OBSERVATION AND SIMULATION OF MID-LATITUDE ICE CLOUDS

BY

WEI WU

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, 2017

Urbana, Illinois

Doctoral Committee:

Professor Greg M. McFarquhar, Chair
Professor Robert M. Rauber
Professor Sonia G. Lasher-Trapp
Associate Professor Steve W. Nesbitt
ABSTRACT

Knowledge of ice crystal particle size distributions (PSDs) is critical for parameterization schemes for atmospheric models and remote sensing retrieval schemes. In-situ observations are commonly used to obtain PSDs and other cloud microphysical properties. In particular, two-dimensional in situ images captured by cloud imaging probes are widely used to derive PSDs in terms of their maximum particle dimension ($D_{\text{max}}$). In the second chapter, different definitions of $D_{\text{max}}$ for non-spherical particles recorded by 2D probes are compared. It is shown that derived PSDs can differ by up to a factor of 6 for $D_{\text{max}} < 200$ µm or $D_{\text{max}} > 2$ mm. The large differences for $D_{\text{max}} < 200$ µm are caused by the strong dependence of sample volume on particle size, whereas differences for $D_{\text{max}} > 2$ mm are caused by the small number of particles detected. Derived bulk properties can also vary depending on the definition of $D_{\text{max}}$ because of discrepancies in the definition of $D_{\text{max}}$ used to characterize the PSDs and that used to describe the properties of individual ice crystals. For example, the mass-weighted mean diameter can vary by 2 times, the ice water content ($IWC$) by 3 times, and the mass-weighted terminal velocity by 6 times. Therefore, a consistent definition of $D_{\text{max}}$ should be used for all measurements and single particle properties. As an invariant measure with respect to the orientation of particles in the imaging plane for 2D probes, the diameter of the smallest circle enclosing the particle ($D_S$) is recommended as the optimal definition of $D_{\text{max}}$. If the 3D structure of a particle is observed, then the technique can be extended to determine the minimum enclosing sphere.

The ice clouds in various weather systems from polar to equator have been sampled using aircraft equipped with in-situ probes in the past several decades and plenty of datasets are available, thus the comparison of observed and modeled PSDs using the parameters of gamma distribution function is investigated next to evaluate and potentially improve the numerical
modeling of ice clouds. The Weather Research and Forecasting (WRF) model is used to represent cloud microphysical features observed in a mesoscale convective system (MCS) sampled on 20 May 2011 during the Mid-latitude Continental Convective Clouds Experiment. Inter-comparison studies are conducted with 3 different spectral bin microphysics schemes: the Caltech-NCAR-NOAA Bin scheme (CNNB), the Fast Spectral Bin Model (FSBM) and the University of Pecs and NCAR Bin scheme (UPNB). The simulated ice cloud PSDs and their variability are compared against those measured in-situ with a two-dimensional cloud probe and a high volume precipitation spectrometer installed on the University of North Dakota Citation aircraft in the trailing stratiform region behind the MCS. The observed and simulated PSDs are fit to gamma distribution functions using the incomplete gamma fit (IGF) routine to determine the intercept ($N_0$), slope ($\mu$) and shape ($\lambda$) parameters. The dependence on environmental conditions of the gamma distribution parameters as ellipsoids of equally realizable solutions in the parameter phase space ($N_0$, $\mu$, $\lambda$) is compared between the three bin schemes and the in-situ observations. Statistically significant differences in PSDs are found among the three bin schemes and between the simulations and observations, including in the median PSD form, the natural variability of PSDs under similar environmental conditions and the dependence of PSDs on temperature. Assumptions about the particle properties (such as mass/terminal velocity-dimensional relations, etc.) and the representations of microphysical processes, such as nucleation, diffusional growth and aggregation growth, in different bin schemes are investigated to explain the differences between models and in-situ observation.

Based on modeling limitations in the above comparison, a final aspect of this work investigates the shape of observed PSDs that cannot be captured by state-of-the-art bin-resolving schemes. Several analytical forms of cloud PSDs have been used in numerical modeling and
remote sensing retrieval studies of clouds and precipitation, including exponential, gamma, lognormal, and Weibull distributions. However, there is no satisfying physical explanation as to why certain distribution forms preferentially occur instead of others. Theoretically, the analytical form of a PSD can be derived by directly solving the general dynamic equation, but no analytical solutions have been found yet. Instead of using a process level approach, the use of the principle of maximum entropy (MaxEnt) for determining the analytical form of PSDs from the perspective of a system is examined. MaxEnt theory states that the probability density function with the largest information entropy among a group satisfying the given properties of the variable should be chosen. Here, the issue of variability under coordinate transformations that arises using the Gibbs/Shannon definition of entropy is identified, and the use of the concept of relative entropy to avoid these problems is discussed. Focusing on cloud physics, the four-parameter generalized gamma distribution is proposed as the analytical form of a PSD using the principle of maximum (relative) entropy with assumptions on power law relations between state variables, scale invariance and a further constraint on the expectation of one state variable. The four-parameter generalized gamma distribution is very flexible to accommodate various type of constraints that could be assumed for cloud PSDs. The exact constraints and distribution parameters need to be further determined using in-situ datasets and idealized numerical models for potential applications in numerical models and remote sensing retrievals.
To My Family
ACKNOWLEDGEMENTS

This work was supported by the U.S. Department of Energy Atmospheric System Research through grants DE-SC0001279, DESC0008500, DE-SC0014065, DE-SC0016476 (through UCAR subcontract Z17-900029) as well as the University of Illinois Dissertation Completion Fellowship from 2016-2017 and by the Advanced Study Program (ASP) at the National Center for Atmospheric Research (NCAR).

I was fortunate to conduct my doctoral research at the Department of Atmospheric Sciences at the University of Illinois at Urbana-Champaign, where I got lots of insightful ideas and numerous support from my advisor Greg McFarquhar and other faculty members, especially those who also work on clouds: Bob Rauber, Sonia Lasher-Trapp, Steve Nesbitt, Nicole Riemer and Larry Di Girolamo. I benefited from lots of discussions from students within the McFarquhar group, especially Bobby Jackson, Joe Finlon, Shichu Zhu, Dan Stechmann, and Sid Gupta.

I also want to acknowledge the opportunity I had to visit NCAR and the collaborations with Hugh Morrison, Lulin Xue, Wojciech Grabowski, Aaron Bansemer, and Andy Heymsfield during the past two years that inspired numerous new ideas. In particular for chapter 3 of this dissertation, I want to acknowledge the simulated fields provided by Lulin Xue from NCAR, Jiwen Fan from PNNL and Zack Lebo from the University of Wyoming.

Finally, I want to thank my wife Le Chang and my daughter for their understanding and mental support during the long and frugal journey of PhD. Family is always my source of courage when facing big challenges.
# TABLE OF CONTENTS

## CHAPTER 1: INTRODUCTION

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>BACKGROUND</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>IN-SITU OBSERVATIONS OF CLOUDS</td>
<td>2</td>
</tr>
<tr>
<td>1.3</td>
<td>NUMERICAL SIMULATIONS OF CLOUDS</td>
<td>4</td>
</tr>
<tr>
<td>1.4</td>
<td>THEORETICAL UNDERSTANDING OF PSDS</td>
<td>6</td>
</tr>
<tr>
<td>1.5</td>
<td>ROLE OF THIS DISSERTATION</td>
<td>7</td>
</tr>
<tr>
<td>1.6</td>
<td>FIGURES</td>
<td>10</td>
</tr>
</tbody>
</table>

## CHAPTER 2: OBSERVATION OF ICE CLOUDS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>INTRODUCTION</td>
<td>12</td>
</tr>
<tr>
<td>2.2</td>
<td>DEFINITIONS OF D_{MAX}</td>
<td>16</td>
</tr>
<tr>
<td>2.3</td>
<td>DATASET</td>
<td>17</td>
</tr>
<tr>
<td>2.4</td>
<td>EFFECT OF D_{MAX} DEFINITIONS ON PSDS</td>
<td>24</td>
</tr>
<tr>
<td>2.5</td>
<td>EFFECT OF D_{MAX} DEFINITIONS ON BULK PROPERTIES</td>
<td>27</td>
</tr>
<tr>
<td>2.5.1</td>
<td>Total number concentration</td>
<td>27</td>
</tr>
<tr>
<td>2.5.2</td>
<td>Ice water content</td>
<td>28</td>
</tr>
<tr>
<td>2.5.3</td>
<td>Mass-weighted terminal velocity</td>
<td>30</td>
</tr>
<tr>
<td>2.5.4</td>
<td>Extinction</td>
<td>30</td>
</tr>
<tr>
<td>2.5.5</td>
<td>Effective diameter</td>
<td>31</td>
</tr>
<tr>
<td>2.6</td>
<td>CONCLUSIONS</td>
<td>32</td>
</tr>
<tr>
<td>2.7</td>
<td>FIGURES AND TABLES</td>
<td>34</td>
</tr>
</tbody>
</table>

## CHAPTER 3: SIMULATION OF ICE CLOUDS

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>INTRODUCTION</td>
<td>46</td>
</tr>
<tr>
<td>3.2</td>
<td>DATA AND MODEL SETUP</td>
<td>49</td>
</tr>
<tr>
<td>3.3</td>
<td>METHODS</td>
<td>51</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Construct flight tracks in numerical simulations</td>
<td>51</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Natural Variabilities of PSDs</td>
<td>53</td>
</tr>
<tr>
<td>3.4</td>
<td>RESULTS</td>
<td>56</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Comparison of median PSDs</td>
<td>56</td>
</tr>
</tbody>
</table>
3.4.2 Comparison of PSD variabilities ................................................................. 57
3.4.3 Comparison of moments of PSDs ............................................................... 59
3.5 CONCLUSIONS ............................................................................................. 60
3.6 FIGURES AND TABLES .................................................................................... 62

CHAPTER 4: STATISTICAL THEORY OF CLOUDS PSD ........................................ 76
4.1 INTRODUCTION ............................................................................................... 76
4.2. MaxEnt and its rationale for cloud physics .................................................. 80
4.3. Problems using Gibbs/Shannon entropy and the concept of relative entropy .......... 85
4.4. Application to cloud PSDs ........................................................................... 87
4.5. Properties of generalized Gamma distribution ............................................. 91
4.6. Testing with in-situ observed liquid and ice PSDs ......................................... 93
4.7. CONCLUSIONS AND DISCUSSIONS ............................................................. 95
4.8 FIGURES AND TABLES .................................................................................... 97

CHAPTER 5: CONCLUSIONS AND DISCUSSIONS ............................................... 102
5.1 SUMMARY ..................................................................................................... 102
5.2 MAIN CONCLUSIONS .................................................................................. 103
5.3 FURTHER STUDIES ..................................................................................... 106

REFERENCES ..................................................................................................... 108
CHAPTER 1: INTRODUCTION

1.1 Background

Ice clouds are an important component in the atmosphere, present in various types of weather systems ranging from summertime deep convective systems to year-round high clouds, such as cirrus, and influence the atmosphere through latent heat release and radiation (Liou 1986; Hartmann et al. 1992; Hansen et al. 1997; Heymsfield and McFarquhar 2002; Wylie et al. 2005). Despite their importance, both the understanding of microphysical processes and properties within ice clouds and the capability to simulate ice clouds in numerical models are still limited.

According to the recent Intergovernmental Panel on Climate Change report (IPCC 2013), the effects of clouds remain the largest uncertainty in climate models due to the complexity of small-scale microphysical processes and the role of aerosols in modulating these processes. It should be noted that the results reported in IPCC (2013) focussed mostly on warm clouds, since most climate models had implemented only very simple ice microphysics. These simple representations have much larger uncertainties compared to warm microphysics, suggesting even larger uncertainties in climate models due to the different radiative effects of warm, ice and mixed clouds (Tan et al. 2016). Compared to warm clouds, the parameterization of ice clouds is very challenging (Heymsfield and Platt 1984; McFarquhar and Heymsfield 1997; Heymsfield et al. 2002; Morrison and Milbrandt 2015) because of the non-spherical shapes of ice particles and the dependence of many important properties, such as fall speed and radiation scattering, on ice crystal shape and size. Previous modeling studies show that simulated ice clouds are very sensitive to the assumed particle size distributions (PSDs) and habits of ice crystals (e.g., Wu 2002; Fridlind et al. 2012a; Furtado et al. 2014). To improve our understanding and numerical
simulation capability of clouds, studies from observations, numerical simulations and theoretical studies are all essential.

1.2 In-situ observations of clouds

In-situ observations collected by various instruments installed on aircraft flying through clouds provide realistic PSDs and bulk properties of clouds, upon which the form of PSDs and other microphysical properties assumed in numerical models and remote sensing retrieval algorithms are usually based (McFarquhar et al. 2011; Baumgardner et al. 2012; Brenguier et al. 2013). Hot-wire probes, such as the King probe (King et al. 1978) and Nevzorov liquid water content (LWC) - total water content (TWC) probe (Korolev et al. 1998), can measure bulk liquid and ice water mass content from the heat transfer of cloud particles impinging on the hot wires. Forward scattering probes, such as the Cloud Droplet Probe (CDP) and the Forward Scattering Spectrometer Probe (FSSP) are used to measure small spherical particles using the theory of Mie scattering (Brenguier et al. 2013). Two-dimensional optical array probes (OAPs), which give images of cloud and precipitation particles using the assumptions of geometric optics that can be used to derive PSDs, were originally developed by Knollenberg (1970). The operating principle of OAPs is illustrated in Fig. 1.1. Over decades of development, three major companies, namely Particle Measurement Systems (PMS), Droplet Measurement Technologies (DMT) and the Stratton Park Engineering Company’s (SPEC), invented different versions of such probes with increasingly faster electronics, higher resolution (up to 10 µm) and/or increased sample volume, including the older 2D-Cloud (2D-C) and 2D-Precipitation (2D-P) probes and the newer Cloud Image Probe (CIP), Precipitation Image Probe (PIP), Two-dimensional Stereo (2D-S) probe and High Volume Precipitation Spectrometer (HVPS). Higher resolution (2.3 µm) images of cloud particles can be further obtained on a charge-coupled device (CCD) or a complementary metal-
oxide-semiconductor (CMOS) camera using SPEC’s Cloud Particle Imager (CPI) or the High Speed Imager (HSI) recently designed by Artium Technologies Inc. in collaboration with Centro Italiano Ricerche Aerospaziali (CIRA) laboratory, even though the sample volume for this type of probe is too small to create PSDs. Recently, the Holographic Detector for Clouds (HOLODEC) (Fugal et al. 2004) was developed to obtain holographic images of cloud particles in a volume with high accuracy.

By using the morphology parameters calculated from two-dimensional images captured by the various in-situ probes mentioned above (Fig. 1.2), cloud PSDs can be constructed. Based on observed PSDs, various bulk cloud properties, such as total number concentration, extinction, liquid and ice water content, mean fall speed, precipitation rate, effective diameter, and single-scattering properties can be determined.

Despite the success of OAPs to measure liquid cloud properties, the uncertainties associated with calculated ice PSDs and bulk microphysical properties is large due to the non-linear shapes of ice particles. The calculation of PSDs and bulk properties are complicated by the fact that different definitions of particle size have been used to characterize PSDs and the functional relationships between single particle properties and the particle size. For example, even though many studies use the maximum diameter ($D_{\text{max}}$) as a measure of particle dimension (Locatelli and Hobbs 1974; McFarquhar and Heymsfield 1998; Petty and Huang 2011; Jackson et al. 2014; Heymsfield et al. 2013; McFarquhar and Heymsfield 1996; Mitchell and Arnott 1994; McFarquhar and Black 2004; Baran et al. 2014; Korolev et al. 2014; Korolev and Field 2015), area-equivalent diameter ($D_{\text{area}}$) (Locatelli and Hobbs 1974; Korolev et al. 2014), and mass-equivalent diameter (or melted diameter $D_m$) (Seifert and Beheng 2006) have also been used. Although all definitions are equivalent for spherical liquid particles, there can be large
differences for nonspherical ice particles, as has been noted for optical array probes (OAPs; Brenguier et al. 2013), imaging disdrometers (Wood et al. 2013), and non-imaging disdrometers (Battaglia et al. 2010). Even if \( D_{\text{max}} \) is used to represent PSDs, there are several different ways \( D_{\text{max}} \) has been calculated for a two-dimensional image (Locatelli and Hobbs 1974; Brown and Francis 1995; McFarquhar and Heymsfield 1996; Mitchell and Arnott 1994; Korolev and Field 2015; Heymsfield et al. 2013). This has important ramifications. For example, McFarquhar and Black (2004) noted that inconsistencies in particle size definitions could have significant impacts on mass conversion rates between different hydrometeor classes used in numerical models. Consistency in the definition used to characterize the PSDs and libraries of particle properties (e.g., mass or scattering properties) is needed to compute bulk or optical parameters.

In chapter 2, the impacts of different definitions of \( D_{\text{max}} \) on PSDs and bulk cloud properties are explored. Differences in bulk properties between the various definitions of \( D_{\text{max}} \) are determined, as are differences in such properties using consistent and inconsistent definitions of \( D_{\text{max}} \) in the derived PSDs and libraries of microphysical and scattering properties.

1.3 **Numerical simulations of clouds**

Numerical models are widely used for weather prediction and climate projection; however, the computing grid size is much larger than the scale at which cloud microphysical processes occur. Therefore, cloud microphysical properties and processes need to be parameterized as functions of the environmental properties that are predicted in these models. There are two major approaches used to represent clouds in numerical models: bulk and bin-resolved schemes, with the principle difference being the necessity of assuming a PSD form explicitly or implicitly in the bulk scheme (Khain et al. 2015). Bin schemes resolve the evolution of PSDs with more
flexibility, while bulk schemes are generally much faster than bin microphysical schemes due to the assumptions of PSD shape. Based on in-situ observations, various probability distribution functions have been proposed to represent cloud PSDs, including exponential (Marshall and Palmer 1948), gamma (Khrgian and Mazin 1952; Ulbrich 1983; Willis 1984; McFarquhar et al., 2007), lognormal (Feingold and Levin 1986) and Weibull distributions (Liu et al. 1995), which are widely used in bulk microphysical schemes in numerical models (e.g., Thompson et al. 2004, 2008; Morrison et al. 2005; Seifert and Beheng 2006; Morrison and Milbrandt 2015).

Many previous studies have compared cloud properties simulated by the bulk schemes with in-situ observations (Brown and Swann 1997; Thompson et al. 2008; Fridlind et al. 2012; Ovchinnikov et al. 2014). However, it is unknown whether bin schemes predict similar PSD forms compared to observations. Furthermore, there can be large variability in PSDs observed under similar environmental conditions in the same system due to the complexity of clouds and the mixing of particles grown at many different locations and environmental conditions. It is not known whether bin models can capture such variability. Recently, a bin microphysical scheme inter-comparison project was conducted to examine the capability of three commonly used bin microphysical schemes to simulate a MCS sampled on 20 May 2011 during the Mid-latitude Continental Convective Clouds Experiment (MC3E) and to improve the bin microphysical schemes through inter-comparison and evaluation against in-situ observations (Xue et al. 2017a, b). Three different bin microphysics schemes were used: the Caltech-NCAR-NOAA Bin scheme (CNNB, Lebo and Seinfeld 2011; Lebo et al. 2012), the Fast Spectral Bin Model (FSBM, Khain et al. 2009; 2010), and the University of Pecs and NCAR Bin scheme (UPNB, Geresdi 1998; Xue et al. 2012). In chapter 3, the PSDs from in-situ observations and simulations using these different microphysical schemes are compared. The differences in PSDs and their variability are
analyzed, linking differences to potential microphysical processes in order to evaluate and potentially improve the bin schemes.

1.4 Theoretical understanding of PSDs

Many empirical functions have been proposed to represent PSDs, such as exponential (Marshall and Palmer 1948), gamma (e.g., Borovikov 1963; Ulbrich 1983), lognormal (e.g., Feingold and Levin 1986; Tian et al. 2010) and Weibull distributions (e.g., Zhang and Zheng 1994; Liu et al. 1995) based on results from various field campaigns. Although many different analytical forms of cloud PSDs have been proposed and widely used in numerical models and remote sensing retrieval algorithms, no adequate physical explanation has been given as to why one particular analytical function should be preferred over another or how that form can be related to cloud microphysics beyond determining the goodness of fit to observed PSDs. Therefore, the choice of a functional form varies from study to study, complicating the comparison of PSD parameters derived from different field campaigns and from model parameterization schemes.

