UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

STATE AND PARAMETER ESTIMATION USING POLARIMETRIC RADAR DATA AND AN ENSEMBLE KALMAN FILTER

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirement for the

Degree of

DOCTOR OF PHILOSOPHY

By

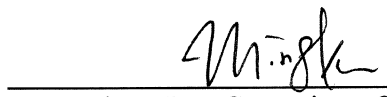
YOUNG SUN JUNG
Norman, Oklahoma
2008

Downloaded from SHAREOK. Persistent Handle URL: https://hdl.handle.net/11244/320231

STATE AND PARAMETER ESTIMATION USING POLARIMETRIC RADAR DATA AND AN ENSEMBLE KALMAN FILTER

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEOROLOGY

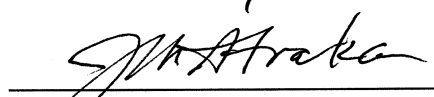
BY



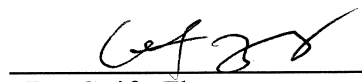
Dr. Ming Xue, Committee Chair



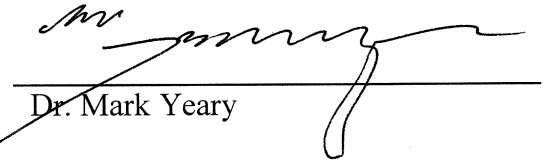
Dr. Kelvin Droegemeier



Dr. Jerry Straka



Dr. Guifu Zhang



Dr. Mark Yeary

© Copyright by YOUNG SUN JUNG 2008
All Rights Reserved.

Acknowledgements

I would like to gratefully and sincerely thank my advisor, Dr. Ming Xue, for the supervision, understanding, and support he provided me over the years. His untiring passion and enthusiasm in research were the motivations that drove me to pursue my academic goals. He stimulated and inspired me to develop independent thinking during my graduate studies. Most of all, he has always given priority to his students and is always accessible when needed, despite his busy schedule.

I am deeply grateful for the expertise and guidance I received from Dr. Guifu Zhang, without which I would have been lost several times. His valuable lectures and my numerous discussions with him helped me understand the theory of wave propagation and scattering, upon which much of my research relies.

I have been especially fortunate to have Dr. Kelvin Droegemeier as a doctoral committee member. His recommendations and suggestions have been invaluable to my work, and he always took time from his busy schedule to help me.

I am also thankful to Dr. Jerry Straka and Dr. Mark Yeary for their continual support and encouragement. Moreover, Dr. Straka helped me adjust to a new place in which I had no acquaintances, and Dr. Yeary carefully read and commented on my draft.

My deepest gratitude goes to my parents, Yunsub Jung and Hyeja Kim, for their unconditional love and faith in me and for their always being supportive of every decision I have made. They encouraged me to be ambitious, and they supported me throughout my life. I owe my parents everything I have accomplished. I also wish to thank my sisters, Youngmi, Sujin, and Hyunjun, for sharing my joy and excitement and for cheering me up when I encountered difficulties.

I wish to thank Dr. Frederick Carr for his counsel and his willingness to help me with difficulties due to cultural differences. I extend my sincere thanks to the staff at the

School of Meteorology and CAPS—Celia Jones, Marcia Pallutto, Eileen Hasselwander, Debra Farmer, Nancy Campbell, and Krysta Bruehl—for all of their help. I also thank Yunheng Wang, Scott Hill, Donald Giuliano, and the OSCER operations team for helping me with the ARPS model, computers, and servers.

I want to thank my good friends for their support throughout my time at OU. I thank Mingjing for her great help with the initial use of her code. Haixia, Jess, and Elaine were always there for me when I needed friends. I also thank Dan for many great conversations that often give me new ideas. I would like to acknowledge other officemates, Jili, Nate, and Charlotte, and many other friends for sharing their learning and encouraging each other. These made my graduate experience something that I will cherish forever.

I greatly value my friendships with Eunha and Youngeun, and I deeply appreciate their helping me when I went through the toughest time in my life. I also wish to thank my managers and colleagues at the Korea Meteorological Administration for their support and encouragement. I extend my gratitude to OU for awarding me the Alumni Fellowship, which allowed me to fully focus on my research and studies.

I owe a special note of gratitude to Dr. Fanthune Moeng, for his encouragement that I attend graduate school. He has been a true friend to me; he celebrated my successes with me and admonished me when I became too easy on myself.

This work was primarily supported by the NSF Engineering Research Center (ERC) for Collaborative Adaptive Sensing of the Atmosphere (CASA) under grants EEC-0313747, and by grants NSF ATM-0608168, ATM-0530814, and ATM-0331594. The computations were performed at the Pittsburgh Supercomputing Center supported by the NSF and at the OU Supercomputing Center for Education and Research at the University of Oklahoma.

Table of Contents

Acknowledgements ..... **iv**

Table of Contents ..... **vi**

List of Tables ..... **x**

List of Figures ..... **xii**

Abstract ..... **xxi**

Chapter 1 Introduction and Overview **1**

1.1 Background and Motivation 1

1.2 Dissertation Overview 4

Chapter 2 Observation Operators for Reflectivity and Polarimetric Variables **8**

2.1 Introduction 8

2.2 The model and convective storm simulations 12

2.3 The observation operators and simulation of observations 13

2.3.1 The shape, orientation, and drop size distribution of hydrometeors 14

2.3.2 Melting ice (snow-hail) model 17

2.3.3 Observation operators 20

2.4 Applications to convective storms 27

2.4.1 Simulated radar fields for the squall-line case 27

2.4.2 Simulated radar fields for a supercell storm 37

2.5 Summary and conclusions 42

Chapter 3 Assimilation of Simulated Polarimetric Radar Data for a Convective Storm Using Ensemble Kalman Filter: Impact of Polarimetric Data on Storm Analysis **44**

3.1 Introduction 44

3.2 Assimilation system and experimental design 49
3.2.1 Simulation of observations and the error model 49
3.2.2 Data assimilation procedure 55
3.2.3 Experimental design 56
3.3 The impact of assimilating polarimetric variables 61
3.4 Summary and further discussion 72
Chapter 4 Simultaneous Estimation of Microphysical Parameters and Atmospheric State using Simulated Polarimetric Radar Data and Ensemble Kalman Filter in the Presence of Observation Operator Error 75
4.1 Introduction 75
4.2 Model and experimental design 79
4.2.1 Forecast model and filter configuration 79
4.2.2 Simulation of observations 82
4.2.3 Parameters to estimate 83
4.2.4 Parameter estimation procedure 84
4.2.5 Design of parameter estimation experiments 85
4.3 Sensitivity analysis 87
4.3.1 Response function 87
4.3.2 Results of sensitivity experiments 89
4.4 Results of parameter estimation 92
4.4.1 Results of single-parameter estimation 92
4.4.2. Results of five-parameter estimation 103
4.4.3. Results of three-parameter estimation 110
4.5. Summary and conclusions 115
Chapter 5 Polarimetric Signatures of a Supercell Storm Simulated Using a Two-Moment Microphysics Scheme 119
5.1 Introduction 119

5.2 Polarimetric radar data emulator 123

5.3 Model configurations 127

5.4 Polarimetric signatures of a simulated convective storms 129

5.4.1 Storm evolution and simulated reflectivity 129

5.4.2 Simulated polarimetric radar fields for a supercell storm 131

5.4.2.1 Z_{DR} and K_{DP} columns 131

5.4.2.2 Z_{DR} arc 134

5.4.2.3 Mid-level Z_{DR} and ρ_{hv} rings 136

5.4.2.4 Hail signature in the forward flank downdraft 139

5.5 Summary and discussions 140

Chapter 6 State Estimation of Convective Storms with Two-Moment Microphysics Scheme and Ensemble Kalman Filter: Experiments with Simulated Data 145

6.1 Introduction 145

6.2 Model and experimental setup 148

6.2.1 Prediction model and EnSRF assimilation procedure 148

6.2.2 Transform of the total number concentration in the filter 151

6.2.3 Observations 152

6.3 Perfect model experiments 153

6.3.1 Assimilation of conventional radar data 153

6.3.2 Assimilation of polarimetric radar data 166

6.4 Imperfect model experiments 170

6.4.1 Forecast model error 172

6.4.2 Forecast model and observation operator errors 174

6.5 Summary and conclusions 178

Chapter 7 Summary and Future work 182

7.1 Summary 182

7.2 Future work 187

References 191

Nomenclature 206

List of Tables

Table 3.1. List of experiments testing the impact of polarimetric variables..... 57

Table 3.2. The improvement over the experiment Zh for the experiments ZhZdr, ZhZdp, ZhKdp, and ZhZZK and over the experiment VrZh for the experiments VrZhZdr, VrZhZdp, VrZhKdp, and VrZhZZK averaged over the last nine cycles (60-100 min of model time). The improvement is expressed in percentages relative to the corresponding control experiment. 67

Table 4.1. Microphysical parameters and their uncertainty ranges used in the sensitivity experiments, and their initial guesses as used in the parameter estimation experiments. 86

Table 4.2. The mean estimated parameter values in logarithmical form for single-parameter estimation experiments, averaged over 15 experiments with three different initial guesses presented in Table 4.1 and five different perturbation realizations for each initial guess, over the last 5 cycles (80 minutes to 100 minutes of model time). Their truth values in logarithmical form are given inside parentheses. 102

Table 4.3. The percentage improvement in state estimation for experiments para5_ZhZdr, para5_ZhKdp and para5_ZhZdrKdp over experiment para5_Zh, averaged over 160 experiments with 32 different initial guesses with 5 parallel runs for each initial guess and over the last 5 cycles (80 minutes to 100 minutes of model time). Prefix 'para5_' is omitted in the experiment names. 108

Table 4.4. Same as in Table 4.2 but for five-parameter experiment para5_Zh and three-parameter estimation experiments para3_Zh, para3_ZhZdr, para3_ZhKdp and para3_ZhZdrKdp, in which n_{0S} and ρ_S were kept as their incorrect initial value throughout the assimilation cycles while other three parameters were estimated. The experiments start from two sets of parameter values, namely, (n_{0R}, n_{0S}, n_{0H},

ρ_S, ρ_H) = ($3x10^6$ m^{-4}, $7x10^5$ m^{-4}, $4x10^5$ m^{-4}, 50 kg m^{-3}, 400 kg m^{-3}) and ($3x10^6$ m^{-4}, $3x10^7$ m^{-4}, $4x10^5$ m^{-4}, 300 kg m^{-3}, 400 kg m^{-3}). 112

Table 4.5. The percentage improvement of state estimation for three-parameter estimation experiments para3_ZhZdr, para3_ZhKdp, and para3_ZhZdrKdp over experiment para3_Zh. The percentage improvements are ensemble mean of those computed from ten experiments listed in Table 4.4. Prefix 'para3_' is omitted in the experiment names. 115

Table 6.1. List of data assimilation experiments. 154

Table 6.2. List of experiments testing the impact of polarimetric variables..... 167

Table 6.3. The improvement over the experiment VrZh for the experiments VrZhZdr, VrZhKdp, and VrZhZdrKdp averaged over the last five cycles (80-100 min of model time). The improvement is expressed in percentages relative to the control experiment. 169

Table 6.4. List of data assimilation experiment in the presence of forecast model error. 172

List of Figures

Fig. 2.1. Backscattering amplitudes as a function of the effective diameter of particle along (a) the major axis and (b) the minor axis. 22

Fig. 2.2. Backscattering amplitudes as a function of the fraction of water within the mixture along the major axis (solid) and the minor axis (dashed) for (a) rain-snow and (b) rain-hail mixtures. 23

Fig. 2.3. The west-east vertical cross sections of the simulated (a) Z_H, (b) Z_{DR}, (c) Z_{dp}, and (d) K_{DP} at 400 minutes into the 2D squall-line simulation.The 0 ºC isotherms are shown as thick black lines. A sequence of cells in (a) is labeled C1-C4. 29

Fig. 2.4. As in Fig. 2.3, but for the model mixing ratios (a) q_r, (b) q_s, and (c) q_h. 30

Fig. 2.5. A modeled vertical profile of total (in both pure and mixture forms) rainwater and total snow-hail mixing ratios and the total amount (mixing ratio) of rain and snow and rain and hail in a mixture form, and the simulated polarization radar signatures at the column labeled **C** in Fig. 2.3: (a) q_r (thick solid), q_h (thin solid), q_s (thin dashed), q_{rh} (thick dashed), and q_{rs} (dash-dotted); (b) reflectivity from the LI model (dashed), reflectivity at horizontal (Z_H, solid) polarization from the MI model; (c) Z_{DR}; and (d) K_{DP}. Here q_{rs} is on the vertical axis. 35

Fig. 2.6. As in Fig. 2.5, but for column **S** in Fig. 2.3. 36

Fig. 2.7. (a) Horizontal wind (vectors; m s-1) and reflectivity, Z_H, (b) rain water q_r (thin solid contours and shading) and q_h (thick dotted at intervals of 1 g kg^{-1}, starting from 0 g kg^{-1}), (c) Z_{DR}, and d) K_{DP}, at z = 2.5 km at 100 minutes of the storm. . 38

Fig. 2.8. The west-east vertical cross-sections of simulated (a) reflectivity at horizontal polarization (Z_H), (b) differential reflectivity (Z_{DR}), (c) specific differential phase (K_{DP}), and (d) rain water (q_r), (e) snow (q_s), and (f) hail (q_h) mixing ratios, (g) reflectivity difference $(Z_{dp})^{0.2}$, and (h) cross-correlation coefficient ρ_{hv} through

the updraft core (maximum vertical velocity) of the simulated supercell storm at 100 minutes, along line AB shown in Fig. 2.7a. It corresponds to y = 28 km. .. 39

Fig. 3.1. Simulated (a),(c),(e) error-free and (b),(d),(f) error-containing observations at the 0.5º elevation at 100 min of the supercell storm simulation of (a),(b) Z_H; (c),(d) Z_{DR}; and (e),(f) K_{DP}. 54

Fig. 3.2. Scatterplot of reflectivity vs (a),(b) Z_{DR}; (c),(d) $(Z_{dp})^{0.2}$, and (e),(f) K_{DP} for (a),(c),(e) truth and (b),(d),(f) observation. The thresholds applied to the observation in the assimilation are overlaid on each plot (thick dashed). 60

Fig. 3.3. The ensemble mean forecast and analysis RMSEs averaged over points at which the true reflectivity is greater than 10 dBZ for (a) u, (b) v, (c) w (d) perturbation potential temperature θ', (e) p', (f) q_c, (g) q_r, (h) q_v (the curves with larger values) and q_i (the curves with lower values), (i) q_s, and (j) q_h for experiments Zh (solid black) and VrZh (dotted black). The vertical straightline segments in the curves correspond to the reduction or increase in RMSEs or ensemble spreads by the data assimilation. 62

Fig. 3.4. The ensemble mean analysis RMSEs of experiments ZhZZK (thick solid), ZhZdr (solid), ZhZdp (dashed), and ZhKdp (dotted) normalized by those of experiment Zh. The reference horizontal line at a unity is overlaid. 63

Fig. 3.5. The vertical profile of RMS analysis errors averaged over points at which the truth reflectivity is greater than 10 dBZ for (a) u, (b) v, (c) w, (d) θ', (e) p', (f) q_c, (g) q_r, (h) q_v, (i) q_s, and (j) q_h at 80 min of experiments Zh (dashed) and ZhZdr (solid). 68

Fig. 3.6. As in Fig. 3.4, but for experiments VrZhZZK (thick solid), VrZhZdr (solid), VrZhZdp (dashed), and VrZhKdp (dotted). 69

Fig. 4.1. Response function difference (a),(c),(e) $RFD^t = J_y(p^t) - J_{y,c}(p^t)$ calculated with correct parameter values and (b),(d),(f) $RFD^w = J_y(p^w) - J_{y,c}(p^t)$ with

incorrect parameter values in the radar observation operators, for (a),(b) reflectivity data; (c),(d) differential reflectivity data; and (e),(f) specific differential phase data. The parameter deviations are in logarithmic space. 91

Fig. 4.2. The time evolution of normalized absolute error (NAE) of the ensemble mean of estimated parameter values from single-parameter estimation experiments, for (a)-(c) n_{0R}; (d)-(f) n_{0S}; (g)-(i) n_{0H}; (j)-(l) ρ_S; and (m)-(o) ρ_H, when the parameter estimation is performed using Z_H alone (thick solid gray), both Z_H and K_{DP} (solid black), and Z_{DR} alone (dashed black). The experiment name starts with the parameter name and is followed by the coefficient and the exponent of the initial guess of (a)-(i) the intercept parameter or (j)-(o) the first two digits of the bulk density presented in Table 4.1. The NAEs are averaged over the five parallel experiments that have the same initial guesses but different realizations of the initial parameter ensemble. 94

Fig. 4.3. The ensemble-mean analysis RMS errors averaged over points at which the true reflectivity is greater than 10 dBZ for (a) u, (b) v, (c) w and (d) perturbation potential temperature θ', (e) water vapor q_v, (f) cloud water q_c, (g) rainwater q_r, (h) cloud ice q_i, (i) snow q_s, and (j) hail q_h, for experiments where n_{0R} alone contains error and is estimated. The experiments use Z_H data alone (thick solid gray), both Z_H and K_{DP} data (solid black), or Z_{DR} data alone (dashed black). The RMS errors are averaged over 15 experiments that start from three initial guesses presented in Table 4.1 and five initial perturbation realizations for each initial guess. 96

Fig. 4.4. The same as Fig. 4.3 but for n_{0S}. 96

Fig. 4.5. The same as Fig. 4.3 but for n_{0H}. 97

Fig. 4.6. The time evolution of the correlation coefficients between parameter n_{0H} and the model prior estimates of Z_{DR} from one of the five parallel experiments named N0h46_Zdr (solid) and those between n_{0H} and Z_H from corresponding

experiment N0h46_Zh (dotted), averaged over the 30 observations used in parameter estimation. 99

Fig. 4.7. Vertical column maximum Z_H and Z_{DR} shown in the horizontal plane (a and c, respectively) and column maximum Z_H and Z_{DR} in the y direction shown in the vertical x-z plane (b and d, respectively), of truth simulation at 45 minutes (shading and thin solid contours). Solid squares indicate the locations of the observations that were used in the single-parameter estimation experiments N0h46_Zh (Z_H observations in a and b) and N0h46_Zdr (Z_{DR} observations in c and d). The data points are projected to the horizontal x-y plane in the left panels and to the vertical x-z plane in the right panels. Thick dotted contours at intervals of 2 g kg^{-1} represent the hail mixing ratio through vertical velocity maximum. 100

Fig. 4.8. As in Fig. 4.7 but for 90 min. 101

Fig. 4.9. The time evolution of the NAE of the mean parameter values from five-parameter estimation experiments for (a) n_{0R}, (b) n_{0S}, (c) n_{0H}, (d) ρ_S, and (e) ρ_H for the experiments para5_Zh (thick solid gray), para5_ZhZdr (dotted black), para5_ZhKdp (dashed black), and para5_ZhZdrKdp (solid black). Initial error level is shown in dashed gray. The average NAE is calculated from the 160 experiments with 32 different initial guesses consisting of the combinations of 5 parameters with 2 initial guesses each, as shown in Table 4.1, where 5 parallel runs with different realizations of initial parameter perturbations are carried out for each initial guess. 104

Fig. 4.10. As Fig. 4.3 but for simultaneous estimation of five parameters for experiments para5_Zh (dashed black) and para5_ZhZdrKdp (solid black). The RMS errors of the no-parameter-estimation experiments with the initial guesses of parameters kept throughout the assimilation cycles are shown in thick solid gray for comparison. The RMS errors are averaged over 160 experiments as in Fig. 4.9. 106

Fig. 4.11. Vertical column maximum Z_H in a *x*-*y* plane (a and b), column maximum Z_H in *y* direction in an *x*-*z* plane (c and d) and in x direction in a y-z plane (e and f), from the truth simulation at 40 min. Black dots in the left panels indicate the locations of Z_H observations used in 5-parameter estimation experiment para5_Zh, and the black dots, triangles and squares in the right panels represent the Z_H, Z_{DR}, and K_{DP} observations, respectively, that were used to estimate n_{0H} in experiment para5_ZhZdrKdp. Initial parameter values for this experiment are (n_{0R}, n_{0S}, n_{0H}, ρ_S, ρ_H) = (3×10^6 m^{-4}, 7×10^5 m^{-4}, 4×10^5 m^{-4}, 50 kg m^{-3}, 400 kg m^{-3}). 109

Fig. 4.12. The mean estimated parameter (left penels) and spread (right panels) for (a),(b) n_{0H}; (c),(d) n_{0H}; and (e),(f) ρ_H for 3-parameter estimation experiments para3_Zh (thick solid gray), para3_ZhZdr (dotted black), para3_ZhKdp (solid black), and para3_ZhZdrKdp (dashed black). The horizontal thick dotted gray lines in the left panels indicate the truth parameter values. In these experiments, n_{0R}, n_{0H} and ρ_H were estimated while n_{0S} and ρ_S were kept at their incorrect initial values throughout the assimilation cycles. The mean and spread are computed from 10 experiments starting from two sets of imperfect parameter values (n_{0R}, n_{0S}, n_{0H}, ρ_S, ρ_H) = (3×10^6 m^{-4}, 7×10^5 m^{-4}, 4×10^5 m^{-4}, 50 kg m^{-3}, 400 kg m^{-3}) and (3×10^6 m^{-4}, 3×10^7 m^{-4}, 4×10^5 m^{-4}, 300 kg m^{-3}, 400 kg m^{-3}). 111

Fig. 4.13. As Fig. 4.3 but for experiments para3_Zh (black dashed) and para3_ZhKdp (solid black). The RMS errors are averaged over 10 experiments starting from two sets of initial parameter values as given in the caption of Fig. 4.12. The RMS errors of experiment para5_Zh are shown in thick solid gray for comparison. 113

Fig. 5.1. Reflectivity (thin solid contours and shading), perturbation potential temperature (dotted contours at 0.5-K intervals for negative potential temperature) and horizontal perturbation wind (vectors, plotted every four grid point; m s^{-1}) at z = 250 m for the simulated storm for (a)-(d) SM and (e)-(h) DM

runs. AA' and BB' show the locations of the vertical cross-sections passing through the updraft core (maximum vertical velocity). 130

Fig. 5.2. The vertical cross-sections of simulated (a) Z_{DR} (dB) along line AA' shown in Fig. 5.1h corresponding to x = 38.2 km and (b) K_{DP} (° km^{-1}) along line BB' corresponding to y = 32.2 km of the simulated supercell storm at 100 min. The 0°C isotherms are shown as thick black lines. 132

Fig. 5.3. (a)-(c) Z_{DR} (shading, dB) and reflectivity (solid contours at 15-dBZ intervals, starting from 10 dBZ), (d)-(f) K_{DP} (shading, ° km^{-1}) and horizontal perturbation wind (vectors, plotted every other grid point; m s^{-1}), and (g)-(i) hail mixing ratio q_h (shading, g kg^{-1}) and rain mixing ratio q_r (solid contours at 1.0-g kg^{-1} intervals, starting from 1.0 g kg^{-1}) at (a),(d),(g) z = 3 km, (b),(e),(h) z = 4 km, and (c),(f),(i) z = 5 km at 100 min of storm. The vertical velocity (dotted contours at 5-m s^{-1} interval from 10 m s^{-1}) is overlaid on each plot. 133

Fig. 5.4. Z_{DR} (shading), q_r (solid contours at 0.5 g kg^{-1} intervals, starting from 0.5 g kg^{-1}), and a) mean-mass diameter of rain drops D_{nr} (dotted black contours at intervals of 0.1 mm, starting from 0.1 mm) for the SM simulation and b) D_{nr} (intervals of 0.3 mm, starting from 0.3 mm) for the DM simulation at z = 500 m at 80 min. of the storm. 135

Fig. 5.5. (a) Z_{DR} (shading), Z_H (solid contours at 15 dBZ intervals from 15 dBZ), q_h (dotted contours at 1.0 g kg^{-1} intervals from 0.5 g kg^{-1}), and D_{nr} (thick dashed contours at intervals of 0.6 mm, starting from 0.9 mm) at z = 4 km, and (b) ρ_{hv} (shading and solid contours at 0.01 intervals starting from 0.98 and lower) and the ratio of rain-hail mixture to the sum of rain and dry hail mixing ratios (dotted contours at 0.1 intervals from 0.2) at z = 3 km at 80 min. of the storm. Ring features are prominent at mid-levels. 137

Fig. 5.6. Simulated ρ_{hv} with size of melting hailstones for the fraction of water, f_w = 0.45 (dotted), 0.65 (solid), and dry hailstone (dashed). 138

Fig. 5.7. The hail signature is characterized by high Z_H (thin solid contours at 15 dBZ intervals, starting from 15 dBZ) and low Z_{DR} (shading, dB) at z = 2.5 km at 60 min. of storm. The hail mixing ratio is overlaid in thick dashed contours at 1 g kg^{-1} intervals from 1 g kg^{-1} 140

Fig. 5.8. As in Fig. 5.5b, but for simulation using the observation operator of Chapter 2. The Mie scattering effect is not included in this simplified version. 142

Fig. 6.1. The perturbation wind (vectors; m s^{-1}, plotted at every four grid point), perturbation potential temperature (thick black lines for 0 K and dotted contours at 0.5-K intervals for negative potential temperature) and simulated reflectivity (solid contours and shading at 10 dBZ interval, starting from 10 dBZ) at z = 250 m at 40, 60, 80, and 100 min of a supercell storm simulation: (a)-(d) truth, EnKF analysis of (e)-(h) experiment Vr; (i)-(l) Zh; and (m)-(p) VrZh. 156

Fig. 6.2. The ensemble mean forecast and analysis RMSEs averaged over points at which the true reflectivity is greater than 10 dBZ for (a) u, (b) v, (c) w, (d) perturbation potential temperature θ', (e) q_v, (f) q_c, (g) q_r, (h) q_i, (i) q_s, (j) q_h, (k) Nt_c, (l) Nt_r, (m) Nt_i, (n) Nt_s, and (o) Nt_h for experiments VrZh (solid, inflation = 20%), Vr (dotted, inflation = 7%), and Zh (dashed, inflation = 20%). The vertical straight-line segments in the curves correspond to the reduction or increase in RMSEs by the data assimilation. 158

Fig. 6.3. Rain water mixing ratio (solid contours and shading at intervals of 0.5 g kg^{-1}, starting from 0.0 g kg^{-1}) and total rain water number concentration (dotted contours at intervals of 40 m^{-3}, starting from 40 m^{-3}) at z = 1 km and t = 100 min for (a) truth simulation, (b) experiment Vr, (c) Zh, and (d) VrZh. 161

Fig. 6.4. The horizontally averaged ensemble mean analysis RMSEs averaged over points at which the truth reflectivity is greater than 10 dBZ for (a) u, (b) v, (c) w, (d) perturbation potential temperature θ', (e) q_v, (f) q_c, (g) q_r, (h) q_i, (i) q_s, (j) q_h, (k) Nt_c, (l) Nt_r, (m) Nt_i, (n) Nt_s, and (o) Nt_h for experiments VrZh (solid, inflation

= 20%), Vr (dotted, inflation = 7%), and Zh (dashed, inflation = 20%) at 100 min. 162

Fig. 6.5. Forecast error correlations (thick solid and dotted contours at interval of 0.2) computed from the ensemble for experiment VrZh at t = 100 min in the x-z plane at y = 33 km, which passes through the updraft core (maximum vertical velocity). Thick solid (dotted) contours represent positive (negative) error correlations between forecast reflectivity Z_H at x = 37.5 km and z = 8 km (indicated by the boldface X) and (a) q_r, (b) q_c, (c) q_h, (d) w, (e) Nt_r, (f)) Nt_c, (g) Nt_h, and (h) θ'. The shading and thin solid contours are model fields from the truth simulation with units of (a)-(c) g kg^{-1}, (d) m s^{-1}, (e)-(g) m^{-3}, and (h) K. Zero correlation lines are suppressed. 163

Fig. 6.6. As in Fig. 6.5, but between forecast w and (a) q_r, (b) q_s, (c) q_h, (d) q_c, (e) Nt_r, (f) Nt_s, (g) Nt_h, and (h) Nt_c. 165

Fig. 6.7. The ensemble mean analysis RMSEs of experiments VrZhZdr (dotted), VrZhKdp (dashed), and VrZhZdrKdp (solid) normalized by those of experiment VrZh. The reference horizontal line at unity is overlaid. 168

Fig. 6.8. As in Fig. 6.2, but for experiments VrZh_ptrα_IM (thin dotted black) and VrZh_cstα_IM (thick solid gray).The RMSEs of the control run, which is VrZh with $\alpha_r = \alpha_h = 0$, are shown in thin solid black for comparison. 173

Fig. 6.9. As in Fig. 6.2, but for experiments VrZh_obs_IM (solid), Vr_obs_IM (dotted), and Zh_obs_IM (dashed). The RMSEs are computed with respect to the true simulation performed using three-moment microphysics scheme. The RMSEs of experiments VrZh (thick dotted gray) and VrZh_ptrα_obs_IM (thin solid gray) are shown for comparison. 176

Fig. 6.10. As in Fig. 6.1, but for a supercell storm simulation: (a)-(d) truth with TM of MY05 and EnKF analysis of (e)-(h) experiment VrZh_obs_IM. Reflectivity is computed using the radar simulator developed in Chapter 5. 178

Fig. 7.1. Reflectivity (solid contours and shading at 10 dBZ interval, starting from 10 dBZ) at 1.25° elevation at (a),(b) 0040 UTC and (c),(d) 0100 UTC from the KTLX's perspective for (a),(c) observations and (b),(d) ensemble mean analysis. 188

Abstract

The US National Weather Service plans to upgrade the entire operational radar network to polarimetric capability early in the next decade. The goal of this dissertation is to develop methodologies that use polarimetric radar data in mesoscale data assimilation systems with the ensemble Kalman filter (EnKF). To directly assimilate polarimetric radar variables into numerical weather prediction models, two sets of comprehensive polarimetric radar simulators are developed. The observed quantities considered include reflectivities at horizontal and vertical polarizations (Z_H and Z_V, respectively), the differential reflectivity Z_{DR}, the specific differential phase K_{DP}, and the cross-correlation coefficient ρ_{hv}. The simulators are applied to a simulated squll-line system and a supercell storm. Simulated fields exhibit realistic polarimetric signatures that include Z_{DR} and K_{DP} columns, Z_{DR} arc, mid-level Z_{DR} and ρ_{hv} rings, and hail signatures along with the bright band signature in the stratiform region and deep reflectivity core in the convective region. These simulators are proven to be useful for validating model microphysics by 1) detecting a problem with the treatment of melting processes and 2) manifesting the limitation of a single-moment microphysics scheme in handling the mechanisms closely linked to the size sorting of precipitation particles.

The simpler but more efficient version of the polarimetric simulators is incorporated into the EnKF data assimilation system to define the observation operators. Various experiments are performed to: (1) assess the impact of assimilating additional polarimetric variables in the EnKF system based on a single-moment scheme, (2) simultaneously estimate fundamental microphysical parameters and atmospheric state in the presence of observation operator error due to uncertainty in the microphysical

parameters, (3) evaluate the capability of the EnKF system to retrieve model state variables using both conventional and polarimetric radar data when a double-moment microphysics scheme is used.

The results based on observing system simulation experiments assuming either a perfect or imperfect prediction model show that the storm analysis is improved when polarimetric variables are assimilated in addition to Z_H or in addition to both Z_H and radial velocity V_r. The additional polarimetic data are more helpful when state and/or parameter estimation is not very successful using conventional radar data alone. The positive impact of polarimetric data is found in all state variables at all levels and more significant improvement is obtained in microphysical variables when the model error originating from the uncertain microphysical parameters is relatively large. The polarimetric variables are generally more beneficial to the analysis of rain water- and hail-related variables, and of the vertical velocity among the wind comonents.

It is also suggested that the polarimetrc variable that contains the most independent information content has the largest positive impact on the analysis. Between Z_{DR} and K_{DP}, Z_{DR} is found to produce a larger positive impact in perfect model experiments because Z_{DR} is less correlated with Z_H than K_{DP} is. In the microphysical parameter estimation experiments, K_{DP} is generally more useful where the data are mostly selected from discrete regions of the storm while many of Z_{DR} data seem to overlap Z_H in location. The microphysical information provided by additional polarimetric radar data appears to help alleviate the solution non-uniqueness problem. This research is, to our knowledge, the first to directly assimilate (simulated) polarimetric radar data into a numerical weather prediction model.

Chapter 1 Introduction and Overview

1.1 Background and Motivation

Because the atmosphere is a chaotic system, the predictability of numerical weather prediction (NWP) is limited by small errors in model initial conditions; they amplify with time, leading to divergence far from the true state (Lorenz 1963; Hohenegger and Schar 2007). Therefore, considerable efforts have been made to find a more accurate representation of the current state of atmosphere to initialize numerical models. Currently, the four-dimensional variational method (4DVAR) and the ensemble Kalman filter (EnKF) are the two leading advanced data assimilation (DA) techniques that can provide "optimal" initial conditions that best fit the observations and the forecast model, subject to their respective uncertainty. These methods can take advantage of model dynamic equations when assimilating observations distributed in space and time.

For convection-resolving NWP, the microphysics scheme is one of the most important physical processes and has a profound impact on the forecast. Caya et al. (2005) compared performance between the 4DVAR and the EnKF when they are used to assimilate simulated radar data using a cloud model with warm rain microphysics; they showed that, in general, the 4DVAR produces better analyses in the early assimilation cycles, while the EnKF performs better in the later cycles [In that study, simulated radial velocity V_r and rainwater mixing ratio q instead of reflectivity Z were assimilated, and the initial guess had no knowledge of the convective storm]. Considering the complexity and highly nonlinear nature of microphysical processes

when ice species are involved, the EnKF method, which uses the full nonlinear model to propagate the ensemble states, appears to be more attractive, while the linearization required by 4DVAR in the minimization process often encounter difficulties (e.g., Xu 1996).

Radar is currently the only platform that can provide observations with high temporal and spatial resolutions suitable for convective storm initialization. Very encouraging results have been obtained by several studies with EnKF and simulated Doppler radar data in recent years. Studies by Snyder and Zhang (2003) and Zhang et al. (2004), using radial velocity data with a warm rain microphysics scheme, and Tong and Xue (2005b) and Xue et al. (2006), using both radial velocity and reflectivity with a 3-ice microphysics, show that the model state variables can be accurately retrieved using the EnKF method under the perfect model assumption although serious degradation of the filter performance can be encountered in real data assimilation where model error inevitably exists. The results of Dowell et al. (2004) show that the wind field retrieved using the EnKF from real single Doppler radar observations is generally comparable to that obtained by conventional dual Doppler wind analysis for a supercell storm case.

More recently, there have been efforts to correct model error using the EnKF method, where the model errors originate from uncertainties in the model parameters (Aksoy et al. 2006a, b; Tong and Xue 2008b, a). In the EnKF framework, the parameters to be estimated are treated in a similar manner as the model state variables are; they are estimated simultaneously with state variables during the assimilation cycles. These studies, in which the only source of model errors is in these parameters and in which the observation operators do not contain error, show that the EnKF is

generally effective in retrieving these uncertain parameters and that parameter estimation is beneficial to state estimation.

Even though the EnKF method has shown great promise for storm-scale DA with simulated reflectivity and radial velocity data from conventional radars, such data may not be sufficient for fully determining the microphysical states in real-data cases where significant model error may degrade the filter's ability to develop reliable covariance. This is especially true when the microphysics involves many cloud and hydrometeor species and many uncertain parameters. Therefore, additional observations available from polarimetric radar (PR) can be very helpful, especially when they contain microphysical information such as the density, phase, shape, and drop size distributions (DSDs) of hydrometeors.

The benefit of polarimetric data has been well demonstrated in quantitative rainfall estimation (e.g., Seliga and Bringi 1976; Ryzhkov et al. 1998; Zrnic and Ryzhkov 1999) and precipitation type classification (e.g., Straka 1993; Vivekanandan et al. 1999). Since PR data contain microphysical information, it is hypothesized that they can be used to improve the state estimation of microphysical fields, and hence, subsequent forecasts resulting from such initial conditions.

The direct assimilation of PR data into NWP models has not gained much attention until recently. This is partly due to the lack of PR data from operational weather radars, and partly because techniques and methods for effectively assimilating such data need development. A good understanding on the PR data by the NWP community has been generally lacking and observation operators needed for the direct assimilation did not exist. To our knowledge, the only paper in the archived literature

that uses PR data for model initialization is that of Wu et al. (2000), which assimilated using a 4DVAR technique rain and ice water mixing ratios pre-derived from the reflectivity and differential reflectivity. However, the planned polarimetric upgrade of the US WSR-88D network later this decade and the PR data now available from the WSR-88D PR upgrade prototype, the KOUN radar, and from the recently established Oklahoma Testbed of the NSF Engineering Research Center (ERC) for Collaborative Adaptive Sensing of the Atmosphere (CASA) (Brotzge et al. 2005) provide important motivations for research in this direction.

1.2 Dissertation Overview

The goal of this study is to develop an atmospheric state and parameter estimation system based on the ensemble Kalman filter method that makes effective use of polarimetric radar data, and to investigate the impact of additional polarimetric variables on convective storm analysis and forecasting. Because of the unavailability of forward observation operators for polarimetric variables that are complete and efficient enough for DA use, a set of observation operators are first developed. As this is the first study that directly assimilates polarimetric data into the model, simulated data are used through OSSEs (Observing System Simulation Experiments), and the impacts of the additional PR data are examined. The Advanced Regional Prediction System (ARPS) is used throughout this study to provide truth simulations and within the EnKF DA. The particular EnKF algorithm used is the ensemble square-root filter (EnSRF) (Whitaker and Hamill 2002).

In Chapter 2, we describe a radar simulator for polarimetric variables that we developed based on the parameterization and fitting of calculations of electromagnetic wave propagation and scattering for S-band radars. This simulator provides the forward observation operators that are needed when directly assimilating polarimetric variables, including reflectivity, differential reflectivity, specific differential phase, and co-polar cross-correlation coefficient. The behavior of the simulator is demonstrated by simulating radar observations for a multicell squall line and a supercell storm. This simulator provides the forward observation operators used in various data assimilation experiments throughout this dissertation.

Chapter 3 discusses the impact of assimilating additional polarimetric observations on convective storm analysis in an OSSE framework. A new error model is introduced for characterizing the errors of conventional and polarimetric radar variables. The necessity of PR data thresholding prior to assimilation to minimize the impact of noise is also discussed. The improvements obtained by assimilating polarimetric measurements in addition to reflectivity or in addition to both reflectivity and radial velocity data are quantitatively assessed.

In Chapter 4, we investigate the impact of polarimetric data on the simultaneous estimation of uncertain model microphysical parameters and atmospheric state where those uncertain parameters are also involved in the observation operators. This implies that the polarimetric data are used to correct both model error and observation operator error where the sources of error are from the incorrect microphysical parameters. The microphysical parameters to be estimated are the intercept parameters of the exponential DSD of rainwater, snow, and hail, as well as the densities of snow and hail,

because these parameters have large uncertainties and have profound impact on the convective-scale forecast. The polarimetric data are believed to be especially valuable for such parameter estimations because of their information content. Sensitivity analyses are performed to help understand the EnKF behavior in the presence of observation operator error. The improvement obtained by using additional polarimetric data in parameter estimation is also quantitatively assessed.

A more general PR emulator is developed in Chapter 5 based on rigorous scattering calculations using the transition matrix (T-matrix) method for PR variables. This emulator is able to simulate polarimetric radar measurements at weather radar frequency bands and can take as input the prognostic variables of high-resolution model simulations using one-, two-, or three-moment (SM, DM, and TM, respectively) microphysics schemes, which were recently implemented in the ARPS. The new emulator is tested at 10.7 cm of wavelength with a model-simulated supercell storm using a DM scheme, and the emulator is found to simulate unique polarimetric signatures reported in the observational studies while the SM simulations could not reproduce signatures closely linked to the size sorting of precipitation drops in the storm. Various potential practical applications of the PR simulator are also discussed.

Results in Chapter 5 suggest that supercell thunderstorms simulated using a two- or higher-moment microphysics scheme is more realistic. Motivated by these results, we explore in Chapter 6 the capability of the EnKF system in estimating state variables, including both the water/ice mixing ratios (third moment of DSD) and the total number concentrations (zeroth moment of DSD), which are also state variables when using a DM scheme. Several sets of experiments are performed to test the performance of the

EnKF system in the presence of model errors in both forecast model and observation operator. The impact of PR data on an analysis employing a DM scheme is also investigated.

Chapter 7 summarizes the dissertation and outlines a future study where the current EnSRF system will be extended to real-data scenarios employing a DM scheme.

Chapter 2 [1] Observation Operators for Reflectivity and Polarimetric Variables

2.1 Introduction

Modern data assimilation (DA) techniques such as 3D and 4D variational data assimilation (3DVAR and 4DVAR, respectively), and ensemble Kalman filter (EnKF) methods are able to assimilate observations directly using the forward observation operators that link the model state variables to the observations (Kalnay 2002). The goal of DA is to minimize, subject to the constraint of observation uncertainty, the difference between the observations and the analysis projected to the observation space using the observation operator. The forward operators also play a role of observation simulator in the Observing System Simulation Experiments (OSSEs) in generating simulated observations (e.g, Xue et al. 2006). The observation operators can also be used to verify model prediction against indirect, often remote-sensed, observations (e.g., Otkin et al. 2007).

For Doppler weather radars like the Weather Surveillance Radar-1988 Doppler (WSR-88D), the radial velocity and equivalent radar reflectivity factor (hereafter reflectivity) data are the two key measurements that can be assimilated into NWP models (e.g., Hu et al. 2006a; Hu et al. 2006b). The observation operators for the radial velocity and reflectivity link the model velocity components to the observed radial velocity and the model hydrometeor fields to the observed reflectivity, respectively

[1] This chapter is an extended version of our paper: Jung, Y., G. Zhang, and M. Xue, 2008: Assimilation of simulated polarimetric data data for a convective storm using ensemble Kalman filter. Part I: Observation operators for reflectivity and polarimetric variables. *Mon. Wea. Rev.,***136**. 2228-2245.

(Tong and Xue 2005b; Xue et al. 2006). They also should take into account other effects that are necessary for realistic observation, such as the Earth curvature effect or the radar beam pattern (Tong and Xue 2005a; Xue et al. 2006).

For reflectivity, the observation operator also depends on the microphysical parameterization schemes used in NWP model. Smith et al. (1975), Smith (1984), Ferrier (1994), Caumont et al. (2006), and Haase and Crewell (2000) all offer formulas that calculate reflectivity from liquid and ice phase hydrometeors present in bulk microphysics schemes. Various assumptions on the drop size distributions (DSDs) and shapes of liquid and ice particles, radar beam pattern and wavelength, and the way that backscattering cross sections are computed are involved in developing those formulas for radar simulators. Some methods are more sophisticated and computationally expensive than others. Among them, Caumont et al. (2006) developed the most general simulator with various options for X-, C-, and S-band radars based on Rayleigh, Rayleigh-Gans, Mie, and T-matrix scattering methods. However, no continuous melting process is considered in these simulators except for Ferrier (1994), which uses the mixing ratios of liquid water on wet precipitation particles that are predicted in the forecast model. May et al. (2007) is a pulse-based radar emulator that emphasizes the simulation of radial velocity and its spectral width.

Even though reflectivity and radial velocity measurements provide key information on convective storms, they are not sufficient to fully describe microphysical states. One of the reasons is that the number of observations is usually much smaller than the degrees of freedom of the forecast model or even the microphysics model alone. This means that we need to determine more model variables with fewer observations.

The other reason has to do with many uncertainties in the bulk microphysics schemes. The microphysics represents one of the most important physical processes at the convective scale. The microphysical processes depend to a large extent on the phase, shape, orientation, density, and DSDs of microphysical species involved, many of which are not fully understood. These properties also directly affect radar measurements within each radar sampling volume. Additional observational parameters available from polarimetric Doppler radars, including differential reflectivity and differential phase measurements can be very helpful here as they contain information about the density, shape, orientation, and DSDs of hydrometeors (Doviak and Zrnic 1993; Bringi and Chandrasekar 2001).

Some polarimetric radar simulators already exist in the literature (Brandes et al. 1995; Brandes et al. 2004b; Zhang et al. 2001; Vivekanandan et al. 1994; Ryzhkov et al. 1998; Huang et al. 2005). However, they are either incomplete in terms of utilizing all available model parameters and state variables or are too expensive for use within DA systems. Within a DA system, the simulation needs to be performed for each observation, and repeated within a variational minimization scheme. Some of the previous studies have focused on single-phase hydrometeor concentration. Brandes et al (1995; 2004b), Zhang et al (2001) offer the expressions for rain. Vivekanandan et al. (1994) and Ryzhkov et al. (1998) propose formulae that can be applied to ice particles ranging from ice crystals to snow aggregates. Because of the lack of sufficient understating of the polarimetric measurements for ice and mixed phases because of their complex behaviors and non-linear interactions, general expressions that are applicable to each of the hydrometeor categories are generally unavailable. More recently, Huang

et al. (2005) proposed a more complete polarimetric radar simulator in which a full radar scattering model is used to simulate polarimetric radar signatures from the data of a model-simulated storm. Such simulators are, however, too expensive for DA use.

In this study, we develop a set of the observation operators consistent with a commonly used three-ice microphysics scheme. The polarimetric variables include reflectivities at the horizontal (Z_H) and vertical (Z_V) polarizations, differential reflectivity (Z_{DR}), reflectivity difference (Z_{dp}), specific differential phase (K_{DP}), and the cross-correlation coefficient at zero lag ($\rho_{hv}(0)$). These operators are applicable to the S-band radar at about 10.7 cm of wavelength and can be extended in the future possibly for other wavelengths. Having such a system of our own enables us to adjust and enhance the simulator to fit our data assimilation needs, and in response to the changes with the microphysics parameterization used in the assimilation and prediction model. In fact, these operators are used in Chapter 3 to test the impact of simulated polarimetric observations on the storm analysis.

In section 2.2, the prediction model used to create the simulation datasets is briefly described. The forward observation operators for the polarimetric radar variables associated with microphysics schemes with varying degrees of assumptions are then developed in section 2.3. These observational operators are then applied to a simulated squall line and a supercell storm in section 2.4. Conclusions and discussions are given in section 2.5.

2.2 The model and convective storm simulations

The Advanced Regional Prediction System (ARPS, Xue et al. 2000; 2001; 2003) is used to produce convective storm simulations, of a squall and a supercell, that are used to test our radar emulator. The reflectivity-related formulas are also closely related to the microphysics scheme used in the model. Briefly, ARPS is a fully compressible and nonhydrostatic atmospheric prediction model. The model state vector consists of three velocity components u, v, w; potential temperature θ; pressure p; and the mixing ratios for water vapor, cloud water, rainwater, cloud ice, snow aggregate, and hail (q_v, q_c, q_r, q_i, q_s, and q_h, respectively) when the ice microphysics scheme based on Lin et al. (1983, hereafter LFO83) is used. The model also predicts the turbulence kinetic energy which is used by the 1.5-order subgrid-scale turbulence closure scheme.

An idealized two-dimensional squall-line system is initiated by a 4-K ellipsoidal thermal bubble with a 10-km horizontal radius and a 1.4-km vertical radius, and the bubble is centered at x = 400 km, z = 1.4 km in the 700 ×19.2 km^2 physical domain. The horizontal grid spacing is 200 m and the vertical grid has a uniform 100-m grid spacing in the lowest 3 km, which then increases to 853 m at the model top. The simulation is run for 12 h with the analytic thermodynamic sounding defined by Weisman and Klemp (1982), where the potential temperature and temperature are 343 and 213 K, respectively, at the 12-km-high tropopause, and the surface potential temperature is 300 K. The mixed-layer mixing ratio is 15 g kg^{-1}, the upper limit of relative humidity is 95 %, and the mixed layer depth is 1.2 km. The environmental wind profile has a constant shear of 17.5 m s^{-1} in the lowest 2.5 km and a constant wind speed of -2.5 m s^{-1} above 2.5 km. These configurations are similar to those used in Xue (2002),

with the main differences being the wind profile and horizontal resolution. This specified environmental condition generally supports long-lived squall lines that sometimes develop a trailing stratiform precipitation region (Thorpe et al. 1982; Rotunno et al. 1988).

For a more intense, isolated supercell storm simulation, ARPS is initialized with the environmental sounding of the 20 May 1977 Del City, Oklahoma, supercell storm (Ray et al. 1981). The CAPE of the sounding is 3300 J kg^{-1} and the storm is initiated by an ellipsoidal thermal bubble with same characteristics as that of the squall-line case except for a vertical radius of 1.5 km. The bubble is centered at x = 48 km, y = 16 km, and z = 1.4 km. The physical domain is 64×64×16 km^3 with a horizontal spacing of 2 km and a vertical separation of 0.5 km. Open conditions are used at the lateral boundaries and free-slip conditions at the top and bottoms of domain. A constant wind of u = 3 m s^{-1} and v = 14 m s^{-1} is subtracted from the original sounding to keep the storm near the center of the domain. These configurations are essentially the same as in the truth simulation of Tong and Xue (2005b), which also briefly describes the initial evolution of the simulated storm. This simulation serves as the truth simulation for the polarimetric data assimilation experiments in Chapter 3. A polarimetric WSR-88D radar is assumed at the southwest corner of the domain, the same location as that assumed in Tong and Xue (2005b).

2.3 The observation operators and simulation of observations

As discussed earlier, a set of forward observation operators that link model state variables with the polarimetric radar variables is required to assimilate the latter into a

numerical model. These operators, together with the radar scanning configurations, ray path, and beam pattern weighting, make up a complete radar simulator. This chapter focuses on the observation operator development. For these operators, a consistency is maintained between the DSD-related parameters of hydrometeors within the operators and within the prediction model. The specific polarimetric radar variables to be considered include reflectivity, differential reflectivity, reflectivity difference, specific differential phase, and the cross-correlation coefficient.

2.3.1 The shape, orientation, and drop size distribution of hydrometeors

The model state variables are projected into the observation space using the observation operators. In this study, we assume that radar observations are taken and available on the original radar elevation levels vertically but are already interpolated onto horizontal model grids, as is done in Xue et al. (2006), which describes the power-gain-based sampling method used in the vertical direction in detail. In the single-moment bulk ice microphysics scheme of LFO83 used in the ARPS, a constant density is assumed for each species and the DSDs of the species are modeled by exponential distributions with fixed intercept parameters (n_0) and variable slopes (Λ). In practice, the slope Λ for each species is diagnosed from the corresponding specified intercept parameter and the predicted mixing ratio. The intercept parameters for rain, snow, and hail used in this study are the default values of $n_{0R} = 8\times10^6$ m^{-4}, $n_{0S} = 3\times10^6$ m^{-4}, and $n_{0H} = 4\times10^4$ m^{-4} (LFO83).

Additional characteristics that affect the radar observables include the shape, orientation and the ice/water fraction of hydrometeors. Unfortunately, these

characteristics are not specified or predicted by the model; therefore, assumptions have to be made. Observations show that larger raindrops (> 1 mm in diameter) are not spherical. Raindrops are normally modeled as oblate spheroids and the oblateness, r, is represented by the axis ratio between minor to major axis, which is related to the equivalent diameter D given by Green (1975) in an equilibrium model. After solving the equilibrium equation and fitting to a polynomial function, Zhang et al. (2001) obtained

$$r = 1.0148 - 2.0465\times10^{-2} D - 2.0048\times10^{-2} D^2 + 3.095\times10^{-3} D^3 - 1.453\times10^{-4} D^4, \quad (2.1)$$

where D is in millimeter. This axis ratio relation has recently been revised based on observations (Brandes et al. 2002), yielding more spherical shapes for smaller drops (1 < D < 4 mm). The potential errors associated with more oblate shapes are about 0.15 dB and 0.2 dB for Z_H and Z_{DR}, respectively, in terms of averaged values (Brandes et al. 2002). However, the revised formula requires the numerical integration over the DSD in the scattering calculations, which significantly increases computational cost while the former allows for analytical integration. Although the revised axis ratio is important in the quantitative precipitation estimation for light rain with many small drops, it is not crucial for assimilation purposes. Also, there is no accepted theory that explains the revised axis ratio relation. Therefore, we use the equilibrium shape in (2.1) in this study. We also assume that the mean and the standard deviation (SD) of the canting angle are 0º, as suggested by observations (Hendry and McCormick 1976), although some observational and theoretical studies suggest that the standard deviation of the canting angles of rain drops is likely not 0° but less than 10° (Beard and Jameson 1983; Bringi and Chandrasekar 2001; Ryzhkov et al. 2002). Assuming 0° SD can lead to the

overestimation of K_{DP} and Z_{DR} by less than 6%, this could be tolerated considering the large uncertainties in DSD (Ryzhkov et al. 2002).

The shape of snow can vary greatly in range/complexity and can be modeled as oblate to prolate spheroids. Nevertheless, in the simplest form, they can be approximated to fall with the major axis aligned horizontally (Straka et al. 2000). The mean canting angle of snow aggregates is assumed to be 0^{o} and the SD of canting angle is assumed to be 20^{o} in this study. A fixed axis ratio of 0.75 for snow is used for the scattering calculation. Also, a fixed density of 100 kg m^{-3} is assumed for dry snow aggregates, consistent with the model parameterization.

Hailstones are observed in many different shapes and the orientation of falling hail is not understood precisely. Yet, ground observations suggest that the majority of hailstones have axis ratios of 0.8 with spongy hail having a lower axis ratio of 0.6 to 0.8 (Knight 1986; Matson and Huggins 1980) and usually wobble and tumble while they fall. Dry hailstones are considered to have random orientations. Therefore, we assume that the axis ratio of hailstones is 0.75 and hailstones fall with their major axes aligned horizontally with a mean canting angle of 0^{o}, although some studies (Aydin and Zhao 1990; Vivekanandan et al. 1993) use other canting angles that are not widely used. As a hailstone melts while falling, meltwater forms a torus around equator and stabilize these wobbling and tumbling motions. The SD (or σ) of canting angle is therefore parameterized as a function of the fractional water content in melting hail, according to $\sigma = 60^{o}(1 - cf_{w})$, where f_{w} is the water fraction within water-hail mixtures (see more

later) and c is a coefficient equaling 0.8 except for very low mixing ratios of the mixture (q_{rh}). This allows dry (wet) hailstones to have large (small) SD of canting angles.

When the hail mixing ratio is low, we expect more small hail, therefore more spherically shaped particles, leading to smaller Z_{DR}. A fixed axis ratio, assumed in our model, can lead to high Z_{DR} for low hail mixing ratio when hail is in the melting phase. To take into account the size dependence of the axis ratio, we set a critical value of q_{rh} (0.2 g kg^{-1}), below which the constant c is decreasing as a function of q_{rh}, thus, effectively reducing Z_{DR}. This gives the same effect by assuming more spherical hail for low hail mixing ratios. When q_{rh} < 0.2 g kg^{-1}, it is therefore assumed that $c = 4\,q_{rh}$. As in the ARPS model, the hail is assumed to have a fixed density of 913 kg m^{-3}. Our hail model, although different in configuration, is consistent with observed Oklahoma hailstones, which show a general trend of decrease in axis ratio with increasing size until reaching a value of about 0.75 (Knight 1986).

