UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

ASSESSMENT OF ONE-MOMENT AND TWO-MOMENT BULK MICROPHYSICS AND SPECTRAL BIN MICROPHYSICS SCHEMES USING IDEALIZED SUPERCELL SIMULATIONS AND REAL DATA CONVECTIVE-SCALE PREDICTIONS

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

MARCUS RYAN JOHNSON
Norman, Oklahoma
2019

ASSESSMENT OF ONE-MOMENT AND TWO-MOMENT BULK MICROPHYSICS AND SPECTRAL BIN MICROPHYSICS SCHEMES USING IDEALIZED SUPERCELL SIMULATIONS AND REAL DATA CONVECTIVE-SCALE PREDICTIONS

A DISSERTATION APPROVED FOR THE
SCHOOL OF METEOROLOGY

BY

Dr. Ming Xue, Chair

Dr. Youngsun Jung, Co-Chair

Dr. Cameron Homeyer

Dr. Guifu Zhang

Dr. Rui Yang

Acknowledgments

I would like to first thank my wife, Rachelle, to whom this dissertation is dedicated. Time to get your own Ph.D. This dissertation is additionally dedicated to our cats, Whiskers and Socks (and Fireball). We finally did it, now both of you will be dining on wet food for the rest of your lives! My parents, Mike and Jill, have supported my academic decisions and celebrated my achievements throughout my entire life, for which I am grateful. There are many friends I have made during my graduate school career, but I will try to list some of them that have helped alleviate the stress of a dissertation: Jon and Laura Labriola, Hristo Chipilski, Eric Loken, Dania Sheaib, Amanda Burke, Bo Huang, and Mauricio Oliveira.

I owe much of my growth as a researcher to my advisor, Dr. Youngsun Jung. There is absolutely no way I would have been able to publish peer-reviewed scientific articles and grown an appreciation for the scientific process without her mentorship. My committee chair, Dr. Ming Xue, has worked tirelessly for the past six years to foster a limitless research environment for me, and has also contributed heavily to my growth as a scientist. I also want to thank the rest of my committee for agreeing to take part in my Ph.D degree: Dr. Cameron Homeyer, Dr. Guifu Zhang, and Dr. Rui Yang. I have learned so much from the five of you over the past few years, and I am truly grateful for that. I would like to additionally thank Dan Dawson (who also helped with my thesis and first publication), Tim Supinie, Jongsook Park, Yong-Hee Lee, Jason Milbrandt, and Hugh Morrison for their co-authorship for the research published in this dissertation. There are several scientists within the National Weather Center that continue to motivate me daily to reach my potential as a scientist: Dr. Ted Mansell, Dr. Lou Wicker, Dr. Nate Snook, Dr. Bryan Putnam, and Dr. Harold Brooks. The entire CAPS and SOM support staff are what keep both of these departments running so smoothly, and I am grateful that they provide an environment primed for

academic success. Finally, I would like to thank Shawn Smith, who took a chance on me as an undergraduate research assistant at Florida State University. I consider my experience at COAPS as my research beginning. The research in this dissertation was supported by a research grant of "Development of a Polarimetric Radar Data Simulator for Local Forecasting Model (II)" by the KMA. Further support was provided by a NOAA Warn-on-Forecast grant (Grant No. NA16OAR4320115) and a National Science Foundation grant (Grant No. AGS-1261776).

Table of Contents

List of Tables

List of Figures

Abstract

Optimal hydrometeor parameterization and their associated processes in microphysics schemes (both spectral bin and bulk) continue to evolve as these schemes attempt to match observed hydrometeor complexity. This dissertation spans several flavors of microphysics schemes: the one-moment Unified Model (UM), the partially-two moment Thompson and Morrison with one rimed ice category, the two-moment Milbrandt-Yau (MY2) and National Severe Storms Laboratory (NSSL) with two rimed ice categories, the Predicted Particle Properties (P3) with multiple mass assumptions within its ice particle size distributions (PSDs), and the spectral bin Hebrew University Cloud Model (HUCM). Microphysical performance (including their bias documentation) is examined in idealized supercell simulations by considering cloud ice and snow moment (and their associated budget) evolution, cloud ice and snow PSDs, low-level classic polarimetric radar signatures (Z_{DR} arc and hail signature in the forward flank downdraft), and ice hydrometeor contoured frequency by altitude diagrams (CFADs). Two test cases over the Korean Peninsula (Changma front and Typhoon Sanba [2012]) are compared to S-band radar observations by applying a dual-polarization radar variable simulator to UM output. 2018 NOAA Hazardous Weather Testbed (HWT) Spring Experiment seasonal forecasts over much of the continental United States (CONUS) and four select convective line cases are both quantitatively and qualitatively compared to observed composite reflectivity, accumulated precipitation, and brightness temperature in the 11.2 μm channel for short-term (t = 1 – 6 forecast hours) and next-day (t = 12 – 36 forecast hours) forecasts.

UM microphysics struggles to match observed dual-pol variables because of its one-moment parameterization of rain, specifically its rain PSD intercept parameter N_0 diagnosis. As N_0 varies inversely with rain mass, the scheme is producing too many small (large) drops in regions

of too weak (intense) reflectivity. Both the fully two-moment MY2 and NSSL schemes are able to simulate a local maximum of Z_{DR} near the forward flank edge and a gradual decrease in the direction of the deep-layer storm relative mean wind vector, but the large, dry hail in the MY2 scheme reduces Z_{DR} on the edge of the supercell, while the NSSL's Z_{DR} arc is less elongated compared to typical observations. The P3 scheme with two ice categories is unable to simulate either signature, due to the restrictive rain and ice PSD slope Λ limiters (both directly and indirectly) preventing larger particles. In idealized supercell simulations, the HUCM and NSSL schemes simulate larger ice crystal moments than snow, while the Thompson scheme simulates more snow mass. This is due to the aggressive cloud ice to snow conversion in the scheme, which is intended given the assumed snow PSD. The flexible spectral bin HUCM PSDs simulate less small snow particles than the rigid bulk NSSL and Thompson schemes, but also sediments too large snow to the surface.

Over the 2018 NOAA HWT Spring Experiment, the Morrison scheme displays a large-scale storm structure ($Z \geq 15$ dBZ) overprediction bias for short-term forecasts that lessens for next-day forecasts, and is likely due to enhanced horizontal graupel advection in the scheme because of its smaller fall speed. Both the NSSL and Thompson schemes underpredict this storm structure. The Morrison and NSSL schemes both overpredict convective storm structure ($Z \geq 40$ dBZ), due to overadvected Morrison graupel melting to large rain, and the "large hail" category design in the NSSL scheme. Each BMP underpredicts light and heavy surface precipitation, indicating that the BMPs underpredict either total column mass and/or its sedimentation to the surface. The documented shortcomings and biases in this dissertation are essential to numerical modelers and their users alike, as users should select the appropriate scheme for their simulated storm, and numerical modelers can optimally tune/construct their microphysics scheme.

Chapter 1 Introduction

1.1 Motivation and Background

Explicit representation of cloud evolution has received much attention lately as simulation and operational numerical weather prediction (NWP) continues to increase in model resolution. Global models (e.g., Global Forecast System [GFS]) that contain coarse model resolutions only need to rely on rudimentary microphysics and cumulus parameterization schemes (e.g., Kessler 1969; Arakawa and Schubert 1974; Kain and Fritsch 1990; Zhao and Carr 1997) to simulate precipitation. Regional models that simulate mesoscale storms (e.g., mesoscale convective systems [MCSs] and supercell thunderstorms) require at a minimum 4 km grid spacing (i.e., "convective-scale") in order to faithfully represent the storms (Weisman et al. 1997). At this grid spacing (or scale), the large-scale assumptions included in the minimalist microphysics and cumulus parameterization schemes begin to lose propriety (Cotton and Anthes 1989). Therefore, explicit microphysics parameterization is necessary to capture convective-scale processes and accurately simulate these storms.

Microphysics parameterizations have traditionally fallen into two categories: spectral bin microphysics (SBM) and bulk microphysics parameterization (BMP) schemes. Spectral bin schemes typically prognose either discretized bins of particle size distributions (PSDs) or their related moments. For example, the Hebrew University Cloud Model (HUCM; Khain et al. 2004) prognoses the PSD using 33 mass-doubling bins for each of the scheme's hydrometeor categories. SBMs offer unmatched flexibility because not only is a PSD form (i.e., gamma) not required, rain and ice properties (e.g., axis ratio, density) may be easily parameterized for each bin without any continuous function assumptions. The lack of rigid PSD parameterization is particularly useful, as observed ice PSDs take on many forms (e.g., Delanoë et al. 2005; Field et al. 2007). The

commonly-used three-parameter gamma distribution does not have exact fits to observations (McFarquhar et al. 2015), complicating both parameterization choice and validation. Still, the computational cost of SBMs is daunting: perhaps the most complex BMP (Milbrandt-Yau three-moment scheme; Milbrandt and Yau 2005b) prognoses 18 variables while the HUCM prognoses 264, representing a prognosed variable increase by a factor of over 14. Given the distinct framework of the two microphysics paradigms, it is difficult to compare SBM and BMP performance differences and specifically attribute them to microphysical design; however, the two have shown qualitatively similar skill (e.g., Xue et al. 2017). For this reason, BMPs are traditionally employed operationally and experimentally, as both tuning of the scheme and adding (ice) complexity is computationally cheaper, and are the main focus of this dissertation (although the performance of one SBM is evaluated; for more details on SBMs see section 4.1).

Bulk microphysics schemes (BMPs), which typically assume a fixed PSD form and predict moments of the PSD, continue to grow in complexity due in part to increasing computing power. One-moment schemes that only predict mass mixing ratio (e.g., Tao and Simpson 1993; Chen and Sun 2002; Straka and Mansell 2005) have lost popularity as two-moment schemes that additionally prognose number concentration N_T (e.g., Ziegler 1985; Seifert and Beheng 2001; Morrison et al. 2005; Thompson et al. 2008; Mansell et al. 2010) are inherently more flexible without prohibitive computational cost. While three-moment schemes that additionally prognose radar reflectivity Z have been developed (e.g., Milbrandt and Yau 2005b), recent microphysics improvement has focused on attempting to match observed ice complexity (Magono and Lee 1966), as liquid hydrometeors are well-approximated with a spherical, constant density model. Currently, both the Milbrandt-Yau (Milbrandt and Yau 2005b) and Thompson BMPs (Thompson et al. 2008) set the snow m-D exponential near 2, which matches observations closer than the typically assumed

spherical relationship (e.g., Mitchell et al. 1990). Recent BMPs are prognosing ice characteristics such as density, rime history, and axis ratio such that microphysical processes can better reflect the hydrometeor's history (e.g., Mansell et al. 2010; Harrington et al. 2013; Milbrandt and Morrison 2013; Morrison and Milbrandt 2015).

The construction of BMPs, particularly ice categories, has a profound impact on simulated storm evolution. Radiative forcing is sensitive to both cloud ice size and number concentration, which is either prognosed through the ice nucleation process or parameterized (e.g., Iacobellis et al. 2003; Xie et al. 2013). Regarding snow, Woods et al. (2007) noted increased surface precipitation accuracy when setting the snow power-law m-D relationship closer to observations (i.e., exponential closer to 2) than assuming spheres. Liu et al. (2011) documented that the Thompson scheme, which assumes a linear combination of exponential and gamma PSDs as well as the previously mentioned realistic m-D relationship, outperformed several other one-moment BMPs (which consistently overpredicted snowfall precipitation) and was similar in seasonal snowfall prediction skill to the two-moment Morrison scheme. Perhaps the most important hydrometeor in deep-convective storms is rimed ice, whose associated storm sensitivity has been studied extensively (e.g., McCumber et al. 1991; Gilmore et al. 2004; van den Heever and Cotton 2004; Franklin et al. 2005; Adams-Selin et al. 2013; Seigel and van den Heever 2013). Specifically in supercell simulations, Dawson et al. (2014) found that the Z_{DR} arc (a microphysical radar signature) is more influenced by the parameterization of rimed ice rather than rain. Johnson et al. (2016) noted improved supercell simulations when employing a BMP with two rimed-ice categories.

Dual-polarization radar observations provide observation validation at a greater temporal and spatial resolution than any other observation instrument. Dual-polarization radar variables also

provide qualitative microphysics information that is useful for simulated hydrometeor validation. For example, reflectivity at the horizontal polarization (Z_H) is sensitive to hydrometeor mass and size, differential reflectivity (Z_{DR}) is sensitive to the axis ratio of hydrometeors, specific differential phase (K_{DP}) is proportional to liquid mass, and cross-correlation coefficient (ρ_{HV}) varies with hydrometeor phase (e.g., Seliga and Bringi 1976; Bringi and Chandrasekar 2001). These qualitative relationships have led to the development of hydrometeor classification algorithms (HCAs; Zrnić et al. 2001; Park et al. 2009; Thompson et al. 2014) and improved QPE relationships (e.g., Ryzhkov et al. 2005b; Thompson et al. 2018). Polarimetric signatures also have been used to infer microphysical and dynamical processes in supercell thunderstorms (e.g., Kumjian and Ryzhkov 2008). Therefore, observed reflectivity and other dual-pol variables provide a popular observation validation for numerical models (e.g., Jung et al. 2012; Snook et al. 2012; Putnam et al. 2014; Yussouf et al. 2015; Supinie et al. 2017; Johnson et al. 2018).

1.2 Dissertation Overview

The purpose of this dissertation is to evaluate several flavors of microphysics schemes including both SBM and BMP microphysics, as well as one-moment and two-moment BMPs, to help identify microphysical biases, to allow numerical modellers to address these biases, and to guide users to employ microphysics best suited for their numerical simulations. We employ the Weather Research and Forecasting (WRF) model and Unified Model (UM) for all numerical simulations in this dissertation. The spectral bin Hebrew University Cloud Model (HUCM), the one-moment Unified Model microphysics, and the two-moment Morrison, Milbrandt-Yau, National Severe Storms Laboratory (NSSL), Predicted Particle Properties (P3), and Thompson bulk microphysics schemes are evaluated in this dissertation to examine a wide span of microphysics parameterization. All simulated storms are at the convective-scale ($\leq$ 4 km; where

microphysics parameterization becomes significant) and include idealized supercell simulations, a tropical cyclone, a changma front, and seasonal and four convective line test case forecasts over the 2018 National Oceanic and Atmospheric Administration (NOAA) Hazardous Weather Testbed (HWT) Spring Experiment.

Chapter 2 compares the simulation of both a Changma front and tropical cyclone over the Korean peninsula by the Unified Model. The Unified Model is employed across several weather agencies across the globe, so motivation exists to outline any biases/shortcomings in the scheme. A polarimetric radar simulator is employed to compare Unified Model simulated polarimetric variables with those of observations. At low scans (i.e., mostly rain), storm structure (horizontal reflectivity Z_H), drop size (differential reflectivity Z_{DR}), and rain mass (specific differential phase (K_{DP}) can be qualitatively compared between the Unified Model and observations. Further, a hydrometeor classification algorithm (HCA) is also applied to compare dominant observed hydrometeors above and below the melting layer with those simulated by the Unified Model.

Chapter 3 determines the sensitivity of simulated polarimetric radar variables to rimed ice parameterization. Two BMPs, the Milbrandt-Yau and NSSL schemes are employed as each scheme has two rimed-ice categories with different parameterizations. This chapter also analyses the performance of the novel P3 scheme, whose ice categories contain a single particle size distribution with up to four mass-diameter relationships (and therefore, four ice modes). Two classic supercell signatures, the Z_{DR} arc and hail signature in the forward flank downdraft (FFD), are chosen to qualitatively measure the performance of each microphysics scheme. Microphysical budgets are analyzed to determine the growth and depletion of rimed ice in each scheme. CFADs are created to examine rain and rimed ice size sorting in each BMP, an important dynamic mechanism necessary to simulate a Z_{DR} arc. Rimed ice size is also analyzed in the updraft to

investigate large rimed ice production in each scheme, which is responsible to simulate the hail signature. Further, the sensitivity of the two classic radar signatures to P3-2 microphysical parameters is also tested.

Chapter 4 compares the ice crystal and snow evolution in the spectral bin HUCM with those in the bulk NSSL and Thompson schemes for an idealized supercell simulation. The HUCM is a unique microphysics scheme as it contains three ice crystal modes (columns, dendrites, and snow) in addition to a separate snow category, which the NSSL and Thompson schemes only contain cloud ice and snow. Ice crystal and snow temporal and vertical prognostic moments in the microphysics schemes are analyzed to determine how each scheme evolves these hydrometeors over the simulation. Microphysical budgets are created to infer which processes contribute most to cloud ice and snow evolution in each scheme. Particle size distributions (PSDs) at hydrometeor mass relative maxima are analyzed to determine how the prognostic bin PSDs compare to the prescribed BMP PSDs. Contoured frequency by altitude diagrams (CFADs) also shed light on the vertical distribution of ice/snow size among the schemes.

In chapter 5, the performance of the Morrison, NSSL, and Thompson BMPs are quantitatively evaluated over the 2018 NOAA HWT Spring Experiment. In line with Warn-on-Forecasts goals, these microphysics schemes are examined within the 1-6 forecast hours, as well as next-day (12-36 h) forecast hours. Contingency table are constructed over the entire season (five weeks predominantly in May) and for four convective line cases using Model Evaluation Tools v6.0 (MET) software. From these contingency tables, which employ a neighborhood approach, performance diagrams, equitable threat score (ETS), fractions skill score (FSS), reliability and relative operating characteristic (ROC) curve diagrams are created to evaluate each microphysics

scheme's ability to simulate composite reflectivity, 1-hour accumulated precipitation, and brightness temperature in the 11.2 μm channel.

In chapter 6, we summarize and discuss the microphysical performance for each BMP across idealized and real data cases. Also, future work about optimal microphysics performance across storm modes and potential observations necessary to realize this goal is explored.

Chapter 2 [1] Performance of the one-moment Unified Model over the Korean Peninsula

2.1 Introduction

As computing power consistently increases, operational centers run numerical weather prediction (NWP) models with convective-scale ($\leqslant$ 4 km; Weisman et al. 1997) grid spacing (e.g., Tang et al. 2013; Goldenberg et al. 2015; Kim 2015; Park et al. 2015a; Ballard et al. 2016). Model microphysical processes become significant at this resolution and drive the evolution of the precipitating system. In order to gain understanding of microphysics complexity, behavior, and potential biases, recent research has focused on microphysics scheme performance (e.g., Cintineo et al. 2014; McMillen and Steenburgh 2015; Morrison et al. 2015; Johnson et al. 2016) and sensitivity (e.g., Morrison and Milbrandt 2011; Morrison et al. 2012; Van Weverberg et al. 2012). Specifically, polarimetric radar data is a powerful tool for comparing model output with observations (e.g., Jung et al. 2012; Dawson et al. 2013; Brown et al. 2016; Putnam et al. 2017b) because radar polarimetry can provide observational microphysics information (i.e., differential reflectivity Z_{DR} is related to hydrometeor shape).

Microphysics schemes in NWP models typically represent particle size distributions (PSDs) using a gamma distribution:

$$N(D) = N_0 D^{\mu} \exp(-\Lambda D) \tag{2.1}$$

where N_0, μ and Λ are the intercept, shape and slope parameters, respectively, and D is the particle diameter. For one-moment (1M) schemes in which one PSD moment (typically mixing ratio) is

[1] This chapter is published as: Johnson, M., Y. Jung, D. Dawson, T. Supinie, M. Xue, J. Park, and Y.-H. Lee, 2018: Evaluation of Unified Model microphysics in high-resolution NWP simulations using polarimetric radar observations. *Adv. Atmos. Sci.*, **35**(7), 771–784, doi: 10.1007/s00376-017-7177-0.

predicted, Λ typically varies freely while N_0 is usually fixed (e.g., Lin et al. 1983; Rutledge and Hobbs 1983; Tao and Simpson 1993; Hong and Lim 2006) or diagnosed (i.e., as a function of temperature or mixing ratio [see Hong and Lim (2006), Thompson et al. (2008), and the Thompson graupel intercept parameter in (Morrison et al. 2015)]. Diagnosing N_0 in 1M schemes allows N_0 to vary, but not independently of its tied parameter (i.e., mass). Diagnostic N_0 has shown improvement over fixed N_0 (Zhang et al. 2008; Wainwright et al. 2014; Pan et al. 2016), while improvement was not as clear in other studies (Straka et al. 2005; Milbrandt and Yau 2006; Van Weverberg et al. 2011). As previous studies have focused mainly on convective events, additional studies regarding the tuning of N_0 to large-scale events would be helpful for regional forecasts.

Two-moment (2M) schemes add an additional predicted variable (typically number concentration) that allows N_0 to vary independently of mass (e.g., Milbrandt and Yau 2005b; Morrison et al. 2009; Mansell et al. 2010). Several studies have compared 1M microphysics scheme performance with multi-moment (2M or higher) schemes. While 1M schemes are theoretically computationally faster than multi-moment schemes with additional predicted moments, one of the largest deterrents of employing 1M schemes are their inability to replicate size sorting (e.g., Dawson et al. 2010; Jung et al. 2010; Dawson et al. 2014; Johnson et al. 2016), a fundamental microphysical process where larger hydrometeors fall faster than smaller ones. This mechanism is presumably more important in deep convective storms within directional wind shear environments, where a strong updraft allows larger hydrometeor growth (and subsequently, size difference) and transport sedimentation spatially distributes them.

It is desirable that microphysics schemes retain consistent performance over various precipitation modes. However, some microphysics schemes have been developed and tested for (and therefore, potentially biased toward) large-scale (e.g., Wilson and Ballard 1999; Hong and

Lim 2006; Thompson et al. 2008) or storm-scale precipitation systems (e.g., Milbrandt and Yau 2005b; Morrison et al. 2009; Mansell et al. 2010). As horizontal grid resolution decreases, the microphysical processes and parameterizations in the scheme do not necessarily change. Ideally, microphysics performance remains unchanged with varying grid scale. In reality, microphysics schemes have the potential to grow model error when they run outside scales for which they are tuned. For example, microphysics schemes tuned for large-scale precipitation systems might favor small rain drops through aggressive breakup and could poorly simulate large raindrops often seen in supercell storms. While some studies have examined model performance sensitivity to horizontal grid resolution (e.g., Bryan and Morrison 2012; Potvin and Flora 2015; Verrelle et al. 2015), more rigorous study across several microphysics schemes is needed to provide guidance for microphysics scheme improvement by modelers and help users choose the best microphysics/resolution combination for their modeling purposes.

While some of the studies previously mentioned have shown the superior performance of multi-moment microphysics schemes over 1M schemes, 1M schemes are still popular for operational models primarily because of their low computational cost. Further, some studies provide optimism that 1M scheme performance could be improved. The Thompson microphysics scheme, whose snow processes use 1M parameterization (i.e., snow mixing ratio is predicted) with a combination of two PSDs, predicts accumulated snowfall closer to observed totals than other 1M schemes and performs similarly to the Morrison 2M scheme in Liu et al. (2011). Further, Bryan and Morrison (2012) demonstrated that increasing grid resolution noticeably increased 1M performance in terms of surface precipitation and storm evolution for a simulated squall line, although the 2M scheme still outperformed the 1M scheme.

This chapter examines the performance of the Unified Model (UM) microphysics scheme using two distinct cases: a Changma front, and Typhoon Sanba (2012) over the Korean peninsula. Polarimetric variables are computed from UM output using the radar simulator based on Jung et al. (2008, 2010) and compared to observations from the Biseul-san radar, which is an S-band polarimetric radar. It is one of six radars operated by the Ministry of Land, Infrastructure and Transport to improve flood forecasts. More details on the radar can be found in Park et al. (2015b). This chapter aims to identify any biases and weaknesses present so that modelers can improve the scheme and help researchers and forecasters to interpret forecasts given microphysical biases. The chapter is organized as follows: section 2.2 details the UM configuration and the polarimetric radar data simulator; section 2.3 compares the structure of polarimetric observations with UM simulated polarimetric variables; section 2.4 analyzes the polarimetric distributions; section 2.5 expands the comparisons to frozen hydrometeors; and section 2.6 provides conclusions.

2.2 Overview of UM microphysics and the polarimetric radar data simulator

2.2.1 UM microphysics

The UM microphysics scheme is rooted in Wilson and Ballard (1999) [itself derived from Rutledge and Hobbs (1983)], although modifications continue to update and improve the scheme (e.g., Wilkinson et al. 2013). The UM microphysics scheme is unique in that it contains many parameterization choices. Here, we list relevant parameterizations in Korea Meteorological Administration (KMA) model simulations. At the most basic level, UM microphysics contains three hydrometeor categories: cloud water, rain, and ice. The default UM microphysics scheme is unique in that cloud ice and snow are contained in the ice category, although an option exists to separate these into individual categories. In our simulations, the ice category is represented by a single generic ice distribution (Field et al. 2005, 2007). Further, the UM microphysics

configuration employed in this chapter contains an additional graupel category, an essential hydrometeor category for deep convection. Briefly, the UM is centered at 37.57°N, 126.97°E over the Korean peninsula on 744 × 928 grid points with a predominant horizontal grid spacing Δx of 1.5 km, and includes 21 grid points of varying grid zones with Δx increasing to 4 km at the lateral boundaries. The model contains 70 terrain-following vertical levels up to 39 km. The model is integrated using a semi-implicit, semi-Lagrangian method, with time step Δt = 50 s and model output every hour. Forecasts are initialized at 0000 UTC using three-dimensional variational data assimilation every 6 h and integrated up to 36 h. More details can be found in Table 2.1.

In our configuration, the PSDs are exponential for rain, gamma for graupel, and a linear combination of exponential and gamma distributions for ice (Table 2.2). As the hydrometeor categories use 1M parameterization, N_0 must be parameterized in the PSDs. Rain and graupel N_0 is a power-law function of the slope parameter Λ:

$$N_0 = n_{\mathrm{ax}} \Lambda^{n_{\mathrm{bx}}} \tag{2.2}$$

where n_{ax} and n_{bx} are constants for hydrometeor x. The rain N_0 relationship is from Abel and Boutle (2012). Ice N_0 is a function of the second and third moments of the ice distribution M_2 and M_3 (Field et al. 2007). Rain and graupel densities are set to 1000 and 500 kg m^{-3}, respectively. Ice assumes a power-law mass relationship that does not assume a spherical shape, resulting in varying bulk density.

2.2.2 Polarimetric radar simulator

The polarimetric radar simulator employed in this study to compute polarimetric variables is well-documented in literature (Jung et al. 2008; Jung et al. 2010; Dawson et al. 2014). Briefly, the simulator constructs particle size distributions (PSDs) from model output consistent with UM microphysical assumptions (or more generally, consistent with the microphysical assumptions of

the employed scheme). However, we do note that the generic ice category is treated as snow (and will be referred to as snow for the remainder of the paper) with constant density of 100 kg m^{-3} for scattering purposes. This should not be a large source of error, as high-density ice crystals are quite small in the PSD used in the scheme, and thus contribute little to polarimetric calculations. Because polarimetric variables are sensitive to axis ratio, orientation, and water fraction that are not typically provided with model information, we make assumptions for these hydrometeor characteristics in the polarimetric simulator that may introduce parameterization error. The axis ratio of rain follows the quartic fit from Brandes et al. (2002), while rain is assumed to have a stable orientation as both mean and standard deviation of canting angle is 0° (Hendry and McCormick 1976). Rimed ice axis ratio is set to 0.75 following Knight (1986), which is also the assumed axis ratio for snow in the radar simulator. Snow and rimed ice mean canting angle is also set to 0°, while the standard deviation of snow canting angle is set to 0°. Rimed ice canting angle standard deviation is assumed 60° for dry rimed ice to model tumbling hail, which decreases linearly to 0° as rimed ice water fraction approaches 1 to model stable, wet rimed ice particles. Water fraction is diagnosed using a linear relationship as a function of temperature. As temperature increases from –2.5° to 2.5°C for snow and from –5° to 0°C for graupel, water fraction increases from 0 to 0.8 for snow and from 0 to 0.4 for graupel. These temperature and water fraction ranges are chosen to tune the simulated melting layer to match observations in terms of depth and intensity. Finally, scattering amplitudes are retrieved from precomputed T-matrix tables that vary with particle diameter and water fraction. Polarimetric variables reflectivity at the horizontal polarization (Z_H), differential reflectivity (Z_{DR}), specific differential phase (K_{DP}), and correlation coefficient (ρ_{HV}) are computed from these scattering amplitudes (Zhang et al. 2001; Jung et al. 2010). For more details on the polarimetric simulator, we refer the reader to the above publications.

2.2.3 Hydrometeor classification algorithm (HCA)

A hydrometeor classification algorithm (HCA) is employed in this study to further compare simulated hydrometeors with those in the observed storms. The model dominant hydrometeor type is defined as the hydrometeor type that contributes most to linear reflectivity. The observed dominant hydrometeor type is determined by an HCA similar to those used in Park et al. (2009) and Putnam et al. (2017b) using the same membership functions, weights, and additional constraints to remove possible hydrometeor diagnosis based on Z, Z_{DR}, ρ_{HV}, and radial velocity V. For simplicity, only Z, Z_{DR}, and ρ_{HV} member functions are considered. Melting level information is incorporated by defining the bottom of the melting level as the lowest model height where simulated snow mass is present, while the top is defined as the lowest model height where ambient temperature reduces below 0°C. Snow, ice crystals, and graupel are not allowed below the melting layer, ground clutter, biological scatterers, dry snow (except in the upper half of the melting layer), and ice crystals are not allowed in the melting layer, while only dry snow, ice crystals, graupel, and hail (including wet hail) are allowed above the melting layer. Finally, once the observed dominant hydrometeor is diagnosed, the resulting hydrometeor types are: ground clutter/anomalous propagation (GC/AP), biological scatterers (BS), dry snow/ice crystals (DS/CR), wet snow (WS), rain (RA; including light/moderate rain, heavy rain, and big drops), rain/rimed ice (RR), and rimed ice (RI; including graupel and hail).

2.3 Polarimetric structure for the model and observations

In this section, the polarimetric simulator is applied to two distinct weather cases–a Changma front, and Typhoon Sanba (2012)–and compared to observed polarimetric variables measured by the Biseul-san radar. For forecast verification, it is typical to compare observations and model forecasts valid at the same time. However, position error that grows fast during the

Table 2.1. UM configuration.

UM configuration	Description	
Grid	Horizontal	Arakawa C-grid
	Vertical	Charney–Phillips grid staggering
Map projection		Rotated latitude and longitude coordinate system
Time integration		Semi-implicit, semi-Lagrangian
Model version		Vn 8.2
Domain	Horizontal	Varying grid spacing: 4 km → 1.5 km; 744 × 928 grid points
	Vertical	70 levels (up to 39 km); terrain-following η system
Δt		50 s (short-time step: 30 s)
Radiation		Spectral band radiation (general 2-stream and radiance)
Surface physics		JULES land-surface scheme
PBL scheme		Non-local scheme with revised diagnosis of K profile depth
Cumulus parameterization		None
Microphysics		Single-moment UM microphysics with graupel

Table 2.2. Summary of the hydrometeor PSDs in the UM microphysics examined in this chapter. Note: M_2 and M_3 are the second and third moments of the ice distribution, respectively. D is the particle diameter.

Hydrometeor	Distribution $N(D)$			
Ice	$\frac{M_2^4}{M_3^3}\left[141\exp\left(-16.8\frac{M_2}{M_3}D\right)+102\left(\frac{M_2}{M_3}D\right)^{2.07}\exp\left(-4.82\frac{M_2}{M_3}D\right)\right]$			
		n_{ax}	n_{bx}	μ
Rain	$N_0 D^{\mu}\exp(-\Lambda D)$	0.22	2.2	0.0
Graupel	$N_0 D^{\mu}\exp(-\Lambda D)$	5×10^{25}	−4.0	2.5

forward model integration makes it difficult to compare precipitation systems in the same domain, especially when the verification data have limited coverage (i.e., the Biseul-san radar utilized in this chapter only covers 150 km). As radar beam height increases with increasing distance, precipitation systems that are being compared should be at a similar distance from the radar. In addition, polarimetric measurements exhibit large dynamic ranges depending on widely varying PSDs within the precipitation system. It would be more practical to compare the variables in the radar echoes that show similar characteristics in both observations and model simulations. Therefore, we compare precipitation systems showing similar qualitative reflectivity structure (i.e., rainbands) and distance from the radar in the observed and model reflectivity to allow for temporal errors.

Because the prevailing geography of Korea is mountainous, many radars are placed on the tops of mountains and the lowest elevation angles are below 0° (e.g., −0.5°). The Biseul-san radar used in this chapter is located on the top of Biseul Mountain (1085 m above mean sea level). We analyze model and observational plots of polarimetric and microphysical variables at the 0.5° elevation angle to ensure the primary precipitating hydrometeor is rain, eliminating substantial ground clutter contamination or beam blockage. The beam height of this elevation angle (taking into account radar altitude and Earth's curvature) within the radar coverage is below 4000 m.

2.3.1 Changma front (10 July 2012)

At 0000 UTC 10 July 2012, an east–west Changma front attached to a low-pressure system was positioned west of the Korean peninsula (not shown). While the low gradually detached from the Changma front and moved northeast, the Changma front settled just over the southern part of

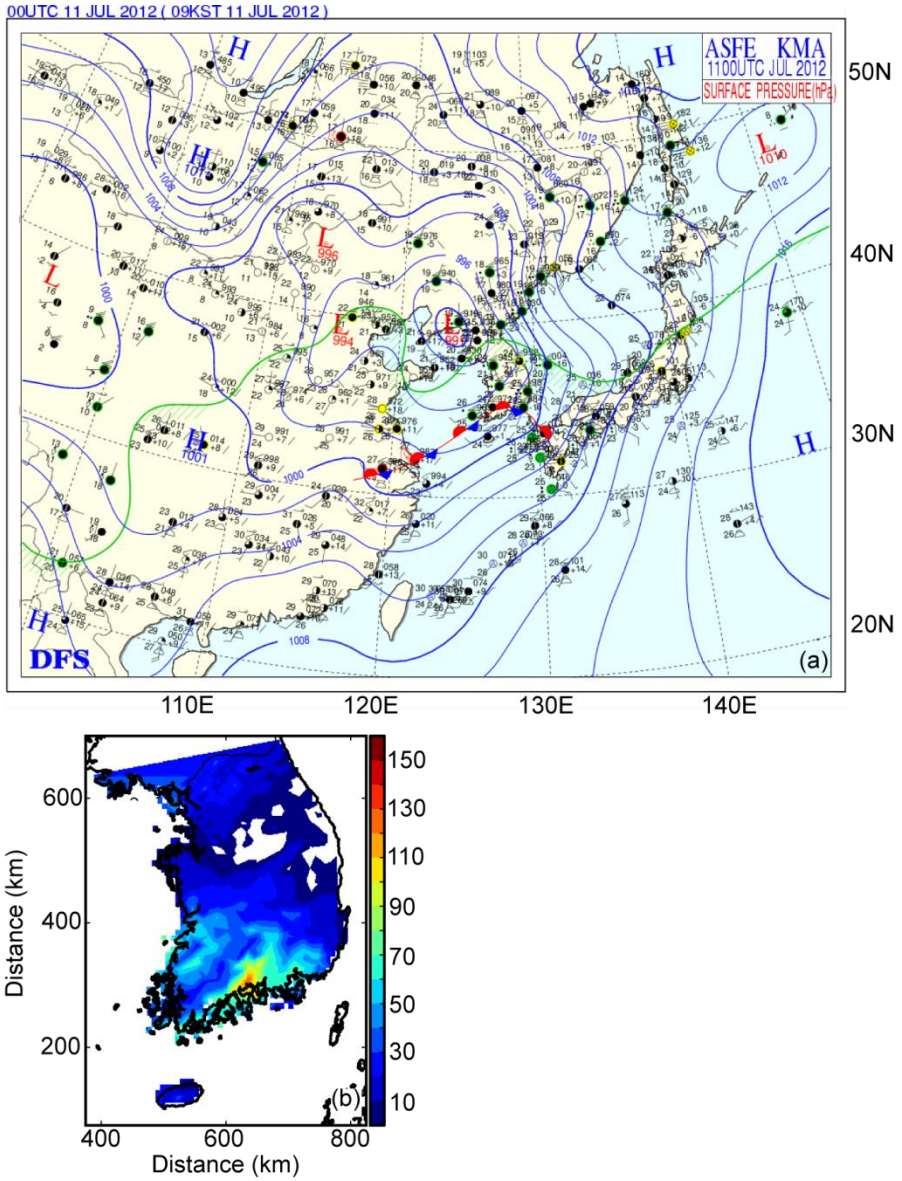

Fig. 2.1. The (a) surface chart at 0000 UTC 11 July 2012 and (b) rain accumulation (units: mm) from AWS gauge data over 12 h ending at 2330 UTC 10 July 2012 for the Changma front case.

the Korean peninsula near the coastline at 0000 UTC 11 July 2012 (Fig. 2.1a). The green line shows the 20°C isodrosotherm, and the Changma front is often located south of this line. Figure 2.1b depicts the accumulated rainfall from AWS rain gauges on the peninsula over 12 h, interpolated linearly over the domain. Accumulated rainfall exceeds 100 mm sparingly in the

middle of the Korean southern coast, which is where the Changma front is positioned. Further north, accumulated rainfall continually decreases and even falls below 5 mm, notably in the central and eastern parts of South Korea. These regions possibly received little rainfall because the stationary front stayed to the south and the low-pressure system moved to the northeast.

Figure 2.2 shows the radar observations of the Changma front at 2130 UTC 10 July 2012, and model output for the simulated Changma front, analyzed at the 3-h forecast valid at 2100 UTC. Z_{DR} and K_{DP} contain 0.3 dB and 0.02° km^{-1} thresholds, respectively, to suppress noise. Overall, Z_H shows widespread precipitation up to 55 dBZ, with rather smooth gradients (Fig. 2.2a). Z_{DR} (Fig. 2.2c) is generally below 2 dB, indicating the main precipitating hydrometeors are likely small to medium-sized. K_{DP} (Fig. 2.2e) is noisy over the radar coverage domain, except for high reflectivity cores in the south. Small areas of K_{DP} exceed 1° km^{-1} in the main precipitation cores (Z_H > 35 dBZ), indicating heavy precipitation. Observed ρ_{HV} is very high (Fig. 2.2g) over the entirety of the precipitation area, except for sparse reduced values in low signal-to-noise ratio regions. The overall large ρ_{HV} values indicate the precipitating system at low levels is dominated by the presence of pure rain.

The storm structure of the simulated Changma front (Fig. 2.2b) is more detached compared to observations. Isolated high Z_H cores with narrow stratiform rain are scattered within the radar coverage. Model underprediction of precipitation coverage is evident, as observed precipitation coverage (defined as $Z_H \geq 5$ dBZ) is 75% of the radar coverage area, compared to the model's 42% (Fig. 2.2b). From the microphysics perspective, the diagnostic intercept parameter is likely one of the main reasons for the fragmented storm organization. As rain mass decreases toward the storm edges, drop size decreases rapidly and inversely proportional to the mixing ratio, and becomes

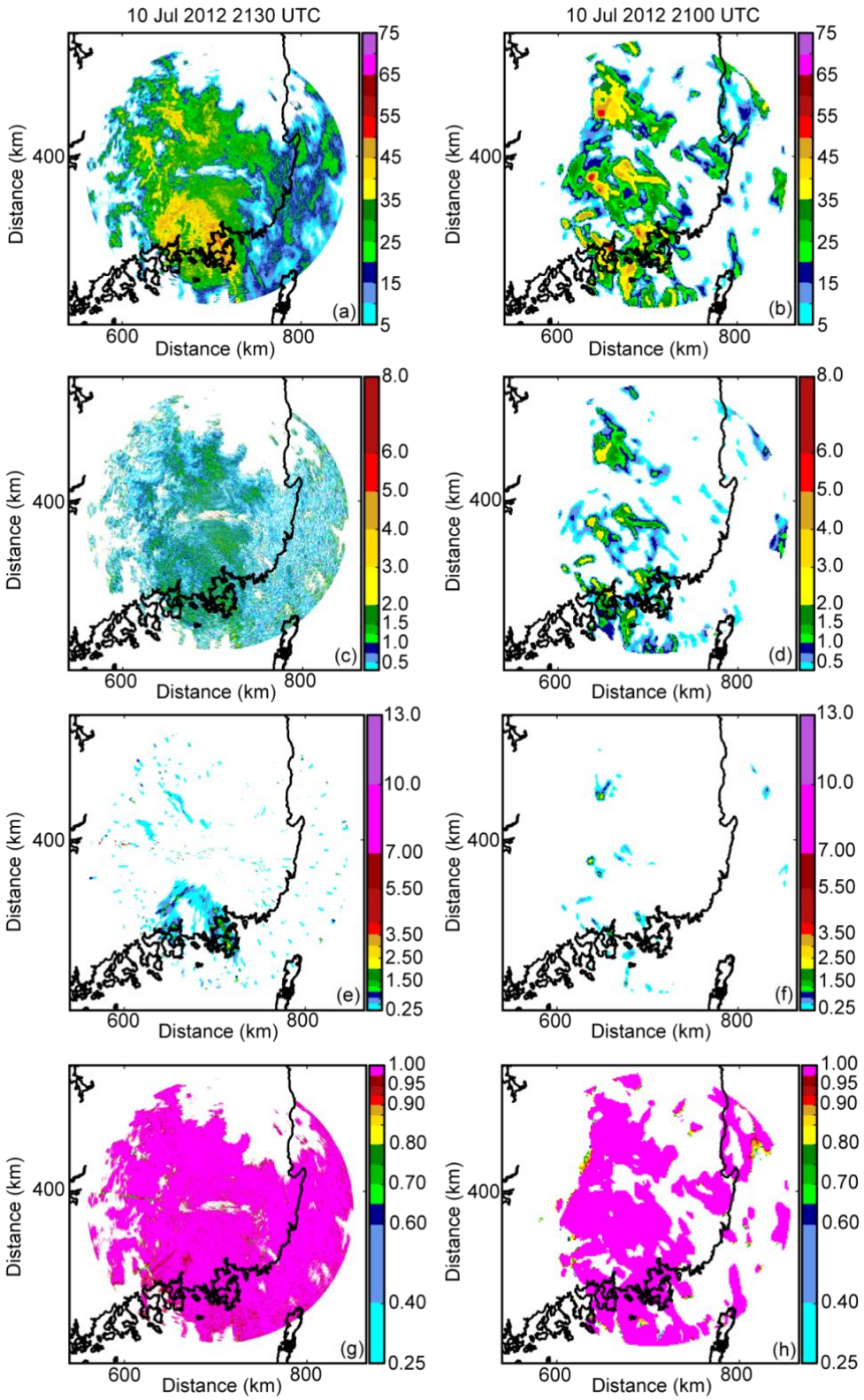

Fig. 2.2. The (a, b) Z_H (units: dBZ), (c, d) Z_{DR} (units: dB), (e, f) K_{DP} (units: ° km^{-1}) and (g, h) ρ_{HV} at 2130 UTC 10 July 2012 in observations (left-hand panels) and the model (right-hand panels) for the Changma front valid as a 3-h forecast at 2100 UTC 10 July 2012 at the 0.5° elevation angle.

increasingly prone to evaporation which is proportional to the rain intercept parameter. Subsequent timesteps with reduced rain mass result in an increased diagnosed rain N_0, which further increases evaporation and decreases drop size. Further, other studies using the UM with grid spacing smaller than a few km have identified the model's struggle to adequately resolve convection, both in size and intensity (e.g., Tang et al. 2013). There are other studies that have tried to attribute overly intense cores and a lack of precipitation coverage to the local non-conservation associated with semi-Lagrangian advection and/or deficits in the subgrid turbulence scheme (e.g., Hanley et al. 2015; Nicol et al. 2015). While those studies revealed sensitivity, the main issue of overly intense updrafts and too little light rain remained.

The model also produces higher Z_{DR} values (> 2 dB; Fig. 2.2d) in the precipitation cores compared to observations. Given the 1M nature of UM microphysics, D_{0r} is monotonically related to rain mass. Thus, increasing reflectivity (increasing rain rate) corresponds to increasing median drop size, resulting in larger Z_{DR} in the reflectivity cores. Outside of the cores, Z_{DR} is frequently below 1 dB, which is somewhat similar to observations. Similar to observed K_{DP}, significant K_{DP} values (Fig. 2.2f) are found only in the precipitation cores with large drops. Although the spatial coverage of simulated K_{DP} is significantly underpredicted similar to Z_H and Z_{DR}, the range of simulated K_{DP} (0.25–3° km^{-1}) is very similar to that of observations. The model ρ_{HV} (Fig. 2.2h) is generally near 1 over the analyzed domain, indicating the primary precipitating hydrometeor at low levels is rain, matching observations. The sparse model ρ_{HV} reduction is caused by interpolation error of model polarimetric variables to the radar elevation angle.

2.3.2 Typhoon Sanba (17 September 2012)

Typhoon Sanba (2012) made landfall on the southern coastline of the Korean peninsula. Surface charts at 0000 UTC 17 September 2012 reveal an intense low of 955 hPa with maximum

winds reaching 148 km h^{-1} as the typhoon's center was positioned just south of the Korean coastline (Fig. 2.3a). The UM forecast is able to simulate the large-scale structure of Typhoon Sanba (2012) reasonably well, capturing the location of the heavy precipitation in the eyewall and rainbands (Fig. 2.3b). The highest 12-h accumulated rainfall over the peninsula is concentrated over the south (Fig. 2.3c), where the eyewall made landfall. The maximum substantially exceeds that in the Changma front case. A few stations reported over 400 mm of rain accumulation, with one exceeding 500 mm. The typhoon weakened as it made landfall but heavy rain continued to cause substantial damage while the typhoon moved northeast.

The observations considered for this case are at 0100 UTC 17 September 2012. The eye of the typhoon is just south of the coastline, with the typhoon's rainbands in the north covering much of the radar coverage area (Fig. 2.4a). Z_H reveals widespread moderate to heavy precipitation. Further from the eyewall, precipitation becomes lighter at the edge of the radar coverage. It appears drops are relatively larger close to the typhoon's eye, where high Z_{DR} is found (Fig. 2.4c). Given the convective nature of the rainbands, drops have more potential to grow before falling out of the updrafts. Significant observed K_{DP} is found in the inner part of the eyewall, with maximum K_{DP} exceeding 2° km^{-1} (Fig. 2.4e). The ρ_{HV} is generally near 1 over the entire radar coverage area, which makes sense given the warm-rain processes that dominate typhoons (Fig. 2.4g).

The isolated convective cores seem less problematic for the UM simulated typhoon compared to the previous case, suggesting that the performance of the microphysics scheme may depend on the scale of the precipitation system (Fig. 2.4b). In fact, the microphysics scheme was originally developed for large-scale systems. The simulated high reflectivity near the coastline is consistent with observations that show the northern part of the eyewall/rainbands. Still, precipitation coverage is underpredicted, as the model precipitation encompasses 76% of the radar

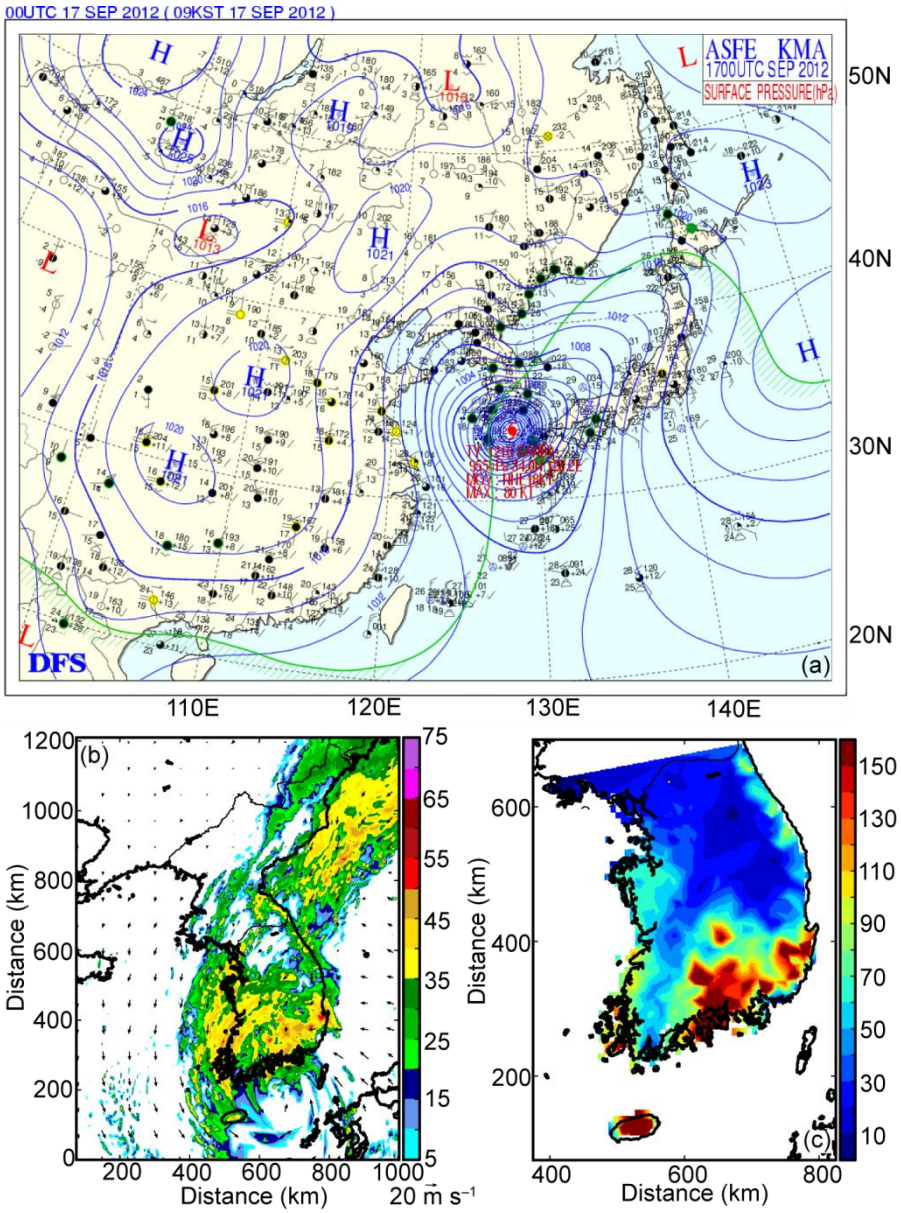

Fig. 2.3. The (a) surface chart at 0000 UTC 17 September 2012, (b) simulated reflectivity Z_H for the UM 6-h forecast valid at 0000 UTC 17 September 2012 and z = ~166 m, and (c) rain accumulation (units: mm) from AWS gauge data over 12 h ending at 0300 UTC 17 September 2012 for Typhoon Sanba.

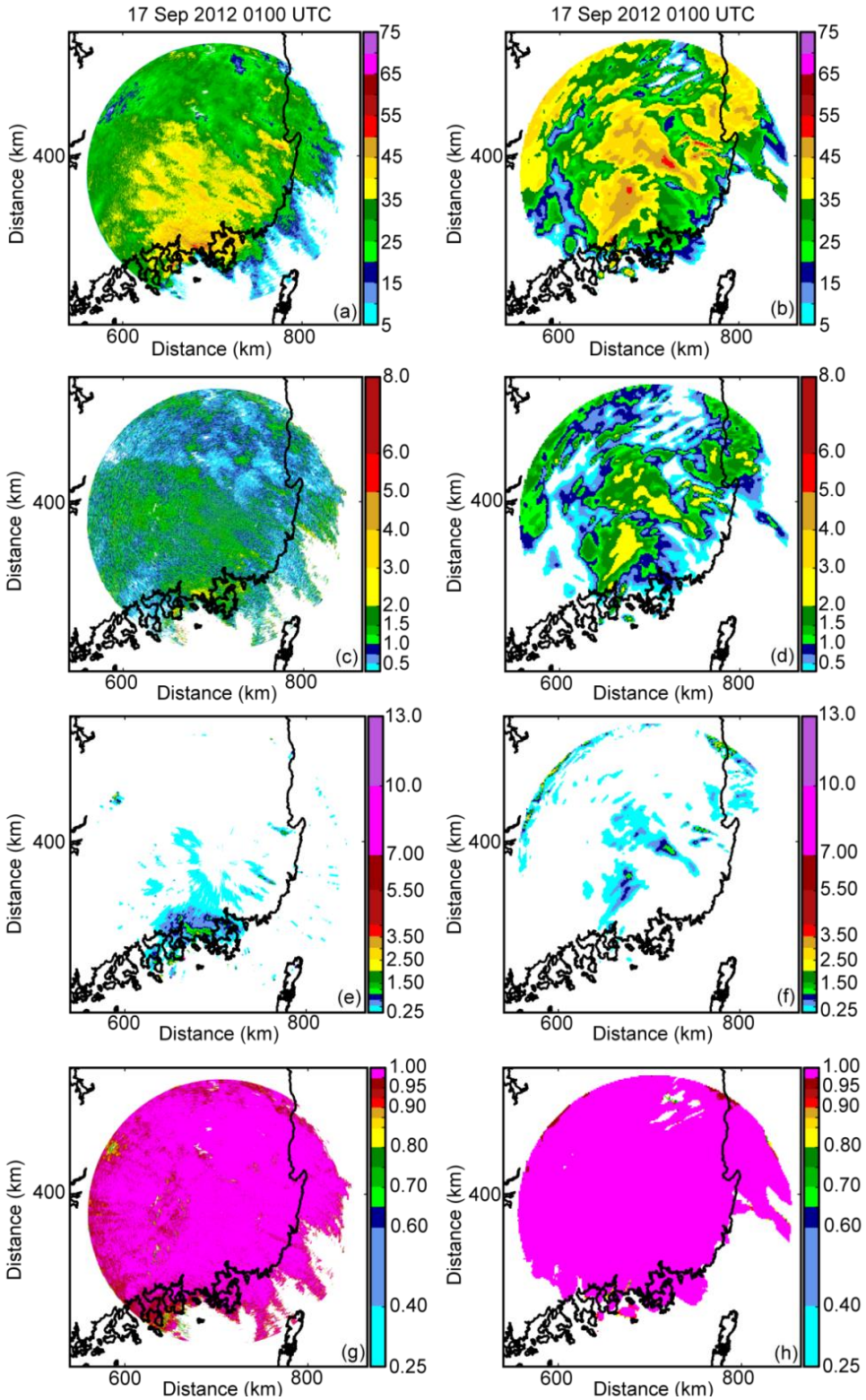

Fig. 2.4. The (a, b) Z_H (units: dBZ), (c, d) Z_{DR} (units: dB), (e, f) K_{DP} (units: ° km^{-1}) and (g, h) ρ_{HV} at 0100 UTC 17 September 2012 in observations (left-hand panels) and the model (right-hand panels) for Typhoon Sanba valid as a 7-h forecast at 0100 UTC 17 September 2012 at the 0.5° elevation angle.

coverage area compared to the observational 90%, possibly due to the N_0 relationship and semi-Lagrangian/subgrid turbulence dynamic reasons previously mentioned. While larger observed Z_{DR} appears in the inner side of the eyewall, where drops can grow large in strong convection, large model Z_{DR} coincides with high reflectivity throughout the domain (Fig. 2.4d). Enhanced model K_{DP} is also found further away from the eye, collocating with high reflectivity and Z_{DR} (larger drops; Fig. 2.4f) because of their monotonic relationships with q_r. The larger K_{DP} found near the edges of the radar domain is due to snow. The model ρ_{HV} is generally near 1, and only a minor reduction is found in sparse areas near the edges of the domain due to snow and graupel, and interpolation error (Fig. 2.4h). Thus, the primary model hydrometeor is pure rain, which matches observations and the expected hydrometeor behavior of a typhoon given its dominant warm-rain processes.

2.4 Polarimetric distributions

2.4.1 Changma front

In order to evaluate the ability of UM microphysics to capture the natural variation of PSDs, observed, model, and model rank histograms of polarimetric variables are constructed for the two test cases (Figs. 2.5, 2.6). All of the histograms use orange, green, blue, purple and violet shaded areas to denote the 0.2 percentiles in the distribution. The observation and model rank plots denote the observation 0.2 percentiles (to facilitate comparison between the model and observations), while the model histograms denote the model 0.2 percentiles. The observations and UM contain different resolutions and thus a different number of data points; therefore, a raw comparison is not valid. The model rank histograms are constructed by sorting observed data, finding locations of percentiles spaced at 0.1, and then distributing the model data into these observed percentiles. By analyzing how the model data fills the observed percentiles, a direct comparison between the

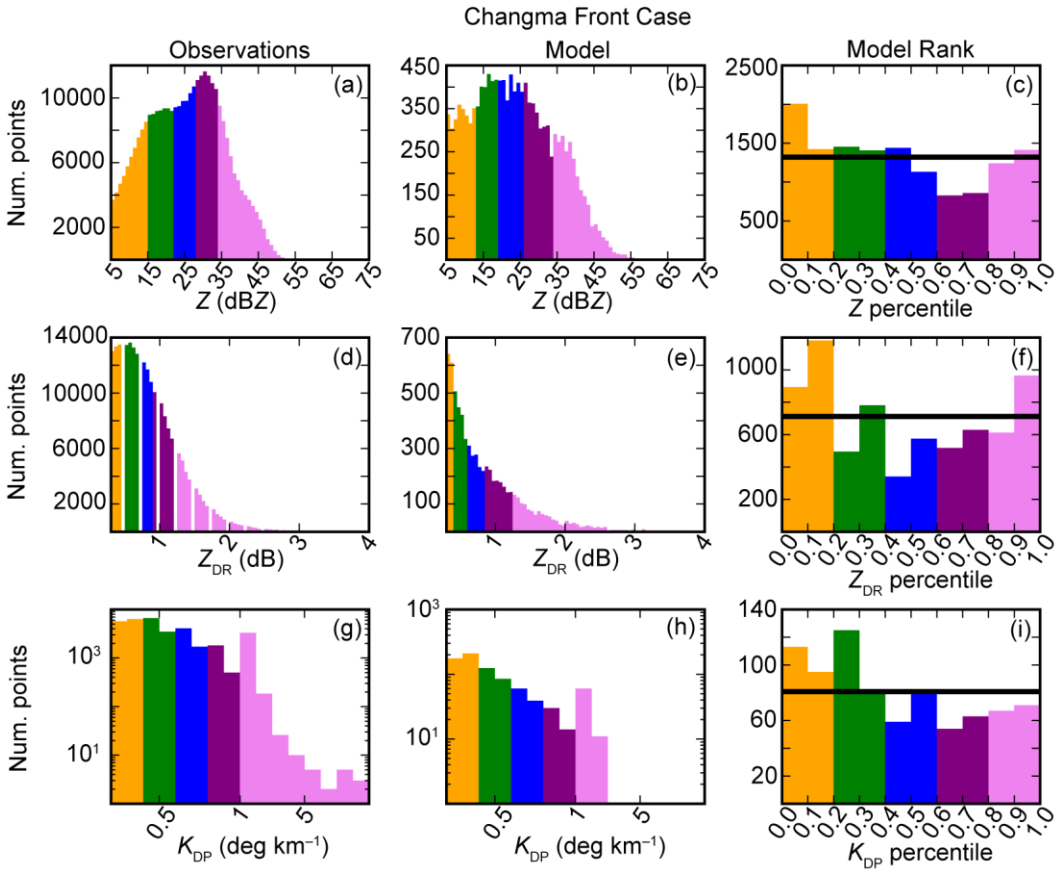

Fig. 2.5. Observed (left-hand panels), model (middle panels), and model rank (right-hand panels) of (a–c) Z (units: dBZ), (d–f) Z_{DR} (units: dB) and (g–i) K_{DP} (units: ° km^{-1}) for the Changma front case. Percentiles at 0.2 intervals are denoted by color shifts in the plots. The observation and model rank plots display observation percentiles, while the model plots show model percentiles. The black lines in the model rank column denote a theoretical UD in which model data are distributed in the same manner as observations.

model and observed distributions is possible, and model biases are readily apparent. The black solid lines in the plots represent model uniform distributions (hereafter UDs), which would occur if the model data perfectly matched the observed distribution. The 5-dBZ threshold in all plots, and additional 0.3-dB and 0.25° km^{-1} thresholds, are included in the histograms to filter noisy data. Additionally, observations are removed when the ρ_{HV} is less than 0.9, because precipitation is mostly pure rain at the 0.5° elevation angle. It is important to mention that comparisons among

cases should be taken with caution because of the small coverage of the radar used in this chapter. It only captures part of the precipitation, and therefore the results may not be representative of the entire storm system.

Observation and model reflectivity histograms are binned at 1 dBZ (Figs. 2.5a,b). The model Changma front reflectivity distribution tends to contain a larger frequency of smaller reflectivity values (< 15 dBZ) than observations, while missing the larger peak of observed reflectivity (~30 dBZ). This is also reflected in the model rank histogram, as the 0.0–0.5 percentile bins exceed the model UD and dip below this line between the 0.5–0.9 percentiles where the observation reflectivity peak is centered (Fig. 2.5c). The model rank histogram slightly rises above the UD line in the largest percentile bin, but there is a clear underprediction of overall reflectivity.

