High-resolution NU-WRF simulations of a deep convective-precipitation system during MC3E: Part I: Comparisons between Goddard microphysics schemes and observations

Wei-Kuo Tao¹, Di Wu¹,², Stephen Lang¹,², Jiundar Chern¹,³, Christa Peters-Lidard⁴, Ann Fridlind⁵, and Toshihisa Matsui¹,⁶

¹Mesoscale Atmospheric Processes Laboratory,
NASA Goddard Space Flight Center, Greenbelt, MD

²Science Systems and Applications, Inc., Lanham, MD

³Goddard Earth Sciences Technology and Research Program,
Morgan State University, Baltimore, MD

⁴Hydrological Sciences Laboratory,
NASA Goddard Space Flight Center, Greenbelt, Maryland

⁵NASA Goddard Institute for Space Studies, New York, NY

⁶Earth System Science Interdisciplinary Center
University of Maryland, College Park, Maryland

J. Geophys. Rev.
Submitted February 20, 2015
Revised July 23, 2015, 2nd Revision October 15, 2015

Corresponding author address: Dr. W.-K. Tao, Code 612,
Mesoscale Atmospheric Processes Laboratory, NASA/GSFC, Greenbelt, MD 20771

Email: Wei-Kuo.Tao-1@nasa.gov
Key Points

- A new Goddard 4-class ice microphysics scheme is implemented into and modified in a regional scale model.
- Radar reflectivities and rain rate intensities are sensitive to the microphysics scheme.
- The modified 4ICE scheme produces radar structures and distributions superior to the unmodified 4ICE and the 3ICE schemes with either graupel or hail alone.

Abstract

The Goddard microphysics were recently improved by adding a 4th ice class (frozen drops/hail). This new 4ICE scheme was developed and tested in the Goddard Cumulus Ensemble model (GCE) for an intense continental squall line and a moderate, less-organized continental case. Simulated peak radar reflectivity profiles were improved in intensity and shape for both cases, as were the overall reflectivity probability distributions versus observations. In this study, the new Goddard 4ICE scheme is implemented into the regional-scale NASA Unified - Weather Research and Forecasting model (NU-WRF), modified and evaluated for the same intense squall line, which occurred during the Midlatitude Continental Convective Clouds Experiment (MC3E). NU-WRF simulated radar reflectivities, total rainfall, propagation, and convective system structures using the 4ICE scheme modified herein agree as well as or significantly better with observations than the original 4ICE and two previous 3ICE
with the modified 4ICE, the bin microphysics-based rain evaporation correction improves propagation and in conjunction with eliminating the unrealistic dry collection of ice/snow by hail can replicate the erect, narrow and intense convective cores. Revisions to the ice supersaturation, ice number concentration formula, and snow size mapping, including a new snow breakup effect, allow the modified 4ICE to produce a stronger, better organized system, more snow, and mimic the strong aggregation signature in the radar distributions. NU-WRF original 4ICE simulated radar reflectivity distributions are consistent with and generally superior to those using the GCE due to the less restrictive domain and lateral boundaries.

Index terms
3310 Clouds and cloud feedbacks
3354 Precipitation
3329 Mesoscale meteorology

Key words
Ice Microphysics, MCS, NU-WRF

1. Introduction

Many new and improved microphysical parameterization schemes have been developed over the past few decades [e.g., Ferrier, 1994; Meyers et al., 1997; Reisner et
al., 1998; Hong et al., 2004; Walko et al., 1995; Colle et al., 2004; Zhu and Zhang, 2004; Morrison et al., 2005; Straka and Mansell, 2005; Milbrandt and Yau, 2005; Morrison and Grabowski, 2008; Thompson et al., 2004, 2008; Dudhia et al., 2008, Morrison and Milbrandt, 2014, 2015; Morrison et al., 2015; and many others]. Please see Levin and Cotton [2008] and Tao and Moncrieff [2009] for a review of the microphysics used in cloud-resolving models. as well as Table 1 in Tao et al. [2011a] and Lang et al. [2014] for a brief review of microphysics parameterizations. Table 1 lists the major characteristics for a range of previously published modeling papers in terms of model used, microphysics schemes (number of ice classes, number of moments), model resolution, integration time and case(s). They include one- and two-moment bulk schemes with two or more ice classes, three-moment bulk schemes, and spectral bin microphysics schemes. Different approaches have been used to examine the performance of new schemes. One approach is to examine the sensitivity of precipitation processes to different microphysical schemes. This approach can help to identify the strength(s) and/or weakness(es) of each scheme in an effort to improve their overall performance [e.g., Ferrier et al., 1995; Straka and Mansell, 2005; Milbrandt and Yau, 2005]. Idealized simulations have also been used to test new microphysical schemes by showing their behavior in a setting that is open to simpler interpretation. In addition, another approach has been to examine specific microphysical processes (e.g., turning melting/evaporation on or off, reducing the auto-conversion rate from cloud water to rain, etc.) within one particular microphysical scheme. This approach can help to identify the dominant microphysical processes within a particular scheme (i.e., evaporation, melting of large precipitating ice particles, etc.) responsible for determining the organization and structure
of convective systems [e.g., Tao et al., 1995; Wang, 2002; Colle et al., 2005; Zhu and Zhang, 2006a; and many others]. In this paper, the main focus is on the first approach wherein the performance of several different Goddard microphysical schemes is examined; however, the sensitivity to some individual processes is also presented.

Cloud-resolving models (CRMs) are typically run at a horizontal grid spacing of 1-2 km or finer and can simulate the dynamical and microphysical processes associated with deep, precipitating atmospheric convection. One advantage of using CRMs is that they allow for explicit interactions between cloud-microphysics, radiation and surface processes. Another advantage is that each model grid is either fully clear or cloudy, so that no cloud (maximum, random) overlap assumption is required.

Simulations using the Goddard Cumulus Ensemble (GCE) model with a new 4ICE (cloud ice, snow, graupel and frozen drops/hail) scheme for an intense squall line observed over central OK during the Midlatitude Continental Convective Clouds Experiment (MC3E) and loosely organized moderate convection observed over Amazonia during the Tropical Rainfall Measuring Mission Large-Scale Biosphere-Atmosphere Experiment in Amazonia (TRMM LBA) [Lang et al., 2014] produced peak reflectivity profiles that were superior to previous iterations of the Goddard 3ICE graupel microphysics scheme [Tao et al., 2003; Lang et al., 2007, 2011] with peak intensities closer to the observed and that monotonically decreased with height also as observed. The 4ICE scheme was able to match the infrequent but relatively rare occurrence of intense echoes within the convective cores. Simulated reflectivity distributions versus height were also improved versus radar in both cases compared to the earlier 3ICE versions. The main reason for developing the 4ICE scheme was to expand the ability of
the microphysics to include more intense convection without the need to switch schemes (i.e., from 3ICE-graupel to 3ICE-hail) a priori. Furthermore, hail and graupel can occur in real weather events simultaneously. Therefore, a 4ICE scheme with both graupel and hail is useful for numerical weather prediction, especially for high-resolution prediction of severe local thunderstorms, mid-latitude squall lines and tornadoes. Current and future global high-resolution CRMs need the ability to predict/simulate a variety of weather systems from weak to intense (i.e., tropical cyclones, thunderstorms) over the entire globe; a 4ICE scheme can respond appropriately to such a variety of environmental conditions.

GCE model simulations are typically forced with the observed large-scale advective tendencies for temperature and water vapor using cyclic lateral boundary conditions [i.e., Tao et al., 2003; Moncrieff et al., 1996], as was the case for the simulations of the intense MC3E squall line in Lang et al. [2014]. However, the horizontally uniform forcing and cyclic boundaries can influence the simulated spatial structures of the squall line. Therefore, the performance of the 4ICE scheme needs to be further assessed with different types of numerical models and initial/lateral boundary conditions. Improved versions of the Goddard bulk microphysics with different options (3ICE and 4ICE) have been implemented into the NASA Unified - Weather Research and Forecasting model (NU-WRF). The major objectives of this study are to examine the performance of these different Goddard schemes in terms of their simulated radar structures, reflectivity distributions and precipitation characteristics versus observations and their vertical distributions of cloud species. Data collected during the joint NASA/DOE MC3E field campaign will be used for this study. The paper has the
following organization. Section 2 describes NU-WRF, the Goddard microphysics and a synopsis of the modifications made to it, the MC3E case, and the numerical experiments. Section 3 presents the simulation results and their evaluation versus observations, and the summary and conclusions are given in section 4.

2. NU-WRF, Goddard microphysics and case descriptions

2.1 NU-WRF

To better represent/simulate cloud-aerosol-precipitation-land surface processes and their interactions on satellite-resolvable scales (~1 km grid spacing), several physical process parameterizations developed for NASA, including CRM-based microphysics and radiation [Tao et al., 2003; Lang et al., 2007, 2011], have been implemented into WRF (versions 3.1.1 up through 3.5.1), collectively known as the NASA Unified - WRF or NU-WRF [Peters-Lidard et al., 2014], which is available to non-NASA users. These physical processes have been tested on convective systems in different environments, including a linear convective system in Oklahoma from the International H2O project (IHOP-2002) [Santanello et al., 2009], an Atlantic hurricane (Hurricane Katrina, 2005) [Tao et al., 2011a], high latitude snow events from the Canadian CloudSat CALIPSO Validation Project (C3VP) in 2007 [Shi et al., 2010; Iguchi et al., 2012a,b, 2014], a Pacific typhoon (Typhoon Morakot, 2009) [Tao et al. 2011b], and mesoscale convective systems (MCSs) in Africa [Shi et al., 2013] and the Southern Great Plains (MC3E in 2011 [Tao et al., 2013]). In addition, two other major NASA modeling components have been coupled with NU-WRF representing land surfaces (i.e., the Land Information System (LIS) [Kumar et al., 2007]) and aerosols (i.e., the WRF Chemistry Model and
Goddard Chemistry Aerosol Radiation and Transport Model (GOCART) [Chin et al., 2000, 2002, 2004]).

2.2 Goddard microphysics

Several versions of the one-moment (1M) two-class liquid and three-class ice microphysics scheme developed and coded at Goddard for the GCE model [Tao and Simpson, 1993] mainly based on Lin et al. [1983] with additional processes from Rutledge and Hobbs [1984] have been implemented into NU-WRF, including the 3ICE scheme with graupel [Tao and Simpson, 1989, 1993; Lang et al., 2007, 2011] and the 3ICE scheme with hail [McCumber et al., 1991; Tao et al., 2003]. A new 1M Goddard 4-class ice (4ICE) scheme built upon previous, successive revisions [Lang et al., 2007, 2011] to the Goddard 3ICE scheme with graupel has recently been developed [Lang et al., 2014]. This new 4ICE scheme, which requires ~20% more CPU time than the improved 3ICE graupel scheme, has prognostic variables for cloud ice, snow, graupel and hail and has just now been implemented into NU-WRF based on WRF 3.4.1. In this study, this new 4ICE scheme (referred to hereafter as the “original” 4ICE scheme) is further enhanced with the addition of a simple hail size mapping, a snow break-up effect and revisions to the pre-scribed snow size mapping, saturation adjustment scheme and number concentration formula (and is referred to hereafter as the “modified” 4ICE scheme) and then evaluated in NU-WRF versus the Goddard 3ICE graupel, 3ICE hail and the original version of the new 4ICE scheme.

2.2.1) The improved 3ICE graupel scheme
Lang et al. [2007] eliminated the dry collection of ice/snow by graupel in the
Goddard 3ICE-graupel scheme to reduce the unrealistic presence of graupel in simulated
anvils. However, radar comparisons using contoured frequency with altitude diagrams
(CFADs, Yuter and Houze 1995) revealed that the resulting snow contents were too
large. These were reduced mainly by lowering the collection efficiency of cloud water
by snow and resulted in further agreement with the radar observations. Overall, the
transfer of cloud-sized particles to precipitation-sized ice appeared to be too efficient in
the original scheme. The resulting changes lead to more realistic model precipitation ice
contents and as a consequence more physically realistic hydrometeor profiles for radiance
calculations for remote sensing applications.