2.3.2 Melting ice (snow-hail) model

As the snow aggregate melts, the water forms a thin layer on the surface of snow aggregate and/or distributes either evenly or nonuniformly within the snow aggregate, effectively forming snow aggregate-liquid water mixture, where the snow aggregate itself is a mixture of solid ice and air. We allow continuous melting of low-density dry snow to rain in the melting layer where both rain and snow mixing ratios are nonzero. We denote the mixing ratio of the rain-snow mixture as $q_{rs.}$ Within $q_{rs,}$ a fraction, f_w, is water and a fraction, f_s, is snow, and of course, $f_w + f_s = 1$. However, as with most microphysics schemes in use today, the LFO83 microphysics scheme used in the ARPS

does not allow or track species in the mixture form. For example, the melt part of snow aggregate is immediately removed from q_s and added to q_r. Therefore, the amount and composition of mixture-form species have to be modeled in a way that allows for realistic radar observation simulations.

In this study, we model the rain-snow mixture in the following way. First, as mentioned earlier, the mixture is assumed to exist only when q_s and q_r coexist. We assume that the fraction of rain-snow mixture reaches a maximum when the snow and water mixing ratios are equal and decreases to zero when one of the two vanishes. Furthermore, we assume this fraction, denoted as F, is the same for snow aggregates and rainwater. The fraction F is then given by

$$F = F_{max}\left[\min(q_s/q_r, q_r/q_s)\right]^{0.3}, \tag{2.2}$$

where F_{max} is the maximum fraction of snow or rainwater mixing ratio existing in the mixture form, or the maximum value of F. In this study, we set $F_{max} = 0.5$. A power of 0.3 is taken of $\min(q_s/q_r, q_r/q_s)$. In the case that q_s decreases linearly downward and rainwater increases linearly upward through the melting layer, the F profile has a bracelike shape, with its value peaking near the middle of the melting layer where $q_s = q_r$; otherwise the function has a triangular shape with an apex in the middle. With more realistic model-simulated profiles of mixing ratios, this gives parabolically shaped profiles of mixtures, which is reasonable (thick dashed lines in Figs. 2.5a and 2.6a).

Once F is determined, the mixing ratio of rainwater in the mixture form is then Fq_r, and that in the pure water form is $(1 - F)q_r$. For snow aggregates, the corresponding mixing ratios are Fq_s and $(1 - F)\, q_s$. The total mixing ratio of the mixture is then $q_{rs} = F\,(q_r + q_s)$ and within this mixture, the water fraction is

$$f_w = (Fq_r) / (Fq_r + Fq_s) = q_r / (q_r + q_s) . \quad (2.3)$$

According to (2.3), the water fraction (f_w) within the snow-water mixture increases from 0 to 1 as snow completely melts after descending through the melting layer while f_s decreases from 1 to 0; this behavior is reasonable.

A fixed density of 100 kg m^{-3} is assumed for the dry snow aggregate. However, the snow aggregate density varies during melting. The density of wet snow aggregates increases from 100 kg m^{-3} to 1000 kg m^{-3} as the fraction of the melted portion increases from 0 to 1. At the very early stage of melting, the size of the snow aggregate does not change much with increasing f_w so that the density increases slowly. As melting progresses, f_w further increases, the snow particle collapses inducing the shrinkage of the particle, and the density increases more rapidly. To simulate this melting process as the snow aggregate particles descend, the density of melting snow aggregate is parameterized as a quadratic function of f_w:

$$\rho_m = \rho_s (1 - f_w^{\,2}) + \rho_w f_w^2 , \quad (2.4)$$

which is used in our reflectivity calculations.

A dry snow aggregate is a mixture of air and ice whose density is 913 kg m^{-3} and a melting snow aggregate is a mixture of air, ice, and water. The dielectric constant for melting snow aggregate is calculated with a two-step procedure using the Maxwell-Garnett mixing formula (Maxwell-Garnett 1904). In the first step, the ice is considered within the enclosure of air. Because it is reasonable to assume that the melting starts from the surface of ice particle, the air-ice mixture (snow aggregate) is considered within the enclosure of meltwater in the second step.

A similar melting model is used for hail with corresponding density and dielectric constant for hail. The density of wet hail increases from 913 kg m^{-3} to 1000 kg m^{-3}. The dielectric constant for melting hail is calculated with the ice in the water matrix.

2.3.3 Observation operators

Reflectivities in linear scale at horizontal (Z_h) and vertical (Z_v) polarizations are obtained as integrations over the DSD weighted by the scattering cross section depending on density, shape, and DSD. For rain, dry snow, dry hail, rain-snow mixture, and rain-hail mixture, we have (Zhang et al. 2001) the following:

$$Z_{h,x} = \frac{4\lambda^4}{\pi^4 |K_w|^2} \int n(D) \left[A|f_a|^2 + B|f_b|^2 + 2C|f_a||f_b| \right] dD \ (\text{mm}^6\ \text{m}^{-3}) \text{ and} \tag{2.5}$$

$$Z_{v,x} = \frac{4\lambda^4}{\pi^4 |K_w|^2} \int n(D) \left[B|f_a|^2 + A|f_b|^2 + 2C|f_a||f_b| \right] dD \ (\text{mm}^6\ \text{m}^{-3}), \tag{2.6}$$

where

$$A = \left\langle \cos^4 \phi \right\rangle = \frac{1}{8}\left(3 + 4\cos 2\bar{\phi} e^{-2\sigma^2} + \cos 4\bar{\phi} e^{-8\sigma^2}\right),$$

$$B = \left\langle \sin^4 \phi \right\rangle = \frac{1}{8}\left(3 - 4\cos 2\bar{\phi} e^{-2\sigma^2} + \cos 4\bar{\phi} e^{-8\sigma^2}\right),$$

and

$$C = \left\langle \sin^2 \phi \cos^2 \phi \right\rangle = \frac{1}{8}\left(1 - \cos 4\bar{\phi} e^{-8\sigma^2}\right),$$

and x can be r (rain) and rs (rain-snow mixture), ds (dry snow), rh (rain-hail mixture), or dh (dry hail). Here f_a and f_b are backscattering amplitudes for polarizations along the major and minor axes, respectively. Here $\bar{\phi}$ is the mean canting angle and σ is the standard deviation of the canting angle. As defined in section 2.3.1, $\bar{\phi} = 0^o$ is assumed

for all hydrometeor types and $\sigma = 20^{\circ}$ for snow and $\sigma = 60^{\circ}(1 - cf_w)$ for hail. Here c = 0.8, where $q_{rh} \geq 0.2$ g kg^{-1} and $c = 4q_{rh}$ otherwise. The latter is to make the hail shape more spherical for low mixing ratios, as discussed in section 2.3.1. Here $\langle ... \rangle$ represents an ensemble average over canting angles and *n(D)* defines the DSD and is the number of particles per unit volume of air and increment diameter.

Integration over DSD can be easily performed if the backscattering amplitudes are expressed in the power-law form of the particle size D (mm):

$$|f_a| = \alpha_{xa} D^{\beta_{xa}} \text{ (mm)} \quad \text{and} \tag{2.7}$$

$$|f_b| = \alpha_{xb} D^{\beta_{xb}} \text{ (mm)}, \tag{2.8}$$

Here $|f_a|$ and $|f_b|$ are the magnitudes of f_a and f_b, respectively.

For rain, we first calculate the scattering amplitude of oblate rain drops with the dielectric constant of water evaluated at 10 ºC based on the T-matrix method following Zhang et al. (2001). We perform a new fitting because their coefficients produce negative differential reflectivity for small drops. The scattering amplitudes from the T-matrix and the fitting results are plotted in Fig. 2.1. New fits generally agree well with those in Zhang et al. (2001) and with the T-matrix results over the entire range, except for the slightly larger values at the larger drop end. In (2.7) and (2.8), $\alpha_{ra} = \alpha_{rb} = 4.28 \times 10^{-4}$, $\beta_{ra} = 3.04$, and $\beta_{rb} = 2.77$ for rain are adopted from the T-matrix calculation and fitting results.

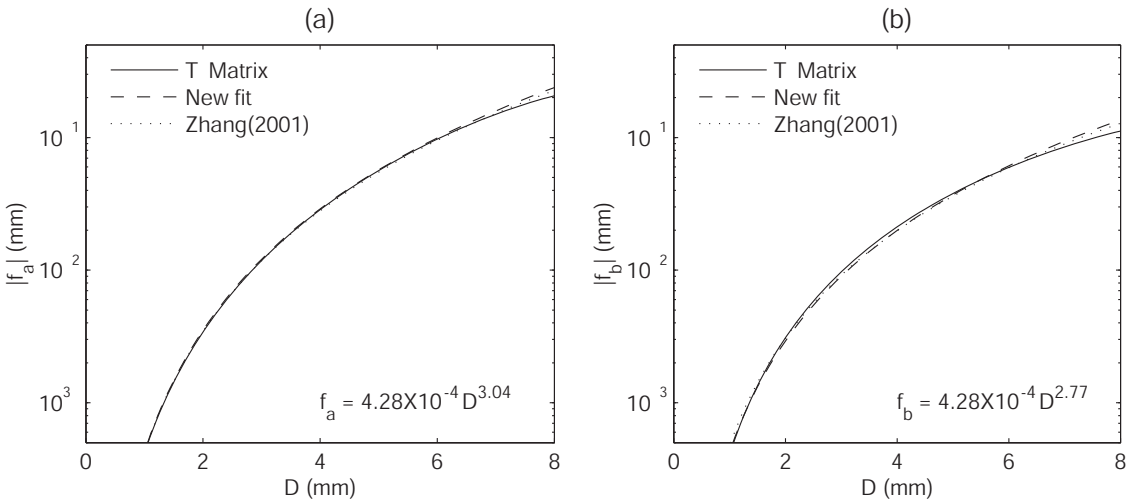


Fig. 2.1. Backscattering amplitudes as a function of the effective diameter of particle along (a) the major axis and (b) the minor axis.

For snow and hail, we calculate the scattering amplitudes as a function of the dielectric constants, which is a function of f_w, and fit the results to the power-law functions given in (2.7) and (2.8). The resultant α_s and α_h for snow and hail are obtained based on the Rayleigh scattering approximation for oblate spheroids, and fitted to polynomial functions of f_w:

$$\alpha_{rsa} = (0.194 + 7.094 f_w + 2.135 f_w^2 - 5.225 f_w^3) \times 10^{-4},$$

$$\alpha_{rsb} = (0.191 + 6.916 f_w - 2.841 f_w^2 - 1.160 f_w^3) \times 10^{-4},$$

$$\alpha_{rha} = (0.191 + 2.39 f_w - 12.57 f_w^2 + 38.71 f_w^3 - 65.53 f_w^4 + 56.16 f_w^5 - 18.98 f_w^6) \times 10^{-3}, \quad \text{and}$$

$$\alpha_{rhb} = (0.165 + 1.72 f_w - 9.92 f_w^2 + 32.15 f_w^3 - 56.0 f_w^4 + 48.83 f_w^5 - 16.69 f_w^6) \times 10^{-3}. \qquad (2.9)$$

The β_s for snow and β_h for hail are equal to 3 at both polarizations. The equations in (2.9) give $\alpha_{dsa} = 0.194 \times 10^{-4}$ and $\alpha_{dsb} = 0.191 \times 10^{-4}$ for dry snow, and $\alpha_{dha} = 0.191 \times 10^{-3}$ and $\alpha_{dhb} = 0.165 \times 10^{-3}$ for dry hail. The scattering amplitudes from Rayleigh scattering approximation and the fitting results as a function of f_w are plotted in Fig. 2.2. As

discussed in section 2.3.2, melting is likely to start from the surface so that the backscattering amplitude increases more rapidly in the early stage of melting and the slope gradually decreases. Snow shows a lower slope for the low fraction of water than that of hail. This is consistent with our density model given in (2.4).

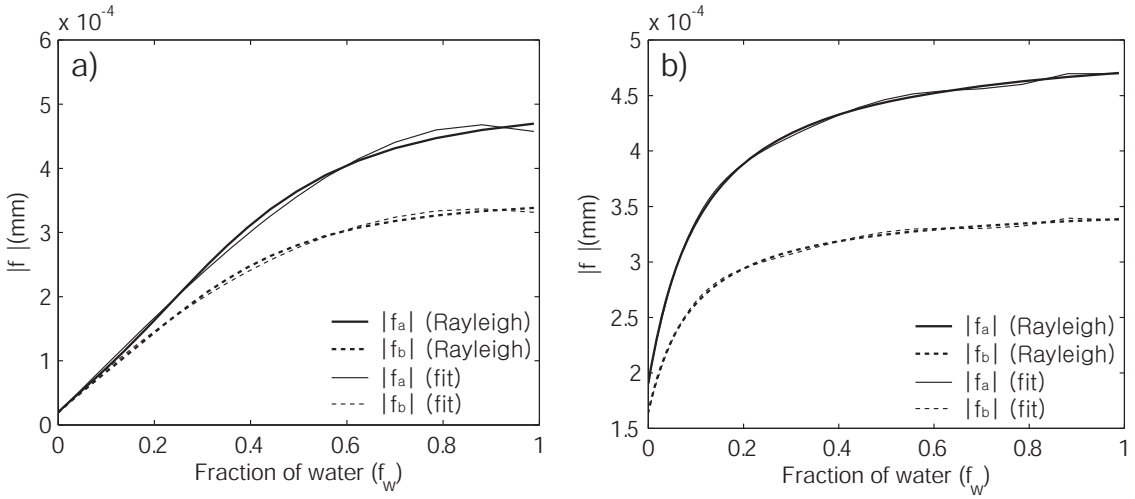


Fig. 2.2. Backscattering amplitudes as a function of the fraction of water within the mixture along the major axis (solid) and the minor axis (dashed) for (a) rain-snow and (b) rain-hail mixtures.

In the current study, the non-Rayleigh scattering effect, which is known to be important for large hailstones with a diameter larger than 10 mm for a S-band radar, is neglected due to the high computational demand by the T-matrix calculation. Instead, the Rayleigh scattering approach is used for the simplicity and efficiency necessary for the data assimilation purpose. The limitation of this assumption is the overestimation of the radar cross section for large hailstones and, therefore, the somewhat overestimation of reflectivity. The non-Rayleigh scattering effect will be included in the new forward operator discussed in Chapter 5, which utilizes the full T-matrix scattering calculation.

After performing integration over the exponential DSD, (2.5) and (2.6) yield simple forms of rain reflectivities at horizontal and vertical polarizations, as follows (Zhang et al. 2001):

$$Z_{h,r} = \frac{4\lambda^4 \alpha_{ra}^2 n_{0R}}{\pi^4 \left|K_w\right|^2} \Lambda_r^{-(2\beta_{ra}+1)} \Gamma(2\beta_{ra}+1) \ (\mathrm{mm^6 m^{-3}}) \tag{2.10}$$

and

$$Z_{v,r} = \frac{4\lambda^4 \alpha_{rb}^2 n_{0R}}{\pi^4 \left|K_w\right|^2} \Lambda_r^{-(2\beta_{rb}+1)} \Gamma(2\beta_{rb}+1) \ (\mathrm{mm^6 m^{-3}}), \tag{2.11}$$

where λ is the radar wavelength, which is approximately 10.7 cm for the WSR-88D radars, The defaults value for the intercept parameter for rain in the LFO08 microphysics scheme is $n_{0R} = 8\times10^6$ $\mathrm{m^{-4}}$, but other values can be used (see discussion in Tong and Xue 2008b). The slope parameter Λ_r can be diagnosed from the rain mixing ratio once the intercept parameter is specified. Here $K_w = 0.93$ is the dielectric factor for water, $\Gamma(\ldots)$ is the complete gamma function.

Integrals for other species in the same way are straightforward. For completeness, they are listed below:

$$Z_{h,x} = \frac{2880\lambda^4 n_{0x}}{\pi^4 \left|K_w\right|^2} \Lambda_x^{-7} \left(A\alpha_{xa}^2 + B\alpha_{xb}^2 + 2C\alpha_{xa}\alpha_{xb} \right) \tag{2.12}$$

and

$$Z_{v,x} = \frac{2880\lambda^4 n_{0x}}{\pi^4 \left|K_w\right|^2} \Lambda_x^{-7} \left(B\alpha_{xa}^2 + A\alpha_{xb}^2 + 2C\alpha_{xa}\alpha_{xb} \right). \tag{2.13}$$

The reflectivities in the linear scale for different species are combined to give logarithmic reflectivity at the horizontal and vertical polarizations (Z_H and Z_V, respectively) and differential reflectivity (Z_{DR}) as

$$Z_H = 10\log_{10}\left(Z_{h,r} + Z_{h,rs} + Z_{h,ds} + Z_{h,rh} + Z_{h,dh}\right) \text{ (dBZ)}, \tag{2.14}$$

$$Z_V = 10\log_{10}\left(Z_{v,r} + Z_{v,rs} + Z_{v,ds} + Z_{v,rh} + Z_{v,dh}\right) \text{ (dBZ), and} \tag{2.15}$$

$$Z_{DR} = 10\log_{10}\left(\frac{Z_h}{Z_v}\right) = 10\log_{10}\left(\frac{Z_{h,r} + Z_{h,rs} + Z_{h,ds} + Z_{h,rh} + Z_{h,dh}}{Z_{v,r} + Z_{v,rs} + Z_{v,ds} + Z_{v,rh} + Z_{v,dh}}\right) \text{ (dB)}. \tag{2.16}$$

The reflectivity difference, another useful polarimetric variable, is defined as

$$Z_{dp} = Z_h - Z_v \quad (\text{mm}^6\,\text{m}^{-3}). \tag{2.17}$$

While Z_{DR} contains the information about the shape of hydrometeor such as the axis ratio, Z_{dp} was proposed to handle mixed-phase precipitation concentration as dry ice phases tend to have less polarization signatures (Seliga and Bringi 1976; Straka et al. 2000; Golestani et al. 1989; Tong et al. 1998; Zrnic and Ryzhkov 1999). With the reflectivity difference, the dry snow and hail contributions are minimized so that rain is better represented. A power of 0.2 is taken of Z_{dp} in our data assimilation experiments so that $(Z_{dp})^{0.2}$ has a more normal-like distribution. Doing so also reduces the dynamic range of data and, therefore, $(Z_{dp})^{0.2}$ is more appropriate than Z_{dp} for the assimilation purpose.

The specific differential phases for rain, rain-snow aggregate mixture, dry snow aggregate, hail-snow mixture, and dry hail are calculated, following Zhang et al. (2001), from

$$K_{DP,x} = \frac{180\lambda}{\pi}\int n(D)C_k Re\left[f_a - f_b\right]dD \ (^\circ\,\text{km}^{-1}), \tag{2.18}$$

$$C_k = \langle \cos 2\phi \rangle = \cos 2\bar{\phi} e^{-2\sigma^2}.$$

As above, integral of (2.18) over DSD can be simplified for rain as following:

$$K_{DP,r} = \frac{180\lambda}{\pi} n_{0R} \alpha_{rk} \Lambda_r^{-(\beta_{rk}+1)} \Gamma(\beta_{rk}+1) \ (^\circ \text{ km}^{-1}), \tag{2.19}$$

where nondimensional coefficients $\alpha_{rk} = 1.30\times10^{-5}$ and $\beta_{rk} = 4.63$ for rain. We can find the α_{xk} for K_{DP} from (2.9) to be α_{xa} - α_{xb} for rain-snow aggregate and rain-hail mixture. Here $\alpha_{dsk} = 0.3\times10^{-6}$ for dry snow and $\alpha_{dhk} = 0.26\times10^{-4}$ for dry hail. β_{xk} values for ice species and water-ice mixtures are equal to 3. Because K_{DP} calculation involves $\langle f \rangle$ while reflectivities involve $\langle |f|^2 \rangle$ [note that the power of D is 4.63 for rain and 3 for ice particles in (2.7) and (2.8), where the mass of the spherical particle is proportional to D^3 while reflectivity is often assumed to be proportional to D^6 in the Rayleigh regime], K_{DP} is more linearly proportional to the rainfall rate (Zrnic and Ryzhkov 1999; Bringi and Chandrasekar 2001).

The specific differential phases for different species are combined in the same manner as the reflectivity to give the total K_{DP}:

$$K_{DP} = K_{DP,r} + K_{DP,rs} + K_{DP,ds} + K_{DP,rh} + K_{DP,dh}. \tag{2.20}$$

The cross-correlation coefficient is defined as

$$\rho_{hv} = \frac{Z_{hv,r} + Z_{hv,s} + Z_{hv,h} + Z_{hv,rs} + Z_{hv,rh}}{\left\{\left[Z_{h,r} + Z_{h,s} + Z_{h,h} + Z_{h,rs} + Z_{h,rh}\right]\left[Z_{v,r} + Z_{v,s} + Z_{v,h} + Z_{v,rs} + Z_{v,rh}\right]\right\}^{1/2}}, \tag{2.21}$$

where the numerator is given as a product of two orthogonal co-polar components of the radar return and computed as

$$Z_{hv,x} = \frac{4\lambda^4}{\pi^4 \mid K_w \mid^2} \int n(D)\left[E\left(\mid f_a \mid^2 + \mid f_b \mid^2\right) + 2F \mid f_a \mid \mid f_b \mid \rho_{0,x}\right] dD$$

$(\mathrm{mm}^6\mathrm{m}^{-3})$, (2.22)

where

$$E = \left\langle \sin^2\phi \cos^2\phi \right\rangle = \frac{1}{8}\left(1 - \cos 4\bar{\phi} e^{-8\sigma^2}\right) \text{ and}$$

$$F = \left\langle \cos^4\phi + \sin^4\phi \right\rangle = \frac{1}{8}\left(3 + \cos 4\bar{\phi} e^{-8\sigma^2}\right).$$

Here $\rho_{0,x}$ = 1.0 for rain and dry snow, and 0.97 for dry hail, but varies as a function of its concentration in a range of 0.82 – 0.95 for rain-snow mixture and 0.85 – 0.95 for rain-hail mixture. ρ_{hv} is very useful in detecting the melting layer since its value is sensitive to the presence of mixture. ρ_{hv} is very high for pure rain while the presence of randomly-oriented large wetted hailstones can significantly lower the correlation.

2.4 Applications to convective storms

To demonstrate that the observation operators in our radar simulator produce reasonable results, they were applied to the 2D squall-line and 3D supercell storm simulations, described in section 2.2. In this section, we examine the simulated radar fields on the model grid before any simulated observation error is added. The error modeling for the polarimetric variables will be discussed in Chapter 3.

2.4.1 Simulated radar fields for the squall-line case

The west-east vertical cross sections of reflectivity at horizontal polarization (Z_H), differential reflectivity (Z_{DR}), reflectivity difference [$(Z_{dp})^{0.2}$], and specific

differential phase (K_{DP}) at 400 min into the 2D squall-line simulation are shown in Fig. 2.3. The 0°C isotherm is overlaid on each plot in thick black lines. The squall-line system is in its mature stage and propagates slowly eastward while the low level flow is from the right. The low level shear vector points rightward therefore the upshear direction is toward the left.

The simulated mature squall line is similar to the multi-cellular squall line discussed by Lin et al. (1998) and Fovell and Tan (1998) in which new cells are periodically regenerated at the leading edge of the gust front. They reach their maximum intensity while propagating rearward, and then weaken as they move into a region of weaker convective instability and turn into more stratiform clouds. Figure 2.3a shows that at the mature stage, the deepest cell, labeled C2, is located near x = 360 km and its echo top reaches nearly 14 km. To its right are two newer cells, labeled C3 and C4, with C3 trying to establish itself and C4 still in its developing stage. To the left or rear of the deepest cell is a much weaker cell, labeled C1, which has passed through the most intense stages and is transitioning into more stratiform clouds (Fig. 2.3a). A deep column of high reflectivity of over 65 dBZ in the deepest convective cell, C2, is mainly associated with the large hail core extending to 9-km height (Fig. 2.4c). A small local maximum of over 70 dBZ at about 4-km height (right below the 0 ºC level above the boldface **C** in Fig. 2.3a) where high rainwater and hail mixing ratios co-exist (Figs. 2.4a,c). Another local maximum close to the 0 ºC level (Fig. 2.3a) is also associated with the coexistence of high hail and rainwater content at that location (Figs. 2.4a,c).

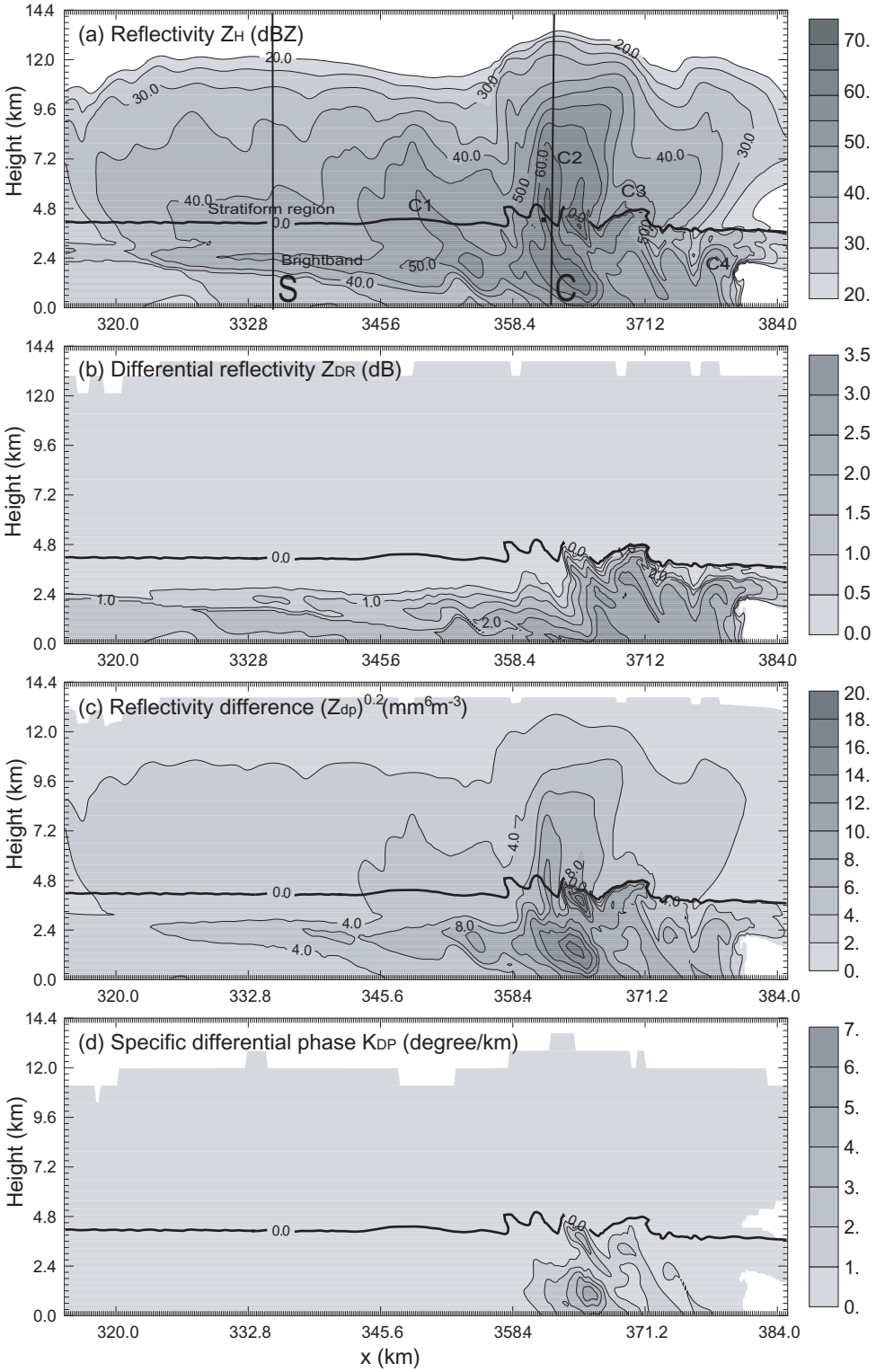


Fig. 2.3. The west-east vertical cross sections of the simulated (a) Z_H, (b) Z_{DR}, (c) Z_{dp}, and (d) K_{DP} at 400 minutes into the 2D squall-line simulation.The 0 ºC isotherms are shown as thick black lines. A sequence of cells in (a) is labeled C1-C4.

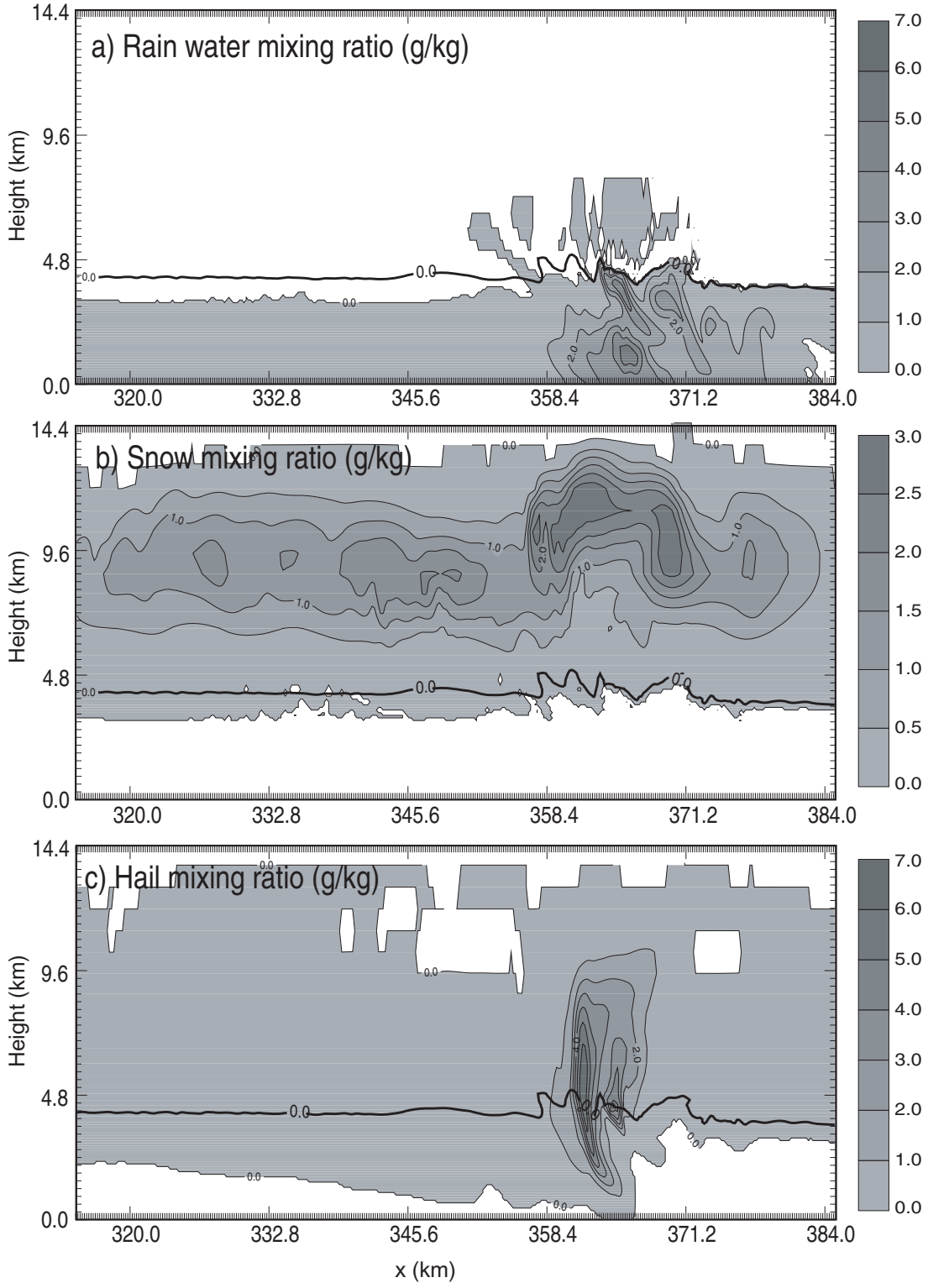


Fig. 2.4. As in Fig. 2.3, but for the model mixing ratios (a) q_r, (b) q_s, and (c) q_h.

The region of high Z_{DR} (Fig. 2.3b) is located off the hail core (Fig. 2.4c) to its right, where rainwater content is significant (Fig. 2.4a). In fact, there is a local minimum, as indicated by the Z_{DR} 'trough', at the location of low-level hail core (Fig. 2.3b). The region of significant Z_{DR} is also caped by the 0°C temperature contour (Fig. 2.3b). This is because that the strongest Z_{DR} signature is associated with rain, whose drops become flattened when their sizes increase. Because of the tumbling, statistically, hailstones appear mostly spherical to the radar beams, resulting in similar reflectivities at horizontal and vertical polarizations. The reflectivity due to hail is large, however, and the large and almost equal contributions of $Z_{h,h}$ and $Z_{v,h}$ to Z_H and Z_V, respectively, reduce the relative importance of $Z_{h,r}$ and $Z_{v,r}$, resulting in small Z_{DR} values according to (2.16). The Z_{DR} values are also significant (1.5 ~ 2.0 dB) in a horizontally elongated region below the brightband in the stratiform precipitation region. Again, this is a region where rain water dominates, and is below the melting layer (Fig. 2.3a).

Z_{dp} is insensitive to ice and is highly correlated with $Z_{h,r}$, showing sensitivity only to the oriented oblate raindrops so that it makes a good indicator of the presence of water within the rain-ice mixture, which enables the use of the concept of the deviation analysis from the rain line (Bringi and Chandrasekar 2001). When analyzed alone, Z_{dp} may be less useful because of its high proportionality to reflectivity. Figures 2.3c and 2.4a show that the $(Z_{dp})^{0.2}$ pattern agrees well with the pattern of rainwater mixing ratio below the melting layer. Snow and hail that have melted sufficiently can be seen as big raindrops to the radar. This is shown as a horizontally elongated enhanced $(Z_{dp})^{0.2}$ band, which matches well with the bright band in Fig. 2.3a. Among polarimetric variables, only $(Z_{dp})^{0.2}$ show some signatures above the 0 ºC level in the convective region. Z_{dp} can

be of moderate strength in the region with high concentration of hail, where Z_H is large no matter the hail is dry or wet (Figs. 2.4a,c). However, these are rather weak signals considering the dynamic range of raw Z_{dp} observations before we take the power of 0.2.

The region of high K_{DP} is mostly confined in the convective rain region (Fig. 2.3d). In fact, its pattern matches that of rainwater mixing ratio very well. This is because K_{DP} is not affected much by the presence of hail. Both Z_{DR} and K_{DP} signatures are rather weak and essentially uniform above the 0°C level.

While examining the simulated radar variables, we noticed that in the stratiform precipitation region, the actual melting level in the model is significant offset from the 0 ºC isotherm. The level of maximum brightband found in Fig. 2.3a is almost 1.7 km below the 0 ºC isotherm (at about 4.2 km level) and consequently significant Z_{DR} signatures is found at lower levels below the bright band. The mixing ratio fields in Fig. 2.4 show that rainwater does not start to appear until about ~900 m below the 0 ºC level, while snow manages to survive below the 0ºC level for a similar depth. Such a discrepancy appears odd, for slowly falling snow in the stratiform precipitation region. To explain this peculiar behavior, we further investigated the microphysics scheme used in this study.

The Lin microphysics parameterization in the ARPS is based on the code from the National Aeronautis and Space Administration (NASA) Goddard Space Flight Center (GSFC, Tao and Simpson 1989). Our investigation reveals that the melting in this scheme does not occur until several degrees above 0ºC (see Figs. 2.4a,b) because the *potential* cooling associated with the evaporation of water at the surface of ice particle exceeds the heating associated with the conduction and convection of heat to

the particle from its environment [see Eq. (32) of LFO83]. Although some delay in the melting due to evaporative cooling is physical, we believe the amount of delay we are observing is too much. For instance, snow and hail do not start to melt until around 7ºC at the location of x = 320 km. We tested another implementation of the LFO83 ice microphysics scheme by Gilmore et al. (2004b) and found the same behavior. The issue is therefore common with the Lin scheme. We found that the single-moment WRF 6-category microphysics scheme (WSM6, Hong and Lim 2006) and the Rutledge and Hobbs (1983) scheme on which WSM6 is based, do not have the same problem because they have a somewhat different treatment of the melting processes. However, they have other issue. Our initial attempts to modify the Lin scheme in this aspect did not lead to satisfactory results, and we will leave this microphysics parameterization issue for future studies because our current study is primarily focused on producing realistic radar simulations given reasonably realistic microphysical fields. On the other hand, we have a good example of how a realistic radar simulator can be used to validate model microphysics, and it will be even more valuable when we simulate and predict real cases and compare the results against real radar data.

To further examine the behaviors of our forward observation operators that include the melting model, two columns of mixing ratios are extracted at x =362.2 km (labeled **C** in Fig. 2.3a) and 336 km (labeled **S**) from the simulated squall-line system, corresponding to the convective and stratiform regions, respectively. The profiles of radar variables are calculated from these mixing ratios and are shown in Figs. 2.5 and 2.6. The 0°C temperature line is overlaid as a straight line on each plot.

Figure 2.5a shows the vertical profiles of q_r, q_s, q_h, q_{rs} and q_{rh} in the convective rain. This region shows the highest mixing ratio of hail. The mixed rain-snow mixing ratio q_{rs} reaches its maximum where the sum of the coexisting rain and snow mixing ratios has a maximum but its peak value is so small as to be hardly identifiable in the plot.

The reflectivity at horizontal polarization incorporating our melting ice (MI) model is plotted in Fig. 2.5b as the solid black curve. In between the levels of q_r and q_h maxima, q_{rh} has its maximum, providing high reflectivity values that together with q_r and q_h yields a deep reflectivity core at the convective region. The result of a previously used simple linear interpolation (LI) model for melting-layer reflectivity (Jung et al. 2005) is also shown for comparison (dashed curve in Fig. 2.5b). With this LI model, the snow is considered 100% wet at or above air temperature of 0 °C and 100% dry at or below air temperature of – 5 °C. In between these two temperatures, the reflectivity is calculated as the weighted average of those given by the wet and dry snow formulas, with weights defined as linear functions of the temperature. The same method is applied to dry and wet hail reflectivity formulas with a corresponding air temperature range of 2.5 °C and –2.5 °C. When the LI model is used, the reflectivity is more directly linked to the temperature and less so to the model microphysics. Figure 2.5b shows that the reflectivity thus calculated keeps decreasing with decreasing height below an air temperature of 2.5 °C. A maximum value is found close to the 0°C level while in the melting model case, the similar local enhancement is found at the deep layer of actual melting. Figures 2.5c,d show the differential reflectivity and specific differential phase,

which are found to slowly increase to their maximum values near the surface as the amount of hail decreases. Their values above 0 °C are small.

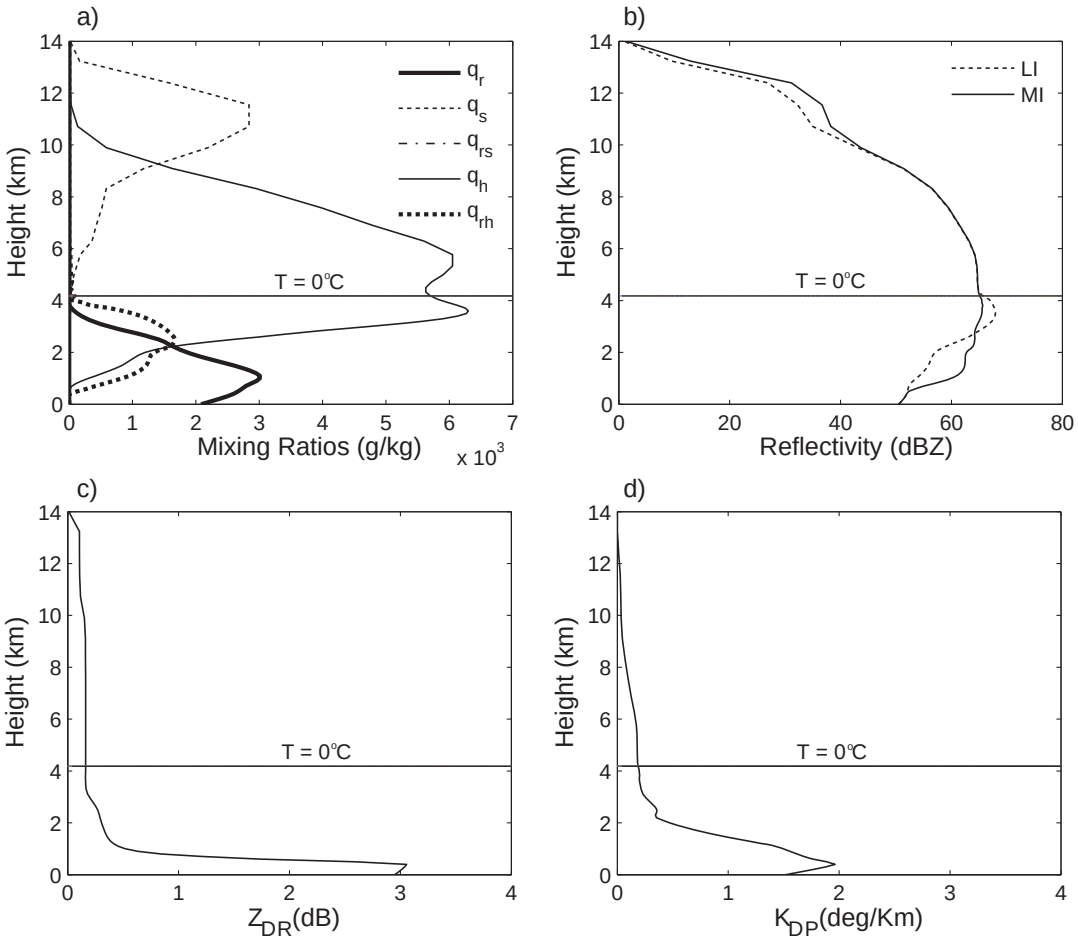


Fig. 2.5. A modeled vertical profile of total (in both pure and mixture forms) rainwater and total snow-hail mixing ratios and the total amount (mixing ratio) of rain and snow and rain and hail in a mixture form, and the simulated polarization radar signatures at the column labeled **C** in Fig. 2.3: (a) q_r (thick solid), q_h (thin solid), q_s (thin dashed), q_{rh} (thick dashed), and q_{rs} (dash-dotted); (b) reflectivity from the LI model (dashed), reflectivity at horizontal (Z_H, solid) polarization from the MI model; (c) Z_{DR}; and (d) K_{DP}. Here q_{rs} is on the vertical axis.

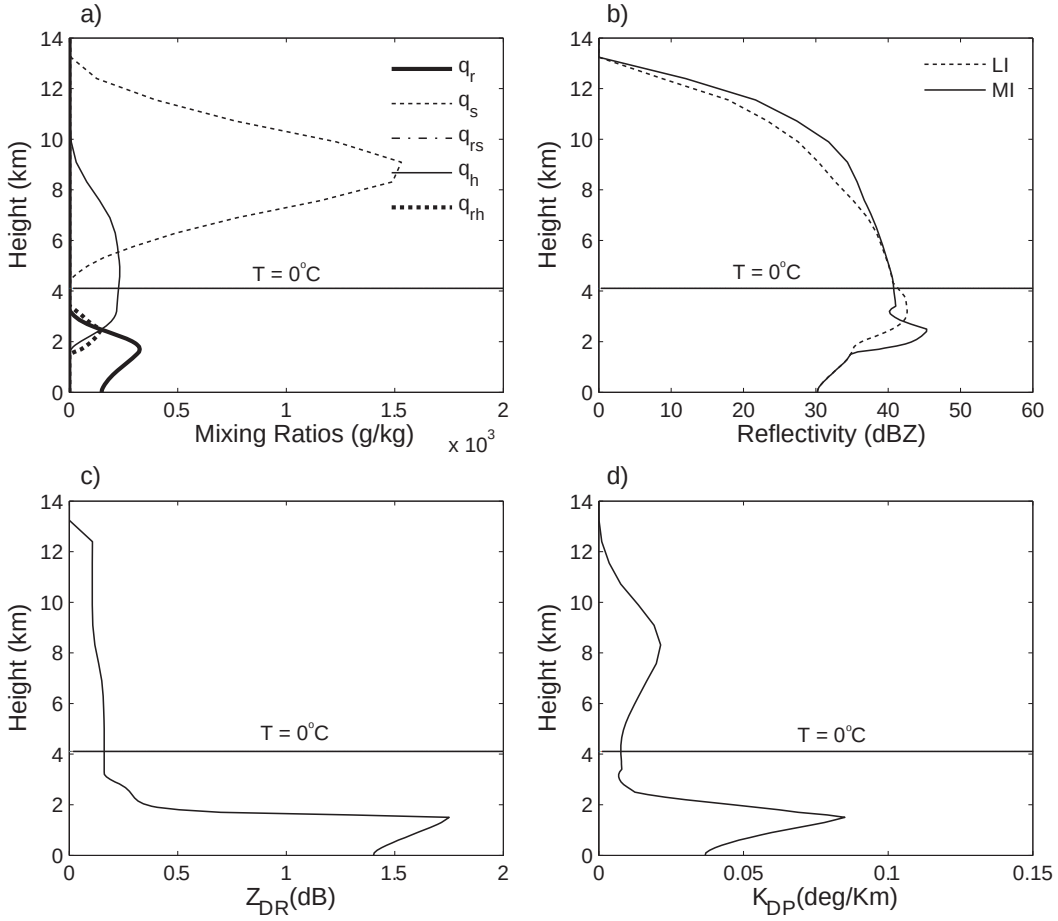


Fig. 2.6. As in Fig. 2.5, but for column **S** in Fig. 2.3.

In the stratiform region where the snow mixing ratio is the largest and is found at the upper levels, it can be seen that the current melting model produces realistic nonpolarimetric and polarimetric radar signatures with a bright band associated with the melting layer shown in both Z_H and Z_{DR} profiles (Figs. 2.6b,c). The reflectivity increase in the melting layer of the MI model is more prominent and shallower than that of the interpolation model. The differential reflectivity peak shows slightly below the reflectivity peak. These characteristics in reflectivity and differential reflectivity agree well with observed profiles (Fig. 8 of Brandes et al. (2004a)) and the composite range-

height indicator plot (Fig. 13 of Ryzhkov et al. (2005a)). The better handling of the radar variables by the MI model is because of the presence of snow-rain and hail-rain mixtures and the better representation of their effects on the dielectric constants of melting snow and hail. The interpolation model does not take into account the change in the dielectric constants directly.

2.4.2 Simulated radar fields for a supercell storm

Next, we apply our observation operators to the simulated supercell storm, which will also be used in Chapter 3 to test the impact of assimilating additional polarimetric variables. Figure 2.7 shows the simulated polarimetric variables at the 2.5-km altitude at 100 minutes of the storm. The storm splits at around 55 minutes into two cells; one moving towards the left of the storm motion vector that ends up near the upper-left corner of domain by 100 min and the other (right mover) stays close to the center of the model domain (Fig. 2.7a). Reflectivity pattern matches well with the hail field and the reflectivity core is collocate with hail maximum in the left-moving storm and with the common area in q_r and q_h maxima in the right-moving storm (Figs. 2.7a,b). The Z_{DR} shows a minimum near the reflectivity maxima, collocated with hail cores (Figs. 2.7b,c). This is consistent with the Z_{DR} hole observed in the microburst studied by Wakimoto and Bringi (1988) and the convective storm studied by Bringi et al. (1986a). These observations also show that Z_{DR} values increase rapidly around the Z_{DR} hole and reach more representative values for melting ice. As discussed in the squall-line case, the K_{DP} field is consistent with that of q_r.

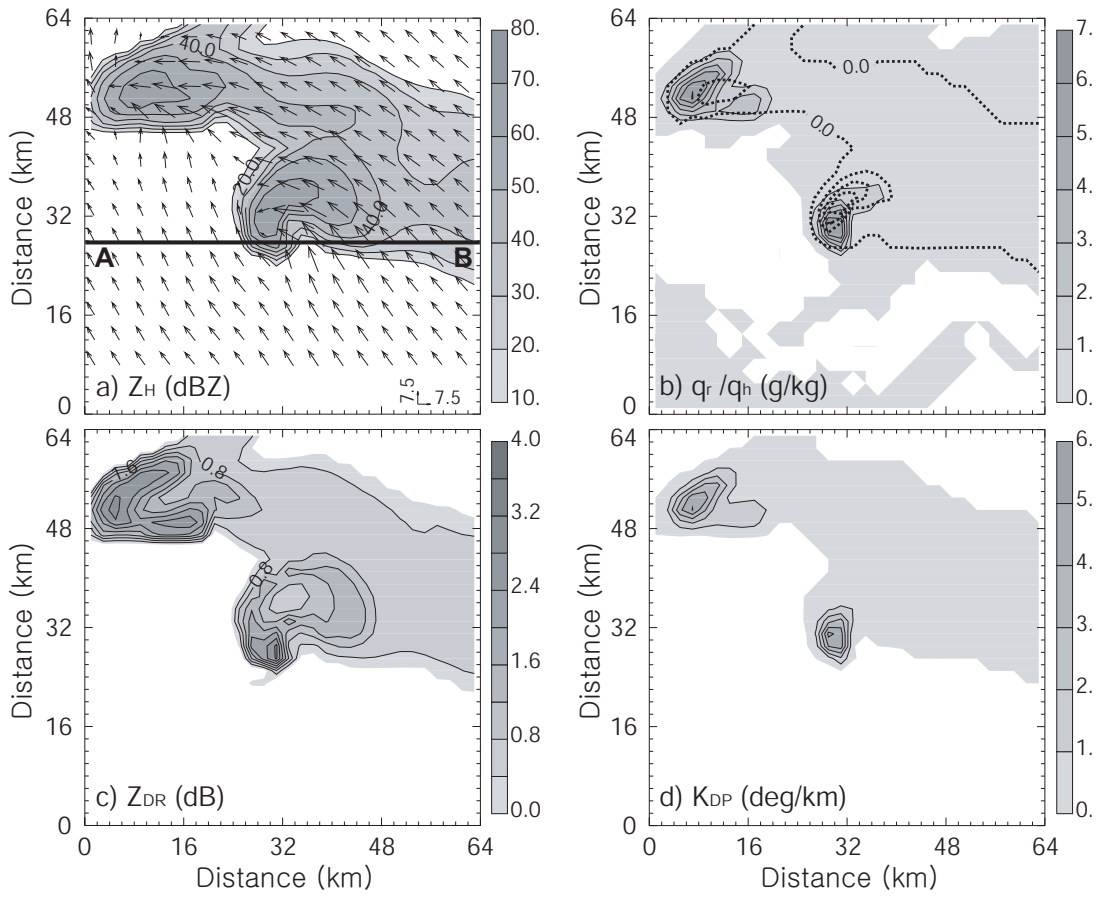


Fig. 2.7. (a) Horizontal wind (vectors; m s-1) and reflectivity, Z_H, (b) rain water q_r (thin solid contours and shading) and q_h (thick dotted at intervals of 1 g kg^{-1}, starting from 0 g kg^{-1}), (c) Z_{DR}, and d) K_{DP}, at z = 2.5 km at 100 minutes of the storm.

Figure 2.8 shows the vertical structure of the supercell storm at line AB shown in Fig. 2.7, which passes through the updraft core and the weak echo region (WER) in the storm. In this case, the reflectivity maximum through the updraft core is found at about 4.5 km above ground (Fig. 2.8a) because of a high concentration of hail in melting phase there (Fig. 2.8f). The high Reflectivity region exceeding 60 dBZ extends to 8.5 km, corresponding to the region of high hail. The fully developed overhang signature is consistent with the patterns of hail and snow (Figs. 2.8a,e,f).

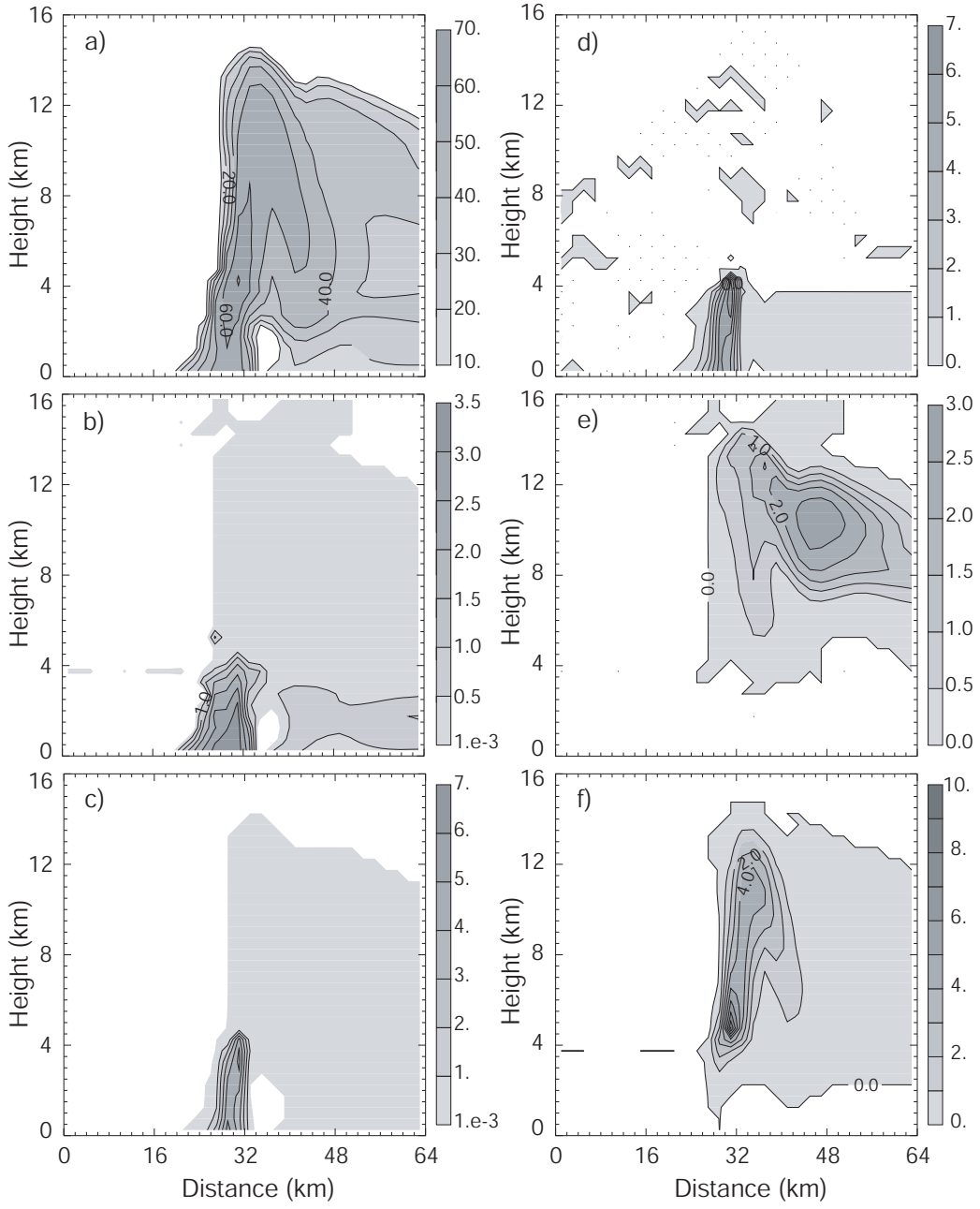


Fig. 2.8. The west-east vertical cross-sections of simulated (a) reflectivity at horizontal polarization (Z_H), (b) differential reflectivity (Z_{DR}), (c) specific differential phase (K_{DP}), and (d) rain water (q_r), (e) snow (q_s), and (f) hail (q_h) mixing ratios, (g) reflectivity difference $(Z_{dp})^{0.2}$, and (h) cross-correlation coefficient ρ_{hv} through the updraft core (maximum vertical velocity) of the simulated supercell storm at 100 minutes, along line AB shown in Fig. 2.7a. It corresponds to y = 28 km.