This problem could be solved theoretically by solving the general dynamic equation describing the particle system to find an analytical form of a cloud PSD. However, no analytical solution has been found even when a geometric collection kernel is used for the simplest case of liquid clouds without nucleation, precipitation and breakup (Drake 1972). When more complex processes (e.g., sublimation, aggregation, melting, riming, deposition, etc.) acting in ice or mixed phase clouds are included, the equation is even more difficult to solve. Because analytic solutions have not been possible, numerical methods have been used to determine PSDs in bin resolved models. In contrast to the process-level approach, a system approach using statistical
theory is also viable when the mass or size of every particle is considered as a random variable acting under stochastic processes. One promising statistical theory for determining cloud PSDs is the principle of maximum entropy (MaxEnt, Jaynes 1957a, b), which states that for a group of probability density functions (PDFs) that satisfy given properties of the variable, the PDF with largest information entropy for this variable should be chosen. These given properties usually serve as constraints to the PDFs. The problem of determining PSDs in cloud physics is indeed to find a PDF with certain constraints, and Zhang and Zheng (1994) and Liu et al. (1995) derived the Weibull distribution as the analytical form of PSDs using constraints on the bulk surface area and total mass content using MaxEnt, respectively. Their derived PSD forms differ only on the parameters characterizing the Weibull distribution. Yano et al. (2016) extended the assumptions about the PSDs to include constraints on the mean diameter and mass flux, and examined the impact of these assumptions using idealized simulations, and laboratory and observational datasets. However, these studies applying MaxEnt to cloud PSDs used the Gibbs/Shannon form of entropy, which can only be used for discrete distributions and is not invariant under coordinate transformation for continuous distributions. Contradictory results can be derived using the same assumptions if Gibbs/Shannon entropy is used to derive cloud PSDs. To solve these problems, a new formalism of entropy is needed (Jaynes 1963, 1968), and it is applied to develop a theory on the analytical form of PSDs in chapter 4.

1.5 Role of this dissertation

This dissertation describes the investigation of three research questions under the general theme of cloud PSDs. First, the limitations of a particular measurement derived from cloud in-situ observations are examined and the uncertainties of PSDs and bulk properties associated with
in-situ measurements are quantified. Second, the capability of state-of-the-art bin microphysical parameterizations utilized in a 3D numerical model to replicate the observed form of cloud PSDs in the trailing stratiform region behind a mesoscale convective system is examined. Third, the underlying physical explanations as to why PSDs exhibit certain forms, such as exponential, lognormal and gamma distributions, is explored.

The dissertation is organized as follows. Chapter 2 describes a new algorithm developed to calculate the maximum dimension of an ice crystal, and introduces the in-situ dataset from MC3E used to examine the impact of varying definitions of maximum dimension on the derived PSDs and calculated bulk properties, including number concentration, ice water content, extinction, mass-weighted fall speed, and precipitation rate. This chapter is adapted from a paper published in the *Journal of Atmospheric and Oceanic Technology*.

Chapter 3 examines the use of three bin-resolving schemes by comparing the simulated PSDs against those measured in-situ using a 2DC and a HVPS on the University of North Dakota Citation aircraft in the trailing stratiform region behind the MCS sampled on May 20 during Midlatitude Continental Convective Clouds Experiments (MC3E). The differences in PSDs between the three bin schemes and observations are quantified, including the differences in median PSD form and their natural variabilities. Chapter 3 is based on a paper to be submitted to *Monthly Weather Review*.

Chapter 4 discusses the use of the principle of maximum entropy (MaxEnt) for determining the analytical form of PSDs from the system perspective instead of using a process level approach. A critical issue related to the definition of entropy used in previous studies is solved and the four-parameter generalized gamma distribution is proposed as the analytical form of a PSD, with appropriate assumptions on power law relations between state variables, scale
invariance and a further constraint on the expectation of one state variable. This material is included in a paper submitted to the *Journal of Atmospheric Sciences*.

Chapter 5 summarizes the principal conclusions of this study and provides recommendations for future studies relating to the understanding of cloud physics and development of new models to examine the theoretical basis of and to make improvement to the representation of cloud microphysical processes.
1.6 Figures

**Figure 1.1:** Illustration of operating principle of OAPs (upper, from eol.ucar.edu) and the sample particle images collected by HVPS during Midlatitude Continental Convective Clouds Experiment (lower)
Figure 1.2: The morphology measure of an ice particle example captured by HVPS during MC3E: (a) the traditional maximum dimension along photodiode array or time direction; (b) the diameter of smallest enclosing circle ($D_S$); (c) the major and minor axis of smallest enclosing eclipse; (d) the width and length of smallest enclosing rectangle.
CHAPTER 2: OBSERVATION OF ICE CLOUDS

Copyright Notice: This chapter is a paper in the *Journal of Atmospheric and Oceanic Technology* (Wu and McFarquhar 2016), a journal of the American Meteorological Society. The differences in this chapter from Wu and McFarquhar (2016) only include changes to the formatting of the tables and figures. Wu and McFarquhar (2016) is © Copyright 2016 American Meteorological Society (AMS). Permission to use figures, tables, and brief excerpts from this work in scientific and educational works is hereby granted provided that the source is acknowledged. Any use of material in this work that is determined to be “fair use” under Section 107 of the U.S. Copyright Act September 2010 Page 2 or that satisfies the conditions specified in Section 108 of the U.S. Copyright Act (17 USC §108, as revised by P.L. 94-553) does not require the AMS’s permission. Reproduction, systematic reproduction, posting in electronic form, such as on a web site or in a searchable database, or other uses of this material, except as exempted by the above statement, requires written permission or a license from the AMS. Additional details are provided in the AMS Copyright Policy, available on the AMS Web site located at (http://www.ametsoc.org/) or from the AMS at 617-227-2425 or copyrights@ametsoc.org.

2.1. Introduction

Ice clouds play important roles in the atmosphere through latent heat release and radiative transfer, which are determined by the underlying microphysical processes. Cirrus, the most common form of ice clouds, covers around 20% of the Earth, and hence its influence on radiation is essential for the Earth’s energy balance (Heymsfield and McFarquhar 2002). To understand the properties of and processes occurring in ice clouds, realistic particle size distributions (PSDs)
and bulk properties of ice clouds are needed and are typically obtained from in-situ observations. Assumptions about the form of PSDs based on the in-situ observations are then made in parametrization schemes that are used in atmospheric models and remote sensing retrievals.

The PSDs are typically derived from two-dimensional images obtained in-situ by probes installed on aircraft flying through clouds and by disdrometers on the ground. Two-dimensional optical array probes (OAPs), which give such images of cloud and precipitation particles, were originally developed by Knollenberg (1970). Different versions of such probes are now available, such as Particle Measurement Systems (PMS) 2D-Cloud (2D-C) and 2D-Precipitation (2D-P) probes (Knollenberg 1981), Droplet Measurement Technologies (DMT) Cloud Imaging Probe (CIP) and Precipitation Imaging Probe (PIP) (Baumgardner et al. 2001), and the Stratton Park Engineering Company’s (SPEC) Two-dimensional Stereo (2D-S) probe (Lawson et al. 2006) and High Volume Precipitation Spectrometer (HVPS) (Lawson et al. 1993). These probes provide information on the sizes, shapes and projected areas of ice particles with dimensions greater than about 10mm, with the exact size range of each probe depending on the magnification, resolution and distance between probe arms. High-resolution (2.3 mm) images of cloud particles can also be obtained from a charge-coupled device (CCD) camera on SPEC’s Cloud Particle Imager (CPI) (Lawson et al. 2001). Further, holographic images can be constructed from two-dimensional images obtained by the Holographic Detector for Clouds (HOLODEC) (Fugal et al. 2004). Various kinds of disdrometers, such as 2D Video Disdrometer (2DVD) (Kruger and Krajewski 2002) and Snow Video Imager (SVI) (Newman et al. 2009), also obtain two dimensional images of particles at the ground.

From observed PSDs, various bulk cloud properties, such as total number concentration, extinction, liquid and ice water content, mean fall speed, precipitation rate, effective diameter
and single-scattering properties can be determined. However, calculation of these parameters is complicated by the fact that different definitions of particle size have been used to characterize PSDs and the functional relationships between single particle properties and the particle size, and these definitions of particle size are not always consistent with definitions used to describe the properties of individual crystals. For example, even though many studies use the maximum diameter ($D_{\text{max}}$) as a measure of particle dimension (Locatelli and Hobbs 1974; McFarquhar and Heymsfield 1998; Petty and Huang 2011; Jackson et al. 2014; Heymsfield et al. 2013; McFarquhar and Heymsfield 1996; Mitchell and Arnott 1994; McFarquhar and Black 2004; Baran et al. 2014; Korolev et al. 2014; Korolev and Field 2015), area-equivalent diameter ($D_{\text{area}}$) (Locatelli and Hobbs 1974; Korolev et al. 2014), and mass-equivalent diameter (or melted diameter, $D_{\text{m}}$) (Seifert and Beheng 2006) have also been used (Table 2.1). Although all definitions are equivalent for spherical liquid particles, there can be large differences for non-spherical ice particles, as has been noted for OAP probes (Brenguier et al. 2013), imaging disdrometers (Wood et al. 2013) and non-imaging disdrometers (Battaglia et al. 2010). This has important ramifications. For example, McFarquhar and Black (2004) noted that inconsistencies in particle size definitions could have significant impacts on mass conversion rates between different hydrometeor classes used in numerical models. Consistency in the definition used to characterize the PSDs and libraries of particle properties (e.g., mass or scattering properties) is needed to compute bulk or optical parameters.

Even if maximum dimension is used to represent PSDs, there are different methods that have been used to calculate $D_{\text{max}}$. Depending upon the definition of maximum dimension used in the library of particle properties, these differences can be problematic. The $D_{\text{max}}$ should be the longest dimension in any direction across the 3D volume of the particle, which is equivalent to
the diameter of smallest sphere enclosing the particle. However, due to the limits of measurement technologies, only two-dimensional projections of a particle in one specific orientation can be observed by 2D imaging probes. Therefore, the true maximum dimension of a particle is not directly available, and can be only estimated from the 2D projected images until new technologies that provide the 3D structure of a single particle emerge. In this study, the focus is on the calculation of a maximum dimension describing the two-dimensional projection of a particle, and ramifications of differences in these definitions. This extends previous studies that examined how the calculation of $D_{\text{max}}$ affected properties derived by disdrometers (Wood et al. 2013), and that identified how differences in $D_{\text{max}}$ definitions could impact PSDs from OAPs (Brenguier et al. 2013). There are several different ways $D_{\text{max}}$ has been calculated for a two-dimensional image (Locatelli and Hobbs 1974; Brown and Francis 1995; McFarquhar and Heymsfield 1996; Mitchell and Arnott 1994; Korolev and Field 2015; Heymsfield et al. 2013). In this study, the impacts of different definitions of $D_{\text{max}}$ on PSDs and bulk cloud properties are explored. Differences in bulk properties between the various definitions of $D_{\text{max}}$ are determined, as are differences in such properties using consistent and inconsistent definitions of $D_{\text{max}}$ in the derived PSDs and libraries of microphysical and scattering properties.

The methods for calculating the $D_{\text{max}}$ for 2D imaging probes are described in section 2.2. Section 2.3 describes the field campaigns and probes from which the in-situ data were acquired. The differences in the derived PSDs from one flight are discussed in section 2.4, and the differences in bulk cloud parameters calculated from the PSDs, such as ice water content (IWC), mass-weighted terminal velocity, precipitation rate, extinction and effective radius are discussed in section 2.5. Section 2.6 summarizes the findings and their significance.
2.2. Definitions of Dmax

In this section, different methods of determining $D_{\text{max}}$ from two-dimensional particle images are discussed. In addition, a new method for computing $D_{\text{max}}$ as the diameter of the smallest enclosing circle ($D_S$) is described. To understand how the maximum dimension of a non-spherical ice particle is defined, it is helpful to first review how cloud imaging probes work. This is done in the context of OAPs since they are commonly used for measuring PSDs, but the results apply to any probe acquiring a two-dimensional image. OAPs consist of an array of photodiode detectors that record light emitted from a laser beam. When a particle passes through the laser beam, a number of diodes proportional to the size of the particle are shadowed along the direction of the photodiode array (Brenguier et al. 2013). Because the diodes are clocked by fast response electronics at a rate proportional to the width of photodiode element, the particle is also measured along the direction of aircraft flight (time direction). Thus, a two-dimensional image is recorded. These two directions are shown in Fig. 2.1 and are designated $D_T$ for the maximum dimension in the time direction and $D_P$ for the maximum dimension in the photodiode array direction.

In the past, at least five different definitions of $D_{\text{max}}$ have been commonly used. The maximum particle dimensions previously used include the maximum dimension in the time direction ($D_T$), maximum dimension in the photodiode array direction ($D_P$), the larger of $D_T$ and $D_P$ ($D_L$) (McFarquhar and Heymsfield 1996), the mean of $D_T$ and $D_P$ ($D_A$) (Brown and Francis 1995), and the length of the hypotenuse of a right-angled triangle ($D_H$) constructed from the 2 dimensions and calculated as $\sqrt{D_T^2 + D_P^2}$ (Mitchell and Arnott 1994). Fig. 2.1 illustrates the four definitions of $D_{\text{max}}$ (excluding $D_L$ and $D_A$ for the particles depicted) for an ice particle example.
imaged by the HVPS probe during the Mid-latitude Continental Convective Cloud Experiment (MC3E).

A new algorithm for determining the maximum dimension of a two-dimensional projected image of a particle as the diameter of a minimum enclosing circle is also used in this study. The problem of finding a minimum enclosing circle of a non-spherical particle is a classical computational geometry problem, and solutions are readily available. The origin of such a circle is the perfect location for a public service, such as a hospital or a post office, because it minimizes the distance from the service for all residents (De Berg et al. 2008). The more general problem of finding a minimally enclosing sphere in N dimensions is the so-called Euclidean 1-center problem (Gärtner 1999). There have been many efforts to derive a fast algorithm to determine the smallest enclosing N-spheres, with time complexity ranging from O(n^4) to O(n). Because of the large number of ice crystals that are typically measured during a flight, it is important to implement the fastest possible algorithm in probe processing software. Historically, the optimal algorithm was thought to have time complexity of O[nlog(n)] (Shamos and Hoey 1975), until the first linear-time algorithm was proposed by Megiddo (1982) using a linear programming method. More recently, Welzl (1991) developed a simple randomized linear-time algorithm, with a subsequent implementation developed by Gärtner (1999). This algorithm employs a stochastic method to rapidly determine the minimum surface for dimensions less than 10. This algorithm has been adopted in the University of Illinois OAP processing software (UIOPS) for the two-dimensional cloud particle images, and is used to compute D_5 in this study.

2.3. Dataset

To test the newly implemented Gärtner (1999) algorithm for calculating D_5, in-situ measurements acquired by airborne probes during MC3E (Petersen and Jensen 2012, Jensen et
al. 2016), jointly sponsored by the National Aeronautics and Space Administration (NASA) and the Department of Energy (DOE) Atmospheric Radiation Measurement (ARM), are used. MC3E was conducted in April and May of 2011 in the vicinity of the DOE-ARM Southern Great Plains (SGP) Central Facility in northern Oklahoma. During the field campaign, the University of North Dakota (UND) Citation sampled clouds in 12 weather systems, including fronts, squall lines, and MCSs.

For this study, the MCS that passed over the ARM SGP site from the west on 20 May 2011 was chosen for analysis because all of the in-situ probes worked well. On this day, a deep trough in upper levels was collocated with the low level jet stream, providing a favorable synoptic setting for the development of convection. At the same time, the lower level jet stream supplied a large amount of moisture from the Gulf of Mexico to fuel the convection. Vertical wind shear was also present, which allowed the MCS to persist for a longer time period compared to the low shear condition. During the 4-hour flight which started at 1255 central daylight time from Ponca City, OK, the UND Citation sampled the stratiform region behind the convective line. As shown in Fig. 2.2, the UND Citation executed several constant-altitude stepped legs, and one upward and one downward spiral over the SGP. The UND Citation ascended as high as 7.6 km, and sampled clouds with temperatures ranging from -23 °C to 20 °C. In this study, only data in ice clouds are used.

A variety of particle habits was sampled during the flight. Figure 2.3 shows representative particles as a function of temperature imaged by the 2D-C. Most particles were classified as “irregular” by a habit identification routine (Holroyd 1987). The roundish shape of many of the particles suggests that they might have experienced some riming during their growth history. But, since there was little or no liquid water content measured at temperatures below 0°C
during the flight, their masses and areas were calculated using Brown and Francis (1995) mass- and area-dimensional relationships that were derived for midlatitude cirrus that also consisted of predominantly quasi-spherical irregular particles, with some bullet rosettes and columns mixed in.

The CIP, 2D-C and HVPS were installed on the UND Citation, and nominally sampled particles between 25 µm to 19.2 mm. In this study, data from the 2D-C and HVPS are combined to give a composite PSD. The 2D-C are used to characterize particles smaller than 1 mm, while the HVPS is used for sizes larger than 1 mm. As is shown in Fig. 2.4, the 1 mm cutoff was chosen since it is around the center of the size range where $N(D_{\text{max}})$ for the 2D-C and HVPS agree within 50% for this flight. It should be noted that a small and poorly known depth of field for particles with $D_{\text{max}} < 150$ µm can cause a substantial uncertainty in $N(D_{\text{max}})$ for $D_{\text{max}} < 150$ µm (Heymsfield 1985; Baumgardner and Korolev 1997). In this study, the smallest bin is set to be 150 µm to eliminate this uncertainty. Further, given the 30 µm resolution of the 2D-C, any particles with $D_{\text{max}} < 150$ µm would have at most 5 photodiodes shadowed, meaning there would be poor resolution for looking at the effects of particle shape on computation of $D_{\text{max}}$.

The 2D-C was used for the analysis instead of the CIP because the 2D-C was equipped with anti-shattering tips while the CIP was not. Large numbers of small ice crystals can be produced when a large crystal shatters on the tips of an OAP, therefore, anti-shattering tips have been developed to deflect such shattered particles away from the probe sample volume (Korolev et al. 2011). Korolev et al. (2011, 2013a) and Jackson et al. (2014) have shown that some particles with $D_{\text{max}} < 500$ µm are shattered artifacts, even when anti-shattering tips are used. Shattered artifacts are identified as those particles with inter-arrival times below some threshold. Typically there are two peaks in the inter-arrival times, with the smaller peak corresponding to
the shattered remnants and the larger peak corresponding to real particles (Field et al. 2006; Korolev and Field 2015). However, the time evolution of the frequency distribution of inter-arrival times in Fig. 2.5 illustrated only a single mode in the distribution of inter-arrival times for the 2D-C and HVPS. Therefore, there is no peak in the inter-arrival time analysis corresponding to shorter times, suggesting few artifacts were generated by the shattering of large crystals on the probe tips for conditions sampled during this flight. Therefore, no shattering removal algorithm was used for both the 2D-C and HVPS.

The UIOPS determines various measures of particle morphology, such as particle dimension, projected area, particle habit, particle mass, rejection status, area ratio, aspect ratio, and inter-arrival times. The code is modified to include the calculation of $D_s$ and other definitions of $D_{max}$ so that alternate versions of PSDs were generated. The projected area of a single particle can be directly determined if the particle image is entirely within the photodiode array. However, many particles touch the edges of the photodiode array and therefore additional assumptions are needed to estimate the projected area. To get projected area for particles touching the edge of the photodiode array, various area-dimensional ($A$-$D$) relations can be used to calculate single particle projected area, where $D$ is the reconstructed dimension of the particle (Heymsfield and Parrish 1978). In this study, the projected area for each particle is determined using $A$-$D$ relations as well as the directly imaged area. Traditionally, $A$-$D$ relations are represented by power laws, where:

$$A = aD_{max}^b$$  \hspace{1cm} (2.1)

with $a$ and $b$ habit-dependent parameters. Particle mass ($m$) is not observed by the imaging probes. To get the particle mass, mass-dimensional ($m$-$D$) relations are also assumed, and are again represented by power laws such as:
\[ m = \alpha D_{\text{max}}^\beta \]  
(2.2)

where \( \alpha \) and \( \beta \) are habit-dependent parameters. In this study, Holroyd III (1987)’s habit classification is used to determine ice particle habits. Subsequently, the appropriate \( A-D \) and \( m-D \) relations from Mitchell (1996) are used to give the habit-dependent \( a, b, \alpha \) and \( \beta \) parameters listed in Table 2.2.

To calculate the PSDs, an assumption must be made about the probe sample volume, since the number distribution function \( N(D_{\text{max}}) \) is calculated as the observed number of counts in each bin divided by the sample volume and bin width. The sample volume is calculated using

\[ SV = \min(DOF; W_{\text{arm}}) \times W_{\text{eff}} \times TAS \]  
(2.3)

where \( DOF \) is depth of field, \( W_{\text{arm}} \) is the distance between the probe arms, \( W_{\text{eff}} \) is the effective width of the photodiode array, and \( TAS \) is the true airspeed. The \( DOF \) is calculated as

\[ DOF = 1.5 \frac{D_{\text{max}}^2}{\lambda} \]  
(2.4)

for the 2D-C and CIP (Knollenberg 1970), where \( \lambda \) is the wavelength of laser. The use of \( D_p \) to calculate the \( DOF \) has been suggested (A. Korolev, 2015, personal communication) since the OAPs only take one slice of a particle at a time, but there is currently no consensus about which definition of size should be used for calculating the \( DOF \) for irregular ice particles (Brenguier et al. 2013). This uncertainty contributes to the uncertainty in the probe sample volume. Another source of uncertainty comes from the determination of \( W_{\text{eff}} \). There are three ways to calculate \( W_{\text{eff}} \): entire-in, center-in and the Heymsfield and Parrish (1978) extension. Figure 2.6 shows the sample volumes calculated using these three methods for the 2D-C, CIP and HVPS. These three methods are different in their treatment of partially imaged particles. The entire-in technique only uses the fully imaged particles, but reduces the sample volume linearly as the particle size increases. On the other hand, Heymsfield and Parrish (1978) utilizes all particles, calculating
their size based on the assumption of spherical particles, and thus extends the sample volume but having more uncertainty in the estimated particle area. The center-in technique represents a trade-off between the volume of data and the quality of the images. This method uses partially imaged particles whose center is inside the sample volume, in addition to the fully imaged particles. A particle is determined to be a center-in particle when its maximum dimension in the time direction does not touch the edge of photodiode array. There is a smaller uncertainty in the imaged area for a center-in particle because a greater fraction of the particle was imaged. In this study, the center-in technique is used. In addition to the number distribution function \( N(D_{\text{max}}) \), the area size distribution functions \( A(D_{\text{max}}) \) and the mass distribution function \( M(D_{\text{max}}) \) are also derived.