Observational Z_{DR} produces a smooth distribution at 0.05-dB bin intervals, albeit with several missing bins (Fig. 2.5d). This is due to observational Z_{DR} rounding, where observations are stored at 0.06–0.07-dB intervals. As a result, the missing bins repeat for bins ending at 0.25 dB intervals. One notable difference between the two histograms is that the model Z_{DR} frequency continually decreases with increasing Z_{DR} (Fig. 2.5e), while observed Z_{DR} peaks in the 0.2–0.4 percentile area. Similar to reflectivity, the model is producing more small Z_{DR} values compared to observations. This is reflected in the model rank Z_{DR} histogram, where the 0.0–0.2 percentiles exceed the UD (Fig. 2.5f). The repeating low–high step shape of the histogram is thought to be due to the observational Z_{DR} rounding previously mentioned. The model rank histogram also reveals a longer model Z_{DR} tail, as the 0.9–1.0 percentile exceeds the UD. Much of the middle Z_{DR} percentiles (0.2–0.8) are below the UD line, lending to the relatively higher number of low/high model Z_{DR} frequency. The observation and model K_{DP} histograms are similar in that small K_{DP} dominates the frequency (Figs. 2.5g,h). For this reason, the histograms are displayed logarithmically. Compared

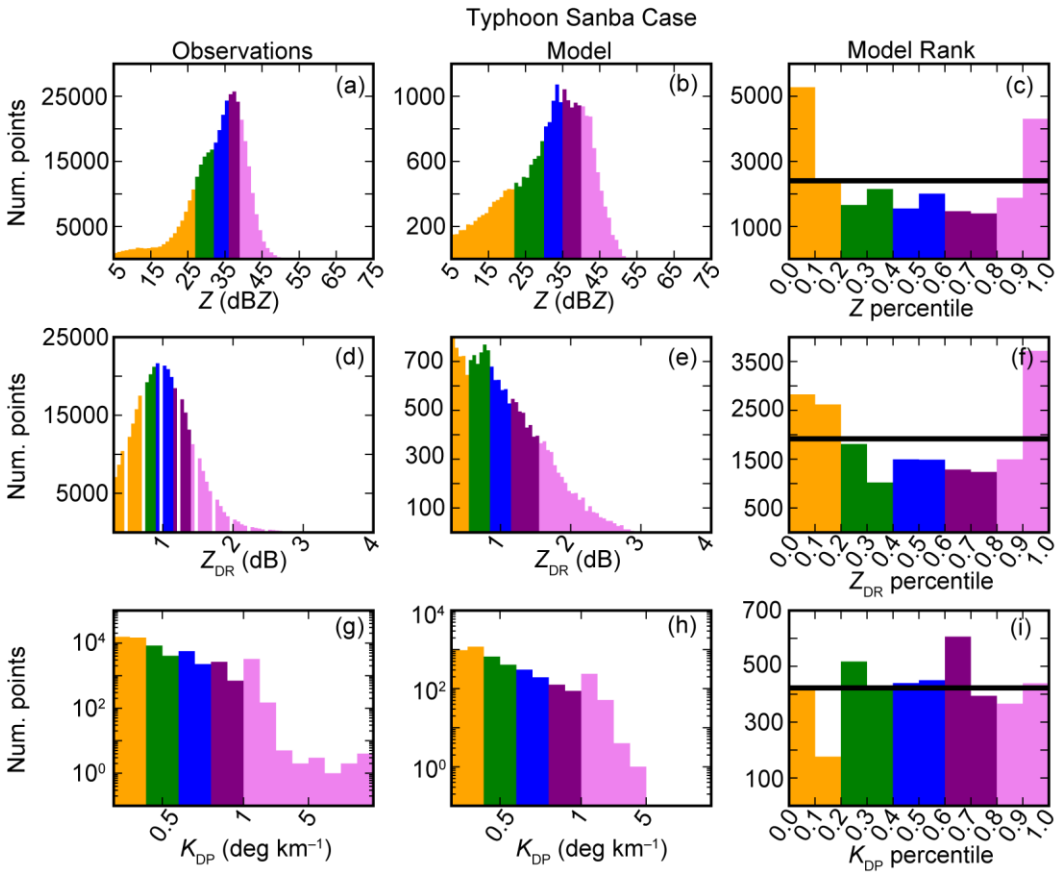

Fig. 2.6. As in Fig. 2.5 but for the typhoon case.

to the model, observations contain a much longer K_{DP} tail. This is reflected in the model rank histogram, where the 0.6–1.0 percentiles are below the UD (Fig. 2.5i). On the other hand, the 0.0–0.4 percentiles are at, or exceed, the UD. Combined with the Z and Z_{DR} histograms, large model drop size is primarily responsible for high Z and Z_{DR} percentiles, while the rainfall amount may be underestimated.

2.4.2 Typhoon Sanba

For Typhoon Sanba (2012), both the observation and model reflectivity distributions seem to be negatively skewed Gaussian (Figs. 2.6a,b), and have narrower distributions compared to the Changma case. However, the model Z contains higher frequencies of smaller and larger reflectivity compared to observations (Fig. 2.6c). A U-shaped model rank histogram is prominent in this case,

in which the smallest and largest percentiles exceed the UD, while middle percentiles stay under the line. Unlike reflectivity, the shapes of the Z_{DR} distributions are quite different. Observational Z_{DR} produces a smooth normal distribution with a peak at around 1 dB, which is larger than the above case. Conversely, the model Z_{DR} does not have a Gaussian distribution and peaks at smaller values (Figs. 2.6d,e). A U-shaped model rank histogram suggests that the model overpredicts the frequency of both the smallest and largest raindrops (Fig. 2.6f). The model Z_{DR} also clearly has a longer tail than observations. Similar to the Changma case, the largest Z and Z_{DR} values are rather small (Z < 55 dBZ and Z_{DR} < 3 dB) (Figs. 2.6b,e). In this Z_{DR} range, the size effect on K_{DP} is not dominant, and thus the model K_{DP} shows a shorter tail compared to observations (Figs. 2.6g,h). As a result, the model rank histogram shows a rather flat distribution (Fig. 2.6i).

2.5 Hydrometeor properties

2.5.1 Simulated vertical hydrometeor profiles

In this section, we expand the scope of the analysis to include upper levels, the aim being to examine the sensitivity of simulated frozen hydrometeors to precipitation systems. Vertical profiles of horizontally averaged hydrometeor water content (HWC) over grid points where HWC is greater than 0 kg m^{-3} are plotted for each case in Fig. 2.7. UM's generic ice category, which contains both ice crystals and snow, is generally favored over graupel at each height for the Changma front case (Fig. 2.7a). Graupel water content extends up to about 12 km. The UM's propensity for ice/snow over graupel is reasonable, as Changma fronts are less convective than systems that favor rimed ice, such as supercells or squall lines. The melting level seems to be near z = ~5 km, as ice water content quickly decreases near this level and rain water content increases. The model graupel did not reach the surface, suggesting complete melting before reaching the surface. Ice/snow is similarly favored over graupel at each height for Typhoon Sanba (2012)

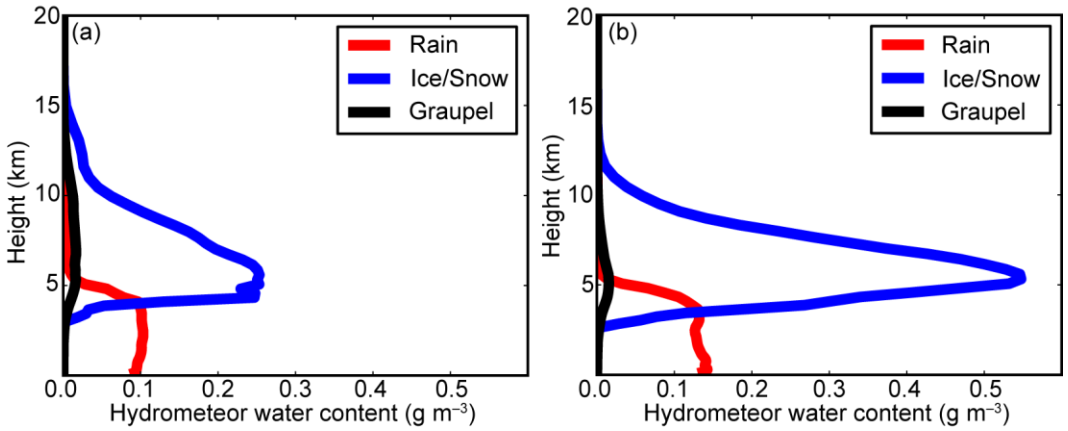

Fig. 2.7. Vertical plots of model rain, ice/snow and graupel horizontally averaged HWC (units: g m^{-3}) for the (a) Changma front and (b) Typhoon Sanba cases.

(Fig. 2.7b), but by far more than the Changma case. The ice/snow peak is more than twice that of the Changma front, exceeding 0.5 g m^{-3}. Graupel is found within a limited layer around the freezing level, where it can grow through riming. However, heavy graupel falls out quickly, while most ice produced above the freezing level comprises small ice particles and aggregates in the typhoon (Houze 2010). The typhoon melting level also appears to be near z = ~5 km, as frozen water content decreases and rain water increases below this height.

2.5.2 Hydrometeor classifications

Hydrometeor classification algorithms (HCAs, Park et al. 2009) are applied to the Changma front and Typhoon Sanba (2012) observations at the 1.6° elevation angle (Fig. 2.8), which captures higher altitudes (i.e., more frozen hydrometeors) than the elevation angle used in sections 2.3 and 2.4. The hydrometeor types included in this chapter are: ground clutter/anomalous propagation (GC/AP), biological scatterers (BS), dry snow/ice crystals (DS/CR), wet snow (WS), rain (RA), rain/rimed ice (RR), and rimed ice (RI). Readers are referred to Putnam et al. (2017b) for further details on the simulated and observed HCA diagnoses.

Observationally, lower levels of the Changma front are typically composed of rain (Fig. 2.8a). A melting layer transition region is mostly composed of rain, wet snow, and dry snow to the east of the radar. To the south, the transition zone is deeper and consists of two layers: an upper layer with rain, dry snow, wet snow and rimed ice; and a lower layer with rain and rain/rimed ice. Sparse melting is found to the west. Finally, in the upper levels of the radar scan, crystals and dry snow are prominent. Similar to observations, lower levels in the UM are typically composed of rain (Fig. 2.8b). Model transition regions are primarily composed of rain/rimed ice and wet snow in all directions. Rimed ice is most prominent to the east of the radar domain, but also present to the north and south. In observations, it is typically contained to both the west and south of the radar. This suggests that the model tends to overpredict the presence of graupel compared to observations. The upper levels of the simulated Changma front in the radar domain are typically dry snow/ice crystals, which matches observations well. Still, the presence of rimed ice above the melting layer is greater than in the observations.

Typhoon Sanba (2012) is mostly rain at lower levels (Fig. 2.8c). This lower-level rain coverage is smaller than that in the Changma front case because of the autumn season. Distinct melting occurs to the west of the radar site, with rain, dry snow and wet snow populating these regions, along with rimed ice. Elsewhere, hydrometeors are typically rain and wet snow in the melting layer. Heights above the melting layer are primarily composed of ice crystals and dry snow. Simulated UM hydrometeors are typically rain at low levels (Fig. 2.8d), in agreement with the observed predominant hydrometeor type. The melting transition region between frozen and liquid hydrometeors is primarily composed of rimed ice and wet snow in all directions, which is similar to the model Changma case. Model levels above the transition region are primarily composed of dry snow/ice crystals, in agreement with observations. Similar to the Changma case, rimed ice

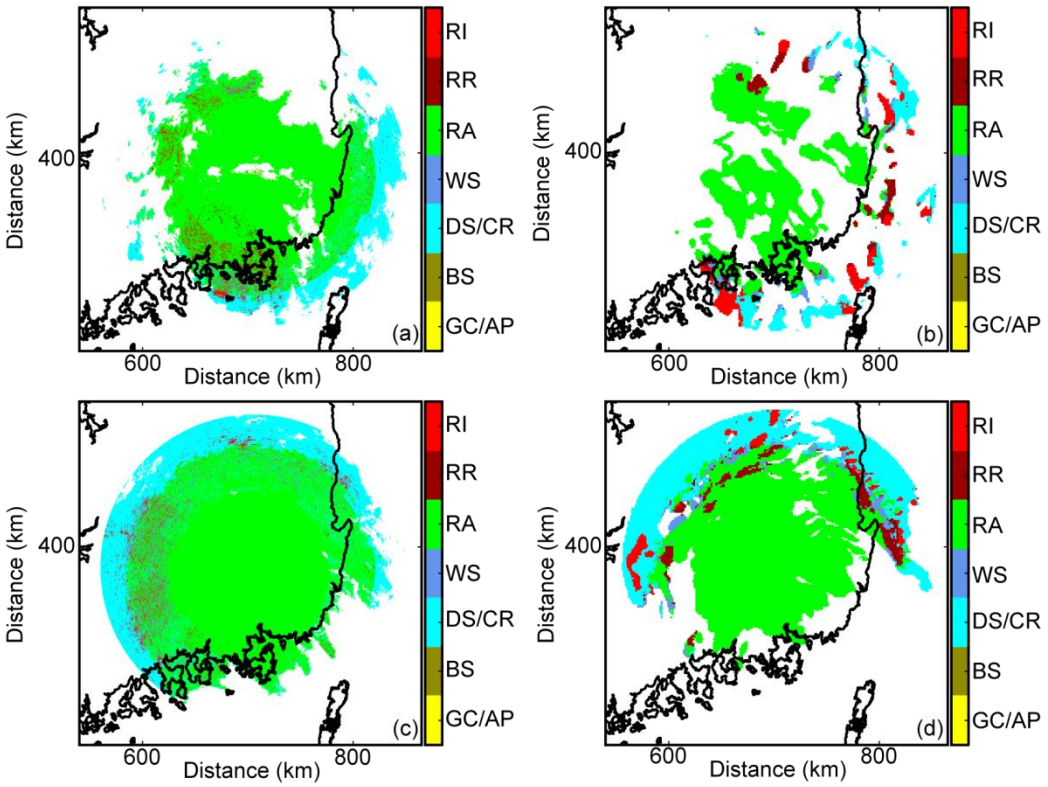

Fig. 2.8. HCAs for the (a, c) observations and (b, d) model in the (a, b) Changma front case and (c, d) typhoon case at the 1.6° elevation angle. The hydrometeors considered are: ground clutter/anomalous propagation (GC/AP), biological scatterers (BS), dry snow/ice crystals (DS/CR), wet snow (WS), rain (RA), rain/rimed ice (RR), and rimed ice (RI).

populates upper levels more frequently than observations.

2.6 Summary and discussion

This chapter examines UM microphysics for two convective cases at 1.5-km grid spacing over the Korean peninsula: a Changma front, and Typhoon Sanba (2012). Simulated polarization radar variables are compared to observations from the S-band Biseul-san radar. Clearly, the model struggles with convection, as reflectivity Z_H gaps are present in each case. The consequences of inadequately resolving convection are significant to the model's forecast, as the model is underpredicting precipitation coverage. The diagnostic relation between the rain intercept parameter and mixing ratio results in a rapidly increasing (decreasing) drop size for an increasing

(decreasing) mixing ratio. This is partially responsible for the large Z_H gradients present in the cases compared to observations. The ability to correctly predict PSDs in microphysics schemes is potentially important, as they have a significant impact on the evolution of precipitating systems through their feedback to thermodynamics and dynamics. Thus, it is imperative to understand microphysics biases and address them, especially for convective-scale modeling where the microphysics error dominates the forecast errors. Quantitatively, the model generally produces more small and large Z_{DR} than observations, which is consistent with the diagnostic relationship discussed above. K_{DP} is less affected by the diagnostic relationship, as it is proportional to a lower order moment than Z and Z_{DR}. Thus, the model drop size bias should be taken into account when interpreting simulated radar variables. The UM's generic ice category, which contains both crystals and snow, overshadowed graupel as the dominant hydrometeor, with neither particle type appreciably reaching the surface. This is reasonable because neither storm system has a strong updraft. Still, graupel might be incorrectly parameterized, as the model graupel is more prominent above the melting level compared to observations.

Many UM microphysics shortcomings in this chapter stem from the 1M nature of the scheme. Schemes with 1M categories have the potential to perform reasonably well for large-scale storm systems (Liu et al. 2011). In that regard, the UM microphysics raindrop size distribution is tuned with aircraft observations collected from stratocumulus and trade-wind cumulus (Abel and Boutle 2012), which may not be adequate for midlatitude deep convective systems where the Korean peninsula is located. As a result, the model struggles with typical mid-latitude weather systems such as the Changma front in this chapter. An incorrect N_0 parameterization by many orders of magnitude can adversely affect the rain PSD and its moments, which is directly linked to many microphysical processes. Further, radar variables are sensitive to drop size distributions.

For example, radar reflectivity is dependent on drop size to the sixth power, and differential reflectivity is related to the axis ratio of hydrometeors. As radar observations are typically used to monitor severe weather, large biases in simulated radar variables can misguide forecasters as well as introduce large errors in assimilation. A well-calibrated regional rain N_0 relationship could result in improved model microphysics and forecasts. For improved flexibility and model performance, it is desirable that a 2M version of the UM microphysics scheme be developed. As an example, the 2M version would potentially be able to simulate both heavy rain dominated by small drops with relatively uniform size in warm-rain processes, and heavy rain with many large raindrops in cold-rain processes. While adjusting the rain N_0 configuration for different weather systems could mitigate this problem to some degree, evolving the scheme to 2M would allow more freedom to simulate a wider range of weather systems.

Chapter 3 [2] Evaluation of bulk rimed ice parameterization and its influence on simulated supercell polarimetric signatures

3.1 Introduction

Cloud microphysics is the collection of all hydrometeor processes spanning cloud particle formation to precipitation. These processes are very complex and highly non-linear. Liquid phase processes are less complex and drops are well-represented as constant density spheres, with the caveat that larger rain drops do slightly deviate from this approximation with decreasing axis ratio (e.g., Pruppacher and Pitter 1971; Chandrasekar et al. 1988; Brandes et al. 2002). On the other hand, ice processes lack adequate observations and theoretical understanding that can explain their highly variable evolution in shape and density depending on temperature and ice supersaturation. This deficiency in complex ice phase, habit, and evolution information inevitably leads to large ice microphysical parameterization uncertainty in numerical weather prediction (NWP) models. Since ice processes play a crucial role in simulations due to their impact on the radiation budget, thermodynamic fields, and storm dynamics, there have been many efforts to improve microphysical parameterization schemes in atmospheric models (e.g., Johnson et al. 1993; Liu et al. 2007; Milbrandt and Morrison 2013).

Although spectral bin microphysics schemes (SBMs) allow greater particle size distribution (PSD) characteristic flexibility (e.g., Takahashi 1976; Khain et al. 2004), current NWP models continue to employ bulk microphysics schemes (BMPs) because of their significant computational cost advantage and current lack of clear SBM advantage in convective simulation skill (e.g., Fan et al. 2017; Xue et al. 2017). Rather than predicting the binned PSD as in SBMs,

[2] This chapter has been conditionally accepted for publication as: Johnson, M., Y. Jung, J. A. Milbrandt, H. Morrison, and M. Xue, 2019: Effects of the representation of rimed ice in bulk microphysics schemes on polarimetric signatures. *Mon. Wea. Rev.*

BMPs assume analytic PSDs, general gamma distributions, and predict moments of the distribution. Predicted moments update model information because they are typically linked to physical quantities, e.g., number concentration is the 0^{th} moment of a PSD. BMPs are typically categorized by their number of predicted moments, i.e., one-moment schemes generally predict mass mixing ratio q_x, two-moment schemes add complexity by further predicting number mixing ratio (N_x), and so on. Because two-moment schemes add particle distribution information, they can provide an improvement over one-moment schemes in convective scale simulations in terms of cold pool structure, simulated polarimetric signatures, and hydrometeor size sorting (e.g., Milbrandt and Yau 2005a; Dawson et al. 2010; Jung et al. 2010; Jung et al. 2012; Putnam et al. 2017a), although the two-moment improvement is less clear for simulated cloud structure and surface precipitation (e.g., Wang et al. 2009; Varble et al. 2011; Van Weverberg et al. 2014). Three-moment schemes that additionally predict reflectivity Z have also been developed (Milbrandt and Yau 2005a, b; Dawson et al. 2014), and simulate hydrometeor size sorting more accurately than two-moment schemes with a fixed shape parameter (e.g., Milbrandt and Yau 2005a; Kumjian and Ryzhkov 2012; Dawson et al. 2014).

Traditional BMPs represent the wide variety of ice habits, with different shapes and densities, using two or more predefined ice hydrometeor types (i.e., pristine ice, snow, graupel, and/or hail). Specifically, the number of rimed-ice categories and their corresponding density and fall speed assumptions can differ considerably. Some BMPs contain a single rimed-ice category, employing either intermediate-density graupel or high-density hail (e.g., Wisner et al. 1972; Rutledge and Hobbs 1984; Ziegler 1985; Murakami 1990; Tao and Simpson 1993; Thompson et al. 2008; Morrison et al. 2009), while others contain both (e.g., Ferrier 1994; Walko et al. 1995; Milbrandt and Yau 2005b; Mansell et al. 2010). In regard to rimed-ice parameterization, Morrison

and Milbrandt (2011) observed weaker cold pools, less surface precipitation, and enhanced anvils using two-moment graupel-like rimed ice rather than two-moment hail-like rimed ice in idealized supercell simulations. Newer BMPs have incorporated additional complexity by predicting rimed-ice particle volume in order to predict particle density (Mansell et al. 2010; Milbrandt and Morrison 2013), which might simulate supercells more accurately (Johnson et al. 2016). Recently, Morrison and Milbrandt (2015) proposed a new approach to represent ice-phase hydrometeors in microphysics schemes as "free" ice phase categories, rather than partitioning ice into several pre-defined categories, which requires assumptions about physical properties, the parameterization of ad hoc artificial conversion between categories, and other limitations.

Several previous studies have shown that simulated polarimetric variables can be used in evaluating microphysics scheme accuracy (e.g., Jung et al. 2010; Putnam et al. 2014; Johnson et al. 2016; Putnam et al. 2017a; Putnam et al. 2017b; Johnson et al. 2018), and are especially useful in a supercell framework due to their sensitivity to rimed-ice treatment. Dawson et al. (2014) showed that the Z_{DR} arc, a low-level maximum of differential reflectivity on the supercell's edge caused by the size sorting of rain and rimed ice (Ryzhkov et al. 2005a; Kumjian and Ryzhkov 2008, 2009), might be more sensitive to the parameterization of rimed ice rather than rain. Further, Johnson et al. (2016) found that the hail signature in the forward flank downdraft (FFD), a region of low differential reflectivity near the surface due to dry, tumbling hail (Kumjian and Ryzhkov 2008), was best simulated when the medium-density rimed-ice category (i.e., graupel) acts as a feeder category to the high-density rimed-ice category (i.e., hail).

This chapter attempts to evaluate two different approaches to representing frozen hydrometeors (i.e., predefined and free ice-phase categories) in idealized supercell simulations, specifically the performance of the P3 scheme. This study focuses on a supercell storm because of

its well-documented observed polarimetric signatures (e.g., Ryzhkov et al. 2005a; Kumjian and Ryzhkov 2008; Romine et al. 2008). As rimed ice has a greater influence on the polarimetric signatures near the surface than pristine ice and snow, discussion will mainly focus on rimed ice. Two two-moment BMPs with two traditional rimed-ice categories (graupel and hail), the Milbrandt-Yau (MY2) and National Severe Storms Laboratory (NSSL) schemes in the Weather Research and Forecasting (WRF) model (Skamarock et al. 2008) are examined due to their improved polarimetric signature accuracy relative to other two-moment BMPs with a single rimed ice category (Johnson et al. 2016). The P3 BMP contains at least one free ice category, which, under appropriate conditions, can represent rimed ice (with a wide range of densities). Such a study provides useful information to BMP developers on how to improve microphysical representations and helps modelers make educated selections of microphysical schemes for their applications.

The remainder of this chapter is organized as follows. Section 3.2 details the numerical model setup, BMPs, and polarimetric radar data simulator; section 3.3 analyzes simulated polarimetric variables; section 3.4 quantifies rimed-ice properties in the BMP schemes; and section 3.5 summarizes and discusses ramifications of rimed-ice parameterization in supercell simulations.

3.2 Numerical simulations

3.2.1 Model setup

WRF v3.9.1 (Skamarock et al. 2008) is used in this study to produce idealized supercell simulations. The model configuration is similar to Morrison and Milbrandt (2011), and is further detailed in Table 3.1. Convection is initiated with a thermal bubble with potential temperature perturbation θ' of 3 K on a 200 x 200 km grid with horizontal grid spacing of 1 km. The Weisman and Klemp (1982) thermodynamic sounding and a wind profile with a veering "quarter-circle"

Table 3.1. WRF model input.

WRF Model Configuration	
Run time	180 min
Δt	6 s
Sound wave Δt	1 s
Model output interval	10 min
Horizontal domain	200 km x 200 km
Model lid	20 km
Δx	1 km
Δy	1 km
Δz	~500 m
Time integration scheme	3rd order Runge-Kutta
Horizontal momentum advection	5th order
Vertical momentum advection	3rd order
Horizontal scalar advection	5th order
Vertical scalar advection	3rd order
Upper level damping	5000 m below model top
Rayleigh damping coefficient	0.003
Turbulence	1.5 order turbulent kinetic energy (TKE)
Horizontal boundary conditions	Open

shear of 3.11×10^{-3} s^{-1} up to 2.3 km and unidirectional shear of 5.70×10^{-3} s^{-1} above to 7 km defines the storm's environment. The supercell is integrated for 3 h with 6 s time step size using the third-order Runge-Kutta scheme. The domain contains a rigid lid at z = 20 km with 40 vertical levels. Rayleigh damping is applied in the top 5 km of the domain, and a 1.5 order turbulent kinetic energy (TKE) scheme is employed for turbulence. As is typical for idealized supercell simulations, we neglect radiation, surface flux, planetary boundary layer, and cumulus parameterizations.

3.2.2 Microphysics schemes

Based on the previous study of Johnson et al. (2016) that showed a two-moment BMP with both predefined graupel-like and hail-like rimed ice categories produced a more realistic supercell

structure compared to other two-moment schemes with one rimed ice category, this study uses the two-moment Milbrandt-Yau (MY2, to distinguish from the one-moment and three-moment versions of the scheme; Milbrandt and Yau 2005a, b) and National Severe Storms Laboratory (NSSL; Mansell et al. 2010) schemes. The MY2 and NSSL schemes are similar in that the mixing ratio q_x and total number mixing ratio N_x are prognosed for each hydrometeor category; i.e., cloud water, rain, cloud ice, snow, graupel, and hail. The default MY2 BMP employed here assumes constant densities of 400 kg m^{-3} and 900 kg m^{-3} for graupel and hail, respectively. Recently, Milbrandt and Morrison (2013) developed a version of the MY2 scheme that prognoses the graupel bulk volume mixing ratio, which allows the graupel density to be predicted; however, this version of the scheme was not available in WRF at the time of this study. The NSSL scheme prognoses both graupel and hail bulk volumes (and hence densities).

This study also uses the Predicted Particle Properties (P3) scheme (Morrison and Milbrandt 2015; Milbrandt and Morrison 2016). The P3 scheme has a similar two-moment two-category representation of the liquid phase as MY2 and NSSL, but has a fundamentally different representation of ice-phase particles, with a user-specified number of "free" ice categories represents all ice particles. For each ice category n, there are four prognostic variables: the total ice mass mixing ratio $q_{i\_tot}(n)$, the rime mass mixing ratio $q_{i\_rim}(n)$, the total number mixing ratio $N_{i\_tot}(n)$, and the rime volume mixing ratio, $B_{i\_rim}(n)$, and a complete gamma size distribution describes each category. This specific choice of prognostic variables allows for the smooth evolution of several important physical properties, such as bulk density, size, fall speed, etc. From the specific properties at a given point in time and space, the dominant type of ice particle for a given category can be determined (e.g. lightly-rimed aggregate), and this can evolve as the properties change. However, there is no "conversion" from one type of particle to another, as in

traditional predefined category schemes like MY2 and NSSL. If the properties of two P3 categories are deemed to be sufficiently similar, based on mean size as a proxy for overall similarity in the current version, ice categories merge into a single category in order to "free up" a category into which new ice can subsequently be initiated. Note that the use of mean size as a proxy for similarity has some utility (see Milbrandt and Morrison 2016) but is limited; the use of other physical properties (such as bulk density and rime fraction) to determine conditions for merging will be examined in future studies. Even with a single category, the P3 scheme has been shown to be able to simulate realistically a wide range of types of ice and performs well for the simulation of deep convective storms and orographic precipitation compared to traditional schemes in WRF (Morrison et al. 2015).

3.2.3 Polarimetric radar data simulator

To facilitate comparisons in BMP rimed-ice parameterization, we compute simulated polarimetric radar variables from WRF model output. The Center for Analysis and Prediction of Storms – Polarimetric Radar data Simulator (CAPS-PRS; Jung et al. 2008; Jung et al. 2010; Dawson et al. 2014) utilizes pre-computed scattering amplitude tables created by the T-matrix method (Waterman 1969; Vivekanandan et al. 1991; Zhang et al. 2001) that vary with particle diameter, water fraction, and now particle density. In the two-moment framework of this paper, the CAPS-PRS uses model prognosed q_x and N_{tx} to construct hydrometeor PSDs, with the shape parameter α consistent with microphysical assumptions. The CAPS-PRS calculates snow water fraction following Jung et al. (2008), which also details the axis ratio/canting angle assumptions made by the PRS. The PRS diagnoses rimed-ice water fraction following Dawson et al. (2014). Briefly, rain soaks rimed ice (and freezes) until its density reaches 910 kg m^{-3}, after which subsequent water collects as liquid on the rimed ice hydrometeor. Depending on the particle's

diameter, the wet rimed ice either completely melts or forms a liquid water torus (Rasmussen and Heymsfield 1987). We note here that the water fraction decreases and canting angle standard deviation increases as rimed ice diameter increases (see Dawson et al. 2014 Fig. 2), which allows the PRS to model both stable, small wet rimed ice and tumbling, relatively dry rimed ice. This added water reshapes the distribution, and iterates until the distribution converges. Finally, the PRS calculates horizontal reflectivity Z_H, differential reflectivity Z_{DR}, and cross-correlation coefficient ρ_{HV} from the relevant indexed scattering amplitudes following Zhang et al. (2001) and Jung et al. (2010). Currently, we neglect cloud water and cloud ice in the CAPS-PRS due to their small sizes.

We list where the CAPS-PRS differs from the above description when calculating P3 scattering amplitudes here. Since P3 microphysics employs a lookup table approach to calculate ice PSD parameters (see Morrison and Milbrandt 2015 for details), we utilize the same approach in the CAPS-PRS. The PRS calculates P3 ice PSDs for two rime fractions f_r, rime densities ρ_r, and normalized mass mixing ratios q_i/N_{ti} by accessing pre-computed intercept N_0, slope Λ, and shape α parameters, which allows for the calculation of linear polarimetric variables as a function of the simulated rime fraction, rime density, and normalized mass mixing ratio through multi-dimensional linear interpolation at each grid point. The ice PSDs are appropriately partitioned into m-D relationships (i.e., ice mode) by accessing pre-computed critical diameters. Small ice and larger, non-spherical ice from depositional growth and/or aggregation of small ice (which constitutes the smallest ice PSD sections when riming is present) follow Rayleigh scattering and are assumed dry, while fully- and partially-rimed ice follow the same T-matrix scattering method mentioned above. Water fraction for rimed ice follows Dawson et al. (2014), except without iteration as added water is assumed to not reshape the distribution. The PRS trilinearly interpolates

linear horizontal reflectivity Z_h, vertical reflectivity Z_v, and the correlation between horizontal and vertical reflectivity Z_{hv} for ice (or water from fully melted ice) calculated for each f_r, ρ_r, and q_i/N_{ti} combination to compute final linear ice polarimetric contributions. The PRS iterates P3 rain PSDs from model prognosed q_r and N_{tr} as the rain size distribution shape parameter α is diagnosed from rain slope parameter Λ following Cao et al. (2008). Then, horizontal reflectivity Z_H, differential reflectivity Z_{DR}, and cross-correlation coefficient ρ_{HV} are calculated from their relevant linear components.

3.3 Simulated supercell radar structure

We examine simulated dual-pol variables at t = 100 min, when the supercell has reached a stable, mature state. We note here that the polarimeteric supercell signatures examined in this paper are persistent at sampled model times after t = 100 min, as they often are in observations (though observed polarimetric signatures are subject to environmental conditions that influence signature evolution, such as updraft intensity strengthening a ZDR column, or hail disruption of a ZDR arc; e.g., Kumjian and Ryzhkov 2009; Palmer et al. 2011; Van Den Broeke 2016). Because of the veering wind sounding, subsequent microphysical analysis focuses on the dominant right-moving cell (e.g., Klemp and Wilhelmson 1978; Rotunno and Klemp 1982; Klemp 1987). Reflectivity plots at the lowest model level (z = ~280 m) for each BMP contain a vertical black line denoting the location where later vertical cross sections are taken, coinciding with the location of maximum vertical velocity (i.e., updraft). The MY2 simulation has somewhat noisy reflectivity (although the melting option used in this paper amplifies reflectivity noise, which depletes rain without a compensating reflectivity increase from wet rimed ice), but the overall structure is well-defined (Fig. 3.1a). Rain, rather than rimed ice, (Figs. 3.2a, 3.3a, 3.4a) dominates horizontal reflectivity Z_H in the forward flank. Z_H peaks above 60 dBZ in the reflectivity core near the updraft, which is

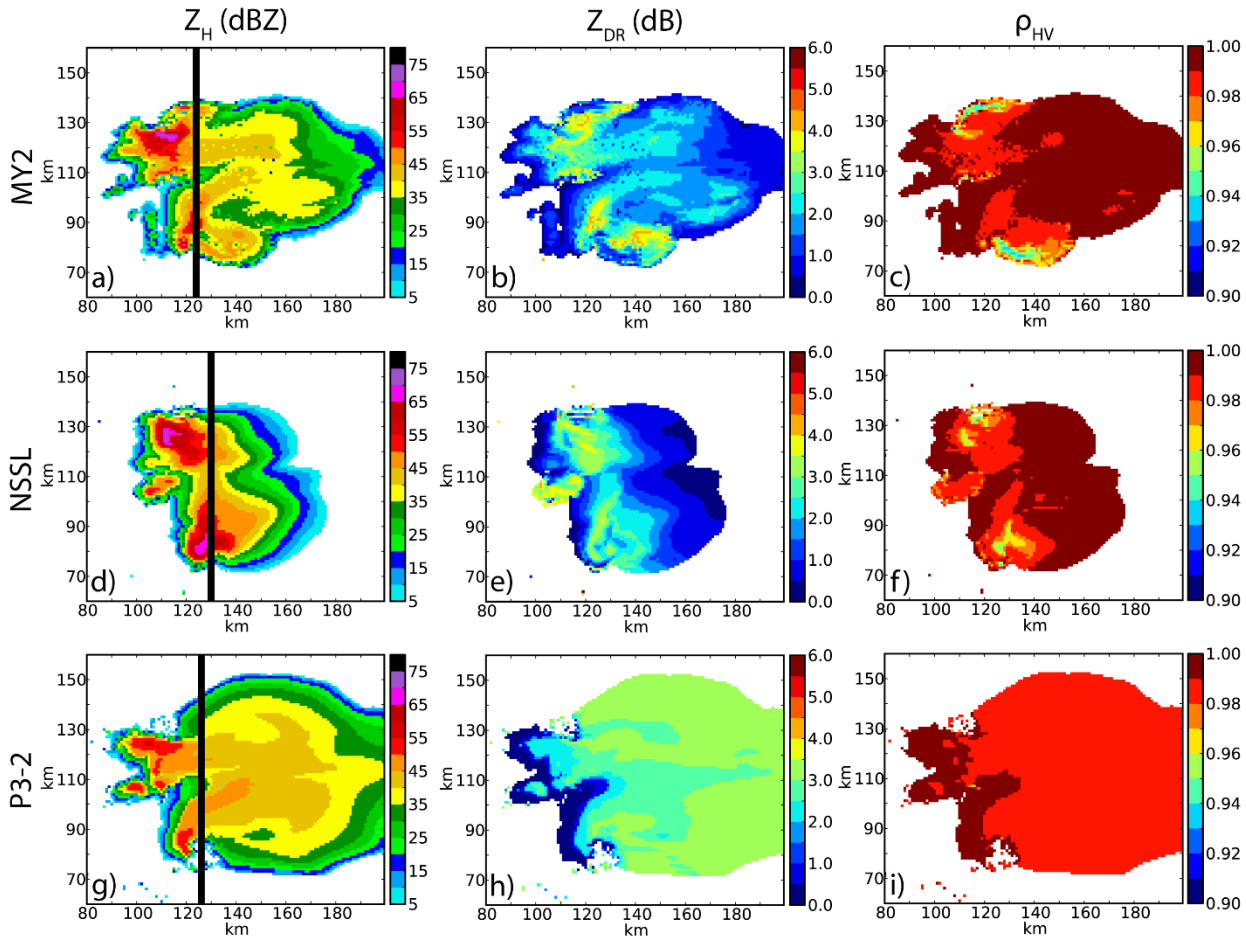

Fig. 3.1. Horizontal reflectivity Z_H (dBZ), differential reflectivity Z_{DR} (dB), and cross-correlation coefficient ρ_{HV} near z = ~280 m for the (a,b,c) MY2, (d,e,f) NSSL, and (g,h,i) P3-2 microphysics schemes at t = 100 min. Vertical black lines in the reflectivity plots denote where vertical cross sections are taken.

dominated by a small amount (hail mixing ratio $q_h \leq 0.50$ g kg^{-1}) of relatively wet (hail water fraction f_{wh} ranging between 0.5 and 1), large hail (hail mass-weighted mean diameter D_{mh} ranging from 4 to 14 mm; Figs. 3.4a,c,e) rather than graupel (Figs. 3.3a,c,e). The NSSL scheme produces the smallest forward flank among the three simulations (Fig. 3.1d), and also the largest reflectivity core, with Z_H exceeding 65 dBZ. Large hail ($q_h \geq 0.5$ g kg^{-1}; $D_{mh} \geq 14$ mm) in the NSSL scheme with varying degrees of wetness (Figs. 3.4b,d,f) dominates the reflectivity core over graupel (Figs. 3.3b,d,f). Z_H in the forward flank follows raindrop size closer than rain mixing ratio, which in the

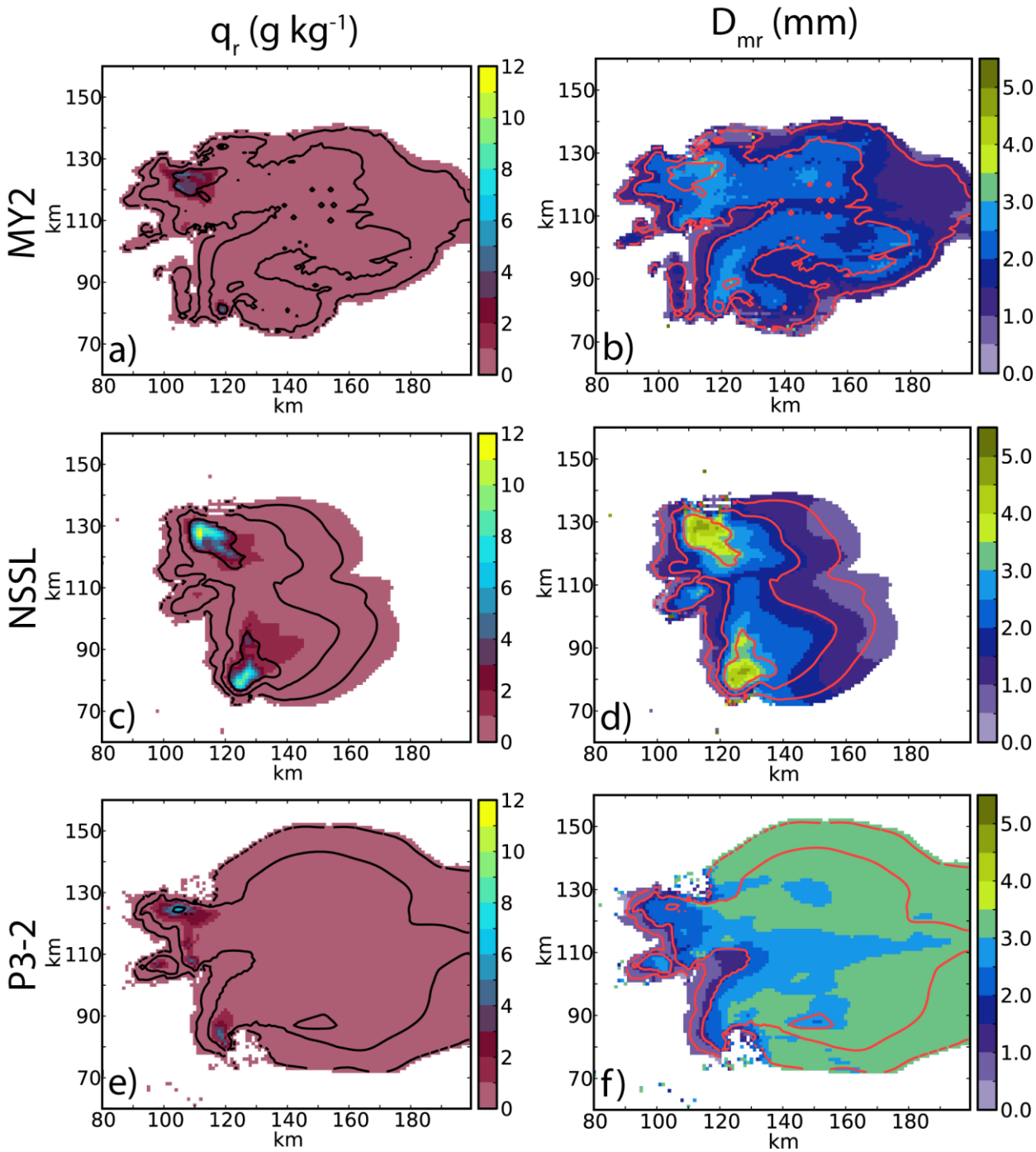

Fig. 3.2. Rain mixing ratio q_r (g kg^{-1}) and rain mass-weighted mean diameter D_{mr} (mm) near z = ~280 m for the (a,b) MY2, (c,d) NSSL, and (e,f) P3-2 BMPs at t = 100 min. Horizontal reflectivity Z_H contours are overlaid in 20 dBZ intervals starting at 15 dBZ.

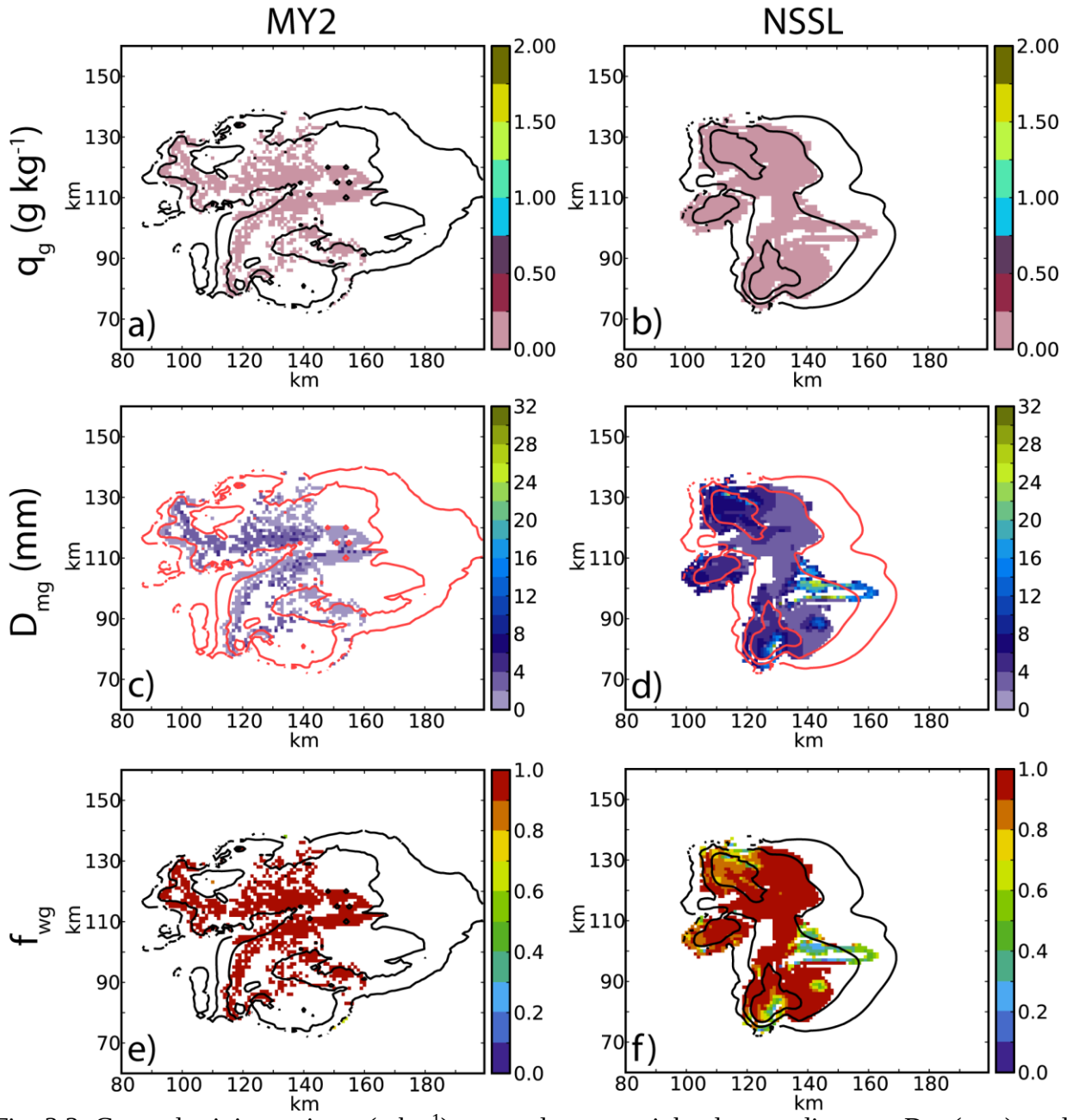

Fig. 3.3. Graupel mixing ratio q_g (g kg^{-1}), graupel mass-weighted mean diameter D_{mg} (mm), and graupel water fraction f_{wg} near z = ~280 m for the (a,c,e) MY2 and (b,d,f) NSSL BMPs at t = 100 min. Horizontal reflectivity Z_H contours are overlaid in 20 dBZ intervals starting at 15 dBZ.

NSSL BMP at this time typically ranges from low to moderate (rain mass-weighted mean diameter D_{mr} between 0 and 3 mm; Figs. 3.2c,d).

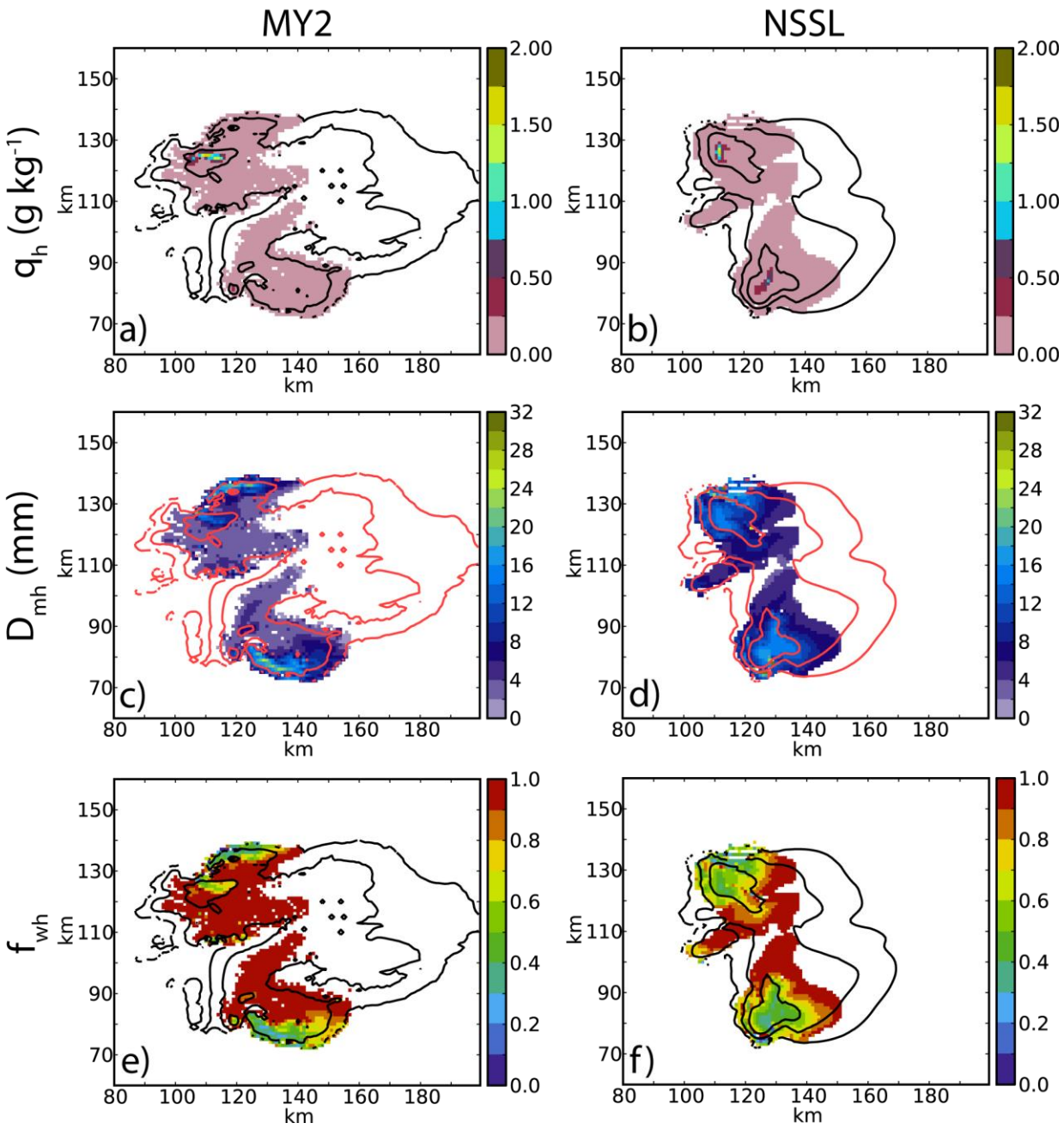

Fig. 3.4. As in Fig. 3.3 with hail.

While P3 with one free ice category is highly flexible with ice mode representation, it has the distinct disadvantage that it cannot fully represent different modes of ice particles at the same point in time and space. In the one-category version, the mixing of different populations of ice, for example by differential sedimentation or local initiation of new ice, creates a "dilution" of the

physical properties (Milbrandt and Morrison 2016). Although polarimetric variables between P3 with one and two ice categories show qualitative similarity near the surface, we focus on the two ice category configuration of P3 (hereafter referred to as P3-2) to allow for similar degrees of freedom to the other examined BMPs. We emphasize that the two free ice categories in P3-2 should not be interpreted as corresponding directly to the two rimed-ice categories in the MY2 and NSSL BMPs; both of the P3-2 ice categories can represent any type of frozen particle. Although comparison of the P3-2 ice categories with MY2 and NSSL BMP rimed ice categories is not a true "apples-to-apples" comparison, we choose to retain the entire P3-2 ice PSDs to prevent discontinuity error arising from separating the rimed ice section (where riming is present) from the ice PSD. Further, the MY2 and NSSL cloud ice and snow categories are not analyzed in this study as these processes contribute little to supercell polarimetric signatures (Kumjian and Ryzhkov 2008).

The simulated supercell's forward flank is largest in P3-2 (Fig. 3.1g). Moderately wet P3-2 (ice category 1 and 2 water fraction $f_{wi1,2}$ between 0.4 and 0.6) ice is much smaller (ice category 1 and 2 mass-weighted mean diameter $D_{mi1,2} \leq 8$ mm) than MY2 hail and NSSL rimed ice near the surface and is also sparse (ice category 1 and 2 mixing ratios $q_{i1,2} \leq 0.25$ g kg^{-1}; Fig. 3.5), subsequently resulting in the weakest reflectivity core among the BMPs examined. We note here that water fraction in the P3-2 is generally smaller than in the other two BMPs due to the lack of water fraction iteration (when computing water fraction for polarimetric variables as discussed in section 3.2.3) and because we assume not all sections of P3-2 ice PSDs are wet. Rain in the scheme easily dominates ice's contribution to reflectivity, which is larger in the forward flank (D_{mr} exceeds 3 mm) than the other two BMPs (Figs. 3.2e,f). Consequently, Z_H exceeds 40 dBZ in the forward flank over a greater area compared to the other BMPs. Near the surface, few differences arise

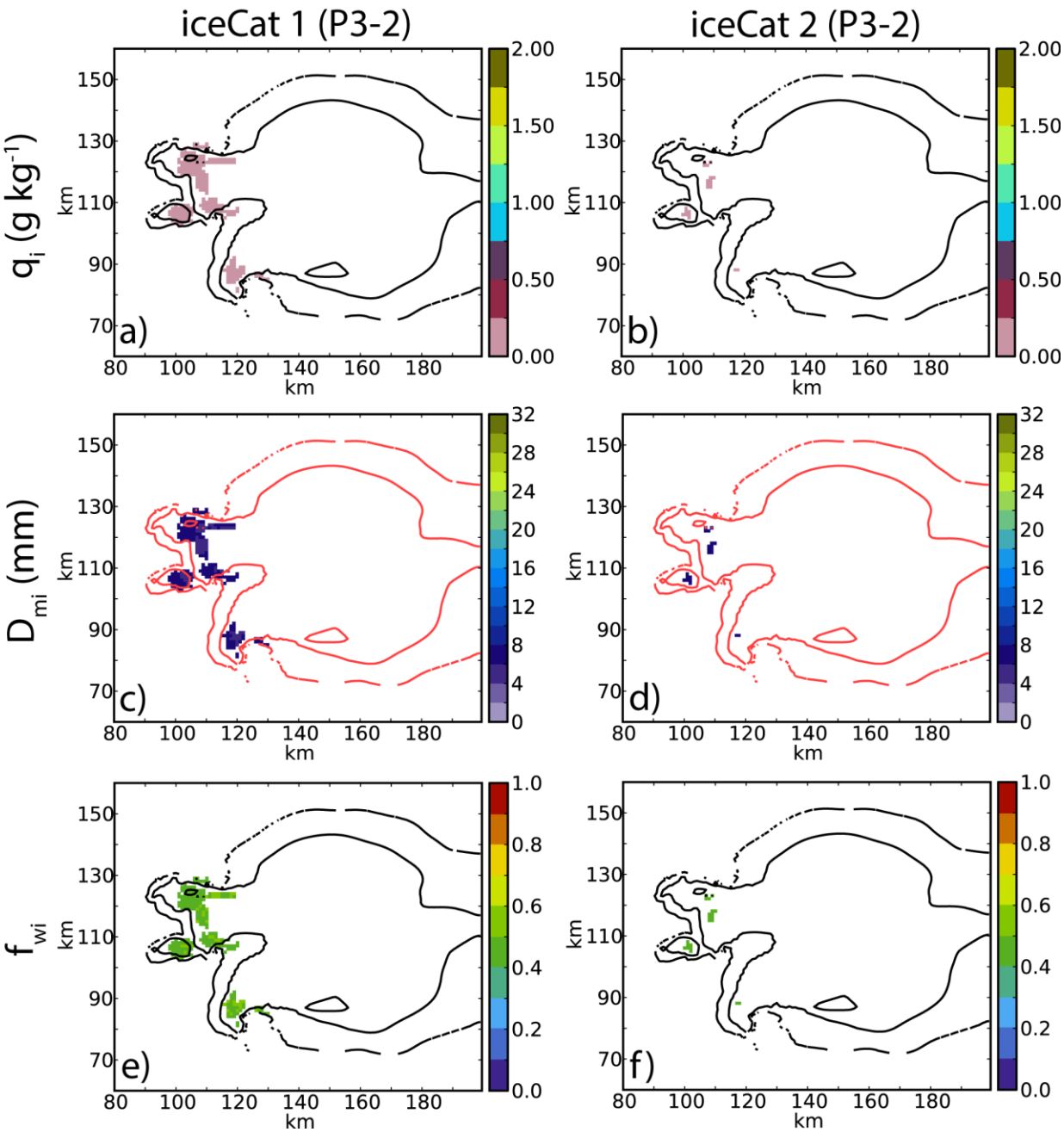

Fig. 3.5. Ice mixing ratio q_i (g kg^{-1}), mass-weighted mean diameter D_{mi} (mm), and ice water fraction f_{wi} near z = ~280m for ice category (a,c,e) 1 and (b,d,f) 2 in the P3-2 BMP at t = 100 min. Horizontal reflectivity Z_H contours are overlaid in 20 dBZ intervals starting at 15 dBZ.

between the two ice categories in P3-2. Mixing ratio, mean particle size, and water fraction are very similar between the two categories. The first category (hereafter iceCat 1) is more abundant

spatially than the second category (hereafter iceCat 2) because iceCat 2 is always merged with iceCat 1 when the differences in their mass-weighted mean diameters are within 0.5 mm.

3.3.1 Z_{DR} arc

Observationally, the size sorting of hydrometeors results in a local Z_{DR} maximum on the southern flank of the right-moving storm called the Z_{DR} arc. Ryzhkov et al. (2005a), Kumjian and Ryzhkov (2008), and Kumjian and Ryzhkov (2009) speculated that Z_{DR} decreases away from this maximum as raindrop size decreases. Through numerical simulations, Dawson et al. (2014) showed that simulated Z_{DR} arcs are potentially more influenced by the size sorting of rimed ice rather than rain, and Dawson et al. (2015) explained rain and rimed ice size sorting was the result of storm-relative winds. Large rimed ice exiting the updraft will fall to the surface faster than smaller rimed ice particles, and therefore, will be less prone to horizontal advection by storm-relative winds. Therefore, a gradient of large to small rimed ice particles will form in the direction of the deep-layer storm-relative mean wind vector (~0.7-12 km in Dawson et al. (2014) simulations). The low-level storm-relative mean wind vector (~0.7-3 km in Dawson et al. (2014) simulations) will further modulate the horizontal spatial distribution of hydrometeors in its direction, but since this layer is more shallow than the deep-layer, it has a smaller impact on the final surface distribution of hydrometeors. In our simulations, the deep layer (~0.3-12 km) storm-relative mean wind (typically southwesterly, but also northwesterly near the updraft; not shown) advects hydrometeors to the east of the updraft (both north and south), the horizontal extent of which is determined by particle size. Low-level storm-relative mean winds in our simulations (~0.3-3 km) are typically north- or southeasterly, which advects particles to the west and shortens the eastern extent of rain and rimed ice hydrometeors at the surface. Therefore, large drops from melted rimed ice and large (dry and partially-wet) rimed ice in our simulations with westerly

momentum quickly fall out and settle in a Z_{DR} maximum on the southern flank to the east of the updraft. Increasingly smaller particles will advect further to the northeast of the southern flank and create a hydrometeor size gradient that is largest near the southern flank and decreases in the direction of the deep-layer storm-relative mean wind vector (that is partially modulated to the west by the low-level storm-relative wind vector).

Similar to Johnson et al. (2016), both the MY2 and NSSL schemes are able to simulate a general Z_{DR} decrease from the southern edge of the supercell into the forward flank (i.e., in the general direction of the deep-layer storm-relative mean wind vector; Figs. 3.1b,e). However, neither scheme is able to simulate a Z_{DR} arc entirely consistent with observations (e.g., Ryzhkov et al. 2005a; Kumjian and Ryzhkov 2008). The MY2 scheme simulates an elongated region of high Z_{DR} near x = 140 km and y = 85 km exceeding 4 dB, and is primarily due to wet hail (Figs. 3.4a,c,e). Z_{DR} on the southern edge of the supercell is low ($\leq$ 1.5 dB) and is due to large (D_{mh} exceeding 14 mm), relatively dry (f_{wh} below 0.5) hail in the scheme (Figs. 3.4c,e). Still the MY2 scheme is unable to simulate relatively large drops in the forward flank on the southern edge of the supercell and shows a weak drop size gradient from the supercell edge to the northeast into the forward flank, consistent with Johnson et al. (2016). The NSSL BMP shows a clear Z_{DR} decrease from the southern flank of the supercell into the forward flank to the northeast. This Z_{DR} pattern is primarily due to rain, which also shows an enhancement of raindrop size on the southern edge of the forward flank that decreases to the northeast into the forward flank (Fig. 3.2d). However, the southern flank of the NSSL simulated supercell is much smaller than that in Johnson et al. (2016); therefore, there is not an elongated Z_{DR} enhancement in the expected Z_{DR} arc region, only a small area of large Z_{DR}. As mentioned previously, rain dominated the P3-2 BMP near the surface. Therefore, it is unsurprising that a Z_{DR} arc is not seen in Fig. 3.1h where the Z_{DR} field in the forward

flank is nearly as homogenous as the scheme's D_{mr} field (Fig. 3.2f), although raindrop size does decrease to the west on the western flank of the storm, consistent with the low-level winds and rain size sorting in Dawson et al. (2014). Z_{DR} in the P3-2's forward flank is much larger ($\geq$ 2.5 dB) than the other two schemes, a result of D_{mr} exceeding 3 mm in large areas.

