UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

OBSERVATIONAL STUDY OF RAIN MICROPHYSICS AND RAIN RETRIEVAL USING POLARIMETRIC RADAR DATA

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

QING CAO
Norman, Oklahoma
2009

OBSERVATIONAL STUDY OF RAIN MICROPHYSICS AND RAIN RETRIEVAL

USING POLARIMETRIC RADAR DATA

A DISSERTATION APPROVED FOR THE
SCHOOL OF ELECTRICAL AND COMPUTER ENGINEERING

BY

Dr. Guifu Zhang, Chair

Dr. Ming Xue

Dr. Robert Palmer

Dr. Yan Zhang

Dr. Terry Schuur

Acknowledgments

I am sincerely grateful of my advisor, Dr. Guifu Zhang, for his professional supervision throughout my Ph.D. study at university of Oklahoma. I have benefitted a lot from his insightful guidance and, moreover, his spirits of diligence and rigorousness set up a good paradigm for my future career. Personally, he and his family gave me lots of helps, especially in the hardest time of my life.

I would like to appreciate Dr. Robert D. Palmer, Dr. Ming Xue, Dr. Terry J. Schuur, and Dr. Rockee Y. Zhang who served on my doctoral committee. They have provided valuable expertise and helps for my research. I would also like to extend my thanks to faculty, staffs and students at ARRC, ECE and SoM, who have helped me in the past four years. Without their assistance, I would have experienced great difficulties. My special thanks are given to all my friends, with whom I have had a joyful life here at OU.

I would like to acknowledge that my research was supported by National Science Foundation Grant ATM-0608168. I am thankful of experts at NSSL and NCAR who have provided useful discussions, suggestions, data and programs. Particularly, I would like to thank Dr. Edward A. Brandes, Dr. Alexander V. Ryzhkov, Dr. Richard J. Doviak, Dr. Dusan S. Zrnic and Ms. Kyoko Ikeda.

Finally, I would like to express my heartfelt appreciation and love to my family for their encouragements, supports and love in my life. I would also like to cherish the memory of my dear father who gave me the sincerest love.

Contents

List Of Tables

List Of Figures

Abstract

Accurate quantitative precipitation estimation (QPE) and quantitative precipitation forecasts (QPF) require accurate microphysical parameterization/modeling of the precipitation. This study presents the modeling of rain microphysics and the application of polarimetric weather radar on rain retrievals. There are three topics addressed in this study. The first topic is the study of raindrop size distribution (DSD) through in-situ disdrometer observations. The observational error of disdrometer is quantified and corresponding error effects in developing DSD model are analyzed. The second topic is the characterization of rain microphysics with the linkage to radar observations. Empirical relations of rain-radar variables are developed for X-, C- and S-band polarimetric radars. The third topic is the retrieval of DSD parameters from polarimetric radar data (PRD). Three different approaches: direct retrieval, Bayesian retrieval and variational retrieval, are introduced. The latter two methods are promising with the optimal use of radar data. Their performance and potential are analyzed and discussed.

Chapter 1

Introduction

1.1 Motivation

Accurate quantitative precipitation estimation (QPE) and quantitative precipitation forecasts (QPF) require accurate characterization and parameterization of precipitation microphysics. For decades, the Doppler weather radar has played an important role in QPE and QPF. With the development of the dual-polarization technique, precipitation can be better studied through polarimetric radar observations. The knowledge and approaches in quantifying polarimetric radar measurements for precipitation have been highly demanded. However, research on this topic is seriously lacking so far. If the microphysical process of the precipitation can be well characterized/modeled, QPE and QPF would essentially benefit from the application of polarimetric radar measurements. Fig. 1.1 shows the schematic diagram, which links the related research area for QPE and QPF. It was proposed by Zhang et al. (2006b) in the National Science Foundation (NSF) grant "Improving microphysics parameterizations and quantitative precipitation forecast through optimal use of video disdrometer, profiler and polarimetric radar observations". The current study focuses on the areas with the blue shadow in the figure. Briefly speaking, they are observations, models, parameterizations, and retrievals of rain microphysics.

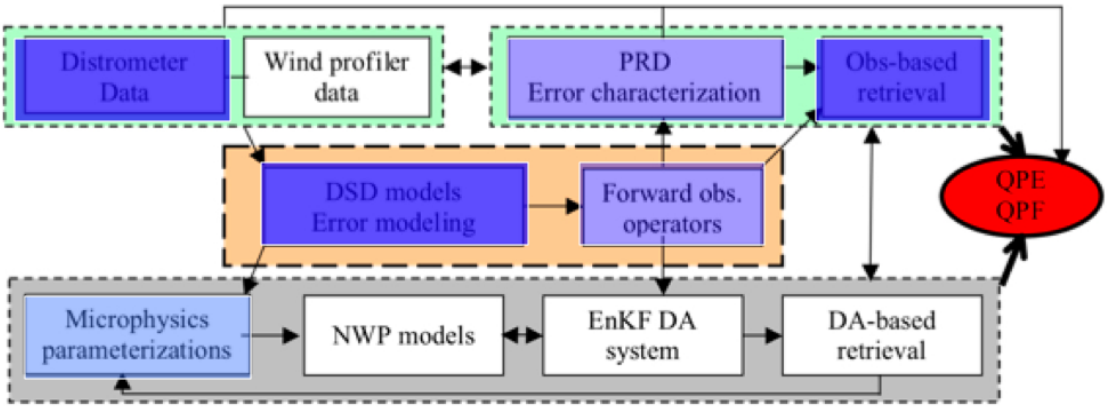

Figure 1.1: Schematic diagram linking together related area of research, from Zhang et al. (2006b).

Observational study is the basis of model developments and microphysical parameterizations, which are of great importance and required for rain retrievals. As we know, in-situ measurements from disdrometer provide information on individual drop sizes, shapes, and terminal velocities at a specific location. The disdrometer observations are related to those at the upper levels, helping to understand the dominant microphysical properties of the rainfall. Remote measurements from radar provide information on the bulk precipitation characteristics over a wide area. The microphysical parameterization/modeling based on disdrometer observations provides the basis for the application of radar measurements.

Raindrop size distribution (DSD) modeling is the major part of the model study. The DSD reveals the fundamental properties of rain microphysics. With a accurate DSD model, the rain properties could be well retrieved from the radar measurements. Generally, a good DSD model is required to represent natural variations well. It is known that the larger the order of a model freedom, the better the variation could be represented. However, the number of independent information from radar observations is limited and the error effect of radar observations could affect the DSD

retrieval. Complicated DSD models might not work in practical retrievals. An appropriate DSD model should keep a good balance between its flexibility and the capability of its parameters to be well resolved. The other necessary part of the model study is error analysis during the DSD modeling. Because the modeling is primarily based on in-situ measurements, observational errors should be taken into account and the error effect should be reduced as far as possible.

Numerical weather prediction (NWP) requires the microphysical parameterization of the precipitation. NWP is a key part for QPF, which is based on the data assimilation (DA) system. Generally, the characterization of rain microphysics through radar variables provides empirical but useful relations, which serve as good physical constraints in the NWP. The development of these relations is practically meaningful. Currently, S-, C-, and X-band polarimetric weather radars are popular platforms for weather services and researchers. The characterization of rain microphysics at these bands facilitates the application of polarimetric radar data (PRD) of different systems.

Rain estimation is one of major purposes of weather radar applications. Previously, rainfall rate was directly estimated with empirical radar-rain relations. With the development of DSD modeling, many efforts have been put into DSD retrieval. The DSD retrieval is obviously superior to the retrieval of integral parameters because DSD is more informative. Much progress has been made in retrieving DSD parameters from polarimetric radar measurements recently [e.g., Haddad et al. (1997); Bringi et al. (2002, 2003); Gorgucci et al. (2002); Brandes et al. (2004a,b); Vivekanandan et al. (2004); Zhang et al. (2001, 2006a)]. However, there are still issues to be resolved. For example, optimal use of radar observations in DSD retrieval is demanded given the fact that various radar variables have different error characteristics. For the

purpose of improving rain retrieval, this study presents three different DSD retrieval approaches and analyzes their performances.

1.2 Organization of Dissertation

The major purpose of this study is to investigate the properties of rain microphysics and develop appropriate physical models and algorithms for rain retrievals from polarimetric radar data. The rest of this study is organized as follows:

- Chapter 2 describes basic instruments, concepts and theories used in this study. The emphases are put on the raindrop size distribution and the polarimetric radar data, as well as the scattering theory linking them. This chapter addresses in-situ and remote sensing instruments, observed DSDs, conventional DSD models, rain and radar variables, raindrop shape models, scattering properties at different frequencies, and the simulation of radar variables based on the scattering theory.

- Chapter 3 discusses the error analysis and DSD parameter estimation. Firstly, observation errors of the disdrometer are quantified through side-by-side comparison. The error effect on DSD parameter estimation is analyzed, and different estimators are evaluated. A method of processing disdrometer data is introduced to reduce the error effect. The constraint-gamma (C-G) DSD model is refined and investigated on issues such as error propagation, low-end DSD truncation, and physical significance.

- In Chapter 4, the characterization of rain microphysics is given for S-, C- and X-band frequencies. The derivation of empirical relations, linking rain properties with radar variables, is based on the raindrop scattering model and DSDs measured by the disdrometer. A direct DSD retrieval approach is described

by applying the C-G DSD model. An adjustment of constraining relation is introduced to improve the DSD retrieval at the leading edge of convection.

- Chapter 5 introduces a Bayesian approach to retrieve DSD from PRD. This algorithm takes the observation error of PRD into account and gives an estimate of the uncertainty of retrieved results. The prior information of rain properties is derived from long-term disdrometer observations. The validity of this algorithm is demonstrated by the comparison of retrievals with in-situ measurements.

- In Chapter 6, a variational approach is applied to the DSD retrieval from polarimetric radar observations in the presence of attenuation. This algorithm applies the variational method to optimize the use of PRD and minimize the error effect. For an advantage, the attenuation correction is embedded into the forward operator and also optimized from the observations. The preliminary study gives a promising result. Besides, this chapter discusses the limitations and potential improvements of this algorithm.

- Chapter 7 summarizes works in this study and outlines future research.

Chapter 2

Basic Knowledge of Observational Study

2.1 Observational Instruments

Remote sensing instruments, such as weather radars, provide large coverage of precipitation observations. For example, the current National Weather Service (NWS) WSR-88D (Weather Surveillance Radar 1988 Doppler) network provides nationwide observations for weather services such as storm detection, forecast, and warning. The in-situ instruments, such as rain gauge and disdrometer, are helpful in studying rain microphysics and developing microphysical models for weather radar applications. In this study, the major in-situ instruments used are three two-dimensional video disdrometers (2DVD), which are operated by the University of Oklahoma (OU), National Center for Atmospheric Research (NCAR) and National Severe Storms Laboratory (NSSL), respectively. A disdrometer dataset of more than 3-years of observations are used to characterize rain microphysics and develop radar-rain retrieval algorithms. Weather radars used for this study are an S-band prototype polarimetric WSR-88D (KOUN), two X-band polarimetric IP1 radars of Center for Collaborative Adaptive Sensing of the Atmosphere (CASA), and Polarimetric Radar for Innovations in Meteorology and Engineering (OU-PRIME) operated by Atmospheric Radar Research Center (ARRC) at OU. In addition, the surface rain gauge measurements of Oklahoma Mesonet are used to verify rain retrieval algorithm as well.

2.1.1 In-situ instruments

Disdrometers, regarded as an effective tool for study of rain microphysics, are widely used by researchers. Currently there are three types of disdrometers popular in the world. The traditional one is the Joss-Waldvogel disdrometer (JWD). JWD is an impact-type disdrometer (Fig. 2.1a), which is designed based on the measurement of raindrop momentum. It has a $\sim$50 cm^2 sampling area and measures raindrop size from 0.3 to 5.5 mm with 20 size intervals. Its accuracy in measuring drop size is about 5% [Tokay et al. (2001)]. JWD has several shortcomings. The foremost one is that JWD has a "dead time" after the impact of a drop. In a heavy rain, JWD may miss some raindrops due to the dead time, leading to an underestimate of raindrop concentration. Its dynamic range of size measurement is also limited. Too small (< 0.3 mm) or too large (> 5.5 mm) raindrops are not measurable by a JWD. The concentration of small raindrops is normally underestimated.

Figure 2.1: (a)Joss-Waldvogel disdrometer RD-80 (From manufacturer's official website www.disdromet.com). (b) Laser optical Parsivel disdrometer (From manufacturer's official website www.ott-hydrometry.de).

The Parsivel disdrometer is a one-dimensional laser optical disdrometer (as shown in Fig. 2.1b). Besides size and raindrops, it can measure the fall velocity of hydrometeors. Its sampling area is also about 50 cm^2. The measurable particle size is 0.06–24.5 mm. The velocity range is 0.05–20.8 m s^{-1}. The Parsivel disdrometer still

has several shortcomings, though it performs better than JWD. It is only an optical device, which captures the one-dimensional projection of a particle. This may cause the biased estimation of particle size when the particle has a non-spherical shape. Moreover, a particle might be blocked by the other particle ahead of it in the light sheet, leading to the miscounting of particles. Generally, the Parsivel disdrometer may underestimate the fall velocity for middle size raindrops (1–3 mm) and numbers for small raindrops (< 1 mm).

Figure 2.2: The OEU (including sensor and computer) of OU 2DVD.

With the development of disdrometer technology, recently a two-dimensional video disdromer (2DVD) was introduced to the research community [Kruger and Krajewski (2002)]. It is capable of measuring particle size, shape, and falling velocity with an unprecedented accuracy. The 2DVD has a fine resolution. The resolution of prototype model (e.g., NSSL 2DVD) is 0.195 mm and the latest model (e.g., OU 2DVD) is 0.132 mm. For raindrop measurement, the 2DVD applies 41 size bins with 0.2 mm bin width. Compared to JWD and Parsivel disdrometer, the 2DVD has a relatively

larger sampling area ($\sim$100 cm^2). The typical sampling volume within a 1-min interval is about 3-5 m^3. Similar to other disdrometers, the 2DVD still has a relatively larger uncertainty in measuring small drops than mid-size drops. Typically, it underestimates the concentration of small raindrops ($<$ 1 mm). Compared to JWD and Parsivel disdrometer, its accuracy has been much improved.

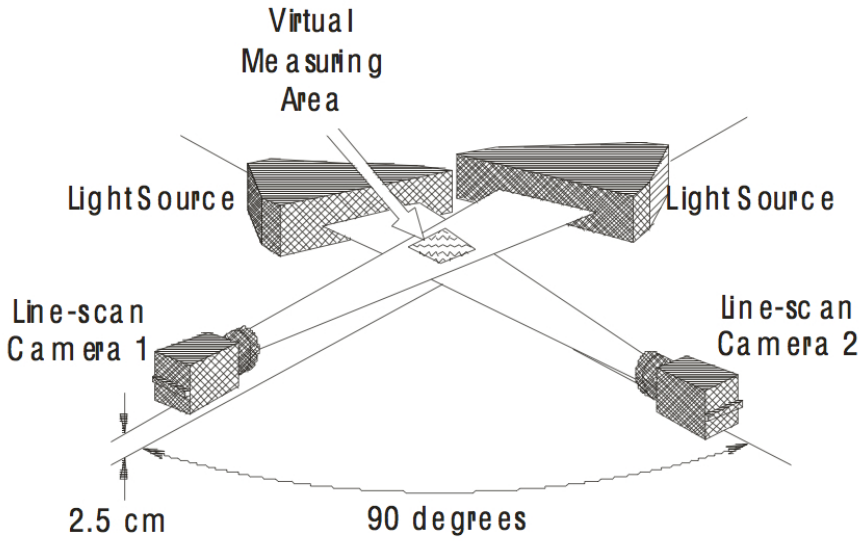

Figure 2.3: Raindrop measurement by 2DVD depends on two orthogonal horizontal light sheets, from Kruger and Krajewski (2002).

Fig. 2.3 shows the principle of 2DVD in measuring particles. There are two sets of light sources and cameras inside 2DVD's outdoor electronic unit (OEU), producing two orthogonal horizontal light sheets. Two lights sheets are spaced at 6–7 mm and the view area is about 10×10 mm^2. When a particle falls down through two light sheets. It blocks the light ways and two cameras capture its projections in two directions. The size and shape of a particle are then retrieved from its shadows captured by the cameras. The time, falling from one light sheet to the other light sheet, is used to calculate the fall velocity. Superior to Parsivel disdrometer, 2DVD can resolve two particles falling in the light sheets at the same time, though one particle might block

the other one in one direction.

Other in-situ instruments used for precipitation research include rain gauge. Standard rain gauge of Oklahoma Mesonet, used in this study, is an unheated tipping-bucket gauge with a sensor of 30.4 cm diameter sampling area (about 700 cm^2). The rainfall is measured at every 0.254 mm. The accumulated rainfall is recorded every 5 minutes. Therefore, rainfall might be recorded at every other 15-minute interval if the rainfall is very light. On the other hand, if the rain is very heavy, the rain gauge is likely to underestimate the total amount of rainfall because the rain might splash outside the bucket. The measurement error of rainfall is about $\pm 5\%$ with the rainfall rate range of 0-50 mm per hour.

2.1.2 Remote sensing instruments

Remote sensing instruments used in this study are three types of weather radars. KOUN, located at the north research campus of OU, is an S-band polarimetric radar with simultaneous transmission and receiving at horizontal (H) and vertical (V) channels. CASA IP1 network consists of 4 X-band polarimetric radars—KSAO, KCYR, KRSP, KLWE—located at 60-70 km southwest of KOUN. OU-PRIME is a C-band polarimetric radar made with the latest radar technique. It is located at the south research campus of OU. OU-PRIME finished its installation and began data collection in early 2009.

The key parameters of three radar systems are listed in Table 2.1. All three radars apply the simultaneous dual-polarization mode. Compared to KOUN and OU-PRIME, CASA IP1 radar is much smaller (with an antenna aperture of 1.2 m) and its antenna scanning is more agile (with an antenna accelerate rate of 50 $^\circ s^{-2}$).

Figure 2.4: Radar on the tower is KOUN, operated by NSSL. OU 2DVD and NCAR 2DVD are deployed inside the red fence in a field experiment. NSSL 2DVD is deployed aside in a basement covered by a shutter.

Figure 2.5: OU-PRIME, operated by ARRC, began data acquisition in January 2009 (From website http://en.wikipedia.org/wiki/OU-PRIME)

Table 2.1: System specifications of three dual-pol radars: KOUN, OU-PRIME, and CASA IP1

	KOUN	OU-PRIME	CASA IP1
Operating frequency (GHz)	2.7-3.0	5.51	9.3
Frequency band	S	C	X
Antenna aperture (m)	8.53	8.5	1.2
Antenna gain (dB)	45.5	50	38
3-dB beam width (degree)	1.0	0.5	1.8
Max. rotation rate (deg s^{-1})	36	36	35
Accelrate rate (deg s^{-2})	15	18	50
Transmitter type	Klystron	Magnetron	Magnetron
Peak power (kW)	750	1000	7.5
PRF (Hz)	320-1300	500-2000	1600, 2400
Pulse width (μs)	1.6, 4.5-5.0	0.4, 0.8, 1, 2	0.18, 0.66
Range resolution (m)	250, 1000	60, 125	48, 100
Dual-polarization	simultaneous	simultaneous	simultaneous

However, it has much less transmitter power and the peak power is only 1% of KOUN. OU-PRIME has the most narrow antenna beam width of 0.5 °, which is half of KOUN (1 °) and almost a quarter of CASA IP1 (1.8 °). Moreover, OU-PRIME has a digital-adjustable PRF with range from 500–2000Hz while KOUN and CASA IP1 only have several fixed PRFs to choose from. The typical PRF used by OU-PRIME is 1180 Hz, corresponding to a maximum unambiguous range of 127 km. KOUN and CASA IP1 radar generally have about 300 and 60 km unambiguous ranges, respectively. In addition, CASA IP1 can transmit pulses with the shortest width, having the finest range resolution of 48 m. Typically, range resolutions applied in practice for three radars

are 250 m, 125 m, and 100 m, respectively.

2.2 Raindrop Size Distribution (DSD)

2.2.1 DSD models

Functional relations are always applied to model DSD in the study/research of rain microphysics. The following models are DSD models generally used by researchers of the meteorological community in the world.

- M-P model [Marshall and Palmer (1948)]

$$N(D) = 8000\exp(-\Lambda D) \tag{2.1}$$

  It is a single parameter model with a slope parameter Λ. This well-known model was proposed by Marshall and Palmer (1948) and was widely used in the past 50 years. It was helpful in bulk-scheme rain parameterization and radar-rain estimation when single-polarization weather radars were prevailing.

- Exponential model

$$N(D) = N_0\exp(-\Lambda D) \tag{2.2}$$

  It is a two parameters model with a concentration parameter N_0 and Λ. It is more flexible than M-P model since the latter is equivalent to the exponential model with a fixed N_0. Besides the DSD, the exponential model can be used to model snow particle size distribution. It is especially appropriate for applications of dual-frequency/dual-polarization weather radars.

- Gamma model [Ulbrich (1983)]

$$N(D) = N_0 D^{\mu}\exp(-\Lambda D) \tag{2.3}$$

In addition to N_0 and Λ, gamma distribution introduces a shape parameter μ. It has been widely accepted that the gamma model is flexible to represent the variability of natural DSDs. As a result, a lot of researchers have applied gamma model to study the DSD since it was introduced to the metrological community by Ulbrich (1983). Some recent studies have applied the normalized gamma DSD. For example, Bringi et al. (2002) applied the gamma DSD as:

$$N(D) = N_w f(\mu)\Big(\frac{D}{D_0}\Big)^{\mu}\exp(-\Lambda D) \tag{2.4}$$

with

$$f(\mu) = \frac{6}{3.67^4}\frac{(3.67+\mu)^{\mu+4}}{\Gamma(\mu+4)} \tag{2.5}$$

$$\Lambda = \frac{3.67+\mu}{D_0} \tag{2.6}$$

where Λ is the slope parameter of gamma function. N_w is a normalized parameter, the value of which equals to the intercept parameter of an exponential DSD that has the same water content and the median volume diameter D_0. The N_0 in Eq. 2.3 is then equal to

$$N_0 = N_w f(\mu) D_0^{-\mu} \tag{2.7}$$

The advantage of a normalized gamma DSD is that parameters (N_w and D_0) have specific physical meaning.

- Log-normal model

$$N(D) = \frac{N_T}{\sqrt{2\pi}\sigma D}\exp\left(\frac{-[\ln(D)-\eta]^2}{2\sigma^2}\right) \tag{2.8}$$

where N_T is the total number concentration. η and σ are the mean and standard deviation of Gaussian distribution, respectively. This model follows the assumption that parameters of the DSD can be modeled as random variables from a multivariate Gaussian distribution. It has a good explanation of DSD based on the probability theory, and the mathematical calculation is not that complicated. However, it might not be the best one matching observed DSDs.

DSD models mentioned above have their own advantages and limitations. By comparison, the gamma distribution generally has the best performance in modeling observed DSDs. Fig. 2.6 shows an example of DSD models. The asters denote a DSD observed by 2DVD. Four lines represent fitted distributions by four models. It is evident that the M-P model has the worst performance. The DSD model, which has more freedom (i.e., more parameters), represent the observation better. The observed DSD in the figure is well modeled by the gamma function. This study focuses on the gamma model and its modification. Although the gamma distribution is flexible for modeling observed DSD, it is inevitable that the model error will be introduced. Next chapter will mention the error issues of DSD observations and its modeling.

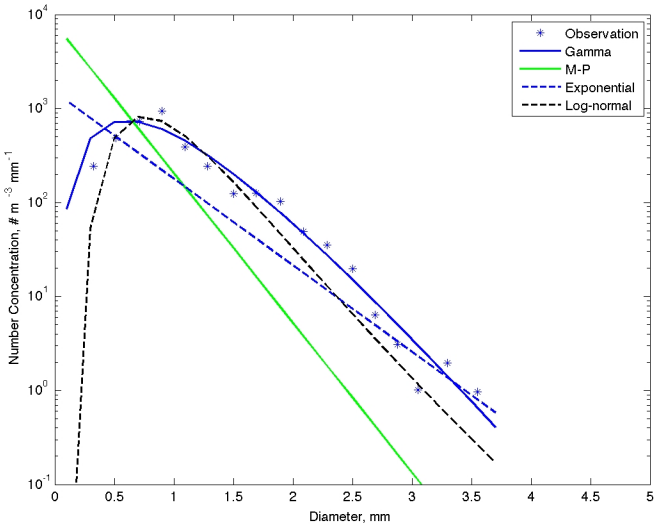

Figure 2.6: Example of an 2DVD-observed DSD, which is represented by different DSD models.

2.2.2 DSD observations

Different types of DSDs represent different rains. With the help of 2DVD, the DSD can be obtained with satisfactory accuracy. There are some examples of 1-min DSDs observed by the 2DVD. Fig. 2.7 shows several typical DSDs chosen from rain events

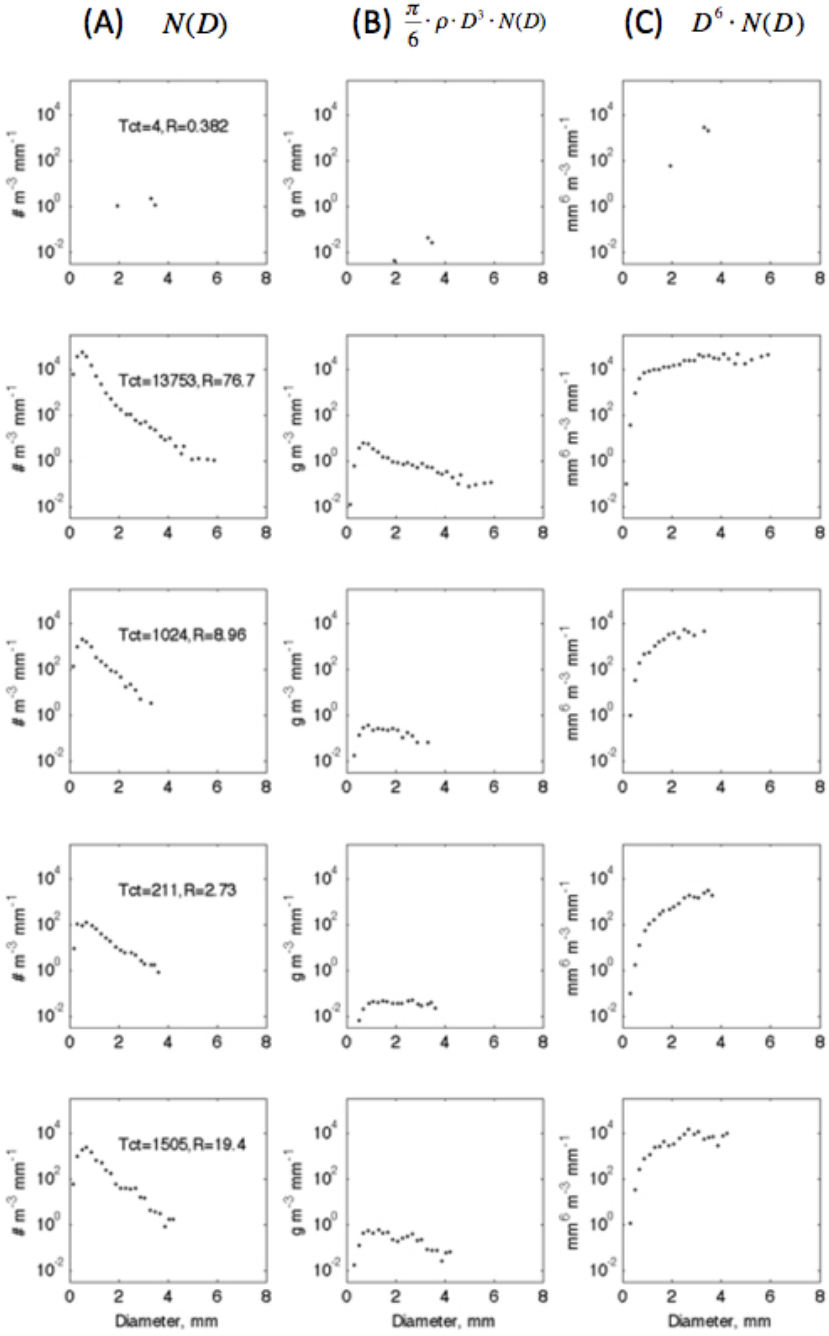

Figure 2.7: Five DSDs (column A) observed by the 2DVD at 1-min interval and bin size resolution of 0.2 mm. Column B represent the liquid water content distribution (see the label, where ρ is the water density). Column C is the 6^{th} order distribution of DSD. Five rows represent DSD1, DSD2, DSD3, DSD4, and DSD5, from top to bottom, respectively.

occurring on 21 July 2006 and 6 November 2006 [Zhang et al. (2008), Figs. 5 and 6]. In addition to each DSD, the corresponding liquid water content (proportional to the 3^{rd} order moment) and 6^{th} order moment per unit drop size interval are shown in the second and third columns. The first row shows a DSD (this type of DSD is referred to as "DSD1"), attributed to the raindrop sorting and usually observed during the leading edge of the convection. DSD1 generally has a very low total number concentration (N_T), total raindrop count (Tct) and water content (W). However, the mean volume diameter (D_m) is not small, sometimes even larger than 3 mm. This type of DSD is classified as a "big drop" DSD by the hydrometeor classification algorithm [e.g., Ryzhkov et al. (2005c)]. The second row shows a DSD observed in the region of a convective core (referred to as "DSD2"). DSDs in this region generally have a broad shape and very high N_T. The observed maximum diameters (D_{max}) are usually larger than 5 mm. DSD2 has many raindrops in all size categories so that rain-rate (R) and reflectivity factor (Z) generally have large values. The third row shows a DSD observed in the trailing stage of the convection (referred to as "DSD3"). This type of DSD still contains many small raindrops (<1 mm) but few large raindrops (>3 mm) such that the DSD appears to be truncated. The D_m is usually small (around 1 mm), although the R might be moderate (i.e., N_T is large). The DSD in the region of weak convection (not shown here) for this event also has similar properties to DSD3. The fourth row shows a DSD observed in the stratiform region (referred to as "DSD4"). DSD4 usually does not have small raindrops as many as DSD3 but tends to contain more large raindrops than DSD3, given the same R. Consequently, when both have the same R, DSD4 generally has a broader distribution but a lower N_T. The D_{max} of DSD4 is usually between 3 mm and 5 mm. The D_{max} of DSD3 is usually not larger than 4 mm. The fifth row shows a DSD (referred to as "DSD5"), which tends to have a bimodal distribution (although the bimodal distribution is not obvious for the DSD, it is evident for the water content distribution). In the region ahead of the convective

core, the DSDs observed within 1-min intervals sometimes have the tendency to be bimodal.

2.2.3 Rain variables

The raindrop size distribution is fundamental for rain microphysics because all the rain properties can be learned from the DSD. The following rain variables are all associated with the DSD.

$$\text{The } n^{th} \text{ moment} : M_n = \int_{D_{min}}^{D_{max}} D^n N(D) dD, [\text{mm}^n\,\text{m}^{-3}] \tag{2.9}$$

$$\text{Total number concentration} : N_T = M_0, [\text{m}^{-3}] \tag{2.10}$$

$$\text{Water content} : W = \frac{\pi}{6} \times 10^{-3} M_3, [\text{g}\,\text{m}^{-3}] \tag{2.11}$$

$$\text{Reflectivity factor} : Z \approx M_6, [\text{mm}^6\,\text{m}^{-3}] \tag{2.12}$$

$$\text{Rainfall rate} : R = 6 \times 10^{-3} \pi \int_{D_{min}}^{D_{max}} D^3 v(D) N(D) dD, [\text{mm}^n\,\text{h}^{-1}] \tag{2.13}$$

$$\text{Evaporate rate} : R_e = \int_{D_{min}}^{D_{max}} \frac{\delta M_e}{\delta t} N(D) dD, [\text{g}\,\text{m}^{-3}\,\text{s}^{-1}] \tag{2.14}$$

$$\text{Accretion rate} : R_a = \int_{D_{min}}^{D_{max}} \frac{\delta M_a}{\delta t} N(D) dD, [\text{g}\,\text{m}^{-3}\,\text{s}^{-1}] \tag{2.15}$$

$$\text{Mean diameter} : D_a = \frac{M_1}{M_0}, [\text{mm}] \tag{2.16}$$

$$\text{Effective diameter} : D_e = \frac{M_3}{M_2}, [\text{mm}] \tag{2.17}$$

$$\text{Mean volume diameter} : D_m = \frac{M_4}{M_3}, [\text{mm}] \tag{2.18}$$

$$\text{Median volume diameter [mm]} : \int_{D_{min}}^{D_0} D^3 N(D) dD = \int_{D_0}^{D_{max}} D^3 N(D) dD \tag{2.19}$$

$$\text{Mass-weighted terminal velocity} : v_{tm} = \frac{\int_{D_{min}}^{D_{max}} D^3 v(D) N(D) dD}{\int_{D_{min}}^{D_{max}} D^3 N(D) dD}, [\text{m}\,\text{s}^{-1}] \tag{2.20}$$

where, D is the equivalent diameter of raindrop. D_{max} and D_{min} are maximum and minimum diameters of the raindrop, respectively. $v(D)$ is the falling velocity of the raindrop. $\frac{\delta M_e}{\delta t}$ and $\frac{\delta M_a}{\delta t}$ are surface evaporation rate and accretion rate of a raindrop, respectively.

2.3 Polarimetric Radar Data

Polarimetric radar measurements are informative for interpreting properties of hydrometeors. Given a spheroid raindrop illuminated by the radar wave, the scattering energies at horizontal and vertical axes are different, leading to a measurement difference between the radar reflectivity at H and V channels. The dual-polarization difference helps to estimate the shape of the raindrop. This section addresses polarimetric radar data (PRD) used in the rest of this study. They are reflectivity factor ($Z_{H,V}$), differential reflectivity (Z_{DR}), co-pol correlation coefficient (ρ_{HV}), and specific differential phase (K_{DP}).

1. Reflectivity factor

   The reflectivity factor is a part of the radar reflectivity, which includes the term of radar constant as well. For convenience, the reflectivity factor is sometimes called the reflectivity. Reflectivity factors for horizontal (Z_H) and vertical polarization (Z_V) are defined as [Zhang et al. (2001)]:

   $$Z_{H,V} = \frac{4\lambda^4}{\pi^4 \left|K\right|^2} \left\langle n \left| F_{hh,vv}(\pi) \right|^2 \right\rangle \tag{2.21}$$

   where λ is the wavelength and $|K|^2$ is a dielectric term. $K = (\epsilon - 1)/(\epsilon + 2)$ and ϵ is the dielectric constant. $|K|^2$ for water is generally varies between 0.91 to 0.93 for a wavelength between 0.01 and 0.1 m. F_{hh} and F_{vv} are complex scattering amplitudes of a particle at horizontal and vertical directions, respectively. π means the direction of wave scattering is 180°, i.e., backscattering. n is the number concentration and the notation "$\langle\,\rangle$" denotes the expected value. There is $\langle nX \rangle = \int N(D) X dD$, which means the average of all particles, given a distribution within the sampling volume.

   Eq. 2.21 suggests that the reflectivity factor is proportional to the number concentration of hydrometeors within a radar resolution volume. Besides, it is

sensitive to the particle size. For example, for Raleigh scattering the reflectivity factor is about the 6^{th} moment of DSD. If the particle size is doubled, the reflectivity will have an increase of about 18 dB. The reflectivity factor is normally given in the unit of dBZ [$10\log_{10}(\mathrm{mm}^6\,\mathrm{m}^{-3})$]. The radar measurement error of reflectivity is normally about 1–2 dB.

2. Differential reflectivity

Differential reflectivity represents the ratio between reflectivity at the horizontal and vertical polarizations. It is given by (normally in unit of dB)

$$Z_{DR} = 10\log\frac{Z_H}{Z_V} = 10\log\frac{\left\langle n\left|F_{hh}(\pi)\right|^2\right\rangle}{\left\langle n\left|F_{vv}(\pi)\right|^2\right\rangle} \tag{2.22}$$

Since the differential reflectivity is the ratio of reflectivity measurements between H and V channels, it is insensitive to the absolute radar calibration of reflectivity. It is also insensitive to partial radar beam blockage. Moreover, differential reflectivity is independent of the concentration of scatters, but it is affected by the propagation effects such as the attenuation.

The differential reflectivity is informative for determining the raindrop size. The raindrop generally has an oblate shape. The larger the raindrop size, the more oblate the shape (e.g. see raindrop shape model in Fig. 2.10 of next section). Therefore, the reflectivity at horizontal direction is normally higher than the reflectivity at vertical direction. For light rains, raindrops are approximately spheric. The differential reflectivity, therefore, is small ($\sim$ 0 dB). For heavy rain with many large raindrops, the differential reflectivity increases. The differential reflectivity of rain is generally between 0 and 5 dB. Given a raindrop axis ratio model, the median size can be retrieved quantitatively through the

differential reflectivity measurement.

The differential reflectivity is also helpful for the classification of scatters other than raindrops. For example, dry hail, graupel and aggregated snow may tumble during their falling. Their differential reflectivities are approximately 0 dB with small positive or negative values. Melting hail and wet snow, less likely to tumble and having a larger dielectric constant because of the coated water, would have a larger differential reflectivity, generally 1–7 dB and 0.5–2.5 dB, respectively. Others scatters, such as insects and birds, may have an uncommon differential reflectivity for hydrometeors. The differential reflectivity is an important input parameter for current fuzzy-logic-based hydrometer classification algorithms [e.g., Ryzhkov et al. (2005b); Lim et al. (2005)].

3. Co-pol correlation coefficient

The co-pol correlation coefficient is an indicator of de-correlation between backscattering signals at horizontal and vertical polarizations. It is given by

$$\rho_{HV} = \frac{\left\langle nF_{hh}(\pi)F_{vv}^{*}(\pi)\right\rangle}{\sqrt{\left\langle n\left|F_{hh}(\pi)\right|^{2}\right\rangle\left\langle n\left|F_{vv}(\pi)\right|^{2}\right\rangle}} \tag{2.23}$$

Generally, the correlation coefficient would decrease when particles have irregular shapes or much uncertainty of canting angles. In addition, if there are mixed particles with different phases, the correlation coefficient would be reduced as well. The value is normally high for hydrometers that are oriented and smooth. For example, the correlation coefficient for raindrops is high, about 0.98–1. Small raindrops have a higher value than large drops with more oblateness. Hail and snow mostly have a lower correlation coefficient than the water. Correlation coefficients of graupel, dry hail, and wet hail are about 0.97–0.995,

0.85–0.97, 0.75–0.95, respectively. The crystal, dry aggregated snow and wet aggregated snow have values of 0.98–1, 0.97–1, 0.9–0.97, respectively. In the mixture of rain and snow, the correlation coefficient is normally < 0.95. The tumbling hail and snow aggregate might have a value < 0.9. Ground clutter of neutral sources such as tree and grass has a relatively low correlation coefficient as 0.5–0.9. Ground clutter of man-made structures has a relatively high value as 0.9–1.0. The correlation coefficient is independent of the concentration of hydrometers and the radar calibration. It is also immune to propagation effects.

4. Specific differential phase

Differential phase (ϕ_{DP}) is the accumulated phase difference between horizontal and vertical polarizations along a propagation path. It is a variable associated with the forward scattering. The specific differential phase (K_{DP}) is defined as the range derivative of the one-way differential phase. Typically, it is computed by

$$K_{DP} = \frac{2\pi}{k}\mathrm{Re}\Big(\Big\langle n\big|F_{hh}(0)\big|\Big\rangle - \Big\langle n\big|F_{vv}(0)\big|\Big\rangle\Big) \tag{2.24}$$

where, $F_{hh,vv}(0)$ indicates the forward scattering amplitude of one particle. The unit of K_{DP} is $^\circ\mathrm{km}^{-1}$. The value of K_{DP} increases with increasing the particle oblateness. It is dependent on the hydrometer number concentration but less sensitive to the size distribution than Z_H and Z_{DR}. K_{DP} is independent of the radar calibration and the partial beam blockage. It is a good variable used for the estimation of precipitation, especially for rain [Doviak and Zrnic (1993)]. For S-band radar echo of rain, the K_{DP} is normally 0–3 $^\circ\mathrm{km}^{-1}$. Snow and hail have a lower K_{DP} because their dielectric constants are much smaller than for liquid water. The dry hail (or dry snow) has a smaller K_{DP} than the melting hail (or wet snow), typically -0.5–0.5 $^\circ\mathrm{km}^{-1}$ (or 0–0.05 $^\circ\mathrm{km}^{-1}$). Practically, the measurement error of K_{DP} is large. The estimation of K_{DP} is usually done by

an gradient calculation with averaging phase measurements of multiple range gates. For example, 9 gates (or 25 gates) average was suggested by Ryzhkov et al. (2005a) as "lightly filtered" (or heavily filtered) for KOUN radar data processing.

2.4 Scattering Theory

2.4.1 Description

When electromagnetic waves radiate on a particle, the particle may absorb a part of radiation energy and scatter the rest of radiation energy. Given that angular radiation density of electromagnetic wave is an constant, absorbed and scattered energy can be judged in terms of cross section. Typically, there are four terms used in scattering theory: scattering cross section (σ_s), absorption cross section (σ_a), total (or extinction) cross section (σ_t) and backscattering (or radar) cross section (σ_b). The total cross section is the summation of the absorption cross section and scattering cross section, i.e., $\sigma_t = \sigma_a + \sigma_s$. The backscattering cross section represents the backward scattering energy, which is usually used for mono-static radar applications. In order to assess particles with different sizes, these cross sections are generally normalized by the geometry section of the particle, $\sigma_g = \pi a^2$, and produce corresponding Q factors as

$$\begin{aligned} Q_t = \sigma_t/\sigma_g, \quad Q_s = \sigma_s/\sigma_g \\ Q_a = \sigma_a/\sigma_g, \quad Q_b = \sigma_b/\sigma_g \end{aligned} \tag{2.25}$$

For those particles with diameter much less than radiation wavelength, the scattering effect can be explained by the Rayleigh theory, according to which the scattering energy is proportional to the sixth power of the particle size and inversely proportional to the fourth power of the wavelength. But for particles whose sizes are comparable

with the wavelength, the Rayleigh scattering theory is not applicable. In such a case, the scattering issue can be solved by Mie theory, which is a complete mathematical-physical scattering theory of electromagnetic radiation by spherical particles.

Fig. 2.8 shows Q factors of particles based on Mie theory. When particle sizes are relatively small when compared to the wavelength (e.g., $D/\lambda < 0.1$), the absorption cross section is much larger than the scattering section, which means the loss of wave propagation is dominated by the particle absorption instead of the scattering. When the particle sizes are relatively large, the loss is greatly affected by scattering. The total cross section is approximately twice as large as the geometric cross section of particles. The radar backscattering cross section is almost linearly increasing with the size for small particles ($D/\lambda < 1/16$), while for large particles, it varies periodically with the decrease of amplitude. It is approaching the constant when the size is approaching the infinite.

2.4.2 Raindrop model

Many studies have shown that raindrops are not spherical but more likely oblate, especially for large raindrops [e.g., Pruppacher and Beard (1970)]. The oblateness increases with increasing raindrop size. The axis ratio of a raindrop can be defined as the ratio of the diameter of vertical axis to the diameter of horizontal axis (see Fig. 2.9). Since the raindrop is oblate, the scattering of a raindrop would be different at horizontal and vertical directions, causing the non-zero differential reflectivity. The differential reflectivity of a raindrop can be quantitatively determined by the raindrop oblateness. In addition, the raindrop oblateness is also important to determine other PRD such as the differential phase and the correlation coefficient.

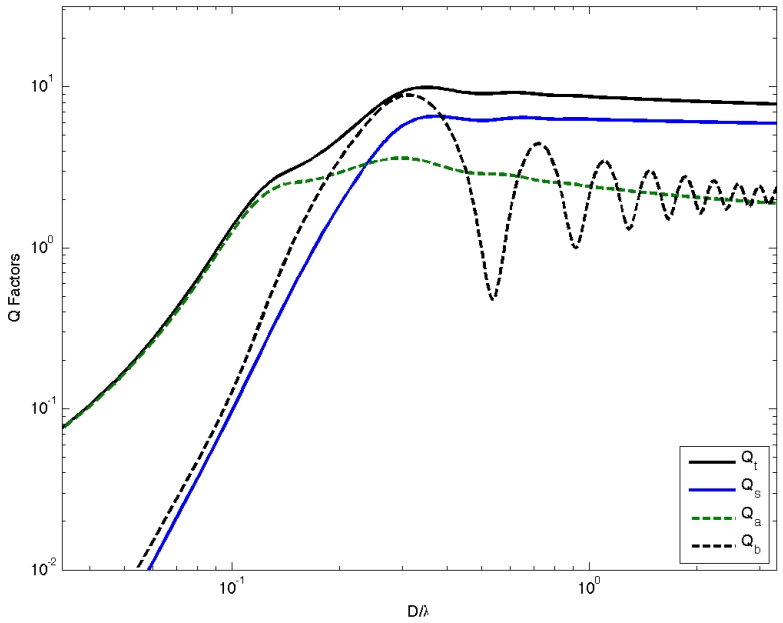

Figure 2.8: Q factors in terms of particle sizes based on Mie theory, from lecture notes of course ECE/METR 6613, Spring 2006.

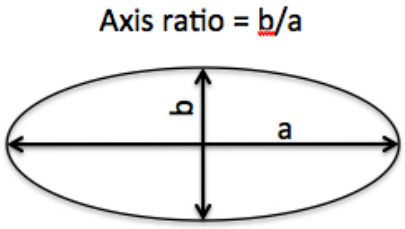

Figure 2.9: An oblate raindrop model and its axis ratio.

Efforts have been made to investigate the shape of a falling raindrop. Keenan et al. (2001) gave a review of observed raindrop axis ratio relations. Generally, there are three kinds of relations. The empirical relations introduced by Pruppacher and Beard (1970); Green (1975); Chuang and Beard (1990) focuses on the raindrop shape under an equilibrium condition. However, other studies such as Pruppacher and Pitter (1971); Beard et al. (1983); Beard and Jameson (1983); Beard and Tokay (1991) found that collision, wind shear and turbulence could lead to the oscillation of raindrops, whose shapes would be more spheric than shapes under equilibrium condition. Their shape models are referred to as the oscillation model. Keenan et al. (2001); Brandes et al. (2002) introduced empirical relations derived from previous observations or relations with a regression procedure.

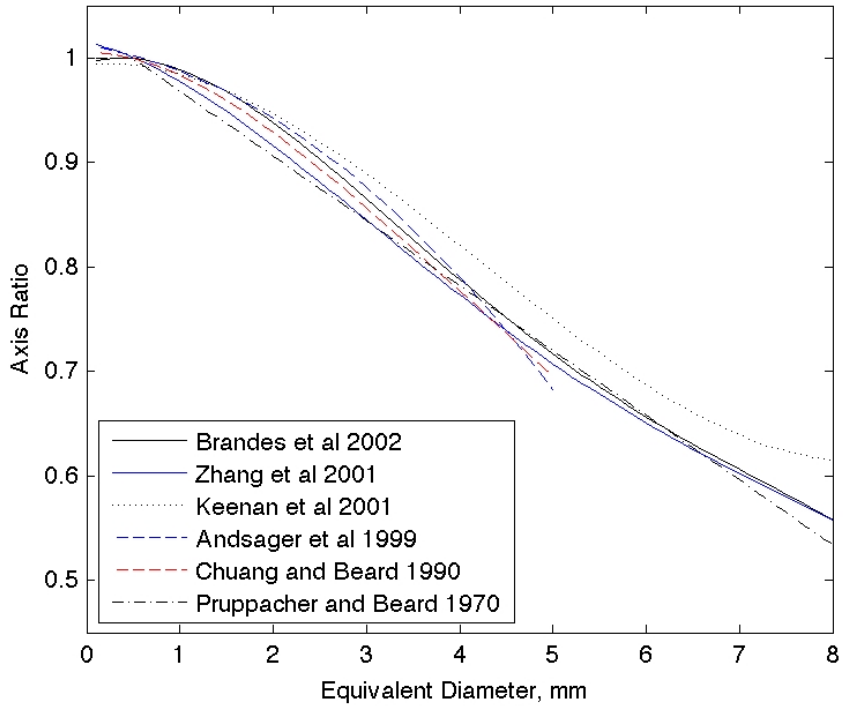

Figure 2.10: Different raindrop axis ratio models.

Even with the same DSD, different raindrop shape models might lead to a distinct value of estimated radar variables. Brandes et al. (2002) illustrated with a specific example that the simulated Z_{DR} using the equilibrium shape model is 0.2 dB larger than the corresponding value calculated using the experimental shape model. It is worth noting that currently there is no consensus on the use of the raindrop shape model in the community. In this study, estimated radar variables are based on the experimental shape model introduce by Brandes et al. (2002). The axis ratio r is given as:

$$r = 0.9951 + 0.0251D - 0.03644D^2 + 0.005303D^3 - 0.0002492D^4 \tag{2.26}$$

2.4.3 Scattering amplitude and phase

In addition to the particle shape, the particle size, canting angle, composition, frequency and temperature are factors that determine particle's scattering amplitude and phase. This study puts efforts on the study of rainfall, therefore, the scattering of raindrop is the major interest of this section.

The major effect of the temperature is on the dielectric constant ϵ, thereafter the parameter K. Table 2.2 gives the dielectric terms of water for different frequencies and temperatures, following the equations introduced by Ray (1972). The variations of dielectric constant ϵ and term $|K|^2$ are small, especially for lower frequency (e.g., S-band 3 GHz). As a result, this study has neglected the effect of temperature and assumed the temperature to be 10 °C for the calculation of raindrop scattering.

The frequency and the size are other two major factors determining the scattering of raindrops. Fig. 2.11 gives the results of raindrop scattering amplitude and phase calculated from T-Matrix method [Zhang et al. (2001)]. The green, blue and black lines represent results for S-band (10 cm), C-band (5.4 cm) and X-band (3.2 cm),

Table 2.2: Dielectric terms of water in terms of temperature and frequency

		0 °C	10 °C	20 °C
3 GHz	ϵ	79.6919+25.1976 i	79.6690+18.2257 i	77.9014+13.2354 i
	$\|K\|^2$	0.9342	0.9313	0.9283
5.5 GHz	ϵ	65.1406+37.1941 i	70.9023+29.4124 i	72.7890+22.4553 i
	$\|K\|^2$	0.9331	0.9307	0.9279
9.3 GHz	ϵ	44.7967+41.4592 i	55.4394+37.8489 i	62.3358+31.9111 i
	$\|K\|^2$	0.9305	0.9291	0.9269

respectively. Solid lines and dashed lines indicate the scattering at horizontal (H) and vertical (V) directions, respectively. As Fig. 2.11 shows, both backscattering and forward scattering amplitudes increase with increasing raindrop size, but the increase is not linear in logarithmic domain. For example, a raindrop with $D > 5$ mm might have a scattering amplitude of four orders larger than a raindrop with $D < 0.5$ mm. The higher the frequency, the larger the scattering amplitude. The difference of scattering amplitude between two frequencies also depends on the size. Due to the oblateness of raindrop, there are differences between scattering amplitudes at horizontal and vertical directions. The larger the raindrop, the more the scattering difference. The difference between two directions is little for raindrops with $D < 1.5$ mm. The scattering phase is a little more irregular than the scattering amplitude. Generally, the forward scattering phase has an increasing tendency with increasing the raindrop size. The backscattering phases between two directions have a difference about 180°.

