications on local and global scales including accurate quantification of fresh water supply and monitoring extreme precipitation events. While providing local or even regional measurements of precipitation can be performed using rain gauge networks and ground based radars, consistent global coverage, especially over the oceans, is unrealistic from a surface perspective. Providing global monitoring of precipitation is the motivation behind the Global Precipitation Measurement (GPM; Hou et al. 2014) mission, which represents the next-generation of space based precipitation monitoring. GPM seeks to provide global coverage at 3 hour resolution using a constellation of international and interagency platforms. 28 February 2014 marks the successful launch date of the joint National Aeronautics and Space Administration (NASA) and Japanese Aerospace Exploration Agency (JAXA) GPM Core Satellite. As followup to the successful Tropical Rainfall Measurement Mission (TRMM; Kummerow et al., 1998), the GPM Core Observatory will serve as the reference satellite for the GPM constellation (Hou et al., 2014).

There are two instrument platforms on the GPM Core satellite, each providing an improvement over the TRMM instrumentation. The GPM Microwave Imager (GMI) is a conical scanning passive microwave radiometer with 13 channels ranging from 10 to 183 GHz; the additional higher frequency channels (at 166 and 183 GHz) were included to detect scattering signals from small ice particles in order to estimate light snowfall and rainfall. The Dual-Frequency Precipitation Radar (DPR) consists of a Ku-band radar at 13.6 GHz with a minimum detectable reflectivity of 18 dBZ (or 0.5 mm h\(^{-1}\)) and a Ka-band radar at 35.5
GHz which has a minimum detectable reflectivity of 12 dBZ (or 0.2 mm h\(^{-1}\)). The addition of the active frequency Ka-band represents the new capability to measure snowfall and light precipitation from a spacebased perspective. Further, with both DPR frequencies sampling precipitation in an overlapping swath, quantitative estimates of characteristics of the particle size distribution (PSD), including number concentrations, can be used to improve the associated GPM retrieval algorithms (Hou et al., 2014).

1.1 GPM Retrieval Algorithms

PSDs within GPM algorithms are assumed to follow the gamma distribution (Ulbrich, 1983), which has three unknown mathematical parameters. Details of the gamma distribution and their role in the GPM algorithm will be explored in later chapters. With DPR providing two independent measurements, retrievals of number concentrations can be completed from a two equations with two unknowns scenario assuming a constant value or constraint on the third. There has been extensive work on constraining one of the three gamma parameters both from a ground-based radar perspective (Zhang et al., 2001, 2003, 2006; Moisseev and Chandrasekar, 2007; Bringi and Chandrasekar, 2001) and as part of satellite retrieval algorithms (Seto et al., 2013; Williams et al., 2014; Grecu et al., 2011; Liao et al., 2014). Currently, GPM assumes a constant value for the third parameter (Seto et al., 2013) with recent studies showing success in reducing biases in the algorithms by producing adaptive constraints (Williams et al., 2014). One of the objectives of this project is to expand the current adaptive constraints beyond rainfall and examine snowfall, which is relatively unexplored.
1.2 GPM Ground Validation

Validation of satellite retrieval algorithms is traditionally approached using ground-based measurements as reference. In support of the GPM mission, NASA funded the GPM Ground Validation (GPM-GV) program with a mission to compile comprehensive datasets needed for validation. The majority of these data are collected over a series of field campaigns and national networks (Petersen and Schwaller, 2008). With TRMM focusing on low to mid latitudes (Kummerow et al., 1998), quality and comprehensive datasets are particularly important for regions poleward of 45° due to the lack of available validation data at these higher-latitudes (Leinonen et al., 2012).

Another component of GPM-GV consists of cloud-resolving model simulations. These simulations provide platforms for testing the algorithm assumptions as well as gaining insights into the dynamical processes within precipitating systems (Petersen and Schwaller, 2008). Within the simulations, parameterizations of microphysics, convective precipitation, land-surface processes, and radiation are created to represent processes that cannot be explicitly simulated. More specifically, microphysical parameterizations are constructed in similar fashion to satellite retrievals in terms of the PSD. Most microphysical parameterizations assume the exponential form of the PSD, different from the GPM algorithms; however, results from simulation studies can still be used to determine if PSD assumptions capture the magnitude and variability seen by measurements.

The overarching goal of the GPM-GV program is to produce a comprehensive dataset of airborne and ground-based sampled precipitation types that have previously been difficult to validate via satellite measurements. GPM-GV field campaigns span from warm-to cold-season precipitation across high- to mid-latitudes using dual-polarimetric radars, ground-based disdrometers and rain gauges, and aircraft microphysical probes and radar (Petersen and Schwaller, 2008). This dissertation will present results from two GPM-GV field campaigns that focused on data collected at high-latitudes (poleward of 45°N): the
GPM Cold Season Precipitation Experiment (GCPEx) and the Light Precipitation Validation Experiment (LPVEx). General descriptions of these campaigns follow, with more details found in their respective chapters.

GCPEx collected data in cooperation with Environment Canada near Barrie, Ontario, Canada from 15 January to 1 March 2012, including several aircraft intensive observation periods (IOPs). The main goal of GCPEx is to characterize snowfall using in-situ microphysical data and ground-based remote sensing (Skofronick-Jackson et al., Accepted). LPVEx collected data around Helsinki, Finland from September to October 2010 with several aircraft IOPs. The main goal of LPVEx was to sample light rainfall events in the high-latitudes. The rationale for the field project states that comparisons of rainfall products from CloudSat, Aqua, and TRMM show disagreement in rainfall amounts at latitudes poleward of 30° (L’Ecuyer et al., 2010). These disagreements may result from the majority of rainfall at these higher latitudes occurring in the form of light rainfall with shallow freezing levels, making it difficult for satellite precipitation sensors to capture accurately (L’Ecuyer et al., 2010).

1.3 Science objectives

This project is organized to address three objectives related to the GPM retrieval algorithms. Objectives use aircraft and ground disdrometer data collected from two GPM-GV field campaigns and seek to answer specific science questions within each campaign as well as general GPM algorithm improvement. The document is organized with the following three chapters addressing these specific objectives:

1. Adaptation of a new framework to characterize ice phase PSDs from GCPEx.
2. Investigating variability in the framework.
3. LPVEx case studies as an investigation of how current microphysical assumptions within cloud-resolving model simulations represent high-latitude light precipitation.

A final chapter presents overall conclusions and implications from the results as well as
future directions for this research topic.
Chapter 2

PSD relationships for mid-latitude ice phase precipitation

With the successful launch of the Global Precipitation Measurement mission (GPM) Core Satellite in February 2014 (Hou et al., 2014), the next-generation of space based precipitation measurement has begun. GPM incorporates major improvements over its predecessor the Tropical Rainfall Measurement Mission (TRMM; Kummerow et al. 1998) by including the Dual-wavelength Precipitation Radar (DPR), which boasts both Ka- and Ku-band radar, allowing for improved detection of snowfall and measurements of snowfall rate. The associated GPM retrieval algorithms also show improvements over their TRMM predecessors; however, the DPR algorithms still present uncertainty due to the necessary assumptions of the particle size distributions (PSD; Hou et al. 2014). To mitigate uncertainty in algorithm assumptions, the GPM Ground Validation (GPM-GV) program has collected precipitation measurements over the course of multiple field campaigns focusing on precipitation that has been previously difficult to measure from space (Petersen and Schwaller, 2008). As part of this program, the GPM Cold Season Experiment (GCPEx) primary objective was to collect measurements of mid- to high-latitude snowfall in January-March, 2012 near Barrie, Ontario, Canada (Skofronick-Jackson et al., Accepted).

Snowfall presents more complex uncertainties than rainfall for retrieval algorithms due to the variety of ice crystal habits and orientations. These uncertainties present themselves when quantifying particle size and mass characteristics. Studies have looked at how to quantify the snow PSD and ice particle mass by relating it to the particle maximum unmelted diameter (Brown and Francis, 1995; Heymsfield et al., 2008, 2010). These studies express the mass-diameter ($m(D)$) relationship in the form $m = aD^b$, where $a$ and $b$ are empirically
derived terms. Brown and Francis (1995) derived $m = 0.00294D^{1.9}$ which has been widely cited and used in algorithm development. Heymsfield et al. (2010) shows that while the Brown and Francis relationship works well as an average representation, it fails to capture the dependence of ice particles on temperature and cloud type. Biases in the mass-diameter relationship were documented to propagate and in some cases amplify through the estimates of ice water content (IWC), effective radius, and precipitation rate. Thus, the results of Heymsfield et al. (2010) refine $m(D)$ to four relationships ($m = aD^{2.1}$) with leading coefficient values dependent on precipitation regime and crystal types and a constant value for the exponent that was representative of all cloud types presented (Heymsfield et al., 2010). One of the determined $m(D)$ comes from measurements over the same region as GCPEX and will be used in this study.

As stated in the introductory chapter, PSDs within the GPM algorithms are assumed to follow the gamma distribution first introduced by Ulbrich (1983), which has three unknown mathematical parameters: the intercept ($N_o [\text{mm}^{-1}\mu \text{ m}^{-3}]$), shape ($\mu [\text{unitless}]$), and slope ($\lambda [\text{mm}^{-1}]$) and is represented by the following equation.

$$N(D) = N_o D^\mu e^{-\lambda D} \quad (2.1)$$

where $N(D)$ is particle concentration $[\text{mm}^{-1} \text{ m}^{-3}]$, and $D$ is particle diameter $[\text{mm}]$ (Seto et al., 2013). A well documented caveat with this formulation of the PSD is the statistical dependence of the parameters (Williams et al., 2014; Moisseev and Chandrasekar, 2007). Using the statistical dependence as an advantage, Zhang et al. (2001, 2003) reduce the number of free parameters within the gamma PSD to two by introducing the $\mu - \lambda$ relationship. These studies show improvement in rain rate estimation due to retrieving 2 parameters and estimating the third though the $\mu - \lambda$. However, there have been multiple studies debating whether the statistical dependence of the gamma distribution parameters can be used to discern physical properties and improve rain rate estimation, or if they are only statistical
artifacts because of the initial assumption of the gamma size distribution (Atlas and Ulbrich, 2006; Moisseev and Chandrasekar, 2007; Zhang et al., 2003; Cao and Zhang, 2009).

As a solution for GPM to avoid the statistical dependence of the gamma distribution parameters, the DPR algorithm (Seto et al., 2013) calculates the PSD using a normalized gamma function (Testud et al., 2001) which presents the rain PSD in terms of two physically measurable quantities and the $\mu$ parameter. The physical quantities are the normalized intercept ($N_w \text{ [mm}^{-1}\text{ m}^{-3}]$), which is a function of the liquid water content (LWC) and the mass-weighted mean diameter ($D_m \text{ [mm]}$).

$$N(D; N_w, D_m, \mu) = N_w f(D; D_m, \mu)$$  \hspace{1cm} (2.2)

where

$$f(D; \mu, D_m) = \frac{6(\mu + 4)^{(\mu+4)}}{4^4 \Gamma(\mu + 4)} \left(\frac{D}{D_m}\right)^{\mu} e^{-(\mu+4)\frac{D}{D_m}}$$  \hspace{1cm} (2.3)

$$D_m = \frac{\sum_{D_{\text{max}}}^{D_{\text{min}}} N(D)D^4dD}{\sum_{D_{\text{min}}}^{D_{\text{max}}} N(D)D^3dD}$$  \hspace{1cm} (2.4)

and

$$N_w = \frac{4^4}{\pi \rho_w} \left(\frac{q}{D_m^4}\right)$$  \hspace{1cm} (2.5)

Where $\rho_w$ is the density of liquid water.

With the dual-frequency radar of DPR, algorithms can retrieve two parameters ($N_w$ and $D_m$) given an assumption for the third ($\mu$). Previous studies have shown that retrieval algorithms are sensitive to the $\mu$ assumption and small changes in the value lead to large changes in precipitation estimates. Grecu et al. (2011) uses Eq (2.1) for the PSD and shows that changing the $\mu$ value from 0 to 1 leads to differences of precipitation water content by
values as high as 0.5 g m$^{-3}$. Williams et al. (2014) shows constant $\mu$ values of 0, 3, 5, and 10 produce a mean normalized biases of 20% on average for rain-rate estimates with each $\mu$ value producing a different range of bias values. The current algorithm retrieves $N_w$ and $D_m$ while assuming a constant $\mu$ value of 3 (Seto et al., 2013).

Williams et al. (2014) proposed a method to improve rainfall retrievals for the current DPR algorithm by calculating a statistically independent $\mu - D_m$ constraint for the $\mu$ parameter. Their results show a bias never exceeding 3.5% in rain-rate estimates; much smaller than the cases with a constant $\mu$ which had an average bias of 20%. This study seeks to replicate the Williams et al. (2014) methodology by determining a statistically independent $\mu - D_m$ relationship for ice phase precipitation using aircraft and ground measurements from GCPEX. This chapter is organized to present the methodological framework, including a discussion on incorporating snowfall uncertainty into the framework, followed by an overview of dataset, results, and discussion.
2.1 Framework

The methodology is derived from the equation framework outlined in Williams et al. (2014) which determined a statistically independent $\mu - D_m$ relationship for rainfall utilizing the mass spectrum. To adapt the methods for snowfall, additional steps must be taken to account for the nonuniform density of ice particles. As stated earlier, the $m(D)$ relationship is used to quantify ice particle mass. Because the framework is based on rainfall methodology outlined in Ulbrich (1983), the variability of snow density cannot be incorporated when using the measured diameter, so the $m(D)$ relationship is used to adapt particle measurements to the equivalent melted diameter.

$$D_{melted} = \left( \frac{6m(D)}{\pi \rho_w} \right)^{1/3}$$ (2.6)

with $D$ as the maximum unmelted diameter and $m(D)$ taken from the Heymsfield et al. (2010) relationship derived for snowfall in a similar climate regime over the same region as GCPEX.