As a major source of large raindrops is melting rimed ice, we consider horizontal ice and rimed-ice mass-weighted mean diameter plots near z = ~4.5 km (Fig. 3.6) to examine the impact of rimed ice melting on rain size sorting near the surface. z = ~4.5 is chosen because the melting layer is primarily near z = ~4 km, as shown in later vertical cross sections (Fig. 3.7). P3-2 ice plots are shown because the rime fraction f_r is typically above 0.8, which can increase the rimed-ice allocation of the ice PSDs. Graupel in MY2 displays a size maximum ($D_{mg} \geq 4$ mm) in the middle of the storm, away from the updraft (Fig. 3.6a). Larger hail ($D_{mh} \geq 10$ mm) in the scheme produces a more distinct size sorting distribution, with large particles on the southern edge of the right-moving cell and hail size decreasing into the forward flank to the northeast (Fig. 3.6b). In NSSL, both graupel and hail ($D_m \geq 10$ mm, which is larger than MY2 graupel) display larger particles near the updraft, where rimed ice initiates and grows (Figs. 3.6c,d). The large difference in graupel distribution between the MY2 and NSSL schemes is partly due to different assumptions for graupel. NSSL graupel spans a greater rimed ice property range (i.e., continuous riming growth to hail through prognostic bulk volume), while MY2 graupel is more constrained as a smaller, medium-density rimed ice category (and has a relatively flat graupel fall speed curve that is typically slower than NSSL graupel fall speed, which allows for greater MY2 graupel horizontal advection into the forward flank). NSSL hail is larger than NSSL graupel, and therefore this scheme displays much stronger size sorting of hail, with a stronger D_{mh} gradient into the forward flank than D_{mg}. Though MY2 graupel size sorting is weak (D_{mg} maxima in the center of the storm

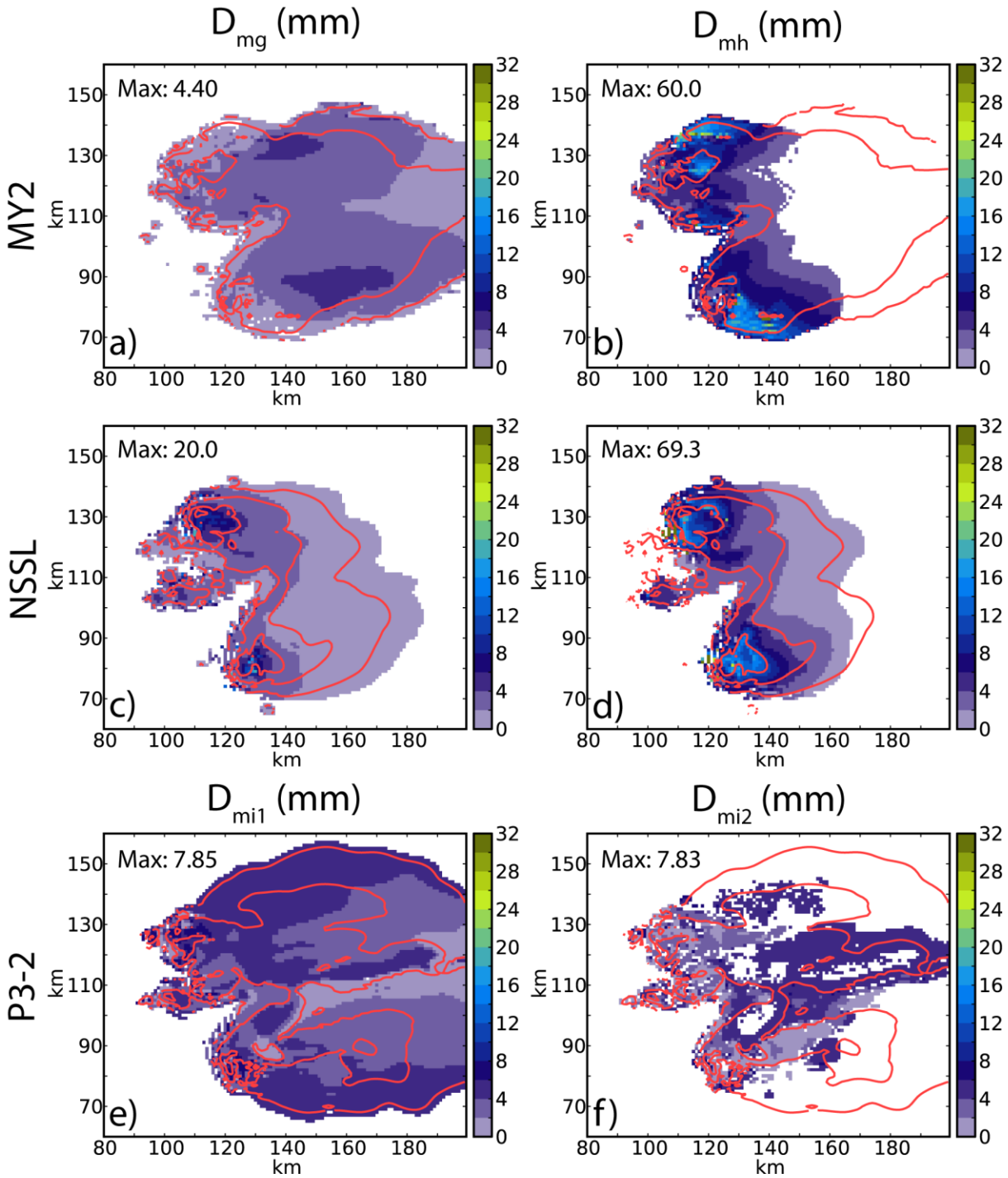

Fig. 3.6. Mass-weighted mean diameter of graupel D_{mg} (mm) and hail D_{mh} (mm) near z = ~4.5 km for the (a,b) MY2 and (c,d) NSSL BMPs and ice category 1 D_{mi1} (mm) and 2 D_{mi2} (mm) near z = ~4.5 km for the (e,f) P3-2 scheme at t = 100 min. Included in each panel is the maximum D_m. Horizontal reflectivity Z_H contours are overlaid in 20 dBZ intervals starting at 15 dBZ.

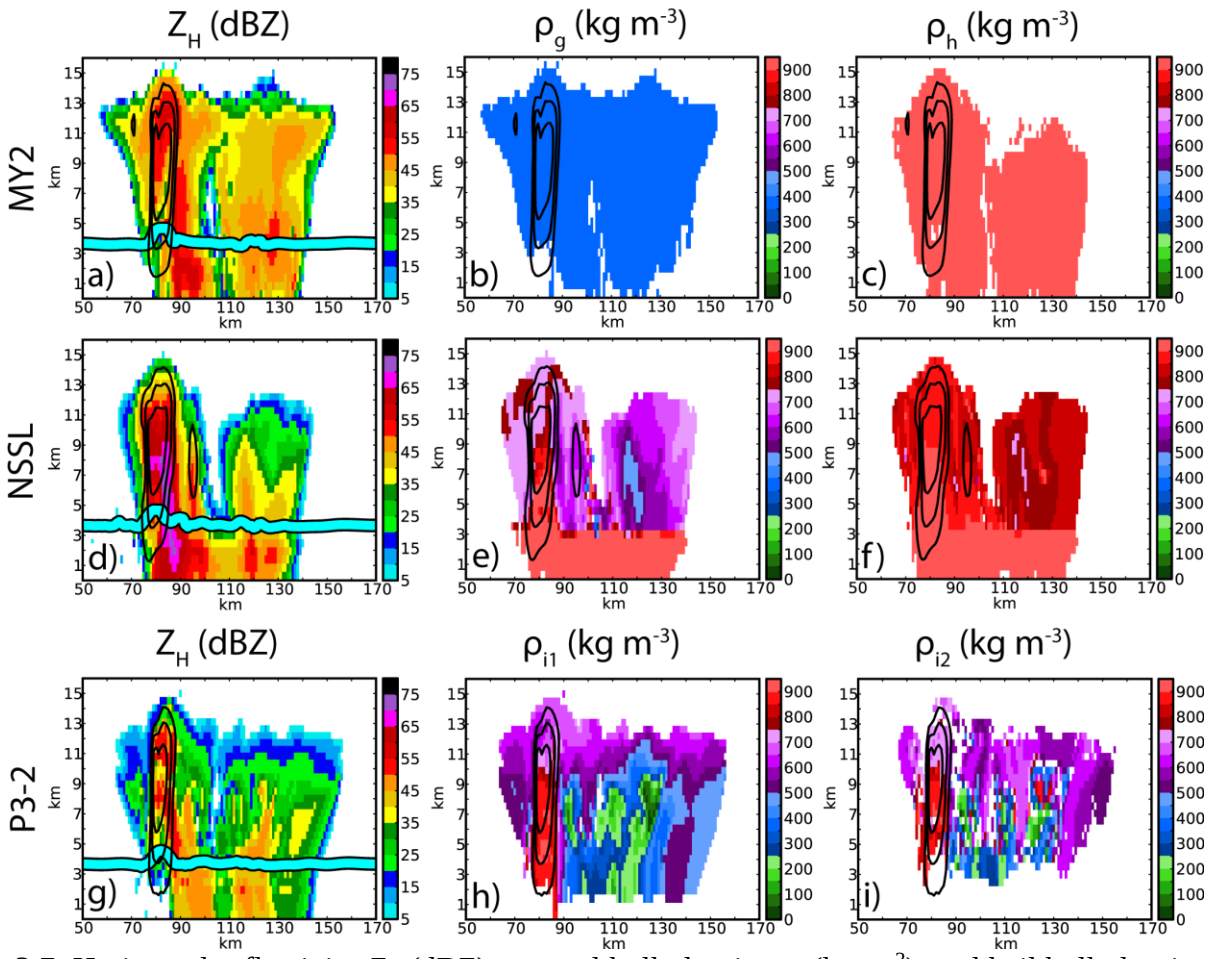

Fig. 3.7. Horizontal reflectivity Z_H (dBZ), graupel bulk density ρ_g (kg m^{-3}), and hail bulk density ρ_h (kg m^{-3}) for the (a,b,c) MY2 and (d,e,f) NSSL schemes, and Z_H (dBZ) and ice bulk density for ice categories 1 ρ_{i1} (kg m^{-3}) and 2 ρ_{i2} (kg m^{-3}) in the (g,h,i) P3-2 scheme through the updraft at t = 100 min. The melting level is depicted as a 0°C isotherm blue line, and vertical velocity contours are shown with 15 m s^{-1} interval starting at 10 m s^{-1}.

that decreases away), both the MY2 and NSSL schemes display rimed ice size sorting as D_m decreases to the northeast of the storm, an important precursor for observed supercell low-level Z_{DR} arcs and their associated gradient in the direction of the storm-relative winds (e.g., Dawson et al. 2014).

Mass-weighted mean diameter of iceCat 1 D_{mi1} exhibits evidence of size sorting as particle size increases toward the edges of the storm (Fig. 3.6e). Maximum ice size is smaller in P3-2

(except for MY2 graupel), resulting in a weaker gradient. Further, D_{mi2} in P3-2 depicts a rather noisy pattern with size generally increasing toward the middle of the splitting storm (Fig. 3.6f). The inverse size sorting gradient in iceCat 2 should not affect the rain distribution near the surface, as the mixing ratio of iceCat 2 is typically smaller than that of iceCat 1 (not shown). Although the predefined rimed-ice categories in the MY2 and NSSL schemes are fundamentally different from those in P3-2, comparison to these predefined schemes can help ensure that BMPs with more ice mode flexibility indeed improve the representation of the wide range of ice modes. The striking differences in mid-level ice distributions between the P3-2 and other BMPs suggest that some processes related to ice size sorting may need improvement in P3.

3.3.2 Hail signature in forward flank downdraft (FFD)

Hailstones that grow sufficiently large in a strong supercell updraft fall out quickly and relatively dry, tumbling as they fall. These large hailstones appear spherical to radar, substantially reducing Z_{DR} (Kumjian and Ryzhkov 2008). As rimed ice canting angle standard deviation increases linearly with decreasing water fraction in the polarimetric simulator employed in this study, simulated rimed ice does not need to be completely dry to reduce Z_{DR}. Another consequence of this parameterization is that the tumbling rimed ice that reduces Z_{DR} may also reduce cross correlation coefficient ρ_{HV} due to the resonance effect in the Mie scattering regime (Kumjian and Ryzhkov 2008, 2009), which occurs as $D|\varepsilon|^{1/2}/\lambda$ approaches 1 (here D is equivolume diameter, ε is hydrometeor dielectric constant, and λ is radar wavelength). Therefore, an area of reduced ρ_{HV} overlapping reduced Z_{DR} can be reasonably attributed to rimed ice in the supercell framework of this study.

Near the surface in the MY2 scheme, large, relatively dry hail (D_{mh} exceeding 14 mm, f_{wh} below 0.5; Figs. 3.4c,e) reduces Z_{DR} substantially and ρ_{HV} (as the hail particles are mixed-phase)

below 0.94 near x = 130 km and y = 75 km (Figs. 3.1b,c). However, the location of the MY2 Z_{DR} reduction is extended too far south into the southern flank compared to typical hail signature observations (Kumjian and Ryzhkov 2008). The MY2 Z_{DR} reduction span and location is likely due to a smaller mass-weighted mean hail fall speed over the deep layer (~0.3-12 km) compared to the NSSL scheme (not shown), which would allow MY2 hail to advect further to the right of the updraft following the deep-layer storm relative wind, rather than quickly falling out close to the updraft. In the NSSL scheme, large, relatively dry hail ($D_{mh} \geq 12$ mm, f_{wh} below 0.5) reaches the surface (Figs. 3.4d,f) and reduces Z_{DR} north of the storm's hook where the observed hail signature is often found (Fig. 3.1e), but not as much as the hail signature produced in Johnson et al. (2016). These hail signature differences underscore how different BMP versions can manifest themselves in simulated polarimetric signatures. Hail in the NSSL scheme also reduces ρ_{HV} in the same location as the Z_{DR} reduction below 0.96 (Fig. 3.1f), which is a smaller ρ_{HV} reduction than in the MY2 scheme. While hail in the maximum ρ_{HV} reduction region is drier in the MY2 scheme (and therefore decreases ρ_{HV} reduction potential), hail is also larger in the MY2 scheme and consequently contains more particles prone to the resonance effect than NSSL hail.

In P3-2, ice barely reaches the surface and there is no simulated Z_{DR} reduction from large dry rimed ice (Fig. 3.1h). Ice is only present in small quantities near the supercell's hook appendage ($q_{i1,2} \leq 0.25$ g kg^{-1}; Fig. 3.5), and iceCat1's relatively small ($D_{mi1} \leq 8$ mm) and moderately wet (f_{wi1} between 0.4 and 0.6) ice particles actually increase Z_{DR} where present. This is primarily due to the scheme's restrictive maximum ice number-weighted mean diameter D_{ni}, which limits the growth of ice. We emphasize here that maximum number-weighted mean diameter in this paper is only examined as a tuning parameter for large ice production, which will be discussed later; it is never analyzed as a simulated microphysics PSD characteristic as in mass-

weighted mean diameter D_m. Based on its water fraction f_{wi}, ice in the P3-2 scheme is mixed-phase and therefore has the potential to reduce ρ_{HV} near the surface. However, ρ_{HV} near the surface is generally large (≥ 0.98; Fig. 3.1i), as wet ice particles in the P3-2 scheme are much smaller than hail in the MY2 and NSSL schemes, and therefore are less prone to the resonance effect. In fact, much of the ρ_{HV} reduction in the P3-2 scheme is due to relatively large, oblate drops ($D_{mr} \geq 2.5$ mm; Fig. 3.2f), which slightly reduces the correlation between horizontal and vertical reflectivity.

3.3.3 Updraft hail core

We analyze vertical cross sections through the updraft (marked by black lines in Fig. 3.1) to examine vertical distributions of rimed ice (Fig. 3.7). Black lines denote vertical velocity contours, while blue lines in Z_H plots denote the melting level represented by the 0°C isotherm. Upper level Z_H in MY2 peaks high in the updraft, near z = 12 km (Fig. 3.7a). While dry graupel is prominent in this area (Figs. 3.8a,e), it is also relatively small ($D_{mg} \leq 4$ mm; Fig. 3.8c). Therefore, a small amount (q_h between 0 and 4 g kg^{-1}) of relatively large, dry hail ($D_{mh} \geq 8$ mm, $f_{wh} \leq 0.1$; Figs. 3.9a,c,e) is responsible for the enhanced Z_H. Z_H in the NSSL scheme peaks lower in the updraft, below z = 8.5 km (Fig. 3.7d) on the northern edge of the updraft. Low amounts (q_g typically below 4 g kg^{-1}) of graupel with a sharp water fraction gradient (Figs. 3.8b,d,f) in this area are smaller than moderate amounts (max q_h of 10 g kg^{-1}) of both dry and wet hail in the area (f_{wh} between 0-0.6; Figs. 3.9b,d,f). The largest hail produced in the NSSL scheme ($D_{mh} \geq 30$ mm) does not correlate with the Z_H maximum, which is mainly due to lower q_h (typically less than 2 g kg^{-1}) and water fraction (drier; typically less than 0.2). Included in the reflectivity figures are rimed-ice densities, which are constant in MY2 (Figs. 3.7b,c) and predicted in NSSL (Figs. 3.7e,f). In the NSSL scheme, graupel and hail densities are both large (generally above 700 kg m^{-3}) and similar in the updraft and below the melting level. Otherwise, graupel exhibits a larger density range than

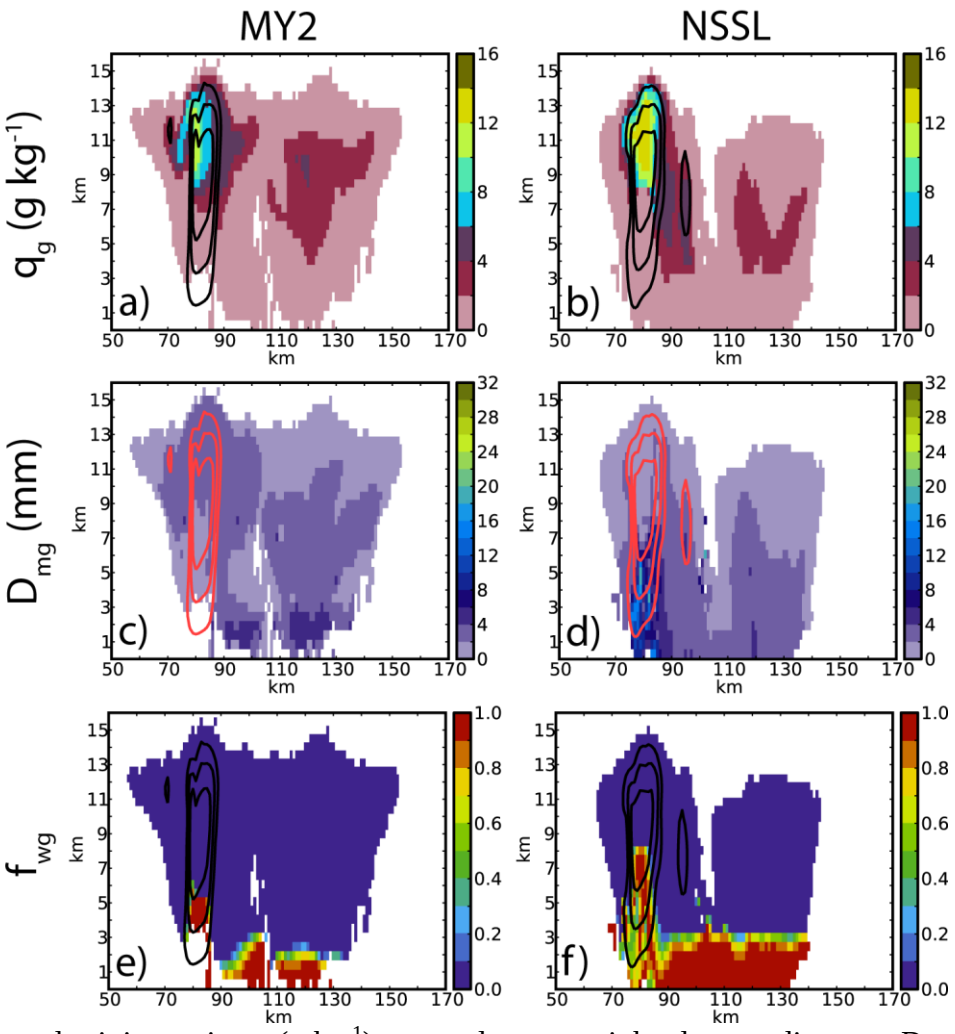

Fig. 3.8. Graupel mixing ratio q_g (g kg^{-1}), graupel mass-weighted mean diameter D_{mg} (mm), and graupel water fraction f_{wg} through the updraft for the (a,c,e) MY2 and (b,d,f) NSSL BMPs at t = 100 min. Vertical velocity contours are shown with 15 m s^{-1} interval starting at 10 m s^{-1}.

hail over the supercell, which makes sense as graupel grows and feeds to hail (large, high-density particles) in the scheme. Compared to the prescribed constant density values in the MY2 scheme of 400 and 900 kg m^{-3} for graupel and hail respectively, NSSL graupel density is typically larger while hail density is slightly smaller, reflective of the NSSL BMP's predicted rimed-ice bulk volume simulating wet growth in a smooth, continuous manner.

The P3-2 scheme produces two distinct Z_H maxima within the updraft, one near z = 12 km and one below z = 9.5 km (Fig. 3.7g). While the upper level maximum is primarily caused by

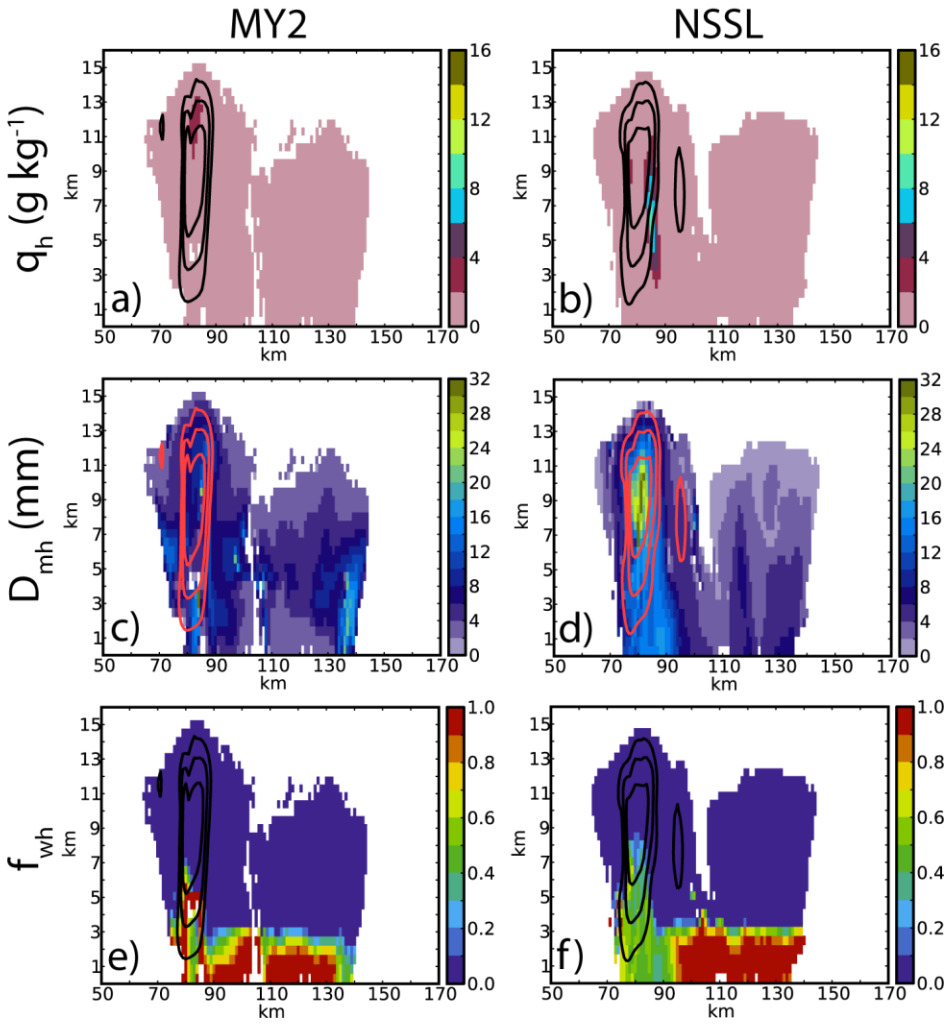

Fig. 3.9. As in Fig. 3.8 with hail.

moderate to large amounts (q_{i1} generally between 4 and 12 g kg^{-1}) of relatively large, dry ($f_{wi1} \leq$ 0.1) iceCat 1 (Figs. 3.10a,c,e), dry iceCat 2 is prominent where iceCat 1 reduces to a local minimum (Figs. 3.10b,d,f). The iceCat 1 and 2 juxtaposition near z = 12 km helps illustrate P3 ice category interaction, as both categories initiate, grow independent of each other, and are kept separate as long as they are not similar in size. The combination of the two separate categories helps produce a smooth, continuous reflectivity field, while the individual ice categories themselves can become noisy. The Z_H maximum below z = 9.5 km is predominantly caused by both moderately dry and wet iceCat 1 (f_{wi1} between 0 and 0.5), though iceCat 2 also contributes to

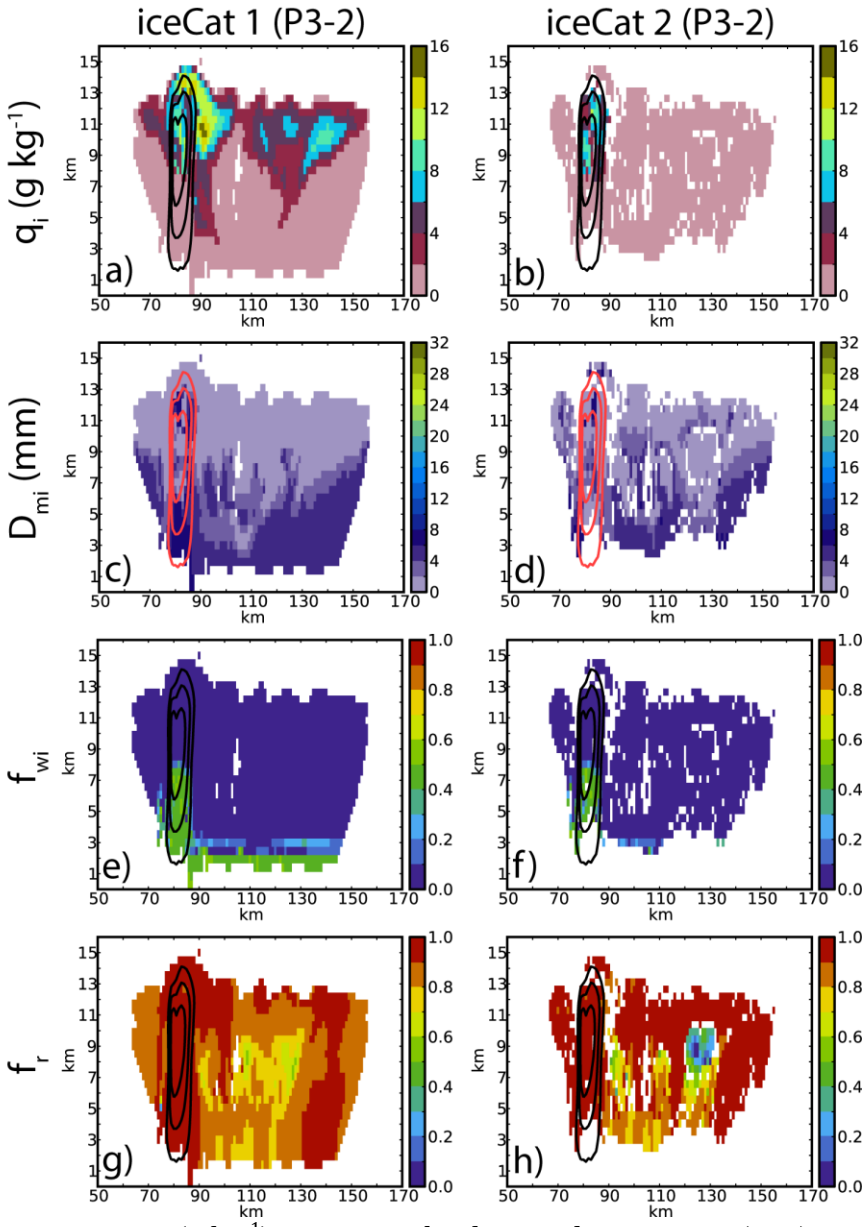

Fig. 3.10. Ice mixing ratio q_i (g kg^{-1}), mass-weighted mean diameter D_{mi} (mm), ice water fraction f_{wi}, and rime fraction f_r through the updraft for ice category (a,c,e,g) 1 and (b,d,f,h) 2 in the P3-2 BMP at t = 100 min. Vertical velocity contours are shown with 15 m s^{-1} interval starting at 10 m s^{-1}.

this maximum. The mass-weighted mean diameters of both iceCat1 and iceCat2 are typically smaller ($D_{mi1,2} \leq 8$ mm) than MY2 and NSSL hail in the updraft, but the P3-2 has a reflectivity core similar in magnitude to the MY2. While P3-2 typically produces more mass than MY2 hail in the upper Z_H maximum, it is also important to remember that P3-2 ice mass-weighted mean

diameter spans the entire PSD, including smaller non-rimed ice. In terms of bulk density (which includes small spherical and large non-spherical ice), densities are generally very similar between the two ice categories (Figs. 3.7h,i). Bulk density is largest (generally above 500 kg m^{-3}) in the updraft where riming processes dominate and decrease away. While P3 ice categories are not exclusively rimed ice, rime fraction f_r is typically greater than 0.8 in the vertical cross section (Figs. 3.10g,h) indicating the dominance of riming growth over vapor deposition. For medium-large densities (i.e., $\rho_i \geq 500$ kg m^{-3}) in the updraft and near the top of the supercell, the majority of the ice PSD is likely rimed and partially-rimed ice, similar to the graupel and hail predefined categories in the MY2 and NSSL BMPs. Small spherical and large non-spherical ice become more prominent away from the updraft as density and rime fraction decrease.

3.4 Rimed-ice properties

3.4.1 Microphysical tendencies

To understand better the production and depletion of rimed-ice categories, we analyze average mixing ratio q_x and number concentration N_{tx} tendencies in the MY2 and NSSL schemes, and both free ice categories in the P3-2 scheme at t = 100 min (Figs. 3.11, 3.12). We sum the tendencies over 5 min at each grid point, and then over the horizontal domain at each model height level. The 5-minute process sums are normalized by the number of seconds over the temporal integration period to produce average tendencies. Tendencies are only included in Figs. 3.11 and 3.12 if the maximum average tendencies (which are horizontally and temporally summed before normalization) that vary with height exceed q and N_t tendency thresholds of 0.25 g kg^{-1} s^{-1} and 1 # m^{-3} s^{-1}, respectively, to limit discussion to tendencies that contribute most to these bulk quantities. The naming convention for the microphysical processes is based on Milbrandt and Yau (2005b). The first letter denotes the relevant bulk quantity related to the tendency (Q for mixing ratio, N for

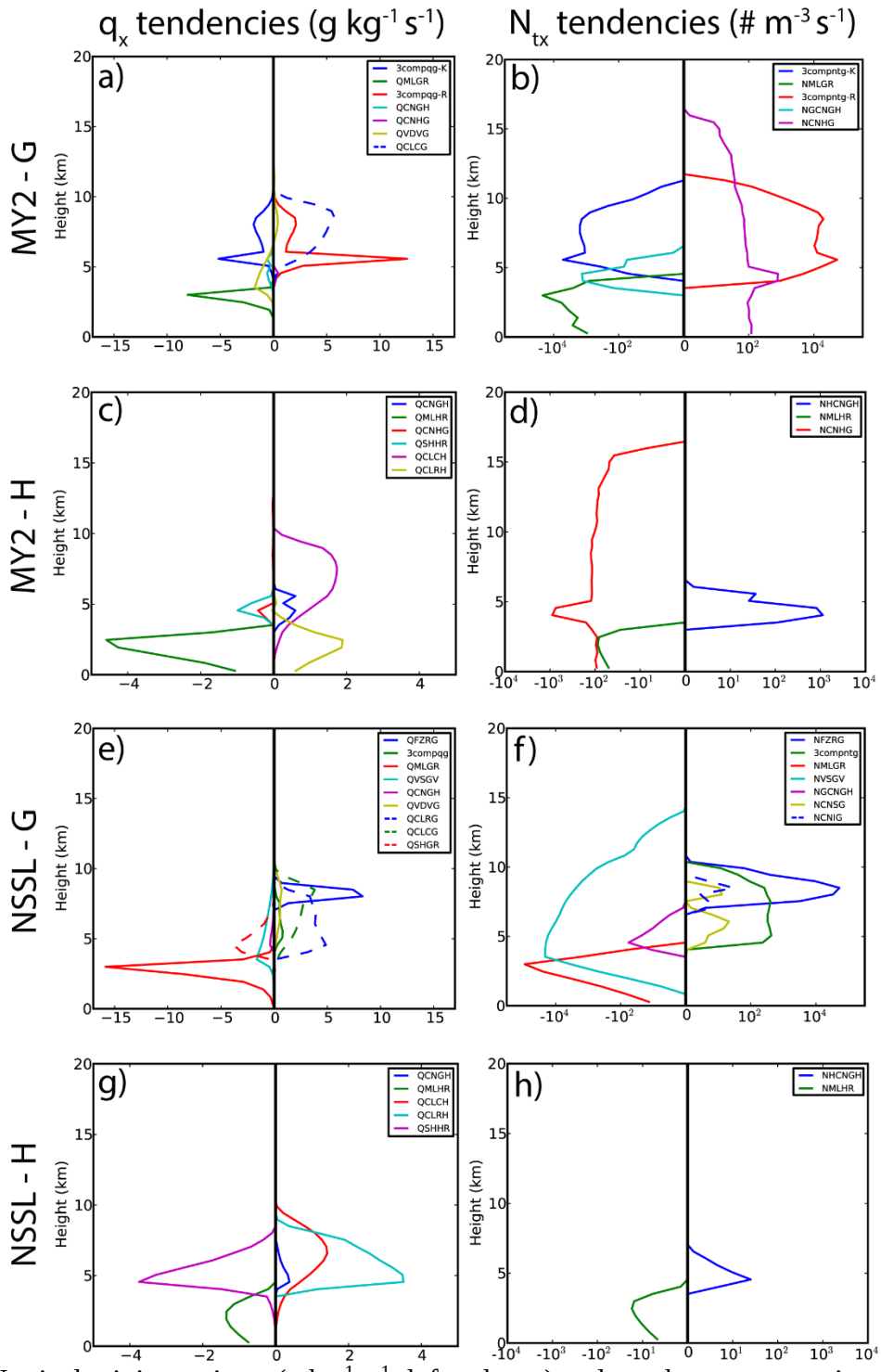

Fig. 3.11. Vertical mixing ratio q_x (g kg^{-1} s^{-1}; left column) and number concentration N_{tx} (# m^{-3} s^{-1}; right column) average microphysical tendencies for MY2 (a,b) graupel and (c,d) hail, and NSSL (e,f) graupel and (g,h) hail. The vertical black line in each plot denotes the zero line separating source and sink terms.

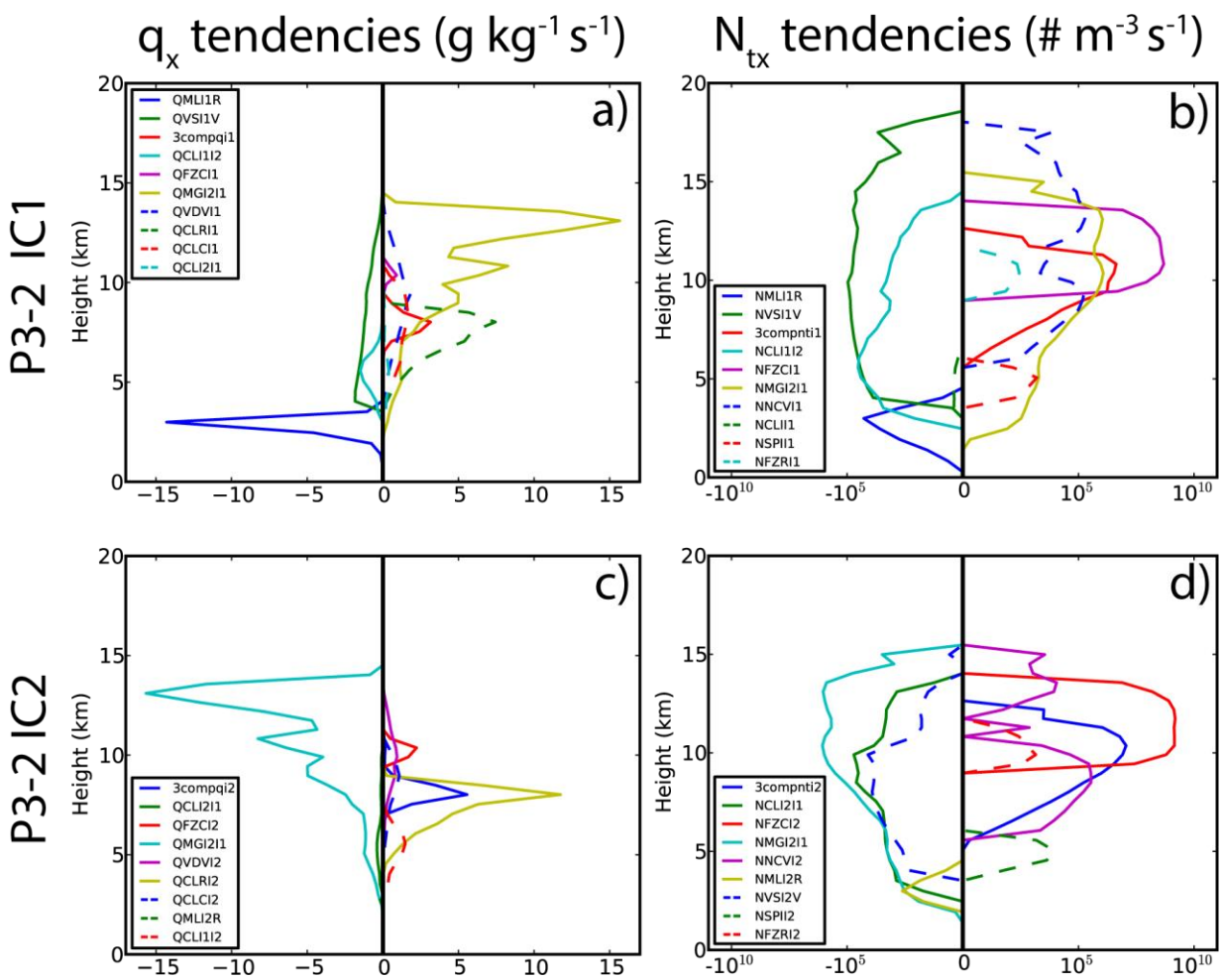

Fig. 3.12. As in Fig. 3.11, for P3-2 (a,b) iceCat 1 and (c,d) iceCat 2.

number concentration). The next two letters denote the process itself (CL: collection, MG: merging, ML: melting, CN: conversion, VD: vapor deposition, SH: shedding, FZ: freezing, VS: vapor sublimation, NC: nucleation, SP: splintering). The last two letters indicate the hydrometeor sink and source respectively (V: vapor, R: rain, C: cloud water, I: cloud ice, S: snow, G: graupel, H: hail, I1: P3-2 ice category 1, I2: P3-2 ice category 2). The exception to this notation is 3-component freezing, which sums all tendencies that contribute to this process. A description of each microphysical process examined is available in Table 3.2.

Three-component freezing primarily initiates MY2 graupel right above the freezing level, while graupel collecting cloud water provides the largest source of q_g between z = 6-10 km (Figs. 3.11a,b). 3-component freezing of water and graupel in the MY2 scheme can either go to graupel or hail, which is why the scheme has both source (3compqg-R) and sink (3compqg-K; although we emphasize here that 3-component freezing graupel "sink" can freeze with water to graupel "source") terms. On the other hand, NSSL graupel primarily initiates from freezing rain (Figs. 3.11e,f). This reveals a subtle graupel initiation difference between the MY2 and NSSL BMPs for these idealized supercell simulations, as MY2 graupel prefers an ice seed when freezing water to graupel while NSSL graupel does not. Similar to the MY2 scheme, NSSL graupel grows by collecting liquid mass (cloud water and rain), which is expected given the convective updraft in the supercell. Hail converting to graupel (which is is an artificial process, present only for optimization) provides a larger source of graupel number aloft in the MY2 scheme and below the melting level but not mass, which is expected as only small hail converts to graupel (hail mean mass diameter D_h < 1 mm). This process is not present in the default NSSL scheme, as NSSL hail is designed to primarily originate from larger and heavily-rimed graupel. Melting provides a large sink of graupel below the melting level in both the MY2 and NSSL schemes, although it is more aggressive in the NSSL scheme. 3-component freezing provides a large graupel number sink above the melting level in the MY2 scheme, while graupel number is typically depleted by sublimation of small graupel particles in the NSSL scheme.

Graupel conversion to hail is the primary origin (i.e., creation rather than particle growth) of hail in the MY2 and NSSL schemes. Graupel mass conversion to hail in the MY2 scheme is larger than the mass conversion in the NSSL scheme (Figs. 3.11c,g), but is also nearly two orders of magnitude larger in terms of the peak number (Figs. 3.11d,h). Collection of rain and cloud water

Table 3.2. Ice microphysical processes in the MY2, NSSL, and P3-2 schemes. Graupel or hail processes are from the MY2 and/or NSSL schemes, while iceCat 1 or iceCat 2 processes are from the P3-2 scheme. Processes that have been filtered are not included.

Process name	Description
3compqg	3-component mass freezing to graupel
QCLCG	Graupel mass collection of cloud water
QCLCH	Hail mass collection of cloud water
QCLRG	Graupel mass collection of rain
QCLRH	Hail mass collection of rain
QCNGH	Graupel mass conversion to hail
QCNHG	Hail mass conversion to graupel
QFZRG	Freezing rain mass to graupel
QMLGR	Graupel mass melting to rain
QMLHR	Hail mass melting to rain
QSHGR	Graupel mass shedding rain
QSHHR	Hail mass shedding rain
QVDVG	Graupel mass depositional growth
QVSGV	Graupel mass sublimation
3compqi1	3-component mass freezing to iceCat 1
3compqi2	3-component mass freezing to iceCat 2
QCLCI1	iceCat 1 mass collection of cloud water
QCLCI2	iceCat 2 mass collection of cloud water
QCLI1I2	iceCat 2 mass collection of iceCat 1
QCLI2I1	iceCat 1 mass collection of iceCat 2
QCLRI1	iceCat 1 mass collection of rain
QCLRI2	iceCat 2 mass collection of rain
QFZCI1	Freezing cloud water mass to iceCat 1
QFZCI2	Freezing cloud water mass to iceCat 2
QMGI2I1	iceCat 2 mass merging to iceCat 1
QMLI1R	iceCat 1 mass melting to rain
QMLI2R	iceCat 2 mass melting to rain
QVDVI1	iceCat 1 mass depositional growth
QVDVI2	iceCat 2 mass depositional growth
QVSI1V	iceCat 1 mass sublimation
3compntg	3-component number freezing to graupel
NGCNGH	Graupel sink number conversion to hail

NHCNGH	Graupel number conversion to hail source
NCNHG	Hail number conversion to graupel
NCNIG	Cloud ice number conversion to graupel
NCNSG	Snow number conversion to graupel
NFZRG	Freezing rain number to graupel
NMLGR	Graupel number melting to rain
NMLHR	Hail number melting to rain
NVSGV	Graupel number sublimation
3compnti1	3-component number freezing to iceCat 1
3compnti2	3-component number freezing to iceCat 2
NCLI1I2	iceCat 2 number collection of iceCat 1
NCLI2I1	iceCat 1 number collection of iceCat 2
NCLII1	iceCat 1 number self-collection
NFZCI1	Freezing cloud water number to iceCat 1
NFZCI2	Freezing cloud water number to iceCat 2
NFZRI1	Freezing rain number to iceCat 1
NFZRI2	Freezing rain number to iceCat 2
NMGI2I1	iceCat 2 number merging to iceCat1
NMLI1R	iceCat 1 number melting to rain
NMLI2R	iceCat 2 number melting to rain
NNCVI1	iceCat 1 number nucleation
NNCVI2	iceCat 2 number nucleation
NSPII1	iceCat 1 number splintering
NSPII2	iceCat 2 number splintering
NVSI1V	iceCat 1 number sublimation
NVSI2V	iceCat 2 number sublimation

are the largest sources of hail mass growth in MY2 and NSSL schemes, which again makes sense given the expected wet growth in a supercell updraft. However, the large amount of hail collecting rain below the melting level in the MY2 scheme is rather excessive, and has been documented in Labriola et al. (2017). It is clear that the NSSL scheme is able to simulate wet growth processes in the updraft in a more physical manner than the MY2 scheme. Melting is the main mass sink below

the melting level in the MY2 and NSSL schemes but is much larger in the MY2 scheme. Since large hail reaches the surface in both schemes, MY2 hail collecting rain below the melting level seems to be compensating for the stronger melting. NSSL hail sheds much more water than MY2 hail over a greater depth, which is reasonable considering shedding can occur during wet growth in the updraft and wet growth in the MY2 scheme primarily occurs below the melting level.

The merging of iceCat2 to iceCat 1 is the primary source of iceCat 1 mass in the P3-2 scheme (Fig. 3.12a). We emphasize here that the merging of similar free ice categories in P3 is purely a computational process to free up an additional ice category, and is not a "conversion" process as in traditional BMPs (e.g., conversion of cloud ice to snow). It is reasonable for this merging to increase at higher altitudes, as newly initiated ice or small ice advected through the updraft does not have time to grow uniquely. iceCat 1 collecting rain is the second largest source of q_{i1}, echoing the dominant rimed ice growth mechanisms in the MY2 and NSSL schemes. Because free ice categories in P3-2 are not predefined rimed ice categories, processes such as nucleation and freezing cloud water add iceCat 1 number (Fig. 3.12b). Such an ice PSD dilution could be problematic in a deep convective framework, as this small nucleated ice could shift the process rates of existing rimed ice. iceCat 1 number contains several sources that vary with height, including nucleation and freezing cloud water at upper levels, and splintering at lower levels. In terms of mass depletion, melting is the dominant process and very sudden below the melting level. Consequently, while melting hail remains at the lowest model level in both the MY2 and NSSL schemes, strong melting leaves little iceCat 1 at this level. Again, this rapid ice depletion is likely a consequence of the P3-2's restrictive D_{ni} limit prohibiting large particle growth; this is discussed in more detail below. Similar to N_{ti1} sources, N_{ti1} sinks vary considerably with height: sublimation dominates upper levels, collection is prominent above the melting level, while melting dominates

below. IceCat 2 mass in P3-2 is largely affected by two processes: collecting rain for growth, and merging to iceCat 1 as a sink (Fig. 3.12c). Despite little constraint on the ice type with its free category approach, P3-2 seems to capture the important rimed ice process of collecting liquid well. iceCat 2 is similar to iceCat 1 in that the number sources and sinks vary with height, though merging to iceCat 1 largely dominates depletion of iceCat 2 (Fig. 3.12d). Nucleation and freezing cloud water provide sources in the upper levels, while splintering increases number closer to the melting level.

3.4.2 Size sorting

Rain and ice mass-weighted mean diameter D_m contoured frequency by altitude diagrams (CFADs; Yuter and Houze 1995) are created to illustrate the size sorting of rain and ice hydrometeors (Fig. 3.13). We normalize these frequencies by the number of grid points with a non-zero D_m. D_{mr} bins are spaced by 0.25 mm, while the analysis for ice uses a 0.5 mm interval. The approximate melting level of 4 km is denoted by a horizontal black line. Rain size sorting is clear in the MY2 and NSSL schemes, as D_{mr} frequency increases with decreasing height (Figs. 3.13a,d). MY2 D_{mr} frequencies favor larger rain diameters at low altitudes due to excessive size sorting (e.g., Wacker and Seifert 2001; Milbrandt and Yau 2005a; Milbrandt and McTaggart-Cowan 2010), while the NSSL scheme modifies N_t to prevent spurious large reflectivity (Mansell 2010). Rain frequencies above the melting level in the schemes are typically near the updraft, revealing that both the MY2 and NSSL schemes carry supercooled water above the melting level. Rimed-ice size sorting potentially drives the Z_{DR} arc signature (Dawson et al. 2014), which is apparent in MY2 and NSSL graupel and hail frequency plots (Figs. 3.13b,c,e,f). Both rimed-ice categories in both schemes show a clear D_m increase with decreasing height. MY2 rimed ice size sorts at a greater rate, a consequence of the previously mentioned NSSL N_t modification limiting

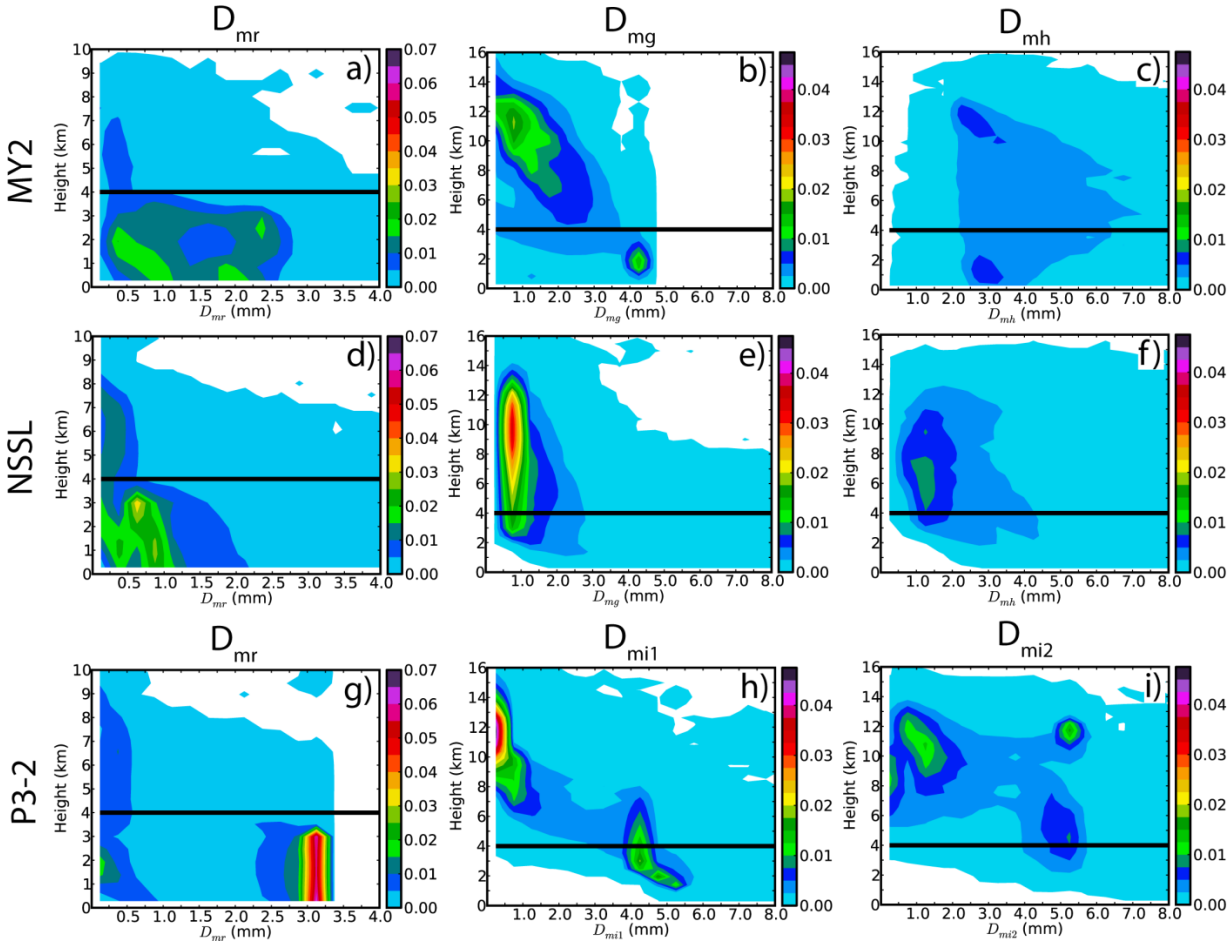

Fig. 3.13. Contoured frequency by altitude diagrams (CFADs) with rain, graupel, and hail mass-weighted mean diameter D_m (mm) for the (a,b,c) MY2 and (d,e,f) NSSL BMPs, and rain, ice category 1, and ice category 2 for the (g,h,i) P3-2 scheme at t = 100 min. The approximate melting level is denoted by a horizontal black line.

rimed ice size in the scheme. One major difference between this MY2 hail CFAD and the one in Johnson et al. (2016) is the absence of small hail particles. For this study, the default MY2 hail-to-graupel conversion threshold is 1 mm, while the previous threshold in Johnson et al. (2016) was reduced from the default 5 mm to 0 mm. Allowing small hail to be converted to graupel in this study likely led to less small hail particles produced, and allows the scheme to produce more realistic, larger hail.

P3-2 ice overwhelmingly melts to a narrow rain diameter range, typically between 3.0-3.25 mm below the melting level (Fig. 3.13g). The relative maxima at smaller diameters are the result of small raindrops in and near the updrafts due to warm rain processes (e.g., autoconversion). The D_{mr} dominance in one diameter bin is consistent with the P3-2 scheme's inability to simulate a Z_{DR} arc due to lack of rain size sorting, which is evident from near constant D_{mr} with decreasing height. On the other hand, D_{mi1} and D_{mi2} both show clear size increase with decreasing height (Figs. 3.13h,i), suggesting that melting produces uniform raindrop size regardless of ice size. Both P3-2 ice categories show a larger mean size range than graupel and hail in the other two schemes, a reflection of the numerous ice modes contained in their PSDs and lack of large ice in the scheme that results in a more homogenous distribution of ice particle size (and therefore enhanced frequency) in the scheme. While iceCat 1 contoured frequency is more continuous, iceCat 2 displays more discrete maxima. This is due to the discrete nature of iceCat 2, which can merge with iceCat 1. At z = 12 km, iceCat 2 contains a second frequency maximum in the D_{mi2} 5.0-5.5 mm bin, as the particles are prominent in the supercell's forward flank. Their occurrence decreases directly below this height, before increasing frequency again near z = 8 km.

P3-2 constricts the rain PSD slope parameter Λ_r to prevent unrealistically small (large Λ_r) or large (small Λ_r) mean drop sizes by explicitly setting Λ_r bounds. Because small drops on the western flank in the P3-2 supercell show evidence of size sorting while larger drops in the forward flank do not (Fig. 3.2f), the default minimum rain PSD slope bound $\Lambda_{r,min}$ is likely disrupting P3-2 surface drop size gradient by preventing the presence of large raindrops on the southern flank of the storm. The sensitivity of rain size gradient to $\Lambda_{r,min}$ is explored by reducing this parameter by adding a factor of 0.2 (referred to as P3-2_0.2$\Lambda_{r,min}$; Fig. 3.14). Reducing $\Lambda_{r,min}$ increases the potential for smaller rain PSD slope near the surface, which shifts the rain distribution to larger

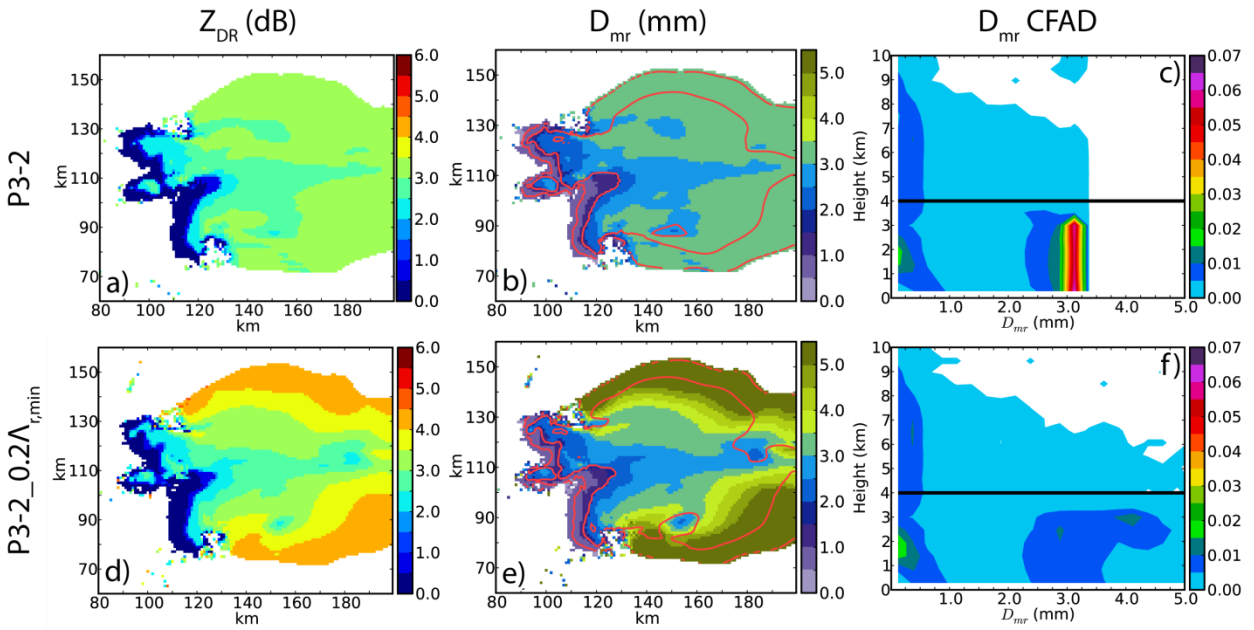

Fig. 3.14. Differential reflectivity Z_{DR} (dB) and rain mass-weighted mean diameter D_{mr} (mm) near z = ~280 m, and contoured frequency by altitude diagrams (CFADs) with rain mass-weighted mean diameter D_{mr} for the P3-2 (a,b,c) default scheme and (d,e,f) with minimum rain lambda $\Lambda_{r,min}$ reduced by 0.2 at t = 100 min. Horizontal reflectivity Z_H contours are overlaid in 20 dBZ intervals starting at 15 dBZ on D_{mr} plots, and the approximate melting level in rain CFAD plots is denoted by a horizontal black line.

drops. Near the surface, D_{mr} in P3-2_0.2$\Lambda_{r,min}$ is larger than 5 mm on the supercell's southern flank, and displays a sharper drop size gradient into the forward flank than in the default P3-2 scheme. The resulting Z_{DR} field exceeds 4 dB near the southern flank, and generally decreases following rain size. While the Z_{DR} gradient in P3-2_0.2$\Lambda_{r,min}$ is much improved compared to the default P3-2 scheme, enhanced Z_{DR} on the southern flank unrealistically extends far to the east of the updraft and does not monotonically decrease along the deep-layer storm-relative wind vector (to the northeast). The surface rain size distribution in P3-2_0.2$\Lambda_{r,min}$ resembles that of iceCat 1 above the melting level (Fig. 3.6e), where the largest ice particles (which are smaller than MY2 hail and NSSL rimed ice) are also advected well to the east of the updraft. Rain CFADs demonstrate that relatively large rain size ($D_{mr} \geq 2$ mm) below the melting level does not change much in either the

default P3-2 or P3-2_0.2$\Lambda_{r,min}$ (Figs. 3.14c,f). Rather, P3-2_0.2$\Lambda_{r,min}$ is able to simulate a more broad rain size range (and therefore, less concentrated D_{mr} frequency and an improved surface rain size gradient) as reducing $\Lambda_{r,min}$ allows for larger raindrops.

3.4.3 P3-2 large ice treatment

One key difference between the simulation with the P3-2 scheme and those with MY2 and NSSL is the absence of large ice in the P3-2 run sedimenting to the lowest model level. Typically, bulk ice size in P3-2 does not reach the large sizes of hail found in the MY2 and NSSL simulations near the surface. The default P3 scheme limits the number-weighted mean ice size D_{ni} to 2 mm, which potentially limits the ice growth and accelerates melting. Therefore, we analyze additional sensitivity tests with D_{ni} limit set to 7 mm (D_{ni}_7) and 12 mm (D_{ni}_12) to examine the relationship between this limit and ice sedimenting to the surface (Fig. 3.15). Indeed, the gradual relaxation of this limit continually increases both the size and frequency of ice reaching the surface. Therefore, the default P3-2 is not removing rimed ice at a greater rate than other BMPs by design, but rather not producing large enough ice at the surface due to the restrictive D_{ni} limit. Mass-weighted mean diameter of iceCat 1 D_{mi1} and iceCat 2 D_{mi2} in both D_{ni} tests regularly exceed 12 mm, which is closer to the surface hail size in MY2 and NSSL than the default P3-2 D_{mi1} and D_{mi2}. As a result, this larger ice reduces Z_{DR} near the observed hail signature location (Kumjian and Ryzhkov 2008), but also unphysically sediments into the forward flank.