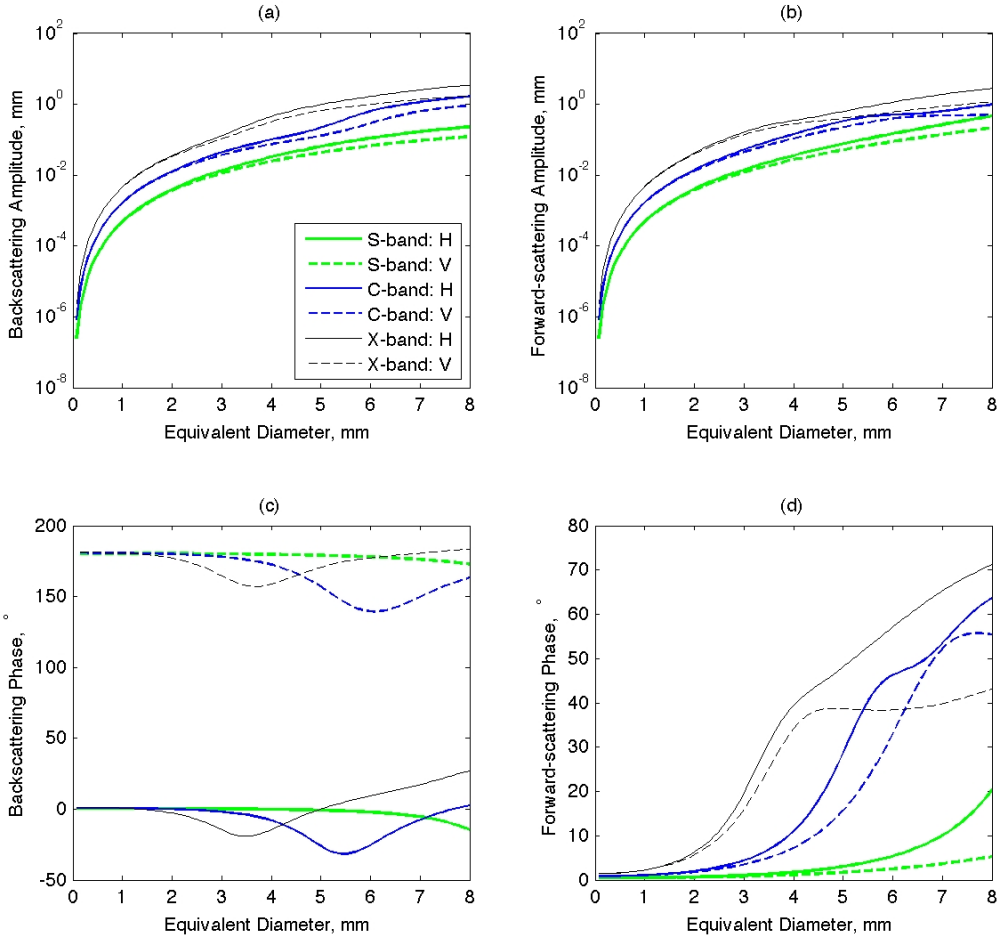

Figure 2.11: Calculated scattering parameters with respect to the equivalent raindrop size based on T-matrix method: (a) backscattering amplitude, (b) forward scattering amplitude, (c) backscattering phase, (d) forward scattering phase.

2.4.4 Simulation of radar variables

According to the scattering theory, radar variables can be calculated given a DSD. Equations used for the calculation are Eqs. 2.21–2.24. In addition, the specific attenuation ($A_{H,V}$), and specific differential attenuation (A_{DP}) are two frequently applied variables representing the precipitation attenuation. Similar to specific differential phase, specific attenuation and specific differential attenuation are variables also associated with the forward scattering. Their calculations are given as:

$$A_{H,V} = \frac{4\pi}{k}\mathrm{Im}\Big(\Big\langle n\Big|F_{hh,vv}(0)\Big|\Big\rangle\Big) \tag{2.27}$$

$$A_{DP} = \frac{4\pi}{k}\mathrm{Im}\Big(\Big\langle n\Big|F_{hh}(0)\Big|\Big\rangle - \Big\langle n\Big|F_{vv}(0)\Big|\Big\rangle\Big) \tag{2.28}$$

The following figures show some examples of simulated radar variables (S-band) based on the procedure mentioned previously. Figs. 2.12–2.14 show the dependence of five radar variables on the rainfall rate, given the assumption of M-P DSD. Results of T-Matrix method are shown by solid lines. As a reference, results of Rayleigh scattering are plotted as well. It is worth noting that the canting angle of raindrops also play a role in the calculation of scattering. Generally, a raindrop is assumed to have a random canting angle during its falling process but the mean canting angle is zero. The probability distribution of canting angle is assumed to follow the Gaussian function with a standard deviations σ_ϕ. Figure results also show the comparison between $\sigma_\phi = 0^\circ$ and $\sigma_\phi = 20^\circ$.

As Fig 2.12 shows, both Z_H and Z_{DR} increase with increasing rainfall rate. They increase quickly for rainfall rate less than 20 mm h^{-1} while slowly for larger rainfall rate. The bigger the uncertainty of canting angle, the smaller the Z_H and Z_{DR}. The difference attributed to σ_ϕ is very small for Z_H but moderate for Z_{DR}. Rain with rate of 100 mm h^{-1} might have a 0.5 dB difference between $\sigma_\phi = 0^\circ$ and $\sigma_\phi = 20^\circ$. Z_H estimated by Rayleigh theory has a larger value than Z_H by T-matrix method.

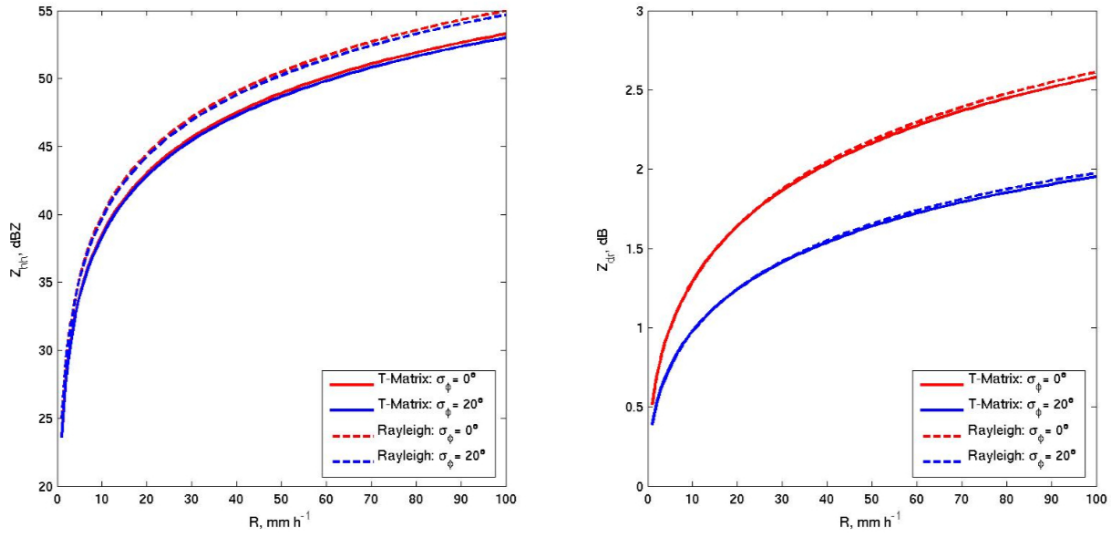

Figure 2.12: Simulated S-band (a) Z_H and (b) Z_{DR} based on scattering theory versus rainfall rate.

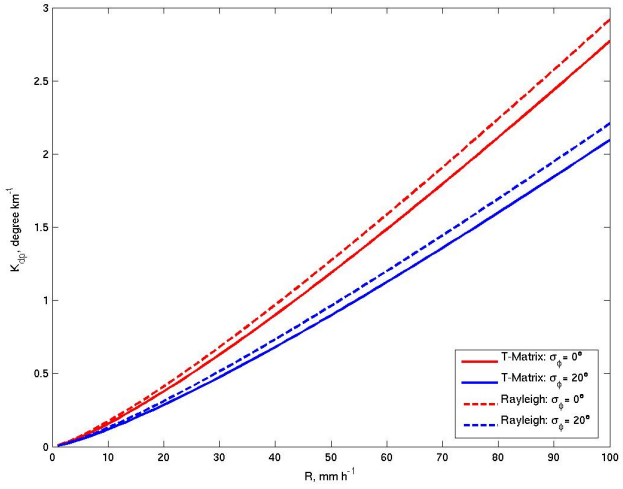

Figure 2.13: Simulated S-band K_{DP} based on scattering theory versus rainfall rate.

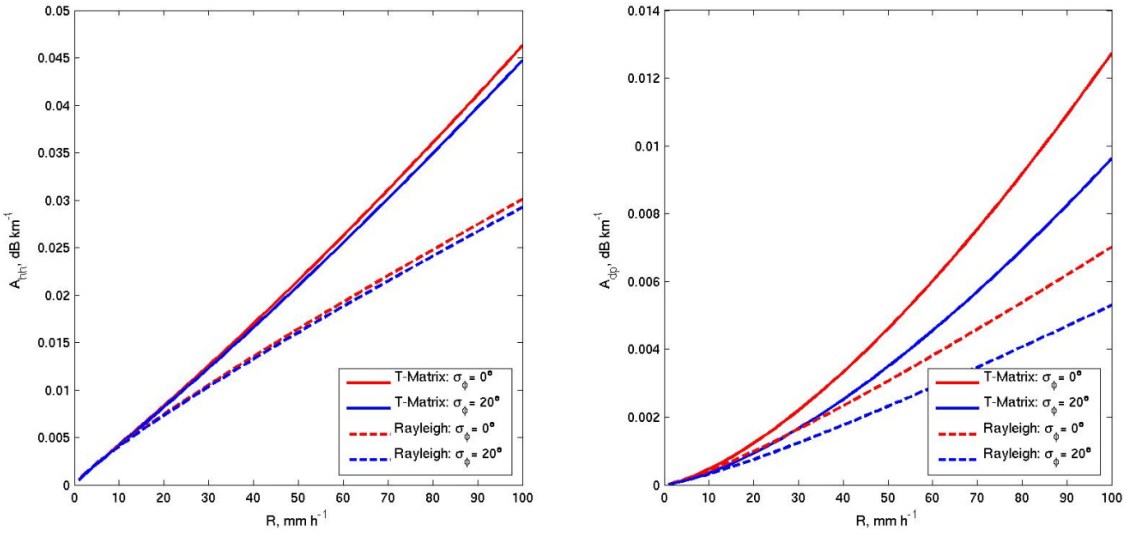

Figure 2.14: Simulated S-band (a) A_{hh} and (b) A_{DP} based on scattering theory versus rainfall rate.

With increasing rainfall rate, the difference becomes larger. The Z_{DR} difference is very small for Rayleigh and T-matrix calculation.

Figs 2.13 and 2.14 show the similar tendency to Fig. 2.12. K_{DP}, A_H and A_{DP} all increase with increasing rainfall rate. A_H increases almost linearly. The bigger the σ_ϕ, the smaller the K_{DP}, A_H and A_{DP}. The difference of A_H attributed to σ_ϕ is very small while the difference of A_{DP} or K_{DP} is notable for moderate and heavy rain. When rainfall rate increases, differences attributed to σ_ϕ become larger for all three variables. There are some explanations for these results. When canting angle increases, the horizontal scattering amplitude would decrease and the vertical amplitude would increase. When the drop size or drop number concentration increases, the attenuation and phase shift would also increase. Horizontal components increase more quickly than vertical components. When rainfall rate increases, the number concentration of raindrops increases and more large drops may exist. As a result, the difference between two directions would become larger, especially for Z_{DR}, K_{DP} and A_{DP}.

Chapter 3

Error Analysis and DSD Parameter Estimation

3.1 Introduction

Disdrometers are usually used to measure natural DSDs, study precipitation microphysics, and verify radar-rain estimation. However, disdrometer measurements themselves are not without error, which are caused by i) undersampling [e.g., Gertzman and Atlas (1977); Wong and Chidambaram (1985)]; ii) physical variations [e.g., Ulbrich (1983); Jameson and Kostinski (1998)]; iii) instrument limitations (resolution and sensitivity); and iv) environmental factors such as wind effect [e.g., Nespor et al. (2000)]. In addition, there might be discrepancies between a natural DSD and its model representation (i.e., model error). Measurement and model errors both cause errors in the estimation of gamma DSD parameters and integral physical parameters [e.g., Zhang et al. (2003); Smith et al. (2005)].

The first issue examined in this chapter is the quantification of disdrometer measurement error related to small sampling volume and limited sampling time. Disdrometer observations contain not only physical variations but also measurement errors. Gertzman and Atlas (1977) and Wong and Chidambaram (1985) presented a detailed analysis of sampling error based on the assumption of independent Poisson distributions of raindrop count. Rain events, however, may not be independent stationary

random processes. Physical variations and sampling errors coexist [e.g., Jameson and Kostinski (1998); Schuur et al. (2001)]. It is difficult to separate sampling errors from physical variations with a single instrument. Side-by-side comparisons, on the other hand, provide information that allows sampling errors to be quantified. Tokay et al. (2001) compared video disdrometer (i.e., 2DVD) and impact disdrometer (i.e., Joss-Waldvogel disdrometer, JWD) measurements. However, their study focused mainly on the comparison of DSD parameters and rain variables and did not quantify errors. To my knowledge, error quantification for 2DVD observations through side-by-side comparison has not yet been reported. Knowing observational errors and their error correlations for different DSD moments, the error propagation can be estimated for any rain variable estimator based on rain moments [e.g., Zhang et al. (2003)]. Moreover, error quantification helps to introduce advanced processing techniques to reduce error effects on DSD modeling or retrieval.

The second issue is to examine error effects on the DSD parameter estimation. Different DSD moments have been used to estimate DSD parameters in previous studies. Waldvogel (1974) first suggested using both the 3^{rd} (M3) and the 6^{th} moments (M6) (the estimator using the 3^{rd} and 6^{th} moments is referred to as M36, the same method of notation will be used for other estimators throughout this dissertation) to determine two parameters of the exponential DSD model. This is a reasonable choice because M3 and M6 are exactly (or approximately) proportional to water content (or attenuation for Rayleigh scattering) and radar reflectivity (for Rayleigh scattering). In the succeeding decades, however, different combinations of DSD moments were also used to estimate DSD parameters. For example, the 3^{rd} and 4^{th} moments were applied by Smith et al. (1976) and Testud et al. (2000). The 0^{th}, 1^{st}, and 2^{nd} moments (i.e., an M012 estimator) were used by Smith (1993). The 3^{rd}, 4^{th} and 6^{th} moments were applied by Kozu and Nakamura (1991); Tokay and Short (1996); Tokay et al.

(2001). The 2^{nd}, 4^{th}, and 6^{th} moments (i.e., an M246 estimator) were used by Ulbrich and Atlas (1998); Zhang et al. (2001, 2003); Brandes et al. (2004a,b); Vivekanandan et al. (2004). Which moments should be used to better estimate DSD parameters? According to Gertzman and Atlas (1977) and Wong and Chidambaram (1985), higher DSD moments have greater errors. Their conclusion was based on the assumption of Poisson probability theory for raindrop numbers. However, as indicated by Smith (2003), low-order moments may have even larger errors than middle-order moments because radar measurements do not accurately represent the dependence of radar-measured variables on small raindrops. Smith et al. (2005) proposed using the M234 estimator and showed that it gave the least error for the estimate of DSD parameters. Kliche et al. (2008) recently argued that the L-Moment (LM) and maximum likelihood (ML) estimators, which obtain the estimate from DSD spectrum instead of integral moments, have better performance than moment estimators used previously (e.g., M234) for retrieving DSD parameters of a gamma distribution. Consequently, it is necessary to evaluate performances of all these estimators.

The third issue is the correlation between DSD parameters. Several researchers [e.g., Ulbrich (1983); Chandrasekar and Bringi (1987); Haddad et al. (1997)] have shown that the retrieved three DSD parameters (N_0, μ, and Λ) of gamma model are not mutually independent. Ulbrich (1983) introduced a $N_0 - \mu$ relation whereby the three DSD parameters could be retrieved using measurements of radar reflectivity (Z_H) and attenuation. However, the $N_0 - \mu$ relation, depending on the fitting procedure, is unstable and fluctuates by several orders of magnitude. Therefore, its utility is limited. Chandrasekar and Bringi (1987) attributed the $N_0 - \mu$ relation to the statistical error. Haddad et al. (2006) further showed that even in the absence of observational noise, the dual-frequency retrieval using a $N_0 - \mu$ relation could be ambiguous. Through disdrometer observations, Zhang et al. (2001) and Brandes et al.

(2004a) found that μ is highly related to Λ. Because the resulting $\mu - \Lambda$ relationship helps to facilitate the DSD retrieval from dual-polarization or dual-frequency radar measurements, it is necessary to know whether the $\mu - \Lambda$ relation represents physical property of rain or just a spurious relation.

In Section 3.2, disdrometer sampling errors are quantified by analyzing measurements from two instruments placed side-by-side. Error effects of measurement and DSD model are investigated for seven different estimators in Section 3.3. The constraint-gamma (C-G) DSD is verified in Section 3.4 by analyzing error effects and quantitative comparison. Section 3.5 introduces a sorting and averaging procedure based on two parameters (SATP) to mitigate the error effect. The C-G DSD model is then refined with this procedure and used for the rest part of this study.

3.2 Quantification of 2DVD Observation Error

Disdrometer observations contain sampling errors that arise from a limited sampling volume ($\sim$0.01 m^2 sampling area and $\sim$3 m^3/minute sampling volume). In practice, it is difficult to separate sampling errors from physical variations based on single disdrometer measurements. Measurements by two similar disdrometers deployed side-by-side, however, can be treated as two realizations, x_1 and x_2, of the same process, which have the same expected value. That is,

$$\begin{cases} x_1 = \langle x \rangle + \varepsilon_1 \\ x_2 = \langle x \rangle + \varepsilon_2 \end{cases} \tag{3.1}$$

where ε_1 and ε_2 denote sampling errors of two disdrometers. Assuming two disdrometers measure the same DSD, differences between two measurements would be due to statistical fluctuations. The physical variation can be cancelled by subtracting

one measurement from the other. Assuming sampling errors of two disdrometer are independent, the sampling error can be quantified by

$$\left\langle |x_1 - x_2|^2 \right\rangle = \left\langle |\varepsilon_1 - \varepsilon_2|^2 \right\rangle = 2\sigma_x^2 \tag{3.2}$$

where σ_x is the standard deviation of sampling error. The error can also be represented by the fractional standard deviation (FSD) as

$$\mathrm{FSD}_x = \frac{\sigma_x}{\langle x \rangle} \tag{3.3}$$

The expected value of x is not known in practice. However, if the rain process is assumed to be a stationary random process, the expected value of x could be estimated using the time-average of all available data based on the ergodic theorem. In this study, the expected value is estimated by taking the time-averaged value of 14200-minute samplings. It is worthwhile to note that this kind of estimation introduces some uncertainties because in practice the rain process is not an ergodic process. Nevertheless, the uncertainty attributed to the time-average is not the emphasis of this study.

Differences between measurements of two 2DVDs arise from spatial inhomogeneity in rain and slight differences in the spatial and temporal resolutions between the two units. Measurement bias is reduced by the calibration. Measured number concentrations within each bin were averaged for both 2DVDs. For each bin, the difference between two mean number concentrations was regarded as a measurement bias. The measurement bias was then subtracted from given 2DVD measurements. Although the measurement bias can not be perfectly tuned for bins with a size less than 0.6 mm, bias effects are insignificant for integral parameters. FSDs of physical parameters, such as drop count, mass, and reflectivity distributions, had similar error characteristics. Taking drop count measurements as an example, FSDs estimated using Eqs. 3.2 and 3.3 are denoted as side-by-side and are shown as a function of bin size

(solid line) in Fig. 3.1. If only measurements of a single 2DVD are used, the standard deviation σ_x in Eq. 3.3 is calculated by $\sigma_x^2 = \langle |x - \langle x \rangle|^2 \rangle$. The dashed line represents the result of the NCAR 2DVD measurements and the dash-dotted line represents the result of the NSSL 2DVD measurements. The dashed and dash-dotted lines agree well for the medium-sized drops. Based on the assumption that the observed raindrops obey the Poisson process, the theoretical FSDs, shown by the dotted-solid line, are derived by $\langle N_i \rangle^{-0.5}$, where N_i is the total number of drops within the i^{th} bin.

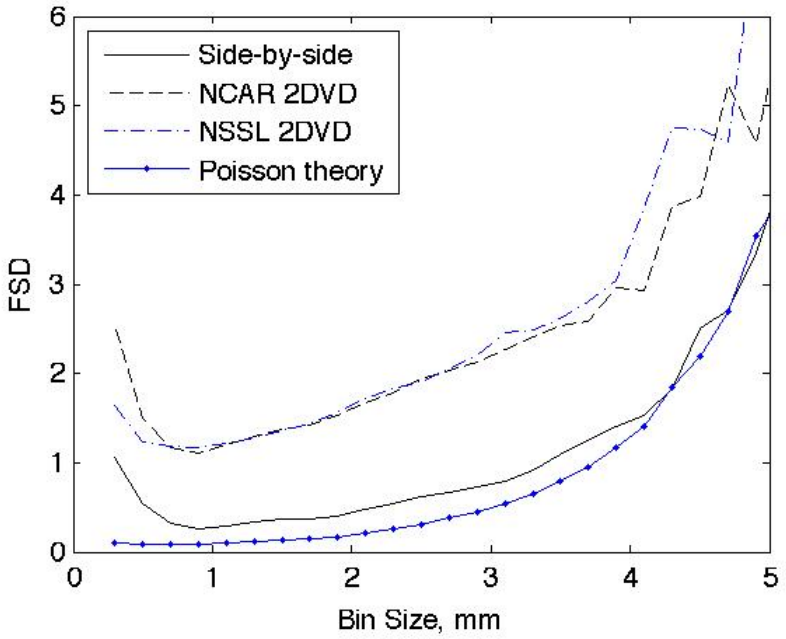

Figure 3.1: FSD of observed drop counts over bin spectrum. The solid line represents the calculation based on side-by-side comparison. The dashed line represents the result of NCAR 2DVD measurements and the dash-dotted line represents the result of NSSL 2DVD measurements. The solid-dotted line represents the calculation based on Poisson assumption.

As shown in Fig. 3.1, FSDs estimated from single 2DVD measurements (dashed and dash-dotted lines) give an overestimation for sampling errors. Considering that side-by-side 2DVD measurements eliminate the physical variation, sampling errors

estimated in this way are more accurate than from single 2DVD estimates. The difference between the solid and dashed/dash-dotted lines demonstrates the DSD's physical variability, which is large compared to the Poisson error (within bin sizes less than 3.5 mm). The solid line also shows that the sampling error increases with decreasing drop size for drops smaller than 0.6 mm. This is mainly attributed to the 2DVD's inability to accurately measure small drops. The dotted-solid line represents the Poisson error. The Poisson error increases considerably when drop sizes are larger than 3 mm. This is due to the undersampling of large raindrops. It is also noticed that the sampling error indicated by the solid line is close to the Poisson error for raindrop sizes greater than 3 mm. It implies that the sampling error is associated primarily with the undersampling for these sizes.

FSDs of DSD moments were estimated by applying Eqs. 3.2 and 3.3, in which the variable x represents the DSD moment (not bin drop counts). The results for moments from the 0^{th} order to the 6^{th} order (M_0, M_1, , M_6) are given in the columns of Table 3.1. The first row contains the result estimated from the side-by-side comparison. The second row, denoted as "Theoretical", contains the result estimated from the same dataset but based on the Poisson statistical model [Appendix of Schuur et al. (2001)]. Because the theoretical result assumes sampled raindrop counts within one-minute obey the Poisson distribution, which is independent random process, it gives a lower limit to the actual FSDs. The theoretical result indicates that moment errors are generally larger for higher moments. This can be explained by the moment estimation. That is, because large drops carry more weight in the calculation of higher moments, their sampling errors are greater contribututors to the total error. The estimate from side-by-side comparison generally agrees with this tendency except the error is somewhat larger. The difference between the "Side-by-side" and "Theoretical" results might be explained by (i) dependent measurement errors, i.e.,

samples from two adjacent times or bins have correlated errors, (ii) instrumental bias, and (iii) non-Poisson distributed raindrops. Because error estimates in Table 3.1 were obtained directly from disdrometer measurements without any assumption regarding DSD shape, they are more realistic than results of simulations.

Table 3.1: FSD of different DSD moments

Moment	M_0	M_1	M_2	M_3	M_4	M_5	M_6
Side-by-side	0.1029	0.0965	0.0906	0.0901	0.1025	0.1311	0.1746
Theoretical	0.0379	0.0350	0.0408	0.0550	0.0767	0.1045	0.1372

Correlations exist not only between DSD moments [e.g., Jameson and Kostinski (1998)], but also between sampling errors of DSD moments [e.g., Zhang et al. (2003)]. Since the sampling errors have been quantified by a side-by-side comparison in this paper, the correlation between sampling errors can be quantified as well. The correlation coefficient is formulated by

$$\rho = \frac{\langle \varepsilon_m \varepsilon_n \rangle}{\sqrt{\langle \varepsilon_m^2 \rangle \langle \varepsilon_n^2 \rangle}} \tag{3.4}$$

where m, n represent two moments (m, n = 0,, 6) and ε denotes the error of observed moments. In practice, the $\langle \varepsilon_m \varepsilon_n \rangle$ is estimated from two 2DVDs measurements by $\frac{1}{N}\sum_{k=1}^{N} 0.5(x_{1,k}^{(m)} - x_{2,k}^{(m)})(x_{1,k}^{(n)} - x_{2,k}^{(n)})$, where x denotes the measured moment, subscript 1, 2 represent two 2DVDs and N is the number of side-by-side samples. Correlation coefficients between sampling errors for DSD moments ranging from the 0^{th} to the 6^{th} order are given in Table 3.2. It is obvious that sampling errors are less correlated if two moments are widely spaced. The correlation will be less than 10% if the order difference is larger than 6. Given the same order difference (e.g., correlation of M_0 and M_3 versus correlation of M_3 and M_6), the correlation tends to be a little larger for higher moments, which is caused by sampling errors. As for

the 2^{nd}, 4^{th}, and 6^{th} moments, which were chosen to retrieve DSD parameters, the correlation coefficients for pairs M_2-M_4, M_4-M_6, and M_2-M_6 are 0.74, 0.82 and 0.40, respectively. It is important to know these FSDs and correlation coefficients because they determine the standard errors of DSD parameter estimation [Zhang et al. (2003)].

Table 3.2: Correlation coefficient of sampling errors for different moments

	M_0	M_1	M_2	M_3	M_4	M_5	M_6
M_0	1	0.8927	0.6805	0.4996	0.3385	0.1927	0.0531
M_1	0.8927	1	0.9142	0.7202	0.5116	0.3449	0.2058
M_2	0.6805	0.9142	1	0.9160	0.7371	0.5594	0.3988
M_3	0.4996	0.7202	0.9160	1	0.9346	0.7927	0.6215
M_4	0.3385	0.5116	0.7371	0.9346	1	0.9493	0.8244
M_5	0.1927	0.3449	0.5594	0.7927	0.9493	1	0.9570
M_6	0.0531	0.2058	0.3988	0.6215	0.8244	0.9570	1

3.3 DSD Parameter Estimation

3.3.1 Estimation method

There are seven estimators, i.e., five moment estimator as well as L-Moment (LM) and maximum likelihood (ML) estimators, evaluated in this study. They are used to estimate three gamma DSD parameters μ, Λ and N_0. The M012 estimator is based on the 0^{th}, 1^{st}, and 2^{nd} order moments. Similarly, the other four estimators are referred to as "M234", "M246", "M346", and "M456". The moment estimators are

i) M012:

$$\hat{\eta} = \frac{\hat{M}_1^2}{\hat{M}_0\hat{M}_2}, \quad \hat{\mu} = \frac{1}{1-\hat{\eta}} - 2,$$
$$\hat{\Lambda} = \frac{\hat{M}_0}{\hat{M}_1}(\hat{\mu}+1), \quad \hat{N}_0 = \frac{\hat{M}_0\hat{\Lambda}^{(\hat{\mu}+1)}}{\Gamma(\hat{\mu}+1)}. \tag{3.5}$$

ii) M234:

$$\hat{\eta} = \frac{\hat{M}_3^2}{\hat{M}_2\hat{M}_4}, \quad \hat{\mu} = \frac{1}{1-\hat{\eta}} - 4,$$
$$\hat{\Lambda} = \frac{\hat{M}_2}{\hat{M}_3}(\hat{\mu}+3), \quad \hat{N}_0 = \frac{\hat{M}_2\hat{\Lambda}^{(\hat{\mu}+3)}}{\Gamma(\hat{\mu}+3)}. \tag{3.6}$$

iii) M246:

$$\hat{\eta} = \frac{\hat{M}_4^2}{\hat{M}_2\hat{M}_6}, \quad \hat{\mu} = \frac{(7-11\hat{\eta}) - (\hat{\eta}^2+14\hat{\eta}+1)^{0.5}}{2(\hat{\eta}-1)},$$
$$\hat{\Lambda} = \left[\frac{\hat{M}_2}{\hat{M}_4}(\hat{\mu}+3)(\hat{\mu}+4)\right]^{0.5}, \quad \hat{N}_0 = \frac{\hat{M}_2\hat{\Lambda}^{(\hat{\mu}+3)}}{\Gamma(\hat{\mu}+3)}. \tag{3.7}$$

iv) M346:

$$\hat{\eta} = \frac{\hat{M}_4^3}{\hat{M}_3^2\hat{M}_6}, \quad \hat{\mu} = \frac{(8-11\hat{\eta}) - (\hat{\eta}^2+8\hat{\eta})^{0.5}}{2(\hat{\eta}-1)},$$
$$\hat{\Lambda} = \frac{\hat{M}_3}{\hat{M}_4}(\hat{\mu}+4), \quad \hat{N}_0 = \frac{\hat{M}_3\hat{\Lambda}^{(\hat{\mu}+4)}}{\Gamma(\hat{\mu}+4)}. \tag{3.8}$$

v) M456:

$$\hat{\eta} = \frac{\hat{M}_5^2}{\hat{M}_4\hat{M}_6}, \quad \hat{\mu} = \frac{1}{1-\hat{\eta}} - 6,$$
$$\hat{\Lambda} = \frac{\hat{M}_4}{\hat{M}_5}(\hat{\mu}+5), \quad \hat{N}_0 = \frac{\hat{M}_4\hat{\Lambda}^{(\hat{\mu}+5)}}{\Gamma(\hat{\mu}+5)}. \tag{3.9}$$

The ML and LM estimators are similar to the ones described by Kliche et al. (2008). The only difference is that Kliche et al. have assumed the drop number in different size categories follows the gamma distribution whereas assumption here is that the

number concentration (i.e., the DSD) follows the gamma distribution. It is worth noting that the number concentration is the drop number in a fixed size sampling volume. However, for Kliche's approach different size categories have different sampling volumes. Consequently, two approaches are not exactly equivalent. Suppose the number concentration of each size category D_i $(i = 1 \ldots 41)$ has an integer number N_i, the summation of N_i is N. The LM and ML estimator are described as follows.

vi) LM:

$$\hat{l}_1 = b_0, \quad \hat{l}_2 = 2b_1 - b_0,$$
$$b_0 = \frac{1}{N}\sum_{k=1}^{N} D_k, \quad b_1 = \frac{1}{N(N-1)}\sum_{k=1}^{N-1} kD_{(k+1)}, \tag{3.10}$$
$$\frac{\hat{l}_2}{\hat{l}_1} = \frac{\Gamma(\hat{\mu}+1.5)}{\sqrt{\pi}\Gamma(\hat{\mu}+2)}$$

where, the l_1 and l_2 are the first two L-moments; $D_{(k)}$ is the k^{th} size category with N_i from small to large in sequence. The estimate of μ is calculated by nonlinear iteration. After obtaining the estimate of μ, Λ can be calculated by

$$\hat{\Lambda} = \frac{\hat{\mu}+1}{\hat{l}_1}. \tag{3.11}$$

vii) ML: The estimate of μ is calculated by iteration from following formula

$$\ln(\hat{\mu}+1) - \Psi(\hat{\mu}+1) = \ln\left[\frac{\frac{1}{N}\sum_{k=1}^{N} D_k}{\left(\prod_{i=1}^{N} D_i\right)^{1/N}}\right], \tag{3.12}$$

where, Ψ is the "psi" function defined by $\Psi(x) = \frac{\Gamma'(x)}{\Gamma(x)}$. The estimate of Λ has the similar form as Eq. 3.11 and is

$$\hat{\Lambda} = \frac{\hat{\mu}+1}{\frac{1}{N}\sum_{k=1}^{N} D_k} \tag{3.13}$$

As described above, ML and LM estimators give estimates of μ and Λ. The third DSD parameter, N_0, is estimated from the 0^{th} moment. The estimate of N_0 is given by

$$\hat{N_0} = \frac{\hat{M}_0 \hat{\Lambda}^{\hat{\mu}+1}}{\Gamma(\hat{\mu}+1)} \tag{3.14}$$

3.3.2 Error analysis in estimating DSD parameters

3.3.2.1 Simulation of disdrometer observations

In order to evaluate estimators mentioned above, simulated disdrometer observations are applied for the reason that the truth can be controlled in the simulation. The first simulation—simulation A—assumes that the true DSD follows the gamma distribution. The simulations apply the similar procedure as described in Chandrasekar and Bringi (1987), Smith et al. (1993), and Moisseev and Chandrasekar (2007), with some modifications (described later in this subsection). Simulation A uses the normalized gamma distribution (Eq. 2.4). Three parameters N_w, μ and D_0 were randomly generated within specific ranges as described by Ulbrich (1983), and Bringi and Chandrasekar (2001), which were believed to represent most natural rain DSDs. It was assumed that μ has a uniform in the range from -1 to 4, D_0 was uniform from 0.5 to 2.5 mm, and $\log_{10} N_w$ was uniform from 3 to 5. The simulated gamma distributions with these parameters were regarded as the "truth". Sampling errors were then added to the "truth" and assumed to be observation errors of disdrometer. Here, errors were not just the Poisson errors applied by Chandrasekar and Bringi (1987) and Moisseev and Chandrasekar (2007). A modification has been made according to the result of error quantification in Section 3.2. The detailed steps are explained as follows. Firstly, the ratio of estimated error to Poisson error, r_i $(i = 1, \ldots, 41)$, is estimated by the ratio of the solid line to the dotted line for each size category in Fig.

3.1. It is found that the ratio is almost unit for sizes greater than 4 mm. Considering the side-by-side comparison may overestimate the error because of instrument factor (e.g., different sensitivity) and environment factor (rain spatial variability over two disdrometers), the value of r_i is reduced by using $r_i' = 1 + 0.6(r_i - 1)$ to better represent the measurement error for small drops. For each size category, the simulated "true" bin number concentration N_i ($i = 1, \ldots, 41$) is assumed as the mean. The error with standard deviation $r_i' N_i^{0.5}$ is then added to N_i to simulate the observed bin number concentration. As a result, the simulated observation would have a larger error than Poisson error. In particularly, the sampling error for small raindrops is much higher (about 1–5 times for $D < 1.0$ mm) than predicted by the Poisson theory.

The natural DSD is not necessary the gamma distribution though most previous studies have evaluated DSD estimators based on this assumption. It is known that the model error exists for representing observed DSDs with a gamma function. In order to investigate the effect that the model error may have on the moment estimation, the second simulation—simulation B—assumes the true DSD follows a perturbed gamma distribution. Simulation B creates the model error by adding random perturbation to the gamma distribution. The perturbation is simulated using correlated random errors. The procedure is described as follows. First, a sequence of correlated random values $\mathbf{x} = [x_1, x_2, \ldots, x_4]$ is simulated (the procedure is the same as that for generating time-series radar echoes of weather signals; Zrnic (1975)). The length of the sequence is 41, which corresponds to 41 size categories. Too small correlation length (e.g., 1 or 2) tends to introduce independently statistical errors without changing the intrinsic gamma shape. Too long correlation length (e.g., 20 or more) would likely to yield a different DSD shape. In this study, the correlation length is appropriately chosen as 5. The random variable x has a Gaussian distribution with a zero mean and an arbitrary standard deviation of 0.05. A perturbed gamma distribution is then

generated by multiplying the i^{th} bin of a simulated gamma distribution by 10^{x_i}. The perturbed distribution consists of some perturbation and a gamma distribution, and it is assumed to be the truth of a "non-gamma" distribution. It is worthwhile to note that the simulated non-gamma distributions are not far from gamma distribution because the uncertainty of perturbation is only 0.5 dB ($\sim 12\%$), which is large enough to simulate most of the observed DSDs. Similar to the procedure of simulation A, sampling errors are then added to the truth for the simulated observation. Although the simulation of non-gamma DSD seems somewhat artificial, the simulation aids in the understanding of the characteristics of moment errors and moment estimators when a DSD does not follow a gamma distribution.

The unrealistic simulation may deteriorate the statistics. Some physical constraints have been applied for data quality control. The simulated spectrum is assumed good and valid for statistical analysis if its integral parameters fall into ordinary ranges of most observed spectra. According to disdrometer observations, very narrow spectra ($D_{max} < 0.8$ mm) account for 0.9% of total dataset, data of Z_H larger than 55 dBZ are less than 0.5%, and data of R less than 0.1 mm h^{-1} and larger than 100 mm h^{-1} account for only 0.4% and 0.15%, respectively. In addition, spectra with few drops ($Tct < 10$) do not have meaningful statistical information. All these constraints are applied to eliminate the "bad" spectra for statistical analysis.

3.3.2.2 Uncertainties of estimated DSD parameters for a gamma DSD

According to Eqs. 3.5–3.9 of moment estimators, the value of estimated μ depends on the ratio η of moments. The value of η is always less than 1 according to the mathematical calculation. Fig. 3.2a shows the dependence of μ on η for different moment estimators. The parameter r is defined as the ratio of estimated η to the true value, i.e., $r = \hat{\eta}/\eta$. Fig. 3.2b shows the histogram of r for M234 and M246 based on

simulation A. Both distributions of r are skewed towards the right of value 1 because the error of moments increases the uncertainty of r. The distribution for M234 is narrower than the distribution for M246, implying that M234 has fewer errors. It is also found that the larger the moment error, the more skewed the shape is towards values greater than 1 (not shown). It is noted that all derivatives of μ with respect to η increase with η. If estimated η deviates from the true value, estimated μ will be biased; moreover, the value of μ has more overestimation than underestimation. Fig. 5 of Zhang et al. (2003) illustrates the tendency for overestimation prevails over underestimation by showing that the $\mu - \Lambda$ scatter points have a larger upward extension. Figs. 6 and 7 of Smith et al. (2005) also demonstrate these kinds of biases in estimating DSD parameters.

Fig. 3.3 shows the histograms of $\Delta\mu$ (the difference between the M246 estimation and the truth) within several rain-rate ranges. These histograms indicate that μ is overestimated rather than underestimated; the smaller the rain-rate, the larger the overestimation. Fig. 3.4 shows the normalized frequency of $\Delta\mu$, for different ranges of μ. The thin dashed line denotes the result of M234. In each subplot, the thin dashed line is almost symmetrical and has a peak at the zero point, implying that M234 has the smallest estimate bias of μ among all the estimators. M456, represented by the dotted line, tends to have the largest bias of μ. Kliche et al. (2008) demonstrated by simulation that ML and LM estimators have much less bias in estimating μ than moment estimators (e.g., M234). As shown in Fig. 3.4, however, we could not obtain the same result. Both ML and LM estimators depend heavily on the accuracy of measuring the number of small drops because the number of large drops is normally small. This implies that ML and LM estimators are sensitive to error at the low end of DSD. This tendency is supported by the results in Kliche et al. (2008), which show that truncating the low end of DSD deteriorates the performance of ML and

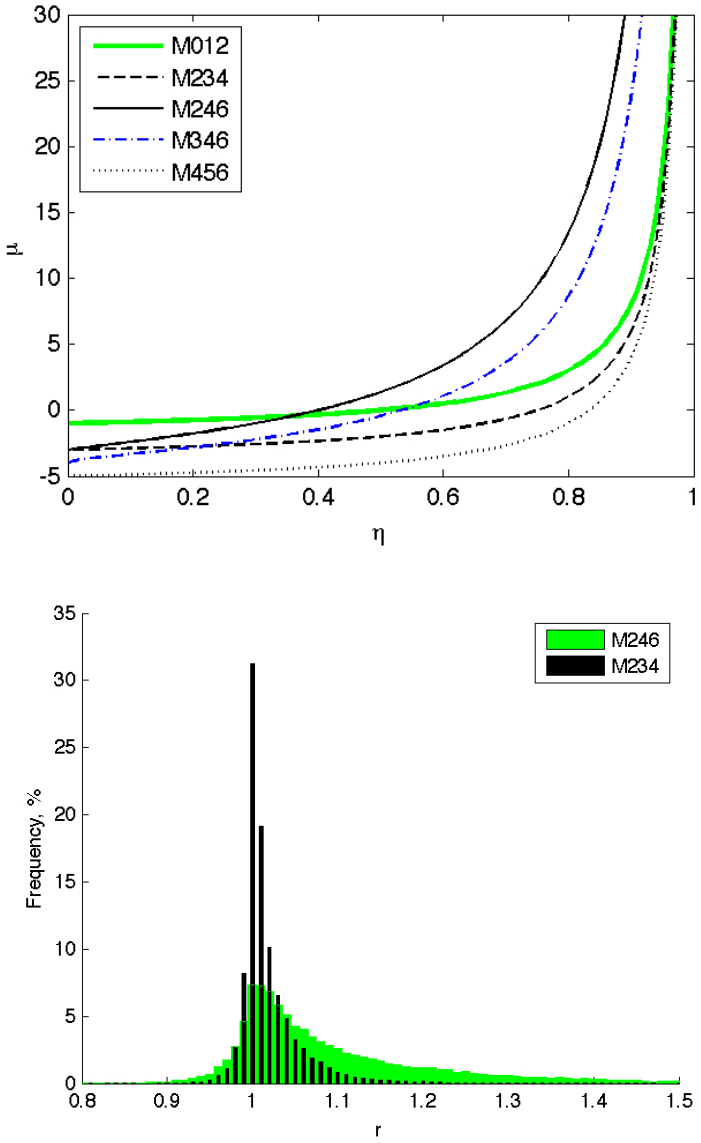

Figure 3.2: (a) The dependence of μ on η for five moment estimators. (b) The histograms of ratio r distributions for M246 and M234 estimators. The dataset is the simulation A, which simulates the gamma DSDs.

LM estimators much more than for high moment estimators. For simulation A, we have enlarged the Poisson error about 1-5 times at the low end ($D \leq 1.0$ mm) of DSD by introducing the modified error. As a result, it is not surprising that the ML and LM estimators have larger biases than M234 in the range $-1 < \mu \leq 0$. In fact, as shown in Fig. 3.4, ML and LM estimators have similar performance to

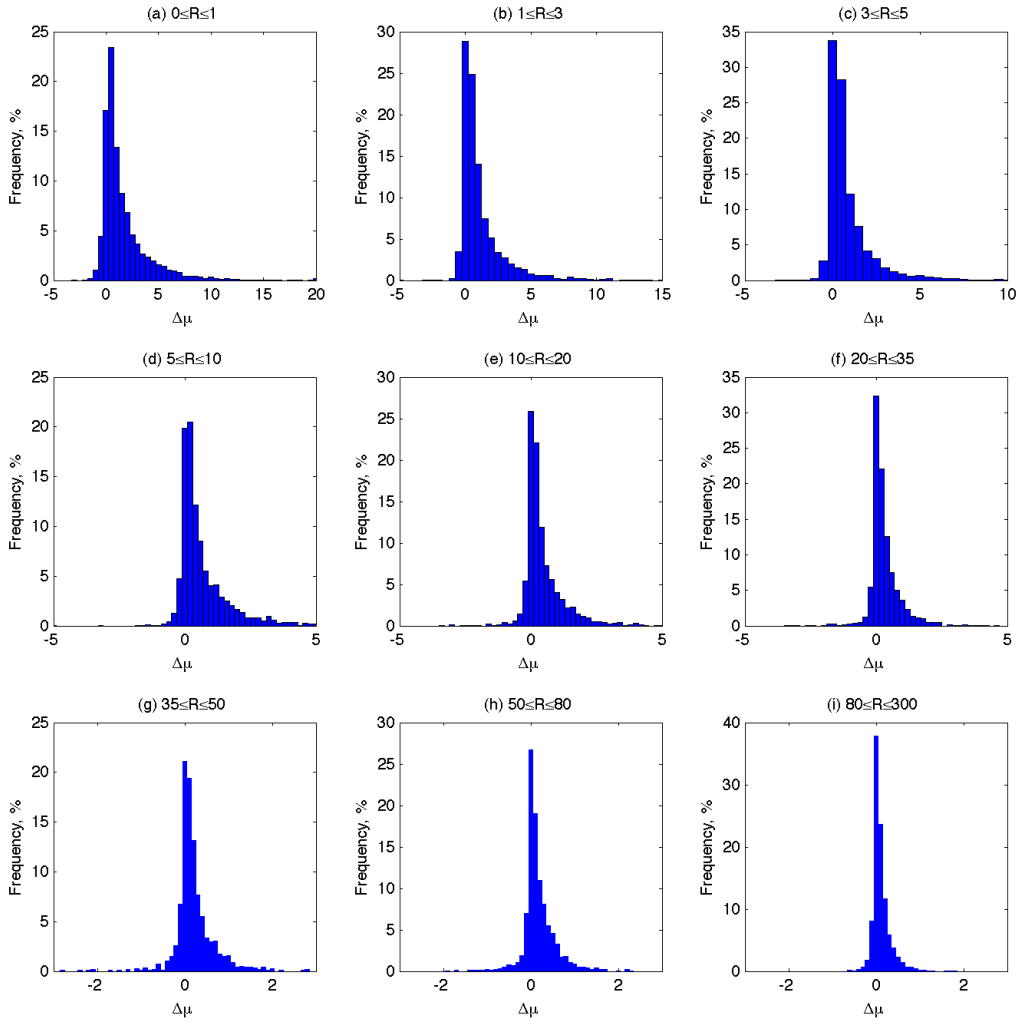

Figure 3.3: The histograms of $\Delta\mu$ distributions for different rain-rate ranges. The dataset is the simulation A and the estimator is M246.

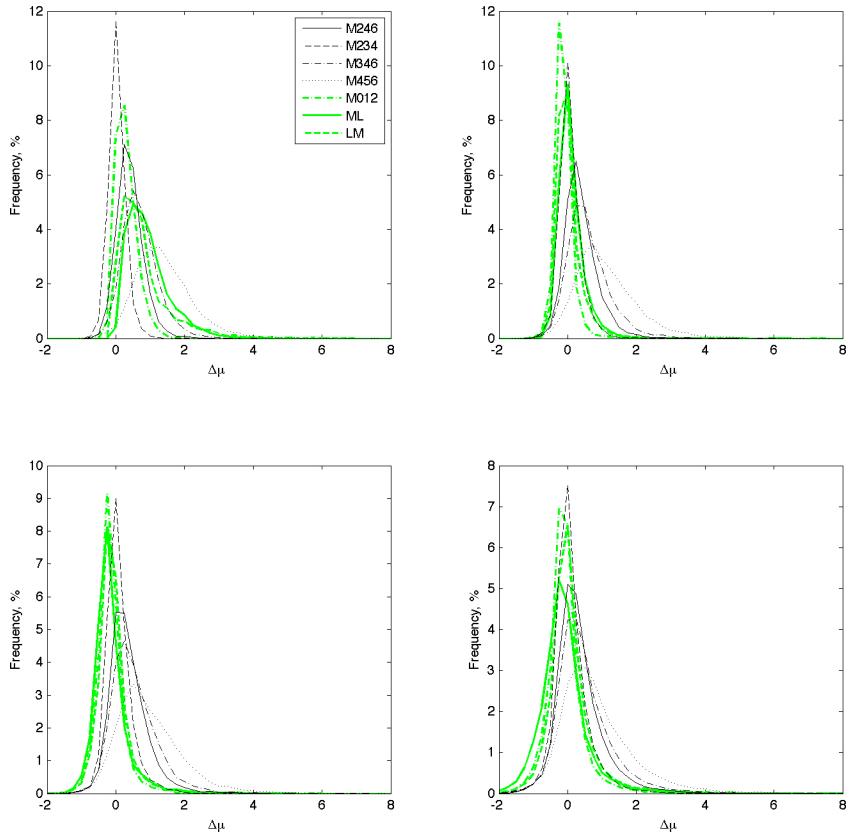

Figure 3.4: The frequency of $\Delta\mu$, normalized by total number of minutes, for different estimators based on simulation A, given the range of true μ: a) $-1 < \mu \leq 0$; b) $0 < \mu \leq 1$; c) $1 < \mu \leq 2$; d) $2 < \mu \leq 4$.

M012. This is reasonable because the low moment estimator M012 also depends largely on the low end of DSD. It should be emphasized here that smaller estimate bias of DSD parameters does not guarantee better estimates of integral parameters because the DSD model is highly nonlinear. In the next subsection, these estimators will be evaluated by considering the estimation of integral parameters. For the rest of this subsection, more details of DSD parameter estimation are to be addressed.

Table 3.3: Fractional Errors (%) for different moments of simulated Gamma DSDs

	M_2	M_3	M_4	M_5	M_6
$0 < R < 1$	7.39	8.52	11.00	14.63	19.41
$1 < R < 3$	4.67	5.66	7.65	10.76	15.13
$3 < R < 5$	3.81	4.75	6.61	9.54	13.82
$5 < R < 10$	3.36	4.20	5.88	8.63	12.79
$10 < R < 20$	2.78	3.60	5.19	7.81	11.84
$20 < R < 35$	2.28	2.94	4.24	6.46	9.91
$35 < R < 50$	1.90	2.45	3.53	5.30	8.11
$50 < R < 80$	1.58	2.08	3.05	4.62	7.02
$80 < R$	1.23	1.56	2.21	3.30	5.00
All	3.87	4.71	6.43	9.14	13.08

Table 3.4: Biases of μ for different estimators of simulated Gamma DSDs

	M246	M234	M246 *	M234 *
$0 < R < 1$	0.72	0.31	10.78	13.18
$1 < R < 3$	0.56	0.24	4.73	6.04
$3 < R < 5$	0.54	0.23	2.97	3.76
$5 < R < 10$	0.51	0.21	2.10	2.60
$10 < R < 20$	0.49	0.20	1.49	1.77
$20 < R < 35$	0.42	0.17	1.11	1.30
$35 < R < 50$	0.35	0.14	0.93	1.10
$50 < R < 80$	0.28	0.11	0.76	0.92
$80 < R$	0.21	0.07	0.54	0.66

$*$ means truncation of raindrops < 0.6 mm

As discussed above, the moment error is a key factor in accurately estimating DSD parameters. As shown in Figs. 3 and 4 of Zhang et al. (2003), the standard deviation of the estimated μ and Λ parameters increases with increasing moment errors or with decreasing correlation between moments. For simulated data, the relative moment errors are calculated by

$$F = \overline{\left| \frac{X - \bar{X}}{X} \right|} \tag{3.15}$$

where the parallel bars "$|\cdot|$" stand for the absolute value, the upper bar "$-$" denotes the average of data, X represents the integral variable of interest and F represents the fractional error. Table 3.3 lists the fractional error for moments from the 2^{nd} order to the 6^{th} order based on dataset of simulation A. It is seen that the higher the R, the smaller the moment error and the higher the moment order, the larger the moment error. Based on these moments, the DSD parameters are estimated. The biases of μ estimated by M234 and M246 are listed in the two left-most columns of Table 3.4. The estimation bias decreases with increasing R, and M234 has a smaller bias than M246. It is noted that the bias of μ is not large. This implies that M246 and M234 estimators have the potential to retrieve physical information from the observations despite their biases.

3.3.2.3 Uncertainties of estimated integral variables for a gamma DSD

The goodness of an estimators performance depends on the definition of "goodness" of estimation. In this study, the "goodness" of an estimator is examined by the fractional error in estimating five integral parameters N_T, R, D_m, Z_H, and Z_{DR}. The relative performances of seven estimators are investigated for each parameter. Given that all five evaluated parameters have physical significances, the better estimator is suggested according to the overall performance. The formulas of N_T, R, D_m, Z_H and

Z_{DR} refer to Eqs. 2.10, 2.13, 2.18, 2.21 and 2.22.

Table 3.5 lists the estimation biases for five integral variables. Table 3.6 shows the corresponding fractional errors of the estimates. R and D_m are evaluated in linear units while Z_H, Z_{DR} and N_T are evaluated in logarithmic units. It is seen that M234 and M246 have a similar performance and should be considered as the two best estimators. The ML and LM estimators, like M012, have the best estimates for N_T but the worse estimates for other parameters. As expected, M456 has the worst estimates for N_T. It is found that the performance of an estimator is related to rain intensity. The dataset was sorted for several ranges of R. Within each range, the fractional error of the estimation is calculated according to Eq. 3.15. The fractional error curves for different estimators are plotted in Fig. 3.5. These estimators generally provide a better estimation with increasing R than the ML and LM estimators except for N_T. This result implies that the latter two estimators are not practical when observed DSD have large errors for small drop sizes. M246, M234, M346, and M456 perform similarly when estimating Z_H and R, implying that the error of an estimated variable is determined mainly by the error of the moment with the closest moment order to it. For example, the R is approximately proportional to the 3.67^{th} moment. The uncertainty of the estimated R comes mainly from the error of the 4^{th} moment applied by estimators. M234 provides a relatively better estimate than other estimators (except M012 for N_T), though the difference is slight compared with M246. Compared to M246, the improvement of M234 is not much. This is consistent with the error analysis of M234 by Smith et al. (2005), who believed that middle moment estimators are better than the estimators with low or high moments. According to the simulation, this conclusion is appropriate if the true DSD is a gamma distribution.