The bulk cloud properties are related to specific moments of the PSDs or are obtained by integrating the area or mass distribution functions. For example, the total number concentration \( N_t \) is calculated as

\[
N_t = \int_0^\infty N(D_{\text{max}}) dD_{\text{max}} \tag{2.5}
\]

and the ice water content (\( IWC \)) is given by

\[
IWC = \int_0^\infty N(D_{\text{max}}) m dD_{\text{max}} \tag{2.6}
\]

The Nevzorov total water content (\( TWC \)) sensor was also installed on UND Citation to measure both liquid and ice water content (Korolev et al. 1998), which can be used to constrain and validate the assumptions used for calculating \( IWC \) from OAPs. For ice clouds, the \( TWC \) is the same as the \( IWC \). The mass flux, or precipitation rate (\( PR \)), is expressed using equation:

\[
PR = \int_0^\infty v_i m N(D_{\text{max}}) dD_{\text{max}} \tag{2.7}
\]
where \( v_t \) is the terminal velocity of an individual particle calculated from particle area, mass, temperature and pressure following Heymsfield and Westbrook (2010). Following McFarquhar and Black (2004), the mass-weighted terminal velocity \( (V_m) \) is expressed as

\[
V_m = \frac{\int_0^\infty v_t m N(D_{\text{max}})dD_{\text{max}}}{\int_0^\infty m N(D_{\text{max}})dD_{\text{max}}} = \frac{PR}{IWC} \tag{2.8}
\]

The extinction \( (\beta_{\text{ext}}) \) at visible wavelength is twice the total projected area \( (A_c) \) since the ice particles are large enough that geometric optics applies (Um and McFarquhar 2015). The \( A_c \) is the integrated projected area of particles over all sizes, given by

\[
\beta_{\text{ext}} = 2A_c = 2\int_0^\infty N(D_{\text{max}})AdD_{\text{max}} \tag{2.9}
\]

The effective diameter \( (D_e) \) is commonly used for parameterization of single scattering properties needed for calculation of shortwave radiative transfer (Fu 1996; McFarquhar and Heymsfield 1998; Mitchell 2002). Although several different definitions of \( D_e \) have been used (McFarquhar and Heymsfield 1998), \( D_e \) is defined here following Fu (1996) as

\[
D_e = \frac{2\sqrt{3}IWC}{3\rho_iA_c} \tag{2.10}
\]

because the ratio of \( IWC \) to \( A_c \) is closely related to ice radiative properties. In Eq (2.11), \( \rho_i \) is the bulk density of ice, assumed to be 0.91 g cm\(^{-3}\). Number-weighted mean dimension \( (D_{nm}) \), or the average particle dimension, is defined as

\[
D_{nm} = \frac{\int_0^\infty N(D_{\text{max}})D_{\text{max}}dD_{\text{max}}}{\int_0^\infty N(D_{\text{max}})dD_{\text{max}}} \tag{2.11}
\]

whereas the mass-weighted mean diameter \( (D_{mm}) \) is calculated as

\[
D_{mm} = \frac{\int_0^\infty N(D_{\text{max}})D_{\text{max}}mdD_{\text{max}}}{\int_0^\infty N(D_{\text{max}})mdD_{\text{max}}} \tag{2.12}
\]

23
Besides the bulk properties addressed here, there are also differences in the radar reflectivity derived from PSDs using alternate definitions of $D_{\text{max}}$. In the Rayleigh scattering regime, the radar reflectivity is a higher order moment of the PSDs than the quantities discussed here (Smith 1984). However, different models exist for the calculation of radar reflectivity from ice particles, and consideration of all these different models is beyond the scope of this study.

For calculating quantities in Eq (2.5) - (2.12), the measured particles are first sorted into bins of varying width with $D_{\text{max}}$ ranging from 150 µm to 1.92 cm. Then the integrations are converted to summations, with the relevant particle properties computed at the mid-point of the bin. The midpoint rule gives a better estimate of integration for the concave-down shape of the PSDs compared with the trapezoidal rule. Given that the number of bins are not large, more complex numerical methods (e.g. Simpson’s rule) are not needed. Based on these calculations, the impacts of different definitions of $D_{\text{max}}$ on PSDs and bulk cloud properties are explored in the next 2 sections.

### 2.4. Effect of $D_{\text{max}}$ definitions on PSDs

The composite PSDs computed using six different definitions of $D_{\text{max}}$ for the MCS on 20 May 2011 are compared in this section. The $N(D_{\text{max}})$ were first determined for each 1s of flight time, and averaged PSDs were subsequently computed in three different temperature ranges. Figure 2.7 compares the $N(D_{\text{max}})$ determined using the six different definitions, with the upper panels showing $N(D_{\text{max}})$ and the lower panels showing the ratio of $N(D_{\text{max}})$ to $N(D_S)$, with the ratio being one when $D_S$ is used as $D_{\text{max}}$. The $N(D_{\text{max}})$ all show a peak between 200 µm to 400 µm for all temperature ranges. For $D_{\text{max}} > 300$ µm, $N(D_{\text{max}})$ decrease sharply to 8 orders of magnitude smaller, with the rate of decrease depending on temperature. The $N(D_{\text{max}})$ using the
different definitions of $D_{\text{max}}$ can vary by up to one order of magnitude with, for example, $N(D_A)$ and $N(D_H)$ at about 1 mm between -10 °C and 0 °C varying by this amount.

The trends in how the different definitions of $D_{\text{max}}$ vary are systematic in that $D_T$, $D_P$, $D_A$ and $D_L$ are always smaller than or equal to $D_S$, while $D_H$ is always greater than or equal to $D_S$, as shown for the example particle in Fig. 2.1. Consequently, $N(D_T)$, $N(D_P)$, $N(D_A)$, and $N(D_L)$ are larger than $N(D_S)$ for smaller $D_{\text{max}}$ and smaller for larger $D_{\text{max}}$. The trend for $N(D_H)$ compared to $N(D_S)$ is opposite. For all the definitions of $D_{\text{max}}$, $N(D_L)$ is the closest to $N(D_S)$. The differences between $N(D_{\text{max}})$ using different definitions of $D_{\text{max}}$ increase when $D_{\text{max}}$ is farther away from the mode diameter, both for smaller and larger sizes.

Since the number distribution function is determined by the number of counts in each bin and by the sample volume for particles with the given size, both factors contribute to the differences in the PSDs. For $D_{\text{max}} < 200$ µm, the large difference between PSDs is due to the dependence of the depth of field, and therefore the sample volume, on particle size. This increases $N(D_{\text{max}})$ if the particles are moved from a larger bin to a smaller bin because of the different definitions of $D_{\text{max}}$. The effect is larger as the particle size, and therefore the DOF and the sample volume, decreases. For $D_{\text{max}} > 200$ µm, the sample area is constant (red solid lines in Fig. 2.6), so that the changes in PSDs for the different definitions are due to the number of particles sorted into each bin, which is determined by the definition of $D_{\text{max}}$. The PSDs are more sensitive to the choice of $D_{\text{max}}$ definitions for larger particles than for smaller particles because the number of particle counts per bin decreases sharply as the particle size increases; thus the classification of even a single particle into a different bin can have a big impact.

When comparing the behavior of PSDs for different temperatures in Fig. 2.7, it is apparent that the differences in $N(D_{\text{max}})$ for different $D_{\text{max}}$ definitions are larger for lower
temperatures than for higher temperatures. This might be explained by more circular particles at higher temperatures due to the action of riming and aggregation. It is also noticeable that the slope of the PSDs increases for the lower temperatures, meaning fewer large particles exist when the temperature is lower. This would again lead to larger differences in the PSDs.

The impact of the different definitions of $D_{\text{max}}$ is also seen when comparing the number weighted mean diameter ($D_{\text{nm}}$) and the mass-weighted mean diameter ($D_{\text{mm}}$) computed from $N(D_{\text{max}})$ determined using different definitions of $D_{\text{max}}$. Figure 2.8a compares $D_{\text{nm}}$ calculated using five definitions of $D_{\text{max}}$ to that calculated using $D_S$. The differences are also summarized in Table 2.3. The $D_{\text{nm}}$ range from 300 $\mu$m to 1200 $\mu$m, with the differences varying between 56% to 140% due to the different definitions of $D_{\text{max}}$. Figure 2.8b shows the comparisons for $D_{\text{mm}}$. The differences in $D_{\text{mm}}$ vary from 65% to 125% and the values range from 300 $\mu$m to 8 mm. Using $D_T$ and $D_P$ gives the smallest $D_{\text{nm}}$ and $D_{\text{mm}}$, with $D_{\text{nm}}$ and $D_{\text{mm}}$ about 74.5-79.6% and 83.9-87.5% of those computed with $D_S$. The $D_{\text{nm}}$ and $D_{\text{mm}}$ determined using $D_A$ are similar, with median ratio of 77.9% and 81.2% of those determined using $D_S$, respectively. The $D_{\text{nm}}$ and $D_{\text{mm}}$ determined using $D_L$ provide the closest estimate to those determined using $D_S$, with average differences of 91.7% and 93.3%, respectively. On the contrary, both $D_{\text{nm}}$ and $D_{\text{mm}}$ computed using $D_H$ are systematically larger than the $D_{\text{nm}}$ and $D_{\text{mm}}$ computed using $D_S$, with up to a 140% difference.

Two factors contribute to the differences between $D_{\text{nm}}$ and $D_{\text{mm}}$ determined using different definitions of $D_{\text{max}}$. First, the differences in $N(D_{\text{max}})$ are large as shown in Fig. 2.7. Since there are more smaller particles when using $D_T$, $D_P$, $D_A$ and $D_L$ to define $D_{\text{max}}$ than when using $D_S$, the $D_{\text{nm}}$ and $D_{\text{mm}}$ using these definitions are smaller than $D_{\text{nm}}$ and $D_{\text{mm}}$ calculated using $D_S$, respectively. The second reason for the difference in the computed $D_{\text{mm}}$ is that the assumed $m$-$D$ relations are not applicable with certain definitions of $D_{\text{max}}$. If the same coefficients in the
m-D relations are used for different $D_{\text{max}}$, then large differences in derived particle mass exist. For example, by using Brown and Francis (1995), $D_A$ gives the estimate of mass that is most consistent with the derivation from the original relationship, while $D_T$ and $D_P$ underestimate the particle mass and $D_L$, $D_S$ and $D_H$ overestimate the particle mass. It is also important to note that there are no consistent definitions of $D_{\text{max}}$ used in different studies giving m-D relations. This point will be discussed in more detail in the subsequent section.

2.5. Effect of $D_{\text{max}}$ definitions on bulk properties

The differences in PSDs translate into differences in bulk properties. In this section, the influence of different definitions of $D_{\text{max}}$ on bulk cloud properties, such as $N_t$, $IWC$, $V_m$, $PR$, $\beta_{\text{ext}}$ and $r_e$, is examined.

2.5.1. Total number concentration

The $N_t$ is obtained by integrating the number distribution function using Eq (2.3). The definition of $D_{\text{max}}$ affects the derived $N_t$ as shown in Fig. 2.9a, with $N_t$ varying from 80% to 140% of that estimated using $D_S$. Even though the same number of particles is recorded by the probe regardless of the definition of $D_{\text{max}}$ used, the $N_t$ changes with the definition of $D_{\text{max}}$ because of the dependence of the estimated probe $DOF$ on $D_{\text{max}}$. In general, definitions that give larger values of $D_{\text{max}}$ than $D_S$ for the same particle, such as $D_H$, produce smaller $N_t$, with values ranging from 94.7% to 99.8% of those obtained using $D_S$ within the 5th to 95th percentiles, and a median of 98.0%. On the other hand, definitions that give smaller values of $D_{\text{max}}$ than $D_S$, such as $D_T$, $D_P$, $D_A$ and $D_L$, have larger $N_t$. For example, $N_t$ derived using $D_T$ and $D_P$ range from 101.8% to 126.0% and 104.0% to 119.5% within the 5th-95th percentiles of that derived using $D_S$, and a median of 106.8% and 107.7%, respectively. The $N_t$ derived using $D_A$ and $D_L$ give closer values to those derived using $D_S$, with a median of 103.1% (99.9%-109.6% within the 5th-95th
percentiles) and 101.3% (100.4%-103.3% within the 5th-95th percentiles) of those obtained with $D_S$, respectively.

### 2.5.2. Ice water content

The $IWC$s calculated using different definitions of $D_{max}$ and the total water content ($TWC$) measured by the Nevzorov probe (black dots) are shown in Fig. 2.9b as a function of the $IWC$ calculated using $D_S$. When the $IWC$s derived using different definitions of $D_{max}$ are compared with the $TWC$ observed by Nevzorov probe, it is seen that the Nevzorov $TWC$s are less than the $IWC$s computed from $N(D_{max})$ when the $IWC$s get larger. This could be explained by the difficulties associated with the Nevzorov’s probes ability to sample larger ice particles (Korolev et al. 2013b). In addition, power law fits are less likely to perform well at the extremes for the estimate of bulk properties. The $IWC$s determined using alternate $D_{max}$ definitions vary between 50% to 150% of those determined using $D_S$. The $IWC$s calculated using $D_H$ gives the largest estimate, ranging from 137.2% to 149.3% within the 5th-95th percentiles of those determined using $D_S$ with a median of 142.9%. In addition to factors leading to varying $N(D_{max})$, differences of $IWC$s are caused by the use of different $D_{max}$ in the mass-dimensional relations, that are inconsistent with the definitions of $D_{max}$ originally used to develop the relations. The $IWC$s determined using $D_H$ are larger than those determined using other definitions since $D_H$ is the largest value of any $D_{max}$ and hence gives the largest estimated particle mass given the use of the same $a$ and $b$ coefficients in the $m$-$D_{max}$ relations. Defining $D_{max}$ as $D_T$ and $D_P$ produced the smallest $IWC$s because those particle dimensions are only measured in one direction, and hence they and their associated masses are smallest. The $IWC$s calculated using $D_T$ are closer to those calculated using $D_S$ than those calculated using $D_P$ (median ratio of 70.0% versus 56.9%). This occurs because the full dimension of particles that touch the edge of photodiode array is not
recorded, but the longer dimension can be recorded in the time direction. For similar reasons, $IWC$s estimated based on $D_A$ are likely underestimated. The $IWC$s derived using DL are closest to $IWC$s derived using $D_S$, with median ratio of 81.9%.

For the implementation of a $m$-$D_{\text{max}}$ relation, it is important that the definition of $D_{\text{max}}$ used be consistent with the definition used in the relevant study that derived the relations. However, this is not always the case. For example, Brown and Francis (1995) used $D_A$ to calculate the mass of aggregates with the $m$-$D$ relation that is originally documented by Locatelli and Hobbs (1974, p. 2188) for “aggregates of unrimed radiating assemblages of plates, side planes, bullets, and column”. In spite of the large differences, possible conversions between different definitions of $D_{\text{max}}$ may provide a way to correct the $m$-$D_{\text{max}}$ derived using different definitions. But, it is difficult and non-trivial to correct the $m$-$D$ relations so that they apply to alternate definitions of $D_{\text{max}}$ because conversions between different definitions of $D_{\text{max}}$ depend upon morphological features of ice crystals that are not always reported in original studies. For example, Fig. 2.10 shows that there is large scatter between the different definitions of $D_{\text{max}}$ on a particle by particle basis, and no simple relations can be found between different definitions of $D_{\text{max}}$. For the same particle with different orientations in the two-dimensional imaging plane, the $D_S$ is invariant; however, other definitions show wider scatter due to different orientations, especially for $D_T$ (Fig. 2.10a) and $D_P$ (Fig. 2.10b). The combination of both $D_T$ and $D_P$ can reduce the scatter significantly, even though there are still systematic differences among different methods. As a result, it appears that methods that involve consideration of the particle dimension in at least two different directions (for example, $D_T$ and $D_P$) are needed to get a reasonable estimate of the $D_{\text{max}}$. This is similar to the findings of Wood et al. (2013) for the Two-
Dimensional Video Disdrometer, which allows particle dimensions to be measured from two perpendicular views.

2.5.3. Mass-weighted terminal velocity

Figure 2.9c shows that $V_m$ determined using varying definitions of $D_{\text{max}}$ can vary from 28% to 180% compared to those determined using $D_S$. As with IWC, larger $V_m$ occur for $D_{\text{max}}$ definitions that give larger particle sizes, with median $V_m$ determined using $D_H$ at 134.0% (120.1% to 166.1% within the 5th-95th percentiles) of those determined using $D_S$. Similarly, the smallest $V_m$ were associated with the $D_{\text{max}}$ definitions giving smaller particle sizes, namely $D_P$ and $D_T$, with median ratio of 54.6% (28.7%-82.0% within 5th-95th percentiles) and 74.5% (66.3%-85.4% within 5th-95th percentiles) of those determined using $D_S$, respectively. In general, the differences are larger when $V_m < 0.4$ m s$^{-1}$. When $V_m$ exceeds 0.6 m s$^{-1}$, the differences decrease to be within 60% to 145% and the $V_m$ computed using the other definitions converge to $V_m$ calculated using $D_S$. The large spread in $V_m$ is contributed by both the variations in particle area and mass used for the calculation of $V_m$, since both particle area and mass estimated using the power laws vary due to the different definitions of $D_{\text{max}}$. The scatter of precipitation rate is also presented in Fig. 2.9d, which shows a similar pattern for $V_m$ since $PR$ is the combined effects of $V_m$ and IWC, and the uncertainties in $V_m$ are much greater than that in IWC.

2.5.4. Extinction

Figure 2.9e shows the variation in extinction determined using different definitions of $D_{\text{max}}$ and the directly imaged area determined from the OAPs (black dots). The extinction determined using different definitions of $D_{\text{max}}$ can vary from 60% to 133% of those calculated using $D_S$. Similar to the result shown for IWC, $\beta_{\text{ext}}$ determined using $D_H$ ranges from 126.7% to
133.4% of that determined using $D_S$, with a median of 130.4%. The uncertainties in both $N(D_{\text{max}})$ and the $A$-$D$ relations contribute to the differences. When using definitions of $D_{\text{max}}$ that produce smaller particle sizes, the $\beta_{\text{ext}}$ values are smaller than those determined using $D_S$ with, for example, medians of 65.4% and 71.9% for $D_P$ and $D_T$, respectively, compared to those determined using $D_S$. Because the OAPs directly measure particle area, there is some measure of truth for particle area, or equivalently extinction. Computations of $\beta_{\text{ext}}$ using $D_L$ are closest to the $\beta_{\text{ext}}$ estimated from the OAP directly imaged area. This is not surprising because many of the $A$-$D_{\text{max}}$ relations are based on the use of $D_L$ as the maximum diameter. The patterns of differences in extinction between definitions are quite similar to those seen in $IWC$. However, the differences are smaller, because $\beta_{\text{ext}}$ is based on a lower order moment compared to $IWC$.

### 2.5.5. Effective diameter

Figure 2.9f shows the effective diameter calculated using the different definitions of $D_{\text{max}}$ compared to that determined using $D_S$. The $D_e$ computed using the different definitions range from 82% to 120% of the $D_e$ computed using $D_S$. Since $D_e$ is related to the ratio of $IWC$ to $A_c$, the dependence of both these variables on the definition of $D_{\text{max}}$ influences $D_e$. In general, the $D_e$ computed using $D_H$ are largest, ranging from 106.9% to 113.7% of the $D_e$ calculated using $D_S$ within the 5th-95th percentiles, with a median of 109.7%. However, using definitions that give smaller particle sizes can also give higher $D_e$ estimates, especially for larger values of $D_e$. When $D_e$ is less than 80 mm, using $D_P$, $D_T$ and $D_A$ still give smaller estimates than those calculated using $D_{\text{max}}$. The $D_e$ calculated using $D_P$ and $D_T$ produced the smallest estimates, with medians of 86.9% (82.3%-94.7% within the 5th-95th percentiles) and 96.7% (92.8%-105.7% within the 5th-95th percentiles) of those determined using $D_S$. Using $D_A$ and $D_L$ give estimates of $D_e$ very close to those determined using $D_S$, with medians of 88.9% (84.2%-96.1% within the 5th-95th percentiles).
percentiles) and 96.4% (94.9%-98.2% within the 5th-95th percentiles) of those calculated using $D_{\text{max}}$, respectively.