To quantify the relationship between P3-2 large ice production and its D_{ni} limit, Fig. 3.16 shows a time series of domain-averaged maximum rimed-ice size ("mean max diameter") in the updraft (vertical velocity w > 10 m s^{-1}) for the default BMPs and D_{ni} limit sensitivity tests. We define maximum rimed-ice size as the largest diameter that the PSD number density N(D) exceeds 10^{-4} m^{-3} mm^{-1} (Snook et al. 2016; Labriola et al. 2017). We additionally impose a constraint of

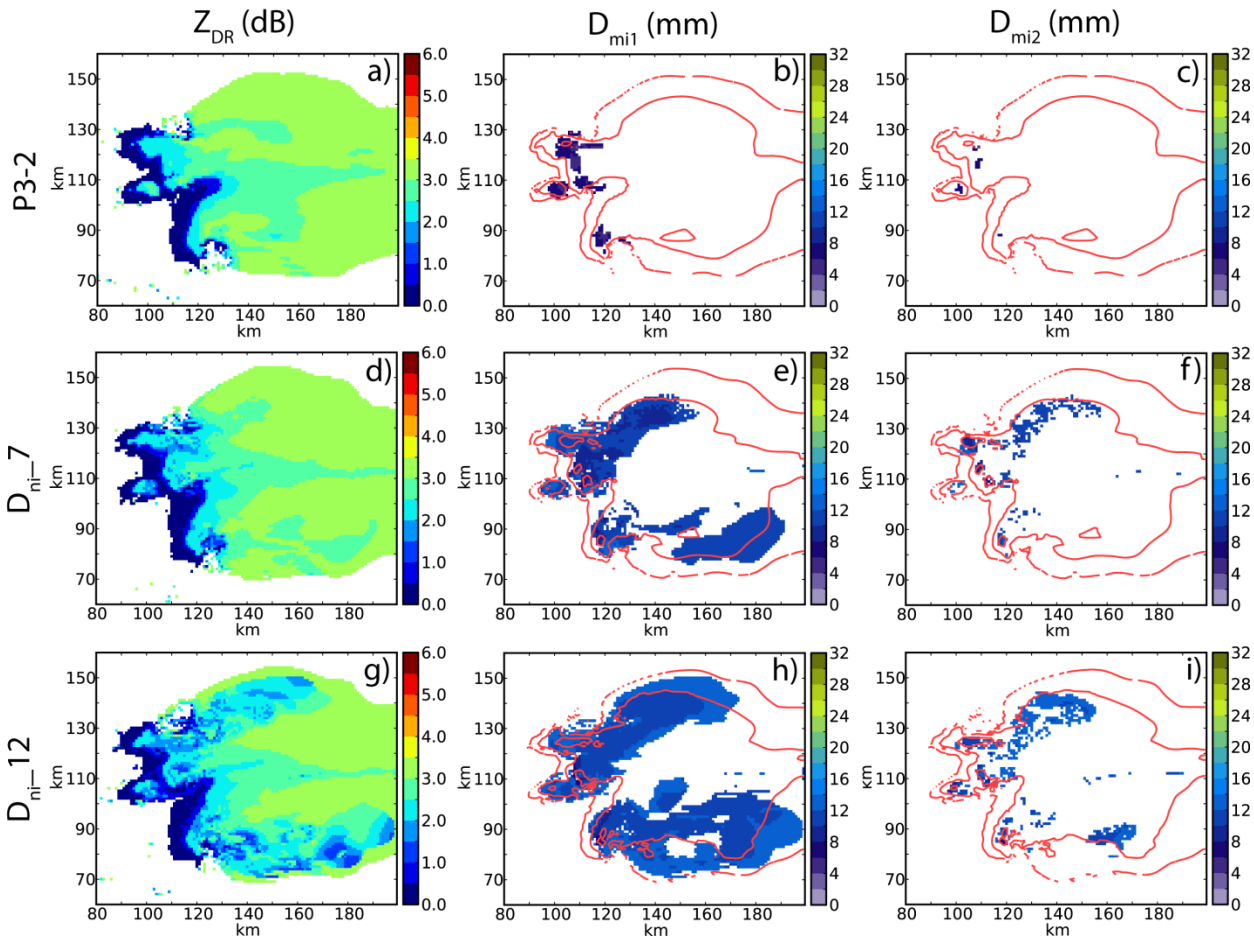

Fig. 3.15. Differential reflectivity Z_{DR} (dB), and mass-weighted mean diameter of iceCat 1 D_{mi1} (mm) and iceCat 2 D_{mi2} (mm) near z = ~280 m for the P3-2 (a,b,c) default scheme, (d,e,f) with number-weighted mean diameter limit set to 7 mm and (g,h,i) 12 mm at t = 100 min. Horizontal reflectivity Z_H contours are overlaid in 20 dBZ intervals starting at 15 dBZ on D_{mi} plots.

rime fraction f_r > 0.9 in P3-2 to ensure the ice is sufficiently rimed to be labeled "rimed ice" in traditional BMP notation, allowing for greater consistency with MY2 and NSSL large ice production. A natural separation exists between graupel and hail in the MY2 and NSSL schemes, with maximum hail size ranging between 18-27 mm and graupel between 8-14 mm after 40 min. Indeed, the default P3-2 categories also follow this separation with iceCat 1 tending to resemble the hail categories in MY2 and NSSL while iceCat 2 tends to resemble graupel (Fig. 3.16a) in these simulations. As iceCat 2 always merges to iceCat 1, it appears iceCat 2 does not have

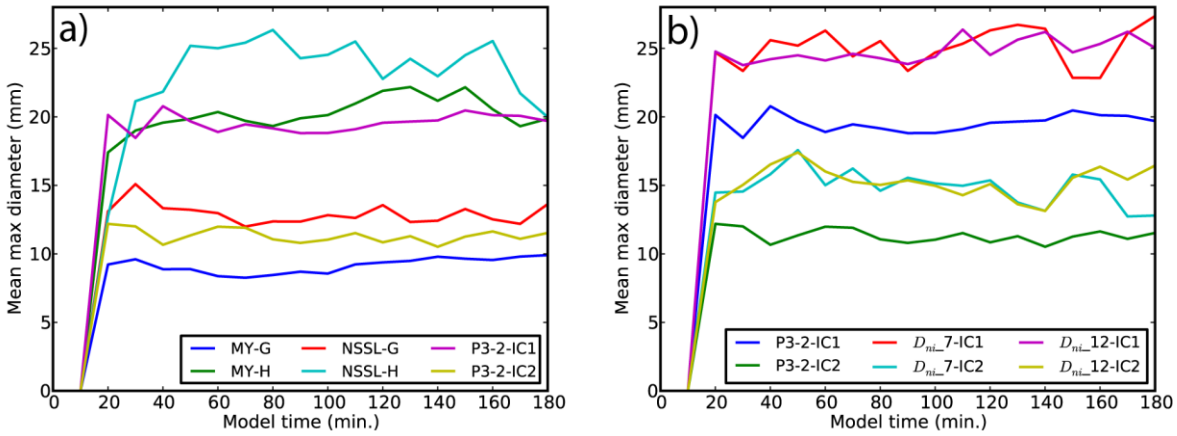

Fig. 3.16. Mean max diameter (mm) of rimed ice in the updraft for the (a) MY2, NSSL, and P3-2 BMPs and (b) P3-2 number-weighted mean diameter limit tests over the duration of the model run.

sufficient time to reach iceCat 1 sizes in the updraft. Maximum iceCat 1 size in the default P3-2 scheme is generally smaller than hail in the other bulk schemes, limiting the sedimentation of ice to the surface. When relaxing the D_{ni} limit, both iceCat 1 and 2 increase in size, with D_{ni}_7 and D_{ni}_12 iceCat 2 exceeding MY2 and NSSL maximum graupel size, and D_{ni}_7 and D_{ni}_12 iceCat 1 following NSSL hail more closely after t = 40 min (Fig. 3.16b). Therefore, relaxing the D_{ni} limit clearly increases the production of large rimed ice in P3-2, which in turn is then able to sediment to the surface. This relationship is not simply one-to-one, as the D_{ni}_7 iceCat 1 and 2 exceed D_{ni}_12 iceCat 1 and 2 at times during the simulations despite having a more stringent D_{ni} limit. This indicates that while relaxing the D_{ni} limit clearly increases the potential for rimed ice to grow in the scheme, the degree to which ice grows depends more on the ice growth processes themselves rather than the D_{ni} limit. Nonetheless, it is clear that the melting formulation in P3-2 is not by itself aggressively depleting ice in the default configuration, but rather the melting of relatively small particles produced by this scheme explains the lack of large ice near the surface.

3.5 Summary and discussion

We perform idealized supercell simulations using WRF v3.9.1 to determine how the representation of rimed ice in bulk microphysics schemes can explain the presence or lack of polarimetric signatures. The BMPs examined are the two-moment versions of Milbrandt-Yau (MY2), National Severe Storms Laboratory (NSSL), and the two-category configuration of the P3 scheme (P3-2). The simulated polarimetric signatures considered are the Z_{DR} arc and the hail signature in the forward flank downdraft, which are particularly sensitive to rimed-ice representation. Both the MY2 and NSSL BMPs are generally able to simulate a Z_{DR} decrease in the direction of the deep-layer storm-relative mean wind vector near the surface (though weakly in MY2), but are unable to simulate a Z_{DR} arc entirely consistent with observations. Large, relatively dry hail (i.e., low Z_{DR}) reduces Z_{DR} on the southern flank in the MY2 supercell, while the NSSL scheme simulates a compact supercell unable to replicate an elongated Z_{DR} arc. The P3-2 scheme simulates a weak Z_{DR} gradient in the forward flank, despite containing two-moment rain and two two-moment free ice categories. Frequency plots suggest that while iceCat 1 and iceCat 2 exhibit clear differential sedimentation, the rimed ice in P3-2 overwhelmingly melts homogenously to larger raindrops with a limited size range compared to the other two BMPs. This is due to the scheme's restrictive minimum rain PSD slope $\Lambda_{r,min}$ bound, which prevents rain in the scheme from reaching larger sizes and simulating a Z_{DR} arc and its associated gradient at the surface. Reducing this parameter facilitated a Z_{DR} gradient more consistent with observations than that in the default P3-2 scheme, but also simulated a Z_{DR} arc extending well into the forward flank and an associated Z_{DR} gradient that failed to decrease monotonically in the direction of the deep-layer storm-relative mean wind vector (to the northeast). Therefore, aspects of P3-2 that limit size sorting may require future improvement to better simulate the Z_{DR} arc.

Only the NSSL scheme correctly simulates the location of the hail signature in the FFD most consistent with observations, with large, dry hail in the scheme falling out of the updraft and quickly reaching the surface. The MY2 scheme does produce large hail that easily reaches the surface but it generally appears on the southern flank of the supercell. The smaller MY2 hail fall speed assumption compared to NSSL is the likely explanation for such a behavior. Still, MY2 simulates large hail better than earlier versions due to an improved (more restrictive than in Johnson et al. (2016)) hail-to-graupel conversion threshold in the default scheme. The MY2 scheme reduces ρ_{HV} more than the NSSL scheme in the location of the hail signature. Although hail is more wet in the NSSL scheme, MY2 hail is larger and therefore contains more particles prone to the resonance effect. Ice in the default P3-2 is generally unable to reach the surface, and the small, wet particles that do reach the surface increase Z_{DR}. Consequently, minor surface ρ_{HV} reductions in the scheme are typically caused by large, oblate drops. The default P3-2 restricts ice number-weighted mean size D_{ni}, and produces larger ice that reaches the surface and reduces Z_{DR} when this limiter is relaxed. However, these particles also unphysically sediment into the forward flank, implying that optimal D_{ni} tuning remains future work. In terms of simulated rimed-ice density, NSSL graupel is typically prognosed as more dense than the constant graupel density in MY2, while NSSL hail is prognosed as less dense than the constant MY2 hail density. This represents a more continuous wet growth process and transfer between the two categories in the NSSL scheme. With the only discriminator between P3-2 ice categories being the mass-weighted mean diameter, P3-2 bulk ice densities are very similar between the two ice categories.

Further analyses of microphysical tendencies reveal differing ice treatment philosophies between the schemes. Whereas MY2 graupel generally forms by 3-component freezing, NSSL graupel is typically formed from freezing rain. In other words, MY2 prefers an ice seed for rimed-

ice initiation in these idealized supercell simulations. The primary hail origin in the MY2 and NSSL schemes is graupel. While MY2 and NSSL hail origin are similar, the NSSL hail category is more representative of a large hail category, reflected by the scheme's prognostic graupel density (i.e., only large, growing graupel may transfer to the hail category) and process rates. Although Milbrandt and Morrison (2013) added prognostic graupel bulk volume to the MY2 scheme, the rimed ice process rates and parameterizations would also need to reflect a spectrum of growing rimed ice to large hail (as in the NSSL scheme) to improve MY2 supercell simulated polarimetric signatures. In the free category approach of the P3-2, several sources (i.e., nucleation, freezing cloud water, 3-component freezing) act to populate ice. Integral to the growth of ice in each scheme is the collection of liquid, though MY2 hail seems to excessively collect rain below the melting level. Hail melting is offset by rain collection in the MY2 and more gradual in the NSSL scheme, representing a weaker net process than graupel melting. Melting of P3-2 iceCat 1 is more comparable to MY2 and NSSL graupel melting, a reflection of its restrictive D_{ni} limiter (default value of 2 mm) producing relatively small ice particle mean sizes.

At this point, it is worth reiterating the P3's treatment of ice, particularly rimed ice. In the scheme, a single PSD contains all ice modes, including small spherical, larger non-spherical, and rimed ice. The degree to which rimed ice occupies the ice PSD depends on the riming fraction, which for the idealized supercell case examined in this paper is frequently large due to the storm's dominant mode of riming growth. This BMP framework can be problematic when newly initiated small ice (such as the previously mentioned nucleation and freezing cloud water sources) is assigned to an ice category dominated by rimed ice. Such a process would "dilute" the category by increasing the amount of small ice within the category's PSD, and the category's processes (i.e., riming growth, melting) would subsequently become more reflective of small ice (Milbrandt

and Morrison 2016). This limitation is not explicitly present in traditional bulk schemes, which allows predefined rimed ice to grow independently of traditionally smaller ice. Still, this P3 rimed ice dilution could be mitigated by increasing the number of ice categories or modifying new ice category destination based on its initiation process (i.e., separating nucleation from freezing liquid).

The P3-2's default ice number-weighted mean diameter limit of 2 mm is overly restrictive for modeling hail. In fact, having a fairly restrictive limiter when riming is unimportant may be necessary as there is no explicit representation of snow particle breakup in P3 (although most BMPs neglect explicit snow breakup). Hence, relaxing the limit improved the performance of the scheme in terms of its simulation of large rimed ice near the surface, but has the potential to grow unrimed or lightly-rimed non-spherical ice in the scheme too large. Future work should investigate this problem further. It may be possible to apply a more sophisticated size limiter that varies with ice properties such as rimed fraction. Another possibility could be to add an explicit parameterization of snow particle breakup and relax the size limiter. A three-moment ice version of P3 evolving the size distribution shape parameter explicitly is currently in development, which may also obviate the need for a restrictive size limiter. Note, a three-moment rain version of P3 has already been developed (Paukert et al. 2019), though the rain size limiter in the two-moment version is not an issue since drop break-up is already parameterized. Also, a prognostic liquid fraction on ice has recently been developed for P3 (Cholette et al. 2019); future work will examine the effects of this on the simulation of melting hail and the impacts on simulated polarimetric signatures. Overall, the results of this study suggest it is important to compare newly designed microphysics schemes with existing, state-of-the-art schemes to understand their behaviors and performance, and to ensure that the intended improvements are realized in these schemes.

Chapter 4 Ice crystal and snow representation in an idealized supercell thunderstorm using spectral bin and bulk microphysics schemes

4.1 Introduction

While many studies have shown that ice microphysical parameterizations play an important role in the evolution of convective storms in numerical studies (e.g., McCumber et al. 1991; McFarquhar et al. 2006; Morrison and Milbrandt 2011; Van Weverberg et al. 2012), the impact of different ice parameterizations in these complex, highly non-linear microphysics schemes on the prediction of storms is less understood and non-trivial to evaluate. In numerical weather prediction (NWP) microphysics, frozen particles are largely categorized into two groups: rimed ice (e.g., graupel and hail) and unrimed ice (e.g., ice crystals and snow). Considering unrimed ice exhibits a wide and complex range of habits as a function of temperature and supersaturation over ice (see Fig. 5 in Bailey and Hallett 2009), which is not well-understood, it is extremely difficult to effectively and accurately parameterize complex physical processes involving unrimed ice. Therefore, unrimed ice habits and processes are inevitably oversimplified in NWP models and their parameterizations largely rely on each microphysics developer's own philosophy and understanding of those processes.

The evolution of a hydrometeor's particle size distribution (PSD) in NWP models is typically constructed in either a bulk or spectral bin framework. While bulk microphysics parameterization schemes (BMPs) assume a fixed form of an underlying hydrometeor PSD (e.g., Lin et al. 1983; Tao and Simpson 1993; Straka and Mansell 2005), spectral bin microphysics schemes (SBMs) discretize the PSD into bins, generally predicting the evolution of either the PSD itself or moments in each bin (e.g., Hall 1980; Reisin et al. 1996; Geresdi 1998). While BMPs gained popularity due to their high computational efficiency, SBMs allowed the addition of

particle attributes (i.e., axis ratio, density) with greater flexibility, albeit with a steep computational price. As an example, a modern 2M BMP with six hydrometeors contains 12 prognostic variables, while the spectral bin Hebrew University Cloud Model with eight hydrometeors (including cloud condensation nuclei) and thirty-three PSD bins contains 264 prognostic variables, an increase by a factor of 22. In fact, SBMs with ice categories were overwhelmingly two-dimensional (2-D) model runs, eventually expanding to three dimensions (e.g., Lynn et al. 2005). For more details on microphysical design, readers are referred to a comprehensive review of bulk and spectral bin schemes by Khain et al. (2015).

In real clouds, ice can grow into various modes (i.e., needles, rosettes, dendrites) for small changes in temperature and ice supersaturation (e.g., Magono and Lee 1966; Bailey and Hallett 2009) and possess different deposition/sublimation rates (e.g., Pruppacher and Klett 1997; Westbrook et al. 2008; Sulia and Harrington 2011). BMPs have generally assumed spherical ice crystals (e.g., Ferrier 1994; Milbrandt and Yau 2005a, b), but other BMPs have either parameterized or diagnosed ice habit (e.g., Walko et al. 1995; Hong et al. 2004). Recently, new microphysics schemes prognose ice crystal axis ratio in order to predict ice habit (e.g., Harrington et al. 2013; Chen and Tsai 2016). While computationally cheaper and adding needed ice complexity, the listed studies are still treating aspect ratio as a bulk quantity. On the other hand, the bin nature of SBMs allows for greater flexibility to represent various ice crystal modes within a single ice category. Young (1974) and Takahashi (1976) partitioned ice crystal size bins, respectively, into x- and z- directions, while separate hydrometeor categories (i.e., columns, dendrites) with binned hydrometeor information (i.e., density) were added to SBMs (Khain and Sednev 1996; Khain et al. 2004). Perhaps most fundamental of all, SBMs can inherently capture

observed bimodal (Lawson et al. 2006), normalized (Delanoë et al. 2005), and other ice distribution types rather than relying on an assumed PSD (as in BMPs).

Snow in BMPs is usually assigned a constant bulk density and an assumed spherical shape. The Thompson (Thompson et al. 2008) and Milbrandt-Yau (Milbrandt and Yau 2005a, b) schemes have broken from this pervasive parameterization, instead setting the snow mass exponential such that density decreases with increasing snow particle diameter to better match observed snow mass-diameter relationships (Mitchell et al. 1990; Brandes et al. 2007). The Thompson scheme also represents snow using a linear combination of an exponential and gamma distribution based on snow moments following observed snow PSDs in Field et al. (2005). These snow improvements have allowed the Thompson scheme to outperform other schemes in regard to snowfall prediction (Liu et al. 2011). Again, this added complexity is cheaper than a spectral bin framework, but is also less flexible. Specifically, Field et al. (2005) demonstrated this PSD fit is not applicable to every snow PSD. SBMs, on the other hand, are not limited by a smooth distribution approximation. Khain et al. (2012) further added snow complexity by computing rime snow mass for each mass bin (compared to a single bulk, continuous bulk, or offline bin rime mass lookup table limitation in BMPs), aiming for a more realistic representation of snow density and conversion of riming snow to graupel. Predicting snow density (i.e., including riming density changes) would be less costly in a BMP, and has the potential to produce reasonable results.

As supercell microphysical processes are typically dominated by rimed ice (e.g., Musil et al. 1986; Illingworth et al. 1987; Kumjian and Ryzhkov 2008), ice crystals/snow are underrepresented in supercell hydrometeor analyses in literature. Although simulated deep-convection precipitation may be less sensitive to snow (Gilmore et al. 2004; Van Weverberg et al. 2011), motivation still exists to accurately parameterize ice crystals and snow (aggregates) due to

their role in lightning initiation from rimed-ice – ice collisions (e.g., Reynolds et al. 1957; Takahashi 1978; Keith and Saunders 1989; Berdeklis and List 2001). Ice crystal and snow parameterization differences can also manifest themselves when assimilating satellite brightness temperature observations, which are sensitive to the hydrometeor PSDs (particularly cloud top ice). In this study, the behavior of spectral and bulk microphysics schemes' representation of ice crystal and snow categories using the Weather Research and Forecasting-Advanced Research WRF (WRF-ARW) model (Skamarock et al. 2008) are examined predominantly by moment distributions in a supercell framework. Particularly, the current increased complexity and flexibility of the HUCM scheme is compared to more simplified bulk schemes in terms of ice behavior. Such a comparison can drive the improved representation (both parameterization and processes) of ice crystal modes and snow in microphysics schemes by helping to clarify the benefit of additional ice complexity. The rest of the paper is organized as follows: section 4.2 details the model setup, section 4.3 compares ice crystal and snow moments and distributions in the simulated supercell, section 4.4 quantifies hydrometeor mean particle size, and section 4.5 summarizes and discusses ice representation.

4.2 Simulation design

4.2.1 Numerical model

Idealized supercell thunderstorms are simulated using the compressible, nonhydrostatic WRF-ARW model version 3.7.1 (Skamarock et al. 2008). Model configuration is similar to Johnson et al. (2016) and is detailed in Table 4.1. Storms are simulated for 2 hours using a convective-scale horizontal spacing of 1 km. The Weisman and Klemp (1982) thermodynamic sounding with a veering quarter-circle wind profile is employed for the atmospheric environment. Convection initiates using an ellipsoidal thermal bubble with maximum potential temperature

Table 4.1.WRF model input.

WRF Model Configuration	
Run time	120 min.
Δt	6 s
Sound wave Δt	1 s
Model output interval	10 min.
Horizontal domain	200 km x 200 km
Model lid	20 km
Δx	1 km
Δy	1 km
Δz	~500 m
Time integration scheme	3rd order Runge-Kutta
Horizontal momentum advection	5th order
Vertical momentum advection	3rd order
Horizontal scalar advection	5th order
Vertical scalar advection	3rd order
Upper level damping	5000 m below model top
Rayleigh damping coefficient	0.003
Horizontal boundary conditions	Open

perturbation θ' of 3 K. For simplicity, radiation, land surface, cumulus, and planetary boundary layer parameterization is neglected. For more configuration details, the reader is also referred to Morrison and Milbrandt (2011).

4.2.2 Microphysics schemes

Three microphysics schemes are analyzed in supercell simulations: the spectral bin HUCM microphysics scheme (Khain and Sednev 1996; Khain et al. 2004), and the bulk National Severe Storms Laboratory (NSSL; Mansell et al. 2010) and Thompson (Thompson et al. 2008) microphysics schemes. HUCM microphysics contains 33 mass-doubling bins, ranging from 3.35×10^{-11} to 1.44×10^{-1} g, to solve for the discretized PSDs of 7 hydrometeor categories (water,

ice crystals [plates, columns, dendrites], snow, graupel, and hail). The NSSL scheme predicts mass mixing ratio q_x and number concentration N_{tx} of cloud droplets, rain, cloud ice, snow, graupel, and hail, and additionally particle volume of graupel and hail. The Thompson scheme predicts mass mixing ratio of cloud droplets, rain, cloud ice, snow, and graupel, and number concentration of rain and cloud ice. Therefore, Thompson snow number concentration is diagnosed. Among many available BMPs, the NSSL scheme is one of the most sophisticated two-moment BMPs and also proven to outperform other BMPs in supercell simulations (Johnson et al. 2016), while the Thompson scheme is employed in operational storm-scale CAMs (e.g., High-Resolution Rapid Refresh).

We would like to point out that hydrometeors in each microphysics schemes may contain different assumptions of PSDs, attributes, and properties. Cloud ice in the NSSL and Thompson schemes are assumed spherical with constant bulk densities of 900 kg m^{-3} and 890 kg m^{-3} respectively. Snow is assumed spherical in the NSSL scheme with bulk density of 100 kg m^{-3}, while snow in the Thompson scheme contains a mass exponent of 2, resulting in snow bulk density:

$$\rho_s(D) = \frac{6}{\pi} \times 0.069 D^{-1} \tag{4.1}$$

i.e., density decreases with increasing diameter. Finally, the NSSL scheme assumes mass distributions for cloud ice and snow as gamma distributions with shape parameter $\alpha = 0$ and –0.8 respectively. The Thompson scheme assumes an exponential ($\alpha = 0$) cloud ice distribution and a linear combination of exponential and gamma distributions for snow (Eq. 1 in Thompson et al. 2008). For details on the rimed ice parameterizations of the bulk schemes, the reader is referred to Johnson et al. (2016).

In HUCM microphysics, ice crystal mode from nucleation is driven by the ambient temperature (Table 4.2). Ice crystal and snow hydrometeor densities generally vary with diameter

Table 4.2. HUCM ice crystal mode during nucleation based on temperature.

Temperature (°C)	Ice crystal mode
$T_c \leq -22.4°$	Column
$-22.4° < T_c \leq -17.8°$	Plate
$-17.8° < T_c \leq -12.7°$	Dendrite
$-12.7° < T_c \leq -8.1°$	Plate
$-8.1° < T_c \leq -4°$	Column
$T_c > -4°$	Plate

(though ice plates are assumed to have constant density of 900 kg m^{-3}; Fig. 4.1). Column bulk density is typically large ($\rho_{i,col} > 800$ kg m^{-3}), while dendrites and snow span both low (~100 kg m^{-3}) and high (600–900 kg m^{-3}) bulk densities. Column, dendrite, and snow bulk densities decrease for larger particle sizes. The Thompson snow mass relationship is also shown in comparison to HUCM mass relationships, and typically produces lower bulk densities than snow and dendrites in the scheme for melted diameter > 0.2 mm.

4.3 Simulated ice crystal and snow moments

The HUCM's ability to simulate a supercell is examined by comparing horizontal reflectivity Z_H at the lowest model level (z = ~280 m) and through the updraft with the 2M bulk NSSL and Thompson schemes (Fig. 4.2). Z_H is calculated using the CAPS polarimetric radar data simulator (e.g., Jung et al. 2008; Dawson et al. 2014; Johnson et al. 2016), which utilizes the T-

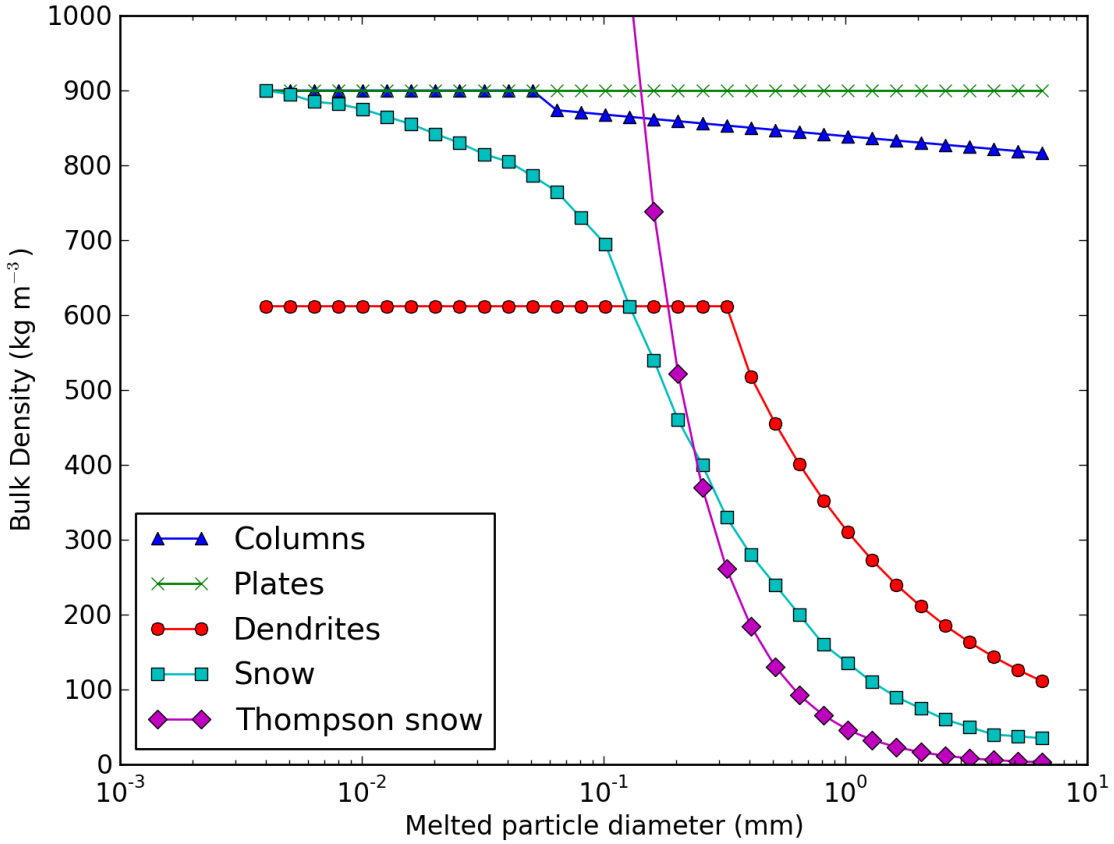

Fig. 4.1. Bulk densities (kg m^{-3}) of columns, plates, dendrites, and snow in HUCM microphysics as a function of their melted particle diameter (mm). Also shown is the bulk density of snow in the Thompson scheme. Adapted from Khain et al. (2004) Fig. 1.

matrix method (Waterman 1969; Vivekanandan et al. 1991) to calculate scattering amplitudes as a function of particle diameter and water fraction. Particle diameter is assumed equivolume and calculated using the mass bins of hydrometeors in the HUCM, and q_x and N_{tx} in the 2M bulk schemes, while water fraction is calculated as in Jung et al. (2008). For the HUCM scheme, wet snow equivolume diameter is recalculated based on the mixed-phase hydrometeor's new density taking into account the shrinking of the horizontal dimension of snow with progressive melting. Given the ice crystal complexity of the HUCM, ice crystals are now included in the simulator for the scheme and are assumed to melt when ambient temperature exceeds 0°C. Columns, plates, and

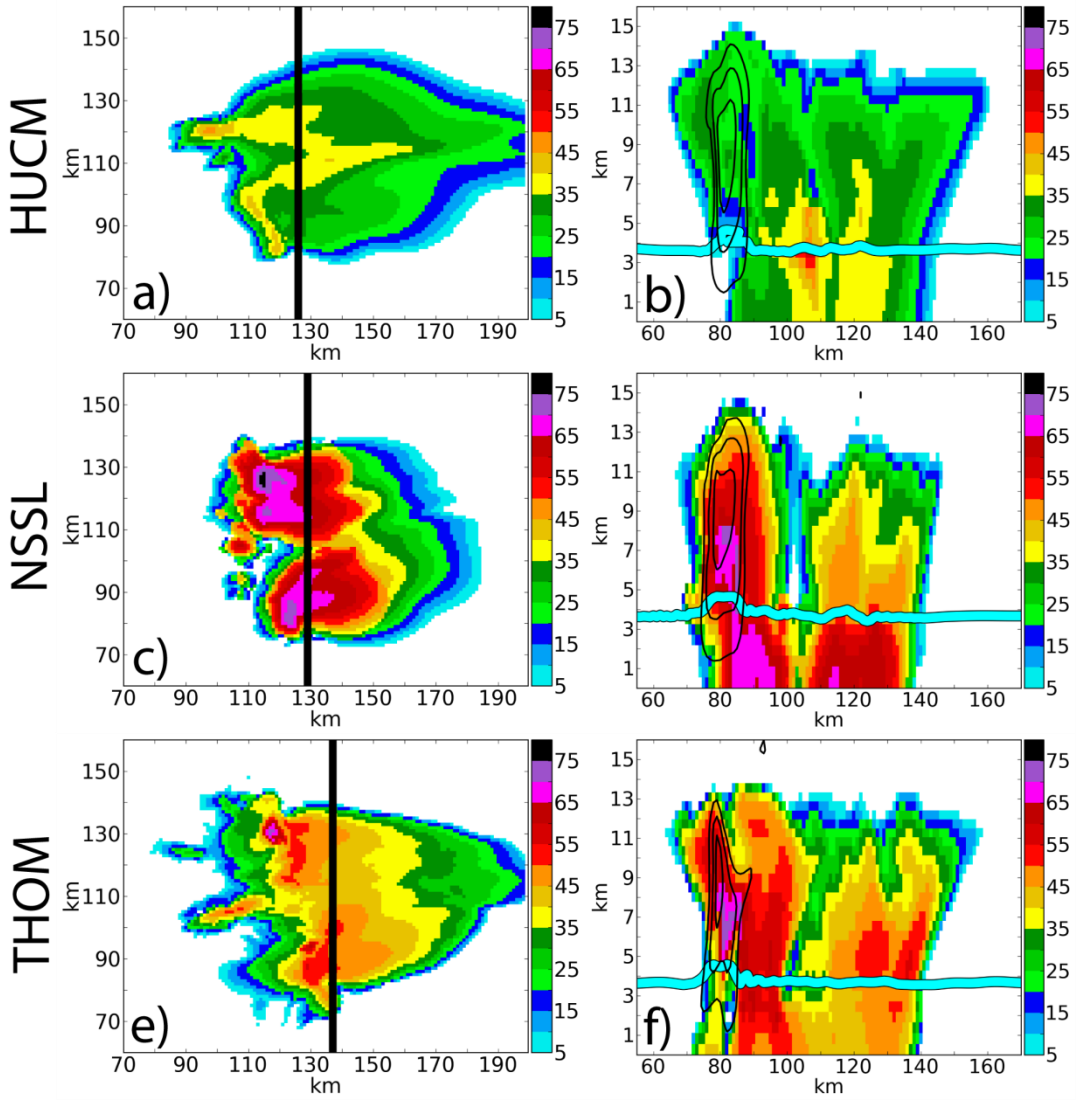

Fig. 4.2. Horizontal reflectivity Z_H for the (a,b) HUCM, (c,d) NSSL, and (e,f) Thompson schemes (left column) at z = ~280 m and (right column) through the updraft at t = 100 min. The black lines in the left column display the location of the vertical cross sections, while blue lines in vertical cross sections denote the 0°C isotherm. Black contours in the right column are vertical velocity starting at 10 m s^{-1} with 15 m s^{-1} interval.

dendrites are assumed to have aspect ratios of observed solid columns ($L/d \leq 2$), solid thick plates, and dendrites respectively (Matrosov et al. 1996). Oblate crystal orientation is assumed to follow a two-dimensional axisymmetric Gaussian distribution, while prolates follow horizontal random orientation (Ryzhkov et al. 2011).

The HUCM is properly simulating a splitting supercell with a prominent hook echo and

forward flank (Fig. 4.2a), similar to the NSSL and Thompson BMPs (Figs. 4.2c,e). However, low-level reflectivity is generally much smaller than the other two schemes. The default rigid HUCM mass bins dictate a maximum rain, graupel, and hail equivolume diameter of 6.50, 8.82, and 6.73 mm, respectively. These rain and wet rimed ice particles at the surface are unable to grow further and produce smaller reflectivites than in the NSSL and Thompson schemes, which do not contain the rigid maximum diameter limitations. Similarly, vertical Z_H through the updraft in the HUCM (Fig. 4.2b) is generally smaller both above and below the melting level compared to the 2M bulk schemes (Figs. 4.2d,f). In fact, the HUCM underpredicts maximum Z_H in the updraft by about 40 dBZ. Again, large rain and wet rimed ice are limited even in the updraft by the specification of HUCM mass bins. Because of their small sizes, the inclusion of HUCM ice crystals in regard to reflectivity is generally small.

4.3.1 Time evolution

Temporal evolution of ice crystal and snow hydrometeor moments in the HUCM, NSSL, and Thompson schemes over the model run are analyzed to compare how much cloud ice and snow each microphysics scheme produces (Fig. 4.3). Microphysical tendencies for ice crystals and snow are analyzed to explain which processes contribute to moment behavior (Figs. 4.4; 4.5). Tendency nomenclature borrows from Milbrandt and Yau (2005b) and is as follows: the first letter denotes the relevant moment, the next two letters denote the process (VD: vapor deposition, CL: collection [collision in HUCM nomenclature], NC: nucleation, FZ: freezing, ML: melting, SP: splintering, CN: conversion, VS: vapor sublimation), the next one (or two for three-component processes) letter denotes hydrometeor sinks, while the last letter denotes the hydrometeor source (L: HUCM liquid, W: cloud water, R: rain, P: plate, C: column, D: dendrite, I: ice crystal, S: snow, G: graupel, H: hail). Because the Thompson scheme only predicts snow mixing ratio, there are no

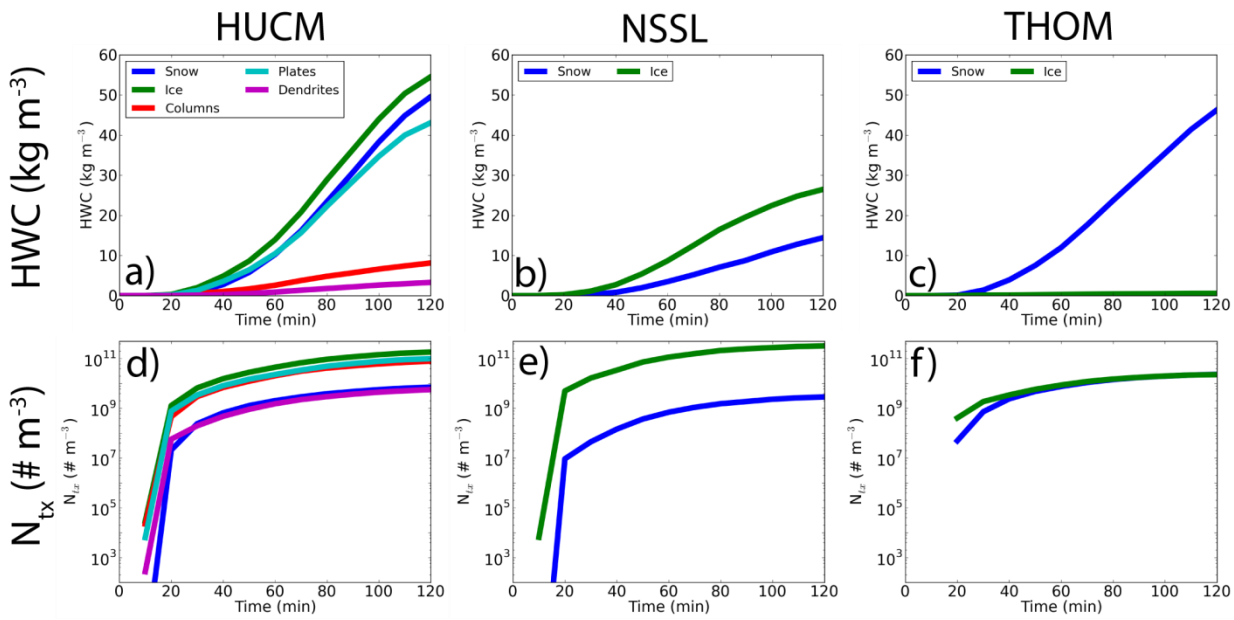

Fig. 4.3. Time series of domain-integrated ice crystal and snow hydrometeor water content (kg m^{-3}; top row) and number concentration N_{tx} (# m^{-3}; bottom row) for the (a,d) HUCM, (b,e) NSSL, and the (c,f) Thompson microphysics schemes for the duration of the model run.

corresponding snow number tendencies in the scheme. A description of each microphysical term examined is available in Table 4.3.

In the HUCM scheme, plates are the dominant ice crystal habit and close to total ice water content (sum of plates, columns, and dendrites) and number concentration (green, Figs. 4.3a,d). This is because of the large amount of freezing liquid into the category (Figs. 4.4c,d), which far surpasses the mass source terms for columns (Fig. 4.4a) and dendrites (Fig. 4.4e). Plates (Fig. 4.4d) and columns (Fig. 4.4b) also nucleate more than dendrites (Fig. 4.4f), likely because they encompass a larger nucleation temperature range (Table 4.2). Snow water content is similar to plate water content, and much larger than column and dendrite water content (Fig. 4.3a). This is reasonable as the largest snow source is the collection of crystals (i.e., aggregation; Figs. 4.4g,h). While column water content is generally much smaller than that of plates, its number concentration is almost as large as that of plates. This suggests the overall size of columns are much smaller

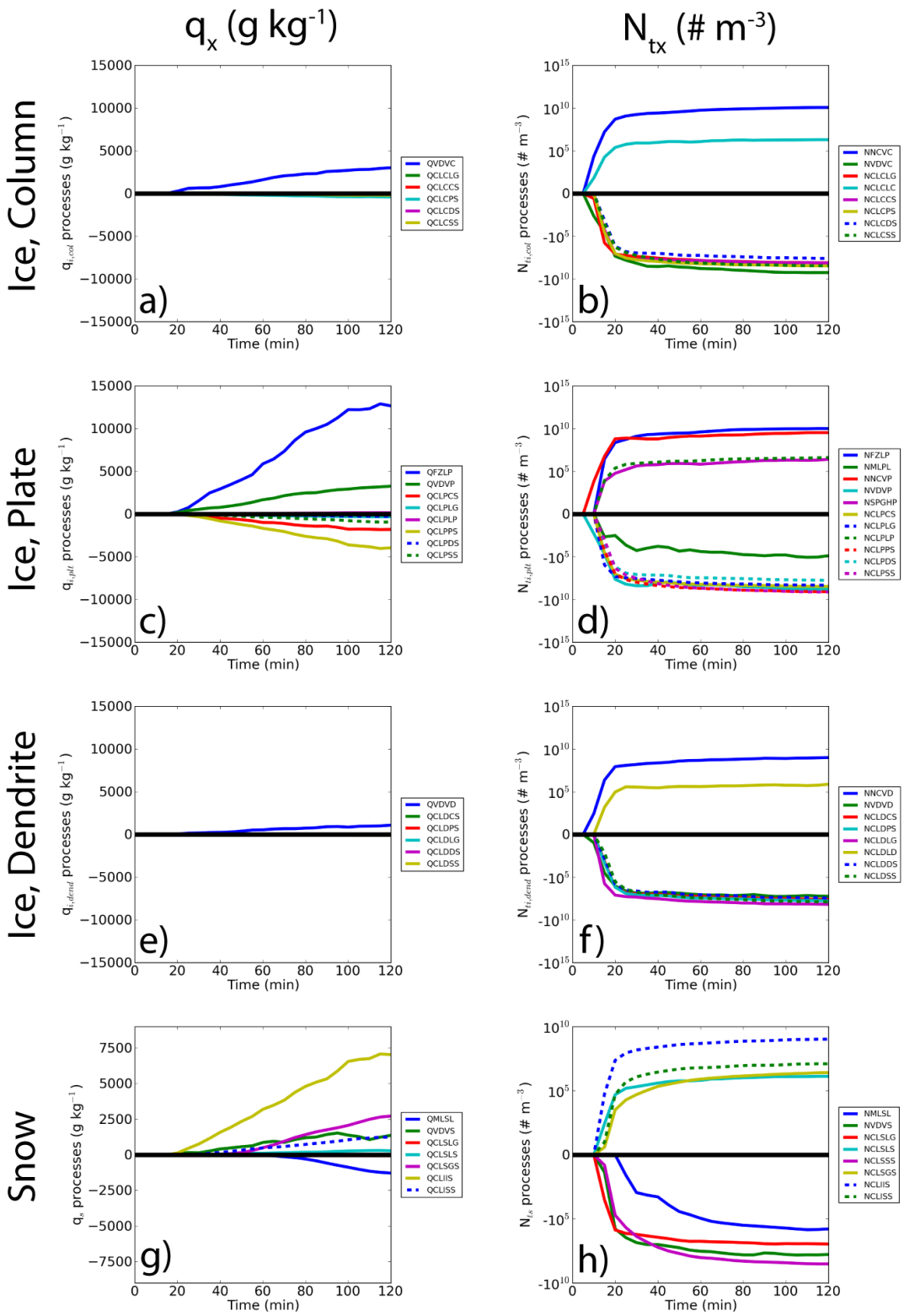

Fig. 4.4. Time series of mixing ratio q_x (g kg^{-1}; left column) and number concentration N_{tx} (# m^{-3}; right column) microphysical tendency sums for HUCM (a,b) columns, (c,d) plates, (e,f) dendrites, and (g,h) snow for the duration of the model run. The horizontal black line in each plot denotes the zero line separating source and sink terms.

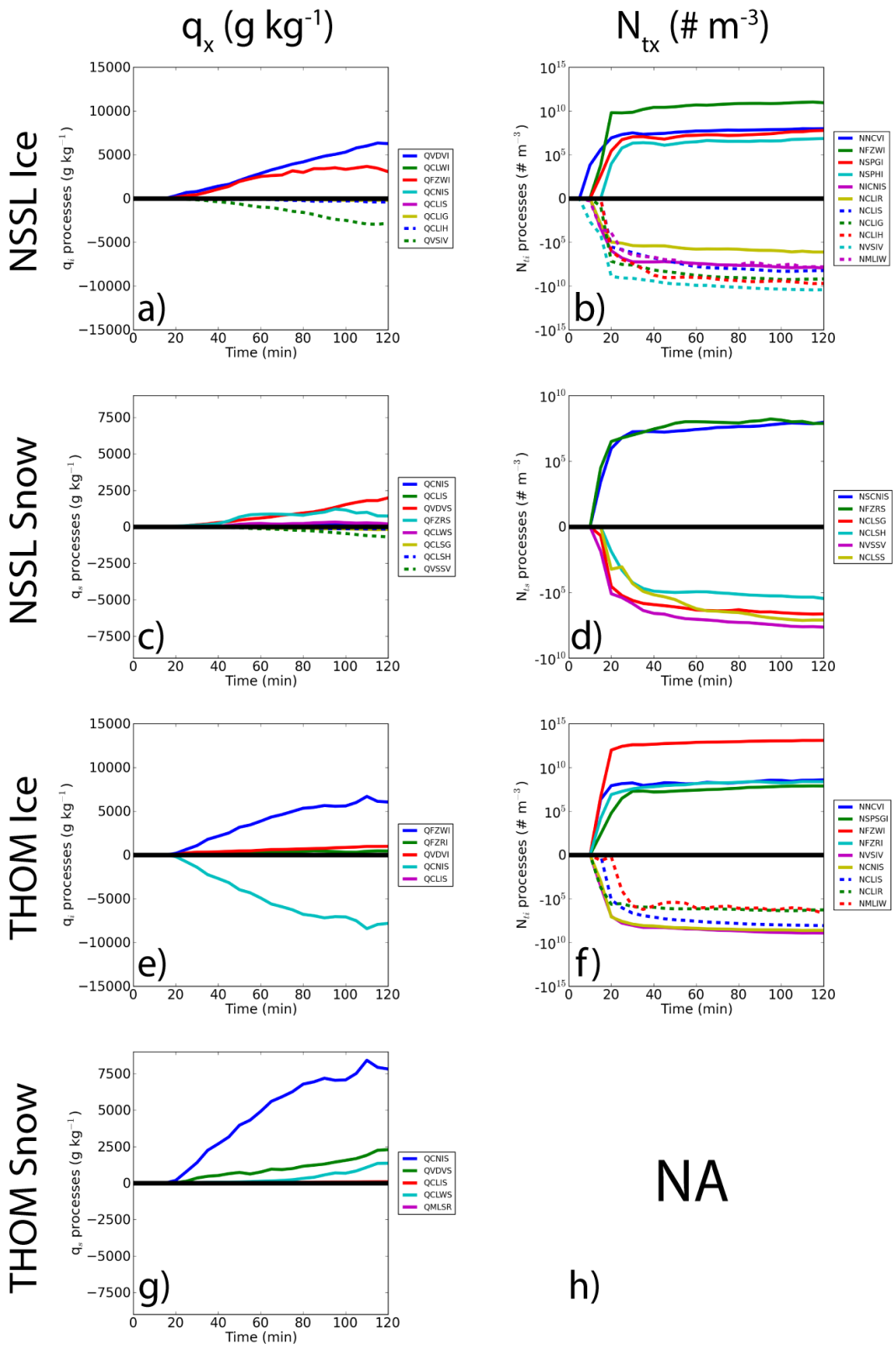

Fig. 4.5. As in Fig. 4.4, for NSSL (a,b) ice crystals and (c,d) snow, and Thompson (e,f) ice crystals and (g,h) snow.

Table 4.3. Ice crystal and snow microphysical processes in the HUCM, NSSL, and Thompson schemes. Column, plate, and dendrite processes are specifically for the HUCM scheme. Tendencies that have been filtered are not included. Parentheses in the process name denote interchangeable HUCM crystals.

Process name	Description
QVDV(C)(P)(D)	Crystal mass depositional growth
QCL(C)(P)(D)LG	Crystal-liquid mass collision to graupel
QCL(C)(P)(D)/(C)(P)(D)S	Crystal-crystal mass collision to snow
QCL(C)(P)(D)SS	Crystal-snow mass collision to snow
QCL(C)(P)L(C)(P)	Crystal-liquid mass collision to crystal
QFZLP	Liquid mass freezing to plates
QMLSL	Snow mass melting to liquid
QVDVS	Snow mass depositional growth
QCLSLG	Snow-liquid mass collision to graupel
QCLSLS	Snow-liquid mass collision to snow
QCLSGS	Snow-graupel mass collision to snow
QCLIIS	Total crystal-crystal mass collision to snow
QCLISS	Total crystal-snow mass collision to snow
QVDVI	Cloud ice mass depositional growth
QCLWI	Cloud ice mass collection of cloud water
QFZWI	Cloud water mass freezing to cloud ice
QCNIS	Cloud ice mass conversion to snow
QCLIS	Snow mass collection of cloud ice
QCLIG	Graupel mass collection of cloud ice
QCLIH	Hail mass collection of cloud ice
QVSIV	Cloud ice mass sublimation
QFZRS	Rain mass freezing to snow
QCLWS	Snow mass collection of cloud water
QCLSG	Graupel mass collection of snow
QCLSH	Hail mass collection of snow
QVSSV	Snow mass sublimation
QFZRI	Rain mass freezing to cloud ice
QMLSR	Snow mass melting to rain
NNCV(C)(P)(D)	Crystal number nucleation
NVDV(C)(P)(D)	Crystal number sublimation
NCL(C)(P)(D)LG	Crystal-liquid number collision to graupel

NCL(C)(P)(D)LH	Crystal-liquid number collision to hail
NCL(C)(P)(D)L(C)(P)(D)	Crystal-liquid number collision to crystal
NCL(C)(P)(D)/(C)(P)(D)S	Crystal-crystal number collision to snow
NCL(C)(P)(D)SS	Crystal-snow number collision to snow
NFZLP	Liquid number freezing to plates
NML(C)(P)(D)L	Crystal number melting to liquid
NSPGHP	Graupel/hail number splintering to plates
NMLSL	Snow number melting to liquid
NVDVS	Snow number sublimation
NCLSLG	Snow-liquid number collision to graupel
NCLSLS	Snow-liquid number collision to snow
NCLSSS	Snow-snow number collision to snow
NCLSGS	Snow-graupel number collision to snow
NCLIIS	Total crystal-crystal number collision to snow
NCLISS	Total crystal-snow number collision to snow
NNCVI	Cloud ice number nucleation
NFZWI	Cloud water number freezing to cloud ice
NSPGI	Graupel number splintering to cloud ice
NSPHI	Hail number splintering to cloud ice
NICNIS	Cloud ice number conversion sink to snow
NCLIR	Rain number collection of cloud ice
NCLIS	Snow number collection of cloud ice
NCLIG	Graupel number collection of cloud ice
NCLIH	Hail number collection of cloud ice
NCNIG	Cloud ice number conversion to graupel
NVSIV	Cloud ice number sublimation
NMLIW	Cloud ice number melting to cloud water
NSCNIS	Cloud ice number conversion to snow source
NFZRS	Rain number freezing to snow
NCLSG	Graupel number collection of snow
NCLSH	Hail number collection of snow
NVSSV	Snow number sublimation
NCLSS	Snow number self-collection
NSPSGI	Snow/graupel number splintering to cloud ice
NFZRI	Rain number freezing to cloud ice

NCNIS	Cloud ice number conversion to snow
NCLIRS	Rain-ice number freezing to snow
NCNSG	Snow number conversion to graupel
NMLSR	Snow number melting to rain

than plates. Initially, ice crystal and snow number rapidly grows until t = 20 min., then slows down in logarithmic space.

Similar to HUCM microphysics, the NSSL scheme produces more ice crystal water content and number concentration than snow (Figs. 4.3b,e). However, much less ice crystal water content and snow water content is produced over the model run compared to the HUCM. The NSSL scheme also produces more ice crystal and fewer snow particles than the HUCM. The HUCM's primary ice source of freezing liquid is sent to either plates or hail. In the NSSL scheme, this process is spread across cloud ice, snow, and graupel. While this process adds relatively more ice number in the scheme, the ice mass added is much smaller than in the HUCM (Figs. 4.5a,b), indicating smaller NSSL cloud ice particles. The conversion of ice to snow in the scheme is also much smaller than HUCM aggregation (Figs. 4.5c,d). This represents differing philosophies between each scheme's snow category, where HUCM snow is primarily aggregates, while NSSL snow is primarily freezing rain.

Unlike the other two schemes, snow water content in the Thompson scheme far exceeds ice crystal water content (Fig. 4.3c). However, number concentration of the two categories is similar to each other (Fig. 4.3f). The reason for snow mass dominance is detailed in Van Weverberg et al. (2013), where the Thompson scheme had a smaller ice crystal number concentration limit and cloud ice-snow conversion threshold compared to the Morrison scheme. Indeed, Thompson cloud ice converts to snow mass greater than in the other microphysics schemes

(Figs. 4.5e,g). As cloud ice number in the scheme has larger sources and smaller sinks than other schemes (Fig. 4.5f), the number concentration limit is likely hindering growth. Easily converted snow grows further by outcompeting ice crystals for depositional mass growth. The Thompson scheme does not show the rapid ice crystal and snow number growth rates at early model times present in the HUCM and NSSL schemes, likely because of the scheme's more stringent ice crystal nucleation requirements and snow 1M parameterization.

4.3.2 Vertical profiles

Vertical profiles of ice crystal and snow moments are taken at a mature supercell stage (t = 100 min.; Fig. 4.2) to show the moments' vertical distributions (Fig. 4.6). Moment values at each height level denote a horizontal average. As in the previous section, tendencies are computed at each height to elaborate on moment distribution. In the HUCM scheme, plate water content dominates overall ice crystal water content (Fig. 4.6a), due to the large amount of liquid mass freezing to plates (Fig. 4.7c). Plates also produce more water content than snow for z > ~9.5 km. Peak ice crystal (all modes) water content occur near z = 11.5 km. Ice crystal mass growth peaks below z = 10 km (Figs. 4.7a,c,e), implying that these particles are lofted up by the updraft while growing through vapor deposition. Snow water content peaks near z = 9 km, the result of crystal aggregation at this height (Fig. 4.7g). All three crystal modes' number concentrations peak near z = 12 km (Fig. 4.6d). While column number peaks near this height, other crystal nucleation and freezing liquid to plates peak at lower levels (Figs. 4.7b,d,f), again indicating that crystal number is primarily vertically advected to this height level Snow number concentration peaks below ice crystal near z = 9.5 km (due to ice crystal aggregation) and exceeds ice crystal number below z = ~4 km (marking the effects of crystal melting), and continues to melt snow particles at the lowest model level.

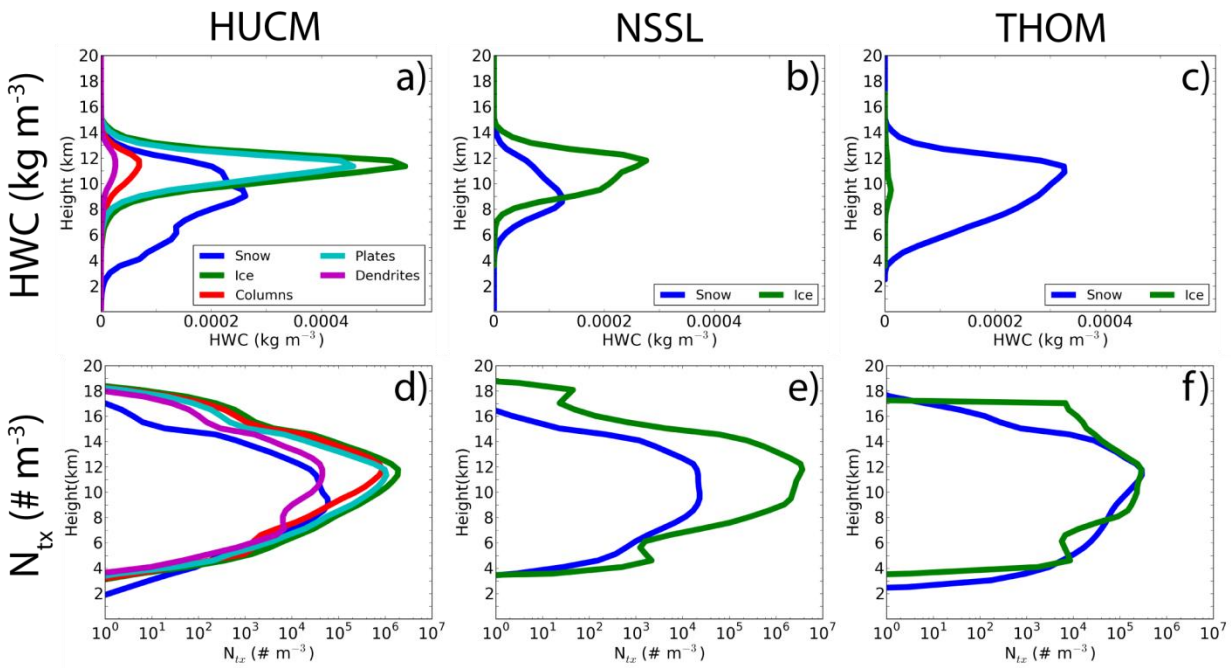

Fig. 4.6. Vertical profiles of horizontally-averaged hydrometeor water content (kg m^{-3}) and total number concentration N_{tx} (# m^{-3}) for the (a,d) HUCM, (b,e), NSSL, and (c,f) Thompson microphysics schemes at t = 100 min.