The performance of the 3ICE graupel scheme was further improved by reducing
the bias in over penetrating 40-dBZ echoes at higher altitudes due mainly to excessively
large contents and/or sizes of graupel particles at those altitudes [Lang et al., 2011]. This
was achieved primarily by introducing size mappings for snow/graupel as functions of
temperature and mass. Other improvements were made and include: accounting for RH
and cloud ice size in the vapor growth of ice to snow, adding Hallett-Mossop rime
splintering, replacing the Fletcher curve for the number of active ice nuclei (IN) with the
Meyers et al. [1992] curve in the cloud ice nucleation, depositional growth and Bergeron
growth parameterizations, allowing ice supersaturations of 10% in the saturation scheme,
adding contact nucleation and immersion freezing, including cloud ice fall speeds, and
allowing for graupel/snow sublimation. These changes both reduced excessive 40 dBZ
penetrations aloft while significantly improving the overall model reflectivity CFADs.
2.2.2) The new 4ICE scheme

The improved 3ICE graupel scheme was then enhanced by the addition of hail processes and further modified to produce a new 4ICE scheme (cloud ice, snow, graupel, and frozen drops/hail) capable of simulating both intense and moderate convection [Lang et al., 2014]. Hail processes taken from the 3ICE hail scheme based on Lin et al. [1983] included hail riming, accretion of rain, deposition/sublimation, melting, shedding and wet growth. However, hail dry collection was eliminated to prevent the same excessive buildup of hail as had occurred previously with graupel [Lang et al., 2007], but hail near wet growth is allowed to efficiently collect other ice particles. Processes that freeze rain now initiate hail not graupel, and upon reaching wet growth, graupel is transferred to hail. Four new hail processes were added: wet hail accretion of graupel, rime splintering via hail riming, hail conversion to snow via deposition at colder temperatures (also applied to graupel), and hail conversion to graupel due to riming under non wet growth conditions. Besides the addition of hail processes, further modifications were made to the 3ICE processes, including the allowance of ice supersaturations of 20%, mitigating spurious evaporation/sublimation, the inclusion of a bin microphysics-based [Li et al., 2009] rain evaporation correction but with physical raindrop size constraints, and a vapor diffusivity factor. The 3ICE snow/graupel size-mapping schemes were adjusted for more stability at higher mixing rations and to increase the aggregation effect for snow. A snow density mapping [Brandes et al., 2007] was also added.

The resulting 4ICE scheme was shown to perform well not only for the intense MC3E 20 May squall line case presented in this study but also for less organized moderate convection observed during TRMM LBA. Not only were the 4ICE radar
CFADs as good or better than the previous 3ICE graupel versions, but peak reflectivity profiles even for the moderate case were superior to the 3ICE in overall intensity despite the addition of a frozen drops/hail category by realistically decreasing monotonically with height above the freezing level as observed due to the greater fall speeds of hail, which allowed higher density precipitation ice to remain near the freezing level.

2.2.3) Additional modifications to the 4ICE scheme

Several additional modifications are added to further improve the flexibility and performance of the 4ICE scheme. First, although ice supersaturations on the order of tens of percent are commonly observed [Jensen et al., 2001; Stith et al., 2002; Garrett et al., 2005], average values are much lower [Heymsfield and Miloshevich 1995; Fu and Hollars 2004]. The maximum allowable ice supersaturation was increased to 20% in the original 4ICE scheme. But, as will be shown, when applied everywhere, results in a weak convective system overall. Therefore, a new formulation is used that allows for a background ice supersaturation of 5%, which increases linearly up to a maximum of 21% as the updraft intensity increases above a background value of 2 m/s. Second, the autoconversion of cloud ice to snow (Psaut) follows a Kessler-style formulation wherein a threshold ice amount must be exceeded before the excess is converted to snow based on a specified timescale and efficiency. The previous configuration for Psaut was quite weak and although strengthened in the original 4ICE, still appears too weak and contributes to having a patchy anvil. The threshold is therefore lowered from 0.6 to 0.06 g m\(^{-3}\) to improve the homogeneity of the simulated anvils. The Meyers et al. [1992] curve for the number of active IN is also replaced by the Cooper curve [Cooper, 1986].
Being a 1M scheme, the previous ice concentration is not stored, which, using the Meyers curve, results in the number of IN decreasing as excess vapor is absorbed. In conjunction with this change, the IN concentration is constrained so the mean cloud ice particle size cannot exceed the specified minimum snow size of 100 microns.

Next, the snow mapping scheme was reconfigured to account for the effects of snow breakup via interactions with graupel and hail. Although dry collection is turned off in the original 4ICE such that graupel and hail do not collect snow, their interaction can affect the distribution of snow particle sizes. Over the years much effort [e.g., Hallett and Mossop, 1974; Hobb and Rangno, 1985; Oraltay and Hallett, 1989] has been devoted to explaining the mechanisms by which ice crystal concentrations can be observed well in excess of the background IN concentration [e.g., Mossop et al., 1968, 1972; Hobbs, 1969, 1974]. These secondary ice multiplication studies have focused primarily on the enhancement of ice crystal concentrations. Less research has been done in the area of mechanical ice breakup via ice-ice collisions [Yano and Phillips, 2011] and very little regarding the impact on the larger parent particles. In addition to the potential for interactions between various sizes of snow particles themselves to produce a self-limiting snow size distribution [Lo and Passarelli, 1982], larger aggregates are unlikely to coexist with faster falling graupel or hail particles as they would likely breakup as a result of such collisions. Vardiman [1978] performed early laboratory measurements of ice fragmentation and demonstrated the potential efficacy of mechanical fracturing especially of rimed dendrites by graupel. Griggs and Choularton [1986] also conducted a laboratory study on ice fragmentation and reported that vapor-grown dendrites are fragile and that their collision with graupel could produce a substantial number of ice crystals.
Using the laboratory data of Takahashi et al. [1995], Yano and Phillips [2011] constructed an idealized model to demonstrate that mechanical break up due to ice-ice collisions involving graupel can substantially contribute to the ice multiplication effect. Though these studies again focused on the production of ice fragments, it is apparent that such collisions would have an impact on the parent snow particle sizes. The snow mapping scheme that was carried over and modified in the original 4ICE scheme has been further modified to allow a more robust aggregation effect in the absence of graupel and hail. However, when graupel and/or hail are present, a simple scaling (Shgx) based on the local graupel/hail mixing ratio(s) is used to increase the snow intercept obtained from the mapping scheme to reduce snow particle size where:

\[ Shgx = \max(1, q_h \times 125.) + \max(1, q_g \times 25.) \] when \( q_h > 0.008 \text{ g m}^{-3}, q_g > 0.04 \text{ g m}^{-3} \)

and \( q_h \) and \( q_g \) are the hail and graupel mixing ratios, respectively. This formulation produces convective snow sizes that remain small but allows anvil snow sizes to become large using a common snow mapping and thus improves the effective mapping in each region rather than utilize a single compromise mapping for both.

Finally, a simple hail mapping scheme was introduced. Lang et al. [2014] demonstrated the performance of the original 4ICE scheme for both moderate and intense convection; however, because the scheme still retained the use of a fixed intercept, a series of experiments was conducted for each case using different hail intercepts (i.e., equivalent to smaller-, medium-, and larger-sized hail). It was found that smaller hail performed the best for the moderate case, while medium hail performed best for the
intense. As noted in Lang et al. [2014], it is not optimal to have to choose the hail intercept for each case *a priori*. Therefore, a simple hail mapping scheme has been devised based on the peak hail profiles from the moderate and intense cases in the Lang *et al.* [2014] study. In the mapping, a starting intercept appropriate for smaller hail (i.e., $0.240 \text{ cm}^{-4}$) is scaled down (i.e., hail size increases) as the hail mixing ratio increases beyond a minimum threshold. It then reaches a minimum value (i.e., $0.0048 \text{ cm}^{-4}$) upon reaching a maximum threshold beyond which it no longer changes. Figure 1 shows the two thresholds as a function of the local (i.e., in cloud not environmental) temperature.

2.3 *The 20 May 2011 squall line*

MC3E was a joint field campaign between the DOE ARM Climate Research Facility and NASA’s Global Precipitation Measurement (GPM) mission ground validation (GV) program [Petersen *et al.*, 2009]. It took place in central Oklahoma from 22 April to 6 June 2011. Some of its major objectives involve the use of high-resolution CRMs in precipitation science and include: (1) testing the fidelity of CRM simulations via intensive statistical comparisons between simulated and observed cloud properties and latent heating (LH) fields for a variety of case types, (2) establishing the limits of CRM space-time integration capabilities for quantitative precipitation estimates, and (3) supporting the development and refinement of physically-based GPM microwave imager (GMI), dual-frequency precipitation radar (DPR), and DPR-GMI combined retrieval algorithms using ground-based observations, aircraft measurements, airborne radar and radiometer, and CRM simulations. The focus of this study will be the intense squall line case presented in Lang *et al.* [2014].
On 20 May 2011 a deep, upper-level low over the central Great Basin moved across the central and southern Rockies and into the central and northern Plains. A surface low pressure center in southeastern Colorado drew warm, moist air from the southern Plains to a warm front over Kansas, while a dry line extended southward from the Texas/Oklahoma Panhandle. As a result, several convective lines formed over the Great Plains and propagated eastward. The northern portion of a long convective line began to enter the MC3E sounding network around 07 UTC 20 May and by 09 UTC had merged with ongoing convection near the KS-OK border to form a more intense convective segment with a well-defined trailing stratiform region that then propagated through the network between 09 and 12 UTC. The convection along the leading edge of this intense squall line exited the network around 13 UTC leaving behind a large area of stratiform rain. For further details see Lang et al. [2014]. This case was also simulated with NU-WRF by Tao et al. [2013] as part of a post mission case study to examine the performance of the NU-WRF, real-time forecasts during MC3E. They found propagating precipitation features and their associated cold-pool dynamics were important for the diurnal variation of precipitation. Terrain effects were also found to be important during initial MCS development with surface fluxes and radiation processes having only a secondary effect for short-term simulations. Differences between Tao et al. [2013] and this study include the model configuration (18, 6 and 2 km vs. 9, 3 and 1 km grid spacing) and initial conditions (North American Regional Reanalysis or NARR vs. Final Analysis by GFS or FNL).

2.4 Model set-up
Figure 2 shows the model grid configuration, which includes an outer domain and two inner-nested domains having a horizontal grid spacing of 9, 3 and 1 km using 524×380×61, 673×442×61, 790×535×61 grid points, respectively. Time steps of 18, 6 and 2 seconds were used in these nested grids, respectively. The Grell-Devenyi cumulus parameterization scheme [Grell and Devenyi, 2002] was used for the outer grid (9 km) only. For the inner two domains (3- and 1-km), the convective scheme was turned off. The PBL parameterization employed the Mellor-Yamada-Janjic [Mellor and Yamada, 1992] Level-2 turbulence closure model through the full range of atmospheric turbulent regimes. The scheme was coded/modified by Dr. Janjic for the NCEP Eta model.

The Goddard broadband two-stream (upward and downward fluxes) approach was used for the short- and long-wave radiative flux and atmospheric heating calculations [Chou and Suarez, 1999, 2001] and its explicit interactions with clouds (microphysics). Model terrain is smoothed from the 5-m (~5-km), 2-m (~4-km) and 30-second (~0.9 km) USGS terrain database for the three nested domains, respectively. Simulations start at 00 UTC 20 May 2014 and are integrated for 48 hours. Initial conditions are from the GFS-FNL (Global Forecast System Final global gridded analysis archive) as are the lateral boundary conditions, which are updated every 6 hours.