Table 3.5: Bias for different integral variables based on simulated Gamma DSDs

	M246	M234	M346	M456	M012	ML	LM
R(mm h^{-1})	-0.094	-0.185	-0.112	-0.156	-0.248	-1.221	-1.197
D_m(mm)	-0.032	-0.022	-0.022	0.001	-0.03	-0.102	-0.094
Z_H(dBZ)	-0.44	-0.284	-0.441	-0.443	-0.439	-1.838	-1.709
Z_{DR}(dB)	-0.093	-0.048	-0.1	-0.117	-0.048	-0.143	-0.13
N_T(dB)	-0.057	-0.018	-0.096	-0.179	-0.001	-0.001	-0.001

Table 3.6: Fractional errors (%) for different integral variables based on simulated Gamma DSDs

	M246	M234	M346	M456	M012	ML	LM
$F[R$(mm h^{-1})] in %	5.49	5.51	5.49	5.5	7.65	18.25	14.75
$F[D_m$(mm)] in %	2.85	2.56	2.56	2.54	4.38	10.48	8.43
$F[Z_H$(dBZ)] in %	2.23	1.99	2.23	2.23	3.43	9.17	7.75
$F[Z_{DR}$(dB)] in %	10.5	7.69	11.13	12.5	11.62	22.79	18.41
$F[N_T$(dB)] in %	2.29	1.27	3.54	6.19	0.93	0.93	0.93

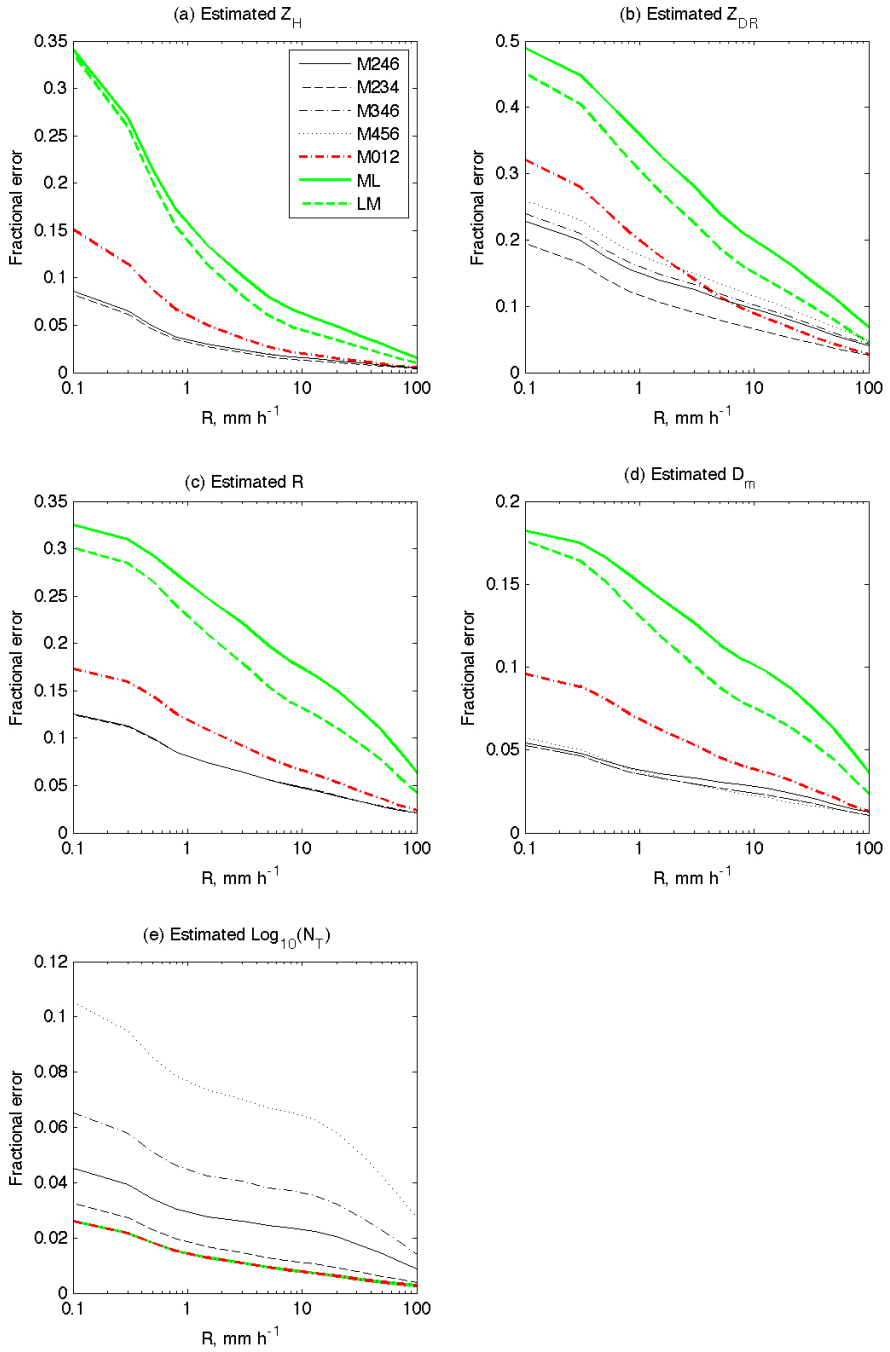

Figure 3.5: The dependences of fractional errors (in %) of estimated integral variables, Z_H (dBZ), Z_{DR} (dB), R (mm h^{-1}), D_m (mm) and $\log_{10}[N_T$ (m^{-3})], on the rain-rate. Seven estimators are evaluated. The dataset is obtained using simulation A.

3.3.2.4 Uncertainties in the estimates for a non-gamma DSD

In Simulation B, the DSD is assumed to have a non-gamma distribution, and the same analysis is made as the previous subsection. As shown in Table 3.7, moment errors for Simulation A and B are very similar, implying that the following results are probably affected not by the change of moment errors but rather by the change of DSD shape. The bias and fractional error results are shown in Table 3.8 and Fig. 3.6. Compared to the results based on the gamma distribution assumption, the estimation uncertainties of R, D_m and N_T have few changes for almost all estimators. One evident change is that M234 is no longer the best estimator for Z_H and Z_{DR}. Although the measured middle moments still have fewer errors (as shown in Table 3.7), the middle moment estimator is degraded when the DSD has model errors. As shown in Fig. 3.6, the estimate uncertainty of Z_{DR} is apparently enlarged for M234, M012, ML and LM but reduced for M246, M346, and M456, especially if R is less than 10 mm h^{-1}. M456, instead of M234, becomes most accurate in estimating Z_{DR}, indicating that high moment estimators are less sensitive to model error in estimating Z_{DR} than low moment estimators. Because the natural DSD is not an exact gamma distribution, low and middle moment estimators might not be the best ones to estimate high-moment parameters. Looking at Tables 3.7–3.9 and Figs. 3.5 and 3.6, there is no fundamental difference for estimators such as M234, M246, and M346 in estimating N_T, R, D_0, Z_H, and Z_{DR}.

Table 3.7: Fractional Errors (%) of different moments based on simulated DSDs

	M_0	M_1	M_2	M_3	M_4	M_5	M_6
Gamma	6.32	4.05	3.87	4.71	6.43	9.14	13.08
Non-Gamma	6.32	4.02	3.85	4.70	6.27	8.45	11.09

Table 3.8: Same as Table 3.5, but based on simulated Non-gamma DSDs

	M246	M234	M346	M456	M012	ML	LM
R(mm h^{-1})	0.024	-0.067	0.006	-0.038	-0.116	-1.041	-1.037
D_m(mm)	-0.013	-0.004	-0.004	0.019	-0.011	-0.083	-0.075
Z_H(dBZ)	-0.08	0.077	-0.081	-0.083	-0.075	-1.466	-1.341
Z_{DR}(dB)	0.019	0.063	0.011	-0.006	0.065	-0.03	-0.017
N_T(dB)	-0.058	-0.018	-0.097	-0.179	-0.001	-0.001	-0.001

Table 3.9: Same as Table 3.6, but based on simulated Non-gamma DSDs

	M246	M234	M346	M456	M012	ML	LM
$F[R$(mm h^{-1})] in %	5.52	5.46	5.5	5.48	7.56	18.17	14.63
$F[D_m$(mm)] in %	2.36	2.28	2.28	2.8	4.28	10.26	8.21
$F[Z_H$(dBZ)] in %	1.84	1.88	1.84	1.84	3.31	8.88	7.41
$F[Z_{DR}$(dB)] in %	8.25	10.86	8.13	7.87	15.07	24.15	19.13
$F[N_T$(dB)] in %	2.36	4.88	3.61	6.26	0.93	0.93	0.93

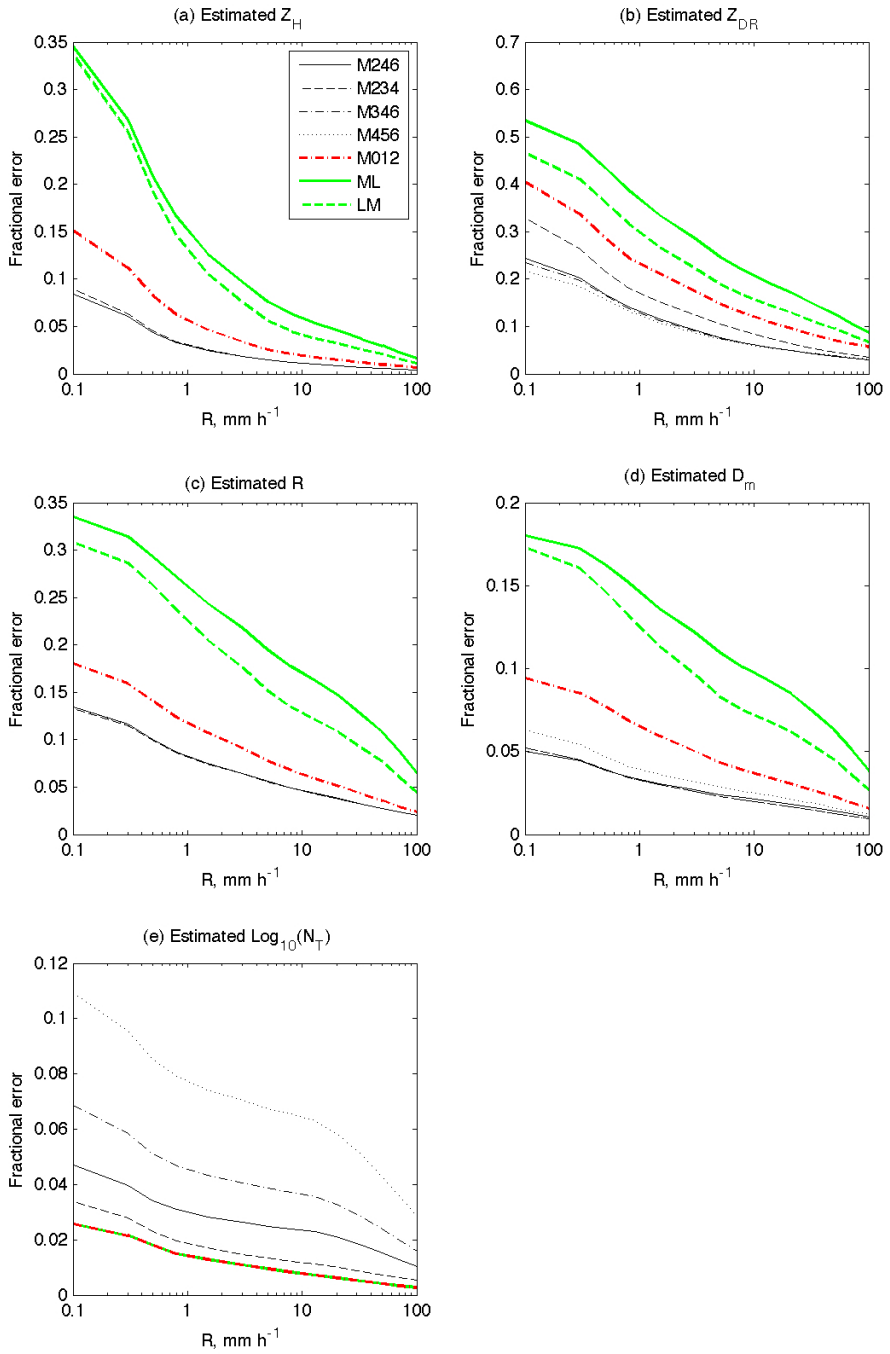

Figure 3.6: The same as Fig. 3.5 except the dataset is from simulation B, which simulates the non-gamma DSDs.

3.3.3 Evaluating estimators using 2DVD dataset

Evaluation using real DSD data gives similar results to the analysis made in previous section based on simulation B, i.e., non-gamma distribution assumption. Following the analysis in the previous section, N_T, R, D_m, Z_H, and Z_{DR} are directly calculated using the observed DSDs, as well as from gamma DSD parameters estimated with each of the seven estimators. Tables 3.10 and 3.11 lists the biases and the fractional errors of estimation, assuming the observed DSDs to be the reference. It is shown that ML and LM estimators, close to M012, have the largest biases and errors in estimating R, D_m, Z_H, and Z_{DR}. Estimators M246, M234, M346 and M456 have small biases and little uncertainty in estimating the five variables, with a few exceptions. Similar to results in previous subsection, M012 has a slightly better performance than ML and LM, but still has large fractional errors in estimating R, D_m, Z_H, and Z_{DR}. M234 is not as accurate as M246, M346 and M456 in estimating Z_H and Z_{DR}. From Tables 3.10 and 3.11, it is possible to conclude that M234 may be more sensitive to model errors than high moment estimators, despite the fact that middle moments have lower sampling errors. Moreover, ML and LM estimators are not accurate in estimating high moment integral parameters for real data because they are sensitive to measurement error at the low end of DSD, as well as to model error.

Table 3.10: Same as Table 3.5, but based on observed DSDs

	M246	M234	M346	M456	M012	ML	LM
R(mm h^{-1})	0.017	0.02	0.015	-0.008	-0.694	-1.307	-1.336
D_m(mm)	-0.009	0	0	0.029	-0.084	-0.1468	-0.151
Z_H(dBZ)	-0.003	0.104	-0.002	-0.002	-0.965	-1.941	-2.070
Z_{DR}(dB)	0.045	0.076	0.04	0.02	-0.086	-0.171	-0.172
N_T(dB)	0.369	0.059	0.018	-0.072	0	0	0

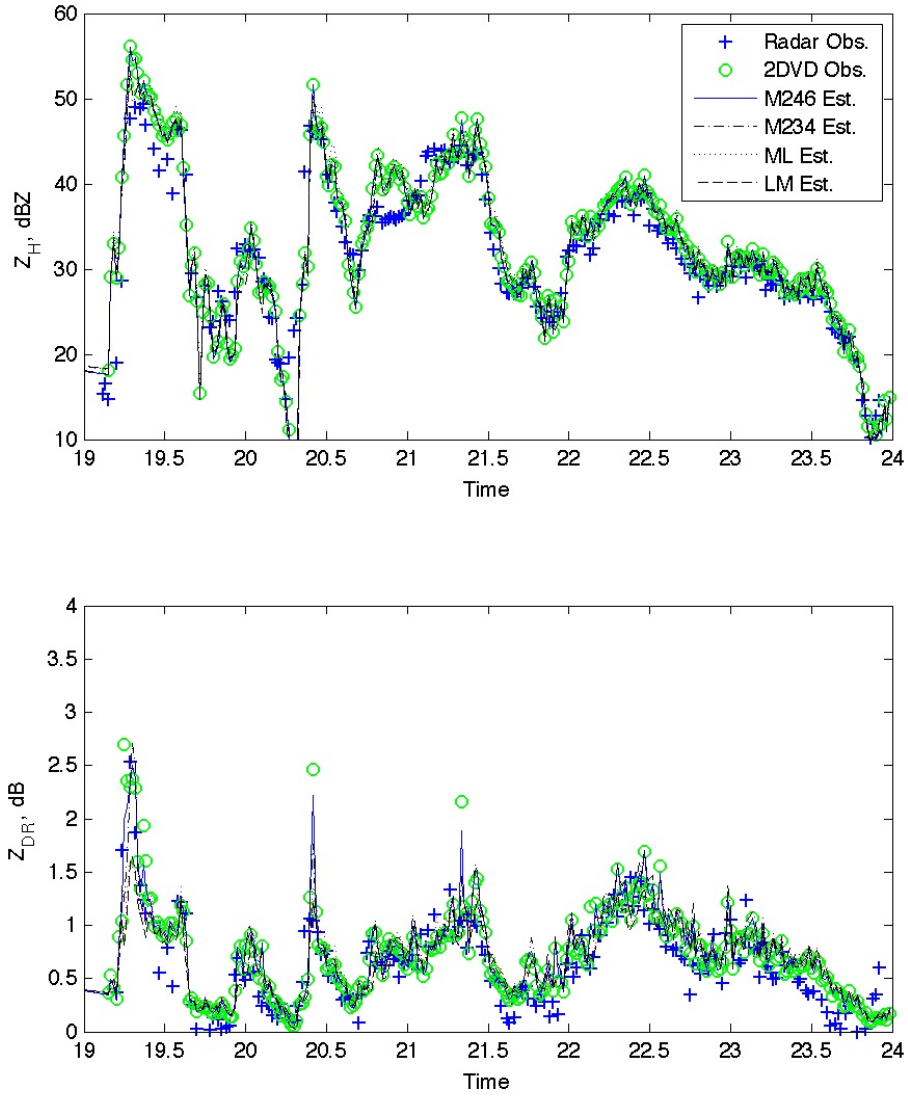

Figure 3.7: The comparisons of (a) Z_H and (b) Z_{DR}. The cross points indicate the observations by the polarimetric radar. The circle points represent the calculations based on DSDs observed by the 2DVD. Thin solid, dash-dotted, dotted and thin dashed lines indicate the calculation based on DSDs estimated by M246, M234, ML and LM estimators. The statistic values of these curves are listed in Tables 3.12 and 3.13

Table 3.11: Same as Table 3.6, but based on observed DSDs

	M246	M234	M346	M456	M012	ML	LM
$F[R(\text{mm h}^{-1})]$ in %	0.4	0.45	0.29	0.05	7.38	16.0	15.30
$F[D_m(\text{mm})]$ in %	1.01	0	0	2.18	7.33	11.76	11.41
$F[Z_H(\text{dBZ})]$ in %	0.07	0.96	0.06	0.03	5.52	9.03	8.73
$F[Z_{DR}(\text{dB})]$ in %	8.64	16.97	7.4	4.07	32.23	39.89	37.71
$F[N_T(\text{dB})]$ in %	4.5	3.46	6.23	8.96	0	0	0

In practice, disdrometer observations are usually used as the ground truth to verify radar observations. However, there are always discrepancies between ground and radar observations. One example is given in Fig. 3.7. The time-series values of Z_H and Z_{DR} observed by NCAR S-Pol radar [Brandes et al. (2002, 2004a); Zhang et al. (2003)] are plotted in cross points. The disdrometer was deployed about 38 km from the radar and the elevation angle of radar beam was 0.5°. The circle points represent the Z_H and Z_{DR} calculations from disdrometer observations underneath the radar resolution volume. By comparison, estimates based on M246, M234, ML and LM estimators are plotted in thin solid, dash-dotted, dotted and thin dashed lines, respectively. The statistics of mean difference and standard deviation for these curves versus radar measurements are listed in Tables 3.12 and 3.13. As they show, the results from four moment estimators are very close to disdrometer observations, though ML and LM estimators have the worst performance. The moment estimators cannot compensate for, nor reduce, the discrepancy between the radar and the disdrometer observations. This implies that the discrepancy is mainly due to instrumental factors and physical reasons, and not to the DSD moment error measured by the disdrometer. From the result, it can be concluded that the choice of moment estimator is not an effective solution for improving radar-disdrometer comparison.

Table 3.12: Mean difference between 2DVD calculations and radar observations

	2DVD Obs. Vs. Radar Obs.	M246 Est. Vs. Radar Obs.	M234 Est. Vs. Radar Obs.	ML Est. Vs. Radar Obs.	LM Est. Vs. Radar Obs.
Z_H(dBZ)	0.78	0.64	0.74	1.05	0.52
Z_{DR}(dB)	0.11	0.12	0.14	0.09	0.05

Table 3.13: Standard deviation of 2DVD calculations versus radar observations

	2DVD Obs. Vs. Radar Obs.	M246 Est. Vs. Radar Obs.	M234 Est. Vs. Radar Obs.	ML Est. Vs. Radar Obs.	LM Est. Vs. Radar Obs.
Z_H(dBZ)	3.07	3.09	3.04	3.42	3.17
Z_{DR}(dB)	0.17	0.18	0.21	0.21	0.20

3.4 Analysis of C-G DSD Model

3.4.1 C-G DSD model

Although gamma distribution is flexible to represent natural DSDs, its three degrees of freedom increase the complexity of DSD retrieval from radar observations. For example, Bringi et al. (2002) proposed using polarimetric radar measurements of Z_H, Z_{DR} and specific differential phase (K_{DP}) to retrieve a normalized gamma DSD. However, Brandes et al. (2004a) have shown that this approach is sensitive to the K_{DP} noise. In addition, K_{DP} is derived from measurements made over many range gates and does not always match Z_H and Z_{DR} measurements well at every range gate. Therefore, the addition of K_{DP} may result in a deterioration of DSD retrieval at a specific range gate, especially if it is not used optimally. Zhang et al. (2001) introduced a constraint-gamma (C-G) DSD model, which had a constraint $\mu - \Lambda$ relationship derived from disdrometer dataset. This model reduces the freedom of

DSD model to two and facilitates the DSD retrieval from dual-polarization or dual-frequency radar measurements. The C-G DSD model has been validated by several studies of DSD retrieval [e.g., Vivekanandan et al. (2004); Brandes et al. (2004a,b); Zhang et al. (2006a); Cao et al. (2008); Cao and Zhang (2009)].

3.4.2 Truncation issues on parameter correlation

Moisseev and Chandrasekar (2007) argued that the relation between the estimated μ and Λ parameters may not represent physics. However, their results were based on an unrealistic simulation with artificially raised moment errors. For example, their main conclusions were based on the truncation of the low end (< 0.6 mm) of a DSD spectrum. This kind of truncation greatly enlarges moment errors, especially for low moments. The moment error for the truncation of 0.6 mm is listed in Table 3.14. The bias of estimated μ is listed in the two right-most columns of Table 3.4. It is shown that moment errors are significantly increased for low Rs and the moment M_2. The bias of estimated μ is also increased. However, this kind of truncation does not realistically reflect disdrometer observations. The 2DVD has a good resolution of 0.195 mm for the prototype model or 0.132 mm for the latest model, from which data are used for this study. This means that the 2DVD could measure raindrops at least as small as 0.2 mm. Although small raindrops might be under-sampled due to environmental effects, it is not possible to miss all the raindrops smaller than 0.6 mm. In addition, Moisseev and Chandrasekar's simulation used a sampling volume of 1.5 m^3 that is equivalent to a shorter sampling time of 30 seconds and a uniform terminal velocity 5 m s^{-1}, leading to a larger error for high moments. These limitations enlarged the sampling error, as did the truncation of size less than 0.6 mm. Consequently, the biases of estimated μ and Λ (not shown) are significantly increased. The histograms of r and $\Delta\mu$ are shown in Fig. 3.8. It is shown that μ may even be overestimated by 20, which is beyond the meaningful range. The artificial overestimations lead to a

spurious correlation between μ and Λ parameters [as shown in Fig. 5, Moisseev and Chandrasekar (2007)]. This correlation is unrealistic and cannot verify the presumption that the observed $\mu - \Lambda$ relation by Zhang et al. (2001) is mainly attributed to error effects.

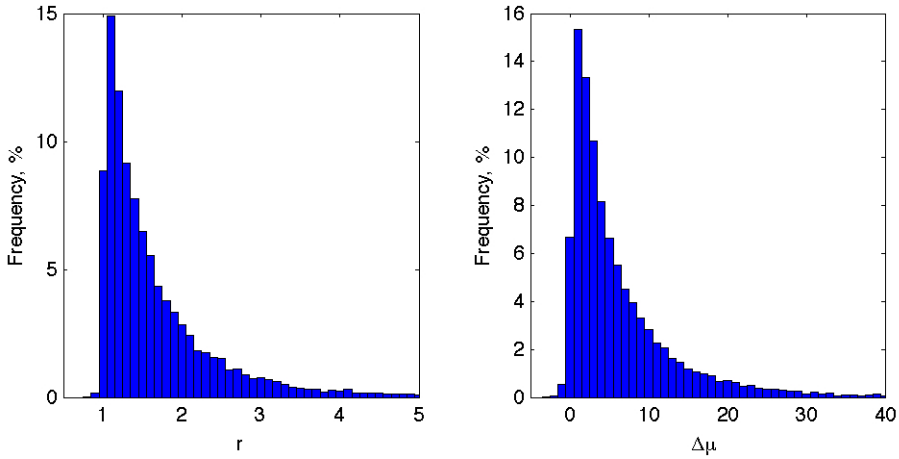

Figure 3.8: The histograms of (a) r and (b) $\Delta\mu$ distributions. The estimator is M246. The dataset is the simulation A, but the low end of DSDs (< 0.6 mm) is truncated.

3.4.3 Refinement of $\mu - \Lambda$ relation

3.4.3.1 DSD sorting and averaging based on two parameters (SATP)

It is well known that DSD variability can be reduced by averaging. For example, Joss and Gori (1978) demonstrated that random, time-sequential, and rain-rate sequential averaging will lead to exponential DSDs. Sauvageot and Lacaux (1995), considering "instantaneous" DSDs having strong variability, further studied averaged DSDs of JWD data within a set of rain-rate intervals, finding that the rain-rate reflectivity ($R - Z_H$) relations obtained from averaged DSDs are close to those calculated from non-averaged data and compatible with those proposed in previous studies. Lee and Zawadzki (2005) introduced the sequential intensity filtering technique (SIFT), which

Table 3.14: Same as Table 3.3, but the low end of DSD (< 0.6 mm) is truncated.

	M_2	M_3	M_4	M_5	M_6
$0 < R < 1$	48.36	32.08	22.60	19.79	21.63
$1 < R < 3$	31.25	18.73	13.29	12.93	15.89
$3 < R < 5$	22.30	12.56	9.48	10.42	14.08
$5 < R < 10$	16.20	8.85	7.35	9.04	12.92
$10 < R < 20$	11.68	6.32	5.87	7.97	11.89
$20 < R < 35$	8.52	4.52	4.58	6.53	9.93
$35 < R < 50$	7.08	3.53	3.72	5.35	8.13
$50 < R < 80$	5.82	2.87	3.17	4.64	7.03
$80 < R$	3.74	1.93	2.26	3.31	5.01
All	21.86	13.24	10.12	10.63	13.67

was to be used for processing a single rain event, sorting DSDs within a time window (typically one hour) by the reflectivity and averaging consecutive DSDs (typically 10). They found that averaging DSDs within an interval of reflectivity could reduce observational errors of disdrometer measurements and yield more stable $R - Z_H$ relations. The averaging methods mentioned above, however, apply a coarse filtering technique that result in a significant loss of physical variations. A better method to process disdrometer data is needed—one that can preserve the physical variability while reducing the impact of observational error.

As shown in the previous subsection, sampling error is an unavoidable problem for 2DVD measurements, resulting in the degradation of the DSD fitting [Zhang et al. (2003)]. Since the DSD models are based on 2DVD observations (e.g., the C-G DSD model depends on fitted shape and slope parameters), the challenge is to reduce the sampling errors so that fitted shape and slope parameters are less affected by errors.

Here, we suggest a sorting and averaging method based on two parameters (SATP). The SATP method is proposed due to the fact that numerous 2DVD measurements are available with which the constraining shape-slope relation of C-G DSD model could be developed. Unlike the SIFT method introduced by Lee and Zawadzki (2005). SATP is applied to a whole dataset rather than a single event. With SATP, two parameters are used to characterize the DSD. Physical variability is therefore preserved much better than with SIFT. The SATP procedure is briefly described as follows.

(i) Select two characteristic parameters to build two-dimensional grids.

(ii) Calculate both characteristic parameters based on 1-min DSD measurements.

(iii) Sort the whole dataset and find DSDs with similar physical characteristics according to their two characteristic parameters.

(iv) Average observed DSDs located in the same grid to obtain a new DSD.

(v) Process averaged DSDs (i.e., fit to a gamma distribution) to develop the shape-slope relation of C-G DSD model.

The characteristic parameters can be any two rain variable (e.g., DSD moments, characteristic sizes of DSD, etc.). In general, the high moments, which have relatively larger measurement errors, and the low moments, which are determined by small drops and susceptible to disdrometer measurement uncertainty, do not represent rain physics well. On the other hand, the middle moments are both more representative of rain physics and more accurately measured. In this study, the middle moment-related parameters, R and D_0, are therefore chosen for processing. The sorting grids used in this study are shown in Fig. 3.9. Each grid in the $R - D_0$ plane is defined by variations of $\pm 5\%$ for R and ± 0.025 mm for D_0. The bar length indicates the number of observed DSDs. DSDs within each grid pixel are characterized by small variations

of R and D_0 and are assumed to represent the similar rain physics. For example, Fig. 3.10 shows three groups of observed DSDs (thin solid lines) within three different grids. The dashed, bold solid and dash-dotted lines represent three mean DSDs. For the three groups of DSDs in Fig. 3.10, FSDs of R are 5.58%, 11.82% and 5.1%, FSDs of D0 are 1.53%, 1.86% and 0.76%, FSDs of N_T are 1.75, 1.73 and 1.46 dB, and FSDs of Z_H are 1.57, 1.85 and 0.72 dB, respectively. These FSDs imply that SATP has the potential to identify similar DSDs. If we assume that the sorted DSDs within each grid pixel have the same DSD (expected value) and similar sampling error, the latter can surely be reduced by averaging the sorted DSDs. The fit to the averaged DSD is therefore less affected by errors and represents the physics better than the fit to non-averaged DSDs. Compared to one-parameter filtering methods, SATP better preserves the physical variation.

SATP is a procedure for processing large amounts of raw data that contain errors. Because it is hard to exactly quantify the error reduction of each individual grid, SATP is not intended to analyze "instantaneous" integral rain parameters. SATP is designed to provide the mean property of the DSD (i.e., the mean $\mu - \Lambda$ relation) rather than improve upon the individual observation. Because the physics is preserved with less error, the mean $\mu - \Lambda$ relation processed by SATP is obviously better than that obtained directly from error-contaminated measurements. Determining the error reduction for a specific DSD is not the major concern. On the other hand, the frequency distribution (Fig. 3.9) shows that light rain data ($R < 3$ mm h^{-1}) account for a large portion of the dataset. The $\mu - \Lambda$ relation derived from the data prior to SATP filtering will be largely controlled by those data. Fitted DSDs of light rains always have large μs and Λs, and unfiltered results will raise the slope of mean $\mu - \Lambda$ relation and cause the retrieval using $\mu - \Lambda$ relation to deteriorate. SATP reduces the effect of light rain events and represents other rain events well. It is worthwhile

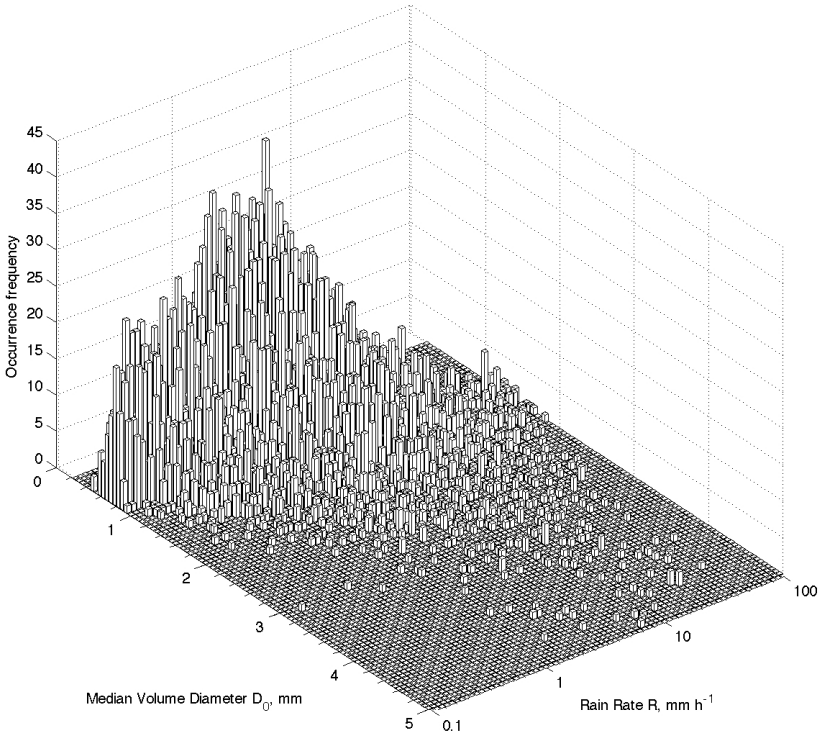

Figure 3.9: Occurrence frequency of sorted rain data based on rain rate (step 10%) and median volume diameter (step 0.05 mm). Each pixel of the $R-D_0$ plane represents a specific DSD. The bar over the pixel denotes the number of observed DSDs sorted for one specific DSD. Observed DSDs within a pixel are averaged to obtain the specific DSD.

to note that SATP could be improved if more parameters (e.g. three parameters) are used to sort the similar DSDs. In this study, the dataset is not sufficient so that only two parameters have been applied. In next subsection, SATP is applied to derive a constrained $\mu - \Lambda$ relation.

3.4.3.2 Refined $\mu - \Lambda$ relation

The shape-slope relation of the C-G DSD model may vary in different climate regions. Previous studies [e.g., Zhang et al. (2001); Brandes et al. (2004b)] have shown that

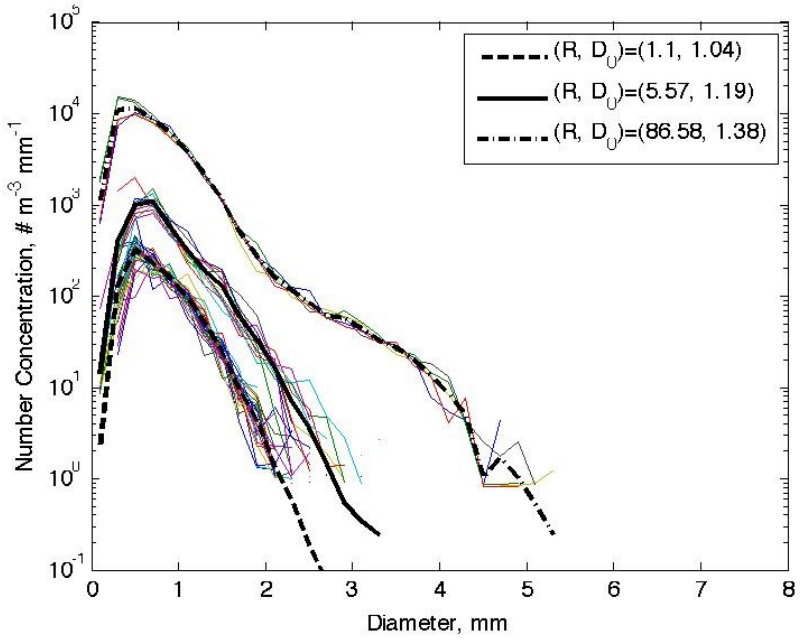

Figure 3.10: Example of sorted DSDs and their mean DSDs. Thin solid lines represent the observed DSDs, which are sorted into three grids using SATP method. DSDs within one grid have the similar distribution. Three mean DSDs are denoted as bold dashed line, bold solid line, and bold dash-dotted line, respectively. Those mean the DSDs, which represent (R, D_0)=(1.1 mm h^{-1}, 1.04 mm), (5.57 mm h^{-1}, 1.19 mm), and (86.58 mm h^{-1}, 1.38 mm), respectively.

the relation for the Southern Great Plains (i.e., Oklahoma) is a little different than the one for a subtropical region (i.e., Florida). Using the SATP method, 2DVD data were processed to refine the $\mu - \Lambda$ relation for rains in Oklahoma. First, the data were grouped on a $R - D_0$ grid and averaged. Averaged DSDs were then fitted to a gamma distribution by the truncated moment fit [TMF, Vivekanandan et al. (2004)], a modified M246 estimator. After that, the second order polynomial least-square fit was used to obtain the mean $\mu - \Lambda$ relation. The fitted μs and Λs for sorted and averaged DSDs are plotted in Fig. 3.11. The solid line is the fitted curve of circle points. A dashed line depicts the Florida $\mu - \Lambda$ relation from Zhang et al. (2001). The dashed line generally has larger μs than the solid line, given the same Λs. It implies

that DSDs in Florida tend to have a narrower shape than DSDs in Oklahoma. The solid line in Fig. 3.11 is the refined $\mu - \Lambda$ relation of C-G DSD model used for DSD retrieval and is given by

$$\mu = -0.0201\Lambda^2 + 0.902\Lambda - 1.718 \tag{3.16}$$

Eq. 3.16 is applicable for a Λ within a range from 0 to 20. Larger Λ values are thought to be the result from measurement errors rather than storm physics [Zhang et al. (2003)].

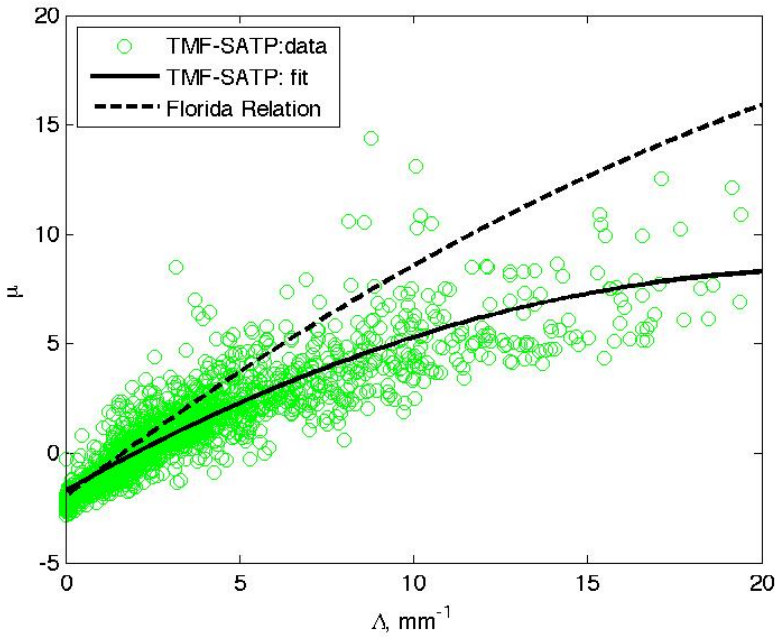

Figure 3.11: Scatter diagram of $\mu - \Lambda$ with DSD sorting. Circles denote DSD data fitted by TMF method. Solid line is the mean curve fitted to circle points by two-order polynomial fit. Dashed line corresponds to the Florida relation [Zhang et al. (2001)].

To verify the refined $\mu - \Lambda$ relation, we examine the D_m and standard deviation of mass-weighted diameter distribution (σ_m) because both can be directly derived from observations and are independent of sorting and fitting procedures. If relation 3.16

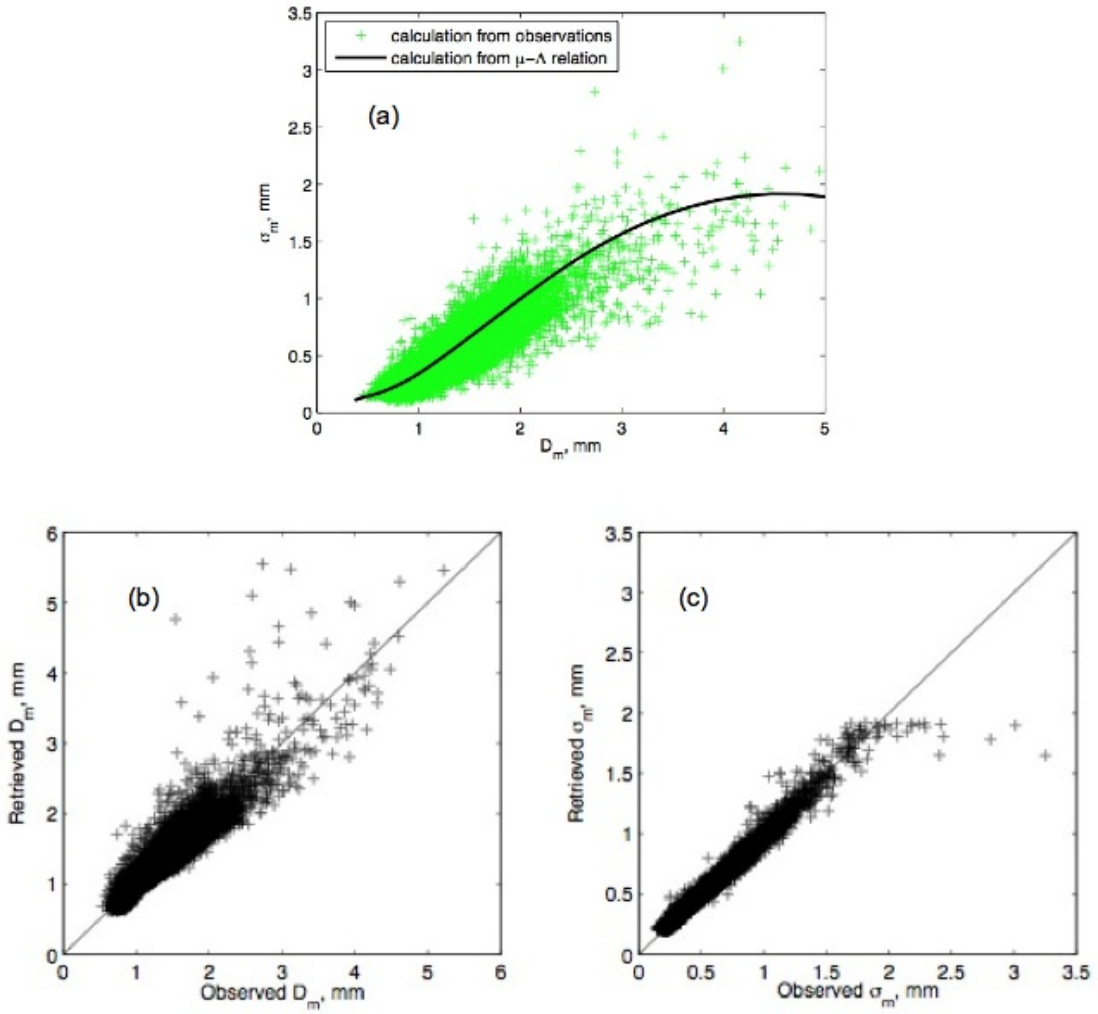

Figure 3.12: (a) Scatter diagram of D_m versus σ_m. Circles denote that D_m and σ_m are calculated from observed DSDs of 14200 minutes. The solid line denotes that D_m and σ_m are calculated from gamma DSDs with $\mu - \Lambda$ constrained by relation 3.16. (b) Scatter diagram plot of retrieved D_m versus observed D_m that using relation 3.16. Crosses denote the data points, and solid line corresponds to the unit slope. The bias is -2.18% and the correlation coefficient is 0.915. (c) The same as (b) but for σ_m. The bias is -1.15% and the correlation coefficient is 0.985.

represents rain physics, the $D_m - \sigma_m$ relation derived from observations and from relation 3.16 should be consistent. Fig. 3.12a shows the result of these calculations. Crosses denote calculations of D_m and σ_m from observed 1-min DSDs. The solid line, derived from relation 3.16, agrees with observations. Further verification was done by examining retrieved D_m and σ_m using the relation 3.16. Scatter diagram of retrieved and observed D_m and σ_m are shown in Fig. 3.12b and c. Some data points deviate substantially from the solid line (e.g., observed σ_m larger than 2.4 mm). These data points typically are DSDs with long and poorly sampled tails (not shown). However, the outliers are few in number and don't contaminate the result. The bias of retrieved D_m using relation 3.16 is only -0.03 mm (or -2.18%) and the correlation coefficient is 0.915 (Fig. 3.12b). The bias of retrieved σ_m is only -0.007 mm (or -1.15%) and the correlation coefficient is 0.985 (Fig. 3.12c). The same comparison procedure can be applied for rain variables of W, R, D_0, and N_T. For observations with $R < 100$ mm h^{-1}, Table 3.15 lists bias and correlation values of retrievals versus observations for several rain variables. Except for N_T, all these variables have a small bias and a correlation coefficient close to 1. These results show that the refined $\mu - \Lambda$ relation is valid for the rain DSD retrieval.

Table 3.15: Bias and correlation coefficient for retrieved rain variables versus observations

	D_m	σ_m	W	R	D_0	N_T
Bias (%)	-2.18	-1.15	2.52	3.37	8.73	14.16
Corr. Coef.	0.915	0.985	0.967	0.986	0.819	0.763

3.4.3.3 Further study of physical information of $\mu - \Lambda$ relation

According to the analysis of simulated data, M234 and M246 have an overall good performance among seven estimators in estimating integral variables (Z_H, Z_{DR}, R,

D_m and N_T), whether the true DSD follows a gamma distribution exactly or approximately. This means that the parameters N_0, μ, and Λ estimated by M246 can represent the physical property of a DSD. Consequently, it is reasonable to think that the $\mu - \Lambda$ relation (e.g., Eq. 3.16), is at least partially due to rain microphysics. In this subsection, it is demonstrated that there is an equivalence between the $\mu - \Lambda$ relation and the mean relation derived through the normalization approach proposed by Testud et al. (2001).

Normalization [e.g., Testud et al. (2001); Bringi et al. (2002)] has been widely accepted as an approach of studying rain physics from observed DSDs. Testud et al. (2001) demonstrated that the "average" of normalized DSD spectra is sufficiently stable that only two parameters are needed to describe the DSD. The normalized distribution $F_N(X)$ is calculated by

$$F_N(X) = N(D)/N_0^* \tag{3.17}$$

with

$$N_0^* = \frac{4^4 W}{\pi \rho_w D_m^4} \tag{3.18}$$

$$X = D/D_m \tag{3.19}$$

where, N_0^* is the normalization term and ρ_w is the density of water. Following Testud's approach, the normalization procedure is performed to process the disdrometer data. The "average" and the standard deviation of normalized DSDs are shown in Fig. 3.13. The solid line denotes the "average" curve (referred to as the "T-function"). The standard deviation (represented by the error bar) is moderate for median X ($0.4 < X < 1.5$). For $X > 1.5$, there is a large fluctuation for the normalized distribution. It is seen that Fig. 3.13 is very similar to Figs. 3-6 in Testud et al. (2001), though the DSD spectra were not classified into different rain categories in

Fig. 3.13 because rain categories were irrelevant to our goal. The purpose of deriving this mean curve is to evaluate the performance of the $\mu - \Lambda$ relation derived from observed DSDs. Retrievals based on the T-function were compared with those based on the $\mu - \Lambda$ relation 3.16.

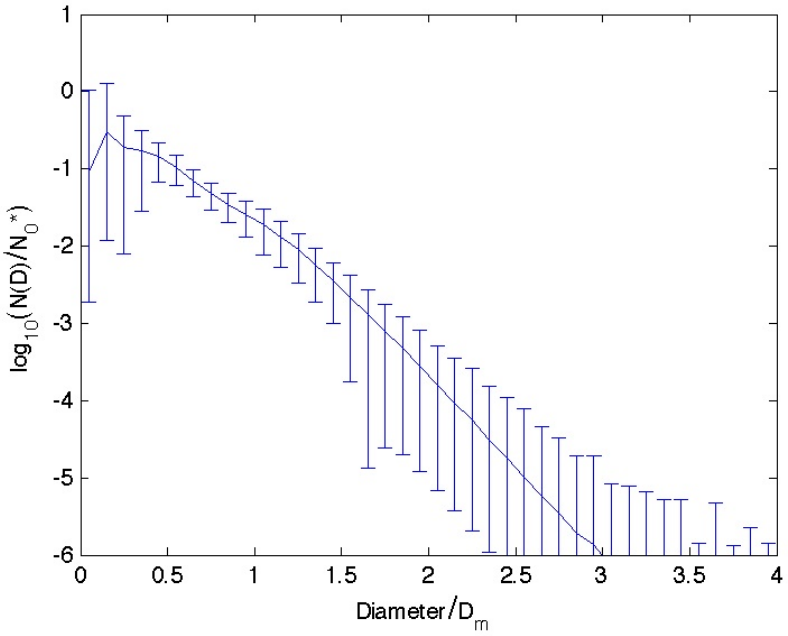

Figure 3.13: The statistics of normalized DSDs following Testud's approach. The solid line is the "average" of the normalized DSDs (referred to as the "T-function"). The error bars represent the standard deviation.

A simulation of normalized DSDs based on the $\mu-\Lambda$ relation is given by Fig. 3.14. The thick solid line is the curve of T-function. The thin solid lines represent a set of normalized DSDs based on the $\mu-\Lambda$ relation with Λ varying from 0 to 10. It is shown that the T-function could be viewed as the mean of thin lines—results of the $\mu - \Lambda$ relation. The $\mu - \Lambda$ relation allows for more flexible shapes of the DSD, especially for the low end. The thin lines have a similar variability to those observed DSDs, which are normalized in Fig. 3.13. It can be seen that there is no essential difference between the two approaches for modeling the natural DSD, though Testud argued

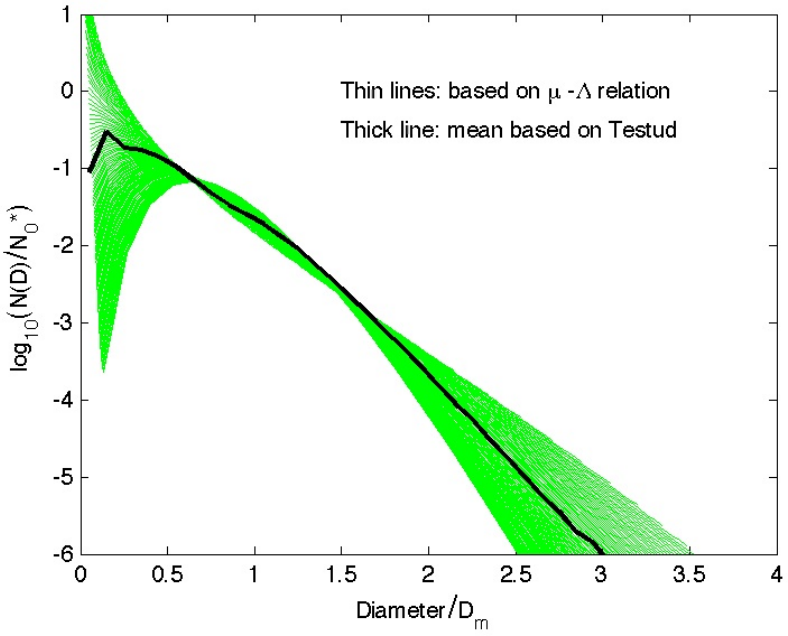

Figure 3.14: The simulation of normalized DSDs (thin lines) that are based on the $\mu - \Lambda$ relation with Λ varying from 0 to 10. The reference is the T-function (thick solid line).

that his approach was free of any assumed DSD forms. Both approaches lead to a DSD model with two parameters. The comparisons of retrievals using two approaches also demonstrate their equivalence [Cao and Zhang (2009)].

3.5 Summary and Conclusion

In this chapter, the model error and measurement error of DSD are analyzed. Measurement errors of disdrometer are quantified through a side-by-side disdrometer comparison. Error effects on DSD parameter estimation are investigated based on simulated and real disdrometer data. The $\mu - \Lambda$ relation (i.e., C-G DSD model) is examined and refined using a sorting and averaging method (SATP) for the sake of minimizing error effect.

Compared to other filtering methods applied to disdrometer observations [Lee and Zawadzki (2005); Sauvageot and Lacaux (1995)], SATP has the following advantages: (i) identifies (or isolates) similar DSDs; (ii) statistical errors of observed DSDs can be reduced while physical variations are preserved; (iii) is applicable to more than one rain event; and (iv) the performance of SATP improves as the size of the dataset increases. There are also limitations to the application of SATP: (i) two parameters are not accurate enough to characterize some extreme cases (e.g., non-gamma distributions); (ii) different combinations of two parameters may affect the retrieved DSD; (iii) averaging the DSDs may reduce the physical variation if the grid pixels are not small enough; and (iv) the dataset is small for extremely heavy rain which limits the utility of SATP. In spite of these limitations, SATP is a promising method for processing disdrometer data. SATP could be improved if more parameters (e.g. three parameters) are used to characterize DSDs.