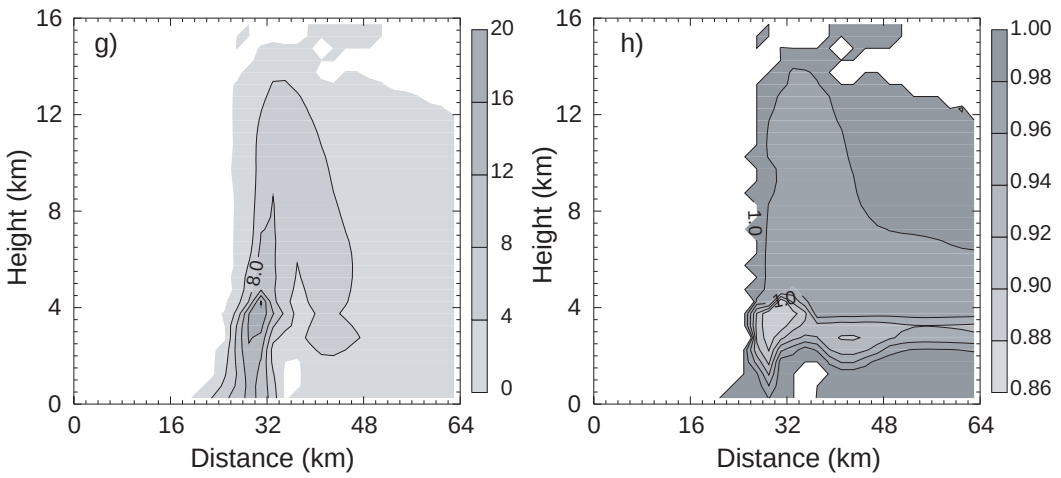


Fig. 2.8. *(Coutinued)*

It can be seen that our emulator produces reasonably weak reflectivity for dry snow compared with that for hail as in the real storm (Figs. 2.8a,d,f). In contrast to the reflectivity field, Z_{DR} and K_{DP} remain low at this level in this precipitation core (Figs. 2.8b,c). The core of the Z_{DR} column (Figs. 2.7c and 2.8b) is located southwest of the center of the WER, where the reflectivity hook is located (Fig. 2.7a), similar to those in Fig. 2 of Hubbert et al. (1998). The top of Z_{DR} column is bounded by the 0°C line while the observational study of Hubbert et al. (1998) shows that it rises above the 0°C line because raindrops or water-coated ice particles are carried by a strong updraft. In the simulated storm, super-cooled water quickly converts to ice phase so that Z_{DR} quickly drops accordingly. The K_{DP} pattern shown in Fig. 2.8c indicates that it has useful information content only for heavy rain, as observational and theoretical studies have shown earlier (Chandrasekar et al. 1990; Balakrishnan et al. 1989). There is hardly any K_{DP} signal in the light-rain region. The maximum values of specific differential phase occur where the rainwater mixing ratio is the greatest, between the 3.5-km level and the surface (Fig. 2.8c). However, the maximum differential reflectivity appears near the

surface because the large hail values at the higher levels reduce the relative contribution of rain, as discussed earlier for the squall-line case. Their signatures are very weak at high altitudes where the hydrometeor density is low, dielectric constant is small, and their effective shapes are spherical.

The patterns of Z_{DR} and K_{DP} are similar for different physical reasons. The Z_{DR} is greater where larger q_r is found because more larger drops with more oblate shapes are expected there. The K_{DP} is more linearly proportional to the amount of rain as discussed in section 2.3.3. Both Z_{DR} and K_{DP} remain low at the middle level to the right of the storm where the hail dominates among the hydrometeors (Figs. 2.8b,c,f). The Z_{DR} increases towards the surface as most hail completely melts before reaching the ground. This behavior is consistent with the Z_{DR} and K_{DP} equations in (2.16) and (2.19) and also agrees well with observations (Hubbert et al. 1998; Illingworth et al. 1987; Zrnic and Ryzhkov 1999), and with the study of Huang et al. (2005) in which a full radar scattering model is used to simulate polarimetric radar signatures of a model-simulated storm. For the purpose of data assimilation, simple formulas like those discussed in this chapter have to be used. At this time, full scattering calculations are prohibitively expensive for data assimilation purposes.

As in the squall line case, $(Z_{dp})^{0.2}$ pattern is highly correlated with the reflectivity pattern below melting layer while its signal is much weaker above melting layer (Fig. 2.8g). The presence of mixture is clear from the pattern of ρ_{hv} (Fig. 2.8h). ρ_{hv} is very high for pure rain near the surface and for pure snow aggregate in the anvil. The correlation drops where the melting of ice particles occurs. The lowest correlation core is shown in the region with high concentration of hail, yet restricted by the presence of

rain water, due to the large number of hailstones in the melting phase. The presence of melting layer is clear between 2 and 3.5 km altitude.

2.5 Summary and conclusions

In this chapter, the forward observation operators that link model state variables with polarimetric radar measurements are developed based on scattering calculations with the T-matrix method for rain and the Rayleigh scattering approximation for snow aggregates and hail. These operators, together with proper handling of the radar beam geometry and beam weighting-functions, form a radar simulator. The operators are developed mainly for the purpose of assimilating the corresponding measurements into storm-scale numerical models; at the same time, they can be used to verify model predictions against radar observations.

The radar measurements considered include the reflectivities of the horizontal and vertical polarizations (Z_H and Z_V), differential reflectivity (Z_{DR}), reflectivity difference (Z_{dp}), specific differential phase (K_{DP}), and the cross-correlation coefficient (ρ_{hv}). The work is necessitated by the unavailability of existing observation operators for most of the polarimetric variables that are efficient for data assimilation purposes and make use of all microphysical information available in a numerical model. Because of the lack of information in typical bulk microphysics schemes on the liquid water fraction of ice, a new melting model is developed that assumes a function for the water fraction based on known rainwater and snow/hail mixing ratios. The effects of varying density due to the melting snow and hail are also included.

The observation operators developed are tested with a model-simulated mature squall-line system that includes both deep convection and stratiform precipitation regions, and a supercell storm with high hail content. Realistic nonpolarimetric and polarimetric radar signatures are produced in the simulated fields, including a bright band and realistic spatial distributions of Z_{DR} and K_{DP} signatures. The simulated radar fields suggest a problem with the treatment of snow and hail melting processes in the Lin-type microphysics scheme, which will be examined in more depth in the future.

Our simulated reflectivity seems generally higher than observed one for ice phases. This is partly because non-Rayleigh scattering effects have been neglected in the calculation. This could have a larger impact in the convective rain than in the stratiform rain. The fixed DSD intercept parameter of hail is probably responsible for high reflectivity in the stratiform precipitation region where we expect mostly small ice particles. The hail intercept parameter is two orders of magnitude smaller than those of rain and snow and can lead to high reflectivity. Last, the lack of raindrop breakup, which is neglected in our microphysical parameterization, is another source of high reflectivity. When the DSD is not properly truncated, a few unrealistically large drops can significantly increase reflectivity.

The observation operators have been implemented in our ensemble Kalman filter data assimilation system, and the impact of additional polarimetric measurements on the analysis of a supercell storm will be examined in Chapter 3.

Chapter 3 [2] Assimilation of Simulated Polarimetric Radar Data for a Convective Storm Using Ensemble Kalman Filter: Impact of Polarimetric Data on Storm Analysis

3.1 Introduction

For convective-scale NWP, microphysics represents perhaps one of the most important physical processes with both direct and indirect influences. The microphysical processes depend to a large extent on the phase, density, and the drop size distributions (DSDs) of the microphysical species involved. These properties also directly affect radar measurements within each radar sampling volume. For these reasons, equivalent radar reflectivity factor (hereafter reflectivity) and radial velocity measurements from conventional Doppler weather radars are usually insufficient to fully describe the microphysical states in a convective storm. Additional observational parameters available from polarimetric Doppler radars, including differential reflectivity and differential phase measurements can be very helpful here as they contain information about the density, shape and DSDs of hydrometeors (Doviak and Zrnic 1993; Bringi and Chandrasekar 2001).

The use of differential reflectivity for meteorological applications, in particular for rainfall estimation, was first proposed by Seliga and Bringi (1976); many studies have shown that polarimetric measurements can improve precipitation-type classification and quantitative rainfall estimates (Straka et al. 2000). Ryzhkov et al. (1998) and Vivekanandan et al. (1994) have proposed that polarimetric methods can

[2] This chapter is published as: Jung, Y., M. Xue, G. Zhang, and J. M. Straka, 2008: Assimilation of simulated polarimetric radar data for a convective storm using ensemble Kalman filter. Part II: Impact of polarimetric data on storm analysis. *Mon. Wea. Rev.,* **136**, 2246-2260.

estimate ice water content more accurately than the one that only uses reflectivity (Z_H). Wu et al. (2000) used differential reflectivity (Z_{DR}) indirectly (rain and ice mixing ratios were derived from reflectivity and Z_{DR} before assimilation) in a cloud-scale four-dimensional variational data assimilaton (4DVAR) system and obtained somewhat encouraging results. Moreover, the planned polarimetry upgrade starting later this decade or early next decade (D. Zrnic 2006, personal communication) by the National Weather Services (NWS) of the entire operational Weather Surveillance Radar-1988 Doppler (WSR-88D) radar network will undoubtedly motivate more active research on the utilization of polarimetric radar data.

An accurate estimate of the amounts of hydrometeors and DSDs using polarimetric radar data can contribute to the improvement and verification of microphysical parameterizations in cloud and mesoscale models. Such estimations can also help enhance our understanding of the interactions between microphysics and kinematics in severe storms and in the mesoscale system (Straka et al. 2000). Polarimetric radars also should be helpful for storm-scale model initialization, especially of the microphysical and related thermodynamic fields, through data assimilation.

The accuracy of NWP depends on the model initial condition. The error in the initial state grows with time and makes the predicted state diverge from its true state. Therefore, a lot of effort has been given to determine more accurate initial conditions that can lead to more accurate weather forecasts. Currently, the two most promising data assimilation techniques for obtaining the atmospheric initial condition or the best estimate of the atmospheric state are the 4DVAR (Le Dimet and Talagrand 1986;

Courtier and Talagrand 1987) and the ensemble Kalman filter (EnKF) method (Evensen 1994; Evensen and Leeuwen 1996; Burgers et al. 1998; Houtekamer and Mitchell 1998; Anderson 2001; Bishop et al. 2001; Whitaker and Hamill 2002; Evensen 2003; Tippett et al. 2003), because of their ability to make effective use of the dynamic model equations and observations distributed in space and time, and to provide the best estimate that is also consistent with the prediction model. Because of its ability in handling complex, nonlinear, physical processes (e.g., ice microphysics) in the assimilation model, and in the forward observation operators (e.g., those for reflectivity), the EnKF method appears to be more suitable for convective-scale data assimilation, which is the main interest of this study.

The EnKF technique was introduced into the meteorological community about a decade ago and has become very popular in recent years. It is an attractive alternative to the more mature 4DVAR method. Very encouraging results have been obtained by a number of researchers for large-scale models (e.g., Houtekamer et al. 2005; Whitaker et al. 2004). Tests with perfect prediction models with simulated Doppler radar data at the convective scale with EnKF have also produced very encouraging success in recent studies. The first paper to investigate the potential of EnKF for assimilating Doppler radar data was Snyder and Zhang (2003). The study used a cloud model with warm rain microphysics and assimilated simulated radial velocity data assumed to be available on the model grid. The studies of Tong and Xue (2005b, hereafter TX05) and Xue et al. (2006, hereafter XTD06) further demonstrated that the cloud fields associated with a three-ice microphysics scheme (cloud ice, snow aggregates, and hail) can be accurately retrieved using the EnKF method. Moreover the inclusion of reflectivity data improves

the results even though its observation operator is highly nonlinear. XTD06 also removed the assumption that radar data are available on the model grid and used more realistic radar-beam-pattern-based forward observation operators.

More recently, Tong and Xue (2008b; 2008a, hereafter TX08a and TX08b) applied the ensemble Kalman filter technique to the problem of simultaneous estimation of the atmospheric state of a convective storm and uncertain DSD-related microphysics parameters associated with a single-moment microphysics scheme, from radar radial velocity and reflectivity data. It was found that the parameter estimation can always be successful when only one of the parameters contains error. The difficulty of parameter estimation increases when multiple parameters contain error and have to be estimated simultaneously. The fact that the errors in some of the parameters produce compensating responses in terms of the observed radar reflectivity, causing solution nonuniqueness, is believed to be the reason for the difficulties. The study suggests that additional polarimetric radar measurements that provide the microphysics and DSD information can help alleviate the solution's nonuniqueness problem. Even when microphysics parameter estimation is not performed, the additional polarimetric measurements are expected to improve the microphysical state estimation. When the microphysics scheme predicts more than one moment (i.e., the mixing ratios), then more microphysical state variables (e.g., the total number concentration and reflectivity factor, as in the three-moment scheme of Milbrandt and Yau 2005a) have to be estimated. If the radial velocity and conventional reflectivity are the only two storm-scale observations, the full state estimation is likely to be very difficult.

In this chapter, we report on the results of our initial efforts in developing capabilities to assimilate polarimetric radar data into a storm-scale NWP model, and in studying the impact of these variables on the analysis or model state estimation. We extended the ensemble Kalman filter data assimilation framework of TX05, XTD06 and TX08a, by adding the ability to assimilate differential reflectivity (Z_{DR}), reflectivity difference (Z_{dp}), the specific differential phase (K_{DP}), and the correlation coefficient $\rho_{hv}(0)$. In Chapter 2, the development of the observation operators for these parameters are described, together with an examination of their applications to a simulated squall line and supercell storm. These observation operators are used in the EnKF Observing System Simulation Experiment (OSSE) system to produce the simulated observation and to assimilate the data.

In section 3.2, the simulation of the radar observations to be used in the OSSEs is discussed, together with their error models. The supercell simulation used in the Chapter 2 is used as the truth simulation from which error-containing observations are generated. It is followed by the design and configurations of the OSSE data assimilation experiments. The impact of assimilating additional polarimetric variables is examined in section 3.3 based on the OSSE results. In section 3.4, we conclude this study and discuss some practical issues in the use of polarimetric radar data for the data assimilation purposes. We believe the study reported herein represents the first attempt to *directly* assimilate polarimetric radar data into a numerical model.

3.2 Assimilation system and experimental design

The prediction model and the truth simulation of a supercell storm used for OSSEs are described in Chapter 2. In the following, we first describe the simulation of the observations from this truth simulation and the error modeling for the reflectivity and polarimetric variables.

3.2.1 Simulation of observations and the error model

Real observations are usually contaminated by measurement and sampling errors, and can contain representativeness error also. In our radar simulator, error-free observations are first generated at model grid points using the observation operators developed in Chapter 2, with the state variables of the truth simulation as input. The results are then brought to the radar elevation levels through interpolation and necessary beam-pattern weighting. We assume that the radar data are at the model grid columns, which is also an assumption made in XTD06. The effective earth radius model is used to take into account the effect of beam bending due to the surface curvature of the earth and the vertical change of refractive index (Doviak and Zrnic 1993). A Gaussian beam weighting function described in XTD06 is used in the vertical direction to simulate Z_h, Z_v, V_r, and K_{DP} observations on the radar elevation planes.

Noise is then added to the error-free observations to simulate observation errors. Operational polarimetric WSR-88D radars transmit and receive horizontally and vertically polarized waves simultaneously, which measures Z_h and Z_v from the same pulses. Because the errors in Z_h and Z_v are mostly correlated for weather echoes in this configuration, the error in Z_{DR} is usually small as a ratio between Z_h and Z_v. To more

realistically model the errors, correlated and uncorrelated random errors having Gaussian distributions are added to uncontaminated Z_h^t and Z_v^t in the linear domain (before the logarithmic transform) and converted to logarithmic reflectivity, Z_H and Z_V (Xue et al. 2007), so that

$$Z_H^o = 10\log_{10}(Z_h^t + \varepsilon_{corr} + \varepsilon_h) \quad \text{and} \tag{3.1}$$

$$Z_V^o = 10\log_{10}(Z_v^t + \varepsilon_{corr} + \varepsilon_v), \tag{3.2}$$

where superscripts t and o denote the uncontaminated (truth) and error-containing simulated observations, respectively. Here ε_{corr} represents the correlated part of the error and ε_h and ε_v are uncorrelated errors for Z_h and Z_v, respectively. They are randomly generated Gaussian errors with zero means and standard deviations proportional to the (uncontaminated) reflectivity (Z_h), as real sampling errors should behave (Doviak and Zrnic 1993; Xue et al. 2007).

Briefly, the actual sizes of the standard deviation (hereafter effective error SD) of the error are experimentally determined in the following way. First, errors ε_{corr}, ε_h and ε_v are simulated by multiplying Z_h^t by a specified factor representing the relative error magnitude for each of them, and by a Gaussian-distributed random number with a zero mean and standard deviation of one. The errors are then used in (3.1) and (3.2) to give Z_H^o and Z_{DR}^o. These error-containing data are collected over the points where $Z_H^o > 0$ dBZ and $Z_{DR}^o > 0$ dB, respectively, for all data sampling times; the effective error SD for each dataset are then calculated. To obtain desired levels of SD of data for the purpose of data assimilation experiments, these steps are repeated through trials with different combinations of ε_{corr} and ε_h (and ε_v) until they are obtained. With this error

model, errors are Gaussian distributed in the linear domain but become non-Gaussian when they are transformed to the log domain. For further details and discussions on the error model, the reader is referred to Xue et al. (2007).

The observations of Z_{DR}^o and Z_{dp}^o are generated from Z_h^o and Z_v^o. The errors in Z_{DR}^o and Z_{dp}^o are simply propagated from errors in the reflectivity observations, as in the real data. Simulated observations of V_r^o and K_{DP}^o are obtained from the error-free V_r^t and K_{DP}^t by adding Gaussian errors of zero mean and specified SDs:

$$V_r^o = u\cos\alpha\sin\beta + v\cos\alpha\cos\beta + w\sin\alpha + \text{ a random error} \quad \text{and} \tag{3.3}$$

$$K_{DP}^o = K_{DP}^t + \text{ a random error}. \tag{3.4}$$

In (3.3), we neglected the effect of the hydrometeor terminal velocity, which is also done in this chapter when assimilating V_r data. The same is done in XTD06 although their more recent studies have included the terminal velocity effect (TX08a).

We note that in (3.1)-(3.4), only the typical radar sampling error is simulated. Other typed of measurement errors associated with mismatched sidelobes, clutter contamination, partial beam filling, range effect, etc., are not taken into account in our error model. In our radar emulator, the SDs or variances can be specified by the user. For operational WSR-88D radars, the reasonable range of the standard deviations of reflectivity and differential reflectivity are 1-2 dBZ and 0.1-0.3 dB, respectively (Ryzhkov et al. 2005b; Doviak and Zrnic 1993). The standard error of K_{DP} in the range of 0.24-0.48° km^{-1} is expected for lightly filtered estimates of K_{DP} from differential phase ϕ_{DP} for operational WSR-88Ds (Ryzhkov et al. 2005b). The V_r error can be assumed to be 1 m s^{-1} (Doviak and Zrnic 1993).

Default error SDs used in our simulation and assimilation experiments are given here. The default values of ε_{corr} and ε_h (and ε_v) are set to be 36% and 2% of Z_h^t so as to yield an effective error SD of about 2 dBZ for Z_H^o and close to 0.2 dB for Z_{DR}^o. Gaussian errors with zero mean and SDs of 1 m s^{-1} for V_r^o and 0.5° km^{-1} for K_{DP}^o, which is reasonable for a 2-km resolution (Ryzhkov et al. 2005b), are added to V_r^t and K_{DP}^t. $\left(Z_{dp}^o\right)^{0.2}$ error is determined by the errors in Z_H^o, and is about 1.0 mm^6 m^{-3}. These errors approach the large end of errors suggested in the literature. Also, Torres and Zrnic (2003) proposed a technique that can significantly reduce statistical errors while maintaining the same level of current WSR-88D radar capabilities such as the scan rate. We assume large errors in the observations to account for the worst cases. The errors in the real observation can be reduced by implementing new techniques in the future, and then the impacts could be larger than those shown later in this chapter. The same SDs (or their squared version, i.e., the error variances) are specified in the filter for the corresponding observations in all experiments presented in this chapter.

As an example, Fig. 3.1 shows the error-containing (Figs. 3.1b,d,f) observations at the lowest radar elevation of 0.5 º that are compared with the error-free observations (Figs. 3.1a,c,e), for the simulated supercell storm. Observations below 250 m in height, which is the first level of the scalar variables in the model for the 500-m vertical grid resolution, are not plotted near the radar at the lower–left-hand corner of each panel. With the default SD errors for Z_H, Z_V, Z_{DR}, and K_{DP} as given above, the overall patterns of error-containing observations are not much affected by the errors. Of course, the error-containing observation fields appear noisy and the values at specific points differ

from the truth values. Some local extrema introduced by the errors, like those at x = 25 and y = 43 km and at x = 45 and y = 47 km in the reflectivity field, are evident and resembles real observations (Fig. 3.1b). In our previous OSSE studies, negative Z_H is set to 0. This is done here also.

The errors in Z_{DR} are simply propagated from the errors in reflectivity at horizontal and vertical polarizations. Even though a large reflectivity error generally leads to a large Z_{DR} error in most cases, their errors are not necessarily strongly correlated at every point because of the uncorrelated part of error. The noise in the data is particularly noticeable for small values of Z_{DR} and most of this noise is removed in our assimilation by data thresholding. Negative Z_{DR} is also set to 0 as we assume that the differential attenuation is small for S-band radars at both polarizations, which could cause negative Z_{DR} by attenuating Z_H more than Z_V. Also, the negative Z_{DR} observed from hail (Bringi et al. 1986a; Illingworth et al. 1987), either from prolate or conical shape particles or three-body scattering (Hubbert and Bringi 1997), is not simulated in this study; they are less important because they will most likely fall below our threshold. We also note here that setting the negative value to zero is also a form of data thresholding; we believe doing so is desirable and can be done with real data also.

We keep the negative values of K_{DP} in the error-containing field (Fig. 3.1f). An SD of 0.5° km^{-1} that is used here is quite large considering the dynamic range of data. However, the fact that K_{DP} error does not scale with the signal (as those of reflectivity do) means that the signal-to-noise ratio of K_{DP} is actually high in heavy precipitation regions.

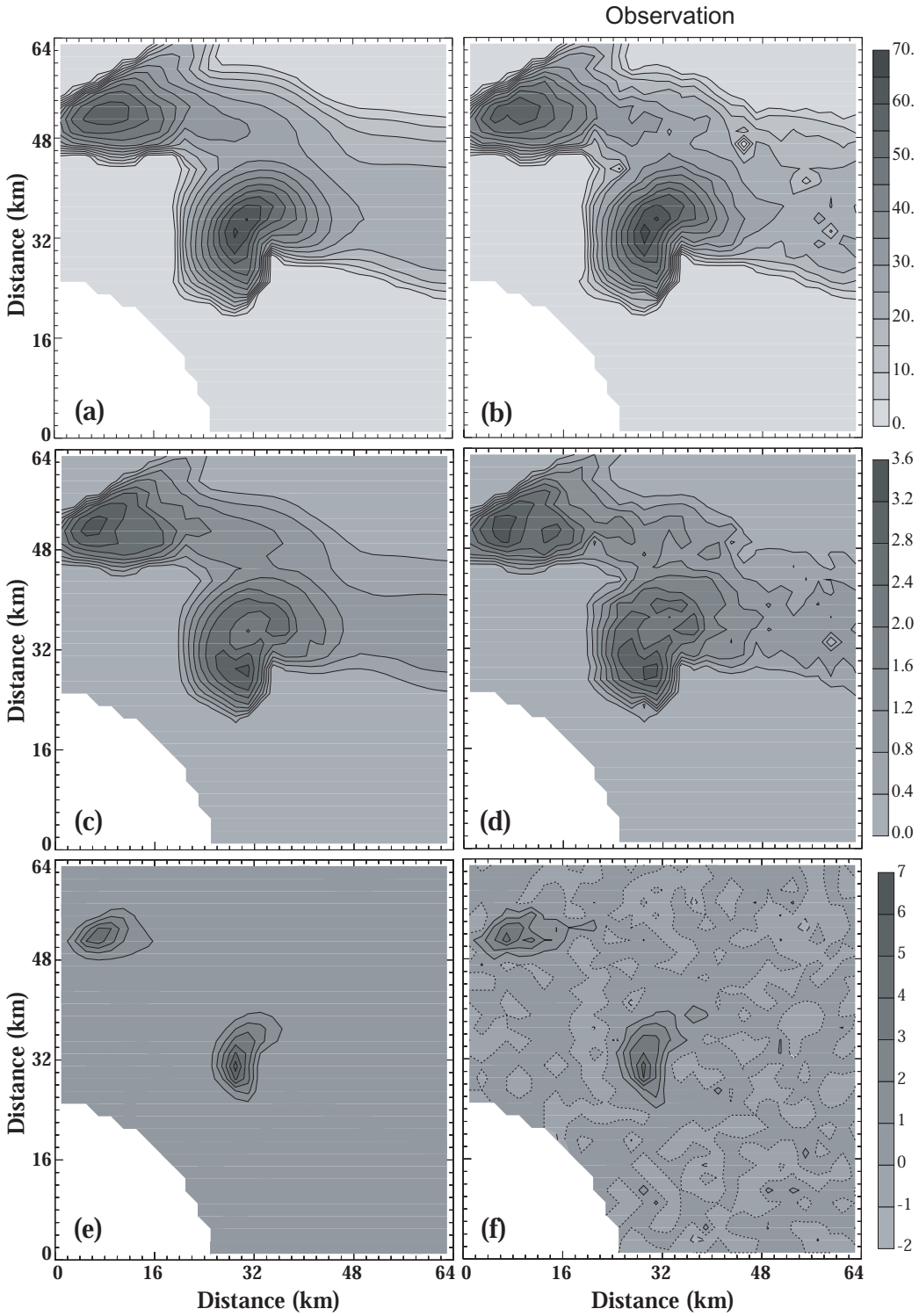


Fig. 3.1. Simulated (a),(c),(e) error-free and (b),(d),(f) error-containing observations at the 0.5º elevation at 100 min of the supercell storm simulation of (a),(b) Z_H; (c),(d) Z_{DR}; and (e),(f) K_{DP}.

3.2.2 Data assimilation procedure

As mentioned earlier, the EnKF radar data assimilation framework of XTD06, which was based on TX05 and further enhanced in TX08a, is used as the basis of our data assimilation work. This framework is enhanced by adding additional capabilities to assimilate the polarimetric radar variables. The observation operators developed in Chapter 2 are used, with our formula for the reflectivity at horizontal polarization [(2.5) in Chapter 2] replacing the reflectivity formula described in TX05. The new error model described above is used.

Our EnKF assimilation system employs the ensemble square-root filter (EnSRF) after Whitaker and Hamill (2002), which is a particular variant of ensemble-based filters. A full description of the filter can be found in XTD06 and TX08a. The experiment environment is largely inherited from XTD06 and TX08a, with the differences noted above.

Following TX08a, initial ensemble members are initialized at t = 20 min of model time by adding spatially smoothed perturbations to the initially horizontally homogeneous first guess defined by the Del City, Oklahoma, sounding. The standard deviations of the perturbations added to each variable are 2 m s^{-1} for u, v, and w; 2 K for θ; and 0.6 g kg^{-1} for mixing ratios of hydrometeors (q_v, q_c, q_r, q_i, q_s, and q_h,). The perturbations are added to the velocity components, potential temperature, and specific humidity, in the entire domain excluding grids comprising the lateral boundaries. For the mixing ratios, the perturbations are added only to the grid points located within 6 km horizontally and 2 km vertically from the observed precipitation. Negative values of

mixing ratios after the perturbations are added are reset to zero. The pressure variable is not perturbed. These configurations are the same as in TX08a.

The first assimilation of simulated observations is performed at 25 min of model time and the analyses are repeated every 5 min until 100 min. The filter uses 40 ensemble members and a covariance inflation factor of 15% and a covariance localization radius of 6 km (Anderson 2001; Xue et al. 2005; Houtekamer and Mitchell 1998, 2001; Hamill et al. 2001). A single virtual polarimetric WSR-88D radar that scans the model atmosphere is located at the southwest corner of the model domain, as is the nonpolarimetric radar in XTD06. For more detailed information on the configuration of the assimilation experiment, the reader is referred to XTD06 and TX08a.

3.2.3 Experimental design

To examine the impact of assimilating polarimetric variables (Z_{DR}, Z_{dp}, and K_{DP}), in addition to the reflectivity at horizontal polarization (Z_H, which is what conventional WSR-88D radars observe) or in additional to both Z_H and V_r, on the analysis of the convective storm, we designed 10 experiments, which are listed in Table 3.1. Experiment Zh serves as the control run for the first set of the data impact experiments that include itself, ZhZdr, ZhZdp, ZhKdp and ZhZZK. Experiments ZhZdr, ZhZdp, ZhKdp test the impact of Z_{DR}, Z_{dp}, and K_{DP} data individually when assimilated in addition to Z_H. Experiment ZhZZK tests the combined impact of all three variables (Z_{DR}, Z_{dp}, and K_{DP}) together. Experiment VrZh is the control run for the second set of experiments that consists of itself, VrZhZdr, VrZhZdp, VrZhKdp, and VrZhZZK. In this set, the impact of polarimetric variables in addition to both radial velocity and

conventional reflectivity data is examined.

Table 3.1. List of experiments testing the impact of polarimetric variables.

Experiment	Observation(s) assimilated
Zh	Z_H (everywhere)
ZhZdr	Z_H and Z_{DR} (Z_{DR} > 0.3 dB)
ZhZdp	Z_H and Z_{dp} ($(Z_{dp})^{0.2}$ > 1.7 mm^6 m^{-3})
ZhKdp	Z_H and K_{DP} (K_{DP} > 0.9 degree km^{-1})
ZhZZK	Z_H, Z_{DR}, Z_{dp}, and K_{DP}
VrZh	V_r (Z_H > 10dBZ) and Z_H
VrZhZdr	V_r, Z_H, and Z_{DR}
VrZhZdp	V_r, Z_H, and Z_{dp}
VrZhKdp	V_r, Z_H, and K_{DP}
VrZhZZK	V_r, Z_H, Z_{DR}, Z_{dp}, and K_{DP}

TX05 shows that the Z_H data from echo-free regions help suppress spurious cells in those areas. The Z_H data within the entire radar range are therefore assimilated in all of our experiments. For the polarimetric variables, thresholds that are experimentally determined are applied to each variable. We performed experiments ZhZdr, ZhZdp, ZhKdp without thresholding and with various thresholds based on their SDs and found that applying thresholds can lead to better analyses. The thresholds used for Z_{DR}, $(Z_{dp})^{0.2}$, and K_{DP} in this study are 0.3 dB, 1.7 mm^6 m^{-3}, and 0.9° km^{-1}, respectively. In other words, we assimilate polarimetric variables only when their values are greater than their respective thresholds.

To help understand the need for thresholding for polarimetric variables, we investigate the effect of observational errors on the analysis in the current assimilation framework. In our one-moment microphysics scheme, all polarimetric variables including Z_{DR} are uniquely determined by the mixing ratios only, with assumed fixed values of DSD parameters. Therefore, they are to some extent correlated with each other. In practice, assimilating two (or more) observations taken at the same point and time that should be correlated may result in the deterioration of the analysis if the noise level is high in one or both observations. When a signal is weak, as is often the case with polarimetric data in many parts of a storm (see examples given in Chapter 2 and here in Fig. 3.1), it is possible that the noise dominates over the signal. In such a case, the assimilation of noise-dominated data may interfere with the assimilation of signals contained in other variables that are less susceptible to the noise (e.g., reflectivity). This can be inferred from the scatterplots of polarimetric variables versus reflectivity in Fig. 3.2. Figures 3.2a,c,e show the scatter diagram between truth (error free) reflectivity and truth (error free) polarimetric variables and Figs. 3.2b,d,f show the same plots between error-containing observations. It is clear from the plots that the relative errors are larger for small values and smaller for large values. In Figs. 3.2a,e, there are several lines showing high population densities of observation points that pack together. When a single hydrometeor dominates in many of the radar sampling volumes, such as snow at the upper levels and rain at the low levels, the functional relation between the reflectivity and the polarimetric variable stands out as a densely clustered curve. In Fig. 3.2a, the straight steeply sloped line corresponds to raindrops. In the error-free cases, all scatter away from the identifiable curves is due to the co-existence of more than one

hydrometeor species in the sampling volumes. For those who are interested in more detailed information on the impacts of noise on signals, many past studies are well documented in Doviak and Zrnic (1993) and Bringi and Chandrasekar (2001).

When the simulated errors are added to the error-free observations, the clearly defined lines become blurred, and overall there is much more scatter with the plots (Figs. 3.2b,d,f). For reflectivity difference $(Z_{dp})^{0.2}$ (Figs. 3.2c,d), the line broadening due to noise is more severe where the slope is low below a certain threshold. As a result, the reflectivity shows a much larger variability for small values of $(Z_{dp})^{0.2}$ in Fig. 3.2d. For K_{DP}, the effect of noise at low K_{DP} values is even more severe – below K_{DP} = 0.9° km^{-1}, no signal is perceivable due to noise (Fig. 3.2f). For this reason, the thresholding of polarimetric variables is clearly necessary, and their values are chosen based on the scatter plots in combination with sensitivity experiments, at levels below which noise dominates, as indicated by the horizontal dashed lines in the plots. These thresholds are applied to the simulated data. When the thresholds are increased above these levels, we found that the quality of analysis starts to decline because some useful signal is excluded. With the given thresholds, only 34.5%, 53.6% and 13.9% of Z_{DR}, Z_{dp}, and K_{DP} observations collected from the echo region (where observed reflectivity is greater than 0 dBZ) are assimilated. If more data could be used, the impact of polarimetric data to be shown later might have been larger.

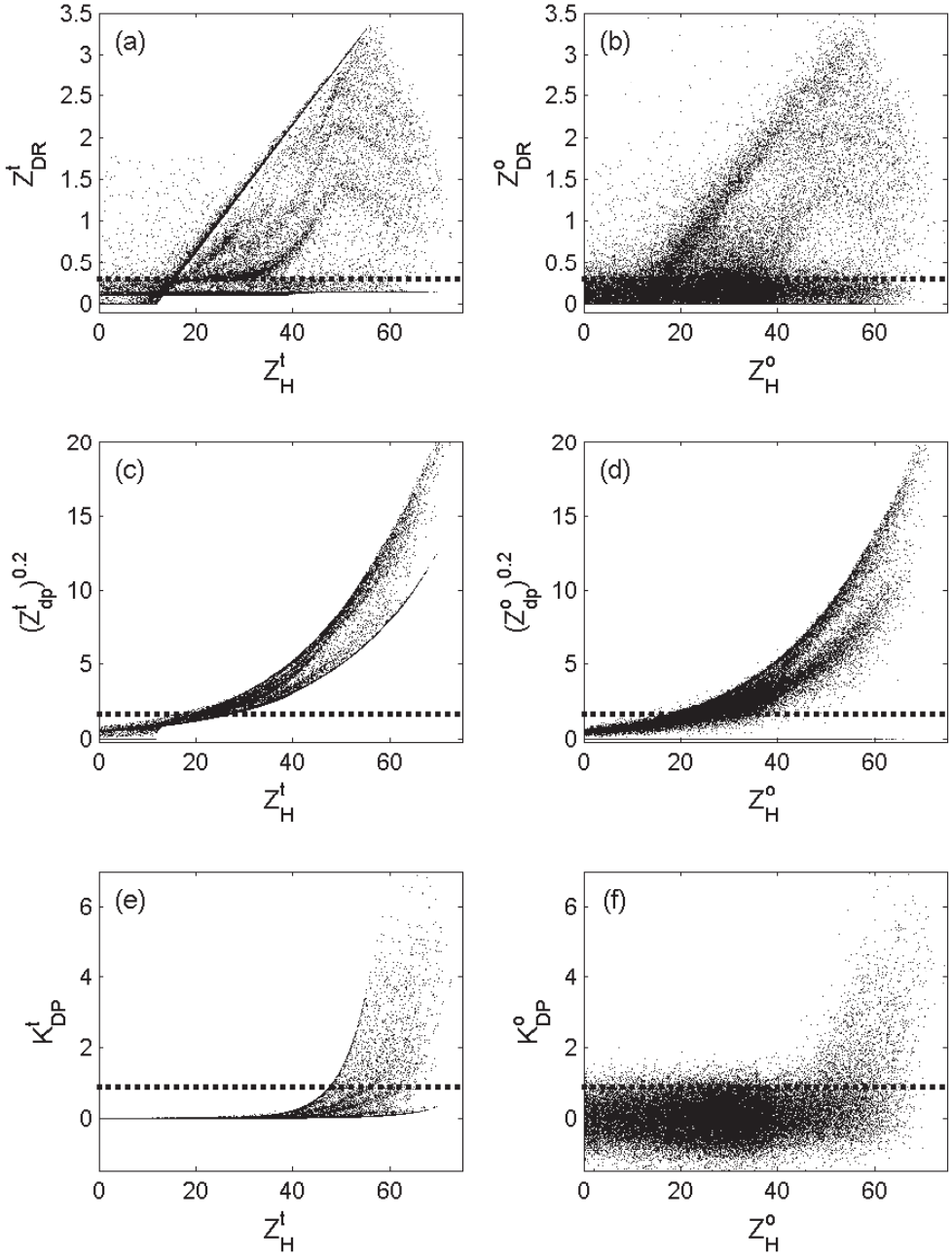


Fig. 3.2. Scatterplot of reflectivity vs (a),(b) Z_{DR}; (c),(d) $(Z_{dp})^{0.2}$, and (e),(f) K_{DP} for (a),(c),(e) truth and (b),(d),(f) observation. The thresholds applied to the observation in the assimilation are overlaid on each plot (thick dashed).

3.3 The impact of assimilating polarimetric variables

We examine, through the two sets of experiments listed in Table 3.1, the impact of Z_{DR}, Z_{dp}, and K_{DP} data when only Z_H is or when both V_r and Z_H are assimilated. The V_r data are only available in precipitation regions where reflectivity is greater than 10 dBZ following TX05.

Figure 3.3 shows the ensemble mean analysis and forecast RMSEs of model state variables during the assimilation cycles of experiments Zh and VrZh, which are our control runs. As in TX05 and XTD06, these errors are calculated in the regions where the truth reflectivity (Z_H^t) is no less than 10 dBZ. Additional details on the plots can be found in those papers. As mentioned earlier, the experiment names are self-descriptive. For example, experiment Zh assimilates Z_H data only and ZhZdr assimilates Z_H and Z_{DR} while experiment VrZhZZK assimilates V_r, Z_H, Z_{DR}, Z_{dp}, and K_{DP}.

Under the perfect model assumption, the solid curves in Fig. 3.3 show that reflectivity data alone can successfully reduce the RMSEs over the first 40 min or so of the assimilation window period to rather low levels. After t = 60 min., the RMSEs more or less stabilize. At the end of the assimilation window, the RMSEs of u and v are between 1 and 1.2 m s^{-1}, while that of w is about 0.6 m s^{-1}. The RMSEs of the hydrometeors are all below 0.1 g kg^{-1} except for q_v. On average over all assimilation cycles, additional 30 to 48 % of analysis error reduction in u, v, and w, and 17 to 27 % in the rest of variables except for q_v, which show about 36 % of error reduction, are achieved with the addition of V_r data. These results are consistent with those of TX05.

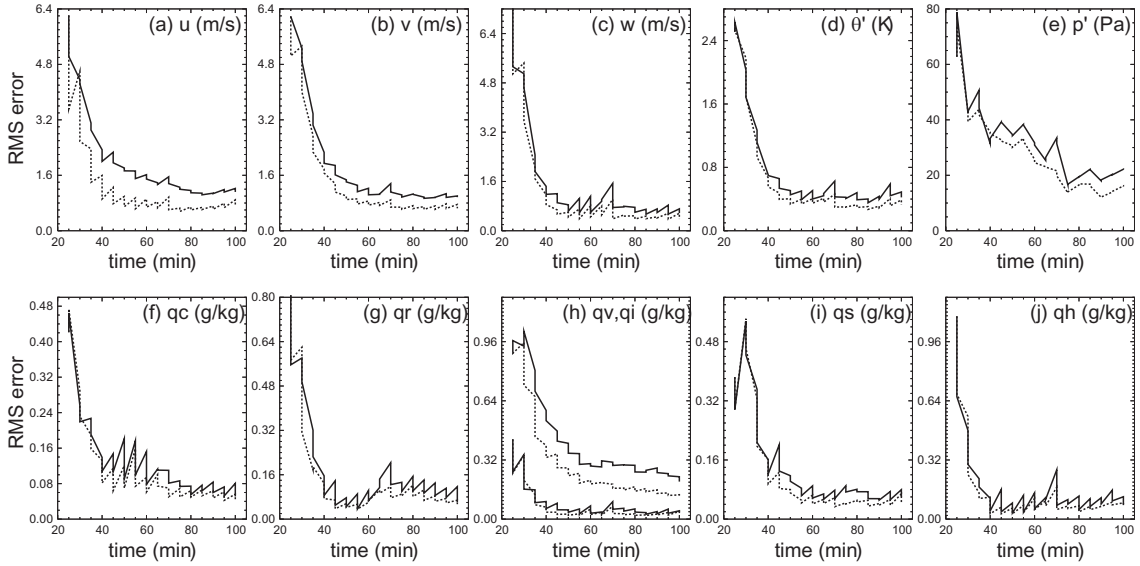


Fig. 3.3. The ensemble mean forecast and analysis RMSEs averaged over points at which the true reflectivity is greater than 10 dBZ for (a) u, (b) v, (c) w (d) perturbation potential temperature θ', (e) p', (f) q_c, (g) q_r, (h) q_v (the curves with larger values) and q_i (the curves with lower values), (i) q_s, and (j) q_h for experiments Zh (solid black) and VrZh (dotted black). The vertical straightline segments in the curves correspond to the reduction or increase in RMSEs or ensemble spreads by the data assimilation.

Because we are interested in if and how much the polarimetric data can further improve the analyses when they are assimilated in addition to reflectivity or both reflectivity and radial velocity data, we normalize the ensemble mean analysis RMSE of the data impact experiments using those of the corresponding control. Namely, the RMSEs of ZhZdr, ZhZdp, ZhKdp, and ZhZZK are normalized by the errors of Zh, and the errors of VrZh are used to normalize those of VrZhZdr, VrZhZdp, VrZhKdp and VrZhZZK. [These normalized RMSEs (NRMSEs) are shown in Figs. 3.4 and 3.6]. A smaller NRMSE suggests a larger improvement through the assimilation of additional variable(s).

Figure 3.4 shows that every polarimetric variable shows a degree of positive impact when assimilated individually in addition to reflectivity (Fig. 3.4), at least during the later assimilation cycles when the filter stabilizes. Generally, ZhZdp (dashed in Fig. 3.4) and ZhKdp (dotted in Fig. 3.4) produce better analyses than ZhZdr (solid in Fig. 3.4) during early-to-intermediate cycles and ZhZdr shows a bigger improvement than ZhZdp and ZhKdp during intermediate-to-later cycles. These results show that different observations may have different relative impact at the different times. At the early stage of assimilation when the forecast error is relatively large, the intensity information carried by Z_{dp} and K_{DP} seems to be more beneficial. Later in the assimilation period, Z_{DR} seem to provide additional information other than intensity.

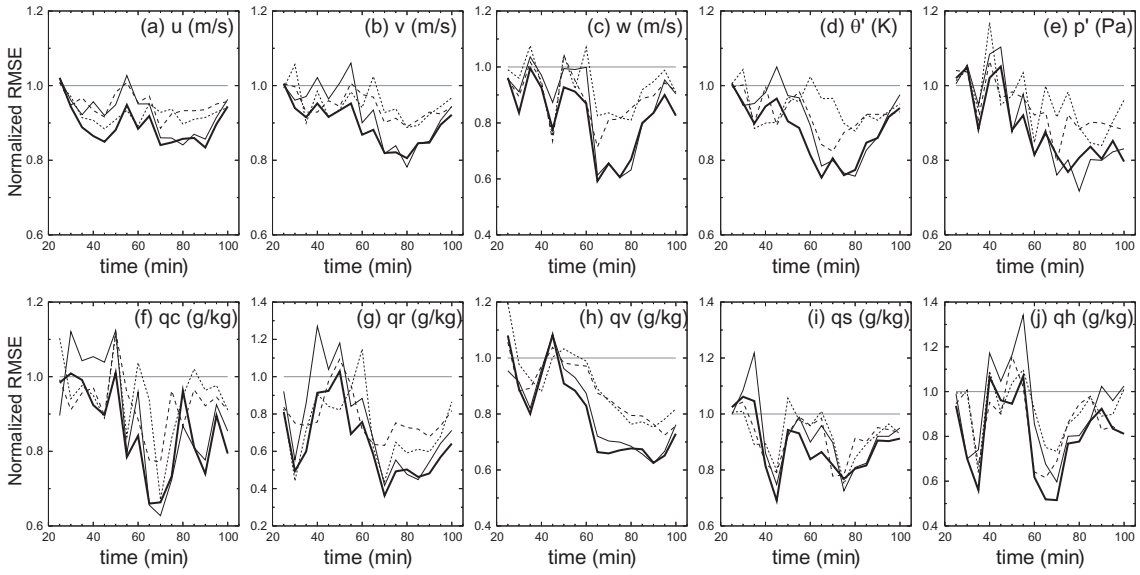


Fig. 3.4. The ensemble mean analysis RMSEs of experiments ZhZZK (thick solid), ZhZdr (solid), ZhZdp (dashed), and ZhKdp (dotted) normalized by those of experiment Zh. The reference horizontal line at a unity is overlaid.

From experiments ZhZdp, ZhZdr, ZhKdp, with the help of any one of the polarimetric variables, the normalized analysis RMSEs stay lower than those of experiment Zh after 60 min of model time for all variables except for q_h but there is a tendency for such error reductions to become smaller in the later assimilation cycles for many of the variables. This is believed to be due to the fact that by the time of the later cycles, the reflectivity data have had more time to correct the model state error while during the intermediate cycles, there is more room for the polarimetric variables to contribute, by accelerating the error reduction. During the earlier cycles, the positive impact of the polarimetric variables is questionable according to Fig. 3.4, which suggests that when the model state estimation is relatively poor (during the earlier cycles), and the positive impact of the polarimetric variables is harder to realized.

After 60 min of model time, in general, ZhKdp shows the smallest error reduction among ZhZdr, ZhZdp and ZhKdp on average. Their error reduction behaviors are all similar to each other with the exception of q_c and p' during the later assimilation cycles. Experiment ZhZdp shows generally larger RMSEs than ZhZdr, but slightly smaller than or similar to ZhKdp in most of variables. The polarimetric variables are more beneficial to w, q_v, and q_r, with the reduction of error in q_r being the largest. This is probably not surprising because rainwater mixing ratio, q_r, is directly involved in the calculation of Z_{DR}, Z_{dp}, and K_{DP}, and the signatures of these variables are strongest where rain mixing ratio is larger (see Chapter 2). These variables are related to w and q_v through their direct connection to the updraft/downdraft intensities and microphysics. For example, q_v converts to q_r through condensation in the updraft and is created from q_r by evaporation in the downdraft. Among the other state variables, the improvements

to u, v, and q_s are rather smaller. Even though q_s is directly related to polarimetric variables, the polarimetric signatures related to dry-ice-phase hydrometeors are generally weak so that most of the observations containing information on q_s are screened out by the observation thresholding. An interesting point is that the analysis error reduction is relatively large in q_h. This is because a considerable amount of q_h information is available from wet hail, which survives the thresholding in the deep layer below the melting level.

Among experiments ZhZdp, ZhZdr, and ZhKdp, experiment ZhZdr has the greatest impact. This may not be intuitively obvious because Z_{DR} mainly carries information on the difference between reflectivity at horizontal and vertical polarization; however, it does not provide much information on the intensity of the reflectivity. On the contrary, K_{DP} and Z_{dp} are directly related to mixing ratios and are expected to be more useful for quantification. This behavior may be explained in terms of independent information content. The Z_{DR} contains information on the mean shape and orientation of hydrometeors and is proportional to the median diameter of precipitation particles in the radar resolution volume. The Z_H is mainly related to the hydrometeor concentration. For rain drops, the shape is a strong function of size and, therefore, Z_{DR} and Z_H share some information in common. Both K_{DP} and Z_{dp} contain the information on both hydrometeor concentration and shape. As discussed earlier in section 3.2.3, with a single-moment scheme, all polarimetric variables are correlated to the reflectivity, with the correlation between Z_{DR} and Z_H being the smallest; the independent information content in Z_{DR} can therefore have a larger impact. The intensity information should have already been well captured by the Z_H data. Another

perhaps more important issue is that, with the current single-moment microphysics scheme used, the DSD parameters, including intercept parameters and hydrometeor densities are fixed and cannot be adjusted using the information contained in the polarimetric radar data. The impact of polarimetric data may increase when adjustments to these parameters are allowed, via, for example, parameter estimation (TX08a, TX08b) or if a multimoment scheme is used. In those cases, the response of the data assimilation system to the polarimetric data may become more physical.

When all three polarimetric variables are assimilated together, the analysis improvement is seen to further increase in general, although ZhZdr does do better temporarily after 60 min of model time for p', ZhZdp does better for q_r and q_h, and ZhKdp does better for p'. It is encouraging that experiment ZhZZK successfully reduces the analysis RMSEs even when individual polarimetric parameters show little or no positive impact. For instance, the normalized RMSEs of ZhZZK stay low at 65 min of model time for q_s, at 85 min for q_c, and at 90 min and 100 min for q_h, while the corresponding RMSEs of ZhKdp, ZhZdr are greater than 1.

From Fig. 3.4j, we see that K_{DP} and Z_{dp} help reduce RMSE up to 60 minutes, Z_{DR} from 60 min to 85 min, and Z_{dp} helps reduce the RMSE after 85 min. Similar behaviors are seen in many other variables (Figs. 3.4a-c,f,g).

The percentage improvement over experiment Zh averaged over the last nine cycles is summarized in Table 3.2. From Table 3.2, we can see that all model state variables experience analysis error reduction when assimilating polarimetric data. The improvement is greatest in q_r, which has an approximately 29 ~ 41 % improvement in ZhZdr, ZhZdp, and ZhKdp and more than 50% improvement when all three variables

are assimilated. As discussed in Chapter 2, K_{DP} is more linearly proportional to rain mixing ratio and has little sensitivity to other species. Therefore, it is expected to be more useful for determining q_r than other variables, including Z_H, even if we take the thesholding into account. Actually, only 14 % of available K_{DP} observations are used in the analysis, which is about 40% of Z_{DR} and about 25% of Z_{dp} observations. Considering this, the impact of K_{DP} on q_r analysis is rather large.

Table 3.2. The improvement over the experiment Zh for the experiments ZhZdr, ZhZdp, ZhKdp, and ZhZZK and over the experiment VrZh for the experiments VrZhZdr, VrZhZdp, VrZhKdp, and VrZhZZK averaged over the last nine cycles (60-100 min of model time). The improvement is expressed in percentages relative to the corresponding control experiment.

Variables	Improvement (%)				Improvement (%)			
	ZhZdr	ZhZdp	ZhKdp	ZhZZK	VrZhZdr	VrZhZdp	VrZhKdp	VrZhZZK
u	10.4	6.2	7.8	14.6	5.0	2.8	4.4	7.1
v	13.1	7.2	6.1	17.1	0.7	0.8	3.0	1.2
w	22.3	14.4	9.7	28.3	12.3	7.3	7.1	18.7
θ'	15.7	10.6	6.2	19.6	3.2	2.6	4.6	7.4
p'	19.7	10.8	9.1	20.0	8.6	2.0	4.9	11.9
q_v	29.0	18.3	17.5	33.5	9.5	5.1	3.9	9.9
q_c	20.0	11.0	7.8	23.4	8.1	5.2	4.7	12.9
q_r	40.7	28.7	31.1	51.9	31.1	22.1	30.5	45.9
q_i	15.0	8.4	9.4	21.9	4.0	4.2	7.8	9.3
q_s	12.2	9.5	9.0	18.4	3.0	2.2	1.1	6.7
q_h	15.4	19.9	11.8	26.1	3.2	8.8	7.2	14.4
tot	123.5	145.0	125.5	274.8	88.7	63.1	79.2	145.4

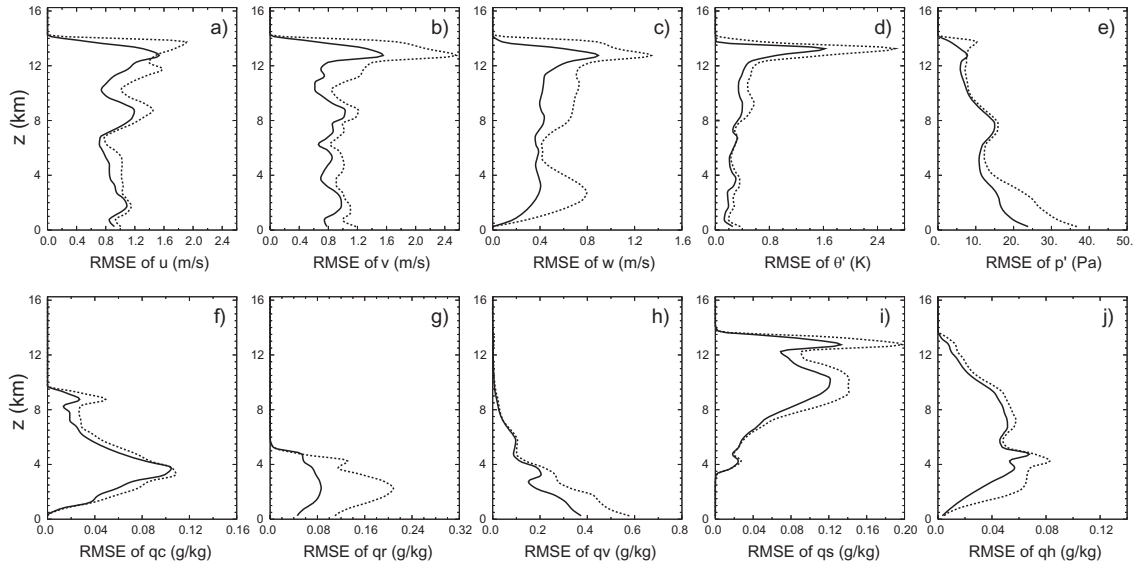


Fig. 3.5. The vertical profile of RMS analysis errors averaged over points at which the truth reflectivity is greater than 10 dBZ for (a) u, (b) v, (c) w, (d) θ', (e) p', (f) q_c, (g) q_r, (h) q_v, (i) q_s, and (j) q_h at 80 min of experiments Zh (dashed) and ZhZdr (solid).