2.5 Numerical experiments

Four main numerical experiments and one sensitivity test are conducted for the 20 May 2011 MC3E case using various versions of the Goddard microphysics: 3ICE-graupel [or Graupel, Lang et al., 2007, 2011], 3ICE-hail [or Hail, McCumber et al. 1990; Tao et al. 2003], the original 4ICE scheme [or 4ICE_v0, Lang et al. 2014], the modified
4ICE scheme [4ICE], and the modified 4ICE scheme but without the rain evaporation correction [4ICE_nec]. Table 2 lists the numerical experiments used in this study.

3. Results

3.1 Radar structures, reflectivity comparisons, and vertical velocity characteristics

The National Mosaic and Multi-Sensor QPE (NMQ) system is a multi-radar, multi-sensor system, ingesting base-level data from all available radars (NEXRAD, Canadian Radar, TDWR, and gap radars) at any given time; it performs quality control and combines reflectivity observations from individual radars onto a unified 3D Cartesian frame. The data have a spatial and temporal resolution of 1 km and 5 min, respectively [Zhang et al., 2011]. Simulated radar reflectivities are calculated from model rain, snow, graupel and hail contents following the model inverse exponential size distributions and accounting for all size and density mappings and assuming a Rayleigh approximation using the formulation of Smith et al. [1975] and Smith [1984]. Figure 3 shows horizontal cross sections of observed and simulated composite radar echoes for the 20 May MCS at 10 UTC. NEXRAD data (Fig. 3a) show a well-developed squall line with an intense, slightly bowed convective leading edge and prominent trailing stratiform region separated by a distinct transition region, extending southwestward from the Kansas-Oklahoma border down into central West Texas. A vertical cross section taken normal to the line (Fig. 4a) shows a classic continental uni-cellular squall line structure [Rutledge et al., 1998; Johnson and Hamilton, 1988; see review by Houze, 1997] with deep, erect leading convective cell(s) followed by a wide trailing stratiform region, featuring a distinct high radar reflectivity bright band near the melting level separated from the
convective core(s) by a transition area with a less prominent bright band. Each of the four NU-WRF simulations captures the basic squall line organization; however, there are several notable differences between the schemes, namely variations in the continuity and intensity of their leading edge convection as well as the size and consistency of their stratiform areas, as well as various discrepancies with the observations.

The Graupel scheme (Fig. 3b) produces a wide trailing stratiform region as observed but with too many moderate reflectivities and leading edge reflectivities that are too weak, although continuous. Without hail, the Graupel scheme simply cannot match the intense radar returns associated with such large, solid, dense, ice particles while too much moderately-falling graupel is transported rearward into the stratiform region. A vertical cross section (Fig. 4b) confirms the weak leading edge reflectivities as well as a tendency for peak values there to be elevated due to moderately falling graupel being easily carried aloft (L2014). Stratiform echoes are maximized near and below the melting level but are slightly too intense and with more vertical undulations above the freezing level than observed. The system is well organized, but the convective leading edge propagates too quickly across central Oklahoma (Fig. 5a) as a result of a very strong surface cold pool due largely to excessive rain evaporation, a typical problem with 1M schemes [Morrison et al. 2009].

The Hail scheme (Fig. 3c) captures the intense nature of the leading convective reflectivities, but their intensity tends to be localized, as the overall convective leading edge is not as continuous as with the observed MCS. This also results in a somewhat disjointed stratiform region though the composite magnitudes appear to the match the observed rather well. However, a vertical cross section through the Hail MCS (Fig. 4c)
shows the highest reflectivities in the stratiform region are elevated in the upper
troposphere not near or below the melting layer as observed, which is completely
unrealistic. This is due to having a fixed snow intercept (i.e., no size mapping) such that
size is maximized with the mass. Also, the convective structure tends to be more multi-
cellular than was observed. Overall the system lacks organization as a result of a weak
and fragmented surface cold pool (Fig. 5b). Though the Hail scheme does not have a rain
evaporation correction, its stratiform region is too small and the structure too poor to
generate a large cold pool. This is due in part to the inclusion of dry collection whereby
hail collects ice and snow too efficiently, reducing their transport into the stratiform
region.

The original 4ICE scheme (Fig. 3d) is somewhat of a blend between the Graupel
and Hail schemes with locally intense reflectivities similar to the Hail scheme but with a
somewhat more continuous convective line though not as organized as the Graupel
scheme and a stratiform region that is slightly more coherent than the Hail scheme but
not as well developed as the Graupel scheme. A vertical cross section (Fig. 4d) through
the 4ICE_v0 MCS, however, shows there are some notable improvements in its simulated
structures relative to the 3ICE schemes. First, in terms of the leading edge convection, its
simulated convective reflectivity core(s) are closer to the observed overall, being both
erect due to the inclusion of the rain evaporation correction [Li et al. 2009] and narrow
and intense due to the inclusion of hail in conjunction with eliminating dry collection.
Second, in terms of the trailing stratiform region, the scheme produces a broad, well-
developed stratiform area with a more vertically-stratified (i.e., more horizontally
uniform) radar structure with values maximized near and below the melting level and
overall values that closely match the observed. This is a result of the revisions to the
snow mapping, namely an enhanced aggregation effect via a greater temperature
dependency. The 4ICE_v0 convective and especially stratiform features are much closer
to the observed than the 3ICE schemes. Overall the original 4ICE scheme has the
essential elements but appears to lack the overall intensity and organization of the
observed system. This is confirmed by the extremely weak surface cold pool (Fig. 5c)
and corresponding lack of forward propagation.

In contrast, the modified 4ICE scheme (Fig. 3e) produces a more organized
system with a longer, more continuous line of leading convection that has a slightly
bowed structure and a broader stratiform area with a more defined transition region
separating it from the leading convection. All of which are in good or better agreement
with the observed. Based on the series of specific modifications made to the 4ICE
scheme (individual results not shown), the main reason for having a stronger, more
organized system is the change in the saturation formulation from a global 20% ice
supersaturation value to one that varies starting at just 5%. The smaller value allows
more water vapor to be converted to ice, which releases more heat aloft over a broad area.
A vertical cross section through the modified 4ICE MCS (Fig. 4e) shows features that are
generally similar to the original 4ICE except for a more pronounced transition zone aloft
as a result of introducing the snow breakup effect and even more vertically-stratified
echoes with a sharper vertical gradient in the trailing stratiform region due to an even
larger aggregation effect in the prescribed snow size mapping scheme. The modified
4ICE scheme produces a moderate intensity cold pool (Fig. 5d) relative to the other
schemes, and its simulated convective line is closest to the observed propagation,
especially over central Oklahoma. Without the rain evaporation correction, the surface cold pool becomes stronger, causing the center of the simulated convective line to propagate too fast (Fig. 5e) and develop an excessive bow structure over central Oklahoma (Fig. 3f) similar to the Graupel scheme (Fig. 3b). This also results in the leading convective cells having a more tilted structure at low levels (Fig. 4f). Overall, the modified 4ICE scheme with the rain evaporation correction best captures more of the observed features and has the most realistic structures compared to the other schemes.

In addition to the structural comparisons, which provide the necessary context, a combination of time varying and comprehensive statistical quantities are crucial for an overall evaluation. Figure 6 shows time-height cross sections of maximum reflectivity both observed by NEXRAD radar and simulated for each of the four NU-WRF simulations within the innermost model domain from 06 to 12 UTC 20 May 2011. Over this period within the analysis domain (shown in Fig. 2), peak reflectivities associated with this intense squall line frequently exceeded 50 dBZ up to 10 km and 60 dBZ below about 7 km with 40 dBZ echoes reaching as high as 15 km. Maximum echo values do fluctuate but overall are fairly steady with only a slight decrease after ~08 UTC (Fig. 6a).

The Graupel scheme (Fig. 6b) greatly underestimates the peak 50 to 60 dBZ intensities of the observed squall line above the freezing level and simply cannot produce such intense echoes due to the smaller size and lower density of graupel. The Hail scheme does produce intense reflectivities (Fig. 6c) due to large-sized hail as a result of its fixed low intercept value (i.e., 0.01 cm$^{-4}$) for hail with peak values near 70 dBZ at ~3-4 km and 55 dBZ values regularly reaching above 12 km, which are more intense near the melting level and above 10 km than was observed. The scheme also produces an
unrealistic elevated secondary maximum near 11 km. In contrast, the original 4ICE simulation (Fig. 6d) produces peak reflectivity values that decrease monotonically with height as observed. Its maximum intensities are fairly good at most levels though somewhat too weak at the lowest levels and too strong near the freezing level. It also uses a fixed but slightly larger intercept for hail (i.e., 0.02 cm$^{-4}$) based on the results from L2014 for this case. Peak reflectivities for the modified 4ICE scheme (Fig. 6e) are fairly similar to 4ICE_v0 in general but slightly weaker near the freezing level, where they are slightly too weak as opposed to slightly too strong for 4ICE_v0. However, although rather simple, the new 4ICE hail mapping performs comparably without having to choose the appropriate hail intercept a priori for the given environment, a notable advantage.

In addition to comparing the peak reflectivities, which are representative of the convective cores, statistical comparisons in the form of CFADs (see Yuter and Houze, [1995]) are performed to evaluate the overall performance of each simulation with respect to reflectivity. This technique computes the probability of a field as a function of height. To achieve the most meaningful comparisons, the CFADs must be computed as similarly as possible between the model and radar-derived fields. Comparisons between the model and observations are based on a 10 min temporal resolution for each. Reflectivity CFADs were constructed by binning the reflectivities into 1-dBZ bins from 0 to 70 dBZ at each level.

Figure 7a shows the observed CFAD. The highest probabilities follow a coherent pattern with the peak density steadily decreasing with height from between 20 and 35 dBZ near the melting level to between 5 and 15 dBZ above 12 km, indicative of a robust sedimentation/aggregation effect. Maximum reflectivities at the lowest frequency
contour of 0.001 % are just over 60 dBZ from the surface up to 6 km and drop off steadily aloft to around 45 dBZ at 14 km. The Graupel scheme simulated CFAD (Fig. 7b) has some notable discrepancies with the observed. First, it lacks all of the reflectivity values higher than 45 dBZ above the freezing level. Second, although it captures some of the aggregation effect evident in the observed CFAD, it is too weak with too few echoes in the 20-25 dBZ range between 4 and 8 km. In contrast, the Hail scheme (Fig. 7c) can simulate the rare high reflectivity values above the freezing level as was observed, though its peak values at the lowest contour of ~65 dBZ near the melting level are higher than the observed peak at this frequency of ~61 dBZ. However, when it comes to the most common echoes, the Hail scheme has an unrealistic aggregation signature quite unlike the observed with the area of highest probabilities shifted too low (< 10 dBZ) aloft, too high (~30 to 35 dBZ) at midlevels and which then decrease down toward the melting level. This is a direct result of having a fixed snow intercept where size varies only with mass with no temperature dependence, causing snow size to peak at midlevels.

The original 4ICE scheme (Fig. 7d), on the other hand, contains the best features of both the Graupel and Hail schemes, only better. It produces a very realistic radar reflectivity CFAD with a more robust and coherent aggregation signature than the Graupel scheme that much more closely resembles the observed as well as peak reflectivities similar to the Hail scheme only closer to the observed and which realistically monotonically decrease with height as observed. With a further enhanced aggregation effect in the snow mapping (also see Fig. 4 for comparison), the modified 4ICE scheme (Fig. 7e) produces an even better aggregation signature than the original 4ICE at mid and upper levels, though the effect appears slightly too strong right above
the freezing level. The distributions of rare but intense echoes are quite similar between the two 4ICE schemes with peak values slightly weaker in the modified scheme. Below the melting level, all schemes having hail maintain higher peak reflectivities due to melting hail in agreement with the observations, though they still decrease too quickly near the surface. Figure 8 shows the individual contribution of precipitating particles (rain, snow, graupel and hail) to the modified 4ICE CFAD; snow is largely responsible for the high occurrence of low dBZ values aloft and hail for the low occurrence of high dBZ values aloft.