Using $D_{melted}$, the framework based on the mass spectrum $w(D)$ can now be represented by the formulation for liquid precipitation

$$w(D) = \frac{\pi}{6} \rho_w N(D) D_{melted}^3$$ (2.7)

The first moment of the mass spectrum is the mass-weighted mean diameter $D_m$ [mm] is determined using Eq. (2.4) and (2.7)

$$D_m = \frac{\sum_{D_{min}}^{D_{max}} w(D)(D_{melted})dD}{\sum_{D_{min}}^{D_{max}} w(D)dD}$$ (2.8)

Mass spectrum variance $\sigma_m^2$ [mm$^2$] or mass spectrum standard deviation $\sigma_m$ [mm] represents the second moment. Physically, $\sigma_m$ represents the width of the PSD mass spectrum.
\[
\sigma_m = \left[ \frac{\sum_{D_{\text{min}}}^{D_{\text{max}}} (D_{\text{melted}} - D_m)^2 w(D)dD}{\sum_{D_{\text{min}}}^{D_{\text{max}}} w(D)dD} \right]^{\frac{1}{2}}
\]  
(2.9)

Substituting equations 2.2 and 2.3 into 2.9 obtains a relatively simple expression for \(\mu\).

\[
\mu = \frac{D_m^2}{\sigma_m^2} - 4
\]  
(2.10)

Which is a relationship first noted in Ulbrich (1983).

While these methods are an improvement over the traditional gamma PSD by using measurable quantities, statistically, \(\sigma_m\) and \(D_m\) are highly correlated (see Williams et al., 2014). Therefore, in order to determine a statistically independent relationship between variables that will be useful for DPR algorithms, a statistically independent parameter \(\sigma_m'\) is introduced

\[
\sigma_m' = \frac{\sigma_m}{D_m^{b_m}}
\]  
(2.11)

where \(b_m\) is adjusted until the Pearson correlation coefficient between \(\sigma_m'\) and \(D_m\) is zero (Haddad et al., 1996; Williams et al., 2014). Eq. 2.11 can now be rearranged into a statistically independent relationship between the mass spectrum parameters, \(\sigma_m\) and \(D_m\).

\[
\sigma_m = \sigma_m' D_m^{b_m}
\]  
(2.12)

Similarly, Eq. 2.12 can be substituted into Eq. 2.10 to determine a statistically independent constraint on \(\mu\) using \(D_m\).}

\[
\mu = \frac{D_m^{(2-2b_m)}}{\sigma_m^2} - 4
\]  
(2.13)

These equations (2.12, 2.13) represent the main focus for the results of this and the following
Now that the statistically independent relationships between the PSD parameters have been established, it is important to investigate uncertainty within this framework specifically due to the adaptation into ice phase precipitation. The additional uncertainty comes from the \( m(D) \) relationship, introduced earlier. For algorithms, it is important to not only quantify sources of uncertainty, but also how uncertainty propagates through the framework. The following section is an investigation into quantifying the propagation of uncertainty within the \( m(D) \) assumption.

### 2.1.1 Error Propagation due to uncertainty in \( m(D) \) relationship

When adapting this framework to ice phase precipitation, a new level of uncertainty is introduced via the \( m(D) \) assumption. While computing the equivalent melted diameter (Eq. 2.6), the \( m(D) \) relationship is calculated and incorporated into the framework using Eq. 2.7. In this section there is a quantitative investigation on how errors in the \( m(D) \) assumption propagates through the \( \sigma_m \) and \( D_m \) calculations.

Using the methods outlined in Kotulski and Szczepinski (2010), the propagation of error can be quantified through the variance (\( \sigma^2 \)) of the desired function. The \( \sigma^2 \) can be approximated using the sum of the partial derivatives of the independent variables within the function and their \( \sigma^2 \). Therefore, the propagation of uncertainty within \( D_m \) (Eq. 2.8) (\( \sigma_{D_m}^2 \)), is found by calculating the partial derivatives of \( D_m \) with respect to \( w(D) \) and the diameter (D).

\[
\sigma_{D_m}^2 = \frac{\partial D_m}{\partial w(D)} \sigma_{w(D)}^2 + \frac{\partial D_m}{\partial D_{melted}} \sigma_{D_{melted}}^2 \tag{2.14}
\]

When expanding \( w(D) \) (Eq. 2.7), \( m(D) \) is the only component that is not constant and is the only component that remains when derivatives are performed. Therefore, Eq. 2.14 becomes:

\[
\sigma_{D_m}^2 = \frac{\partial D_m}{\partial m(D)} \sigma_{m(D)}^2 + \frac{\partial D_m}{\partial D_{melted}} \sigma_{D_{melted}}^2 \tag{2.15}
\]
expanding the derivatives further to assumed $a$ and $b$ values.

$$\sigma_{D_m}^2 = \frac{\partial D_m^2}{\partial a^2} \sigma_a^2 + \frac{\partial D_m^2}{\partial b^2} \sigma_b^2 + \frac{\partial D_m}{\partial D_{melted}} \sigma_{D_{melted}}^2$$  \hspace{1cm} (2.16)

Because $a$ and $D_{melted}$ are constant coefficients, when derivatives are taken, the expression simplifies to:

$$\sigma_{D_m}^2 = \frac{\partial D_m^2}{\partial b^2} \sigma_b^2$$  \hspace{1cm} (2.17)

Interestingly, the only error that propagates is the $m(D)$ exponent value $b$. Again due to the fact that $a$ and $D_{melted}$ are coefficient values. Repeating the process for $\sigma_m$ includes an additional term due to $D_m$ as an independent variable within in the $\sigma_m$ equation (Eq. 2.9).

$$\sigma_{\sigma_m}^2 = \frac{\partial \sigma_m}{\partial \sigma_{w(D)}} \sigma_{w(D)}^2 + \frac{\partial \sigma_m}{\partial \sigma_{D_{melted}}} \sigma_{D_{melted}}^2 + \frac{\partial \sigma_m}{\partial \sigma_{D_m}} \sigma_{D_m}^2$$  \hspace{1cm} (2.18)

Expanding the $w(D)$ expression:

$$\sigma_{\sigma_m}^2 = \frac{\partial \sigma_m}{\partial \sigma_{m(D)}} \sigma_{m(D)}^2 + \frac{\partial \sigma_m}{\partial \sigma_{D_{melted}}} \sigma_{D_{melted}}^2 + \frac{\partial \sigma_m}{\partial \sigma_{D_m}} \sigma_{D_m}^2$$  \hspace{1cm} (2.19)

Expanding the $m(D)$ expression:

$$\sigma_{\sigma_m}^2 = \frac{\partial \sigma_m}{\partial a^2} \sigma_a^2 + \frac{\partial \sigma_m}{\partial b^2} \sigma_b^2 + \frac{\partial \sigma_m}{\partial D_{melted}} \sigma_{D_{melted}}^2 + \frac{\partial \sigma_m}{\partial D_m} \sigma_{D_m}^2$$  \hspace{1cm} (2.20)

Simplifies to

$$\sigma_{\sigma_m}^2 = \frac{\partial \sigma_m}{\partial b^2} \sigma_b^2 + \frac{\partial \sigma_m}{\partial D_m} \sigma_{D_m}^2$$  \hspace{1cm} (2.21)

Recall Heymsfield et al. (2010) determined the exponent value $b = 2.1$ as representative of all snow environments. Assuming their finding is correct for the $m(D)$ relationship, there would be no additional error propagated through the $D_m$ and $\sigma_m$ calculations as a result of
using ice assumptions. However, if there is an error associated with the Heymsfield et al. (2010) exponent assumption, quantification of the error is shown using Figure 2.1, which gives a visualization of the propagation of error. The intermediate steps of simplifying the expressions in Eq. 2.17 and 2.21 are not shown. The software package MATHEMATICA was utilized to determine the $\sigma_{Dm}^2$ and $\sigma_{m}^2$ expressions. Varying $\sigma_b^2$ from 0 to 0.5 determined the relationship between the $\sigma_b^2$ and the propagated $\sigma^2$ values. The relationship between $\sigma_b^2$ and the propagated values turns out to be linear, with $\sigma_{Dm}^2$ amplified by a factor of approximately 1.64$\sigma_b^2$. $\sigma_m$ shows a damping variance by a factor of approximately 0.69$\sigma_b^2$. This confirms that not only can these methods be adapted to ice assumptions, but also the new uncertainty within the m(D) can be accounted for within the algorithm and shows a linear propagation. The following section describes the GCPEx dataset from which this framework is applied to determine the statistically independent relationships between $\sigma_m - D_m$ and $\mu - D_m$. 

14
2.2 GCPEx

GCPEx (Skofronick-Jackson et al., Accepted) occurred during January and February 2012 near Barrie, Ontario as part of GPM-GV. The overall goal was to collect cold season precipitation microphysics measurements to support the GPM snowfall retrieval algorithms. Data was obtained through coordinated research aircraft flights over five ground measurement sites. Ground based and airborne platforms represent a variety of measurements from both an in-situ and remotely sensed perspectives. Figure 2.2 shows the experiment location within the Great Lakes region and includes location for aircraft flight maneuvers (blue boxes) and the five ground sites including the main ground facility at the Environment Canada Center for Atmospheric Research Experiments (CARE; red star).

Table 2.1 outlines the in-situ microphysical data used in this study. The two sampling platforms highlighted are the University of North Dakota Citation aircraft which utilized multiple in-situ probes collecting particle images ranging from 2 $\mu$m to 19.2 mm. PSDs from the Cloud Imaging Probe (CIP) and High Volume Precipitation Spectrometer (HVPS) probes were combined to create the complete GCPEx PSD dataset (Heymsfield et al., 2008). Particles smaller than 100 $\mu$m are ignored due to uncertainties in the probe’s sample area. The transition from CIP to HVPS data occurs at 1000 $\mu$m. Particles that were larger than the maximum sampling diameter of the HPVS were reconstructed and are included in the dataset (Delene, 2011). Data from these probes was processed to yield binned particle concentrations using techniques outlined in Heymsfield et al. (2008). There were also instruments present to measure the atmospheric state variables and aircraft location (temperature, pressure, 3D wind field, altitude, etc.) The other sampling platform used in this study is the 2D Video Disdrometer (2DVD; Schonhuber et al., 2008) at the CARE ground measurement site (Figure 2.2) which collects particle images that were processed into binned particle concentrations with an imposed fall velocity threshold of 4 m s$^{-1}$ (Tokay et al., 2001; Huang et al., 2010).
During GCPEx there were 9 Intensive Observation Periods (IOPs) involving the Citation, summarized in Table 2.2. Sampling by the aircraft captured a variety of winter precipitation with the majority of cases representing synoptically forced snowfall. The non-synoptic snowfall cases include IOPs 2, 5, 7, and 9. More specifically, IOP 2 was the only freezing rain case sampled by the Citation. IOP 5 was the only lake-effect event sampled by the Citation. A few other IOPs were lake-enhanced or became lake-effect after sampling had been completed. On 18 February, IOP 7, was initiated by a surface cyclone with mixed precipitation falling at the ground sites. The last IOP saw surface precipitation with periods of snowfall and mixed precipitation. Another IOP of note is IOP 6 which was timed to sample precipitation during overpasses by the NOAA-18 and 19 satellites. Because of this, the main precipitation event, an upper level feature, was missed. Instead, light snowfall produced by weak forcing was sampled.
2.3 Results

Results are organized to present data over all GCPEx IOPs, followed by a comparison of the individual events, and completed with a comparison of the measured PSDs to the CARE 2DVD observations. Some additional data filtering occurred in the form of removing data at temperatures $> 0^\circ C$ as well as removing times during instrument malfunction not caught during the initial data processing. This filtering removed approximately 15% of the over 22,000 data points. CIP instrument malfunction was detected when initial results showed little to no data at the smallest diameter values.

Figure 2.3 illustrates the $\sigma_m - D_m$ data distribution for the entire GCPEx UND Citation dataset. Panel (a) is the distribution of $\sigma_m - D_m$ with colors displaying the 2D histogram for the data distribution which shows how many data points are located within each $\sigma_m - D_m$ value. The data distribution that is represented by the color contours is shown in panel (b). The density of points are contoured with the indicated color scale. Overall, the shape of the distribution in panel (a) is as expected with smaller $D_m$ values at smaller $\sigma_m$ and as $D_m$ increases so do the values of $\sigma_m$. The largest concentration of data (red pixels) occur in the $D_m$ range of 0 – 0.25 mm. The black lines display the statistically independent relationship between $\sigma_m$ and $D_m$ calculated from Eq. 2.12 (solid) and the statistically independent relationship presented in Williams et al. (2014) (dashed). The two relationships show no similarity for the entire range of $D_m$ with the GCPEx relationship showing larger $\sigma_m$ values for all $D_m$. The different relationships are an expected result due to the differing densities for rain and snowfall. Therefore, in order to apply these methods to the GPM algorithm, different relationships are needed for rainfall, snowfall, and possibly other precipitation types.