Figure 3.5 shows the vertical profiles of the RMSEs averaged over points at which the truth reflectivity is greater than 10 dBZ for experiments Zh (dotted) and ZhZdr (solid) at 80 min. It is seen that the errors of all variables are reduced at almost all levels by assimilating Z_{DR}, with the exceptions being with u in a shallow layer between 12.5- and 13.0-km height. Considering that most Z_{DR} observations at the high altitudes are excluded by the threshold constraints (see Table 3.1) because Z_{DR} values are typically small for ice phase particles (see Table 3.1), the fact that improvements are found at all levels is encouraging specially with the large error reduction at the upper levels in v, w, θ', and q_s. Also, the error reduction is generally largest where the RMSE profiles peak. Apparently, direct improvement to the analysis at the low levels is propagated upwards, or throughout the computational domain, through the dynamic

prediction model. The reduction of errors in q_r and q_h below 5 km where the melting occurs is also noticeable at the time shown.

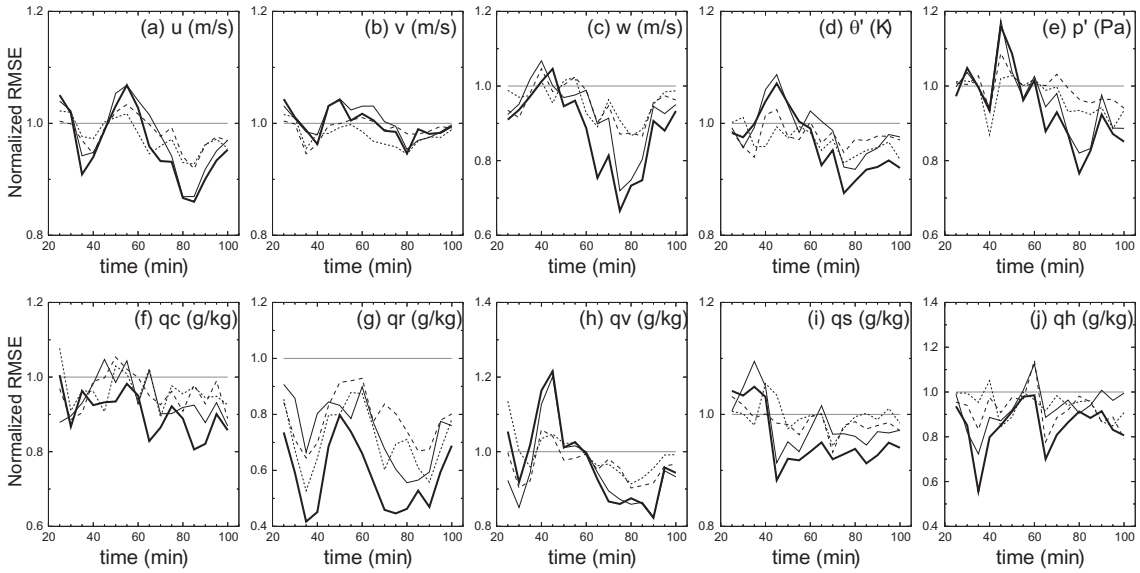


Fig. 3.6. As in Fig. 3.4, but for experiments VrZhZZK (thick solid), VrZhZdr (solid), VrZhZdp (dashed), and VrZhKdp (dotted).

In the next set of experiments (VrZh, VrZhZdr, VrZhKdp, and VrZhZZK), we examine the impact of Z_{DR}, Z_{dp}, and K_{DP} data when both V_r and Z_H are assimilated. From Fig. 3.6, we see that in such a case, the impact of polarimetric variables is rather small though still positive in general during middle to later cycles in most of the state variables although temporary deterioration can occur with q_s and q_h. Variables u and w show decreasing error reduction starting around 80 min of model time and the RMSE reduction is minimized at the end of assimilation cycles while the improvement is very small in v. Such a diminishing impact of the additional polarimetric variables appears again due to the very accurate analysis that one can already achieve by using reflectivity

and radial velocity data, especially after they have had a sufficiently long time to contribute to the state estimation.

The gross improvement can be assessed more easily from Table 3.2. The error reduction characteristics are generally similar to but somewhat different from those of the previous set of experiments. As in previous cases, the improvement is generally larger in w, q_v, and q_r and smaller in u, v, and q_s either when polarimetric variables are assimilated individually or when all are assimilated together. Again, q_h shows large improvement, even larger than those of w and q_v in VrZhZdp and VrZhKdp, and that of q_v in VrZhZZK. Another interesting point is that the gross error reduction by VrZhZdp is smaller than that of VrZhKdp in contrast to the experiments without V_r. When V_r is assimilated, the percentage improvements by polarimetric variables relative to the control experiment are significantly reduced with the percentage reduction by K_{DP} being the smallest in general comparing to the corresponding experiments without V_r. Moreover, the percentage improvement in q_r is similar between VrZhKdp and ZhKdp.

From Table 3.2, we can see that the NRMSE reduction by VrZh for the 11 model state variables range from 1% to 46 % when all polarimetric data are assimilated together. However, these additional error reductions may not be very meaningful in practice. Within the current OSSE framework using a perfect prediction model, the analysis obtained using V_r and Z_H alone is already very good; the RMS analysis errors in u and v, for example, are quickly reduced to below 1 m s^{-1} within 4-5 cycles (Figs. 3.3a-c), therefore there is little room for further improvement (the 1 m s^{-1} analysis error is already at or below the level of V_r RMSE, which is 1 m s^{-1} as defined in section 3.2). For real data cases where model error tends to be rather large, the extra information

content afforded by the polarimetric data may produce a larger impact, especially when the polarimetric data are used to correct microphysics-related model error. For the single-moment scheme used here, many uncertainties exist with the values of the intercept parameters associated with the assumed exponential DSDs, and with snow and hail densities. TX08b shows that large analysis error can result when errors exist in these DSD parameters and the resultant analysis errors tend to be larger than the amount of error reduction achieved here through the assimilation of additional polarimetric variables. TX08b also shows that the errors in the DSD parameters can often be corrected through EnKF-based parameter estimation, although nonuniqueness in the solution does seem to exist. The final parameter estimation was found to be sensitive to the initial guess when multiple parameters are estimated together. It was suggested there that additional polarimetric data could impose additional constraints that may improve the uniqueness of the solution, given the fact that the polarimetric data contain DSD information. Using additional polarimetric parameters to improve the DSD parameter retrieval whereby reducing microphysical uncertainties and model error is discussed in Chapter 4, where polarimetric data assimilation plays an even greater role.

Last, we also performed additional experiments assimilating combinations of any two of Z_{DR}, Z_{dp}, and K_{DP} and these experiments exhibit lower analysis errors for most variables than experiments assimilating any one of the two variables involved in terms of time-averaged RMSEs after 60 min. One exception is found in q_h of ZhZdrZdp whose time-averaged RMSE is smaller than that of ZhZZK. For example, ZhZdrKdp result in better analyses than those of ZhZdr and ZhKdp in terms of RMSE but worse analyses than that of ZhZZK. This is also true when V_r is assimilated. In this case,

VrZhZZK produces the best analyses among all experiments including those assimilating any two combinations of polarimetric variables with one exception in v of VrZhZdpKdp. So in general, it is better to assimilate more polarimetric variables.

3.4 Summary and further discussion

In this chapter, an ensemble Kalman filter system that incorporates the ability to assimilate polarimetric radar variables is described. It employs the observation operators developed in Chapter 2. The polarimetric variables considered include the differential reflectivity, Z_{DR}, reflectivity difference Z_{dp}, and specific differential phase K_{DP}. A new error model for reflectivities at horizontal and vertical polarizations is developed in this study that includes both correlated and uncorrelated errors, and the relative errors of which are assumed to have Gaussian distributions in the linear domain based on Xue et al. (2007). This model gives realistic errors for the derived quantities, such as Z_{DR} and Z_{dp}. The simulated error-containing radar observations are shown, for example, for the truth simulation of a supercell.

The enhanced EnKF assimilation system is used to assimilate radar data sampled from a simulated supercell storm, to examine the impact of additional polarimetric measurements, including Z_{DR}, Z_{dp}, and K_{DP}, on the quality of storm analysis under the perfect model assumption. It is found that the assimilation of these variables, in addition to the reflectivity at horizontal polarization (reflectivity measurement of non-polarimetric radars), helps further reduce the analysis error and the improvement during the intermediate and later assimilation cycles can be quite significant for some state variables. The results also show that the analyses for all model state variables are

improved at all vertical levels in general. Although Z_{DR} does not direct reflect the magnitude of hydrometeor concentration, it gives the largest impact among the three polarimetric variables examined. When both V_r and Z_H are assimilated, the impact of additional polarimetric variables becomes smaller, partly because the analyses obtained with V_r and Z_H alone are already very good.

The result shows that applying data thresholding when assimilating polarimetric variables leads to better analysis. When two or more observations are taken at the same point and time that are somewhat correlated, the assimilation of noise-containing observations may interfere with the assimilation of signals contained in other observations. This noise effect can limit the improvement to the analysis or even harm the analysis, especially when the noise level is high, which is likely to be true where a signal is weak. Applying thresholding also reduces the assimilation cost.

It is suggested that polarimetric radar data can be very useful for estimating DSD parameters, such as the intercept parameters and hydrometeor densities used in single-moment microphysics schemes, because of their information content on DSDs. The DSD parameter estimation experiments using our EnKF framework are presented in Chapter 4, following the work of TX08a and TX08b.

We expect to see a larger impact for real data cases where the state estimation using V_r and Z_H is generally not as good. In such cases, additional polarimetric data are expected to play a larger role.

We also pointed out earlier that the error levels assumed for the polarimetric variables are on the larger side. The data thresholding necessitated by the relatively

larger errors caused the discarding of large fractions of the simulated polarimetric observations. If the actual errors are smaller, larger impacts may be expected.

Last, we point out that even though correlations among the reflectivity-related observation variables and their errors are expected, in the EnSRF used here, which assimilates observations serially, one at a time, all observations are assumed to be uncorrelated. The ideal way of processing correlated observations is to either transform the observation variables into a space where the assimilated quantities are no longer correlated (this may or may not be possible) or to use an algorithm that can take the observation error covariance into account. Their practical implementations are often nontrivial, however. To have an idea how much our observation errors are correlated, we calculated the observational error correlation coefficients between Z_H and the polarimetric variables and found the coefficients to be 3.8×10^{-2}, 0.37, and -3.2×10^{-3} for those between Z_H and Z_{DR}, $(Z_{dp})^{0.2}$, and K_{DP}, respectively. These correlations suggest that the results of our serial algorithm are probably reasonable.

Chapter 4 [3] Simultaneous Estimation of Microphysical Parameters and Atmospheric State using Simulated Polarimetric Radar Data and Ensemble Kalman Filter in the Presence of Observation Operator Error

4.1 Introduction

The accuracy of the numerical weather prediction (NWP) is subject to two factors - error in the initial condition and deficiency of the NWP model. A considerable amount of research has focused on developing more advanced techniques to minimize the errors in the initial condition (Le Dimet and Talagrand 1986; Courtier and Talagrand 1987; Evensen 1994; Evensen and Leeuwen 1996; Burgers et al. 1998; Houtekamer and Mitchell 1998; Anderson 2001; Bishop et al. 2001; Whitaker and Hamill 2002; Evensen 2003; Tippett et al. 2003; Gao and Xue 2007; Liu et al. 2007). Among these, the ensemble Kalman filter (EnKF) techniques are thought to be attractive because of their ability to make effective use of prediction models and to deal with complex and highly nonlinear processes in the assimilation process. Previous studies using the EnKF method have achieved encouraging success for applications at large scale through convective scale (e.g., Houtekamer et al. 2005; Whitaker et al. 2004; Snyder and Zhang 2003; Tong and Xue 2005b, TX05 hereafter; Xue et al. 2006, hereafter XTD06).

On the other hand, the deficiency in the NWP models, which is commonly referred to as model error, has received less attention until recently, because the

[3] This chapter is conditionally accepted as: Jung. Y., M. Xue, and G. Zhang, 2008: The estimation of microphysical parameters and atmospheric state using simulated polarimetric radar data and ensemble Kalman filter in the presence of observation operator error. *Mon. Wea. Rev.*

characteristics of model error are little known and its statistical properties are poorly understood (Dee 1995; Houtekamer et al. 2005). Model error can arise from many sources such as insufficient resolution in time and/or space, misrepresentation of the physical and sub-grid scale processes, and the use of non-physical model boundaries and/or external forcing.

It has been observed in certain EnKF studies that model error can dominate the error growth in data assimilation cycles and must be parameterized to prevent the filter from diverging from its truth state (Houtekamer et al. 2005). One way to account for model error within the EnKF system is to add the so-called additive error to the model state by assuming an error model (Lawson and Hansen 2005). Houtekamer et al. (2005) used additive errors by assuming model error covariance that has the same functional form as the forecast error covariance used in a 3DVAR system. Their experiments using a global model showed that the added model errors increased the ensemble spread to the level of ensemble mean error. Hamill and Whitaker (2005) performed several experiments to account for the model error due to unresolved scales using a global spectral model. They compared the two most popular methods for parameterizing model error – covariance inflation and additive error models. Additive error was randomly sampled from the time series of the difference between two runs at different resolutions. Their results performed at the global scale show that the additive error model outperformed the covariance inflation method and produced more accurate analyses. The ability of the additive error approach in increasing the space spanned by the existing ensemble perturbations is an advantage but the added errors are usually flow independent and therefore inconsistent with the actual flow.

Difficulties can arise when we attempt to apply these methods to the convective-scale where model error is very flow- and situation-dependent. For this reason, the estimation of tunable model parameters, which often have a profound impact on the forecast, using the data assimilation scheme appears to be an attractive alternative or addition to the aforementioned methods for dealing with convective-scale model error. Early work using adjoint-based parameter estimation can be found in fields such as hydrology that solves the problem of aquifer identification (e.g., Yakowitz and Duckstein 1980). In meteorology, such studies include the estimation of nudging coefficients using the four-dimensional variational assimilation (4DVAR) method (Zou et al. 1992), statistical model error parameters using a maximum-likelihood method (Dee and Silva 1999), and estimation of wind-stress coefficient using the extended Kalman filter method (Hao and Ghil 1995). The relative importance of optimal parameter values versus optimal initial condition of state is discussed by Zhu and Navon (1999) using a 4DVAR system of a full-physics global spectral model. Their results show that the impact of optimal parameters on the forecast persists even after the impact of the optimal initial condition has been lost. A comprehensive review on parameter estimation studies in meteorology and oceanography up to the mid 1990s can be found in Navon (1997).

Anderson (2001) proposed using EnKF for simultaneous estimation of parameters and state. Several studies have since shown that EnKF is capable of successfully estimating parameters through the data assimilation process and may therefore help improve the subsequent forecast (Annan et al. 2005a; Annan and Hargreaves 2004; Annan et al. 2005b; Hacker and Snyder 2005; Aksoy et al. 2006b, a).

More recently, Tong and Xue (2008b; 2008a, hereafter TX08a and TX08b) applied the EnKF method to the estimation of fundamental microphysical parameters in a storm-scale model. In TX08a, parameter identifiability is addressed through an investigation of correlation fields and a detailed sensitivity analysis. TX08b performed simultaneous estimation of up to 5 microphysical parameters using simulated radar data and found, as in Aksoy et al. (2006a), that a single imperfect parameter can be successfully estimated while the accuracy of estimation declines as the number of error-containing parameters increases. Another common conclusion of both studies is that the parameter estimation is beneficial in reducing errors in both estimated parameters and state. The studies also indicate that the parameter estimates are sensitive to the filter configuration and significant nonlinearities exist between model parameters and state variables, so that an attempt to improve one parameter may influence the estimate of other parameters.

The matter of simultaneous parameter and state estimation is further complicated when the very same parameters to be estimated are involved in the forward observation operators that link the model state to the observations. In past studies, either the parameters to be estimated were not involved in the observation operators, or the observation operators were assumed to be perfect. In the case of radar reflectivity-related observations, the model microphysical parameters also appear in the observation operators. In TX08b that estimates microphysical drop size distribution (DSD) parameters from simulated reflectivity data, the observation operators were assumed to be perfect, i.e., correct parameter values were used in the operators. In that study, difficulties were encountered when estimating multiple DSD parameters and this arose from the fact that the responses to error in different parameters compensate each other

in terms of the observed radar reflectivity, causing solution non-uniqueness. This result suggests that additional constraints provided by polarimetric radar measurements may help improve the well-posedness of the problem (Jung et al. 2008b; Jung et al. 2008c). Chapter 3 showed positive impacts of directly assimilating polarimetric variables on state estimation in a perfect model scenario.

In this chapter, we extend the earlier studies of TX08a and TX08b that performed simultaneous DSD parameter and state estimation from reflectivity only and assuming perfect observation operators, and the studies of Chapters 2 and 3 that assimilated simulated polarimetric radar data with a perfect model, and perform simultaneous state and parameter estimation from polarimetric radar data whose observation operators also contain DSD parameter error. We attempt to quantitatively assess how additional polarimetric data can improve the parameter and state estimation using the EnKF approach. The forecast model, EnKF assimilation system and the design of OSSEs are first described in section 4.2, along with a discussion on the characteristics of the parameters to be estimated. Section 4.3 discusses the results of the sensitivity analysis and section 4.4 examines the impact of polarimetric radar data on the parameter and state estimation. A summary and conclusions are given in section 4.5.

4.2 Model and experimental design

4.2.1 Forecast model and filter configuration

The same as the OSSE studies of TX08a,b, capter 2 and 3, a truth simulation is created using the Advanced Regional Prediction System (ARPS, Xue et al. 2000; 2001; 2003) for a supercell storm. The ARPS is a fully compressible and nonhydrostatic

atmospheric prediction model. The ARPS prognostic variables include three velocity components u, v, and w; potential temperature θ; pressure p; and mixing ratios of water vapor, cloud water, rainwater, cloud ice, snow aggregate, and hail (q_v, q_c, q_r, q_i, q_s, and q_h, respectively) with the Lin et al. (1983, hereafter LFO83) ice microphysics scheme. The turbulence kinetic energy is another prognostic variable used by the 1.5-order subgrid-scale turbulence closure scheme. The ARPS model is also used for the sensitivity analysis and in the state and parameter estimation.

The configurations of the forecast model and assimilation system used here are very similar to those used in Tong and Xue (2005; 2008a; 2008b), except for one major modification: the forward observation operator for reflectivity uses the one developed in Chapter 2 instead. The capabilities to assimilate polarimetric data were developed in Chapters 2 and 3, although the data are also used for parameter estimation here. The size of ensemble is 80 and no covariance inflation is applied. The effect of terminal velocity is assumed to have been removed from the radial velocity data in this study.

The sounding of the 20 May 1977 Del City, Oklahoma, supercell storm (Ray et al. 1981) is used by the truth storm simulation. The CAPE of this sounding is 3300 J kg^{-1}. The grid spacing is set to 2 km horizontally and 0.5 km vertically. The dimension of the model domain is 64×64×16 km^3 and a virtual polarimetric Weather Surveillance Radar-1988 Doppler (WSR-88D) radar is located at the south-west corner of the domain. The storm is initiated by a 4-K ellipsoidal thermal bubble with a 10-km horizontal radius and a 1.5-km vertical radius centered at x = 48 km, y = 16 km, and z = 1.4 km. The time step for model integration is 6 seconds with 3 seconds for the acoustically-

active model equation terms. These configurations are essentially the same as used in TX05, TX08a, and Chapter 2.

The ensemble square-root filter (EnSRF) proposed by Whitaker and Hamill (2002) is employed, in which the observations are serially assimilated. More detailed information on the filter implementation can be found in XTD06 and TX08a.

Following TX08a, and TX08b, spatially smoothed stochastic perturbations with standard deviations of 2 m s^{-1} for velocity components (u, v, and w), 2 K for potential temperature (θ), and 0.6 g kg^{-1} for mixing ratios of hydrometeors (q_v, q_c, q_r, q_i, q_s, and q_h,) are added to the initially horizontally homogeneous first guess defined by the Del City sounding to initialize the ensemble members at t = 20 minutes of model time. The perturbations are added at the grid points located within 6 km horizontally and 3 km vertically from observed reflectivity. As in previous studies of TX08a,b, pressure is not perturbed. The covariance localization radius is set to 6 km.

A 80-minute assimilation window is used with the first analysis at 25 minutes of model time and the last at 100 minutes. Radar volume scan data are available and assimilated every 5 min. Reflectivity data from the entire domain, including the non-precipitation regions, are assimilated and used to update all state variables while radial velocity data, from regions where reflectivity is greater than 10 dBZ, are used to update wind variables (u, v, and w) only; it is found in our experiments that updating thermodynamic and microphysical variables using radial velocity does not further the analysis.

4.2.2 Simulation of observations

Detailed information on the forward observation operators that link model state variables with the polarimetric radar variables can be found in Chapter 2; these operators are used to generate error-free observations. The error models described in Xue et al. (2007) and Chapter 3 are used to generate simulated observation errors with slightly different error statistics. In this study, we assume that a basic quality control process has been applied to the observations prior to the assimilation. The effect is achieved by limiting the modeled reflectivity error samples to within 5 times their standard deviation, which correspond to 10 dBZ (larger error samples are dropped). To accommodate this change while keeping the error standard deviations (SDs) at a similar level as in Chapter 3, the correlated and uncorrelated parts of error for reflectivity are increased to 40% and 2.7% of the truth reflectivity, respectively. The resultant error distribution is similar to that of Xue et al. (2007) (solid line in their Figure 1) except for a shorter tail on the negative end (not shown). Therefore, the effective error SDs of the simulated observations are 1 m s^{-1} for V_r, about 2 dBZ for reflectivity at the horizontal polarization (Z_H), close to 0.2 dB for differential reflectivity (Z_{DR}), and 0.5 degree km^{-1} for specific differential phase (K_{DP}). The same SDs are specified in the filter for the corresponding observations. Reflectivity difference Z_{dp} is not examined here since it exhibits the highest correlation to Z_H among the polarimetric variables (Chapter 3), and hence is believed to contain the least independent information.

4.2.3 Parameters to estimate

LFO83 used in the ARPS model is a single-moment 5-class (cloud water, cloud ice, rainwater, snow and hail) bulk microphysics scheme, in which the DSD is described by an exponential function with a fixed intercept parameter. The water amount of hydrometeors in each category is represented by the corresponding mixing ratio, and it changes through interactions with the other categories. Such interactions include condensation or deposition, collection, breakup, freezing, evaporation or sublimation, melting, and precipitation sedimentation. DSD-related parameters including bulk density and intercept parameter of the DSD of each category explicitly appear in the equations for microphysical processes and can greatly influence the magnitude and relative importance of those processes. Briefly, the intercept parameter is the product of the total number concentration and the slope parameter of the exponential distribution (see Eqs. 1 through 6 of LFO83). Significant uncertainties exist in them because these parameters, which vary significantly both in time and space in nature, are usually predefined as constants in single-moment microphysics schemes. TX08a demonstrated through sensitivity analysis that the error in the intercept parameters and the bulk densities considerably influence the storm evolution. In this study, the same set of parameters are selected for parameter estimation under the assumption of imperfect observation operators; these parameters are the intercept parameters for rain n_{0R}, snow n_{0S}, and hail n_{0H}, and the bulk densities for snow ρ_S and hail ρ_H.

4.2.4 Parameter estimation procedure

The parameters to be estimated are given first-guess values at the beginning of assimilation cycles that typically deviate from the truth values; they are further perturbed for each of the ensemble members to form an ensemble of parameter values. Their values are adjusted/updated during the EnKF assimilation cycles. The update of these parameters in the early assimilation cycles when the errors in the estimated state is still very large is found to hurt rather than help parameter estimation; the estimated parameter values easily drift away from the truth, because the covariance between the parameters and observations at this early stage is very unreliable. Since the success of parameter estimation and the convergence rate depend on the filter performance of the previous assimilation cycles and the error is cumulative, larger error in the early cycles can significantly slow down the parameter estimation process (TX08a). As error in the estimated state can usually be significantly reduced in the first 2 to 3 cycles, we delay the parameter estimation until 40 minutes of model time or the time of the fourth EnKF analysis. During the assimilation period, parameter values estimated in the previous assimilation cycle are used in the forecast model as well as the observation operators of the following cycle. To prevent the collapse of the parameter variance because of the lack of dynamic error growth in the parameters, a covariance inflation procedure following Aksoy et al. (2006a) and TX08b is applied, which restores the parameter spread to predefined minimum value after each analysis cycle, when the prior parameter spread is smaller than this. For the logarithmically transformed intercept parameters, this predefined minimum spread is set to 1 m^{-4}, for logarithmically transformed snow and hail densities, it is set to 0.5 kg m^{-3}.

4.2.5 Design of parameter estimation experiments

We first perform five sets of single-parameter estimation experiments that examine the capability of the EnKF when only a single parameter contains error. We then perform a set of experiments in which 5 parameters are unknown. However, our main focus is on the improvement that can be obtained by using additional polarimetric data. Following TX08b, the radial velocity is not used in the parameter estimation due to its small response to the change in parameter values as well as the fact that it is not a direct function of hydrometeors. The radial velocity data are used for state estimation, however.

In the single-parameter estimation experiments, one of the five parameters starts with an incorrect first-guess value while in the five-parameter experiments all five parameters start incorrect. In the experiments where the parameter error is involved in the observation operators, the forecast and analysis trajectory is found to be very sensitive to the initial perturbations of the parameters. To increase the robustness of our estimation, we perform five parallel experiments that only differ in the sampling of initial parameter perturbations; the same was also done in Aksoy et al. (2006a).

As in TX08b we sample the random perturbations in the log domain (with 10 log(x) transform) which avoids negative values of intercept parameters and bulk densities. With this procedure, unrealistically small or large parameter values can occur occasionally, causing forecast instability. Such experiments were rerun using reduced large and small time step sizes of 2 and 0.5 seconds respectively. Table 4.1 lists the true and first guess values of the parameters. Because the Guassian random perturbations are sampled in the log space, after the ensemble mean of the parameters after being

converted back to their original space is usually not the same as the ensemble mean in the log space. As in TX08b, the parameter estimation is performed in the log space of the parameters while the ensemble prediction uses their values in the original scale.

Table 4.1. Microphysical parameters and their uncertainty ranges used in the sensitivity experiments, and their initial guesses as used in the parameter estimation experiments.

Parameter p_i		Rain intercept n_{0R} (m^{-4})	Snow intercept n_{0S} (m^{-4})	Hail/graupel intercept n_{0H} (m^{-4})	Density of snow ρ_S (kg m^{-3})	Density of hail/graupel ρ_H (kg m^{-3})
Control (assumed true) values of parameter p_i^t		8×10^6	3×10^6	4×10^4	100	913
Parameter values used in the sensitivity test	Lower bound	3×10^6	5×10^5	4×10^2	20	400
	Upper bound	8×10^7	1×10^8	4×10^6	400	913
Three initial guesses used in the single-parameter estimation experiments		3×10^6	7×10^5	4×10^3	50	400
		2×10^7	3×10^7	4×10^5	300	700
		8×10^7	1×10^8	4×10^6	400	750
Two initial guesses used in the five-parameter estimation experiments		3×10^6	7×10^5	4×10^5	50	400
		2×10^7	3×10^7	4×10^6	300	700

Within the first few cycles when the error covariance is still poor, the error in the estimated parameters often grow too large to prevent successful estimation in later cycles or even cause instability in the model integration. To avoid this problem, we

constrain the parameters within their respective lower and upper bounds, which are the same bounds, used in the sensitivity experiments (see Table 4.1).

A data selection procedure developed by TX08b is used here. At each analysis time, 30 observations are chosen based on the correlation between the estimated parameter and the prior estimate (model version) of Z_H, Z_{DR}, K_{DP} observations, when only one of the observed quantities is used for parameter estimation. When more than one observed quantities are used, 15 observations from each data set are chosen based on correlation. For polarimetric variables, data thresholding is found necessary as in Chapter 3. However, we apply lower thresholds to allow for the use of more observations, especially at the upper levels. For Z_{DR} and K_{DP}, the thresholds are 0.05 dB and 0.05 degree km^{-1}, respectively; data values lower than the thresholds are discarded.

4.3 Sensitivity analysis

4.3.1 Response function

Before we perform parameter estimation, we first carry out a set of sensitivity experiments to examine if the model output, in the form of polarimetric variables, is sensitive to the DSD parameters to be estimated. This issue is ultimately related to the identifiability of each parameter with given observations (Yakowitz and Duckstein 1980; Tong and Xue 2008b).

Table 4.1 lists the uncertainty ranges and initial guesses used in our sensitivity and the parameter estimation experiments, respectively; these values were also used in TX08a,b. These choices are based on observed ranges of values although they are not necessarily all encompassing (Joss and Waldvogel 1969; Houze et al. 1979; Mitchell

1988; Gunn and Marshall 1958; Gilmore et al. 2004a; Pruppacher and Klett 1978; Brandes et al. 2007).

The sensitivity analysis procedure follows TX08a. First, EnKF data assimilation cycles are performed using perfect model parameters. The EnKF analyses are performed every 5 min with the first and last analysis being at 25 and 100 min. Forty ensemble members are used and the covariance inflation factor is 15%. Other configurations are as described in section 4.2.1. Five min forecasts are then launched from the ensemble mean analyses with one of the DSD parameters set to an 'incorrect' value sampled within its uncertainty range (Table 4.1). This is done for 16 analysis cycles for several sampled values for the individual parameters. These 5-min forecasts are used to calculate the response function, J, as defined in TX08a:

$$J_y(p^s) = \frac{1}{\sigma_y^2} \sum_{m-1}^{M} \left(y_m(p^s) - y_m^o \right)^2 , \tag{4.1}$$

where p denotes the parameter and superscript s is either w for incorrect value or t for true value. With p^t, the correct parameter value is used in the observation operator. y_m^o denotes the m^{th} observation and $y_m(p)$ is a prior estimate based on the model forecast. Here the observations consist of Z_H, Z_{DR}, and/or K_{DP}. σ_y is the SD of the observation error.

The response functions for each type of observation are averaged over the 16 cycles for each 'incorrect' value of a given DSD parameter. Since we are interested in the change in the model response to the error in the parameter, we compute the response function difference (*RFD*), $RFD^s = \overline{J_y(p^s)} - \overline{J_{y,c}(p^t)}$, where the bar represents the

average over the assimilation cycles. RFD^t is essentially same as ΔJ_y in TX08a where true parameter value is used in the response function calculation. Here $J_{y,c}(p^t)$ is the response function calculated from the forecasts of the control experiment with the truth parameter value.

The difference between RFD^t and RFD^w presents some hints about the amount of error that can be attributed to the error in the observation operator. As in any modern data assimilation system, in the EnKF system, the amount of correction made to the forecast is proportional to the difference between the observations and the forecast projected to the observation space using the observation operator, which is the quantity in the parenthesis in (4.1). Therefore, RFD^t represents the total root-mean-square (RMS) difference between the forecast and observations if the forecast is project into the observations without error while RFD^w represents the total RMS difference the filter would see in the presence of both forecast and observation operator error. When RFD^w is larger than RFD^t, the observation operator error acts to amplify the total error when measured against the particular observation.

Another practical significance of the sensitivity analysis is its ability to rank the relative importance of model parameters so that more important ones can be chosen for estimation. A higher sensitivity implies that the parameter in question has more impact on the forecast than that with a smaller sensitivity (Navon 1997).

4.3.2 Results of sensitivity experiments

From a response function point of view, a necessary condition for a parameter to be identifiable is that it has a unique minimum within its bounds and the response

function has to be sensitive to the parameter (TX08a). To investigate the parameter identifiablity with polarimetric radar data, we plot RFD^t and RFD^w against the deviation of parameter values from their truth in Fig. 4.1.

With respect to (*wrt*) reflectivity observation, both RFD^t and RFD^w curves are concave with their minima located at or near the zero deviation points of individual parameters (Figs 4.1a,b), it is therefore very likely that the truth value can be found by using reflectivity observations when only one of the parameters has error. For Z_{DR}, The *RFD wrt* n_{0S} exhibits very small sensitivity for positive deviations, indicating potential difficulty of estimating n_{0S} in that range. The *RFD*s of n_{0R} and n_{0H} have clear concave shapes with their minima at zero deviation (Figs. 4.1c,d) while the bulk densities, ρ_S and ρ_H, show rather small sensitivity. The *RFD wrt* K_{DP} are even smaller (Figs. 4.1d,f) for all parameters and no unique minimum is apparent for n_{0S} and ρ_S due to the lack of sensitivity *wrt* to positive deviations.

The parameter identification problem is more complex in the presence of observation operator error. When the DSD parameters are involved in the observation operators, incorrect parameter values result in under or over-correction to the parameter that can lead to larger analysis error. In other words, a large difference between RFD^w and RFD^t indicates a large impact of the parameter error through the observation operator. Generally, these differences are moderate for moderate sensitivity and very small when the overall sensitivity is small, but can be very large when the total sensitivity is large (e.g., n_{0H} and ρ_S for Z_H) (Fig. 4.1).

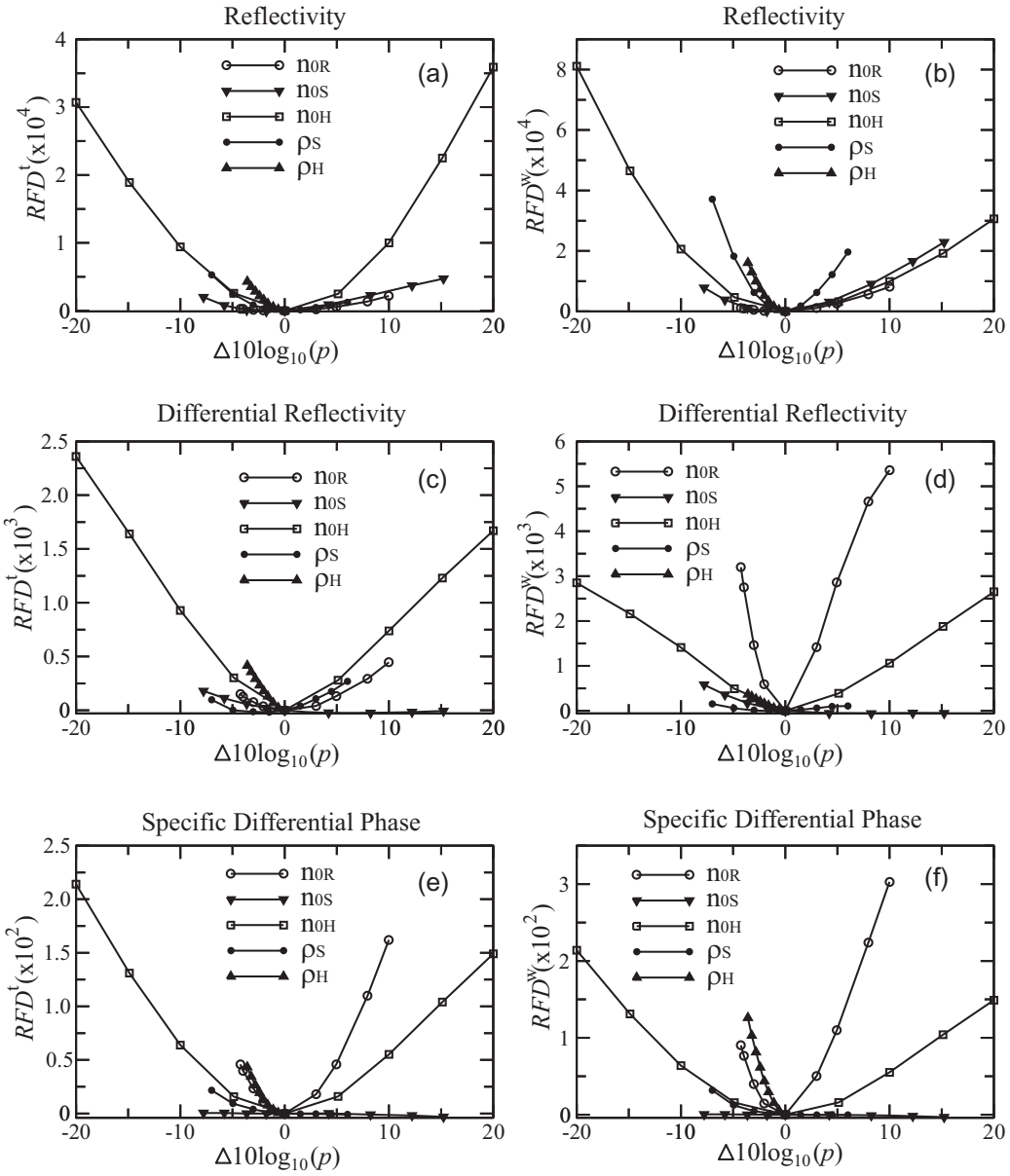


Fig. 4.1. Response function difference (a),(c),(e) $RFD^t = J_y(p^t) - J_{y,c}(p^t)$ calculated with correct parameter values and (b),(d),(f) $RFD^w = J_y(p^w) - J_{y,c}(p^t)$ with incorrect parameter values in the radar observation operators, for (a),(b) reflectivity data; (c),(d) differential reflectivity data; and (e),(f) specific differential phase data. The parameter deviations are in logarithmic space.

The problem becomes even more complicated when multiple observation datasets are used due to complex nonlinear interactions within the filter. For example, the *RFD* for K_{DP} might be too small for successful estimation of n_{0H} while the estimation of n_{0H} using Z_H might also be challenging due to the large difference between RFD^w and RFD^t. However, when K_{DP} and Z_H are used together, the estimation can be successful as we will see in section 4. While the sensitivity results are not sufficient to determine if certain parameters can be estimated successfully, they can still provide useful guidance for interpreting the estimation results.

4.4 Results of parameter estimation

4.4.1 Results of single-parameter estimation

We first investigate the impact of polarimetric data on the estimation of individual DSD parameters. Such estimation assumes that one of the microphysical parameters is the dominant source of error. Because DSD parameters can change over several orders of magnitude, following TX08b, we perform the parameter estimation in the logarithmic space of these parameters. Because we take the average over a number of 'parallel' experiments (see section 4.2.5), we define "normalized absolute error" (NAE) as follows:

$$NAE_i = \frac{1}{N}\sum_{k=1}^{N}\frac{\left|p_{i,k} - p_i^t\right|}{p_i^t}, \tag{4.2}$$

where $p_{i,k}$ is the ensemble mean of the ith parameter in linear space of kth experiment out of a total of N.

Figure 4.2 show the NAEs of estimated parameters from single-parameter experiments. These errors are averaged over 5 parallel experiments that start from three different initial guesses. The experiment names are made up of the parameter name, and the coefficient and exponent of the initial guess of intercept parameter or the first two digits of the bulk density shown in Table 4.1. Observations used in the parameter estimation are indicated after "_". For example, experiment N0r36_ZhKdp estimates n_{0R} from an initial guess of 3×10^6 m^{-4} using both Z_H and K_{DP} data. In most cases, reflectivity data alone can reduce the initial parameter errors (thick solid gray) but the results are not as good as those of TX08b obtained with perfect observation operators. As observed in TX08b, the parameter value can depart far from the truth in the first 1 or 2 cycles (e.g. Figs. 4.2a,b,d,e,g,h,k) and oscillates (around its truth values in log space). The error in the final estimate is larger than the initial error in such experiments as N0h43, N0h45, and Rhos05. Generally, an increase of NAE is observed in the later cycles of the intercept parameters (e.g., Figs. 4.2a-d,e) while the bulk densities converge to their truth values except for Rhos05. These results are quite different from those of TX08b where all parameters eventually converge to their truth values in their single-parameter experiments that use only reflectivity data.

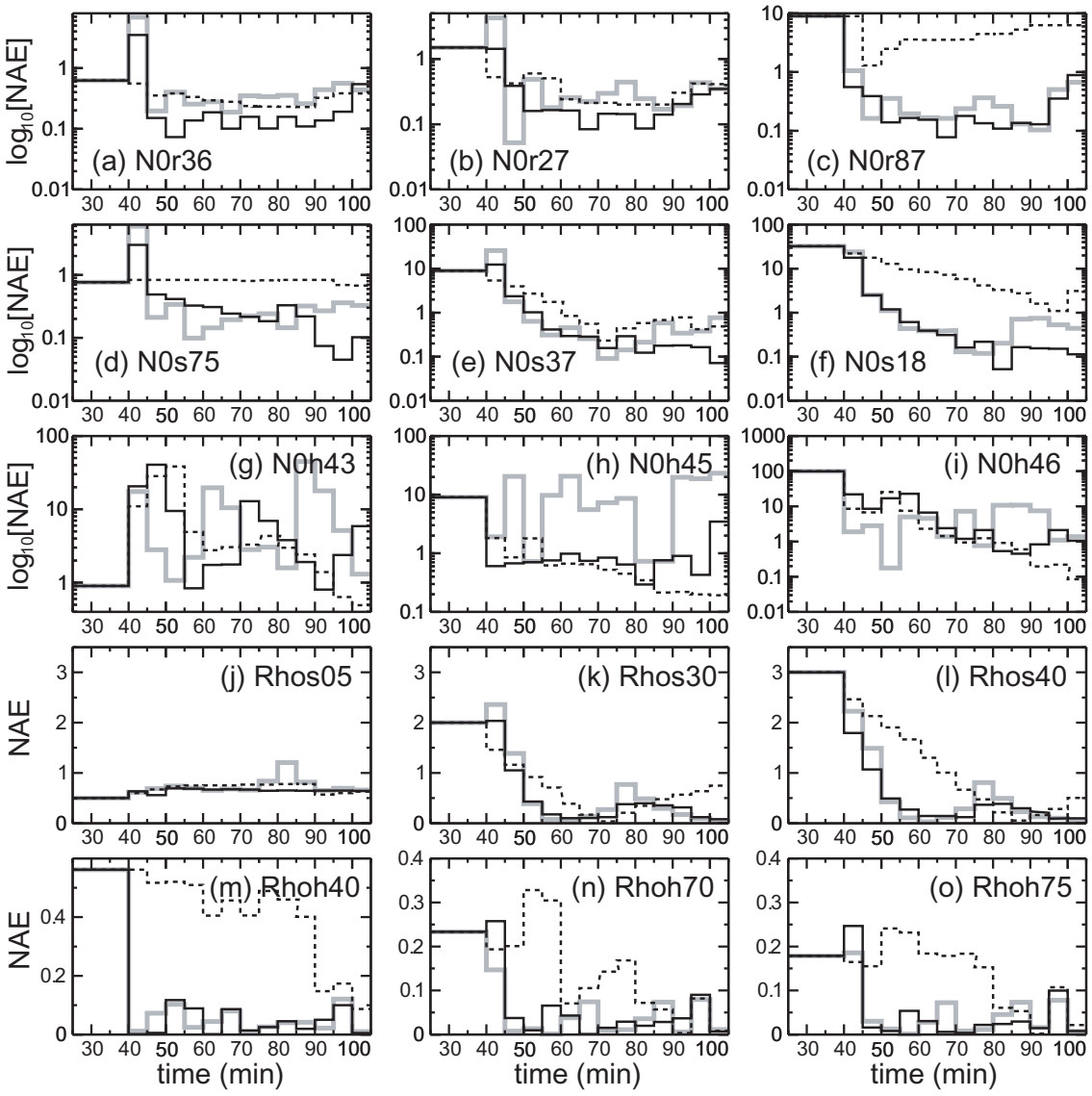


Fig. 4.2. The time evolution of normalized absolute error (NAE) of the ensemble mean of estimated parameter values from single-parameter estimation experiments, for (a)-(c) n_{0R}; (d)-(f) n_{0S}; (g)-(i) n_{0H}; (j)-(l) ρ_S; and (m)-(o) ρ_H, when the parameter estimation is performed using Z_H alone (thick solid gray), both Z_H and K_{DP} (solid black), and Z_{DR} alone (dashed black). The experiment name starts with the parameter name and is followed by the coefficient and the exponent of the initial guess of (a)-(i) the intercept parameter or (j)-(o) the first two digits of the bulk density presented in Table 4.1. The NAEs are averaged over the five parallel experiments that have the same initial guesses but different realizations of the initial parameter ensemble.

Figure 4.2 shows that the estimation of intercept parameters is generally improved when K_{DP} is used in addition to Z_H (solid black curves in Fig. 4.2). For n_{0R}, the NAEs stay lower than those of experiments using Z_H alone (thick solid gray) at most times (Figs. 4.2a-c). Figure 3 shows the ensemble mean analysis RMSEs of state variables from experiments N0r_Zh (thick solid gray), N0r_ZhKdp (solid black), and N0r_Zdr (dashed black). They are averaged over 15 experiments that start from three initial guesses (corresponding to Figs. 4.2a-c), which each initial guesses having 5 parallel experiments with different initial ensemble parameter perturbations. In this case, the benefit of K_{DP} to the estimation of state is rather small because the state obtained with Z_H alone is already rather good. The overall RMSE levels of state are lower than those in Figs. 4.4 and 4.5, which are for experiments estimating n_{0S} and n_{0H}, respectively. In Figs. 4.2d-f, the NAE of n_{0S} experiences a clear reduction in the later cycles when K_{DP} is used in addition to Z_H and the estimated q_s and q_i are improved in response (Fig. 4.4). The positive impact of K_{DP} on the estimation of n_{0H} may not be apparent from Fig. 4.2. However, significant improvement is obtained in the state estimation (Fig. 4.5). It is believed that the smaller variability of NAEs during the assimilation cycles (Figs. 4.2g-i) and the significantly smaller NAEs compared to that of N0h45_Zh (Fig. 4.2h) contribute to the large improvement in the analysis of state. Additional use of K_{DP} in the estimation of bulk densities yields slightly smaller errors in the parameter estimation but exhibits little impact on the state estimation (not shown).

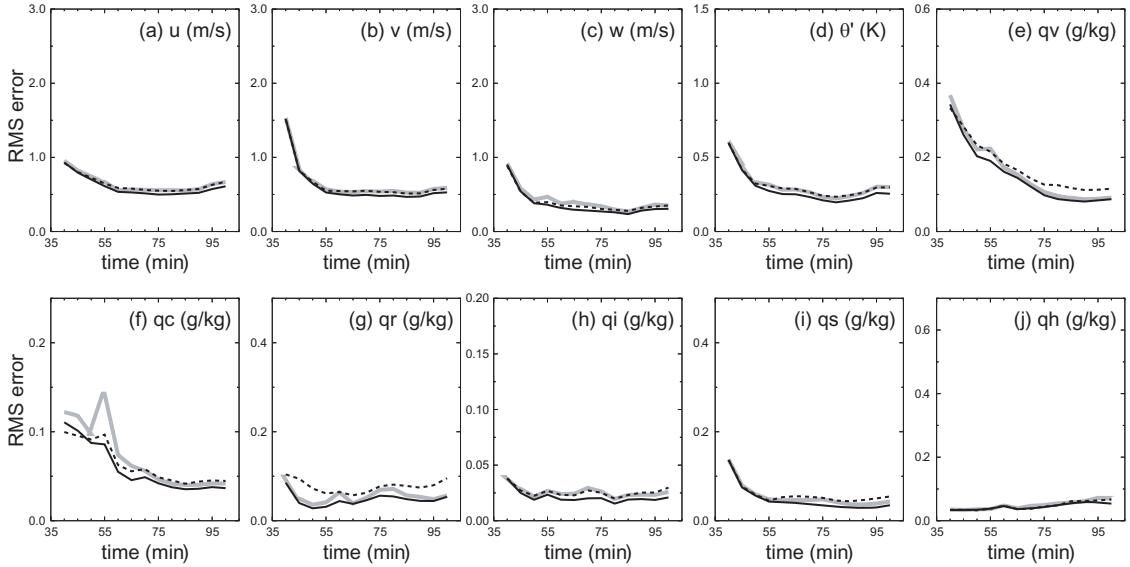


Fig. 4.3. The ensemble-mean analysis RMS errors averaged over points at which the true reflectivity is greater than 10 dBZ for (a) u, (b) v, (c) w and (d) perturbation potential temperature θ', (e) water vapor q_v, (f) cloud water q_c, (g) rainwater q_r, (h) cloud ice q_i, (i) snow q_s, and (j) hail q_h, for experiments where n_{0R} alone contains error and is estimated. The experiments use Z_H data alone (thick solid gray), both Z_H and K_{DP} data (solid black), or Z_{DR} data alone (dashed black). The RMS errors are averaged over 15 experiments that start from three initial guesses presented in Table 4.1 and five initial perturbation realizations for each initial guess.

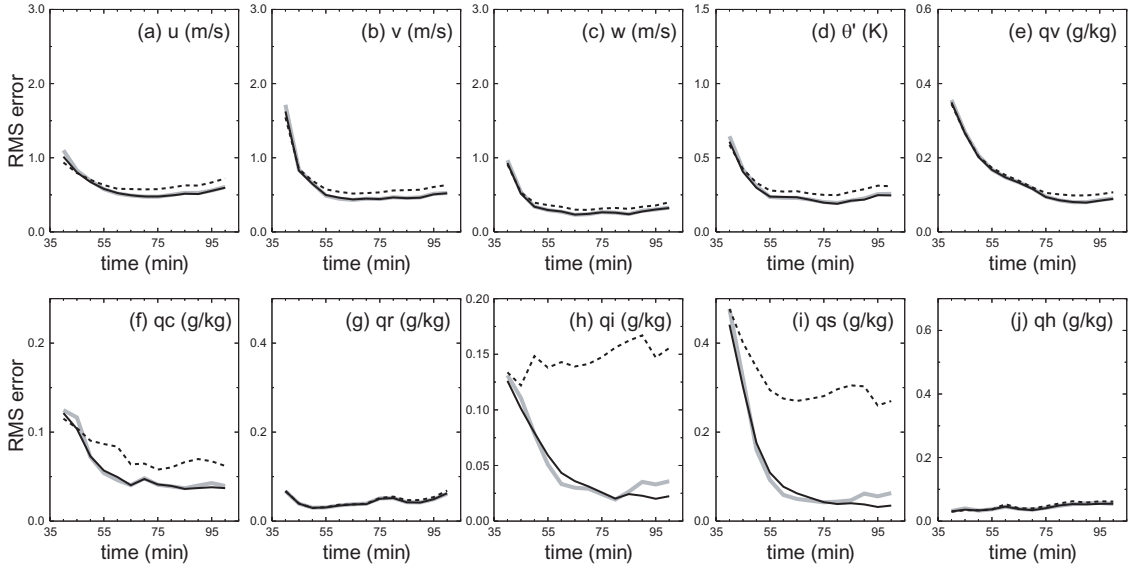


Fig. 4.4. The same as Fig. 4.3 but for n_{0S}.

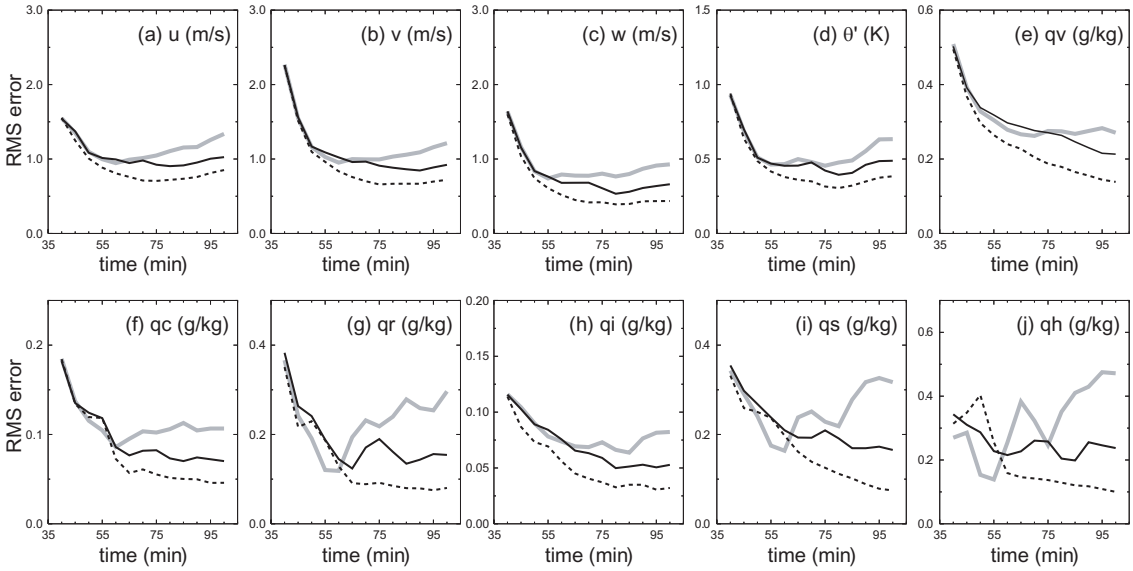


Fig. 4.5. The same as Fig. 4.3 but for n_{0H}.

The largest benefit of polarimetric data is obtained in the estimation of n_{0H} when Z_{DR} is used alone without reflectivity data in the parameter estimation. The NAEs exhibit a steady trend of reduction in general with the exception of the large deviation found in the early assimilation cycles in N0h43_Zdr (black dashed) while the NAEs of N0h_Zh (thick solid gray) show large oscillations with time (Figs. 4.2g,h,i). The estimation of all state variables, including microphysical variables as well as dynamic and thermodynamic variables is significantly improved as the parameter estimation improves (Fig. 4.5). However, the use of Z_{DR} alone in the parameter estimation has a negative impact on both state and parameter estimation for the other 4 parameters (Figs. 4.2, 4.3e,g, and 4.4).

The reason why Z_{DR} outperforms Z_H in the estimation of n_{0H} may be explained by the sensitivity analysis. In section 4.3, it is found that the difference between $RFD(n_{0H}^{w})$ and $RFD(n_{0H}^{t})$ for reflectivity (solid lines with square symbols in Figs.

4.1a,b) is larger while those for Z_{DR} have similar shape and magnitude (solid lines with squares in Figs. 4.1c,d). As discussed in section 4.3, the amount of correction made to the forecast is proportional to the difference between the observations and the forecast projected to the observation space using the observation operator. Therefore, a large (RFD^w - RFD^t) implies that the analysis may deteriorate due to large uncertainty in the observation operators and hence in the observed quantities themselves.