Figure 9 shows the normalized degree of overlap between the observed and simulated probability distribution functions (PDFs) at each level where unity represents perfect overlap and zero indicates no overlap between the observed and simulated reflectivity PDFs at a given level. The two 4ICE simulated PDFs are consistently better than the Graupel between the surface and ~11 km, which is itself vastly better than the Hail for all levels above 5 km. Between the two 4ICE schemes, the modified scheme is better overall, being consistently better at mid and upper levels but not so just above the freezing level (~5 km) and at lower levels. These improvements to the radar distributions were gained largely through the introduction of the snow mapping (i.e., the Graupel scheme) and its subsequent revisions with a stronger (i.e., 4ICE_v0) and stronger (i.e., 4ICE) aggregation effect. With its fixed snow intercept, the Hail scheme suffers from a lack of an aggregation effect, resulting in its low scores aloft. Overall, the Hail scheme has the poorest overall performance in terms of CFADs, while the modified 4ICE clearly performs the best overall due to its ability to replicate the observed aggregation effect the best, especially above 6 km. Also, the original 4ICE scheme scores in NU-WRF are
better than those using the GCE model for this case (see Figure 7 in Lang et al. 2014); this is likely due primarily to the smaller domain and cyclic lateral boundaries used in the GCE model, which can inhibit the size and continuity of the stratiform region and therefore the proportion of stratiform echoes and possibly the structure of the stratiform region itself. Differences between the large-scale forcing imposed in the GCE model and the updated lateral boundary conditions used in NU-WRF could contribute to the differences, but the smaller domain size and cyclic boundaries are first order issues.

Figure 10 shows CFADs of in-cloud vertical air velocity over the total, convective and stratiform regions, which are determined based on the Steiner et al. [1995] convective-stratiform separation method (please see the following section for further details), somewhat similar to those shown in Tao et al. [1987, their figure 10]. The velocity CFADs characterize the cloud dynamics, which both drive and respond to the microphysics. The general features are similar for all the simulations, with upward velocities exceeding 40 m/s in the mid-to-upper troposphere in the convective regions, peak convective updrafts about twice as strong as the downdrafts, and higher probabilities of moderate (~10 to 20 m/s) updrafts in the convective regions than in the stratiform. Differences in the microphysics schemes lead to relatively minor variations in the velocity CFADs. For example, the Graupel and modified 4ICE schemes, which have the strongest cold pools and most organization, also have slightly broader total velocity CFADs aloft (Fig. 10a). Stratiform PDFs for the Graupel scheme (Fig. 10c) are appreciably wider than the other schemes with stronger updrafts/downdrafts classified as stratiform. The scheme also has a higher percentage of weak-to-moderate updrafts (~5-10 m/s) in the lower troposphere in the convective region (Fig. 10b) but a somewhat
reduced proportion aloft compared to the others. The combination of more moderate
reflectivities and more sheared updraft structures due to stronger cold pool dynamics in
the absence of a rain evaporation correction [cf. Li et al., 2009] makes it more difficult to
cleanly separate the convective and stratiform regions in the Graupel simulation. This
causes low-level updrafts to be included in the convective region but the upper portion of
some of those more tilted updrafts to be assigned to the stratiform region. Overall, the
fact that the total distributions are quite similar for all the schemes suggests that the large-
scale shear and instability dominate microphysics scheme differences in determining the
updraft intensities and distribution, especially for such a strongly unstable and sheared
environment.

3.2 Surface rainfall and its convective and stratiform characteristics

Surface rainfall and its characteristics are important for hydrological applications,
including hydrological as well as ocean mixed layer models, and surface processes and
are a key model component in the development of satellite rain retrieval algorithms.
Another key component of the NMQ system is the next generation quantitative
precipitation estimation (Q2). However, despite its high temporal and spatial resolution,
radar-only Q2 rainfall has its own limitations. As noted in Tang et al. [2012], daily
average Q2 rainfall has a positive bias compared to gauge-corrected Stage IV and NCEP
Climate Prediction Center (CPC) rain gauge estimates during summer (JJA, 2010).
Therefore, Stage IV bias-corrected surface rainfall estimates (Q2bias) [Tang et al., 2012],
which incorporate rain-gauge data to correct the radar product bias, are used to compare
with the model simulations. Figure 11 shows one-hour accumulated surface rainfall at 10
UTC for the NU-WRF simulations and the NMQ Q2 Stage IV bias-corrected estimates.

All of the NU-WRF simulations produce areas of heavy rain with trailing lighter rain areas consistent with their radar signatures (Fig. 3). There are, however, some notable differences between the simulations in terms of the size, organization and intensity of their heavy as well as light rain areas. The Graupel scheme (Fig. 3b) produces a broad, coherent area of trailing light to moderate rainfall, but the extensive areas of moderately intense rainfall there are not observed. Though well organized, the heavy rainfall at the leading edge of the Graupel system appears too narrow and somewhat weak. Without hail, moderate-falling graupel is more easily transported rearward in the Graupel simulation, reducing convective rain rate intensities while intensifying those in the trailing stratiform region. Locally, heavy rainfall in the Hail simulation (Fig. 3c) appears to capture the intensity and breadth of the observed but lacks the coherent extended arc structure of the observations. It also produces a slightly narrow, less coherent light rain area, but its intensity appears similar to the observed estimates. Dry collection in the Hail scheme allows some slow-falling snow to be collected and fall out as hail in the convective leading edge as opposed to being transported rearward and inhibits stratiform development. The original 4ICE scheme (Fig. 3d) is similar to the Hail in that it produces locally heavy rainfall but lacks in the overall intensity and organization of the heavy rain areas associated with the convective leading line; its trailing light rain area is also too narrow compared to the NMQ estimates. In terms of both light and heavy rain features and the overall rainfall pattern, the modified 4ICE scheme (Fig. 3e) best matches the observations. It is generally able to replicate both the coherent arc of heavy rain along the leading edge as well as the width, coherence and intensity of the trailing
The 4ICE schemes allow only ice that was formed in a manner that would produce a high particle density (e.g., freezing drops or extreme riming) to fall out as hail in the convective leading edge and therefore more slow-falling snow to be transported rearward. Although the enhanced snow autoconversion in the modified 4ICE scheme helps to produce a slightly broader, more uniform anvil and thus light rain area, as previously noted, a key difference between the original and modified 4ICE schemes is the amount of ice supersaturation permitted with the smaller global value of 5% in the modified scheme leading to a much better developed and organized and realistic MCS.

Tables 3 and 4 show the quantitative rainfall amount and area coverage (at 1-km grid spacing) between 06 and 12 UTC on 20 May 2011. Data from the first 6 hours were not used since the simulations use a cold start. Only areas with surface rain rates greater than 0.15 mm/h, the minimum Q2 rain rate, are partitioned. Unclassified rain due to light rain areas in the model and a mismatch between the rain and classification time intervals in both the model, which requires 3D data for the partitioning, and NMQ estimates causes the totals to exceed the sum of the convective and stratiform parts. The results show that the total rainfall amount in the modified 4ICE run is significantly more (~10%) than the 3ICE runs (both Graupel and Hail) and vastly more (~18%) than the original 4ICE. When compared to the bias-corrected Q2 estimate, the modified 4ICE, Graupel and Hail rainfall totals are all relatively close to the bias corrected, just 5.7% higher, 4.4% lower and 3.7% lower, respectively. The original 4ICE, however, is 10.4% lower, which may indicate a possible low bias. In terms of the total convective plus stratiform rain (i.e., not including the model light rain rate areas below 0.15 mm/h), the modified 4ICE is closest to the bias corrected (4.1% higher), while the others are 12.3% (original 4ICE), 7.4%
(Graupel) and 6.3% (Hail) less than the bias corrected, which suggests they may have a slight low bias, but there are no error estimates for the bias-corrected estimates so it remains uncertain. Clear sky initiation (i.e., a cold start), initial/lateral boundary conditions, and the fact that the observed line extended farther south than the innermost domain could account for the differences. **Consistent with Fig. 11**, in terms of total rain area coverage (Table 4), all of the schemes are very close to that of the bias corrected (within 3% relative to the total area). The models appear to miss the complete extent of the observed light rain area over western Oklahoma. This could be due to the real squall line extending further south than in the simulations, resulting in the stratiform region in the simulations being under developed at the southern end of the analysis domain relative to the observed.

Rainfall can also be separated into convective and stratiform regions. There are several reasons for the distinction [Houze, 1997]. Precipitation rates are generally much higher in the convective region where ice particles tend to be rimed as opposed to the stratiform region where aggregates dominate. Microphysics and, as a result, surface rainfall rates and the vertical distribution of latent heating are also different in the two regions [see reviews by Houze, 1997 and Tao, 2003]. The convective-stratiform partitioning method used in this study is based on the horizontal radar reflectivity gradient with the criteria for identifying convective regions based on intensity, “peakedness”, and the surrounding area as described by Steiner et al. [1995]. Because the scheme was originally developed for tropical convection, several parameters have been tuned for mid-latitudes [Feng et al., 2011]. A 2-km mean sea level (MSL) height (versus 3 km in Steiner et al. [1995]) is used as the analysis level to avoid bright band
contamination, and the convective reflectivity threshold is 43 dBZ (versus 40 dBZ in Steiner et al. [1995]), according to the Z-R relationship in mid-latitudes. Echoes that exceed 10 dBZ but not identified as convective are designated as stratiform [Feng et al., 2011] (see Lang et al. [2003] for a review of convective-stratiform observational and modeling studies and different separation methods). The same separation method is applied to both the observations and model results.

Tables 3 and 4 also show the observed and simulated rainfall amounts and area coverage in the convective and stratiform regions with their corresponding conditional rain rates listed in Table 5. The simulations reproduce the observed convective area coverage to within a factor of 0.8-1.2 but underestimate the total convective rainfall, ranging from a factor of 0.94 to 0.62 relative to the bias-corrected value. Thus, on average, the simulated conditional convective surface rain rates are too weak. However, while the Graupel scheme has an average conditional convective surface rain rate of just 55% of the bias-corrected rate (Table 5), which is no surprise given the moderate fall speeds of graupel, the conditional convective surface rain rates for the Hail and modified 4ICE schemes are 91% and 81% of the bias-corrected rate, respectively. Though much closer, aliasing in the rain/hail sedimentation and in the hail melting due to the use of a 1M scheme could be factors. As rain or hail begins to fall towards the surface, their initial mass in the next lowest grids cells will be small, which in a 1M scheme will force their sizes to be too small, slowing their fall speeds and overdoing hail melting. The rain evaporation correction has a significant impact on the stratiform region where it can help to overcome this effect by boosting drop sizes but very little on the convective. Overall, the modified 4ICE has the most convective rainfall with more intense convective rain
rates than the Graupel scheme and a larger convective area as a result of having a longer continuous length of leading edge convection than the Hail scheme due to being better organized.

All of the NU-WRF simulations are within a factor of 0.73 to 0.82 of the bias-corrected stratiform area coverage (Table 4) and thus all produce too little stratiform rain area in comparison. However, all of the simulated stratiform rainfall amounts are equal to or greater than the bias corrected (Table 3), meaning the model conditional stratiform surface rain rates appear too intense (Table 5). Though vastly lower than the conditional convective surface rain rates, which agrees well with the observed trend of having much higher conditional rain rates in the convective region (Table 5), the Hail and both 4ICE conditional stratiform surface rain rates are still 38 and 58% higher than the bias corrected, respectively, while the conditional Graupel rate is 84% higher. The degree of over-bias is in rough proportion to the amount of graupel present in each scheme’s respective stratiform region (see Fig. 15 and the related discussion in Section 3.3). In terms of the overall stratiform percentage, the Hail scheme is quite close to the observed value of ~33% followed by the modified 4ICE. The original 4ICE and Graupel scheme stratiform percentages are too high (~50-53%) due to both too little convective and too much stratiform rainfall.