The error analysis and DSD estimation can be concluded as follows:

- As for 2DVD measurements, sampling errors of large drops have been found to be substantial and dominated by statistical errors while sampling errors of very small drops are mainly attributed to system limitations. Middle DSD moments (e.g., the 3^{rd} moment), on the other hand, have a comparatively smaller sampling error than outer moments (e.g., the zero moment or the 6^{th} moment).

- DSD moments essentially represent rain physics though they contain errors. Integral parameters of interest can be estimated/retrieved with small bias and fractional error if corresponding moments or the closest moments are included in the estimator. For the purpose of retrieval, it is important in model development to include the moments that are close to observed variables.

- Middle moment estimators of gamma DSD parameters produce fewer errors than the lower and higher moment estimators. If the true DSD follows the gamma distribution, M234 has the best overall performance in estimating the evaluating variables (N_T, R, D_0, Z_H, and Z_{DR}). However, if the true DSD does not follow the gamma distribution, the performance of M234 will degrade, especially for estimating Z_H and Z_{DR}. In this case, the selection of a moment estimator depends on the integral parameters of interest.

- ML and LM estimators have less bias in estimating gamma parameters according to the study of Smith and Kliche. The performance, however, depends significantly on the error at the low end of DSD and the assumption of gamma distribution. If the measurement error at the low end is enlarged or the DSD deviates from the gamma distribution (model error), ML and LM estimators will degrade and perform worse than estimators such as M234, which directly use middle moments. Generally, ML and LM estimators perform similarly to low moment estimators such as M012.

- The overall differences among M234, M246, and M346 are not substantial for the five evaluated parameters. M246 is recommended as the one for practical use. Although there are differences existing among seven estimators mentioned in this study, the choice of estimators could not account for the discrepancies between radar and disdrometer observations. The discrepancies between radar and disdrometer observations are mainly due to instrumental difference.

- Measurement and model errors cause estimated DSD parameters to be biased. Moment estimators tend to overestimate the DSD shape and slope parameters μ and Λ. The overestimation depends on the moment errors and may lead to a spurious relation between μ and Λ if the moment errors are more than 20%. However, the observed moments are believed to have errors around 10%

or less and the overestimation of μ is mostly not more than 5. Consequently, the $\mu - \Lambda$ relations found in previous studies are not determined entirely by the error effect. Notably, integral physical parameters are not substantially biased. Therefore, it is speculated that the $\mu - \Lambda$ relations are related to microphysical processes.

- The $\mu - \Lambda$ relation of constrained-gamma DSD model is demonstrated to be equivalent to the mean function of normalized DSDs derived through Testud's approach. Both approaches construct a DSD model with two parameters. There is no essential difference between these two DSD models in estimating integral parameters. But the $\mu - \Lambda$ relation allows more flexibility of DSD shape and performs better in estimating N_T. Furthermore, the constrained-gamma DSD model, like the exponential DSD model, is more convenient to apply in practice.

Chapter 4

Characterization of Rain Microphysics and Direct DSD Retrieval

4.1 Introduction

Parameterization of rain microphysics is essentially important in the numerical weather prediction model for quantitative precipitation forecasts. The previous microphysical parameterizations were mostly model-based, i.e., assuming a specific DSD model. For example, Kessler (1969) introduced a single-moment scheme, which was based on the exponential DSD model with a fixed parameter N_0. The single-moment parameterization scheme has been widely used in forecast models [e.g., weather research and forecast (WRF) model] due to its simplicity and computational efficiency. The two-moment parameterization schemes, intrinsically assuming a two-parameter DSD, provide better representation of rain microphysics and have attracted a lot of research interests [e.g., Hong et al. (2004); Chen and Liu (2004); Zhang et al. (2006a)]. The two-moment schemes are appropriate for the application of polarimetric radar data and it is found that rain microphysics can be well characterized in terms of polarimetric radar variables, Z_H and Z_{DR}. In this chapter, the disdrometer observations are used to characterize the rain microphysics in Oklahoma. Different from previous studies that have assumed a specific DSD model, this study applies observed DSDs,

free of the model assumption. Seven rain variables, N_T, R, W, D_0, R_e, R_a, and V_{tm}, are characterized in terms of Z_H and Z_{DR}. For convenient applications, this study addresses the microphysical characterization for S-, C- and X-band, which are major frequencies for weather radars. It is worth noting that simulated radar variables, based on disdrometer observations and the scattering theory, might have discrepancies with the radar observations. The calibration of radar data by disdrometer observations might be required in practical applications.

The second issue of this chapter is the DSD retrieval from PRD. Generally, the retrieval applies a two-parameter DSD model and two radar variables Z_H and Z_{DR}, which normally have better data quality than other PRD. This chapter focuses on the retrieval of C-G DSD model that has been verified by several studies [Brandes et al. (2004a,b); Vivekanandan et al. (2004); Zhang et al. (2006a)]. The retrieval mainly follows the idea of Zhang et al. (2001) but applies the refined constraint $\mu - \Lambda$ relation described in Chapter 3. Possible adjustment of $\mu - \Lambda$ relation is also investigated. As we know, the C-G DSD model is developed based on disdrometer observations. Although it is applicable for most radar observations, it is worth noting the fact that there are discrepancies between radar and disdrometer observations due to factors such as rainfall inhomogeneity, sampling volume differences, limitations in radar measurements (e.g., contamination, sampling error, miscalibration), limitations in disdrometer measurements (e.g., undersampling, splashing, wind effects), and non-stationary rain processes (e.g., drop sorting, clustering, and evaporation). According to previous studies, if radar-measured reflectivity and differential reflectivity are close to the $Z_H - Z_{DR}$ relation derived from disdrometer measurements, retrieved rain variables (e.g., R, D_0, and N_T) generally agree with in-situ measurements. However, when radar measurements deviate greatly from disdrometer $Z_H - Z_{DR}$ relation, the result of DSD retrieval is not satisfactory. For example, if an DSD only has a

few large drops, the Z_{DR} would be large while the Z_H might be small. In this case, the DSD is narrow and has a higher μ value than the estimation from $\mu - \lambda$ relation. This kind of DSD can be observed at the leading edge of a convection and is identified as the "Big Drop" type. In such a case, the adjustment to the $\mu - \lambda$ relation is recommended in the retrieval to represent the real physical process better.

4.2 Characterization of Rain Microphysics

4.2.1 S-Band

This section gives the characterization of rain microphysics at S-band frequency. Brandes et al. (2004b) and Zhang et al. (2006a) characterized the rain microphysics by assuming a C-G DSD model. Errors due to the DSD model, however, may propagate through their procedure. The fitting directly with observations (without using the constraining $\mu - \Lambda$ relation) should reduce the error propagation. In order to further reduce the observation error, SATP method (described in Chapter 3) is applied to process DSDs observed by the disdrometer. Rain variables, N_T (m^{-3}), R (mm h^{-1}), W (g m^{-3}), and D_0 (mm), were calculated according to Chapter 2 for each data point. Radar variables, Z_H (in linear units of mm^6 m^{-3}) and Z_{DR} (dB) were calculated according to Eqs. 2.21 – 2.22 as well. Using a polynomial fit for all data points, rain variables were expressed in terms of radar variables as:

$$N_T = Z_H \times 10^{(-0.0837 Z_{DR}^3 + 0.702 Z_{DR}^2 - 2.062 Z_{DR} + 0.794)} \tag{4.1}$$

$$R = Z_H \times 10^{(-0.0363 Z_{DR}^3 + 0.316 Z_{DR}^2 - 1.178 Z_{DR} - 1.964)} \tag{4.2}$$

$$W = Z_H \times 10^{(-0.0493 Z_{DR}^3 + 0.430 Z_{DR}^2 - 1.524 Z_{DR} - 3.019)} \tag{4.3}$$

$$D_0 = 0.0436 Z_{DR}^3 - 0.216 Z_{DR}^2 + 1.076 Z_{DR} + 0.659 \tag{4.4}$$

In the similar way, microphysical processes can be estimated given a DSD. Evaporation rate (R_e, g m^{-3}s^{-1}), accretion rate (R_e, g m^{-3}s^{-1}), and mass-weighted terminal velocity (V_{tm}, m s^{-1}), were calculated. Using the polynomial least-square fit with the weight of rainwater content, mean curves (solid lines) for each parameter were derived in terms of W (g m^{-3}) and D_0 (mm) as:

$$R_e = W\big(0.0923D_0^{-3} - 0.309D_0^{-2} + 1.056D_0^{-1} - 0.0082) \times 10^{-3} \tag{4.5}$$

$$R_a = W\big(- 0.014D_0^3 + 0.211D_0^2 - 1.50D_0 + 7.04) \times 10^{-3} \tag{4.6}$$

$$V_{tm} = 0.0916D_0^3 - 1.088D_0^2 + 4.754D_0 + 0.525) \tag{4.7}$$

As Fig. 4.1 shows, Eqs. 4.2–4.4 represent rain variables well for Z_{DR} ranging from 0.15 to 4 dB while Eqs. 4.5–4.7 are good empirical relations for R_e, R_a and V_{tm} with D_0 less than 4 mm. To evaluate these empirical relations, results of these relations are compared with observations. DSDs with a Z_{DR} from 0.15 to 4 dB were used to validate Eqs. 4.1–4.4 while DSDs with a D_0 from 0.5 to 4 mm were used for Eqs. 4.5–4.7. The bias and correlation coefficient values between empirical values and observed values are calculated and listed in Table4.1. Scatter diagrams of these results reveal that Eqs. 4.6 and 4.7 for estimates of R_a (proportional to the 2.67^{th} moment of DSD) and V_{tm} have a very small bias ($< 1\%$). Eqs. 4.2–4.4 for estimates of R, W, and D_0 have a bias less than 10%, and Eqs. 4.2 and 4.5 have a bias larger than 10%. This implies that empirical relations for low moments N_T (zero moment) and R_e ($\sim 1.6^{th}$ moment) are not as good as relations for high moments. This is probably because the control variables (Z_H, Z_{DR}, W and D_0) are mainly determined by meddle size or large drops and they are not good at representing low moments. Middle moments (e.g., R_a and V_{tm}), on the other hand, are well represented. It is worth noting that $D_0 - Z_{DR}$ plot (Fig. 4.1c) has a large scatter. This is likely due to the fact that the main contributor to D_0 is the median size drop while Z_{DR} is mainly determined by the large drop. This is also the reason why the D_0 estimated from Z_{DR}, when compared to the observed D_0, only has a correlation coefficient about 0.8.

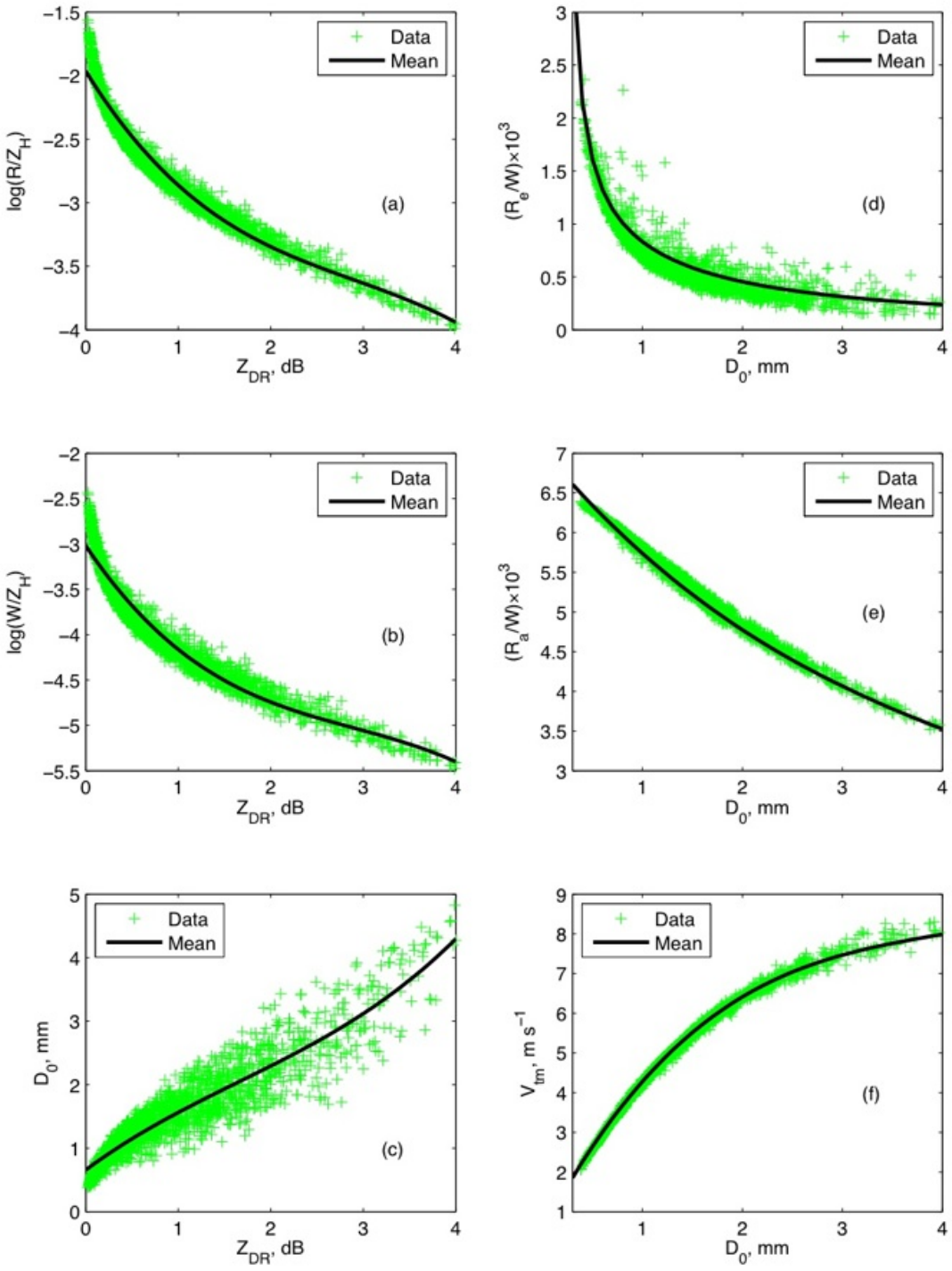

Figure 4.1: Relations between rain and radar variables. Crosses denote data points and solid lines are mean curves by polynomial fitting. (a) Ratio of R to Z_H versus Z_{DR}. (b) Ratio of W to Z_H versus Z_{DR}. (c) D_0 versus Z_{DR}. (d) Ratio of R_e to W versus D_0. (e) Ratio of R_a to W versus D_0. (f) V_{tm} versus D_0.

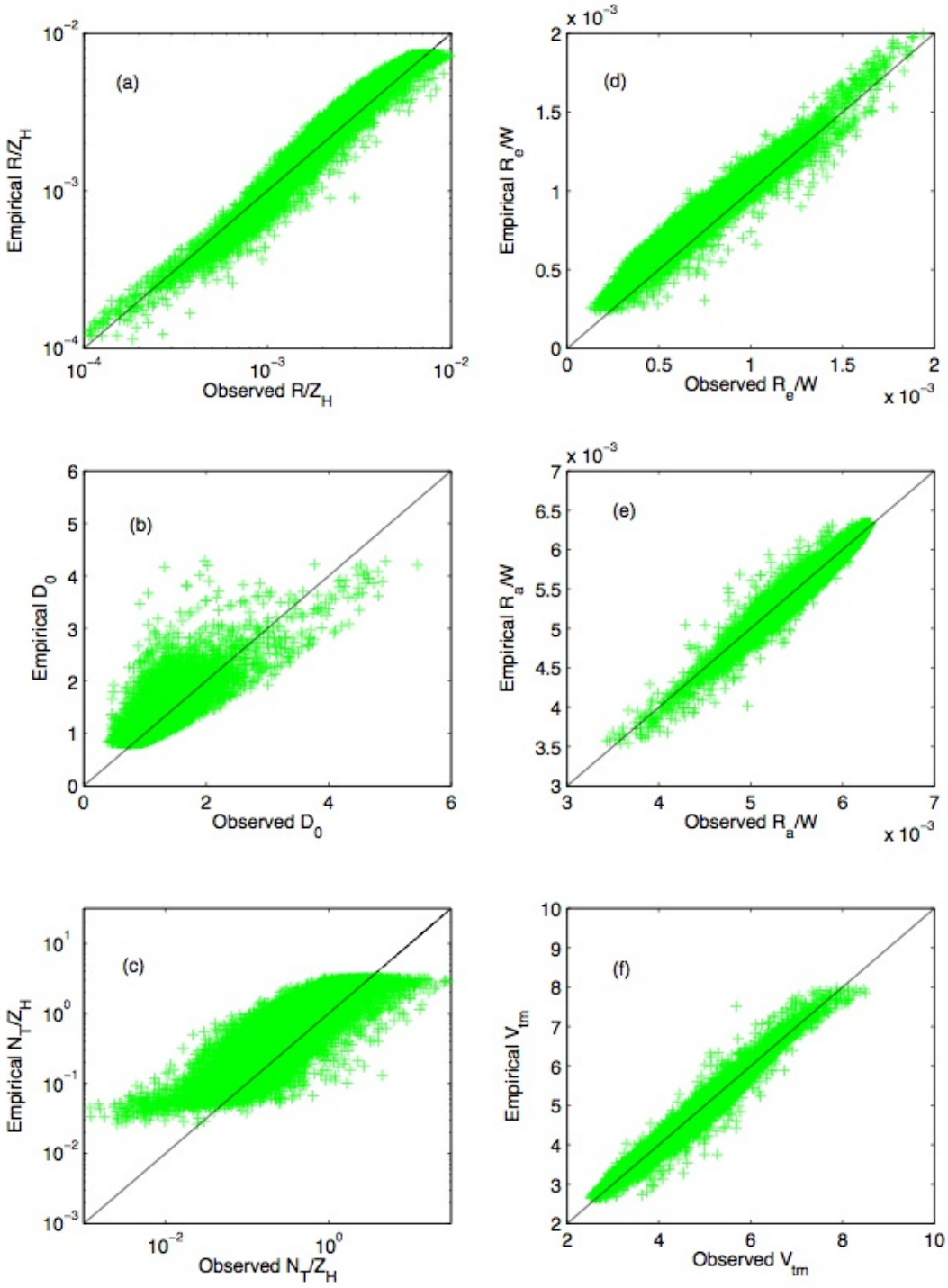

Figure 4.2: Scatter diagram of empirical values versus observations. Crosses denote data points and solid lines correspond to the unit slope. Empirical relations are Eqs. 4.1–4.7. (a) Ratio of R to Z_H. (b) D_0. (c) Ratio of N_T to Z_H. (d) Ratio of R_e to W. (e) Ratio of R_a to W. (f) V_{tm}.

Table 4.1: Bias and correlation coefficient for rain variables using Eqs. 4.1–4.7 versus observations

	N_T	R	W	D_0	R_e	R_a	V_{tm}
bias (%)	20.24	9.09	9.63	7.18	13.80	-0.52	-0.026
corr. coef.	0.63	0.97	0.94	0.81	0.97	0.98	0.98

4.2.2 C-Band and X-Band

The frequency difference only affects relations that include radar variables in the previous section. Radar variables have distinct properties between S-band and C-, X-band. Fig. 4.3 gives the simulated Z_H and Z_{DR} versus the raindrop size. Differences of Z_H and Z_{DR} among three bands are very small for raindrops $D < 2$mm. X-band (C-band) Z_H becomes distinctly larger than S-band Z_H from $\sim$3 mm ($\sim$4.5 mm) and the maximum difference is $\sim$5 dB ($\sim$7 dB). The difference is particularly evident for Z_{DR}. C-band Z_{DR} has a sharp increase starting from $\sim$3 mm and reaches the peak at $\sim$6 mm. The maximum Z_{DR} difference can be $\sim$2.5 dB, given the same raindrop size. Beyond 6 mm, C-band Z_{DR} decreases and its value becomes even lower than S-band Z_{DR}. X-band Z_{DR} has a similar trend to C-band Z_{DR} but its range of variation is much less than C-band Z_{DR}.

Compared to S-band, the attenuation associated with the precipitation is much larger at C, X-band. Fig. 4.4 shows the simulated attenuation factors versus the raindrop size with the same assumption of concentration as in Fig. 4.3. As figure shows, even for small raindrops ($D < 1$ mm) C- or X-band attenuation is distinctly larger than S-band. X-band has the largest attenuation. The maximum A_H difference is seen at $\sim$3.5 mm, where X-band A_H is larger than S-band A_H with an order of

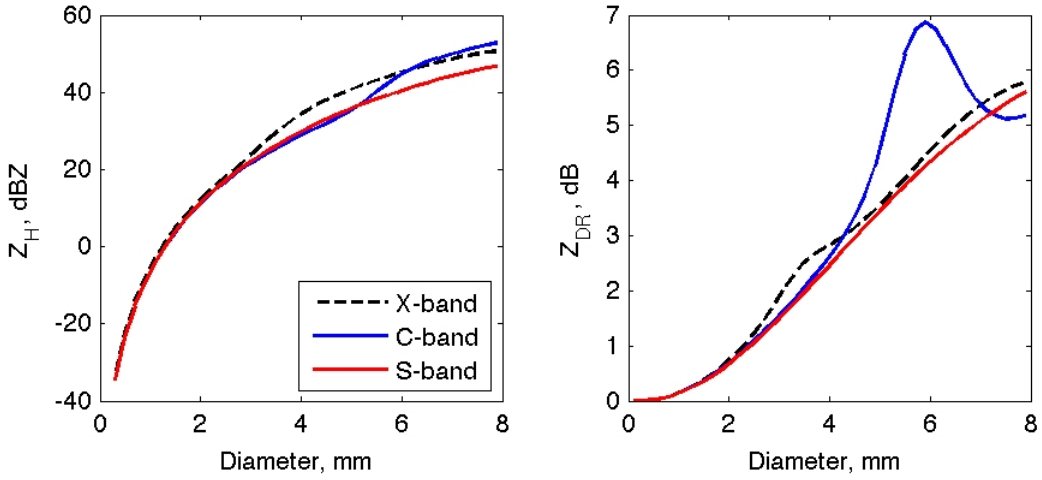

Figure 4.3: Simulation of (a) Z_H and (b) Z_{DR} versus equivalent diameter of raindrop. Given that raindrops have the same size and the same concentration 1 m^{-3} within a bin interval 0.2 mm.

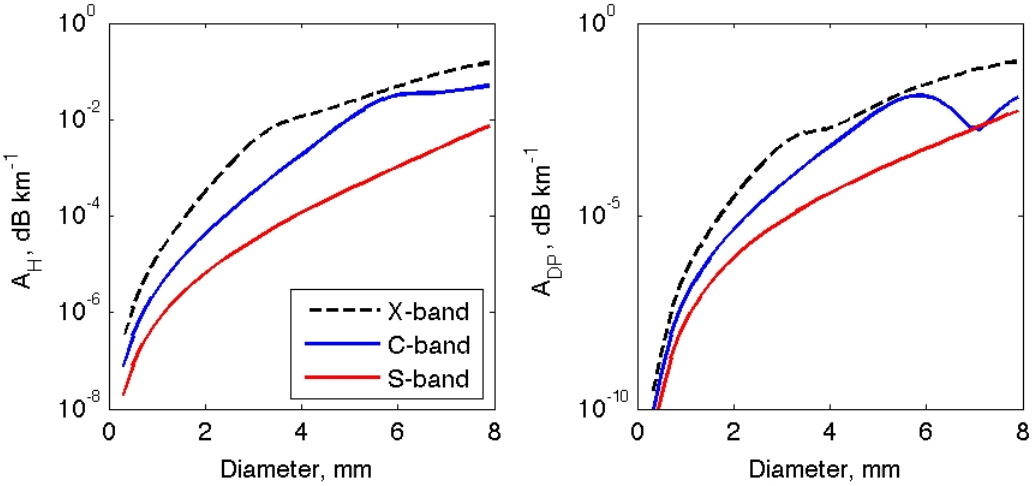

Figure 4.4: Same as Fig. 4.3 but for simulation of (a) A_H and (b) A_{DP}.

magnitude of ∼2. Generally, C- and X-bands A_H values are larger than S-band with a factor of more than 10. The similar trend is seen for A_{DP}. It is worth noting that attenuation factors of C- and X-band have some undulations for middle and/or large drops. The undulation is attributed to the fact that raindrop sizes are comparable with the wavelength and the scattering of raindrops does not follow the linear relation estimated by Rayleigh theory.

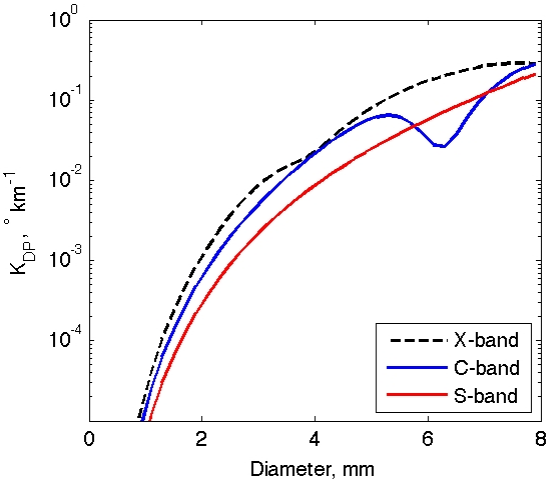

Figure 4.5: Same as Fig. 4.3 but for simulation of K_{DP}.

The phase parameters, similar to attenuation parameters, are also associated with the forward scattering. In practice, phase terms can be used to estimate the attenuation. Fig. 4.5 shows the simulation of K_{DP}. K_{DP} shows a similar variation trend to A_H and A_{DP}, implying that K_{DP} could be used to estimate these two parameters. The empirical relations associated with the attenuation and phase terms are given later in this section.

In the following, characterized relations of rain and radar variables are updated for C- and X-band. The development procedure is as same as the one used for the S-band. Fig. 4.6 presents the result of C-band parameterization. The wavelength is

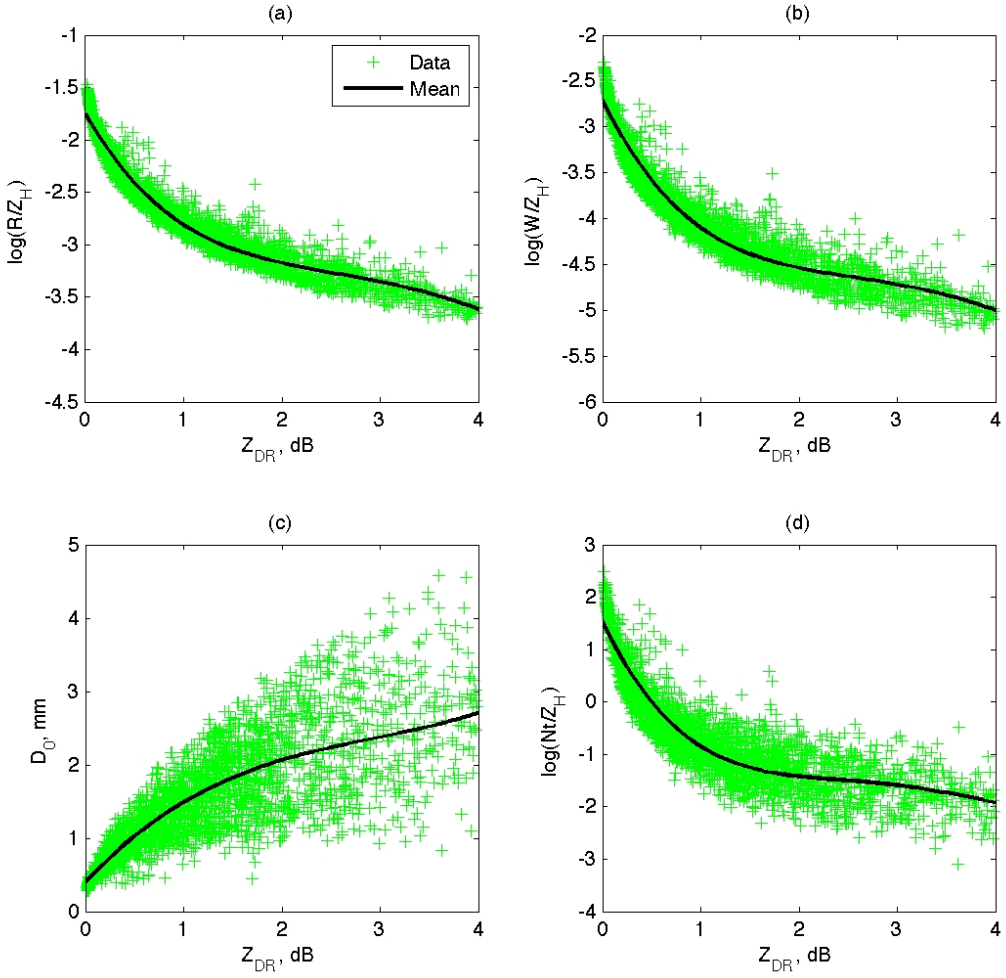

Figure 4.6: C-band relations between rain and radar variables. (a) Ratio of R to Z_H versus Z_{DR}. (b) Ratio of W to Z_H versus Z_{DR}. (c) D_0 versus Z_{DR}. (d) Ratio of N_T to Z_H versus Z_{DR}.

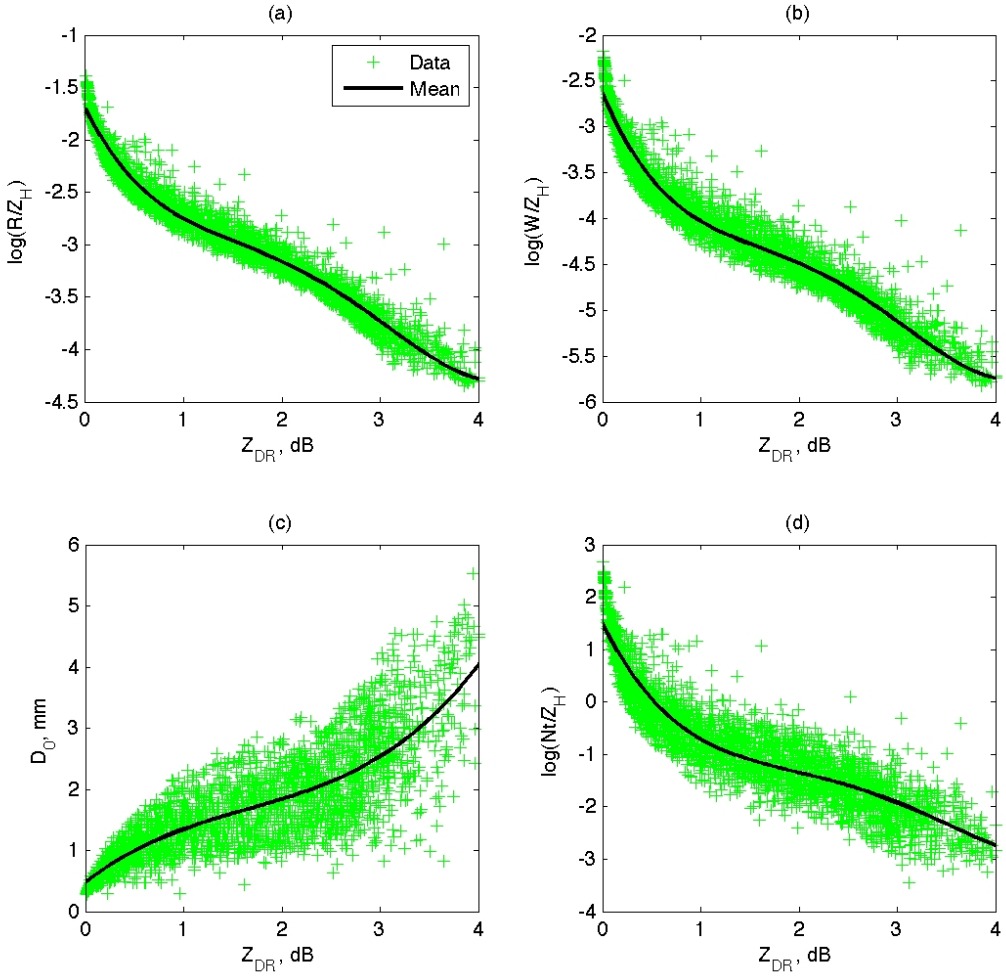

Figure 4.7: X-band relations between rain and radar variables. (a) Ratio of R to Z_H versus Z_{DR}. (b) Ratio of W to Z_H versus Z_{DR}. (c) D_0 versus Z_{DR}. (d) Ratio of N_T to Z_H versus Z_{DR}.

specified as 5.4 cm, which is applied by OU-PRIME radar. In the figure, the crosses denote calculated parameters from 2DVD data, which have been processed by SATP method. Solid lines are mean relations which are fitted to polynomial functions using data points. The mean relations are given by:

C-band (5.4 cm) :

$$N_T = Z_H \times 10^{(0.0355Z_{DR}^4 - 0.450Z_{DR}^3 + 2.012Z_{DR}^2 - 3.990Z_{DR} + 1.541)} \quad (4.8)$$

$$R = Z_H \times 10^{(0.0101Z_{DR}^4 - 0.146Z_{DR}^3 + 0.713Z_{DR}^2 - 1.635Z_{DR} - 1.754)} \quad (4.9)$$

$$W = Z_H \times 10^{(0.014Z_{DR}^4 - 0.20Z_{DR}^3 + 0.980Z_{DR}^2 - 2.186Z_{DR} - 2.711)} \quad (4.10)$$

$$D_0 = 0.0438Z_{DR}^3 - 0.390Z_{DR}^2 + 1.436Z_{DR} + 0.402 \quad (4.11)$$

Similar to Fig. 4.6, Fig. 4.7 shows the mean relations derived for X-band. The wavelength is specified as 3.2 cm, which is used by CASA IP1 radars. The relations are given by:

X-band (3.2 cm) :

$$N_T = Z_H \times 10^{(0.049Z_{DR}^4 - 0.551Z_{DR}^3 + 2.091Z_{DR}^2 - 3.803Z_{DR} + 1.50)} \quad (4.12)$$

$$R = Z_H \times 10^{(0.041Z_{DR}^4 - 0.381Z_{DR}^3 + 1.177Z_{DR}^2 - 1.890Z_{DR} - 1.70)} \quad (4.13)$$

$$W = Z_H \times 10^{(0.055Z_{DR}^4 - 0.517Z_{DR}^3 + 1.642Z_{DR}^2 - 2.585Z_{DR} - 2.628)} \quad (4.14)$$

$$D_0 = 0.0984Z_{DR}^3 - 0.488Z_{DR}^2 + 1.265Z_{DR} + 0.470 \quad (4.15)$$

According to Figs. 4.6 and 4.7, relations 4.8–4.15 are recommended for use with Z_{DR} from 0.1 to 4 dB. As mentioned previously, these relations represent the mean properties. There exist uncertainties attributed to the variation of DSD. For example, parameter Z_H is appropriate the 6^{th} order moment of DSD. R is $\sim 3.67^{th}$ order moment and W and N_T are lower moments. The uncertainty of $\log(R/Z_H)$ is less than uncertainties of $\log(W/Z_H)$ and $\log(N_T/Z_H)$. Data points of D_0 look much more scattering than other three parameters, especially for large Z_{DR}. This fact implies that the dependence of D_0 on Z_{DR} is more sensitive to the DSD shape than the ratio

of moment parameters. C-band relations have more undulations than X-band. It is likely attributed to the fluctuation of C-band Z_{DR} shown in Fig. 4.3b.

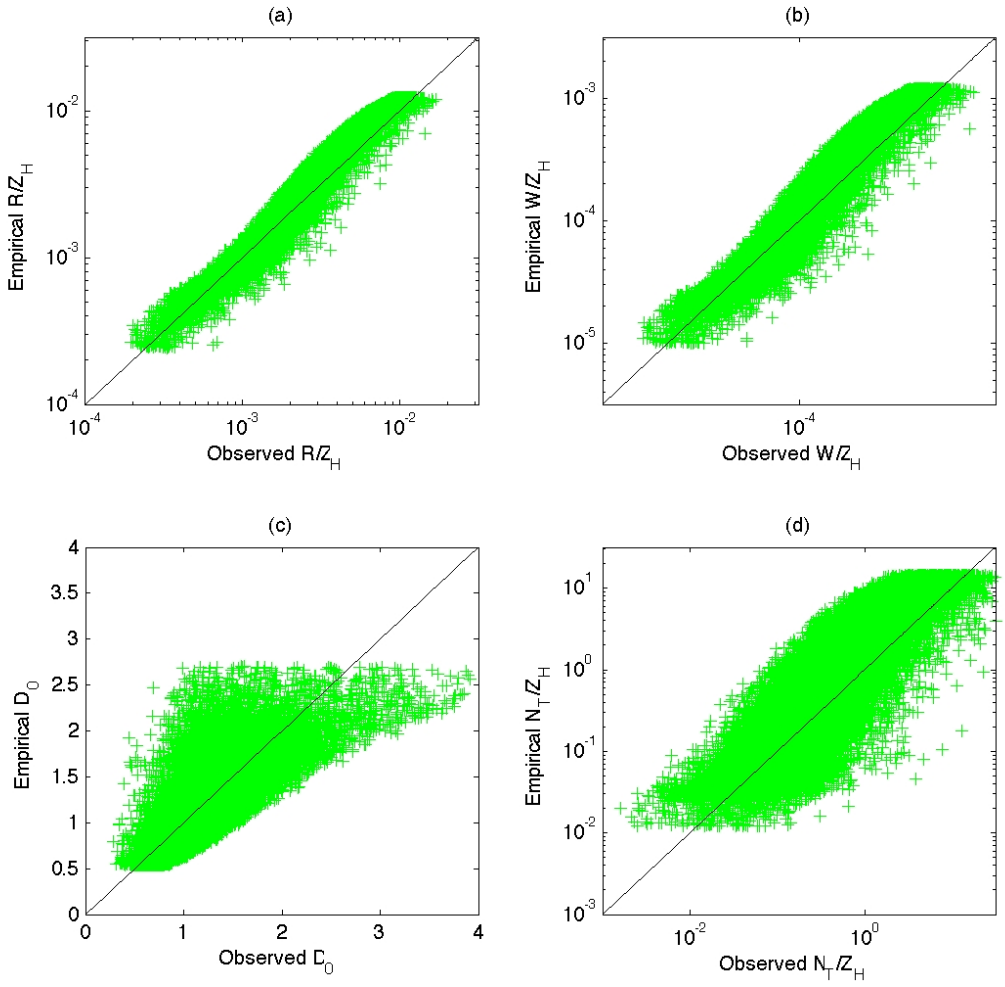

Figure 4.8: Oneone plots of results from C-band relations versus observations. (a) Ratio of R to Z_H. (b) Ratio of W to Z_H. (c) D_0. (d) Ratio of N_T to Z_H.

Performances of these empirical relations are demonstrated in Figs. 4.8 and 4.9. The horizontal axis indicates the parameters calculated from 2DVD data. The vertical axis denotes the results of empirical relations. The bias and root-mean-square error (RMSE) of empirical estimations are shown in Table 4.2. For both bands, parameters $\log(R/Z_H)$ and $\log(W/Z_H)$ have a small bias and RMSE, generally less than

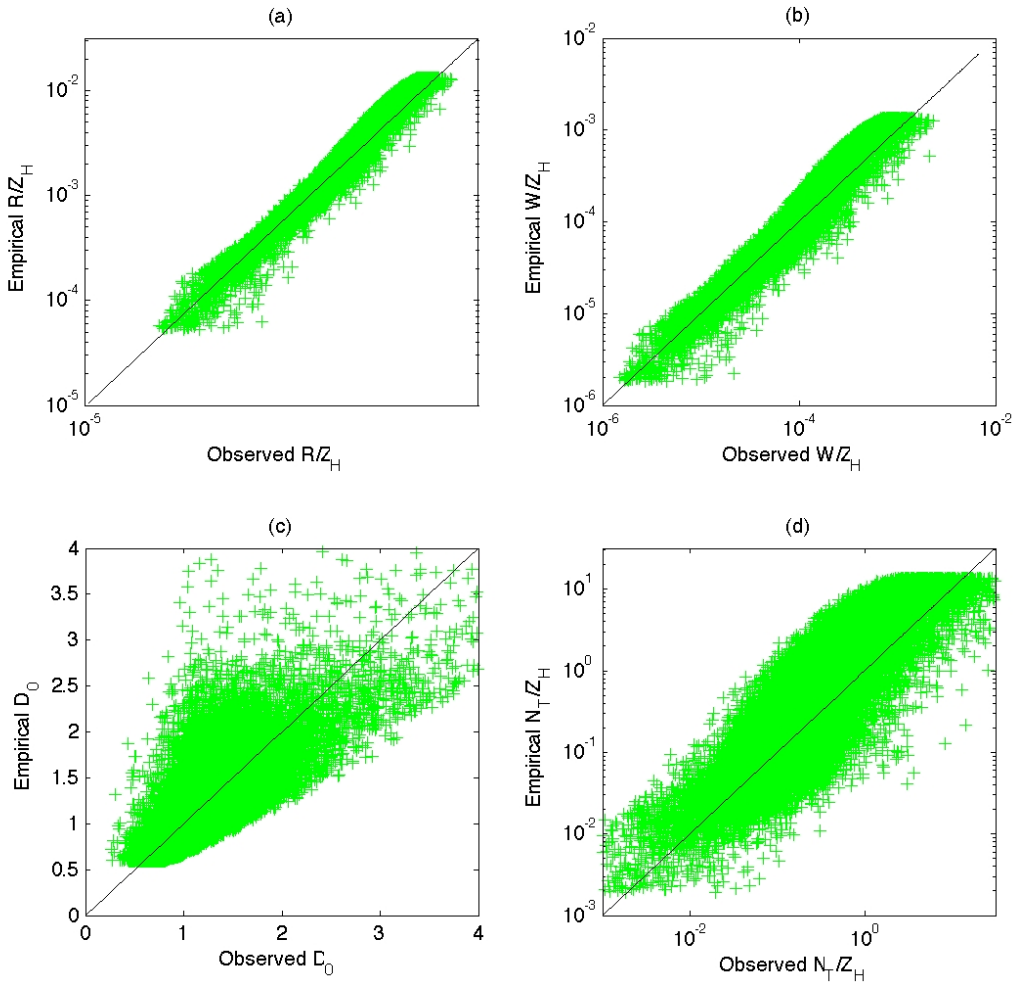

Figure 4.9: Oneone plots of results from X-band relations versus observations. (a) Ratio of R to Z_H. (b) Ratio of W to Z_H. (c) D_0. (d) Ratio of N_T to Z_H.

5%. Parameter $\log(N_T/Z_H)$ has a large bias and RMSE. This fact illustrates that radar variables Z_H and Z_{DR} can not represent low DSD moments well because they are parameters associated with high DSD moments. D_0 has a relatively small bias but has a relatively large RMSE ($\sim 25\%$) for empirical estimation.

Attenuation and phase terms are interest parameters for C- and X-band applications. According to the scattering theory, attenuation and phase parameters are

Table 4.2: Bias and RMSE of estimations by empirical relations for C- and X-band

		$\log(R/Z_H)$	$\log(W/Z_H)$	D_0	$\log(N_T/Z_H)$
X-band	bias (%)	-2.80	-2.77	-4.59	-82.98
	RMSE (%)	4.35	4.57	25.57	111.0
C-band	bias (%)	-2.74	-2.72	-3.77	-87.44
	RMSE (%)	4.28	4.48	24.06	118.5

calculated for all 2DVD dataset. The results are shown in Figs. 4.10 and 4.11 for C- and X-band. The crosses denote data points. Figures show that there are correlations existing among attenuation and phase parameters. Correlation coefficients among those parameters are listed in Table 4.3. Generally, X-band has a stronger correlation between attenuation and phase terms than C-band while its correlation between A_H and Z_H is a little lower. The correlation between A_H and K_{DP} is the strongest one given the same frequency. In addition, parameters A_{DP} and K_{DP} are more sensitive to the DSD shape than A_H and Z_H. Therefore, variations of A_{DP}–K_{DP} pair and A_{DP}–A_H pair are larger with a weaker correlation between two parameters of a pair. Among them, C-band A_{DP} and K_{DP} have the weakest correlation.

Table 4.3: Correlation coefficient among attenuation and phase parameters.

	A_H–K_{DP}	A_{DP}–K_{DP}	A_{DP}–A_H	A_H–Z_H
X-band	0.998	0.968	0.968	0.910
C-band	0.961	0.844	0.941	0.943

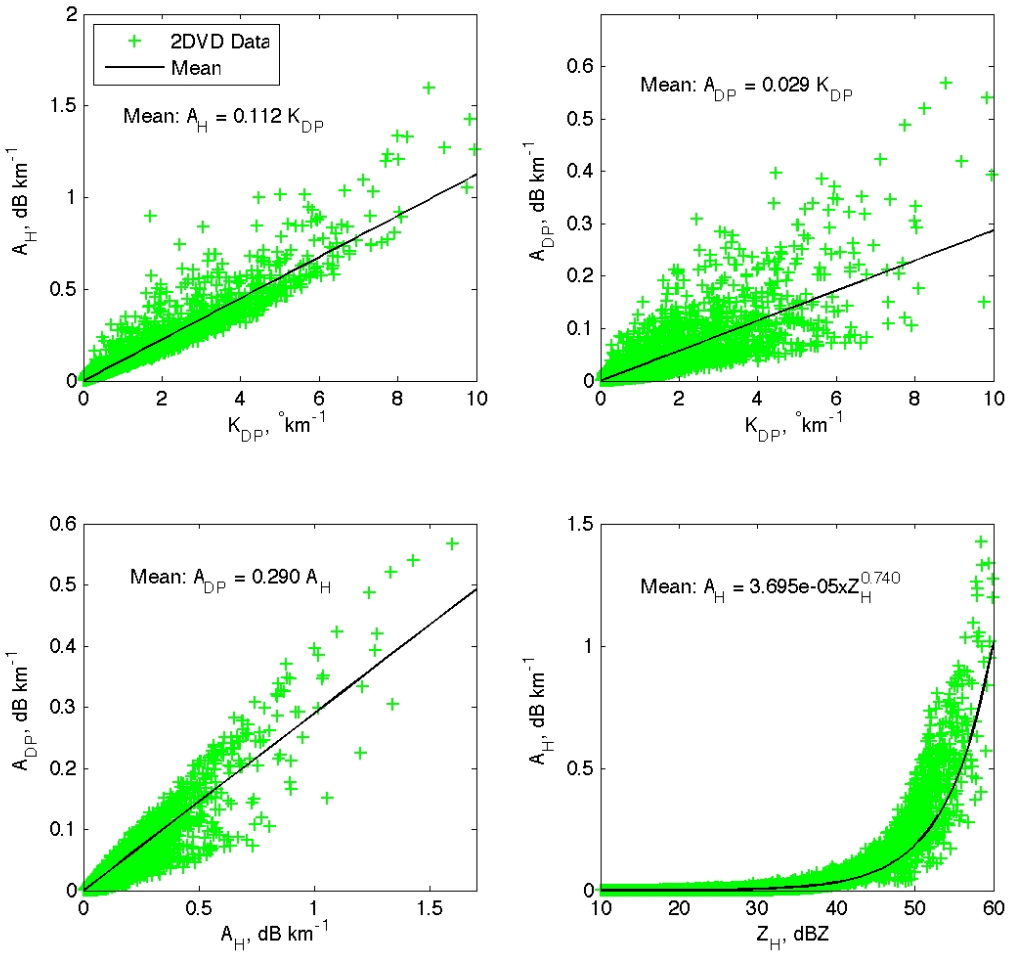

Figure 4.10: C-band parameterization of attenuation and phase terms. (a) A_H versus K_{DP}, (b) A_{DP} versus K_{DP}, (c) A_{DP} versus A_H, (d) A_H versus Z_H.

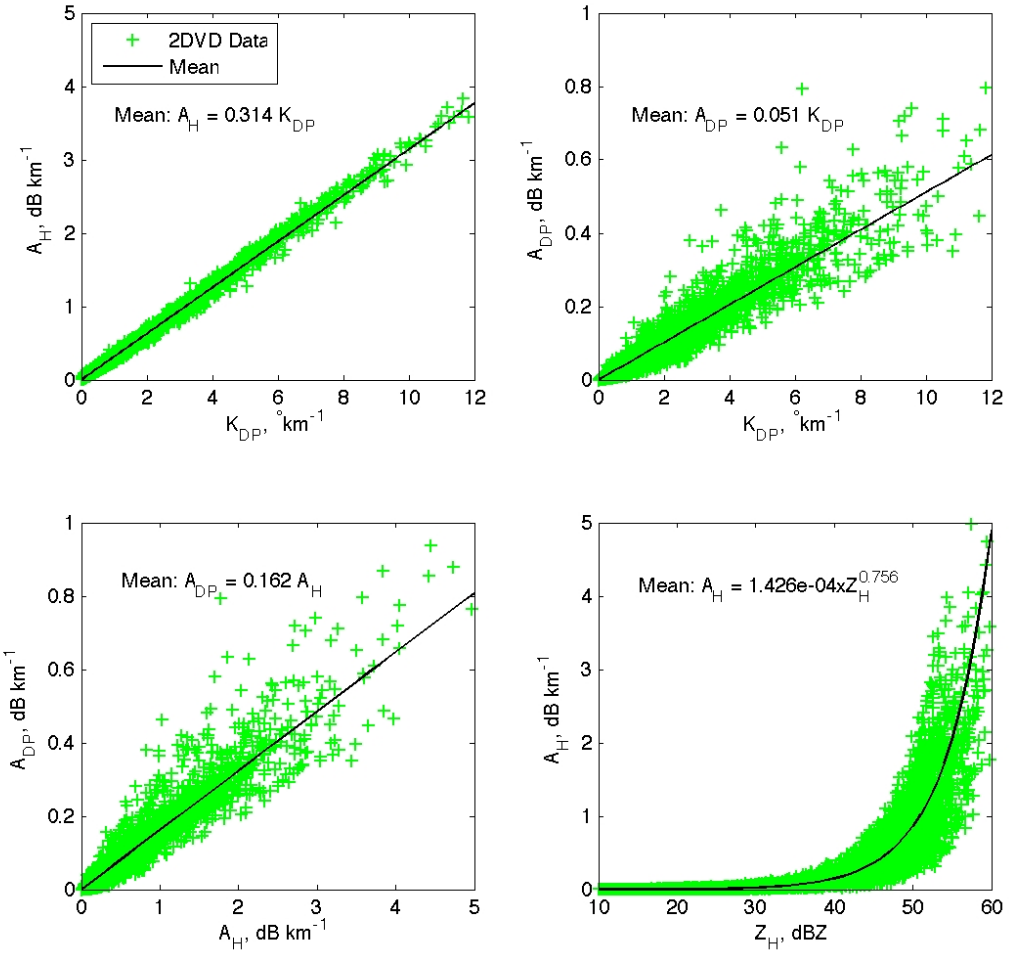

Figure 4.11: Same as Fig. 4.10 but for X-band parameterization.

It is practically useful to develop empirical relations among attenuation and phase parameter. As shown in Figs. 4.10 and 4.11, solid lines are mean relations derived by the power law fitting. These mean relations are given by:

C-band (5.4 cm):
$$A_H = 0.112 K_{DP} \tag{4.16}$$
$$A_{DP} = 0.029 K_{DP} \tag{4.17}$$
$$A_{DP} = 0.290 A_H \tag{4.18}$$
$$A_H = 3.695 \times 10^{-5} Z_H^{0.740} \tag{4.19}$$

X-band (3.2 cm) :
$$A_H = 0.314 K_{DP} \tag{4.20}$$
$$A_{DP} = 0.051 K_{DP} \tag{4.21}$$
$$A_{DP} = 0.162 A_H \tag{4.22}$$
$$A_H = 1.426 \times 10^{-4} Z_H^{0.756} \tag{4.23}$$

Using these relations, the attenuation can be corrected with measurements of differential propagation phase. Practically, it is a simple but efficient way to correct precipitation attenuation given the conditions (i) there is not existing any other precipitation phases than rain, and (ii) phase measurements have a good quality (e.g. with good SNR). Sophisticated attenuation correction methods will be discussed in Chapter 6.