Similar to TX08a, we examine the error correlations to help us understand the filter behavior for the parameter estimation of n_{0H}. This is because the adjustment to parameters is accomplished based on error covariance, the dimensional version of correlation in the filter. Figure 4.6 shows the time series of the correlation coefficient between parameter n_{0H} and the prior estimates of Z_H for one of the five parallel experiments named N0h46_Zh (dotted), and that between n_{0H} and Z_{DR} of the corresponding experiment N0h46_Zdr. The coefficients are averaged over the 30 observations used in the parameter estimation. The correlation coefficient is calculated from the parameter ensemble and the model version (prior estimate) of the observations (Z_H or Z_{DR}) from the forecast ensemble. The correlation coefficient in experiment N0h46_Zdr keeps increasing during early cycles and stays high during the rest of the cycles. On the other hand, the correlation coefficient in N0h46_Zh drops rapidly in the first two cycles. It bounces back in the next two cycles but oscillates during the remaining cycles and stays lower than that of N0h46_Zdr. Since nonlinear feedback exists between parameter and state estimations during the assimilation cycles, large error in parameter estimation due to weak correlation leads to poor state estimation and slow convergence or even parameter estimation failure.

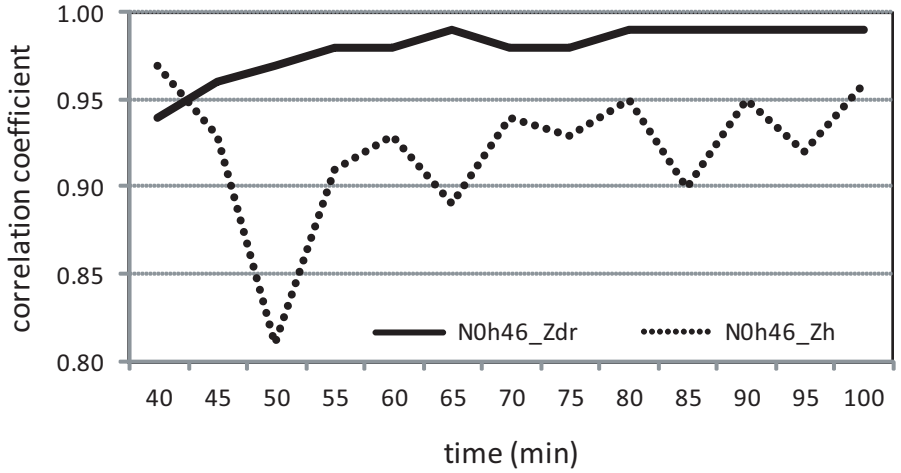


Fig. 4.6. The time evolution of the correlation coefficients between parameter n_{0H} and the model prior estimates of Z_{DR} from one of the five parallel experiments named N0h46_Zdr (solid) and those between n_{0H} and Z_H from corresponding experiment N0h46_Zh (dotted), averaged over the 30 observations used in parameter estimation.

The spatial distribution of observations used for parameter estimation appears to also affect the estimation. From Figs. 4.7 and 4.8, Z_H observations used in N0h46_Zh are clustered at a few locations (Figs. 4.7a,b and 4.8a,b) while Z_{DR} observations used in N0h46_Zdr are scattered in a wider area (Figs. 4.7c,d and 4.8c,d). Observations from the same spatial regions of a storm are likely to carry similar information on the storm. Repeated application of observations with similar information content tends to accelerate the reduction of parameter spread. The covariance inflation procedure used to prevent the collapse of spread can lead to oscillations and over-adjustments (TX08b). In N0r46_Zh, the parameter spread falls to the predefined minimum SD after 2 cycles while it takes 7 cycles in N0r46_Zdr. We also notice that many of Z_H observations are taken from the region where at least three phases (rain, hail, and melting hail) contribute to Z_H. At 45 min, many of the Z_H data chosen are below 4 km, which is about the 0 °C level (Fig. 4.7b). At 90 min they are mostly near the extended hail core region, possibly

near strong updraft (Fig. 4.8b). On the contrary, many of the Z_{DR} observations are taken from the region where dry hail dominates over snow (Figs. 4.7d and 4.8d).

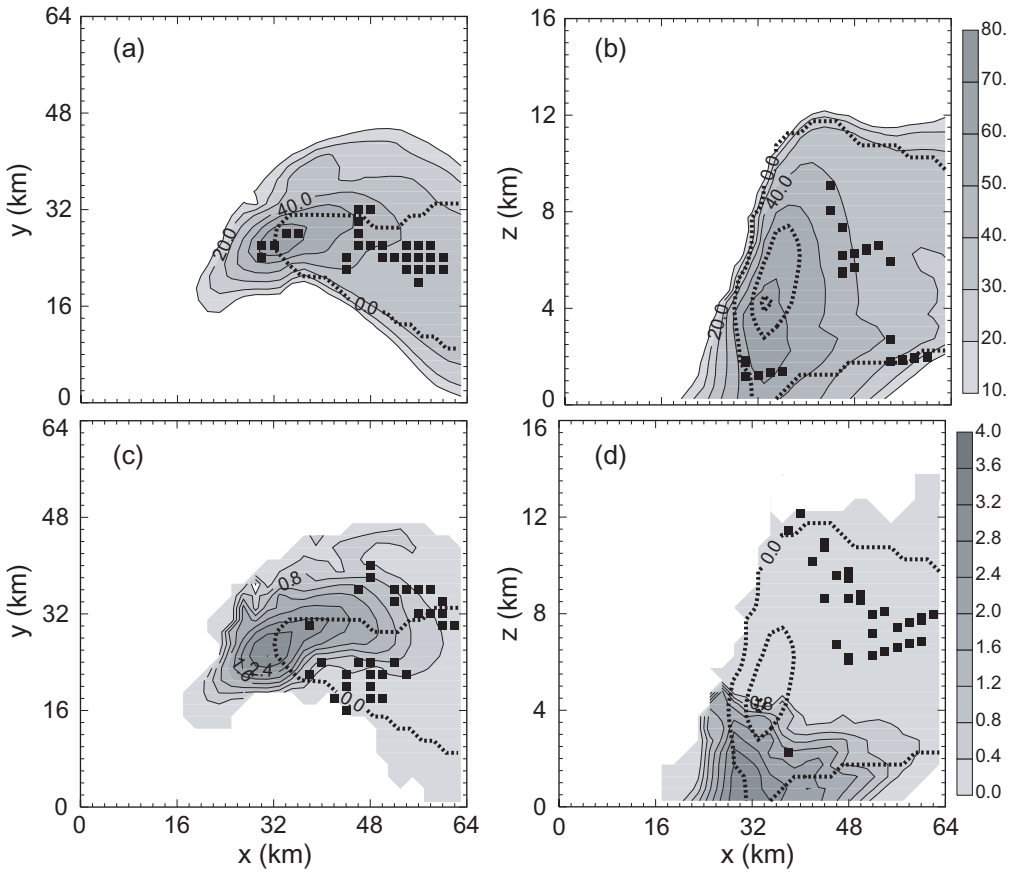


Fig. 4.7. Vertical column maximum Z_H and Z_{DR} shown in the horizontal plane (a and c, respectively) and column maximum Z_H and Z_{DR} in the y direction shown in the vertical x-z plane (b and d, respectively), of truth simulation at 45 minutes (shading and thin solid contours). Solid squares indicate the locations of the observations that were used in the single-parameter estimation experiments N0h46_Zh (Z_H observations in a and b) and N0h46_Zdr (Z_{DR} observations in c and d). The data points are projected to the horizontal x-y plane in the left panels and to the vertical x-z plane in the right panels. Thick dotted contours at intervals of 2 g kg^{-1} represent the hail mixing ratio through vertical velocity maximum.

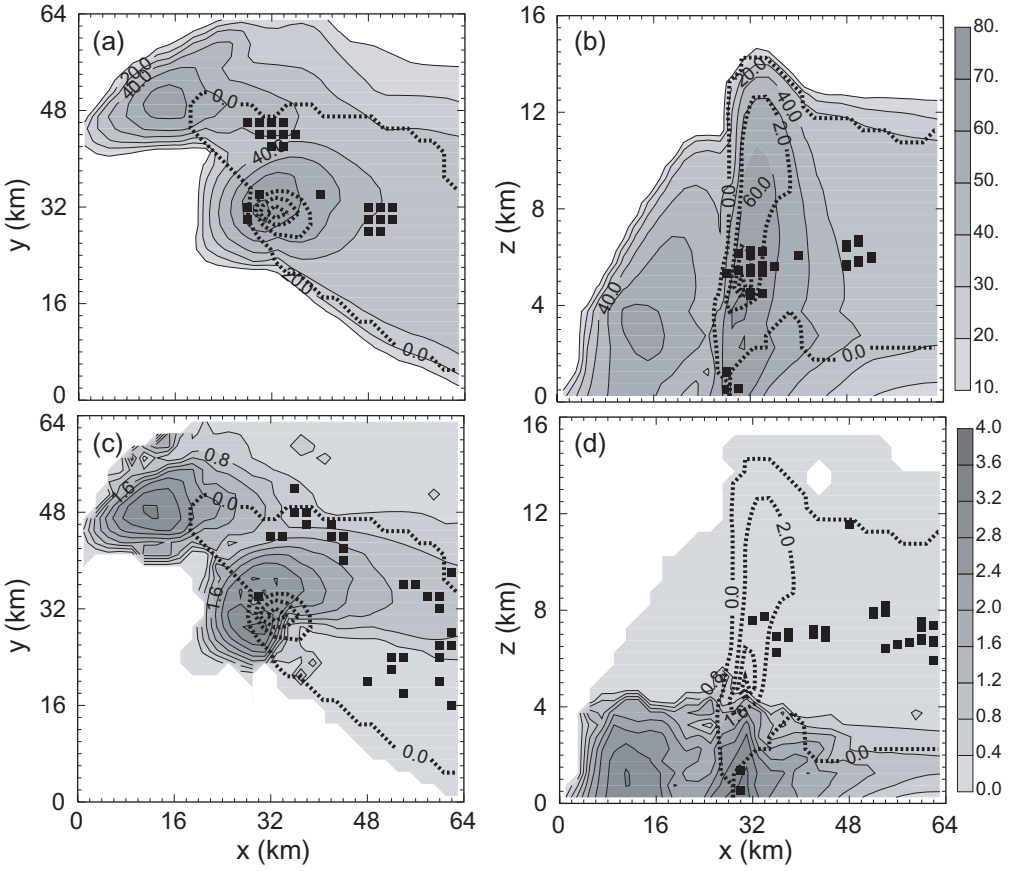


Fig. 4.8. As in Fig. 4.7 but for 90 min.

The mean estimated parameters in logarithmic form from the single-parameter estimation experiments are presented in Table 4.2, together with the true values given in parentheses. The mean values are computed from the 15 experiments with three different initial guesses for each parameter (see Table 4.1) and are averaged over the last 5 cycles. All five parameter estimates are more accurate when both Z_H and K_{DP} are used in the parameter estimation than when only Z_H is used. In the case of n_{0H}, the best estimate is obtained using Z_{DR} data alone. The mean parameter values in a logarithmic form, averaged over 5 runs are 51.2, 46.7, 48.0 for N0h43_Zdr, N0h45_Zdr, N0h46_Zdr, respectively; they are 57.5, 55.8 and 53.8 for N0h43_Zh, N0h45_Zh,

N0h46_Zh, respectively, while the truth is 46. The n_{0H} averaged over runs with different initial guesses is 56.0 for N0h_Zh and 49.1 for N0h_Zdr (Table 4.2). After being converted to the linear domain, these values correspond to a factor of 6 difference; 56.0 is about five times larger than 49.1 in terms of their linear values. We point out that N0h_Zdr produces a more stable estimate of n_{0H} than N0h_Kdp because in the former the estimated parameter shows smaller spread among the experiments with different realizations (not shown) and has almost no oscillation during the assimilation cycles (see Fig. 4.2) even though the averaged values in Table 4.2 appear to be similar. As a result, the state estimation of N0h_Zdr exhibits significant improvement over that of N0h_Kdp.

Table 4.2. The mean estimated parameter values in logarithmical form for single-parameter estimation experiments, averaged over 15 experiments with three different initial guesses presented in Table 4.1 and five different perturbation realizations for each initial guess, over the last 5 cycles (80 minutes to 100 minutes of model time). Their truth values in logarithmical form are given inside parentheses.

Observations used in parameter estimation	n_{0R} (69.0)	n_{0S} (64.8)	n_{0H} (46.0)	ρ_S (20.0)	ρ_H (29.6)
Z_H	68.6	65.0	56.0	20.3	29.4
Z_{DR}	73.2	66.2	49.1	18.0	28.8
Z_H and K_{DP}	69.1	64.8	49.2	19.9	29.4

The best results for certain parameters or state variables are obtained with somewhat different combinations of polarimetric measurements. Based on our results,

the combined use of both Z_H and K_{DP} appears a good choice when estimating one of n_{0R}, n_{0S}, ρ_S, and ρ_H, while the use of Z_{DR} is recommended for the estimation of n_{0H}.

4.4.2. Results of five-parameter estimation

In this subsection, we examine the filter performance when five parameters are estimated simultaneously. Again, errors and estimated DSD parameters are averaged over 160 experiments, as described in section 4.2.5 (Table 4.1).

Figure 4.9 shows the NAEs of the ensemble mean estimated parameters from the five-parameter estimation experiments. Five-parameter estimation experiments reveal difficulties in estimating all five parameters simultaneously in the presence of observation operator error. The initial error level is overlaid for easier comparison (dashed gray). When Z_H is used alone (thick solid gray), the NAEs of n_{0R}, n_{0S}, and ρ_S experience rapid error growth in the first 1 to 2 cycles (Figs. 4.9a,b,d, respectively). These NAEs decrease significantly in the next several cycles but increase again in later cycles. The errors of n_{0S} and ρ_S remain above the initial error level during all assimilation cycles except for a temporary drop at 85 minutes for ρ_S. This result is quite different from that of TX08b that used perfect observation operators. In their study, Z_H alone was able to reduce the errors in all five parameters below their initial errors most times.

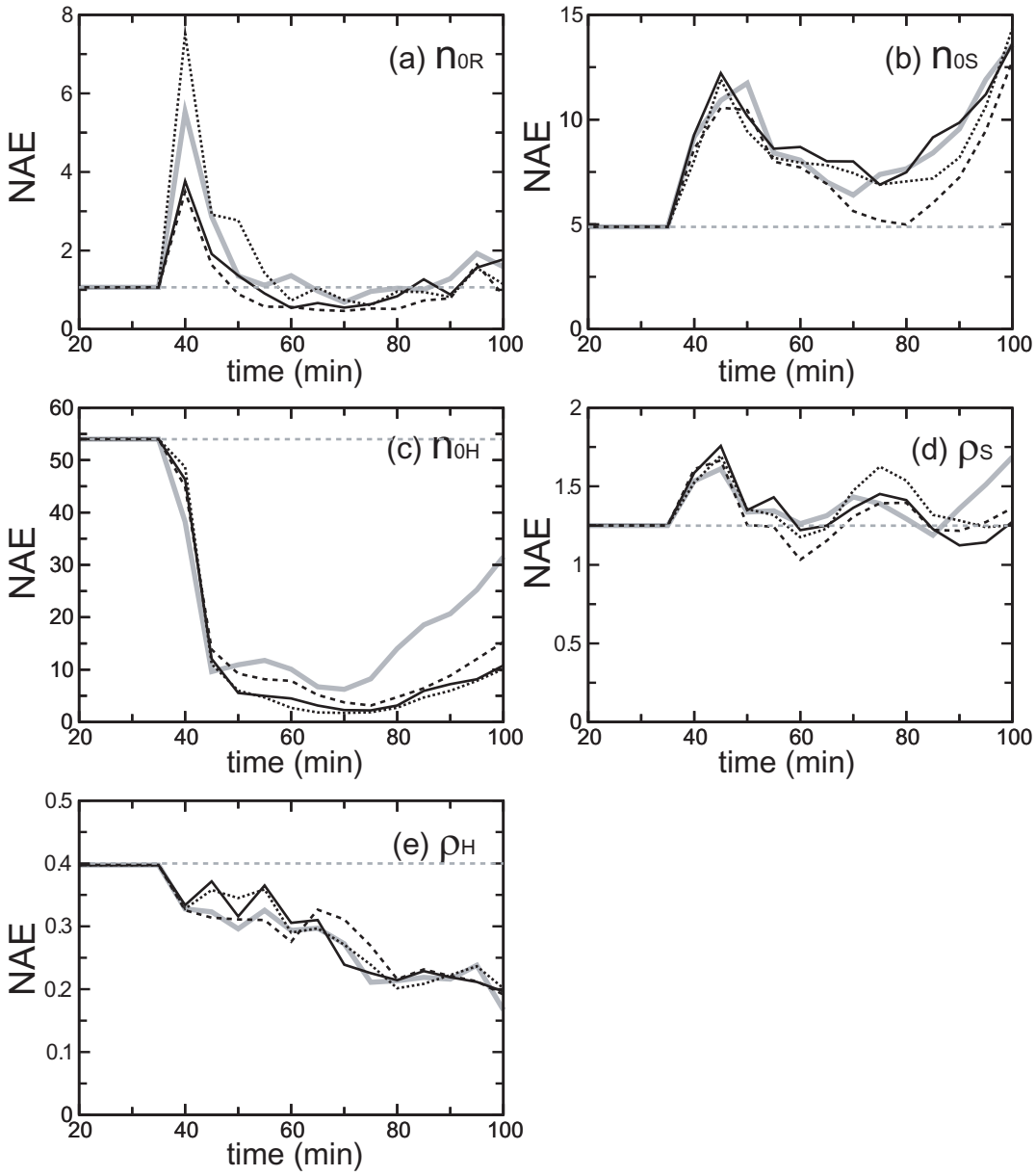


Fig. 4.9. The time evolution of the NAE of the mean parameter values from five-parameter estimation experiments for (a) n_{0R}, (b) n_{0S}, (c) n_{0H}, (d) ρ_S, and (e) ρ_H for the experiments para5_Zh (thick solid gray), para5_ZhZdr (dotted black), para5_ZhKdp (dashed black), and para5_ZhZdrKdp (solid black). Initial error level is shown in dashed gray. The average NAE is calculated from the 160 experiments with 32 different initial guesses consisting of the combinations of 5 parameters with 2 initial guesses each, as shown in Table 4.1, where 5 parallel runs with different realizations of initial parameter perturbations are carried out for each initial guess.

A positive impact of polarimetric data is observed in the estimation of n_{0R}, n_{0H}, and ρ_S during the later assimilation cycles no matter which additional polarimetric parameter is used (Figs. 4.9a,c,d). When either Z_{DR} (dotted black) or K_{DP} (dashed black) is used or when both Z_{DR} and K_{DP} are used (solid black) in addition to Z_H in the estimation of n_{0H}, and ρ_S, the error grows much slower after 80 min; the error, however, grows rapidly when Z_H is used alone. The most significant positive impact of polarimetric data is found with the estimation of n_{0H}, whose error level is significantly lower in all cases that use polarimetric data (Fig. 4.9c).

As in the single-parameter estimation experiments, K_{DP} is slightly more beneficial than Z_{DR} in general but Z_{DR} produces a better estimation of n_{0H} than does K_{DP}. Smaller errors in the estimated parameters during the assimilation cycles help improve state estimation while smaller errors at the end should improve the subsequent forecast.

For the estimated state, the best results are obtained when both Z_{DR} and K_{DP} are used for parameter estimation (solid lines in Fig. 4.10). The RMSEs of experiments para5_ZhZdr and para5_ZhKdp (not shown) are slightly larger than those of para5_ZhZdrKdp but smaller than those of para5_Zh (dashed lines) with the exception of q_s because of the poor performance of para5_ZhZdrKdp in the estimation of n_{0S}. A tendency of error increase is found in most state variables in para5_Zh during later assimilation cycles in response to the error increase in n_{0R}, n_{0S}, n_{0H}, and ρ_H; this error increase is much weaker and the errors stay lower in para5_ZhZdrKdp in all state variables except for q_s.

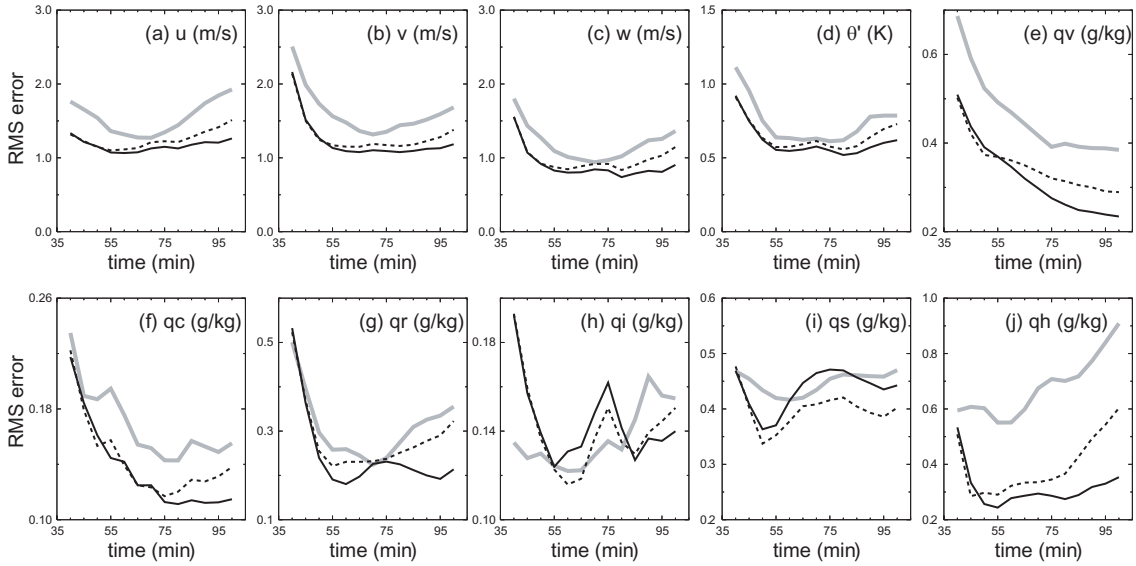


Fig. 4.10. As Fig. 4.3 but for simultaneous estimation of five parameters for experiments para5_Zh (dashed black) and para5_ZhZdrKdp (solid black). The RMS errors of the no-parameter-estimation experiments with the initial guesses of parameters kept throughout the assimilation cycles are shown in thick solid gray for comparison. The RMS errors are averaged over 160 experiments as in Fig. 4.9.

Even though the observation operator error adds extra complication to parameter estimation, the positive impact of parameter estimation on state estimation is clear, even with the failure in estimating n_{0S} and ρ_S. This is seen by comparing the state variable errors with those of no-estimation experiments (thick solid gray in Fig. 4.10) where the initial 'incorrect' parameter values are kept throughout the assimilation cycles. In the latter experiments, the state variable errors increase significantly after 65 minutes of model time, presumably because the parameter errors now dominate.

Since we are interested in how much the polarimetric data can further improve not only the parameter but also the state estimation, we calculate the percentage improvement in state variables according to the following,

$$Improvement(\%) = \left\{ \sum_{i=1}^{N} \left[(\varepsilon_c)_i - (\varepsilon)_i \right] \Big/ \sum_{i=1}^{N} (\varepsilon_c)_i \right\} \times 100 \ , \tag{4.3}$$

where ε_c is the RMSE of the control experiment for a particular variable and N is the number of experiments averaged over. This improvement is further averaged over the last five assimilation cycles.

The improvement amounts of para5_ZhZdr, para5_ZhKdp, para5_ZhZdrKdp over para5_Zh are summarized in Table 4.3. We can see from the table that the improvement is larger in w, q_r, q_v, and q_h and smaller (actually negative) in q_s. This is in general consistent with the finding of Chapter 3. The improvement due to polarimetric data is greatest (between 28 and 35 %) in q_h here, while it was greatest in q_r in Chapter 3. No negative impact was found in any state variable in Chapter 3. The poor performance in estimating q_s is understandable, since polarimetric signatures related to the low-density dry snow are generally very weak.

The spatial distribution of observations used in one of the five–parameter estimation experiments is shown in Fig. 4.11 as an example. As in the single-parameter estimation experiments, the Z_H observations used to estimate n_{0H} in para5_Zh are concentrated in two general areas in the precipitation region (black dots in Figs. 4.11a,c,e) while the Z_H (black dots), Z_{DR} (triangles), and K_{DP} (squares) data in para5_ZhZdrKdp (Figs. 4.11b,d,f) are selected from a broader region. Interestingly, the Z_H data are mostly selected from the lower levels, Z_{DR} mostly from the upper levels, and K_{DP} mostly from the middle levels.

Table 4.3. The percentage improvement in state estimation for experiments para5_ZhZdr, para5_ZhKdp and para5_ZhZdrKdp over experiment para5_Zh, averaged over 160 experiments with 32 different initial guesses with 5 parallel runs for each initial guess and over the last 5 cycles (80 minutes to 100 minutes of model time). Prefix 'para5_' is omitted in the experiment names.

Variables	Improvement (%)		
	ZhZdr	ZhKdp	ZhZdrKdp
u	8.9	9.7	12.1
v	7.6	9.6	10.5
w	12.7	12.2	16.9
θ'	8.2	6.6	11.5
p'	7.0	2.9	7.7
q_v	15.2	18.5	19.4
q_c	9.3	12.0	13.3
q_r	21.8	22.3	25.9
q_i	-0.1	2.4	3.7
q_s	-11.0	-2.9	-11.6
q_h	33.5	28.5	34.9
average	10.3	11.1	13.1

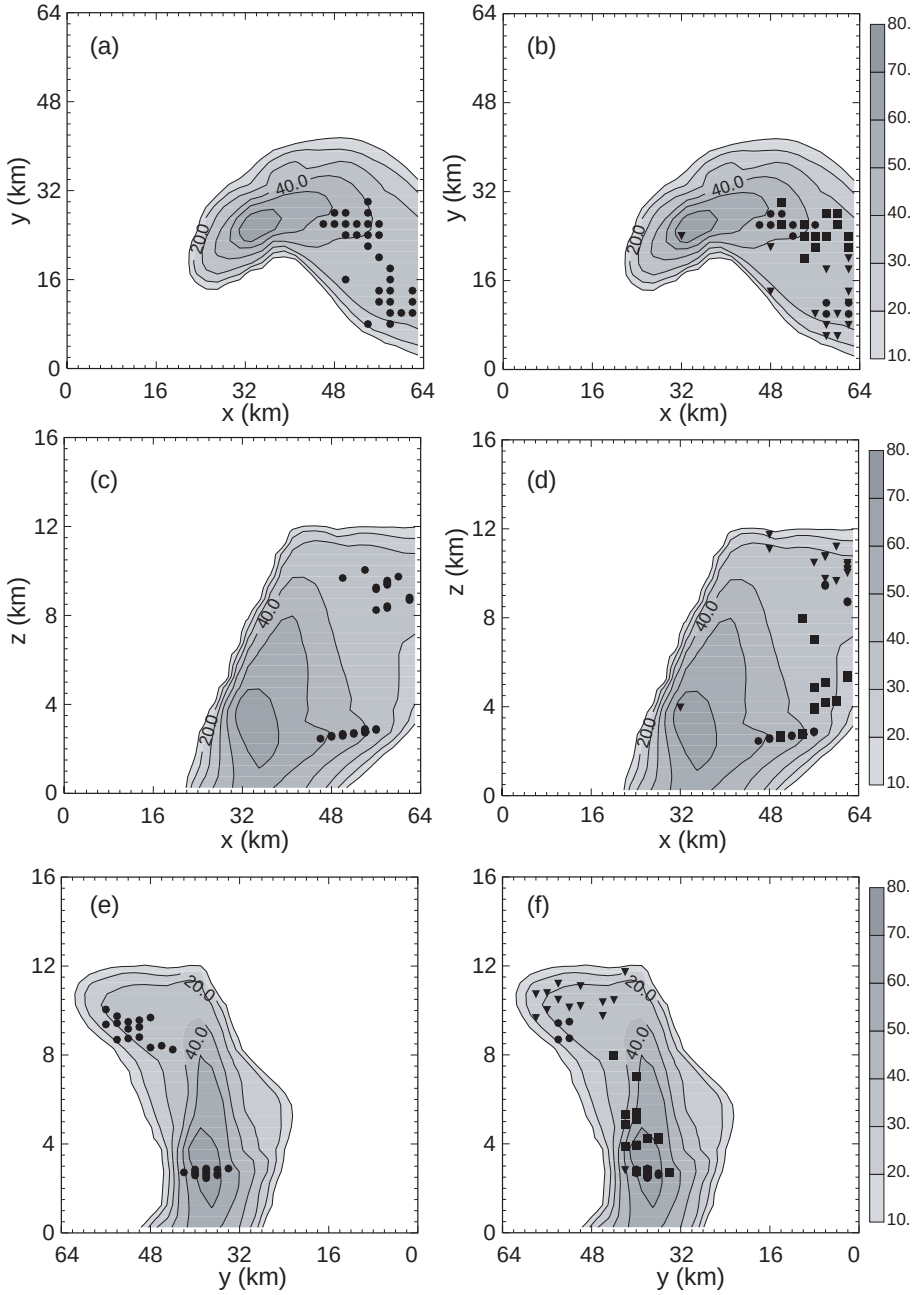


Fig. 4.11. Vertical column maximum Z_H in a *x*-*y* plane (a and b), column maximum Z_H in *y* direction in an *x*-*z* plane (c and d) and in x direction in a y-z plane (e and f), from the truth simulation at 40 min. Black dots in the left panels indicate the locations of Z_H observations used in 5-parameter estimation experiment para5_Zh, and the black dots, triangles and squares in the right panels represent the Z_H, Z_{DR}, and K_{DP} observations, respectively, that were used to estimate n_{0H} in experiment para5_ZhZdrKdp. Initial parameter values for this experiment are (n_{0R}, n_{0S}, n_{0H}, ρ_S, ρ_H) = ($3\text{x}10^6$ m^{-4}, $7\text{x}10^5$ m^{-4}, $4\text{x}10^5$ m^{-4}, 50 kg m^{-3}, 400 kg m^{-3}).

4.4.3. Results of three-parameter estimation

Figure 4.9 shows that the errors in estimated n_{0S} and ρ_S are almost always larger than their initial errors. This suggests that it may be better not to estimate n_{0S} and ρ_S, but to keep their initial values. To test this hypothesis, we perform ten additional experiments starting from incorrect values in all five parameters but estimating only three of them or n_{0R}, n_{0H}, and ρ_H . Two sets of initial guesses are used, they are (n_{0R}, n_{0S}, n_{0H}, ρ_S, ρ_H) = ($3\text{x}10^6$ m^{-4}, $7\text{x}10^5$ m^{-4}, $4\text{x}10^5$ m^{-4}, 50 kg m^{-3}, 400 kg m^{-3}) and ($3\text{x}10^6$ m^{-4}, $3\text{x}10^7$ m^{-4}, $4\text{x}10^5$ m^{-4}, 300 kg m^{-3}, 400 kg m^{-3}). The estimated mean parameter values and spreads computed from ten such experiments are shown in Fig. 4.12. In experiments para3_ZhZdr (dotted black), para3_ZhKdp (solid black), para3_ZhZdrKdp (dashed black), with the help of polarimetric variables, the mean n_{0H} and ρ_H converge nicely to their truth values and exhibit a clear tendency of rapid decrease in spread during middle to later cycles. Meanwhile, the parameters in para3_Zh (thick solid gray) show large oscillations and stay away from the truth, and the spreads remain high. The n_{0R} estimation is most successful with additional K_{DP} data. The mean estimated parameter values averaged over the 10 experiments and over the last 5 cycles are more accurate than those of para5_Zh when polarimetric variables are used except for n_{0R} in para3_ZhZdr and para3_ZhZdrKdp (Table 4.4). Compared to experiment para5_Zh, the largest improvement by not estimating n_{0S} and ρ_S is achieved in n_{0H}. The positive impact of polarimetric data is also greatest in n_{0H} estimation. For example, the estimated n_{0H} in para5_Zh contains about 2,200% error in linear space while the estimate in para3_ZhKdp contains only about 17 % error; for reference, the average initial error is about 5,000 % of the assumed truth in linear space.

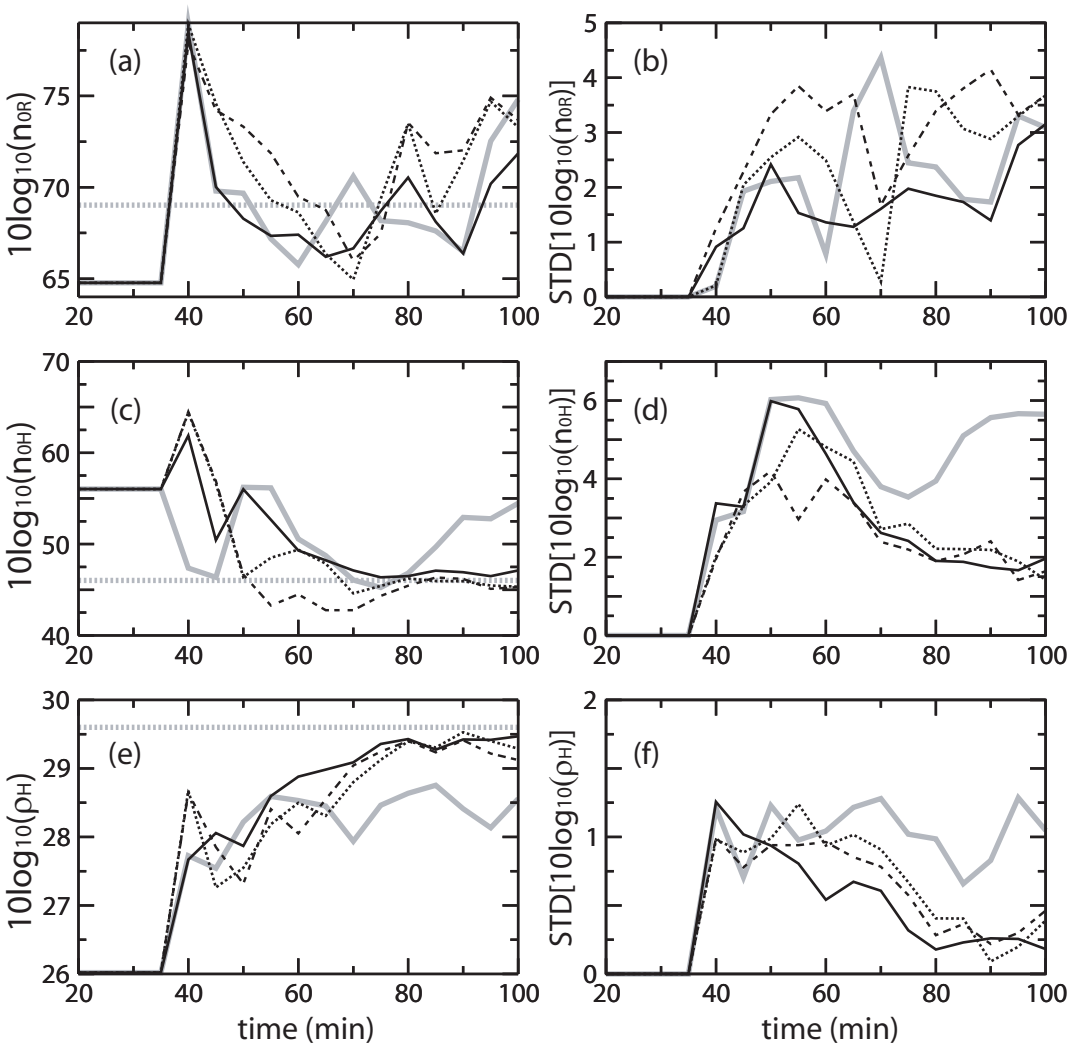


Fig. 4.12. The mean estimated parameter (left penels) and spread (right panels) for (a),(b) n_{0H}; (c),(d) n_{0H}; and (e),(f) ρ_H for 3-parameter estimation experiments para3_Zh (thick solid gray), para3_ZhZdr (dotted black), para3_ZhKdp (solid black), and para3_ZhZdrKdp (dashed black). The horizontal thick dotted gray lines in the left panels indicate the truth parameter values. In these experiments, n_{0R}, n_{0H} and ρ_H were estimated while n_{0S} and ρ_S were kept at their incorrect initial values throughout the assimilation cycles. The mean and spread are computed from 10 experiments starting from two sets of imperfect parameter values (n_{0R}, n_{0S}, n_{0H}, ρ_S, ρ_H) = (3×10^6 m^{-4}, 7×10^5 m^{-4}, 4×10^5 m^{-4}, 50 kg m^{-3}, 400 kg m^{-3}) and (3×10^6 m^{-4}, 3×10^7 m^{-4}, 4×10^5 m^{-4}, 300 kg m^{-3}, 400 kg m^{-3}).

Table 4.4. Same as in Table 4.2 but for five-parameter experiment para5_Zh and three-parameter estimation experiments para3_Zh, para3_ZhZdr, para3_ZhKdp and para3_ZhZdrKdp, in which n_{0S} and ρ_S were kept as their incorrect initial value throughout the assimilation cycles while other three parameters were estimated. The experiments start from two sets of parameter values, namely, (n_{0R}, n_{0S}, n_{0H}, ρ_S, ρ_H) = ($3x10^6$ m^{-4}, $7x10^5$ m^{-4}, $4x10^5$ m^{-4}, 50 kg m^{-3}, 400 kg m^{-3}) and ($3x10^6$ m^{-4}, $3x10^7$ m^{-4}, $4x10^5$ m^{-4}, 300 kg m^{-3}, 400 kg m^{-3}).

Experiment	n_{0R} (69.0)	n_{0S} (64.8)	n_{0H} (46.0)	ρ_S (20.0)	ρ_H (29.6)
para5_Zh	71.4	74.4	59.7	23.2	28.4
para3_Zh	69.1	71.9	50.5	22.4	28.5
para3_ZhZdr	72.1	71.9	45.8	22.4	29.4
para3_ZhKdp	69.0	71.9	46.7	22.4	29.4
para3_ZhZdrKdp	72.6	71.9	45.5	22.4	29.3

The state estimation is also improved when the parameter estimation is improved by not estimating the snow-related parameters (Fig. 4.13). The RMSEs of para3_Zh (black dashed) are generally smaller than those of para5_Zh (thick solid gray), except for q_i, and the RMSE differences increase with time. The percentage improvement over para5_Zh in para3_Zh averaged over 11 model state variables is 23.4 %, with a largest improvement of 42% found in q_h . w, q_r, and q_s experience about 30% improvement.

The RMSEs are further reduced significantly by polarimetric data in the parameter estimation (Fig. 4.13). The q_s estimation is no longer hampered by the additional K_{DP} data (solid black) but rather experiences large RMSE reduction compared to Fig. 4.10i. When Z_H is used alone (black dashed), after large reduction during the first 20 minutes of assimilation cycles (not shown in the plots), the RMSEs

start increasing between 40 and 70 min mostly because of the poor estimation of n_{0R} during early cycles and the poor estimation of n_{0H} between 45 to 60 min (Figs. 4.12a,c). Because the accuracy of the estimated state as well as estimated parameters depends on the history of estimation, large errors in the early assimilation cycles, regardless of their source, impact the state and parameter estimation process. On the contrary, continuous error reductions throughout the assimilation cycles are seen in all state variables in para3_ZhKdp, except for q_i (Fig. 4.13).

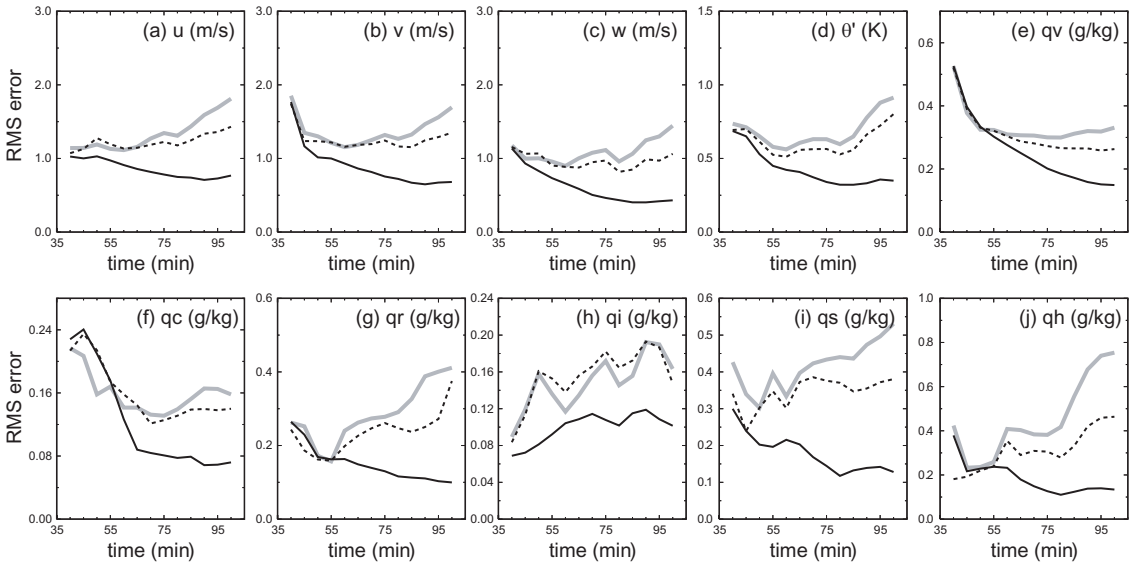


Fig. 4.13. As Fig. 4.3 but for experiments para3_Zh (black dashed) and para3_ZhKdp (solid black). The RMS errors are averaged over 10 experiments starting from two sets of initial parameter values as given in the caption of Fig. 4.12. The RMS errors of experiment para5_Zh are shown in thick solid gray for comparison.

In the early cycles between 40 and 45 min, experiment para3_Zh produces comparable estimate of n_{0R} but better estimate of n_{0H} than para3_ZhKdp (Figs. 4.12a,c). However, the state estimation of para3_Zh is generally poorer than that of para3_ZhKdp. This seemingly contradictory result can be explained by the

compensating model responses described in TX08b. The increase in n_{0R} compensates the decrease in n_{0H} in terms of reflectivity. When the problem is insufficiently constrained by data, multiple solutions can exist. The microphysical information provided by additional polarimetric data appears to help alleviate the non-uniqueness problem.

The gross improvement produced by the polarimetric data in the three-parameter estimation experiment with five incorrect parameter values can be assessed more easily from Table 4.5. Statistically, the overall errors in the analysis are approximately cut in half. All state variables exhibit fairly large improvements ranging from 29.9 to 66.4 %. The best analysis is obtained by using K_{DP} data in addition to Z_H, which is consistent with the parameter estimation results shown in Fig. 4.12. This appears reasonable because K_{DP} data seem to provide different information content than Z_H since they are selected mostly from discrete regions of the storm while many of Z_{DR} data seem to overlap Z_H in location (Fig. 4.11). Another interesting point is that when not estimating snow related parameters, q_s experiences the second largest improvement in para3_ZhZdr and para3_ZhZdrKdp and the third largest improvement in para3_ZhKdp.

Last, in the three- and five-parameter estimation experiments, when the polarimetric data are used alone, individually or together, without Z_H, the estimated states are generally not as good as those using Z_H alone. These results are not presented here.

Table 4.5. The percentage improvement of state estimation for three-parameter estimation experiments para3_ZhZdr, para3_ZhKdp, and para3_ZhZdrKdp over experiment para3_Zh. The percentage improvements are ensemble mean of those computed from ten experiments listed in Table 4.4. Prefix 'para3_' is omitted in the experiment names.

Variables	Percentage improvement (%)		
	ZhZdr	ZhKdp	ZhZdrKdp
u	42.9	46.9	42.5
v	42.6	49.5	42.2
w	49.0	56.1	49.5
θ'	45.7	50.0	45.6
p'	36.4	43.3	36.7
q_v	40.2	46.5	36.1
q_c	46.4	52.0	44.5
q_r	57.6	63.0	54.5
q_i	29.9	45.1	29.9
q_s	60.0	61.6	55.9
q_h	64.3	66.4	57.8
average	46.8	52.8	45.0

4.5. Summary and conclusions

We investigated the impact of additional polarimetric data on correcting errors in DSD-related fundamental parameters in the model microphysics scheme. Such errors also affect the observation operators of all radar observations except radial velocity (in our case at least where reflectivity weighting for radial velocity is ignored). These parameters, namely, the intercept parameters of rain n_{0R}, snow n_{0S}, and hail $n_{0\mathrm{H}}$, and the

bulk densities of snow ρ_S and hail ρ_H, are estimated, individually or all together, simultaneously with the model state using a sequential ensemble square-root Kalman filter. The polarimetric data considered include differential reflectivity Z_{DR} and specific differential phase K_{DP}. To obtain more robust results, single-, five-, and three-parameter estimations are repeated with different initial guesses and different initial ensemble perturbations for each parameter, and the mean and standard deviation statistics are computed and compared. Compared to the earlier parameter estimation work of Tong and Xue (2008b), this study includes the effect of observation operator error, and examines the impact of additional polarimetric data. In Chapter 3, the impact of polarimetric data is examined in the absence of any parameter error. Based on the author's knowledge, no previous parameter estimation study has addressed the issue of parameter error within the observation operators.

Generally, the reflectivity, Z_H, observations alone can effectively reduce the error in n_{0R}, n_{0S}, ρ_S and ρ_H when only one parameter contains error, even in the presence of observation operator error, they, however, perform poorly when estimating n_{0H}. The K_{DP}, in addition to Z_H, is found to help further reduce the errors in the intercept parameters and improve the state estimation through improved parameter estimation. Adding K_{DP} has almost no impact on the estimation of snow and hail densities and their related state variables because the estimation with reflectivity alone is already very successful. The best estimation of n_{0H} is obtained when Z_{DR} is used alone (for parameter estimation) while its estimation using K_{DP} and Z_H is also reasonably good.

Our results reveal some difficulties in simultaneously estimating all five parameters that contain error. Unlike TX08b that assumes perfect observation operators,

our five-parameter estimation experiments show that the errors in n_{0S} and ρ_S are increased during the assimilation cycles by the parameter estimation to above their initial levels with or without using polarimetric data (for parameter estimation). However, the positive impacts of polarimetric data on the state estimation are clear when Z_{DR} or K_{DP} or both Z_{DR} and K_{DP} are used along with Z_H in the parameter estimation. When all five parameters contain initial error, both parameter and state estimations are improved when n_{0R}, n_{0H}, and ρ_S are estimated without n_{0S} and ρ_S. Moreover, the positive impact of polarimetric data is further increased compared to the case when all five parameters are estimated. This behavior can be understood from the fact that the polarimetric signature of snow is very weak and the sensitivity of the polarimetric measurement to the corresponding parameters is also small.

Since it is suggested by previous studies (Aksoy et al. 2006a, TX08b) that a larger ensemble size leads to better parameter estimation, we performed additional five-parameter estimation experiments with a doubled ensemble size of 160. When compared to the 80-member counterparts, the estimated states are improved in general except for experiment para5_ZhKdp, which shows comparable results in a statistical sense. Some of the parameter estimations, however, experience deterioration in some experiments while larger improvements in other parameters seem to more than compensate the negative effect of these parameters on the state estimation.

We point out that the accuracies of the state and DSD parameters estimated through the EnKF system may differ when different polarimetric measurements are used. Certain combinations of polarimetric measurements may yield a better estimated state but less accurate parameter values than other combinations. This variability also

exists among state variables and estimated parameters. A better understanding of the combined impact can help optimize the assimilation/estimation system although in practice nonlinear interactions in the model, which are abundant in the complex microphysical processes, can make it difficult to delineate the effects of one source of input data or parameter value on another. While the sensitivity studies performed here and in TX08a,b are helpful, more effective approaches may be needed to further improve our understanding.

While the polarimetric data are believed to contain much useful information on the microphysics, the use of a single-moment microphysics scheme based on an assumed exponential DSD may limit the ability of polarimetric data in helping estimate the intercept parameters. If a two-moment microphysics scheme is used where both the mixing ratios and total number concentration are predicted, the intercept parameter no longer has to be specified. In this case, our goal would be changed to the estimation of both mixing ratios and total number concentrations that are now state variables. The increased number of state variables needing estimation may demand more observation information and the polarimetric observations may become a more valuable addition to the radial velocity and reflectivity observations of non-polarimetric Doppler radars. The impact of polarimetric data on full microphysical state estimation when a two-moment microphysics scheme is used will be examined in the Chapter 6.

Chapter 5 Polarimetric Signatures of a Supercell Storm Simulated Using a Two-Moment Microphysics Scheme

5.1 Introduction

The supercell thunderstorm has received the attention of the meteorology community because it often provokes serious damage from tornadoes, large hail, strong wind, and/or heavy precipitation. Many observational studies focus mainly on understanding the time evolution of storm structure, microphysical characteristics, and dynamics revealed by radar reflectivity and radial velocity data (e.g., Browning and Donaldson 1963; Browning 1964; Marwitz 1972; Musil et al. 1986; Ray et al. 1981). Numerical studies have tried to simulate such supercell storms and aid the understanding of storm initiation and development (e.g., Klemp and Wilhelmson 1978; Klemp et al. 1981; Lemon and Doswell 1979; Weisman and Klemp 1982). Recently, some research has demonstrated that the storm microphysical processes and properties can be better understood with polarimetric radar data (e.g., Bringi et al. 1986b; Hubbert et al. 1998; Kumjian and Ryzhkov 2008a,b; hereinafter KR08a,b, respectively; Romine et al. 2008).

Although conventional and polarimetric radar observations offer important insights into storms, observations are often insufficient to provide details on storm evolution due to certain limitations. Such limitations include the lack of coverage from the beam blockage or radar cone of silence, insufficient observational frequencies, and insufficient spatial resolution when the radar is located far from the storm. In addition to these external factors, reflectivity and polarimetric measurements provide only bulk

properties of all hydrometeors in the radar resolution volume, and radial velocity offers only the wind component projected in the direction of the line of sight.

On the other hand, numerical models allow us to study small details that are not directly observed within the current observational network with high temporal and spatial resolutions. They can help to substantiate findings made in observational studies. Numerical models also can be used to discover new theory. Through a numerical study by Wilhelmson and Klemp (1978), it was found that the strong low-level shear and strong low-level inflow, enough to inhibit the propagation of gust front away from the storm, are necessary for a storm split. Most of all, the numerical model is of primary importance in modern weather forecasting. However, the numerical solutions have to be validated with appropriate observations, emphasizing the complementary relationship between observations and numerical models.

For a direct comparison between model output and observation, the model variables are often converted into the form of observations using the radar emulator, which is also referred to as the radar forward observation operator in the data assimilation system. The radar emulator should be accurate, consistent with the model microphysics scheme, and make use of all the relevant information available in the model.

While many polarimetric radar emulators are mainly interested in single-phase hydrometeor concentration (Brandes et al. 1995; Brandes et al. 2004b; Ryzhkov et al. 1998; Vivekanandan et al. 1994; Zhang et al. 2001; Capsoni et al. 2001), only a handful of complete polarimetric radar emulators exist in the literature that utilizes all parameters available in the model. Huang et al. (2005) developed a simulator based on

T-matrix scattering calculations (Waterman 1969; Vivekanandan et al. 1991) coupled with the Regional Atmospheric Research Model (RAMS) employing the two-moment microphysics scheme. However, they have only examined the vertical cross-section of a simulated supercell storm, and some details on the dielectric constant model are not given in the short conference paper. They employ a simple melting treatment for ice species with fixed fractions of water and ice (and air for graupel) based on the height or air temperature.

Recently, Jung et al. (2008b) developed a polarimetric simulator combining the power-law fitting of the scattering amplitudes of rain calculated using T-matrix codes and Rayleigh scattering approximation for ice for the single-moment microphysics scheme (see Chapter 2). This method may not be as accurate as a full T-matrix scattering calculation. However, the computation is much faster, which makes it suitable for data assimilation purposes. They introduced a new melting ice model with a continuously varying density of ice particles and the fractional water in the ice.

Pfeifer et al. (2008) also proposed a polarimetric emulator called Synthetic Polarimetric Radar (SynPolRad), based on the T-matrix method. SynPolRad is coupled with a single-moment microphysics scheme with various assumptions for drop size distribution (DSD). The researchers obtained a fixed water fraction for wet ice hydrometeors by fitting the value of simulated polarimetric variables to their respective expected values within a certain range for observations. However, their dielectric constant model is physically questionable because it assumes a high-density melting core with a low-density shell. Additionally, the specific differential phase, which is a very useful polarimetric measurement, is not included in SynPolRad.

Although these emulators have their own strengths and weaknesses, they all show that the polarimetric emulators can be very useful for evaluating model microphysics. Furthermore, a computationally optimized emulator can serve as a forward observation operator in the data assimilation system.

In this chapter, we develop a radar emulator that is more general than that described in Chapter 2 and employs the full T-matrix scattering method for both rain and ice hydrometeors. This emulator has an option for the user to specify any radar wavelength for scattering calculations. In this study, it is set to 10.7 cm, which is the wavelength of Weather Surveillance Radar-1988 Doppler (WSR-88D). Model prognostic variables of single-, two-, or three-moment (hereinafter SM, DM, and TM, respectively) microphysics schemes can be used as inputs. The polarimetric variables include reflectivity at the horizontal polarization Z_H, differential reflectivity Z_{DR}, specific differential phase K_{DP}, and the co-polar cross-correlation coefficient at zero-lag $\rho_{hv}(0)$.

Recent studies of Dawson et al. (2007) suggest that supercell thunderstorms with a more realistic reflectivity structure and cold pool strength can be obtained with a high horizontal resolution (1 km or smaller grid spacing) and the multi-moment microphysics schemes of Milbrandt and Yau (2005b; 2005a), with the most improvement achieved when moving from SM to DM microphysics scheme. Therefore, we apply the new emulator to a simulated supercell storm using a DM scheme to see if this emulator can reproduce characteristic polarimetric signatures that are commonly found in the observations. This will give us an opportunity to cross-examine a DM scheme and the emulator at the same time.

This chapter is organized as follows. Section 5.2 discusses the polarimetric radar data emulator, with assumptions made for DSD-related parameters. In section 5.3, the prediction model and the configurations used to create an idealized supercell storm are briefly described. Section 5.4 presents the simulated polarimetric signatures and compares them with results of the SM scheme. Finally, the results are summarized in section 5.5.

5.2 Polarimetric radar data emulator

The emulator developed in this study is more complex and general than the one in Chapter 2, which was used in our previous studies (Jung et al. 2008c; Jung et al. 2008a). As discussed in Chapter 2, the DSD-related parameters within the emulator should be consistent with those within the numerical model. The DSDs of each species, $n(D)$, are modeled by a gamma distribution to take a maximum of three DSD parameters from a TM scheme:

$$n(D) = n_0 D^{\alpha} e^{-\lambda D} \tag{5.1}$$

where D is the particle size, and n_0, α, and λ are the concentration, shape, and slope parameters, respectively. The fixed densities of 1,000, 100, 913 kg m^{-3} for rain (ρ_R), snow (ρ_S), and hail (ρ_H), respectively, are assumed, as in the prediction model.

Some extra characteristics are required to simulate polarimetric variables, such as the shape, the statistical properties of the drop orientation, and the ice/water composition of hydrometeors. Since they are not specified in the prediction model, assumptions have to be made based on their physical properties. These assumptions are largely inherited from Chapter 2; for more detailed information, see Chapter 2.