2.6. Conclusions

Many previous studies have used alternate definitions and algorithms for computing the maximum dimension ($D_{\text{max}}$) of an ice crystal. A new method for calculating $D_{\text{max}}$ as the diameter of the smallest circle ($D_S$) enclosing a two-dimensional image using a linear-time algorithm is described in this study. To see the effects of different definitions, the particle size distribution (PSDs) and bulk cloud properties are derived using various definitions of $D_{\text{max}}$. Since there is no consensus on the optimum definition of $D_{\text{max}}$ for 2D imaging probes, the uncertainties in PSDs and bulk properties due to different definitions are quantified. Derived bulk properties vary depending on the definitions of $D_{\text{max}}$ because of discrepancies in the definition of $D_{\text{max}}$ used to characterize the PSDs and that used to describe the properties of individual ice crystals. The main findings of this study are as follow:

1). The differences in the number distribution functions $N(D_{\text{max}})$ derived using various definitions of $D_{\text{max}}$ can differ by up to a factor of 6 for $D_{\text{max}} < 200 \ \mu\text{m}$ and $D_{\text{max}} > 2 \ \text{mm}$. The large differences for $D_{\text{max}} < 200 \ \mu\text{m}$ are caused by use of different definitions, as well as the strong dependence of sample volume on the particle size, whereas differences for $D_{\text{max}} > 2 \ \text{mm}$ are caused by the small number of particles detected.

2). Number-weighted and mass-weighted mean diameter ($D_{nm}$ and $D_{mm}$, respectively) calculated using alternate definitions of $D_{\text{max}}$ vary from 56% to 140% and 65% to 125% of those calculated using $D_S$, respectively.

3). The difference in derived $IWC$ can differ from 50% to 150% depending on the definitions of $D_{\text{max}}$ used.
4). The $V_m$ can vary from from 28% to 180%, depending on the definitions of $D_{\text{max}}$ used.

5). The precipitation rate (mass flux) based on the above $IWC$ and terminal velocity calculations can differ from 20% to 250%, depending on the definitions of $D_{\text{max}}$ used.

6). The extinction determined using different definitions $D_{\text{max}}$ can range from 60% to 133% to that computed using $D_S$.

7). The effective diameter computed using different definitions of $D_{\text{max}}$ can range from 82% to 120% of that determined using $D_S$.

8). Higher moments of PSDs have larger differences between the different definitions of $D_{\text{max}}$ than do the lower order moments of the PSDs.

9). Of the six different definitions of $D_{\text{max}}$, $D_P$, $D_T$, $D_A$ and $D_L$ give smaller estimates of particle size than does $D_S$, while $D_H$ yields a larger estimate. Using $D_L$ provides the closest estimate to $D_S$ among the six definitions considered here.

The results presented here apply only to the stratiform regions of MCSs. Further research is needed to determine how the results may vary for other kinds of clouds which may contain a different mixture of habits. In addition, the maximum dimension derived from two dimensional images may not represent a true maximum dimension for a three-dimensional particle, unless the maximum dimension is always in a plane perpendicular to the laser beams of OAPs.

Consideration of the three-dimensional value of particles would make the computation of bulk properties more complex, since the underlying $m$-$D$ and $A$-$D$ relations have been developed using two-dimensional projections of measured particles.

Based on the above-mentioned analysis, consistent definitions of $D_{\text{max}}$ should be used in subsequent studies deriving PSDs and bulk properties of ice clouds, because there is no simple relation converting between different definitions of $D_{\text{max}}$ (Fig. 2.10). Definitions that involve
considerations of maximum dimensions in at least two directions (e.g., $D_T$ and $D_P$) are needed to get a reasonable estimate of the $D_{max}$. The $D_S$ proposed in this study is an attractive choice for $D_{max}$ due to the invariant properties with respect to orientations in the imaging plane. In addition, the linear-time algorithm described in this study make the $D_S$ calculation almost as fast as the calculations for other definitions. If the 3D structure of a single particle is observed in the future, the technique can be naturally extended to determine the minimum enclosing sphere, which represents the true maximum dimension of hydrometeors. Even though it is unlikely there will be a standard definition of $D_{max}$ in the near future, it is strongly suggested that the definition of $D_{max}$ used should be mentioned in subsequent papers as the uncertainties due to different definitions have been shown to be large in this study.

2.7 Figures and Tables

<table>
<thead>
<tr>
<th>Definition</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_T$</td>
<td>PSDs (Lawson et al. 2015)</td>
</tr>
<tr>
<td>$D_P$</td>
<td>PSDs (Korolev and Field 2015)</td>
</tr>
<tr>
<td>$D_A$</td>
<td>m-D relation for aggregate (Brown and Francis 1995)</td>
</tr>
<tr>
<td>$D_L$</td>
<td>PSDs (McFarquhar and Heymsfield 1996)</td>
</tr>
<tr>
<td>$D_H$</td>
<td>A-D relations (Mitchell and Arnott 1994)</td>
</tr>
<tr>
<td>$D_S$</td>
<td>PSDs (Heymsfield et al. 2013)</td>
</tr>
<tr>
<td>$D_{area}$</td>
<td>m-D relation for graupel (Locatelli and Hobbs 1974)</td>
</tr>
<tr>
<td>$D_{in}$</td>
<td>PSDs (Seifert and Beheng 2006)</td>
</tr>
</tbody>
</table>
TABLE 2.2. List of $m$-$D$ relation ($m=\alpha Db$) and $A$-$D$ relation ($A=\alpha D^b$) parameters

<table>
<thead>
<tr>
<th>Habit</th>
<th>$\alpha$ [g cm$^{-\beta}$]</th>
<th>$\beta$</th>
<th>$a$ [cm$^{2-\beta}$]</th>
<th>$b$</th>
<th>Notes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sphere</td>
<td>0.4765</td>
<td>3</td>
<td>0.7854</td>
<td>2</td>
<td>Standard spherical ice particles</td>
</tr>
<tr>
<td>Tiny</td>
<td>0.4765</td>
<td>3</td>
<td>0.7854</td>
<td>2</td>
<td>Small particles are treated as sphere</td>
</tr>
<tr>
<td>Linear</td>
<td>0.001666</td>
<td>1.91</td>
<td>0.0696</td>
<td>1.5</td>
<td>Columns, above for $D_{\text{max}}&lt;0.3$ mm, and below for $D_{\text{max}}\geq0.3$ mm</td>
</tr>
<tr>
<td></td>
<td>0.000907</td>
<td>1.74</td>
<td>0.0512</td>
<td>1.414</td>
<td></td>
</tr>
<tr>
<td>Oriented</td>
<td>0.001666</td>
<td>1.91</td>
<td>0.0696</td>
<td>1.5</td>
<td>Used the same value as Linear</td>
</tr>
<tr>
<td></td>
<td>0.000907</td>
<td>1.74</td>
<td>0.0512</td>
<td>1.414</td>
<td></td>
</tr>
<tr>
<td>Plate</td>
<td>0.00739</td>
<td>2.45</td>
<td>0.65</td>
<td>2</td>
<td>For hexagonal plates</td>
</tr>
<tr>
<td>Graupel</td>
<td>0.049</td>
<td>2.8</td>
<td>0.5</td>
<td>2</td>
<td>Graupels</td>
</tr>
<tr>
<td>Dendrite</td>
<td>0.000516</td>
<td>1.8</td>
<td>0.21</td>
<td>1.76</td>
<td>Dendrites</td>
</tr>
<tr>
<td>Aggregate</td>
<td>0.00294</td>
<td>1.9</td>
<td>0.2285</td>
<td>1.88</td>
<td>Used Brown and Francis (1995) parameters.</td>
</tr>
<tr>
<td>Irregular</td>
<td>0.00294</td>
<td>1.9</td>
<td>0.2285</td>
<td>1.88</td>
<td>Same as Aggregate</td>
</tr>
</tbody>
</table>

TABLE 2.3. Summary of the median ratio of bulk properties derived using various of definitions of $D_{\text{max}}$ to that derived using $D_S$.

<table>
<thead>
<tr>
<th>Definitions</th>
<th>$D_{\text{mm}}$</th>
<th>$D_{\text{mm}}$</th>
<th>$N_T$</th>
<th>$IWC$</th>
<th>$V_m$</th>
<th>PR</th>
<th>Extinction</th>
<th>$D_e$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$D_T$</td>
<td>79.6%</td>
<td>83.9%</td>
<td>107.7%</td>
<td>70.0%</td>
<td>74.5%</td>
<td>52.4%</td>
<td>71.9%</td>
<td>96.7%</td>
</tr>
<tr>
<td>$D_P$</td>
<td>74.5%</td>
<td>87.5%</td>
<td>106.8%</td>
<td>56.9%</td>
<td>54.6%</td>
<td>30.0%</td>
<td>65.4%</td>
<td>86.9%</td>
</tr>
<tr>
<td>$D_A$</td>
<td>77.9%</td>
<td>81.2%</td>
<td>103.1%</td>
<td>58.5%</td>
<td>52.0%</td>
<td>30.0%</td>
<td>66.0%</td>
<td>88.9%</td>
</tr>
<tr>
<td>$D_L$</td>
<td>91.7%</td>
<td>93.3%</td>
<td>101.3%</td>
<td>81.9%</td>
<td>84.1%</td>
<td>68.0%</td>
<td>84.8%</td>
<td>96.4%</td>
</tr>
<tr>
<td>$D_H$</td>
<td>118.1%</td>
<td>113.4%</td>
<td>98.0%</td>
<td>142.9%</td>
<td>134%</td>
<td>191.5%</td>
<td>130.4%</td>
<td>109.7%</td>
</tr>
</tbody>
</table>
Figure 2.1: Schematic of definition of $D_T$, $D_P$, $D_H$, and $D_S$ for an ice particle captured by the HVPS during MC3E. $D_A$ and $D_L$ are not shown in the figure since they are mathematical functions of $D_T$ and $D_P$: $D_A$ is the mean of $D_T$ and $D_P$, while $D_L$ is the larger of $D_T$ and $D_P$. 
Figure 2.2: Flight track of UND Citation on 20 May 2011. The color indicates the time during the flight.
**Figure 2.3:** Representative particles imaged by the 2DC during four flight legs on 20 May 2011 between 13:30 and 15:30 CDT shown as function of average temperature of leg on which they were measured.
Figure 2.4: Overlap of average PSDs measured by 2D-C and HVPS for all time with T < 0°C for MCS sampled on 20 May 2011.
Figure 2.5: The distribution of inter-arrival time for a) 2D-C and b) HVPS for time when constant altitude flight legs were flown by UND Citation with T < 0 °C.
Figure 2.6: The sample volume calculated using center-in, entire-in and HP78 extension for 2D-C, CIP and HVPS, installed on UND Citation during MC3E.
Figure 2.7: \( N(D_{\text{max}}) \) as function of \( D_{\text{max}} \) using six different definitions of \( D_{\text{max}} \) (upper panels), and the ratio of \( N(D_{\text{max}}) = N(D_S) \) for \( D_{\text{max}} = D_T, D_P, D_A, D_L \) and \( D_H \) as indicated in legend for MCS on May 20, 2011. (a) shows the average for \(-10^\circ C < T < 0^\circ C\), (b) for \(-20^\circ C < T < -10^\circ C\), and (c) for \(-30^\circ C < T < -20^\circ C\).
Figure 2.8: Box and whisker diagram showing ratio of a) $D_{nm}$ and b) $D_{mm}$ computed using definition of $D_{max}$ shown on horizontal axis compared to that computed using $D_S$. Notched box extends from 1st to 3rd quartile, with a red line indicating the median value. Blue dots are those past the whiskers, which are defined as 5% to 95% percentile. The data used are averaged over 10 seconds.
Figure 2.9: Box and whisker diagram showing ratio of bulk property: a) total number concentration, b) ice water content, c) mass-weighted terminal velocity, d) precipitation rate, e) extinction, and f) effective diameter computed using definition of $D_{\text{max}}$ or directly observed (only for b) ice water content and e) extinction) shown on horizontal axis compared to that computed using $D_S$. The box extends from 1st to 3rd quartile, with a red line indicating the median value. Red dots are those past the whiskers, which are defined as 5% to 95% percentile. The data used are averaged over 10 seconds.
**Figure 2.10:** Relations between $D_{\text{max}}$ computed using alternate definitions as function of that computed using $D_S$ for all accepted particles measured using 2DC on 20 May 2011. The color indicates frequency of occurrence for different definitions of $D_{\text{max}}$ over the same $D_S$. 
CHAPTER 3: SIMULATION OF ICE CLOUDS

This chapter will be submitted as an article to Monthly Weather Review by Wu, McFarquhar, Xue, Morrison, Grabowski and Poellot (2017). I processed the observation dataset and analyzed the model output data simulated by Xue et al. (2017a) contained in this chapter as part of my work, so this study is included as a chapter of the dissertation.

3.1 Introduction

Clouds are ubiquitous and play important roles in the atmosphere through latent heat release and radiative transfer, influencing both weather and climate (Liou 1986; Hartmann et al. 1992; Wylie et al. 2005). However, our understanding of cloud microphysical properties and processes and their impact on climate is still low, especially for ice clouds (IPCC 2013; Khain et al. 2015). To understand the microphysical processes occurring in clouds, knowledge of realistic cloud particle size distributions (PSDs) and microphysical properties (such as bulk water content, total number concentration and extinction) are needed from observations. In-situ observations obtained by instruments installed on aircraft are especially important because they best determine the sizes and shapes of cloud particles including those of liquid droplets and ice crystals (e.g., McFarquhar et al. 2011; Baumgardner et al. 2012; Brenguier et al. 2013).

In-situ observations can be used to evaluate and improve the representation of cloud microphysical properties and processes in numerical models. A cloud is a collection of large numbers of liquid water droplets and/or ice crystals with varying habits and sizes ranging from micrometers to centimeters. Since it is not practical to track all the cloud particles, two approaches are used to represent their evolution in numerical models: bin microphysics schemes and bulk microphysics schemes. Bin schemes directly simulate the evolution of a number of
different particle sizes of specific shapes and phases so that the shape of a PSD does not need to be assumed. On the other hand, bulk schemes usually assume the PSDs have a specific functional form while predicting one, two or three moments of the PSD (e.g., the total mass content, the total number concentration, radar reflectivity). Assumption about the gamma fit parameters may also be required depending on the number of moments prognosed. Bulk microphysical schemes are generally much faster than bin microphysical schemes, but do not specifically predict the evolution of different particles sizes. Assumptions about the PSD shape and any diagnostic parameters of the PSD are based on observations, typically obtained by probes mounted on aircraft flying through clouds or from ground based in-situ observations. Based on such observations, certain functional forms, such as a gamma distribution, have been found to represent PSDs well (e.g., Ulbrich 1983; Willis 1984; McFarquhar et al., 2007), and are widely used in bulk microphysical schemes in numerical models (e.g., Thompson et al. 2004, 2008; Morrison et al. 2005; Seifert and Beheng 2006; Morrison and Milbrandt 2015).

For numerical models used for weather prediction and climate projection, the computing grid size is much larger than the scale at which cloud microphysical processes occur. Therefore, some cloud microphysical properties and processes need to be parameterized as functions of the environmental properties that are predicted in these models for both bin and bulk microphysical schemes. Compared to warm clouds, the parameterization of cold clouds is very challenging (Heymsfield and Platt 1984; McFarquhar and Heymsfield 1997; Heymsfield et al. 2002; Morrison and Milbrandt 2015) because of the non-spherical shapes of ice particles and the dependence of many important properties, such as fall speed, mass, and radiation scattering, on ice crystal shape and size. Previous modeling studies show that simulated cold clouds are very
Bin schemes are believed to better represent natural processes compared to the bulk schemes because they specifically resolve complex microphysical processes acting on specific sizes of particles (Khain et al. 2015). However, it is unknown whether bin schemes predict similar PSD forms compared to observations. While lots of previous studies have compared the bulk cloud properties and PSDs simulated by the bulk schemes with the in-situ observations (Brown and Swann 1997; Thompson et al. 2008; Fridlind et al. 2012; Ovchinnikov et al. 2014), there are not many comparisons for bin-resolved schemes. In addition, due to the complexity of clouds and the mixing of particles grown at many different locations and environmental conditions, there can be large variability in PSDs observed under similar environmental conditions in the same system. It is not known whether bin models can capture such variability. Recently, a bin microphysical scheme inter-comparison project was conducted to examine the capability of three commonly used bin microphysical schemes to simulate a MCS sampled on 20 May 2011 during the Mid-latitude Continental Convective Clouds Experiment (MC3E) and to improve the bin microphysical schemes through inter-comparison and evaluation against in-situ observations (Xue et al. 2017a, b). The comparison of storm structure and dynamics between simulations and ground-based observations was investigated in Part 1 of a two-part paper series (Xue et al. 2017a). Part 2 compares the microphysical processes between simulations and how these differences will impact the storm dynamics and structure (Xue et al. 2017b). This study focuses on comparing the observed and simulated PSDs and their variability as a part of the inter-comparison project. By comparing against in-situ observations, the capabilities of these three schemes are assessed.
commonly used bin schemes to produce the dependence of observed PSDs on temperature and their variability on the stratiform region of the MCS is investigated.

The remainder of the chapter is organized as follows. Section 3.2 introduces the data and model setup used in this study. Section 3.3 presents the methodology to find locations in simulations corresponding to the regions that sampled by aircraft and to quantify the variability of PSDs under similar environmental conditions. Section 3.4 compares the in-situ observations with numerical simulations, with a focus on PSDs and their variability. The underlying differences in representing ice particle properties (such as particle mass and fall speed) and microphysical processes (such as nucleation and aggregation) are also discussed to explain the differences in PSDs. Section 3.5 summarize the most significant findings and offers discussions for future research.

3.2 Data and Model Setup

In-situ measurements acquired by airborne probes on 20 May 2011 during MC3E (Petersen et al., 2012), jointly sponsored by the National Aeronautics and Space Administration (NASA) and the Department of Energy (DOE) Atmospheric Radiation Measurement (ARM), are compared against the result of three bin–resolved simulations in this chapter. The case on 20 May 2011 is a classical MCS, with broad stratiform clouds trailing behind the strong convective line at its mature stage (Tao et al., 2013; Wu and McFarquhar, 2016; Xue et al., 2017a, b). On 20 May 2011, the MCS passed over the ARM Southern Great Plains (SGP) site from the west, with a deep trough in upper levels collocated with the lower level jet stream carrying a large amount of moisture from the Gulf of Mexico. The UND Citation sampled the stratiform region behind the convective line for 4 hours starting at 12:55 CDT. During the flight, the UND Citation first
executed constant-altitude stepped legs at -6°C, -10°C and -16°C in two sequences (denoted OBS1 and OBS2 thereafter), and then performed an upward and downward spiral in the trailing stratiform region near the SGP site. The aircraft reached as high as 7.6 km, with temperatures ranging from -23 °C to 20 °C.

The University of Illinois Optical Array Probe (OAP) Processing Software (UIOPS) was used to process the raw image data collected by the OAPs installed on the UND Citation. To construct the PSDs, two-dimensional cloud (2D-C) probe and high volume precipitation spectrometer (HVPS) were used to determine the number distribution function $N(D)$ for particles maximum dimension ($D_{\text{max}}$) defined as smallest enclosing circle between 150 µm to 1 mm and 1 mm to 12.8 mm, respectively. The 2D-C can nominally measure particles as small as 30 µm. However, particles with $D_{\text{max}} < 150$ µm are not used due to the poorly defined sample volume for these small particles (Baumgardner and Korolev 1997), and the possibility that some particles at these small sizes may be shattered artifacts even though shattering-mitigating tips and artifact rejection algorithms were used (Korolev et al., 2013; Jackson et al., 2014). The overlap region of the 2D-C and HVPS agreed within 50% for this flight (Fig. 2.4). The microphysical properties, such as total number concentration, ice water content and mass fluxes, were also derived from the PSDs using UIOPS. The Nevzorov total water content sensor was used to measure the bulk water content and provide closure with that derived from PSDs. More detailed information about the case and how the data were processed can be found in chapter 2.

The WRF model was used to simulate an ideal squall line initialized by a sounding acquired over Morris, OK at 1200 UTC on 20 May 2011. Three different bin microphysics schemes were used: the Caltech-NCAR-NOAA Bin scheme (CNNB, Lebo and Seinfeld 2011; Lebo et al. 2012), the Fast Spectral Bin Model (FSBM, Khain et al. 2009; 2010), and the
University of Pecs and NCAR Bin scheme (UPNB, Geresdi 1998; Xue et al. 2012). The model setup is described by Xue et al. (2017a), so only a brief summary is provided here. The simulation is performed over a 3D domain of 612 km x 122 km x 25 km with 1 km horizontal spacing and around 250 m vertical spacing (100 vertical levels). The time step is 3 seconds. The planetary boundary layer scheme, land surface model and radiation schemes were turned off in all idealized simulations for simplicity. The idealized simulation was used instead of a real case simulation as bin schemes are too computationally expensive to conduct such an inter-comparison project. However, according to Xue et al. (2017a), the ideal simulation captures the structure of the MCS compared to the ground and aircraft based observations.