Like the HUCM scheme, both ice water content and number in the NSSL scheme peaks near z = 12 km (Figs. 4.6b,e). Snow water content peaks near z = 8.5 km while number maintains peak values between z = 12 and 9.5 km. Since ice source processes peak near z = ~9-10 km (freezing liquid; Figs. 4.8a-d), ice grows by vapor deposition while being transported upward. The large snow number peak depth is located directly above the maximum number sources (freezing rain and cloud ice to snow conversion), and implies that snow number is transported upward. While there are no significant snow number sources above 11 km in the HUCM (Fig. 4.7h), cloud ice to snow conversion occurs up to 15 km in the NSSL scheme (Fig. 4.9d). Both NSSL peak snow and ice water contents are smaller than those of HUCM. This can be attributed to the HUCM's freezing of liquid to plates and subsequent large aggregation to snow rate, whereas the NSSL's freezing liquid is spread across three hydrometeors (cloud ice, snow, and graupel). In the

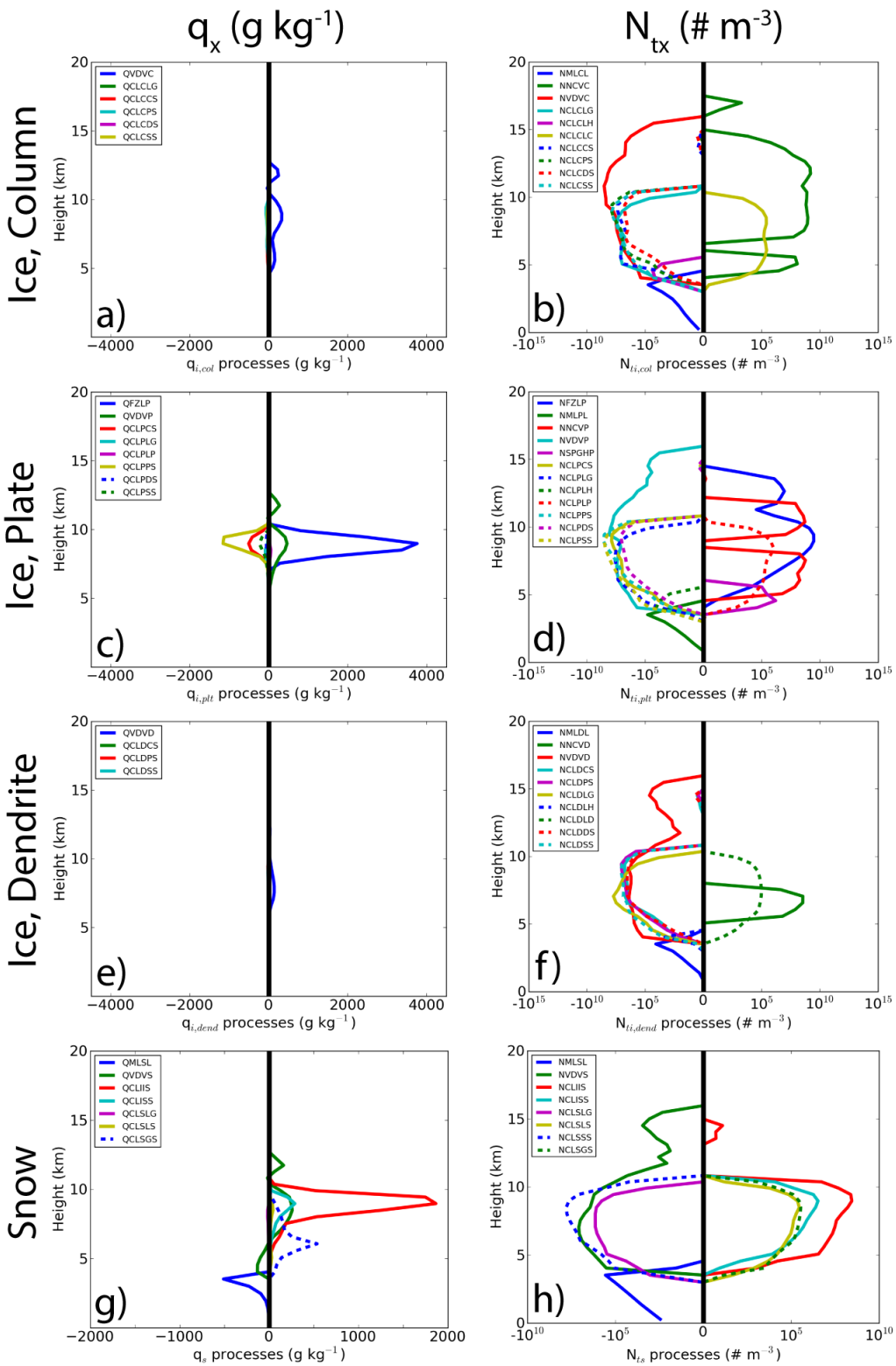

Fig. 4.7. Vertical mixing ratio q_x (g kg^{-1}; left column) and number concentration N_{tx} (# m^{-3}; right column) microphysical tendency sums for HUCM (a,b) columns, (c,d) plates, (e,f) dendrites, and (g,h) snow at t = 100 min. The vertical black line in each plot denotes the zero line separating source and sink terms.

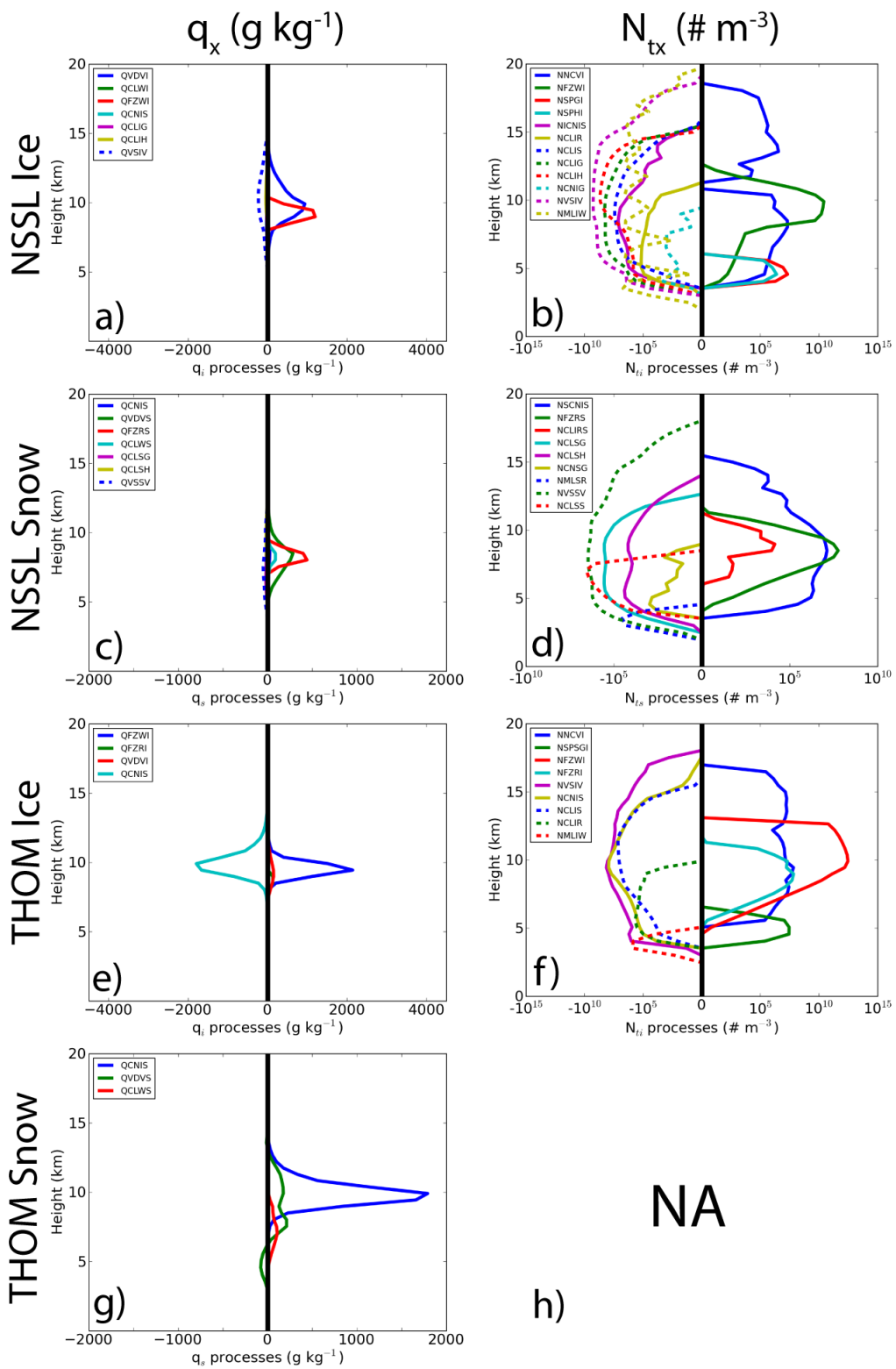

Fig. 4.8. As in Fig. 4.7, for NSSL (a,b) ice crystals and (c,d) snow, and Thompson (e,f) ice crystals and (g,h) snow.

Thompson scheme, the snow category contains both pristine ice (typically assigned to ice crystals in other BMPs) and snow aggregates by design of the snow PSD and due to the aggressive ice crystal-snow conversion in the model (Figs. 4.8e,g). Therefore, the snow field appears to be dominated by pristine ice, as snow water content peaks closer to ice crystals in the other schemes (z = ~11 km; Fig. 4.6c). Ice water content peaks near z = 9.5 km, which is below the snow peak. Easily converted snow outcompetes ice crystals during depositional growth in the Thompson scheme (Van Weverberg et al. 2013), which makes it difficult for ice crystals to reach similar masses as snow. Snow and ice number peak near z = 12 km and their values are similar (Fig. 4.6f). Ice number growth peaks below this height, implying ice crystals lofted above similar to the other two schemes. Snow number is not predicted in the Thompson scheme and, therefore, is diagnosed from the PSD. Very small ice mass combined with large number suggests that the ice category in the Thompson scheme represents only very small particles.

4.3.3 Selected particle size distributions

Based on the heights of moment maxima in the previous section, storm structure at z = ~9.5 km and z = ~12 km is analyzed through reflectivity plots (Fig. 4.9). Included in the plots are circles denoting the location of subsequent PSD comparison. Similar to the lowest model level and vertical cross-section of Z_H (Fig. 4.2), the HUCM simulates the weakest reflectivity core at both z = ~9.5 km and ~12 km (Figs. 4.9a,b). At these model heights, rimed ice in the supercell controls the reflectivity structure. Therefore, as the HUCM has stringent maximum rimed ice sizes, it is unable to simulate the larger reflectivity cores found in the NSSL and Thompson schemes (Figs. 4.9c-f). Across all three schemes, the maximum reflectivity is found near the updraft and decreases away, which is physical given its mass origin in a supercell framework and subsequent advection into the forward flank.

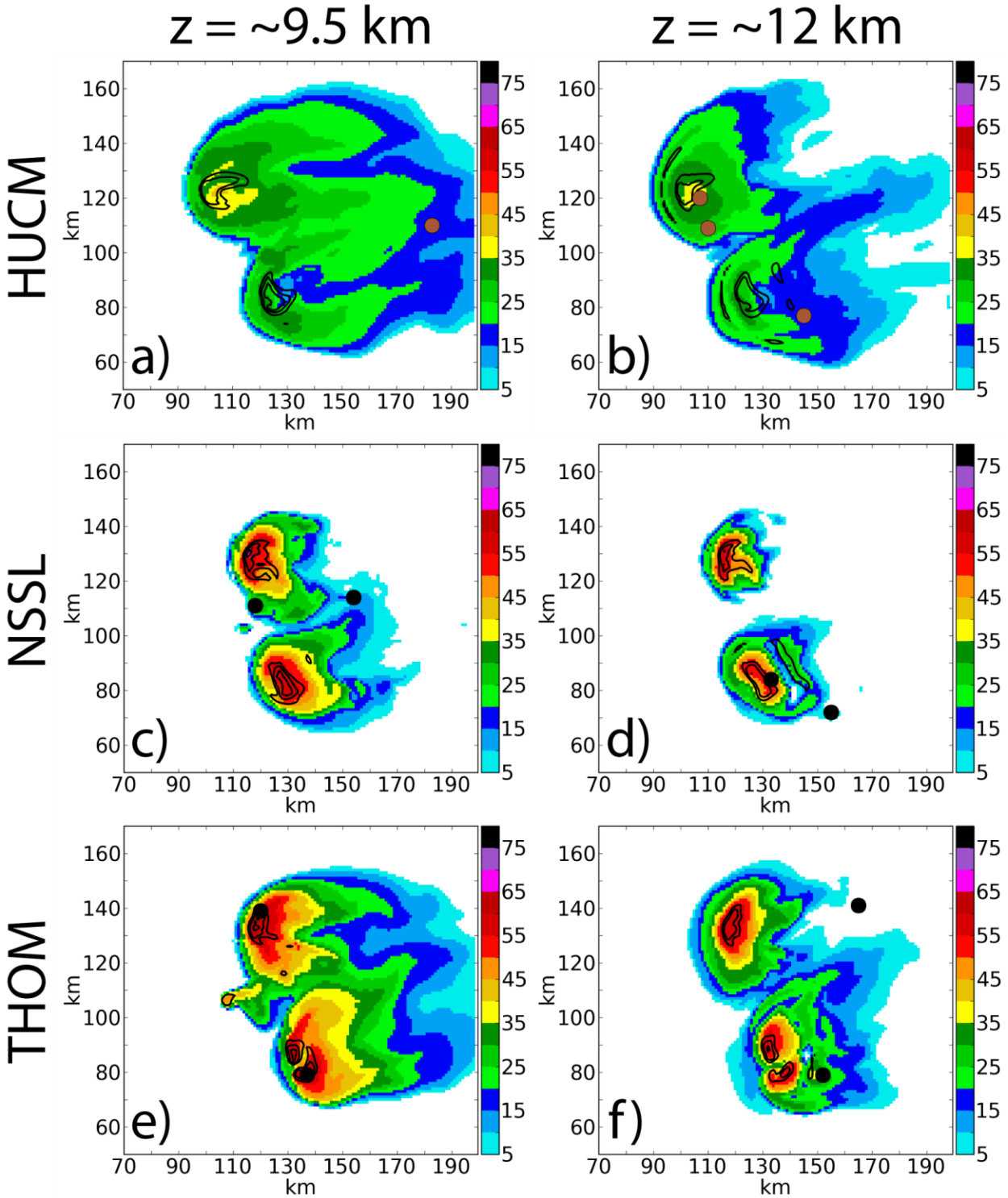

Fig. 4.9. Horizontal reflectivity Z_H (dBZ) for the (a,b) HUCM, (c,d) NSSL, and (e,f) Thompson schemes at z = ~9.5 km (left column) and z = ~12 km (right column). Circles in the plots denote locations of PSD samples, while contours are vertical velocity in 15 m s^{-1} intervals starting at 10 m s^{-1}.

Particle size distributions (PSDs) sampled from locations in Figure 3.9 are examined to directly compare SBM predicted quantities with bulk PSD assumptions. HUCM PSDs are plotted in log-space since the bins are mass-doubling (Fig. 4.10). As plate and snow mass dominates near z = 12 km, PSD sampling locations are strategically chosen such that column (Figs. 4.10a-d) and dendrite mass (Figs. 4.10i-l) are similar to plate mass, and plate (Figs. 4.10e-h) mass dominates at this height. Near z = 9.5 km, the PSD sampling location is chosen where snow mass dominates over ice crystals (Figs. 4.10m-p). These heights are selected based on ice crystal and snow moment maxima across the three microphysics schemes discussed in the previous section. Even with independent sampling locations, several similarities pervade through the HUCM PSDs. Column number density is typically largest for smaller diameters ($3\times10^{-3}-10^{-1}$ mm; Figs. 4.10a,e,i), plates peak at similar or larger size than columns (Figs. 4.10b,f,j,n), and dendrites exhibit narrower PSDs with smaller maximum diameter but typically peak at larger sizes compared to columns and plates (Figs. 4.10c,g,k,o). Snow number density is largest between 10^{-1} and 10^{0} mm (Figs. 4.10d,h,l,p). This is reasonable as smaller ice crystals aggregate to snow particles with larger mass. Column and plate number densities also tend to peak higher (10^{7}–10^{8} m^{-3} mm^{-1}) than dendrite and snow PSDs (10^{5}–10^{7} m^{-3} mm^{-1}) due to the smaller temperature range in which these particles are produced for dendrites (Table 4.2) and ice nucleation/freezing rain (i.e., column/plate number sources) exceeding snow aggregation in terms of particle production (Fig. 4.4). Still, plate number densities peak at similar or larger diameters than columns, so while these ice crystal categories contain similar number concentration over the model run, the larger plate particles contribute to its larger mass compared to columns. Column and plate PSDs show bimodal distributions at some locations, while dendrite and snow PSDs additionally show gamma distributions. Ice crystal and snow PSDs exhibit the inherent greater flexibility of PSD representation in SBMs

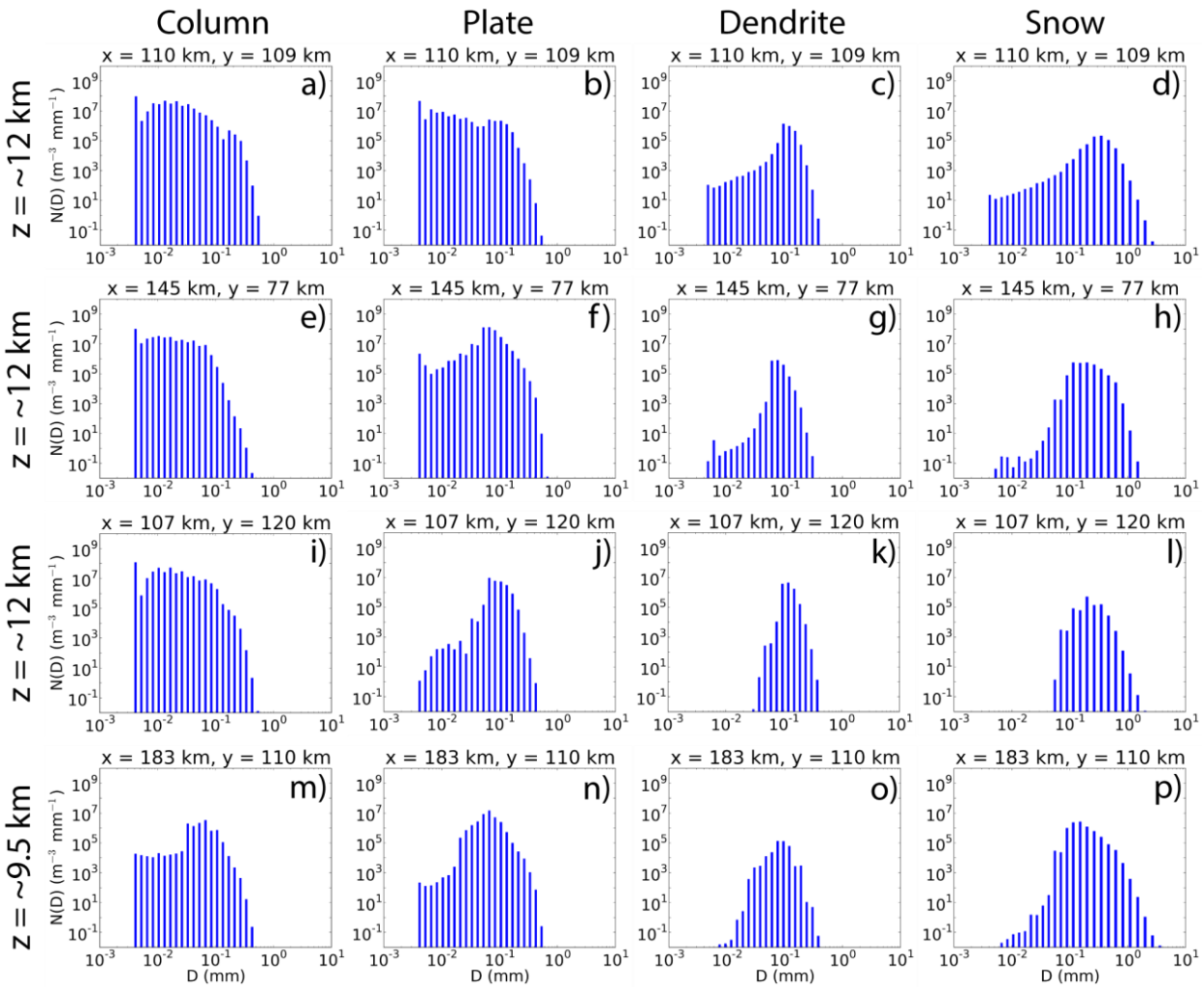

Fig. 4.10. Selected particle size distributions at t = 100 min. of (first column) columns, (second column) plates, (third column) dendrites, and (fourth column) snow particles at (a-l) z = ~12 km and (m-p) z = ~9.5 km in the HUCM scheme. Equivolume diameter unit is mm while the distribution function unit is m^{-3} mm^{-1}.

rather than parameterized bulk PSDs.

Ice crystal and snow PSDs for the NSSL and Thompson scheme are also sampled at z = ~9.5 km and z = ~12 km where either ice crystal or snow mass dominates, or both masses are comparable (Fig. 4.11). The PSDs are presented logarithmically for easier comparison with HUCM PSDs. The NSSL scheme bulk size distributions (generalized gamma distribution with four PSD parameters; readers are referred to Seifert and Beheng (2006) for details) illustrate the

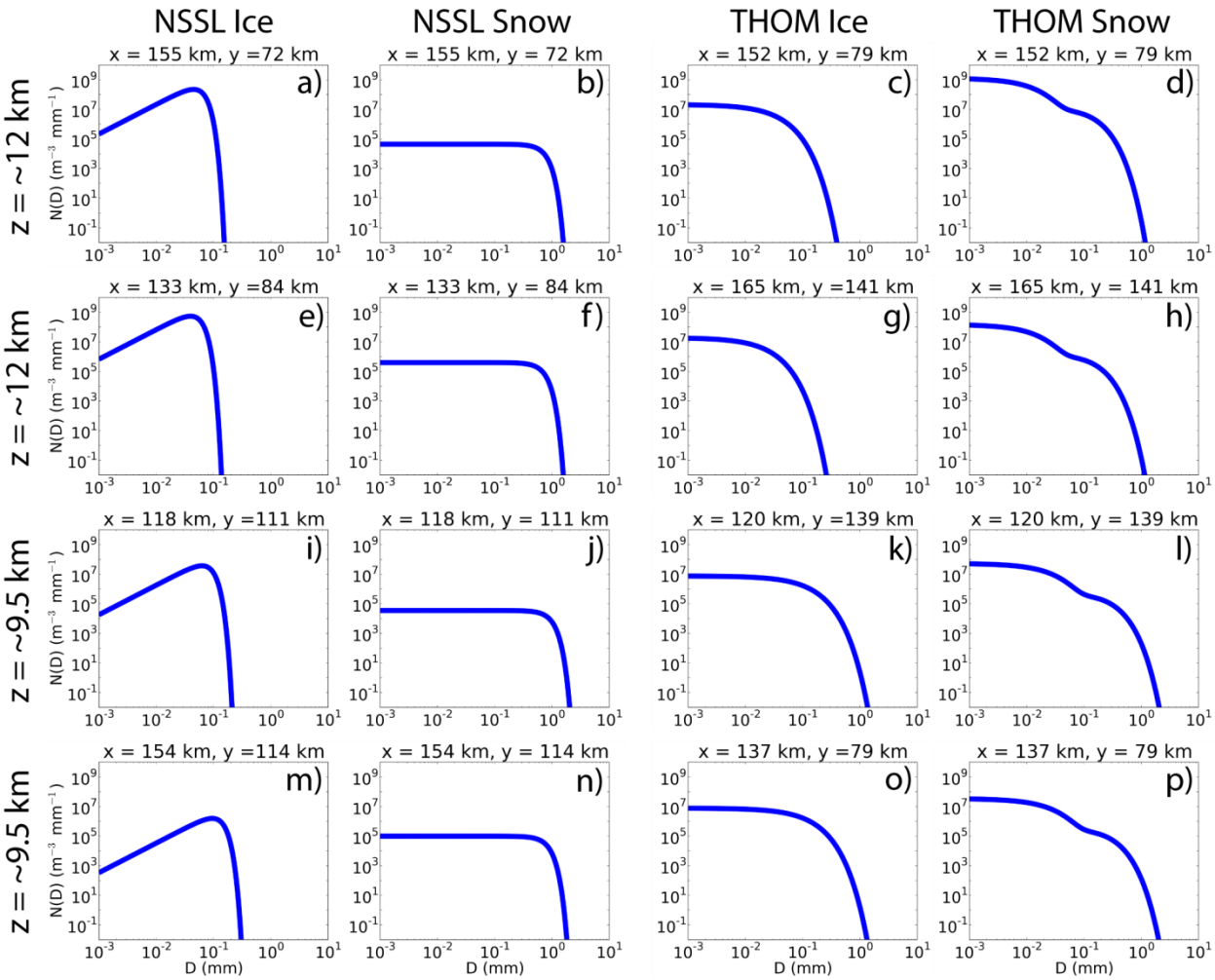

Fig. 4.11. Selected particle size distributions at t = 100 min. of (first column) ice crystal and (second column) snow particles in the NSSL scheme, and (third column) ice crystal and (fourth column) snow particles in the Thompson scheme at (a-h) z = ~12 km and (i-p) z = ~9.5 km. Equivolume diameter unit is mm while the distribution function unit is m^{-3} mm^{-1}.

single functional form of the PSDs across the samples. Ice crystal PSDs (Figs. 4.11a,e,i,m) increase until D = ~10^{-1} mm, then sharply decrease. Peaks (10^{6}–10^{9} m^{-3} mm^{-1}) of the distributions are similar to (if not slightly larger than) HUCM ice crystal peak magnitudes. Snow number densities are approximately constant near 10^{5} m^{3} mm^{-1} until rapidly decreasing near D = ~10^{0} mm (Figs. 4.11b,f,j,n). Unlike the HUCM, snow number densities in the NSSL scheme are relatively larger for small diameters, a consequence of the rigid bulk PSD design. Consistent with previous Thompson BMP discussion, the snow number density (Figs. 4.11d,h,l,p) is typically larger at all

sizes than the ice crystal number density at the same location (Figs. 4.11c,g,k,o), even for small particle diameters due to the strong cloud ice to snow conversion in the scheme. Ice crystal PSDs are largest at the smallest particle diameter because of their assumed exponential distribution. Compared to NSSL, Thompson cloud ice number density peaks are typically smaller, near 10^7 m^{-3} mm^{-1}. Snow is a bimodal distribution, with a PSD reduction near D = ~10^{-1} mm and a secondary peak between 10^{-1} and 10^0 mm. This is the same range that the snow PSD in the HUCM scheme peaks, which contains less number density than columns and plates. Thus, the Thompson snow PSD seems to be a good representation of both ice crystal and snow modes considering HUCM PSDs. Between the bulk schemes, the Thompson schemes favors a larger number of small ice crystal and snow particles due to its PSD assumptions.

4.4 Ice crystal and snow frequency

In order to evaluate macro PSD characteristics of ice crystal and snow in the simulated supercells, contoured frequency by altitude diagrams (CFADs; Yuter and Houze 1995) are created for the HUCM (Fig. 4.12), and NSSL and Thompson schemes (Fig. 4.13) for mass-weighted mean diameter D_m, which uses equivolume diameter. Briefly, the CFAD represents the number of grid points in each size bin, normalized by the total number of grid points that contain a nonzero mass. The bins are spaced logarithmically to capture small ice crystal frequencies. There are two $D_{mi,col}$ maxima in the HUCM, one between 0.08 and 0.09 mm at z = ~12.5 km, and one between 0.1 and 0.2 mm near 10 km (Fig. 4.12a). These column CFAD peaks are higher than the two column mass growth peaks in Fig. 4.7a, suggesting that column particles grow while being carried aloft. Plates and dendrites contain a D_m maximum near z = ~12 km between 0.1 and 0.2 mm (Figs. 4.12b,c). This agrees with the previously-analyzed PSDs, where columns typically peaked at smaller diameter than plates and dendrites at z = 12 km (Fig. 4.10). Otherwise, crystal D_m is similar

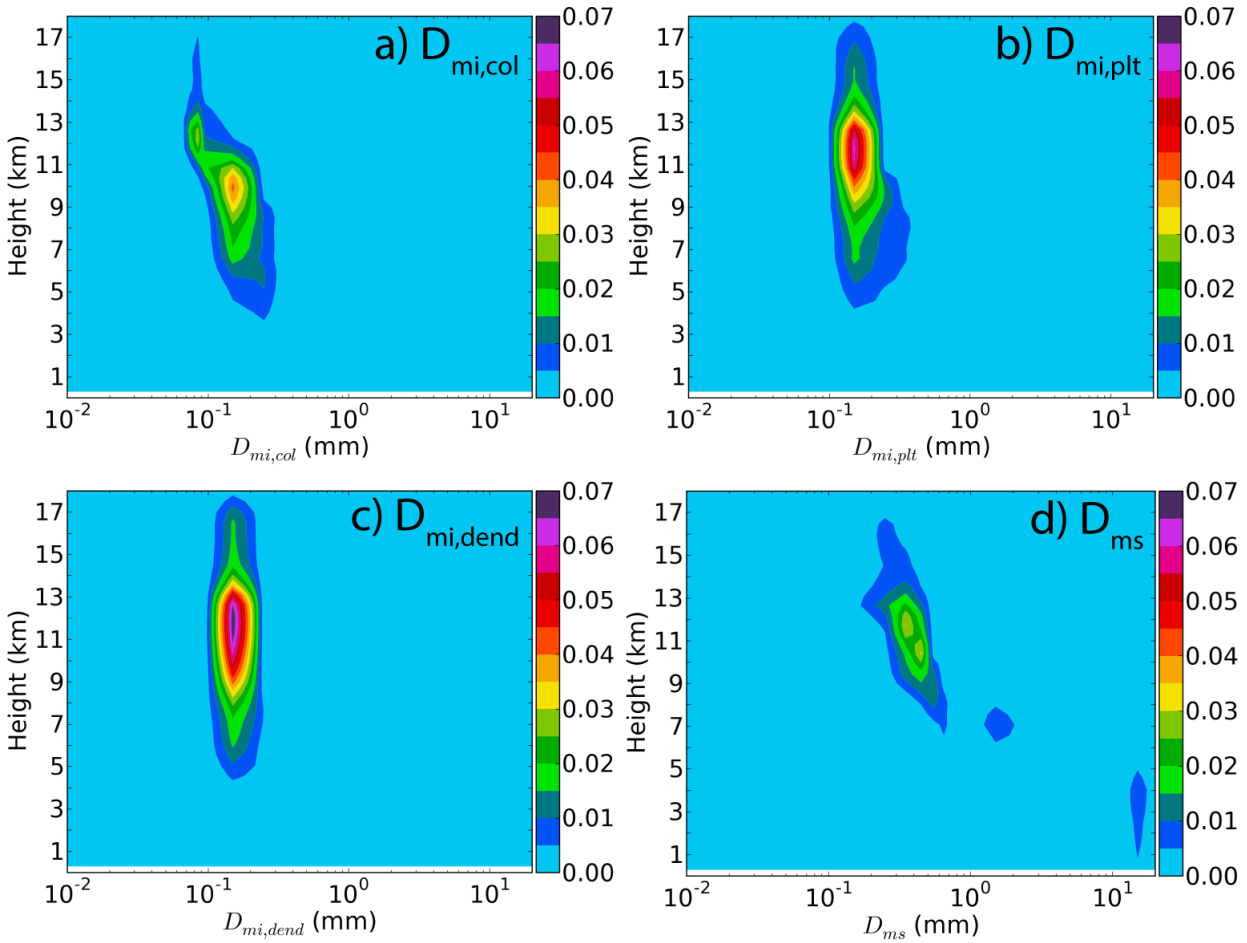

Fig. 4.12. Contoured frequency by altitude diagrams (CFADs) of (a) column, (b) plate, (c) dendrite, and (d) snow mass-weighted mean diameter D_m (mm) in the HUCM scheme at t = 100 min.

between the crystal modes, generally ranging between 0.09 and 0.3 mm. Columns and plates grow below z = ~12 km, mostly through deposition . Snow grows much larger than crystals, exceeding 10 mm near the surface (Fig. 4.12d). This is likely due to the combination of rapid snow particle growth through aggregation approaching the melting layer and the selectively faster melting of smaller snow particles, resulting in the shift of the snow PSD toward larger sizes. This large snow near the surface where the environmental temperature is ~ 25° C seems unreasonable considering

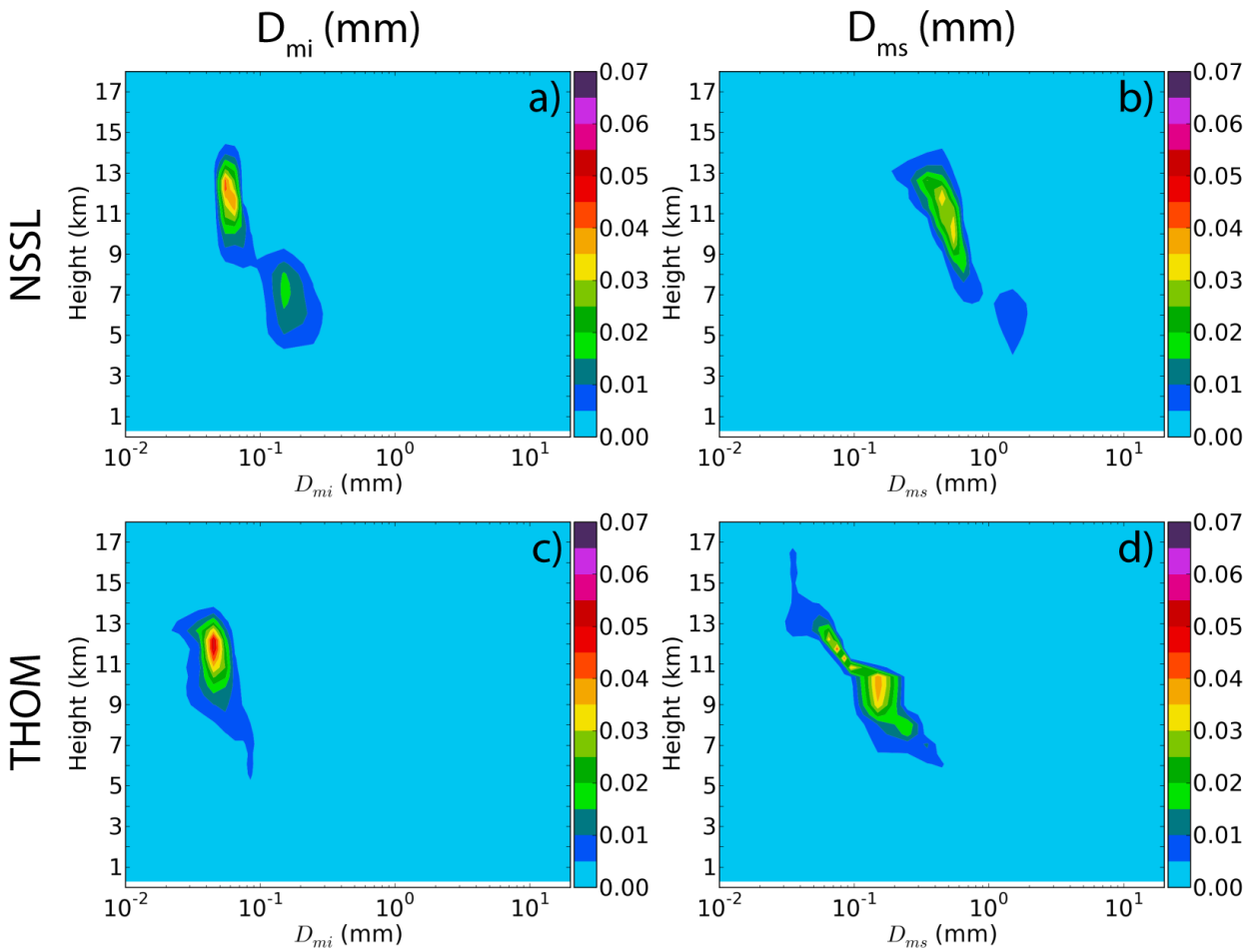

Fig. 4.13. Contoured frequency by altitude diagrams (CFADs) of ice crystal and snow mass-weighted mean diameter D_m (mm) in the (a,b) NSSL and (c,d) Thompson schemes at t = 100 min.

the typical low observed density of large snow (e.g., Magono and Nakamura 1965; Fabry and Szyrmer 1999; Brandes et al. 2007), suggesting that the HUCM melting treatment needs improvement.

NSSL D_{mi} has two frequency peaks, near z = ~12 km between 0.05 and 0.07 mm and between 0.1 and 0.2 mm near z = ~7 km (Fig. 4.13a). The upper peak is likely due to small particles originating from nucleation and freezing cloud water, while the source of the lower frequency peak at larger diameters are likely either freezing cloud water or hail/graupel splintering to cloud ice. The general size of D_{mi} in the NSSL is generally smaller than in the HUCM, as the NSSL secondary larger peak between 0.1 and 0.2 mm has smaller frequency. Like the HUCM, snow in the NSSL

scheme is much larger than ice crystals (Fig. 4.13b) and is continuously growing while falling through depositional growth and self-collection, while sublimation removes smaller snow particles. Ice crystal size in the Thompson scheme is smaller than the other two schemes, with most frequencies contained below 0.1 mm (Fig. 4.13c). As previously discussed, the aggressive cloud ice-snow conversion in the Thompson scheme pushes larger crystals into the snow category. Thus, the Thompson scheme has the smallest snow particles (Fig. 4.13d), with peak frequencies near 0.08 mm and noticeable frequency at snow sizes comparable to smaller ice crystals in the other schemes. Snow frequency is also lower at larger mean diameters than the previous schemes, likely due to the inclusion of small converted pristine ice in the distribution. Snow in both bulk schemes appear to deplete well above the surface, suggesting snow sinks are more properly parameterized than in the HUCM.

4.5 Summary and discussion

Idealized supercell simulations were performed in this study for the HUCM spectral bin, and the NSSL and Thompson bulk microphysics schemes to analyze simulated ice crystal and snow behavior. Overall, plates dominate the ice crystal mass fields in the HUCM scheme, while ice crystal number is primarily composed of plates and columns. Freezing liquid to plates contributes most to plate mass, while nucleation temperature ranges favor plates and columns compared to dendrites. The HUCM and NSSL schemes produce more ice crystal mass and number than snow over the model run, while the Thompson scheme favors snow mass. Ice crystal mass peaks at higher vertical levels than snow in the HUCM and NSSL schemes, while snow mass peaks above ice crystals in the Thompson scheme. Ice crystal number peaks at higher levels than snow in the HUCM and NSSL schemes, while number peaks are at similar levels in the Thompson scheme. The Thompson snow behavior compared to its cloud ice category is due to the

parameterization of the scheme's snow category, both in its snow PSD configuration (linear combination of exponential and gamma distributions) and aggressive ice crystal-snow conversion in the scheme. Small pristine ice typically belonging to the ice crystal category in other microphysics schemes is contained in the snow category in this scheme, and explains its more similar behavior (e.g., higher snow mass peak above ice crystals, more snow mass produced than ice crystals) with HUCM and NSSL ice crystal categories.

In the sampled PSDs, the HUCM column number densities peak at a similar or smaller diameter than plates, while dendrite and snow number densities peak at increasingly larger diameters. Therefore, while columns and plates have similar number concentration over the model run, larger plate particles contribute to its larger mass compared to columns. Ice crystal and snow PSDs exhibit both bimodal and gamma distributions, though PSD mode variability does exist across the PSDs sampled. Compared to the flexible SBM PSDs, the comparatively more rigid NSSL and Thompson bulk PSDs produce more small snow particles than the HUCM scheme. This is a fairly important drawback for both BMPs' snow category, which is driven by the conversion of growing ice crystals (though in the convective framework of this study, freezing rain contributed most to NSSL snow mass). Therefore, it is undesired to construct the snow PSD (through the shape parameter) such that the number density maximum is at or near equivolume diameter $D = 0$ mm, a constraint that the HUCM does not follow. Thompson snow PSDs share qualitative PSD peak location similarities with both HUCM ice crystal and snow PSDs. Quantitatively, HUCM column mass-weighted mean diameter D_m contains a smaller peak than other crystal modes, in agreement with HUCM PSD behavior. Otherwise, D_m is similar among the HUCM crystals. Ice crystals are generally smaller in both bulk schemes than in the HUCM, which could be linked to each scheme's conversion of larger ice crystals to snow. Snow grows large in the HUCM, due to aggregation and

weak melting that fails to deplete snow near the surface. On the other hand, snow in the bulk schemes do not grow to large sizes and deplete well above the surface. Snow in the Thompson scheme is smallest, likely owing to the cloud-ice snow conversion including traditional ice crystals in the category.

The above results reveal how ice crystal and snow parameterization differences in a spectral bin and two bulk schemes manifest in a supercell simulation. As the NSSL scheme qualitatively simulates many of the HUCM ice crystal and snow macro properties, and with the Thompson scheme skillfully forecasting snow compared to other BMPs (Liu et al. 2011), the increased complexity of HUCM ice may not qualitatively provide additional benefit over comparatively simple BMPs due to the much greater observational ice complexity. However, there are quantitative differences in ice and snow moments between the HUCM and BMPs examined in this paper (e.g., more HUCM ice crystal mass than in the NSSL and Thompson schemes); therefore, SBM and BMP simulations that can be verified with observations could provide insight on how increased HUCM ice complexity manifests in storm simulation accuracy. Several ice crystal modes may be necessary for the proper representation of simulated ice crystals to match the wide range of observed crystal type, especially in a microphysics framework where ice crystal deposition varies with ice crystal capacitance (e.g., Pruppacher and Klett 1997; Westbrook et al. 2008; Sulia and Harrington 2011). Though this study employs a deep convective framework, ice crystal mode and snow parameterization sensitivity with different storm systems (especially those where ice crystals and snow dominate hydrometeor mode) would also elaborate on accurate representation of these hydrometeors.

Chapter 5 Comparison of 2M bulk microphysics ensemble members during the 2018 NOAA HWT Spring Experiment

5.1 Introduction

Microphysics parameterization in numerical weather prediction (NWP) still struggles to match the increased observed ice complexity. Bulk microphysics schemes (BMPs) are typically preferred over spectral bin microphysics (SBM) in operational and experimental numerical weather prediction (NWP) models because they only predict moments of the particle size distribution (PSD) rather than the discretized PSD itself, making them computationally cheaper. Given the historic moment-evolution of BMPs, microphysics performance of real storm cases typically compares one-moment and two-moment schemes. The improvement of two-moment BMPs over their one-moment counterparts due to PSD flexibility in NWP (i.e., cold pool structure, differential sedimentation) at the convective scale is well-documented (e.g., Otkin and Greenwald 2008; Dawson et al. 2010; Jung et al. 2012). Specifically, several studies (e.g., Yussouf et al. 2013; Putnam et al. 2014; Wheatley et al. 2014; Putnam et al. 2017a; Wang and Wang 2017) have noted the benefits of improved mesoscale convective system (MCS) and supercell storm structure (i.e., reflectivity) with the use of 2M schemes.

Given the benefits of two-moment microphysics, operational NWP models have steadily increased in microphysics complexity. Operational models at the NOAA National Centers for Environmental Prediction (NCEP) include the Rapid Update Cycle (RUC; Benjamin et al. 2004a; Benjamin et al. 2004b), the Rapid Refresh (RAP; Benjamin et al. 2016b), and the High-Resolution Rapid Refresh (HRRR; Benjamin et al. 2016a). Before its replacement by the RAP model, the RUC employed Reisner et al. (1998) and Thompson et al. (2004) BMPs, which similarly are two-moment only for cloud ice. Earlier versions of the RAP model evolved to the Thompson et al.

(2008) BMP, which added two moment prognosis to the rain category. Currently, the RAP and HRRR both employ the Thompson and Eidhammer (2014) aerosol-aware BMP, which adds an additional two moment prognosis to cloud water. Similarly, the NSSL Experimental Warn-on-Forecast System for ensembles (NEWS-e) expanded microphysics complexity from the partially two-moment Thompson et al. (2008) BMP in 2016 to the fully two-moment NSSL (Mansell et al. 2010) BMP with an additional hail category in 2017 to more accurately simulate supercells (Skinner et al. 2018). While earlier Center for Analysis and Prediction of Storms (CAPS) Storm Scale Ensemble Forecasts (SSEFs) were primarily one- and partially two-moment microphysics (Kong et al. 2007; Xue et al. 2007; Xue et al. 2008), current SSEF configurations overwhelmingly use partially and full two-moment microphysics (Kong 2018).

As initial conditions and NWP physics parameterizations (i.e., land, radiation, microphysics) are not exact, an ensemble of forecasts that vary these parameters are important to construct a skillful probabilistic set of event outcomes (e.g., Tracton and Kalnay 1993; Fritsch and Carbone 2004). While these ensembles can be constructed using 3D or 4D variational data assimilation (3DVAR/4DVAR), an increasingly popular approach to constructing NWP ensembles employs the ensemble Kalman filter (EnKF; Evensen 1994), which does not require static background covariances or costly and complex adjoint models (Lorenc 2003). Specifically at the convective scale, EnKF has been employed to simulate both supercells (e.g., Aksoy et al. 2009, 2010; Dawson et al. 2012; Yussouf et al. 2015; Snook et al. 2016; Labriola et al. 2017) and mesoscale convective systems (MCSs) (e.g., Snook et al. 2011; Putnam et al. 2014; Wheatley et al. 2014; Snook et al. 2015; Putnam et al. 2017a), usually taking advantage of temporally- and spatially-dense radar observation assimilation. One particular advantage in the EnKF framework

is the ability to retrieve and correct microphysical parameters (Tong and Xue 2008a, b), which directly modifies microphysical processes that control storm evolution.

The NOAA Warn-on-Forecast (WoF; Stensrud et al. 2009; Stensrud et al. 2013) research project attempts to improve warning lead times of hazardous convective-scale events through short-term high-resolution observation assimilation and NWP, rather than relying on storm observations. Consistent with this vision, CAPS has collaborated with the Hazardous Weather Testbed (HWT) every spring since 2007 to provide guidance on optimal real-time ensemble configuration based on dynamics core, ensemble design, initial perturbation, physics configuration, microphysics parameterization, etc (e.g., Kong 2015; Kong 2018). Prior to 2016, CAPS designed the HWT SSEFs to reflect microphysical advances (i.e., complexity and moments) and prior sensitivity results (i.e., initial perturbation, boundary conditions). Starting in 2016, the HWT has provided guidance for ensemble design under the Community Leveraged Unified Ensemble (CLUE; Clark et al. 2018) to better facilitate physics comparison and bias/error sources. The SSEF 3DVAR outperformed several other ensembles in the 2015 Spring Experiment in terms of QPF, while the SSEF EnKF was not as successful as the 3DVAR (both objectively and subjectively; Gallo et al. 2017). As convective-scale ensembles continually migrate to EnKF, it is critical to determine and mitigate sources of ensemble biases and error to improve short-term forecasts of hazardous weather events.

The purpose of this chapter is to determine the sensitivity of 2018 Spring Experiment simulations to BMP choice and their performance through radar, precipitation, and brightness temperature forecast evaluation. Because BMPs are highly non-linear and lack observation validation of parameterization choice (e.g., melting rate, shape parameter), motivation exists to evaluate microphysical performance of convective-scale event simulation, which is driven by

microphysics rather than cumulus parameterization. This chapter can help guide forecasters on both optimal ensemble construction (BMP choice/frequency) and bias recognition based on operational ensemble configuration. Along with seasonal forecasts (i.e., over the 2018 NOAA HWT Spring Experiment), four convective line test cases are evaluated to contribute to WoF's emphasis of hazardous weather prediction. Both short-term forecasts (1-6 h) aligned with WoF goals and longer-term forecasts ("next-day") are considered to determine the accuracy of BMPs with forecast length. The rest of the chapter is organized as follows: section 5.2 details SSEF configuration and BMP parameterization, section 5.3 examines short-term microphysical skill, section 5.4 analyzes next-day model performance, and section 5.5 provides the summary and conclusions.

5.2 Simulation design

5.2.1 2018 CAPS EnKF SSEF

CAPS provided several SSEF configurations as part of the 2018 NOAA Hazardous Weather Testbed Spring Experiment. The SSEF suite selected for this chapter is a 10-member EnKF forecast with varied microphysics and planetary boundary layer physics initialized at 0000 UTC and run for 48 hours. Each member in the 10-member ensemble is initialized from a separate 40-member EnKF ensemble (initialized at 1800 UTC with hourly RAP/HRRR assimilation and radar assimilation every 15 minutes during the hour prior to ensemble launch) with Thompson (Thompson et al. 2008) microphysics. The EnKF SSEF is run over the CONUS domain (Fig. 5.1) using the Weather Research and Forecasting (WRF) model version 3.9.1.1 (Skamarock et al. 2008) at convective-scale spacing (Δx = 3 km) with 51 vertical levels. The 2018 Spring Experiment officially ran from April 30 to June 1 during weekdays. This chapter additionally includes April 27th (during the two-week CAPS SSEF maintenance period), as data is available and includes

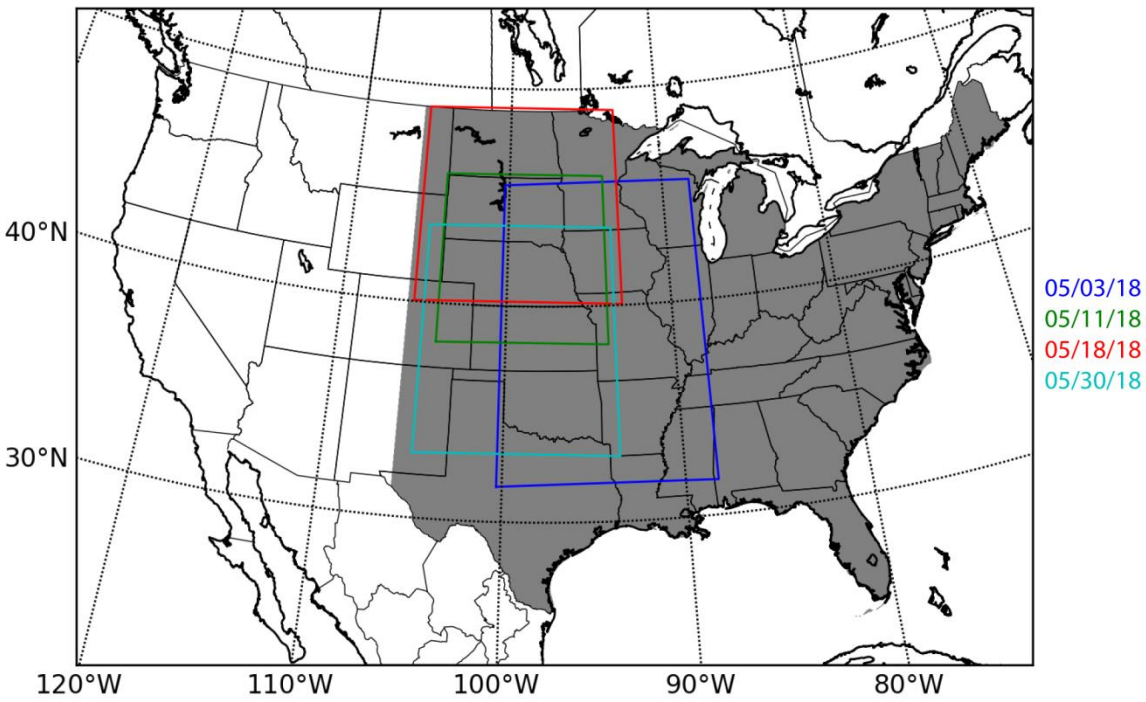

Fig. 5.1. The 2018 CAPS Spring Experiment computational domain. The seasonal analysis domain is shaded, while select test case analysis domains for short-term (1-h to 6-h) forecasts are represented by colored boxes.

observed storms on this day.

The 10-member EnKF SSEF examined in this chapter varies initial/boundary conditions, microphysics, and planetary boundary layer (PBL) schemes among the members. Each member employs the Noah land surface model (Chen and Dudhia 2001). As mentioned earlier, members 1-8 are initialized with 0000 UTC analyses from ensemble members in the 40-member EnKF Thompson microphysics SSEF (consistent with the 10-member SSEF PBL choice for each member), while members 9 and 10 are initialized with the 40-member EnKF Thompson

microphysics SSEF ensemble mean. Microphysics schemes used in the 10-member EnKF SSEF include the Morrison (Morrison et al. 2009), National Severe Storms Laboratory (NSSL; Mansell et al. 2010), Predicted Particle Properties (P3; Morrison and Milbrandt 2015; Milbrandt and Morrison 2016), and Thompson (Thompson et al. 2008) schemes. The PBL schemes include the Mellor–Yamada–Janjic´ (MYJ; Janjic´ 1994), Yonsei University (YSU; Hong et al. 2006), and the Mellor–Yamada–Nakanishi–Niino (MYNN; Nakanishi and Niino 2006). The specific physics configuration for each EnKF member is detailed in Table 5.1.

5.2.2 Microphysics schemes

As this chapter examines the impact of microphysics on model performance, members 1, 4, 6, 9, and 10 are selected as these members contain the same LSM and PBL schemes (although initial and boundary conditions also differ). These members use the Thompson, Morrison, and NSSL bulk microphysics schemes. The Thompson microphysics scheme is two-moment (i.e., predicting mass mixing ratio q and number concentration N_t) for rain and cloud ice, and additionally predicts the mass mixing ratio q of cloud water, snow, and graupel. The Morrison microphysics scheme is two-moment for cloud ice, snow, rain, and graupel, and additionally predicts the mass mixing ratio q of cloud water. Finally, the NSSL scheme is two-moment for cloud water, cloud ice, snow, graupel, hail and rain, and also prognoses the bulk volume of graupel and hail, allowing for the prediction of rimed ice bulk density. We also point out here that an updated version of the Thompson scheme that prognoses cloud water number and updates cloud ice number through nucleation using aerosols (Thompson and Eidhammer 2014) is currently used operationally in the RAP and HRRR models. Further, the graupel category in the Morrison scheme may be modified to be more "hail-like" (i.e., higher density and fall speed assumptions) through a WRF namelist parameter. The P3 BMP is not analyzed extensively here, and the reader is referred

Table 5.1. 2018 CAPS EnKF 10-member SSEF physics configuration. Adapted from Kong (2018) Table 4.

Member	Microphysics	Land Surface Model (LSM)	Planetary Boundary Layer (PBL)
M01	Thompson	Noah	MYJ
M02	NSSL	Noah	YSU
M03	NSSL	Noah	MYNN
M04	Morrison	Noah	MYJ
M05	P3	Noah	YSU
M06	NSSL	Noah	MYJ
M07	Morrison	Noah	YSU
M08	P3	Noah	MYNN
M09	Thompson	Noah	MYJ
M10	NSSL	Noah	MYJ

to Morrison and Milbrandt (2015) and Milbrandt and Morrison (2016) for more information on the scheme's design.

Although hydrometeor category nomenclature is similar across the Morrison, NSSL, and Thompson BMPs, specific hydrometeor design can vary considerably. The Morrison BMP assumes an exponential (i.e., a gamma distribution with shape parameter $\alpha = 0$) distribution for each hydrometeor except cloud water, where the shape parameter varies with cloud water number concentration (default value of 250×10^6 m^{-3}) following Martin et al. (1994). All hydrometeors are assumed spherical with liquid, cloud ice, snow, and graupel densities of 997, 500, 100, and 400 kg m^{-3} respectively. The NSSL BMP assumes a gamma distribution for each hydrometeor. Cloud

water, cloud ice, and snow utilize mass distributions with shape parameter α = 0, 0, and -0.8, respectively, while rain, graupel, and hail follow the standard diameter PSD with shape parameter α = -0.4, 0, and 1 respectively. Cloud ice, snow, graupel, and hail are assumed spherical, with cloud ice and snow bulk densities of 900 and 100 kg m-3 respectively, while graupel and hail have predicted bulk densities. The Thompson snow PSD is a linear combination of an exponential (shape parameter α = 0) and gamma distribution with the intercept N_0 and slope Λ parameters diagnosed from snow second and third moments following the observations of Field et al. (2005). As cloud water in the Thompson scheme is one-moment, the default cloud water number in the scheme is 100×10^6 m^{-3}. Cloud water assumes a gamma distribution, with the shape parameter α diagnosed as a function of its number concentration following Martin et al. (1994). The remaining hydrometeors assume an exponential distribution (shape parameter α = 0). As graupel is also one-moment, the intercept parameter N_0 in the scheme is diagnosed from rain median volume diameter and graupel mass mixing ratio. Rain, graupel, and cloud ice are assumed spherical with bulk densities of 1000, 500, and 890 kg m^{-3}, while the snow mass-diameter relationship is set as $m(D) = 0.069D^2$ following Cox (1988) to better match observations.

5.2.3 Model Evaluation Tools (MET) and verification datasets

Forecast evaluation statistics (e.g., contingency tables, neighborhood fractional coverage) are computed using the Model Evaluation Tools Version 6.0 (METv6.0). As model output and verification data are on the same 2018 Spring Experiment CONUS grid, the METv6.0's "grid_stat" is employed. The neighborhoods employed for short-term and next-day forecasts are squares with half-lengths (i.e., the distance from the middle grid point [on which the neighborhood is centered] to the outermost grid point) of 9 and 39 km, respectively. Neighborhood statistics are not computed for data points that do not have 100% "valid" data (i.e., outside the analysis domain/masking

regions) Contingency tables are built following the "neighborhood maximum" approach (e.g., Sobash et al. 2011). A hit is defined as both model and observation fields exceeding the given threshold anywhere inside the neighborhood, a false alarm is defined as the model field exceeding the given threshold anywhere inside the neighborhood (but not observations), and so on. Ensemble probabilities are calculated using the neighborhood maximum ensemble probability (NMEP) approach (Schwartz and Sobash 2017), where event prediction at each grid point is set to 1 if the event is predicted anywhere in the searching box (in this chapter, set to the same size as the neighborhood).

Ensemble and individual ensemble member performances are analyzed by comparing simulated model and observed composite reflectivity *Z*, 1-hour accumulated precipitation (1-h AP), and brightness temperature in the 11.2 μm channel (corresponding to the Geostationary Operational Environmental Satellites [GOES] channel 14). Model composite reflectivity is calculated using Rayleigh scattering, and is therefore sensitive to microphysics parameterization. Model 1-h AP is diagnosed from the hourly summation of hydrometeor mass sedimenting below the lowest model level at each grid point. Model brightness temperature is calculated from model hydrometeor fields by coupling the Community Radiative Transfer Model (CRTM) with the Advanced Regional Prediction System (ARPS; Xue et al. 2003) to simulate model brightness temperature in the 11.2 μm channel (e.g., Jung et al. 2013). Observed composite reflectivity and 1-h AP are provided by the NSSL's multi-radar/multi-sensor (MRMS) system (e.g., Smith et al. 2016; Zhang et al. 2016), while observed brightness temperatures are provided from the operational GOES-16. Observed *Z*, 1-h AP, and brightness temperature in the 11.2 μm channel are interpolated to the 2018 HWT Spring Experiment grid before any forecast evaluation.

5.3 Short-term forecasts

5.3.1 Ensemble performance

Consistent with Warn-on-Forecast goals of improving short-term forecasts, the 10-member EnKF SSEF is analyzed from 0100 UTC to 0600 UTC for each day (04/27/18 – 06/01/18) of the 2018 NOAA HWT Spring Experiment (model data is output hourly, while 0000 UTC is not included as this is the final assimilation time). The neighborhood maximum ensemble probabilities (NMEP; see section 5.2.3) from the 10 members (or on some dates with missing data, 7 or 9 members) are smoothed using a gaussian filter with standard deviations σ of 39, 60, 90 and 120 km to help determine ideal smoothing lengths for probabilistic forecasts. From these grid-point probabilities, attributes (Hsu and Murphy 1986), sharpness (Murphy 1993), and relative operating characteristic (ROC) curve (Mason 1982) diagrams are constructed with probability bins p = 0, 0.1,...,1 (Fig. 5.2). Ensemble ability to simulate storm structure is analyzed using composite reflectivity $Z \geq 15$ dBZ for overall (stratiform and convective) storm coverage and $Z \geq 40$ dBZ for intense, convective cores following Putnam et al. (2017b). Similarly, ensemble light and heavy precipitation forecast skill is determined using 1-hour accumulated precipitation (1-h AP) thresholds ≥ 0.01" and ≥ 0.5", which roughly (but not exactly) correspond with the prior two Z thresholds in $Z = 300R^{1.4}$, the WSR-88D operational summer convection Z-R relationship.

Attribute diagrams are conditioned on the ensemble forecasts (i.e., what are the observed frequencies when an event is forecast?), and compare forecast probability with observed frequency. The sharpness (i.e., frequency of forecasts in each probability bin) of these forecasts is measured with an accompanying frequency plot in Fig. 5.2. Large-scale storm structure (defined by $Z \geq 15$ dBZ) is forecast with high reliability as forecast probabilities display little deviation from their corresponding observed frequencies (Fig. 5.2a). Each smoothing standard deviation σ contains an underprediction bias in medium probabilites (0.3-0.9) that worsens as σ increases.