4.3 DSD Retrieval from PRD

With empirical relations mentioned in previous two sections, rain variables of interest can be calculated. It is convenient to use these relations as physical models for some practical applications. However, these relations can only be applied with a fixed frequency. If frequency changes, all the relations have to be updated. Moreover, these empirical relations only represent a mean characteristics, for which they sacrifice some physical variations. These issues are not the obstacle for DSD retrieval. On

one hand, physical variations can be well represented given a good DSD model. The DSD retrieval may improve the estimation of rain variables from radar observations. On the other hand, radar variables can be calculated given a DSD and the scattering amplitudes at different frequencies. It is more flexible in different applications by retrieving the DSD. This section presents the DSD retrieval introduced by Zhang et al. (2001) and discusses some issues in practice.

4.3.1 Retrieval based on a C-G DSD model

Following the studies of Zhang et al. (2001) and Brandes et al. (2004b), the DSD retrieval is described as follows. There are two major assumptions. One is that the DSD satisfies the C-G DSD model with a constraining $\mu - \Lambda$ relation described by Eq. 3.16. The other assumption is that the raindrop axis ratio relation is invariant within the radar sampling volume and is given by Eq. 2.26 [Brandes et al. (2002)].

Given the constraining relation and fixed axis ratio relation, the DSD shape parameter μ is uniquely determined by the observed Z_{DR}, according to Eqs. 2.21–2.22. With μ retrieved from Z_{DR}, the other DSD parameter N_0 is then calculated from Z_H. This method of retrieving DSD applies Z_H and Z_{DR} values directly from the observations, without considering the measurement error effect. The measurement error might propagate into the retrieval result. The error issue of DSD retrieval will be discussed in the next chapter, which introduces the Bayesian approach.

4.3.2 Adjustment of $\mu - \Lambda$ relation

The refined $\mu - \Lambda$ relation (Eq. 3.16) enables the retrieval of the gamma distribution parameters (N_0, μ, and Λ) from measurements of Z_H and Z_{DR}. A Z_H–Z_{DR} scatter diagram based on disdrometer observations is presented in Fig. 4.12. Solid and dashed lines denote polynomial fits for Oklahoma and Florida (Zhang et al. 2006),

respectively. There is little difference between these two curves for $Z_H < 30$ dBZ. The mean Z_H–Z_{DR} relation for Oklahoma is given by equation

$$Z_{DR} = 10^{(-2.6857 \times 10^{-4} Z_H^2 + 0.04892 Z_H - 1.4287)} \tag{4.24}$$

where both Z_H and Z_{DR} are expressed on a logarithmic scale. This $Z_H - Z_{DR}$ relation 4.24 is consistent with the $\mu - \Lambda$ relation 3.16 for the fact that both were derived from the same disdrometer dataset.

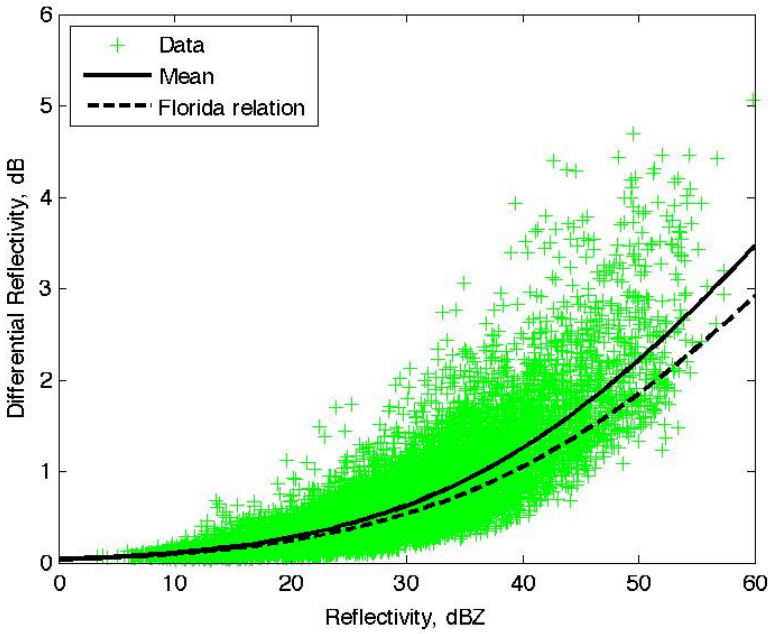

Figure 4.12: Plot of Z_{DR} versus Z_H from 2DVD measurements in Oklahoma. Cross points denote 2DVD measurements. The solid line is the mean curve, which is fitted to all data pointes in logarithmic domain by a two-order polynomial fitting. The dashed line is the Florida relation (Zhang et al. 2001)

Previous studies [e.g., Schuur et al. (2001); Brandes et al. (2003, 2004a); Vivekanandan et al. (2004); Zhang et al. (2006a)] have shown that disdrometer observations are

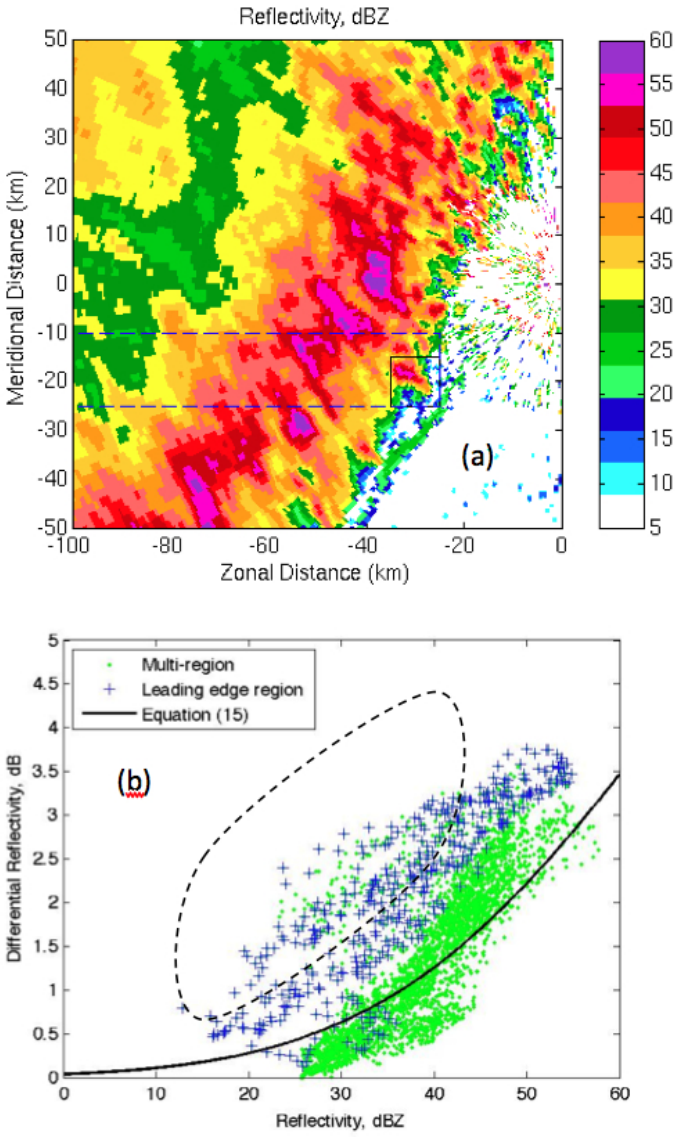

Figure 4.13: (a) PPI of KOUN radar-measured reflectivity (065955 UTC, 13 May 2005). A solid square isolate a strong convective storm at the leading edge of this rain event. The dashed rectangular region is a multiple precipitation-type region that includes portions of the convective leading edge, thunderstorm core, and trailing stratiform precipitation. (b) Plot of Z_{DR} versus Z_H. Dots denote measurements within a multiple precipitation-type region, i.e., dashed-line region of subplot (a). Asterisks denote measurements within leading edge region, i.e., solid-line square of subplot (a). The solid line is the mean curve of disdrometer observations (Eq. 4.24). The region within the dashed-line includes BD cases.

generally consistent with radar observations and that DSD models derived from disdrometer observations generally work well when applied to radar retrieval. However, the sampling volume of a radar is much larger than that of a disdrometer. The KOUN, for example, has a sampling volume of ~ 0.07 km^3 at a 30 km range. Consequently, its sampling volume can be 10^7 or more than that of a 2DVD. The difference between radar measurements and 2DVD measurements might be large, especially for inhomogeneous rains (e.g. at the leading edge of convection). The radar retrieval may not work well if Z_H and Z_{DR} measurements depart significantly from the disdrometer-based mean relation. Fig. 4.13a shows a PPI image of radar reflectivity measured by KOUN on May 13, 2005. A solid square isolates a strong convective cell at the leading edge of squall line. The dashed region includes portions of the leading and trailing convective line. The scatter diagram of Z_H and Z_{DR} within these two regions are plotted in Fig. 4.13b. The disdrometer-based mean $Z_H - Z_{DR}$ relation 4.24 is plotted for reference. Most $Z_H - Z_{DR}$ pairs from the rectangular box cluster well around the line described by Eq. 4.24 except for measurements corresponding to the isolated convective cell, where relatively high Z_{DR} are associated with relatively low Z_H. According to the hydrometeor classification algorithm described by Ryzhkov et al. (2005b), these points are identified as rain dominated by big drops (BD). In the BD region, the DSD tends to be narrower and the total concentration of drops tends to be much lower than in stratiform rain with the same intensity. Retrievals based on relation 3.16 may result in errors in the BD region. For example, such a retrieval would result in a broadly estimated DSD and an unreasonably large total number concentration. To solve this problem, the disdrometer-based relation is adjusted to

$$\mu = \mu'(\Lambda) + C\Delta Z_{DR} \tag{4.25}$$

where the $\mu'(\Lambda)$ is the disdrometer-based relation 3.16 derived in the previous chapter, ΔZ_{DR} is the difference between the radar-measured Z_{DR} and the Z_{DR} estimated from measured Z_H according to Eq. 4.24, and C is an adjustment parameter. It is

reasonable to consider the adjustment term $C\Delta Z_{DR}$ to be related to Z_H and Z_{DR} because Z_H and Z_{DR} give useful information about rain type, intensity, and DSD. Intuitively, the adjustment parameter C should be positive and dependent on Z_H and Z_{DR}. For a given ΔZ_{DR}, it is desired that C increase when Z_{DR} increases or Z_H decreases. It is hard, however, to determine a good adjustment term that fully represents the variability of rain physics. This study focuses on the adjustment of BD region, where the ΔZ_{DR} has a maximum dynamic range of 3. We expect the adjusted μ to fall into the normal dynamic range of 6. Therefore, the value of C was chosen to be 2 for this study. No adjustment is made if $\Delta Z_{DR} < 0.5$ dB. The adjustable $\mu - \Lambda$ relation could improve the retrieval of N_T and D_0 at the leading edge of convective squall line, which is often characterized as the BD region. On the other hand, the effect of adjustment is minor outside of BD regions.

4.3.3 Radar retrieval

The radar retrieval was applied to the rain event illustrated in Fig. 4.13a. The retrieval was performed over the whole storm area (including the BD region) using radar-measured Z_H and Z_{DR} as well as the refined $\mu - \Lambda$ relation with an adjustment. The retrieval without the adjustment was also performed as a comparison. The retrieving procedure is similar to that described in previous studies (e.g., Zhang et al. 2001, Brandes et al. 2004) except for the numerical method used to solve the nonlinear equations and the procedure to estimate the maximum diameter. The regression method used here is the two-dimensional Newton-Raphson method [Press et al. (2001)]. The maximum diameter in previous studies was estimated from an empirical relation fitted to disdrometer observations (e.g., Brandes et al. 2004), which remains to be an issue in radar retrieval. When the $\mu - \Lambda$ relation is used with the adjustment, the impact of maximum diameter is minor. As a result, this study sets

the maximum diameter to a constant of 8 mm, which works for most rain events ($Z_H < 60$ dBZ).

The PPI images of radar measurements and retrieval results are shown in Fig. 4.14. The fields of Z_H, Z_{DR}, and results of hydrometer classification are shown in column (a). Z_H and Z_{DR} measurements classified as NR or RH are filtered out before the rain retrievals are performed. The column (b) contains retrieval results based on the $\mu - \Lambda$ relation without the adjustment, and the column (c) represents retrieval results based on the $\mu - \Lambda$ relation with the adjustment. These two approaches give similar and reasonable retrievals for most duration of the storm. It is noted that retrieved rain-rates through two approaches have little difference. In general, the $\mu - \Lambda$ adjustment mainly reduces the number concentration of small drops while causing less change for median and large drops. Consequently, lower moments are affected more than higher moments. The retrieved rain-rate is less affected by the adjustment. In BD regions with low Z_H, however, retrievals without $\mu - \Lambda$ adjustment produce a much higher N_T and a smaller D_0. They suggest that a large number of small drops exist in the area where a small number of big drops should be. Thus, retrievals with the adjustment give more reasonable results for the developing convective cells.

The refined $\mu - \Lambda$ relation 3.16 has already been verified by disdrometer data in Chapter 3. Next, comparisons between radar retrievals and disdrometer observations are given to demonstrate its validity for radar data. The rain event analyzed was a precipitation system that passed through central Oklahoma on 2 May 2005 when the NCAR 2DVD was deployed 28 km south of KOUN. Fig. 4.15 shows time series Z_H and Z_{DR} from 1100 UTC to 1330 UTC. The asterisks denote radar measurements. To reduce error, radar measurements have been averaged over five range gates. To eliminate a systematic bias between radar and disdrometer, radar-measured Z_{DR} was

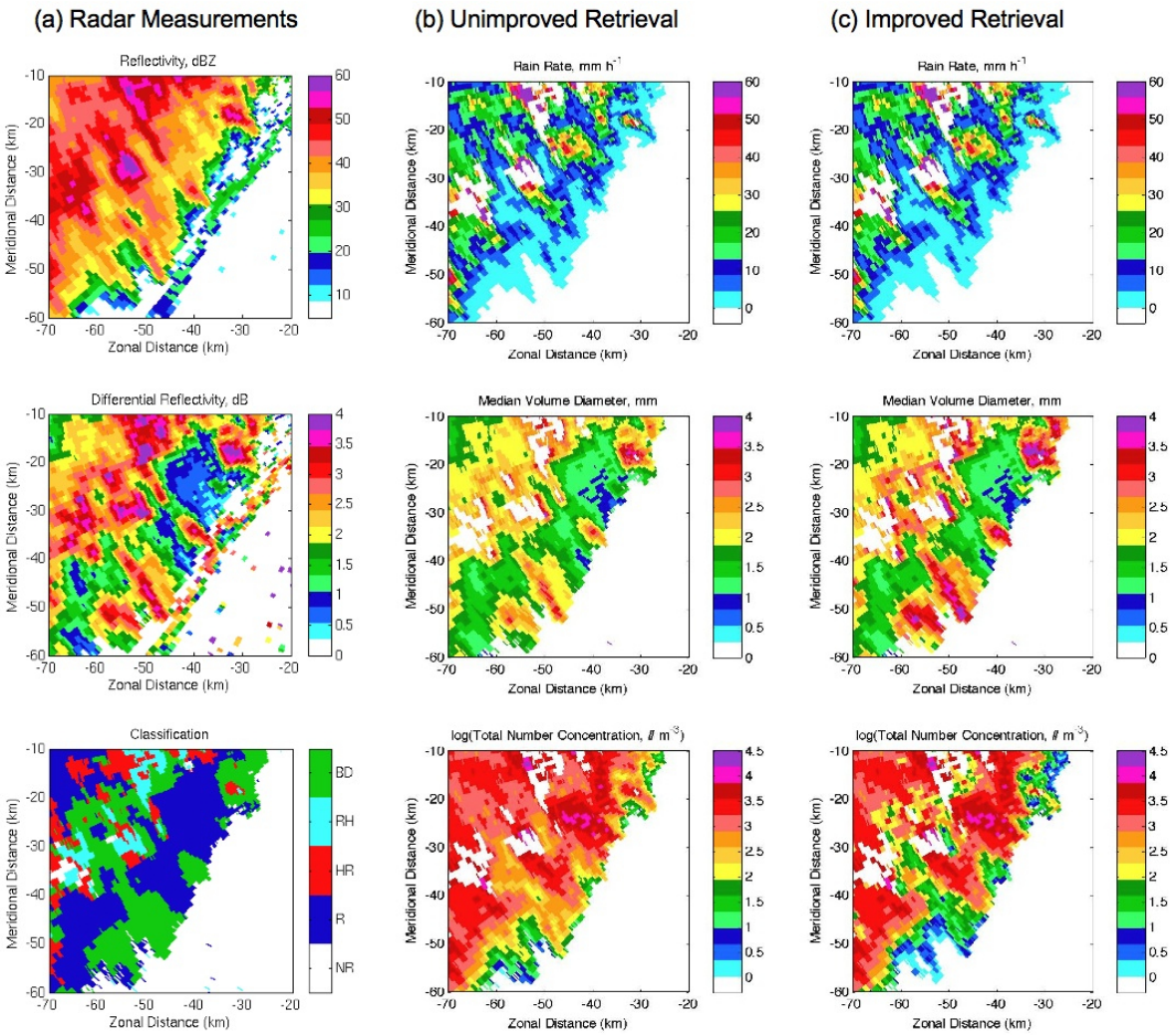

Figure 4.14: Comparison of radar retrievals based on adjusted and unadjusted $\mu - \Lambda$ relation. Column (a) shows radar-measured Z_H and Z_{DR} (065955 UTC on May 13, 2005), and classifications of rain (NR-no rain echo, R-light and moderate rain, HR-heavy rain, RH-rain/hail mixture, BD-big drops). Z_H and Z_{DR} classified as NR and RH have been filtered out before the rain retrievals were performed. Column (b) shows radar retrieval results of R, D_0 and N_T based on the refined $\mu - \Lambda$ relation without adjustment (Eq. 3.16). Column (c) shows radar retrieval results of same variables based on the $\mu - \Lambda$ relation with adjustment (Eq. 4.25).

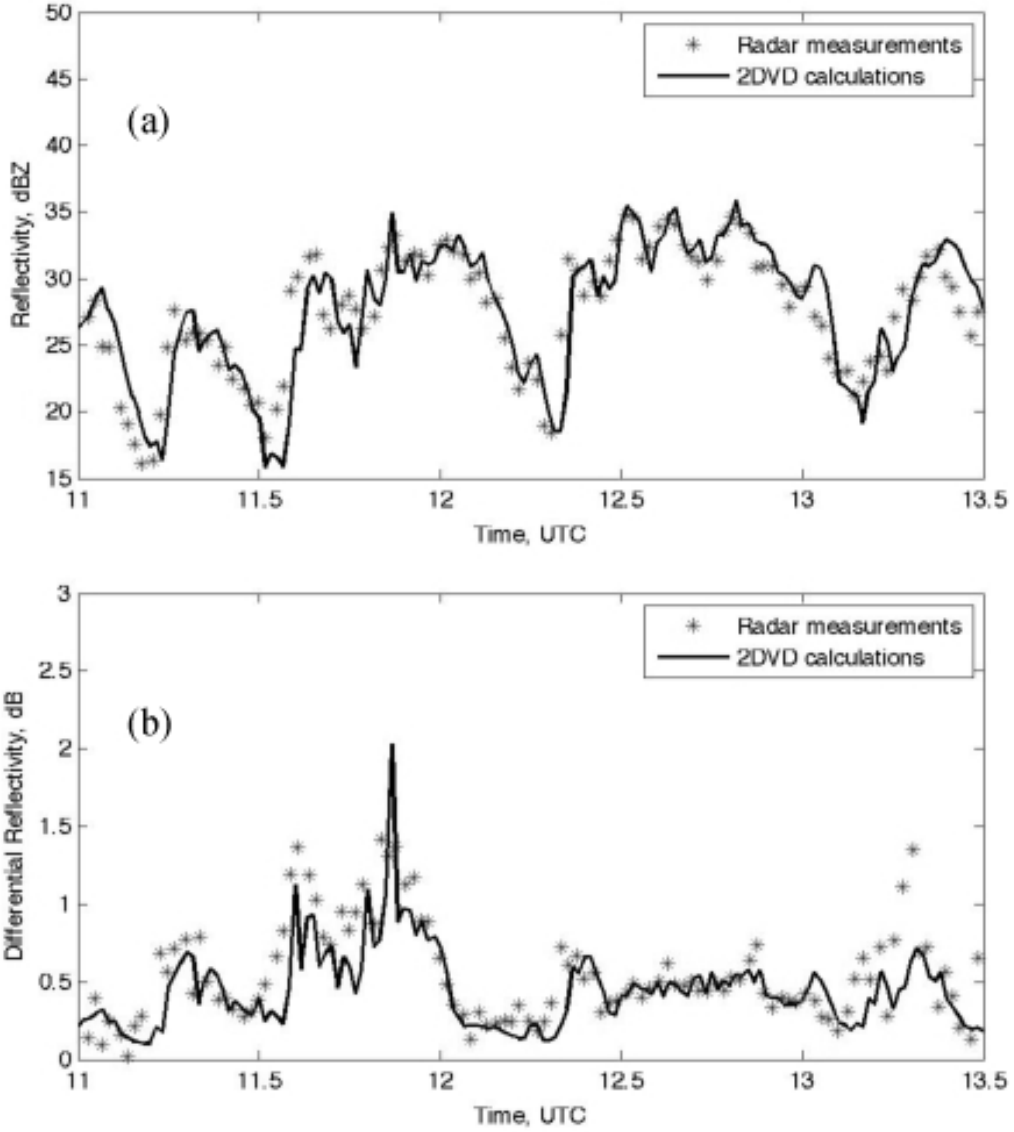

Figure 4.15: Comparison of radar measurements and disdrometer calculations for (a) reflectivity. (b) differential reflectivity. Asterisks denote radar measurements on May 2, 2005. Solid lines denote that reflectivity and differential reflectivity are calculated from observed DSDs by 2DVD.

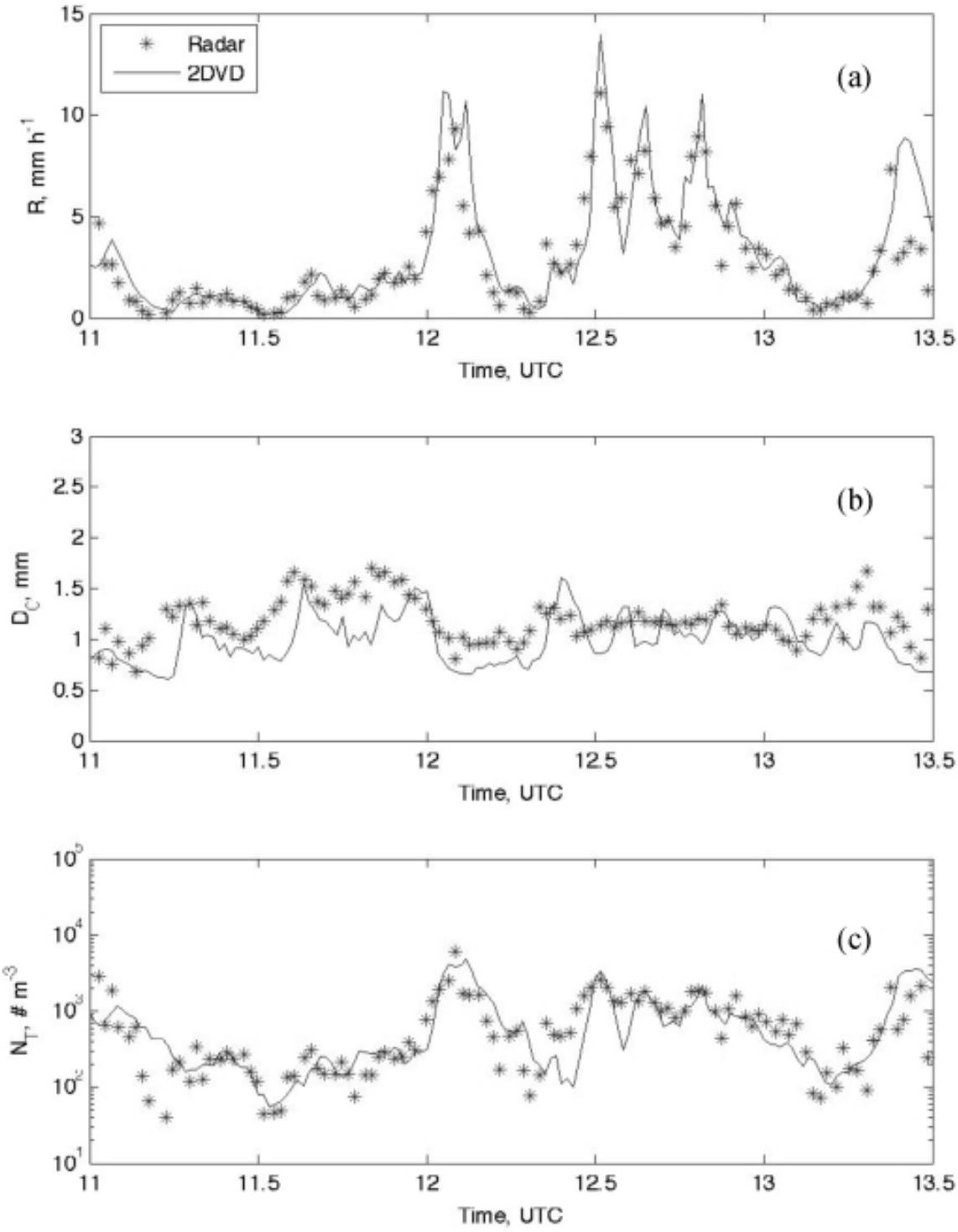

Figure 4.16: Comparison of (a) rain rate, (b) median volume diameter and (c) total number concentration from radar retrievals (asterisks points) and disdrometer observations (solid lines). (May 2, 2005)

adjusted by subtracting 0.3 dB. Possible contamination from a low melting layer and ground targets was mitigated by removing data points with a cross correlation coefficient less than 0.9. The solid lines denote Z_H and Z_{DR} calculated from disdrometer observations. Fig. 4.15 shows disdrometer calculations match radar measurements quite well. The discrepancy between radar measurements and disdrometer calculations during some short periods (e.g., at 1140UTC and at 1320UTC) is attributed to the inhomogeneity of spatial distribution of rain.

Using the C-G DSD model, rain parameters were retrieved from dual-polarization radar measurements. Fig. 4.16 shows the comparison of R, D_0, and N_T between radar retrievals and disdrometer calculations. The asterisks denote radar retrievals and solid lines denote disdrometer calculations from observed DSDs. Referring to Fig. 4.15, if the radar-measured Z_H and Z_{DR} agree with the disdrometer measurements, the retrieved rain variables in Fig. 4.16 match the disdrometer measurements. Compared to retrieval results in a previous study [Cao et al. (2006), Fig. 7], the retrieval of D_0 has been improved, especially for the period of light rain. The N_T for light rain is also close to the disdrometer observation. Considering the sampling volume difference between radar and disdrometer, the refined disdrometer-based $\mu - \Lambda$ relation is believed to give a satisfactory retrieval using polarimetric radar measurements. The adjustment of $\mu - \Lambda$ relation gives a reasonable rain retrieval for BD cases.

4.4 Summary

In this chapter, disdrometer observations in Oklahoma are used to characterize rain microphysics in terms of radar variables at S-, C- and X-band frequencies. The DSDs observed by the disdrometer provide a realistic representation of rain microphysics, free of any assumption of a functional DSD. According to the analysis of disdrometer dataset, N_T, R, W, and D_0 are well characterized in terms of Z_H and Z_{DR}. R_e, R_a,

and V_{tm} are well characterized in terms of W and D_0. Empirical relations between A_H (or A_{DP}) and Z_H (or K_{DP}) are also given. The characterization of rain microphysics provides useful relations for the application of polarimetric radar data.

Many studies have showed that the C-G DSD model is valid to represent rain physics. When it is applied for the radar retrieval, the C-G DSD model yields satisfactory retrievals of rain variables (e.g., N_T, R, and D_0) for most rain events. However, when the DSD is characterized as one with a bimodal distribution, a long tail, or significant big drops (BD), the retrieval from radar measurements may have problem and retrieved variables tend to have large deviations from in-situ observations. For BD DSDs, retrievals based on relation 3.16 would overestimate N_T and underestimate D_0. A simple adjustment of the $\mu - \Lambda$ relation according to Eq. 4.25 is found to resolve this problem, giving better retrievals of N_T and D_0 as observed by the disdrometer. This adjustment improves the retrieval especially for the leading edge of convection. It is also worth noting that the C-G DSD model might be further improved by combining additional information such as temporal and spatial correlations, more observations, and prior statistical information. The DSD retrieval is a promising topic for radar meteorology community and deserves further researches.

Chapter 5

A Bayesian Approach for DSD Retrieval

5.1 Introduction

In the previous chapter, characterization of rain variables and DSD retrieval from PRD have been addressed. This chapter further studies the DSD retrieval and focuses on optimizing the use of radar measurements. It is worth noting that previous studies assume that radar measure the truth. In that case, all the radar measurements have the equal quality and retrievals have the equal reliability. However, it is known that: 1) radar data contain measurement errors; 2) different radar quantities have different measurement errors; and 3) radar does not necessarily measure the rain, i.e., other categories of targets could be included in the radar echo. If we directly apply radar measurements to do rain retrieval (e.g., DSD retrieval in chapter 4), some of the results could be totally wrong. The model error, which is attributed to different species of targets, and the measurement error could mask the truth of rain retrieval. Since the measurement error is an inevitable obstacle to the accuracy of rain estimators, how to quantify the error effect and optimize the use of radar measurements becomes an important issue. The realistic solution is that we quantitatively find different contributions to the radar echo and separate the rain part. It is practically impossible to realize this task. The other solution is that we try to identify how much the reliability of rain retrieval could be. That is to say, we allow the error effect exist

in the process of retrieval, but we should know how much it could affect the result of retrieval. Bayesian theory offers a promising method to realize this by optimizing the use of measurements [e.g., Evans et al. (1995); Mc Farlane et al. (2002); Di Michele et al. (2005); Chiu and Petty (2006)].

The Bayesian method has been applied on the rainfall estimation with radar data. One example is the study of Hogan (2007), who applied Bayesian approach spatially (i.e., variational method). Although the variational method is believed to be advanced in using PRD, it is always complicated and there remain many issues to be resolved. For example, Hogan (2007) applied Z_H as a strong constraint (i.e., Z_H was not in the variational scheme and its error was ignored). If there were not such a constraint, the variational system would be hard to converge with satisfied results. On the other hand, high-order DSD models (e.g., gamma DSD) have been demonstrated to represent rain physics well in many previous studies. However, they are complicated for the application in the variational scheme. For this reason, Hogan (2007) only assumed a simple Z_H-R relation (i.e., intrinsic simple DSD model with only one-parameter). In this chapter, the study is focused on the DSD retrieval using the Bayesian approach. Different with what Hogan did, current study applies Bayesian approach temporally (i.e. using historic data as the prior information). A C-G DSD model, which has been successfully applied in previous studies, is used in this scheme. Efforts are put into the use of DSD parameters, rather than integral parameters, as state parameters of the retrieval scheme.

Because the retrieval is assumed for rain signals, radar echoes other than pure rain signals are considered as rain signals with a model error. As a result, the model error as well as the measurement error jointly affect the retrieval. This retrieval scheme provides not only mean values of state parameters, but also their standard

deviations. The error effect can be determined by the standard deviation of state parameter, which also implies the reliability of retrieval.

It is worth noting that many factors are required for the success of the Bayesian approach. First and foremost is the correct prior information on rain characteristics. This study utilizes a large dataset of disdrometer measurements to construct the prior distribution of state parameter. Since it is reasonable to assume the disdrometer has measured the ground truth of rain (i.e., DSD), the prior distribution from the disdrometer observation is better than the presumed model, such as Gaussian PDF distribution, which is often used in other Bayesian approaches.

In this chapter, a Bayesian approach of rain retrieval is presented by retrieving DSD parameters from Z_H and Z_{DR} at S-band (10.7 cm). In addition to the evaluation by simulated radar data, a rain event on 13 May 2005 in central Oklahoma is analyzed to verify the algorithm by comparing radar retrievals with in-situ measurements of disdrometer and rain gauge. In order to aid the analysis, two popular empirical rain estimators are used to compare with the Bayesian algorithm. One empirical estimator, used by WSR-88D as a default estimator for mid-latitude rain, is only based on Z_H. The other one is an estimator based on Z_H and Z_{DR}. The two relations are, respectively:

$$R(Z_H) = 0.017 \times Z_H^{0.714} \tag{5.1}$$

$$R(Z_H, Z_{DR}) = 0.0142 \times Z_H^{0.77} Z_{DR}^{-1.67} \tag{5.2}$$

Eq. 5.2 was developed by NSSL for polarimetric radar applications. It had an optimum performance for rain during the JPOLE field campaign (Ryzhkov et al. (2005a)).

This chapter is organized as follows. Section 2 gives a detailed description of the Bayesian approach. The retrieval algorithm is evaluated in Section 3 using the disdrometer dataset. Section 4 gives a case study to demonstrate the application of this algorithm. Discussion and conclusion are provided in the last section.

5.2 Bayesian Approach

5.2.1 Theory

Bayesian theory provides a way to estimate state variables by evaluating the *a posterior* probability density function (PDF) of state variables. The mean values and standard deviations of state variable can be obtained with the *a posterior* PDF. According to Bayesian theory, *a posterior* PDF, $P_{post}(\mathbf{x}|\mathbf{y})$, is calculated by

$$P_{post}(\mathbf{x}|\mathbf{y}) = \frac{P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x})}{\int P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x}) \cdot d\mathbf{x}} \tag{5.3}$$

where state variable vector $\mathbf{x}$ represents parameters of interest. Vector $\mathbf{y}$ indicates available observations for the retrieval of physical properties. $P_{pr}(\mathbf{x})$ is the prior PDF of state variable. $P_f(\mathbf{y}|\mathbf{x})$ is a conditional PDF representing the relation between observations and state variables. Given the $P_{post}(\mathbf{x}|\mathbf{y})$, the mean value $E(\mathbf{x})$ and the standard deviation $SD(\mathbf{x})$ of state variable are therefore obtained by integrating $P_{post}(\mathbf{x}|\mathbf{y})$ as:

$$E(\mathbf{x}) = \frac{\int \mathbf{x} \cdot P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x}) \cdot d\mathbf{x}}{\int P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x}) \cdot d\mathbf{x}} \tag{5.4}$$

$$SD(\mathbf{x}) = \sqrt{\frac{\int \left(\mathbf{x} - E(\mathbf{x})\right)^2 \cdot P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x}) \cdot d\mathbf{x}}{\int P_f(\mathbf{y}|\mathbf{x}) \cdot P_{pr}(\mathbf{x}) \cdot d\mathbf{x}}} \tag{5.5}$$

Eqs. 5.3–5.5 are basic equations of the Bayesian approach. In this chapter, they are applied for the rain DSD retrieval. The following subsections address details on the state vector $\mathbf{x}$, observation vector $\mathbf{y}$, $P_{pr}(\mathbf{x})$, and $P_f(\mathbf{y}|\mathbf{x})$.

5.2.2 State parameters

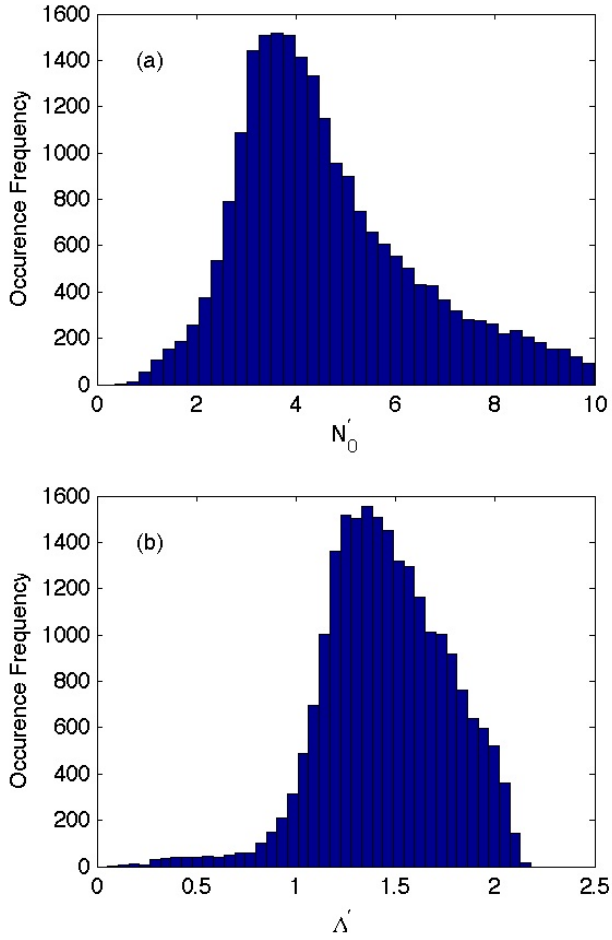

Figure 5.1: Histogram of estimated DSD parameters based on 2DVD data: a) N_0', unit in $log_{10}(\mathrm{mm}^{-1-\mu}\mathrm{m}^{-3})$, and b) Λ', unit in $\mathrm{mm}^{-1/4}$.

In many rain retrieval algorithms, integral variables such as rainfall rate are treated as the state parameter. Since DSD is the interest property of rain microphysics, this

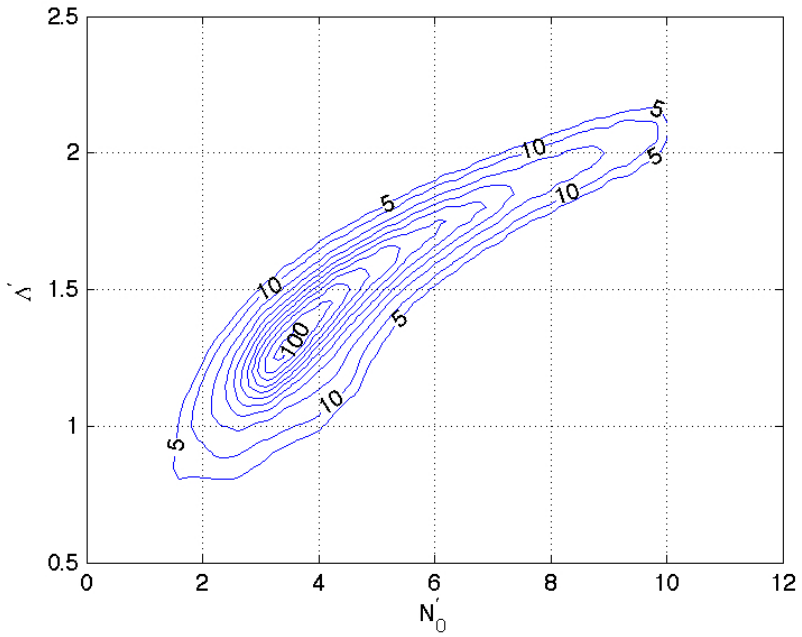

Figure 5.2: Contour of the occurrence frequency of joint estimated DSD parameters, N_0' [unit in $log_{10}(\mathrm{mm}^{-1-\mu}\mathrm{m}^{-3})$], and Λ' (unit in $\mathrm{mm}^{-1/4}$). The interval of unmarked contours between 10 and 100 is 10.

study applies DSD parameters as state parameters. For this study, a constrained-gamma DSD model with an updated constraining relation Eq. 3.16 is applied. Therefore, unknown states required for the retrieval are two DSD parameters, N_0 and Λ.

According to Bayesian theory, the prior PDF of state variable should be known. The prior distribution of DSD parameter can be found by fitting the DSDs observed by 2DVDs. If we assume measured DSDs follow the gamma distribution, gamma function parameters (i.e., N_0, Λ, and μ) can be fitted by a regression method. For each parameter, tens of thousands of values from all the DSDs are used to construct the occurrence distribution. Prior PDF of a parameter is then obtained by normalizing its occurrence distribution. For two parameters required for the retrieval, a joint PDF is constructed for the Bayesian approach. The regression procedure follows the truncated moment fit method described by Vivekanandan et al. (2004). It utilizes the 2^{nd}, 4^{th}, and 6^{th} DSD moments to estimate three gamma parameters. Although DSD parameters estimated in this way (i.e., using the moment method) would have a bias [e.g., Zhang et al. (2003); Smith et al. (2009)], studies in Chapter 3 have shown that the moment method using the 2^{nd}, 4^{th}, and 6^{th} moments has less error effects of model error and measurement error, compared to other moment methods and the maximum likelihood and L-moment methods.

As to N_0 and Λ distributions constructed from 2DVD dataset, a practical problem is that they are greatly skewed and have large dynamic ranges (not shown). It is difficult to discretize them and they are obviously not appropriate for practical use. In order to reduce their dynamic ranges, the original DSD parameters are transformed to new forms, i.e., $N_0' = log_{10}N_0$ and $\Lambda' = \Lambda^{0.25}$. The histograms of occurrence frequency are shown in figure 5.1. The dynamic ranges of N_0' and Λ' are reduced significantly.

With the transformation, the heavy rain, moderate rain and light rain could be represented by new parameters more appropriately. For example, originally Λ has a skewed distribution with its most portion representing the light rain. Now the moderate and heavy rain (e.g., $0 < \Lambda < 3$) account for a larger proportion of the dynamic range than before. Because the large Λ always has a larger bias (typically overestimation) than a small one, the shrinking of Λ range suggests that the bias would be reduced. For this study, state vector $\mathbf{x}$ in Eqs. 5.3–5.5 is set to be $[N_0', \Lambda']$. Their units are $log_{10}(\mathrm{mm}^{-1-\mu}\mathrm{m}^{-3})$, and $\mathrm{mm}^{-1/4}$, respectively. Correspondingly, observation vector $\mathbf{y}$ is set as $[Z_H, Z_{DR}]$, whose units are dBZ and dB, respectively.

The joint PDF of state variables are obtained by normalizing the occurrence frequency histogram. The detailed procedure is described as follows. Firstly, the range of parameter N_0' is discretized with interval 0.1. The range of parameter Λ' is discretized with interval 0.05. Sort the dataset of estimated DSD parameters into the grids of N_0' and Λ', and the occurrence frequency is then counted for each grid. Fig. 5.2 shows the contour of the result. As it shows, the two parameters have a correlation with each other, especially for large values of Λ'. Because the joint distribution is applied as the prior information, the effect of correlation is therefore accounted for in the Bayesian algorithm. Most DSDs have a N_0' between 3 and 5 (i.e., N_0 is about 10^3–10^5 $\mathrm{m}^{-3}\mathrm{mm}^{-1}$) and Λ' between 1.1 and 1.6 (i.e., Λ is about 1.5-6). Finally, the joint PDF is calculated by the normalization of distribution in Fig. 5.2. It is worth noting that this joint distribution entirely comes from observations and is free of any mathematical function. It represents the actual rain physics better than the assumption of a Gaussian distribution, which is always applied as the prior PDF in other Bayesian approaches of rain retrieval.

5.2.3 Conditional PDF

The conditional PDF $P_f(\mathbf{y}|\mathbf{x})$ is a bridge between measurements and state variables, analogous to the forward model used in the optimal-estimation theory. It characterizes how measurements would be with the given DSD parameters. The measurement error is also characterized by this PDF. Generally, it is difficult to find an true PDF to represent this process. Conventional method assumes a Gaussian distribution for the uncertainty of measurements. Following this assumption, the conditional PDF in current study is assumed to be a bivariate-normal distribution as:

$$P_f(Z_H, Z_{DR}|\Lambda', N_0') = \frac{1}{2\pi SD(Z_H)SD(Z_{DR})\sqrt{1-\rho^2}} \exp\left\{\frac{-1}{2(1-\rho^2)}\left[\frac{(Z_H - E(Z_H))^2}{SD^2(Z_H)} - \frac{2\rho(Z_H - E(Z_H))(Z_{DR} - E(Z_{DR}))}{SD(Z_H)SD(Z_{DR})} + \frac{(Z_H - E(Z_{DR}))^2}{SD^2(Z_{DR})}\right]\right\} \tag{5.6}$$

This equation gives a probability model of observed Z_H and Z_{DR}, given two DSD parameters Λ' and N_0'. Z_H and Z_{DR} in this model are assumed to work in the logarithmic units. The variable ρ in Eq. 5.6 denotes the correlation coefficient between Z_H and Z_{DR} errors. The error should include both observation error and model error. Generally, observations of Z_H and Z_{DR} can be considered to have independent observation errors. Most previous studies of Bayesian approach have considered the observation error only and assumed $\rho = 0$. However, the forward operator model might introduce the error correlation between these two variables. In this study, the C-G DSD model is applied to estimate Z_H and Z_{DR} in the forward operator. The model error tends to be correlated. Therefore, ρ should not be zero and vary with different Z_H-Z_{DR} pairs. It is normally hard to estimate this kind of correlation for each Z_H-Z_{DR} pair. Fortunately, it is found that the effect of ρ on the retrieval is not essential. For the simplicity, this study assumes a constant $\rho = 0.5$ for the application

of Eq. 5.5.

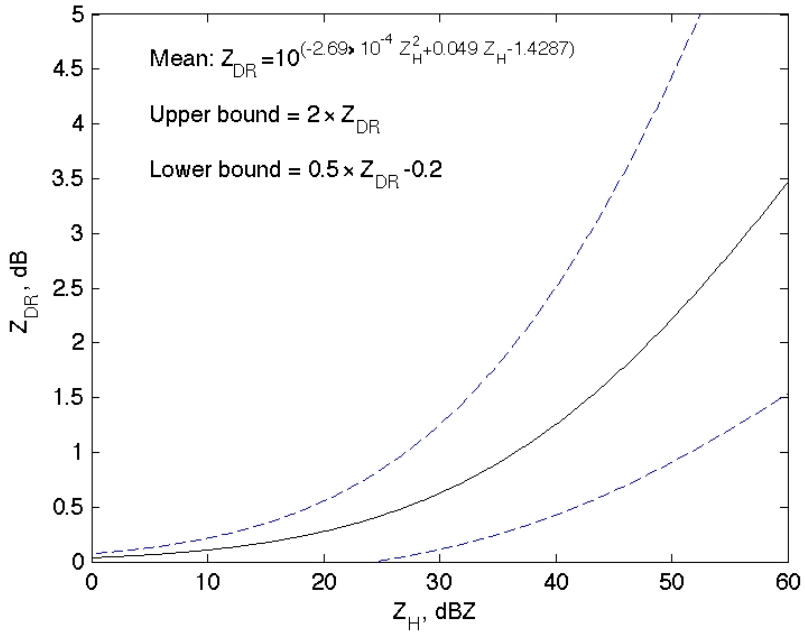

Figure 5.3: Sketch figure of Z_{DR} vs. Z_H from 2DVD measurements. Solid line denotes the mean curve (Eq. 15 of Cao et al. 2008). Upper bound and lower bound are given according to the mean curve.

Theoretically, $SD(\cdot)$ terms in equation 5.5 stand for the error effect (including the model error and the measurement error), which control the probability of "observed" Z_H and Z_{DR}. The real error effect ought to be complicated. As a result, the $SD(\cdot)$ values are not easy to decided for each pair of DSD parameters. Here, a simple model of $SD(\cdot)$ terms is proposed. As we know, radar measurement error for rain signals is normally 1–2 dB for Z_H and 0.1–0.3 dB for Z_{DR}. If radar signals are not for pure rain, Z_H and Z_{DR} values would be different than the ones estimated by the theorem of raindrop scattering, which is the basis of this retrieval algorithm. That is to say, the model error would be enlarged for the rain retrieval. In that case, $SD(\cdot)$ terms should be set to a larger value. Considering the Z_{DR} is much more sensitive to the model error than Z_H, we ignore the variation of $SD(Z_H)$ and only take the $SD(Z_{DR})$

into account.

Firstly, the $SD(Z_H)$ is assumed to be a constant (2 dB) in Eq. 5.6. $SD(Z_{DR})$ is assumed to be a function of Z_H and Z_{DR}. As discussed in previous paragraph, the value of $SD(Z_{DR})$ mainly depends on the measurement error for rain signal. For non-rain signal, the model error would become a significant factor. It is then reasonable to presume $SD(Z_{DR})$ grow with increasing the degree of non-rain. Fig. 10 of Cao et al. (2008) shows a common range of Z_H and Z_{DR} of rain signals, which are estimated from 2DVD observations. It implies that Z_H and Z_{DR} of rain should be limited in a bounded region. The boundaries can be approximately determined by the mean curve of Z_{DR} vs. Z_H (Eq. 15 of Cao et al. 2008). As it is shown in Fig. 5.3, the mean curve is denoted by a solid line with its equation noted on the figure. Two dashed lines give the upper boundary and lower boundary of rain data region. The upper boundary is the twice as much as the mean curve. The lower boundary is the half of mean curve minus 0.2. The following equation gives the $SD(Z_{DR})$ model. Within the bounded region, $SD(Z_{DR})$ is assumed to be a constant (0.3 dB). Outside of the region, $SD(Z_{DR})$ is assumed to increase linearly with the difference between observed Z_{DR} and the upper/lower boundary. The ratio 0.3 is an empirical number, which is determined by the presumption that 3 dB difference brings 1.2 dB $SD(Z_{DR})$. The model of $SD(Z_{DR})$ is given by:

$$SD(Z_{DR}) = \begin{cases} 0.3 \times (Z_{DR} - Z_{DR}^{up}) + 0.3, & \text{above upper boundary} \\ 0.3, & \text{within the region} \\ 0.3 \times (Z_{DR}^{low} - Z_{DR}) + 0.3, & \text{below lower boundary} \end{cases} \tag{5.7}$$

where, Z_{DR}^{up} (Z_{DR}^{low}) denotes the upper (lower) boundary. Eq. 5.5 implies that if an observed Z_{DR} deviated from the normal range of rain data, Z_{DR} would be less reliable

to represent the rain.

5.2.4 Forward model

The forward model is used to calculate polarimetric radar variables from DSD parameters. The basis of the calculation is the scattering theory of raindrops described in Chapter 2. Equations for calculating the Z_H and Z_{DR} are Eqs. 2.21 and 2.22. The backscattering cross sections of H and V directions are calculated as described in Chapter 4 with assumptions of radar wavelength 10.7 cm (S-band), raindrop temperature 10 °C, zero canting angle of raindrops, and experimental shape model (Brandes et al. 2002). In order to evaluate the DSD retrieval using the Bayesian approach, two integral rain variables, R(mm h^{-1}) and D_m (mm) according to Eqs. 2.13 and 2.18, are used.

In brief, the procedure of rain retrieval using this Bayesian approach is described as follows. Given two PRD, Z_H and Z_{DR}, the conditional probability is calculated with Eqs. 5.6 and 5.7 for any state variables N_0' and Λ'. It is integrated in Eqs. 5.5 and 5.4 to obtain mean values and standard deviations of N_0' and Λ' with their prior joint PDF. With retrieved DSD parameters $E(N_0')$ and $E(\Lambda')$, the gamma DSD is constructed. In this way, rain variables of interest such as rainfall rate can be calculated from the retrieved gamma DSD.

5.3 Evaluation of Algorithm

As described in previous subsection, the scattering theory is the basis of this Bayesian retrieval algorithm. Therefore, the success of this algorithm should depend on the presumption that the real PRD would be consistent with the quantities estimated by

the scattering theory. Given the premise, it is interesting to know how the algorithm perform for the rain retrieval. In other words, is the algorithm self-consistent with the results estimated from scattering theory? To evaluate the consistency of this algorithm, 2DVD measurements are used to simulate the PRD. In this way, the performance of this algorithm could be known given that the PRD follow the scattering theory.