3.3 Methods

3.3.1 Construct flight tracks in numerical simulations

To compare in-situ observations with the results of numerical simulations, the location in the modeled storm that most closely corresponds to the location of the observations must be determined. The comparison is complicated by the fact that the aircraft sampled only a limited area in the stratiform region of the MCS. Since there is no exact correspondence between the observed and simulated storms, the area of the simulated storms compared against observations is selected by constructing flight tracks on the simulated fields with similar distances away from the convective lines, the same temperature range and having similar radar reflectivity characteristics. Fig. 3.1a-d are reproduced from Fig. 3 of Xue et al. (2017a) and show the observed and simulated maximum radar reflectivity based on analysis from the gridded NEXRAD product at 12 UTC and the simulated fields at hour 6. Fig. 3.1e shows how the maximum radar reflectivity over the whole column averaged over a 100 km length in the direction of the convective line varies with the distance away from the convective line for
observations and simulations. All three schemes produced strong convective lines, with maximum reflectivity ranging from 53 to 62 dBZ, stronger than the observed $Z_e$ of 45 dBZ. However, the width of the stratiform clouds differ from simulation to simulation, with the downwind edge of stratiform region defined as radar reflectivity larger than 25 dBZ ranging from around 110 km for CNNB to 150 km for UPNB and 170 km for FSBM, all less than 180 km for observations. The red circle indicates the region of the radar reflectivity around 35 dBZ where the aircraft sampled the storm, which is 150-170 km behind the convective line in the extended stratiform region, and well behind the transition zone. For FSBM and UPNB, the aircraft is placed in the same 35 dBZ reflectivity region as observations within the stratiform clouds. The green vertical bar in Fig. 3.1e, which is 145 km behind the convective line, shows the location of the artificial aircraft for FSBM, while the blue vertical bar shows the location of artificial aircraft in UPNB, 135 km behind the convective line. Due to the smaller maximum radar reflectivity of the stratiform region simulated by CNNB, 30 dBZ is chosen as the threshold instead of 35 dBZ and the red vertical bar in Fig. 1e shows the location of the artificial aircraft in CNNB, which is 90 km behind the convective line. All the flight tracks are in the area with radar reflectivity from 30-35 dBZ.

After the distance to the convective lines was determined for all three simulations, the three constant altitude legs at the appropriate location at model levels with temperatures nearest to -6 °C, -10 °C and -16 °C were examined to compare simulated microphysical properties against those observed on the constant altitude flight legs. Figure 3.2 compares the modeled and observed temperature and vertical velocity for the selected legs. Since the model vertical levels are discrete, the mean temperature of the chosen levels may have a +/- 1°C offset compared to that of the observations. However, the standard deviations of temperature for both observations
and simulations are within 0.2 °C except for FSBM for the -6 °C flight leg. The vertical velocities along the constant altitude leg at -6 °C, -10 °C and -16 °C for the observations and simulations are also very similar, as shown in the right panels of Fig. 3.2. The detailed mean and standard deviation of temperature and vertical velocity for both observations and simulations at three flights legs are listed in Table 3.1. Generally, the environmental conditions along the flight track are similar according to Table 3.1.

### 3.3.2 Natural Variabilities of PSDs

Table 3.1 shows the environmental conditions used to categorize the conditions of which the PSDs are sampled. The median PSDs for each condition were determined using all simulated or observed PSDs acquired for flight legs at -6 °C, -10 °C and -16 °C. To define the variability of PSDs for given environmental conditions, the uncertainties in the PSDs must be quantified. The uncertainty in a PSD due to statistical sampling is proportional to the square root of the number of particles counted in each size bin (Hallett 2003; McFarquhar et al. 2015), hereafter $N_i$, so that

$$
\Delta N(D_i) = \frac{\sqrt{N_i}}{SV \cdot \Delta D_i}
$$

(3.1)

with $SV$ being the probe sample volume, $D_i$ the maximum dimension of bin $i$, $\Delta D_i$ the bin width, and $N(D_i)$ the number distribution function at $D_i$.

The quantification of the sampling uncertainty and natural variability of the observed PSDs measured during the -10 °C constant altitude flight leg from 13:53 UTC to 14:00 UTC on May 20, 2011 is shown in Fig. 3.3. The different plots show the PSDs averaged over different time scales. The 1 second averaged PSDs in Fig. 3.3a show that sampling uncertainties and natural variability together contribute to the spread of the PSDs. When the PSDs are averaged over 5 s, 10 s, and 20 s, the counting uncertainties of the PSDs at different time scales are reduced because the fractional error is smaller when more particles are counted in each bin, but
the small scale natural cloud inhomogeneity is also increasingly lost. As shown in Fig. 3.3, the spread is much less upon computing the 10 second averages, with the black error bar denoting the increasingly smaller counting uncertainties. Therefore, the spread of 10 s average PSDs are used to represent the natural variabilities at larger temporal/spatial resolution assuming that the contribution of counting uncertainties to the spread of 10s PSDs are negligible.

As shown in Fig. 3.4, the spread of \( N(D) \) around the median value is not symmetrically distributed for the 10 s averaged PSDs. For example, \( N(D) \) can be around 0.1 \( \text{L}^{-1} \text{µm}^{-1} \) higher than the mean for 0.32 mm particles, while only lower than the mean by 0.05 \( \text{L}^{-1} \text{µm}^{-1} \). Due to the asymmetry of variability in \( N(D) \), the 10\(^{th} \) to 90\(^{th} \) percentiles are used to represent the natural variabilities instead of standard deviation. The minimum and maximum of \( N(D) \) are not used to define the range since this could encompass outliers in the distribution. Hereafter, \( N_{10}(D) \) denotes the 10\(^{th} \) percentile of the PSDs under the given environmental conditions, while \( N_{90}(D) \) denotes the 90\(^{th} \) percentile.

In order to represent the PSDs in a form conducive to parameterization schemes for models or remote sensing schemes, a gamma distribution given by \( N(D) = N_0 D^\mu e^{-\lambda D} \) is commonly used to represent a PSD, where \( N_0 \) is the intercept parameter, \( \mu \) the shape parameter and \( \lambda \) the slope parameter. To fit a single PSD to a gamma distribution, the discrete incomplete gamma fit (DIGF) technique is used to determine the gamma distribution parameters \( (N_0, \mu, \lambda) \) following McFarquhar et al. (2015) by minimizing \( \chi^2 \):

\[
\chi^2 = \sum_{j=1}^{m} \left[ \frac{M_{on(j)} - M_{fn(j)}}{\sqrt{M_{on(j)}M_{fn(j)}}} \right]^2
\]

where \( M_{on} \) is the \( n(j) \)th moment of observed PSD determined as \( M_{on(j)} = \sum_l D_l^{n(j)} N(D_l) \Delta D_l \)

and \( M_{fn} \) is the \( n(j) \)th moment of fitted PSD \( N_f(D) \) determined as \( M_{fn(j)} = \int_{D_{min}}^{D_{max}} D^{n(j)} N_f(D) dD \).
The summation is over the $m$ different moments chosen for the minimization. Usually the $0^{\text{th}}, 2^{\text{nd}}$ and $4^{\text{th}}$ moments (therefore $m=3$) are used to define $\chi^2$, which approximately correspond to number concentration, bulk ice water content and radar reflectivity for ice clouds. The gamma parameters $(\bar{N}_0, \mu, \lambda)$ that minimize $\chi^2$ are the fit parameters. Therefore, a single PSD corresponds to a single point in the 3D gamma parameter phase space, while the range of PSDs create a volume in the gamma parameter phase space representing the natural variability of the PSDs.

The range of PSDs observed under similar environmental conditions can be used to create an ellipsoid in the gamma parameter phase space of solutions that characterize the natural variability that are equally realizable under this environmental condition (McFarquhar et al. 2015). The ellipsoid can be determined by finding the minimum ellipsoid enclosing all points in $(N_0, \mu, \lambda)$ phase space that satisfying $\chi^2 < \Delta \chi^2$, where

$$\Delta \chi^2 = \sum_{j=1}^{m} \left[ \frac{M_{on,10}(j) - M_{fn}(j)}{\sqrt{M_{on,10}(j)M_{fn}(j)}} \right]^2 + \sum_{j=1}^{m} \left[ \frac{M_{on,90}(j) - M_{fn}(j)}{\sqrt{M_{on,90}(j)M_{fn}(j)}} \right]^2 \quad (3.3)$$

and $M_{on,10}$ and $M_{on,90}$ are the $n$th moments of $N_{10}(D)$ and $N_{90}(D)$. The technique used here differs from that used by McFarquhar et al. (2015) in that the “error bars” of the PSDs are determined by applying the spread from the 10$^{\text{th}}$-90$^{\text{th}}$ percentiles of all 10 s averaged PSDs instead of the sampling uncertainties in the individual PSDs. Therefore the volume of the ellipsoid corresponds to natural variabilities in the median PSDs instead of statistical counting uncertainties. For example, the variabilities of PSDs shown in Fig. 3.4 can be represented as the volume in the 3D phase space of gamma distribution parameters $(N_0, \mu, \lambda)$ shown in the inset of Fig. 3.4. The points within the ellipsoid are the possible choice of gamma distribution parameters for this particular flight leg. This variation should be taken into consideration in bulk schemes in
numerical models when using gamma size distribution instead of assuming fixed values of the distribution parameters.

3.4 Results

3.4.1 Comparison of median PSDs

Fig. 3.5 shows the median and spread of the 10 second averaged PSDs for OBS1, OBS2, CNNB, FSBM and UPNB at -6 °C, -10 °C and -16 °C, respectively. The underlying ice, snow and graupel PSDs for CNNB, FSBM and UPNB are also plotted. The median PSDs of OBS1 and OBS2 are representative of single-modal distributions whereas the median PSDs of the simulations generally have a multi-modal character, such as the median PSD for UPNB at -10 °C with two peaks at 300 µm and 1000 µm, which corresponds to the peaks in the graupel and snow PSDs, respectively. The deviation of the simulated PSD from the observed shape suggests that some microphysical processes acting in nature may be absent or inadequately represented in these three bin schemes, especially the conversion between manually separated species: ice, snow and graupel. There are also some similarities. The OBS1 and OBS2 both show a peak in \( N(D) \) at \( D \) between 300 and 500 µm, with both FSBM and UPNB having a similar peak. However, there is no such peak in the CNNB simulations. Instead, the \( N(D) \) is decreasing for those sizes in the CNNB simulations. The \( N(D) \) for particles with \( D < 300 \) µm at -16 °C varies among the three bin schemes. The FSBM has changes of less than 3% in \( N(D) \) between 100 and 300 µm. However, the \( N(D) \) at 100 µm for the CNNB simulations is about 10% of that at 300 µm and for the UPNB \( N(D) \) is about 5 orders of magnitude smaller at 100 µm compared to that at 300 µm. In contrast, the observed \( N(D) \) at 100 µm are within about 10% of those at 300 µm. As the temperature increases from -16 °C to -6 °C, the number concentration at 100 µm for all three bin schemes becomes closer to the observed value of 0.01 L⁻¹µm⁻¹.
3.4.2 Comparison of PSD variabilities

In Fig. 3.5, the spread of the PSDs is larger for the CNNB and UPNB compared to the observations, especially for the CNNB scheme which has spreads of up to 2 orders of magnitude in $N(D)$ over almost all measured size ranges. The spreads of PSDs for FSBM is smaller than that observed over most of the size range compared expect that for $D < 300$ µm. The CNNB does not create the widespread stratiform regions as the FSBM and UPNB schemes do, and its stratiform region is too narrow and inhomogeneous. This can be explained by assumptions about ice particle shape in different schemes. Unlike FSBM and UPNB, CNNB assumes spherical ice particles which have much faster fall speeds at the smallest sizes (Fig. 3.6b). Since the ice particles fall too fast, they cannot be transported too far away from the convective line, and therefore the stratiform region in CNNB is smaller compared to FSBM and UPNB, and the observations. On the contrary, both FSBM and UPNB have simulated extended stratiform region like observations. Due to the large spread of PSDs produced by the CNNB scheme, the ellipsoids for CNNB are not shown because the volumes in CNNB are around 10 times larger compared to observations and those simulated by the FSBM and UPNB schemes.

The range of equally realizable solutions for the PSD gamma parameters in $(N_0, \mu, \lambda)$ phase space for the observations and simulations at the different temperature levels is shown in Fig. 3.7. The volume of the ellipsoids and therefore the variability in UPNB simulations are larger than that of OBS1 and OBS2, while the volume of the ellipsoids in FSBM are smaller than that of OBS1 and OBS2 at -16 °C. In addition, as the temperature increased in the in-situ observations from -16 °C to -6 °C, the volume of the ellipsoids reduced to one order of magnitude smaller volume, from 5.76x10³ to 4.52 x10² for OBS1 and from 7.88x10³ to 1.30x10² for OBS2. That trend is much smaller in all the numerical models, with ellipsoid volume
decreasing from 2.44x10^3 to 2.09 x10^2 for FSBM and increasing from 1.04x10^4 to 1.67x10^4 for UPNB. These trends in OBS1 and OBS2 indicate that clouds are more homogenous in lower altitude, however, the inhomogeneity of clouds in simulations is not reduced as the cloud particles fall into lower altitude.

Furthermore, the slopes of the major axis of the ellipsoids characterizing the observations decrease as the temperature changes; however, this does not happen in all three simulations. The change of slope can be seen better in 2D projections of the ellipsoid. Table 3.2 also lists the major axis vector, and the slopes of the major axis in $\mu$-$\lambda$ and $\mu$-$\log_{10}(N_0)$ planes. The projection in $\mu$-$\lambda$ phase space in Fig. 3.8 shows that the $\mu$-$\lambda$ slope decreases as the temperature increases for observations. The change of slope is smaller in all three bin model simulations compared to that in observations. The projection of ellipsoids on $\mu$-$\lambda$ phase space shows that $\mu$ takes the same range of values, but the range of values that $\lambda$ takes is reduced. Looking at OBS1 for example, $\mu$ is within the range of -1 to 6 for all three flight legs, while the range of values for $\lambda$ is reduced from 0 to 1.5x10^{-4} m^{-1} at -16 °C to 0 to 0.6x10^{-4} m^{-1} at -6 °C, with corresponding slope change from 2.18 x10^3 to 0.84 x10^3 m^{-1}. For FSBM and UPNB, the slope change is much smaller, with the slope change from 1.52 x10^3 to 1.36 x10^3 m^{-1} for FSBM and from 2.74 x10^3 to 2.96 x10^3 m^{-1} for UPNB. The mass-weighted diameter for gamma distribution can be expressed by

$$D_{mm} = \frac{\mu + \beta}{\lambda}$$

(3.4)

where $\beta$ is the parameter for an $m$-$D$ relation given by $m=\alpha D^\beta$. Since $\mu$ does not change with temperature, the decrease in the range of $\lambda$ with increasing temperature indicates that the mass-weighted diameters increase with increasing temperature. This is consistent with ice particle aggregation growth. The aggregation growth in all three bin simulations may underestimate the broadening of PSDs.
The projection in $\mu$-$\log_{10}(N_0)$ phase space in Fig. 3.9 shows that the $\mu$-$\log_{10}(N_0)$ slope also decreases as the temperature increases and the change of slope is smaller in all three bin model simulations. The projection of ellipsoids on $\mu$-$\lambda$ phase space shows that $\mu$ takes the same range of values, but the range of values of $\lambda$ is reduced. Looking at OBS1 for example, $\mu$ is within the range of -1 to 6 for all three flight legs, while the range of values for $N_0$ is reduced from $10^5$ to $10^{19}$ m$^{-4}$ at -16 °C to $10^5$ to $10^{17}$ m$^{-4}$ at -6 °C, corresponding to the change of slope from 1.91 to 1.44. The change of slope in $\mu$-$\log_{10}(N_0)$ plane is from 1.72 to 1.66 for FSBM, while it is from 20.3 to 2.15 for UPNB. For a gamma distribution, $N_0$ can be expressed by

$$N_0 = N_T \frac{\lambda^{\mu+1}}{\Gamma(\mu + 1)}$$

(3.5)

where $N_T$ is the total number concentration. The decrease of $N_0$ is convoluted by both the change of $N_T$ and $\lambda$. Since some bulk models use relations between gamma distribution parameters to reduce the complexity, the change of relations between these parameters should be taken into consideration.

3.4.3 Comparison of moments of PSDs

Moments of PSDs relate to the clouds bulk properties, such as water content relating to $3^{\text{rd}}$ ($2^{\text{nd}}$) moment of PSDs and radar reflectivity relating to $6^{\text{th}}$ ($4^{\text{th}}$) moment for liquid (ice) clouds. Since bin schemes can only predict number and mass of cloud particles at specific mass bins, bulk properties relating to other moments of PSDs are calculated from the predicted PSDs. Fig. 3.10 shows the $0^{\text{th}}$ to $5^{\text{th}}$ moments of the PSDs for both the observations and numerical simulations at -10 °C. Generally, the lower moments (Fig. 3.10a-c) simulated have similar median value and spreads compared to observations, while the spreads for larger moments (Fig. 3.10d-f) are broader than the observations. This is due to the fact that larger variabilities in PSDs will amplify in larger moments of PSDs. With that said, even though all the bin schemes produce
similar number concentration and total water content, the derived radar reflectivity differ much more significantly, as is shown in Fig. 3.1. This has ramification for the diagnostic variables related to higher moments of PSDs, such as the radar reflectivity, as in most numerical models only lower moments such as number and mass are predicted.

3.5 Conclusions

A MCS sampled by the UND Citation during MC3E was used to examine the capability of three different bin microphysical schemes to reproduce the observed PSDs and their variability within the trailing stratiform region of an MCS. These three schemes are the Caltech-NCAR-NOAA Bin scheme (CNNB), the Fast Spectral Bin Model (FSBM) and the University of Pecs and NCAR Bin scheme (UPNB). The observed and simulated PSDs were fit to gamma distributions. The variability of observed and simulated particle size distributions (PSDs) were quantified using ellipsoids in the phase space of the gamma distribution parameter \((N_0, \mu, \lambda)\) to represent volumes of equally realizable solutions. Within this framework, the PSD and its dependence on the environmental conditions were compared between the three bin schemes and the in-situ observations. The main findings of this study are as follows:

1). The simulated ice cloud PSDs are generally multi-modal, unlike the observed distributions which have a more mono-modal shape. Conversion between ice species may not be represented adequately in bin schemes.

2). The variability of PSDs as defined by the 10th to 90th spread of 10s averaged PSDs are generally larger than those measured in-situ using 2DC and HVPS on University of North Dakota Citation aircraft in the trailing stratiform region behind the MCS.
3). The volume of generated equally realizable ellipsoids decreased to 1 order of magnitude smaller volume with increasing temperature from -16 °C to -6 °C in observations, while the trend is much smaller for FSBM and UPNB.

4). The slopes of the major axis of generated equally realizable ellipsoids in observations decrease more with increasing temperatures than that in the simulations. For the projected ellipses on the \( \mu-\lambda \) plane, the slope at -6 °C decreased to less than half of that at -16 °C, while the change of slope in FSBM and UPNB is just around 10%.

5). The simulated lower moments of PSDs are similar to observations, while higher moments in simulations are generally larger compared to observations.

6). The differences in PSDs among the three bin schemes and between the simulations and observations are due to the assumptions about the particle properties, such as mass/terminal velocity-dimensional relations, etc. and representations of microphysical processes in different bin schemes, such as manually defined species, conversion between ice species, nucleation, diffusional growth and aggregation growth.