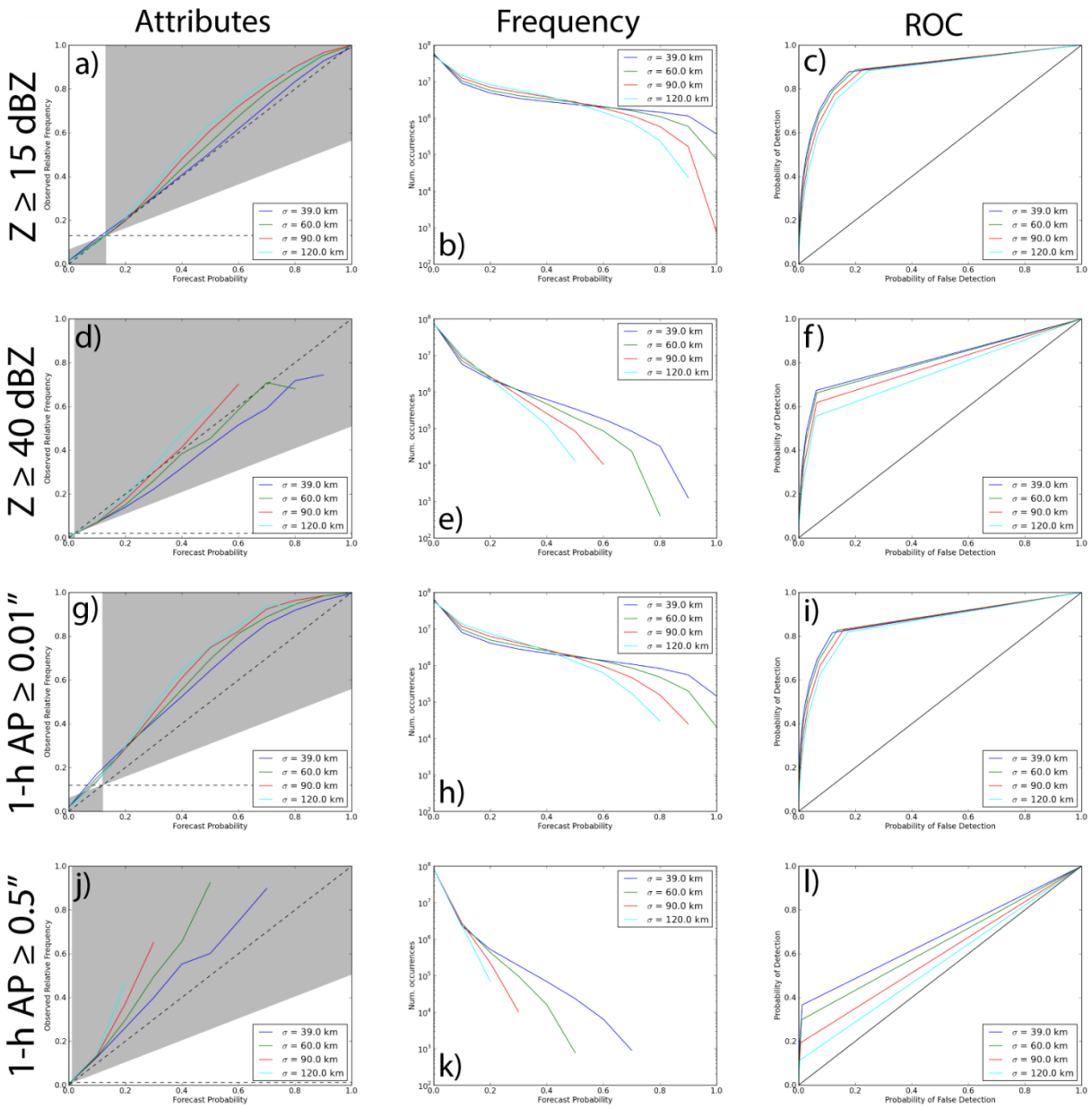

Fig. 5.2. Attributes, frequency, and relative operating characteristic (ROC) curve diagrams for the 10-member EnKF short-term (0100-0600 UTC) forecasts over the 2018 HWT Spring Experiment for composite reflectivity Z (a,b,c) ≥ 15 dBZ, (d,e,f) ≥ 40 dBZ, and 1-hour accumulated precipitation (g,h,i) ≥ 0.01" and (j,k,l) ≥ 0.5". The shaded region in the attributes diagrams indicate forecast skill, similar to forecasts above the black line in ROC curves.

Large-scale (and mesoscale) storm structure does not require the smoothing necessary for isolated, convective cores, and therefore lose reliability as their (especially large) forecast probabilities are

subject to more smoothing (Fig. 5.2b). For convective storm structure ($Z \geq 40$ dBZ), the ideal σ in terms of reliability decreases as forecast probability increases, as increasing σ shifts convective storm structure overprediction at high probabilities to small probability forecasts (Fig. 5.2d). Ensemble predictions with σ = 39 and 60 km contain a persistent overprediction bias, while σ = 90 and 120 km contain both underprediction and overprediction biases. This reveals a microphysics bias among the Morrison, NSSL, Thompson, and P3 BMPs, as the ensemble aggregate is simulating too many large rain and/or rimed ice hydrometeors (those expected to contribute most to high reflectivity). The impact of gaussian filter standard deviation is also apparent in the sharpness diagram (Fig. 5.2e), where σ = 90 km and 120 km fail to produce forecast probabilities above 0.6. As ensembles with large smoothing reduce high probability forecasts, underprediction at high probabilities is expected. Still, ensemble reliability reveals high skill for large-scale and convective storm structure regardless of σ, as most reliability curves are contained within the shaded skill area (between climatology and the perfect diagonal reliability line).

A good counterpart to help explain model performance is the ROC curve, which is conditioned on observations (i.e., how is the model simulating observed and non-observed events?). For $Z \geq 15$ dBZ, the ROC curve (Fig. 5.2c) shows high skill with probability of detection (POD) generally exceeding high probability of false detection (POFD). Therefore, the ensemble is generally able to distinguish between observed and non-observed large-scale storm structure. POD typically decreases and POFD tends to increase as σ increases for small probabilities, while POFD decreases for large probabilities. As σ increases, larger NMEPs are reduced, which correspond to smaller PODs as increased misses at larger forecast probabilities are shifted to smaller probabilities. Further, larger σ smooths NMEPs to greater areal extents than smaller σ, which increases false alarms (increasing POFD) for small probability forecasts. As increasing σ also

reduces the number of high probability forecasts, they are less prone to false alarms. Convective storm structure ($Z \geq 40$ dBZ) is not as skillful, as POD decreases compared to large-scale storm structure. This is expected, as these convective cores are more spatially-isolated and therefore are more prone to temporal and location errors. POD and POFD behavior follows that of large-scale storm structure, i.e., POD decreases and POFD increases as σ increases for small probability forecasts. Therefore, although increasing σ improves the reliability of convective structure forecasts, oversmoothing the forecast storm structure fields (both magnitude and spatially) reduces the model's ability to distinguish between events and non-events.

Light precipitation (1-h AP $\geq$ 0.01") has good reliability, but also suffers from a persistent underprediction bias (Fig. 5.2g) that is larger than large-scale storm structure underprediction bias. While the precipitation threshold used corresponds with the reflectivity threshold for large-scale storm structure, composite reflectivity Z is a column-max value, while accumulated precipitation is at the surface. Therefore, as the ensemble members are underpredicting maximum reflectivity in each grid column, they are likely underpredicting reflectivity over the majority of each grid column, which compounds mass underprediction (and therefore mass flux) over the column. This bias is worst for σ = 39 km for low forecast probability, while σ = 120 km has the highest underprediction bias for medium-high forecast probabilities. Large gaussian smoothing increases the amount of small forecast probabilities (Fig. 5.2h) and their areal extent, allowing these forecasts to capture more observed precipitation events. This smoothing also decreases the number of large forecast probability forecasts, which increases light precipitation underprediction bias as forecast probability increases. Heavy precipitation (1-h AP $\geq$ 0.5") is much less reliable given its low climatology (Fig. 5.2j). Therefore, the model tends to significantly underpredict the occurrence of heavy precipitaiton (and as σ increases, stops predicting larger forecast probabilities;

Fig. 5.2k). Even though convective storm structure is overpredicted for σ = 39 and 60 km, composite reflectivity is still indicative of an isolated large mass core. The model is still underpredicting hydrometeor mass over the entire column, since surface heavy precipitation is underpredicted. Given the infrequency of heavy precipitation prediction, the ROC diagrams for all σ are marginal and not much higher than the "no-skill" line, as less forecasts result in small false alarms but also low POD. In contrast, light precipitation shows much higher skill in its ROC curve (Fig. 5.2i), with similar skill decrease as σ decreases found in the other ROC curves. Still, 1-h AP shows good reliability for the forecast probabilities it predicts (i.e., inside skill area) while simulating observed light precipitation events and non-events with skill.

5.3.2 Microphysical performance

BMP performance of the Morrison, NSSL, and Thompson schemes are examined by analyzing EnKF members 1, 4, 6, 9, and 10 (Table 5.1), which have the same land surface (Noah land surface model; Chen and Dudhia 2001) and PBL (Mellor–Yamada–Janjic´ [MYJ]; Janjic´ 1994) physics. Members 9 (Thompson) and 10 (NSSL) have the same initial/boundary conditions separate from members 1 (Thompson), 4 (Morrison), and 6 (NSSL), which each have different initial/boundary conditions. The same large-scale ($Z \geq 15$ dBZ; 1-h AP ≥ 0.05") and convective ($Z \geq 40$ dBZ; 1-h AP ≥ 0.5") storm structure and precipitation thresholds used in section 5.3.1 are applied here. "Seasonal" forecasts include daily forecasts during weekdays over the entire 2018 NOAA HWT Spring Experiment (04/27 – 06/01). Four convective line test cases (05/03, 05/11, 05/18, 05/30) are further examined to help determine BMP performance in deep-convective storm modes. Like the seasonal forecasts, the performance of these test cases is limited to 01-06 h to assess short-term predictability. Each test case has a limited verification domain compared to the

seasonal eastern CONUS domain (Fig. 5.1), chosen to mitigate verification scores from separate storm systems during the analysis time.

The 05/03 test case contains two convective lines from two different synoptic/mesoscale structures. A stationary front situated over the Oklahoma panhandle forces convection between 1600 and 1700 UTC on 05/02 (Fig. 5.3a), and shows clear squall line structure over central Kansas by 2100 UTC. Below the stationary front is a north-south oriented dry line extending into west Texas. Convection along the dry line initiates between 1800 – 1900 UTC on 05/02, and spawns both supercells and other discrete storm cells near southwestern Oklahoma and the Texas panhandle. By the analysis time (05/03 0100 UTC), the discrete cellular structure initialized along the dry line merges into a single convective line over central Oklahoma and extends into northern Texas, eventually merging with the squall line by 0500 UTC. Between 1900 UTC and 2000 UTC on 05/10 (i.e., 05/11 test case), a dry line is attached to a stationary front and low pressure system over southeastern Wyoming (Fig. 5.3b). Convection also begins during this time frame, with strong, discrete storm structure by 2300 UTC in western Nebraska. By the analysis time, these discrete storms have merged into a convective line that slowly transverses Nebraska over the analysis period.

Convection initiates in western South Dakota between 1800 – 1900 UTC on 05/17 (i.e., 05/18 test case) along a N-S oriented dry line (Fig. 5.3c). Precipitation is present in North Dakota for much of the 05/17 day as a stationary front is present in the state. By 05/18 0000 UTC, the storms organize into a prominent mesoscale convective system (MCS) in western South Dakota and Nebraska, and southern North Dakota. At 0600 UTC, the convective line of the MCS is primarily in Nebraska, while the stratiform region encompasses northwestern Nebraska and most of South Dakota. Precipitation is present over much of Kansas and Nebraska on 05/29 from the

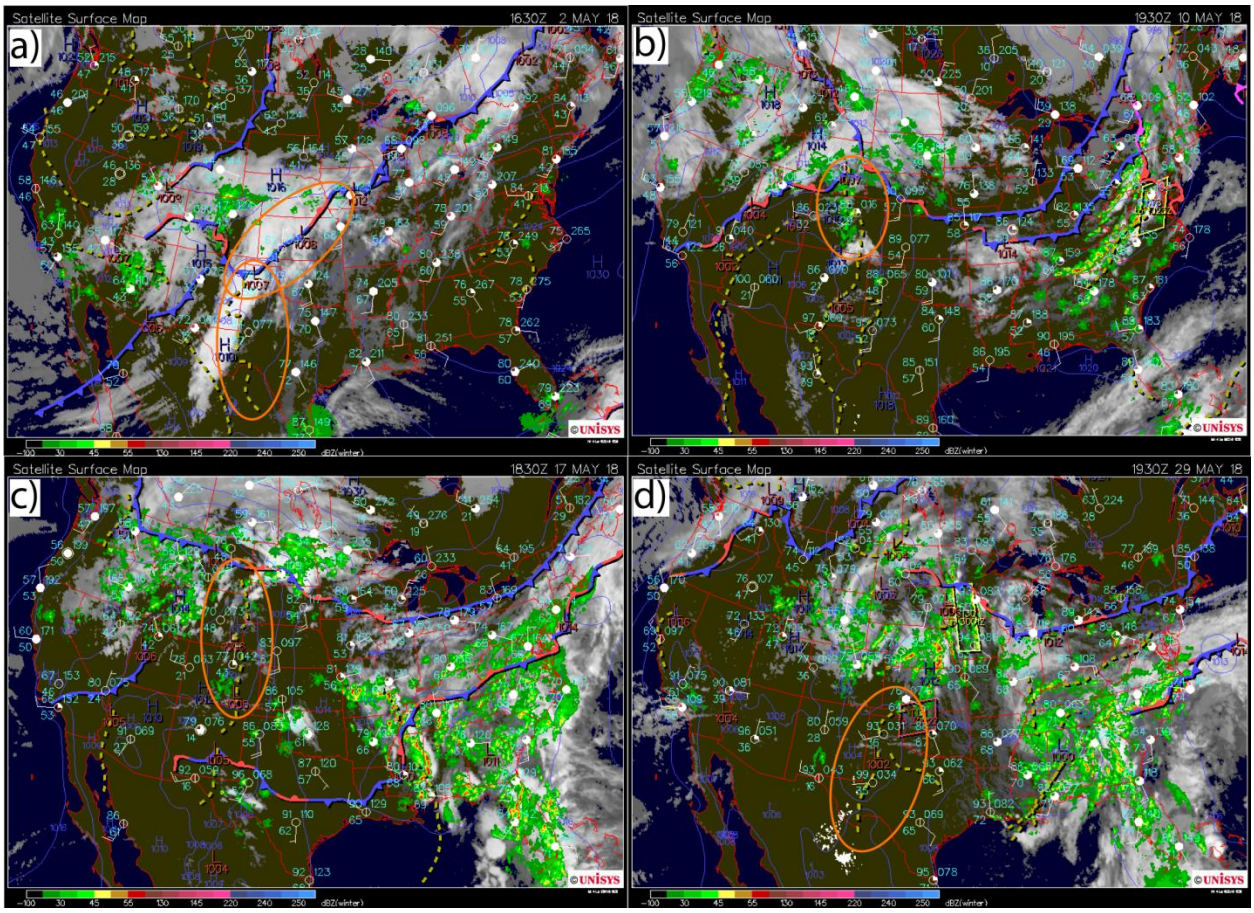

Fig. 5.3. Surface charts for the (a) 05/03, (b) 05/11, (c) 05/18, and (d) 05/30 convective line test cases. Synoptic and mesoscale areas of interest are denoted by orange circles in the plots. Surface charts are courtesy of UCAR, and can be found online at: http://www2.mmm.ucar.edu/imagearchive/.

remnants of a cold front. Between 1900 and 2000 UTC, convection initiates along a N-S dry line on the Kansas/Oklahoma border (Fig. 5.3d). While early storms are more discrete, the storms merge into a convective line by 05/30 0100 UTC (the beginning of the analysis time) centered over central Kansas. The convective line begins to lose a cohesive structure at 0500 UTC. The dry line also spawns a discrete supercell that initiates between 2000 – 2100 UTC in the Texas panhandle. The supercell advects to the east into Oklahoma, and stays separate from the convective line over the analysis period.

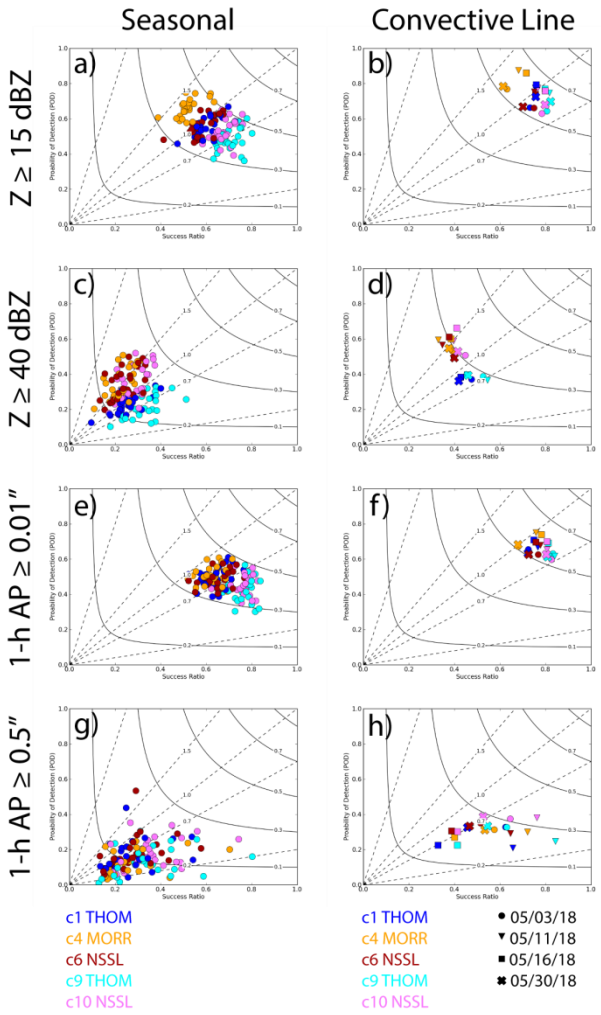

Fig. 5.4. Performance diagrams for seasonal and convective line forecasts of composite reflectivity Z (a,b) ≥ 15 dBZ, (c,d) ≥ 40 dBZ and 1-h accumulated precipitation (e,f) ≥ 0.01" and (g,h) ≥ 0.5". EnKF members 1, 4, 6, 9, and 10 are represented blue, orange, red, cyan, and pink markers, while convective line marker shape corresponds with test cases as denoted in the legend below the subplots.

5.3.2.1 Performance diagrams

Seasonal large-scale storm structure (Z ≥ 15 dBZ) in performance diagrams (Roebber 2009) appears to be more sensitive to initial/boundary conditions, as members 9 and 10 (Thompson and NSSL with same initial/boundary conditions) have more similar bias and critical success index [CSI] with each other than with members 1 and 6 (Thompson and NSSL with different initial/boundary conditions), respectively (Fig. 5.4a). Still, the NSSL and Thompson schemes

show a clear underprediction bias while the Morrison scheme is overpredicting large-scale storm structure. The Morrison scheme has the potential to overpredict storm structure, as graupel in the scheme has a noticeably slower fall speed than in the NSSL and Thompson schemes, and therefore is more prone to horizontal advection. There is no noticeable drop-off in skill among the microphysics schemes, as CSI largely varies between 0.3 and 0.5. Microphysics for the convective line test cases also exhibit these biases, as the Morrison scheme overpredicts large-scale storm structure, while the NSSL and Thompson scheme genearlly undepredict (Fig. 5.4b). There is also a noticeable incrase in CSI (0.5 – 0.7) for the convective line test cases. Because verifications statistics are computed over a limited domain, miss and false alarm potential over regions with strong synoptic/mesoscale forcing are reduced. For convective-scale storm structure, there is a clear difference in bias between the microphysics schemes regardless of initial/boundary conditions (Fig. 5.4c). The Morrison and NSSL schemes tend to overpredict large reflectivity regions ($Z \geq 40$ dBZ), while the Thompson schemes underpredicts these regions. It is important to remember that the Morrison and NSSL members are initialized with ensemble members using Thompson microphysics. Therefore, although the Morrison scheme can overpredict the horizontal extent of convective storm structure (far-advected graupel melting to large raindrops; Johnson et al. 2016) and the NSSL scheme can overpredict reflectivity by simulating large rimed ice (Johnson et al. 2016; Chapter 4), the hydrometeor assumptions built within the schemes are initialized with an entirely set of assumptions (Thompson), and must correct (i.e., more consistent with the scheme's microphysical assumptions) the hydrometeors over the model run. Since these convective regions are more localized than large-scale storm structure, there is also a noticeable decline in forecast skill compared to $Z \geq 15$ dBZ, as both success ratio and POD decreases. The seasonal convective storm-structure biases are also apparent across the convective line test cases

(Fig. 5.4d), as the Morrison and NSSL schemes overpredict convective storm structure while the Thompson scheme underpredicts. Similar to large-scale storm structure, convective storm structure across the convective line test cases are typically more skillful than the seasonal forecasts over most of the CONUS.

Light precipitation accumulation (1-h AP $\geq$ 0.01") is more sensitive to initial/boundary conditions than microphysics, as members 9 and 10 have more similar bias and CSI than members 1 and 6, respectively (Fig. 5.4e). Overall, every microphysics scheme examined here underpredicts (low POD) light precipitation. While the Morrison scheme overpredicts large-scale storm structure, the composite reflectivity examined in this chapter represents a column maximum. Therefore, while the Morrison scheme is overpredicting the column maximum mass, it is clearly either underpredicting total column mass or its sedimentation to the surface. The light preciptation underprediction bias is persistent through the convective line test cases (Fig. 5.4f). There is little CSI difference across the microphysics schemes, as test case light precipitaiton CSI is more related (weakly) to test case rather than microphysics scheme. For heavy precipitation (1-h AP $\geq$ 0.5"), there is a general underprediction bias across all microphysics schemes examined (again, either total column mass or its sedimentation is underpredicted), but also very little bias correlation with microphysics scheme (Fig. 5.4g). In fact, the convective line test cases reveal that heavy precipitation bias and skill is more sensitive to individual cases rather than any microphysics biases (Fig. 5.4h). As success ratio typically increases at a greater rate than probability of detection (POD) across the convective line test cases, convective line test cases with increased CSI are simply forecasting less false alarms. Heavy precipitation skill for both seasonal and test case events is the lowest across storm structure and precipitaiton skill, indicating that microphysics schemes need to capture high intensity precipitaiton events more accurately.

5.3.2.2 Equitable threat score (ETS)

Equitable threat score (ETS) is similar to critical success index, except that "random" hits (proportional to both forecast and observed events) are subtracted from both the numerator (hits) and denominator (the sum of hits, misses, and false alarms). The spread of seasonal ETS is also presented by plotting the 25^{th} and 75^{th} percentiles at each forecast hour in Figure 5.5. At the first two forecast hours, members 9 and 10 (Thompson and NSSL with identical initial/boundary conditions) contain more large-scale storm structure skill ($Z \geq 15$ dBZ) than other members, revealing that ETS during this early model period is more sensitive to initial conditions than microphysics. ETSs are also at their highest value during this period, which is expected as data assimilation ends at 0000 UTC, after which model error begins to increase. The Morrison scheme is the least skillful BMP examined over the first two forecast hours, but is the most skillful from t = 0400 – 0600 UTC as the scheme does not degrade in skill as quickly as the NSSL and Thompson schemes. After t = 0200 UTC, there is not a discernable advantageous BMP between the NSSL and Thompson schemes, consistent with large-scale storm structure performance diagrams which showed little CSI difference with microphysics. For the convective line test cases, large-scale storm structure skill is highly case-dependent (Fig. 5.5b). Still, relative ETS behavior is similar to seasonal ETS. The Morrison scheme typically has the lowest ETS over the first two forecast hours. Because the scheme has a large-scale overprediction bias, false alarms are reducing the scheme's ETS without a compensating score increase from hits. On the other hand, members 9 and 10 (Thompson and NSSL with identical initial/boundary conditions) have the highest ETS at the first two forceast hours. At later time, there is little BMP advantage in regard to BMP, similar to seasonal ETS plots. Each member in the 05/03 convective line case began the forecast period with larger ETS than in other test cases, but rapidly declined until 0400 UTC, where ETS is the lowest

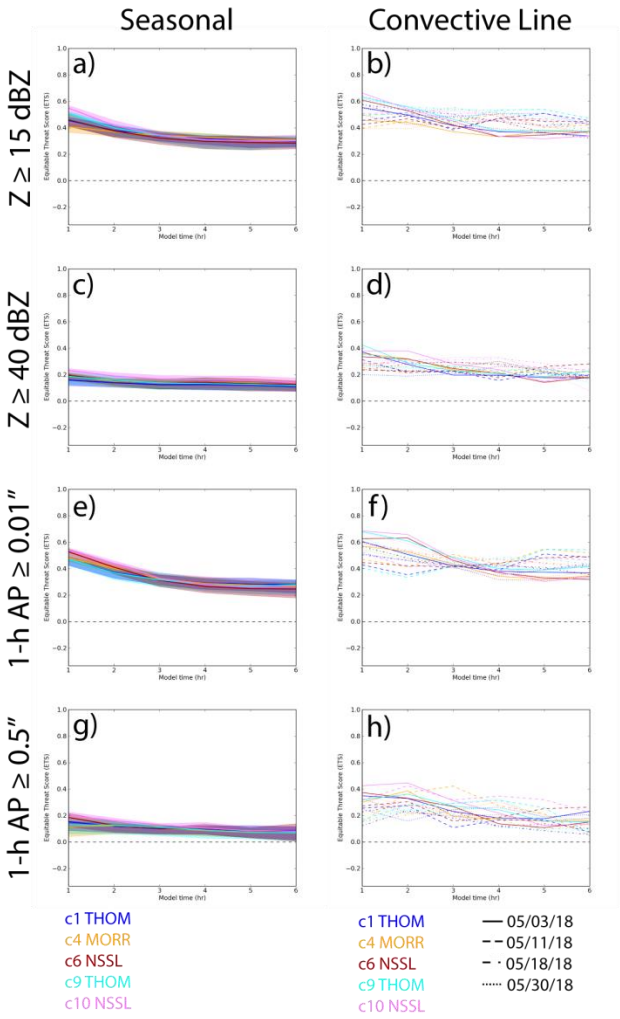

Fig. 5.5. Equitable threat scores (ETSs) over 1-6 forecast hours (0100 UTC – 0600 UTC) for seasonal and convective line of composite reflectivity Z (a,b) ≥ 15 dBZ, (c,d) ≥ 40 dBZ and 1-h accumulated precipitation (e,f) ≥ 0.01" and (g,h) ≥ 0.5". EnKF members 1, 4, 6, 9, and 10 are represented blue, orange, red, cyan, and pink lines, while convective line ETS linestyle corresponds with test cases as denoted in the legend below the subplots. Seasonal lines also include shading representing the 25th and 75th percentile of daily ETS over the season.

at this compared to members in other test cases. Therefore, between 0100 UTC – 0400 UTC, each ensemble member is either missing observed large-scale storm structure (reducing hits and increasing misses) and/or simulating storms in the wrong location (additionally increasing false alarms).

For convective storm structure ($Z \geq 40$ dBZ), the NSSL and Morrison schemes generally outperform the Thompson BMP over the 2018 HWT Spring Experiment (Fig. 5.5c). However, ETS typically hovers between 0.1 and 0.2 for each BMP. In fact, similar to $Z \geq 15$ dBZ, ETS among the BMPs examined is very similar from t = 0300 – 0600 UTC. ETS typically does not correlate with BMP across the convective line test cases for early (0100 – 0200 UTC) and late (0500 – 0600 UTC) forecasts (Fig. 5.5d). Member 10 (NSSL) has the largest ETS over the middle forecast period (0300 – 0400 UTC), while member 1 (Thompson) typically has the smallest ETS. Still, these ETSs cannot be attributed to microphysical bias, as ETS for their corresponding members with identical microphysics (members 9 and 6, respectively) vary across the test cases.

Seasonal ETS is analyzed for light (1-h AP $\geq$ 0.01") and heavy (1-h AP $\geq$ 0.50") precipitation and, similar to storm structure, is typically largest at t = 0100 UTC as this is the closest analysis time to assimilation (Figs. 5.5e,g). At early analysis times (t = 0100 – 0200 UTC), the NSSL scheme has larger ETS than the Morrison and Thompson schemes for light precipitation. However, after t = 4 h when BMP ETS shows little variance, the Thompson and Morrison scheme have larger ETS than the NSSL scheme. Member 10 (NSSL) typically has the largest heavy precipitation ETS, but there is little correlation between BMP and ETS. This is consistent with previous performance diagrams, which reveal little precipitation CSI difference between the schemes. BMP light precipitation ETS over the convective line test cases is highly case dependent (Fig. 5.5f). Members 9 and 10 typically have the highest ETS at the first forecast hours, except for the 05/11 test case. By the end of the analysis, the Thompson scheme generally has the largest light precipitation ETS. Among the test cases, light precipitation ETS for the 05/03 test case declined rapidly until t = 0400 UTC, while ETS for the 05/11 test case increased over the analysis time. Heavy precipitation ETS across the convective line test cases is largest for the NSSL scheme

at t = 0100 UTC (Fig. 5.5h). As heavy precipitation is typically underpredicted, the NSSL scheme is forecasting more hits (and therefore, less misses) at this time. After this time, heavy precipitation ETS across the test cases is highly case dependent, as no scheme consistently shows comparatively higher or lower ETS. Overall, there is weak correlation between BMP and storm structure/accumulated precipitation ETS over the entire analysis period, implying that ETS for both seasonal and deep-convective test cases is not very sensitive to BMP choice.

5.3.2.3 Fractions skill score

Fractions skill score (FSS; Roberts and Lean 2008) is a flavor of root mean square error (RMSE) that accounts for spatial errors in forecasts in its formulation. Instead of comparing forecast and observed output at each grid point, the difference between forecast and observed neighborhood coverage around each grid point is calculated. Therefore, the forecast can exhibit skill as long as the prognosed storm structure/accumulated precipitation examined here is similar in structure and spatially close (but not necessarily exact) to observations. Forecasts are considered "skillful" when the FSS exceeds 0.5+0.5×O, where O is the observed rate of the variable and threshold (i.e., $Z > 15$ dBZ). As long as there is at least one forecast and one observation in the domain, FSS asymptotes to 1 as neighborhood size increases toward the size of the domain.

For large-scale storm structure ($Z \geq 15$ dBZ), all BMPs are considered skillful as the FSS at all neighborhood sizes exceeds the skill line (Fig. 5.6a). Members 9 and 10 have similar FSS at smaller neighborhood square half-lengths, while the Thompson scheme outperforms the remaining BMPs as neighborhood size increases. The Morrison scheme has the lowest FSS at all neighborhood sizes. The scheme is unable to simulate large-scale storms consistent with observations because of its overprediction bias. For convective lines, the Thompson scheme typically has a larger FSS than other BMPs, especially as neighborhood half-length increases

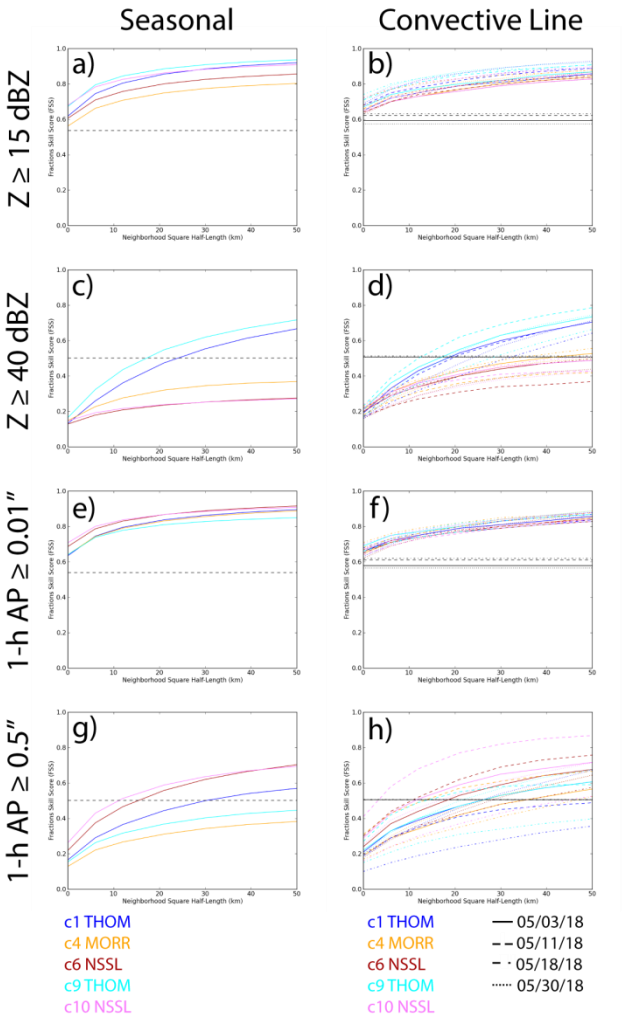

Fig. 5.6. Fractions skill score (FSS) over 1-6 forecast hours for seasonal and convective line forecasts of composite reflectivity *Z* (a,b) ≥ 15 dBZ, (c,d) ≥ 40 dBZ and 1-h accumulated precipitation (e,f) ≥ 0.01" and (g,h) ≥ 0.5" as a function of neighborhood square half-length. EnKF members 1, 4, 6, 9, and 10 are represented blue, orange, red, cyan, and pink lines, while convective line FSS linestyle corresponds with test cases as denoted in the legend below the subplots. The skill line in each plot is either seasonal, or specific to test case as denoted by its linestyle.

(Fig. 5.6b). Although the scheme has a slightly larger underprediction of large-scale storm structure compared to the NSSL scheme, its spatial distribution of forecasted large-scale storm structure is more consistent with observations. Still, all BMPs performed above their large-scale storm structure skill line at all scales with FSS exceeding 0.8 at neighborhood square half-length of 50 km.

The Thompson BMP has a higher FSS than the Morrison and NSSL schemes for convective storm structure, which increases as neighborhood half-length increases (Fig. 5.6c). The Morrison scheme also has a larger FSS than both NSSL members, but neither the Morrison nor the NSSL schemes are able to produce a "skillful" forecast at any neighborhood size. The Thompson scheme is only able to produce convective storm structure skill forecasts when neighborhood half-length exceeds ~25 km. This FSS score behavior is also apparent for the convective line test cases (Fig. 5.6d). The Thompson BMP exceeds the skill line and other BMP FSSs as neighborhood half-length increases, which the other schemes struggle to produce skillful forecasts. The Morrison scheme typically has larger convective storm structure FSS than the NSSL scheme. As the Morrison and NSSL schemes overpredict convective storm structure, these schemes are unable to skillfully match observed convective storm structure as in the Thompson scheme. The NSSL scheme can simulate large reflectivity (compared to observations) due to its "large hail" category design (Johnson et al. 2016). The Morrison scheme advects graupel horizontally to a greater extent than other BMPs due to its small graupel fall speed, which melts to large rain and contributes to large reflectivity (Johnson et al. 2016).

The NSSL BMP produces higher fractional skill score than the remaining BMPs (Fig. 5.6e) for seasonal light accumulated precipitation (1-h AP $\geq$ 0.01"). Even though the NSSL scheme lags the Thompson scheme in terms of quantitatively simulating large-scale storm structure consistent with observations, it is able to better simulate light precipitation total column mass and its sedimentation to the surface. Still, each BMP produces skillful light precipitation forecasts at all neighborhood lengths as FSSs are well above the skill line. Each microphysics scheme examined produces skillful light precipitation forecasts for the four convective line test cases (Fig. 5.6f). There is no clear advantageous BMP that is able to better simulate light precipitation for convective

lines. In fact, even though member 6 (NSSL) has the largest seasonal light precipitation FSS for many of the neighborhood lengths examined, it also simulates the lowest light precipitation FSS for several of the test cases. Therefore, although the scheme is able to generally simulate light precipitation for storm modes across the 2018 Spring Experiment period (May) consistent with observations, it struggles with the underprediction of light precipitation of convective lines (i.e., stratiform), likely due to the scheme's large rimed ice design with little horizontal advection. The NSSL BMP also produces the most skillful seasonal heavy precipitation forecasts (1-h AP $\geq$ 0.5") for all neighborhood sizes (Fig. 5.6g). The Thompson scheme noticeably outperforms the Morrison scheme, but both schemes struggle to produce forecasts with skill. Similar to light precipitation and its relation to composite reflectivity, the NSSL scheme produces the least skillful convective storm structure seasonal heavy precipitation forecasts, but is clearly able to simulate the total column mass well for heavy precipitation events. Except for the 05/30 convective line case, the NSSL scheme is also able to simulate heavy precipitation for convective lines with more skill than the other BMPs examined (Fig. 5.6h). Therefore, even though the BMP struggles with stratiform regions of convective lines (presumably due to the horizontal rimed ice advection deficiency in the scheme), it is able to better simulate the more localized heavy precipitation (i.e., large rimed ice, large rain mass) events. Although the Morrison scheme produces less skillful season heavy precipitation forecasts compared to the NSSL and Thompson schemes, it produces more skillful forecasts than the Thompson scheme for two of the convective lines examined (05/11, 05/18). As the Morrison scheme was developed and tested for a squall line, it might be a more appropriate BMP to simulate convective lines than the Thompson BMP.

5.3.2.4 Convective line test case paintball plots

Paintball plots over the convective line test case domains are analyzed with composite reflectivity and 1-h accumulated precipitation thresholds consistent with previous subsections of section 5.3 to help explain the microphysics behavior examined in the forecast evaluation diagrams (i.e., performance, ETS, FSS). The Morrison BMPs persistent overprediction of large-scale structure evident in the previously-analyzed performance diagrams is apparent in the paintball plots at t = 0100 UTC, where both the Morrison and NSSL scheme typically produce a larger areal coverage of $Z \geq 15$ dBZ across the test cases than the Thompson BMP (Figs. 5.7a-d). As this overprediction is not as widespread at later analysis times (but still present, as in the Morrison scheme; Fig. 5.8), the initial large-scale storm overprediction is likely due to ensemble initialization, which comes from Thompson microphysics (and therefore, introduces a separate set of hydrometeor assumptions and parameterizations into the Morrison and NSSL schemes that must be corrected at later forecast hours). The Morrison BMP typically has lower ETS at early forecast hours because of these false alarms. Similarly, members 9 (Thompson) and 10 (NSSL) have the largest ETS at early forecast hours because these members are less prone to false alarms in the convective lines, which is due to initial/boundary conditions (initialized from the ensemble mean) rather than microphysics parameterization. Still, each BMP is able to match the areal distribution of large-scale storm structure well across the convective line test cases, reflected in the similar FSS across the BMPs.

Similar to large-scale storm structure, the Morrison and NSSL BMPs overpredict convective storm structure for convective line test cases relative to the Thompson BMP (Figs. 5.7e-h). Rimed ice (which typically contributes to this larger reflectivity) is fundamentally different between the Morrison, NSSL and Thompson schemes: Morrison rimed ice is represented by a medium density and slow fall speed, NSSL rimed ice represents growing, rimed graupel

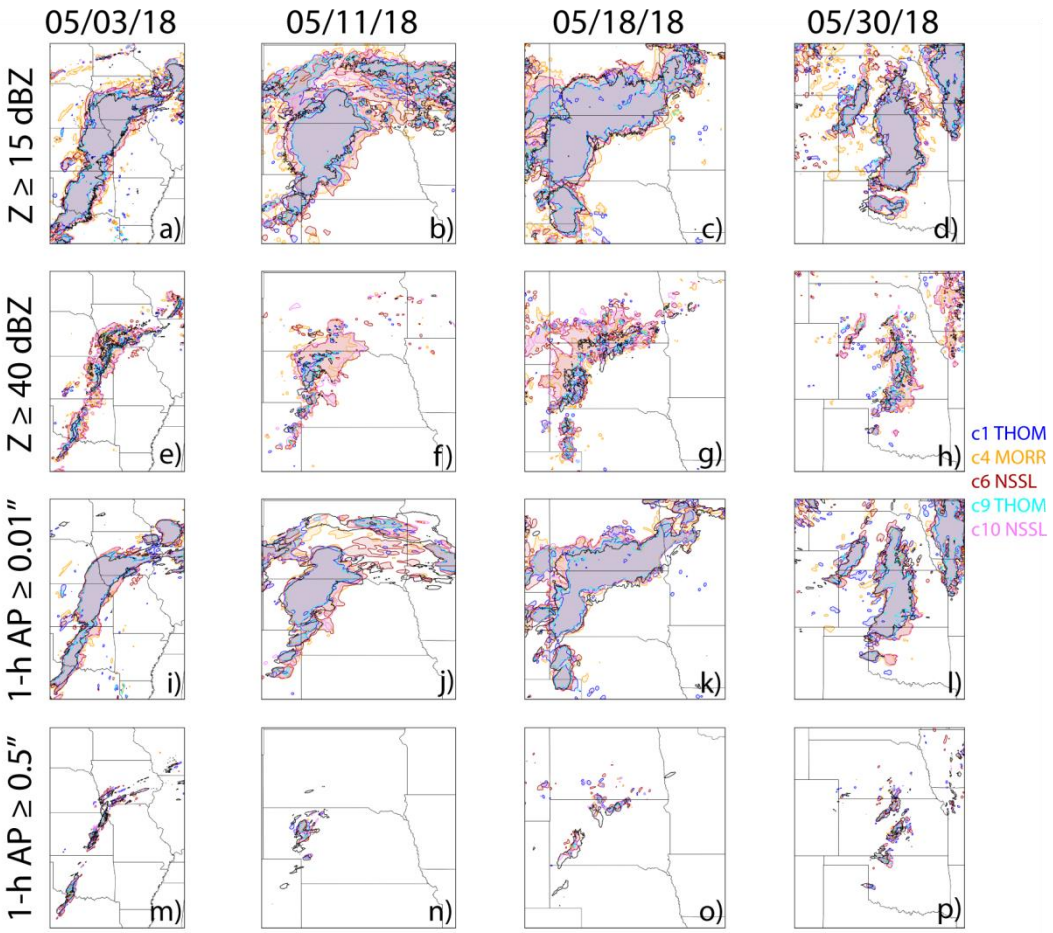

Fig. 5.7. Paintball plots of composite reflectivity *Z* (a,b,c,d) ≥ 15 dBZ, (e,f,g,h) ≥ 40 dBZ and 1-h accumulated precipitation (i,j,k,l) ≥ 0.01" and (m,n,o,p) ≥ 0.5" for the convective line test cases at t = 0100 UTC. EnKF members 1, 4, 6, 9, and 10 are represented blue, orange, red, cyan, and pink shaded regions, while observations are contoured in black.

moving to large hail, while Thompson graupel diagnoses the intercept parameter N_0 to better represent wet growth's effect on the rimed ice distribution. Therefore, it is not completely valid to initialize microphysics with a completely different set of assumptions, which is clear in the convective storm structure paintball plots. These results are consistent with performance diagram biases in section 5.3.2.1, where the Morrison and NSSL schemes have a persistent positive convective structure bias, while the Thompson scheme has a negative convective structure bias across seasonal and test cases. Still, relative BMP ETS is inconsistent at this time for convective

lines because even though the Morrison and NSSL BMPs are penalized for false alarms, underprediction in the Thompson BMP reduces hits and increases misses. Although Thompson convective storm structure is typically underpredicted, its areal coverage is much closer to observations than the massive overprediction in the Morrison and NSSL schemes, reflected in the scheme's typically larger FSS.

Model light precipitation matches observations rather well, despite the underprediction biases seen in previous performance diagrams (with the exception of light precipitation from a stationary front to the north of the 05/11 convective line; Figs. 5.7i-l). The Morrison and NSSL schemes still tend to predict greater light precipitation coverage than the Thompson scheme, but the extent of overprediction (i.e., false alarms) is much less than that of storm structure. Therefore, light precipitation underprediction occurs at later forecast hours. Light precipitation ETS and FSS shows very weak (if any) correlation with BMP, which is reflected in the paintball plots as BMP light precipitation validation is similar across the test cases at this time. On the other hand, heavy precipitation underprediction is apparent across all convective line test cases, and agrees with previous performance diagrams (Figs. 5.7m-p). The BMPs examined here do not demonstrate a spatial displacement bias (except for an eastern bias for two heavy precipitation swaths over Oklahoma on 05/30), but rather simply underpredict 1-h AP $\geq$ 0.5" coverage for the test cases. The NSSL scheme typically has the largest areal extent of heavy precipitation resulting in a larger ETS (i.e., not getting penalized for both increased misses and less hits) and FSS. This is a greater representation of large column mass in the scheme, which is reasonable as the scheme's rimed ice configuration is representative of riming graupel growing to large hail.

As forecast evaluation metrics examined in this chapter are either composite (performance diagram, fractions skill score) or analyzed temporally (ETS), paintball plots are additionally

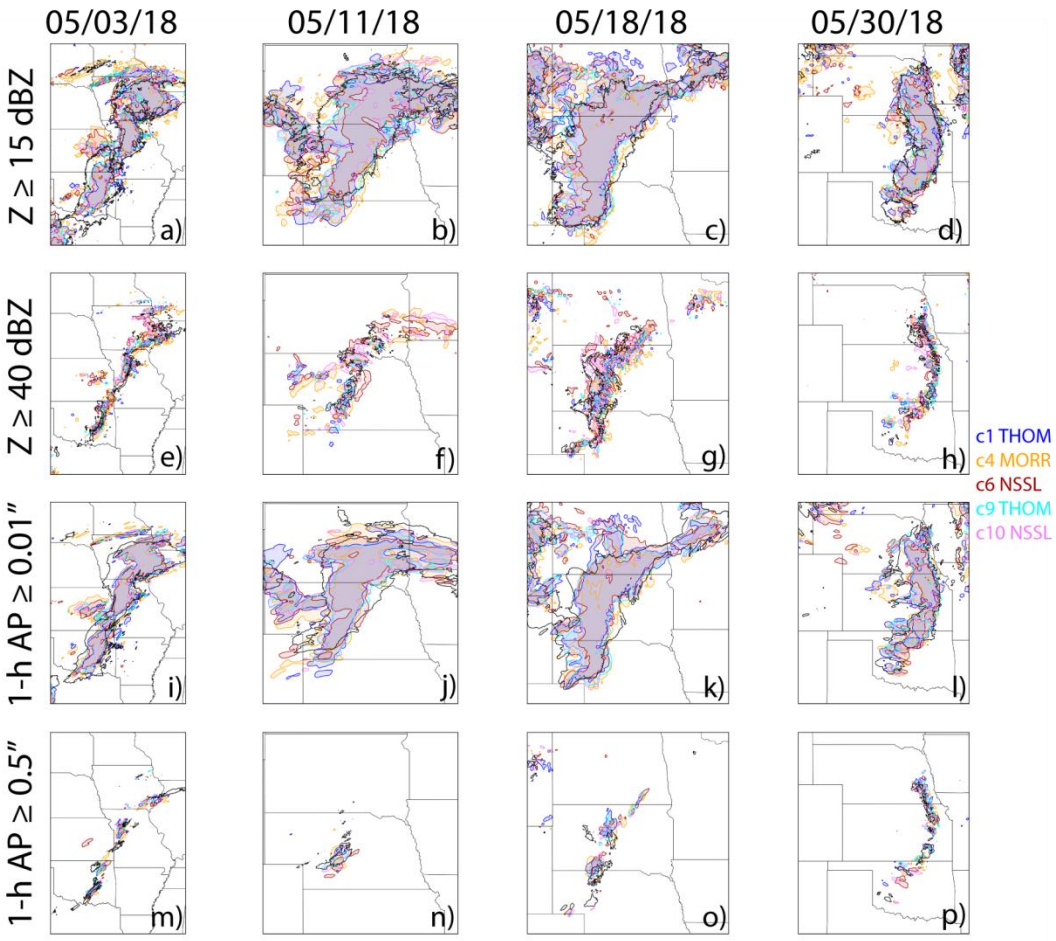

Fig. 5.8. As in Fig. 5.7, but at t = 0400 UTC.

examined at t = 0400 UTC to examine model performance evolution (Fig. 5.8). While members 1 (Thompson) and 6 (NSSL) typically underpredict the west-east coverage of the convective lines, they generally simulate a larger convective line in the N-S direction and match secondary storms better than their member 9 and 10 counterparts, which is why members 1 and 6 have a smaller underprediction bias. The Morrison scheme still continues to simulate the largest areal coverage of large-scale storm structure, consistent with its overprediction bias. Each BMP noticeably underpredicts large-scale storm structure over Oklahoma (convective line triggered by a dry line) for the 05/03 test case. This explains the rapid decline in ETS across all BMPs for this case. The Morrison and NSSL BMP convective structure overprediction is also noticeable in the paintball

plots, as these schemes tend to produce a larger areal coverage than the Thompson scheme (Figs. 5.8e-h). This explains the larger convective storm structure FSS for the Thompson scheme, whose convective storm structure underprediction matches observations closer than the Morrison and NSSL schemes. However, member 10's (and member 6's to an extent; both NSSL members) consistent overprediction of convective structure allows the scheme to have a higher ETS for convective line cases at this time, as the scheme is able to produce more hits (and less misses) than the other BMPs. Except for the convective line initiated from the stationary front over Missouri (model lag error), there is little position error for convective storm structure across the convective line cases, although the Morrison and NSSL convective storm structure overprediction could potentially mask this error.

The model light precipitation underprediction bias is present at t = 0400 UTC (while it is not at t = 0100 UTC; Figs. 5.8i-l). Each BMP underpredicts both the western and eastern light precipitation of the convective lines. ETS noticeably declines for the 05/03 case and increases for the 05/11 case. This is apparent in the paintball plots: there are several misses (and less hits, contributing to smaller ETS) across the BMPs for both convective lines examined in the 05/03 case, while the stationary front precipitation with misses at 0100 UTC continuously leaves the domain and is much less present at 0400 UTC, increasing ETS. The underprediction biases present for heavy precipitation (1-h AP $\geq$ 0.5") in the convective line test cases are evident in the paintball plots at t = 0400 UTC as the model typically both underpredicts the areal extent and misses observed heavy precipitation (Figs. 5.8m-p). Model simulated heavy precipitation areal coverage and position at this time is similar among the BMPs (i.e., no noticeable BMP position error), which is why ETS varies little at this (and later) times. As NSSL FSS is typically larger than other BMPs,

the NSSL's greater heavy precipitation areal coverage at earlier times is more likely linked to the scheme's larger FSS (rather than at this time).

5.3.3 Brightness Temperature

Brightness temperature in the 11.2 μm channel is simulated from 2018 NOAA HWT Spring Experiment model output by coupling the Community Radiative Transfer Model (CRTM; see Jung et al. 2013) with the Advanced Regional Prediction System (ARPS; Xue et al. 2003), and is verified with Geostationary Operational Environmental Satellites [GOES] channel 14 observations. Consistent with the rest of section 5.3, contingency tables are created over short-term forecasts (1-6 h) but only for the prior four convective line test cases analyzed. Two thresholds are chosen consistent with prior simulated satellite brightness temperature studies: brightness temperature T_B < 225 K to isolate deep convection (Jones et al. 2018) and T_B < 270 K for all clouds (Cintineo et al. 2014). Because the brightness temperature threshold of 270 K spans a greater areal coverage than other variables (i.e., composite reflectivity *Z*, 1-h accumulated precipitation) examined in this chapter, the test case domains are expanded to include the entire seasonal analysis domain (Fig. 5.1). While the same BMPs (Morrison, NSSL, and Thompson) are examined here as in section 5.3.2, numerical simulations differ in microphysics only consistent with members 9 and 10 (Table 5.1; although NSSL simulations consistent with member 6 are included to better understand the impact of initial/boundary conditions on short-term forecasts).

There is a noticeable drop-off in upper-level cloud skill (T_B < 225) between the 05/30 convective line case (CSI between 0.2 and 0.4) and the other three test cases (CSI between ~0.5 to 0.7; Fig. 5.9a). Across all cases, the Morrison and Thompson BMPs typically underpredict upper-level clouds, while the NSSL scheme typically overpredicts these clouds (i.e., large number of hits, but also false alarms). Fall speeds for cloud ice particles larger than 20 μm are smaller in the

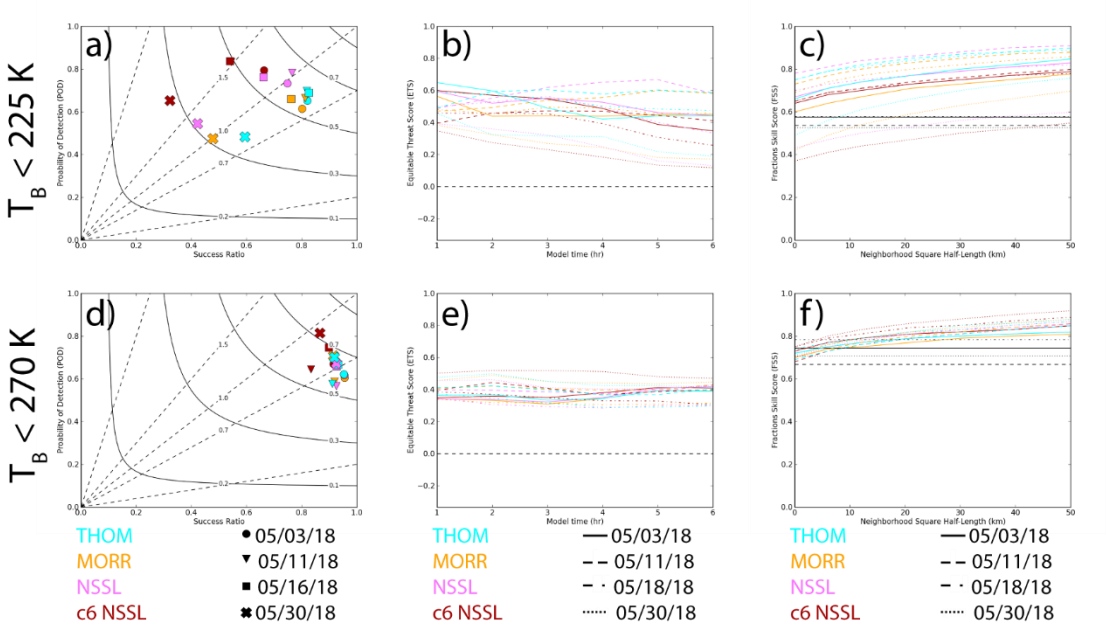

Fig. 5.9. Performance, equitable threat score (ETS), and fractional skill score (FSS) using brightness temperature thresholds T_B (a,b,c) < 225 K and (d,e,f) < 270 K. Thompson, Morrison, and NSSL runs with identical initial/boundary conditions are denoted by cyan, orange, and pink markers/lines, while NSSL simulations consistent with member 6 is red. Each test case is denoted by marker/linestyles as denoted in legends below the subplots.

NSSL scheme (not shown), and therefore have the potential to be lofted higher by the updraft (decreasing T_B) and horizontally advected to a greater extent than Morrison and Thompson cloud ice. In the framework of deep convective storms, the Morrison and Thompson schemes are simulating smaller storm depth compared to observations, while the NSSL scheme is producing more intense storms (i.e., stronger updraft, increased storm depth and cooler brightness temperatures). Specifically comparing the NSSL scheme with different initial/boundary conditions, member 6 has consistently smaller skill and higher overprediction bias than the NSSL member with identical initial/boundary conditions as the Morrison and Thompson runs. In terms of all clouds, every microphysics scheme contains an underprediction bias across the four test cases (Fig. 5.9d). Member 6 (NSSL), which has the strongest overprediction bias of upper-level

clouds, has the smallest underprediction bias of the BMPs examined. This can be attributed to initial/boundary conditions rather than the NSSL cloud ice bias, as NSSL microphysics with identical initial/boundary conditions as the Morrison and Thompson schemes typically has a greater underprediction bias of all clouds than these two schemes. Forecast skill of all clouds is typically moderate (0.5 – 0.7), but the bias of most BMPs are near 0.7. Therefore, while success ratio is typically high (0.8-1.0) with a small number of false alarms, the probably of detection (POD) among the BMPs is centered around 0.7, as the BMPs are missing many clouds of the test case forecasts.

Upper-level clouds (T_B < 225 K) equitable threat score (ETS) for the 05/30 test case is noticeably lower than the other three cases (Fig. 5.9b). This is not unexpected, as previous performance diagrams reveal a noticeable decline in CSI for this case. Member 6 (NSSL) typically has the smallest ETS over the short-term forecast period. This can be attributed to initial/boundary conditions, as its corresponding NSSL member with different initial/boundary conditions (same as the Morrison and Thompson) vary in relative ETS compared to these two BMPs. The Thompson has the highest ETS over the first two forecast hours, while the Morrison scheme generally improves upper-level cloud ETS by the end of the short-term forecast period. While all cloud absolute ETS (T_B < 270 K) is similar across the cases and falls between 0.3 and 0.6, ETS is less sensitive to BMP choice as the test case ETSs often cluster together (Fig. 5.9e). The BMPs have the highest all cloud ETS for the 05/30 test case, which is also the test case with the smallest upper-level cloud ETSs. All cloud ETS is high for member 6 at early forecast hours (0100-0300 UTC). While false alarms penalized member 6 for upper-level clouds, the more aggressive convection in this member results in better overall cloud coverage and more all cloud hits. Similar to upper-level clouds, the Morrison BMP typically forecasts all clouds with more skill as the forecast hour

increases. By the end of the forecast period, the Thompson BMP generally has the lowest ETS for all clouds (even though the scheme simulates convective clouds well relative to the other two schemes), indicating that the BMP increasingly misses observed low-level clouds over the short-term forecast period.

The Thompson BMP generally has the largest upper-level cloud fractions skill score (FSS) especially as neighborhood size increases, except for the 05/11 case where Thompson FSS is the second largest regardless of neighborhood size (Fig. 5.9c). Therefore, the Thompson scheme is able to better match the observed spatial pattern of upper-level clouds (even though the scheme underpredicts [i.e., more misses] these clouds) compared to the other BMPs examined. Member 6 (NSSL) generally has the lowest FSS (except for the 05/03 test case where its FSS is the second-lowest). This is not unexpected, as this member typically had the lowest ETS (false alarms) over the short-term forecast period. FSS across the test cases are typically skillful, except for the 05/30 case in which member 6 struggles to produce skillful upper-level cloud forecasts until neighborhood half-length reaches 40 km. The other three 05/30 simulations produce skillful forecasts once neighborhood half-length exceeds 20 km. This case also contains the lowest ETS and CSI compared to the other three cases. All cloud FSS is largely similar among the four test cases spanning from 0.7 to 0.9 across all neighborhood sizes (Fig. 5.9f). With the exception of the 05/11 test case where FSS is tightly clustered among the BMPs, member 6 has the largest FSS across neighborhood sizes. Again, as this simulation design overpredicts upper-level cloud coverage, the more aggressive convection in the scheme results in a better match between all simulated and observed clouds. This is not a microphysical bias, as the corresponding NSSL member with different initial/boundary conditions varies in relative performance compared to the other two schemes (i.e., both more and less skillful). Once neighborhood size reaches 10 km, each

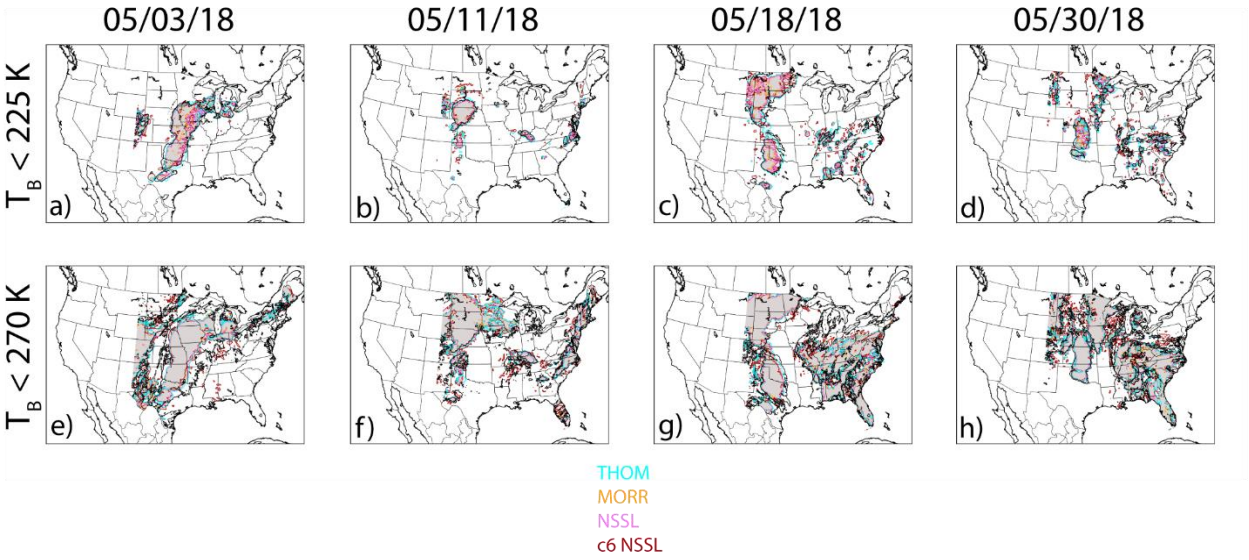

Fig. 5.10. Paintball plots using brightness temperature thresholds T_B (a,b,c,d) < 225 K and (e,f,g,h) < 270 K for the convective line test cases at t = 0100 UTC. Thompson, Morrison, and NSSL runs with identical initial/boundary conditions are denoted by cyan, orange, and pink shaded regions, while NSSL simulations consistent with member 6 is red. Observations are contoured in black.

BMP is able to produce skillful all cloud forecasts. Aside from member 6, BMP performance in simulating all clouds varies inconsistently across the test cases.