To include the Mie scattering effect for all hydrometeor types, we perform the numerical integration of the scattering amplitudes over the DSD. This enables us to deploy the revised axis ratio relation based on the observations for rain (Brandes et al. 2002):

$$r = 0.9951 + 0.02510D - 0.03644D^2 + 0.005303D^3 - 0.0002492D^4 . \quad (5.2)$$

While a few relationships, including the ones given by Green (1975) and Beard et al. (1991), are available in the emulator, new relationships can easily be added as well.

Within the scattering calculations, the maximum sizes of rain drops ($D_{max,r}$), snow aggregates ($D_{max,s}$), and hailstones ($D_{max,h}$) are assumed to be 8, 30, and 70 mm, respectively. These size ranges are partitioned into 100 bins. For rain, dry snow, and dry hail, the forward and backward scattering amplitudes along the major and minor axes with assumed drop shape are calculated at the center of each size bin and stored in the lookup table. For melting species, the same tables are constructed at the uniform water fraction interval, which is 5% in this study. The same melting ice and dielectric constant models developed in Chapter 2 are employed in the scattering calculation. For example, for a melting snow aggregate with a specified water fraction, the density and the dielectric constant of that particle are calculated and used to compute the forward and backward scattering amplitudes at each size bin with that water fraction. These scattering amplitudes are then integrated over the DSD later, when the model mixing ratios (and the total number concentration for a DM scheme and the 6th moment of DSD for a TM scheme) are given as input.

For rain, dry snow aggregate, dry hail, rain-snow aggregate mixture, and rain-hail mixture, radar reflectivity factors at horizontal and vertical polarizations are calculated as follows (Zhang et al. 2001):

$$Z_{h,x} = \frac{4\lambda^4}{\pi^4 |K_w|^2} \int_0^{D_{\max,x}} \left[A\left|f_{a,x}(\pi)\right|^2 + B\left|f_{b,x}(\pi)\right|^2 + 2C\,\mathrm{Re}\left[f_{a,x}(\pi) f_{b,x}^*(\pi) \right] \right] n(D)dD$$

($\mathrm{mm}^6\ \mathrm{m}^{-3}$) and (5.3)

$$Z_{v,x} = \frac{4\lambda^4}{\pi^4 |K_w|^2} \int_0^{D_{\max,x}} \left[B\left|f_{a,x}(\pi)\right|^2 + A\left|f_{b,x}(\pi)\right|^2 + 2C\,\mathrm{Re}\left[f_{a,x}(\pi) f_{b,x}^*(\pi) \right] \right] n(D)dD$$

($\mathrm{mm}^6\ \mathrm{m}^{-3}$), (5.4)

where

$$A = \left\langle \cos^4\phi \right\rangle = \frac{1}{8}\left(3 + 4\cos 2\bar{\phi} e^{-2\sigma^2} + \cos 4\bar{\phi} e^{-8\sigma^2}\right),$$

$$B = \left\langle \sin^4\phi \right\rangle = \frac{1}{8}\left(3 - 4\cos 2\bar{\phi} e^{-2\sigma^2} + \cos 4\bar{\phi} e^{-8\sigma^2}\right),$$

and

$$C = \left\langle \sin^2\phi \cos^2\phi \right\rangle = \frac{1}{8}\left(1 - \cos 4\bar{\phi} e^{-8\sigma^2}\right),$$

and x can be *r* (rain) or *rs* (rain-snow mixture), *ds* (dry snow), *rh* (rain-hail mixture), or *dh* (dry hail). Here, $f_a(\pi)$ and $f_b(\pi)$ are complex backscattering amplitudes for polarizations along the major and minor axes, respectively, and f_a^* and f_b^* are their respective conjugates. Here, $\mathrm{Re}[\cdots]$ represents the real part of the complex number, and $|...|$ implies the magnitude of the value between single bars. Here, $\langle ... \rangle$ means that an ensemble average is taken over canting angles, and *n(D)* defines the DSD and is the

number of particles per unit volume of air and unit bin size. Truncation is applied at raindrop, snow aggregates, and hailstone's maximum size when the integration is performed. Here, $\bar{\phi}$ is the mean canting angle, σ is the standard deviation of the canting angle, λ is the radar wavelength, and $K_w = 0.93$ is the dielectric factor for water.

The reflectivity in linear scale for different species are combined to give logarithmic reflectivity at horizontal and vertical polarizations (Z_H and Z_V, respectively) and differential reflectivity (Z_{DR}) as

$$Z_H = 10\log_{10}\left(Z_{h,r} + Z_{h,rs} + Z_{h,ds} + Z_{h,rh} + Z_{h,dh}\right) \text{ dBZ}, \tag{5.5}$$

$$Z_V = 10\log_{10}\left(Z_{v,r} + Z_{v,rs} + Z_{v,ds} + Z_{v,rh} + Z_{v,dh}\right) \text{ dBZ}, \tag{5.6}$$

$$Z_{DR} = 10\log_{10}\left(\frac{Z_h}{Z_v}\right) = 10\log_{10}\left(\frac{Z_{h,r} + Z_{h,rs} + Z_{h,ds} + Z_{h,rh} + Z_{h,dh}}{Z_{v,r} + Z_{v,rs} + Z_{v,ds} + Z_{v,rh} + Z_{v,dh}}\right) \text{ dB}. \tag{5.7}$$

Z_{DR} is a good indicator of the mean shape of hydrometeors and depends on their relative orientation to the radar beam. Therefore, DSD shifts toward larger or smaller drop sizes can be roughly inferred from the Z_{DR} value.

The cross-correlation coefficient is defined as

$$\rho_{hv} = \frac{\left|Z_{hv,r} + Z_{hv,s} + Z_{hv,h} + Z_{hv,rs} + Z_{hv,rh}\right|}{\left\{\left(Z_{h,r} + Z_{h,s} + Z_{h,h} + Z_{h,rs} + Z_{h,rh}\right)\left(Z_{v,r} + Z_{v,s} + Z_{v,h} + Z_{v,rs} + Z_{v,rh}\right)\right\}^{1/2}}, \tag{5.8}$$

where the numerator is give as a product of two orthogonal co-polar components of the radar signals and computed as

$$Z_{hv,x} = \frac{4\lambda^4}{\pi^4 |K_w|^2}\int_0^{D_{\max,x}}\Big[C\left(|f_{a,x}(\pi)|^2 + |f_{b,x}(\pi)|^2\right) + A\left(f_{a,x}(\pi)f_{b,x}^*(\pi)\right) + B\left(f_{b,x}(\pi)f_{a,x}^*(\pi)\right)\Big]n(D)dD$$

$(\text{mm}^6\ \text{m}^{-3})$. (5.9)

ρ_{hv} is very useful in detecting the melting layer since it is sensitive to the presence of the mixture. For example, ρ_{hv} is very high for pure rain, while much lower in the presence of randomly oriented, large wetted hailstones.

The specific differential phases for rain, rain-snow aggregate mixture, and dry snow aggregate, rain-hail mixture, and dry hail are calculated as follows:

$$K_{DP,x} = \frac{180\lambda}{\pi}\int_0^{D_{\max,x}} C_k Re\left[f_{a,x}(0) - f_{b,x}(0)\right] n(D) dD \ (^\circ \text{ km}^{-1}), \tag{5.10}$$

where

$$C_k = \langle \cos 2\phi \rangle = \cos 2\bar{\phi} e^{-2\sigma^2}$$

and $f_a(0)$ and $f_b(0)$ are forward scattering amplitudes for polarizations along the major and minor axes, respectively. K_{DP} is known as more useful in quantitative precipitation estimation because it is more linearly proportional to the rainfall rate than reflectivity. However, the K_{DP} field is often very noisy in weak rain regions and vulnerable to errors.

When creating observations on the radar elevation planes, the effective earth radius model (Doviak and Zrnic 1993) is used to take into account beam bending, and a Gaussian beam weighting function described in Xue et al. (2006) is used in the vertical direction. The error model described in Xue et al. (2007) and Chapter 3 is optional to simulate observations errors.

5.3 Model configurations

An idealized supercell storm is initialized by the sounding of the May 20, 1977 Del City, Oklahoma, supercell storm (Ray et al. 1981) using the Advanced Regional

Prediction System (ARPS, Xue et al. 2000; 2001; 2003), which is a fully compressible and non-hydrostatic atmospheric prediction model. The multi-moment bulk microphysics scheme of Milbrandt and Yau (2005b; 2005a, hereinafter MY05) is recently implemented in the ARPS microphysics package (Dawson et al. 2007) and is used in this study.

With the DM option, the ARPS predicts three velocity components u, v, and w; potential temperature θ; pressure p; mixing ratios of water vapor q_v; and mixing ratios of cloud water, rainwater, cloud ice, snow aggregate, and hail (q_c, q_r, q_i, q_s, and q_h, respectively), and their total number concentrations (Nt_c, Nt_r, Nt_i, Nt_s, and Nt_h, respectively). The graupel category originally included in the MY05 package is turned off to maintain consistency with our previous experiments (Chapters 3 and 4). The turbulence kinetic energy is also predicted to be used in the 1.5-order subgrid-scale turbulence closure scheme.

A 4-K ellipsoidal thermal bubble 10km long and 1.5km high is centered at x = 48 km, y = 16 km, and z = 1.4 km and used to initiate the storm. A radiation, a rigid wall with a wave-absorbing layer, and a free-slip condition are applied to the lateral, top, and bottom boundaries, respectively.

A few changes are made to the configurations used in Chapters 3 and 4 to accommodate the use of a DM scheme. The model domain size is 63 km x 63 km x 16 km, with a horizontal resolution of 1.5 km and a vertical separation of 0.5 km. This horizontal spacing is reduced here because the model fails to simulate the storm split with a 2.0-km resolution. Constant winds of u = 1 m s^{-1} and v = 13 m s^{-1} are subtracted from the original sounding to maintain the storm near the center of the domain. The

time step sizes are 3 seconds for model integration and 1.5 seconds for acoustic wave. A monotonic scheme is applied to the 4^{th} order computational mixing to prevent Gibbs phenomenon.

The multi-moment bulk microphysics of MY05 assumes that each hydrometeor type has a constant density. The default values for rain, snow, and hail are 1,000, 100, and 913 kg m^{-3}. The DSDs for all hydrometeor types are modeled by exponential distribution.

5.4 Polarimetric signatures of a simulated convective storms

5.4.1 Storm evolution and simulated reflectivity

Figure 5.1 shows the time evolution of the reflectivity field of the simulated supercell storm with the MY05 DM scheme (Figs. 5.1e-h) at the 250-m altitude. Briefly, the updraft quickly intensifies during the first 20 minutes, with the reflectivity core greater than 40 dBZ appearing after 10 minutes of simulation (not shown). While the forward flank regions continue to extend for the next 30 minutes, the storm starts to split into two parts at around 1 hour (Fig. 5.1f). Then, the left-moving cell (relative to the storm motion vector) continues to develop while moving northwest of the right-moving cell. The right-moving storm is at its mature stage at 80 minutes of model time and maintains intensity for the next few hours. In this study, we focus only on the right-moving storm, which is usually a major one.

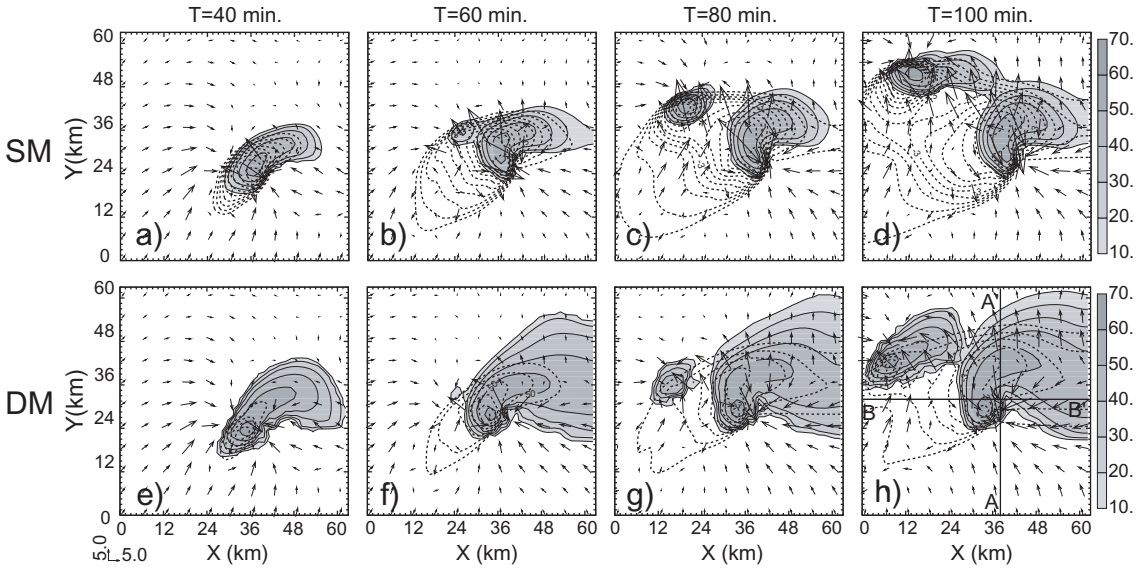


Fig. 5.1. Reflectivity (thin solid contours and shading), perturbation potential temperature (dotted contours at 0.5-K intervals for negative potential temperature) and horizontal perturbation wind (vectors, plotted every four grid point; m s^{-1}) at z = 250 m for the simulated storm for (a)-(d) SM and (e)-(h) DM runs. AA' and BB' show the locations of the vertical cross-sections passing through the updraft core (maximum vertical velocity).

As discussed in Dawson et al. (2007), the simulated reflectivity using the DM scheme (Figs. 5.1e-h) shows a more realistic structure and intensity in the hook echo and forward flank downdraft (FFD) regions compared with the simulated reflectivity with the MY05 SM scheme (Figs. 5.1a-d). The simulated storm using the MY05 SM scheme is very similar to that using the Lin et al. (1983, hereafter LFO83) ice microphysics scheme shown in Fig. 2 of TX05. Compared to the shape of the echo using the DM scheme, that using the SM scheme has a kidney shape with a narrow area of the FFD, while the DM scheme produces a swirl-shaped weak echo region (WER) wrapping around the hook echo and the extending FFD to the far east of the storm (Figs. 5.1d,h). Another significant difference between storms using SM and DM schemes is the cold pool strength. The SM scheme tends to produce a stronger cold pool than the

DM scheme. Dawson et al. (2008) showed that one of the main causes of the strong cold pool with SM schemes is the evaporative cooling of raindrops in the downdraft related to the fixed large n_{0R}, which maintains many small drops that can evaporate quickly through the microphysical heat budget analysis.

The evaluation of the microphysics scheme is beyond the scope of this study. Therefore, our main focus will be maintained if our emulator can produce the characteristic polarimetric signatures when a DM scheme is employed in the prediction model.

5.4.2 Simulated polarimetric radar fields for a supercell storm

There are several very unique polarimetric signatures repeatedly reported in the observational study. They include Z_{DR} and K_{DP} columns, a mid-level Z_{DR} ring, a hail signature (Z_{DR} hole), Z_{DR} arc (Z_{DR} shield), and mid-level ρ_{hv} ring (e.g., Wakimoto and Bringi 1988; Bringi et al. 1986b; Hubbert et al. 1998; Kumjian and Ryzhkov 2008b; a, hereafter KR08a and b; Romine et al. 2008). These signatures appear in the specific locations within a storm as a result of storm dynamics and microphysics. If the numerical model can handle storm dynamics and microphysics properly, the polarimetric radar emulator should be able to reproduce those signatures. In this study, these polarimetric signatures are discussed from the model-simulated storm point of view in the following subsections.

5.4.2.1 Z_{DR} and K_{DP} columns

Figure 5.2 shows the vertical structures of Z_{DR} (Fig. 5.2a) at line AA' shown in Fig. 5.1h and of K_{DP} (Fig. 5.2b) at line BB'. Both cross-sections pass through the

maximum vertical velocity region within the storm. The locations of the Z_{DR} and K_{DP} columns are associated with the updraft because the rainwater has to be carried aloft by a strong updraft. The Z_{DR} and K_{DP} columns extending above 0°C level are clear in Fig. 5.2.

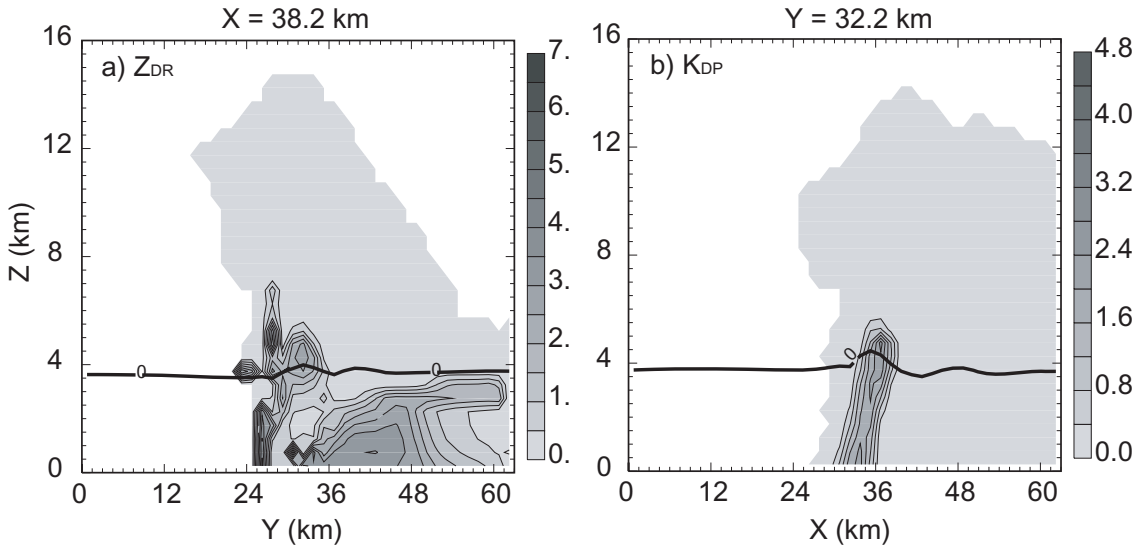


Fig. 5.2. The vertical cross-sections of simulated (a) Z_{DR} (dB) along line AA' shown in Fig. 5.1h corresponding to x = 38.2 km and (b) K_{DP} (° km^{-1}) along line BB' corresponding to y = 32.2 km of the simulated supercell storm at 100 min. The 0°C isotherms are shown as thick black lines.

The offset of their centers is pointed out in the observational study of the polarimetric signatures of KR08a. In the observations, the K_{DP} column is often found in the west or northwest of the Z_{DR} column. A similar offset is observed in the lower mid-level (Figs. 5.3a,d), while they become collocated in the upper mid-level in our simulated storm (Figs. 5.3c,f). At a 3-km altitude, the Z_{DR} column is located in the south of the updraft core near the reflectivity hook. This location agrees well with that of the Z_{DR} column reported in Hubbert et al. (1998). At this level, the K_{DP} column is found in the reflectivity maximum and located in the west of the updraft core. At a 4km altitude,

the Z_{DR} column looks close to the half-ring around the updraft core to the east, while the center of the K_{DP} column is almost collocated with the updraft core. At a 5-km height, the Z_{DR} and K_{DP} columns are collocated with the updraft core.

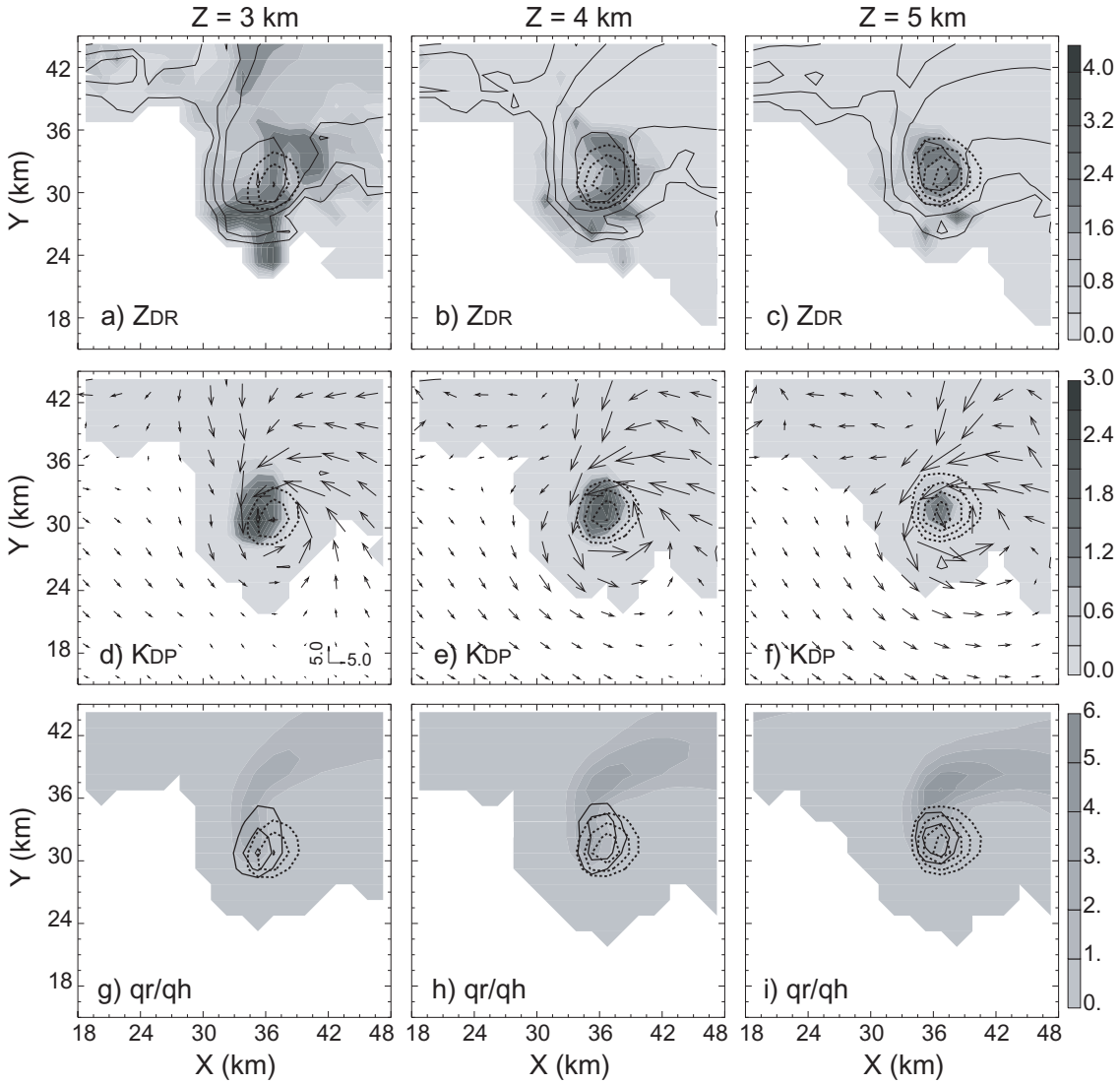


Fig. 5.3. (a)-(c) Z_{DR} (shading, dB) and reflectivity (solid contours at 15-dBZ intervals, starting from 10 dBZ), (d)-(f) K_{DP} (shading, ° km^{-1}) and horizontal perturbation wind (vectors, plotted every other grid point; m s^{-1}), and (g)-(i) hail mixing ratio q_h (shading, g kg^{-1}) and rain mixing ratio q_r (solid contours at 1.0-g kg^{-1} intervals, starting from 1.0 g kg^{-1}) at (a),(d),(g) z = 3 km, (b),(e),(h) z = 4 km, and (c),(f),(i) z = 5 km at 100 min of storm. The vertical velocity (dotted contours at 5-m s^{-1} interval from 10 m s^{-1}) is overlaid on each plot.

The main cause of this spatial offset of the Z_{DR} column from the updraft core is the presence of hail. From Fig. 5.3g,h,i, it is apparent that the production of q_r is strongly related to the updraft. Within the updraft, the low-level q_r is transported to the higher level, and/or q_r is also created through the condensation on site. q_h is also produced in the updraft. Large hailstones fall through the updraft, and small hailstones are carried by the horizontal wind while growing and falling in the FFD north of the updraft core. The high q_h region is overlapped with the high q_r region from the north or northwest. The presence of hail reduces the Z_{DR} because the tumbling nature of the hailstones makes their apparent shape close to spherical. Therefore, the high Z_{DR} column shows up at the south or southeast of the updraft.

On the other hand, the K_{DP} is almost transparent to the hail and only sensitive to the amount of rainwater. The K_{DP} maximum, high q_r region, and updraft core are, hence, almost collocated.

5.4.2.2 Z_{DR} arc

The Z_{DR} arc is the low-level signature repeatedly observed in the southern edge of the FFD along the sharp gradient of reflectivity in the right-moving supercells (KR08a). This is characterized by a horizontally elongated high Z_{DR} band along the right edge of the FFD near the surface. Although this is a quite consistent feature found in the supercells, regardless of seasons and geographic regions, this signature has been noted only recently (KR08a,b, Romine et al. 2008). This signature is analyzed in detail in KR08b, where it is argued that the size-sorting mechanism, due to speed and directional wind shears, is primarily responsible for this signature; large drops

discharged from the updraft fall into the region close to the origin, while smaller drops are advected farther into the FFD. The rain evaporation is likely another source of a modified DSD toward a large D_{nr}. In the location of the Z_{DR} arc, the DSD, initially lacking small drops, loses small drops fast while falling through the dry air. Therefore, a two- or higher moment microphysics scheme has to be used to simulate this feature because the SM is not capable of handling size-sorting mechanisms (Fig. 5.4a).

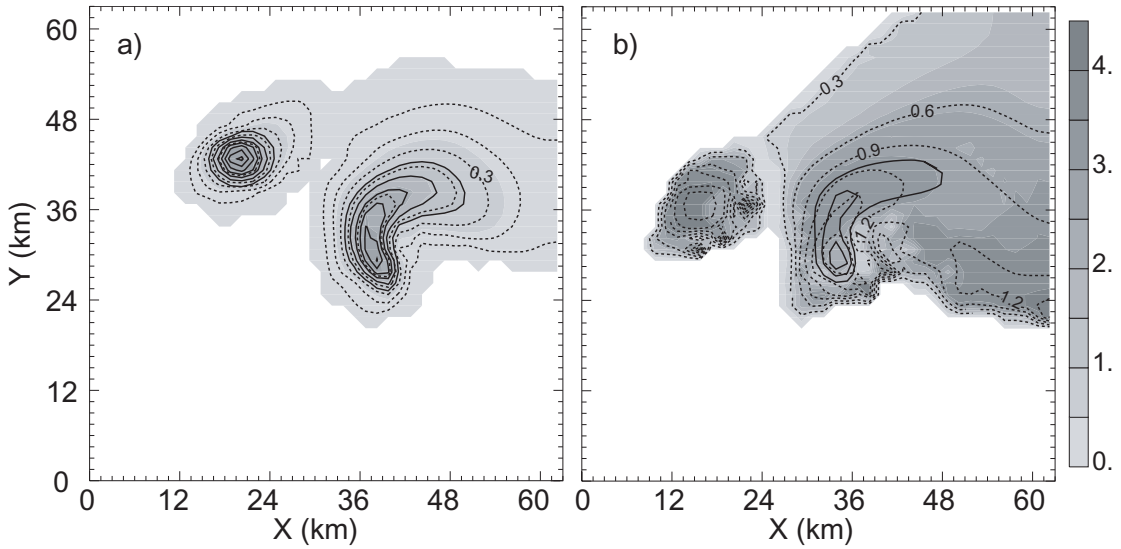


Fig. 5.4. Z_{DR} (shading), q_r (solid contours at 0.5 g kg^{-1} intervals, starting from 0.5 g kg^{-1}), and a) mean-mass diameter of rain drops D_{nr} (dotted black contours at intervals of 0.1 mm, starting from 0.1 mm) for the SM simulation and b) D_{nr} (intervals of 0.3 mm, starting from 0.3 mm) for the DM simulation at z = 500 m at 80 min. of the storm.

The modified DSD of rain as a result of the size sorting can be evaluated easily by examining the mean-mass diameter D_{nr}, where the D_{nr} is calculated for the exponential distribution as

$$D_{nr} = \left(\frac{q_r \rho_{air}}{\pi \rho_r N t_r} \right)^{1/3} . \tag{5.10}$$

where ρ_{air} is the density of air. For the SM run, n_{0R} is used to calculate D_{nr}, rather than Nt_r using $Nt_r = n_{0R} D_{nr}$. The modified DSD directly affects Z_{DR} because Z_{DR} is proportional to the median diameter of precipitation particles in the radar resolution volume. The calculated D_{nr} for the SM and DM simulations is presented in Fig. 5.4 along with simulated Z_{DR}.

The Z_{DR} arc signature is well captured by the DM scheme and polarimetric radar data emulator at a 0.5-km altitude at 80 min. of storm in Fig. 5.4b. This high Z_{DR} region along the southern edge of the FFD is shallow (~ 2-km depth), rather narrow, and persistent in time, as observed, and shifts slightly toward the north with its height. This band becomes weak and broad above 2 km, practically fading away. The shape and location of the high Z_{DR} region agrees with the D_{nr} pattern. The q_r pattern, in conjunction with the D_{nr}, substantiates that this is an area with a small number of large drops and lacking small drops. On the other hand, the simulated storm using the SM scheme completely misses the Z_{DR} arc signature because both D_{nr} and Z_{DR} are proportional only to the q_r (Fig. 5.4a).

5.4.2.3 Mid-level Z_{DR} and ρ_{hv} rings

The mid-level Z_{DR} ring (KR08a) refers to the enhanced Z_{DR} value in the shape of a ring with a depressed Z_{DR} value in the middle. The Z_{DR} ring is sometimes a complete circle and sometimes just a half-ring. KR08a reports that the enhanced Z_{DR} region is always found on the right flank of the updraft when only a half-ring is manifest. In our simulation, it is usually a half-ring on the right flank of the updraft in the mid-level and close to a complete ring in the lower level.

The half-Z_{DR} ring in Fig. 5.5a is mostly overlapping with the high D_{nr} region. The maximum D_{nr} region is collocated with the updraft core. The local maxima found on the south and east of the main D_{nr} core may be explained by the large raindrops falling around the updraft following a cyclonic wind (Fig. 5.3d-f). The half-ring signature is highly correlated with the presence of hail (Fig. 5.5a). As discussed in subsection 5.5.2.1, the presence of hailstones reduced the Z_{DR} value because their tumbling motion and random orientation make their apparent shape spherical to the radar beam. At a 4km altitude, the region with the high hail-mixing ratio locates to the west and northwest of the updraft core. This weakens the Z_{DR} signature on the left flank of the updraft.

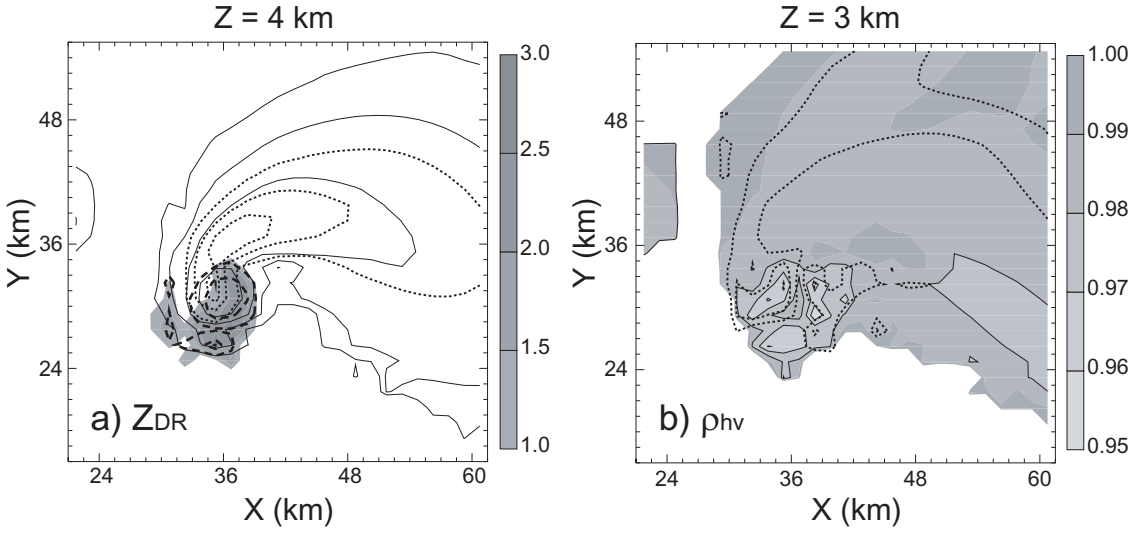


Fig. 5.5. (a) Z_{DR} (shading), Z_H (solid contours at 15 dBZ intervals from 15 dBZ), q_h (dotted contours at 1.0 g kg^{-1} intervals from 0.5 g kg^{-1}), and D_{nr} (thick dashed contours at intervals of 0.6 mm, starting from 0.9 mm) at z = 4 km, and (b) ρ_{hv} (shading and solid contours at 0.01 intervals starting from 0.98 and lower) and the ratio of rain-hail mixture to the sum of rain and dry hail mixing ratios (dotted contours at 0.1 intervals from 0.2) at z = 3 km at 80 min. of the storm. Ring features are prominent at mid-levels.

The ρ_{hv} ring is another mid-level feature with a depressed ρ_{hv} value in the shape of a ring (KR08a). A well-defined ρ_{hv} ring is seen in Fig. 5.5b. The ρ_{hv} values for pure water and ice are very high and decrease when hydrometeors of diverse types are mixed together in the same resolution volume. The dotted contours in Fig. 5.5b show the ratio of rain-hail mixture to the sum of rain and dry hail mixing ratios. The high value of this ratio indicates that three different types coexist, but the mixture is dominant in that area. The low value of this ratio suggests that either pure rain or dry hail is dominant in that region (see Eq. (2.2) and section 2.3.2 for more details on the melting ice model). The pattern of this ratio agrees well with the ring-shaped ρ_{hv} depression.

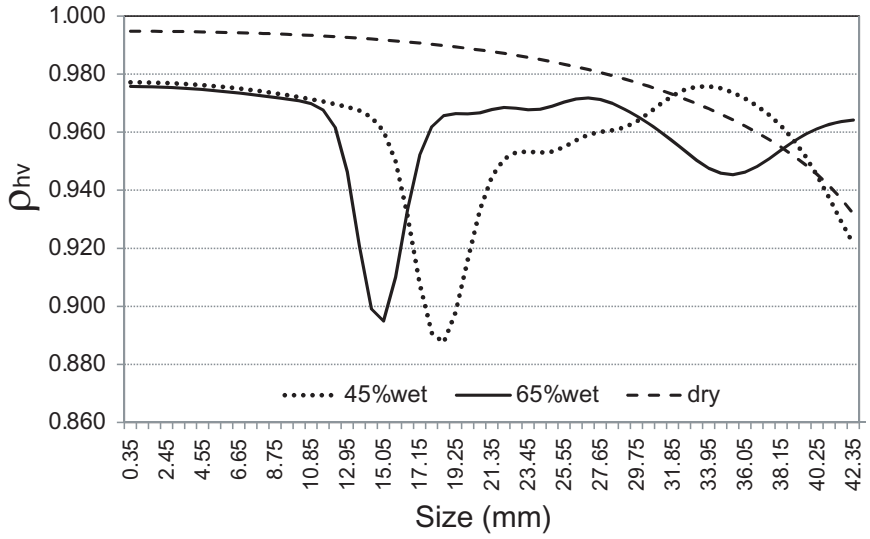


Fig. 5.6. Simulated ρ_{hv} with size of melting hailstones for the fraction of water, f_w = 0.45 (dotted), 0.65 (solid), and dry hailstone (dashed).

When ice is in a melting phase, the resonance effect can contribute to the reduction of the ρ_{hv} value (Fig. 5.6). For dry hail, the ρ_{hv} slowly decreases with size within the range shown in Fig. 5.6, in which the particle size is truncated at 4.235 cm. With the exponential DSD, large drops have little effect because the number of drops is

very small at that size, although the resonance effect can be much more significant in very large drops. It can be seen that the ρ_{hv} shows a sudden drop at a certain size. Both that characteristic size and the maximum amplitude decrease with an increasing water fraction. The size-sorting mechanism is also necessary to simulate this signature so that the DSD can have a sufficient number of hailstones at the characteristic size to reduce the total ρ_{hv} values. This signature is very weak or completely missing when a SM scheme is used (not shown).

We would like to point out that the simulated ρ_{hv} is higher than the observed one, in general. The non-meteorological effects that can also contribute to the reduction of ρ_{hv}, such as noise bias, clutter contamination, dust and bugs, are not included in our emulator. This fact may partially be responsible for the rather high ρ_{hv}. The other source of the difference might be due to the simplified model of randomly orientated spheroids for hail and snow, which does not account for the effect of the irregular shape of the natural hydrometeors.

5.4.2.4 Hail signature in the forward flank downdraft

The hail signature (KR08a) is characterized by a high Z_H and low Z_{DR} at the lowest tilt associated with hail reaching the ground. This feature is also called "Z_{DR}-Hole" (Wakimoto and Bringi 1988), which often stretches from the surface up to a certain height. Our simulated storm does not exhibit this signature near the surface because most of the hail completely melts while falling in this area. However, the emulator is capable of simulating this signature if a significant amount of hail survives the melting and reaches the ground. For example, a hail signature is evident at a 2.5km

height, where the relatively large hail concentration presents (Fig. 5.7). The simulated Z_{DR} pattern with a Z_{DR}-Hole surrounded by high Z_{DR} values shows a remarkable similarity to the observation in Fig. 3a of KR08a.

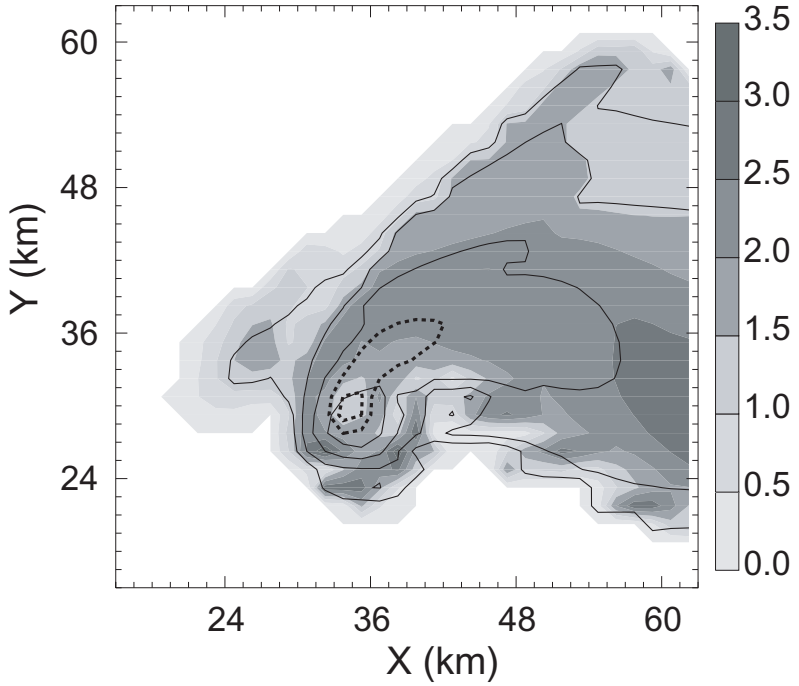


Fig. 5.7. The hail signature is characterized by high Z_H (thin solid contours at 15 dBZ intervals, starting from 15 dBZ) and low Z_{DR} (shading, dB) at z = 2.5 km at 60 min. of storm. The hail mixing ratio is overlaid in thick dashed contours at 1 g kg^{-1} intervals from 1 g kg^{-1}.

5.5 Summary and discussions

In this chapter, a synthetic polarimetric radar emulator based on full T-matrix scattering calculations and accurate formulations for polarimetric radar variables is developed. This emulator takes advantage of the continuous melting ice model developed in Chapter 2. The density of the melting ice and dielectric constant are also allowed to vary continuously. This emulator is flexible enough to take a radar

wavelength as a user input and to choose one of the one-, two-, and three-moment microphysics schemes to plug in.

This emulator can simulate the reflectivity of the horizontal and vertical polarizations (Z_H and Z_V), differential reflectivity (Z_{DR}), specific differential phase (K_{DP}), and cross-correlation coefficient (ρ_{hv}) including radial velocity (V_r). These measurements are what will be observed by operational WSR-88D radar after the polarimetric upgrade and are currently being observed by the research polarimetric WSR-88D radar, KOUN.

The new radar emulator is applied to the idealized supercell storm simulated using a two-moment (DM) microphysics scheme. Another storm with the same configurations but using a single-moment (SM) microphysics scheme is created for comparison. The simulated storm using a DM scheme exhibits the unique polarimetric signatures reported in the literatures, including the Z_{DR} and K_{DP} columns, Z_{DR} arc, mid-level Z_{DR} and ρ_{hv} rings, and hail signature. Some of the signatures associated with the size-sorting mechanisms, however, could not be simulated when a SM scheme was used. These signatures include the Z_{DR} arc and mid-level Z_{DR} and ρ_{hv} rings. These results support that a two- or higher-moment microphysics has to be used to describe storm microphysics and kinematics more realistically. It would be especially important for polarimetric data to be useful when they are directly assimilated in the numerical models.

To evaluate the importance of the accuracy of the simulator, we compared these signatures with those simulated using the simplified version of the emulator developed in Chapter 2, where the efficiency is considered to be equally as important as the

accuracy when developing it for data assimilation purposes. The simulator developed in Chapter 2 can simulate most of these signatures. However, the Z_H, Z_{DR}, and K_{DP} values are somewhat higher than observations and the ρ_{hv} ring could not be simulated correctly because the Mie scattering for the ice species is not included in Chapter 2, which seems to be critical to this feature (Fig. 5.8).

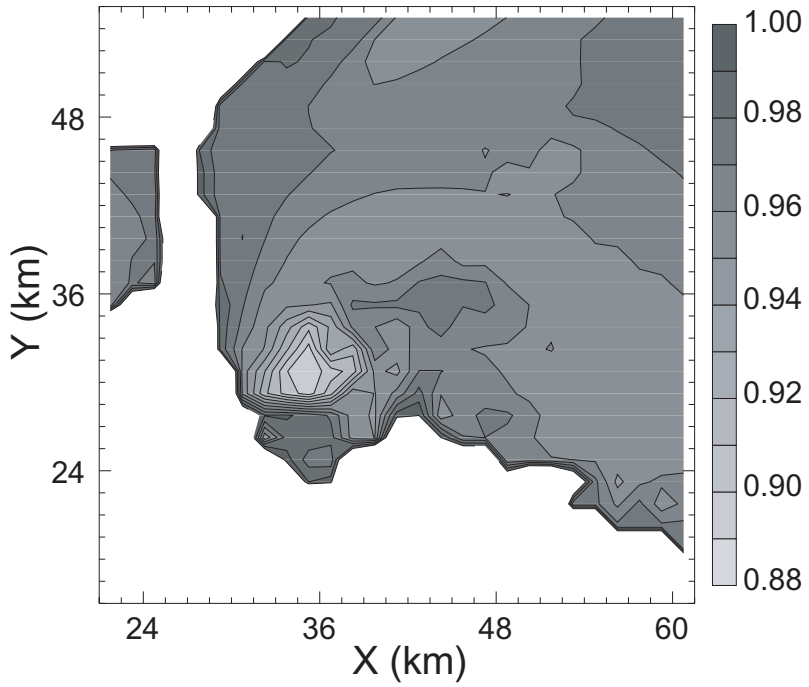


Fig. 5.8. As in Fig. 5.5b, but for simulation using the observation operator of Chapter 2. The Mie scattering effect is not included in this simplified version.

The verification of the numerical weather prediction is challenging because most of the model prognostic variables are not directly observed. In the convective scale forecast, the reflectivity has been used to verify the model prediction for decades. However, reflectivity alone is not sufficient to verify microphysics because many independent variables and uncertain constants based on many assumptions on DSDs are involved in reflectivity simulation. Here, simulated polarimetric variables can help

discriminate against and/or highlight certain variables from others by using their differential sensitivity to the phase. They can be even more useful than reflectivity because they contain extra information on the DSDs and microphysical processes. As an example, Chapter 2 demonstrated that a realistic radar simulator could be useful in evaluating the model microphysics scheme by identifying delayed melting processes in the Lin-type microphysics scheme. Additionally, this simulator could provide a unique opportunity to study storm microphysical processes more in depth.

More importantly, a good radar simulator can help forecasters to determine where attention and monitoring should be directed in a severe weather event. KR08a related some of the polarimetric signatures to the severity of the storms: strengthening the Z_{DR} and K_{DP} columns and ρ_{hv} to the updraft strength, and the Z_{DR} arc to the enhanced storm-relative environmental helicity (SREH). This suggests that polarimetric signatures can be used as an indicator of storm intensity. Although the numerical model provides the vertical velocity as one of the prognostic variables, observed quantities can often be interpreted more intuitively. For instance, analyzing the Z_{DR} field could be as informative as examining the mixing ratio and the number concentration of each hydrometeor type separately in addition to the vertical velocity field. In this regard, the polarimetric emulator can be useful to the forecasters by providing the opportunity to look at the model polarimetric fields in advance of the occurrence of the actual event.

For higher frequency radar such as the X-band radar, attenuation from the severe storm can be significant enough to alter observed polarimetric signatures, which are very different from the theoretical ones. In this case, the attenuation algorithm must be included to produce realistic polarimetric signatures. This will be developed as a part of

the emulator in the future. The evaluation of both the multi-moment microphysics scheme and the simulator developed in this study using KOUN data is also planned for future study.

Chapter 6 State Estimation of Convective Storms with Two-Moment Microphysics Scheme and Ensemble Kalman Filter: Experiments with Simulated Data

6.1 Introduction

Numerical weather prediction (NWP) is an initial-boundary value problem: the accuracy of the final solution depends on the accuracy of the initial condition, given appropriate boundary conditions. Therefore, a significant amount of research has been carried out in the field of the data assimilation (DA) to minimize the errors in the initial condition (Le Dimet and Talagrand 1986; Courtier and Talagrand 1987; Evensen 1994; Evensen and Leeuwen 1996; Burgers et al. 1998; Houtekamer and Mitchell 1998; Anderson 2001; Bishop et al. 2001; Whitaker and Hamill 2002; Evensen 2003; Tippett et al. 2003; Gao and Xue 2007; Liu et al. 2007).

Several past studies have shown that the ensemble Kalman filter (EnKF) techniques can be successfully applied to the convective scale through both Observing System Simulation Experiments (OSSEs) and real-data experiments (e.g., Doswell et al. 2004; Houtekamer et al. 2005; Snyder and Zhang 2003; Tong and Xue 2005b, hereafter TX05; Xue et al. 2006, hereafter XTD06; Whitaker et al. 2004). The fact that the EnKF technique is capable of handling complex and highly nonlinear processes in the assimilation process makes it attractive for convective scale DA.

However, the forecast is also hampered by model errors that can arise from many sources (e.g., insufficient spatial and/or temporal resolution, misrepresentation of the physical and sub-grid scale processes). It is suggested that the model error can dominate the error growth during the forecast, even if it starts from an accurate initial

condition (Zhu and Navon 1999; Houtekamer et al. 2005). Therefore, it is equally important to improve the forecast model while making an effort to improve the model initial condition.

In the convective scale NWP, the supercell thunderstorms have received a significant amount of attention because they often accompany destructive tornados, large hail, strong winds, and/or heavy precipitation. The structure and time evolution of these storms is highly sensitive to the drop-size distributions (DSDs) of the hydrometeors involved in the microphysics scheme, which is used in simulating the microphysical processes in the real atmosphere (McCumber et al. 1991; Ferrier et al. 1995; Gilmore et al. 2004a; van den Heever and Cotton 2004; Tong and Xue 2008b). The hydrometeor in one phase—either vapor, water, or ice—interacts with the others through microphysical processes, condensation or deposition, collection, breakup, freezing, evaporation or sublimation, melting, and precipitation sedimentation. The DSD-related parameters of each hydrometeor explicitly appear in the governing equations of microphysical processes and influence the magnitude and relative importance of those processes. For example, many small drops in the downdraft area will evaporate faster than a few large drops, hence inducing a stronger cold pool. Therefore, modeling accurate DSDs has a profound impact on the success of the simulation of precipitating systems.

The single-moment bulk microphysics scheme is widely used in the current research and in operational NWP models. In the single-moment microphysics scheme, the shape of the DSD is predefined, and the total number concentration is uniquely determined by the mixing ratio of each hydrometeor type. This single-moment

microphysical parameterization is not capable of handling the change in shape of the DSD in time and space as a result of the size-sorting mechanisms, such as the differential conversion to other species, the differential sedimentation, vertical motion, rotation, unidirectional shear, and directional shear (Kumjian and Ryzhkov 2008b). Kumjian and Ryzhkov (2008b) argued that size sorting due to the unidirectional and directional shears is primarily responsible for the unique Z_{DR} arc signature consistently found in the polarimetric data collected from many supercell storms.

The rapid evolution of computing hardware technology during recent decades has enabled us to use finer model grids and more complex microphysics schemes, which can overcome the limitation of the single-moment microphysics schemes in simulating more realistic DSD models (Ferrier 1994; Ferrier et al. 1995; Morrison et al. 2005; Morrison and Pinto 2005; Milbrandt and Yau 2005b, a). Dawson et al. (2007) compared the supercell storms simulated using one-, two-, and three-moment (SM, DM, and TM, respectively) microphysics schemes implemented in the Advanced Regional Prediction System (ARPS) microphysics package and showed that the DM and TM schemes simulate more realistic storm structures than the SM schemes. This is substantiated by the results presented in Chapter 5, where a polarimetric radar data simulator using a rigorous approach in computing scattering matrix is developed and applied to a simulated supercell storm using SM and DM schemes. In Chapter 5, the SM fails to simulate certain characteristic polarimetric signatures such as Z_{DR} arc and mid-level Z_{DR} and ρ_{hv} rings. These studies suggest that the forecast skill is likely improved by employing higher-moment microphysics schemes.

Motivated by Dawson et al. (2007) and Chapter 5, we employ a DM scheme in the prediction model and perform EnKF analysis in this study. The forward observation operator developed in Chapter 2 is modified to accommodate a DM scheme in the forecast model. This chapter is organized as follows: The prediction model and the experimental design used in this study are first briefly described in section 6.2. A special treatment applied to the total number concentration is then discussed. The simulated observations created for both the perfect and imperfect model experiments are also discussed. In section 6.3, we examine the ability of filter in retrieving increased numbers of model state variables through Observing System Simulation Experiments (OSSEs) using the perfect model assumption. The impact of assimilating additional polarimetric variables is also investigated. Section 6.4 discusses the experiment results in the presence of forecast model error with/without observation operator error. Finally, we summarize the results and conclusions in section 6.5.

6.2 Model and experimental setup

6.2.1 Prediction model and EnSRF assimilation procedure

ARPS is used in both simulation and analysis in this study. Briefly, ARPS is a fully compressible and nonhydrostatic atmospheric prediction model (Xue et al. 2000; 2001; 2003). With the DM scheme of Milbrandt and Yau (2005b; 2005a, hereinafter MY05), ARPS predicts three velocity components u, v, and w; potential temperature θ; pressure p; mixing ratios of water vapor q_v; and mixing ratios of cloud water, rainwater, cloud ice, snow aggregate, and hail (q_c, q_r, q_i, q_s, and q_h, respectively) and their total number concentrations (Nt_c, Nt_r, Nt_i, Nt_s, and Nt_h, respectively). Exponential drop size

distributions are assumed for microphysical variables. The graupel category originally included in the MY05 package is tuned off to maintain the consistency with previous experiments (Chapters 3 and 4). The turbulence kinetic energy is another prognostic variable used by the 1.5-order subgrid-scale turbulence closure scheme.

Configurations are largely inherited from Chapters 3 and 4, with some modifications presented later in this section. A truth simulation is created for a supercell storm initialized by the sounding of the 20 May 1977 Del City, Oklahoma, supercell storm (Ray et al. 1981). The ensemble square-root filter (EnSRF) after Whitaker and Hamill (2002) is employed, in which observations are not perturbed. The full description of the filter can be found in XTD06. The assimilation of Z_H and V_r starts at 25 min of model time and is repeated every 5 min until 100 min.

Some modifications to the configurations of the prediction model and assimilation system are made to accommodate the used of the DM scheme in the prediction model. Most of all, the forward observation operator developed in Chapter 2 has been modified to take both mixing ratios and total number concentrations of rain, snow aggregate, and hail as input. The horizontal resolution is 1.5 km because the model fails to simulate the storm split with 2.0-km resolution. Domain size is adjusted to 63 km x 63 km x 16 km in accordance. Constant winds of $u = 1$ m s^{-1} and $v = 13$ m s^{-1} are subtracted from the original sounding to maintain the storm near the center of the domain. The ensemble consists of 80 ensemble members. The filter uses a covariance localization radius of 4.5 km to be consistent with three grid points and covariance inflation factors, as shown in Table 6.1 and 6.2 (Anderson 2001; Houtekamer and Mitchell 1998, 2001; Hamill et al. 2001).

The spatially smoothed stochastic perturbations with the correlation length scale of 4.5 km are added to the initially horizontally homogeneous first guess defined by the Del City sounding to initialize the ensemble members at t = 20 min of model time; for velocity components (u, v, and w), potential temperature (θ), and water vapor (q_v) to the entire domain; and for mixing ratios of hydrometeors (q_c, q_r, q_i, q_s, and q_h,) at the grid points located within 6 km from observed reflectivity, Z, greater than 10 dBZ. The standard deviations of those perturbations are 2 m s^{-1}, 2 K, and 0.6 g kg^{-1}, and 0.6 g kg^{-1}, respectively. The total number concentrations of rain, snow, and hail are diagnosed using their default values of the intercept parameter of the MY05 SM scheme (1.0×10^6 m^{-4}, 1.0×10^7 m^{-4}, and 1.0×10^5 m^{-4}, respectively) and mixing ratios. The Nt_c is set to the constant 1.0×10^8 m^{-4}, and Nt_i is calculated as a function of temperature following Cooper (1986). In this way, the total number concentrations can be considered perturbed.

In the early cycles, the covariance between model state variables and observations is unreliable when the errors in the state estimates are still very large. Assimilating additional polarimetric variables is found to harm the analysis at this stage. The errors in the state estimates are usually significantly reduced in the first 2 to 3 cycles when a SM scheme is used. With a DM option, it takes more cycles to reduce errors to the level of those with a SM scheme, which will be discussed in section 6.3.1. Therefore, we delay the assimilation of Z_{DR} and K_{DP} until 50 min of model time or the time of the sixth EnKF analysis.

The time evolution of truth storm at 250-m altitude is show in Figs. 6.1a-d. Briefly, the updraft quickly intensifies during the first 20 min (not shown), and the

forward flank region keeps extending afterward. The storm split starts at around 1 hour (Fig. 6.1b). The left-moving cell relative to the storm motion vector is completely separated from the right-moving storm at the low level at 80 min, which is at the mature stage. The left-moving cell continues to develop and move northwest of the right-moving cell, while the right-moving storm maintains intensity for the next few hours.