Consistent with previous modeling results, the Hail scheme produces less stratiform but more convective rain than the Graupel scheme. McCumber et al. [1991] suggested that the most important characteristic difference between graupel and hail is the terminal velocity.
Figure 12a,b shows PDFs of the total simulated and observed surface rain rate intensities. Both the Hail and the two 4ICE schemes, which all include faster falling hail, have a higher proportion of heavy precipitation (i.e., > 30 mm/h) as well as less moderate precipitation (i.e., 10-20 mm/h) than does the Graupel scheme, placing them in better agreement with the bias-corrected Q2 radar estimates in both situations. However, the Hail and especially the modified 4ICE scheme are in the best agreement with the observed frequencies of heavy surface rain rates (i.e., > 30 mm/h). In terms of very light (i.e., < 2.5 mm/h) and moderate (i.e., 5 to 20 mm/h) surface rainfall rates, the Hail scheme’s frequencies are consistently closest to the Q2 bias-corrected (Fig. 12b) despite its unrealistic anvil radar structure. For the convective region (Fig. 12c), both the simulated and NMQ-estimated surface rain rate PDFs are shifted to higher intensities relative to the total as expected but with similar biases; the Hail and modified 4ICE schemes are again fairly comparable and reasonably close to the biased-corrected frequencies but slightly underestimate the occurrence of surface rain rates above 30 mm/h. This bias is more apparent with the original 4ICE scheme whereas the Graupel scheme greatly underestimates their occurrence. This is consistent with and shows the source of the low biases in the conditional convective surface rain rates in relation to the intensity spectrum. All of the simulations, but especially the Graupel scheme, tend to produce too high of a proportion of moderate convective rain rates (i.e., 5 to 20 mm/h) compared to the NMQ frequencies. Rain rate PDFs for the stratiform region (Fig. 12d) are shifted to lower surface rain rate intensities as expected. Overall, as with the total PDFs, the Hail scheme performs quite well and clearly agrees the best with the bias-corrected frequencies for the light to moderates rain rates prevalent in the stratiform
region and is consistent with it having the best conditional stratiform rain rate (Table 5). Again as with the convective region, the Graupel scheme greatly overestimates the frequency of moderate surface rain rates (between 5 and 20 mm/h) only this time at the expense of too few light rain rates (versus too few heavy rain rates in the convective region). The two 4ICE schemes also produce too high a fraction of moderate stratiform surface rain rates but mainly from 5 to 10 mm/h. At the weakest rain intensities (i.e., < 2.5 mm/h), they are slightly better than the Graupel scheme but not nearly as good as the Hail.

3.3 Simulated hydrometeor properties

Although simulated hydrometeor profiles have traditionally lacked effective (i.e., comprehensive) validation, in situ and polarimetric radar-based hydrometeor identification (HID) algorithms [Straka et al., 2000] can provide information on the type of species expected in different parts of a convective system. For example, Stith et al. [2002] noted that graupel-dominated, GCE model-simulated stratiform profiles (using an earlier version of the Goddard microphysics) were unrealistic based on their in situ aircraft studies, which were dominated by aggregates with graupel not found in significant amounts. An accurate vertical distribution of cloud species is important for satellite retrievals [i.e., Lang et al., 2007; Olson et al., 2006]. Unrealistic precipitation ice contents (i.e., snow and graupel), for example, can bias the simulated brightness temperatures and make it difficult to infer cloud properties from remote sensing data, which link them with synthetic values from models [Matsui et al., 2013]. Simulated hydrometeor profiles can also be used to confirm or explain specific model behavior.
Figure 13 shows vertical profiles of the total horizontal domain- and time-averaged cloud species (i.e., cloud water, rain, cloud ice, snow, graupel and/or hail). Low-level rain and cloud water mixing ratios have only subtle variations with the Graupel and modified 4ICE schemes having on average slightly stronger low-level rain evaporation signatures due to their better organization and more developed stratiform regions and lack of a rain evaporation correction in the Graupel scheme. Snow is a dominant ice species to varying degrees in all of the schemes, but the modified 4ICE scheme has the most snow and less cloud ice than the Graupel or 4ICE_v0. Snow autoconversion, which was slightly increased in the original 4ICE, was further strengthened in the modified 4ICE, allowing more cloud ice to be converted into snow than in the Graupel and 4ICE_v0 schemes. The combination of a prescribed fixed snow intercept smaller than the snow mappings in the other schemes, which for a given amount of mass yields larger snow sizes aloft and hence less snow deposition, and the allowance of dry collection in the Hail scheme contributes to it having less snow and cloud ice. The Graupel scheme produces a much larger graupel profile than the two 4ICE schemes, as both rimed particles and frozen drops are treated as graupel, which has a moderate fall speed and remains suspended much longer than hail. The modified 4ICE scheme has less graupel than the original as a result of switching from the Meyers to the Cooper curve for the number of active IN and capping cloud ice size to the minimum snow size. The result is more ice and deposition growth and hence snow at the expense of riming and graupel. Hail is much larger and has much faster fall speeds than graupel. This allows it to fall further than graupel below the melting layer before fully melting but also greatly reduces the amount that is suspended aloft as shown by the differences between the Graupel scheme’s graupel profile and the
Hail scheme’s hail profile. In terms of the two 4ICE schemes, the modified 4ICE has more hail due in part to its better organization and slightly larger convective area but also to the new hail size mapping scheme, which produces smaller hail with both a larger surface area and reduced fall speeds than that for the fixed hail intercept used in the original 4ICE until mixing ratios become fairly large. The vertical distribution of snow and hail for the Hail scheme is quite similar to the results of Lin et al. [1983] upon which it is based and which were also for midlatitude convection.

Figures 14 and 15 show the vertical distributions of cloud species in the convective and stratiform regions. The Graupel scheme’s convective region is dominated by both graupel and snow. Without hail, freezing of supercooled water is forced to become graupel, which can remain suspended longer and results in a large proportion of graupel. However, with graupel dry collection turned off, snow and cloud ice are also present in large amounts. In the stratiform region, snow is more dominant; however, graupel is still present in large quantities, having been efficiently transported into the stratiform region due to its moderate fall speed as is cloud ice due to the weak snow autoconversion effect. In contrast to the Graupel, the Hail scheme’s convective region has very little cloud ice, much less snow, and a large proportion of hail considering its high fall velocity. With hail dry collection included, ice and snow are scavenged to become hail. Its hail profile also contains a secondary maximum near 11 km, which coincides with the secondary reflectivity maximum that was not observed. The combination of hail dry collection and deposition likely contribute to this secondary hail peak, which is not present in either 4ICE scheme. The Hail stratiform region is almost completely snow, though the amount is less than the other schemes. Unlike graupel, hail
falls out quickly in the convective region while cloud ice is depleted via an over efficient Psfi term (vapor growth of ice into snow, see Lang et al. [2011]).

As for the two 4ICE schemes, both contain sizeable proportions of cloud ice, snow, graupel and hail in their convective regions. Similar to the Graupel scheme, eliminating dry collection allows for ample cloud ice and snow to be present with the modified 4ICE having a higher proportion of snow relative to cloud ice due to its enhanced snow autoconversion, which is likewise apparent in the stratiform profiles. While the modified 4ICE has a larger convective hail profile due in part to the hail mapping, both have almost no hail in their stratiform regions. As with the total profiles, the Cooper curve leads to less graupel in proportion to snow in the modified scheme compared to the original in both the convective and stratiform regions. Both 4ICE schemes produce much less graupel than the Graupel scheme, as a significant fraction of frozen supercooled water becomes hail. The end result is that the modified 4ICE scheme has very little graupel in its stratiform region, which is largely dominated by snow consistent with both in situ measurements (e.g., Stith et al. [2002]) and radar HID analyses (e.g., Lerach et al. [2004], Guy et al. [2013]) of MCSs and also similar to the Hail scheme except that the total amount of stratiform snow is much greater in the modified 4ICE.

4. Summary and discussion

In this study, NU-WRF was used at a relatively high horizontal resolution (i.e., 1-km for the innermost domain) to examine the performance of a modified version of the new Goddard 4ICE microphysics scheme in relation to the original and two previous
3ICE versions of the Goddard microphysics, a hail scheme and an improved graupel scheme, for a strong, well-organized MCS with intense leading edge convection and a well-developed trailing stratiform region that was observed on 20 May 2011 during the MC3E field campaign. The schemes were evaluated in terms of their radar reflectivity structures and distributions, propagation, rainfall and surface rain rate histograms versus NMQ NEXRAD radar data and gauge-corrected rainfall estimates and also compared in terms of their simulated hydrometeor profiles. The major results are as follows:

- All four schemes reproduce the basic leading convective edge trailing stratiform squall line structure, though individual performance metrics varied significantly from scheme to scheme and between metrics. However, collectively the modified 4ICE scheme clearly performed the best, equaling or outperforming the other schemes in terms of system organization and structure, propagation, horizontal and vertical reflectivity structures, radar echo distributions, peak reflectivity profiles, and total surface rain rate histograms.

- The Hail scheme actually produces conditional surface rain rates that are closest to the observed with the highest convective and lowest stratiform rates. Without graupel, all high-density ice falls out quickly in the convective region, leaving its stratiform region completely dominated by snow. However, the vertical structure of its stratiform region is completely unrealistic. With only a fixed snow intercept, reflectivities are maximized well above the freezing level, which results in it having the worst radar CFAD scores aloft. Having hail, it can produce intense echoes, but dry collection causes some of the slow-falling snow to be collected and fall out pre-maturely as hail in the convective leading edge and may
contribute to an unrealistic secondary echo maximum at upper levels. Overall, its simulated MCS also lacks organization.

- The Graupel scheme produces a well-developed MCS with a large, coherent stratiform rain area. Its radar CFAD scores are much better than the Hail aloft as a result of having a snow mapping scheme. However, without hail, it vastly underestimates peak reflectivities and convective surface rain rates, and too much graupel is carried into the stratiform region causing excessively high stratiform surface rain rates. Also, without a rain evaporation correction, its leading edge convection propagates too fast.

- The original Goddard 4ICE scheme improves upon the Graupel by including hail, which allows it to produce intense echoes and higher convective surface rain rates. It eliminates the biases associated with hail dry collection by allowing only ice that was formed in a manner that would produce a high-density hydrometeor (e.g., freezing drops or extreme riming) to fall out as hail in the convective leading edge and therefore more slow-falling snow to be transported rearward to produce a broader more uniform light rain area. The increased aggregation effect in its revised snow mapping produces radar CFADs that are even better than the Graupel and far better than the Hail and more vertically-stratified stratiform reflectivity features in better agreement with observations. Also, omitting dry collection while including a rain evaporation correction leads to relatively narrow but intense and erect leading convective cells. Unlike the Hail scheme, its peak reflectivities monotonically decrease with height above the freezing level as
observed. However, its simulated MCS lacks overall organization and intensity due to an allowed ice supersaturation value of 20% being applied system wide.