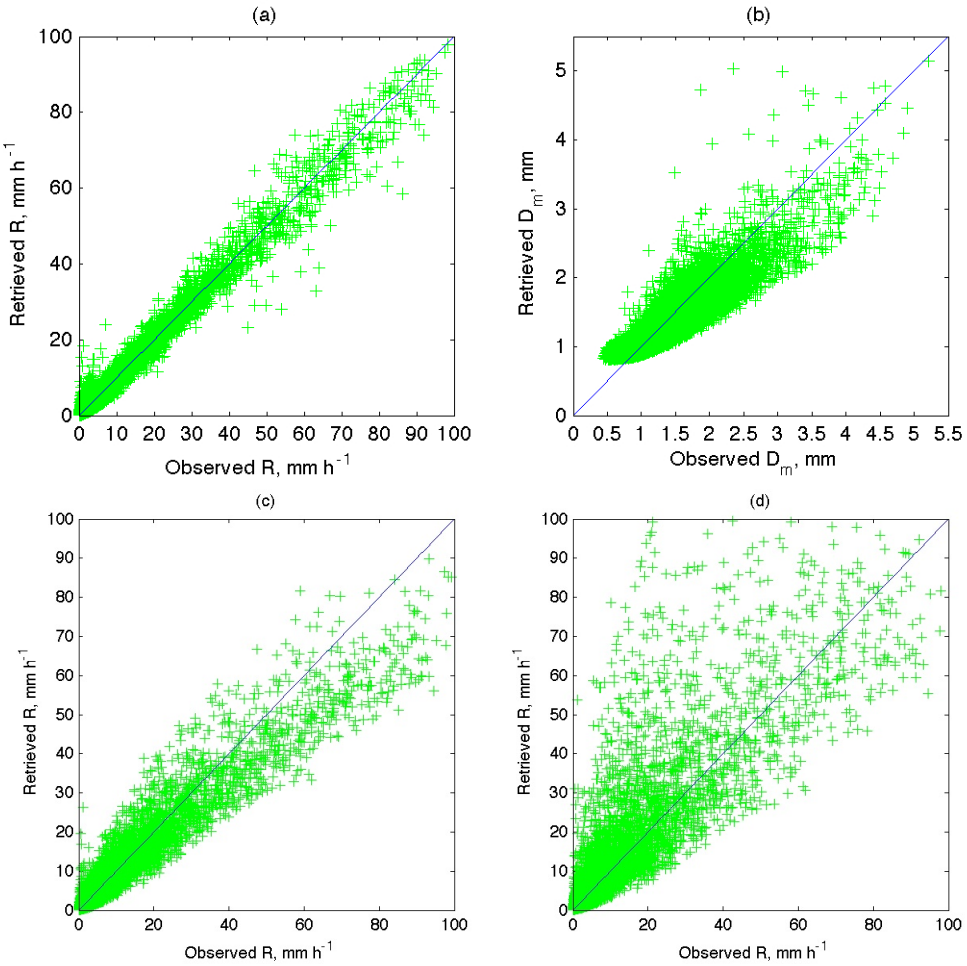

Figure 5.4: Scatter diagram of retrieved values versus observations: a) R, and b) D_m are results from Bayesian algrithm. c) and d) are retrieved rainfall rate using algorithms of $R(Z_H, Z_{DR})$ and $R(Z_H)$, respectively. Crosses represent data points and solid lines denote equal values of axes.

The 2DVD dataset used for the evaluation includes about 24,500 one-minute DSDs with the threshold of 50 raindrops. Z_H and Z_{DR} are simulated from these DSDs based on the scattering theory. The related equations are Eqs. 2.21 and 2.22. The simulated Z_H and Z_{DR} are then used to retrieve DSD parameters following the procedure described in previous subsection. Corresponding rain variables are calculated with retrieved mean values $E(N_0')$ and $E(\Lambda')$. Those rain variables can be compared with the ones calculated from observed DSDs. The statistics of the comparison shows the performance of the retrieval algorithm.

Table 5.1: Bias and RMSE of Bayesian retrievals versus 2DVD measurements

		$0.1 < R < 3$	$3 < R < 15$	$15 < R < 30$	$30 < R < 100$
R(mm h^{-1})	bias (%)	11.9	1.76	-0.64	-1.19
	RMSE (%)	49.7	17.3	11.5	21.5
D_m(mm)	bias (%)	-5.02	-4.43	0.74	8.93
	RMSE (%)	17.3	15.2	13.6	18.7

Fig. 5.4 shows the scatter diagram of retrieved results versus the observations. Each cross points represent the result for a 1-minute DSD. Figs. 5.4a and 5.4b gives the R and D_m comparison for Bayesian algorithm. For a reference, Figs. 5.4c and 5.4d display comparisons of variable R for NSSL's empirical estimators, Eqs. 5.1 and 5.2, respectively. As to variable R, the $R(Z_H)$ algorithm obviously has the worst performance. There is much more scattering of point than other two algorithms, which are based on two polarimetric variables. The reason of the large uncertainty is attributed to the fact that $R(Z_H)$ estimator applies intrinsic one-parameter DSD assumption, which can not well represent the variation of rain physics. Other two

algorithms, which apply the assumption of a two-parameter DSD, have a much better result. Moreover, retrievals by the Bayesian algorithm are a little better than by the empirical estimator $R(Z_H, Z_{DR})$. Estimator $R(Z_H, Z_{DR})$ has an evident underestimation for large rainfall rate (e.g., $R > 60$ mm h^{-1}). The uncertainty of estimator $R(Z_H, Z_{DR})$ is obviously larger too.

As Figs. 5.4a and 5.4b show, retrieved R and D_m by Bayesian algorithm have a good match with the 2DVD observations, with a correlation coefficients 0.98 (or 0.89) for R (or D_m). There are existing a little scattering of data points. The scattering increases especially with increasing D_m. The scattering of data points can likely be attributed to the defect of C-G model, which can not perfectly represent the variation of DSDs. For example, there is obvious large scattering of data points with observed $D_m < 3$ mm but retrieved $D_m > 3.5$ mm. These data points generally have concave DSDs and fewer median size raindrops (or a long tail, i.e., flat high end), which are not sufficient to form a gamma distribution. There are also some data points in Fig. 5.4a with undesired overestimate of R (e.g., observed $R < 8$ mm h^{-1} but retrieved $R > 10$ mm h^{-1}). These data points generally have convex and narrow DSDs with a low concentration of small raindrops ($D < 0.8$ mm). Retrieved DSDs under the constraint of $\mu - \Lambda$ relation always have more small raindrops than observations, leading to a larger R. Nevertheless, Figs. 5.4 shows that the C-G model still perform better than the model presumed by estimator $R(Z_H, Z_{DR})$ or $R(Z_H)$.

The statistics of Figs. 5.4a and 5.4b are shown in Table 5.1. The bias and RMSE of retrievals are used to show the performance of Bayesian algorithm. In order to distinguish different rain types, DSDs are classified with four categories by the range of rainfall rate, i.e., $0.1 < R < 3$ mm h^{-1}; $3 < R < 15$ mm h^{-1}; $15 < R < 30$ mm h^{-1}; $30 < R < 100$ mm h^{-1}. It is shown that R tends to be overestimated for $R < 3$

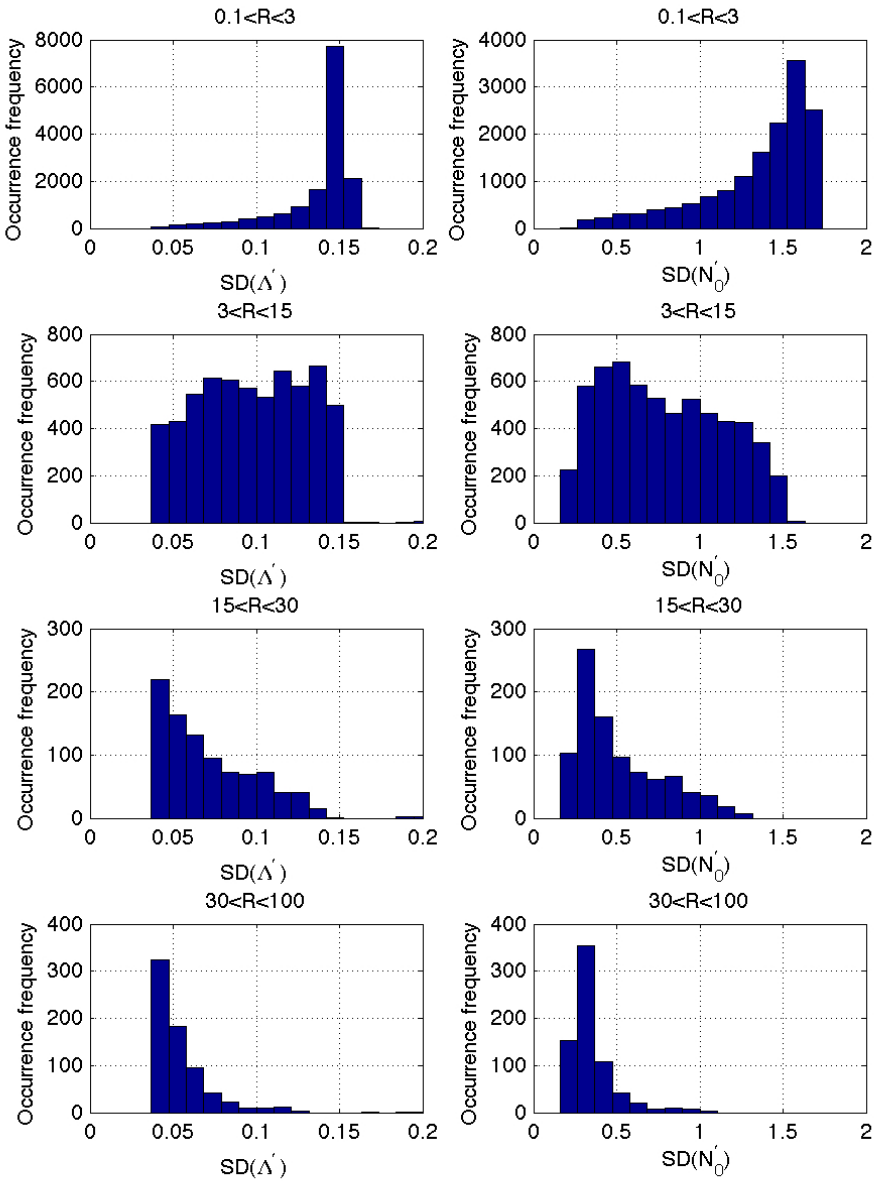

Figure 5.5: Occurrence histogram of retrieved SD values. The left column is for values of $SD(\Lambda')$ and the right column is for $SD(N_0')$. The rows from top to bottom are for data within the ranges of $0.1 < R < 3$, $3 < R < 15$, $15 < R < 30$, and $30 < R < 100$ mm h^{-1}, respectively.

mm h^{-1}. D_m tends to be underestimated for $R < 15$ mm h^{-1} while overestimated for $R > 30$ mm h^{-1}. The estimation of R has a greater uncertainty for light rain (i.e., $R < 3$ mm h^{-1}). The possible reason is that DSDs of light rain, which have fewer raindrops to construct a DSD, might have a larger measurement error or model error than other types of DSDs. Apart from the light rain, the bias of R estimate is less than 2% and the RMSE is lower than 22%. That is to say, the uncertainty of rain estimation is generally no more than 1 dB, which is less than the uncertainty of radar measurements (e.g., the uncertainty of Z_H measurement is generally thought to be ~ 2 dB).

Other variables to be evaluated are $SD(\cdot)$ terms retrieved by the Bayesian algorithm. The dataset of Z_H and Z_{DR} comes from rain data. As a result, statistics of retrieved $SD(\cdot)$ terms help the understanding of retrieval performance for rain data. Similar to the previous analysis, dataset are classified with several categories in terms of rainfall rate. Fig. 5.5 shows the occurrence histogram of estimated $SD(\cdot)$ values of state variables. It is evident that estimated $SD(\cdot)$ values tend to decrease with increasing R, implying that estimated state variables have the less uncertainty. Since the measurement errors have been modeled with $SD(Z_H)$ and $SD(Z_{DR})$ in Eq. 5.6, retrieved $SD(\Lambda')$ and $SD(N_0')$ represent the error effect of Z_H and Z_{DR}. Higher $SD(\cdot)$ values denote a larger error effect in the retrieval. Therefore, the $SD(\cdot)$ value could be regarded as an indicator of the data quality. Given the real radar data, if estimated $SD(\cdot)$ values are beyond the range for corresponding rainfall rate, it is likely that the radar data are not for pure rain and contamination might be included in the data.

Fig. 5.6 shows more analysis of $SD(\cdot)$ values. It gives the dependence of $E(\Lambda)$ and $SD(\Lambda)$ on the ratio of Z_{DR} to Z_H. It is worth noting $SD(\Lambda')$ and $SD(N_0')$ also have the similar characteristics to $SD(\Lambda)$. Fig. 5.6 shows the analysis of $SD(\Lambda)$

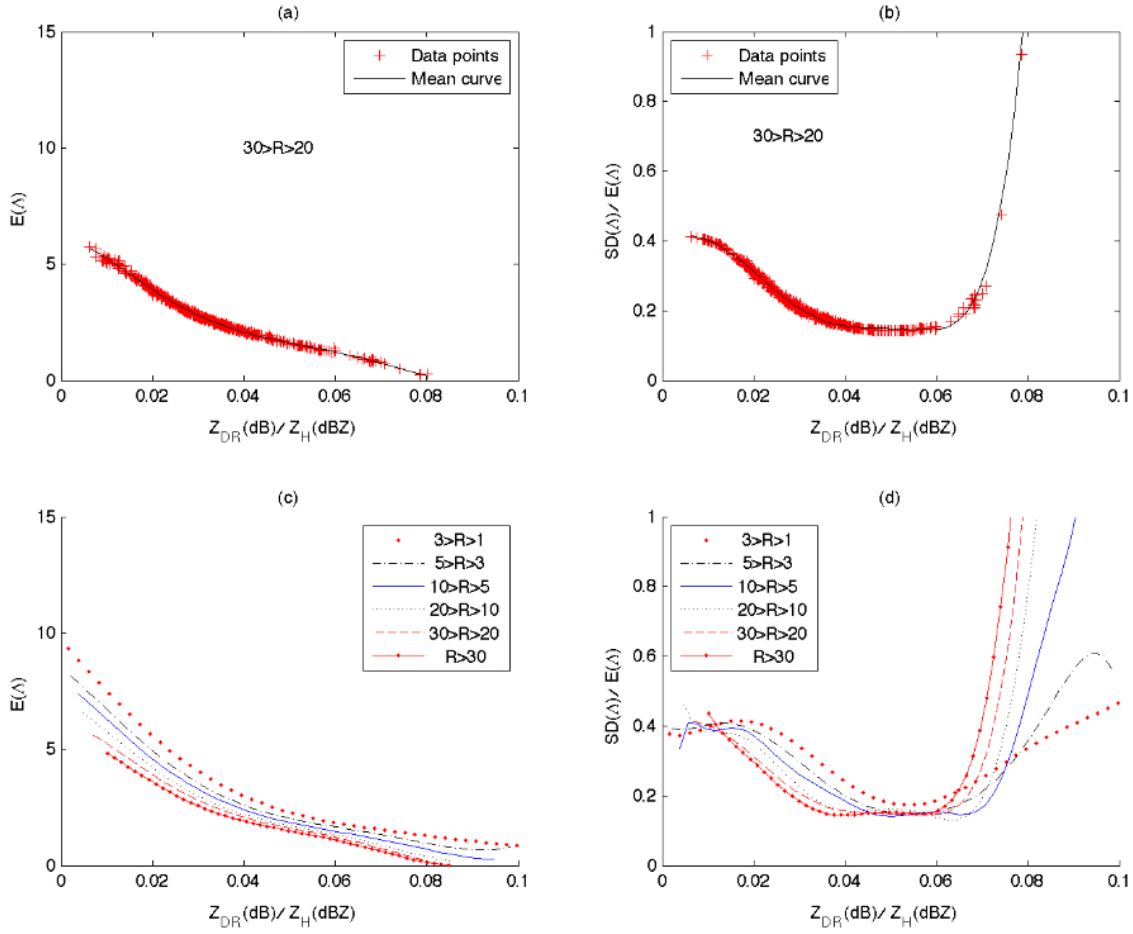

Figure 5.6: The dependence of $E(\Lambda)$ (unit in mm^{-1}) and $SD(\Lambda)$ (unit in mm^{-1}) on the ratio of Z_{DR} to Z_H. (a) and (b) show the mean curve fitted to data points of $20 < R < 30$. (c) and (d) display mean curves associated with data points of $1 < R < 3$, $3 < R < 5$, $5 < R < 10$, $10 < R < 20$, $20 < R < 30$, and $R > 30$ mm h^{-1}, respectively. The variable N_0' has the similar characteristics to Λ (not shown).

only because variable Λ is familiar to the community. In Figs. 5.6a and 5.6b, cross points denote data points with $20 < R < 30$ mm h^{-1}. Solid lines represent mean curves of $E(\Lambda)$ and normalized $SD(\Lambda)$ [i.e., $SD(\Lambda)/E(\Lambda)$]. These two curves are fitted by two multi-order polynomial functions. For $E(\Lambda)$, the general trend is that its value decreases with increasing ratio Z_{DR}/Z_H, implying large ratio would lead to an estimation of wide DSD. For normalized $SD(\Lambda)$, its value has the similar trend but it may increase significantly with increasing Z_{DR}/Z_H beyond a given value. For example, when $Z_{DR}/Z_H > 0.06$, the normalized $SD(\Lambda)$ value increases and the rate of increasing is much faster than its rate of decreasing. These two figures suggest that Z_{DR} and Z_H of rain signals have an intrinsic relation. The uncertainty of retrieved DSD parameter such as $SD(\Lambda)$ would be enlarged if these two PRD were apart from that relation. In other words, if retrieved $SD(\cdot)$ values are small, it is likely that the PRD represent the rain signal. Similar to Figs. 5.6a and 5.6b, Figs. 5.6c and 5.6d show mean curves for dataset with different R ranges. As figures show, all the categories have very similar tendency for curves of $E(\Lambda)$ and normalized $SD(\Lambda)$. Generally speaking, the larger the ratio of Z_{DR}/Z_H, the smaller the $E(\Lambda)$. In addition, Normalized $SD(\Lambda)$ has a low value for a certain range of Z_{DR}/Z_H ratio. Beyond that range the uncertainty of estimation (i.e., normalized SD values) would increase. Figs. 5.5 and 5.6 demonstrate the $SD(\cdot)$ characteristics of rain retrieval. This additional information from $SD(\cdot)$ terms provides the confidence/reliability of rain retrieval. It is one of major advantages of the Bayesian approach for rain retrieval.

5.4 Case Study

In this section, a case of radar-rain retrieval is studied. The Bayesian algorithm is applied to real radar data measured by a polarimetric WSR-88D, named KOUN. In order to verify the algorithm, rain gauge and disdrometer measurements are used to

compared with the retrieved results (i.e., rainfall rate and one-hour rain accumulation). Corresponding analysis are given for the comparison.

5.4.1 Data description

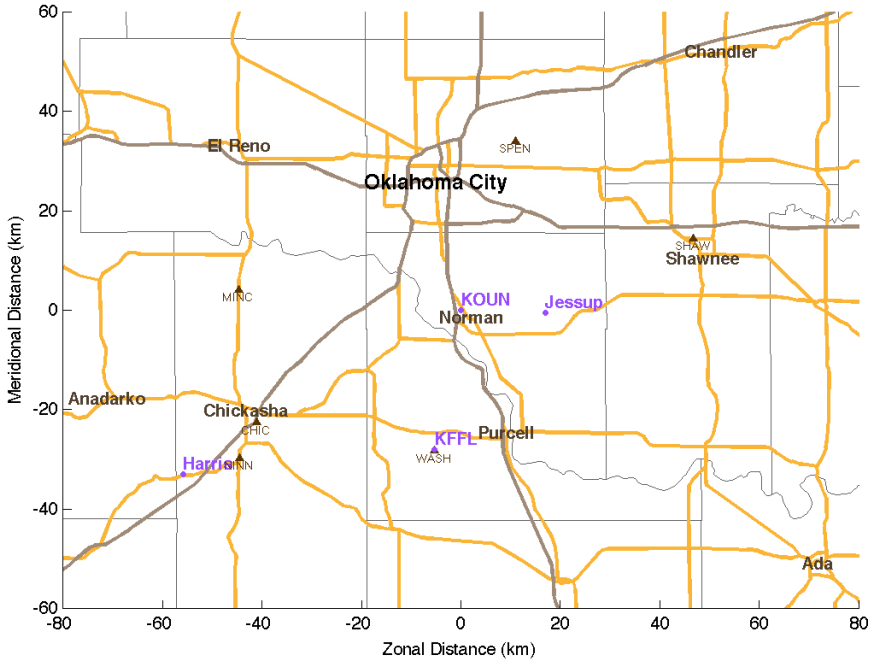

Figure 5.7: Locations of radar, disdrometer and rain gauge. KOUN radar is located in Norman. Six Oklahoma Mesonet sites (dark triangles), named SPEN, MINC, CHIC, NINN, WASH, and SHAW. OU 2DVD is deployed at KFFL.

As shown in Fig. 5.7, KOUN radar is located at Norman in central Oklahoma. Rain gauge measurements from six Oklahoma Mesonet sites are used in this study. The six sites (noted by triangles in the figure), named SPEN, MINC, CHIC, NINN, WASH, and SHAW, are located at 35.7 km north, 45.0 km west, 47.0 km southwest, 53.6 km southwest, 28.7 km south, and 48.8 km east of KOUN, respectively. A disdrometer is also used to do the comparison. It was placed at the site of KFFL, which

is only about 300 m away from the Mesonet site WASH.

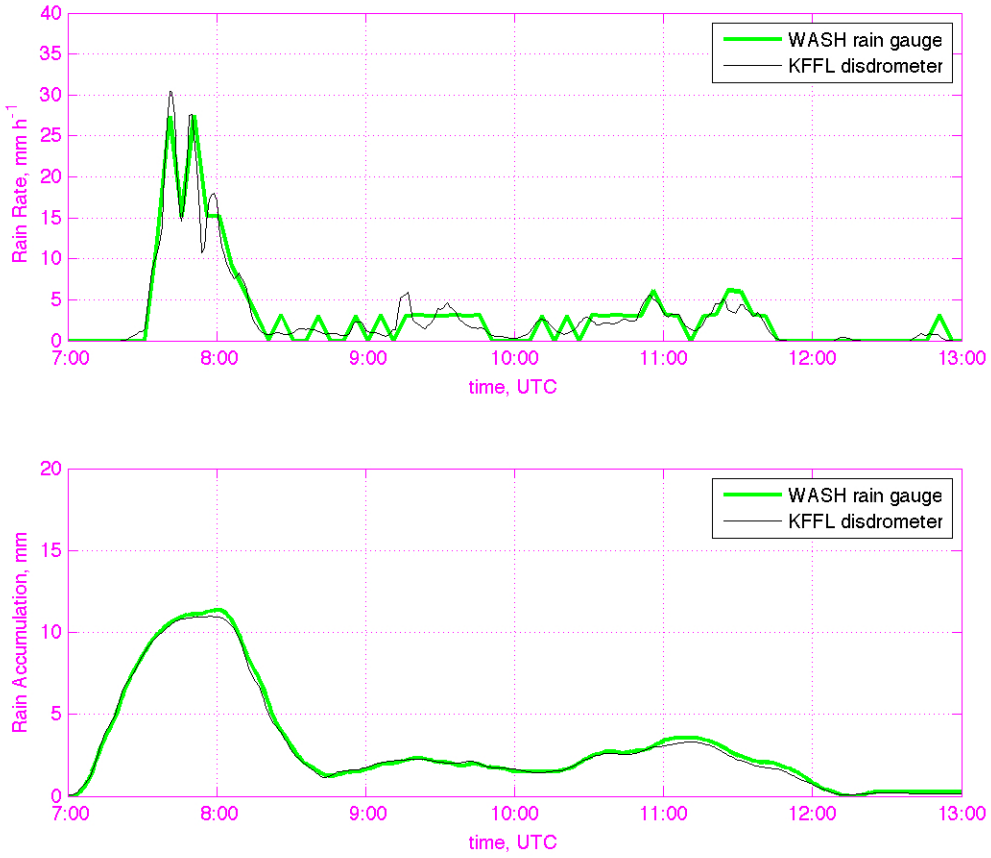

Figure 5.8: Rain comparison between measurements from disdrometer and Mesonet rain gauge: a) rainfall rate; b) one-hour rain accumulation. The one-hour rain accumulation at a given time was calculated by accounting for the rainfall within half hour before and after this time.

The instrumental validation of rain gauge can be examined by comparing its measurements to those of the disdrometer. The rain gauge at WASH was so close to the disdrometer that rain properties measured by both instruments should be similar, though there were effects of the rain inhomogeneity and the measurement error. Fig. 5.8 shows the comparison between disdrometer and rain gauge measurements. Two

subfigures are comparisons of rainfall rate and one-hour rain accumulation, respectively. The rainfall rate of rain gauge is calculated based on total rain of 5 minutes rain. The one-hour rain accumulation at a given time was calculated by accounting for the rainfall within half hour before and after this time. Because the Mesonet rain gauge recorded rain data every five minutes while the 2DVD recorded data every minute, to be consistent, the 2DVD data are smoothed with five minutes' interval. As Fig. 5.8 shows, for both parameters of rainfall rate and one-hour rain accumulation, temporal variations match with each other very well. The fractional difference of measurements between two instruments is only 3.18%, which is less than the allowance for measurement error of rain gauge (i.e., 5%). Fig. 5.8 demonstrates the consistency between two instruments and also suggests that rain gauge and disdrometer are reliable sources to be regarded as measuring the ground truth.

On 13 May 2005 a squall line passed through Oklahoma from the northwest to the southeast. Fig. 5.9 shows PPI images of KOUN radar measurements at 0830UTC. As shown in Z_H and Z_{DR}, the convective rain was followed by a large coverage of stratiform precipitation. The convective core had Z_H as large as 55 dBZ and Z_{DR} as large as 3.5 dB. Radar PPI scans were made at an elevation of 0.5, meaning that the radar beam center was at a height of approximately 250 m at the range of 28.4 km (where disdrometer was deployed). Fig. 5.9 also shows the cross-correlation coefficient (ρ_{HV}) and results of hydrometer classification developed by NSSL (Schuur et al. (2003); Ryzhkov et al. (2005c)). The classification algorithm is based on PRD such as Z_H, Z_{DR}, K_{DP}, and ρ_{HV}. The meaning of the hydrometer classes are NR: non-rain echo, R: rain, HR: heavy rain, RH: rain and hail mixture, and BD: big drop. The following subsection presents the radar data quality control before the application of retrieval algorithm.

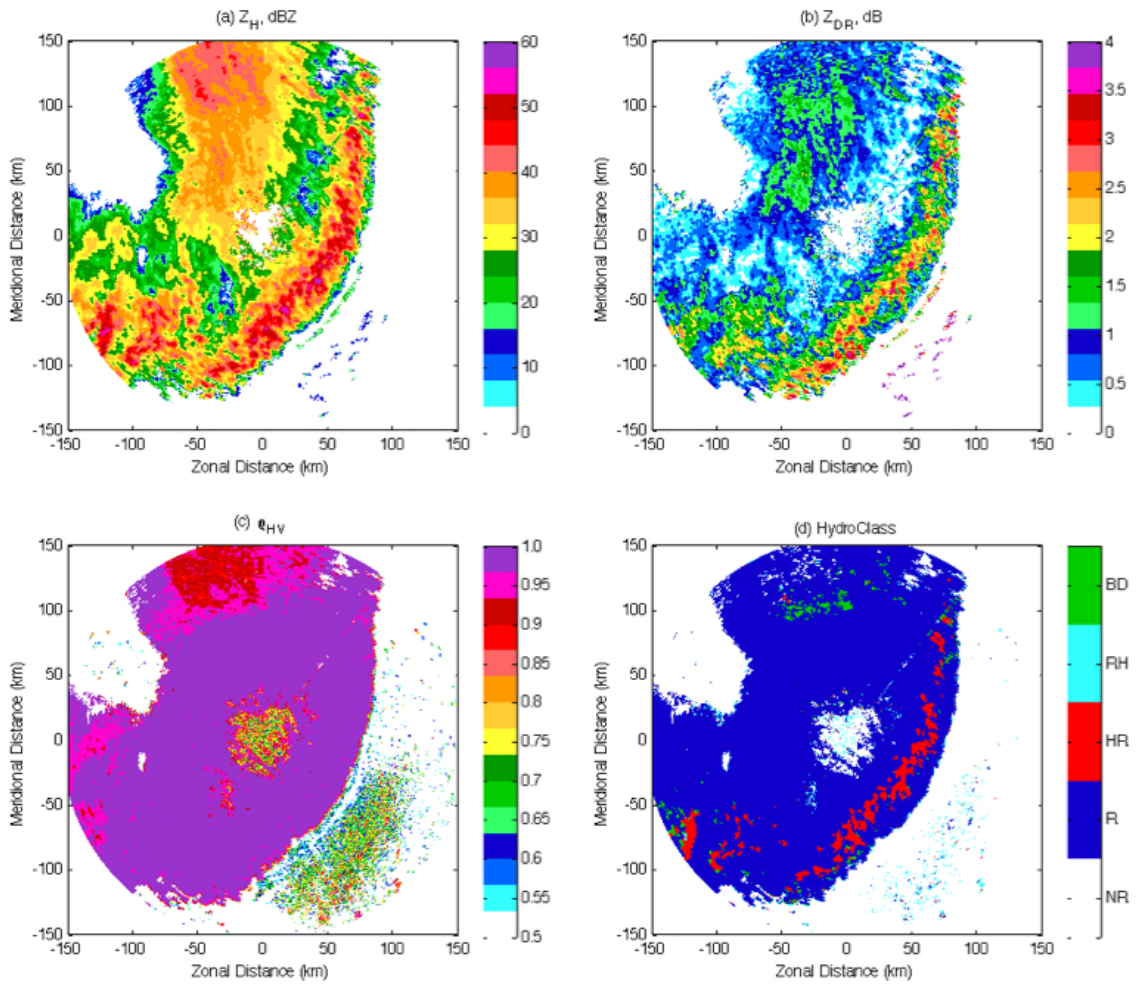

Figure 5.9: PPI image of KOUN radar observations at 0830 UTC on 13 May 2005: a) Z_H, b) Z_{DR}, c) ρ_{HV}, and d) hydroclass. The meaning of hydrometer classification are NR: non-rain echo, R: rain, HR: heavy rain, RH: rain and hail mixture, and BD: big drop.

5.4.2 Radar data quality control

In order to smooth the data and filter out speckle, a speckle filtering procedure was performed on the radar-measured Z_H and Z_V. Filtered Z_{DR} was calculated from filtered Z_H and Z_V. The concept of speckle filter applied in this study is similar to

the one described by Lee et al. (1997) except for minor changes. Assuming radar reflectivity is combined with a multiplicative noise N, the observation is then expressed in logarithmic domain by

$$Z^{O}_{H,V}(dBZ) = Z^{T}_{H,V}(dBZ) + N(dB) \tag{5.8}$$

where subscripts H and V represent horizontal and vertical polarization, respectively; superscripts O and T denote the observation and the truth, respectively; and $N(dB)$ is assumed to be a random white noise with Gaussian PDF. Therefore, radar reflectivity at horizontal or vertical polarization could be estimated by

$$\hat{Z}^{T}_{H,V}(dBZ) = \overline{Z^{O}_{H,V}}(dBZ) + b\Big[Z^{O}_{H,V}(dBZ) - \overline{Z^{O}_{H,V}}(dBZ)\Big] \tag{5.9}$$

where,

$$b = \frac{\mathrm{Var}\big[Z^{T}_{H,V}(dBZ)\big]}{\mathrm{Var}\big[Z^{O}_{H,V}(dBZ)\big]} \tag{5.10}$$

and,

$$\mathrm{Var}\big[Z^{T}_{H,V}(dBZ)\big] = \mathrm{Var}\big[Z^{O}_{H,V}(dBZ)\big] - \mathrm{Var}\big[N(dB)\big] \tag{5.11}$$

The Var[·] means the variance. In this study, the mean and variance of observed $Z_{H,V}$ are estimated from observations of adjacent area of distance 1 km. The measurement error [i.e., standard deviation of $N(dB)$] is assumed to be 2 dB. The minimum variance of $Z_{H,V}$ is assumed to be 1. That is, the minimum value of b is set to 0.2.

Because the verification was conducted only for pure rain, the study has eliminated radar echo pixels contaminated by hail, anomalous propagation, biological scatters, and so on. This task was accomplished based on the hydrometer classification algorithm developed by NSSL. The eliminated region was filled up with estimated Z_H (or Z_{DR}), which was the averaged Z_H (or Z_{DR}) of rain signals at adjacent area within 1 km distance. If contamination area was large (e.g., storm core region with large area of rain/hail echoes), Z_H (or Z_{DR}) would be estimated by spatially interpolation using

Figure 5.10: Time-series figure of un-calibrated Z_H and Z_{DR} from radar, as well as Z_H and Z_{DR} calculated from 2DVD measurements.

rain signals at adjacent area (not necessarily within 1 km distance).

As discussed in the introduction, radar measurements were calibrated using 2DVD measurements. For the rain event on 13 May 2005, Fig. 5.10 shows the time-series figures of un-calibrated Z_H and Z_{DR} from radar (for pure rain), as well as Z_H and Z_{DR} calculated from 2DVD measurements. Being interpolated into a consistent time scale (i.e., with one minute interval from 0700 to 1300 UTC), the difference between two lines was calculated and averaged. The difference was found to be -1.08 dB for

Z_H or 0.36 dB for Z_{DR}. Radar measured Z_H and Z_{DR} of a full scanning were hence calibrated by subtracting these two values before used for the Bayesian retrieval.

5.4.3 Analysis of retrievals

Fig. 5.11 shows one example of retrieval results, which are based on KOUN radar observations shown in Fig. 5.9. Fig. 5.11a gives the PPI image of retrieved R by the Bayesian algorithm. Retrieved R by empirical dual-pol estimator (i.e., Eq. 5.2) is close to this figure (not shown). Fig. 5.11b shows the retrieved R by empirical single-pol estimator (by Eq. 5.1). There are less region of rain output in Fig. 5.11a, which has been filtered by the classification result of Fig. 5.11d. For example, Fig. 5.11b contains some speckles at southeast of the leading edge of the storm, which cannot be filtered by the speckle filter. However, those speckles are normally biological echoes and have been rejected in Fig. 5.11a. Although Figs. 5.11a and 5.11b have similar storm feature and close R estimate within the stratiform region, it is obvious that Fig. 5.11b has a higher R estimate than Fig. 5.11a in the region of strong convection, which extends from northeast through southwest. Figs. 5.11c and 5.11d display images of retrieved $SD(\Lambda')$ and $SD(N_0')$ by Bayesian algorithm, respectively. Both $SD(\cdot)$ images have a similar trend, implying that either one could be used alone. It is worth noting that both Figs. 5.11c and 5.11d have directly applied PRD shown in Figs. 5.11a and 5.11b without the filtering of classification result (i.e., Fig. 5.11d). The good thing is that the $SD(\cdot)$ images are consistent with the classification results. The rain region typically has small $SD(\cdot)$ values. Large $SD(\cdot)$ values, which are beyond their normal ranges for the rain, indicate there is no rain at the region (e.g., $SD(N_0') > 2.5$).

If rain gauge and disdrometer measurements are assumed to be ground truth, the Bayesian algorithm can be verified through the quantitative comparison between

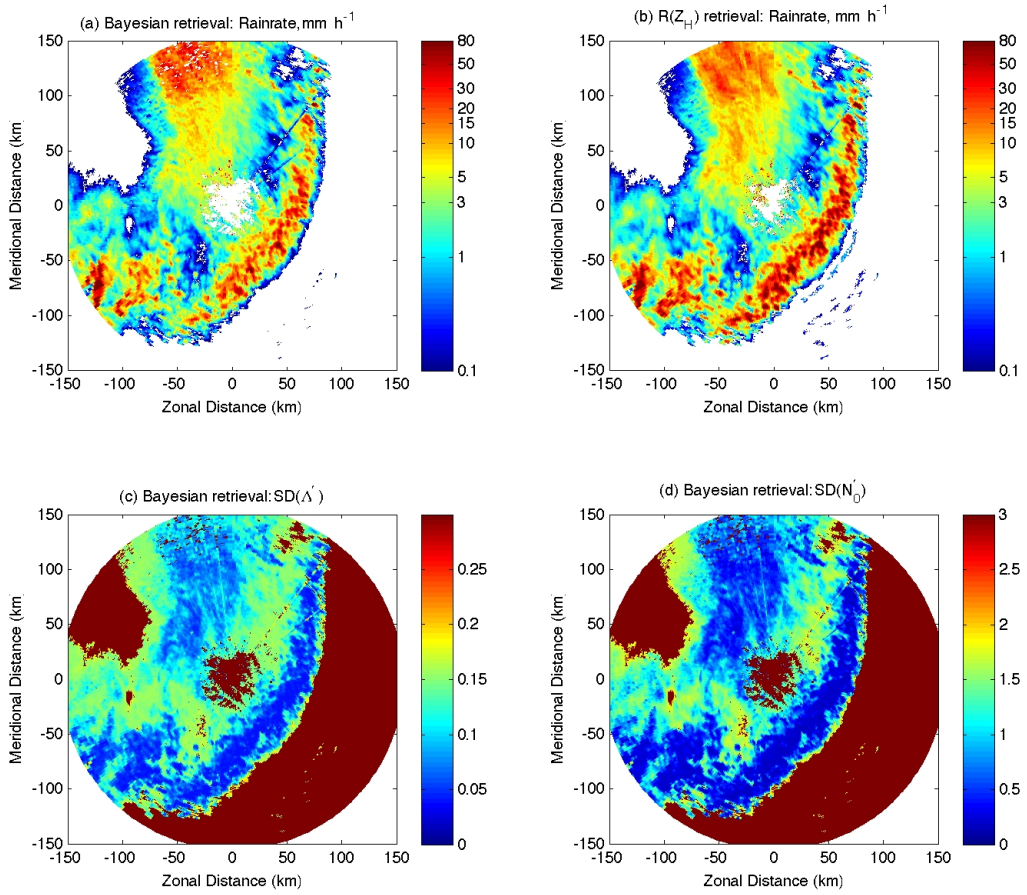

Figure 5.11: Retrieval results from radar observations shown in Fig. 5.9 (i.e., 0830 UTC on 13 May 2005). a) rainfall rate by Bayesian retrieval, b) rainfall rate by $R(Z_H)$ retrieval, c) $SD(\Lambda')$ by Bayesian retrieval, and d) $SD(N_0')$ by Bayesian retrieval.

observations and retrievals. Fig. 5.12 (or 5.13) shows comparisons of rainfall rate (or one-hour rain accumulation) for retrievals and observations at different sites (six rain gauges and one disdrometer). In both figures thick solid lines indicate surface measurements. Thin solid lines denote radar retrievals using the Bayesian algorithm. As a reference, dashed lines give radar retrievals using the empirical dual-pol relation (Eq. 5.2) and dotted lines indicate radar retrievals using the empirical single-pol relation (Eq. 5.1). As figures show, the Bayesian estimator is practically consistent

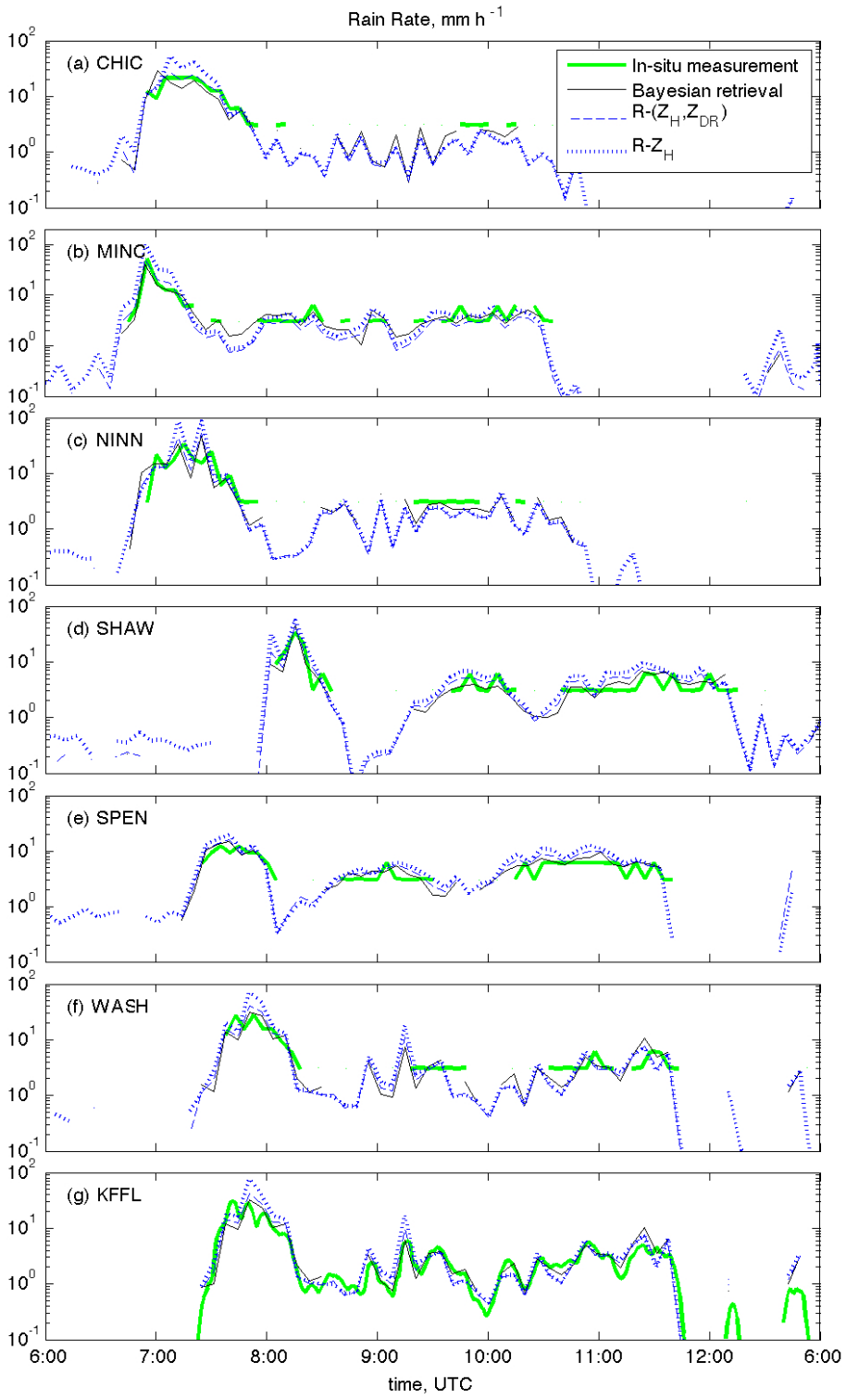

Figure 5.12: Rainfall rate comparisons between radar retrievals and in-situ measurements at seven sites: a) CHIC; b) MINC; c) NINN; d) SHAW; e) SPEN; f) WASH; g) KFFL.

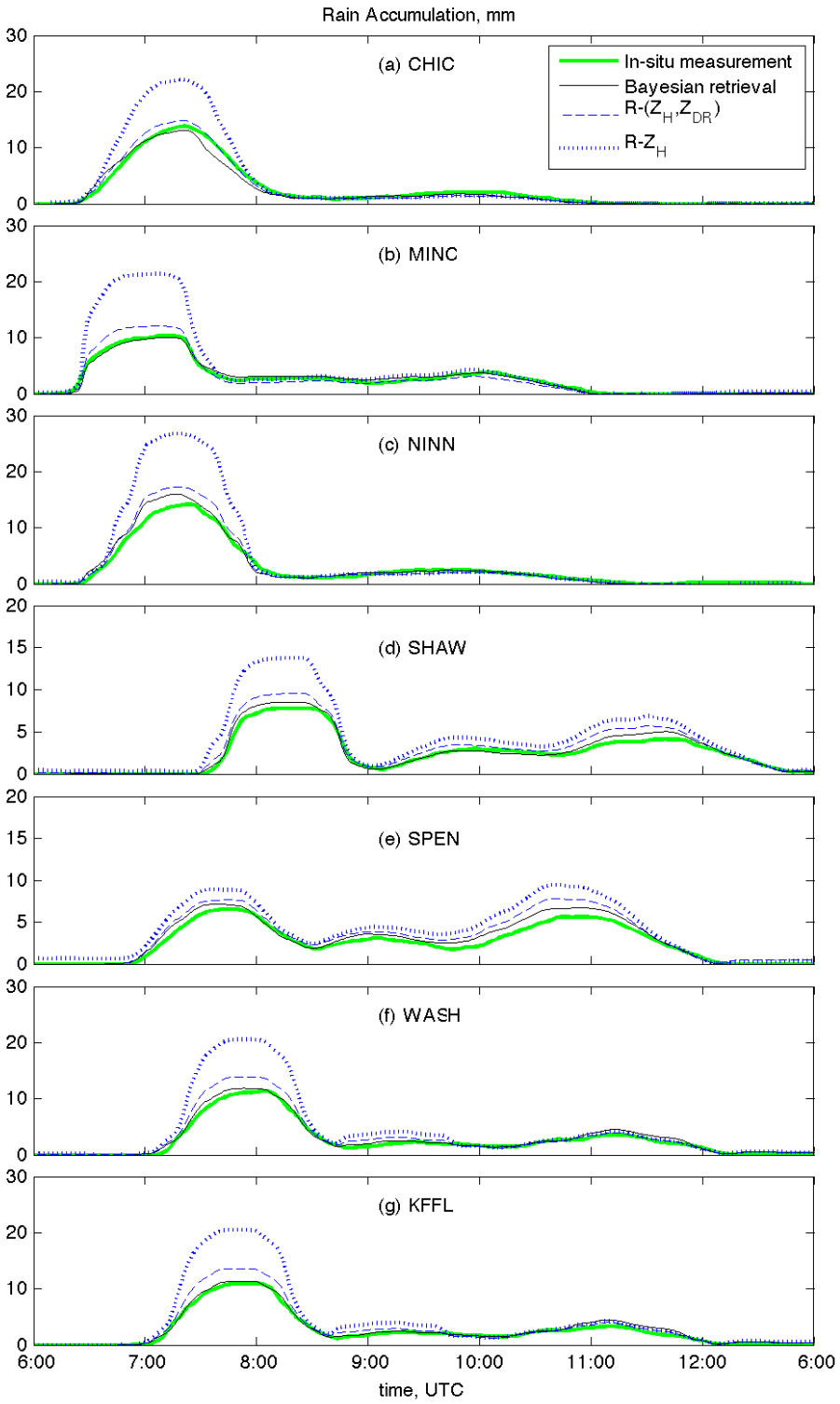

Figure 5.13: One-hour rain accumulation of radar retrievals and in-situ measurements at seven sites: a) CHIC; b) MINC; c) NINN; d) SHAW; e) SPEN; f) WASH; g) KFFL.

Table 5.2: Bias and RMSE of rain retrievals (one-hour rain accumulation) versus in-situ measurements at sites

		CHIC	MINC	NINN	SHAW	SPEN	WASH	KFFL
Bayesian	bias (%)	-7.3	2.5	8.3	7.8	15.5	10.7	10.0
	RMSE (%)	21.0	11.5	21.7	14.7	19.6	16.4	15.5
$R(Z_H, Z_{DR})$	bias (%)	1.1	1.6	12.3	22.4	28.0	19.8	20.2
	RMSE (%)	17.3	26.4	33.5	29.9	33.9	34.4	33.3
$R(Z_H)$	bias (%)	35.4	48.1	48.2	49.5	44.9	48.9	50.4
	RMSE (%)	80.4	103.5	106.8	72.3	56.0	95.3	92.5

with the dual-pol estimator. Both estimators give satisfactory results of rainfall rate (or one-hour rain accumulation), capturing the temporal variation of surface measurements. The single-pol estimator normally overestimates rainfall during the convection while it performs fairly well in the stratiform region. It is worth noting that the mixture of rain/hail might exist near convective cores (e.g., around 0655UTC at MINC). Radar measured Z_H and Z_{DR} are sometimes extremely large (e.g., $Z_H = 55$ dBZ and $Z_{DR} = 3.5$ dB). If no quality control were performed, rainfall rate estimated from contaminated Z_H and Z_{DR} could be much larger than 100 mm h^{-1}. However, surface measurements in Figs. 5.12 and 5.13 show that this is not the case. Radar retrievals in these figures demonstrate that the data quality control [i.e., using radar measurements (classified as rain) from an adjacent area to interpolate into a contaminated region] can give a reasonable rain estimation for the contaminated region.

By comparison, the Bayesian retrieval is superior to the empirical dual-pol retrieval. In additoin, it is reasonable to see that both of them are better than the single-pol retrieval. As Fig. 5.13 shows, the empirical dual-pol retrieval tends to overestimate the one-hour rain accumulation for heavy (sometimes moderate) rain. The empirical single-pol retrieval has even worse performance, especially during the strong convection. Table 5.2 gives the bias and RMSE of one-hour rain accumulation retrievals versus in-situ measurements. As it shows, the empirical single-pol retrieval has the worst results. At seven sites, the Bayesian retrieval has small biases ($\sim 10\%$) and RMSE ($\sim 20\%$) compared to those of empirical retrievals. With the exception of CHIC, the empirical dual-pol retrieval generally has a larger bias and standard deviation than the Bayesian retrieval.

5.5 Conclusion

In this chapter, a Bayesian algorithm is proposed for rain estimation through estimating gamma DSD parameters from Z_H and Z_{DR}. For this algorithm, the prior PDF of state variables is constructed using a large set of 2DVD data collected in central Oklahoma. The conditional PDF of the radar observation is assumed to follow a bivariate Gaussian model. Forward models of radar variable calculation are based on the theory of raindrop backscattering. The evaluation from 2DVD data and verification by a case study have shown that the Bayesian algorithm has the potential to improve the rain estimation from PRD.

There are still issues about the implementation of this algorithm. First, different with empirical estimators, the basis of this algorithm is the raindrop backscattering theory. Radar measurements should be consistent with the estimation from ground observations. Many previous studies have shown such a consistence to some degree.

Given that radar measurements match with disdrometer measurements, the proposed Bayesian algorithm is effective as shown in this study. On the contrary, if radar observations cannot be explained by the scattering theory, i.e., the radar and surface measurements are not consistent, this algorithm won't give the right retrieval. In that case, Bayesian algorithm will consider the mismatch as observation/model error. The corresponding $SD(\cdot)$ terms would reflect such an error effect. This fact also implies that calibration to the raw radar data might be required when this Bayesian algorithm is applied for other radars. In the case study the equivalence has been shown between the rain gauge and disdrometer. This fact suggests that rain gauge, which is widely used nationwide, could be an alternative source for radar calibration in practical applications.

This study also shows that the Bayesian algorithm essentially has the similar performance with the empirical dual-pol estimator (i.e., Eq. 5.2). The similarity comes from the fact that both methods are based on the application of Z_H and Z_{DR} and an intrinsic two-order DSD model. However, the Bayesian retrieval, which does not use deterministic coefficients, performs better. It gives the estimate with the maximum posterior probability, as well as the standard deviation of estimate, which can be used as a good indicator of radar data quality for rain echoes. Moreover, although the Bayesian approach in this study is limited in two PRD, other PRD such as correlation coefficient and differential phase can be incorporated into the observation $\mathbf{y}$ in Eq. 5.3, making the algorithm more effective. However, there are still issues on the appropriate assumption of their conditional distribution model (i.e., Eq. 5.6).

In addition, it is worth noting that the approach introduced in this chapter incorporates the historic information and does not apply the spatial information. As a result, the spatial algorithm such as attenuation correction is beyond the scope of

this chapter. The variational method could be a good candidate to solve this kind of problem. Related research will be shown in next chapter.

Chapter 6

Variational Method for Optimal DSD Retrieval in Presence of Attenuation

6.1 Introduction

Previous chapters mainly focus on the application of radar data without considering the attenuation. The reason is that WSR-88D radar network in the United States works at the S-band frequency ($\sim$ 10 cm wavelength), for which the attenuation effect is minor and can be neglected for most cases. However, most national weather radar networks in the world operate at a higher frequency. For example, many weather radars used in Europe and the TDWR (Terminal Doppler Weather Radars) of the U.S. operate at the C-band ($\sim$ 5 cm wavelength). Recently, X-band ($\sim$ 3 cm wavelength) weather radars such as those in the CASA (Center for Collaborative Adaptive Sensing of the Atmosphere) IP1 network have received more attention. Unlike S-band radars, the propagation effect of precipitation attenuation on C-band and X-band measurements can not be ignored. The attenuation is a significant problem for radar-based rain estimation and precipitation microphysics studies at these shorter wavelengths.

For single-polarization radars, the attenuation correction is mainly based on the Hitschfeld-Bordan (H-B) method and its revised versions [e.g., Delrieu et al. (2000); Zhang et al. (2004); Berne and Uijlenhoet (2006)]. With dual-polarization radars, the measured propagation phases (differential phase or specific differential phase) have been widely used for the attenuation correction. Such algorithms include the direct phase correction (DP) method [e.g., Bringi et al. (1990)], data fitting method [Ryzhkov and Zrnic (1995)], ZPHI algorithm [Testud et al. (2000)], self-consistence (SC) method [Bringi et al. (2001)] and revised SC methods [Park et al. (2005); Vulpiani et al. (2005); Gorgucci and Baldini (2007); Liu et al. (2006); Ryzhkov et al. (2007)]. All these algorithms apply various empirical relations associated with the attenuation. For example, the deterministic power law relation between the attenuation and radar reflectivity is the basis for the H-B and revised H-B algorithms. The power law relations between the attenuation and specific differential phase are essential for phase-based attenuation corrections. These empirical relations should be uniquely known for the DP method while the coefficients could be adjusted by the SC method and its revised versions.