The conclusions here are only based on observations and simulations of the stratiform region from the MCS observed on 20 May. The applicability to other cases is not known. However, the differences in representing the ice particle properties and microphysical processes may impact the simulated weather systems. Further study is warranted using data from other field campaigns conducted in different seasons and geographic locations.
### 3.6 Figures and Tables

**Table 3.1:** The environmental conditions along flight tracks in observations and simulations

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>OBS1</th>
<th>OBS2</th>
<th>CNNB</th>
<th>FSBM</th>
<th>UPNB</th>
</tr>
</thead>
<tbody>
<tr>
<td>leg</td>
<td>mean</td>
<td>mean</td>
<td>mean</td>
<td>mean</td>
<td>mean</td>
</tr>
<tr>
<td>-16°C leg</td>
<td>-16.3</td>
<td>-16.4</td>
<td>-16.4</td>
<td>-15.8</td>
<td>-16.2</td>
</tr>
<tr>
<td></td>
<td>0.13</td>
<td>0.13</td>
<td>0.10</td>
<td>0.09</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>-0.33</td>
<td>0.06</td>
<td>0.26</td>
<td>0.22</td>
</tr>
<tr>
<td>vertical velocity (m/s)</td>
<td>std</td>
<td>0.44</td>
<td>0.67</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td>-10°C leg</td>
<td>-10.45</td>
<td>-10.5</td>
<td>-10.00</td>
<td>-10.51</td>
<td>-9.74</td>
</tr>
<tr>
<td></td>
<td>0.14</td>
<td>0.12</td>
<td>0.17</td>
<td>0.11</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>0.00</td>
<td>-0.39</td>
<td>-0.03</td>
<td>0.24</td>
<td>0.13</td>
</tr>
<tr>
<td>vertical velocity (m/s)</td>
<td>std</td>
<td>0.42</td>
<td>0.30</td>
<td>0.23</td>
<td>0.22</td>
</tr>
<tr>
<td>-6°C leg</td>
<td>-5.57</td>
<td>-5.45</td>
<td>-6.23</td>
<td>-6.86</td>
<td>-6.11</td>
</tr>
<tr>
<td></td>
<td>0.16</td>
<td>0.14</td>
<td>0.14</td>
<td>0.75</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>0.72</td>
<td>-0.61</td>
<td>0.05</td>
<td>-0.42</td>
<td>-0.02</td>
</tr>
<tr>
<td>vertical velocity (m/s)</td>
<td>std</td>
<td>0.54</td>
<td>0.56</td>
<td>0.20</td>
<td>0.39</td>
</tr>
</tbody>
</table>

**Table 3.2:** The volume, major axis vector and slopes of ellipsoid in $\mu$-$\lambda$ and $\mu$-$\log_{10}(N_0)$ planes

<table>
<thead>
<tr>
<th>Temperature (°C)</th>
<th>OBS1</th>
<th>OBS2</th>
<th>FSBM</th>
<th>UPNB</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$V=5.76e3$</td>
<td>$V=7.88e3$</td>
<td>$V=2.44e3$</td>
<td>$V=1.04e4$</td>
</tr>
<tr>
<td></td>
<td>(8.75e-4,4.58e-4,1)</td>
<td>(7.24e-4,3.62e-4,1)</td>
<td>(1.13e-3,6.57e-4,1)</td>
<td>(7.43e-4,3.65e-4,1)</td>
</tr>
<tr>
<td></td>
<td>$S_{\mu-\lambda}=2.18e3$</td>
<td>$S_{\mu-\lambda}=2.76e3$</td>
<td>$S_{\mu-\log_{10}(N_0)}=2.00$</td>
<td>$S_{\mu-\log_{10}(N_0)}=2.74e3$</td>
</tr>
<tr>
<td></td>
<td>$S_{\mu-\log_{10}(N_0)}=1.91$</td>
<td>$S_{\mu-\log_{10}(N_0)}=2.00$</td>
<td>$S_{\mu-\log_{10}(N_0)}=1.72$</td>
<td>$S_{\mu-\log_{10}(N_0)}=2.03$</td>
</tr>
<tr>
<td>-10°C</td>
<td>$V=1.30e3$</td>
<td>$V=4.90e3$</td>
<td>$V=2.59e3$</td>
<td>$V=1.52e4$</td>
</tr>
<tr>
<td></td>
<td>(1.27e-3,3.772e-4,1)</td>
<td>(9.01e-4,4.78e-4,1)</td>
<td>(1.2e-3,7.14e-4,1)</td>
<td>(7.62e-4,3.72e-4,1)</td>
</tr>
<tr>
<td></td>
<td>$S_{\mu-\lambda}=1.30e3$</td>
<td>$S_{\mu-\lambda}=2.09e3$</td>
<td>$S_{\mu-\log_{10}(N_0)}=1.88$</td>
<td>$S_{\mu-\log_{10}(N_0)}=2.69e3$</td>
</tr>
<tr>
<td></td>
<td>$S_{\mu-\log_{10}(N_0)}=1.64$</td>
<td>$S_{\mu-\log_{10}(N_0)}=1.88$</td>
<td>$S_{\mu-\log_{10}(N_0)}=1.68$</td>
<td>$S_{\mu-\log_{10}(N_0)}=2.04$</td>
</tr>
<tr>
<td>-6°C</td>
<td>$V=4.53e2$</td>
<td>$V=1.30e3$</td>
<td>$V=2.09e3$</td>
<td>$V=1.67e4$</td>
</tr>
<tr>
<td></td>
<td>(1.71e-3,1.19e-3,1)</td>
<td>(8.75e-4,4.58e-4,1)</td>
<td>(1.23e-3,7.37e-4,1)</td>
<td>(7.25e-4,3.37e-4,1)</td>
</tr>
<tr>
<td></td>
<td>$S_{\mu-\lambda}=8.41e2$</td>
<td>$S_{\mu-\lambda}=1.17e3$</td>
<td>$S_{\mu-\log_{10}(N_0)}=1.59$</td>
<td>$S_{\mu-\log_{10}(N_0)}=2.96e3$</td>
</tr>
<tr>
<td></td>
<td>$S_{\mu-\log_{10}(N_0)}=1.44$</td>
<td>$S_{\mu-\log_{10}(N_0)}=1.59$</td>
<td>$S_{\mu-\log_{10}(N_0)}=1.66$</td>
<td>$S_{\mu-\log_{10}(N_0)}=2.15$</td>
</tr>
</tbody>
</table>
**Figure 3.1**: Maximum radar reflectivity from NEXRAD (a) and three bin schemes: CNNB (b), FSBM (c), UPNB (d) reproduced from Fig. 3 in Xue et al. (2017a) and (e) The distribution of averaged maximum radar reflectivity along the line perpendicular to convective lines for observation from NEXRAD (black) and three bin schemes: CNNB (red), FSBM (green) and UPNB (blue). The purple rectangle is the regions where the UND Citation aircraft flew.
Figure 3.1 (Cont.)

Constructed flight tracks are denoted in vertical bars for CNNB (red), FSBM (green) and UPNB (blue), respectively.
Figure 3.2: The frequency distribution (left inset plots) and horizontal profile (right inset plots)
Figure 3.2 (Cont.)

of temperature (left) and vertical velocity (right) along the flight track for observations and numerical simulations at three different temperature range: -16 °C (upper), -10 °C (middle) and -6 °C (lower). The different colors denote different data source, with OBS1 in black, OBS2 in gray, CNNB in red, FSBM in green and UPNB in blue.
Figure 3.3: 1 s, 5 s, 10 s and 20 s averages of PSDs during the -10 °C flight leg in OBS1, with different colors indicating their time in the flight leg. The median (black line) and statistical sampling uncertainties (error bars) are overlaid.
Figure 3.4: (a) the raw 10 s averaged PSDs (gray) with median PSD (blue) and the natural variability expressed by the spread from 10th-90th percentile (red error bar) highlighted, and (b) the equally realizable ellipsoid for gamma distribution parameter for the flight leg at -10 °C in OBS1.
Figure 3.5: The raw 10 s averaged PSDs (gray) with median PSDs (blue) and their spread from
Figure 3.5 (Cont.)
10th-90th percentile (red error bar) under similar environmental conditions for OBS1, OBS2, CNNB, FSBM and UPNB at different temperature: -16 °C (upper), -10 °C (middle) and -6 °C (lower). For CNNB, FSBM and UPNB, underlying median ice (red), snow (green), graupel (cyan) PSDs are also plotted.
Fig. 3.6: Mass-diameter relationships for all hydrometeor species in (a), terminal velocity as a function of particle mass for all hydrometeor species in (b) and the ice nucleation parameterizations used by these schemes (Meyers and Cooper) in (c).
Figure 3.7: Ellipsoid of equally realizable solution for the gamma distribution parameters under similar environmental conditions for OBS1, OBS2, FSBM and UPNB simulation at different temperature.
Figure 3.8: The 2D projection of ellipsoids in Fig.6 on the $\mu$-$\lambda$ phase space.
Figure 3.9: The 2D projection of ellipsoids in Fig.6 on the $\mu$-$\log_{10}(N_0)$ phase space.
Figure 3.10: Observed and simulated 0th-5th moments of PSDs, with different color representing different bin microphysical schemes for the flight altitude leg at -10 °C.
CHAPTER 4: STATISTICAL THEORY OF CLOUDS PSD

This chapter is submitted as an article to the Journal of the Atmospheric Sciences by Wu and McFarquhar (2017). I developed the theory and analyzed the observational dataset to examine the theory as a part of my work, so it is included as a chapter of the dissertation. Due to the numerous mathematical symbols used in this section, Table 4.1 lists all the symbols and their definitions.

4.1 Introduction

Various analytical forms of cloud particle size distributions (PSDs), such as exponential (Marshall and Palmer 1948), gamma (e.g., Borovikov 1963; Ulbrich 1983), lognormal (e.g., Feingold and Levin 1986; Tian et al. 2010) and Weibull distributions (e.g., Zhang and Zheng 1994; Liu et al. 1995), have been used in numerical models and remote sensing retrieval algorithms. These functional forms of the distribution and the choice of free parameters characterizing the distribution have been typically determined on the basis of what provides the best match to in-situ observations. The scaling technique, as an alternative approach to describe cloud PSDs, has been used recently to derive parameters characterizing a PSD by assuming a limited number of degrees of freedom and a “universal distribution” without stating its exact analytical form (e.g. Testud et al. 2001; Lee et al. 2004). Without considering the number of degrees of freedom needed to characterize a PSD, determining the analytical form of the “universal distribution” used in the scaling approach is a challenging question. Although many different analytical forms of cloud PSDs have been proposed, no study has yet provided an adequate physical explanation as to why a certain functional form is preferred over another. Therefore, the choice of a functional form varies from study to study, complicating the
A theoretical way to find an analytical form of a PSD is to solve the general dynamic equation describing the particle system, given by

\[ \frac{\partial \bar{n}(v, t)}{\partial t} = -\bar{n}(v, t) \int_0^{+\infty} \bar{K}(v, u)\bar{n}(u, t) \, du + \frac{1}{2} \int_0^v \bar{K}(u, v-u)\bar{n}(u, t)\bar{n}(v-u, t) \, du + \int_0^{+\infty} \bar{L}(v, u)\bar{n}(u, t) \, du - \frac{\bar{n}(v, t)}{v} \int_0^v u\bar{L}(v, u) \, du + \overline{SC}(v, t) - \overline{SK}(v, t) \]  

where \( \bar{n}(v, t) \) is the number distribution function for particles with volume \( v \) at time \( t \), \( \bar{K}(u, v) \) and \( \bar{L}(u, v) \) are the collection kernel and breakup kernel for particles with volumes \( v \) and \( u \), \( \overline{SC}(v, t) \) is the source term, and \( \overline{SK}(v, t) \) the sink term. The vector form of the equation is used here because the particle system may contain several types of particles, such as ice particles with varying shapes and liquid particles. For the particle system of a single species (e.g. purely liquid clouds), a scalar form of Eq (4.1) will be sufficient. If Eq (4.1) is solved for aerosols undergoing coagulation growth with a constant coagulation kernel for Brownian motion and no source or sink terms, a lognormal distribution arises for the number distribution function (Park et al. 1999; Otto et al. 1999). However, even for the simplest case of liquid clouds without nucleation, precipitation and breakup, no analytic form for a cloud PSD has been found when a geometric collection kernel is used (Drake 1972). When more complex processes acting in ice or mixed phase clouds are considered (e.g., sublimation, aggregation, melting, riming, deposition, etc.), the equation is even more difficult to solve and an analytic solution cannot be found at this time. Because analytic solutions have not been possible, numerical methods have been used to determine PSDs in bin resolved models. However, the derived PSDs are very sensitive to even
the representation of processes in liquid-phase clouds, such as the choice of raindrop breakup kernels (Srivastava 1971, 1982; List and McFarquhar 1990; Hu and Srivastava 1995; McFarquhar 2004), with the collision-induced breakup size distribution determining the shape of the modeled PSD. There are sensitivities to the representation of even more processes for ice or mixed-phase clouds.

A statistical theory is another viable way to determine the form of PSDs. Here the mass or size of every particle is considered as a random variable acting under stochastic processes. One promising statistical theory for determining cloud PSDs is the principle of maximum entropy (MaxEnt, Jaynes 1957a, b). MaxEnt theory states that for a group of probability density functions (PDFs) that satisfy given properties of the variable, the PDF with largest information entropy for this variable should be chosen. Thus, a uniform distribution function (most uncertain) is selected if no other properties are specified. But, if the mean of the distribution is prescribed, the exponential distribution is the most probable distribution, following the same logic as used to derive the Maxwell-Boltzmann distribution in statistical mechanics. If both the mean and variance are prescribed, the most probable distribution is the normal distribution. The concept of MaxEnt has been used widely in physics (e.g., Rose et al. 1990; Antoniazzi et al. 2007), mechanical engineering (e.g., Sellens and Brzustowski 1985; Li et al. 1991; Berger et al. 1996), image processing (e.g., Wernecke and D’Addario 1977; Skilling and Bryan 1984), machine learning (e.g., Rosenfeld 1996; Berger et al. 1996), ecology (e.g., Phillips et al. 2004, 2006; Banavar et al. 2010), economics (e.g., Cozzolino and Zahner 1973; Buchen and Kelly 1996), and even in atmospheric sciences for representing cloud microphysics (e.g., Zhang and Zheng 1994; Liu et al. 1995; Yano et al. 2016) to turbulent flows (e.g., Majda and Wang 2006; Craig and Cohen 2006; Verkley and Lynch 2009). Its use in the study of spray PSDs in mechanical and
material engineering is closely related to its use in the study of cloud PSDs. Li and Tankin (1987), Dumouchel (2006), and Lecompte and Dumouchel (2008) employed MaxEnt to derive analytical forms of spray PSDs, and D´echelette et al. (2011) has a comprehensive review on this topic. Some applications of statistical mechanics may not state the MaxEnt explicitly, but similar methods have been employed by Griffith (1943) to explain the particle size distribution in a comminuted system, and by Lienhard (1964) to explain the unit hydrograph in hydrology. Thus, they are considered the same approach.

The problem of determining PSDs in cloud physics is very similar to the problems in these other fields. For numerical models simulating clouds with bulk microphysics schemes, only a number of moments of the PSD are predicted. For example, many schemes prognose the mass and number concentration. Other moments of a PSD are then calculated using the assumed form of the PSD and assumptions about various constants describing these distribution forms (Thompson et al. 2004; Morrison et al. 2005; Seifert and Beheng 2006; Morrison and Milbrandt 2015). These other moments include radar reflectivity and extinction. Thus, for developing parameterizations of cloud microphysics, there are some constraints on the properties of PSDs, exactly the type of scenario where MaxEnt can be used. Using MaxEnt, Zhang and Zheng (1994) and Liu et al. (1995) introduced the Weibull distribution as the analytical form of PSDs using constraints on the surface area and mass, respectively. Their derived PSD forms differ on the parameters characterizing the Weibull distribution. Yano et al. (2016) extended the assumptions about the PSDs to include constraints on the mean diameter and mass flux, and examined the impact of these assumptions using idealized simulations, and laboratory and observational datasets. All prior studies applying MaxEnt to cloud PSDs used the Gibbs/Shannon form of entropy. However, the Gibbs/Shannon entropy is not invariant under coordinate transformation,
and therefore contradictory results can be derived using the same assumptions. To solve these problems, a new formalism of entropy is needed as Jaynes (1963, 1968) noted. This paper applies this new form of entropy to cloud particle PSDs. The problem of Gibbs/Shannon entropy is discussed in Section 4.3 after a brief review of MaxEnt in Section 4.2. Based on the new formalism and several plausible assumptions about the cloud system, the four-parameter generalized gamma distribution is proposed as the most reasonable analytical form of cloud PSDs in section 4.4. The properties of the generalized gamma distribution are summarized in section 4.5. The applications of the four-parameter generalized gamma distribution to in-situ observed liquid and ice clouds PSDs are investigated in section 4.6. The principle findings of the study and directions for future work are summarized in section 4.7.

4.2. MaxEnt and its rationale for cloud physics

MaxEnt theory was first proposed by Jaynes (1957a, b) to explain the classical Maxwell-Boltzmann distribution. The same principle has also been applied to Fermi-Dirac statistics and Bose-Einstein statistics and non-equilibrium statistical mechanics (Jaynes 1963, 1968; Dougherty 1994; Banavar et al. 2010). In statistical mechanics, it is assumed that if there are \( N_i \) particles in the \( i \)th energy state \( E_i \), the total energy of the system \( E \) is given by

\[
E = \sum_{i=1}^{n} N_i E_i
\]  

(4.2)

where there are \( n \) total energy states with the total number of particles in the ensemble \( N \) given by the summation of all particles in each energy state expressed by

\[
N = \sum_{i=1}^{n} N_i
\]  

(4.3)

The number of microscopic configurations in which the \( N \) particles can be distributed over the \( n \) different energy states, \( W \), is given by
\[ W = \frac{N!}{N_1! N_2! \cdots N_n!} \]  

(4.4)

Boltzmann defined the entropy as \( S_B = k_B \ln(W) \), where \( k_B \) is the Boltzmann constant (Pathria and Beale 2011). \( S_B \) monotonically increases with \( W \) and is a measure of disorder: the greater the number of microscopic configurations in the system, the more uncertain the system can be. Using Sterling’s formula

\[
\ln(n!) = n\ln(n) - n + O(\ln(n))
\]

(4.5)

Boltzmann’s entropy becomes

\[
S_B = k_B \ln(W) = k_B \left[ \ln(N!) - \sum_{i=1}^{n} \ln(N_i!) \right] \approx k_B N \sum_{i=1}^{n} N_i \ln \left( \frac{N_i}{N} \right) = k_B N \sum_{i=1}^{n} p_i \ln(p_i) = NS
\]

(4.6)

where \( p_i = \frac{N_i}{N} \) is the probability of particles in every \( i \)th energy state, and \( S = -k_B \sum_{i=1}^{n} p_i \ln(p_i) \) is Gibbs’ form of entropy, which is the same form as Shannon’s information entropy except the inclusion of the Boltzmann constant (Shannon 1948). Assuming that there is a solution, denoted by \( \bar{N}_i \) (or \( \bar{p}_i \)) that maximizes \( W \) and therefore \( S \), it can be shown using Eq (4.6) and the definition of Boltzmann entropy that

\[
\frac{W_{max}}{W} = e^{k_B(S_{max} - S)}
\]

(4.7)

Since \( N \) is a very large number and \( k_B \) is a very small number in the context of statistical mechanics, \( W_{max} \) will be much larger than any other \( W \) achieved with other \( N_i \), indicating any other \( p_i \) that deviates from \( \bar{p}_i \) has significantly fewer microscopic configurations. For example, for a mole of gas, there are \( N_A \) (Avogadro constant, 6.02x10^{23} \text{ mol}^{-1}) particles, and the large ratio of \( W_{max} \) to other \( W \) rules out the possibility of other distributions.
The derivation of $N_i$ (or $\bar{p}_i$) is an optimization problem, expressed mathematically by

$$\text{argmax}_{N_i} \ln(W) \text{ subject to Eq (4.2) and Eq (4.3), that can be derived using the method of Lagrange multipliers, where}$$

$$d\ln(W) - \lambda_0 \left( \sum_{i=1}^{i=n} N_i - N \right) - \lambda_1 \left( \sum_{i=1}^{i=n} N_i E_i - E \right) = 0 \quad (4.8)$$

where $\lambda_0$ and $\lambda_1$ are the Lagrange multipliers. By using Stirling’s approximation, Eq (4.8) becomes

$$\sum_{i=1}^{i=n} -\ln(N_i) dN_i - \lambda_0 \left( \sum_{i=1}^{i=n} dN_i \right) - \lambda_1 \left( \sum_{i=1}^{i=n} dN_i E_i \right) = \sum_{i=1}^{i=n} (-\ln(N_i) - \lambda_0 - \lambda_1 E_i) dN_i = 0 \quad (4.9)$$

so that

$$\bar{p}_i = \frac{\bar{N}_i}{N} = Ce^{-\lambda_1 E_i}, \text{ where } C = C_0 e^{-\lambda_0} \quad (4.10)$$

which is the Maxwell-Boltzmann distribution (Pathria and Beale 2011). The Lagrange multipliers $\lambda_0$ and $\lambda_1$ can be obtained by substituting $N_i$ in Eq (4.2) and Eq (4.3).

Based on the above arguments, Jaynes (1957a, b) argued that for a group of PDFs that satisfy the given properties of a variable $x$, the PDF with largest information entropy (Shannon 1948) should characterize the variable, with statistical mechanics being just one example of this principle applied to an ideal gas. The methodology can be generalized using a continuous distribution to characterize the variable

$$\text{argmax}_{P(x)} \int_0^\infty P(x) \ln P(x) \, dx$$

subject to $(nc + 1)$ constraints:

$$\int_0^\infty f_k(x) P(x) \, dx = F_k \quad (4.11)$$
where $P(x)$ is the probability that state variable $x$ will occur, and the $nc+1$ constraints are expressed in the form of fixed expectation of $f_k(x)$ with $k = 0, 1, 2..., nc$. For $k=0$, $f_0(x)=1$ and $F_0=1$ are chosen as the normalization condition for the PDF. Since the 0th constraint is valid for every PDF, only $nc$ other constraints need to be given explicitly. Therefore, the number of given constraints will be denoted as $nc$. To get the maximum of $S(x) = -\int_0^\infty P(x) \ln P(x) dx$ with these constraints, the method of Lagrange multipliers can be applied as before with the discrete sums so that the Lagrange function $L(x, \lambda_1, \lambda_2, \cdots, \lambda_k)$ is expressed by

$$L(x, \lambda_1, \lambda_2, \cdots, \lambda_k) \equiv -\int_0^\infty P(x) \ln P(x) - \sum_{k=0}^{nc} \lambda_k \left( \int_0^\infty f_k(x) P(x) dx - F_k \right)$$

(4.12)

where $k = 0, 1, 2..., nc$. Then the general result can be solved so that:

$$P(x) = \frac{1}{Z(\lambda_1, \lambda_2, \cdots, \lambda_m)} \exp \left( -\sum_{k=0}^{nc} \lambda_k f_k(x) \right)$$

(4.13)

where the partition function $Z(\lambda_1, \lambda_2, \cdots, \lambda_m) = \int \exp (-\sum_{k=0}^{nc} \lambda_k f_k(x))$.