- Although only the complete set of comprehensive changes is shown (except for the sensitivity test on the rain evaporation correction), based on the series of individual changes that were made (the individual results of which are not show), the modified 4ICE scheme improves upon the original in four ways. Though still allowing locally high ice supersaturations, lowering the background value to 5% strengthened the simulated MCS overall, leading to a better-developed stratiform region, a longer, more coherent leading convective line and improved system organization and propagation and results in it having the most total rainfall, best total surface rain rate histograms and better conditional convective surface rain rates. Continued revisions to the snow mapping with an even greater aggregation effect coupled with the addition of a snow breakup effect via graupel/hail collisions, lead to the highest radar CFAD scores aloft, the most vertically-stratified stratiform radar echoes, and best representation of the weak echo transition region. The use of the Cooper curve for the number of active IN leads to a higher proportion of snow over graupel, which lowers the amount of graupel present in the stratiform region, which lowers and thereby the improves stratiform surface rain rates. Though peak reflectivity values are slightly underestimated, the addition of a simple hail mapping relaxes the need to choose a fixed hail intercept value \textit{a priori}.

- Though conditional convective surface rain rates were too low, the Hail and modified 4ICE schemes had rates that were 91% and 81%, respectively, of the
bias corrected. Despite having hail, the 1M rain could be a factor in their low bias. Conversely, all four schemes were over biased in their conditional stratiform surface rain rates in rough proportion to the amount of graupel in their stratiform regions.

Hail processes were critical for this intense summertime MCS. Hail is essential for producing intense echoes above ~50 dBZ and higher surface rain rates. Without it, the Graupel scheme fails to produce echoes above 45 dBZ above the freezing level and allows too much moderate-falling graupel to be transported rearward. As a result, the 4ICE and Hail schemes produced more heavy (i.e., > 30 mm/h) and less moderate precipitation (i.e., 10-20 mm/h) than the Graupel, in better agreement with observations.

The rain evaporation correction improved system propagation and leading cell structure. Schemes with a well-developed stratiform region and no correction (i.e., Graupel and the modified 4ICE without the correction) had stronger cold pools and tended to propagate too quickly. Leading convective cells also exhibited a greater tilt without the correction.

Snow size mapping greatly improves the vertical variation of the modal values within the reflectivity distributions. Without it, the Hail scheme produced a disjointed weak reflectivity mode quite unlike the robust aggregation mode in the observations. The revised snow mappings in the new and modified 4ICE schemes more realistically reproduce the robust and coherent aggregation signature (i.e., the vertical variation of mode values) in the observed radar reflectivity
distribution (i.e., within the low values from ~5-25 dBZ), respectively, than the original mapping that was first implemented in the improved Graupel scheme.

- PDFs of vertical velocity were largely similar for all four schemes, suggesting the larger-scale shear and instability are more important than the changes made in the microphysics for determining the updraft intensities and distribution in such an unstable and sheared environment.

The 20 May 2011 MC3E case was one of the cases used to develop and evaluate the new 4ICE scheme in *Lang et al.* [2014] using the GCE model. Overall, the 4ICE results here are consistent with those from the GCE model [*Lang et al.*, 2014]. However, as noted previously, those GCE model simulations were forced with observed large-scale advective tendencies for temperature and water vapor requiring the use of cyclic lateral boundary conditions, which can complicate and inhibit (along with the smaller domain) the simulated spatial structures of the squall line, namely the stratiform region, by allowing the leading edge convection to wrap around behind the MCS (see for example Fig. 2 in *Lang et al.* [2014]). Restricting the stratiform area can affect the distribution of radar echoes and hence the agreement between the observed and simulated radar distributions. Accordingly, CFAD scores for the original 4ICE scheme for this same case in the GCE model study are consistently lower (i.e., less than 0.75; see Fig. 7 in *Lang et al.* [2014]) than they are using NU-WRF in this study (i.e., consistently above 0.8) using the same original version of the 4ICE scheme. Also, the double cyclic boundaries made it difficult to see the impact of the rain evaporation correction, which is quite evident in this study. The ability to use a larger domain with open lateral boundaries and non-
uniform horizontal forcing in NU-WRF is less restrictive and produces superior results and is a more realistic evaluation of the 4ICE scheme.,

Simultaneously, the new 4ICE scheme has been implemented and tested in the Goddard Multi-scale Modeling System (MMF), which utilizes the GCE model as the cloud-precipitation parameterization within the Goddard Earth Observing System (GEOS) global model. Matsui et al. [2015] evaluated statistical distributions of convective precipitation type from the Goddard MMF with the new modified 4ICE scheme by contrasting land and ocean regions in the Tropics in comparison with TRMM signal statistics. Chern et al. [2015] studied the impact of different microphysical schemes, including the new modified 4ICE scheme, as well as their performance within the Goddard MMF compared with three CloudSat/CALIPSO retrieval products.

In Part II, the new Goddard 4ICE scheme with the additional modifications presented in this study will be compared with other WRF microphysics schemes (i.e., Morrison, WSM6, and WDM6). This modified version will also be implemented into the NCAR WRF for community use.

Acknowledgments

This research was supported by the NASA Precipitation Measurement Missions (PMM), the NASA Modeling, Analysis, and Prediction (MAP) Program, and the Office of Science (BER), U.S. Department of Energy/Atmospheric System Research (DOE/ASR) Interagency Agreement (No. DE-AI02-04ER63755). NMQ radar and precipitation products were provided by Dr. Xiquan Dong at the University of North Dakota and Carrie Langston at the National Severe Storms Laboratory, while Dr. Yudong
Tian (University of Maryland) at NASA GSFC provided the bias-corrected Q2 data. The authors are grateful to Drs. Ramesh Kakar and David B. Considine at NASA headquarters for their support of this research. Acknowledgment is also made to the NASA Goddard Space Flight Center and NASA Ames Research Center computing facilities and to Dr. Tsengdar Lee at NASA HQ for the computational resources used in this research.

References


Grant, L. D., and S. C. van den Heever, (2015), Cold pool and precipitation responses to


Lang, S., W.-K. Tao, J.-D. Chern, D. Wu, and X. Li (2014), Benefits of a 4\textsuperscript{th} ice class in the simulated radar reflectivities of convective systems using a bulk microphysics scheme, *J. Atmos. Sci.*, 71, 3583-3612.