6.2.2 Transform of the total number concentration in the filter

10log(**x**) or ln(**x**) are often used in the assimilation system to reduce the dynamic range of data and/or to avoid negative values, where **x** is the positive definite state vector (Hogan 2007; Tong and Xue 2008b, Chapter 4). The drawback of the log transform is that it loses its sensitivity with increasing value. As a result, a small overestimation (or underestimation) in the log domain can lead to a large error in the physical domain when the value of a variable subject to the estimation is very large. This necessitated the application of upper and lower bounds in the parameter estimation in Chapter 4 to prevent a large deviation in the early assimilation cycles that can lead to the divergence of the solution. However, applying the upper and lower bounds is not appropriate in the estimation of *Nt* because they can be as small as 0, and the upper limit is not known. Therefore, we use the power of *Nt* where the power is smaller than 1 so that we keep the sensitivity at the large values while reducing the dynamic range of data. The power is experimentally determined and set to 0.4. The state vector of *Nt* for various hydrometeors at one model grid point during assimilation is, therefore,

$$\mathbf{x} = \begin{pmatrix} \left(Nt_c\right)^{0.4} \\ \left(Nt_r\right)^{0.4} \\ \left(Nt_i\right)^{0.4} \\ \left(Nt_s\right)^{0.4} \\ \left(Nt_h\right)^{0.4} \end{pmatrix}. \tag{6.1}$$

6.2.3 Observations

Two sets of observations are created in this study. For the perfect model experiments, model state variables of the truth simulation are first converted to observations at model grid points using the forward observation operator developed in Chapter 2. This observation operator is modified to accommodate the new variables in the DM scheme as discussed in section 6.2.1, which are the total number concentration, in a way that maintains the consistency with the prediction model and used to create simulated observations. Then vertical interpolation using a Gaussian beam weighting (XTD06) and the effective earth radius model (Doviak and Zrnic 1993) is performed to obtain observations on the radar elevation level. Finally, simulated noise following Xue et al. (2007) is added to error-free observations. As in Chapter 4, we drop the sample error that is larger than 10 dBZ. The error for reflectivity is set to 42.7 % of the truth reflectivity to obtain an effective error standard deviation (SD) of about 2 dBZ for Z_H. We lower the maximum fraction of hydrometeors in the mixture form $F_{\max}$ for hail to 0.3, which was 0.5 in the previous experiments, to mitigate unrealistically high reflectivity due to the Rayleigh scattering assumptions. Gaussian errors with zero mean and SDs of 1 m s^{-1} for V_r are added to truth V_r.

For K_{DP} and Z_{DR}, error-free observations are used in the data impact experiments to examine the maximum possible improvement that can be obtained when polarimetric variables are assimilated in addition to the conventional radar data.

To take into account the model error, we create observations using the supercell simulated with the same configurations but with the TM scheme of MY05 and the polarimetric simulator developed in Chapter 5. This comprehensive simulator employs rigorous scattering calculations using the T-matrix method. Details on the simulator can be found in Chapter 5. In this way, the simulation performed using a TM microphysics is considered perfect, and we are dealing with two sources of error: 1) model error originating from misrepresentation of DSDs and 2) forward observation operator error due to the misrepresentation of the scattering properties. The procedure of creating observations and the characteristics of them are the same as those for the perfect model experiments described above. In this experiment, the observations are considered independent of the forecast model.

6.3 Perfect model experiments

6.3.1 Assimilation of conventional radar data

Table 6.1 lists the set of the assimilation experiments performed in this subsection. Three experiments, which assimilate reflectivity and radial velocity individually and together, are conducted under the perfect model assumption. The experiment names are self-descriptive. For example, experiment VrZh assimilates both V_r and Z_H data, while experiment Vr assimilates only V_r data. Experiment VrZh serves

as a control run in subsection 6.3.2, which tests the impact of polarimetric variables and the imperfect model experiment with a perfect observation operator in subsection 6.4.1.

TX05 showed that the reflectivity of the clear echo helps suppress the spurious cells in echo-free region. Therefore, reflectivity data from the entire radar coverage are assimilated, while radial velocity data are assimilated only where $Z \geq 10$ dBZ, as in the previous studies (e.g., TX05; XTD06, Chapters 3 and 4). All model state variables are updated when we assimilate Z_H and/or V_r. The optimal inflation factors are set to 7%, 20%, and 20% for experiments Vr, Zh, and VrZh, respectively.

Table 6.1. List of data assimilation experiments.

Experiments	Observation(s) assimilated	Covariance inflation
Vr	V_r ($Z_H > 10$ dBZ)	7%
Zh	Z_H (everywhere)	20%
VrZh	V_r ($Z_H > 10$ dBZ) and Z_H	20%

In Fig. 6.1, the simulated reflectivity, along with analyzed horizontal perturbation winds and perturbation potential temperature near the surface of experiments Vr, Zh, and VrZh, is compared with their truth at 20-min intervals, starting from 40 min of simulation. When radial velocity is assimilated alone (Figs. 6.1e–h), spurious echoes are not suppressed because radial velocity data are not available in these areas and, therefore, no observations are assimilated. These spurious echoes quickly developed during the forward model integration after we added random perturbations to mixing ratios within 6 km from the grid point, where reflectivity is greater than 10 dBZ. Once they are created, they survive through assimilation cycles

while maintaining a low value of mixing ratios. With a DM microphysics scheme, a low mixing ratio can produce high reflectivity if the total number concentration is low enough. In fact, the high reflectivity shown in echo-free regions in Figs. 6.1e–h are mostly associated with very low mixing ratios (Fig. 6.3), which will be discussed later. The reflectivity and cold pool structure of experiment Vr is poorly estimated in the echo-free region throughout analysis cycles. However, the structure of reflectivity near the core and forward-flank downdraft (FFD) area is retrieved reasonably well at the later time of assimilation cycles, where the reflectivity is strong (Figs. 6.1g–h). The shape and strength of retrieved cold pool cores of both left- and right-moving cells and the location and orientation of the front generally agree to those of the truth at 100 min.

When reflectivity is assimilated alone (Figs. 6.1i–l), reflectivity of the clear air effectively suppresses the spurious echo at the first cycle (not shown). The reflectivity pattern is almost as good as that of VrZh (Figs. 6.1m–p) and is very close to truth (Figs. 6.1a–d). The general cold pool pattern matches that of truth, although the strength tends to be overestimated, and the center of the cold pool is shifted to the west.

Cold pool strength and wind fields of VrZh (Figs. 6.1m–p) show the best match to the truth among Vr, Zh, and VrZh. Improvement over Zh is clear in and around the wind fields of the left-moving storm and in the location and strength of the cold pool after 80 min.

In general, surface convergence in the updraft and divergence in FFD are well captured by retrieved wind fields of all three experiments. The retrieved wind directions and speeds are reasonable compared to the truth wind in the echo region, while the

perturbation wind analysis at the upper left and lower left corners of the domain is rather poor, where is echo-free region and perturbation wind is relatively weak.

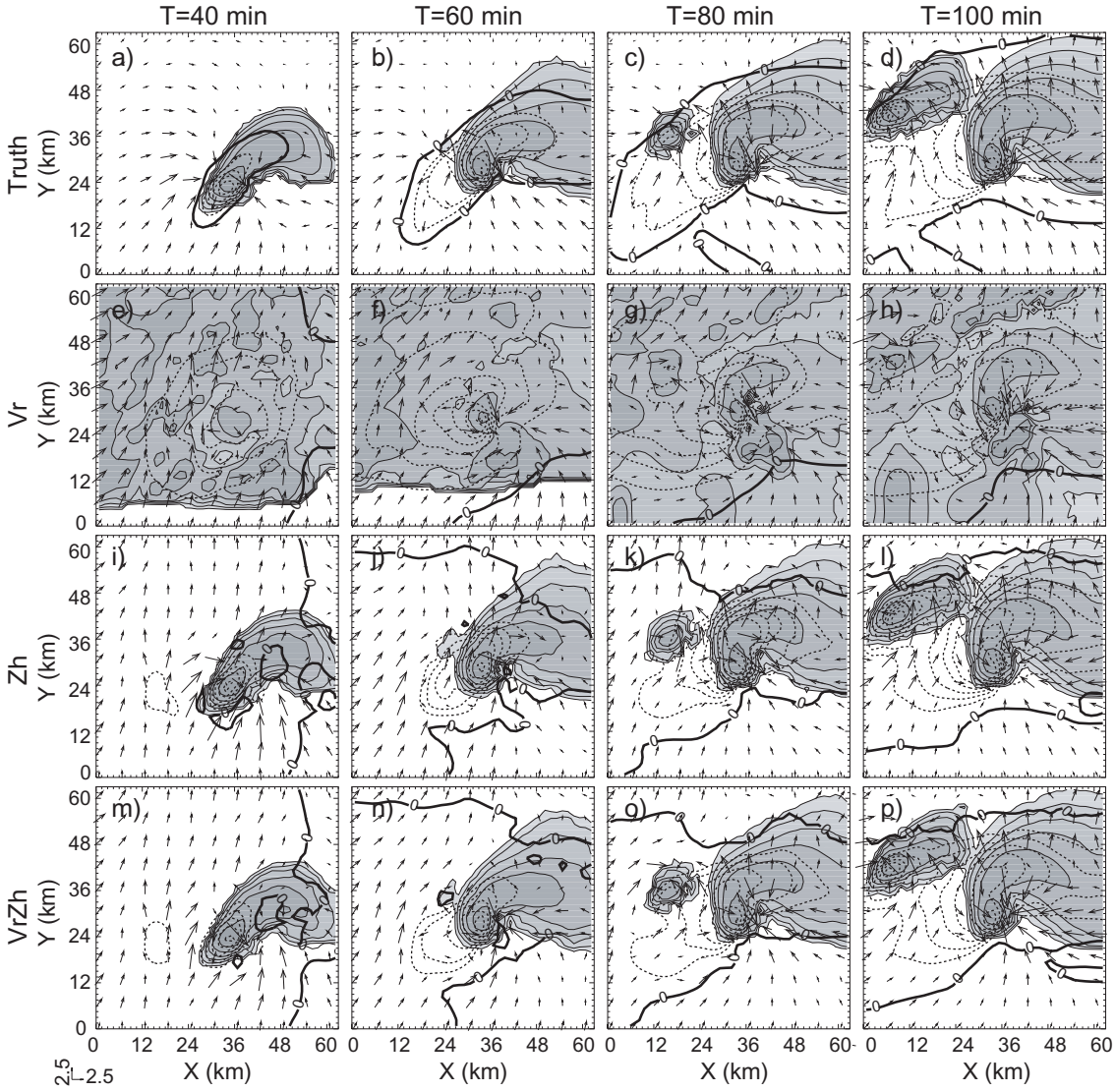


Fig. 6.1. The perturbation wind (vectors; m s^{-1}, plotted at every four grid point), perturbation potential temperature (thick black lines for 0 K and dotted contours at 0.5-K intervals for negative potential temperature) and simulated reflectivity (solid contours and shading at 10 dBZ interval, starting from 10 dBZ) at z = 250 m at 40, 60, 80, and 100 min of a supercell storm simulation: (a)-(d) truth, EnKF analysis of (e)-(h) experiment Vr; (i)-(l) Zh; and (m)-(p) VrZh.

Figure 6.2 shows the change of ensemble mean forecast and analysis root-mean-square errors (RMSEs) of model state variables during the assimilation cycles of Vr, Zh, and VrZh averaged over model grid points at which the true reflectivity is greater than 10 dBZ. For total number concentrations, RMSEs are expressed in the form of $(Nt_x)^{0.4}$, which is our control value in the filter, where x can be c (cloud water), r (rainwater), i (cloud ice), s (snow aggregate), and h (hail). Under the perfect model assumption, the solid curves in Fig. 6.2 show that the filter can successfully reduce RMSEs with both reflectivity and radial velocity data. These results suggest that the filter is able to develop the reliable error covariance matrix for the increased number of state variables if the forecast model is perfect. The final RMSE levels are comparable to those in Fig. 3.3 performed using a SM scheme, although the error reduction rates are generally slower than those of Fig. 3.3. The RMSEs of u and v drop below 1 m s^{-1} around t = 50 min and converge below 0.7 m s^{-1} at the end, while the RMSE of w reaches about 0.3 m s^{-1}. The RMSEs of mixing ratios are below 0.05 g kg^{-1} for q_c, q_r, q_i, and q_s, about 0.8 g kg^{-1} for q_h, and around 0.26 g kg^{-1} for q_v. For Nt_x, after being converted to the linear domain, the RMSEs at the end are about 4.6×10^5, 0.186, 6.89×10^3, 91.9, and 0.177 for Nt_c, Nt_r, Nt_i, Nt_s, and Nt_h, respectively. These are all several orders of magnitude smaller than their dynamic ranges: they are O(3) , O(4), O(5), O(2), and O(4) smaller for Nt_c, Nt_r, Nt_i, Nt_s, and Nt_h, respectively. The rain water experiences the largest benefit from the assimilation considering both mixing ratio and the total number concentration when a DM scheme is used.

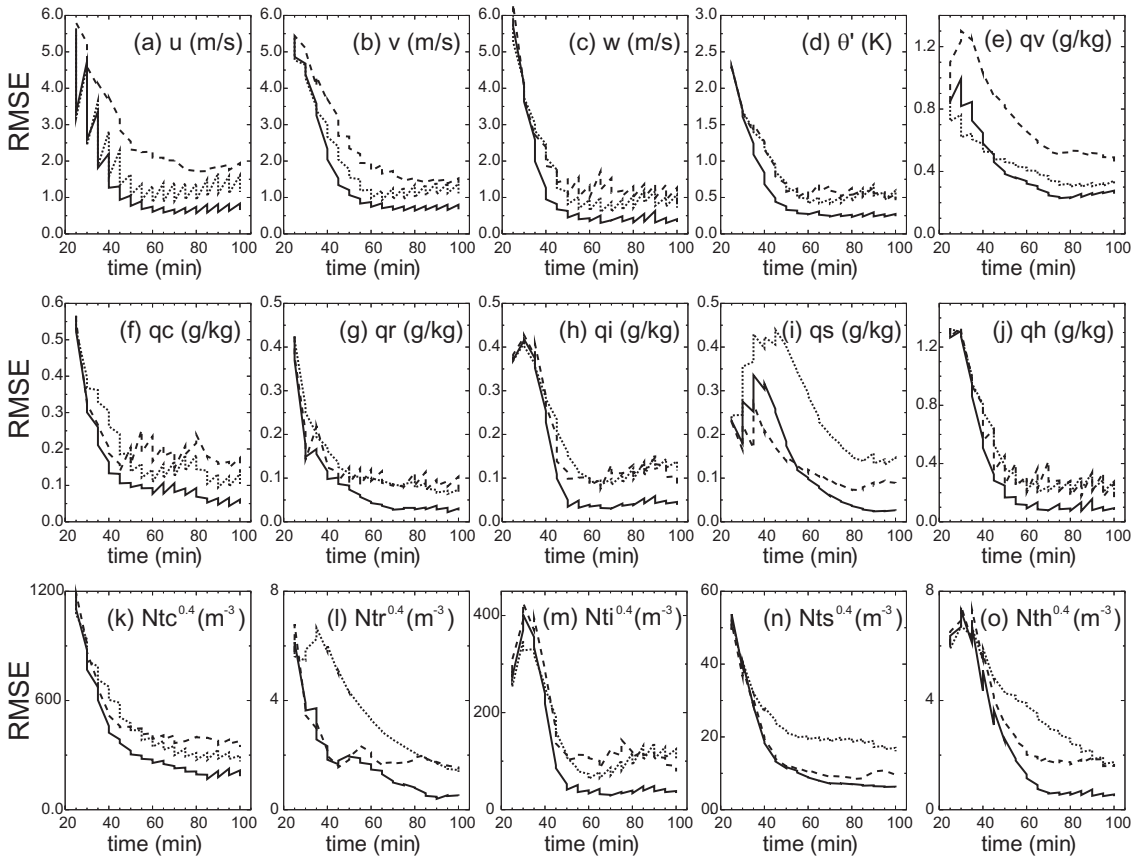


Fig. 6.2. The ensemble mean forecast and analysis RMSEs averaged over points at which the true reflectivity is greater than 10 dBZ for (a) u, (b) v, (c) w, (d) perturbation potential temperature θ', (e) q_v, (f) q_c, (g) q_r, (h) q_i, (i) q_s, (j) q_h, (k) Nt_c, (l) Nt_r, (m) Nt_i, (n) Nt_s, and (o) Nt_h for experiments VrZh (solid, inflation = 20%), Vr (dotted, inflation = 7%), and Zh (dashed, inflation = 20%). The vertical straight-line segments in the curves correspond to the reduction or increase in RMSEs by the data assimilation.

When reflectivity is assimilated alone, the filter is still able to reduce error effectively (Fig. 6.2), but RMSEs at the end of assimilation cycles are larger than those of the corresponding experiment shown in Fig. 3.3. It is reasonable because not all free variables estimated using reflectivity are independent. For example, rain reflectivity is determined by two free variables, q_r and Nt_r. Even though the estimated q_r value is not optimal, the filter would compensate the error in the reflectivity by adjusting Nt_r. As a

result, simulated reflectivity is almost as good as that of VrZh (Fig. 6.1), although the RMSEs of Zh are generally two or three times larger than those of VrZh.

Radial velocity results in smaller RMSEs in wind variables (u, v, and w), q_v, q_c, and Nt_c than reflectivity does when used alone (Fig. 6.2). This is reasonable because radial velocity is a linear function of wind variables. At the end of the assimilation windows, the analysis RMSEs of u and v are around 1.1 m s^{-1} and that of w is about 0.7 m s^{-1}. Cloud ice is directly related to neither reflectivity nor radial velocity; its estimation, therefore, solely depends on the cross-covariance structure of state variables. Reflectivity clearly performs better than V_r for q_s, Nt_r, Nt_s, and Nt_h, which all directly affect the reflectivity.

While comparing results from experiments Zh and Vr, we noticed that the near perfect reflectivity patterns of Zh shown in Figs. 6.1i–l actually have comparable sizes of error at the end of the assimilation window to Vr in q_r, Nt_r, q_h, and Nt_h, which are involved in the reflectivity calculations at the elevation shown in Fig. 6.1. On the other hand, it is hard to infer the truth reflectivity pattern from Fig. 6.1h with similar levels of RMSEs to those of Zh. These results led us to two questions: the first question is why the reflectivity pattern of Vr is noisier than that of Zh with similar levels of RMSE in microphysical variables. First of all, the echo-free region is not included in the RMSE calculations because we are more interested in the error in the storm region. However, spurious reflectivity of Vr in the echo-free region mostly involves low q_x and low Nt_x, which would not have a big impact on the total RMSE calculations even if they are taken into account. This spurious but high reflectivity can be explained by the selectively sensitive reflectivity to the drop size. For example, reflectivity is only

sensitive to the number of large drops, while many small drops have little influence on the reflectivity. Therefore, certain combinations of q_x and Nt_x can lead to high reflectivity even though mixing ratio is low.

The second question is how the simulated reflectivity pattern of Zh can be so good even when RMSEs of relevant state variables are relatively high. The state estimation is essentially an undetermined problem. Reflectivity is nonlinear function of distributions of rain drops, snow aggregate, and hailstones in the atmosphere, while it is only an observation to be used in estimating usually more than one hydrometeor type involved in more than one state variable. The number of hydrometeor types is reduced when one species is dominant in the radar resolution volume. When a SM scheme is used, reflectivity is solely determined by the mixing ratio of interest, and, therefore, it can be retrieved very accurately in this simple case if the observation is error-free and the observation operator is perfect. When a DM scheme is used, the mixing ratio and the total number concentration, which vary independently, are two state variables to be estimated for one hydrometeor type. In this case, the same reflectivity can be obtained using the infinite pair of q_r and Nt_r. When the q_r estimate is not optimal, the filter may fit reflectivity closely by adjusting Nt_r, which essentially increases error in Nt_r. This is illustrated in Fig. 6.3. Figure 6.3 shows estimated q_r and Nt_r at 1-km height for experiments Vr, Zh, and VrZh, along with the truth simulation. At this level, rain water is dominant, and so the estimation problem is simpler. When reflectivity is assimilated alone (Fig. 6.3c), the rain water mixing ratio at the core is overestimated in both left- and right-moving storms compared with truth. Because the filter does not know the truth q_r value, it increases the total number concentration to fit the reflectivity to truth

closely as shown in Fig. 6.1. It is clear that VrZh, and probably Vr at least in the echo region, produce better estimates of q_r and Nt_r than Zh at this level (Figs. 6.3c,d), although it is not obvious in Fig. 6.1. The uncertainties increase when more than one species co-exist in the same resolution volume. This is a good example of how analysis can deteriorate when the problem is undetermined.

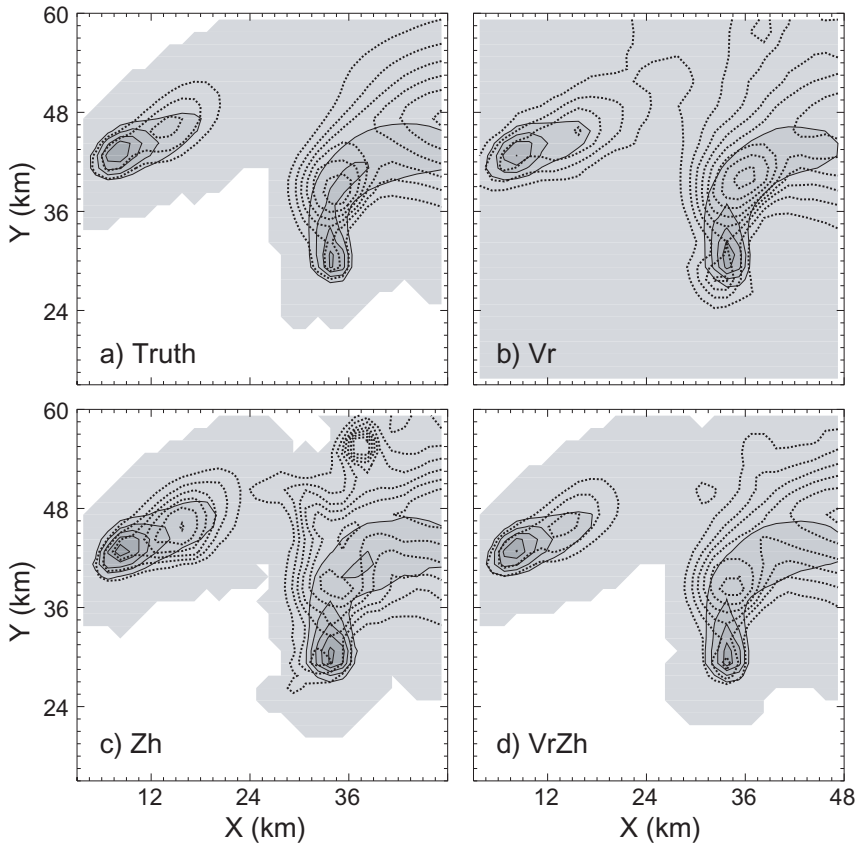


Fig. 6.3. Rain water mixing ratio (solid contours and shading at intervals of 0.5 g kg^{-1}, starting from 0.0 g kg^{-1}) and total rain water number concentration (dotted contours at intervals of 40 m^{-3}, starting from 40 m^{-3}) at z = 1 km and t = 100 min for (a) truth simulation, (b) experiment Vr, (c) Zh, and (d) VrZh.

Figure 6.4 shows the vertical profiles of the RMSEs averaged over points at which the truth reflectivity is greater than 10 dBZ for experiments Vr (dotted), Zh

(dashed), and VrZh (solid) at 100 min. Experiment Vr generally performs better than Zh in u, v, w, q_v, q_c, q_r, and Nt_c at middle to lower levels (Figs. 6.4a–c,e–g,k). Improvements are found at all levels in all state variables when both V_r and Z_H are assimilated compared with experiment Zh and Vr, with the exceptions being u between 1.5- and 3.5-km and q_r between 5.0- and 5.5-km height.

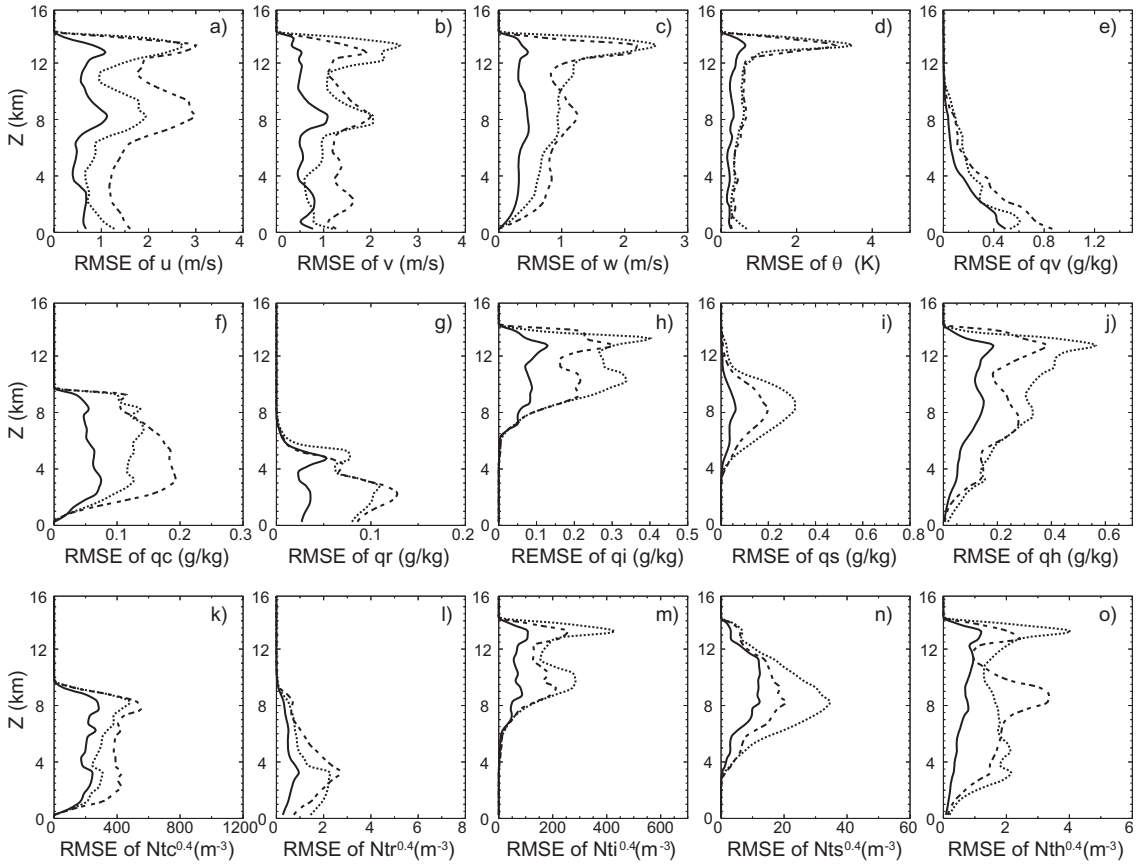


Fig. 6.4. The horizontally averaged ensemble mean analysis RMSEs averaged over points at which the truth reflectivity is greater than 10 dBZ for (a) u, (b) v, (c) w, (d) perturbation potential temperature θ', (e) q_v, (f) q_c, (g) q_r, (h) q_i, (i) q_s, (j) q_h, (k) Nt_c, (l) Nt_r, (m) Nt_i, (n) Nt_s, and (o) Nt_h for experiments VrZh (solid, inflation = 20%), Vr (dotted, inflation = 7%), and Zh (dashed, inflation = 20%) at 100 min.

In the EnKF system, the successful estimation of model state variables depends on the ability of the filter to develop reliable multivariate covariance structures through

assimilation cycles. This makes the EnKF system very attractive because the variables that are not directly related to the observed quantities can be retrieved using the cross-covariance information between the forecast ensemble of a state variable and the prior estimates of observations. Figure 6.5 shows the forecast error correlations between the prior estimate of reflectivity at point x = 38.5 km and z = 8 km and model variables at each grid point in the x–z plane at y = 33 km through the maximum updraft at 100 min.

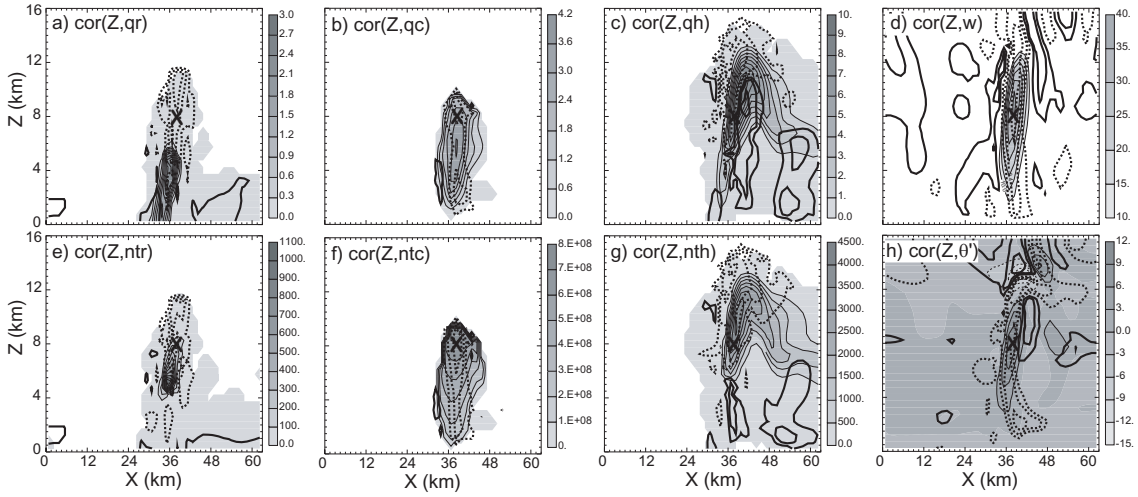


Fig. 6.5. Forecast error correlations (thick solid and dotted contours at interval of 0.2) computed from the ensemble for experiment VrZh at t = 100 min in the x-z plane at y = 33 km, which passes through the updraft core (maximum vertical velocity). Thick solid (dotted) contours represent positive (negative) error correlations between forecast reflectivity Z_H at x = 37.5 km and z = 8 km (indicated by the boldface X) and (a) q_r, (b) q_c, (c) q_h, (d) w, (e) Nt_r, (f)) Nt_c, (g) Nt_h, and (h) θ'. The shading and thin solid contours are model fields from the truth simulation with units of (a)-(c) g kg^{-1}, (d) m s^{-1}, (e)-(g) m^{-3}, and (h) K. Zero correlation lines are suppressed.

Because the observation is located near the strong updraft core, the errors are correlated through much of the deep convection vertically. Compared with the correlation patterns computed using a SM scheme in Fig. 5 of TX05, almost all variables shows the opposite sign of correlation patterns. As seen in Fig. 6.5d, the large

negative correlations between Z_H and w are found in the updraft extended from the surface up to the model top, while this region shows the large positive correlations in Fig. 5b of TX05. This is reasonable considering that a DM scheme is capable of handling an independently varying total number concentration of each hydrometeor type depending on the updraft strength. The supersaturation increases when updraft speeds increase. Then more cloud condensation nuclei (CCN) are activated to decrease the supersaturation and, therefore, Nt_c increases (Rogers and Yau 1989). The number of ice nuclei (IN) to be activated is also a function of supersaturation; the same theory can be applied to the cloud ice. With more drops to grow, the number of rain and ice phase hydrometeors increases. Regarding the size of each drop, increased updraft speed reduces the residence time of drops to experience growing in the atmosphere. This leads to a larger number of small drops and low reflectivity. This is consistent with the correlation patterns shown in Figs 6.5a–g.

Figure 6.5c exhibits a deep layer of a large positive correlation region on the right of the updraft core. This is where large hailstones fall on the edge of the updraft. In the updraft, q_h and Nt_h are negatively correlated with reflectivity; small amounts of hail with the DSD shifted toward the larger sizes in the updraft increase reflectivity in situ, while large amounts of hail with the DSD shifted toward a smaller size decrease it. In the former case, larger mass of hail falls near the updraft because the total mass of hailstones that are too heavy to be advected farther increase. On the other hand, the correlations for Nt_h are still negative in this region (Fig. 6.5g), suggesting that the number of hailstones falling in this region are relatively small.

The correlation pattern for θ' is similar to that of w (Fig. 6.5h). The θ' is positively correlated with mixing ratios because of the latent heat release where the mixing ratio is negatively correlated with reflectivity, as discussed above.

To investigate these physical processes more closely, we compute the correlations between w at the location of Z_H shown in Fig. 6.5 and microphysical quantities discussed above and present in Fig. 6.6. The correlations between w and q_c and Nt_c (Figs. 6.6d,h) show that a stronger updraft produces more cloud drops and more cloud water. It is also shown that Nt_c is more strongly correlated with w than q_c.

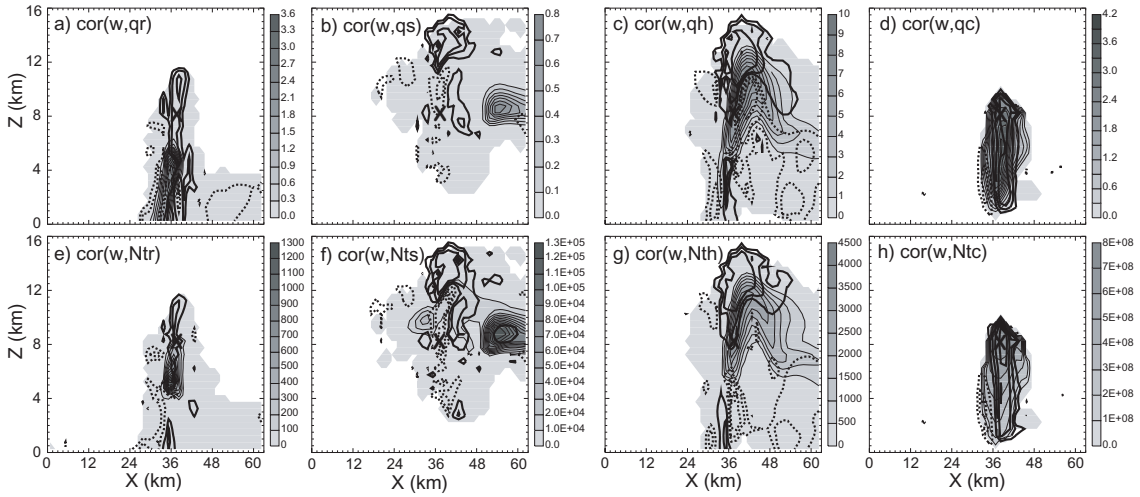


Fig. 6.6. As in Fig. 6.5, but between forecast w and (a) q_r, (b) q_s, (c) q_h, (d) q_c, (e) Nt_r, (f) Nt_s, (g) Nt_h, and (h) Nt_c.

Although stronger updraft transports more rain water to the upper level (Fig. 6.6a), the increased number of rain drops may restrict the increase of reflectivity (Fig. 6.6e). In Fig. 6.6e, the correlation for Nt_r shows that the correlation increases to the upper level. This pattern suggests that more rain drops grow from the increased number of cloud drops along the stronger updraft. q_s and Nt_s are negatively correlated with w,

implying that the production of low density snow is suppressed as w increases. However, the reflectivity change depends on the relative changes in magnitudes of these two fields. If decrease of snow mixing ratio is more rapid than that of the total number concentration of snow, reflectivity would still decrease. The correlation for q_h is weak negative, while the correlation for Nt_h is positive at the location of observation. This is a clear indication of the shifting DSD toward smaller drop size, which lowers reflectivity.

In Fig. 6.5d, the deep layer of negative correlation region is tilted slightly toward the east, with decreasing height, located between x = 36 and 43.5 km in the low level. The region is associated with the negative correlation for q_r and the near-zero correlation with Nt_r (Figs. 6.6a,e). This combination of q_r and Nt_r may be responsible for the negative correlation shown in Fig. 6.5d. However, we would like to point out that the correlations at large distances are not reliable and are not used after applying covariance localization.

6.3.2 Assimilation of polarimetric radar data

In this subsection, we investigate the impact of assimilating polarimetric variables in addition to both reflectivity (everywhere) and radial velocity (Z_H > 10 dBZ) data (Table 6.2). As in section 6.3.1, each experiment name exhibits all observed variables assimilated in the experiment. The same thresholds applied to the polarimetric variables in Chapter 3 are also applied here. They are 0.3 dB and 0.9 ° km^{-1} for Z_{DR} and K_{DP}, respectively. The covariance inflation factor for each experiment is chosen experimentally, and the SDs of Z_{DR} and K_{DP} in the filter are set to 0.2 dB and 0.5 ° km^{-1},

respectively. Experiment VrZh discussed in subsection 6.3.1 is used as the control run in this subsection.

Table 6.2. List of experiments testing the impact of polarimetric variables.

Experiments	Observation(s) assimilated	Covariance inflation
VrZhZdr	V_r, Z_H, and Z_{DR} (Z_{DR} > 0.3 dB)	10%
VrZhKdp	V_r, Z_H, and K_{DP} (K_{DP} > 0.9 ° km^{-1})	15%
VrZhZdrKdp	V_r, Z_H, Z_{DR}, and K_{DP}	15%

Figure 6.7 shows the ensemble mean analysis RMSEs of the data impact experiments listed in Table 6.2 normalized by those of VrZh. A smaller normalized RMSE (NRMSE) suggests that a larger improvement is obtained by assimilating additional polarimetric variables. As shown in Chapter 3, both Z_{DR} and K_{DP} exhibit a degree of positive impact in the later cycles at least after 80 min of model time in most state variables with the exceptions in θ' of VrZhKdp and VrZhZdrKdp, q_v of VrZhKdp, q_s and Nt_s of VrZhZdr, and q_c and Nt_c of VrZhZdr, VrZhKdp, and VrZhZdrKdp, which show negative impact temporarily. Generally, the state variables that are directly related to the polarimetric variables receive larger positive impact from the relatively early cycles (Figs. 6.7g,i–j,l,o), except for Nt_s (Fig. 6.7n). Although q_i and Nt_i are not directly related to the polarimetric variables, they also show relatively large improvements. With the help of polarimetric variables, the NRMSEs stay lower than 1 after 65 min of model time for q_r, q_i, q_h, Nt_r, Nt_i, and Nt_h, but a tendency for error reductions to become smaller in the later assimilation cycles is found in state variables who show large improvement (Figs. 6.7g–j,m,o), except for Nt_r (Fig. 6.7l).

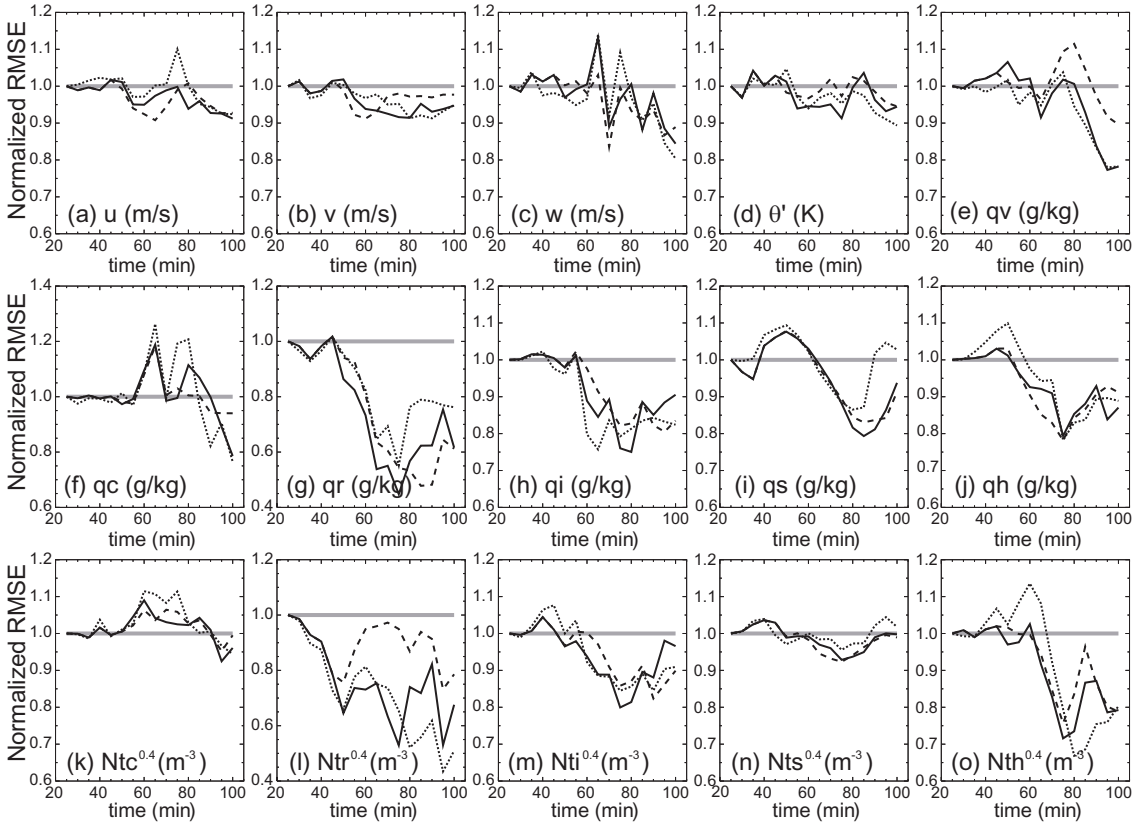


Fig. 6.7. The ensemble mean analysis RMSEs of experiments VrZhZdr (dotted), VrZhKdp (dashed), and VrZhZdrKdp (solid) normalized by those of experiment VrZh. The reference horizontal line at unity is overlaid.

To help assess the gross improvement more easily, the percent improvement over the control experiment averaged over the last five cycles is summarized in Table 6.3. The largest improvement is found in Nt_r in VrZhZdr and q_r in both VrZhKdp and VrZhZdrKdp, which experience approximately 47%, 45%, and 36% improvement, respectively. The improvement is also large in the hail and ice categories.

Table 6.3. The improvement over the experiment VrZh for the experiments VrZhZdr, VrZhKdp, and VrZhZdrKdp averaged over the last five cycles (80-100 min of model time). The improvement is expressed in percentages relative to the control experiment.

Variables	Improvement (%)		
	VrZhZdr	VrZhKdp	VrZhZdrKdp
u	5.1	4.4	6.7
v	7.4	2.6	6.3
w	10.4	9.4	8.0
θ'	6.1	1.5	2.6
p'	-0.3	-2.4	3.1
q_v	15.3	1.1	13.4
q_c	6.6	3.4	2.8
q_r	22.7	45.0	36.3
q_i	17.0	16.5	14.5
q_s	3.4	14.6	15.5
q_h	13.1	10.9	12.6
Nt_c	1.0	-0.1	0.8
Nt_r	47.2	15.2	30.3
Nt_i	11.9	12.6	9.3
Nt_s	-0.6	2.6	2.5
Nt_h	26.6	14.9	19.0
tot	193.0	152.2	183.8

Generally, VrZhZdr produce a better analysis than does VrZhKdp, and the improvement is larger in rain water-, hail-, and cloud ice-related variables, water vapor, and vertical velocity. Among microphysical variables, q_c, q_s, Nt_c, and Nt_s show small or even negative improvement. The negative impact is found in p' when polarimetric variables are assimilated individually. When both Z_{DR} and K_{DP} are assimilated, all model state variables show positive impact, although the overall improvement is slightly smaller than that of VrZhZdr. However, the difference is too small to be

significant. These results generally agree with what were found in Chapters 3 and 4 using SM.

We would like to point out that the current data impact experiments have some limitations to fully examining the impact of the direct assimilation of polarimetric radar data. As discussed in Chapters 3 and 4, polarimetric variables have very little impact when the analysis obtained using conventional data is already very good. Under the perfect model assumption within the OSSE framework, the EnKF system is able to retrieve all model state variables successfully, leaving no room for polarimetric variables to play. This necessitates the use of error-free polarimetric variables while conventional radar data contain errors. Assimilating error-contaminated polarimetric variables can easily produce negative impact where the analysis is already very good because polarimetric variables are more susceptible to observation errors than reflectivity.

We expect to see a larger impact of polarimetric data in real-data situations in which the model error tends to be larger. Reflecticity alone is not sufficient to determine microphysical processes associated with the size-sorting mechanism of a supercell storm revealed by unique polarimetric signatures shown in Chapter 5. The extra information on DSD carried by the polarimetric data may help improve analysis by providing additional constraints.

6.4 Imperfect model experiments

To examine the ability of EnKF system in retrieving model state variables in the presence of model error, we designed five experiments listed in Table 6.4. The first set

of experiments consists of VrZh_ptrα_IM and VrZh_cstα_IM, and the model error originates from the incorrect DSD shape parameter α for rain and hail, while the observation operater is assumed perfect. In the experiment names, _ptrα_ stands for "perturbed α" and _cstα_ for "constant α." Suffix "_IM" denotes "imperfect forecast model experiment." In the experiment with perturbed α, α_r increases from -1 to 3 at a constant interval of 0.05, while α_h decreases from 3 to 1 at intervals of 0.025 for each member, where the correct α value is 0 for the exponential DSD. Variables α_r and α_h are set to 3 and 2, respectively, in VrZh_cstα_IM. Each experiment assimilates both reflectivity and radial velocity with the covariance inflation factor of 20%.

In the second set of experiments, both forecast model error and the observation operator error are taken into account. The truth simulation is created using the TM scheme of MY05, which allows concentration, shape, and slope parameters to vary freely in time and space. In this way, the incorrect DSD shape parameter is a source of the error in both the forecast model and the observation operator. Observations are produced using the polarimetric radar simulator developed in Chapter 5 using the T-matrix method for all species employing a different axis ratio for rain. Therefore, observation operator involves another source of error: misrepresentation of the scattering properties. Three experiments are performed using the same configuration presented in Table 6.1, but with new observations and indicated by "_obs" in the experiment names. Polarimetric variables are not assimilated because of their small impact to the state estimation in OSSEs based on the results in section 6.3.2.

Table 6.4. List of data assimilation experiment in the presence of forecast model error.

Errors in	Experiments	Incorrect values	truth
incorrect shape parameter in the forecast model	VrZh_ptrα_IM	α_r = -1~3, α_h = 3~1	α_r = 0, α_h = 0
	VrZh_cstα_IM	α_r = 3, α_h = 2	
incorrect DSDs in both forecast model and observation operator, misrepresentation of the scattering properties in the observation operator	Vr_obs_IM	Exponential distribution, Rayleigh approx. for ice species	Gamma DSDs, Mie scattering for all species
	Zh_obs_IM		
	VrZh_obs_IM		

6.4.1 Forecast model error

Figure 6.8 shows the ensemble mean forecast and analysis RMSEs of model state variables during the assimilation cycles of experiments VrZh_cstα_IM and VrZh_ptrα_IM, which test the impact of the model error when the shape parameters are set incorrectly in the prediction model. The RMSEs show rapid error reduction in the early few cycles whose rates are almost the same as those of the control experiment's. When each ensemble member uses the same incorrect α values in VrZh_cstα_IM (thick solid gray), the error reductions generally become slower than those of the control between 40 to 60 min of model time, and the RMSEs of state variables stay higher than those of the control experiment, except for q_v, which show almost the same RMSEs during the whole assimilation period. The model state variables directly related to the incorrect model parameters show far larger errors than rest of variables: the errors in q_r, Nt_r, q_h, and Nt_h estimates are around 135%–500% larger than those of the control experiment, while other variables show 10%–110% larger errors.

The RMSEs of u, v, and w estimates are less than 1 m s^{-1}, and those of q_c, q_i, and q_s are less than 0.1 g kg^{-1} with that of θ' being less than 0.5 K in VrZh_cstα_IM. The RMSEs of Nt_c, Nt_i, and Nt_s are also 2–4 orders of magnitude smaller than their respective dynamic ranges. Considering the relatively low errors in the later cycles after the RMSEs stabilize, 10%-110% errors may be insignificant. This result suggests that the impact of the incorrect shape parameters is rather small to the model state variables that are not directly related to them.

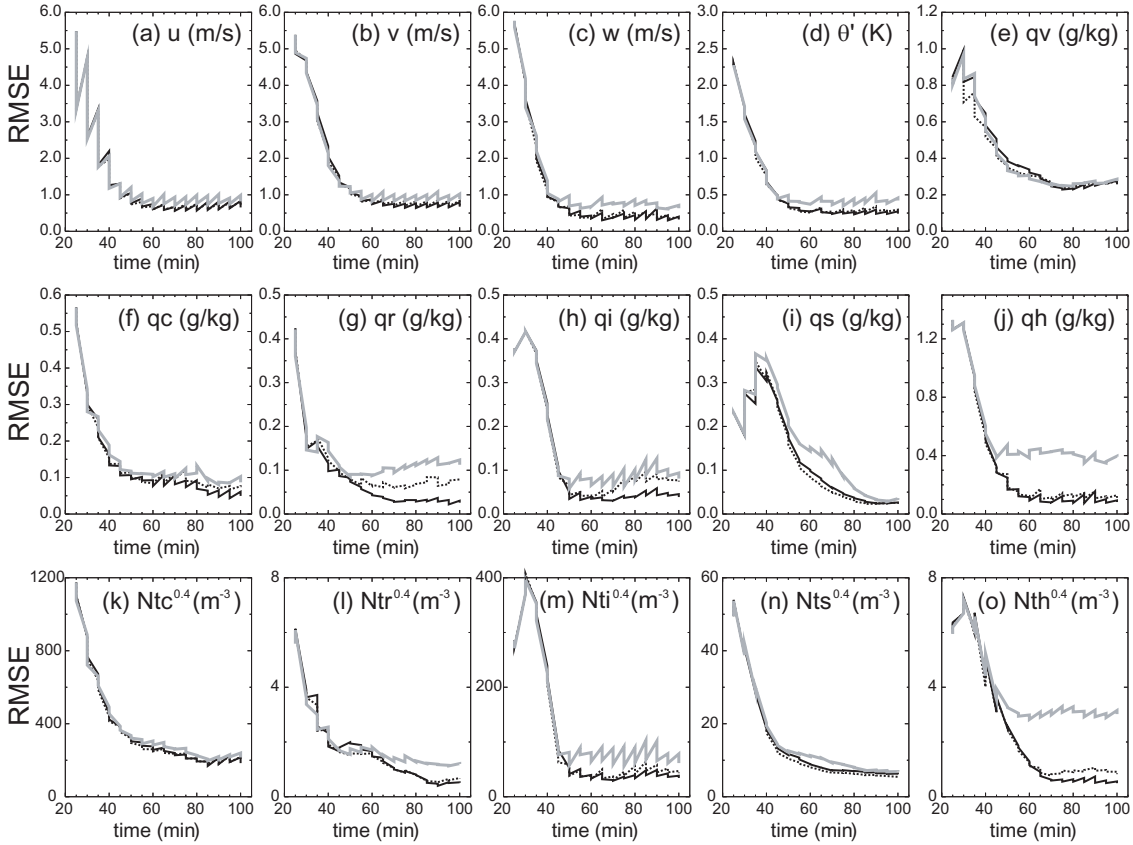


Fig. 6.8. As in Fig. 6.2, but for experiments VrZh_ptrα_IM (thin dotted black) and VrZh_cstα_IM (thick solid gray).The RMSEs of the control run, which is VrZh with $\alpha_r = \alpha_h = 0$, are shown in thin solid black for comparison.

When each member is assigned an α value perturbed in a certain range, the state estimation is significantly improved (dotted line in Fig. 6.8). The RMSEs of most state variables are now almost as low as those of the control experiment VrZh (black solid line in Fig. 6.8), with the exception of q_r, and q_i. It is interesting that q_h and Nt_h can be estimated accurately, although α_h is perturbed between 1 and 3, where the true α_h is 0, while the RMSEs of q_r, and q_i are about two times larger than those of the control.

Several past studies showed that employing different parameterization schemes in different ensemble members improves analysis (e.g., Meng and Zhang 2007, 2008; Zhang and Snyder 2007). When the perturbed α_r and α_h are used in different ensemble members, the ensemble spreads of q_r, Nt_r, q_h, and Nt_h increase compared with those of control experiment's. In the later cycles, the ensemble spreads of Nt_r, q_h, and Nt_h are at almost the same level of their RMSEs, while those in VrZh_cstα_IM are even lower than their corresponding spreads of VrZh (not shown). It can be inferred from these results that the model error arose from the uncertain parameters can partially be accounted for by using perturbed parameters in the ensemble.

6.4.2 Forecast model and observation operator errors

In this subsection, we explore the ability of the EnKF system in producing an accurate analysis in the presence of observation operator error in addition to the forecast model error. This could be the OSSE scenario closest to the real data case. The model state variables produced using a TM scheme are ingested as an input to the radar emulator to create simulated observations, which are independent of the observation

operator used in the data assimilation system. The RMSEs of the state variables shown in Fig. 6.9 are calculated with respect to the true simulation employing the TM scheme.

Although reflectivity data alone can noticeably reduce the RMSEs errors in most state variables (Fig. 6.9), the RMSEs at the end of the assimilation window are higher than those obtained in the perfect model scenarios (Fig. 6.2). For example, the RMSEs of u and v are about 3 m s^{-1}, that of w is around 2 m s^{-1}, and mixing ratios are all greater than 0.1 g kg^{-1}. A tendency of rapid RMSE decrease in the first 4–6 cycles, followed by error increase, is found in many of the microphysical variables, including q_c, q_r, q_i, Nt_c, Nt_r, Nt_i, Nt_s, and Nt_h, which is the indication of filter divergence. In the current OSSE framework, only microphysical variables are associated with observation operator errors, while dynamic and thermodynamic variables are not directly involved in them. As a result, the RMSEs of microphysical variables of experiment Zh_obs_IM in the later cycles are higher than those of Vr_obs_IM, except for q_s, where V_r alone is very poor in estimating q_s. (Figs. 6.9e–h,j–o).

To prevent filter divergence, we performed the same experiment with a larger inflation factor, 25%, and found that the average RMSE is reduced about 25% when it is averaged over state variables over the last five cycles. However, this is not practical because the optimal covariance inflation factor depends on many other factors and is not known in advance in the real-time forecast.