Following this technique, the exponential distribution can be derived as the maximum entropy distribution if only the mean of the variable is known. The Weibull distribution can be further derived as the maximum entropy distribution if the mean of the power function of the variable is known. If both the mean and variance of a variable are known, the normal distribution will be the maximum entropy distribution. Similarly, the lognormal distribution will be derived if the mean and variance of the logarithm of the variable are known. Kapur (1989) describes commonly used PDFs and their constraints.

Statistical mechanics can be used to define the properties of a cloud just as it is used to define the properties of an ideal gas. Just as in thermodynamics where there are variables
describing the microscopic and macroscopic state of the ideal gas, there are variables describing 
the microscopic and macroscopic properties of clouds. The macroscopic states are mainly 
defined by the total number of cloud particles, the cumulative extinction (projected area) of all 
cloud particles, the bulk liquid or ice water content of all particles in a distribution and other bulk 
microphysical properties. The microscopic states are described by the size, area and mass of the 
individual hydrometeors. A key question in the application of statistical mechanics to 
distributions of cloud particles is how many particles are needed to make the method robust, 
because there are inevitably fewer cloud particles than gas molecules. If it is assumed that the 
total number concentration is \( N_t \), then the total number of cloud particles in a sample volume \( V \) 
is \( N = N_t V \). The volume should be sufficiently large to make \( N \) large, but at the same time, not 
so large to exceed the typical volume of a cloud or a scale where there is a lot of horizontal or 
vertical inhomogeneity. Here a unit cloud volume (\( V \)) of 100m x 100m x 10m = 10^5 m^3 is 
proposed as large enough. Assuming a concentration of \( N_t \approx 100 \text{ cm}^{-3} \), then \( N = N_t V = 10^{13} \) 
should be big enough to make the number of particles sufficiently large. This volume is also 
small enough compared to typical model grid volume or radar sample volumes.

In cloud physics, the number distribution function is expressed as \( N(D) \), which can be 
normalized by \( N_t = \int_0^\infty N(D)dD \) to define the number distribution probability density function 
expressed by

\[
\begin{align*}
P(D) &= \frac{N(D)}{N_t}.
\end{align*}
\] 

Thus, the MaxEnt approach can be applied in the study of cloud PSDs, and its use in cloud 
physics has been discussed by Zhang and Zheng (1994), Liu et al. (1995) and Yano et al. (2016). 
However, there are problems directly applying the MaxEnt to cloud PSDs as discussed in section 
4.3. Previous studies chose the particle diameter (\( D \)) or particles mass (\( m \)) as the state variable \( x \),
assuming two constraints: 1) total particle number concentration; and 2) mean diameter (Yano et al. 2016), total surface area (Zhang and Zheng 1994), total bulk water content (Liu et al. 1995), or mass flux (Yano et al. 2016). The derived PSD forms maximizing the entropy are then special cases of Eq (4.13), with \( n_c = 1 \), expressed by

\[
P(x) = \frac{1}{z(\lambda)} \exp\left(-\lambda_1 f_1(x)\right)
\]

(4.15)

here \( x \) could be \( D \) or \( m \), and \( f_1(x) \) could be \( D, A, m \), or \( mv \) (\( v \) is the fall speed of a particle) that is a power function of \( x \). Note that the PDF over size and the PDF over other state variable \( x \) can be converted, so that the number distribution function can be expressed

\[
N(D) = N_t P(D) = N_t P(x) \frac{dx}{dD} = \frac{N_t}{z(\lambda)} \exp\left(-\lambda_1 f_1(x)\right) \frac{dx}{dD}
\]

(4.16)

Usually, the state variable \( x \) and the particle diameter \( D \) are assumed to be related through a power law (e.g., \( x = aD^b \)), so that Eq (4.16) can be rewritten as

\[
N(D) = \frac{N_t ab}{Z(\lambda)} D^{b-1} \exp\left(-\lambda_1 f_1(aD^b)\right)
\]

(4.17)

4.3. Problems using Gibbs/Shannon entropy and the concept of relative entropy

Eq (4.17) is a general solution for the functional form of cloud PSDs maximizing the entropy content as long as one constraint is given explicitly. However, it can be shown that different PSD forms can be derived using the same constraint. For example, below it is shown that the same constraints used in Liu et al. (1995) can be employed to derive a different PSD than the one they derived. It should be noted that Liu et al. (1995) is just chosen as a random example and all the forms derived in Zhang and Zheng (1994) and Yano et al. (2016) suffer the same problems. Assuming that the total bulk number concentration \( N_t \) and total bulk water mass
content $TWC$ are constraints and using mass $m$ as the variable characterizing particles, Liu et al. (1995) showed that the MaxEnt distribution was given by

$$N(m) = C_1 \exp(-\lambda_1 m)$$

(4.18)

where $C_1 = \frac{N_t^2}{TWC}$ and $\lambda_1 = \frac{N_t}{TWC}$ are the distribution parameters. This distribution can be rewritten in term of particle size ($D$) using an assumed mass-dimensional relation $m = \alpha D^\gamma$ as

$$N(D) = \overline{C_1} D^{-\beta-1} \exp(-\overline{\lambda}_1 D^\beta)$$

(4.19)

where $\overline{C_1} = \frac{\alpha \beta N_t^2}{TWC}$ and $\overline{\lambda}_1 = \frac{\alpha N_t}{TWC}$ are the distribution parameters.

However, if the maximum dimension $D$ is instead used to characterize the PDFs, and the same two constraints are applied as expressed by

$$\begin{align*}
\int_0^\infty N(D) \, dD &= \int_0^\infty N_t \, P(D) \, dD = N_T \\
\int_0^\infty \alpha D^\beta N(D) \, dD &= \int_0^\infty \alpha D^\beta N_t \, P(D) \, dD = TWC
\end{align*}$$

(4.20)

(4.21)

the MaxEnt distribution becomes

$$N(D) = \overline{C_1} \exp(-\overline{\lambda}_1 D^\beta)$$

(4.22)

where $\overline{C_1} = N_t \beta \left(\frac{\alpha N_t}{\beta TWC}\right)^{1/\gamma}$ and $\overline{\lambda}_1 = \frac{\alpha N_t}{\beta TWC}$ are the distribution parameters. By comparing Eq (4.19) and Eq (4.22), it is found that two different analytical forms of PSDs can be derived using the same assumption. In fact, a different analytical form of the PSD can be derived whenever the state variable $x$ characterizing the cloud particle changes. This is due to the fact that the Gibbs/Shannon entropy is not invariant under transformation of variables (Jaynes 1963, 1968).

Thus, Jaynes (1963, 1968) proposed another definition of entropy, typically called relative entropy, that makes entropy invariant under variable transformations.
The new definition of entropy proposed by Jaynes (1963, 1968), \( S_r \), is expressed by
\[
S_r(x) = -\int_0^\infty P(x) \ln \frac{P(x)}{I(x)} \, dx, \tag{4.23}
\]
where \( I(x) \) is called the invariant measure, or a prior distribution that represents an initial guess of what the distribution should be. When \( I(x) \) is a uniform distribution, the new definition of entropy is identical to the Gibbs/Shannon entropy minus a constant. The new entropy \( S_r(x) \) has also been called the relative entropy, or Kullback-Leibler divergence. It can be shown that \( S_r(x) \) is invariant under coordinate transformation \((x \to y, \text{where } y = g(x))\), because
\[
S_r(y) = -\int_0^\infty P'(y) \ln \frac{P'(y)}{I'(y)} \, dy = -\int_0^\infty P(x) \ln \frac{P(x)}{I(x)} \, dx = S_r(x), \tag{4.24}
\]
with \( P'(y) = P(x) \frac{dx}{dy} \) and \( I'(y) = I(x) \frac{dx}{dy} \).

To maximize \( S_r(x) \) with given constraints, the method of Lagrange multipliers is again used so that
\[
L \equiv -\int_0^\infty P(x) \ln \frac{P(x)}{I(x)} - \sum_{k=1}^{nc} \lambda_k \left( \int_0^\infty f_k(x) P(x) \, dx - F_k \right), \tag{4.25}
\]
where \( k = 0, 1, 2, \ldots, nc \) and the maximum (relative) entropy distribution is solved in the form
\[
P(x) = \frac{1}{Z(\lambda_1, \lambda_2, \ldots, \lambda_n)} I(x) \exp \left( -\sum_{k=0}^{nc} \lambda_k f_k(x) \right), \tag{4.26}
\]
where the new partition function is \( Z(\lambda_1, \lambda_2, \ldots, \lambda_n) = \int_0^\infty P_0(x) \exp \left( -\sum_{k=0}^{nc} \lambda_k f_k(x) \right) \).

4.4. Application to cloud PSDs

The new concept of entropy is invariant under a coordinate transformation, and the distribution derived maximizing this entity is consistent with the same constraint, regardless of
the variable used to characterize the PDF. However, before the theory can be applied to any system, the appropriate constraints and invariant measure \( I(x) \) are needed. These can only be obtained from an understanding of the system studied. To apply the theory to cloud physics, the first step is to determine the constraints for a cloud. Yano et al. (2016) used observed and simulated datasets to evaluate constraints of mean diameter, bulk extinction, bulk water content, and bulk mass flux. Here the use of different constraints is not examined, but instead the focus is upon the general application of the new definition of entropy. Unlike the Gibbs/Shannon entropy used in previous studies, the choice of state variable \( x \) is not important for \( S_r \), as the invariant measure \( I(x) \) will adjust accordingly. In this study, the particle diameter \((D)\) is chosen as the state variable \( x \) of the cloud. The number distribution function \( N(D) \), following Eq (4.26), can thus be expressed by

\[
N(D) = \frac{N_t}{Z(\lambda_1, \lambda_2, \ldots, \lambda_n)}I(D)\exp \left( -\sum_{k=0}^{n} \lambda_k f_k(D) \right)
\]

where the number of constraints is usually larger than 1. To make the above form simpler, it is assumed that there is only one additional constraint and the constraint function is the power law with particle diameter \((f(D) = aD^b)\) following Zhang and Zheng (1994), Liu et al. (1995), and Yano et al. (2016). Eq (4.27) then becomes

\[
N(D) = \frac{N_t}{Z(\lambda_1)}I(D)e^{-\lambda_1 ab} = \frac{N_t}{Z(\lambda_1)}I(D)e^{-\lambda b}
\]

where \( \lambda = \lambda_1 a \).

The next step applying MaxEnt theory is to determine the invariant measure \( I(D) \), which must be provided from a knowledge of the underlying physics. Jaynes (1968) provided guidelines to choose the invariant measure based on the transformation group, and Jaynes (1973) showed an example using the transformation group. The basic idea is that the shape of the
invariant measure should be invariant in different systems. In particular for this case, the shape of the invariant measure should not change with the volume of cloud studied. Two volumes of the same cloud are considered: cloud A with the total volume $V_A$ and cloud B, a subset of the cloud A, with total volume: $V_B = \kappa_1 V_A$, $\kappa_0 < \kappa_1 \leq 1$. Hereafter, the properties of cloud A and cloud B are denoted with the subscripts A and B respectively. Here, $\kappa_1$ cannot be too small since a large number of particles is needed for the application of statistical mechanics; hence $\kappa_0$ is used for a lower bound instead of 0. In our case, $\kappa_0 = 0.001$ should be enough, which will give the number of cloud particles in cloud B approximately $10^{10}$.

For volume $V_A$, the total mass is $TWC \times V_A$. Therefore, no particles larger than $D_{\text{maxA}}$ are possible, where $\rho \alpha D_{\text{maxA}}^\beta = TWC \times V_A$ with $\alpha$ and $\beta$ the $m-D$ relation parameters and $\rho$ the particle density. Thus, $I_A(D) = 0$ for $D > D_{\text{maxA}}$. For cloud A, the prior probability is $I_A(D)$ with $\int_0^{D_{\text{maxA}}} I_A(D) dD = 1$. For cloud B, the volume will be $V_B = \kappa_1 V_A$, $D_{\text{maxB}} = \kappa D_{\text{maxA}}$ ($\kappa = \kappa_1^{1/\beta}$), and the prior probability $I_B(D)$ satisfies $\int_0^{D_{\text{maxB}}} I_B(D) dD = 1$. A new scaled dimensionless variable $x = \frac{D}{D_{\text{max}}}$ is defined to scale $I_A(D)$ into the range [0,1] so that

$$\int_0^{D_{\text{maxA}}} I_A(D) dD = \int_0^1 I_A(xD_{\text{maxA}}) d(xD_{\text{maxA}}) = \int_0^1 D_{\text{maxA}} I_A(xD_{\text{maxA}}) dx = \int_0^1 f_A(x) dx = 1 \quad (4.29)$$

where $f_A(x) = D_{\text{max}} I_A(xD_{\text{maxA}})$, and similarly $f_B(x) = D_{\text{maxB}} I_B(yD_{\text{maxB}})$ where $y = \frac{D}{D_{\text{maxB}}}$. Because of scale invariance, the scaled PDFs $f_A(x)$ and $f_B(y)$ over the same range of [0, 1] should be the same, implying that

$$f_A(x) = f_B(x) \rightarrow D_{\text{maxA}} I_A(xD_{\text{maxA}}) = D_{\text{maxB}} I_B(xD_{\text{maxB}}) \rightarrow I_A(D) = \kappa I_B(\kappa D) \quad (4.30)$$
This is the scale invariance that the cloud system must satisfy in order for two different volumes to have the same shape of invariant measure. Following the formula for conditional probability, for any $D$ that is within the range of $(0, \kappa D_{maxA})$, it can be shown that

$$I_A(D) = I_B(D) \int_0^{\kappa D_{maxA}} I_A(u) \, du$$

Eq (4.31) is the standard conditional probability formula, and will hold whether or not any transformation invariance is assumed.

Combining the invariance requirement Eq (4.30) and the conditional probability relation Eq (4.31), then it is determined that

$$\kappa I_A(\kappa D) = I_A(D) \int_0^{\kappa D_{maxA}} I_A(u) \, du$$

Differentiating with respect to $\kappa$, and setting $\kappa = 1$ yields

$$I_A(D) + \frac{\partial I_A(D)}{\partial D} D = I_A(D) I_A(D_{maxA}) D_{maxA} \rightarrow \frac{\partial I_A(D)}{\partial D} D = (I_A(D_{maxA}) D_{maxA} - 1) I_A(D)$$

Solving the differential equation Eq (4.33), it can be shown that the most general solution is

$$I_A(D) = \frac{\mu + 1}{D_{maxA}^{\mu+1}} D^\mu$$

where $\mu = I_A(D_{maxA}) D_{maxA} - 1$ is a constant in the range of $1 < \mu < \infty$. The constant $\mu$ cannot be further determined by scale invariance. Using Eq (4.30), the invariant measure of cloud B is

$$I_B(D) = \frac{\mu + 1}{(\kappa D_{maxA})^{\mu+1}} D^\mu = \frac{\mu + 1}{D_{maxB}^{\mu+1}} D^\mu$$

The form of invariant measure provided by Eq (4.34) further satisfies translational, rotational and scale transformations, typical transformations between coordinate systems as suggested by Jaynes (1968). In the case of cloud PSDs, particle diameter is the only variable.
describing the PDF, so no rotational transformation exists. Since no spatial variables are involved, the PDF does not change when the coordinate system is translated. Scale transformation is the last transformation to be satisfied. For two coordinate system R and S, the length relates by $\kappa$ with $D_R = \kappa D_S$, the invariant measure for R is $I_R(D_R)$ with 
\[ \int_0^{D_{\text{max}}^R} I_R(D_R) dD_R = 1 \] 
and the invariant measure for S is $I_S(D_S)$ with 
\[ \int_0^{D_{\text{max}}^S} I_S(D_S) dD_S = 1. \]

Since it is the same cloud observed, the relation $I_R(D_R) dD_R = I_S(D_S) dD_S$ holds, which means $\kappa I_R(\kappa D_S) = I_S(D_S)$. Eq (34) clearly satisfies this relation. The invariant measure provided by Eq (4.34) satisfies all the Abelian group transformations proposed by Jaynes (1968).

If Eq (4.34) is assumed to represent the invariant measure, combined with Eq (4.28), the final $N(D)$ is the four-parameter generalized (or modified) gamma distribution, given by
\[ N(D) = N_0 D^\alpha e^{-\lambda D^\beta} \] 
(4.36)
where $N_0 = \frac{N_{cC}}{z(\lambda_1)}$.

It should be noted that the derived PSD forms from Zhang and Zheng (1994), Liu et al. (1995) and Yano et al. (2016) are all special cases of Eq (4.35), so that this study is consistent with but more general than previous studies. It is also clear now why the two approaches to derive the PSD form in section 3 generate different results. Eq (4.19) and Eq (4.22) differ by $D_\beta$, which is the invariant measure. The first approach assumed a uniform invariant measure over particle mass and the second assumed a uniform invariant measure over particle size, and $\frac{dm}{db} = \alpha \beta D^{\beta-1}$ is the difference.

4.5. Properties of generalized Gamma distribution

The properties of the generalized gamma distribution are summarized in this section. The generalized (or modified) gamma distribution is the most general form of a PSD, which can be
simplified to an exponential, gamma or Weibull distribution in special cases. To the authors’ knowledge, the generalized gamma distribution was first proposed by Amoroso (1925) to study the income distribution, and later independently proposed by Nukiyama and Tanasawa (1939) for fitting of the size distribution of sprays particles in mechanical and material engineering. Stacy (1962) studied the mathematical properties of the generalized gamma distribution, and the properties related to cloud PSDs will be summarized here.

The cumulative distribution function for generalized (modified) gamma distribution in the form of Eq (4.36) is:

\[
F(D; N_0, \mu, \lambda, b) = \frac{N_0}{b\lambda \Gamma \left(\frac{\mu + 1}{b}\right)} \gamma \left(\frac{\mu + 1}{b}, \lambda D^b\right)
\] (4.37)

where \(\gamma(s, x) = \int_0^x t^{s-1} e^{-t} dt\) is the lower incomplete gamma function. The \(n\)th moment can be calculated as:

\[
M_n = E(x^n) = \frac{N_0}{b\lambda \Gamma \left(\frac{\mu + 1 + n}{b}\right)} \Gamma \left(\frac{\mu + 1 + n}{b}\right).
\] (4.38)

For any variable \(x\) that is related to \(D\) through a power law (e.g., \(x = cD^d\)), it also follows a generalized gamma distribution, with the form:

\[
N(x) = \frac{N_0}{c^{(\mu+1)/d}} x^{\frac{\mu+1}{d}-1} e^{-\lambda x^{\frac{b}{d}}} \] (4.39)

One main benefit of the four-parameter generalized gamma distribution is that it is invariant under coordinate transformations of variable characterizing a PSD. The same form applies to all power law variables, such as particle diameter, area and mass. The lognormal distribution also has this property and this is one of the reasons Feingold and Levin (1986) recommended the lognormal distribution for PSDs. This property is not shared by the exponential distribution, gamma distribution and Weibull distribution. For example, Seifert and
Beheng (2006) assumed the commonly used three-parameter gamma distribution over mass, which will turn into a four-parameter generalized gamma distribution. A second benefit is that the generalized gamma distribution can also simplify to a gamma distribution, Weibull distribution or even exponential distribution under certain circumstances. Third, the physical meaning of distribution parameters is more clear than parameters used in some empirical distribution functions used in previous studies. Due to the properties mentioned above, Maur (2001) and Petty and Huang (2011) also proposed the use of generalized (or modified) gamma distribution without stating the underlying physical basis.

4.6. Testing with in-situ observed liquid and ice PSDs

In this section, in-situ observed PSDs are fit to different analytical forms, including the gamma, Weibull, lognormal and generalized gamma distribution. The fitting in this section is used to test the application of four-parameter generalized gamma distribution in real clouds.