Panel (c) of Figure 2.3 tests the goodness of fit for the statistically independent relationship from (a). Plotted are the residuals to the relationship from the GCPEx data, with the colors representing the same 2D histogram as (a). On visual inspection, the relationship
appears to intersect the data near the center of the distribution with near zero values for the largest histogram values. In fact, the relationship splits 62%, 38% with 38% with values greater than the $\sigma_m - D_m$ relationship and 62% with values less than the relationship. Statistically speaking, this demonstrates that the mean and median values of this distribution are different. While the algorithms are interested in the statistically independent relationship, they also require more quantifiable information on the uncertainty. Recall the statistically independent parameter $\sigma'_m$ was used to determine the power law relationship. The next step is to examine this parameter’s data distribution with respect to $D_m$.

Further investigation of how well the $\sigma_m - D_m$ relationship fits the data distribution is shown by displaying the statistically independent parameter $\sigma'_m$ with respect to $D_m$ (Figure 2.4). Figure 2.4(a) is the distribution of the $\sigma'_m$ parameter with respect to $D_m$ with the 2D histogram contoured onto the data (same as Figure 2.3) and shows the majority of data distributed near a value of $\sigma'_m = 0.5$ mm. The mean value of this distribution is $\sigma'_m = 0.61$ which is the leading coefficient value for the statistically independent relationship in Figure 2.3. Panel (b) of Figure 2.4 illustrates the distribution of $\sigma'_m$ (black) with the bounds of the first standard deviation of the observations. Approximately 79% of the distribution falls within the first standard deviation. Noting the data does not follow a Gaussian curve (Gaussian would contain 68% of data within the first standard deviation). This result shows that providing algorithms with the statistically independent constraint based on $\sigma'_m$ and the bounds of the first standard deviation, approximately 79% of ice particles will be represented.

Continuing the examination of the data distributions, Figure 2.5 displays the $\mu - D_m$ distribution using the same setups as Figures 2.3 and 2.4. Panel (a) shows the $\mu - D_m$ distribution has a large range of $\mu$ values ranging from -3 to greater than 20 at small $D_m$. As $D_m$ increases, the range of $\mu$ values decreases with the final range $\mu = -3$ to 2. As noted earlier for the $\sigma_m - D_m$ distribution, largest values on the 2D histogram (red pixels) occur between $D_m = 0$ to 0.25 mm and corresponds to $\mu$ range of -1 to 7 at $D_m = 0.1$ mm to $\mu =
-1 to 1 at $D_m = 0.25$ mm. It should be noted that the current GPM algorithm assumption of $\mu = 3$ is larger than the majority of the $\mu - D_m$ distribution, especially at $D_m$ values greater than 0.3 mm. Similar to Figure 2.3 the solid black line represent the statistically independent relationship between $\mu$ and $D_m$ for the GCPEx UND Citation distribution, calculated from Eq. 2.13. The dashed black line is the relationship for rainfall from the Williams et al. (2014) study, which is does not fall on the presented $\mu - D_m$ distribution, and produces larger magnitudes of $\mu$ for all $D_m$. This confirms the statement above that different relationships are necessary for retrieving rain and snowfall.

Figure 2.5(b) shows the residuals in $\mu$ to the statistically independent relationship from 2.5(a). This relationship splits the data with 58% of the distribution above and 42% below, very similar to Figure 2.3(b) meaning the mean and median of this data distribution are different. Panel (c) is the normalized PDF of the data distribution (black) with the first standard deviation of the $\mu$ distribution (red). The distribution of $\mu$ has a vast majority ($>90\%$) of the distribution falling within the first standard deviation.

Quantifying the bounds of the first standard deviation for algorithm use can be done using information from the data distributions shown above. More specifically, using the knowledge that 79% of the data distribution falls within the first standard deviation of the $\sigma'_m$ parameter (Figure 2.4), two more statistically independent relationships can be calculated for the $\sigma_m - D_m$ and $\mu - D_m$ distributions that represent the bounds of the first standard deviation of $\sigma'_m$ (or 79% of the data). Figure 2.6 displays the same data distributions for $\sigma_m - D_m$ and $\mu - D_m$ as Figures 2.3(a) and 2.5(a), respectively, with the addition of the statistically independent relationships for the first standard deviation of $\sigma'_m$. The leading coefficient is the only value within the relationships to change, since that is the only component dependent on the value of $\sigma'_m$. The exponent represents the value at which $\sigma'_m$ and $D_m$ are statistically independent. These sets of equations represent a mean value and range of uncertainty and compose what is needed within the algorithms to improve upon the current assumptions.

Now that a mean and range of uncertainty has been established for the data distribu-
tion, it is important to provide a deeper investigation into the variability of the $\mu - D_m$ relationship. Beginning with Figure 2.7, which shows the distribution of data for $\mu - D_m$ with the lines representing the $\mu - D_m$ relationship for each of the nine IOPs (as well as the relationship from all GCPEX data). Solid red lines indicate the standard deviation range shown in Figure 2.5c. The inset graph shows the spread of values for the leading coefficient and exponent for the $\mu - D_m$ relationship. The majority of the IOPs fall within the bounds of the first standard deviation for all GCPEX data. The exceptions to this statement are IOPs 2 and 7. Recall from Table 2.2 that these were freezing rain and mixed precipitation, respectively. IOP 6 also shows a different relationship than the rest of the IOPs by producing larger $\mu$ values at $D_m > 0.1$ mm. Overall, the relationship for IOP 6, which falls within the first standard deviation range, shows a much shallower slope for the $\mu - D_m$ relationship with values near $\mu = 4 - 5$. A possible explanation as to the difference in this relationship comes from the fact that the precipitation sampled from this IOP was light and weakly forced, as the main forcing and precipitation had passed through earlier, as stated in the previous section. Chapter 3 will focus more on the variability within these relationships and will look at each IOP in more detail.

2.3.1 Completing the Vertical Column

Since GPM views the atmosphere as a series of vertical columns, it is important to determine if the relationships presented in the previous section are representative of the entire vertical column. To explore this, PSD data from the 2DVD at CARE can be constructed following the same methods as the UND Citation data. Due to instrument issues there are only data available from IOPs 1, 4, 5, and 8, all of which were snow cases. Figure 2.8 shows the (a) $\sigma_m - D_m$ and (b) $\mu - D_m$ distributions with the 2D histograms contoured (same as earlier). The histogram pixels show the largest concentration of data points (red pixels) fall from $D_m$ values of 0 mm to about 0.1 mm, a smaller range than the complete UND Citation data distributions. This is not surprising since there are only four IOPs represented
here, as opposed to the nine IOPs in the entire Citation dataset. For these distributions, the solid black line is the statistically independent relationship calculated from the 2DVD and the dashed line is the relationship from the UND Citation. The red and pink dashed lines represent the bounds for the first standard deviation from Figure 2.6. Overall, these relationships are similar; however, the 2DVD generally shows larger $\sigma_m$ values for all $D_m$ and within the $\mu - D_m$ distribution, the 2DVD relationship is slightly larger than the Citation for all $D_m$ with differences of $> 1 \mu$ for larger $D_m$ and increasing to about 5 for small $D_m$. Some of these differences can be expected due to the lack of sampling at the cold temperatures seen via aircraft. Therefore the cold temperature values seen at small $D_m$ (that resulted in large $\mu$ values) will not be present in the ground sampling. The majority of the 2DVD relationships fall within the bounds of the first standard deviation, the exception to this is for $\mu - D_m$ at $D_m < 0.1$ mm. While these results are promising that a singular relationship and a range of uncertainty can represent the whole column of precipitation, ground observations beyond the single ground site over a wide range of temperatures are needed before conclusions can be made.
2.4 Discussion

Utilizing aircraft data from GCPEx, a statistically independent relationship relating the measurable PSD parameters of \( \sigma_m \) and \( D_m \) to \( \mu \) was developed for snowfall. By using the equivalent melted diameter of the in-situ snowfall measurements, the framework outlined in Williams et al. (2014) was adapted to determine the \( \mu - D_m \) relationship. The snowfall relationships found here were different than the rainfall relationships presented in Williams et al. (2014); this needs to be taken into account when adapted into the algorithm. When adapting this framework to snowfall, a new level of uncertainty was introduced in the form of the \( m(D) = aD^b \) relationship. The propagation of this uncertainty can be quantified and was found to be linear and only dependent on the \( b \) parameter.

The statistically independent relationships were also presented with equations for the bounds of the first standard deviation of the data distribution. The percentage of data contained within those bounds is about 78\%, noting the data distribution does not follow a Gaussian curve. Comparisons of the \( \sigma_m - D_m \) relationship from the Citation to ground 2DVD measurements at the CARE facility show relatively similar relationships, with the 2DVD results mostly falling within the bounds of the standard deviation. Therefore, if the relationships are utilized within the algorithm as a mean plus a range of uncertainty, the relationship derived aloft may be translatable to the surface precipitation. More surface observations are needed to confirm this finding. Finally, a first investigation into the variability of the \( \mu - D_m \) relationship shows the spread of each individual IOP. While most of the IOPs fell within the bounds of the first standard deviation, the freezing rain and one of the mixed precipitation cases fell below the range, demonstrating the need for specific relationships for each precipitation type. The next chapter will continue the investigation of variability of the \( \mu - D_m \) distribution by investigating any environmental influences on the distribution.
## 2.5 Figures and Tables

<table>
<thead>
<tr>
<th>UND Citation Instrument</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cloud Droplet Probe</td>
<td>PSD: 1 – 50 $\mu$m</td>
</tr>
<tr>
<td>2D Cloud Probe</td>
<td>PSD: 30 – 960 $\mu$m</td>
</tr>
<tr>
<td>CIP</td>
<td>PSD: 25 – 1550 $\mu$m</td>
</tr>
<tr>
<td>Cloud Particle Imager</td>
<td>PSD: 2.3 – 2300 $\mu$m</td>
</tr>
<tr>
<td>HVPS-3</td>
<td>PSD: 150 – 19,200 $\mu$m</td>
</tr>
<tr>
<td>King Probe</td>
<td>LWC</td>
</tr>
<tr>
<td>Total Temperature</td>
<td>Temperature</td>
</tr>
<tr>
<td>Laser Hydrometer</td>
<td>Dew/Frost Temperature</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ground Instrumentation</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td>2DVD (CARE only)</td>
<td>PSD: 0.2 – 8.0 mm</td>
</tr>
</tbody>
</table>

Table 2.1: GCPEX Instrumentation

<table>
<thead>
<tr>
<th>IOP</th>
<th>Date</th>
<th>Citation Time [UTC]</th>
<th>Precipitation</th>
<th>Primary Forcing</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Jan. 19</td>
<td>15 – 18</td>
<td>Snow</td>
<td>Surface Synoptic</td>
</tr>
<tr>
<td>2</td>
<td>Jan. 27</td>
<td>01 – 05</td>
<td>Freezing Rain</td>
<td>Surface Synoptic</td>
</tr>
<tr>
<td>3</td>
<td>Jan. 28</td>
<td>16 – 19</td>
<td>Snow</td>
<td>Upper Level Synoptic</td>
</tr>
<tr>
<td>4</td>
<td>Jan. 30-31</td>
<td>23 – 02</td>
<td>Snow</td>
<td>Upper Level Synoptic</td>
</tr>
<tr>
<td>5</td>
<td>Feb. 12-13 (2 flights)</td>
<td>02 – 05 &amp; 03 – 06</td>
<td>Snow</td>
<td>Upper Level Synoptic</td>
</tr>
<tr>
<td>6</td>
<td>Feb. 14</td>
<td>17 – 20</td>
<td>Snow</td>
<td>Lake Effect</td>
</tr>
<tr>
<td>7</td>
<td>Feb. 16</td>
<td>14 – 17</td>
<td>Mixed</td>
<td>Surface Synoptic</td>
</tr>
<tr>
<td>8</td>
<td>Feb. 18</td>
<td>10 – 19</td>
<td>Snow</td>
<td>Surface Synoptic</td>
</tr>
<tr>
<td>9</td>
<td>Feb. 24 (2 flights)</td>
<td>12-15 &amp; 17-19</td>
<td>Snow (Mixed)</td>
<td>Surface Synoptic</td>
</tr>
</tbody>
</table>

Table 2.2: Summary of GCPEX IOPs
Figure 2.1: Propagation of uncertainty within the m(D) relationship. Red line denotes propagation relationship through $D_m$. Blue line represents propagation through $\sigma_m$.

Figure 2.2: Overview of the GCPEX domain. Smaller figure provides reference of where GCPEX domain is located within the Great Lakes region. Blue boxes outline the research aircraft flight boxes. Red markers indicate location of the five ground measurement facilities: CARE (star), Skydive (plus sign), Steamshow (diamond), Mortons (circle), and Huronia (square).
Figure 2.3: Panel (a) shows $\sigma_m - D_m$ distribution for the entire GCPEX Citation dataset. Black lines represent statically independent power relationships from the GCPEX Citation data (solid) and Williams et al. (2014) (dashed). Colors represent the binned histogram of the data with the actual distribution shown in panel (b). This distribution is also represented in the contours in panel (a). (c) The residuals to the $\sigma_m - D_m$ relationship from (a).
Figure 2.4: Panel (a) shows $\sigma_m - D_m$ distribution of data. Color bar contains same information as Figure 2.3. Panel (b) shows the normalized PDF (black) from (a) and the first standard deviation of the distribution (red).
Figure 2.5: Panel (a) shows $\mu - D_m$ distribution of data with statically independent power relationships from the GCPEX Citation data (solid) and Williams et al. (2014). (b) The residuals to the $\mu - D_m$ relationship from (a). Panel (c) shows the normalized PDF from (a) with the first standard deviation of the distribution (red).
Figure 2.6: For the entire GCPEX dataset. Panel (a) shows $\sigma_m - D_m$ distribution of data. Black line represents statistically independent relationship between $\sigma_m$ and $D_m$. Red and pink lines represent the first standard deviation of $\sigma^2_m$ and encapsulate about 79% of the data. Panel (b) is the same as (a) for the $\mu - D_m$ distribution.