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-


<table>
<thead>
<tr>
<th>Study</th>
<th>Model</th>
<th>Microphysics</th>
<th>Resolution/Vertical Layers</th>
<th>Integration Time</th>
<th>Case(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin et al. (1983)</td>
<td>2D</td>
<td>3ICE</td>
<td>200 m/95</td>
<td>48 min</td>
<td>Montana hail event</td>
</tr>
<tr>
<td>Cotton et al. (1982, 1986)</td>
<td>2D</td>
<td>3ICE</td>
<td>500 m/31</td>
<td>5 h</td>
<td>Orographic snow</td>
</tr>
<tr>
<td>Rutledge and Hobbs (1984)</td>
<td>2D kinematic</td>
<td>3ICE</td>
<td>600 m/20</td>
<td>Steady State</td>
<td>Narrow cold front</td>
</tr>
<tr>
<td>Lord et al. (1984) *</td>
<td>2D axisymmetric</td>
<td>3ICE vs Warm Rain</td>
<td>2 km/20</td>
<td>4.5 days</td>
<td>Idealized</td>
</tr>
<tr>
<td>Yoshizaki (1986)#</td>
<td>2D slab-symmetric</td>
<td>3ICE vs Warm Rain</td>
<td>0.5 km/32</td>
<td>4.5 h</td>
<td>12 September GATE squall line</td>
</tr>
<tr>
<td>Nicholls (1987)</td>
<td>2D slab-symmetric</td>
<td>3ICE vs Warm Rain</td>
<td>0.5 km/25</td>
<td>5 h</td>
<td>12 September GATE squall line</td>
</tr>
<tr>
<td>Fovell and Ogura (1988)#*%</td>
<td>2D slab-symmetric</td>
<td>3ICE vs Warm Rain</td>
<td>1 km/31</td>
<td>10 h</td>
<td>Midlatitude squall line</td>
</tr>
<tr>
<td>Tao and Simpson (1989, 1993)#</td>
<td>2D and 3D</td>
<td>3ICE vs Warm Rain</td>
<td>1 km/31</td>
<td>12 h</td>
<td>GATE squall line</td>
</tr>
<tr>
<td>Tao et al. (1990)</td>
<td>2D</td>
<td>3ICE</td>
<td>1 km/31</td>
<td>12 h</td>
<td>GATE squall line</td>
</tr>
<tr>
<td>McCumber et al. (1991)#%$</td>
<td>2D and 3D</td>
<td>3ICE (graupel vs hail, 2ICE vs 3ICE)</td>
<td>1 km/31</td>
<td>12 h</td>
<td>GATE squall line</td>
</tr>
<tr>
<td>Wu et al. (1999)</td>
<td>2D slab-symmetric</td>
<td>2ICE</td>
<td>3 km/52</td>
<td>39 days</td>
<td>TOGA COARE</td>
</tr>
<tr>
<td>Ferrier (1994), Ferrier et al. (1995)#</td>
<td>2D slab-symmetric</td>
<td>2-moment 4ICE</td>
<td>1 km/31</td>
<td>12 h</td>
<td>COHMESS, GATE squall line</td>
</tr>
<tr>
<td>Tao et al. (1995)</td>
<td>2D slab-symmetric</td>
<td>3ICE</td>
<td>0.75 and 1 km/31</td>
<td>12 h</td>
<td>EMEX and PRESTORM</td>
</tr>
<tr>
<td>Reference</td>
<td>Simulation Type</td>
<td>Domain</td>
<td>Resolution</td>
<td>Time Frame</td>
<td>Results</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>-----------------</td>
<td>--------</td>
<td>------------</td>
<td>------------</td>
<td>---------</td>
</tr>
<tr>
<td>Walko et al. (1995)#</td>
<td>2D</td>
<td>4ICE</td>
<td>0.3 km/ 80</td>
<td>30 min</td>
<td>Idealized</td>
</tr>
<tr>
<td>Meyers et al. (1997)#/S</td>
<td>2D</td>
<td>2-moment 4ICE</td>
<td>0.5 km/ 80</td>
<td>30 min</td>
<td>Idealized</td>
</tr>
<tr>
<td>Straka and Mansell (2005)#</td>
<td>3D</td>
<td>10-ICE</td>
<td>0.5 km/ 30</td>
<td>~2 h</td>
<td>Idealized</td>
</tr>
<tr>
<td>Lang et al. (2007)#</td>
<td>3D</td>
<td>3ICE</td>
<td>0.25, 1.15 km/41</td>
<td>8 h</td>
<td>LBA</td>
</tr>
<tr>
<td>Zeng et al. (2008)#</td>
<td>2D and 3D</td>
<td>3ICE</td>
<td>1 km/41</td>
<td>40 days</td>
<td>SCSMEX, KWAJEX</td>
</tr>
<tr>
<td>Milbrandt and Yau (2005)#</td>
<td>1D</td>
<td>3-moment</td>
<td>N.A./ 51</td>
<td>50 min</td>
<td>Idealized hail storm</td>
</tr>
<tr>
<td>Morrison et al. (2005)#</td>
<td>1D</td>
<td>2-moment 2ICE</td>
<td>N.A. / 27</td>
<td>3 days</td>
<td>SHEBA</td>
</tr>
<tr>
<td>Morrison and Grabowski (2008)#</td>
<td>2D</td>
<td>2-moment ICE</td>
<td>50 m/ 60</td>
<td>90 min</td>
<td>Idealized</td>
</tr>
<tr>
<td>Reisner et al. (1998)#</td>
<td>MM5</td>
<td>3ICE and 2-moment ICE</td>
<td>2.2 km/ 27</td>
<td>6 h</td>
<td>Winter storms</td>
</tr>
<tr>
<td>Thompson et al. (2004)#</td>
<td>2D MM5</td>
<td>3ICE</td>
<td>10 km/ 39</td>
<td>3 h</td>
<td>Idealized</td>
</tr>
<tr>
<td>Thompson et al. (2008)#</td>
<td>2D WRF</td>
<td>3ICE</td>
<td>10 km/ 39</td>
<td>6 h</td>
<td>Idealized</td>
</tr>
<tr>
<td>Colle and Mass (2000)</td>
<td>MM5</td>
<td>3ICE</td>
<td>1.33 km/ 38</td>
<td>96 h</td>
<td>Orographic flooding</td>
</tr>
<tr>
<td>Colle and Zeng (2004)#</td>
<td>2D MM5</td>
<td>3ICE</td>
<td>1.33 km/ 39</td>
<td>12 h</td>
<td>Orographic</td>
</tr>
<tr>
<td>Colle et al. (2005)#</td>
<td>MM5</td>
<td>3ICE</td>
<td>1.33 km/ 320</td>
<td>36 h</td>
<td>IMPROVE</td>
</tr>
<tr>
<td>Yang and Ching (2005)#</td>
<td>MM5</td>
<td>3ICE</td>
<td>6.67 km/ 23</td>
<td>2.5 days</td>
<td>Typhoon Toraji (2001)</td>
</tr>
<tr>
<td>Zhu and Zhang (2006b)#</td>
<td>MM5</td>
<td>3ICE</td>
<td>4 km/ 24</td>
<td>5 days</td>
<td>Hurricane Bonnie (1998)</td>
</tr>
<tr>
<td>Wang (2002)*</td>
<td>TCM3-hydrostatic</td>
<td>3ICE</td>
<td>5 km/ 21</td>
<td>5 days</td>
<td>Idealized</td>
</tr>
<tr>
<td>Hong et al. (2004)#</td>
<td>2D WRF</td>
<td>3ICE</td>
<td>250 m/ 80</td>
<td>1 h</td>
<td>Idealized and Korea heavy snow event</td>
</tr>
<tr>
<td>Li and Pu (2008)*</td>
<td>WRF</td>
<td>2ICE and 3ICE</td>
<td>3 km/ 31</td>
<td>1.25 days</td>
<td>Hurricane Emily (2005)</td>
</tr>
<tr>
<td>Jankov et al. (2005, 2007)#</td>
<td>WRF</td>
<td>2ICE and 3ICE</td>
<td>12 km/ 31</td>
<td>1 day</td>
<td>IHOP</td>
</tr>
<tr>
<td>Dudhia et al. (2008)***</td>
<td>WRF</td>
<td>3ICE</td>
<td>5 km/ 31</td>
<td>1.5 days</td>
<td>Korean heavy snow event</td>
</tr>
<tr>
<td>Tao et al. (2009, 2011)</td>
<td>WRF</td>
<td>2ICE and 3ICE, 3ICE and 4ICE</td>
<td>1 km/ 41, 1.667 km/ 31</td>
<td>1.5 days</td>
<td>IHOP and Hurricane Katrina (2005)</td>
</tr>
<tr>
<td>Han et al. (2012)&amp;</td>
<td>WRF</td>
<td>1- and 2-moment 3ICE</td>
<td>1.3 km/ 52</td>
<td>2 days</td>
<td>Northern California winter cyclone</td>
</tr>
<tr>
<td>Iguchi et al. (2012a,b)</td>
<td>WRF</td>
<td>3ICE and SBM</td>
<td>1 km/ 60</td>
<td>36 h</td>
<td>C3VP and MC3E</td>
</tr>
<tr>
<td>Li et al. (2009a,b)&amp;</td>
<td>2D GCE</td>
<td>3ICE and SBM</td>
<td>1 km/ 33</td>
<td>12 h</td>
<td>PRE-STORM</td>
</tr>
<tr>
<td>Del Genio et al. (2012)</td>
<td>WRF</td>
<td>2-moment 3ICE</td>
<td>600 m/ 50</td>
<td>3 days</td>
<td>TWP-ICE</td>
</tr>
<tr>
<td>Gilmore et al. (2004)#</td>
<td>SAM</td>
<td>3ICE</td>
<td>1 km/ 40</td>
<td>2 h</td>
<td>Idealized</td>
</tr>
<tr>
<td>Powell et al. (2012)&amp;</td>
<td>WRF</td>
<td>1- and 2-moment 3ICE</td>
<td>3 km/ 61</td>
<td>24 h and 30 h</td>
<td>AMMA</td>
</tr>
<tr>
<td>Tao et al. (2013)#</td>
<td>WRF</td>
<td>3ICE</td>
<td>2 km/ 41</td>
<td>2 days</td>
<td>MC3E</td>
</tr>
<tr>
<td>Wu et al. (2013)</td>
<td>WRF</td>
<td>2ICE and 3ICE</td>
<td>3 km/ 41</td>
<td>2 days</td>
<td>SGP MCS</td>
</tr>
<tr>
<td>Lang et al. (2011)#</td>
<td>3D GCE</td>
<td>3ICE</td>
<td>250 and 500 m/ 70</td>
<td>6 h and 72 h</td>
<td>TRMM LBA and KWAJEX</td>
</tr>
<tr>
<td>Morrison et al. (2008)#&amp;</td>
<td>2D WRF</td>
<td>1- and 2-moment 3ICE</td>
<td>250 m/ N.A.</td>
<td>7 h</td>
<td>Idealized</td>
</tr>
<tr>
<td>Varble et al. (2011)&amp;</td>
<td>Multiple models</td>
<td>1- and 2-moment 3ICE</td>
<td>917 m and 1 km/ 50 or 120</td>
<td>6 days</td>
<td>TWP-ICE</td>
</tr>
<tr>
<td>Fridlind et al. (2012)&amp;</td>
<td>Multiple models</td>
<td>1- and 2-moment 2ICE / 3ICE</td>
<td>900 m - 3 km/ N.A.</td>
<td>6 days</td>
<td>TWP-ICE</td>
</tr>
<tr>
<td>Van Weverberg et al. (2012)#</td>
<td>ARPS</td>
<td>3ICE</td>
<td>3 km/ 50</td>
<td>30 h</td>
<td>Convective and stratiform cases</td>
</tr>
<tr>
<td>Van Weverberg et al. (2012)&amp;</td>
<td>2D WRF</td>
<td>2-moment 3ICE and 4ICE</td>
<td>1 km/ 40</td>
<td>5 h</td>
<td>Idealized</td>
</tr>
<tr>
<td>Van Weverberg et al. (2013a,b)</td>
<td>WRF</td>
<td>2-moment 3ICE</td>
<td>1 km/ 40</td>
<td>5 h</td>
<td>Idealized and TWP-ICE</td>
</tr>
<tr>
<td>Van Weverberg et al. (2013)&amp;</td>
<td>WRF</td>
<td>1- and 2-moment 2ICE and 3ICE</td>
<td>4 km/ 35</td>
<td>7 days</td>
<td>TWP-ICE</td>
</tr>
<tr>
<td>Bryan and Morrison (2011)&amp;</td>
<td>3D CM1</td>
<td>1- and 2-moment 3ICE</td>
<td>250 m/ 100</td>
<td>9 h</td>
<td>Idealized</td>
</tr>
<tr>
<td>Morrison and Milbrandt (2011)&amp;</td>
<td>WRF</td>
<td>2-moment 3ICE and 4ICE</td>
<td>1 km/ 500 m</td>
<td>2 h</td>
<td>Idealized</td>
</tr>
<tr>
<td>Morrison and Grabowski (2007)&amp;</td>
<td>2D kinematic</td>
<td>2-moment 3ICE</td>
<td>N.A.</td>
<td>N.A.</td>
<td>Idealized</td>
</tr>
<tr>
<td>Luo et al. (2010)#&amp;</td>
<td>WRF</td>
<td>1- and 2-moment 3ICE</td>
<td>3.3 km/ 30</td>
<td>1 day</td>
<td>Mei-Yu front</td>
</tr>
<tr>
<td>Li et al. (2008)</td>
<td>3D</td>
<td>Univ. of Utah CRM</td>
<td>3ICE</td>
<td>500 m / N.A.</td>
<td>3 days</td>
</tr>
<tr>
<td>-----------------</td>
<td>----</td>
<td>-------------------</td>
<td>------</td>
<td>-------------</td>
<td>--------</td>
</tr>
<tr>
<td>Molthan and Colle (2012)</td>
<td>WRF</td>
<td>1- and 2-moment 3ICE</td>
<td>9, 3, 1 km / 34</td>
<td>1 day</td>
<td>C3VP synoptic snow event</td>
</tr>
<tr>
<td>Guy et al. (2013)</td>
<td>3D GCE</td>
<td>3ICE</td>
<td>1 km / 63</td>
<td>&gt; 10 h</td>
<td>AMMA 2 Sahel MCSs</td>
</tr>
<tr>
<td>Lang et al. (2014)</td>
<td>3D GCE</td>
<td>3ICE and 4ICE</td>
<td>200 m and 1 km / 70 and 76</td>
<td>6 h and 96 h</td>
<td>TRMM LBA and MC3E MCS</td>
</tr>
<tr>
<td>Saleeby and Cotton (2004)#</td>
<td>3D RAMS</td>
<td>1- and 2-moment 5ICE, Aerosol scheme &amp; Drizzle mode</td>
<td>2.0 km / 40</td>
<td>2 h</td>
<td>Idealized supercell</td>
</tr>
<tr>
<td>Saleeby et al. (2007)#</td>
<td>3D RAMS</td>
<td>2-moment new ice fall speeds</td>
<td>3.0 km / 35</td>
<td>36 h</td>
<td>Winter synoptic case / orographic</td>
</tr>
<tr>
<td>Saleeby and Cotton (2008)#</td>
<td>3D RAMS</td>
<td>2-moment 5ICE</td>
<td>2.0 km / 45</td>
<td>27 h</td>
<td>Orogenic snowfall</td>
</tr>
<tr>
<td>Saleeby et al. (2010)%</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>1.25 km / 50</td>
<td>48 h</td>
<td>Warm &amp; mixed-phase maritime cumulus</td>
</tr>
<tr>
<td>Saleeby et al. (2009), (2013a)%</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>750m / 45 600m / 45</td>
<td>42 h</td>
<td>Orogenic snowfall</td>
</tr>
<tr>
<td>Saleeby and van den Heever (2013b)#</td>
<td>2D &amp; 3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>200 m / 50 m 1 km / 0.1-1 km 3 km / 75-800 m</td>
<td>1 day 2 h 42 h</td>
<td>Shallow warm rain Deep convection Orogenic snowfall</td>
</tr>
<tr>
<td>Saleeby et al. (2015)%</td>
<td>3D RAMS</td>
<td>2-moment, Aerosol scheme</td>
<td>250 m / 40</td>
<td>36 h</td>
<td>ATEX</td>
</tr>
<tr>
<td>Igel et al. (2013)%</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>3.0 km / 45</td>
<td>48 h</td>
<td>Synoptic warm front</td>
</tr>
<tr>
<td>Igel et al. (2015a)%</td>
<td>3D RAM</td>
<td>1- and 2-moment 5ICE</td>
<td>70-750 m / 65</td>
<td>10 days</td>
<td>RCE</td>
</tr>
<tr>
<td>Igel et al. (2015b)$</td>
<td>2D RAMS</td>
<td>Bin</td>
<td>20 m / 140</td>
<td>0.5 h</td>
<td>Idealized shallow cumulus</td>
</tr>
<tr>
<td>Grant and van den Heever (2015)%</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>25-300 m / 92</td>
<td>3 h</td>
<td>Idealized supercell</td>
</tr>
<tr>
<td>Grant and van den Heever (2014)%</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>100 m - 1.0 km / 57</td>
<td>16 h</td>
<td>Idealized sea breeze convection</td>
</tr>
<tr>
<td>van den Heever and Cotton (2004)&amp;</td>
<td>3D RAMS</td>
<td>5ICE, specified hail diameter</td>
<td>1.0 km / 35</td>
<td>2 h</td>
<td>Idealized supercell</td>
</tr>
<tr>
<td>van den Heever et al. (2006)%</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>500 m / 36</td>
<td>12 h</td>
<td>CRYSTAL-FACE Deep convection</td>
</tr>
<tr>
<td>van den Heever et al. (2007)%</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>1.5 km / 40</td>
<td>26 h</td>
<td>St. Louis urban convection</td>
</tr>
<tr>
<td>van den Heever et al. (2011)%</td>
<td>2D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>1.0 km / 38</td>
<td>100 days</td>
<td>Radiative convective equilibrium</td>
</tr>
<tr>
<td>Herbener et al. (2014)%</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>2.0 km / 56</td>
<td>144 h</td>
<td>Idealized TC</td>
</tr>
<tr>
<td>Seigel et al. (2013)%</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>500 m / 70</td>
<td>7 h</td>
<td>Idealized continental squall line</td>
</tr>
<tr>
<td>Seigel and van den Heever (2013)&amp;</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, varied hail size</td>
<td>500 m / 65</td>
<td>7 h</td>
<td>Idealized continental squall line</td>
</tr>
<tr>
<td>Adams-Selin et al. (2013a)&amp;</td>
<td>WRF</td>
<td>5ICE vs 6ICE</td>
<td>1.0 km / 72</td>
<td>6 h</td>
<td>Oklahoma bow echo</td>
</tr>
<tr>
<td>Adams-Selin et al. (2013b)&amp;</td>
<td>WRF</td>
<td>Multiple WRF schemes</td>
<td>3.0 km / 35</td>
<td>24 h</td>
<td>Oklahoma bow echo</td>
</tr>
<tr>
<td>Storer et al. (2010)%</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>1.0 km / 35</td>
<td>5.5 h</td>
<td>Idealized supercell</td>
</tr>
<tr>
<td>Storer et al. (2013)%</td>
<td>3D RAMS</td>
<td>2-moment 5ICE, Aerosol scheme</td>
<td>1.0 km / 65</td>
<td>100 days</td>
<td>RCE deep convection</td>
</tr>
<tr>
<td>Zhu et al. (2012)&amp;</td>
<td>Multiple models</td>
<td>1- and 2-moment 3ICE</td>
<td>1-2.8 km / 50-92</td>
<td>33 h</td>
<td>TWP-ICE</td>
</tr>
<tr>
<td>Varble et al. (2014a,b)&amp;</td>
<td>Multiple models</td>
<td>1- and 2-moment 3ICE</td>
<td>917 m and 1 km / 76-103</td>
<td>33 h</td>
<td>TWP-ICE</td>
</tr>
</tbody>
</table>
Table 1 Key papers using high-resolution numerical cloud models with bulk microphysics schemes to study the impact of microphysical schemes on precipitation. Model type (2D or 3D), microphysical scheme (one moment or multi-moment), horizontal resolution and number of vertical layers, integration time, and case(s) are listed. Papers with a “*” are used for comparison with the present study, papers with a “#” denote development of a new scheme, papers with a “$” modify/improve existing schemes, papers with a “&” compare different schemes, and papers with a “%” indicate process (budget) studies. TCM3 stands for the “Tropical Cyclone Model with triple nested movable mesh.” RCE is radiative-convective equilibrium, SGP the southern Great Plains, ATEX the, AMMA the African Monsoon Multidisciplinary Analyses, TWP-ICE the Tropical Warm Pool International Cloud Experiment, CRYSTAL-FACE the Cirrus Regional Study of Tropical Anvils and Cirrus Layers - Florida Area Cirrus Experiment, IMPROVE the Improvement of Microphysical PaRameterization through Observational Verification Experiment, EMEX the Equatorial Monsoon Experiment, SCSMEX the South China Sea Monsoon Experiment, KWAIJEX the Kwajalein Experiment, PRE-STORM the Preliminary Regional Experiment for STORM-Central, TOGA-COARE the Tropical Ocean Global Atmosphere Coupled Ocean Atmosphere Response Experiment, GATE the Global Atmospheric Research Program's Atlantic Tropical Experiment, SHEBA the Surface Heat Budget of the Arctic Ocean Experiment, COHMEX the Cooperative Huntsville Meteorological Experiment, and FIRE-ACE the FIRE Arctic Cloud Experiment.
<table>
<thead>
<tr>
<th>Run</th>
<th>Microphysics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graupel</td>
<td>3ICE scheme with graupel option and 1 km horizontal grid</td>
</tr>
<tr>
<td>Hail</td>
<td>3ICE scheme with hail option and 1 km horizontal grid</td>
</tr>
<tr>
<td>4ICE_v0</td>
<td>Original 4ICE scheme and 1 km horizontal grid</td>
</tr>
<tr>
<td>4ICE</td>
<td>Modified 4ICE scheme and 1 km horizontal grid</td>
</tr>
<tr>
<td>4ICE_nec</td>
<td>Modified 4ICE scheme but no rain evaporation correction, 1 km horizontal grid</td>
</tr>
</tbody>
</table>