Paintball plots using the analyzed brightness temperature thresholds (T_B < 225 K, 270 K) are created to further investigate BMP cloud biases and skill for each of the four test cases (Figs. 5.10; 5.11). At t = 0100 UTC, the Thompson BMP typically produces the largest upper-level cloud (Figs. 5.10a-d) coverage across the four test cases. This results in the largest ETS over the first few forecast hours, as the scheme is able to record more hits without the penalization of misses. Upper-level cloud simulation behavior is case-dependent: typically the BMPs reasonably simulate upper-level cloud coverage for the 05/03 case (Fig. 5.10a), each BMP simulated too intense storms (larger cloud depth compared to observations) over the Oklahoma panhandle/southwestern Kansas forced by a dry line for the 05/11 test case (Fig. 5.10b), the Thompson BMP and member 6 suffer from several false alarms in the southeast U.S. forced by a low pressure system attached to a stationary

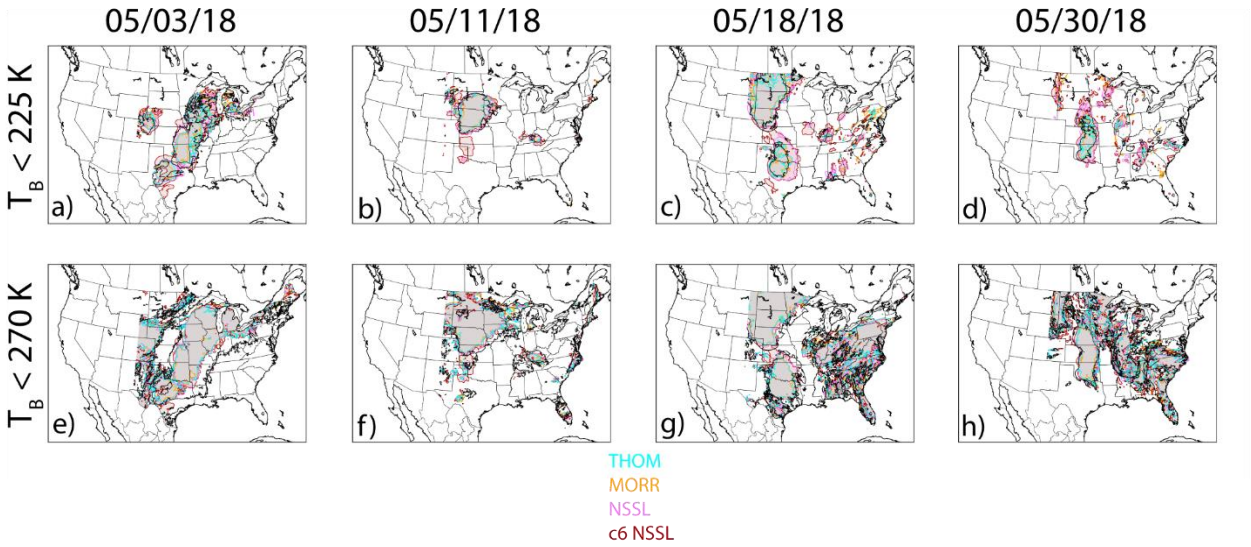

Fig. 5.11. As in Fig. 5.10, but at t = 0400 UTC.

front for the 05/18 case (Fig. 5.10c), and each scheme underpredicts the strength of storms (smaller observed depth) over Iowa forced by a low pressure system attached to a stationary front for the 05/30 case (Fig. 5.10d). The Morrison scheme typically simulates the weakest storms (smaller observed depth) at this time, which is expected as the scheme displayed an upper-level cloud underprediction bias. Each BMP underpredicts the spatial coverage of all clouds ($T_B < 270$ K) and misses isolated observed clouds across the four test cases (Figs. 5.10e,f,g,h). This is expected as each BMP displayed a persistent all cloud underprediction bias across all test cases. Member 6 (NSSL) typically simulates more clouds near isolated cores (i.e., cold front in North Dakota for 05/03 case, stationary front near Arkansas for 05/11 case), which allows the scheme to have large ETSs at this time.

Compared to upper-level cloud ($T_B < 225$ K) coverage at t = 0100 UTC, the NSSL BMP (especially member 6 configuration) consistently overpredicts upper-level clouds across the four test cases, both in coverage span and false alarms (Figs. 5.11a,b,c,d). Therefore, the scheme is overpredicting the intensity of storms, and is more consistent with previous performance diagrams

that revealed the scheme's overprediction bias. This explains the member's typically lowest fractions skill score (FSS), as the model simulates upper-level clouds at a greater frequency than observations. This is likely due to the scheme's smaller cloud ice fall velocity compared to the other two BMPs, which allows cloud ice to be lofted higher by updrafts (decreasing T_B) and horizontally advected to a greater spatial extent than cloud ice in the other two BMPs. For the 05/30 test case, each BMP misses several upper-level clouds over the southeastern U.S. from tropical storm Alberto, and overpredicts storm intensity from a stationary front over Illinois and Minnesota/Wisconsin and a cold front over eastern Montana. Because misses lower CSI and ETS more than false alarms (i.e., both decreased hits and increased misses), the missed storm intensities over the southeastern U.S. is the likely reason why upper-level cloud CSI and ETS for this case are noticeably lower than in the other three cases. Further, these misses and overpredictions illustrate each microphysics' inability to match observed upper-level cloud distributions for this case, and explains the drop in FSS compared to the other three test cases. Thompson simulated upper-level cloud coverage is also noticeably smaller than at t = 0100 UTC, which explains the BMP's underprediction bias. Still, it predicts more upper-level cloud coverage than the Morrison BMP near observed upper-level clouds and does not overpredict upper-level cloud coverage as in the NSSL scheme, which is why the Thompson scheme typically has the highest upper-level cloud FSS.

With the exception of the 05/30 case (which has the highest all cloud ETS), each BMP noticeably underpredicts all cloud coverage, with large-scale incorrect storm width biased on the eastern edge of storms (Figs. 5.11e,f,g,h). As there are no widespread false alarms on the western edge of these storms, this eastern concentration of misses is likely more the result of storm lag bias coupled with underprediction, rather than storm lag bias alone. In fact, all cloud false alarms are

rare, which is why success ratio is so large across the BMPs. While there are many southeastern U.S. upper-level cloud misses for the 05/30 test case, there are few at the all cloud threshold (T_B < 270 K). It is clear that the BMPs are not missing these storms, but rather are underpredicting their intensity. The 05/11 test case has two noticeable all cloud misses in eastern Tennessee from a cold front and a lag bias over central Texas from a dry line. Still, these storms are closer to mesoscale, which is why ETS is not prohibitively penalized at this time. While all cloud coverage is similar across the BMPs, member 6 predicts relatively larger storm coverage (due to initial/boundary conditions, as the corresponding NSSL member periodically simulates smaller all cloud coverage than the Morrison and Thompson schemes), which allows the member to typically have the largest FSS and smallest underprediction bias.

5.4 Next-day forecasts

The ensemble and microphysical performance of next-day forecasts are examined in addition to the previously analyzed short-term forecasts. The Storm Prediction Center (SPC) currently creates Day 1, Day 2, Day 3, and Day 4-8 convective outlooks, which are initially valid from 1200 UTC – 1200 UTC for their relevant days. Therefore, next-day forecasts are defined here using the same definition of the SPC's Day 2 convective outlooks, which is 1200 UTC – 1200 UTC on the day after forecasts are created. As the CAPS 10-member EnKF is initialized at 0000 UTC, next-day forecasts are defined as the 1200 UTC – 1200 UTC period starting on the following day (i.e., the 12-36 forecast hours) as in Loken et al. (2017). Further, as SPC probabilities are defined to predict an event happening within 25 mi (~40 km) of the forecasted probability, the neighborhood used for verification is relaxed from the 9 km square half-length for short-term forecasts to 39 km square half-length for next-day verification.

5.4.1 Ensemble performance

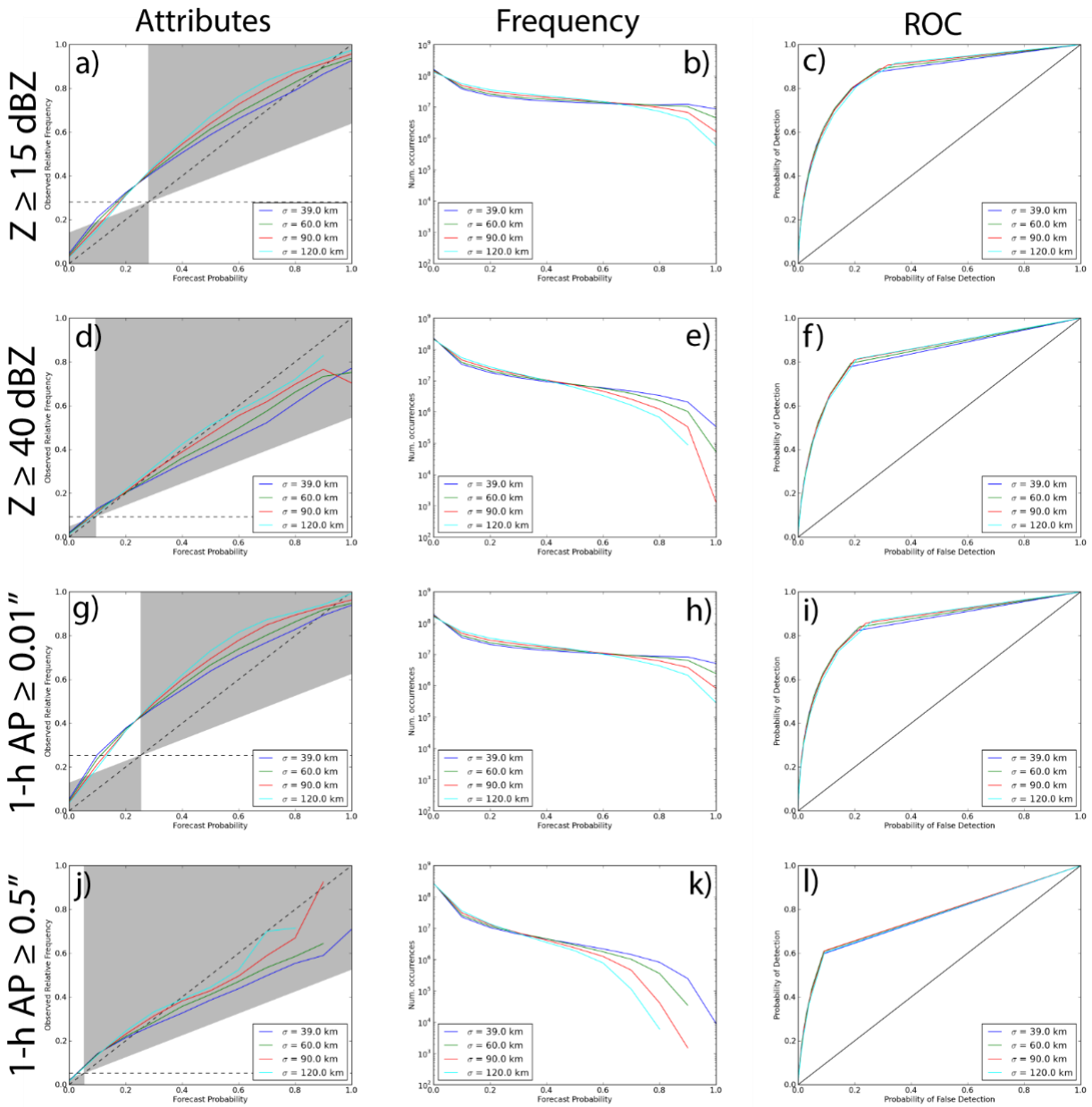

Fig. 5.12. Attributes, frequency, and relative operating characteristic (ROC) curve diagrams for the 10-member EnKF next-day (1200-1200 UTC) forecasts over the 2018 HWT Spring Experiment for composite reflectivity Z (a,b,c) ≥ 15 dBZ, (d,e,f) ≥ 40 dBZ, and 1-hour accumulated precipitation (g,h,i) ≥ 0.01" and (j,k,l) ≥ 0.5". The shaded region in the attributes diagrams indicate forecast skill, similar to forecasts above the black line in ROC curves.

As in section 5.3.1 (which details the construction of the probabilistic forecasts and diagrams), the performance of seasonal (spanning weekdays from 04/27/18 – 06/01/18) forecasts

is analyzed by examining attributes and relative operating characteristic (ROC) diagrams. The CAPS 10-member ensemble underpredicts large-scale storm structure ($Z \geq 15$ dBZ) for all gaussian filter standard deviation σ, except for large forecast probability ($p > 0.7$; Fig. 5.12a). This is similar to the short-term underprediction bias (Fig. 5.2a), and indicates the ensemble is unable to simulate large-scale storm structure similar to observations. The largest σ (120 km) is the most reliable for low-probability forecasts, while it loses reliability for medium-high probabilities ($0.3 < p < 0.7$). This large-scale storm structure underprediction at medium probabilities is expected, given that the largest smoother used (120 km) tends to concentrate its probabilities at lower thresholds than smaller smoothers (Fig. 5.12b). As σ increases, high probability forecasts become more reliable since these forecasts are increasingly smoothed away and concentrated closer to observations. Convective storm structure ($Z \geq 40$ dBZ) is typically underpredicted, except at small forecast probability Fig. 5.12d). This is expected, because rather than predicting large-scale storm structure that can be inferred from synoptic frontal patterns or mesoscale systems, convective cores are highly localized and more difficult to predict. As σ increases, convective storm structures forecasts are increasingly smoothed and forecast less high probability events in favor of lower probability forecasts (Fig. 5.12e), allowing probabilistic forecasts with increasing σ to better capture these highly localized cores. Therefore, convective storm structure forecasts are generally more reliable as σ increases. Still, both large-scale and convective storm structure are typically forecast with skill as the attribute lines typically fall well within the shaded region.

Both large-scale and convective storm structure display similar ROC curve behavior. As gaussian filter standard deviation σ increases, both probability of detection (POD) and probability of false detection (POFD) increase for small probability (Figs. 5.12c,f). This is expected behavior, because increasing σ not only decreases the magnitude of forecast probabilities where $p > 0$, but

also increases the areal coverage of forecast probabilities (i.e., increases p where p = 0 for smaller σ). With a spatial increase of low forecast probabilities, POD increases as ensemble forecasts are able to better match observations and record more hits (rather than misses). The low forecast probability spatial increase also increases POFD, because false alarms increase where forecasts are extended to regions with no corresponding observations. As probability increases, POD and POFD decrease with increasing σ as high probability forecasts are smoothed away. This smoothing tradeoff for the Gaussian lengths examined is not significant for either large-scale storm structure or convective storm structure, as the ROC curve is largely unchanged and well above the no-skill line regardless of σ.

Light precipitation (1-h AP $\geq$ 0.01) is typically underpredicted except for high probabilities, similar to but at a greater rate than large-scale storm structure (Fig. 5.12g). It is important to note that composite reflectivity represents the column-max of reflectivity. Although precipitation thresholds correspond well to the reflectivity thresholds chosen, light precipitation is a better indicator of total mass within each column, the sedimentation of which is prone to BMP fall speed, mass-reflectivity conversion, etc. error. Therefore, the light precipitation underprediction amplification for p < 0.8 is due to total column mass underprediction and its sedimentation to the surface. Light precipitation underprediction increases as σ increases for p > 0.2. As σ increases, the number of high probability forecasts decreases in favor of low probabilities (Fig. 5.12h), which exacerbates the ensemble's light precipitation underprediction. Similar to convective storm structure, heavy precipitation (1-h AP $\geq$ 0.5") is typically overpredicted (especially as p increases), and reliability typically increases as σ increases (Fig. 5.12j). This is reasonable because heavy precipitation events are highly localized, and therefore increasing smoothing allows the ensemble to better capture these events. Heavy precipitation high

probabilities are smoothed at a greater rate (i.e., forecast less often) than convective storm structure (Fig. 5.12k). Therefore, the ensemble has a smaller overprediction bias for total column mass and its sedimentation to the surface than maximum column mass at high forecast probabilities. However, the ensemble shows skill for both light and heavy precipitation, as the attribute lines are typically within the shaded skill region.

Similar to storm structure ROC curves, light precipitation POD and POFD increases as σ increases for low forecast probability (Fig. 5.12i). Again, this is expected because as σ increases, the spatial distribution of low probabilistic forecasts increases. Therefore, large σ will increase the number of forecasted events, which will increase both hits and false alarms, and increase POD and POFD. POD and POFD decrease as σ increases for high forecast probability, because increasing σ continuously removes high probability forecasts (resulting in less hits and false alarms). Still, each smoothing standard deviation σ shows similar forecast skill for light precipitation, as each ROC curve is similar and well above the no-skill line. Heavy precipitation ROC is largely similar and not sensitive to σ (Fig. 5.12l). Both POD and POFD are typically low, indicating that heavy precipitation is not forecast often, so no σ will greatly improve the ROC score as smoothing the forecast heavy precipitation probabilities will not significantly increase the horizontal extent of these forecasts.

5.4.2 Microphysical performance

As in section 5.3.2, the microphysical performance of the Morrison, NSSL, and Thompson schemes are analyzed for the entire 2018 NOAA HWT Spring Experiment (04/27-06/01, members 1, 4, 6, 9, 10; Table 5.1) over the seasonal analysis domain (Fig. 5.1). Storm structure is analyzed using composite reflectivity *Z* thresholds of 15 and 40 dBZ, while surface precipitation is examined using 1-hour accumulated precipitation (1-h AP) thresholds of 0.01” and 0.5”. Three supercell test

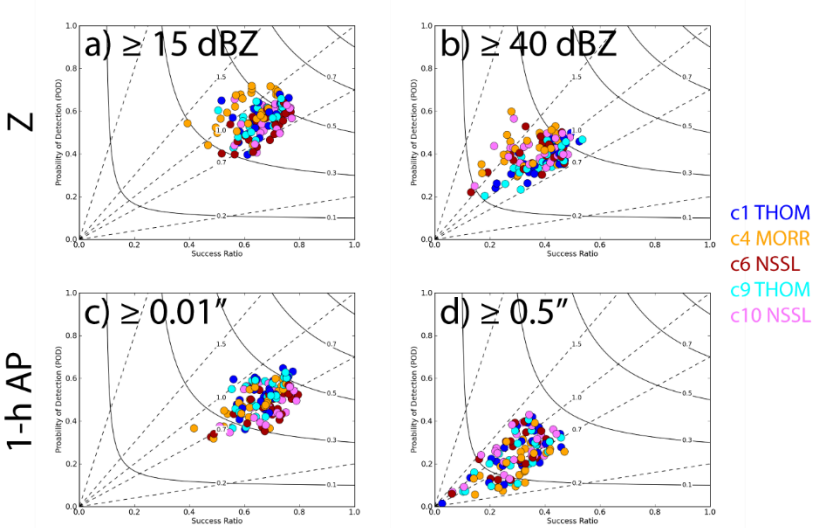

Fig. 5.13. Performance diagrams for seasonal next-day forecasts (12-36h) for composite reflectivity *Z* (a-d) ≥ 15 dBZ, (e-h) ≥ 40 dBZ and 1-h accumulated precipitation (i-l) ≥ 0.01” and (m-p) ≥ 0.5”. EnKF members 1, 4, 6, 9, and 10 are represented blue, orange, red, cyan, and pink markers.

cases (initialized at 05/02, 05/10, 05/29 0000 UTC) are additionally analyzed to determine how well each microphysics scheme can qualitatively match their observed initiation time and storm mode.

5.4.2.1 Performance diagrams

Each microphysics scheme shows moderate skill (CSI between 0.3-0.5) for next-day large-scale storm structure (Fig. 5.13a). Both the Thompson and NSSL microphysics schemes underpredict large-scale storm structure regardless of initial/boundary conditions, while the Morrison scheme is typically centered on the no-bias line. NSSL (6 and 10) and Thompson (1 and 9) members have more similar bias/CSI than members with similar initial/boundary conditions (members 9 and 10). Therefore, while large-scale storm structure bias/CSI display sensitivity to initial/boundary conditions for short-term forecasts, they are less sensitive for next-day forecasts as the effect of initial conditions lessens with increasing forecast time. For next-day convective

storm structure, the Thompson scheme shows little bias, as the cases are centered on the no-bias line (Fig. 5.13b). On the other hand, the NSSL and Morrison BMPs have several cases centered on the no-bias line, while others show a clear overprediction bias. Therefore, these schemes are simulating more and/or larger hydrometeors (i.e., small graupel advecting to a greater areal extent in the Morrison scheme, large hail in the NSSL scheme) than observations, which is consistent with their short-term forecasts. There is a noticeable decline in CSI for convective storm structure (CSI between 0.1 and 0.3) compared to large-scale storm structure. However, this is expected, as convective cells are much more isolated than overall storm structure.

There is a clear underprediction bias for next-day light precipitation (1-h AP $\geq$ 0.01") across each BMP examined (Fig. 5.13c), consistent with seasonal short-term forecasts. While this is expected for the NSSL and Thompson BMPs which display a large-scale storm structure underprediction bias, the Morrison scheme generally displays no bias for large-scale storm structure. Although the Morrison scheme neither under- nor overpredicts the maximum column mass in stratiform/light precipitation regions, it is underpredicting the total column mass and its sedimentation to the surface similar to the NSSL and Thompson BMPs. There is a similar underprediction bias for next-day heavy precipitation (1-h AP $\geq$ 0.5"), although each microphysics scheme overpredicts next-day heavy precipitation for certain cases (Fig. 5.13d). Although the Morrison and NSSL schemes either show no bias or an overprediction of convective storm structure, these schemes are still underpredicting the total column mass in these heavy precipitation regions. Microphysical skill of heavy precipitation noticeably decreases from light precipitation skill, which is expected as heavy precipitation is more sparse than light precipitation. Unlike storm structure, there is little correlation between BMP choice and precipitation bias. Therefore, even though each BMP employs different hydrometeor representation (e.g., large hail in NSSL, graupel-

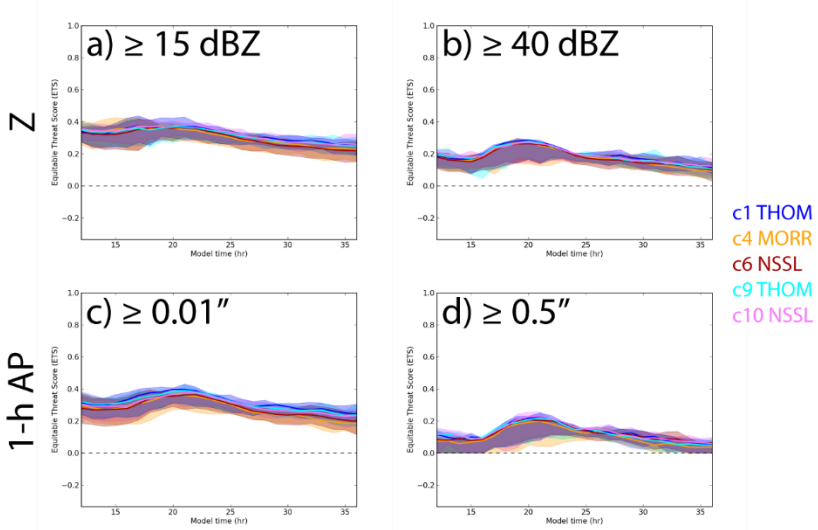

Fig. 5.14. Equitable threat scores (ETSs) over 12-36 forecast hours for seasonal forecasts of composite reflectivity *Z* (a) ≥ 15 dBZ, (b) ≥ 40 dBZ and 1-h accumulated precipitation (c) ≥ 0.01” and (e) ≥ 0.5”. EnKF members 1, 4, 6, 9, and 10 are represented blue, orange, red, cyan, and pink lines. Seasonal lines also include shading representing the 25th and 75th percentile of daily ETS over the season.

hail hybrid in Thompson), the surface accumulated precipitation (total column mass and its flux) is less sensitive to BMP choice than reflectivity structure (maximum column mass).

5.4.2.2 Equitable Threat Score (ETS)

There is little storm structure or precipitation ETS difference between the BMPs examined, indicating that storm structure and precipitation forecast skill is not sensitive to BMP choice when accounting for random forecasts (Fig. 5.14). This is not unexpected, as previous composite reflectivity and 1-h accumulated precipitation performance diagrams displayed little correlation between CSI and BMP choice. Large-scale storm structure ($Z \geq 15$ dBZ) has the largest ETS between the 17th (1700 UTC) and 22nd (2200 UTC) forecast hours. This indicates a forecast period with an increased number of observed storms, as the potential for hits (and therefore increased ETS) is larger than other forecast hours. The largest ETS spread (i.e., difference between the daily ETS 25th and 75th percentiles) tends to occur at or near this ETS peak. This could be indicative of

storm initiation variability among the BMPs, as correctly simulating the storm initiation results in high ETS, while incorrectly simulating storm initiation would increase either misses or false alarms, and decrease ETS. After the 22nd forecast hours (presumably when storm activity decreases), large-scale storm structure ETS spread narrows compared to earlier forecasts as model opportunity for hits and misses/false alarms decreases. Convective storm structure ($Z \geq 40$ dBZ) shows similar behavior as large-scale storm structure (Fig. 5.14b). ETS across BMPs noticeably peaks between the 16th and 23rd forecast hours, indicating that over the 2018 NOAA HWT Spring Experiment, observed storms had the most potential for strong, convective cores over this time period as the potential for hits is greatest. This period is also when the largest ETS spread occurs. Even though BMP ETS potential is higher due to increased hit opportunities, BMPs are also prone to more misses and false alarms that decrease ETS. Similar to absolute ETS magnitude, there is little relative ETS differences between the BMPs, as no BMP has a persistent higher or lower storm structure ETS over the next-day period.

While absolute ETS is very similar among the BMPs examined for light precipitation (1-h AP $\geq$ 0.01"; Fig. 5.14c), the Thompson scheme has a slightly larger ETS than the Morrison and NSSL schemes over much of the next-day forecast period. As each BMP contains a persistent light precipitation underprediction bias (Fig. 5.13c), the Morrison and NSSL schemes contain less hits and more misses than the Thompson scheme for light precipitation. ETS skill is largest between the 17th and 24th forecast hours, slightly lagging large-scale reflectivity. This is reasonable, as composite reflectivity represents a reflectivity column maximum, while accumulated precipitation represent column mass sedimenting to the surface. While there is no BMP that shows consistently high heavy precipitation ETS score relative to other BMPs, the Morrison scheme typically has the smallest ETS among the BMPs examined over the next-day forecast period (Fig. 5.14d). Heavy

precipitation ETS is generally largest between the 17th and 25th forecast hours. This lags convective storm structure, which is expected as large hydrometeors in the storm sediment to the surface at later times. ETS spread is also largest at this enhanced ETS time, indicating that while hit potential is largest at this time, so is miss/false alarm potential.

5.4.2.3 Fractions skill score (FSS)

Large-scale storm structure ($Z \geq 15$ dBZ) fractions skill score (FSS) is largely similar between the Thompson members and member 10 (NSSL with same initial/boundary conditions as member 9), and between the Morrison member and member 6 (NSSL BMP; Fig. 5.15a). Generally, both Thompson members have more forecast skill over next-day forecasts than the other BMPs (although the difference between the FSS of the Thompson members and member 10 (NSSL) is small). Therefore, the Thompson members and member 10 (NSSL) simulate the observed large-scale storm structure coverage better than the other two EnKF members (even though the Morrison scheme displays less next-day bias than the other BMPs), which is consistent with their short-term forecast skill. The Thompson members and member 10 produce large-scale storm structure skillful forecasts for neighborhood square half-length greater than about 18 km, while the Morrison and member 6 produce skillful forecasts for neighborhood square half-length greater than 25 km. The lack of skillful forecasts at smaller neighborhoods is expected, as the model is simulating next-day forecasts (12-36 forecast hours) and therefore is quite removed from the assimilation cycle. No BMP scheme is able to produce skillful forecasts for next-day convective storm structure ($Z \geq 40$ dBZ; Fig. 5.15b). The Thompson BMP typically produces the highest FSS for increasing neighborhood square half-length (which is also consistent with short-term forecasts), but only reaches FSSs near 0.3 at neighborhood half-length of 50 km. At this length, the other two BMPs have FSSs near 0.2. The overall FSS decline for convective storm structure compared to large-

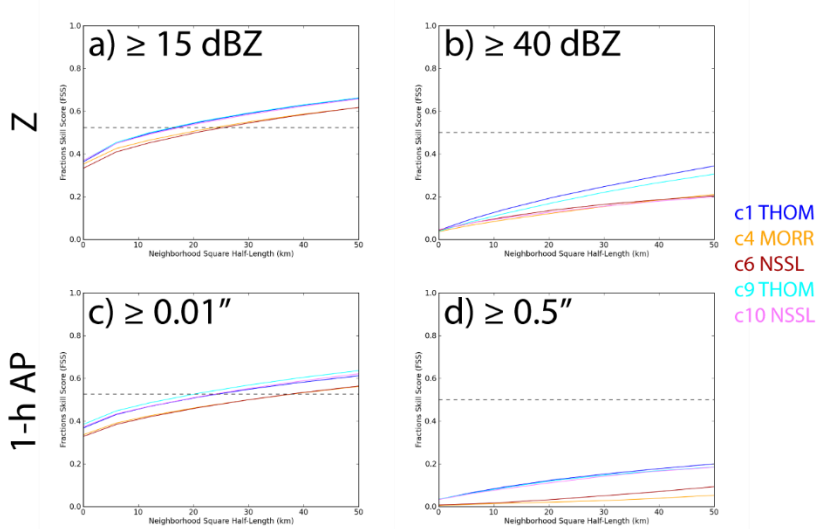

Fig. 5.15. Fractions skill score (FSS) over 12-36 forecast hours for seasonal forecasts of composite reflectivity *Z* (a) ≥ 15 dBZ, (b) ≥ 40 dBZ and 1-h accumulated precipitation (c) ≥ 0.01” and (d) ≥ 0.5” as a function of neighborhood square half-length. EnKF members 1, 4, 6, 9, and 10 are represented blue, orange, red, cyan, and pink lines. The skill line in each plot is represented by a black dashed line.

scale storm structure is expected as these spatially smaller reflectivity maxima are more isolated and more difficult to correctly simulate far from assimilation time.

No BMP is able to better simulate light precipitation (1-h AP ≥ 0.01”) that is consistent with observations relative to the other schemes, as FSS exhibits little sensitivity to microphysics parameterization (Fig. 5.15c). Similar to large-scale storm structure, the Thompson members and member 10 (NSSL) exhibit higher FSS than the Morrison member and member 6 (NSSL), but do not produce skillful forecasts until neighborhood square half-length exceeds 25 km. The remaining two members do not produce skillful forecasts until neighborhood square half-length is close to 40 km. While the NSSL BMP produces light precipitation forecasts with high skill for short-term forecasts, the BMP clearly loses skill with increasing forecast time as it is unable to match next-day observed light precipitation. No BMP is able to simulate heavy precipitation (1-h AP ≥ 0.5”) with skill for the neighborhood sizes examined (half-length up to 50 km; Fig. 5.15d). While the

decline in skill from light precipitation is expected given the more sparse nature of heavy precipitation, the highest FSS over the neighborhood sizes occurs at half-length 50 km for member 1 (Thompson) with an FSS of 0.2, well below the skillful forecast line. The NSSL BMP also displays a heavy precipitation forecast skill decline relative to the other BMPs, as it previously displayed the largest heavy precipitation skill for short-term forecasts. The Morrison scheme contains the smallest FSS for most neighborhood sizes (similar to its short-term forecast skill), which is persistently below 0.1.

5.4.2.4 Qualitative next-day supercell forecasts

Three of the four convective line test cases previously examined (05/03, 05/11, 05/30) contain supercell thunderstorms prior to the short-term forecast evaluation period (t = 0100 UTC – 0600 UTC), hereafter referred to as the 05/02, 05/10, and 05/29 test cases given their convective initiation on the previous day. Therefore, each BMP's ability to simulate the observed supercell mode and convection initiation time (which is crucial for proper risk [e.g., large hail vs. tornado] forecast) is examined. As model output is available hourly, convection initiation time is defined as the latter time in the initiation range (i.e., if convection creates a storm between 1200 – 1300 UTC, initiation time is at 1300 UTC). Because storm structure is examined using composite reflectivity (i.e., the column-max of reflectivity), storm mode is further diagnosed using 0-3 km updraft helicity (UH) thresholds of 50 (reduced from the 2-5 km UH value of 75 $m^2 s^{-2}$ as in Sobash et al. (2016)) and 125 $m^2 s^{-2}$ (to diagnose strong supercells). Each supercell examined in this chapter is initiated along a dry line, as detailed in section 5.3.2.

Observed convection initiates storms at the following times for each test case: at t = 19 h (forecast hour) in the eastern Texas panhandle for the 05/02 case (Fig. 5.16a), at t = 20 h in southeastern Wyoming for the 05/10 case (Fig. 5.16b), and at t = 22 h in the eastern Texas

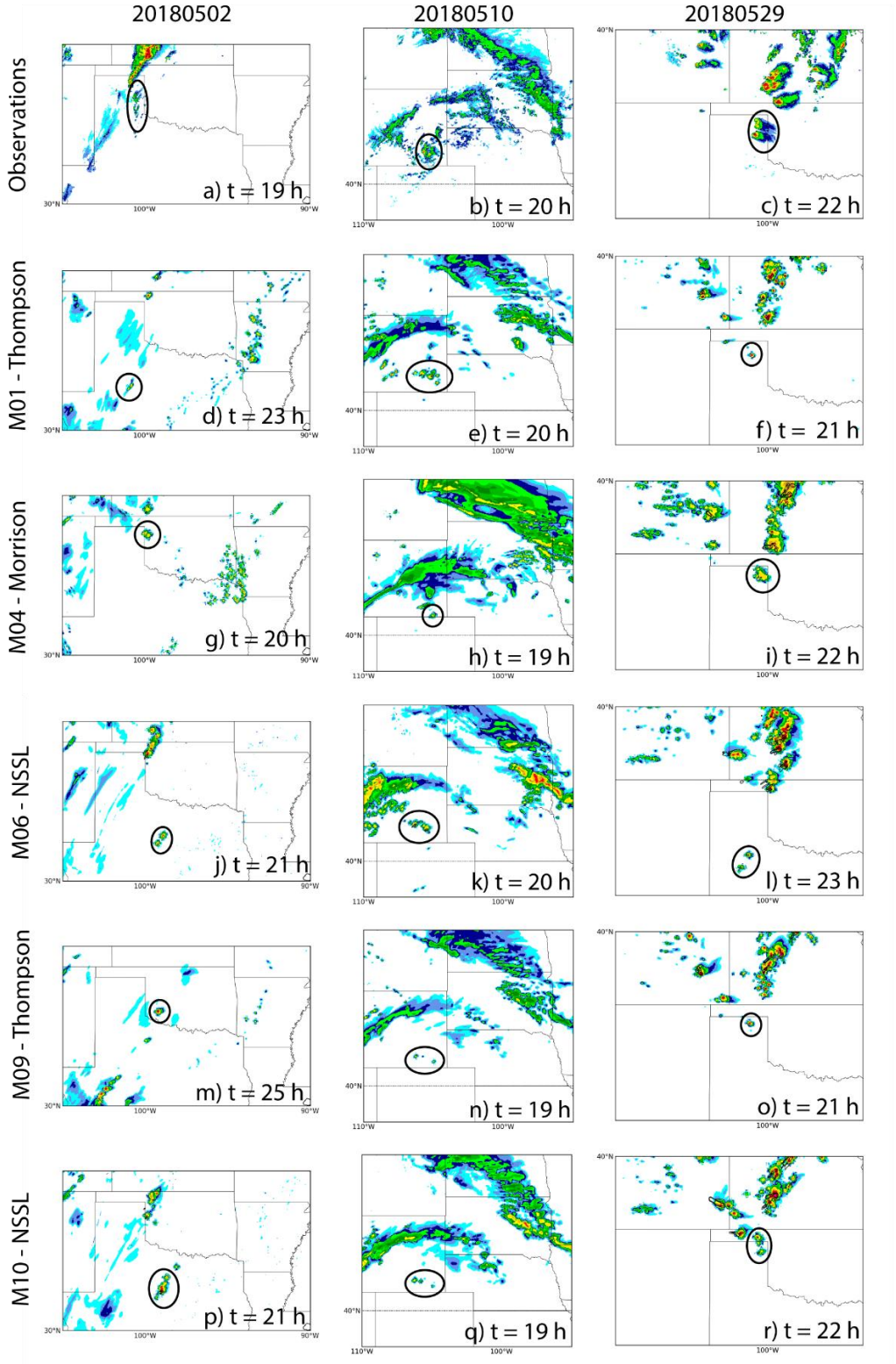

Fig. 5.16. Observed (a,b,c) convection initiation for the 05/02, 05/10, 05/29 supercell test cases, and their simulated initiation in EnKF members (d,e,f) 1, (g,h,i) 4, (j,k,l) 6, (m,n,o) 9, and (p,q,r) 10. Storms of interest are highlighted by a black circle, and initiation times are included in each subplot. 0-3 km updraft helicity black contours are overlaid at 50 $m^2 s^{-2}$ and 125 $m^2 s^{-2}$.

panhandle for the 05/29 case (Fig. 5.16c). Both Thompson members struggle to correctly initiate the 05/02 supercell until later in the forecast period (t = 23, 25 h for members 1 and 9, respectively; Figs. 5.16d,m). This is primarily due to the particular synoptic setup for this case: a stationary front north of the dry line creates a squall line, which delays storm initiation members as the squall line uses much of the atmosphere's available potential energy. While a storm is denoted in the Morrison subplot (Fig. 5.16g), it quickly merges into the squall line to the north; in fact, this member did not simulate dry line convection as in observations. The NSSL scheme simulates initiated storms closer to their observed time (t = 21 h; Figs. 5.16j,p), although they are south of the Texas panhandle and closer to central Texas. Therefore, dry line strength/position error or inaccurate available potential energy location is likely simulating storms in the wrong location for the NSSL scheme.

Each BMP improves its convection initiation time for the 05/10 and 05/29 test cases. In fact, storms are created in southeastern Wyoming for all members either at observed time (t = 20 h) or at the hour prior (t = 19 h; Figs. 5.16e,h,k,n,q). This is reasonable as the 05/10 synoptic conditions are relatively simple (i.e., stationary fronts in the 05/10 test case do not initiate deep convection). Each member typically simulates the 05/29 test case in the eastern Texas panhandle (Figs. 5.16f,i,o,r), with the exception of member 6 which initializes storms further to the south (Fig. 5.16l). This is potentially due to several reasons: the inversion cap precluding dry line convection could be the most shallow in this location, the dry line gradient could be the strongest at this point (resulting in a greater moisture flux upwards), cloud cover to the north could reduce solar flux, etc. This is not an NSSL bias, as this southern displacement is not present in member 10 (NSSL with different initial/boundary conditions). Still, each microphysics member is able to simulate convection initiation within an hour of the observed time (t = 21-23 forecast hours).

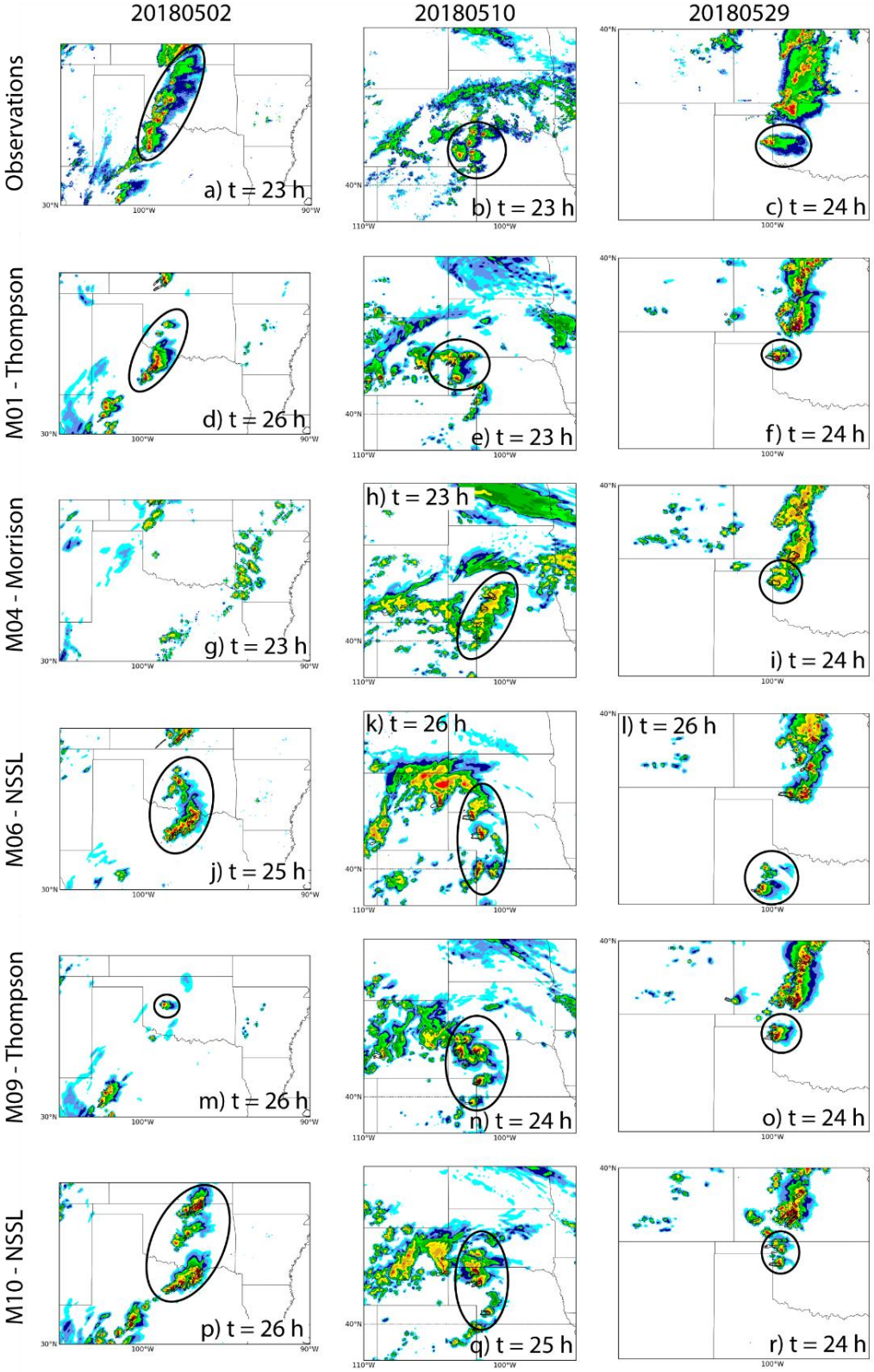

Fig. 5.17. As in Fig. 5.16, but at simulated and observed mature storm times.

Next-day storm mode is analyzed when observed and simulated storms have reached a mature state, and remain discrete (i.e., super-/multicell storms) in observations (simulated discrete cells are more prone to merging). This occurs at t = 23 h (forecast hour) as a cellular line over western Oklahoma/north Texas for the 05/02 case (Fig. 5.17a), at t = 23 h as a cellular cluster over western Nebraska for the 05/10 case (Fig. 5.17b), and at t = 24 h as a discrete supercell over the eastern Texas panhandle for the 05/29 case (Fig. 5.17c). As previously mentioned, the Morrison scheme is unable to simulate dry line convection for the 05/02 case, and therefore fails to simulate its associated mature storms (Fig. 5.17g). Storms simulated by the Thompson scheme are discrete but weak over Oklahoma (i.e., non-supercellular) for both members (Figs. 5.17d,m). Updraft helicity reveals that member 1 is likely simulating a supercell is western Texas, while observed discrete supercells are more constrained to Oklahoma. Each NSSL member simulates a convective line near the Oklahoma/Texas border with little cellular structure (Figs. 5.17j,p). However, member 10 does simulate a discrete cell with enhanced updraft helicity over central Oklahoma, although it is further to the east at a later time compared to observations. The convection initiation lag present for the case is also apparent at simulated mature storm times, which succeed observed mature storms by at least 2 hours (t = 25-26 h).

Both Thompson members simulate discrete storms with high updraft helicity (0-3 km UH > 125 m^2 s^{-2}) over western Nebraska for the 05/10 case, which are likely indicative of supercells (Figs. 5.17e,n). These simulated storms are also present near the observation time (t = 23-24 h), indicating the Thompson scheme is able to faithfully simulate supercell evolution for this case (as Thompson convective initiation time is also similar to observations). The Morrison scheme simulates storms over western/central Nebraska with high updraft helicity (> 125 m^2 s^{-2}) at the observed mature stage (t = 23 h), but these storms might be more representative of a convective

line rather than discrete supercells (although composite reflectivity represents a column max; Fig. 5.17h). NSSL storms are both discrete and contain long, intense updraft helicity swaths over western Nebraska (Figs. 5.17k,q). However, these storms lag mature states (t = 25-26 h) compared to the other microphysics schemes and observations. As convection initiation time is similar among the BMPs, the NSSL scheme is evolving these storms slower than the other BMPs examined. Both Thompson members simulate discrete storms separate from the convective line to the north in the eastern Texas panhandle/western Oklahoma for the 05/29 case (Figs. 5.17f,o). Each storm has a long swath of enhanced 0-3 km UH, indicative of supercell storms. The Morrison scheme simulates a partially-merged storm with high updraft helicity (> 125 m^2 s^{-2}) over the eastern Texas panhandle (Fig. 5.17i), but is not fully discrete as in observations. The NSSL scheme simulates discrete cells with long swaths of updraft helicity, but simulates a storm south of observations in member 6 (Figs. 5.17l,r). As member 6 initiates convection at a later time than the other members, it does not simulate a mature supercell until t = 26 h, while the other members are more consistent with observations (t = 24 h). Still, this is not a microphysical bias, but rather a bias isolated to member 6 (initial/boundary conditions), which could be incorrectly simulating dry line gradient, available potential energy, solar flux, inversion cap strength, etc.

5.5 Summary and discussion

In this chapter, we evaluate the forecast skill of the 2018 CAPS 10-member EnKF SSEF ensemble and its members containing the Morrison, NSSL, and Thompson microphysics schemes for short-term (1-6 forecast hour) and next-day (12-36 forecast hour) forecasts of composite reflectivity *Z* and 1-h accumulated precipitation (1-h AP). Short-term and next-day forecasts are validated over much of the CONUS and span weekdays from 04/27 to 06/01 (i.e., "seasonal"), while short-term forecasts are additionally verified for four convective lines. Short-term forecasts

also contain simulated brightness temperature in the 11.2 μm channel (T_B) to assess each BMP's skill in simulating cloud depth, while next-day forecasts include qualitative storm initiation and mature storm mode skill analysis for three supercell cases.

The ensemble simulates short-term storm structure and accumulated precipitation with high reliability. Large-scale storm structure ($Z \geq 15$ dBZ) and light precipitation (1-h AP ≥ 0.01") reliability typically decreases with increasing Gaussian filter standard deviation σ, indicating that these large-scale features do not require heavy smoothing. The ideal σ for convective cores ($Z \geq$ 40 dBZ) in terms of reliability decreases as forecast probability increases, as the number of high probability forecasts consequently increases (i.e., less smoothing). Heavy precipitation (1-h AP $\geq$ 0.5") is underforecasted over the ensemble such that smoothing worsens this bias. Precipitation displays a higher underprediction bias than storm structure, indicating that while maximum column mass can be either overpredicted or underpredicted, total column mass and its flux to the surface is consistently underpredicted. Storm-structure and precipitation ROC curves reveal high ensemble skill in differentiating between events and non-events, with the exception of heavy precipitation. Seasonal heavy precipitation is underforecast such that both probability of detection (POD; less hits) and probability of false detection (POFD; less false alarms) are low regardless of σ.

Both the NSSL and Thompson schemes underpredict large-scale storm structure for short-term forecasts, while the Morrison scheme overpredicts large-scale storm structure over seasonal and convective line test cases. The Morrison scheme has a smaller graupel fall velocity than the other two BMPs examined, which can advect graupel to a greater horizontal extent (i.e., increasing the horizontal coverage of $Z \geq 15$ dBZ). Still, there is minimal difference in critical success index (CSI) and therefore equitable threat score (ETS; which is CSI accounting for random forecasts)

across the short-term evaluation period, indicating that the increased hits in the Morrison scheme are unable to offset the scheme's false alarm penalties relative to the NSSL and Thompson BMPs. The Thompson scheme generally has the largest fractions skill score (FSS) for large-scale storm structure for both seasonal and convective line test cases, indicating that even though the scheme generally misses this observed storm structure, it quantitatively matches its observed horizontal distribution better than the other two schemes examined. Both the Morrison and NSSL schemes overpredict convective storm structure ($Z \geq 40$ dBZ), while the Thompson scheme underpredicts convective storm structure for both seasonal and convective line short term forecasts. The NSSL scheme is known to simulate large hail by design of its category, while the enhanced graupel advection in the Morrison scheme has the potential to melt to large drops, increasing reflectivity. Again, there is little CSI or ETS difference between the schemes, while the Thompson scheme consistently matches observed convective storm structure (in terms of FSS) better than the other schemes, as the Morrison and NSSL scheme are penalized for false alarms and struggle to produce forecasts with skill.

Each microphysics scheme underpredicts both light and heavy accumulated precipitation for short-term forecasts. EnKF members 9 (Thompson) and 10 (NSSL) have more similar bias and CSI than their counterparts with different initial/boundary conditions (members 1 and 6, respectively). Therefore, the extent of light precipitation underprediction bias seems to be more related to initial/boundary conditions rather than microphysics. Heavy precipitation displays little correlation with BMP choice. Although the Morrison scheme overpredicts storm structure and the NSSL scheme overpredicts convective storm structure, the total column mass and its sedimentation to the surface (i.e., accumulated precipitation) is underpredicted in each scheme. As precipitation CSI is similar across the BMPs, precipitation ETS exhibits little variance for the BMPs. The NSSL

scheme simulates accumulated precipitation most consistently with observations, as it displays the largest FSS for light precipitation seasonal and heavy precipitation seasonal and convective line test case short-term forecasts. Convective line light precipitation forecasts are more sensitive to the test case rather than the BMPs themselves. Even though the NSSL scheme is unable to simulate maximum column mass with more skill than the other BMPs, it is generally able to better simulate total column mass and its surface flux. All BMPs simulate light precipitation with skill, while only the NSSL scheme consistently forecasts heavy precipitation with skill (in terms of FSS).

The NSSL scheme typically overpredicts upper-level ($T_B < 225$) cloud coverage for the four convective line cases, while the Morrison and Thompson schemes underpredict these convective clouds. The NSSL scheme has a smaller cloud ice fall speed than the Morrison and Thompson schemes for ice particles larger than 20 μm, which allows these particles to get lofted higher (smaller T_B) and to a greater horizontal extent than cloud ice in the Morrison and Thompson schemes. Each microphysics member underpredicts all ($T_B < 270$) cloud coverage with small false alarm ratio, indicating that each scheme is too aggressive in suppressing convection (or its cloud formation through nucleation). While there is not a clear advantageous BMP for upper-level or all clouds in terms of ETS, member 6 typically has the smallest ETS for upper-level clouds, indicating that the false alarms in the scheme (from cloud ice's low fall velocity) is penalizing the member. The Thompson scheme typically has the largest upper-level cloud FSS, while member 6 (NSSL) has the lowest upper-level cloud FSS and the highest all cloud FSS. Again, while false alarms are skewing the predicted upper-level cloud distribution away from observations, the more aggressive convection (and cloud formation) in the NSSL scheme allows the member to better match all observed clouds. This is one factor for poor member 6 (NSSL) upper-level cloud performance but not the only factor, as this overprediction bias lessens with different initial/boundary conditions.

Next-day large-scale storm structure and light precipitation reliability increases for decreasing σ at medium probabilities (~0.3 – 0.7), while the opposite is true at low and high probabilities. Similar to short-term forecasts, the ensemble's underprediction of maximum column mass in stratiform regions (composite reflectivity) is amplified in its underprediction of total column mass (1-h accumulated precipitation). As both of these stratiform quantities are typically underpredicted, increasing σ concentrates forecasts into small probability bins, improving the ensemble's undeprediction bias. Increasing σ also removes high probability forecasts, which alleviates the ensemble's overprediction bias at these probability bins. Both next-day convective storm structure and heavy precipitation are overpredicted. Therefore, increasing σ smooths these local, isolated features and creates a more reliable forecast. Next-day storm structure and precipitation are forecast with high skill, both in attribute diagrams and ROC curves. There is a skill reduction for convective properties compared to stratiform regions, reflective of the isolated nature (and therefore, more difficult to predict) of convective storm structure/precipitation.

Next-day seasonal large-scale storm structure is generally underpredicted by the Thompson and NSSL members (consistent with short-term forecasts), while the Morrison scheme is centered on the no-bias line. Convective storm structure typically follows the biases of short-term forecasts: the Morrison and NSSL schemes display an overprediction bias while the Thompson scheme underpredicts convective storm structure, although these biases are smaller than in short-term forecasts. Therefore, as the forecast extends beyond 12 hours, storm structure biases begin losing sensitivity to BMP choice. Next-day seasonal precipitation is underforecast across all BMPs, consistent with short-term forecasts. There is little sensitivity between BMP choice and storm structure/precipitation ETS, as each scheme has very similar ETS across the forecast period. ETS peaks near the 20th forecast hour, which is also the forecast hour at which the BMPs exhibited the

most ETS variability. While there is a greater potential for storm structure/precipitation hits as storm activity increases near this time, there is also a greater potential for misses, increasing ETS variance. The Thompson scheme is able to better match observed storm structure and heavy precipitation compared to the other BMPs in terms of FSS. While this is consistent with short-term storm structure, the NSSL scheme loses precipitation forecast skill with increasing forecast hour as it has the highest FSS for short-term precipitation forecasts. Still, each BMP struggles to forecast seasonal next-day storm structure/precipitation with skill in terms of FSS, as stratiform storm-structure and precipitation require a neighborhood square half-length of at least 25 km, while no BMP is able to simulate skillful convective storm structure/precipitation forecasts (FSS below the skill line at all neighborhood sizes examined).

In terms of the next-day supercell cases examined, each BMP is able to simulate convection initiation within 1 hour of observed initiation, with the exception of the 05/02 case. Each BMP struggles with the synoptic setup for this case, as a stationary front directly north of the dry line responsible for the supercells spawns a squall line. The Morrison member misses dry line convection in favor of the squall line, while the Thompson scheme delays initiation (t = 23-25 h) well after observations (t = 19 h). The NSSL scheme delays convection initiation by two hours for this case. With the persistent convection struggles across the BMPs for the 05/02 supercell case, it is unsurprising that the members also struggle to simulate a discrete, mature supercell. In fact, only member 1 (Thompson) and 10 (NSSL) are able to simulate discrete storms with enhanced 0-3 km updraft helicity (UH > 50 m^2 s^{-2}), and is therefore not consistent across microphysical members (i.e., members 9 and 6 respectively). Each scheme simulates storms with enhanced UH for the 05/10 and 05/29 cases. NSSL and Thompson storms are more discrete and therefore likely match the observed supercell storm mode. The Morrison scheme exhibits a merged convective line

structure for each case, although since composite reflectivity is examined (i.e., reflectivity column maximum), it is difficult to truly ascertain whether there storms have merged at their mature state.

It is critical to continue to analyze and document biases in microphysics schemes, which are frequently updated, especially in the framework of ensemble forecasts. Both operational and numerical modelers should take great care in constructing mixed-microphysics ensembles, such that the biases in the microphysics schemes chosen can either attempt to be balanced through scheme selection or alleviated. For example, the NSSL and Thompson schemes (which constitute 6 of the 10 EnKF members) display a short-term large-scale storm structure underprediction bias, while all BMPs examined in this chapter (which constitute 8 of the 10 EnKF members) display a short-term light precipitation underprediction bias. Unsurpisingly, the 10-member EnKF ensemble also displays a large-scale storm structure underprediction bias, and an even larger light precipitation underprediction bias. For large-scale storm structure, the Morrison scheme has an overprediction bias; therefore, constructing the ensemble with more members using the Morrison BMP might alleviate the ensemble's large-scale storm structure underprediction bias. Relaying each BMP's persistent light precipitation underprediction bias would allow BMP authors to either tune their scheme or provide guidance for optimal tuning parameters for operational use to help alleviate this bias. Still, these are just two metrics (storm structure and precipitation) for forecast evaluation; optimal ensemble design to best mitigate individual BMP biases remains future work.

Chapter 6 Summary and Future Work

6.1 Dissertation summary

Microphysics parameterization remains a large source of numerical weather prediction (NWP) error, as it is difficult to directly verify in-situ hydrometeor distributions and their process rates. Further, while observed liquid is typically spherical or oblate with high density, ice exhibits a wide range of crystal habit, density (including mass-diameter relationship), etc. that non-linearly vary with both current thermodynamic conditions (e.g., temperature and ice supersaturation) and particle history (e.g., degree of riming, aggregation). Relative microphysical performance is driven by how each author chooses to parameterize hydrometeors in their scheme (e.g., bulk vs. bin scheme, number of hydrometeor categories, process rates, particle size distribution parameterization, etc.). In this dissertation, we evaluate and compare the performance of several flavors of microphysics parameterization: the operational one-moment Unified Model (UM), the spectral bin Hebrew University Cloud Model (HUCM) with three ice crystal categories, the partially two-moment Thompson (whose updated scheme is employed operationally by the Rapid Refresh [RAP] and High-Resolution Rapid Refresh [HRRR]), the partially two-moment Morrison, the two-moment Milbrandt-Yau (MY2) with two rimed ice categories, the two-moment National Severe Storms Laboratory (NSSL) with two rimed ice categories that prognose bulk volume, and the novel Predicted Particle Properties (P3) with ice PSDs spanning multiple mass-diameter (and therefore, ice mode) relationships. Each numerical simulation either uses the UM for test cases over Korea, or idealized or real data simulations using the Weather Research and Forecasting (WRF) model. Microphysical performance is analyzed through the utilization of simulated dual-polarization radar variables, microphysical process budget analyses, hydrometeor moment

evolution, and composite reflectivity, 1-h accumulated precipitation, and simulated brightness temperature in the 11.2 μm channel contingency tables.

Two convective cases over the Korean peninsula simulated by the UM, a Changma front and Typhoon Sanba (2012), are compared to observed S-band dual-polarization variables by employing a dual-polarization radar simulator in chapter 2 (here we note that the remaining chapters employ the WRF model). As UM microphysics are one-moment, the PSD intercept parameter N_0 must either be diagnosed (e.g., as a function of mixing ratio) or set to a constant. Consistent with previous UM analyses, the microphysics struggles to adequately simulate convection, as storm structure gaps (i.e., reflectivity) exist for both cases examined (the definitive reason for this shortcoming has yet to be determined). In terms of simulated dual-pol variables, the UM simulates both too intense reflectivity cores and reflectivity gaps, as well as too large drops inside the cores and too small drops elsewhere. This is due to the one-moment nature of the UM, specifically the rain N_0 diagnosis. In the UM, rain N_0 is inversely proportional to rain mixing ratio q_r. Therefore, as q_r decreases, rain N_0 decreases exponentially, resulting in small drops (with little reflectivity contribution) in low mass regions, and large drops (i.e., intense reflectivity cores) in high mass regions. The rain N_0 parameterization might explain reflectivity gaps in the scheme, as large rain N_0 increases evaporation, which further decreases q_r and continues to increase rain N_0.

The performance of three complex BMPs (MY2, NSSL, P3 with two ice categories [P3-2]) in an idealized supercell simulation is examined by analyzing simulated classic low-level supercell dual-pol signatures in chapter 3. Both the MY2 and NSSL schemes are able to simulate a general Z_{DR} decrease in the direction of the deep-layer storm relative mean wind vector, but neither scheme is able to fully replicate a Z_{DR} arc near the surface. Large hail in the MY2 scheme reduces Z_{DR} on the southern flank of the supercell, while the simulated forward flank in the NSSL

is too small to facilitate a classic, elongated Z_{DR} arc. Z_{DR} at the surface in the P3-2 is homogenous and displays a very weak gradient. While the scheme properly handles rimed ice size sorting (which facilitates the Z_{DR} arc), subsequent sensitivity tests reveal that the minimum rain PSD slope bound Λ is too large to allow rain size sorting. When this parameter is relaxed, the P3-2 scheme displays a raindrop and Z_{DR} gradient more consistent with observations. Both the MY2 and NSSL schemes are able to simulate large, relatively dry hail that reaches the surface and reduces Z_{DR}. Only the NSSL scheme is able to simulate this hail in the classic hail signature location, as MY2 hail has a lower fall speed than NSSL hail, and therefore is more prone to horizontal advection. The P3-2 scheme only sediments sparse (and not large) ice to the surface in its default configuration, which is due to the restrictive ice number-weighted diameter D_{ni} in the scheme. When relaxing the default D_{ni} from 2 mm to 7 and 12 mm, the scheme grows ice to a larger size and sediments it to the surface, reducing Z_{DR}. However, this large ice also unphysically sediments well into the forward flank, indicating that optimal D_{ni} remains future work.

Ice crystal and snow behavior in the spectral bin HUCM with three ice crystal categories (columns, plates, dendrites) and the bulk NSSL and Thompson schemes are analyzed in idealized supercell simulations in chapter 4. Both the HUCM and NSSL schemes prognose more ice crystal mass and number concentration than snow moments, while the Thompson scheme prognoses more snow mass than ice crystals. The Thompson snow category is constructed to contain both pristine ice and snow, both through its snow PSD construction (linear combination of exponential and gamma distribution favoring small and medium-sized snow particles) and its ice crystal to snow conversion, which is more aggressive than in the NSSL scheme. In terms of ice crystal complexity, the HUCM favors plate mass as this is the ice crystal mode to which water freezes. Both column and plate number exceed dendrite number because of their greater span of nucleation temperature.

Although snow is primarily sourced from cloud ice or freezing rain in the NSSL and Thompson schemes, both snow PSDs are illogically set up such that the snow PSD maximum is at or near particle diameter D = 0 mm. This is a bulk microphysics parameterization disadvantage, as the spectral bin HUCM snow PSDs do not contain this limitation. Still, ice crystal aggregation to snow is too aggressive in the HUCM, as snow particles in the scheme exceed mass-weighted mean diameter D_m of 10 mm near the surface, which is unexpected behavior (i.e., large surface snow) in a supercell framework.