Figure 2.7: $\mu - D_m$ data distribution with statistically independent relationships for each individual IOP. Inset graph shows the spread of values for the leading coefficient and exponent within the statistically independent relationship.
Figure 2.8: Panel (a) shows $\sigma_m - D_m$ distribution of data from CARE 2DVD for IOPs 1, 4, 5, and 8. Colors represent the 2D histogram of the 2DVD data distribution. Solid black line represents statistically independent relationship between $\sigma_m$ and $D_m$ for the 2DVD instrument. Dashed black line is the statistically independent relationship between $\sigma_m$ and $D_m$ from the UND Citation for IOPs 1, 4, 5, and 8. Pink and red dashed lines are same range of first standard deviation as Figure 2.6. Panel (b) is the same as (a) for the $\mu - D_m$ distribution.
Chapter 3
PSD parameters and environmental influences

Within this chapter, the variability and organization of the $\sigma_m - D_m$ and $\mu - D_m$ data distributions will be further explored using environmental parameters measured during GCPEX. Temperature, liquid water content (LWC), and relative humidity (RH) represent the environment measured by the Citation during GCPEX (Table 2.1). Ice water content (IWC) is another variable that will be explored. While IWC is not a measured quantity, it can be calculated using Heymsfield et al. (2004):

$$IWC = \int_0^{D_{max}} N(D)m(D)dD$$

where $m(D) = (\pi/6)D^3\rho_e$ and $\rho_e$ is the effective density of the PSD distribution.

The ultimate goal of exploring the variability in the data distributions involves two components. The first involves identifying any discernible organization by the environmental parameters, which is completed by computing the average temperature, LWC, IWC, and RH values for each of the data points within the $\sigma_m - D_m$ and $\mu - D_m$ distributions (Figure 2.6). If organization is present, the thought is to provide refined $D_m - \mu$ relationships that are based on the environment (i.e. different relationships for ranges in temperature, LWC, IWC, or RH). The second part explores the implementation of any environmental constraints on the $D_m - \mu$ relationship. In order to utilize any environmental constraints, the measurements need to be accurately reproduced within a readily available global dataset. This leads the investigation to the incorporation of the Global Forecasting System (GFS; EMC, 2003) and European Centre for Medium-Range Weather Forecasting (ECMWF; Wang and Zeng, 2014)
models. To provide the comparisons of the simulated platforms to the GCPEx measurements, the GFS and ECMWF analyses are averaged using the four analysis times surrounding the IOP sampling. These analyses are also spatially averaged over the GCPEx experimental region (Figure 2.2) from 44° to 45°N and 79° to 80°W. Noting there are some biases when comparing the measurements to the simulated data due to aircraft flight track focusing on precipitating features whereas the simulated environment may not be precipitating.
3.1 Results

3.1.1 General Environmental Organization

A look at the variability within $\sigma_m - D_m - \mu$ distributions shown in the previous chapter begins with Figure 3.1. This figure shows the $\sigma_m - D_m$ (panel (a)) and $\mu - D_m$ (panel (b)) data distribution with contours of altitude. Average altitude is calculated as the average value for the binned histogram counts from the data distributions from the previous chapter. As expected, data with the smallest $D_m$ characterizes the PSDs at higher altitudes, and as $D_m$ increases, there is a general decrease in altitude. Moving forward within the distributions, special attention will be paid to two separate data clusters that deviate from the general environmental organization of the data. The first can be seen in the group of data points showing low attitudes (1-3 km) at small $D_m$. This data cluster presents slightly larger $\sigma_m$ values ($<0.1 \text{ mm}$) at small $D_m$ and smaller $\mu$ values (ranging 0 to -3) than the general data cluster. The second data cluster occurs only in the $\mu - D_m$ distribution at $D_m$ values near 0.2 mm where there is a second spike in the range of $\mu = 5$ to 20. Investigation into the distributions will be undertaken by examining environmental measurements from the Citation aircraft.

Figure 3.2 is the same $\sigma_m - D_m$ configuration as Figure 3.1 with the average temperature contoured for each pixel. The general organization for both panel (a) $\sigma_m - D_m$ and (b) $\mu - D_m$ is as expected with colder temperatures at the higher altitudes at smaller $D_m$ values with temperatures warming with lower altitudes and larger $D_m$. Due to the first isolated data cluster (at small $D_m$) occurring at low altitudes (Figure 3.1), it is no surprise that it shows warmer temperatures of around -15° to 0°C. The second data cluster at $D_m$ near 0.2 mm also occurred at low altitudes so it shows warmer temperatures of -10° to 0°C.

Continuing investigations in the $\mu - D_m$ distribution with LWC measured by the Citation King Probe (Figure 3.3) shows interesting results with respect to the first nonconforming data cluster (at $D_m < 0.2 \text{ mm, } \mu < 0$). The data cluster shows the highest values of LWC.
measured by the aircraft at values from $10^{-0.5}$ to $1 \text{ g m}^{-3}$. These measurements along with
the warm temperatures may indicate a region of mixed-phase particles. Overall, the general
distribution of measured LWC shows some organization with respect to $\mu$ with lower values
of around $10^{-2}$ to $10^{-1.5} \text{ g m}^{-3}$ at $\mu > 3$, this includes the second data cluster at $D_m$ near
0.2 mm. Below $\mu$ of 3 are the larger LWC measurements with values around $10^{-1}$ to $1 \text{ g m}^{-3}$.

Looking at the distribution for calculated IWC (Heymsfield et al., 2004) (Figure 3.4) organization
is with respect to $D_m$ with IWC increasing with increasing $D_m$. Specifically, for $D_m < 0.2 \text{ mm}$ IWC values range from $10^{-2}$ to $10^{-1} \text{ g m}^{-3}$. $D_m$ between 0.2 and 0.3 mm
IWC values range from $10^{-1}$ to $10^{-0.5} \text{ g m}^{-3}$. With the exception of the data cluster spike
at $D_m$ near 0.2 mm which shows IWC values between $10^{-2}$ and $10^{-1.5} \text{ g m}^{-3}$. About half to
one order of magnitude less than the data at $D_m = 0.2$ and $\mu < 5$. For $D_m > 0.3 \text{ mm}$, the
IWC are greatest with values ranging from $10^{-0.5}$ to $>1 \text{ g m}^{-3}$.

RH with respect to ice (Figure 3.5) shows little organization within the $\mu - D_m$ distribution. Values throughout range from approximately 100% to 120% with a few pixels showing
larger values mostly at values near $D_m < 0.2 \text{ mm}$. The special data clusters also do not display
much difference than the overall distribution. However, the data spike at $D_m = 0.2$
mm has a smaller value of RH (100 – 110%) than the similar $\mu$ values at $D_m < 0.15 \text{ mm}$.
Given this distribution, it is unlikely for RH to be useful in further constraining $\mu$.

As a general observation, temperature and LWC were the environmental parameters
showing promise for further constraint on the $\mu$ distribution, with temperature showing
organization with respect to $D_m$ as well as indicating the separate data clusters. LWC
shows organization with respect to $\mu$ with $\mu$ near 3 focusing as the transition from low LWC
to higher LWC values. The data cluster at small $D_m$ and $\mu$ stood out within the LWC
distribution with the highest values, indicating possible melting or particle riming. Next
section will investigate the individual IOPs to further explain and confirm the findings from
the total data distribution.
3.1.2 Individual IOPs

Now that the general organization of the $\mu - D_m$ distribution by environmental variables has been displayed, a deeper investigation begins by examining which specific IOPs contribute to what regions of the data distribution. Figure 3.6 displays the entire $\mu - D_m$ data distribution with IOP number contoured (see Table 2.2 for IOP details). At general inspection, the distribution is quite messy; however, there is some organization present, especially when considering the two special data clusters. The data spike at $D_m = 0.2$ and $\mu > 5$ consists almost entirely of IOP 6, which occurred on 14 February. Recall from the previous chapter that sampling for IOP 6 occurred after the main precipitation event had passed and sampled the weakly forced snowfall timed under a few satellite overpasses. Other IOPs within the data spike include IOP 2, 4, 5, 7, and 9. With the exception of IOP 4, these represent the cases that were not synoptically forced snowfall. IOP 2 was the freezing rain case, IOP 5 was lake-effect snowfall, IOPs 7 and 9 were cases with mixed precipitation.

The other data cluster at $D_m < 0.2$ and $\mu < 0$ is more difficult to determine what IOPs contribute, but it appears to be mainly a combination of data points from IOPs 2, 5, 7, 8, and 9. The following discussion separates the data into the individual IOPs to gain a better picture of what each IOP’s distribution of data looks like.

Figure 3.7 displays the same information as Figure 3.2(b) except separated into the nine individual IOPs. Separation of the data provides some new insights into how the $\mu - D_m$ distribution is organized. IOPs 1, 3, 8, and 9 show the most stratification by temperature, with coldest temperatures at small $D_m$ and increasing as $D_m$ increases. IOP 4 also shows this organization, but has fewer data points in the -30 to -20°C range. Recall from Table 2.2 that IOPs 1, 3, 4, 8, and 9 represent the synoptically forced snowfall cases from GCPEX, noting the IOP 9 was snow at the beginning of sampling and switched to mixed precipitation, so it is not surprising that these cases show the clearest organization by temperature. This figure also shows exactly which IOPs contribute to the data clusters first noted in Figure 3.1. The
data spike at $D_m$ near 0.2 mm, which was composed mostly of IOP 6 with contributions from IOPs 2, 4, 5, 7, and 9, contains the warmest temperatures from -15° to 0°C. Again, with the exception of IOP 4, these cases represent the non-synoptically forced snowfall IOPs. Separating the IOPs also shows which IOPs contribute to the data cluster at small $D_m$ and $\mu$. The majority of the data cluster is from IOP 2, with contributions from IOPs 5, 7, 8, and 9. The temperatures for each of the IOPs near the data cluster ranges from -10° to 0°C. Along with the higher LWC, seen in Figure 3.3, it would suggest that this data cluster is due to frozen drops or riming, especially when considering IOP 2, the freezing rain case, contributes to the majority of this data (as well as the lake-effect case (IOP 5), and both mixed precipitation cases (IOPs 7 and 9)).

Looking at LWC separated by IOP in Figure 3.8 reveals an organization different than what was seen for temperature. The snowfall IOPs (IOP 1, 3, 4, 8, and 9) appear to be organized by $D_m$, with LWC increasing with $D_m$, not by $\mu$ as the entire dataset distribution suggested. The freezing rain, lake-effect, and mixed precipitation cases show no organization with respect to either $D_m$ or $\mu$. Overall, LWC may be valuable in providing a filter for removing mixed phase or rimed particles but probably not as a way to further constrain the $\mu$ relationship.

Relating the individual IOPs back to the UND Citation probe images may provide some further context toward the types of particles sampled within each data distribution. Figures 3.9 – 3.17 display the IOP panels from Figure 3.7 with selected images from two of the in-situ probes, the CIP and HVPS. The images display a representative 10 second image of what was sampled by the probes for the arrow pointed data region. For particles too small to discern from the HVPS image alone, CIP imagery is also provided. Resolution of the images are 150 $\mu$m for HVPS and 25 $\mu$m for CIP. Finally, the width of the image panels, also termed the buffer width, equal the maximum measured diameter for each probe (19.2 mm for HVPS, 1600 $\mu$m for CIP). IOP 1 (Figure 3.9) imagery matches the data distribution with small particles at colder temperatures and larger particles at warmer temperatures.
For synoptically forced snowfall, this is the expected relationship. It also shows the use of the equivalent melted diameter, as opposed to the unmelted maximum particle dimension, within the framework has not altered the expected distribution.

Images from the freezing rain IOP 2 in Figure 3.10 show different particles within the data distribution than IOP 1. The HVPS image within the cooler temperatures (−20°C) shows small ice particles and small aggregates. Moving into the warmer temperatures, specifically the small $D_m - \mu$ data cluster shows small round particles, possibly liquid drops which would be consistent with the LWC values seen in Figure 3.8. At larger $D_m (>0.2$ mm) images HVPS show particles mainly consisting of small aggregates. With this IOP, imagery is consistent with the data distribution.

Figure 3.11 is the synoptically forced snowfall event of IOP 3. Imagery shows similar distribution to IOP 1 (Figure 3.9) with small particles and small aggregates at colder temperatures and larger particles at warmer temperatures. IOP 4 is also a synoptic snowfall case with images shown in Figure 3.12. Probe images do not show the clear distribution as the other snowfall cases. The is generally uniform distribution of small particles and small aggregates similar to what was seen in IOP 1 and 3, but they are present throughout the entire distribution. Some possible explanations for this IOP include the presence of higher liquid water contents which may have contaminated the framework within the use of the equivalent melted diameter. Mission reports also talk about high reflectivity banding occurring during sampling so convective snow growth with supercooled water could contribute to changes in the PSD shape.

IOP 5 (Figure 3.13) is the lake-effect case from GCPEX. Images shown for this IOP are from the small $D_m - \mu$ data cluster and from the main data group at similar temperatures. The separate data cluster has slightly smaller particles than the main data distribution. Overall the data distribution consists of smaller aggregates, similar to IOP 4 with sampling only occurred at temperatures above −20°C. This would contribute to the conclusion that the convective cases have different PSD characteristics, so may need to be treated differently.
within this framework.