Table 2 List of numerical experiments.

<table>
<thead>
<tr>
<th>Run</th>
<th>Total Rainfall (mm)</th>
<th>Convective Rainfall (mm)</th>
<th>Stratiform Rainfall (mm)</th>
<th>Stratiform %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2bias</td>
<td>9.74</td>
<td>6.21</td>
<td>3.25</td>
<td>33.5</td>
</tr>
<tr>
<td>Graupel</td>
<td>9.31</td>
<td>3.86</td>
<td>4.90</td>
<td>52.7</td>
</tr>
<tr>
<td>Hail</td>
<td>9.38</td>
<td>5.59</td>
<td>3.27</td>
<td>34.9</td>
</tr>
<tr>
<td>4ICE_v0</td>
<td>8.73</td>
<td>3.95</td>
<td>4.35</td>
<td>49.9</td>
</tr>
<tr>
<td>4ICE</td>
<td>10.30</td>
<td>5.83</td>
<td>4.02</td>
<td>39.1</td>
</tr>
</tbody>
</table>

Table 3 Total rainfall and its convective and stratiform components from an NMQ bias-corrected observational radar network estimate and four NU-WRF simulations using different Goddard bulk microphysical schemes.
### Table 4

Total rainfall coverage (for rain rates greater than the Q2 minimum of 0.15 mm/h) and its convective and stratiform components from the NMQ bias-corrected observational radar network estimate and four NU-WRF simulations using different Goddard bulk microphysical schemes.

<table>
<thead>
<tr>
<th>Run</th>
<th>Total Rainfall Area Coverage in %</th>
<th>Convective Area Coverage in %</th>
<th>Stratiform Area Coverage in %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2bias</td>
<td>26.1</td>
<td>4.4</td>
<td>18.7</td>
</tr>
<tr>
<td>Graupel</td>
<td>25.1</td>
<td>5.0</td>
<td>15.4</td>
</tr>
<tr>
<td>Hail</td>
<td>25.6</td>
<td>4.4</td>
<td>13.7</td>
</tr>
<tr>
<td>4ICE_v0</td>
<td>23.3</td>
<td>4.0</td>
<td>15.4</td>
</tr>
<tr>
<td>4ICE</td>
<td>24.0</td>
<td>5.1</td>
<td>15.0</td>
</tr>
</tbody>
</table>
Figure 1 Hail mapping size thresholds as a function of hail mixing ratio (horizontal axis) and local in cloud temperature (vertical axis). Hail mixing ratios less than the dashed line use a larger intercept (i.e., 0.240 cm$^{-4}$) representative of smaller hail while those greater than the solid line use a smaller intercept (i.e., 0.0048 cm$^{-4}$) representative of larger hail at each given temperature. Intercept values are interpolated for mixing ratios between the two thresholds.
Figure 2 NU-WRF grid configuration. The outer domain (labeled 1 at the center) has a horizontal resolution of 9 km. The middle domain (labeled 2) has a horizontal resolution of 3 km, and the inner domain (labeled 3) has a horizontal resolution of 1 km and covers the southern Plains.
Figure 3 Composited radar reflectivity from (a) NEXRAD observations and the (b) Graupel, (c) Hail, (d) original 4ICE, (e) modified 4ICE, and (f) modified 4ICE with no rain evaporation correction NU-WRF simulations at 10 UTC on 20 May 2011. The precipitation analysis area is indicated by the red boundary. Longitude and latitude values are shown along the horizontal and vertical edges, respectively.
Figure 4  Vertical cross-sections of (a) NEXRAD-observed radar reflectivity and NU-WRF-simulated reflectivity from the (b) Graupel, (c) Hail, (d) original 4ICE, (e) modified 4ICE, and (f) modified 4ICE with no rain evaporation correction simulations at 10 UTC on 20 May 2011. Positions of the cross-sections are shown by the lines in Figure 3 for the radar observations and WRF simulations, respectively. The vertical axes show height in km and the horizontal axes the horizontal distance in km along the cross section.
Figure 5  Surface perturbation potential temperature (color shade) overlaid with radar reflectivity contours from the model simulations (black) and NEXRAD (red). The radar reflectivity contours are for 45 dBZ. Longitude and latitude values are shown along the horizontal and vertical edges, respectively.
Figure 6 Maximum radar reflectivities for (a) NEXRAD and NU-WRF with the (b) Graupel, (c) Hail, (d) original 4ICE, and (e) modified 4ICE microphysics schemes. Vertical axes are heights in km; horizontal axes indicate time from 06 to 12 UTC on 20 May 2011.

Figure 7 Radar reflectivity CFADs from (a) NEXRAD observations and NU-WRF simulations with the (b) Graupel, (c) Hail, (d) original 4ICE, and (e) modified 4ICE microphysics schemes from 06 to 12 UTC on 20 May 2011. Right axes are heights in km; horizontal dashed lines indicate the level of the 0 °C environmental temperature. The thicker solid black lines are overlays of the observed 0.001 and 2.0 frequency contours; the thinner black lines highlight the simulated 2.0 frequency contours.
Figure 8 Components of the modified 4ICE radar reflectivity CFAD shown in Figure 7 (d) from (a) rain, (b) snow, (c) graupel and (d) hail. Horizontal dashed lines indicate the level of the 0 °C environmental temperature.
Figure 9 PDF matching scores for the CFADs in Figure 7. The score indicates the amount of overlap between the simulated and observed PDF at each level.

Figure 10  Vertical velocity CFADs of in-cloud up- and downdrafts in the (a) total, (b) convective and (c) stratiform regions from 06 to 12 UTC on 20 May 2011. Solid lines indicate 0.005% frequencies and dashed lines 1.0% frequencies.
Figure 11 Surface one hour accumulated rainfall from (a) NMQ Q2 Stage IV bias-corrected radar rain estimates and the (b) Graupel, (c) Hail, (d) original 4ICE and (e) modified 4ICE NU-WRF simulations ending at 10 UTC on 20 May 2011. The precipitation analysis area is indicated by the red boundary shown in (a). Longitude and latitude values are shown along the horizontal and vertical edges, respectively.
Figure 12 PDFs of NMQ-observed and NU-WRF-simulated rainfall intensity in mm/h from four different variations of the Goddard microphysical schemes for the (a) total region using a logarithmic scale and (b) total, (c) convective, and (d) stratiform regions using a linear scale. The observed rain rates are estimated from the Stage IV bias-corrected Q2 radar estimates. PDFs were calculated every 10 minutes from both the observed and simulated datasets from 06 to 12 UTC on 20 May 2011 within the analysis domain shown in Figure 3.
Figure 13  Domain- and time-averaged hydrometeor profiles from the (a) Graupel, (b) Hail, (c) original 4ICE, and (d) modified 4ICE schemes from 06 to 12 UTC on 20 May 2011. The horizontal axes show mixing ration in g kg\(^{-1}\).

Figure 14  Same as Figure 13 except for the convective regions.
Figure 15  Same as Figure 13 except for the stratiform regions.