Both short-term and next-day forecasts of storm structure and precipitation are verified with observations over the 2018 NOAA HWT Spring Experiment for the Morrison, NSSL, and Thompson schemes in chapter 5. The Morrison scheme overpredicts large-scale storm structure ($Z \geq 15$ dBZ) for short-term forecasts, while the NSSL and Thompson schemes underpredict large-scale storm structure. For next-day forecasts, the NSSL and Thompson schemes' underprediction bias lessens, while the Morrison scheme loses its bias. The Morrison scheme has a smaller graupel fall velocity than in the NSSL and Thompson schemes, which allows for greater horizontal advection (and therefore, greater large-scale storm structure coverage). In terms of convective storm structure ($Z \geq 40$ dBZ), both the Morrison and NSSL schemes overpredict these intense reflectivity regions, while the Thompson scheme underpredicts. This bias is more apparent for short-term forecasts, but still present for next-day forecasts. The NSSL scheme is known to produce large reflectivities due to its rimed ice configuration (riming graupel growing to large hail), while the graupel advected to a greater horizontal extent in the Morrison scheme can melt to large raindrops and increase Z. Each microphysics scheme displays a precipitation underprediction bias for short-term and next-day forecasts. Therefore, although the Morrison scheme overpredicts maximum column mass in stratiform regions, and the Morrison and NSSL schemes overpredict

maximum column mass in convective regions, these schemes are still underpredicting total column mass. Despite these biases, both critical success index (CSI) and equitable threat score (ETS) are very similar across the schemes, indicating that the documented biases do not significantly modify forecast skill. For short-term forecasts, the Thompson scheme typically has the highest storm structure FSS, while the NSSL scheme has the highest precipitation FSS, indicating that these schemes are able to better simulate horizontal storm structure and precipitation distributions consistent with observations. The Thompson scheme has the highest storm structure and heavy precipitation FSS for next-day forecasts, indicating that the NSSL scheme loses precipitation forecast skill as the forecast hour increases. Still, each scheme struggles to produce skillful next-day forecasts in terms of FSS.

6.2 Future work

Explicit microphysics representation has evolved considerably from early primitive cumulus parameterization and liquid-only microphysics to complex and sophisticated ice representation (i.e., number of categories, PSD configuration). Such complexity is necessary to attempt to match the multitude of observed ice modes, which drastically vary in shape with small changes to air temperature, ice supersaturation, etc. Operational numerical weather prediction (NWP) models frequently employ either one- or partially two-moment schemes. Especially relevant to deep convection, one-moment schemes cannot replicate size sorting as q and N_t cannot evolve independently of each other. Further, N_0 must be set constant or diagnosed in one-moment schemes, the parameterization of which might not be consistent across storm modes (e.g., synoptic frontal precipitaiton vs. mesoscale deep convection). Therefore, it is possible to tune N_0 in one-moment operational (and research) microphysics such that simulated hydrometeor distributions are more consistent with regional and seasonal forecasts (e.g., Great Plains summer convection

where large rain mass can imply large rain drops). A more prudent approach would be to take advantage of persistent computational power increase and expand operational NWP models to fully two-moment schemes (as is becoming more common for operational research), prognosing number concentration and allowing N_0 to evolve independently of the hydrometeor's other PSD parameters or ambient thermodynamic and microphysical information. A two-moment BMP would not require any a priori knowledge of N_0's seasonal, regional, and environment thermodynamic/microphysical conditions, nor constrain N_0 to a single value for a given q, ambient temperature, etc. This would allow for more PSD flexibility in accurately simulating the wide observed range of hydrometeor spectra, but also increase the difficulty of moment verification (i.e., number concentration within a storm) without inserting observation operator error (e.g., converting mass and number to simulated reflectivity).

Recent bulk microphysics schemes have attempted to match the observed complexity of ice by predicting the properties of ice distributions in addition to their moments. Microphysics schemes should continue to evolve and try to match the observed complexity of ice, i.e., development of a BMP that incorporates ice crystal habit prediction (and its modified deposition/sublimation rate due to the crystal's geometry), prognostic bulk volume across all ice particles (e.g., snow riming, aggregation), ice water fraction, etc. It is imperative to evolve ice representation in microphysics schemes because its associated process rates typically vary non-linearly with the physical properties of these ice particles. Observed dual-polarization radar variables could provide qualitative hydrometeor process rates for all storm modes over the CONUS throughout the year, which provides a useful tool for assessing the performance of and could guide ice property/process rate implementation in BMPs. Laboratory studies quantitatively relating ice crystal habit (i.e., particle's axes) to ambient temperature and ice supersaturation, or ice particle

density evolution through aggregation/wet growth/etc. are example future studies that would benefit and guide such an implementation into BMPs. More in-cloud measurements of ice particles (e.g., optical probes, imaging) could expand on both observed ice habit and population (PSD) evolution, their dominant process based on storm sector, and quantitative process rates. Current BMPs and those that continue to grow in ice complexity should frequently be tested across many storm modes to determine the ideal process rate/dominant process parameterization for a certain BMP for a regional/seasonal case (or to extract more accurate parameterizations from the relevantly-tuned BMP for implementation in other BMPs).

Future NWP ensembles should strive for optimal ensemble configuration. As such, it is imperative to continually test ensemble microphysical configuration (as is done in the annual Spring Forecast Experiment) to eliminate inaccurate microphysics scheme possibilities from future ensemble configurations. The addition of accurate, well-performing microphysics schemes to ensembles can not only improve the forecast analysis, but also increase ensemble spread, a desired characteristic to prevent deterministic forecasts. The introduction of additional microphysics into an ensemble also introduces the potential for individual member biases (e.g., too small storm coverage, too intense precipitation) degrading the quality of the forecast analysis. Therefore, future ensemble configuration should also focus on attempting to balance the biases relevant to the ensemble (e.g., updraft helicity/hail for deep convection storms, overall storm structure/precipitation for synoptic-scale storms) through microphysics member frequency and selection. For systematic biases (such as the underprediction of precipitation in this dissertation), machine learning can provide a useful tool to extract individual member/ensemble biases from past forecasts to guide bias correction during real-time forecasts.

Spectral bin schemes differ from bulk microphysics schemes in that the PSD is discretized into bins which are predicted, rather than moments related to the PSD. This is particularly advantageous as there is currently no consensus for observed hydrometeor PSD form representation. Further, it is currently not possible to definitively extract observed PSD information using the commonly assumed gamma PSD to compare with simulated model PSDs. Future work could include exploring the benefit of spectral bin PSD flexibility (and their more straightforward implementation of complex ice representation) in research NWP microphysics. The primary disadvantage of SBMs is that they are extremely computationally expensive (each bin is predicted rather than the entire distribution; therefore, they are unfit for operational NWP) and difficult to verify due to a lack of spectral process rates, and therefore do not currently provide improved storm simulations over BMPs. Observational PSD measurements usually span several locations across several time periods; therefore, determining PSD bin evolution is challenging without an abundance of PSD observations in different cloud sections (e.g., melting near the surface, aggregation the upper levels, rimed ice growth in the updraft, etc.). If SBM is further explored for NWP microphysics, more in-situ PSD measurements (e.g., aircraft probes, in-cloud probes) are needed to help explain how PSD bins evolve. Such a field campaign would need to include PSD samples across all types of storm modes. As an example, squall lines might be able to provide process rates for snow aggregation, but winter storms can help shed light on the full extent of snow aggregation, as well as providing more samples to derive robust process rates.

References

Abel, S. J., and I. A. Boutle, 2012: An improved representation of the raindrop size distribution for single-moment microphysics schemes. *Q. J. R. Meteorol. Soc.*, **138,** 2151–2162, doi:10.1002/qj.1949.

Adams-Selin, R. D., S. C. van den Heever, and R. H. Johnson, 2013: Impact of graupel parameterization schemes on idealized bow echo simulations. *Mon. Wea. Rev.*, **141,** 1241–1262, doi:10.1175/MWR-D-12-00064.1.

Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137,** 1805–1824, doi:10.1175/2008MWR2691.1

Aksoy, A., D. C. Dowell, and C. Snyder, 2010: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part II: Short-range ensemble forecasts. *Mon. Wea. Rev.*, **138,** 1273–1292, doi:10.1175/2009MWR3086.1

Arakawa, A., and W. H. Schubert, 1974: Interaction of a cumulus cloud ensemble with the large-scale environment, part I. *J. Atmos. Sci.*, **31,** 674–701, doi:10.1175/1520-0469(1974)031<0674:IOACCE>2.0.CO;2

Bailey, M. P., and J. Hallett, 2009: A comprehensive habit diagram for atmospheric ice crystals: Confirmation from the laboratory, AIRS II, and other field studies. *J. Atmos. Sci.*, **66,** 2888–2899, doi:10.1175/2009JAS2883.1.

Ballard, S. P., Z. Li, D. Simonin, and J.-F. Caron, 2016: Performance of 4D-var NWP-based nowcasting of precipitation at the met office for summer 2012. *Q. J. R. Meteorol. Soc.*, **142,** 472–487, doi:10.1002/qj.2665.

Benjamin, S. G., and Coauthors, 2004a: An hourly assimilation-forecast cycle: The RUC. *Mon. Wea. Rev.*, **132,** 495–518, doi:10.1175/1520-0493(2004)132,0495:AHACTR.2.0.CO;2.

Benjamin, S. G., G. A. Grell, J. M. Brown, T. G. Smirnova, and R. Bleck, 2004b: Mesoscale weather prediction with the RUC hybrid isentropic-terrain-following coordinate model. *Mon. Wea. Rev.*, **132,** 473–494, doi:10.1175/1520-0493(2004)132,0473:MWPWTR.2.0.CO;2.

Benjamin, S. G., J. M. Brown, and T. G. Smirnova, 2016a: Explicit precipitation-type diagnosis from a model using a mixed-phase bulk cloud–precipitation microphysics parameterization. *Wea. Forecasting*, **31,** 609–619, doi:10.1175/WAF-D-15-0136.1.

Benjamin, S. G., and Coauthors, 2016b: A north american hourly assimilation and model forecast cycle: The rapid refresh. *Mon. Wea. Rev.*, **144,** 1669–1694, doi:10.1175/MWR-D-15-0242.1

Berdeklis, P., and R. List, 2001: The ice crystal–graupel collision charging mechanism of thunderstorm electrification. *J. Atmos. Sci.*, **58,** 2751–2770, doi:10.1175/1520-0469(2001)058<2751:TICGCC>2.0.CO;2

Brandes, E. A., G. Zhang, and J. Vivekanandan, 2002: Experiments in rainfall estimation with a polarimetric radar in a subtropical environment. *J. Appl. Meteor.* , **41,** 674–685, doi:10.1175/1520-0450(2002)041<0674:EIREWA>2.0.CO;2.

Brandes, E. A., K. Ikeda, G. Zhang, M. Schönhuber, and R. M. Rasmussen, 2007: A statistical and physical description of hydrometeor distributions in colorado snowstorms using a video disdrometer. *J. Appl. Meteor. Climatol.*, **46,** 634–650, doi:10.1175/JAM2489.1

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

Brown, B. R., M. M. Bell, and A. J. Framback, 2016: Validation of simulated hurricane drop size distributions using polarimetric radar. *Geophys. Res. Lett.*, **43,** 910–917, doi:10.1002/2015GL067278.

Bryan, G. H., and H. Morrison, 2012: Sensitivity of a simulated squall line to horizontal resolution and parameterization of microphysics. *Mon. Wea. Rev.*, **140,** 202–225, doi:10.1175/mwr-d-11-00046.1.

Cao, Q., G. Zhang, E. Brandes, T. Schuur, A. Ryzhkov, and K. Ikeda, 2008: Analysis of video disdrometer and polarimetric radar data to characterize rain microphysics in Oklahoma. *J. Appl. Meteor. Climatol.*, **47,** 2238–2255, doi:10.1175/2008JAMC1732.1.

Chandrasekar, V., W. A. Cooper, and V. N. Bringi, 1988: Axis ratios and oscillations of raindrops. *J. Atmos. Sci.*, **45,** 1323–1333, doi:10.1175/1520-0469(1988)045<1323:ARAOOR>2.0.CO;2

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129,** 569–585, doi:10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Chen, J.-P., and T.-C. Tsai, 2016: Triple-moment modal parameterization for the adaptive growth habit of pristine ice crystals. *J. Atmos. Sci.*, **73,** 2105–2122, doi:10.1175/JAS-D-15-0220.1.

Chen, S.-H., and W.-Y. Sun, 2002: A one-dimensional time dependent cloud model. *J. Meteor. Soc. Japan*, **80,** 99–118, doi:10.2151/jmsj.80.99.

Cholette, M., H. Morrison, J. A. Milbrandt, and J. M. Thériault, 2019: Parameterization of the bulk liquid fraction on mixed-phase particles in the predicted particle properties (P3) scheme: Description and idealized simulations. *J. Atmos. Sci.*, **76,** 561–582, doi:10.1175/JAS-D-18-0278.1.

Cintineo, R., J. A. Otkin, M. Xue, and F. Kong, 2014: Evaluating the performance of planetary boundary layer and cloud microphysical parameterization schemes in convection-permitting ensemble forecasts using synthetic GOES-13 satellite observations. *Mon. Wea. Rev.*, **142,** 163–182, doi:10.1175/MWR-D-13-00143.1.

Clark, A. J., and Coauthors, 2018: The community leveraged unified ensemble (CLUE) in the 2016 NOAA/hazardous weather testbed spring forecasting experiment. *Bull. Amer. Meteor. Soc.*, **99,** 1433–1448, doi:10.1175/BAMS-D-16-0309.1

Cotton, W. R., and R. A. Anthes, 1989: *Storm and Cloud Dynamics. International Geophysics Seres vol. 44*, Academic Press, 883 pp.

Cox, G. P., 1988: Modelling precipitation in frontal rainbands. *Quart. J. Roy. Meteor. Soc*, **114,** 115–127, doi:10.1002/qj.49711447906.

Dawson, D. T., M. Xue, J. A. Milbrandt, and M. K. Yau, 2010: Comparison of evaporation and cold pool development between single-moment and multimoment bulk microphysics schemes in idealized simulations of tornadic thunderstorms. *Mon. Wea. Rev.*, **138,** 1152–1171, doi:10.1175/2009MWR2956.1.

Dawson, D. T., L. J. Wicker, E. R. Mansell, and R. L. Tanamachi, 2012: Impact of the environmental low-level wind profile on ensemble forecasts of the 4 May 2007 Greensburg, Kansas, tornadic storm and associated mesocyclones. *Mon. Wea. Rev.*, **140,** 696–716, doi:10.1175/MWR-D-11-00008.1.

Dawson, D. T., L. J. Wicker, E. R. Mansell, Y. Jung, and M. Xue, 2013: Low-level polarimetric radar signatures in EnKF analyses and forecasts of the May 8, 2003 Oklahoma City tornadic supercell: Impact of multimoment microphysics and comparisons with observation. *Adv. Meteor.*, **2013,** 818394, doi:10.1155/2013/818394.

Dawson, D. T., E. R. Mansell, Y. Jung, L. J. Wicker, M. R. Kumjian, and M. Xue, 2014: Low-level ZDR signatures in supercell forward flanks: The role of size sorting and melting of hail. *J. Atmos. Sci.*, **71,** 276–299, doi:10.1175/JAS-D-13-0118.1.

Dawson, D. T., E. R. Mansell, and M. R. Kumjian, 2015: Does wind shear cause hydrometeor size sorting?. *J. Atmos. Sci.*, **72,** 340–348, doi:10.1175/JAS-D-14-0084.1.

Delanoë, J., A. Protat, J. Testud, D. Bouniol, A. J. Heymsfield, A. Bansemer, P. R. A. Brown, and R. M. Forbes, 2005: Statistical properties of the normalized ice particle distribution. *J. Geophys. Res.*, **110,** D10201, doi:10.1029/2004JD005405.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99,** 10143–10162, doi:10.1029/94JC00572.

Fabry, F., and W. Szyrmer, 1999: Modeling of the melting layer. Part II: Electromagnetic. *J. Atmos. Sci.*, **56,** 3593–3600, doi:10.1175/1520-0469(1999)056<3593:MOTMLP>2.0.CO;2

Fan, J., and Coauthors, 2017: Cloud-resolving model intercomparison of an MC3E squall line case: Part I - Convective updrafts. *J. Geophys. Res.*, **122,** 9351–9378, doi:10.1002/2017JD026622.

Ferrier, B. S., 1994: A double-moment multiple-phase four-class bulk ice scheme. Part I: Description. *J. Atmos. Sci.*, **51,** 249–280, doi:10.1175/1520-0469(1994)051<0249:ADMMPF>2.0.CO;2.

Field, P. R., R. J. Hogan, P. R. A. Brown, A. J. Illingworth, T. W. Choularton, and R. J. Cotton, 2005: Parametrization of ice-particle size distributions for mid-latitude stratiform cloud. *Q. J. R. Meteorol. Soc.*, **131,** 1997–2017, doi:10.1256/qj.04.134.

Field, P. R., A. J. Heymsfield, and A. Bansemer, 2007: Snow size distribution parameterization for midlatitude and tropical ice clouds. *J. Atmos. Sci.*, **64,** 4346–4365, doi:10.1175/2007JAS2344.1.

Franklin, C. N., G. J. Holland, and P. T. May, 2005: Sensitivity of tropical cyclone rainbands to ice-phase microphysics. *Mon. Wea. Rev.*, **133,** 2473–2493, doi:10.1175/MWR2989.1

Fritsch, J. M., and R. E. Carbone, 2004: Improving quantitative precipitation forecasts in the warm season: A USWRP research and development strategy. *Bull. Amer. Meteor. Soc.*, **85,** 955–966, doi:10.1175/BAMS-85-7-955

Gallo, B. T., and Coauthors, 2017: Breaking new ground in severe weather prediction: The 2015 NOAA/hazardous weather testbed spring forecasting experiment. *Wea. Forecasting*, **32,** 1541–1568, doi:10.1175/WAF-D-16-0178.1

Geresdi, I., 1998: Idealized simulation of the Colorado hailstorm case: Comparison of bulk and detailed microphysics. *Atmos. Res.*, **45,** 237–252, doi:10.1016/S0169-8095(97)00079-3.

Gilmore, M. S., J. M. Straka, and E. N. Rasmussen, 2004: Precipitation uncertainty due to variations in precipitation particle parameters within a simple microphysics scheme *Mon. Wea. Rev.*, **132,** 2610–2627, doi:10.1175/MWR2810.1

Goldenberg, S. B., S. G. Gopalakrishnan, V. Tallapragada, T. Quirino, F. Marks Jr., S. Trahan, X. Zhang, and R. Atlas, 2015: The 2012 triply nested, high-resolution operational version of the hurricane weather research and forecasting model (HWRF): Track and intensity forecast verifications. *Wea. Forecasting*, **30,** 710–729, doi:10.1175/WAF-D-14-00098.1.

Hall, W. D., 1980: A detailed microphysical model within a two-dimensional dynamic framework: Model description and preliminary results. *J. Atmos. Sci.*, **37,** 2486–2507, doi:10.1175/1520-0469(1980)037<2486:ADMMWA>2.0.CO;2

Hanley, K. E., R. S. Plant, T. H. M. Stein, R. J. Hogan, J. C. Nicol, H. W. Lean, C. Halliwell, and P. A. Clark, 2015: Mixing-length controls on high-resolution simulations of convective storms. *Q. J. R. Meteorol. Soc.*, **141,** 272–284, doi:10.1002/qj.2356.

Harrington, J. Y., K. Sulia, and H. Morrison, 2013: A method for adaptive habit prediction in bulk microphysical models. Part I: Theoretical development. *J. Atmos. Sci.*, **70,** 349–364, doi:10.1175/JAS-D-12-040.1.

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, doi:10.1029/JC081i030p05353.

Hong, S.-Y., J. Dudhia, and S.-H. Chen, 2004: A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Wea. Rev.*, **132,** 103–120, doi:10.1175/1520-0493(2004)132<0103:ARATIM>2.0.CO;2.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134,** 2318–2341, doi:10.1175/MWR3199.1

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.

Houze, R. A., 2010: Clouds in tropical cyclones. *Mon. Wea. Rev.*, **138,** 293–344, doi:10.1175/2009MWR2989.1.

Hsu, W.-R., and A. H. Murphy, 1986: The attributes diagram: A geometrical framework for assessing the quality of probability forecasts. *Int. J. Forecasting*, **2,** 285–293, doi:10.1016/0169-2070(86)90048-8.

Iacobellis, S. F., G. M. McFarquhar, D. L. Mitchell, and R. C. Somerville, 2003: The sensitivity of radiative fluxes to parameterized cloud microphysics. *J. Climate*, **16,** 2979–2996, doi:10.1175/1520-0442(2003)016<2979:TSORFT>2.0.CO;2

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, doi:10.1002/qj.49711347604.

Janjic´, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122,** 927–945, doi:10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2

Johnson, D. E., P. K. Wang, and J. M. Straka, 1993: Numerical simulation of the 2 August 1981 CCOPE supercell storm with and without ice microphysics. *J. Appl. Meteor.*, **32,** 745–759, doi:10.1175/1520-0450(1993)032<0745:NSOTAC>2.0.CO;2

Johnson, M., Y. Jung, D. T. Dawson II, and M. Xue, 2016: Comparison of simulated polarimetric signatures in idealized supercell storms using two-moment bulk microphysics schemes in WRF. *Mon. Wea. Rev.*, **144,** 971–996, doi:10.1175/MWR-D-15-0233.1.

Johnson, M., Y. Jung, D. Dawson, T. Supinie, M. Xue, J. Park, and Y.-H. Lee, 2018: Evaluation of unified model microphysics in high-resolution NWP simulations using polarimetric radar observations. *Adv. Atmos. Sci.*, **35,** 771–784, doi:10.1007/s00376-017-7177-0.

Jones, T. A., P. Skinner, K. Knopfmeier, E. Mansell, P. Minnis, R. Palikonda, and W. Smith, 2018: Comparison of cloud microphysics schemes in a Warn-on-Forecast system using synthetic satellite objects. *Wea. Forecasting*, **33,** 1681–1708, doi:10.1175/WAF-D-18-0112.1.

Jung, Y., G. Zhang, and M. Xue, 2008: Assimilation of simulated polarimetric radar data for a convective storm using the ensemble Kalman filter. Part I: Observation operators for

reflectivity and polarimetric variables. *Mon. Wea. Rev.*, **136,** 2228–2245, doi:10.1175/2007MWR2083.1.

Jung, Y., M. Xue, and G. Zhang, 2010: Simulations of polarimetric radar signatures of a supercell storm using a two-moment bulk microphysics scheme. *J. Appl. Meteor. Climatol.*, **49,** 146–163, doi:10.1175/2009JAMC2178.1.

Jung, Y., M. Xue, and M. Tong, 2012: Ensemble Kalman filter analyses of the 29–30 May 2004 Oklahoma tornadic thunderstorm using one- and two-moment bulk microphysics schemes, with verification against polarimetric radar data. *Mon. Wea. Rev.*, **140,** 1457–1475, doi:10.1175/MWR-D-11-00032.1.

Jung, Y., M. Xue, and L. Grasso, 2013: Assimilation of simulated high-resolution all-sky radiance and radar data for storm-scale ensemble forecasts. *Sixth WMO Symp. on Data Assimilation*, College Park, MD, WMO, G-p20. [Available online from http://das6.umd.edu/program/Posters/uploads/Gp20-Jung_Youngsun.pdf.]

Kain, J. S., and J. M. Fritsch, 1990: A one-dimensional entraining/detraining plume model and its application in convective parameterization. *J. Atmos. Sci.*, **47,** 2784–2802, doi:10.1175/1520-0469(1990)047<2784:AODEPM>2.0.CO;2

Keith, W. D., and C. P. R. Saunders, 1989: Charge transfer during multiple large ice crystal interactions with a riming target. *J. Geophys. Res.*, **94,** 13103–13106, doi:10.1029/JD094iD11p13103.

Kessler, E., 1969: *On the Distribution and Continuity of Water Substance in Atmospheric Circulations*. *Meteor. Monogr.*, No. 32, Amer. Meteor. Soc., 84 pp.

Khain, A., A. Pokrovsky, M. Pinsky, A. Seifert, and V. Phillips, 2004: Simulation of effects of atmospheric aerosols on deep turbulent convective clouds using a spectral microphysics

mixed-phase cumulus cloud model. Part I: Model description and possible applications. *J. Atmos. Sci.*, **61,** 2963–2982, doi:10.1175/JAS-3350.1.

Khain, A. P., and I. Sednev, 1996: Simulation of precipitation formation in the eastern Mediterranean coastal zone using a spectral microphysics cloud ensemble model. *Atmos. Res.*, **55,** 77–110, doi:10.1016/S0169-8095(96)00005-1.

Khain, A. P., V. Phillips, N. Benmoshe, and A. Pokrovsky, 2012: The role of small soluble aerosols in the microphysics of deep maritime clouds. *J. Atmos. Sci.*, **69,** 2787–2807, doi:10.1175/2011JAS3649.1.

Khain, A. P., and Coauthors, 2015: Representation of microphysical processes in cloud-resolving models: Spectral (bin) microphysics versus bulk parameterization. *Rev. Geophys.*, **53,** 247–322, doi:10.1002/2014RG000468.

Kim, D.-J., 2015: Center report from KMA forecasting system operation & research. *WGNE-30*, College Park, MD. [Available online from http://polar.ncep.noaa.gov/conferences/WGNE-30/pdfs/day3/05-centre_report_KMA.pdf.]

Klemp, J. B., and R. B. Wilhelmson, 1978: Simulations of right- and left-moving storms produced through storm splitting. *J. Atmos. Sci.*, **35,** 1097–1110, doi:10.1175/1520-0469(1978)035<1097:SORALM>2.0.CO;2.

Klemp, J. B., 1987: Dynamics of tornadic thunderstorms. *Annu. Rev. Fluid Mech.*, **19,** 369–402, doi:10.1146/annurev.fl.19.010187.002101.

Knight, N. C., 1986: Hailstone shape factor and its relation to radar interpretation of hail. *J. Climate Appl. Meteor.*, **25,** 1956–1958, doi:10.1175/1520-0450(1986)025<1956:HSFAIR>2.0.CO;2.

Kong, F., and Coauthors, 2007: Preliminary analysis on the real-time storm-scale ensemble forecasts produced as a part of the NOAA hazardous weather testbed 2007 spring experiment. *22nd Conf. on Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction*, Park City, UT, Amer. Meteor. Soc., 3B.2. [Available online from https://ams.confex.com/ams/22WAF18NWP/techprogram/paper_124667.htm.]

Kong, F., 2015: 2015 CAPS Spring Forecast Experiment program plan. NOAA/NWS/SPC, 25 pp. [Available online from http://forecast.caps.ou.edu/SpringProgram2015_Plan-CAPS.pdf.]

Kong, F., 2018: 2018 CAPS Spring Forecast Experiment program plan. NOAA/NWS/SPC, 29 pp. [Available online from http://forecast.caps.ou.edu/SpringProgram2018_Plan-CAPS.pdf.]

Kumjian, M. R., and A. V. Ryzhkov, 2008: Polarimetric signatures in supercell thunderstorms. *J. Appl. Meteor. Climatol.*, **47,** 1940–1961, doi:10.1175/2007JAMC1874.1.

Kumjian, M. R., and A. V. Ryzhkov, 2009: Storm-relative helicity revealed from polarimetric radar measurements. *J. Atmos. Sci.*, **66,** 667–685, doi:10.1175/2008JAS2815.1.

Kumjian, M. R., and A. V. Ryzhkov, 2012: The impact of size sorting on the polarimetric radar variables. *J. Atmos. Sci.*, **69,** 2042–2060, doi:10.1175/JAS-D-11-0125.1.

Labriola, J., N. Snook, Y. Jung, B. Putnam, and M. Xue, 2017: Ensemble hail prediction for the storms of 10 May 2010 in south-central Oklahoma using single- and double-moment microphysical schemes. *Mon. Wea. Rev.*, **145,** 4911–4936, doi:10.1175/MWR-D-17-0039.1.

Lawson, R., B. Baker, B. Pilson, and Q. Mo, 2006: In situ observations of the microphysical properties of wave, cirrus, and anvil clouds. Part II: Cirrus clouds. *J. Atmos. Sci.*, **63,** 3186–3203, doi:10.1175/JAS3803.1.

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, doi:10.1175/1520-0450(1983)022<1065:BPOTSF>2.0.CO;2.

Liu, C., K. Ikeda, G. Thompson, R. Rasmussen, and J. Dudhia, 2011: High-resolution simulations of wintertime precipitation in the Colorado Headwaters region: Sensitivity to physics parameterizations. *Mon. Wea. Rev.*, **139,** 3533–3553, doi:10.1175/MWR-D-11-00009.1.

Liu, X., J. E. Penner, S. J. Ghan, and M. Wang, 2007: Inclusion of ice microphysics in the NCAR community atmospheric model version 3 (CAM3). *J. Climate*, **20,** 4526–4547, doi:10.1175/JCLI4264.1

Loken, E. D., A. J. Clark, M. Xue, and F. Kong, 2017: Comparison of next-day probabilistic severe weather forecasts from coarse- and fine-resolution CAMs and a convection-allowing ensemble. *Wea. Forecasting*, **32,** 1403–1421, doi:10.1175/WAF-D-16-0200.1

Lorenc, A. C., 2003: The potential of the ensemble Kalman Filter for NWP—a comparison with 4D-Var. *Quart. J. Roy. Meteor. Soc*, **129,** 3183–3203, doi:10.1256/qj.02.13.

Lynn, B. H., A. P. Khain, J. Dudhia, D. Rosenfield, A. Pokrovsky, and A. Seifert, 2005: Spectral (bin) microphysics couples with a mesoscale model (MM5). Part I: Model description and first results. *Mon. Wea. Rev.*, **133,** 44–58, doi:10.1175/MWR-2840.1.

Magono, C., and T. Nakamura, 1965: Aerodynamic studies of falling snowflakes. *J. Meteor. Soc. Japan*, **43,** 139–147, doi:10.2151/jmsj1965.43.3_139.

Magono, C., and C. W. Lee, 1966: Meteorological classification of natural snow crystals. *J. Fac. Sci. Hokkaido Univ. Ser. 7*, **2,** 321–335.

Mansell, E. R., 2010: On sedimentation and advection in multimoment bulk microphysics. *J. Atmos. Sci.*, **67,** 3084–3094, doi:10.1175/2010JAS3341.1.

Mansell, E. R., C. L. Ziegler, and E. C. Bruning, 2010: Simulated electrification of a small thunderstorm with two-moment bulk microphysics. *J. Atmos. Sci.*, **67,** 171–194, doi:10.1175/2009JAS2965.1.

Martin, G. M., D. W. Johnson, and A. Spice, 1994: The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds. *J. Atmos. Sci.,* **51,** 1823–1842, doi:10.1175/1520-0469(1994)051,1823:TMAPOE.2.0.CO;2.

Mason, S. J., 1982: A model for assessment of weather forecasts. *Aust. Meteor. Mag.*, **30,** 291–303.

Matrosov, S. Y., R. F. Reinking, R. A. Kropfli, and B. W. Bartram, 1996: Estimation of ice hydrometeor types and shapes from radar polarization measurements. *J. Atmos. Oceanic Technol.*, **13,** 85–96, doi:10.1175/1520-0426(1996)013<0085:EOIHTA>2.0.CO;2

McCumber, M., W.-K. Tao, J. Simpson, and R. S. Penc, Su-Tzai, 1991: Comparison of ice-phase microphysical parameterization schemes using numerical simulations of tropical convection. *J. Appl. Meteor.*, **30,** 985–1004, doi:10.1175/1520-0450-30.7.985

McFarquhar, G. M., H. Zhang, G. Heymsfield, J. B. Halverson, R. Hood, J. Dudhia, and F. Marks, 2006: Factors affecting the evolution of Hurricane Erin (2001) and the distributions of hydrometeors: Role of microphysical processes. *J. Atmos. Sci.,* **63,** 127–150, doi:10.1175/JAS3590.1.

McFarquhar, G. M., T. Hsieh, M. Freer, J. Mascio, and B. F. Jewett, 2015: The characterization of ice hydrometeor gamma size distributions as volumes in N0–λ–μ phase space: Implications for microphysical process modeling. *J. Atmos. Sci.,* **72,** 892–909, doi:10.1175/JAS-D-14-0011.1

McMillen, J. D., and W. J. Steenburgh, 2015: Impact of microphysics parameterizations on simulations of the 27 October 2010 Great Salt Lake-effect snowstorm. *Wea. Forecasting*, **30,** 136–152, doi:10.1175/WAF-D-14-00060.1.

Milbrandt, J. A., and M. K. Yau, 2005a: A multimoment bulk microphysics parameterization. Part I: Analysis of the role of the spectral shape parameter. *J. Atmos. Sci.*, **62,** 3051–3064, doi:10.1175/JAS3534.1.

Milbrandt, J. A., and M. K. Yau, 2005b: A multimoment bulk microphysics parameterization. Part II: A proposed three-moment closure and scheme description. *J. Atmos. Sci.*, **62,** 3065–3081, doi:10.1175/JAS3535.1.

Milbrandt, J. A., and M. K. Yau, 2006: A multimoment bulk microphysics parameterization. Part IV: Sensitivity experiments. *J. Atmos. Sci.*, **63,** 3137–3159, doi:10.1175/JAS3817.1.

Milbrandt, J. A., and R. McTaggart-Cowan, 2010: Sedimentation-induced errors in bulk microphysics schemes. *J. Atmos. Sci.*, **67,** 3931–3948, doi:10.1175/2010JAS3541.1.

Milbrandt, J. A., and H. Morrison, 2013: Prediction of graupel density in a bulk microphysics scheme. *J. Atmos. Sci.*, **70,** 410–429, doi:10.1175/JAS-D-12-0204.1.

Milbrandt, J. A., and H. Morrison, 2016: Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part III: Introduction of multiple free categories. *J. Atmos. Sci.*, **73,** 975–995, doi:10.1175/JAS-D-15-0204.1.

Mitchell, D. L., R. Zhang, and R. L. Pitter, 1990: Mass-dimensional relationships for ice particles and the influence of riming on snowfall rates. *J. Appl. Meteor.*, **29,** 153–163, doi:10.1175/1520-0450(1990)029<0153:MDRFIP>2.0.CO;2.

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, doi:10.1175/JAS3446.1.

Morrison, H., G. Thompson, and V. Tatarskii, 2009: Impact of cloud microphysics on the development of trailing stratiform precipitation in a simulated squall line: Comparison of one- and two-moment schemes. *Mon. Wea. Rev.*, **137,** 991–1007, doi:10.1175/2008MWR2556.1.

Morrison, H., and J. Milbrandt, 2011: Comparison of two-moment bulk microphysics schemes in idealized supercell thunderstorm simulations. *Mon. Wea. Rev.*, **139,** 1103–1130, doi:10.1175/2010MWR3433.1.

Morrison, H., S. A. Tessendorf, K. Ikeda, and G. Thompson, 2012: Sensitivity of a simulated midlatitude squall line to parameterization of raindrop breakup. *Mon. Wea. Rev.*, **140,** 2437–2460, doi:10.1175/MWR-D-11-00283.1.

Morrison, H., and J. A. Milbrandt, 2015: Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part I: Scheme description and idealized tests. *J. Atmos. Sci.*, **72,** 287–311, doi:10.1175/JAS-D-14-0065.1.

Morrison, H., J. A. Milbrandt, G. H. Bryan, K. Ikeda, S. A. Tessendorf, and G. Thompson, 2015: Parameterization of cloud microphysics based on the prediction of bulk ice particle properties. Part II: Case study comparisons with observations and other schemes. *J. Atmos. Sci.*, **72,** 312–339, doi:10.1175/JAS-D-14-0066.1.

Murakami, M., 1990: Numerical modeling of dynamical and microphysical evolution of an isolated convective cloud–the 19 July 1981 CCOPE cloud. *J. Meteor. Soc. Japan*, **68,** 107–128, doi:10.2151/jmsj1965.68.2_107.

Murphy, A. H., 1993: What is a good forecast? An essay on the nature of goodness in weather forecasting. *Wea. Forecasting*, **8,** 281-293.

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, doi:10.1175/1520-0450(1986)025<1037:MCOAWD>2.0.CO;2

Nakanishi, M., and H. Niino, 2006: An improved Mellor–Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Bound.-Layer Meteor.*, **119,** 397–407, doi:10.1007/s10546-005-9030-8.

Nicol, J. C., R. J. Hogan, T. H. M. Stein, K. E. Hanley, P. A. Clark, C. E. Halliwell, H. W. Lean, and R. S. Plant, 2015: Convective updraught evaluation in high-resolution NWP simulations using single-Doppler radar measurements. *Q. J. R. Meteorol. Soc.*, **141,** 3177–3189, doi:10.1002/qj.2602.

Otkin, J. A., and T. J. Greenwald, 2008: Comparison of WRF model-simulated and MODIS-derived cloud data. *Mon. Wea. Rev.*, **136,** 1957–1970, doi:10.1175/2007MWR2293.1

Palmer, R. D., and Coauthors, 2011: Observations of the 10 May 2010 tornado outbreak using OU-PRIME: Potential for new science with high-resolution polarimetric radar. *Bull. Amer. Meteor. Soc.*, **92,** 871–891, doi:10.1175/2011BAMS3125.1.

Pan, Y., M. Xue, and G. Ge, 2016: Incorporating diagnosed intercept parameters and the graupel category within the ARPS cloud analysis system for the initialization of double-moment microphysics: Testing with a squall line over south China. *Mon. Wea. Rev.*, **144,** 371–392, doi:10.1175/MWR-D-15-0008.1.

Park, H. S., A. V. Ryzhkov, D. S. Zrnić, and K.-E. Kim, 2009: The hydrometeor classification algorithm for the polarimetric WSR-88D: Description and application to an MCS. *Wea. Forecasting*, **24,** 730–748, doi:10.1175/2008WAF2222205.1.

Park, J.-S., Y. H. Lee, M. Suk, K. Nam, Y. Jung, and J. Ko, 2015a: Evaluation of UM microphysics using dual-polarised radar simulator. *37th Conference on Radar Meteorology*, Norman, OK, American Meteorological Society: Boston, MA. [Available online from https://ams.confex.com/ams/37RADAR/webprogram/Handout/Paper275815/37th-RadarMet-JSpark-Poster-final.pdf.]

Park, S., S.-H. Jung, and G. Lee, 2015b: Cross validation of TRMM PR reflectivity profiles using 3D reflectivity composite from the ground-based radar network over the Korean peninsula. *J. Hydrometeor.*, **16,** 668–687, doi:10.1175/JHM-D-14-0092.1.

Paukert, M., J. Fan, P. J. Rasch, H. Morrison, J. A. Milbrandt, J. Shpund, and A. Khain, 2019: Three-moment representation of rain in a bulk microphysics model. *J. Adv. Model. Earth Syst.*, **11,** 257– 277, doi:10.1029/2018MS001512.

Potvin, C. K., and M. L. Flora, 2015: Sensitivity of idealized supercell simulations to horizontal grid spacing: Implications for warn-on-forecast. *Mon. Wea. Rev.*, **143,** 2998–3024, doi:10.1175/MWR-D-14-00416.1.

Pruppacher, H. R., and R. L. Pitter, 1971: A semi-empirical determination of the shape of cloud and rain drops. *J. Atmos. Sci.*, **28,** 86–94, doi:10.1175/1520-0469(1971)028<0086:ASEDOT>2.0.CO;2

Pruppacher, H. R., and J. D. Klett, 1997: *Microphysics of Clouds and Precipitation*. 2nd ed. Kluwer Academic Publishers, 954 pp.

Putnam, B. J., M. Xue, Y. Jung, N. A. Snook, and G. Zhang, 2014: The analysis and prediction of microphysical states and polarimetric radar variables in a mesoscale convective system using double-moment microphysics, multinetwork radar data, and the ensemble Kalman filter. *Mon. Wea. Rev.*, **142,** 141–162, doi:10.1175/MWR-D-13-00042.1.

Putnam, B. J., M. Xue, Y. Jung, N. A. Snook, and G. Zhang, 2017a: Ensemble probabilistic prediction of a mesoscale convective system and associated polarimetric radar variables using single-moment and double-moment microphysics schemes and EnKF radar data assimilation. *Mon. Wea. Rev.*, **145,** 2257–2279, doi:10.1175/MWR-D-16-0162.1.

Putnam, B. J., M. Xue, Y. Jung, G. Zhang, and F. Kong, 2017b: Simulation of polarimetric radar variables from 2013 CAPS spring experiment storm-scale ensemble forecasts and evaluation of microphysics schemes. *Mon. Wea. Rev.*, **145,** 49–73, doi:10.1175/MWR-D-15-0415.1.

Rasmussen, R. M., and A. J. Heymsfield, 1987: Melting and shedding of graupel and hail. Part I: Model physics. *J. Atmos. Sci.*, **44,** 2754–2763, doi:10.1175/1520-0469(1987)044<2754:MASOGA>2.0.CO;2.

Reisin, T., Z. Levin, and S. Tzivion, 1996: Rain production in convective clouds as simulated in an axisymmetric model with detailed microphysics. Part I: Description of the model. *J. Atmos. Sci.*, **53,** 497–519, doi:10.1175/1520-0469(1996)053<0497:RPICCA>2.0.CO;2.

Reisner, J., R. M. Rasmussen, and R. T. Bruintjes, 1998: Explicit forecasting of supercooled liquid water in winter storms using the MM5 mesoscale model. *Quart. J. Roy. Meteor. Soc*, **124,** 1071–1107, doi:10.1002/qj.49712454804.

Reynolds, S. E., M. Brook, and M. F. Gourley, 1957: Thunderstorm charge separation. *J. Meteor.*, **14,** 426–436, doi:10.1175/1520-0469(1957)014<0426:TCS>2.0.CO;2

Roebber, P. J., 2009: Visualizing multiple measures of forecast quality. *Wea. Forecasting*, **24,** 601–608, doi:10.1175/2008WAF2222159.1.

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, doi:10.1175/2008MWR2330.1.

Rotunno, R., and J. B. Klemp, 1982: The influence of the shear-induced pressure gradient on thunderstorm motion. *Mon. Wea. Rev.*, **110,** 136–151, doi:10.1175/1520-0493(1982)110<0136:TIOTSI>2.0.CO;2.

Rutledge, S. A., and P. V. Hobbs, 1983: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. VIII: A model for the ‘‘seeder-feeder’’ process in warm-frontal rainbands. *J. Atmos. Sci.*, **40,** 1185–1206, doi:10.1175/1520-0469(1983)040<1185:TMAMSA>2.0.CO;2.

Rutledge, S. A., and P. V. Hobbs, 1984: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. XII: A diagnostic modeling study of precipitation development in narrow cold-frontal rainbands. *J. Atmos. Sci.*, **41,** 2949–2972, doi:10.1175/1520-0469(1984)041<2949:TMAMSA>2.0.CO;2.

Ryzhkov, A., M. Pinsky, A. Pokrovsky, and A. Khain, 2011: Polarimetric radar observation operator for a cloud model with spectral microphysics. *J. Appl. Meteor. Climatol.*, **50,** 873–894, doi:10.1175/2010JAMC2363.1.

Ryzhkov, A. V., T. J. Schuur, D. W. Burgess, and D. S. Zrnic, 2005a: Polarimetric tornado detection. *J. Appl. Meteor.*, **44,** 557–570, doi:10.1175/JAM2235.1.

Ryzhkov, A. V., S. E. Giangrande, and T. J. Schuur, 2005b: Rainfall estimation with a polarimetric prototype of WSR-88D. *J. Appl. Meteor.*, **44,** 502–515, doi:10.1175/JAM2213.1

Schwartz, C. S., and R. A. Sobash, 2017: Generating probabilistic forecasts from convection-allowing ensembles using neighborhood approaches: A review and recommendations. *Mon. Wea. Rev.*, **145,** 3397–3418, doi:10.1175/MWR-D-16-0400.1.

Seifert, A., and K. D. Beheng, 2001: A double-moment parameterization for simulating autoconversion, accretion and selfcollection. *Atmos. Res.*, **59–60,** 265–281, doi:10.1016/S0169-8095(01)00126-0.

Seifert, A., and K. D. Beheng, 2006: A two-moment cloud microphysics parameterization for mixed-phase clouds. Part 1: Model description. *Meteor. Atmos. Phys.*, **92,** 45–66, doi:10.1007/s00703-005-0112-4.

Seigel, R. B., and S. C. van den Heever, 2013: Squall-line intensification via hydrometeor recirculation. *J. Atmos. Sci.*, **70,** 2012–2031, doi:10.1175/JAS-D-12-0266.1.

Seliga, T. A., and V. N. Bringi, 1976: Potential use of radar differential reflectivity measurements at orthogonal polarizations for measuring precipitation. *J. Appl. Meteor. Climatol.*, **15,** 69–76, doi:10.1175/1520-0450(1976)015<0069:PUORDR>2.0.CO;2

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 125 pp. [Available online from http://www2.mmm.ucar.edu/wrf/users/docs/arw_v3.pdf.]

Skinner, P. S., and Coauthors, 2018: Object-based verification of a prototype warn-on-forecast system. *Wea. Forecasting*, **33,** 1225–1250, doi:10.1175/WAF-D-18-0020.1.

Smith, T. M., and Coauthors, 2016: Multi-radar multi-sensor (MRMS) severe weather and aviation products: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97,** 1617–1630, doi:10.1175/BAMS-D-14-00173.1.

Snook, N., M. Xue, and Y. Jung, 2011: Analysis of a tornadic mesoscale convective vortex based on ensemble Kalman filter assimilation of CASA X-band and WSR-88D radar data. *Mon. Wea. Rev.*, **139,** 3446–3468, doi:10.1175/MWR-D-10-05053.1

Snook, N., M. Xue, and Y. Jung, 2012: Ensemble probabilistic forecasts of a tornadic mesoscale convective system from ensemble Kalman filter analyses using WSR-88D and CASA radar data. *Mon. Wea. Rev.*, **140,** 2126–2146, doi:10.1175/MWR-D-11-00117.1

Snook, N., M. Xue, and Y. Jung, 2015: Multiscale EnKF assimilation of radar and conventional observations and ensemble forecasting for a tornadic mesoscale convective system. *Mon. Wea. Rev.*, **143,** 1035–1057, doi:10.1175/MWR-D-13-00262.1

Snook, N., Y. Jung, J. Brotzge, B. Putnam, and M. Xue, 2016: Prediction and ensemble forecast verification of hail in the supercell storms of 20 May 2013 *Wea. Forecasting*, **31,** 811–825, doi:10.1175/WAF-D-15-0152.1.

Sobash, R. A., J. S. Kain, D. R. Bright, A. R. Dean, M. C. Coniglio, and S. J. Weiss, 2011: Probabilistic forecast guidance for severe thunderstorms based on the identification of extreme phenomena in convection-allowing model forecasts. *Wea. Forecasting*, **26,** 714–728, doi:10.1175/WAF-D-10-05046.1.

Sobash, R. A., G. S. Romine, C. S. Schwartz, D. J. Gagne, and M. L. Weisman, 2016: Explicit forecasts of low-level rotation from convection-allowing models for next-day tornado prediction. *Wea. Forecasting*, **31,** 1591–1614, doi:10.1175/WAF-D-16-0073.1.

Stensrud, D. J., and Coauthors, 2009: Convective-scale Warn-on-Forecast system: A vision for 2020. *Bull. Amer. Meteor. Soc.*, **90,** 1487–1499, doi:10.1175/2009BAMS2795.1.

Stensrud, D. J., and Coauthors, 2013: Progress and challenges with Warn-on-Forecast. *Atmos. Res.*, **123,** 2–16, doi:10.1016/j.atmosres.2012.04.004.

Straka, J. M., and E. R. Mansell, 2005: A bulk microphysics parameterization with multiple ice precipitation categories. *J. Appl. Meteor.*, **44,** 445–466, doi:10.1175/JAM2211.1.

Straka, J. M., M. S. Gilmore, K. M. Kanak, and E. N. Rasmussen, 2005: A comparison of the conservation of number concentration for the continuous collection and vapor diffusion growth equations using one- and two-moment schemes. *J. Appl. Meteor.*, **44,** 1844–1849, doi:10.1175/JAM2314.1.

Sulia, K. J., and J. Y. Harrington, 2011: Ice aspect ratio influences on mixed-phase clouds: Impacts on phase partitioning in parcel models. *J. Geophys. Res.*, **116,** D21309, doi:10.1029/2011JD016298.

Supinie, T. A., N. Yussouf, Y. Jung, M. Xue, J. Cheng, and S. Wang, 2017: Comparison of the analyses and forecasts of a tornadic supercell storm from assimilating phased-array radar and WSR-88D observations. *Wea. Forecasting*, **32,** 1379–1401, doi:10.1175/WAF-D-16-0159.1

Takahashi, T., 1976: Hail in an axisymmetric cloud model. *J. Atmos. Sci.*, **33,** 1579–1601, doi:10.1175/1520-0469(1976)033<1579:HIAACM>2.0.CO;2.

Takahashi, T., 1978: Riming electrification as a charge generation mechanism in thunderstorms. *J. Atmos. Sci.*, **35,** 1536–1548, doi:10.1175/1520-0469(1978)035<1536:REAACG>2.0.CO;2

Tang, Y., H. W. Lean, and J. Bornemann, 2013: The benefits of the Met Office variable resolution NWP model for forecasting convection. *Meteorol. Appl.*, **20,** 417–426, doi:10.1002/met.1300.

Tao, W.-K., and J. Simpson, 1993: Goddard cumulus ensemble model. Part I: Model description. *Terr. Atmos. Oceanic Sci.*, **4,** 35–72.

Thompson, E. J., S. A. Rutledge, B. Dolan, V. Chandrasekar, and B. L. Cheong, 2014: A dual-polarization radar hydrometeor classification algorithm for winter precipitation. *J. Atmos. Oceanic Technol.*, **31,** 1457–1481, doi:10.1175/JTECH-D-13-00119.1

Thompson, E. J., S. A. Rutledge, B. Dolan, M. Thurai, and V. Chandrasekar, 2018: Dual-polarization radar rainfall estimation over tropical oceans. *J. Appl. Meteor. Climatol.*, **57,** 755–775, doi:10.1175/JAMC-D-17-0160.1

Thompson, G., R. M. Rasmussen, and K. Manning, 2004: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part I: Description and sensitivity analysis. *Mon. Wea. Rev.*, **132,** 519–542, doi:10.1175/1520-0493(2004)132<0519:EFOWPU>2.0.CO;2.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136,** 5095–5115, doi:10.1175/2008MWR2387.1.

Thompson, G., and T. Eidhammer, 2014: A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *J. Atmos. Sci.*, **71,** 3636–3658, doi:10.1175/JAS-D-13-0305.1.

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 I: Sensitivity analysis and parameter identifiability. *Mon. Wea. Rev.*, **136,** 1630–1648, doi:10.1175/2007MWR2070.1

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

II: Parameter estimation experiments. *Mon. Wea. Rev.*, **136,** 1649–1668, doi:10.1175/2007MWR2071.1

Tracton, M. S., and E. Kalnay, 1993: Operational ensemble prediction at the National Meteorological Center: Practical aspects. *Wea. Forecasting*, **8,** 379–398, doi:10.1175/1520-0434(1993)008<0379:OEPATN>2.0.CO;2

Van Den Broeke, M. S., 2016: Polarimetric variability of classic supercell storms as a function of environment. *J. Appl. Meteor. Climatol.*, **55,** 1907–1925, doi:10.1175/JAMC-D-15-0346.1.

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, doi:10.1175/1520-0469(2004)061<1596:TIOHSO>2.0.CO;2

Van Weverberg, K., N. P. van Lipzig, and L. Delobbe, 2011: The impact of size distribution assumptions in a bulk one-moment microphysics scheme on simulated surface precipitation and storm dynamics during a low-topped supercell case in Belgium. *Mon. Wea. Rev.*, **139,** 1131–1147, doi:10.1175/2010MWR3481.1

Van Weverberg, K., A. M. Vogelmann, H. Morrison, and J. A. Milbrandt, 2012: Sensitivity of idealized squall-line simulations to the level of complexity used in two-moment bulk microphysics schemes. *Mon. Wea. Rev.*, **140,** 1883–1907, doi:10.1175/MWR-D-11-00120.1.

Van Weverberg, K., and Coauthors, 2013: The role of cloud microphysics parameterization in the simulation of mesoscale convective system clouds and precipitation in the tropical western Pacific. *J. Atmos. Sci.*, **70,** 1104–1128, doi:10.1175/JAS-D-12-0104.1.

Van Weverberg, K., E. Goudenhoofdt, U. Blahak, E. Brisson, M. Demuzere, P. Marbaix, and J.-P. v. Ypersele, 2014: Comparison of one-moment and two-moment bulk microphysics for high-resolution climate simulations of intense precipitation. *Atmos. Res.*, **147-148,** 145–161, doi:j.atmosres.2014.05.012.

Varble, A., and Coauthors, 2011: Evaluation of cloud-resolving model intercomparison simulations using TWP-ICE observations: Precipitation and cloud structure. *J. Geophys. Res.*, **116,** D12206, doi:10.1029/2010JD015180.

Verrelle, A., D. Ricard, and C. Lac, 2015: Sensitivity of high-resolution idealized simulations of thunderstorms to horizontal resolution and turbulence parametrization. *Q. J. R. Meteorol. Soc.*, **141,** 433–448, doi:10.1002/qj.2363.

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, doi:10.1175/1520-0450(1991)030<1053:RATPRM>2.0.CO;2.

Wacker, U., and A. Seifert, 2001: Evolution of rain water profiles resulting from pure sedimentation: Spectral vs. parameterized description. *Atmos. Res.*, **58,** 19–39, doi:10.1016/S0169-8095(01)00081-3.

Wainwright, C. E., D. T. Dawson II, M. Xue, and G. Zhang, 2014: Diagnosing the intercept parameters of the exponential drop size distributions in a single-moment microphysics scheme and impact on supercell storm simulations. *J. Appl. Meteor. Climatol.*, **53,** 2072–2090, doi:10.1175/JAMC-D-13-0251.1.

Walko, R. L., W. R. Cotton, M. P. Meyers, and J. Y. Harrington, 1995: New RAMS cloud microphysics parameterization part I: The single-moment scheme. *Atmos. Res.*, **38,** 29–62, doi:10.1016/0169-8095(94)00087-T.

Wang, Y., C. N. Long, L. R. Leung, J. Dudhia, S. A. McFarlane, J. H. Mather, S. J. Ghan, and X. Liu, 2009: Evaluating regional cloud-permitting simulations of the WRF model for the Tropical Warm Pool International Cloud Experiment (TWP-ICE), Darwin, 2006. *J. Geophys. Res.*, **114,** D21203, doi:10.1029/2009JD012729.

Wang, Y., and X. Wang, 2017: Direct assimilation of radar reflectivity without tangent linear and adjoint of the nonlinear observation operator in the GSI-based EnVar system: methodology and experiment with the 8 May 2003 Oklahoma City tornadic supercell. *Mon. Wea. Rev.*, **145,** 1447–1471, doi:10.1175/MWR-D-16-0231.1.

Waterman, P. C., 1969: Scattering by dielectric obstacles. *Alta Freq.*, **38,** 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, doi:10.1175/1520-0493(1982)110<0504:TDONSC>2.0.CO;2.

Weisman, M. L., W. C. Skamarock, and J. B. Klemp, 1997: The resolution dependence of explicitly modeled convective systems. *Mon. Wea. Rev.*, **125,** 527–548, doi:10.1175/1520-0493(1997)125<0527:TRDOEM>2.0.CO;2.

Westbrook, C. D., R. J. Hogan, and A. J. Illingworth, 2008: The capacitance of pristine ice crystals and aggregate snowflakes. *J. Atmos. Sci.*, **65,** 206–219, doi:10.1175/2007JAS2315.1.

Wheatley, D. M., N. Yussouf, and D. J. Stensrud, 2014: Ensemble Kalman filter analyses and forecasts of a severe mesoscale convective system using different choices of microphysics schemes. *Mon. Wea. Rev.*, **142,** 3243–3263, doi:10.1175/MWR-D-13-00260.1

Wilkinson, J. M., A. N. F. Porson, F. J. Bornemann, M. Weeks, P. R. Field, and A. P. Lock, 2013: Improved microphysical parametrization of drizzle and fog for operational forecasting

using the Met Office Unified Model. *Q. J. R. Meteorol. Soc.*, **139,** 488–500, doi:10.1002/qj.1975.

Wilson, D. R., and S. P. Ballard, 1999: A microphysically based precipitation scheme for the UK meteorological office unified model. *Q. J. R. Meteorol. Soc.*, **125,** 1607–1636, doi:10.1002/qj.49712555707.

Wisner, C., H. D. Orville, and C. Myers, 1972: A numerical model of a hail-bearing cloud. *J. Atmos. Sci.*, **29,** 1160–1181, doi:10.1175/1520-0469(1972)029<1160:ANMOAH>2.0.CO;2

Woods, C. P., M. T. Stoelinga, and J. D. Locatatelli, 2007: The IMPROVE-1 storm of 1–2 February 2001. Part III: Sensitivity of a mesoscale model simulation to the representation of snow particle types and testing of a bulk microphysical scheme with snow habit prediction. *J. Atmos. Sci.*, **64,** 3927–3948, doi:10.1175/2007JAS2239.1.

Xie, S., X. Liu, C. Zhao, and Y. Zhang, 2013: Sensitivity of CAM5-simulated arctic clouds and radiation to ice nucleation parameterization. *J. Climate*, **26,** 5981–5999, doi:10.1175/JCLI-D-12-00517.1

Xue, L., and Coauthors, 2017: Idealized simulations of a squall line from the MC3E field campaign applying three bin microphysics schemes: dynamic and thermodynamic structure. *Mon. Wea. Rev.*, **145,** 4789–4812, doi:10.1175/MWR-D-16-0385.1

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. Phys.*, **82,** 139–170, doi:10.1007/s00703-001-0595-6.

Xue, M., and Coauthors, 2007: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2007 spring experiment. *22nd Conf. on*

Weather Analysis and Forecasting/18th Conf. on Numerical Weather Prediction, Park City, UR, Amer. Meteor. Soc., 3B.1. [Available online from https://ams.confex.com/ams/22WAF18NWP/techprogram/paper_124587.htm.]

Xue, M., and Coauthors, 2008: CAPS realtime storm-scale ensemble and high-resolution forecasts as part of the NOAA Hazardous Weather Testbed 2008 Spring Experiment. *24th Conf. on Severe Local Storms*, Savannah, GA, Amer. Meteor. Soc., 12.2. [Available online from https://ams.confex.com/ams/24SLS/techprogram/paper_142036.htm.]

Young, K. C., 1974: A numerical simulation of wintertime, orographic precipitation. Part I: Description of model microphysics and numerical techniques. *J. Atmos. Sci.*, **31,** 1735–1748.

Yussouf, N., E. R. Mansell, L. J. Wicker, D. M. Wheatley, and D. J. Stensrud, 2013: The ensemble kalman filter analyses and forecasts of the 8 May 2003 Oklahoma City tornadic supercell storm using single- and double-moment microphysics schemes. *Mon. Wea. Rev.*, **141,** 3388–3412, doi:10.1175/MWR-D-12-00237.1

Yussouf, N., D. C. Dowell, L. J. Wicker, K. H. Knopfmeier, and D. M. Wheatley, 2015: Storm-scale data assimilation and ensemble forecasts for the 27 April 2011 severe weather outbreak in Alabama. *Mon. Wea. Rev.*, **143,** 3044–3066, doi:10.1175/MWR-D-14-00268.1

Yuter, S. E., and R. A. Houze, Jr., 1995: Three-dimensional kinematic and microphysical evolution of Florida cumulonimbus. Part II: Frequency distributions of vertical velocity, reflectivity, and differential reflectivity. *Mon. Wea. Rev.*, **123,** 1941–1963, doi:10.1175/1520-0493(1995)123<1941:TDKAME>2.0.CO;2.

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, doi:10.1109/36.917906.

Zhang, G., M. Xue, Q. Cao, and D. Dawson, 2008: Diagnosing the intercept parameter for exponential raindrop size distribution based on video disdrometer observations: Model development. *J. Appl. Meteor. Climatol.*, **47,** 2983–2992, doi:10.1175/2008JAMC1876.1.

Zhang, J., and Coauthors, 2016: Multi-radar multi-sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities. *Bull. Amer. Meteor. Soc.*, **97,** 621–638, doi:10.1175/BAMS-D-14-00174.1.

Zhao, Q., and F. H. Carr, 1997: A prognostic cloud scheme for operational NWP models *Mon. Wea. Rev.*, **125,** 1931–1953, doi:10.1175/1520-0493(1997)125<1931:APCSFO>2.0.CO;2

Ziegler, C. L., 1985: Retrieval of thermal and microphysical variables in observed convective storms. Part I: Model development and preliminary testing. *J. Atmos. Sci.*, **42,** 1487–1509, doi:10.1175/1520-0469(1985)042<1487:ROTAMV>2.0.CO;2.

Zrnić, D. S., A. Ryzhkov, J. Straka, Y. Liu, and J. Vivekanandan, 2001: Testing a procedure for automatic classification of hydrometeor types. *J. Atmos. Oceanic Technol.*, **18,** 892–913, doi:10.1175/1520-0426(2001)018<0892:TAPFAC>2.0.CO;2