There are problems in attenuation correction algorithms mentioned above. The attenuation estimated from empirical relations may be affected by the strong constraints that sacrifice a lot of physical variabilities. Moreover, the measurement error, which can further deteriorate the attenuation estimation, are not fully taken into account in these algorithms. There are possible ways to solve these problems. Since drop size distribution of hydrometeors (e.g., DSD) is fundamental for precipitation microphysics, the sacrifice of physical variability can be mitigated by estimating DSD parameters, which are used to estimate the attenuation [e.g., Meneghini and Liao (2007)]. In addition, the effect of measurement error can be minimized through a

variational approach by optimizing the use of all available measurements with error-based weights [e.g., Hogan (2007); Xue et al. (2009)]. The combination of these two approaches should have great potential to improve attenuation correction and QPE. To do this work, DSD parameters need to be estimated as part of the state vector. Because of the involvement of DSD parameters in the observation operator, the variational scheme becomes highly non-linear. The forward model of radar measurements (i.e., attenuated observations) and the corresponding partial derivatives are complicated functions of DSD parameters. The development of adjoint codes is a problem.

In this chapter, two approaches mentioned above are combined for the first time to correct the attenuation and estimate DSD parameters from X-band (or C-band) PRD. The DSD is retrieved through a two-dimensional variational scheme. Attenuation effects are built into the forward observation operator and the attenuation correction is accomplished adaptively during the iterative optimization/estimation process. The rest of this chapter is organized as follows. The methodology is described in Section 6.2. The algorithm is evaluated in Section 6.3 using simulated X-band PRD from S-band radar measurements. Retrievals based on real PRD are analyzed in Section 6.4. Limitations and potentials of the algorithm are discussed in Section 6.5.

6.2 Methodology

6.2.1 Variational approach

This variational algorithm applies three PRD, Z_H, Z_{DR}, and K_{DP}, for the DSD retrieval. The optimal use of measurements involves the minimization of a cost function as:

$$J(\mathbf{x}) = J_b(\mathbf{x}) + \mathbf{w}_{Z_H} J_{Z_H}(\mathbf{x}) + \mathbf{w}_{Z_{DR}} J_{Z_{DR}}(\mathbf{x}) + \mathbf{w}_{K_{DP}} J_{K_{DP}}(\mathbf{x}) \tag{6.1}$$

where,

$$J_b(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}_b) \tag{6.2}$$

$$J_{Z_H}(\mathbf{x}) = \frac{1}{2}\left[H_{Z_H}(\mathbf{x}) - \mathbf{y}_{Z_H}\right]^T \mathbf{R}_{Z_H}^{-1}\left[H_{Z_H}(\mathbf{x}) - \mathbf{y}_{Z_H}\right] \tag{6.3}$$

$$J_{Z_{DR}}(\mathbf{x}) = \frac{1}{2}\left[H_{Z_{DR}}(\mathbf{x}) - \mathbf{y}_{Z_{DR}}\right]^T \mathbf{R}_{Z_{DR}}^{-1}\left[H_{Z_{DR}}(\mathbf{x}) - \mathbf{y}_{Z_{DR}}\right] \tag{6.4}$$

$$J_{K_{DP}}(\mathbf{x}) = \frac{1}{2}\left[H_{K_{DP}}(\mathbf{x}) - \mathbf{y}_{K_{DP}}\right]^T \mathbf{R}_{K_{DP}}^{-1}\left[H_{K_{DP}}(\mathbf{x}) - \mathbf{y}_{K_{DP}}\right] \tag{6.5}$$

The cost function J is composed of four parts. J_b is the background part. The other three terms correspond to observations of Z_H, Z_{DR}, and K_{DP}, respectively. Superscript T denotes the matrix transpose. Matrix $\mathbf{w}$ represents relative weights of the observation terms and is associated with the signal-to-noise ratio (SNR). Vector $\mathbf{x}$ is the state vector and $\mathbf{x}_b$ is the background or first guess. Vector $\mathbf{y}$ indicates the radar observations. H denotes the nonlinear observation operator of radar variables. $\mathbf{B}$ is the background error covariance matrix. $\mathbf{R}$ is the observational error covariance matrix. Matrix $\mathbf{w}$ can be regarded as a part of $\mathbf{R}$. This study separates them for the convenience of defining a simple $\mathbf{w}$ in term of SNR. Subscripts Z_H, Z_{DR} and K_{DP} are used to denote the terms for the corresponding observations. Above equations try to follow the standard notations used in the modern data assimilation literature, as defined in Ide et al. (1997).

The size of matrix $\mathbf{B}$ is n^2 where n is the size of state vector $\mathbf{x}$. The full matrix is usually huge. The matrix computation and storage, especially for the inversion of $\mathbf{B}$, can be a major problem during the iterative minimization of the cost function. To solve this problem, a new state variable $\mathbf{v}$ is introduced, written as,

$$\mathbf{v} = \mathbf{D}^{-1}\delta\mathbf{x} \tag{6.6}$$

with $\delta\mathbf{x} = \mathbf{x} - \mathbf{x}_b$ and $\mathbf{D}\mathbf{D}^T = \mathbf{B}$ [Parrish and Derber (1992)]. Notation δ indicates the increment. $\mathbf{D}$ is the square root of the background error covariance matrix $\mathbf{B}$. The cost function is then rewritten as follows,

$$\begin{aligned} J(\mathbf{v}) =& \frac{1}{2}\mathbf{v}^T\mathbf{v} + \frac{1}{2}\mathbf{w}_{Z_H}\big[H_{Z_H}(\mathbf{x}_b + \mathbf{D}\mathbf{v}) - \mathbf{y}_{Z_H}\big]^T\mathbf{R}_{Z_H}^{-1}\big[H_{Z_H}(\mathbf{x}_b + \mathbf{D}\mathbf{v}) - \mathbf{y}_{Z_H}\big] + \\ & \frac{1}{2}\mathbf{w}_{Z_{DR}}\big[H_{Z_{DR}}(\mathbf{x}_b + \mathbf{D}\mathbf{v}) - \mathbf{y}_{Z_{DR}}\big]^T\mathbf{R}_{Z_{DR}}^{-1}\big[H_{Z_{DR}}(\mathbf{x}_b + \mathbf{D}\mathbf{v}) - \mathbf{y}_{Z_{DR}}\big] + \\ & \frac{1}{2}\mathbf{w}_{K_{DP}}\big[H_{K_{DP}}(\mathbf{x}_b + \mathbf{D}\mathbf{v}) - \mathbf{y}_{K_{DP}}\big]^T\mathbf{R}_{K_{DP}}^{-1}\big[H_{K_{DP}}(\mathbf{x}_b + \mathbf{D}\mathbf{v}) - \mathbf{y}_{K_{DP}}\big] \end{aligned} \tag{6.7}$$

In this way, the inversion of $\mathbf{B}$ is avoided. The minimization of cost function J is achieved by searching the minimum gradient of cost function $\nabla_v J$, which is given by:

$$\begin{aligned} \nabla_v J =& \mathbf{v} + \mathbf{w}_{Z_H}\mathbf{D}^T\mathbf{H}_{Z_H}^T\mathbf{R}_{Z_H}^{-1}\big(\mathbf{H}_{Z_H}^T\mathbf{D}\mathbf{v} - \mathbf{d}_{Z_H}\big) + \\ & \mathbf{w}_{Z_{DR}}\mathbf{D}^T\mathbf{H}_{Z_{DR}}^T\mathbf{R}_{Z_{DR}}^{-1}\big(\mathbf{H}_{Z_{DR}}^T\mathbf{D}\mathbf{v} - \mathbf{d}_{Z_{DR}}\big) + \\ & \mathbf{w}_{K_{DP}}\mathbf{D}^T\mathbf{H}_{K_{DP}}^T\mathbf{R}_{K_{DP}}^{-1}\big(\mathbf{H}_{K_{DP}}^T\mathbf{D}\mathbf{v} - \mathbf{d}_{K_{DP}}\big) \end{aligned} \tag{6.8}$$

$\mathbf{H}$ represents the Jacobian operator, a matrix containing the partial derivative of observation operator H with respective to each element of the state vector. Vector $\mathbf{d}$ is the innovation of observation, i.e., $\mathbf{d} = \mathbf{y} - H(\mathbf{x}_b)$.

The spatial influence of the observation is determined by the background error covariance matrix $\mathbf{B}$. Huang (2000) showed that the element b_{ij} of matrix $\mathbf{B}$ could be modeled as a spatial filter,

$$b_{ij} = \sigma_b^2 \exp\left[-\frac{1}{2}\Big(\frac{r_{ij}}{r_L}\Big)^2\right] \tag{6.9}$$

where subscripts i, j denotes two grid points in the analysis space; σ_b^2 is the background error covariance; r_{ij} indicates the distance between the i^{th} and j^{th} grid points; r_L is the decorrelation length of the observed physical quantity. In this chapter, r_L is assumed to be a constant in the two-dimensional analysis space, i.e., the error covariance is spatially homogeneous at horizontal plane, as is for the isotropic covariance

option in Liu and Xue (2006). The square root of **B**, i.e., **D**, can be computed by applying a recursive filter described by Gao et al. (2004) and Liu et al. (2007). In this way, the cost of computation and storage can be reduced significantly (by a factor of **B** dimension), compared to the computation of inversion of **B**.

6.2.2 Forward observation operator

The C-G DSD model has been successfully applied in rain retrieval in previous two chapters. Therefore, it is appropriate for testing the variational retrieval here. In this chapter, $N_0' = \log_{10}(N_0)$ and Λ are chosen as state variables. The state vector $\mathbf{x}$ is composed of N_0' and Λ at all grid points. Given two DSD parameters at each grid point, the DSD can be determined. Therefore, rain properties, including intrinsic Z_H and Z_{DR}, as well as K_{DP} can be estimated as well. The forward operators of Z_H and Z_{DR} follow Eqs. 2.21 and 2.22, respectively. The forward operator of K_{DP} is given by

$$K_{DP} = \frac{180\lambda}{\pi} \int_0^\infty \mathrm{Re}\big[f_a(0) - f_b(0)\big] N(D) dD \quad (^\circ \mathrm{km}^{-1}) \tag{6.10}$$

where $f_a(0)$ and $f_b(0)$ represent the forward scattering amplitudes at horizontal and vertical polarizations, respectively. $\mathrm{Re}(\cdot)$ denotes the real part of a complex value. The scattering amplitudes $f_{a,b}(0)$ are also calculated based on the T-matrix method as described in Chapter 2. Specific attenuations at horizontal (A_H) and vertical (A_V) polarizations can be calculated by

$$A_{H,V} = 4.343 \times 10^3 \int_0^\infty \sigma_{ext}^{H,V}(D) N(D) dD \quad (\mathrm{dB\ km}^{-1}) \tag{6.11}$$

where $\sigma_{ext}^{H,V}$ is the extinction cross section at horizontal or vertical polarizations. The specific differential attenuation A_{DP} is defined as:

$$A_{DP} = A_H - A_V \quad (\mathrm{dB\ km}^{-1}) \tag{6.12}$$

If specific attenuations are known, the attenuated reflectivity (Z_H^a) and attenuated differential reflectivity (Z_{DR}^a) at each range gate can be calculated by

$$Z_H^a(n) = Z_H(n) - 2\sum_{i=1}^{n-1} A_H(i)\Delta r \tag{6.13}$$

and,

$$Z_{DR}^a(n) = Z_{DR}(n) - 2\sum_{i=1}^{n-1} A_{DP}(i)\Delta r \tag{6.14}$$

where numbers i and n denote the i^{th} and n^{th} range gates from the radar location, respectively. Δr is the range resolution.

6.2.3 Lookup table method

In Eq. 6.8, it is expensive to directly compute the transpose of linearized operator $\mathbf{H}$, which is the matrix of partial derivatives. In general, the adjoint method is applied to compute $\mathbf{H}^T$ efficiently without storing the full matrix. However, this method is not appropriate in this study. Here the calculation of radar variables is based on the pre-calculation of scattering amplitude. Without the approximation (using a empirical relation to model the scattering amplitude), it is difficult to represent the derivatives functionally in terms of DSD parameters. In such a case, it is a problem to develop an adjoint for the calculation of $\mathbf{H}^T$. In order to solve this problem, the lookup table method is applied.

The partial derivative of each polarimetric radar variable, i.e., Z_H, Z_{DR}, or K_{DP}, with respect to each state variable, Λ or N_0', is needed at every grid point. Therefore, there are total six tables of derivative for the observation operator H (i.e., $\partial Z_H/\partial\Lambda$, $\partial Z_{DR}/\partial\Lambda$, $\partial K_{DP}/\partial\Lambda$, $\partial Z_H/\partial N_0'$, $\partial Z_{DR}/\partial N_0'$, and $\partial K_{DP}/\partial N_0'$). In each lookup table, the derivative values are pre-calculated for parameter Λ varying from 0 to 50 and parameter N_0' varying from 0 to 10. To ensure the accuracy, the range of each parameter is discretized at an interval of 0.1. Consequently, each lookup table has

$nv \times 501 \times 101$ elements (nv is the dimension of Λ and N_0'). In this way, the partial derivative of operator H is found from these tables for any given values of Λ and N_0'. The interpolation can be performed for values between the discrete Λ or N_0' to further improve the accuracy. Generally, parameter values in lookup tables are sufficiently accurate for the iterative minimization of cost function because parameter ranges are wide. For state variables out of the table range, they are restricted at the edge of the table range although this rarely happens in practice.

With introducing the lookup table, the cast of derivative calculation is saved. In the similar way, given any two state parameters, the calculation of intrinsic (non-attenuated) Z_H, Z_{DR}, K_{DP}, A_H, and A_{DP} can be made efficiently as well with using the lookup table method. As a result, the observation operator H is computed as a combination of different values found in various lookup tables, avoiding the integral calculation in the forward model. Preliminary results in following sections have demonstrated that the lookup table is an efficient tool to deal with non-linear forward models of complicated functions.

6.2.4 Iteration procedure

The iteration procedure of minimizing the cost function is shown in Fig. 6.1. At the beginning of the program, necessary data files such as all lookup tables, the background, radar measured Z_H, Z_{DR}, K_{DP}, and SNR are loaded. In the mean time, initial parameters of the variational scheme are configured. Radar measurements are then preprocessed. Within the analysis region, only radar measurements with SNR > 1 dB are used. Moreover, observational weights are set differently. The weight (i.e., element of matrix $\mathbf{w}$) is set to 1 for SNR > 20 dB, 1/2 for SNR > 10 dB, 1/4 for SNR > 5 dB and 1/8 for SNR < 5 dB, respectively. With the initial state vector (e.g., set $\mathbf{v}$=0), intrinsic variables, Z_H, Z_{DR}, K_{DP}, A_H, and A_{DP}, are found for each grid

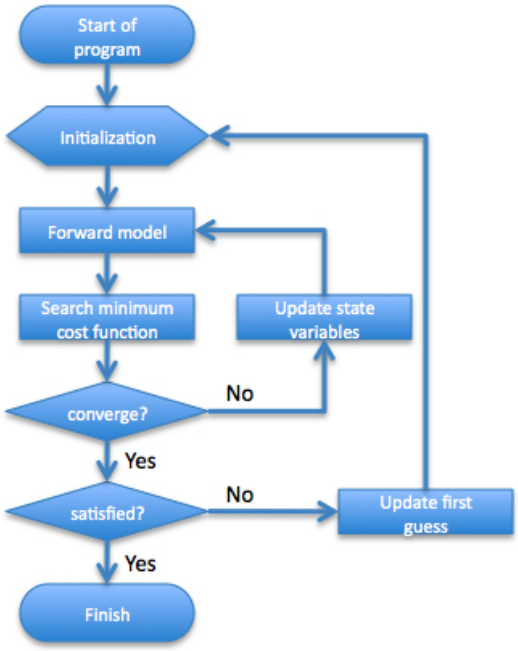

Figure 6.1: Flowchart of variational retrieval scheme.

point from the lookup tables. Corresponding Jacobian matrices (**H**) are constructed based on the lookup tables as well. After the interpolation from grid points to the observation points, attenuated Z_H and Z_{DR} are calculated according to Eqs. 6.13 and 6.14. Calculated polarimetric variables, Z_H, Z_{DR} and K_{DP}, and measured variables are used in Eq. 6.8 to calculate the gradient of cost function. The initial first guess is always assumed to be the background. During the minimization process, the state vector is updated at each loop until the iteration is converged. If the background contains no useful information (e.g., the constant background), the analysis field based on the first guess may not be satisfactory enough. In such a case, the analysis result is considered as a new first guess and used to repeat the minimization process. In general, several outer loops would give the satisfactory result, which has a relatively small cost function.

6.3 Evaluation by Simulation

The advantage of using simulated data is that the truth is known and can be compared to the retrieved result. In this section, the variational approach is evaluated by X-band PRD that are simulated from real measurements of an S-band radar. The S-band measurements come from KOUN radar. It is assumed that the simulated PRD are measured by two CASA IP1 radars [see Xue et al. (2006)], i.e., the radars located at Cyril (KCYR) and Lawton (KLWE) of Oklahoma, which are about 80 and 100 km southwest of the KOUN, respectively.

6.3.1 Simulation of X-band PRD

On 8 May 2007 a convective system passed through Oklahoma from west to east. PPI images of Z_H and Z_{DR} measured by KOUN at 1230 UTC are shown in Fig. 6.2. Two asterisks located at the southwest part of the image denote the locations of KLWE and KCYR. Two 20 km$\times$20 km regions indicated by two square boxes in Fig. 6.2a are analysis regions used to test the variational algorithm. It is worth noting that these two regions include a part of storm core, where the attenuation can be notable at the X-band frequency. The simulation procedure is described as follows. Let's take the simulation of KCYR measurements for an example. Firstly, assume the KCYR makes full 360° azimuth scans at 1° increment. The maximum range is 30 km and the range resolution is 48 meters. Secondly, interpolate KOUN measurements at the lowest elevation into every radar range gate of KCYR, ignoring the effect of radar elevation differences. Thirdly, interpolated Z_H and Z_{DR} are used to retrieve the "true" DSD for each radar range gate, assuming the contribution completely comes from the rain. Next, intrinsic PRD (Z_H, Z_{DR}, K_{DP}, A_H, and A_{DP}) are calculated based on the "true" DSD. After intrinsic PRD are obtained for all range gates, attenuated PRD are then calculated along each beam path. Finally, random noises and biases are added to the attenuated PRD to simulate measurement errors and system biases.

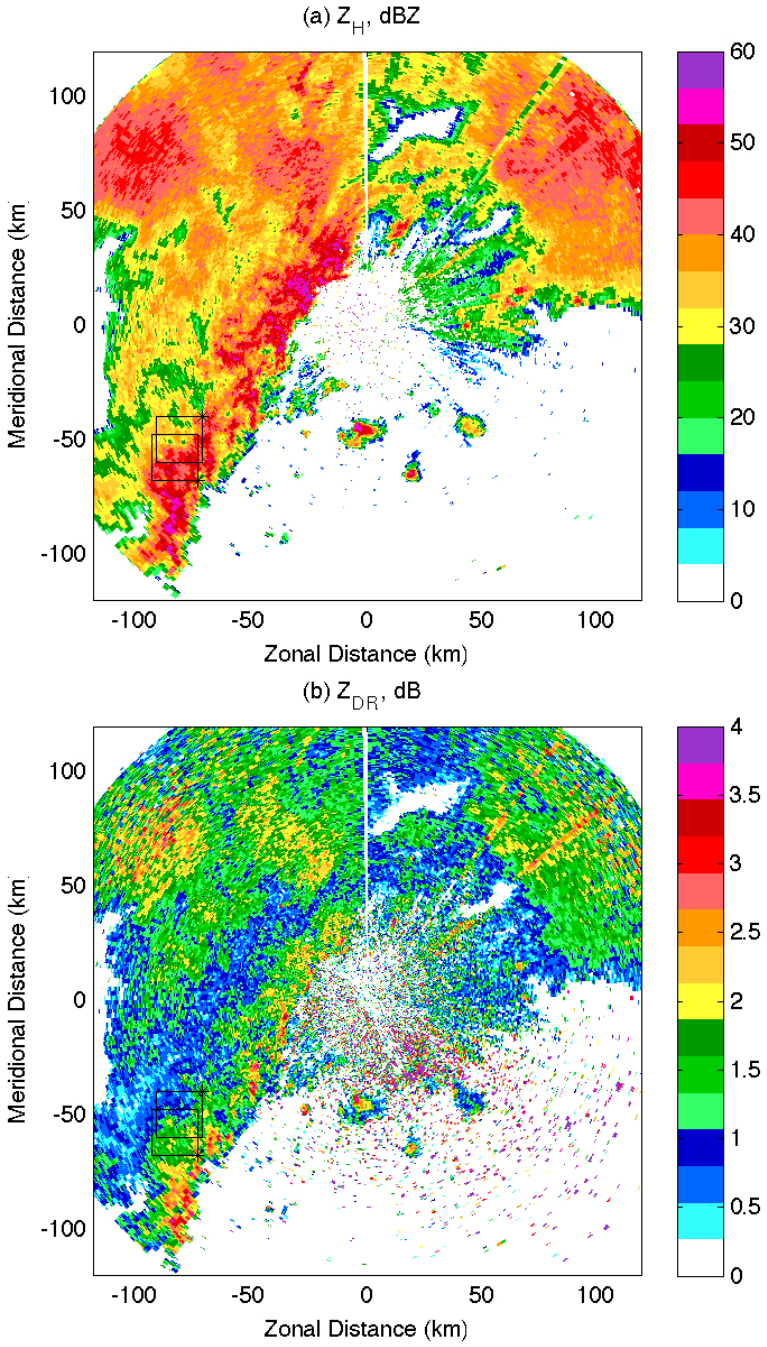

Figure 6.2: (a) Z_H, (b) Z_{DR} measured by KOUN (elevation angle 0.5°, range resolution 250 m, 1230 UTC, 8 May 2007). Two solid line boxes indicate the regions used for the simulation.

6.3.2 Test of algorithm consistency

Does the variational method have a self-consistent retrieval? The following experiments test the performance of the algorithm under ideal conditions, for which only statistical errors are considered while biases and other errors are not added. Measurement errors are assumed to be Gaussian random noises with standard deviations of 2 dB for Z_H (dBZ), 0.2 dB for Z_{DR} (dB), and 0.1 $^\circ\text{km}^{-1}$ for K_{DP} ($^\circ\text{km}^{-1}$), respectively. In the third step of X-band PRD simulation, the "true" DSD is retrieved from two S-band PRD, Z_H and Z_{DR}. The retrieval follows the procedure described in Chapter 4. It is worth noting that the "true" DSD is assumed to follow the C-G model, which is the same as the DSD model used in the variational retrieval. That is to say, when X-band PRD are simulated in this way, there is no DSD model error (it is worth noting that DSD parameters are not known before hand) in the variational retrieval.

The variational analysis applies configurations described as follows. The analysis region is a 20 km×20 km square shown by the box in Fig. 6.2. It is covered by 251×251 analysis points at an interval of 80 meter. The initial background is set to constant values over the whole analysis domain (N_0' is 3 and Λ is 4). These rough guesses may be far from the truth. In the variational scheme, the decorrelation scale L is set to be 20 grids, i.e., 1.6 km, which is reasonable for the spatial property of a storm. Default observation errors are 2 dB for Z_H, 0.2 dB for Z_{DR}, and 0.1$^\circ\text{km}^{-1}$ for K_{DP}, the same as those of simulated observations. Since the background is constant, the background error is set as 2, which is rather a large value.

The first experiment is performed for KCYR radar. Simulated PRD and retrieved results are shown in Fig. 6.3. Three columns from left to right show the images of Z_H, Z_{DR}, and K_{DP}, respectively. Three rows indicate different properties of PRD. The third row denotes the "true" PRD, which are simulated with the C-G DSD model,

i.e., without model error for the variational retrieval. The second row represents the simulated observations. The attenuations have been applied to the simulated observations. Measurement errors have been added to them as well. The first row shows analysis results using the variational algorithm. The input data of the variational algorithm are the simulated PRD shown in the second row.

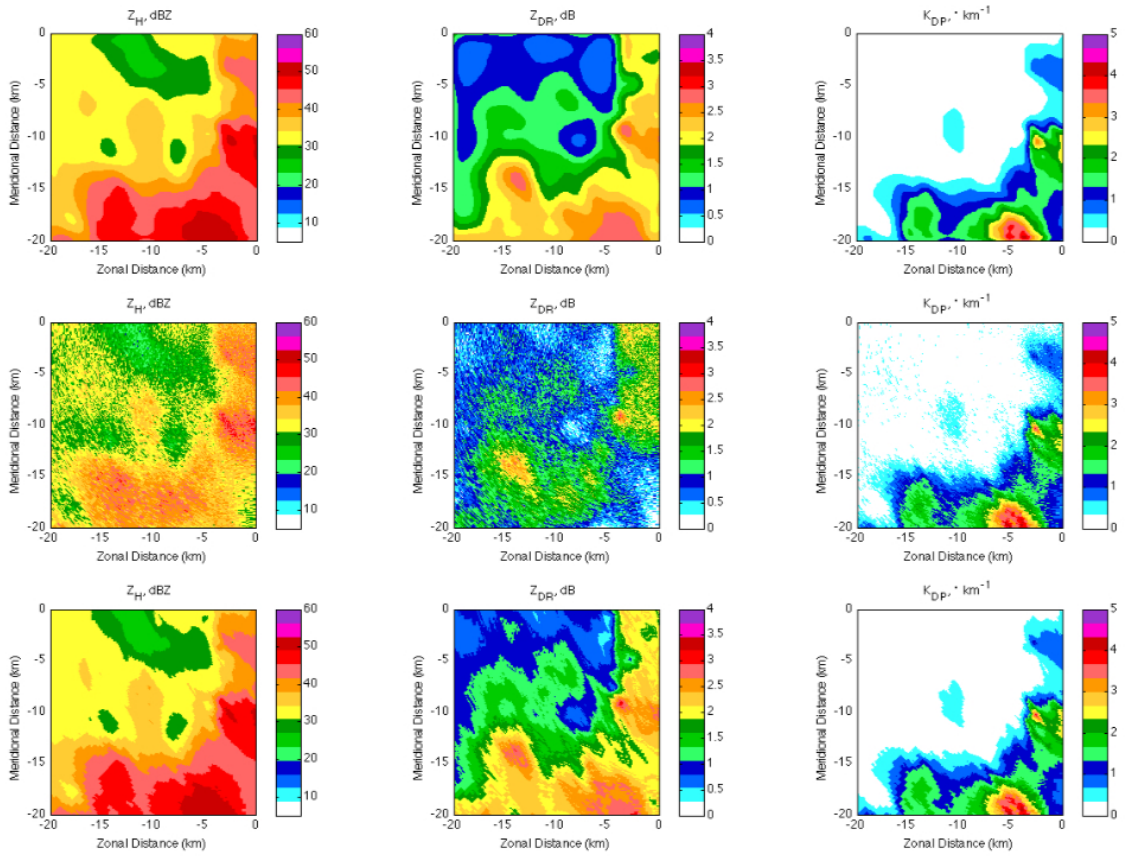

Figure 6.3: Simulated PRD and retrieved results for KCYR. Three rows from top to bottom denote the retrieval results, the simulated PRD (with attenuation effect) and the truth fields, respectively. Three columns from left to right show the Z_H, Z_{DR}, and K_{DP}, respectively. True DSDs are assumed to follow C-G DSD model.

As Fig. 6.3 shows, the variational algorithm successfully retrieves Z_H, Z_{DR}, and K_{DP} even though observed PRD contain attenuations and noises. The analysis results match the truth very well except for some smoothing. The true PRD are interpolated

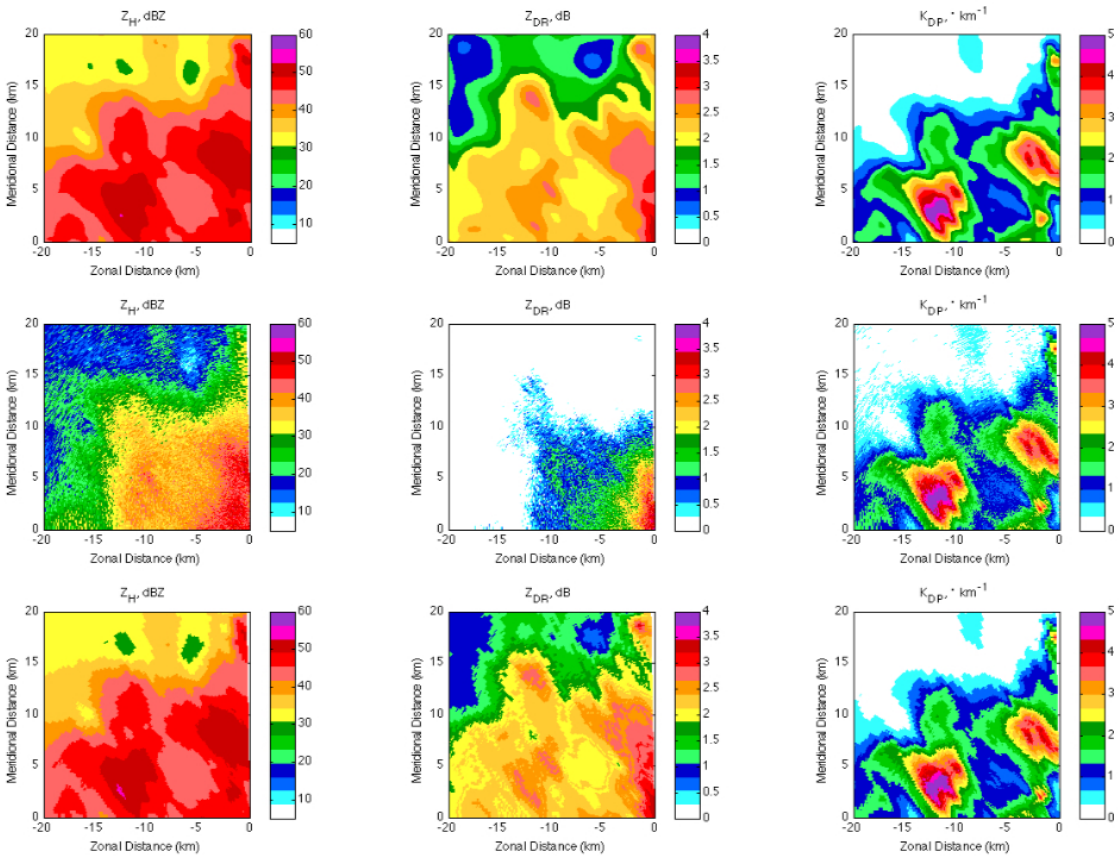

Figure 6.4: The same as Fig. 6.3 but for KLWE

into the grid points and compared to the analysis results. The biases of retrievals with respect to the true PRD are 0.11 dB, 0.01 dB and less than 0.001 $^{\circ}\text{km}^{-1}$ for Z_H, Z_{DR}, and K_{DP}, respectively. Accordingly, RMSEs of retrieval are 0.47 dB, 0.10 dB and 0.06$^{\circ}\text{km}^{-1}$. These results demonstrate the excellent performance of the variational algorithm in a perfect condition, i.e., with controlled measurement errors and without DSD model errors and biases. Moreover, the lookup table method, as well as adaptive attenuation correction integrated in the forward model, has been proven to be effective in this situation.

Similar analysis is performed for simulated PRD of KLWE. The results are shown in Fig. 6.4. In the analysis region, there are heavy rains around the radar so that the attenuation effect is more severe than the KCYR case shown in Fig. 6.3. This is obvious from the second row of Fig. 6.4. Simulated Z_H and Z_{DR} have a very low value in the far distance. The strong attenuation close to the radar can negatively affect the minimization process of cost function because the retrieval at far range is sensitive to the attenuation correction at near range. However, the variational algorithm still gave nearly perfect results. The biases of retrieval in Fig. 6.4 are 0.13 dB, 0.01 dB and 0.006$^{\circ}\text{km}^{-1}$. The RMSEs of retrieval are 0.40 dB, 0.09 dB and 0.07$^{\circ}\text{km}^{-1}$. In the overlapping region of KCYR and KLWE radars, Figs. 6.3 and 6.4 show a good match for all three radar variables.

6.3.3 Test of DSD model error

Because natural DSDs are not necessarily represented by the C-G DSD model, this subsection tests the effect of DSD model error on the variational algorithm. During the simulation, the true DSD is assumed to follow the exponential model instead of the C-G DSD model. There is an evident difference between the C-G model and

the exponential model. The exponential model is equivalent to a gamma model with a shape parameter μ=0 while the C-G model is the gamma model with its shape parameter depending on the slope parameter. Since the variational algorithm applies the C-G model, simulated X-band PRD using exponential model can bring a notable DSD model error.

The simulation and retrieval procedures are similar to those described in previous subsection. The same S-band data are used so that the effect of model error can be perceived through the comparison. Corresponding results are shown in Fig. 6.5 for KCYR and Fig. 6.6 for KLWE. Although the true PRD in Fig. 6.5 (or Fig. 6.6) do not have much difference with those in Fig. 6.3 (or Fig. 6.4), the intrinsic DSD are different. However, both figures show that the retrieval results (first row) still match the truth (third row) very well even though DSD model errors have been introduced. For Fig. 6.5, the biases of retrieval are 0.09 dB, 0.01 dB, and 0.001°km^{-1} for Z_H, Z_{DR}, and K_{DP}, respectively. The RMSEs of retrieval are 0.46 dB, 0.10 dB and 0.06 °km^{-1}. For Fig. 6.6, the corresponding biases are 0.16 dB, 0.03 dB, and 0.006 °km^{-1}. The corresponding RMSEs are 0.46 dB, 0.11 dB and 0.08 °km^{-1}. Compared to the biases and RMSEs in Figs. 6.3 and 6.4, there are no fundamental difference from them. That is to say, the performance of this variational algorithm has not been notably deteriorated with the addition of DSD model error. It implies that the assumption of C-G DSD model should be reasonable and practicable for this variational algorithm.

6.3.4 Test of measurement error and bias

There are two considerations for practical application of this algorithm. First, measurement errors are always unknown and might be overestimated/underestimated in the observation error covariance matrix. The other consideration is that there might

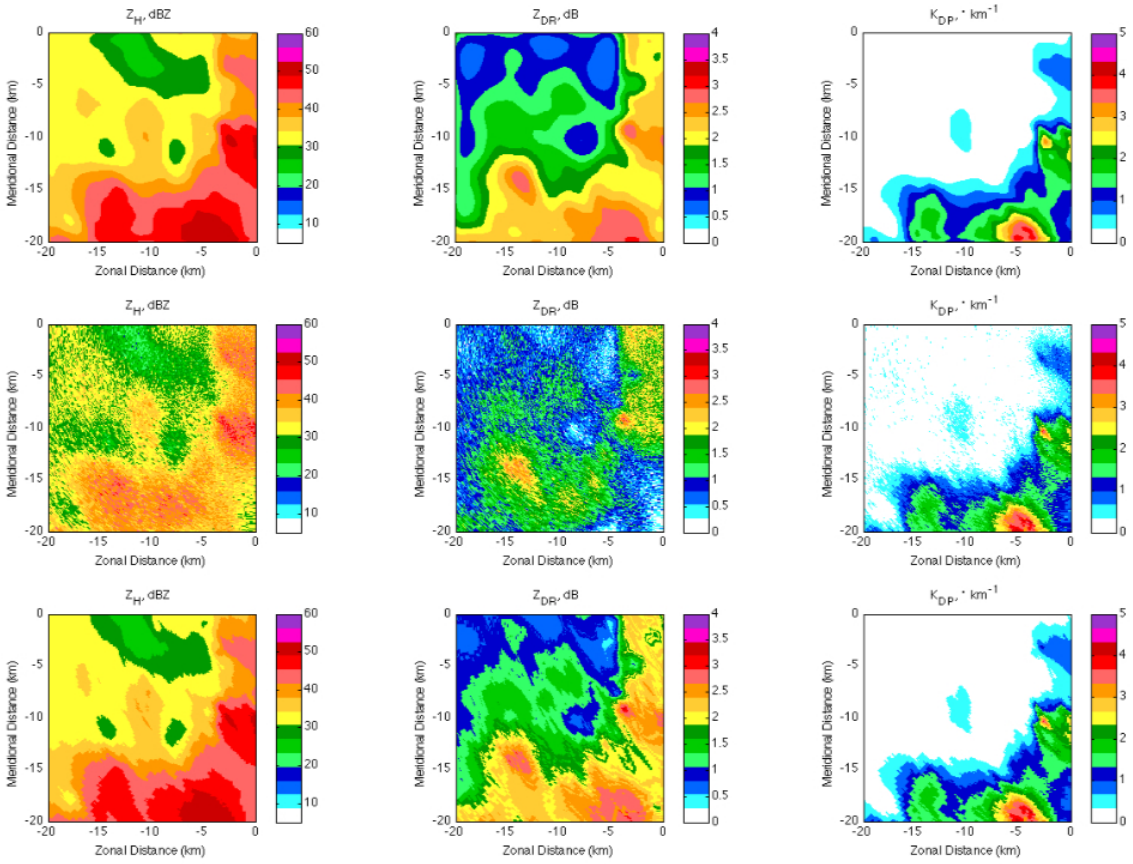

Figure 6.5: The same as Fig. 6.3 but true DSDs are assumed to follow exponential DSD model

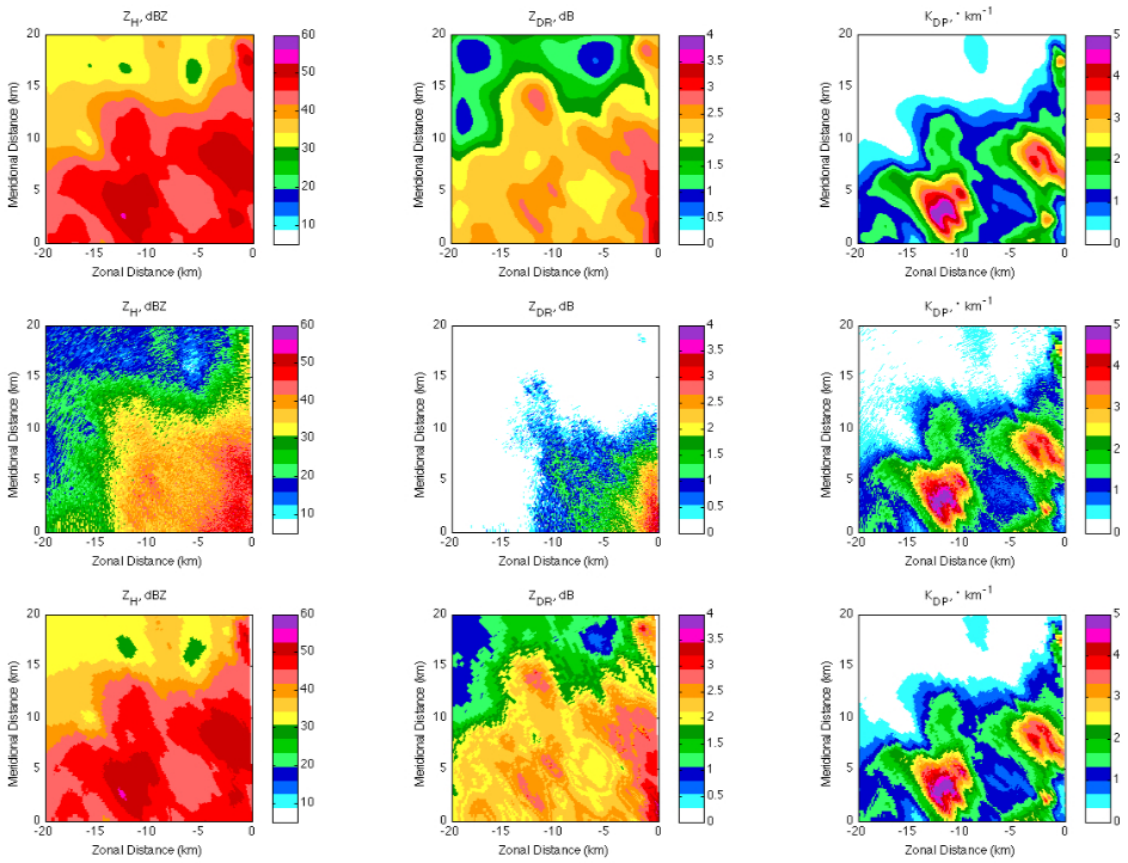

Figure 6.6: The same as Fig. 6.5 but for KLWE

exist intrinsic biases attributed with the forward model error, which leads to the difference between radar measurements and observational model outputs. This subsection examines the algorithm's sensitivity to the error and the bias that can not been well characterized in the retrieval.

The following tests are based on the simulated "truth" shown in Fig. 6.6. It is worth noting that the DSD model error exists for all the following tests. To simulate the "observation", the "true" radar parameters are added with different measurement errors and biases for a total 12 tests. Tests 1-4 assume no bias but different measurement errors for simulated "observations". Tests 5-8 assume the same measurement errors but different biases. Tests 9-12 assume the same biases as tests 5-8 except that measurement errors are different. The detailed configurations of simulated data and statistics of retrievals are shown in table 6.1. In each cell of the table, right values of notation "slash" are simulated biases/measurement errors. Left values are biases/RMSE of retrievals compared to the truth. In the ideal condition, measurement errors should be characterized by error covariance matrices ($\mathbf{R}_{Z_H}$, $\mathbf{R}_{Z_{DR}}$, $\mathbf{R}_{K_{DP}}$) of the variational scheme. However, it is impossible because true errors are always unknown. Therefore, a mismatched error structure is assumed for these tests. For those error covariance matrices, the standard deviations of Z_H, Z_{DR}, and K_{DP} are assumed to be 0.5 dB, 0.1 dB and 0.1 $^{\circ}\mathrm{km}^{-1}$, respectively. That is to say, the "measurement" errors only match the "truth" in test 1. In other tests, the "true" error are generally larger than the error assumption in the retrieval.

As to tests 1-4, the retrieval RMSEs are generally less than the "true" measurement errors. This means the algorithm has the capability to smooth the observation with a less statistical error. However, the measurement error may introduce the retrieval bias and the bias increases with increasing the measurement error. It is noted

that the bias and the RMSE are not large, implying that the statistical error does not have much effect on the retrieval if there is no intrinsic bias in the measurement.

Table 6.1: Bias and RMSE of variational retrieval for different experiments

	Retrieval bias / Simulated bias			Retrieval RMSE / Simulated error		
TEST	Z_H (dB)	Z_{DR} (dB)	K_{DP} ($^\circ$km^{-1})	Z_H (dB)	Z_{DR} (dB)	K_{DP} ($^\circ$km^{-1})
1	0.091 / 0	0.027 / 0	0.004 / 0	0.393 / 0.5	0.107 / 0.1	0.084 / 0.1
2	0.083 / 0	0.009 / 0	0.006 / 0	0.409 / 1.0	0.108 / 0.2	0.083 / 0.2
3	0.178 / 0	0.023 / 0	0.012 / 0	0.476 / 1.5	0.110 / 0.3	0.088 / 0.3
4	0.267 / 0	0.036 / 0	0.019 / 0	0.537 / 2.0	0.120 / 0.4	0.093 / 0.4
5	0.444 / 0.125	0.115 / 0.025	0.020 / 0.025	0.597 / 0.5	0.159 / 0.1	0.084 / 0.1
6	0.841 / 0.25	0.219 / 0.05	0.037 / 0.05	0.952 / 0.5	0.253 / 0.1	0.092 / 0.1
7	1.575 / 0.5	0.411 / 0.1	0.067 / 0.1	1.687 / 0.5	0.445 / 0.1	0.114 / 0.1
8	2.879 / 1.0	0.755 / 0.2	0.118 / 0.2	3.037 / 0.5	0.807 / 0.1	0.160 / 0.1
9	0.448 / 0.125	0.113 / 0.025	0.022 / 0.025	0.606 / 0.75	0.157 / 0.15	0.085 / 0.15
10	0.862 / 0.25	0.216 / 0.05	0.040 / 0.05	0.979 / 1.0	0.250 / 0.2	0.095 / 0.2
11	1.604 / 0.5	0.408 / 0.1	0.071 / 0.1	1.724 / 1.25	0.443 / 0.25	0.117 / 0.25
12	2.940 / 1.0	0.747 / 0.2	0.122 / 0.2	3.117 / 1.5	0.801 / 0.3	0.165 / 0.3

Tests 5-8 have the same measurement errors as test 1 except they have different biases. Compared to test 1 results, tests 5-8 show notable biases and RMSEs in retrieval resutls. Except some values of K_{DP}, all retrieval biases or RMSEs are larger than simulated biases or errors in tests 5-8. Test 8 shows that 1 dB bias in Z_H measurements leads to about 3 dB bias and 3 dB RMSE in Z_H retrievals. This fact implies that the variational algorithm is more sensitive to the measurement bias than to the measurement error. The measurement bias not only introduces a larger bias in the retrieval but also enlarges the retrieval RMSE. Moreover, the large the measurement bias, the larger the retrieval bias and RMSE. Test 9-12 apply the similar simulation to tests 5-8 except measurement errors are different. For example, simulated data in test 12 have three times measurement error as large as in test 8.

However, retrieval biases and RMSEs of test 12 are almost the same as those of test 8. This fact also demonstrates that the algorithm's sensitivity to the measurement bias is greater than to the measurement error.

These 12 tests give a good understanding of the algorithm's sensitivity. In practice, the situation would be much more complicated. For example, measurement error might not be equal for every measurement. However, this would not be a serious issue according to the previous analysis. The serious problem might exist when the radar data are inconsistent with each other. For example, according to the radar forward model used in the retrieval algorithm, three parameters, Z_H, Z_{DR}, and K_{DP}, should be intrinsically consistent. Any inconsistency is equivalent to introducing measurement biases, which might lead to large biases and RMSEs in the retrieval. Moreover, the data inconsistency might not exist everywhere equally. For example, radar measurements might not be reliable due to low SNRs. Within low SNR regions, measurement biases and errors might be very large while they might be small in other regions. The performance of the variational algorithm would also degrade in such a situation. With the understanding of algorithm's sensitivity, the next subsection will show some results from the real radar data and discuss corresponding issues for practical implementation of the algorithm.

6.4 Retrieval Based on Real PRD

The previous section tests the variational algorithm using simulated data and shows promising results. The simulations also show that the performance of the algorithm depends on the data quality. It makes sense because previous tests applies a constant background, which contains useless information for the retrieval. The simulated data have a good data quality so that the background information is not necessary for the

retrieval to get a good result. However, the data quality is a major problem when this algorithm is applied with the real data. As long as the observations have a bad data quality, the retrieval would not be able to get a reasonable result. In that case the useful information is required to compensate the degradation of the data quality. The rest part of this section gives two cases of real data retrievals to address this issue.

6.4.1 Case 1: X-band data

This case applies the X-band data collected by two CASA radars described in the previous section. Figs. 6.7 and 6.8 show the PPI images of Z_H, Z_{DR}, K_{DP}, and SNR, measured by KCYR and KLWE, respectively. The square boxes in the figures represent the analysis region of retrieval. As Fig. 6.7d shows, there exist some regions with a low SNR (<10 dB) in the analysis area. Particularly, the low SNR region is much larger in the the analysis area of KLWE (Fig. 6.8). Within the low SNR region, the data quality of radar data is problematic. Z_{DR}, which is noisy in both figures, obviously has a worse quality than Z_H. K_{DP}, having many positive and negative values, is least reliable among three variables. To mitigate the effect of poor data quality, some observational weighs in terms of SNR have been introduced in the cost function (as shown in Eq. 6.1). In addition, the default measurement errors are set to 2 dB for Z_H(dBZ), 0.4 dB for Z_{DR}(dB), and 0.2 $^\circ$km^{-1} for K_{DP} ($^\circ$km^{-1}), respectively. The measurement errors for Z_{DR} and K_{DP} are twice as much as those assumed in the simulation case because for real data Z_{DR} and K_{DP} usually have more uncertainty than Z_H. It can be obviously seen that the real Z_H and Z_{DR} in Figs. 6.7 and 6.8 are much noisier than simulated Z_H and Z_{DR} in Figs. 6.3 and 6.4.

Let's first test the constant background in the real data retrieval. The background parameters are set to be the same values as in previous tests (i.e., $N_0' = 3$ and $\Lambda = 4$).

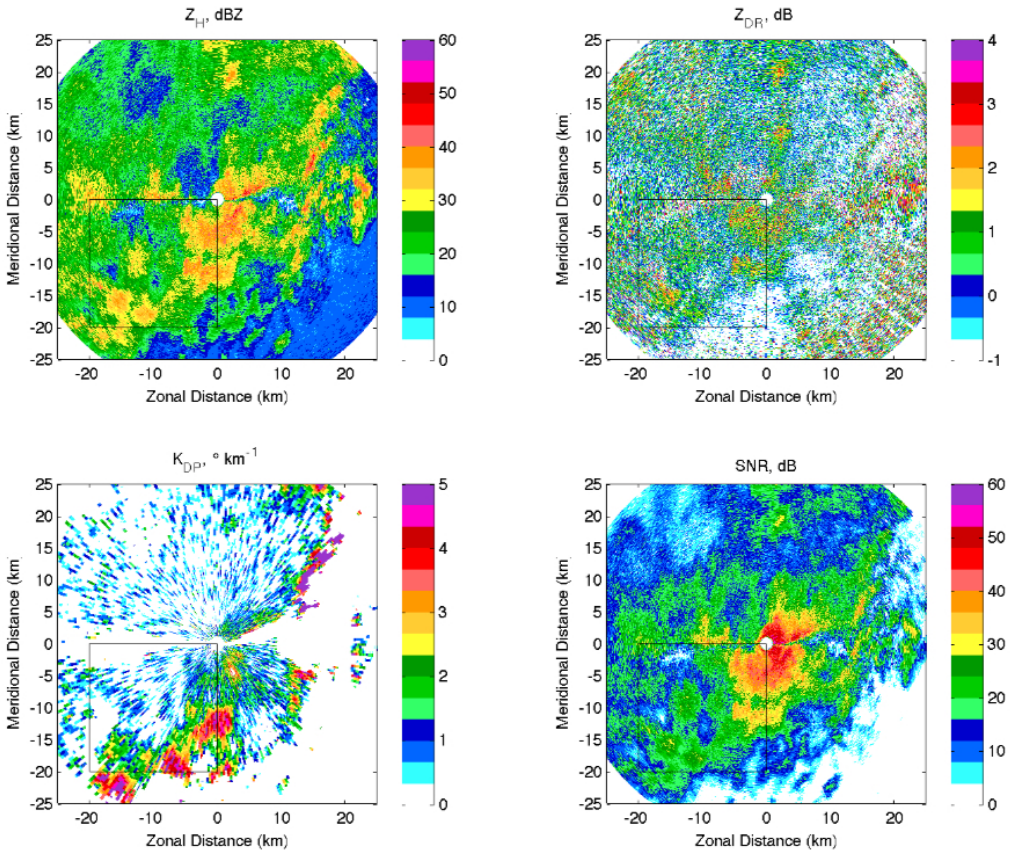

Figure 6.7: (a) Z_H, (b) Z_{DR}, (c) K_{DP}, and (d) SNR as measured by KCYR at the elevation angle of 2° at 1230 UTC on May 8th, 2007. The square box region is the retrieval domain.