An in-situ dataset collected by a two-dimensional cloud probe (2DC) and high sample volume spectrometer (HVPS) during the Midlatitude Continental Convective Clouds Experiment (MC3E, Jensen et al. 2016) are used for the fitting. Wu and McFarquhar (2016) describes how the data were collected and how the binary data were processed to generate cloud PSDs for the fit PSD. Two different distribution were used in the analysis: a one-minute time period in liquid clouds and another one-minute period in ice clouds. The particle images are all manually checked to make sure no mixed cloud particles exist in these two periods. Liquid PSDs measured between 13:20:00-13:20:59 at a temperature of around 4 °C are averaged, and best fits to the different analytical functions listed in Fig. 4.1 were performed. Following McFarquhar et al. (2015), the fitting technique minimized the $\chi^2$ difference between the fitted and observed moments of $N(D)$ defined by
\[ \chi^2 = \sum_{i=1}^{nm} \left( \frac{M_{fit,i} - M_{obs,i}}{\sqrt{M_{fit,i}M_{obs,i}}} \right)^2 \]

(4.40)

where \( M_{obs,i} \) is the \( i \)th moment of the observed PSD, and \( M_{fit,i} \) is the \( i \)th moment of the fit PSD. Here the 0th, 3rd and 6th moments corresponding to total number concentration, bulk liquid water content and radar reflectivity were used in the fitting procedure to determine the parameters of the gamma, Weibull and lognormal distribution. To determine the parameters of the generalized gamma distribution, the first moment, representing the mean particle size, was also used because four moments are required to describe the four parameters of the generalized gamma distribution. All the fitted functions have \( \chi^2 \) less than 0.001 in Eq (4.40), indicating all fits provide good agreement between fit and measured moments. Further, the fit gamma, Weibull and generalized gamma distribution all appear visibly similar to the observed PSD, while the lognormal fit seems to deviate further from the observed PSDs. The fit generalized gamma distribution has a \( b \) parameter very close to 1 (0.99), so the fit curve is very close to the gamma distribution. This implies that mean diameter is the constraint for liquid clouds in this time period. However, due to the discrete nature of the observed PSDs, directly deriving the parameters by fitting observed PSDs having large uncertainties into a generalized gamma distribution carries forward those uncertainties. As McFarquhar et al. (2015) showed, a volume of generalized gamma distribution parameters may fit the observed PSD equally well.

Fits to PSDs measured in ice clouds from 15:55:00-15:55:59 at a temperature of around -10 °C from the same flight was also conducted, shown in Fig. 4.2. Here the 0th, 2nd and 4th moments were used in Eq (4.40) to determine the fit parameters. For ice clouds, these approximately correspond to the total number concentration, bulk ice water content and radar reflectivity, respectively. Similarly, an additional moment, the 1st moment, is used to find the
generalized gamma distribution fit parameters. The $b$ parameter in the generalized gamma distribution is 0.39 here, and the fitted curve is closer to the observed PSDs compared to the gamma distribution and Weibull distribution.

4.7. Conclusions and discussions

Several analytical forms of cloud PSDs, such as exponential and gamma distributions, have been assumed in numerical models and remote sensing retrievals in past studies. However, no satisfying physical basis has yet been provided for why any of these characterize PSDs. The use of the principle of maximum entropy (MaxEnt) to find analytical forms of PSDs was examined here, building upon its use in prior studies (Zhang and Zheng 1994; Liu et al. 1995; Yano et al. 2016). The main findings of this study are summarized as follows:

1). A new definition of entropy, $S_r = -\int_0^\infty P(x)\ln \frac{p(x)}{l(x)} dx$ which is invariant under coordinate transformations, was used to resolve an inconsistency in previous studies. The previous use of Gibbs/Shannon entropy allowed different PSD to be derived using the same constraint by simply using a different state variable $x$.

2). The new definition of relative entropy used in this study to determine a physical basis for a cloud PSD requires an assumption about an invariant measure $I(D)$, which is obtained from a physical understanding of the system studied. Here, it was shown that $I(D)$ can be obtained if invariance regarding group transformation is assumed.

3). Assuming that the microscopic state variables that characterize the properties of cloud particles (e.g., particle diameter, area, mass, fall speed) are related to each other through power laws, it was shown that if one constraint related to any state variable was assumed, a four-parameter generalized gamma distribution can be derived. The state variable that needs to be used as a constraint is not yet well determined.
4). It was shown that if one state variable follows the generalized gamma distribution, all state variables must also follow it.

5). Using the in-situ observed PSDs from optical array probes (OAPs), reasonably good fits to the observed PSDs can occur for all the analytical forms of PSDs, even though the fit of generalized gamma distribution is visibly better for the selected PSDs. Due to the discrete nature of observed PSDs and large uncertainties for OAPs, parameters derived by directly fitting have large uncertainties.

Although the MaxEnt approach provides a physical basis for the form of the generalized four-parameter gamma distribution, it does not determine the values of parameters \(N_0, \mu, \lambda\) and \(b\). These can only be determined using observational datasets. Among the four parameters, \(b\) is particularly interesting, since it implicitly implies what the constraint for the system is. Yano et al. (2016) provides a good approach to examine the assumptions of the constraint (and therefore the value of \(b\)) using the observational data; however, the results were inconclusive due to large uncertainties in the dataset used. Due to the large uncertainties in the OAPs, newer probes, such as the Holographic Detector for Clouds (HOLODEC, Fugal and Shaw 2009), may be more useful for this study due to accurate sample volume. The authors will investigate the HOLODEC dataset when it becomes available in the future study.

It should be noted that the generalized gamma distribution is derived when only one constraint of the power function of particle dimension is used. It is possible that more than one constraint exists in the clouds or that the constraint functions \(f_k(D)\) are not in power law form under certain circumstances. In those cases, the more general form of the cloud PSD (Eq 4.26) should be used. The full potential of MaxEnt will be realized after more understanding of the physical systems is gained. Due to the limitations of in-situ observed PSDs, the authors are also
developing idealized models to simulate the growth of cloud particles, which will be used to examine the theory from a process-oriented perspective.

4.8 Figures and Tables

Figure 4.1: Sample in-situ liquid PSD \(N(D_{\text{max}})\) as function of \(D_{\text{max}}\) (black) and fitted for gamma distribution (red), Weibull distribution (blue), lognormal distribution (cyan) and generalized gamma distribution (purple). The fitted parameters are listed in the legend.
Figure 4.2: Same as Fig. 4.1, but for ice PSDs.
### Table 4.1 List of symbols and their definitions

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>Prefactor of m-D relations</td>
</tr>
<tr>
<td>$\beta$</td>
<td>Power factor of m-D relations, also the power parameter in generalized gamma distribution in Eq (36)</td>
</tr>
<tr>
<td>$\gamma(s,x)$</td>
<td>Lower incomplete gamma function</td>
</tr>
<tr>
<td>$\Gamma(x)$</td>
<td>Gamma function</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>Scale factor between two length in two clouds</td>
</tr>
<tr>
<td>$\kappa_0$</td>
<td>Lower limit of $\kappa$</td>
</tr>
<tr>
<td>$\kappa_1$</td>
<td>Scale factor between two volume in two clouds</td>
</tr>
<tr>
<td>$\kappa_2$</td>
<td>Scale factor between two coordinate system</td>
</tr>
<tr>
<td>$\lambda$</td>
<td>The slope parameter in generalized gamma distribution in Eq (36)</td>
</tr>
<tr>
<td>$\lambda_1$</td>
<td>The Lagrangian multiplier for the first constraint</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>The Lagrangian multiplier relating to $\lambda_1$ by $\lambda_2 = \alpha \lambda_1$</td>
</tr>
<tr>
<td>$\lambda_3$</td>
<td>The Lagrangian multiplier relating to $\lambda_1$ by $\lambda_3 = \beta \lambda_1$</td>
</tr>
<tr>
<td>$\lambda_k$</td>
<td>The Lagrangian multiplier for the kth constraint</td>
</tr>
<tr>
<td>$\mu$</td>
<td>The shape parameter in generalized gamma distribution in Eq (36)</td>
</tr>
<tr>
<td>$\rho$</td>
<td>Particle density</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>The measure of goodness for a fit in Chi-square statistic</td>
</tr>
<tr>
<td>$a$</td>
<td>Prefactor of a general power law relations</td>
</tr>
<tr>
<td>$A$</td>
<td>The projected area of a cloud particle</td>
</tr>
<tr>
<td>$b$</td>
<td>Power factor of a general power law relations</td>
</tr>
<tr>
<td>$C$</td>
<td>The constant that relates to $\lambda_1$ through $C = C_0 exp(-\lambda_3)$ in Eq (10)</td>
</tr>
<tr>
<td>$C_0$</td>
<td>The constant in Eq (10)</td>
</tr>
<tr>
<td>$C_1$</td>
<td>Constant in Eq (18)</td>
</tr>
<tr>
<td>$\tilde{C}_1$</td>
<td>Constant in Eq (19)</td>
</tr>
<tr>
<td>$\tilde{C}_2$</td>
<td>Constant in Eq (22)</td>
</tr>
<tr>
<td>$D$</td>
<td>The maximum dimension of a cloud particle</td>
</tr>
<tr>
<td>$e$</td>
<td>Euler’s number, approximately equals 2.71828</td>
</tr>
<tr>
<td>$E_i$</td>
<td>The $i$th kinetic energy state</td>
</tr>
<tr>
<td>$E$</td>
<td>Total kinetic energy of the particle system</td>
</tr>
</tbody>
</table>
Table 4.1 (Cont.)

<table>
<thead>
<tr>
<th>Symbols</th>
<th>Definitions</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_A(x)$</td>
<td>The scaled invariant measure for $I_A(x)$</td>
</tr>
<tr>
<td>$f_B(x)$</td>
<td>The scaled invariant measure for $I_B(x)$</td>
</tr>
<tr>
<td>$f_i(x)$</td>
<td>The $i$th constraint as a function of $x$</td>
</tr>
<tr>
<td>$F_k$</td>
<td>The expected value of $f_k(x)$</td>
</tr>
<tr>
<td>$IWC$</td>
<td>Ice water content</td>
</tr>
<tr>
<td>$I(x)$</td>
<td>The invariant measure</td>
</tr>
<tr>
<td>$I_A(x)$</td>
<td>The invariant measure for cloud A</td>
</tr>
<tr>
<td>$I_B(x)$</td>
<td>The invariant measure for cloud B</td>
</tr>
<tr>
<td>$I_R(x)$</td>
<td>The invariant measure for coordinate system R</td>
</tr>
<tr>
<td>$I_S(x)$</td>
<td>The invariant measure for coordinate system S</td>
</tr>
<tr>
<td>$k$</td>
<td>Constraint number</td>
</tr>
<tr>
<td>$L(x, \lambda_1, \lambda_2, ..., \lambda_n)$</td>
<td>Lagrangian function</td>
</tr>
<tr>
<td>$LWC$</td>
<td>Liquid water content</td>
</tr>
<tr>
<td>$m$</td>
<td>The mass of a cloud particle</td>
</tr>
<tr>
<td>$n$</td>
<td>Number of energy state in the ideal gas system</td>
</tr>
<tr>
<td>$nc$</td>
<td>The number of constraints</td>
</tr>
<tr>
<td>$nm$</td>
<td>The number of moments used for fitting</td>
</tr>
<tr>
<td>$M_{obs,i}$</td>
<td>The $i$th moment of the observed PSD</td>
</tr>
<tr>
<td>$M_{fit,i}$</td>
<td>The $i$th moment of the fit PSD</td>
</tr>
<tr>
<td>$N$</td>
<td>Total number of ideal gas molecules</td>
</tr>
<tr>
<td>$N_0$</td>
<td>Generalized gamma distribution parameter in Eq (36)</td>
</tr>
<tr>
<td>$N(D)$</td>
<td>Number distribution function over size</td>
</tr>
<tr>
<td>$N(m)$</td>
<td>Number distribution function over mass</td>
</tr>
<tr>
<td>$N_i$</td>
<td>Total number of ideal gas molecules in energy state $E_i$</td>
</tr>
<tr>
<td>$N_c$</td>
<td>Total number concentration</td>
</tr>
<tr>
<td>$P_i$</td>
<td>Probability of ideal gas particles in energy state $E_i$</td>
</tr>
<tr>
<td>$P(x)$</td>
<td>Probability of $x$ state</td>
</tr>
<tr>
<td>Symbols</td>
<td>Definitions</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>$S$</td>
<td>Gibbs/Shannon entropy</td>
</tr>
<tr>
<td>$S_B$</td>
<td>Boltzmann entropy</td>
</tr>
<tr>
<td>$S_r$</td>
<td>Relative entropy</td>
</tr>
<tr>
<td>$TWC$</td>
<td>Total water content</td>
</tr>
<tr>
<td>$\nu$</td>
<td>Cloud particle fall speed</td>
</tr>
<tr>
<td>$W$</td>
<td>The multiplicity representing the number of microscopic configurations</td>
</tr>
<tr>
<td>$x$</td>
<td>A random state variable that describe the cloud particle</td>
</tr>
<tr>
<td>$Z(\lambda_1, \lambda_2, \ldots, \lambda_m)$</td>
<td>Partition function</td>
</tr>
</tbody>
</table>
CHAPTER 5: CONCLUSIONS AND DISCUSSIONS

5.1 Summary

This dissertation explores three scientific questions about clouds. First, uncertainties in the derived morphological parameters of cloud particles imaged by in-situ cloud probes installed on aircraft are studied and quantified. A new algorithm to calculate the maximum dimension of an ice crystal is developed, and an in-situ dataset collected using 2DC and HVPS installed on University of North Dakota Citation aircraft in the trailing stratiform region behind a MCS sampled during the Midlatitude Continental Convective Clouds Experiments (MC3E) is used to examine the impact of different definitions of maximum dimension on the derived PSDs and calculated bulk properties, including number concentration, ice water content, extinction, mass-weighted fall speed, and precipitation rate. Derived bulk properties vary depending on the definitions of $D_{\text{max}}$ because of discrepancies in the definition of $D_{\text{max}}$ used to characterize the PSDs and that used to describe the properties of individual ice crystals.

Second, the capability of state-of-the-art bin-resolving numerical models to replicate observed cloud microphysical properties in the trailing stratiform region of a MCS are examined. The simulated PSDs using three bin-resolved microphysical schemes are compared to those measured in-situ on 20 May 2011 using a two-dimensional cloud probe and a high volume precipitation spectrometer installed on the University of North Dakota Citation aircraft during MC3E. These three schemes are the Caltech-NCAR-NOAA Bin scheme (CNNB), the Fast Spectral Bin Model (FSBM) and the University of Pecs and NCAR Bin scheme (UPNB). The observed and simulated PSDs were then fit to gamma distributions. The variability of observed and simulated particle size distributions (PSDs) were quantified using ellipsoids in phase space of the gamma distribution parameter ($N_0, \mu, \lambda$) to represent volumes of equally realizable
solutions. The underlying microphysical processes that contribute to the differences between observed and simulated cloud properties are discussed.

Third, physical explanations as to why PSDs exhibit certain forms, such as exponential, lognormal and gamma distributions, are explored. Several analytical forms of cloud PSDs, such as exponential and gamma distributions, have been proposed based on in-situ observations and assumed in numerical models and remote sensing retrieval algorithms in the past. These analytical forms are empirical functions, and no corresponding physical explanations have been provided to support their use to characterize cloud PSDs. The use of the principle of maximum entropy (MaxEnt) to find analytical forms of PSDs is examined in chapter 4 building upon the use of this principle in prior studies (Zhang and Zheng 1994; Liu et al. 1995; Yano et al. 2016). Based on the principle of maximum entropy, the four-parameter generalized gamma distribution is proposed to represent cloud PSDs, with assumptions of power law relations between state variables, scale invariance and a further constraint on the expectation of one state variable.

5.2 Main conclusions

The main conclusions in this dissertation are:

1). The differences in the number distribution functions \( N(D_{\text{max}}) \) derived using various definitions of \( D_{\text{max}} \) can differ by up to a factor of 6 for \( D_{\text{max}} < 200 \ \mu m \) and \( D_{\text{max}} > 2 \ mm \). The large differences are caused by the use of different definitions, the strong dependence of sample volume on the particle size for small particles, as well as the small number of large particles collected.

2). Number-weighted and mass-weighted mean diameter calculated using alternate definitions of \( D_{\text{max}} \) vary from 56% to 140% and 65% to 125% of those calculated using the diameter of the smallest enclosing circle \( (D_{S}) \), respectively. The difference in derived ice water
content can differ from 50% to 150% and the mass-weighted fall speed can vary from 28% to 180% depending on the definitions of $D_{\text{max}}$ used. The precipitation rate (mass flux) based on the above ice water content and terminal velocity can differ from 20% to 250. The extinction determined using different definitions $D_{\text{max}}$ can range from 60% to 133% to that computed using $D_s$. The effective diameter computed using different definitions of $D_{\text{max}}$ can range from 82% to 120% of that determined using $D_s$. Higher moments of PSDs have larger differences between the different definitions of $D_{\text{max}}$ than do the lower order moments of the PSDs.

3). Definitions that involve considerations of maximum dimensions in at least two directions are needed to get a reasonable estimate of the $D_{\text{max}}$. The $D_s$ proposed in this study is an attractive choice for $D_{\text{max}}$ due to the invariant properties with respect to orientations in the imaging plane.

4). The simulated ice cloud PSDs are generally multi-modal, unlike the observed distributions which have singular modes. The multimodality in the PSDs simulated by the CNNB, FSBM and UPNB microphysical schemes is due to the artificially separated ice species in numerical models.

5). The variability of PSDs as defined by the $10^{th}$ to $90^{th}$ spread of 10s averages are generally larger than those measured in-situ. The volume of generated equally realizable ellipsoids changes with altitude in the observations, while the trend is much smaller in all the simulations. The slope of the major axis of generated equally realizable ellipsoids also changes with altitude in the observations, while the slopes are quite consistent in different temperatures in all simulations.

6). The simulated lower moments of the PSDs are close to the observations, while higher moments in the simulations are larger compared to the observations. The differences in PSDs
among the three bin schemes and between the simulations and observations are due to the assumptions about the particle properties, such as mass/terminal velocity-dimensional relations, and representations of microphysical processes in different bin schemes, such as manually separated ice species, conversion between the ice species, nucleation, diffusional growth and aggregation growth.

7). A new definition of entropy with invariant measure \( S_r = \int_0^\infty P(x) \ln P(x) P_0(x) dx \) was used to resolve an inconsistency in previous studies, whereby the use of the Gibbs/Shannon entropy gives will derive different PDFs using the same constraints by using a different state variable \( x \).

8). The new definition of relative entropy used in this study to determine a physical basis for a cloud PSD requires an assumption about an invariant measure \( I(D) \), which is obtained from a physical understanding of the system studied. Here, it was shown that \( I(D) \) can be obtained if invariance regarding group transformation is assumed.

9). Since the microscopic state variables characterizing the properties of cloud particles (e.g., particle diameter, area, mass, fall speed) are assumed to be related to each other through power laws, it was shown that if one constraint related to any state variable was assumed, a four-parameter generalized gamma distribution can be derived. The state variable that needs to be used as a constraint is not yet well determined.

10). It is shown that if one state variable follows the generalized gamma distribution, all state variables must follow it. Direct fits to the in-situ observed PSDs using optical array probes (OAPs) are slightly better for the generalized gamma distribution, but the uncertainties in the in-situ observations are large to determine a better analytical form of cloud PSDs. Probes with newer technology, such as HOLODEC may be useful in the future.
5.3 Further studies

Due to the limitations of observation technology, the maximum dimension derived from two-dimensional images collected by OAPs may not represent a true maximum dimension for a three-dimensional particle, unless the maximum dimension is always in a plane perpendicular to the laser beams of OAPs. If the 3D structure of a single particle is observed in the future, the technique developed in chapter 2 for image processing to derive the particle morphological parameters can be naturally extended to three-dimensional space to determine the minimum enclosing sphere, minimum enclosing rectangle and minimum enclosing ellipsoids, which provide more accurate estimates of maximum dimension and aspect ratio of hydrometeors. The impact of particles falling with an angle between the horizontal plane and maximum dimension may need to be further investigated.

As for the statistical theory of PSDs, even though the MaxEnt approach provides a physical basis for the form of the generalized four-parameter gamma distribution, it does not determine the values of the parameters ($N_0$, $\mu$, $\lambda$ and $b$). These can only be determined using observational datasets. In particular, the values of parameters $\mu$ and $b$ that have influences on the theoretical models need to be determined accurately by using datasets from advanced probes. Yano et al. (2016) provides a good approach to examine the assumptions of constraint (therefore the value of $b$) using the observational data, but their results were inconclusive due to large uncertainties in the dataset used. Datasets from different sources should be explored to determine the value of these generalized gamma distribution parameters, such as the Holographic Detector for Clouds (HOLODEC, Fugal and Shaw 2009). It should be noted that the generalized gamma distribution is derived when only one constraint of a power function for particle diameter is used. It is possible that more than one constraint exists in nature or that the constraint functions $f_k(D)$
cannot be approximated well enough in a power law form. In either case, the more general form of cloud PSD (Eq 4.26) should be used then. The full potential of MaxEnt will be realized after a more thorough understanding of the physical systems is gained.

Due to the limitation of in-situ observed PSDs, the author is also developing an idealized model named ParticleSimulator to simulated the growth of large numbers of individual cloud droplets with diffusional growth, collision-coalescence growth, and break up processes acting. The future intent is to use this new model to examine the four-parameter generalized gamma distribution from a process-oriented perspective.
REFERENCES