Following the lake-effect case is IOP 6 which sampled weakly forced snowfall after the main precipitation event. Figure 3.14 shows temperatures warmer for this case, so it is no surprise that the images show smaller aggregates throughout the distribution similar to the other snowfall IOPs, including the data spike at $D_m = 0.2$ mm which provide no insight into the cause of this data cluster. Overall, images do show slightly larger particles at larger $D_m$, which is expected and contributes to showing the assumptions within the framework do not change the basic data distribution for snowfall.

The warmest case is IOP 7 (Figure 3.15). This case has the majority of data within the separate data cluster at $D_m < 0.2$ and $\mu < 0$. Looking at the images confirms the mixed precipitation with small round particles and small ice particles. The last synoptically forced snowfall case is IOP 8 (Figure 3.16) showed similar distributions to IOPs 1 and 3, with the exception of some data points within the small $D_m - \mu$ cluster. Images are as expected with aggregates at warmer temperatures within the main data distribution and small particles within the separate data cluster.

Finally, IOP 9 (Figure 3.17) images show a combination of the snowfall and mixed precipitation cases, which makes sense since this case was snowfall that switch to mixed precipitation. Images show small ice particles and aggregates at the colder temperatures and small round particles at warmer temperatures. The image at $D_m = 0.4$ mm is a good example of the mixture of particle types sampled within a 15 second window. Larger $D_m$ values also had higher LWC values (Figure 3.8), so the fact that particle images show relatively similar particle sizes at different $D_m$ values may come from contamination of liquid particles.

With the previous figures illustrating how different the freezing rain and mixed precipitation cases are from the snowfall IOPs, it will be important from an algorithm perspective to create different PSD parameter interrelationships for the different precipitation types. This was also shown in the previous chapter in Figure 2.7 where the IOPs 2 and 7 partially fell outside the first standard deviation bounds. Beginning the separation by precipitation
types, Figure 3.18 is similar to Figure 2.5(a) with the IOPs separated by snowfall and non-snowfall cases. Comparing the panels shows the snowfall cases are producing larger $\mu$ values for $D_m < 0.2$ mm and similar $\mu$ for $D_m > 0.2$ mm. This is not surprising given the separate data cluster at small $D_m - \mu$ was associated with higher LWC and the non-snowfall cases. Future directions should include recalculation of the framework for the other precipitation types and filtering of high LWC values within the snowfall relationship.

In summary, the synoptically forced snowfall IOPs show the same general organization with respect to temperature. Variability within the general data distribution comes from the other cases, with the weakly forced snowfall (IOP 6) as the main contributor to the $\mu$ spike at $D_m = 2$ and the freezing rain case (IOP 2) as the main contributor to the data cluster at small $D_m$ and $\mu$. Other contributors to those data clusters include both mixed precipitation IOPs (7 and 9) as well as the lake-effect case (IOP 5). Moving forward, utilization of environmental constraints on $\mu$ will only be useful if they can be implemented on a global scale. With sparse global measurements, the logical platform for implementing environmental constraints is through the use of global model simulations. Ideally, the platform would be cloud-resolving simulations at a global scale, but given the vast computation expense required for that undertaking, it is unrealistic on a GPM operational scale. Therefore, the current global analyses would be the logical choice to provide environmental parameters. The next section will investigate whether these global analyses are reliable in capturing the magnitude and variability in the GCPEx environment as measured by the Citation aircraft.

3.1.3 Global Model Analysis and Environment

The GFS and ECMWF model analyses represent currently available global platforms for simulating environmental parameters for algorithm assumptions. Using these platforms, the environmental parameters available for comparison to the Citation are limited to temperature and RH. As seen in Figure 3.5, there was no organization of the $D_m - \mu$ data distribution, so investigating whether these simulation platforms can be used to further
constrain $\mu$ is limited to temperature only.

Temperature differences between the Citation measurements and the GFS and ECMWF analyses are shown in Figure 3.19 for each IOP. The blue line is the GFS temperature minus the Citation measurement, and the red line is the ECMWF temperature minus the Citation value. Looking at each IOP individually, IOPs 1, 6, and 8 shows the smallest range of difference between the simulated and measured values. The other IOPs show various ranges of differences with neither the GFS or ECMWF showing better temperature profiles than the other. IOP 2 Citation temperature is approximately $1^\circ$ to $3^\circ$C warmer than both GFS and ECMWF at low levels. With this being the freezing rain case, the simulated environments being colder shows that they produced the wrong surface precipitation type. IOP 3 has differences ranging from $0^\circ$C at 500 hPa to approximately $6^\circ$C below 900 hPa. IOP 4 shows differences ranging from $0^\circ$ to $5^\circ$C with temperature differences decreasing from 700 hpa to 500 hpa then increasing again when the simulated temperature becomes larger than the Citation above 500 hpa. IOP 5, the lake-effect case, shows better comparison with the ECMWF than GFS analyses below 900 hpa. The GFS is colder than the measurements by $3^\circ$C near 750 hPa and becomes warmer (about $2^\circ$C) than the measurements below 850 hPa. ECMWF is cooler than the measurements by generally less than $1^\circ$C. Both simulated platforms have differences near $0^\circ$C above 700 hPa and below 850 hPa. Comparisons to IOP 7 show cooler temperatures than measurements for the GFS, with the ECMWF showing differences ranging from $1^\circ$ to $2^\circ$C warmer than the Citation below 800 hpa and colder above 600 hpa. Overall, the simulated platforms capture the magnitude of temperature (within a few $^\circ$C), but do not capture the vertical variability if there is an isothermal or inversion within the vertical profile. Temperature may be useful to discern general $\mu$ behavior but case by case uncertainty may lead to uncertainties in precisely determining $\mu$. In order to implement this data in way useful to constrain $\mu$, future investigations should include quantitative error analyses of the simulated variables to the measurements, and how they may propagate through the GPM algorithms.
3.2 Discussion

Investigations of the organization and variability of the $\sigma_m - D_m$ and $\mu - D_m$ data distributions introduced in the previous chapter reveal a few possible ways to further constrain the $\mu$ PSD parameter. Among the environmental parameters measured by the UND Citation during GCPEx, temperature and LWC show the most promise for discerning organization. The $D_m - \mu$ distributions show temperature organization with temperature increases as $D_m$ increases. LWC organization is less stratified but shows values increasing as $\mu$ decreases, with the main transition to between values occurring near $\mu = 5$.

Breaking the $D_m - \mu$ distribution into individual IOPs reveals precipitation type and forcing are important for determining what the over shape and organization of the distribution. The weak forcing, lake-effect, freezing rain, and mixed precipitation cases accounted for almost all of the separate data clusters first noted in Figure 3.1. These clusters also had the warmest temperatures. The cluster located at small $D_m$ and small $\mu$ is the result of melting or riming, as evidence to the freezing rain and mixed precipitation cases contributing to the data cluster and the warm temperatures and larger LWC values. The other data cluster, at $D_m = 0.2$ mm is more difficult to explain, with the majority of the data coming from IOP 6. This precipitation sampled from this case was weakly forced as the main precipitation event had already occurred, so it may be these data points are produced by warm sector, weakly forced precipitation.

Separation of IOPs show that different precipitation types produce different data distributions. High LWC and liquid particle presence may be contaminating the results; therefore, the snowfall and non-snowfall cases must be separated for use in the algorithms. The non-snowfall cases contribute to the majority of the $D_m < 0.2$ mm, $\mu < 5$ values which were shown to consist mostly of small non-dendritic particles. Overall, the snowfall distribution show the expected results of small ice particles at colder temperatures and small $D_m$ with aggregates and larger particles at larger $D_m$ and warmer temperatures.
While some environmental influences have been documented, it is unlikely to be useful to the algorithm beyond very coarse temperature vertical structure information. Current global simulation platforms present the only logical dataset to provide consistent global environmental variables. It was shown that the only variable within those analyses useful to the $D_m - \mu$ distribution is temperature. Overall, the magnitude of environmental temperature was captured within a few °C, but they missed any vertical variability in terms of isothermal layers and inversions. Further investigation would be required to assess any biases seen in the GFS and ECMWF analyses and how they will affect the $\mu$ constraint.
3.3 Figures and Tables

Figure 3.1: Panel (a) shows the $\sigma_m - D_m$ distribution with the average altitude within each pixel of the 2D histogram from Figure 2.3. (b) same organization as (a) for the $\mu - D_m$ with average altitude within each pixel from Figure 2.5.
Figure 3.2: Panel (a) $\sigma_m - D_m$ distribution with the average temperature within each pixel. (b) $\mu - D_m$ with average temperature within each pixel.

Figure 3.3: $\mu - D_m$ with average LWC within each pixel as measured by the Citation King probe.
Figure 3.4: $\mu - D_m$ with average IWC (Heymsfield et al., 2004) within each pixel.

Figure 3.5: $\mu - D_m$ with average RH with respect to ice within each pixel
Figure 3.6: $\mu - D_m$ distribution of data with IOP number contoured on each data point.
Figure 3.7: Figure 3.2 (b) separated into individual IOP distributions.
Figure 3.8: Figure 3.3 (b) separated into individual IOP distributions.
Figure 3.9: IOP 1 panel from Figure 3.7 with in-situ probe images. Each row of images equates to 5 seconds of sampling time. HVPS resolution is 150 µm with buffer width of 19.2 mm. CIP resolution is 25 µm with a buffer width of 1600 µm.
Figure 3.10: IOP 2 panel from Figure 3.7 with in-situ probe images. Each row of images equates to 5 seconds of sampling time. HVPS resolution is 150 µm with buffer width of 19.2 mm. CIP resolution is 25 µm with a buffer width of 1600 µm.
Figure 3.11: IOP 3 panel from Figure 3.7 with in-situ probe images. Each row of images equates to 5 seconds of sampling time. HVPS resolution is 150 µm with buffer width of 19.2 mm. CIP resolution is 25 µm with a buffer width of 1600 µm.
Figure 3.12: IOP 4 panel from Figure 3.7 with in-situ probe images. Each row of images equates to 5 seconds of sampling time. HVPS resolution is 150 µm with buffer width of 19.2 mm. CIP resolution is 25 µm with a buffer width of 1600 µm.
Figure 3.13: IOP 5 panel from Figure 3.7 with in-situ probe images. Each row of images equates to 5 seconds of sampling time. HVPS resolution is 150 µm with buffer width of 19.2 mm. CIP resolution is 25 µm with a buffer width of 1600 µm.
Figure 3.14: IOP 6 panel from Figure 3.7 with in-situ probe images. Each row of images equates to 5 seconds of sampling time. HVPS resolution is 150 µm with buffer width of 19.2 mm. CIP resolution is 25 µm with a buffer width of 1600 µm.
Figure 3.15: IOP 7 panel from Figure 3.7 with in-situ probe images. Each row of images equates to 5 seconds of sampling time. HVPS resolution is 150 µm with buffer width of 19.2 mm. CIP resolution is 25 µm with a buffer width of 1600 µm.
Figure 3.16: IOP 8 panel from Figure 3.7 with in-situ probe images. Each row of images equates to 5 seconds of sampling time. HVPS resolution is 150 µm with buffer width of 19.2 mm. CIP resolution is 25 µm with a buffer width of 1600 µm.
Figure 3.17: IOP 9 panel from Figure 3.7 with in-situ probe images. Each row of images equates to 5 seconds of sampling time. HVPS resolution is 150 µm with buffer width of 19.2 mm. CIP resolution is 25 µm with a buffer width of 1600 µm.
Figure 3.18: Separation of entire GCPEX database into snow only and mixed/freezing rain cases. Colors represent the 2D histogram and black dashed line shows the $\mu - D_m$ relationship when separating the cases.
Figure 3.19: Comparison of Citation measured temperature to averaged GFS and ECMWF analyses. Blue line is the GFS simulated temperature minus the Citation measurement. Red line is the ECMWF simulated temperature minus the Citation measurement.
Chapter 4

Evaluation of WRF microphysics schemes using microphysical measurements during LPVEx

As the ability to measure global precipitation using space-based instrumentation improves, the need for comprehensive validation datasets also increases. This is the basis for the Global Precipitation Measurement mission Ground Validation (GPM-GV) program (Petersen and Schwaller, 2008). GPM-GV compiles datasets of global precipitation from a series of field campaigns using ground-based and airborne radars, microwave radiometers, ground-based disdrometers, precipitation gauges, airborne in-situ microphysical probes, and cloud-resolving model simulations for use in validating and improving the GPM Core and constellation satellite retrieval algorithms. The GPM field campaigns have and will sample a variety of precipitation types across a variety of meteorological regimes that GPM will sample in mid- and high-latitudes. New high-quality integrated validation datasets are particularly important for algorithm improvement in high-latitude regions poleward of 50° latitude due to a lack of satellite validation data (Leinonen et al., 2012; Swenson, 2010). In fact, accurate satellite retrievals to date are considered to be scarce over high latitudes (Kummerow et al., 2011) and NASA’s Tropical Rainfall Measurement Mission (TRMM) satellite, GPM’s predecessor, has been limited to tropical and subtropical regions. Thus, the legacy retrieval algorithms require substantial development and validation in mid- and high latitude regions, particularly in the cool seasons where low freezing levels and falling snow are prevalent (Hou et al., 2014). While some progress has been made using new technologies, most notably the recently launched CloudSat satellite, the associated algorithms are new and the resulting high-latitude precipitation estimates remain largely unverified (Haynes et al., 2009; Mitrescu et al., 2010; Wood et al., 2015).
The use of high-resolution simulations for satellite validation has become a prevalent data source, especially for such regions with limited observational precipitation data (Matsui et al., 2013; Shi et al., 2010). With increasing computational resources, simulations are now able to provide better parameterizations that can be used to simulate precipitation processes with increasing realism and may be compared with forward computations of remote sensing measurements (e.g., Matsui et al. (2013)). The Weather Research and Forecasting model (WRF) is a well-documented mesoscale simulation platform and is currently being used for both weather forecasting (Janjic, 2003) and research applications (Skamarock et al., 2008). Recent studies have shown the WRF-Advanced Research WRF (ARW) provides reliable high-resolution simulations for use in satellite validation studies at mid- to high-latitudes (Molthan et al., 2010; Shi et al., 2010; Otkin et al., 2007; Jankov et al., 2009; Han et al., 2013; Cassano et al., 2011).