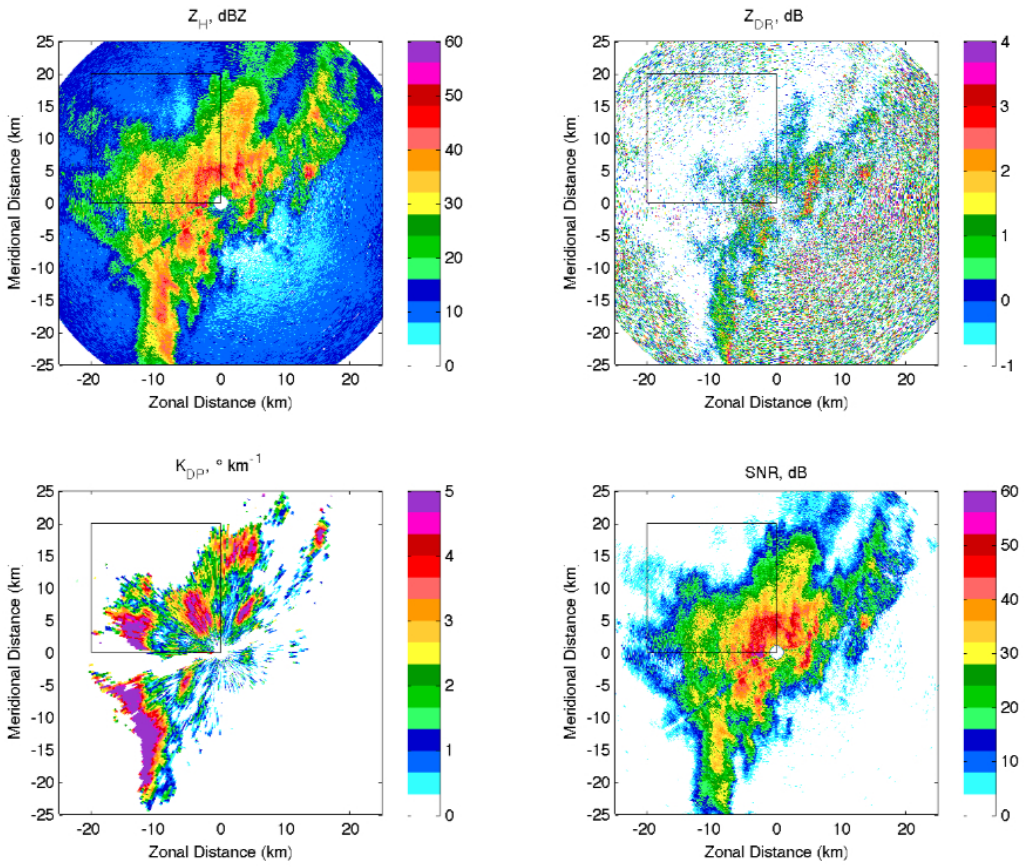

Figure 6.8: The same as Fig. 6.7 but data were measured by KLWE.

The background error is set to 4, denoting a large dynamic range of N_0' or Λ. The retrieval results are shown in Figs. 6.9 and 6.10 for KCYR and KLWE, respectively. Unfortunately, because the constant background cannot provide any helpful information within the region of low SNR, the retrieval results there are bad. Moreover, because the attenuation correction process makes the far range retrieval have a substantial dependence on the near range retrieval, the low SNR region could actually affect the retrieval almost at the entire region. This impact is evident in Figs. 6.9 and 6.10, especially for Z_{DR} and K_{DP}. The variational algorithm failed in this experiment when a constant background is applied. It is reasonable to have such a result because the variational retrieval is a global optimization system. If satisfactory retrieval were desired, good physical information (no matter from data or background) of the entire region should be provided. Otherwise, incorrect retrieval at one point might happen. Its negative effect could be spread through spatial correlation and incorrect attenuation correction, resulting in potential degradation of the entire system.

The S-band radar measurements can be an additional source in providing useful information to compensate for the bad data quality of X-band radar measurements. The following experiment applies the same data but a different background obtained from S-band radar measurements. In simulation section, Z_H and Z_{DR} measured by KOUN have been used to simulate a "truth" field (Figs. 6.3 and 6.4, the third row). Here the simulated "truth" field is used as the background. Generally, the S-band measurements should be close to the truth though there exist model error and measurement error effects. As a result, using S-band based background should have a smaller background error than using constant background. In this experiment the background error is set to 0.5, representing a moderate error. For example, given the same Λ, N_0' error of 0.5 introduces 5 dB error for Z_H.

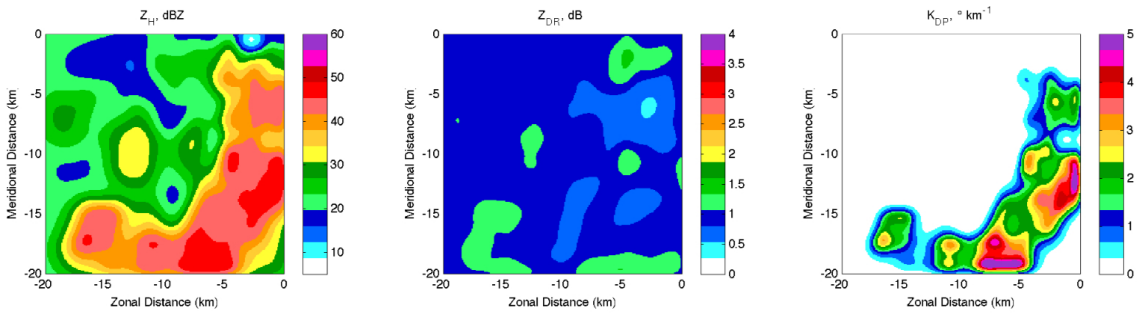

Figure 6.9: Retrieved results based on KCYR radar measurements. The background was set to be constant. From left to right: (a) Z_H, (b) Z_{DR}, (c) K_{DP}.

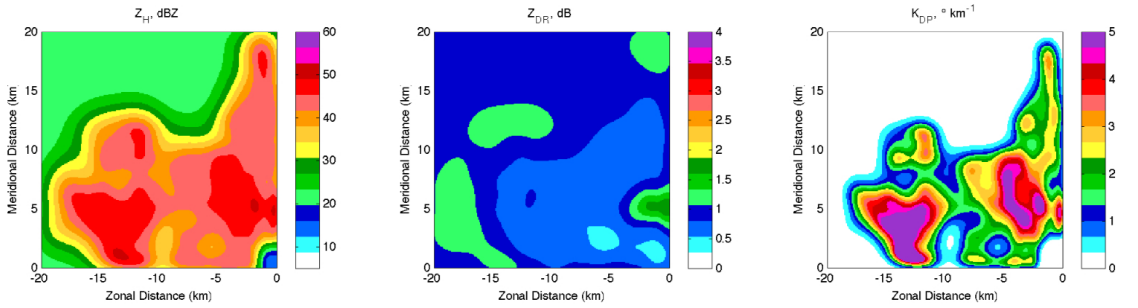

Figure 6.10: The same as Fig. 6.9 but for the retrieval of KLWE radar measurements.

The variational retrieval results are shown in Figs. 6.11 and 6.12 for KCYR and KLWE, respectively. As expected, the background has compensated for the X-band data with a low quality so that the performance of the variational algorithm is stable and satisfactory. Since we do not know the truth, the reasonableness can be examined by comparing the results of two radars at the overlapped region. As Figs. 6.11 and 6.12 show, the major features of all three radar varaibles match very well at the overlapped region (refer to overlapped region of two square boxes in Fig. 6.2). In addition, compared to background images (Figs. 6.3 and 6.4, the third row), Figs. 6.11 and 6.12 show more details. The details are due to the fact that the X-band data have better range resolution and have contributed to the retrieval. The detailed features of two retrievals also have a good match. Considering two X-band radars are two independent sources, this result convinces the validity of the variational retrieval

algorithm introduced in this study.

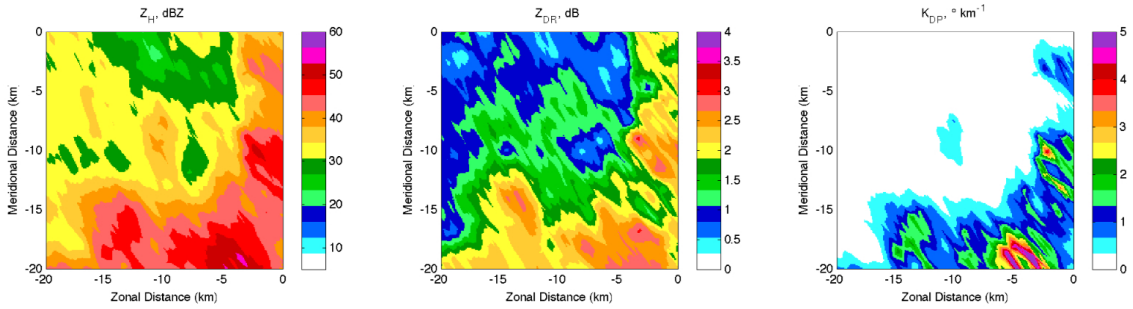

Figure 6.11: Retrieved results based on KCYR radar measurements. The background was based on the retrieval of S-band radar (KOUN) measurements. From left to right: (a) Z_H, (b) Z_{DR}, (c) K_{DP}.

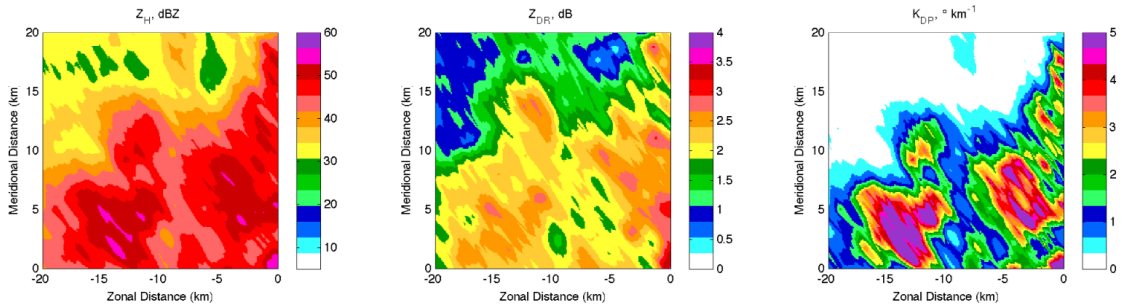

Figure 6.12: The same as Fig. 6.11 but for the retrieval of KLWE radar measurements.

6.4.2 Case 2: C-band data

The second case is the C-band radar retrieval. Fig. 6.13 shows PPI images of four variables measured by OU-PRIME on 12 April 2009. The elevation angle of radar PPI scan is 1.3°. This case is a stratiform precipitation event, having a melting layer located at the height of 1.8-3 km. Within 80 km distance, the radar echoes generally come from raindrops. The precipitation is quite quasi-uniform and has a narrow spectrum width of 0.5-2 m s^{-1}. The square box region in Fig. 6.13, which is 50×50 km^2 region, is chosen for the variational analysis.

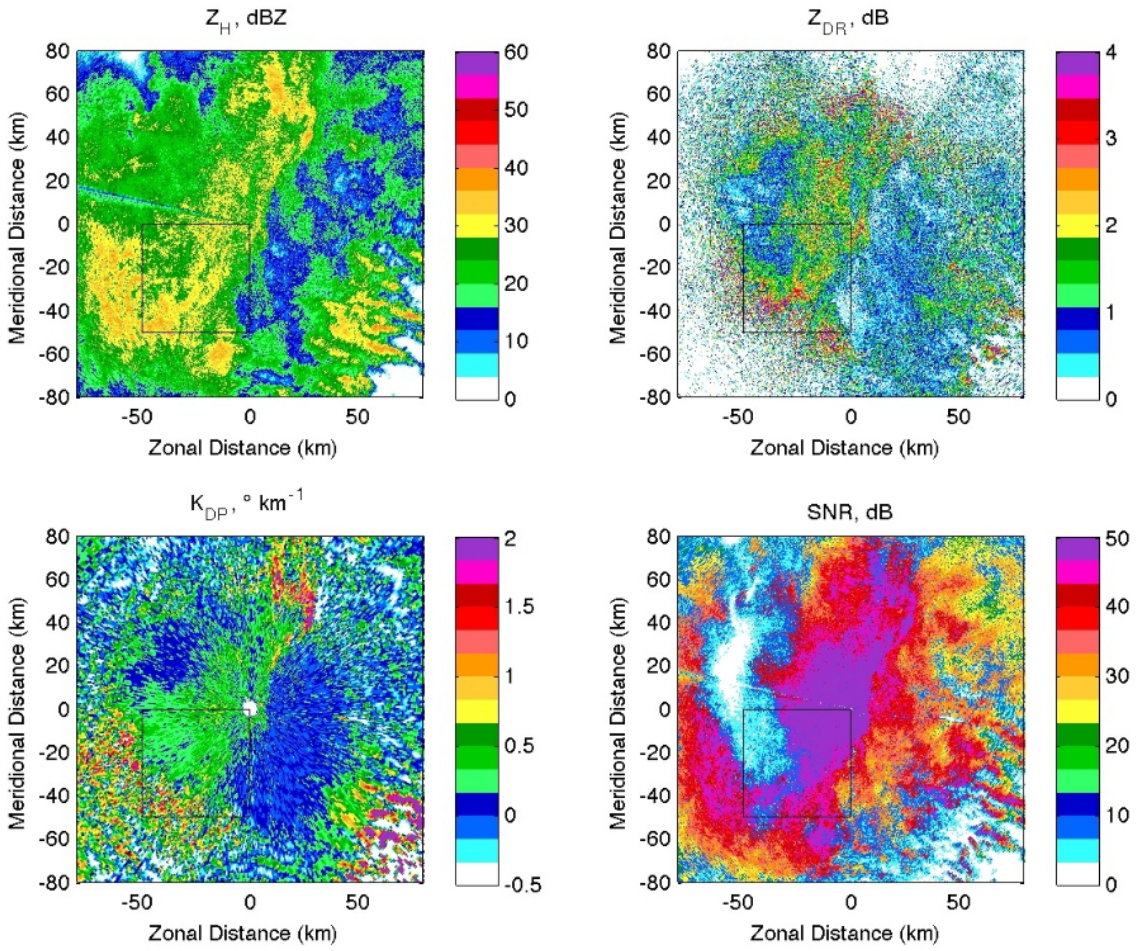

Figure 6.13: (a) Z_H, (b) Z_{DR}, (c) K_{DP}, and (d) SNR as measured by OU-PRIME on 12 April 2009, 1.3°, 0704 UTC. The square box region is the retrieval domain.

It is noted that the northwest part of the analysis region has a low SNR less than 10 dB. The retrieval using a constant background can not output satisfactory results (not shown). To improve the retrieval, additional information is required to use. The NEXRAD radar measurements is a possible source for the supplement of C-band radar data. The advantage is that NEXRAD has a wide coverage and its data can be available any time. The limitation is that it only has the single-polarization data, which are insufficient to provide good background with two DSD parameters. Fig. 6.14 shows the reflectivity measured by KTLX radar of NEXRAD. Its axis has been set to be the analysis region shown in Fig. 6.13. There is a problem with using Z_H to estimate the background because the background requires two DSD parameters for each grid. In the following variational retrieval, the M-P DSD model is assumed to estimate the background from the Z_H of KTLX. It is worth noting that using M-P parameters as C-G parameters obviously introduces the background error. However, it is expected that this procedure could, at least partially, bring physical information and contribute to the retrieval at low SNR regions.

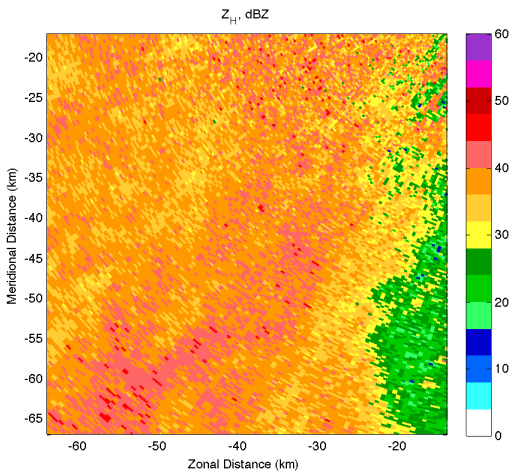

Figure 6.14: Z_H measured by KTLX on 12 April 2009, 0.5°, 0704 UTC.

Fig. 6.15 gives retrieval results as well as radar measurements in the analysis region. Although the results can not be verified, the retrieval looks reasonable. Attenuations in measured Z_H and Z_{DR} have been corrected in the retrieval. C-band has a shorter wavelength than S-band, therefore, retrieved Z_H has a little higher value than KTLX's Z_H. Retrieved K_{DP} matches observed K_{DP} well but has a better smoothing effect. More important thing is that three retrieved variables are consistent with each other according to the scattering theory. The retrieval results should be satisfactory given these radar measurements. Limitations of this variational algorithm will be discussed in the next section.

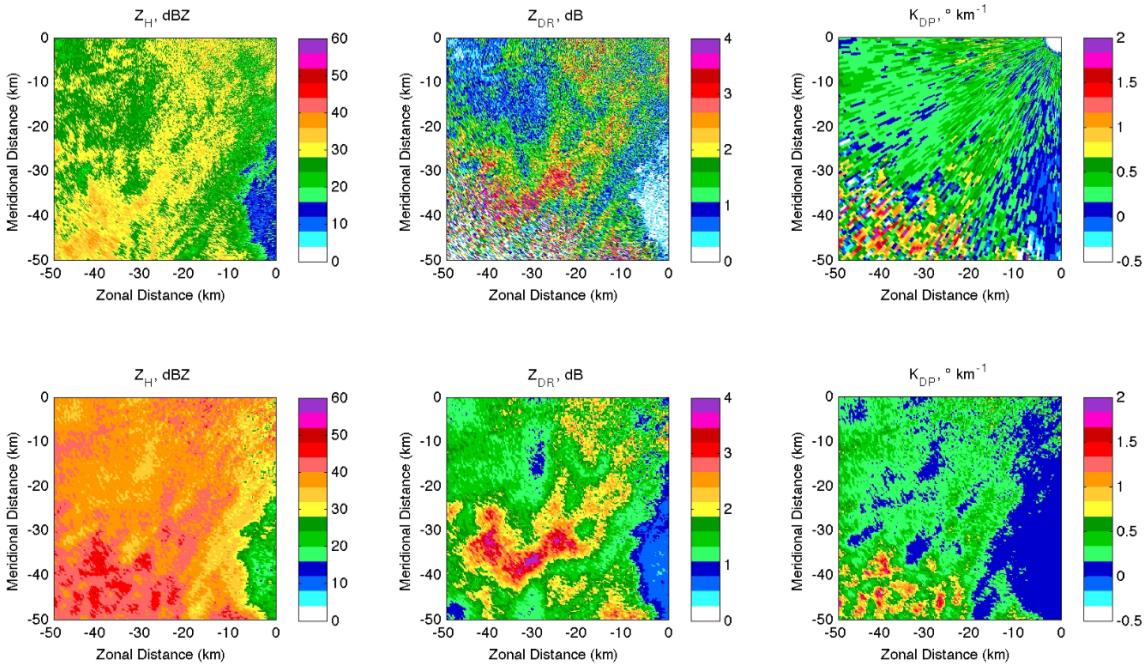

Figure 6.15: Retrieved results (lower row) based on OU-PRIME radar measurements (upper row). The background is based on Z_H measured by KTLX. From left to right: (a) Z_H, (b) Z_{DR}, (c) K_{DP}.

6.5 Discussion

This chapter proposes a variational retrieval algorithm based on attenuated radar measurements. The C-G DSD parameters are treated as the state variables in the variational scheme. Three radar variables (Z_H, Z_{DR}, and K_{DP}) are optimized to correct the attenuation and do the retrieval by mitigating the effect of their measurement errors. The proposed lookup table method is demonstrated effective for the computation of complicated forward model and its partial derivatives. Preliminary results based on simulated and real radar data show the effectiveness of this variational algorithm. The possible error sources for the variational retrieval algorithm may come from following factors.

- The major source of the uncertainty comes from the data quality. According to the analysis of simulation and real data, radar data with low SNRs would deteriorate the retrieval remarkably if there were no useful information to correct them. At the region where the data quality is poor, the background is required to provide complementary physical information.

- The second one is the forward model. As to the variational approach, observations and analysis fields are connected through the forward model. In this study, the forward model is based on the backscattering theory of raindrops. There are two possible situations. One is that the real radar measurement of raindrop do not match the forward model output. The other one is that the radar might measure other species such as snow and hail. In either situation, the forward model would have a model error in the retrieval. This kind of error acts like introducing a bias in the radar measurement and lead to a large bias and RMSE in the retrieval.

- The third source is the model error associated with the DSD model. The variational algorithm treats the two parameters of C-G DSD model as state variables.

In reality, assumption of the C-G DSD may not be valid and is a possible source of error. However, simulation results using exponential DSD suggest that the algorithm is not very sensitive to the assumed DSD model. The analysis of integral parameters (e.g., Z_H, Z_{DR}, and K_{DP}) is robust. The retrieval using real data also implies this suggestion when radar data quality is good or compensated by a reasonable background (e.g., S-band measurements).

- Another source of error is the estimation of error spatial structure. The true magnitudes and correlations of the error covariance are never exactly known. In this algorithm, the spatial structure of background error covariance is modeled by a two-dimensional isotropic Gaussian function. Each measurement error covariance matrix is assumed to be diagonal with an empirical standard deviation. According to the tests, it seems that these two models work for the modeling of the background error and the measurement error. The tests also show that the algorithm is less sensitive to the measurement error than to the measurement bias although the measurement error would also introduce a bias and RMSE in the retrieval.

The algorithm introduced in this chapter, for the first time, retrieves DSD parameters of a two-parameter model in a variational scheme. Previous studies of variational rain retrieval generally assumed a simple DSD model [e.g., Jung et al. (2008a) applied a M-P model] or its equivalence [e.g., Hogan (2007) applied a $Z - R$ relation]. The two-parameter DSD model gives more flexibility to model the natural DSD, improving the estimation of attenuation and the succedent retrieval. However, a complicated DSD model bring another issue. In addition to the rain, radar might measure the snow, the hail or their mixture. More state parameters are required in the variational scheme to distinguish these species. Consequently, additional observations are needed to retrieve these unknown parameters. In this algorithm, only three radar observations are available, restricting the number of state parameters. For example,

if the scheme includes three species of precipitation and uses a two-parameter model for each species, at least six observations are required. It is difficult to obtain so many reliable observations from one or two radars. The physical variability should be sacrificed if multiple phases of precipitation were taken into account. So far, multiple sources rather than radar observations have been utilized in a variational scheme [e.g., Jung et al. (2008a,b)]. However, this topic is related to the data assimilation and beyond the scope of current study, which focuses on the radar retrieval.

According to the previous analysis, the major barrier for the implementation of this algorithm is the data quality. This problem can be overcome by applying an appropriate background. In this study, different backgrounds have been tried. Case one uses S-band dual-polarization observations to generate the background. This background has less errors than the one used in case two, which is obtained from NEXRAD radar observations. However, considering the NEXRAD has a national coverage, the latter case should have more practical significance. It is known that the NEXRAD is being upgraded to the dual-polarization capacity. The updated NEXRAD will provide a good data source in the variational scheme. In such a case, the data quality issue is no longer a problem for the application of this variational DSD retrieval algorithm.

Chapter 7

Conclusions

7.1 Summary

In this dissertation, the rain microphysics in central Oklahoma has been well investigated based on the long-term 2DVD observations. Useful microphysical characterizations, the C-G DSD model, and retrieval algorithms are developed for applications of polarimetric radar data in the QPE and QPF. This work is briefly summarized as follows.

- Observation errors of 2DVD are quantified. The low-end and high-end of DSD observed by the 2DVD have larger error than the middle part of the DSD, which results in the middle moments having relatively lower errors than other moments.

- Estimations of DSD parameters are analyzed. In term of estimating integral rain parameters, the moment estimator M246 is better than other estimators, such as M012, M234, ML, and LM, though it has unavoidable estimate biases for DSD parameters.

- The C-G DSD model is re-analyzed and refined. Observed DSDs can be well modeled by the C-G model. The constraining $\mu - \Lambda$ relation represents the physical truth of rain properties although it is sensitive to the observation error.

The SATP method improves the development of $\mu - \Lambda$ relation with less error effect.

- Rain microphysics is characterized for S-, C- and X-band radar applications. The empirical relations of rain-radar variables are useful in the QPE and QPF.

- The DSD retrieval proposed by Zhang et al. (2001) is re-studied. It is illustrated again that the C-G DSD model can be successfully used in the DSD retrieval from polarimetric radar observations. An adjustment of the $\mu - \Lambda$ relation improves the retrieval (e.g., N_T and D_0) at the leading edge of convective storm.

- A Bayesian approach of DSD retrieval is proposed. The DSD retrieval combined with the Bayesian approach provides an estimate of retrieval uncertainty, aiding in the understanding of data quality of radar data and the reliability of the retrieval. The validity of this algorithm is demonstrated by the quantitative comparison between in-situ measurements and radar retrievals.

- A variational algorithm for DSD retrieval is introduced. For the first time, two DSD parameters are retrieved from a variational method when applied to attenuated radar observations. The preliminary results are encouraging. The attenuation correction can be computed iteratively through the optimal use of observations. However, the current method significantly depends on whether the useful information covers the whole region of analysis.

7.2 Major Achievements

The first achievement in this study is the analysis of the C-G DSD model in depth. As we know, accurate radar-rain retrieval requires an appropriate model to represent DSDs. A three-parameter model, though more flexible, is seldom applied because of the trade-off between the model complexity and the reliability of the retrieval.

In recent studies, the C-G DSD model, a two-parameter model with a constraining $\mu - \Lambda$ relation, has been shown to be superior to the traditional M-P model or an exponential model. However, suspicions still exist about the validity of the $\mu - \Lambda$ relation. This study quantifies the measurement error of 2DVD and analyzes the error effect on the development of the $\mu - \Lambda$ relation. The results reveal a large physical significance of the $\mu - \Lambda$ relation, providing the validity of C-G DSD model. The refined $\mu - \Lambda$ relation in this study also facilitates the application of C-G DSD model in the radar-rain retrieval.

The second achievement is the development of forward radar observation operators and their applications to microphysical characterization and radar-rain retrieval. Based on the disdrometer-radar comparison study, the raindrop scattering model is demonstrated as an efficient tool to estimate radar parameters. As a result, rain microphysics can be characterized by radar variables. This study gives the empirical radar-rain relations for the S-, C- and X-band, providing an efficient/convenient approach for weather radar applications.

The third achievement is the study of DSD retrieval with polarimetric radar measurements. Recently the DSD retrieval has attracted a lot of research interests in radar meteorology community. This study introduces three DSD retrieval methods. The direct retrieval method is straightforward and convenient for operational use. The Bayesian retrieval and variational retrieval methods optimize the use of radar measurements. Both are promising methods for operational radar-rain retrieval. By comparison, the variational retrieval is more complicated, but it also has a greater potential than the Bayesian retrieval. So far, the Bayesian retrieval method is easier to carry out for the operational purpose. Later the variational retrieval method would prevail with its improved version.

7.3 Suggestions for Future Research

My subsequent work will primarily focus on improving the variational DSD retrieval from PRD. The major tasks needed to accomplish this are described as follows:

- The first task is to improve the variational retrieval through better data quality control. Preprocessing algorithms of radar data are about to be developed, making the data to be consistent and more reliable.

- The second task is to make the variational algorithm appropriate for operational use with S-band dual-polarization data. The S-band data are less affected by the attenuation, causing the variational method is less sensitive to the data quality. Even without an informative background, the S-band variational retrieval could output satisfactory results by neglecting the attenuation effect. With the dual-polarization upgrade of the NEXRAD, this work would be great practical significance for the application on the NEXRAD.

- The third task is to solve the problem of multiple-phases of precipitation. The current algorithm is limited in a small region (e.g., distance is less than several tens of kilometers) where snow/hail echoes are rarely observed. If the algorithm is applied to a large coverage (e.g., distance is more than 100 km), different precipitation phases are likely to exist, leading to the forward model error. Therefore, multiple-phases of precipitation should also be retrieved in the variational scheme.

The study of the rain microphysics and the rain retrieval is a very significative work. In this dissertation, I only address a limited study of this topic. Many interesting issues are worth of further researches. I hope my future research could have a great contribution to the operational use of advanced radar-rain algorithms and could be recognized by the radar meteorology community.

Bibliography

Beard, K. and A. Jameson, 1983: Raindrop canting. *J. Atmos. Sci.*, **40**, 448–454.

Beard, K., D. Johnson, and A. Jameson, 1983: Collisional forcing of raindrop oscillations. *J. Atmos. Sci.*, **40**, 455–462.

Beard, K. and A. Tokay, 1991: A field study of raindrop oscillations: Observations of size spectra and evaluation of oscillation causes. *Geophys. Res. Lett.*, **18**, 2257–2260.

Berne, A. and R. Uijlenhoet, 2006: Quantitative analysis of X-band weather radar attenuation correction accuracy. *Nat. Hazards Earth Syst. Sci.*, **6**, 419–425.

Brandes, E., G. Zhang, and J. Vivekanandan, 2002: Experiments in rainfall estimation with a polarimetric radar in a subtropical environment. *J. Appl. Meteor.*, **41**, 674–685.

Brandes, E., G. Zhang, and J. Vivekanandan, 2003: An evaluation of a drop distribution–based polarimetric radar rainfall estimator. *J. Appl. Meteor.*, **42**, 652–660.

Brandes, E., G. Zhang, and J. Vivekanandan, 2004a: Comparison of polarimetric radar drop size distribution retrieval algorithms. *J. Atmos. Oceanic Technol.*, **21**, 584–598.

Brandes, E., G. Zhang, and J. Vivekanandan, 2004b: Drop size distribution retrieval with polarimetric radar: Model and application. *J. Appl. Meteor.*, **43**, 461–475.

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

Bringi, V., V. Chandrasekar, N. Balakrishnan, and D. Zrnić, 1990: An examination of propagation effects in rainfall on radar measurements at microwave frequencies. *J. Atmos. Oceanic Technol.*, **7**, 829–840.

Bringi, V., V. Chandrasekar, J. Hubbert, E. Gorgucci, W. Randeu, and M. Schoenhuber, 2003: Raindrop size distribution in different climatic regimes from disdrometer and dual-polarized radar analysis. *J. Atmos. Sci.*, **60**, 354–365.

Bringi, V., G. Huang, V. Chandrasekar, and E. Gorgucci, 2002: A methodology for estimating the parameters of a gamma raindrop size distribution model from polarimetric radar data: Application to a squall-line event from the TRMM/Brazil campaign. *J. Atmos. Oceanic Technol.*, **19**, 633–645.

Bringi, V., T. Keenan, and V. Chandrasekar, 2001: Correcting C-band radar reflectivity and differential reflectivity data for rain attenuation: A self-consistent method with constraints. *IEEE Trans. Geosci. Remote Sens.*, **39**, 1906–1915.

Cao, Q. and G. Zhang, 2009: Errors in estimating raindrop size distribution parameters employing disdrometer and simulated raindrop spectra. *J. Appl. Meteor. Climatol.*, **48**, 406–425.

Cao, Q., G. Zhang, E. Brandes, T. Schuur, A. Ryzhkov, and K. Ikeda, 2008: Analysis of video disdrometer and polarimetric radar data to characterize rain microphysics in Oklahoma. *J. Appl. Meteor. Climatol.*, **47**, 2238–2255.

Cao, Q., G. Zhang, T. Schuur, A. Ryzhkov, E. Brandes, and K. Ikeda, 2006: Characterization of rain microphysics based on disdrometer and polarimetric radar observations. *Proceedings. IGARSS 2006, IEEE*, Denver, Colorado.

Chandrasekar, V. and V. N. Bringi, 1987: Simulation of radar reflectivity and surface measurements of rainfall. *J. Atmos. Oceanic Technol.*, **4**, 464–478.

Chen, J. and S. Liu, 2004: Physically based two-moment bulkwater parameterization for warm-cloud microphysics. *Quart. J. Roy. Meteor. Soc.*, **130**, 51–78.

Chiu, J. and G. Petty, 2006: Bayesian retrieval of complete posterior pdfs of oceanic rain rate from microwave observations. *J. Appl. Meteor. Climatol.*, **45**, 1073–1095.

Chuang, C. and K. Beard, 1990: A numerical model for the equilibrium shape of electrified raindrops. *J. Atmos. Sci.*, **47**, 1374–1389.

Delrieu, G., H. Andrieu, and J. Creutin, 2000: Quantification of path-integrated attenuation for X- and C-band weather radar systems operating in mediterranean heavy rainfall. *J. Appl. Meteor.*, **39**, 840–850.

Di Michele, S., A. Tassa., A. Mugnai, F. Marzano, P. Bauer, and J. Baptista, 2005: Bayesian algorithm for microwave-based precipitation retrieval: Description and application to TMI measurements over ocean. *IEEE Transactions on Geoscience and Remote Sensing*, **43**, 778–791.

Doviak, R. J. and D. S. Zrnic, 1993: *Doppler Radar and Weather Observations.* Academic Press, San Diego, CA.

Evans, K., J. Turk, T. Wong, and G. Stephens, 1995: A bayesian approach to microwave precipitation profile retrieval. *J. Appl. Meteor.*, **34**, 260–279.

Gao, J., M. Xue, K. Brewster, and K. Droegemeier, 2004: A three-dimensional variational data analysis method with recursive filter for doppler radars. *J. Atmos. Oceanic Technol*, **21**, 457–469.

Gertzman, H. S. and D. Atlas, 1977: Sampling errors in the measurements of rain and hail parameters. *J. Geophys. Res.*, **82**, 4955–4966.

Gorgucci, E. and L. Baldini, 2007: Attenuation and differential attenuation correction of C-band radar observations using a fully self-consistent methodology. *IEEE Geoscience and Remote Sensing Letters*, **2**, 326–330.

Gorgucci, E., V. Chandrasekar, V. N. Bringi, and G. Scarchilli, 2002: Estimation of raindrop size distribution parameters from polarimetric radar measurements. *J. Atmos. Sci.*, **59**, 2373–2384.

Green, A., 1975: An approximation for the shape of large raindrops. *J. Appl. Mereor.*, **14**, 1578–1583.

Haddad, Z. S., J. P. Meagher, S. L. Durden, E. A. Smith, and E. Im, 2006: Drop size ambiguities in the retrieval of precipitation profiles from dual-frequency radar measurements. *J. Atmos. Sci.*, **63**, 204–217.

Haddad, Z. S., D. A. Short, S. L. Durden, E. Im, S. H. Hensley, M. B. Grable, and R. A. Black, 1997: A new parameterization of the rain drop size distribution. *IEEE Trans. Geosci. Remote Sens.*, **35**, 532–539.

Hogan, R., 2007: A variational scheme for retrieving rainfall rate and hail reflectivity fraction from polarization radar. *J. Appl. Meteor. Climatol.*, **46**, 1544–1564.

Hong, S.-Y., J. Dudhia, and S.-H. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev.*, **132**, 103–120.

Huang, X., 2000: Variational analysis using spatial filters. *Mon. Wea. Rev.*, **128**, 2588–2600.

Ide, K., P. Courtier, M. Ghil, and A. Lorenc, 1997: Unified notation for data assimilation: Operational, sequential and variational. *J. Meteor. Soc. Japan*, **75**, 181–189.

Jameson, A. R. and A. B. Kostinski, 1998: Fluctuation properties of precipitation. Part II: Reconsideration of the meaning and measurement of raindrop size distributions. *J. Atmos. Sci.*, **55**, 283–294.

Joss, J. and E. G. Gori, 1978: Shapes of raindrop size distributions. *J. Appl. Meteor.*, **17**, 1054–1061.

Jung, Y., M. Xue, G. Zhang, and J. Straka, 2008a: Assimilation of simulated polarimetric radar data for a convective storm using the ensemble Kalman filter. Part II: Impact of polarimetric data on storm analysis. *Mon. Wea. Rev.*, **136**, 2246–2260.

Jung, Y., G. Zhang, and M. Xue, 2008b: Assimilation of simulated polarimetric radar data for a convective storm using the ensemble Kalman filter. Part I: Observation operators for reflectivity and polarimetric variables. *Mon. Wea. Rev.*, **136**, 2228–2245.

Keenan, T., L. Carey, D. Zrnić, and P. May, 2001: Sensitivity of 5-cm wavelength polarimetric radar variables to raindrop axial ratio and drop size distribution. *J. Appl. Mereor.*, **40**, 526–545.

Kliche, D. V., P. L. Smith, and R. W. Johnson, 2008: L-moment estimators as applied to gamma drop size distributions. *J. Appl. Meteor. Climatol.*, **47**, 3117–3130.

Kozu, T. and K. Nakamura, 1991: Rain parameter estimation from dual-radar measurements combining reflectivity profile and path-integrated attenuation. *J. Atmos. Oceanic Technol.*, **8**, 259–270.

Kruger, A. and W. F. Krajewski, 2002: Two-dimensional video disdrometer: A description. *J. Atmos. Oce. Tech.*, **19**, 602–617.

Lee, G. and I. Zawadzki, 2005: Variability of drop size distributions: Time-scale dependence of the variability and its effects on rain estimation. *J. Appl. Meteor.*, **44**, 241–255.

Lee, J., M. Grunes, and G. D. Grandi, 1997: Polarimetric SAR speckle filtering and its impact on classification. *Geoscience and Remote Sensing, IGARSS '97*, Vol. 2, 1038–1040.

Lim, S., V. Chandrasekar, and V. Bringi, 2005: Hydrometeor classification system using dual-polarization radar measurements: Model improvements and in situ verification. *IEEE Trans. Geosci. Remote Sensing*, **43 (4)**, 792–801.

Liu, H. and M. Xue, 2006: Retrieval of moisture from slant-path water vapor observations of a hypothetical GPS network using a three-dimensional variational scheme with anisotropic background error. *Mon. Wea. Rev.*, **134**, 933–949.

Liu, H., M. Xue, R. J. Purser, and D. F. Parrish, 2007: Retrieval of moisture from simulated GPS slant-path water vapor observations using 3DVAR with anisotropic recursive filters. *Mon. Wea. Rev.*, **135**, 1506–1521.

Liu, Y., V. Bringi, and M. Maki, 2006: Improved rain attenuation correction algorithms for radar reflectivity and differential reflectivity with adaptation to drop shape model variation. *Proceedings. IGARSS 2006*, Denver, Colorado.

Marshall, J. and W. Palmer, 1948: The distribution of raindrops with size. *J. Meteor.*, **5**, 165–166.

Mc Farlane, S., K. Evans, and A. S. Ackerman, 2002: A bayesian algorithm for the retrieval of liquid water cloud properties from microwave radiometer and millimeter radar data. *J. Geophys. Res.*, **107**, 4317.

Meneghini, R. and L. Liao, 2007: On the equivalence of dual-wavelength and dual-polarization equations for estimation of the raindrop size distribution. *J. Atmos. Oceanic Technol.*, **24**, 806–820.

Moisseev, D. N. and V. Chandrasekar, 2007: Examination of $\mu - \Lambda$ relation suggested for drop size distribution parameters. *J. Atmos. Oceanic Tech.*, **24**, 847–855.

Nespor, V., W. F. Krajewski, and A. Kruger, 2000: Wind-induced error of raindrop size distribution measurement using a two-dimensional video disdrometer. *J. Atmos. Oce. Tech.*, **17**, 1483–1492.

Park, S., V. Bringi, V. Chandrasekar, M. Maki, and K. Iwanami, 2005: Correction of radar reflectivity and differential reflectivity for rain attenuation at X band. Part I: Theoretical and empirical basis. *J. Atmos. Oceanic Technol.*, **22**, 1621–1632.

Parrish, D. F. and J. C. Derber, 1992: The national meteorological center's spectral statistical-interpolation analysis system. *Mon. Wea. Rev.*, **120**, 1747–1763.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, 2001: *Numerical Recipes in FORTRAN 77.* 2d ed., Cambridge University Press.

Pruppacher, H. and K. Beard, 1970: A wind tunnel investigation of the internal circulation and shape of water drops falling at terminal velocity in air. *Quart. J. Roy. Meteor. Soc.*, **96**, 247–256.

Pruppacher, H. and R. Pitter, 1971: A semi-empirical determination of the shape of cloud and raindrops. *J. Atmos. Sci.*, **28**, 84–86.

Ray, P. S., 1972: Broadband complex refractive indices of ice and water. *Appl. Opt.*, **11**, 1836–1844.

Ryzhkov, A., S. Giangrande, and T. Schuur, 2005a: Rainfall estimation with a polarimetric prototype of WSR-88D. *J. Appl. Meteor.*, **44**, 502–515.

Ryzhkov, A., T. Schuur, D. Burgess, P. Heinselman, S. Giangrande, and D. Zrnic, 2005b: The joint polarization experiment: Polarimetric rainfall measurements and hydrometeor classification. *Bull. Amer. Meteor. Soc.*, **86**, 809–824.

Ryzhkov, A., T. J. Schuur, D. Burgess, and D. Zrnić, 2005c: Polarimetric tornado detection. *J. Appl. Mereor.*, **44**, 557–570.

Ryzhkov, A., P. Zhang, D. Hudak, J. L. Alford, M. Knight, and J. Conway, 2007: Validation of polarimetric methods for attenuation correction at C band. *33rd Conference on Radar Meteorology*, Cairns, Queensland, Australia, A.M.S.

Ryzhkov, A. and D. Zrnic, 1995: Precipitation and attenuation measurements at a 10-cm wavelength. *J. Appl. Meteor.*, **34**, 2121–2134.

Sauvageot, H. and J. Lacaux, 1995: The shape of averaged drop size distributions. *J. Atmos. Sci.*, **52**, 1070–1083.

Schuur, T., A. Ryzhkov, P. Heinselman, D. Zrnic, D. Burgess, and K. Scharfenberg, 2003: Observations and classification of echoes with the polarimetric WSR-88D radar. Tech. rep., National Severe Storms Laboratory.

Schuur, T., A. Ryzhkov, D. Zrnić, and M. Schönhuber, 2001: Drop size distributions measured by a 2D video disdrometer: Comparison with dual-polarization radar data. *J. Appl. Meteor.*, **40**, 1019–1034.

Smith, J. A., 1993: Marked point process model distributions. *J. Appl. Meteor.*, **32**, 284 –296.

Smith, P., 2003: Raindrop size distributions: Exponential or gamma - does the difference matter? *J. Appl. Meteor.*, **42**, 1031–1034.

Smith, P., D. Kliche, and R. Johnson, 2005: The bias in moment estimators for parameters of drop-size distribution functions: sampling from gamma distributions. *Proc. of 32nd conference on radar meteorology*, Albuquerque, NM.

Smith, P., D. Kliche, and R. Johnson, 2009: The bias and error in moment estimators for parameters of drop size distributions: Sampling form gamma distributions. *J. Appl. Mereor.*, early online release.

Smith, P., Z. Liu, and J. Joss, 1993: A study of sampling-variability effects in raindrop size observations. *J. Appl. Meteor.*, **32**, 1259–1269.

Smith, P., D. Musil, S. Weber, J. Spahn, F. N. Johnson, and W. R. Sand, 1976: raindrop and hailstone size distributions inside hailstorms. *Int. Conf, on cloud Physics*, Boulder, CO.

Testud, J., E. L. Bouar, E. Obligis, and M. Ali-Mehenni, 2000: The rain profiling algorithm applied to polarimetric weather radar. *J. Atmos. Oceanic Technol.*, **17**, 332–356.

Testud, J., S. Oury, R. Black, P. Amayenc, and X. Dou, 2001: The concept of "normalized" distribution to describe raindrop spectra: A tool for cloud physics and cloud remote sensing. *J. Appl. Mereor.*, **40**, 1118–1140.

Tokay, A., A. Kruger, and W. F. Krajewski, 2001: Comparison of drop size distribution measurements by impact and optical disdrometers. *J. Appl. Meteor.*, **40**, 2083–2097.

Tokay, A. and D. Short, 1996: Evidence from tropical raindrop spectra of the origin of rain from stratiform versus convective clouds. *J. Appl. Meteor.*, **35**, 355–371.

Ulbrich, C., 1983: Natural variations in the analytical form of the raindrop size distribution. *J. Climate Appl. Meteor.*, **22**, 1764–1775.

Ulbrich, C. and D. Atlas, 1998: Rainfall microphysics and radar properties: Analysis methods for drop size spectra. *J. Appl. Meteor.*, **37**, 912–923.

Vivekanandan, J., G. Zhang, and E. Brandes, 2004: Polarimetric radar rain estimators based on constrained gamma drop size distribution model. *J. Appl. Meteor.*, **43**, 217–230.

Vulpiani, G., F. Marzano, V. Chandrasekar, and S. Lim, 2005: Constrained iterative technique with embedded neural network for dual-polarization radar correction of rain path attenuation. *IEEE Trans. Geosci. Remote Sensing*, **43**, 2305–2314.

Waldvogel, A., 1974: The N_0 jump of raindrop spectra. *J. Atmos. Sci.*, **31**, 1067–1078.

Wong, R. and N. Chidambaram, 1985: Gamma size distribution and stochastic sampling errors. *J. Appl. Mereor.*, **24**, 568–579.

Xue, M., M. Tong, and K. Droegemeier, 2006: An OSSE framework based on the ensemble square-root Kalman filter for evaluating impact of data from radar networks on thunderstorm analysis and forecast. *J. Atmos. Ocean Tech.*, **23**, 46–66.

Xue, M., M. Tong, and G. Zhang, 2009: Simultaneous state estimation and attenuation correction for thunderstorms with radar data using an ensemble Kalman filter: tests with simulated data. *Quart. J. Roy. Meteor. Soc.*, **135 (643)**, 1409–1423.

Zhang, G., J. Sun, and E. Brandes, 2006a: Improving parameterization of rain microphysics with disdrometer and radar observations. *J. Atmos. Sci.*, **63**, 1273–1290.

Zhang, G., J. Vivekanandan, and E. Brandes, 2001: A method for estimating rain rate and drop size distribution from polarimetric radar. *IEEE Trans. Geosci. Remote Sensing*, **39 (4)**, 830–840.

Zhang, G., J. Vivekanandan, E. Brandes, R. Menegini, and T. Kozu, 2003: The shape–slope relation in observed gamma drop size distributions: Statistical error or useful information? *J. Atmos. Oceanic Technol.*, **20**, 1106–1119.

Zhang, G., J. Vivekanandan, and M. Politovich, 2004: Radar/radiometer combination to retrieve cloud characteristics for icing detection. *11th Conference on Aviation, Range, and Aerospace.*, Hyannis, MA, AMS.

Zhang, G., M. Xue, Q. Cao, and D. Dawson, 2008: Diagnosing the intercept parameter for exponential raindrop size distribution based on video disdrometer observations: Model development. *J. Appl. Meteor. Climatol*, **47**, 2983–2992.

Zhang, G., M. Xue, T. Schuur, P. Chilson, R. Palmer, and D. Zrnic, 2006b: Improving microphysics parameterizations and quantitative precipitation forecast through optimal use of video disdrometer, profiler and polarimetric radar observations. *NSF Grant, ATM-0608168*.

Zrnic, D., 1975: Simulation of weather-like Doppler spectra and signals. *J. Appl. Mereor.*, **14**, 619–629.

Appendix A - List Of Symbols

A_{hh}	Specific Horizontal Attenuation
A_{DP}	Specific Differential Attenuation
$\mathbf{B}$	Error Covariance Matrix of Background
cm	Centimeter, 10^{-2} meters
C	Adjustment Parameter
D	Equivalent Diameter
$\mathbf{D}$	Square Root of Background Error Covariance Matrix
D_0	Median Volume Diameter
D_a	Mean Diameter
D_e	Effective Diameter
D_i	Diameter of i^{th} Bin
D_m	Mean Volume Diameter
D_{max}	Maximum Diameter
D_{min}	Minimum Diameter
$f_{a,b}$	Scattering Amplitude at Major (or Minor) Axis of a Drop without Canting
F_{hh}	Scattering Amplitude at Horizontal Direction
F_{vv}	Scattering Amplitude at Vertical Direction
$F_N(X)$	Normalized Distribution
GHz	Gigahertz, 10^9 Hz
$H(\mathbf{x})$	Nonlinear Observation Operator

$\mathbf{H}$	Linearization of Observation Operator
Hz	Hertz, unit increment of frequency
I/Q	In-phase/Quadrature Signal
J	Cost Function of Variational Scheme
k	Wave Number
km	Kilometer, 10^3 meters
kW	Kilowatts, 10^3 Watts
K_{DP}	Specific Differential Phase
l_1, l_2	First Two L-Moments
m	Meter, unit increment of length
mm	Milimeter, 10^{-3} meters
MHz	Megahertz, 10^6 Hz
M_a	Mass of Accretion
M_e	Mass of Evaporation
M_n	DSD n^{th} Moment
N_0	Concentration Parameter of DSD
N_0'	Transformation of N_0
$N(D)$	Raindrop Size Distribution
N_i	Total Number of Drops Within i^{th} Bin
N_T	Total Number Concentration
N_w	Concentration Parameter of Normalized Gamma DSD
N^*	Normalization Term of DSD
P_f	Conditional PDF of Observation
P_{post}	Posterior PDF
P_{pr}	Prior PDF
P_t	Peak Transmit Power

r	Raindrop Axis Ratio
R	Rainfall Rate
$\mathbf{R}$	Error Covariance Matrix of Observation
R_a	Accretion Rate
R_e	Evaporation Rate
T	Temperature
$\mathbf{v}$	Transformed State Vector
v_a	Aliasing Velocity
$v(D)$	Falling Velocity
$V_{tm},\ v_{tm}$	Mass-weighted Terminal Velocity
$\mathbf{w}$	Weighting Function of Variational Scheme
W	Water Content
$\mathbf{x}$	State Vector
$\mathbf{y}$	Observation Vector
Z	Radar Reflectivity Factor
Z_{DR}	Differential Reflectivity
Z_{DR}^a	Attenuated Z_{DR}
Z_{DR}^{up}, (or Z_{DR}^{low})	Upper (or Lower) Boundary of $Z_H - Z_{DR}$ for Rain Data
Z_{hh}	Horizontal Radar Reflectivity
Z_{vv}	Vertical Radar Reflectivity
Z_H (or Z_V)	Radar Reflectivity Factor for Horizontal (or Vertical) Polarization
Z_H^a	Attenuated Z_H
$Z_{H,V}^O$	Observed $Z_{H,V}$
$Z_{H,V}^T$	Ground Truth of $Z_{H,V}$
Γ	Gamma Function

Δ	Increment
Δr	Range Resolution
δ	Increment
ε	Error
ϵ_r	Complex Dielectric Constant
Λ	Slope Parameter of DSD
Λ'	Transformation of Λ
λ	Wavelength
μ	Shape Parameter of DSD
ρ	Correlation Coefficient
ρ_w	Water Density
σ	Standard Deviation of Gaussian Distribution
$\sigma_{ext}^{H,V}$	Extinction Cross Section at Horizontal or Vertical Polarization
σ_H (or σ_V)	Back Scattering Cross Section for Horizontal (or Vertical) Polarization
σ_m	Standard Deviation of Mass-weighted Diameter Distribution
σ_ϕ	Standard Deviation of Canting Angle
τ	Pulse Duration
Ψ	The "psi" Function
∇	Partial Derivative
$^\circ$	Degree
$^\circ$C	Centidegree

Appendix B - List Of Acronyms and Abbreviations

2DVD	Two-Dimensional Video Disdrometer
ARRC	Atmospheric Radar Research Center
BD	Big Drops
CASA	Collaborative Adaptive Sensing of the Atmosphere
C-G	Constrained-Gamma Model
DA	Data Assimilation
DP	Direct Phase Correction Algorithm
DSD	Raindrop Size Distribution
FSD	Fractional Standard Deviation
H	Horizontal Channel
H-B	Hitschfeld-Bordan Algorithm
IUT	Indoor User Terminal
JPOLE	Joint Polarization Experiment
JWD	Joss-Waldvogel Disdrometer
KFFL	Kesseler Farm Field Laboratory
LM	L-Moment
M012	Moment Estimator Using Moments 0, 1, 2
M234	Moment Estimator Using Moments 2, 3, 4
M246	Moment Estimator Using Moments 2, 4, 6
M346	Moment Estimator Using Moments 3, 4, 6
M456	Moment Estimator Using Moments 4, 5, 6

M-P	Marshall and Palmer Model
ML	Maximum Likelihood
NCAR	National Center for Atmospheric Research
NEXRAD	Next-Generation Radar
NSF	National Science Foundation
NSSL	National Severe Storms Laboratory
NWP	Numerical Weather Prediction Model
NWS	National Weather Service
OEU	Outdoor Electronics Unit
OU	University of Oklahoma
OU-PRIME	Polarimetric Radar for Innovations in Meteorology and Engineering
PDF	Probability Density Function
PPI	Plane Position Indicator
PRD	Polarimetric Radar Data
PRF	Pulse Repetition Frequency
PRT	Pulse Repetition Time
QPE	Quantitative Precipitation Estimation
QPF	Quantitative Precipitation Forecast
RMSE	Root Mean Square Deviation of Error
SATP	Sorting and Averaging Procedure Based on Two Parameters
SC	Self-Consistence Algorithm
SD	Standard Deviation
SIFT	Sequential Intensity Filtering Technique
SNR	Signal-to-Noise Ratio
S-Pol	S-band Polarimetric Radar

TDWR	Terminal Doppler Weather Radar
TMF	Truncated Moment Fit
V	Vertical Channel
WRF	Weather Research and Forecasting Model
WSR-88D	Weather Surveillance Radar–1988 Doppler