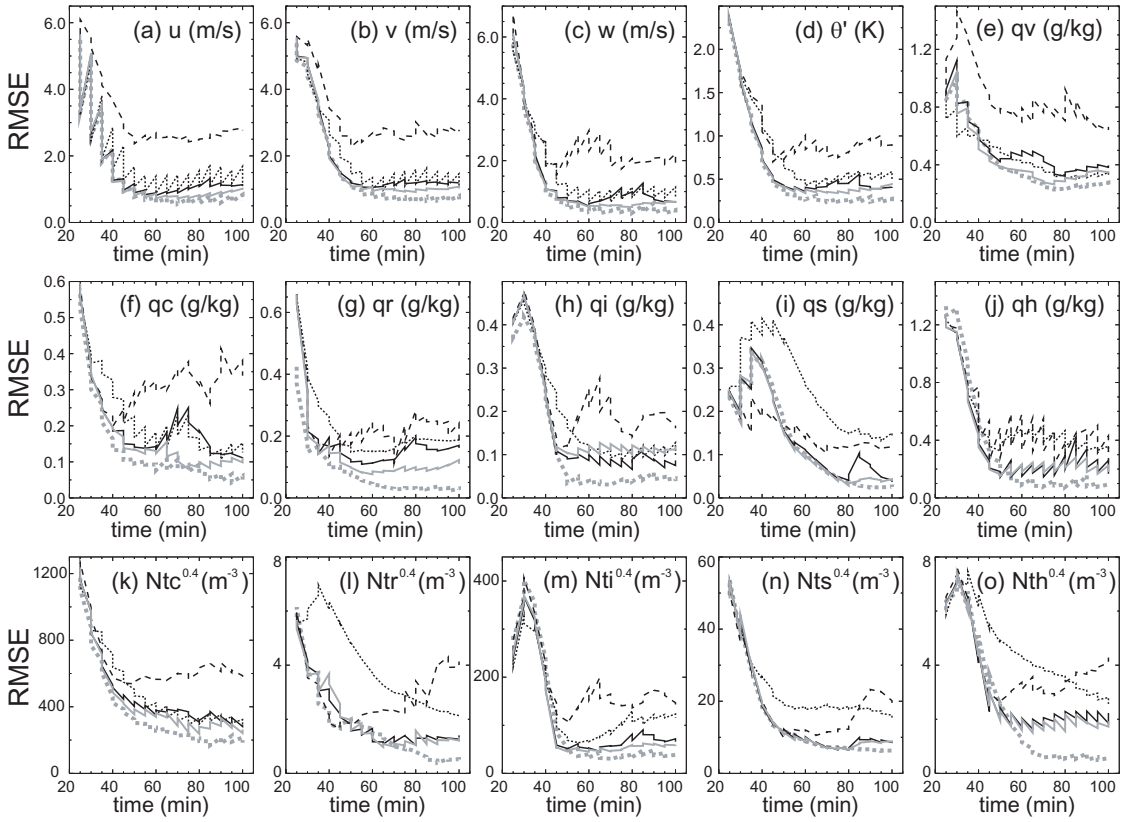


Fig. 6.9. As in Fig. 6.2, but for experiments VrZh_obs_IM (solid), Vr_obs_IM (dotted), and Zh_obs_IM (dashed). The RMSEs are computed with respect to the true simulation performed using three-moment microphysics scheme. The RMSEs of experiments VrZh (thick dotted gray) and VrZh_ptrα_obs_IM (thin solid gray) are shown for comparison.

Experiment Vr_obs_IM (dotted line in Fig. 6.9) produced a much better analysis than did Zh_obs_IM in general. In the early assimilation cycles, V_r is more beneficial than Z_H to u and v, while Z_H is more beneficial than V_r to most microphysical variables except q_v, q_h, Nt_i. Although the error reduction rates of q_c, q_r, q_i, Nt_c, Nt_r, Nt_s, and Nt_h in early cycles of Vr_obs_IM are slower than those of Zh_obs_IM, a trend of error increase in the later cycles is weak or not found in these variables. Considering the fact that the wind variables are not associated with observation operator error, good performance of V_r compared with Z_H is reasonable.

When Z_H is assimilated in addition to V_r, the positive impact of assimilating additional Z_H data is clear in θ', q_r, q_i, q_s, q_h, Nt_r, Nt_i, Nt_s, and Nt_h (Figs. 6.9 d,g–j,l–o). Because of the model error, the overall RMSE levels of state in VrZh_obs_IM (thin solid black) are higher than those of VrZh (thick dotted gray). The larger deterioration of state estimation is found in w, q_h, Nt_r, Nt_h, and q_r, which show between 150%-500% larger errors than those in VrZh. This result suggests that the model microphysical processes are more sensitive to correctly modeling DSDs of rain and hail than they are cloud water, cloud ice, and snow aggregates. The perturbed α used in the ensemble of experiment VrZh_ptrα_obs_IM helps further reduce the overall analysis error (thin solid gray in Fig. 6.9), although the improvement is relatively small. In this experiment, α_r increases from -1 to 3, and α_h decreases from 4 to 0 at a constant interval of 0.05 for each member.

At 40 min or after four analysis cycles, the basic pattern of reflectivity and cold pool structure of VrZh_obs_IM is retrieved reasonably well, but the intensity of reflectivity and cold pool strength of the analyzed storm is weaker than that of truth (Fig. 6.10a,e). The low-level horizontal perturbation wind pattern also appears reasonable if the mean southerly wind is subtracted from the perturbation wind of VrZh_obs_IM. The stronger southerly wind seems to be partly due to the lack of Vr observation in the echo-free region so that the wind variables could not be corrected there. The wind errors are gradually corrected during subsequent analysis cycles by 100 min, while cold pool strength and the intensity of reflectivity are consistently underestimated throughout assimilation cycles.

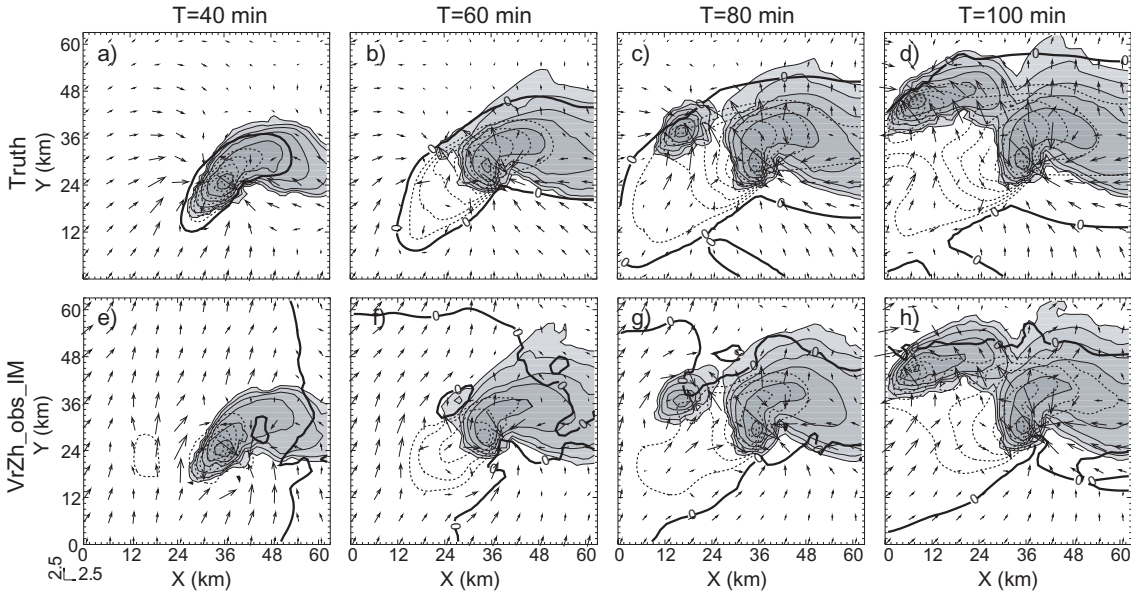


Fig. 6.10. As in Fig. 6.1, but for a supercell storm simulation: (a)-(d) truth with TM of MY05 and (e)-(h) EnKF analysis of experiment VrZh_obs_IM. Reflectivity is computed using the radar simulator developed in Chapter 5.

6.5 Summary and conclusions

When both V_r and Z_H are assimilated, the EnSRF system employing a two-moment (DM) microphysics scheme is able to accurately estimate the increased number of model state variables of a simulated supercell storm with a sufficiently large ensemble under the perfect model assumptions. However, a significant deterioration of the analysis was encountered when Z_H is assimilated alone due to the solution non-uniqueness of this undetermined problem. Where the reflectivity is determined by two independent variables, mixing ratio q_x, and the total number concentration Nt_x for hydrometeor type x, not-optimally estimated q_x increases the error in Nt_x estimate in a way in which they compensate the reflectivity difference due to the errors in q_x. As a result, simulated reflectivity computed from estimated q_x and Nt_x fits observed

reflectivity very closely while failing to reduce RMSEs of q_x and Nt_x. The results, which show that the model states can be retrieved more accurately with both V_r and Z_H than with any one of them, suggest that additional observations will likely improve the well-posedness of this problem.

The impact of additional error-free polarimetric variables on the quality of storm analysis is generally positive, though rather small under the perfect model assumption, partly because the analysis obtained using V_r and Z_H is already accurate, leaving little room for the polarimetric variables to play a role. A larger impact is expected in a real-data situation in which the model error is larger, and reflecticity alone is not sufficient to fully determine the microphysical variables accounting for the entire spectrum of variability in drop size distributions (DSDs). The extra information on DSDs carried by the polarimetric data may help improve analysis by providing additional constraints in this case.

The forecast error correlations between model state variables and simulated reflectivity taken from the updraft core and among model states computed from the ensemble show that the simulation employing the DM scheme produces a more realistic pattern compared with those using the single-moment (SM) microphysics scheme. The signs of correlation are opposite for these two in the regions where significant correlations exist. The opposite sign of the SM run is mainly because a SM scheme is not capable of handling the DSD shift toward larger or smaller drop sizes because of size sorting, while a DM scheme is at least able to account for this, although it is still incapable of addressing the change in shape of the DSDs.

The filter performance in the presence of forecast model error with/without the observation operator error is also investigated using V_r and Z_H data. The results show that model error can noticeably deteriorate the quality of state estimates of microphysical variables when the prediction model is subject to the microphysical parameterization error due to the assumed incorrect DSD shapes and when the observation operator is suffering from incorrect DSD shapes and a misrepresentation of the scattering properties. Nevertheless, the retrieved dynamic, thermodynamic, and microphysical fields are still reasonably good. As the observation operator for radial velocity is not involved in any errors in the current OSSE framework, retrieved wind variables are relatively accurate when V_r is assimilated with/without Z_H. However, Z_H is necessary to suppress spurious cells in the echo-free regions. When the shape parameter α of a certain hydrometeor type was perturbed and used in different ensemble members, the ensemble spreads of state variables of corresponding type were increased, and overall analysis was improved accordingly.

In the real-data scenario, we may encounter severe model errors arising from many sources, such as insufficient resolution in time and space, misrepresentation of physical processes, and the non-physical model boundaries. These model errors can dominate the error growth during the assimilation cycles and lead to filter divergence (Houtekamer et al. 2005). To account for model error, various methods have been proposed, such as additive error model and simultaneous model parameter estimations. However, using a more accurate parameterization scheme could be a more direct, better solution. The previous study of Dawson et al. (2007), a comparison of simulated polarimetric signatures using SM and DM schemes presented in Chapter 5, and the

correlation analysis performed in this chapter all support the fact that a DM scheme may describe storm microphysics and kinematics more realistically than does a SM scheme. Motivated by these results, the assimilation of real data using EnSRF and a DM scheme is underway. Effective ways to deal with other source of model error will be studied along the way.

Chapter 7 Summary and Future work

7.1 Summary

Radar data are indirect measurements of the model state. In modern data assimilation (DA) systems using variational data or ensemble Kalman filter (EnKF) techniques, the direct assimilation of such indirect measurements require forward observation operators that convert model state variables into the form of observations. These observation operators often involve more than one state variable and can be very complex and nonlinear. While the formulas that link all available state variables and observations for reflectivity and radial velocity can be found in the literature, those for polarimetric radar (PR) data either focus on a single-phase hydrometeor or are too complex and expensive for DA purposes. Moreover, melting ice particles are often ignored or calculated by using a very simple linear interpolation between two temperature layers chosen on an ad hoc basis even for the calculation of conventional radar reflectivity, which significantly affects reflectivity calculated within the melting layer. In this dissertation, we first developed a comprehensive radar simulator for PR variables at the S band of wavelength 10.7 cm. The PR variables include reflectivity at horizontal (Z_H) and vertical (Z_V) polarizations, differential reflectivity Z_{DR}, reflectivity difference Z_{dp}, specific differential phase K_{DP}, and the cross-correlation coefficient ρ_{hv}. The transition matrix (T-matrix) method was used for the scattering calculations of raindrops, and Rayleigh scattering approximation was applied to snow and hail particles. The PR variables are expressed as functions of the hydrometeor mixing ratios as well as their densities and drop size distribution (DSD) parameters. This simulator served as a

testbed for developing and testing forward observation operators of PR variables. New features include the continuous melting ice (snow/hail) model that allows varying dielectric constants and density due to melting and error models accounting for correlated and uncorrelated error components for reflectivity at horizontal and vertical polarizations.

Those forward operators were incorporated into the Advanced Regional Prediction System (ARPS) ensemble square-root Kalman filter (EnSRF) data assimilation system and used to assess the impact of assimilating additional PR variables. Results show that storm analysis is improved when PR variables are assimilated in addition to Z_H or in addition to both Z_H and radial velocity V_r. Positive impact is largest when Z_{DR}, Z_{dp}, and K_{DP} are assimilated all together. Improvement is generally larger in vertical velocity, water vapor, and rainwater mixing ratios. The rainwater field provides the most benefits, while the impacts on horizontal wind components and snow mixing ratio are smaller. Improvement is found at all model levels even though the PR data, after the application of thresholds, are mostly limited to the lower levels. Among Z_{DR}, Z_{dp}, and K_{DP}, Z_{DR} is found to have the largest positive impact on the analysis. It is suggested that Z_{DR} provides more independent information than the other variables. The impact of PR data is also expected to be larger when they are used to retrieve drop size distribution parameters. PR data thresholding prior to assimilation is found to be necessary to minimize the impact of noise.

Encouraged by the result of data impact experiments, we applied PR data to the simultaneous estimation of five uncertain DSD-related microphysical parameters which are also involved in the observation operators. These parameters are the intercept

parameters of rain, snow, and hail (n_{0R}, n_{0S}, and n_{0H}, respectively) and the bulk densities of snow (ρ_S) and hail (ρ_H), and they are estimated individually or collectively using EnSRF. Z_{DR}, K_{DP}, and Z_H are used individually or in combination for parameter estimation, while V_r and Z_H are used for state estimation. In these experiments, the latest estimated parameter values are used in the forecast model and observation operators in the subsequent cycle.

Both single- and five-parameter estimation experiments reveal difficulties in estimating certain parameters in the presence of observation operator error using Z_H alone. It is found that the PR data are more helpful when the parameter estimation is not very successful with Z_H only. Between Z_{DR} and K_{DP}, K_{DP} is found to have a larger positive impact on parameter estimation, while Z_{DR} is more useful in the estimation of n_{0H}. In the five-parameter estimation, the filter fails to recover snow-related parameters n_{0S} and ρ_S with or without PR data, probably because of the combined effects of mis-specification of the amount of correction made to the forecast fields due to observation operator error and the insensitivity of model response. When snow-related parameters are not estimated but are kept at their initial wrong guesses, both the estimation of the other three parameters and the stimation of state are significantly improved and the positive impact of PR data is larger than when all five parameters are estimated.

The parameter estimation experiments discussed in Chapter 4 show promising results, implying that additional PR measurements, which provide microphysics and DSD information, can help alleviate solution non-uniqueness of the sometimes ill-posed inverse problem. The fact that PR data can help in the estimation of intercept parameters also suggests that they might help the estimation of total number

concentrations predicted with higher-moment microphysics schemes because they contain similar information. Very recent studies by Dawson et al. (2007) suggest that supercell thunderstorms with more realistic reflectivity structure and cold pool strength can be obtained with a high horizontal resolution (1 km or smaller grid spacing) and multi-moment microphysics schemes of Milbrandt and Yau (2005b; 2005a), with the most improvement being achieved when moving from a single moment (SM) to a two-moment (DM) microphysics scheme.

Evaluating the performance of a DM scheme prior to use is necessary but challenging because detailed microphysical observations are very limited. One way to validate the model forecast is compare synthetic observations simulated from model output using an observation simulator with real observations. For verification purposes, the accuracy of the simulator is more important than its efficiency. Therefore, we developed another more complex PR data simulator based on rigorous scattering calculation using the T-matrix method for all hydrometeor types in Chapter 5. This emulator is general, can be applied to any weather radar frequency band, and can take as input the prognostic variables of NWP model simulations predicting single, double, or triple moments of microphysics DSDs.

This complex emulator was applied to an idealized supercell storm simulated using a DM scheme assuming 10.7 cm wavelength. The simulation results show that realistic polarimetric signatures in terms of general location, shape, and strength can be reproduced when a DM microphysics scheme is employed. These signatures include Z_{DR} and K_{DP} columns; Z_{DR} arc, mid-level Z_{DR}, and ρ_{hv} rings; and hail signature. We compared this simulation to one employing a single-moment microphysics scheme and

found that signatures including the Z_{DR} arc and mid-level Z_{DR} and ρ_{hv} rings cannot be reproduced with the single moment scheme; these signatures are closely linked to precipitation particle size sorting. Such emulations suggest that the higher-moment schemes better represent the related microphysical processes.

In Chapter 6, we explored the capability of the EnSRF system in estimating state variables associated with a DM scheme that includes not only the water/ice mixing ratios (the third moment of DSD) but also the total number concentrations (zeroth moment of DSD). The results show that these state variables can be accurately estimated using both V_r and Z_H observations with a perfect prediction model. In this case, additional polarimetric variables have a small and generally positive impact on the quality of analysis, partly because the analysis obtained using V_r and Z_H is already very good. Imperfect model experiments with forecast model error and with/without observation operator error were also performed to test the filter performance. Two types of model errors were considered: microphysical parameterization error due to incorrectly assumed DSD shape and a misrepresentation of the scattering properties of hydrometeors. The results showed that model error can noticeably deteriorate the estimates of microphysical state variables. Perturbing the shape parameter α of microphysical DSDs and using them in different ensemble members was found to improve overall analysis; doing so increases the ensemble spread of the state variables directly related to those species.

7.2 Future work

As this is the first study to directly assimilate PR data, the potential of PR data for convective storm analysis in the EnKF system is primarily explored in an OSSE framework in this study. Our OSSE experiments show that polarimetric data can provide extra information about storm microphysical processes and the evolution of DSDs. The results also indicate that double-moment microphysics schemes can produce much more relistic simulations of supercell thunderstorms than the most commonly used single-moment schemes. Based on these OSSE results, we plan to extend our EnSRF analysis to a long-lived tornadic thunderstorm case on May 29-30, 2004, employing a double-moment scheme. A previous study of the same case using SM showed that the analyzed V_r and Z_H match observations well but the predicted storm is generally weaker and moves faster than the observed storm (Tong 2006).

Preliminary experiments have been performed with different filter configurations to help design the assimilation system and investigate potential issues that may be encountered when using a DM microphysics scheme. The preliminary results using a DM scheme reveal severe flow divergence problems in some parts of the storm, such as at the southern edge of the forward flank downdraft region. One example of an analyzed reflectivity field is shown in Fig 7.1 as compared to the observed one. Briefly, 60 ensemble members are initialized at 0000 UTC on 30 May. V_r and Z_H data of Oklahoma City WSR-88D radar (KTLX) and V_r of the Vance Air Force radar (KVNX) are assimilated at 0010, 0020, 0040, and 0100 UTC in this experiment. In Fig. 7.1b, it is clear that a noticeable portion of storm is missing south of the main storm [the reference line at y = 32 km is overlaid to help with visual comparison] and west of the small

storm to the north of the main one. These areas are associated with regions where sharp gradients of reflectivity are observed. It seems that the filter experiences difficulty in retrieving mixing ratios and total number contentraions where the reflectivity abruptly decreases. The storm exhibits a more severe loss of echoes near sharp reflectivity gradient at 0100 UTC (Fig. 7.1d). The smallest storm, located at the north boundary between x = 64 and 80 km in Fig. 7.1c, is almost completely missing in Fig. 7.1d.

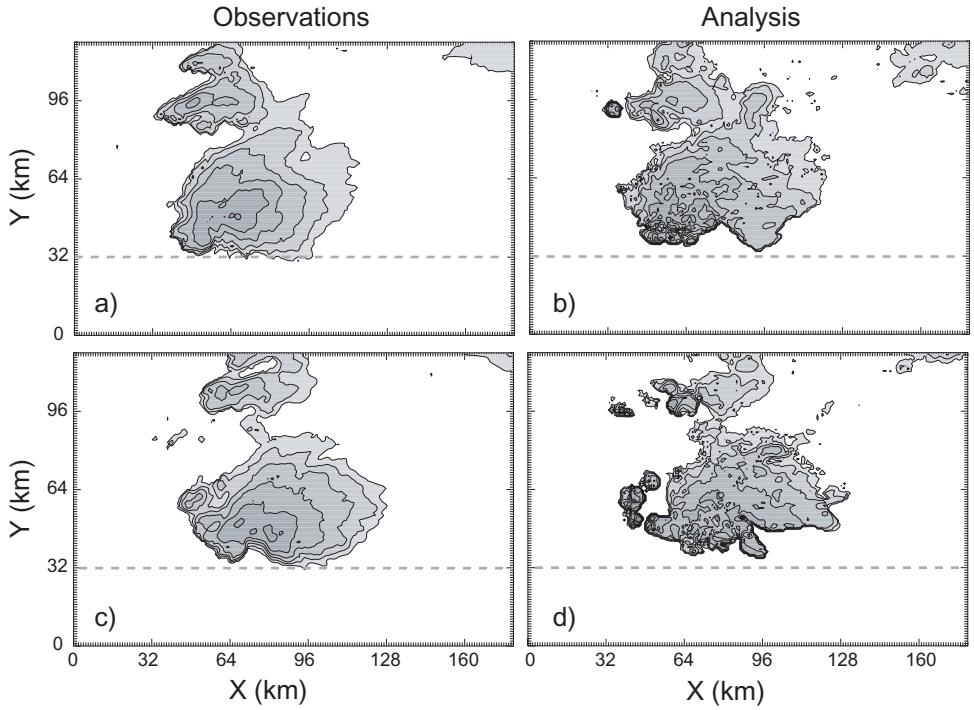


Fig. 7.1. Reflectivity (solid contours and shading at 10 dBZ interval, starting from 10 dBZ) at 1.25° elevation at (a),(b) 0040 UTC and (c),(d) 0100 UTC from the KTLX's perspective for (a),(c) observations and (b),(d) ensemble mean analysis.

The analysis intervals were significantly increased in the current experiment compared with previous studies because the ensemble spreads of a prior estimate of reflectivity could not grow enough in 5 min. This should be reduced with appropriate techniques to account for the loss of spread. The covariance inflation factor was

experimentally determined as 60%; the storm more quickly dissipates or diverges from the observations with smaller covariance inflation factors. However, too large covariance inflation is undesirable because it induces imbalance among ensemble members (Anderson 2001). The large inflation factor creates another problem with the current application: discontinuity in the fields at the cut-off radius in which the inflation is applied only to the grid points that are directly influenced by the observations. The total number concentrations was never updated to prevent fast shrinking of echoes. During assimilation, over-adjustment is common, especially in the early stage of assimilation when errors in the state variables are high. With the DM scheme, when either the mixing ratio or the number concentration becomes negative at a grid point for a particular species, both are set to zero. Updating only the mixing ratio reducing the chances of this happening, but doing so is not necessarily the right solution. One possible remedy for this could be a log transform of those values during assimilation.

The model bias is another source of storm dissipation. The echo is suppressed in some part of storm, such as the southern part of FFD, in many ensemble members during the forward model integration. Covariance inflation is not effective in such a case because the spread is still small after the inflation is applied. One possible solution is to add positive noise within the observed echo region before or after assimilation. In this case, the error should be carefully designed so as not to disturb the balance between the mixing ratio and the total number concentration. Clipping at predefined thresholds applied to mixing ratio and total number concentration during forward model integration could also contribute to the diminishing echo. Reducing the thresholds might help to alleviate this problem.

In the EnKF system, mixing ratio and total number concentration are estimated independently; therefore, the balance between them is not guaranteed in the analysis unless the covariance is very reliable. Covariance inflation introduced to account for model error can cause other problems if overdone. To prevent filter divergence and effectively reduce small-scale perturbations within the echoes, better solutions have to be found. Much research is still needed in this area.

References

Aksoy, A., F. Zhang, and J. W. Nielsen-Gammon, 2006a: Ensemble-based simutaneous state and parameter estimation in a two-dimensional sea breeze model. *Mon. Wea. Rev.*, **134,** 2951-2970.

Aksoy, A., F. Zhang, and J. W. Nielsen-Gammon, 2006b: Ensemble-based simultaneous state and parameter estimation with MM5. *Geophy. Res. Letters*, **33,** L12801, doi:10.1029/2006GL026186.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129,** 2884-2903.

Annan, J. D. and J. C. Hargreaves, 2004: Efficient parameter estimation for a highly chaotic system. *Tellus*, **56A,** 520-526.

Annan, J. D., J. C. Hargreaves, N. R. Edwards, and R. Marsh, 2005a: Parameter estimation in an intermediate complexity earth system model using an ensemble Kalman filter. *Ocean Modelling*, **8,** 135-154.

Annan, J. D., D. J. Lunt, J. C. Hargreaves, and P. J. Valdes, 2005b: Parameter estimation in an atmospheric GCM using the ensemble Kalman filter. *Nonlinear Processes in Geophysics*, **12,** 363-371.

Aydin, K. and Y. Zhao, 1990: A computational study of polarimetric radar observables in hail. *IEEE Trans. Geosci. Remote Sens.*, **28,** 412-422.

Balakrishnan, N., D. Zrnic, J. Goldhirsh, and J. Rowland, 1989: Comparison of simulated rain rates from disdrometer data employing polarimetric radar algorithms. *J. Atmos. Ocean. Tech.*, **6,** 476-486.

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

Beard, K. V., R. J. Kubesh, and H. T. Ochs, 1991: Laboratory measurements of small raindrop distortion. Part I: Axis ratios and fall behavior. *J. Atmos. Sci.*, **48,** 698-710.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129,** 420.

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

Brandes, E. A., G. Zhang, and J. Vivekanandan, 2004a: Drop size distribution retrieval with polarimetric radar: Model and application. *Journal of Applied Meteorology*, **43,** 461-475.

Brandes, E. A., G. Zhang, and J. Vivekanandan, 2004b: Comparison of polarimetric radar drop size distribution retrieval algorithms. *Journal of Atmospheric and Oceanic Technology*, **21,** 584-598.

Brandes, E. A., J. Vivekanandan, J. D. Tuttle, and C. J. Kessinger, 1995: A study of thunderstorm microphysics with multiparameter radar and aircraft observations. *Mon. Wea. Rev.*, **123,** 3129-3143.

Brandes, E. A., K. Ikwda, G. Zhang, M. Schonhuber, and R. M. Rasmussen, 2007: A statistical and physical description of hydrometeor distributions in Colorado snowstorm using a video disdrometer. *J. Appl. Meteor. Climatol.*, **46,** 634-650.

Bringi, V. N. and V. Chandrasekar, 2001: *Polarimetric Doppler Weather Radar*. Cambridge, 636 pp.

Bringi, V. N., J. Vivekanandan, and J. D. Tuttle, 1986a: Multiparameter radar measurements in Colorado convective storms. Part II: Hail detection studies. *J. Atmos. Sci.*, **43,** 2564-2577.

Bringi, V. N., R. M. Rasmussen, and J. Vivekanandan, 1986b: Multiparameter radar measurements in Colorado convective storms. Part I: Graupel melting studies. *J. Atmos. Sci.*, **43,** 2545-2563.

Brotzge, J. A., K. Brewster, B. Johnson, B. Philips, M. Preston, D. Westbrook, and M. Zink, 2005: CASA's first testbed: Integrated project #1 (IP1). Preprints, 32nd Conf. Radar Meteor., Albuquerque, New Mexico, Amer. Meteor. Soc., CDROM 14R.2.

Browning, K. A., 1964: Airflow and precipitation trajectories within severe local storms which travel to the right of the mean winds. *J. Atmos. Sci.*, **21,** 634-639.

Browning, K. A. and R. J. Donaldson, 1963: Airflow and structure of a tornadic storm. *J. Atmos. Sci.*, **20,** 533-545.

Burgers, G., P. J. v. Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, **126,** 1719-1724.

Capsoni, C., M. D'Amico, and R. Nebuloni, 2001: A Multiparameter polarimetric radar simulator. *Journal of Atmospheric and Oceanic Technology*, **18,** 1799-1809.

Caumont, O., V. Ducrocq, G. Delrieu, M. Gosset, J.-P. Pinty, J. P. d. Chatelet, H. Andrieu, Y. Lemaitre, and G. Scialom, 2006: A radar simulator for high-resolution nonhydrostatic models. *J. Atmos. Ocean. Tech.*, **23,** 1049-1067.

Caya, A., J. Sun, and C. Snyder, 2005: A comparison between the 4D-VAR and the ensemble Kalman filter techniques for radar data assimilation. *Mon. Wea. Rev.*, **133,** 3081–3094.

Chandrasekar, V., V. N. Bringi, N. Balakrishnan, and D. S. Zrnic´, 1990: Error structure of multiparameter radar and surface measurements of rainfall. Part III: Specific differential phase. *J. Atmos. Oceanic Technol.*, **7,** 621-629.

Cooper, W. A., 1986: Ice initiation in natural clouds. *Precipitation Enhancement-A Scientific Challenge*, No. 43, Amer. Meteor. Soc., 29-32.

Courtier, P. and O. Talagrand, 1987: Variational assimilation of meteorological observations with the adjoint equation. Part II: Numerical results. *Quart. J. Roy. Meteor. Soc.*, **113,** 1329-1347.

Dawson, D., M. Xue, J. A. Milbrandt, M. K. Yau, and G. Zhang, 2007: Impact of multi-moment microphysics and model resolution on predicted cold pool and reflectivity intensity and structures in the Oklahoma tornadic supercell storms of the 3 May 1999. 22nd Conf. Num. Wea. Pred.

Dawson, D. T., M. Xue, and J. A. Milbrandt, 2008: Improvements in the treatment of evaporation and melting in multi-moments versus single-moment bulk microphysics: results from numerical simulations of the 3 May 1999 Oklahoma tornadic storms. 24th Conf. Severe Local Storms

Dee, D. P., 1995: On-line estimation of error covariance parameters for atmospheric data assimilation. *Mon. Wea. Rev.*, **123,** 112-81145.

Dee, D. P. and A. M. d. Silva, 1999: Maximum-likelihood estimation of forecast and observation error covariance parameters. Part I: Methodology. *Mon. Wea. Rev.*, **127,** 1822-1834.

Doswell, D. C., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132,** 1982-2004.

Doviak, R. and D. Zrnic, 1993: *Doppler Radar and Weather Observations*. 2nd ed. Academic Press, 562 pp.

Dowell, D., F. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132,** 1982-2005.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99(C5),** 10 143-10 162.

Evensen, G., 2003: The ensemble Kalman filter: Theoretical formulation and practical implementation. *Ocean Dynamics*, **53,** 343-367.

Evensen, G. and P. J. v. Leeuwen, 1996: Assimilation of geosat altimeter data for the Agulhas current using the ensemble Kalman filter with a quasigeostrophic model. *Mon. Wea. Rev.*, **124,** 85-96.

Ferrier, B. S., 1994: A double-moment multiple-phase four-class bulk ice scheme. Part I: Description. *J. Atmos. Sci.*, **51,** 249-280.

Ferrier, B. S., W.-K. Tao, and J. Simpson, 1995: A double-moment multiple-phase four-class bulk ice scheme. Part II: Simulations of convective storms in different large-scale environments and comparisons with other bulk parameterizations. *J. Atmos. Sci.*, **52,** 1001-1033.

Fovell, R. G. and P.-H. Tan, 1998: The temporal behavior of numerically simulated multicell-type storms. Part II: The convective cell life cycle and cell regeneration. *Mon. Wea. Rev.*, **126,** 551-577.

Gao, J. and M. Xue, 2007: An efficient dual-resolution approach for ensemble data assimilation and tests with assimilated Doppler radar data. *Mon. Wea. Rev.*, Conditionally accepted.

Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004a: Precipitation uncertainty due to variations in precipitation particle parameters within a simple microphysics scheme. *Mon. Wea. Rev.*, **132,** 2610-2627.

Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004b: Precipitation and evolution sensitivity in simulated deep convective storms: Comparisons between liquid-only and simple ice and liquid phase microphysics. *Monthly Weather Review*, **132,** 1897-1916.

Golestani, Y., V. Chandrasekar, and V. N. Bringi, 1989: Intercomparison of multiparameter radar measurements. 24th Conf. on Radar Meteorology, Tallahassee, FL, Amer. Meteor. Soc., 309-314.

Green, A. V., 1975: An approximation for shape of large raindrops. *J. Appl. Meteor.*, **14,** 1578-1583.

Gunn, K. L. S. and J. S. Marshall, 1958: The distribution with size of aggregate snow flakes. *J. Meteor.*, **15,** 452-461.

Haase, G. and S. Crewell, 2000: Simulation of radar reflectivity using a mesoscale weather forecast model. *Water Resour. Res.*, **36,** 2221-2231.

Hacker, J. P. and C. Snyder, 2005: Ensemble Kalman filter assimilation of fixed screen-height observations in a parameterized PBL. *Mon. Wea. Rev.*, **133,** 3260-3275.

Hamill, T. M. and J. S. Whitaker, 2005: Accounting for the error due to unresolved scales in ensemble data assimilation: A comparison of different approaches. *Mon. Wea. Rev.*, **133,** 3132–3147.

Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, **129,** 2776-2790.

Hao, Z. and M. Ghil, 1995: Sequential parameter estimation for a coupled ocean-atmosphere model. Proceeding, WMO 2nd International Symposium on

Assimilatoin of Observations in Meteorology and Oceanography, Tokyo, Japan, 181-186.

Hendry, A. and G. C. McCormick, 1976: Radar observations of the alignment of precipitation particles by electrostatic fields in thunderstorms. *J. Geophys. Res.*, **81,** 5353-5357.

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

Hohenegger, C. and C. Schar, 2007: Atmospheric preditability at synoptic versus cloud-resolving scales. *Bull. Amer. Meteor. Soc.*, **88,** 1783-1793.

Hong, S.-Y. and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *J. Korean Meteor. Soc.*, **42,** 129-151.

Houtekamer, P. L. and H. L. Mitchell, 1998: Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, **126,** 796-811.

Houtekamer, P. L. and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129,** 123-137.

Houtekamer, P. L., H. L. Mitchell, G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, **133,** 604-620.

Houze, R. A. J., P. V. Hobbs, P. H. Herzegh, and D. B. Parsons, 1979: Size distributions of precipitation particles in frontal clouds. *J. Atmos. Sci.*, **36,** 156-162.

Hu, M., M. Xue, and K. Brewster, 2006a: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of Fort Worth tornadic thunderstorms. Part I: Cloud analysis and its impact. *Mon. Wea. Rev.*, **134,** 675-698.

Hu, M., M. Xue, J. Gao, and K. Brewster, 2006b: 3DVAR and cloud analysis with WSR-88D level-II data for the prediction of Fort Worth tornadic thunderstorms. Part II: Impact of radial velocity analysis via 3DVAR. *Mon. Wea. Rev.*, **134,** 699-721.

Huang, G.-J., V. N. Bringi, S. v. d. Heever, and W. Cotton, 2005: Polarimetirc radar signatures from RAMS microphysics. 32nd Int. Conf. Radar Meteor., Albuquerque, NM, Amer. Meteor. Soc., CDROM P11R.6.

Hubbert, J., V. N. Bringi, L. D. Carey, and S. Bolen, 1998: CSU-CHILL polarimetric radar measurements from a severe hail storm in eastern Colorado. *Journal of Applied Meteorology*, **37,** 749-775.

Hubbert, J. V. and V. N. Bringi, 1997: The effects of 3-body scattering on differential reflectivity. *28th Conf. on Radar Meteorology*, Amer. Meteor. Soc., 11-12.

Illingworth, A. J., J. W. F. Goddard, and S. M. Cherry, 1987: Polarization radar studies of precipitation development in convective storms. *Quart. J. Roy. Meteor. Soc.*, **113,** 469-489.

Joss, J. and A. Waldvogel, 1969: Raindrop size distribution and sampling size errors. *J. Atmos. Sci.*, **26,** 566-569.

Jung, Y., M. Xue, and J. M. Straka, 2005: Assimilation of polarimetric radar data using ensemble Kalman filter: Experiments with simulated data. Extended abstract, 17th Conf. Num. Wea. Pred., Washington DC, Amer. Meteor. Soc., 13A.3.

Jung, Y., G. Zhang, and M. Xue, 2008a: The estimation of microphysical parameters and atmospheric state using simulated polarimetric radar data and ensemble Kalman filter in the presence of observation operator error. *Mon. Wea. Rev.***,** Conditionally accepted.

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

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

Kalnay, E., 2002: *Atmospheric modeling, data assimilation, and predictability*. Cambridge University Press, 341 pp.

Klemp, J. B. and R. B. Wilhelmson, 1978: Simulations of right- and left-moving thunderstorms produced through storm splitting. *J. Atmos. Sci.*, **35,** 1097-1110.

Klemp, J. B., R. B. Wilhelmson, and P. S. Ray, 1981: Observed and numerically simulated structure of a mature supercell thunderstorm. *J. Atmos. Sci.*, **38,** 1558-1580.

Knight, N. C., 1986: Hailstone shape factor and its relation to radar interpretation of hail. *J. Climate Appl. Meteor.*, **25,** 1956-1958.

Kumjian, M. R. and A. V. Ryzhkov, 2008a: Storm-relative helicity revealed from polarimetric radar measurements. *J. Atmos. Sci.***,** In print.

Kumjian, M. R. and A. V. Ryzhkov, 2008b: Polarimetric signatures in supercell thunderstorms. *J. Appl. Meteor. Climatol.*, **47,** 1940-1961.

Lawson, W. G. and J. A. Hansen, 2005: Alignment error models and ensemble-based data assimilation. *Mon. Wea. Rev.*, **133,** 1687-1709.

Le Dimet, F. X. and O. Talagrand, 1986: Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects. *Tellus*, **38A,** 97-110.

Lemon, L. R. and C. A. Doswell, 1979: Severe thunderstorm evolution and mesocyclone structure as related to tornadogenesis. *Mon. Wea. Rev.*, **107,** 1184-1197.

Lin, Y.-L., R. D. Farley, and H. D. Orville, 1983: Bulk parameterization of the snow field in a cloud model. *J. Climate Appl. Meteor.*, **22,** 1065-1092.

Lin, Y.-L., R. L. Deal, and M. S. Kulie, 1998: Mechanisms of cell regeneration, development, and propagation within a two-dimensional multicell storm. *J. Atmos. Sci.*, **55,** 1867-1886.

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.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20,** 130-141.

Marwitz, J. D., 1972: The structure and motion of severe hailstorms. Part I: Supercell storms. *J. Appl. Meteor.*, **11,** 166-179.

Matson, R. J. and A. W. Huggins, 1980: The direct measurement of the sizes, shapes, and kinematics of falling hailstones. *J. Atmos. Sci.*, **37,** 1107-1125.

Maxwell-Garnett, J. C., 1904: Colors in metal glasses and in metallic films. *Philos. Trans. Roy. Soc. London*, **Ser. A,** 203.

May, R. M., M. I. Biggerstaff, and M. Xue, 2007: A Doppler radar emulator with an application to the detectability of tornadic signatures. *J. Atmos. Ocean Tech.***,** *J. Atmos. Ocean Tech.*, **24**, 1973-1996.

McCumber, M., W.-K. Tao, and J. Simpson, 1991: Comparison of ice-phase microphysical parameterization schemes using numerical simulations of tropical convection. *J. Appl. Meteor.*, **30,** 985-1004.

Meng, Z. and F. Zhang, 2007: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Imperfect model experiments. *Mon. Wea. Rev.*, **135,** 1403–1423

Meng, Z. and F. Zhang, 2008: Tests of an ensemble Kalman filter for mesoscale and regional-scale data assimilation. Part II: Comparison with 3DVAR in a real-data case study. *Mon. Wea. Rev.*, **136,** 522-540.

Milbrandt, J. A. and M. K. Yau, 2005a: A multi-moment bulk microphysics parameterization. Part II: A proposed three-moment closure and scheme description. *J. Atmos. Sci.*, **62,** 3065-3081.

Milbrandt, J. A. and M. K. Yau, 2005b: A multi-moment bulk microphysics parameterization. Part I: Aanlysis of the role of the spectral shape parameter. *J. Atmos. Sci.*, **62,** 3051-3064.

Mitchell, D. L., 1988: Evolution of snow-size spectra in cyclonic storms. Part I: Snow growth by vapor deposition and aggregation. *J. Atmos. Sci.*, **45,** 3431-3451.

Morrison, H. and J. O. Pinto, 2005: Mesoscale modeling of springtime arctic mixed-phase stratiform clouds using a new two-moment bulk microphysics scheme. *J. Atmos. Sci.*, **62,** 3683-3704.

Morrison, H., J. A. Curry, and V. I. Khvorostyanov, 2005: A new double-moment microphysics parameterization for application in cloud and climate models. Part I: Description. *J. Atmos. Sci.*, **62,** 1665-1677.

Musil, D. J., A. J. Heymsfield, and P. L. Smith, 1986: Microphysical characteristics of a well-developed weak echo region in a High Plains supercell thunderstorm. *J. Climate Appl. Meteor.*, **25,** 1037-1051.

Navon, I. M., 1997: Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dyn. Atmos. Oceans*, **27,** 55-79.

Otkin, J. A., D. J. Posselt, E. R. Olson, H.-L. Huang, J. E. Davies, J. Li, and C. S. Velden, 2007: Mesoscale numerical weather prediction models used in support of infrared hyperspectral measurement simulation and product algorithm development. *J. Atmos. Ocean Tech.*, **24,** 585-601.

Pfeifer, M., G. C. Craig, M. Hagan, and C. Keil, 2008: A polarimetric radar forward operator for model evaluation. *J. Appl. Meteor. Climatol.***,** DOI: 10.1175/2008JAMC1793.1.

Pruppacher, H. R. and J. D. Klett, 1978: *Microphysics of Clouds and Precipitation*. D. Reidel Publishers, 714 pp.

Ray, P. S., B. Johnson, K. W. Johnson, J. S. Bradberry, J. J. Stephens, K. K. Wagner, R. B. Wilhelmson, and J. B. Klemp, 1981: The morphology of severe tornadic storms on 20 May 1977. *J. Atmos. Sci.*, **38,** 1643-1663.

Rogers, R. R. and M. K. Yau, 1989: *A Short Course in Cloud Physics*. 3rd ed. Pergamon Press, 293 pp.

Romine, G. S., D. W. Burgess, and R. B. Wilhelmson, 2008: A dual-polarization-radar-based assessment of the 8 May 2003 Oklahoma City area tornadic supercell. *Mon. Wea. Rev.*, **136,** 2849-2870.

Rotunno, R., J. B. Klemp, and M. L. Weisman, 1988: A theory for strong long-lived squall lines. *J. Atmos. Sci.*, **45,** 463-485.

Rutledge, S. A. and P. V. Hobbs, 1983: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. Part VIII: A model for the feeder-seeder process in warm frontal rainbands. *J. Atmos. Sci.*, **40,** 1185-1206.

Ryzhkov, A. V., D. S. Zrnic, and B. A. Gordon, 1998: Polarimetric method for ice water content determination. *J. Appl. Meteor.*, **37,** 125-134.

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

Ryzhkov, A. V., S. E. Giangrande, V. M. Melnikov, and T. J. Schuur, 2005b: Calibration issues of dual-polarization radar measurements. *J. Atmos. Ocean. Tech.*, **22,** 1138-1155.

Ryzhkov, A. V., D. S. Zrnic, J. C. Hubbert, V. N. Bringi, J. Vivekanandan, and E. A. Brandes, 2002: Polarimetric radar observations and interpretation of co-cross-polar correlation coefficients. *J. Atmos. Ocean. Tech.*, **19,** 340-354.

Seliga, T. A. and V. N. Bringi, 1976: Potential use of radar differential reflectivity measurements at orthogonal polarizations for measuring precipitation. *J. Appl. Meteor.*, **15,** 59-76.

Smith, P. L., 1984: Equivalent radar reflectivity factors for snow and ice particles. *J. Climate Appl. Meteor.*, **23,** 1258-1260.

Smith, P. L., Jr., C. G. Myers, and H. D. Orville, 1975: Radar reflectivity factor calculations in numerical cloud models using bulk parameterization of precipitation processes. *J. Appl. Meteor.*, **14,** 1156-1165.

Snyder, C. and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131,** 1663-1677.

Straka, J. M., D. S. Zrnic, and A. V. Ryzhkov, 2000: Bulk hydrometeor classification and quantification using polarimetric radar data: Synthesis of relations. *J. Appl. Meteor.*, **39,** 1341-1372.

Straka, J. M. a. D. S. Z., 1993: An algorithm to deduce hydrometeor types and contents from multiparameter radar data. Preprints. 26th int. Conf. on Radar Meteorology, Norman, OK, Amer. Meteor. Soc., 513-515.

Tao, W.-K. and J. Simpson, 1989: Modeling study of a tropical squall-type convective line. *J. Atmos. Sci.*, **46,** 177-202.

Thorpe, A. J., M. J. Miller, and M. W. Moncrieff, 1982: Two-dimensional convection in non-constant shear: A model of midlatitude squall lines. *Quart. J. Roy. Meteor. Soc.*, **108,** 739-762.

Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003: Ensemble square root filters. *Mon. Wea. Rev.*, **131,** 1485-1490.

Tong, H., V. Chandrasekar, K. R. Knupp, and J. Stalker, 1998: Multiparameter radar observations of time evolution of convective storm: Evaluation of water budgets and latent heating rates. *J. Atmos. Ocean. Tech.*, **15,** 1097-1109.

Tong, M., 2006: Ensemble Kalman filter assimilation of Doppler radar data for the initialization and prediction of convective storms, Ph.D. Dissertation, School of Meteorology, University of Oklahoma, 243 pp.

Tong, M. and M. Xue, 2005a: Simultaneous retrieval of microphysical parameters and atmospheric state variables with radar data and ensemble Kalman filter method. Preprint, 17th Conf. Num. Wea. Pred., Washington DC, Amer. Meteor. Soc., CDROM P1.30.

Tong, M. and M. Xue, 2005b: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS Experiments. *Mon. Wea. Rev.*, **133,** 1789-1807.

Tong, M. and M. Xue, 2008a: Simultaneous estimation of microphysical parameters and atmospheric state with simulated radar data and ensemble square-root Kalman filter. Part II: Parameter estimation experiments. *Mon. Wea. Rev.*, **136,** 1649-1668.

Tong, M. and M. Xue, 2008b: Simultaneous estimation of microphysical parameters and atmospheric state with simulated radar data and ensemble square-root Kalman filter. Part I: Sensitivity analysis and parameter identifiability *Mon. Wea. Rev.*, **136,** 1630-1648.

Torres, S. M. and D. S. Zrnic, 2003: Whitening in range to improve weather radar spectral moment estimates. Part I: Formulation and simulation. *J. Atmos. Ocean. Tech.*, **20,** 1433-1448.

van den Heever, S. C. and W. R. Cotton, 2004: The impact of hail size on simulated supercell storms. *J. Atmos. Sci.*, **61,** 1596-1609.

Vivekanandan, J., W. M. Adams, and V. N. Bringi, 1991: Rigorous approach to polarimetric radar modeling of hydrometeor orientation distributions, *J. Appl. Meteor.,* **30**, 1053-1063.

Vivekanandan, J., R. Raghavan, and V. N. Bringi, 1993: Polarimetric radar modeling of mixtures of precioitation particles. *IEEE Trans. Geosci. Remote Sens.*, **31,** 1017-1030.

Vivekanandan, J., V. N. Bringi, M. Hagen, and G. Zhang, 1994: Polarimetric radar studies of atmospheric ice particles. *IEEE Trans. Geosci. Remote Sens.*, **32,** 1-10.

Vivekanandan, J., S. M. Ellis, R. Oye, D. S. Zrnic, A. V. Ryzhkov, and J. Straka, 1999: Cloud microphysics retrieval using S-band dual-polarization radar measurements. *Bull. Amer. Meteor. Soc.*, **80,** 381-388.

Wakimoto, R. M. and V. N. Bringi, 1988: Dual-polarization observations of microbursts associated with intense convection: The 20 July storm during the MIST Project. *Mon. Wea. Rev.*, **116,** 1521-1539.

Waterman, P. C., 1969: Scattering by dielectric obstacles, *Alta Frequenza*, (Speciale), 348-352.

Weisman, M. L. and J. B. Klemp, 1982: The dependence of numerically simulated convective storms on vertical wind shear and buoyancy. *Mon. Wea. Rev.*, **110,** 504-520.

Whitaker, J. S. and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130,** 1913-1924.

Whitaker, J. S., G. P. Compo, X. Wei, and T. M. Hamill, 2004: Reanalysis without radiosondes using ensemble data assimilation. *Mon. Wea. Rev.*, **132**.

Wilhelmson, R. B. and J. B. Klemp, 1978: A numerical study of storm splitting that leads to long-lived storms. *J. Atmos. Sci.*, **35,** 1975-1986.

Wu, B., J. Verlinde, and J. Sun, 2000: Dynamical and microphysical retrievals from Doppler radar observations of a deep convective cloud. *J. Atmos. Sci.*, **57,** 262-283.

Xu, Q., 1996: Generalized adjoint for physical processes with parameterized discontinuities - Part I: Basic issues and heuristic examples. *J. Atmos. Sci.*, **53,** 1123-1142.

Xue, M., 2002: Density currents in shear flows: Effects of rigid lid and cold-pool internal circulation, and application to squall line dynamics. *Quart. J. Roy. Met. Soc.*, **128,** 47-74.

Xue, M., K. K. Droegemeier, and V. Wong, 2000: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part I: Model dynamics and verification. *Meteor. Atmos. Phy.*, **75,** 161-193.

Xue, M., M. Tong, and K. K. Droegemeier, 2005: Impact of radar configuration and scan strategy on assimilation of radar data using ensemble Kalman filter. 9th Symp. Integrated Obs. Assim. Sys. Atmos. Oceans Land Surface, San Diego, California, Amer. Meteor. Soc., CDROM, 9.3.

Xue, M., M. Tong, and K. 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., Y. Jung, and G. Zhang, 2007: Error modeling of simulated reflectivity observations for ensemble Kalman filter data assimilation of convective storms. *Geophy. Res. Letters*, **34,** L10802, doi:10.1029/2007GL029945.

Xue, M., D.-H. Wang, J.-D. Gao, K. Brewster, and K. K. Droegemeier, 2003: The Advanced Regional Prediction System (ARPS), storm-scale numerical weather prediction and data assimilation. *Meteor. Atmos. Phy.*, **82,** 139-170.

Xue, M., K. K. Droegemeier, V. Wong, A. Shapiro, K. Brewster, F. Carr, D. Weber, Y. Liu, and D.-H. Wang, 2001: The Advanced Regional Prediction System (ARPS) - A multiscale nonhydrostatic atmospheric simulation and prediction tool. Part II: Model physics and applications. *Meteor. Atmos. Phy.*, **76,** 143-165.

Yakowitz, S. and L. Duckstein, 1980: Instability in aquifer identification: Theory and case studies. *Water Resour. Res.*, **16,** 1045-1064.

Zhang, F. and C. Snyder, 2007: Ensemble-based data assimilation. *Bull. Amer. Meteor. Soc.*, **88,** 565-568.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observations on the convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132,** 1238-1253.

Zhang, G., J. Vivekanandan, and E. Brandes, 2001: A method for estimating rain rate and drop size distribution from polarimetric radar measurements. *IEEE Trans. Geosci. Remote Sens.*, **39,** 830-841.

Zhu, Y. and I. M. Navon, 1999: Impact of parameter estimation on the performance of the FSU global spectral model using its full-physics adjoint. *Mon. Wea. Rev.*, **127,** 1497-1517.

Zou, X., I. M. Navon, and F. X. Le Dimet, 1992: An optimal nudging data assimilation scheme using parameter estimation. *Quart. J. Roy. Meteor. Soc.*, **118,** 1163-1186.

Zrnic, D. S. and A. V. Ryzhkov, 1999: Polarimetry for weather surveillance radars. *Bull. Amer. Meteor. Soc.*, **80,** 389-406.

Nomenclature

List of Symbols

D	Equivalent diameter of a precipitation particle
$D_{max,x}$	Maximum drop size of category x
D_{nr}	Mean-mass diameter of raindrop
f_w	Water fraction within a water-ice mixture
f_a	Backscattering amplitude for polarization along the major axis
f_b	Backscattering amplitude for polarization along the minor axis
J	Response function
K_w	Dielectric factor for water
K_{DP}	Specific differential phase
n_{0H}	Intercept parameter of hail/graupel size distribution
n_{0R}	Intercept parameter of rain size distribution
n_{0S}	Intercept parameter of snow size distribution
n_{0x}	Intercept parameter of category x
N	Ensemble size
Nt_x	Total number concentration of category $x \in (c, r, i, s, h)$
p	Pressure
p'	Pressure perturbation
p_i	Model microphysical parameter (Chapter 4)
p_i^t	Control (assumed true) value of parameter p_i (Chapter 4)
q_x	Mixing ration for category $x \in (c, r, i, s, h)$
q_{rh}	Mixing ration for rain-hail mixture

q_{rs}	Mixing ration for rain-snow aggregate mixture
q_v	Water vapor specific humidity
r	Axis ratio between minor to major axis of a precipitation particle
u	Horizontal velocity in the x direction
v	Horizontal velocity in the y direction
V_r	Radial velocity
w	Vertical velocity
x, y	Horizontal coordinates
$\mathbf{x}$	Model state vector
y	Observation vector
z	Vertical coordinate
Z_{dp}	Reflectivity difference
Z_{DR}	Differential reflectivity
Z_H	Equivalent reflectivity factor at the horizontal polarization
Z_V	Equivalent reflectivity factor at the vertical polarization
α	Shape parameter
β	Radar azimuth
ε_{corr}	Correlated part of the observation error
ε_h	Uncorrelated observation error for Z_H
ε_v	Uncorrelated observation error for Z_V
ϕ	Canting angle of a precipitation particle
$\overline{\phi}$	Mean canting angle of precipitation particles
λ	Radar wavelength

λ_x	Slope parameter of category x
θ	Potential temperature
θ'	Perturbation potential temperature
σ	Standard deviation of canting angle (Chapter 2)
ρ_{air}	Density of air
ρ_H	Density of hailstone
ρ_{hv}	Co-polar cross-correlation coefficient
ρ_m	Density of rain-snow/hail mixture (Chapter 2)
ρ_S	Density of snow aggregate

Acronyms

ARPS	Advanced Regional Prediction System
CASA	Center for Adaptive Sensing of the Atmosphere
CAPE	Convective available potential energy
DSD	Drop size distribution
EnKF	Ensemble Kalman filter
EnSRF	Ensemble Square Root Filter
GSFC	Goddard Space Flight Center
KOUN	Research polarimetric WSR-88D radar
NASA	National Aeronautics and Space Administration
NWP	Numerical Weather Prediction
OSSE	Observing System Simulation Experiments

PRD	Polarimetric radar data
RMS	Root-mean-square
SD	Standard deviation
SREH	storm-relative environmental helicity
UTC	Coordinated Universal Time
WSR-88D	Weather Surveillance Radar 88 Doppler
WRF	Weather Research and Forecasting
3DVAR	Three-dimensional variational
4DVAR	Four-dimensional variational