WRF-ARW simulations within GPM-GV have been used previously to provide validation for the model parameterizations. Molthan et al. (2010) and Shi et al. (2010) utilized in-situ aircraft measurements to evaluate model parameterizations capture precipitation processes for the Canadian CloudSat/Cloud – Aerosol Lidar and Infrared Pathfinder Satellite Observations Validation Project (C3VP). They sought to validate how WRF simulates the vertical structure of snowfall and how the microphysical parameterization assumptions compare to calculated precipitation exponential size distribution (SD) parameters. Both studies show that WRF was able to reproduce the general structures of clouds and precipitation as well as capture a comparable amount of ice and liquid water content, but Molthan et al. (2010) showed the assumptions within the microphysical parameterizations tested did not capture the vertical variability measured by the aircraft. This study seeks to use the methodology introduced in Molthan et al. (2010) to examine how well microphysical parameterizations perform at simulating light precipitation at high-latitudes.

As an international effort with the Finnish Meteorological Institute (FMI), University of Helsinki, the NASA CloudSat and GPM-GV projects, the Light Precipitation Validation
Experiment (LPVEx) focused on high-latitude light precipitation around Helsinki, Finland (around 60°N). The main goal of the field experiment was to provide a comprehensive dataset of in-situ airborne and ground-based data for improvement of satellite detection algorithms of high latitude light rainfall (L’Ecuyer et al., 2010). Campaign aircraft operations occurred in September and October 2010 in the vicinity of the Gulf of Finland (Figure 4.1).

Using data collected before and during LPVEx, Leinonen et al. (2012) presents a five-year climatology of rainfall measurements near Helsinki at the Jarvenpää field site using Joss-Waldvogel RD-69 impact disdrometer (JWD, Joss and Waldvogel (1967)). For rainfall-dominated months, usually June-October, between 2006-2010, Leinonen et al. (2012) reveals that rainfall near Helsinki is dominated by small raindrop size and low average rain rate with mean values of rainfall (R), median drop volume diameter ($D_o$), and normalized intercept parameter of the gamma fitted distribution ($N_w$; Bringi and Chandrasekar (2001)) as 1.34 mm h$^{-1}$, 1.02 mm, and 4900 mm$^{-1}$m$^{-3}$ respectively. They also show that the assumption of an exponential SD is valid for drops diameters from 0.32 mm to 3 mm. Leinonen et al. (2012) also provided a comparison of the JWD and 2-D Video Disdrometers (2DVD; Schonhuber et al. (2008)) for a rainfall case during LPVEx on 20 October, 2010 showing the JWD is more sensitive to the smallest drops ($\leq 0.5$mm).

Expanding on current LPVEx analyses, this study begins to examine the exponential size distribution at and above the surface, including ice assumptions. The goal is to provide a comparison of microphysical characteristics for two LPVEx light rainfall case studies using high-resolution WRF simulations, ground, and airborne in-situ measurements. This study is outlined as follows: section 2 describes the data sources and methodology for this study, section 3 presents the results beginning with comparisons of the simulations to aircraft in-situ measurements, followed by comparisons to ground disdrometers. Finally, section 4 provides concluding summary and discussion.
4.1 Data Sources and Methodology

4.1.1 The Light Precipitation Validation Experiment

LPVEx was an international field campaign as part of the NASA GPM-GV program. Data collected during the campaign ranges from a variety of ground and airborne measurements, as well as several collocated satellite overpasses. This study focuses on the in-situ aircraft data collected during spiral aircraft maneuvers and the 2DVD at three ground measurement sites. In-situ aircraft microphysical data was collected by the University of Wyoming King Aircraft using the Cloud Imaging Probe (CIP) (for particles of diameters 25 \(\mu\text{m} - 1550 \mu\text{m}\)) and the 2-D Particle (2DP) probe (diameters of 200 \(\mu\text{m} - 6400 \mu\text{m}\)) capable of measuring particle size diameters and number concentrations. Data from these probes was processed to yield binned particle concentrations, with a range of bin diameters of 0.025 mm to 3 cm. Particles larger than the maximum 2DP diameter had to be reconstructed, techniques of which are outlined in Heymsfield and Parrish (1978). Further information about water contents, both ice (IWC) and liquid (LWC) were calculated from the probe data using Heymsfield et al. (2008). The aircraft concentrations were fit to an exponential PSD,

\[
N(D) = N_0 e^{-\lambda D}
\]  

(4.1)

where the intercept \((N_0)\) and slope \((\lambda)\) parameters were extracted following methods in Molthan et al. (2010). To provide a comparison with the CIP/2DP derived water content values, LWC and total water contents (TWC) were also measured via the Nevzorov hot-wire probe. Nevzorov probe data was processed using techniques outlined in Korolev et al. (1998). IWC from the CIP/2DP probes and TWC from the Nevzorov probes were used to calculate effective bulk density \((\rho_e)\) values above the freezing level. Heymsfield et al. (2004) provides the relationship for estimating \(\rho_e\) using IWC and LWC from aircraft measurements.
To calculate \( \rho_e \), water content is divided by the total volume of equivalent diameter spheres, with diameter size representing center of each bin.

\[
\rho_e = \frac{IWC_{or\,TWC}}{V}
\]  

During LPVEx, ground instrumentation consisting of 2DVD (Schonhuber et al., 2008), rain gauges, and ceilometers were stationed at three ground sites: Harmaja, Jarvenpää, and Emäsaló (Figure 4.1). The Harmaja ground site is stationed on Harmaja island located just south of Helsinki at 60.10°N 24.98°E. Jarvenpää ground site is more inland, located northeast of Helsinki at 60.46°N 25.10°E. Finally, the Emäsaló ground site is located on the coastline 20km east of Helsinki at 60.20°N 25.63°E. As stated earlier, the 2DVD is the ground instrument chosen in this study to characterize the surface precipitation characteristics. The data was processed using techniques outlined in Tokay et al. (2001) to extract binned concentration, fall speed, LWC, and R. From the concentrations, with bin diameters ranging from 0 mm to 10 cm, \( N_o \) and \( \lambda \) were calculated from the exponential SD using the same fitting technique as for the aircraft measurements.

While LPVEx had many successful missions, the 21 September and 20 October intensive observation periods (IOPs) were deemed the best-sampled precipitation events of the campaign and are the IOPs presented in this study. On 21 September, a surface cyclone moved into the region around 02 UTC, with the warm sector producing widespread uniform stratiform rain within the LPVEx experimental region beginning around 06 UTC. Aircraft sampling occurred from 0750 to 1200 UTC with spiral maneuvers located over the Research Vessel (RV) Aranda, from 0835 to 0915 UTC, and under a CloudSat overpass at 1106 UTC with a collocated spiral occurring from 1035 to 1100 UTC, (Figure 4.1). The freezing level for this IOP was observed to be around 2 km. The 20 October IOP sampled widespread light and moderate precipitation east of Helsinki as a warm front located northeast of the LPVEx region slowly backed into area with precipitation beginning around 5 UTC. In-situ
aircraft sampling occurred from 0719 to 1117 UTC and included two aircraft spirals over the Emäsaloground site from 0800 to 0815 UTC and Harmaja ground site from 0910 to 0935 UTC (Figure 4.1). Figure 4.2 displays the timing and location of CloudSat on 21 September. The top panel traces the CloudSat location overlain on the Moderate Resolution Imaging Spectroradiometer (MODIS) blackbody temperature at 11.02 nm. Sampling over Northern Europe began at 1106 UTC with sampling directly over the LPVEx region occurring between 1108 and 1110 UTC. MODIS imagery shows precipitation over the majority of Baltic Sea region, with blackbody temperatures around 220 to 230 K. The bottom CloudSat profile displays a stratiform rain profile while over the LPVEx region with the brightband at the observed freezing level near 2 km. Near the spiral, reflectivity values range from -16 dBZ near cloud top at 9 km to 8 dBZ with small areas of 12 dBZ below the brightband.

The precipitation structures for 21 September are also examined via ground-based radars in figure 3. Utilizing two ground radars, Vantaa PPI scans (Figure 4.3 (a) and (c)) and Kumpula RHI scans (Figure 4.3 (b) and (d)) are shown over the LPVEx region during the aircraft spirals at 09 UTC (Figure 4.3 (a) and (b)) with reflectivity values near 15-20 dBZ over the spiral location and at 1050 UTC (Figure 4.3 (c) and (d)) with lower reflectivity values near 5 dBZ. The brightband is also prevalent in both RHI scans at around 2 km, also seen in the CloudSat profile.

The second LPVEx case on 20 October IOP and sampled widespread light and moderate precipitation east of Helsinki as a warm front located northeast of the LPVEx region slowly backed into area. Precipitation began around 05 UTC with in-situ aircraft sampling from 0719 to 1117 UTC. Aircraft maneuvers included two aircraft spirals. The first spiral occurred over the Emäsaloground site from 08:00 to 08:15 UTC and second spiral occurred over the Harmaja ground site from 0910 to 0935 UTC (Figure 4.1). An examination of the Vantaa RHI (Figure 4.4(a) and (c)) scans shows the precipitation located mainly in the eastern portion of the experiment region. The first spiral (Figure 4.4(a) and (b)) was located in the center of the LPVEx region, outside of the main region of precipitation. Kumpula RHI
(Figure 4.4(b)) scan shows low reflectivity values of -20 dBZ to 5 dBZ, the lowest values of the four spirals presented. The second spiral (Figure 4.4(c) and (d)) was located further west within the main region of precipitation. The Kumpula RHI scan (Figure 4.4(d)) shows reflectivity ranging from 5 dBZ to 25 dBZ. For this IOP, the freezing level and brightband was observed to be lower than 21 September at around 1 km.

4.1.2 WRF

WRF-ARW version 3.2.1 was used as the simulation platform to provide high-resolution simulations of the chosen LPVEx IOPs. The simulation set-up includes a triply nested domain (Figure 4.5) with the outermost domain at 25 km resolution, which encompasses most of Northeastern Europe; the second domain has a 5 km resolution and contains much of Southern Finland and Northern Europe. The innermost domain has a resolution of 1 km centered on the LPVEx experimental domain and encompasses the majority of the Gulf of Finland. Simulations for both IOPs were initialized at 1200 UTC the day before IOP operations and run for 36 hours. Initial and boundary conditions supplied every 6 h from the National Center for Environmental Prediction Global Forecast System analysis at 1° spatial resolution (EMC, 2003). Unless otherwise specified, results of the simulations used in this study are from the 1 km domain.

Within the WRF-ARW model, there are many options for parameterizing atmospheric processes, including for precipitation microphysics. To evaluate WRF across multiple parameterization options, IOP simulations are completed as an ensemble with each member representing one of the two different microphysical parameterizations: the Goddard scheme (Tao et al., 2003) and the WRF Single Moment 6-class (WSM6) scheme (Hong and Lim, 2006). Common to both microphysical parameterization schemes is the use of a set of equations developed by Lin et al. (1983) and Rutledge and Hobbs (1984) to provide SD and density assumptions. Within both microphysical packages there are six classes of hydrometeors that can be represented as mixing ratios: water vapor \( q_v \), cloud water \( q_c \), cloud ice
(q_i), snow (q_s), rain (q_r), and graupel (q_g). Simulations were also completed using the WRF Double Moment 6-class scheme (WDM6) (Lim and Hong, 2010), but are not presented here as the simulations show similar results to the WSM6. The Goddard single moment six-class microphysics scheme (Tao et al., 2003) utilizes an exponential SD with fixed best-fit N_o and effective bulk-density (\( \rho_x \)) based on the fundamental equations from Lin et al. (1983) and Rutledge and Hobbs (1984). The exponential SD for WRF is the same as equation 4.1 except the SD is calculated for each hydrometeor class:

\[
N_x(D) = N_{ox} e^{-\lambda_x D}
\]  

(4.3)

where D is diameter, \( N_{ox} \) is the fixed intercept for a specified hydrometeor class, and \( \lambda_x \) is the slope of the exponential size distribution for a specified hydrometeor class. \( \lambda_x \) is determined using the fixed \( N_{ox}, \rho_x, \) and total mass of the hydrometeor population (\( \rho_d q_x \)):

\[
\lambda_x = \left( \frac{\pi \rho_x N_{ox}}{\rho_d q_x} \right)^{0.25}
\]  

(4.4)

It is important to note that with the \( N_{ox} \) and \( \rho_x \) values set as constants, any variability within \( \lambda_x \) comes from the hydrometeor water content. The Goddard scheme was chosen for use in this study due to previous association with GPM-GV simulation studies (Molthan et al., 2010; Shi et al., 2010) and the processes within this microphysical scheme have been specifically developed for the production rain retrieval algorithms for TRMM (Tao et al., 2003), GPM’s predecessor precipitation measurement mission.

To further compare the assumptions within the Goddard scheme, simulations were also run using the WSM6 scheme (Hong and Lim, 2006). The WSM6 scheme uses the same fixed \( N_{or} \) and \( \rho_r \) assumptions as the Goddard scheme with respect to liquid precipitation, but seeks to improve the approach to ice microphysics. Within WSM6, ice microphysical processes are represented by assuming a temperature-dependent \( N_o \) (Hong and Lim, 2006). Additional parameterizations used in the simulations and their associated references are
As a check of whether the WRF simulations are producing spatially similar precipitation to observations, reflectivity from the 5 km WRF domain is compared to composite radar images provided by the Finnish Meteorological Institute (Figure 4.6). Overall, the spatial extent of precipitation qualitatively agrees for 21 September (top panel) with WRF showing the warm sector precipitation over the LPVEx region. For 20 October, however, the spatial extent of WRF reflectivity shows an eastward bias of precipitation displacement of around 100 km. The mitigation of this bias, which is present for all the WRF schemes on 20 October, is described in the following section.

### 4.1.3 Constructing vertical columns

As with most satellite sampling, GPM DPR views the atmosphere as a vertical column at about a 250 m vertical and 5 km horizontal resolution (Hou et al., 2014). LPVEx aircraft sampling sought to replicate this vertical column viewpoint using spiral maneuvers. For a more complete vertical column of precipitation, aircraft spirals were planned over ground instrumentation to gain collocated measurements at the surface and aloft. This complete sampling scenario occurred for both spirals on 20 October. The 21 September spirals were repositioned to coincide with the CloudSat overpass and data collection on the RV Aranda (Figure 4.1). To further the vertical column perspective and to provide a realistic comparison to aircraft sampling, the ensemble of WRF simulations were constructed to be spatially representative of what was measured via aircraft. Aircraft spirals had a diameter on the order of 10 km; therefore, WRF vertical columns are created by averaging all grid points at each height level within a 25 km² (5 km x 5 km) box surrounding the grid point nearest to the center of the aircraft spiral. However, as stated in the previous section, there is an eastward bias of precipitation on 20 October. To mitigate the bias of precipitation by the WRF ensemble, the averaged WRF grid points were shifted 100 km eastward from the aircraft spiral center.
To determine whether the WRF vertical columns are creating a representative precipitation environment to what was sampled by aircraft, the thermodynamic variables are compared. Figure 4.7 depicts the comparison of environmental variables from WRF to the in-situ aircraft measurements for relative humidity (RH) with respect to ice, with respect to water, and temperature. Figure 8a depicts hydrometeor water content. Simulated temperature is in good agreement with the aircraft measurements, showing maximum differences of 2°C with a root mean square error (RMS) fit percentage of around 75% for all spirals. It should also be noted that for all spirals, the freezing level height was also simulated accurately (2 km for 21 September, 1 km for 20 October). WRF RH has a much lower RMS fit percentage of less than 40% for all spirals. However, simulations show agreement with the vertical trends measured by the aircraft below the freezing level for both cases. The exception to this is in the Emäsallo spiral on 20 October where the simulations are too moist at above freezing temperatures. Above the freezing level aircraft RH tends to decrease with altitude for all spirals. WRF RH vertical trends vary above the freezing level with the WSM6 scheme remaining relatively constant for ice and decreasing for liquid on 21 September, Goddard increasing slightly for both liquid and ice until 4.5 km when they both decrease. 20 October WRF RH remains relatively constant for both ice and liquid and both schemes during the Harmaja Spiral and decreases above the freezing level for the Emäsallo Spiral.

For the aircraft hydrometer total water content measurements, CIP/2DP IWC is used at temperatures below 2°C, with LWC, determined by presence of round particles in the CIP/2DP imagery, calculations used at or above 2°C. Simulated hydrometeor water content has a poor RMS fit percentage of below 30% for all spirals; however, simulated values are within an order of magnitude with aircraft measurements (Figure 4.8a). showing simulated values within an order of magnitude, the exception is with the Emäsallo spiral on 20 October where the Goddard microphysics scheme underestimates the water content above 1.5km. This is probably due to a lack of cloud and precipitation formation within the Goddard scheme from the displacement of precipitation location by WRF for this IOP. To check vari-
ability resulting from the location of the WRF vertical column selection, the 8 surrounding vertical columns were also calculated for hydrometeor water content with results shown in Figure 4.8b. The exception is with the Emäsalö spiral on 20 October where the Goddard microphysics scheme underestimates the water content above 1.5km. Focusing on the variability within the ensemble WRF simulations, the 8 surrounding WRF vertical columns were shown for hydrometeor water content in Figure 8b. Differences in the vertical columns are minimal, less than an order of magnitude thus showing a spatially consistent environment throughout the simulations. Overall, the ensembles of WRF simulations were able to capture a consistent representative thermodynamic environment for what was sampled by the aircraft.

An examination of the hydrometeor types produced within WRF (Figure 4.8c) shows both schemes producing relatively similar mixing ratios of snow above the freezing level for the 21 September Aranda and 20 October Harmaja spirals, with the Goddard scheme producing smaller mixing ratios of cloud ice for all spirals except the 20 October Emäsalö spiral near the freezing level. Within the 20 October Emäsalö spiral the WSM6 scheme is producing higher snow mixing ratios with lesser amounts of cloud ice mixing ratios above the freezing level. The Goddard scheme is producing more snow mixing ratios than cloud ice mixing ratios between 1.5 and 2 km and more cloud ice mixing ratios than snow mixing ratios above 1.75 km. Below the freezing level both schemes are producing similar rain mixing ratio profiles for all spirals with small layers of cloud water mixing ratios. On 21 September there are also cloud water mixing ratios present within the Goddard scheme above the freezing level, while on 21 October cloud water mixing ratios amounts decrease rapidly above the freezing level. Overall, above the freezing level snow has the largest mixing ratio for both schemes with rain and small layer of cloud water dominating below the freezing level.
4.2 Results

4.2.1 Aircraft comparison

As stated earlier, both of the chosen WRF microphysics schemes use a fixed $N_o$ for rain ($N_{or} = 8.0 \times 10^3 \text{ mm}^{-1} \text{ m}^{-3}$). For snow, the Goddard scheme also uses a fixed $N_o$ ($N_{os} = 1.6 \times 10^4 \text{ mm}^{-1} \text{ m}^{-3}$) while WSM6 utilizes a temperature dependent $N_o$ ($N_{os} = 2 \times 10^3 e^{0.12(T - T_o)} \text{ mm}^{-1} \text{ m}^{-3}$, where $T$ is simulated temperature in Kelvin and $T_o = 273.15$K).

Using equation 4.1, $\lambda$ and $N_o$ values were extracted from the CIP/2DP SD. Comparing the WRF assumed $N_o$ and the aircraft $N_o$ (Figure 4.9a) reveals some large differences in the assumption value to what was measured.

21 September aircraft $N_o$ is highly variable with altitude, especially from 1.5 to 3 km with values ranging from $10^2$ to $10^5 \text{ mm}^{-1} \text{ m}^{-3}$. When compared to the WRF rain assumption, the aircraft $N_o$ below the freezing level is generally less than WRF for the Aranda spiral and for the CloudSat spiral between 1 and 2 km. Because the rain assumption is a constant value, it fails to capture any of the vertical variability seen in the measurements. Snow assumptions from the Goddard scheme on the 21 September intersect the measured aircraft values at a few altitudes between 1.5 and 3.5 km, but the majority of the measured values are less than this assumption. The differences between the Goddard snow assumption and measured values reach approximately one order of magnitude in the CloudSat spiral between 3 and 4 km. The WSM6 assumption intersects the aircraft measurements throughout the vertical column for the CloudSat spiral, but the majority of aircraft data above the freezing level is at least one order of magnitude larger than the assumption. The only intersection between the WSM6 assumption and the aircraft data during the Aranda spiral occurs at the top of the spiral near 4.5 km. Otherwise, measurements are around an order of magnitude greater than the assumption value.

20 October aircraft $N_o$ values show less vertical variability than the 21 September case with values generally consistent at $10^3 \text{ mm}^{-1} \text{ m}^{-3}$ below the freezing level and $10^4 \text{ mm}^{-1}$
m$^{-3}$ above. Even with the small vertical variability, the WRF rain assumption does not match the aircraft values below the freezing level. Emäsalo spiral is closest to the rain assumption, but is about half an order of magnitude less than the assumption from the bottom of the spiral to approximately 0.75 km when aircraft is less than the assumption by approximately an order of magnitude. The Harmaja spiral is approximately one order of magnitude less than the WRF rain assumption entirely below the freezing level. Above the freezing level, the Goddard assumption is relatively close to the aircraft measurements especially for the Emäsalo spiral. For the Harmaja spiral, measured values are near the WRF rain assumption, not the Goddard snow assumption. The WSM6 assumption is consistently less than the measurements for both spirals on 20 October by at least an order of magnitude. In total, the No values calculated from the aircraft measurements are not well represented in the WRF assumptions, both in the rain and snow regimes. The exception to this is the Goddard snow assumption during the Emäsalo spiral. Many differences between the measured and simulated values were on an order of magnitude different, thus demonstrating the inaccuracy of the current N$_o$ assumptions for high latitude light rain.

Figure 4.9b shows the similar comparisons as Figure 4.9a for $\lambda$. The aircraft measurements of $\lambda$ show the similar vertical profiles as the N$_o$ profiles with 21 September showing more vertical variability with values ranging from $10^{-1}$ to $10$ mm$^{-1}$. 20 October profiles are relatively consistent at values of $10^{-1}$ mm$^{-1}$. When comparing the aircraft measurements to the WRF assumptions for 21 September the WRF assumptions are larger than measured throughout both spirals. Below the freezing level, the WSM6 assumption is slightly closer to the measurements for the Aranda spiral, but for the CloudSat spiral, the Goddard assumption is better. Above the freezing level, both WRF assumptions intersect the aircraft observations especially the WSM6 assumption for the Aranda spiral, but fail to capture the vertical variability. Also, with the exception of the WSM6 assumption in the Aranda Spiral, measurements are approximately half an order of magnitude less than the WRF snow assumptions. On 20 October, below the freezing level, the Emäsalo spiral measurements
are slightly less than the WRF assumptions. WRF values for the Harmaja spiral below the freezing level are larger than all the measurements. The maximum difference is on an order of magnitude larger than the aircraft. Above the freezing level, the WSM6 best captures the magnitude of the aircraft measurements for both the Emäsalō and Harmaja spirals but are still larger than the measurements. Goddard λ values are similar to the WSM6 for the Emäsalō Spiral and consistently an order of magnitude larger for the Harmaja spiral. Since the λ calculation is dependent on $N_o$ and hydrometeor water content (Figure 4.8a), differences in assumptions of magnitude and vertical variability will propagate through the λ difference between the WRF and aircraft values.

Aircraft $\rho_e$ calculations utilize the IWC derived from the CIP/2DP probes and TWC measured by the Nevzorov probe for temperature measurements less than 2°C. Both WRF microphysics schemes represent $\rho$ as constant values for both snow and rain ($\rho_s = 100 kg m^{-3}$ and $\rho_r = 1000 kg m^{-3}$). Figure 4.10 shows the calculated aircraft effective densities display little vertical variability above the freezing level. The WRF assumptions show larger densities for ice than what was sampled by the aircraft by roughly less than an order of magnitude. There is more vertical variation in the Nevzorov measurements where the values approach the WRF assumptions near the freezing level.

## 4.2.2 2DVD comparison

To complete the vertical column perspective, $N_o$ and λ were calculated for the 2DVD measurements using the same methods as the aircraft. Since the 2DVD is a ground-based instrument, comparisons to WRF output are in the form of time series that span the length of the IOP period (735-1130 UTC for 21 September and 705-1130 UTC for 20 October). As shown in Figure 4.1, there were three ground sites where 2DVD measurements were collected: Emäsalō, Jarvenpää, and Harmaja. For the 21 September case, the Harmaja ground site will be ignored due instrumentation issues during the IOP period. Surface precipitation for both cases is light to moderate rainfall; so only the liquid water assumptions will be presented
from WRF. Comparing the exponential SD parameters (Figure 4.11), $N_o$ values calculated at the ground sites show values around two orders of magnitude smaller than the fixed $N_o$ WRF value ($8.0 \times 10^3 \text{mm}^{-1} \text{m}^{-3}$). This large difference in $N_o$ is larger than any differences found when comparing the WRF assumptions to the aircraft measurements. It is important to note that the 2DVD underestimates drops less than 0.5 mm (Tokay et al., 2013). This contributes to the large difference in values when comparing 2DVD measurements to WRF assumptions since an underestimation of small drops leads to a smaller $N_o$ value. However, this would not account for two orders of magnitude difference between the assumptions and measurements. $\lambda$ values are also different than the aircraft measurements (Figure 4.11). For all three ground sites on both IOP days the values from WRF match quite closely in magnitude and variability to the measurements. The maximum difference of approximately $10^{0.2} \text{mm}^{-1}$ is between the WSM6 assumption and measurements on 21 September at the Jarvenpää site from 7:35 to 9:35 UTC. As stated earlier, calculations of $\lambda$ are dependent not only on $N_o$, but also hydrometeor water content which may provide insights into why the WRF estimates of $\lambda$ were closely matched to the 2DVD measurements.

A comparison of the WRF LWC to the 2DVD measurements shows that unlike the SD variables, WRF captures the range of values measured by the instrument (Figure 4.12), ranging from less than 0.01 g m$^{-3}$ at Harmaja on 20 October to 0.22 g m$^{-3}$ at Jarvenpää on 21 September with the WSM6 scheme capturing the variability during that IOP at the Jarvenpää site from 845 to 945 UTC. There are a few other time periods where the general trends in LWC is represented in the WRF schemes, including the Goddard increasing LWC at both Emäsalo and Jarvenpää on 21 September from 1030 to 1130 UTC and the Goddard scheme on 20 October for all three sites. The fall speeds and R-values produced by WRF are largely different from what was measured by the 2DVD (Figure 4.12). 2DVD fall speeds were measured around 2-3 m s$^{-1}$ with WRF producing values less than 1 m s$^{-1}$. Similarly the R measured by the 2DVD are greater than what is produced by WRF, with WRF values just above zero for each site and 2DVD values as high as 2.5 mm hr$^{-1}$ on 21 September at the
Jarvenpää site (Figure 4.12). Based on these results, it is not surprising the SD parameters differ, as it would seem both WRF microphysics schemes are assuming a large number of small drops which would reflect the lower fall speeds and rain rates. Instead the 2DVD and aircraft probes are capturing smaller numbers of larger drops, creating the larger fall speeds and R. This discrepancy can be seen in images from the aircraft probes that show aggregates remain as larger drop sizes when melted (Figure 4.13) not the small particles assumed in WRF for liquid precipitation.

Comparing these results with the climatology values in Leinonen et al. (2012) ($R = 1.34$ mm h$^{-1}$ and $N_w = 4900$ mm$^{-1}$ m$^{-3}$) is about half of the WRF assumed value of $N_{or} = 8.0 \times 10^3$ mm$^{-1}$ m$^{-3}$ and an order of magnitude larger than what was measured by the 2DVD. It should be noted that the mean $D_o$ for these two IOPs was 0.82 mm, which is less than the Leinonen et al. (2012) climatology value of $D_o$ of 1.02 mm. WRF R-values never approach the climatology value reported while the 2DVD R show similar values to climatology on 21 September at the Jarvenpää ground site, otherwise reported values are lower than the climatology value.
4.3 Summary and Discussion

WRF simulations of light rainfall events on 21 September and 20 October, 2010 during the LPVEx field campaign were performed to provide a platform for testing microphysical assumptions for use in satellite algorithm validation. Simulations were run using the Goddard single moment six-class and the WSM6 microphysical schemes and were compared to in-situ aircraft and ground-based 2DVD measurements. Results from the simulations show both schemes produce a reasonable environment in terms of temperature, RH (with respect to both liquid water and ice), and hydrometeor water content, but had some difficulty accurately capturing both the vertical variability and magnitudes of the exponential SD parameters.

Fixed $N_{or}$ values within both schemes provided overestimates on the order of one to two magnitudes of what was measured from both the aircraft and ground-based perspective. The fixed $N_{os}$ from the Goddard scheme generally overestimated the aircraft measurements by as much as an order of magnitude. The temperature dependent $N_{os}$ assumption within WSM6 provided the closest estimate of aircraft $N_o$ by intersecting the aircraft measurements throughout the vertical column; however, the majority of aircraft measurements were larger than what WSM6 estimated. $\lambda$ also had displacements between the WRF schemes and aircraft measurements, specifically below the freezing level. Above the freezing level WSM6 better captured the magnitude of $\lambda$ values but was unable to account for the vertical variability of the measurements. The Goddard scheme was generally less than an order of magnitude different than the aircraft values, and also failed to capture the vertical variability. The use of fixed bulk density values ($\rho_s = 100 \text{kg m}^{-3}$ and $\rho_r = 1000 \text{kg m}^{-3}$) seem to produce an accurate representation of vertical variability but were displaced, generally less than an order of magnitude, than what was measured. The WRF assumed $N_{or}$ are even more displaced when compared to the ground-based 2DVD instrument at all three ground sites. Showing at least two orders of magnitude difference between the simulation and mea-
measurements. This large difference seems to be caused by WRF assuming liquid precipitation is mainly by small drops, where the aircraft images shows precipitation above the freezing level contains large aggregates, leading to drop sizes larger than assumed in WRF. λ 2DVD measurements and assumptions showed the closet comparisons in terms of both magnitude and variability. Overall, the assumptions currently being used do not depict an entirely accurate representation for high-latitude light precipitation both in terms of magnitudes of PSD characteristics and their vertical variability. This is probably due to the fact that precipitation in this region is dominated by low freezing level, light stratiform precipitation events with relatively large snow aggregates and drops. Future work will investigate the ramifications of these assumptions on spaceborne retrieval algorithms.
## 4.4 Figures and Tables

<table>
<thead>
<tr>
<th>Physical Process</th>
<th>Parameterization Scheme</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planetary Boundary Layer</td>
<td>YSU</td>
<td>Hong et al. (2006)</td>
</tr>
<tr>
<td>Radiation (Longwave and Shortwave)</td>
<td>Updated Rapid Radiative Transfer</td>
<td>Iacono et al. (2008)</td>
</tr>
<tr>
<td>Land-Surface</td>
<td>Noah Land Surface</td>
<td>Ek et al. (2003)</td>
</tr>
<tr>
<td>Cumulus (25 km domain only)</td>
<td>Kain-Frisch</td>
<td>Kain (2004)</td>
</tr>
<tr>
<td>Cloud Microphysics</td>
<td>Goddard Scheme</td>
<td>Tao et al. (2003)</td>
</tr>
<tr>
<td></td>
<td>WSM6</td>
<td>Hong and Lim (2006)</td>
</tr>
<tr>
<td></td>
<td>WDM6</td>
<td>Lim and Hong (2010)</td>
</tr>
</tbody>
</table>

Table 4.1: Configuration of the Advanced Research WRF, version 3.2.1 for the LPVEx IOP simulations.
Figure 4.1: Overview of LPVEx aircraft spiral locations and ground sites. Helsinki Finland is labeled for geographic reference in future figures.

Figure 4.2: CloudSat overpass for 21 September IOP. Top panel displays CloudSat trajectory (red line) with timestamps and MODIS swath for the 11.02 nm Blackbody Temperature. Bottom panel shows the radar reflectivity factor (dBZe) for the CloudSat profile from 11:07:30 to 11:09:30 UTC.
Figure 4.3: For 21 September 2010, Vantaa radar 0.3 elevation PPI scan of radar reflectivity (dBZ) at 09:00 and 10:50 UTC (panel a and c respectively). From the Kumpula radar, RHI scans near the center of the RV Aranda and CloudSat aircraft spiral from 09:01 and 10:52 UTC (panels b and d respectively). The track of the University of Wyoming King Air in the plane of the PPI and RHI is plotted on all panels in red.
Figure 4.4: For 20 October 2010, Kerava radar 0.5 elevation PPI scan of radar reflectivity (dBZ) at 08:07 and 9:15 UTC (panel a and c respectively). From the Kumpula radar, RHI scans near the center of the Emäsaló and Harmaja aircraft spirals from 08:03 and 09:16 UTC (panels b and d respectively). The track of the University of Wyoming King Air in the plane of the PPI and RHI is plotted on all panels in red.

Figure 4.5: Domains for LPVEx WRF Simulations
Figure 4.6: Comparison of WRF reflectivity using Goddard microphysics within the 5km domain (a,c) to observed reflectivity from Finnish Meteorological Institute (b, d) for 21 September 2010 09 UTC (a,b) and 20 October 2010 10 UTC (c,d). Helsinki, Finland is marked with a red star.
Figure 4.7: Comparisons of model vertical columns to aircraft spirals for (a) Relative Humidity with respect to ice (dashed) and water (solid) (b) Temperature
Figure 4.8: (a) Hydrometeor water content. (b) Comparing 8 WRF vertical columns surrounding original WRF column from (a). Original column is solid line. Surrounding columns are dashed lines. (c) WRF hydrometeor types for each spiral.
Figure 4.9: Comparison of model vertical columns to aircraft spirals for the exponential SD (a) \( N_o \) and (b) \( \lambda \) for snow (solid) and rain (dash).

Figure 4.10: Comparison of \( \rho \) from aircraft spirals (when \( T \leq 2^\circ C \)) to WRF assumptions of \( \rho_s \) and \( \rho_r \). Nevzorov \( \rho_s \) calculated using TWC.
Figure 4.11: Time series of comparison of 2DVD exponential size distribution parameters to WRF assumptions for (a) Intercept parameter ($N_o$) and (b) Slope parameter ($\lambda$).
Figure 4.12: Comparison of LWC, Fall Speed, and Rain Rate from 2DVD measurements to WRF values for the Jarvenpää, Emäsalo, and Harmaja (20 October only) ground sites.
Figure 4.13: Sample particle images from the CIP probe during Aranda spiral on 21 September and Harmaja spiral on 20 October. Images selected are representative of particle behaviors throughout aircraft sampling.
5.1 General Summary

This project addressed three objectives for improving GPM retrieval algorithms using GPM-GV data. The first two look at adapting a new framework to characterize ice phase PSDs and their variability. The third objective deals with case studies of high latitude light rainfall. Data comes from two field campaigns that focused on higher latitude precipitation. GCPEX sampled mid to high-latitude snowfall in Ontario, Canada in January – March, 2012. LPVEX is the other campaign presented and sampled high-latitude light rainfall around the Gulf of Finland during September – October, 2010.

GCPEX sampled PSDs through aircraft IOPs that encompassed a variety of precipitation types and forcing. While the majority of cases were synoptically forced snowfall, there was one freezing rain case, two mixed precipitation cases, and one lake-effect IOP. Using this data, statistically independent relationships between PSD parameters of $\sigma_m, D_m$, and $\mu$ were determined utilizing a new framework, introduced by Williams et al. (2014). In order to adapt this framework to snowfall, the equivalent melted diameter is used due to the framework’s origins within the rainfall regime. Quantifying the uncertainty within the framework due to the diameter assumptions was accomplished by calculating how errors in the $m(D)$ relationship propagated through the framework. Results show a linear propagation with uncertainty only dependent on the exponent value of $m(D) = aD^b$.

The statistically independent relationships for GCPEX were found to be different than the rainfall estimate presented in Williams et al. (2014) for both the $\sigma_m - D_m$ and $\mu - D_m$.
data distributions. Based on this result, it can be concluded that separate relationships are necessary for the rain and ice phase precipitation. Further error quantification is based on providing the bounds of the first standard deviation of the statistically independent parameter $\sigma'_m$. It was shown that approximately 79% of the data distribution falls within the bounds. The IOPs that fell outside of the first standard deviation represent mixed and freezing rain cases. Separation of the IOPs also show the mixed and freezing rain cases with higher LWC and contribute to the majority of the data at small $D_m - \mu$. The use of the equivalent melted diameter within the framework may be an issue within the non-snowfall cases and demonstrates the need for separate relationships for each precipitation types. Surface 2DVD measurements at the CARE facility mostly fell within the bounds of the first standard deviation. However, the lack of measurements at cold temperatures lead to the surface relationship to fall outside of the bounds at $D_m < 0.1\text{mm}$. Therefore, more ground measurements are necessary before any conclusions can be made about whether the $D_m - \mu$ relationship derived from the aircraft is representative of the entire vertical column of precipitation. However, since most moderate to heavy snow falls above a median diameter of this value, the statistics presented here can be a valuable constraint for the retrievals under these conditions.

Variability within the data distributions is examined using measured environmental parameters of temperature, liquid and ice water content, and relative humidity. While temperature and ice water content show organization within the data distributions, the current global analyses from the GFS and ECMWF were unable to capture the variability in temperature seen from the aircraft. Therefore, application of the environmental influence on the relationship is unlikely to be useful within the GPM algorithm beyond broad changes in the temperature profile with height. Structures such as fronts are not captured by these analyses.

The 21 September and 20 October LPVEx IOPs demonstrate that current microphysical assumptions within WRF do not represent the measurements of PSD parameters. Aircraft
environment was simulated reasonably well with respect to temperature, RH, and hydrometeor water content. Differences arise when comparing the measure and assumed exponential SD parameters. Results shows large differences, some exceeding an order of magnitude, between assumed and aircraft measured PSD characteristics. The differences are even larger when compared to the 2DVD measurements at the three ground sites and also include differences in particle fall speeds between measurements and simulation results. With surface measurements of LWC reasonable, the conclusion is that the microphysics parameterizations do not capture high-latitude light rainfall processes, which are shown to be caused by large aggregates, not small particles as assumed by WRF.

5.2 Implications

The work presented in this dissertation have a wide range of implications that are focused on improving the GPM algorithms to improve GPM monitoring. Some examples include improved forecasting as both the European and American forecast models currently incorporate microwave-based rainfall information. With GPM’s new microwave snowfall measurements, operational models can now incorporate snowfall information that will lead to improved snowfall forecasts. As stating earlier, GPM is continuing the legacy of TRMM which has been providing tropical precipitation measurements since 1997. The new and continuing measurements means consistent and comprehensive global monitoring of precipitation for use in both climate simulations and climatology studies.

5.3 Future Investigations

Future steps mainly focus on testing the $D_m - \mu$ relationship within an algorithm setting to determine if there is a similar reduction of bias in snowfall retrievals as seen in rainfall retrievals presented in Williams et al. (2014). Other future studies include defining the
relationship into precipitation types for mixed precipitation and freezing rain, removing high LWC values within the snowfall relationship, deriving the framework without using the equivalent melted diameter, and incorporating more surface measurements mainly at colder temperatures. Also, forward models of the radiative properties of snowflakes need to be considered in order to fully assess the impacts of these PSD constraints on retrievals.


L’Ecuyer, T., W. Petersen, and D. Moiseev (2010). Light precipitation validation experiment (LPVEx). *National Aeronautics and Space Administration*.


