Comparisons of Modeled and Observed Reflectivity and Fall Speeds for Snowfall of Varied Riming Degree During Winter Storms on Long Island, NY

Andrew L. Molthan[1], Brian A. Colle[2], Sandra E. Yuter[3], David Stark[4],

Submitted to the *Monthly Weather Review*

November 20, 2015

Revised:

22 July 2016

[1] NASA Marshall Space Flight Center / Earth Science Office, Huntsville, Alabama

[2] School of Marine and Atmospheric Sciences, Stony Brook University/SUNY, Stony Brook, NY

[3] Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University, Raleigh, NC

[4] NOAA National Weather Service, New York City, NY

*Corresponding author address:
Dr. Andrew L. Molthan
andrew.molthan@nasa.gov
NASA Marshall Space Flight Center / Earth Science Office
320 Sparkman Drive
Huntsville, Alabama, 35805

Abstract

Derived radar reflectivity and fall speed for four Weather Research and Forecasting model bulk microphysical parameterizations (BMPs) run at 1.33 km grid spacing are compared with ground-based, vertically-pointing Ku-band radar, scanning S-band radar, and in situ measurements at Stony Brook, NY. Simulations were partitioned into periods of observed riming degree as determined manually using a stereo microscope and camera during nine winter storms. Simulations were examined to determine whether the selected BMPs captured the effects of varying riming intensities, provided a reasonable match to the vertical structure of radar reflectivity or fall speed, and whether they produced reasonable surface fall speed distributions. Schemes assuming non-spherical mass-diameter relationships yielded reflectivity distributions closer to observed values. All four schemes examined in this study provided a better match to the observed, vertical structure of reflectivity during moderate riming than light riming periods. The comparison of observed and simulated snow fall speeds had mixed results. One BMP produced episodes of excessive cloud water at times, resulting in fall speeds that were too large. However, most schemes had frequent periods of little or no cloud water during moderate riming periods and thus underpredicted the snow fall speeds at lower levels. Short, 1-4 hour periods with relatively steady snow conditions were used to compare BMP and observed size and fall speed distributions. These limited data suggest the examined BMPs underpredict fall speeds of cold-type snow habits and underrepresent aggregates larger than 4 mm diameter.

1. Introduction

As operational numerical weather prediction continues a trend towards finer spatial resolution, bulk microphysics schemes (BMPs) are relied upon to capture numerous microphysical processes and characteristics of resulting precipitation. Several assumptions are made within these schemes, including the shape and related parameters of the particle size distribution, various size-fall speed relationships, and mechanisms for the production of dry or rimed snow, and graupel. Several studies have examined the performance of BMPs by comparing characteristics of simulated ice classes against surface, aircraft, and remote sensing acquired during winter storms. Observations on 3-4 December 2001 during IMPROVE-II showed that BMPs in the Weather Research and Forecasting (WRF, Skamarock et al. 2008) model available at that time tended to overpredict the snow aloft in the snow growth region (Garvert et al. 2005; Lin and Colle 2009). The snow fall speed was found to be too fast in the Purdue Lin (Lin and Colle 2009) and WRF six-class, single-moment scheme (WSM6, Hong et al. 2006) when compared to the Thompson et al. (2004) scheme. The revised Thompson et al. (2008, THOM2) scheme incorporated a new mass-diameter relationship and particle size distribution for snow. Lin and Colle (2011) developed the single-moment Stony Brook scheme (SBU-YLIN), which combines the snow and graupel categories into a single precipitating ice class with corresponding riming factor. When compared against the THOM2 and Morrison et al. (2009, MORR) schemes, Lin and Colle (2011) found that the Stony Brook scheme reduced snow amounts aloft, which compared more favorably with in situ observations acquired over the Oregon Cascades.

Other examinations of simulated and observed snowfall were performed using observations from the Canadian CloudSat/Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) Validation Project (C3VP) in Ontario (Skofronick-Jackson et al. 2015, Petersen et al. 2007). Snowfall observed during the 22 January 2007 event was comprised primarily of lightly-rimed dendrites and their aggregates and sampled by ground-based and aircraft observations. Molthan et al. (2010) used C3VP in situ observations and radar observations from 22 January 2007 to evaluate the Goddard six-class scheme with graupel (Tao et al. 2003; Lang et al. 2007) and Molthan et al. (2012) extended the evaluation to include the WSM6, THOM2, MORR, and SBU-YLIN schemes. Their studies generally concluded that schemes providing greater flexibility in size distribution parameters, density, or additional moments improved performance over the use of constant, assumed parameters. Surface measurements of particle fall speeds during the C3VP event suggested that diameter-velocity parameterizations of the THOM2, MORR, and WSM6 overestimated fall speeds for sizes larger than 1 mm, while the SBU-YLIN scheme produced fall speeds closest to observations (Molthan and Colle 2012). Meanwhile, the Goddard scheme tended to underestimate fall speeds for all sizes (Molthan et al. 2010). Whereas the Goddard scheme tended to underestimate fall speeds in Molthan et al. (2010), Han et al. (2013) found it provided the best agreement with observations acquired in snowfall located above the melting layer of a broader region of stratiform rainfall, which preceded a cold front affecting western California. Shi et al. (2010) evaluated the performance of the Goddard scheme for lake-effect snow observed during C3VP through comparisons of observed and simulated C- and W-band radar reflectivity and AMSU-B brightness temperatures. Comparisons of simulated and observed radar reflectivity demonstrated that for lake-effect bands, the WRF simulation underestimated the echo top height of the observed band and failed to identify numerous, smaller cores of reflectivity. For broader regions of synoptic-scale precipitation, there was a tendency to overestimate the coverage of reflectivity above 20 dBZ. Additional analysis of reflectivity CFADs revealed an overall ability for their simulation to capture the overall large-scale cloud structures but additional refinements to microphysics and smaller scale features were needed.

Regional differences in scheme performance highlighted in the aforementioned studies warrant additional evaluations for other phenomena. Studies have evaluated snow and graupel characteristics within BMPs for events in California, the Pacific Northwest, Appalachians, and southern Ontario, but no known studies to date have examined BMP assumptions in simulations of winter storms in the northeastern United States. Recent studies have documented the evolution of snowfall microphysics in such storms as a precursor to model comparisons. Stark et al. (2013) observed the evolution of ice crystal habits through stereo microscope observations of snow obtained at the surface, corroborated by WSR-88D cross-sections and a vertically-pointing Doppler radar. The degree of riming for ice crystals was assessed from stereo microscope particle images (Mosimann et al. 1994). The degree of riming and prevalence of dendrites increased with snow band maturity and intensity, corresponding to an increase in snow-to-liquid ratio, precipitation, and fall speed. As snow bands passed, weaker ascent and lower relative humidity values corresponded to plate-like crystals, an overall decrease in dendrites, and less riming. Colle et al. (2014) surveyed a dozen winter cyclones across three seasons that impacted the northeastern United States and related snow-to-liquid ratios to predominant crystal habits and degree of riming. Dominant crystal habits and variability in riming were noted in relation to frontal zones and distance from the cyclone center, thus, a single event is likely to be comprised of periods of varying habit, degree of riming, and snow-to-liquid ratio.

In this study, model simulations of events documented by Stark et al. (2013) and Colle et al. (2014) are categorized by the degree of riming present in surface observations of snowfall. For each 15-minute period, an average degree and range of riming was determined by visual inspection of stereo microscope images. This time series is then used to partition radar observations and model output to represent times when light and moderate riming occurred over the observation site. Model performance is then assessed with respect to observed degree of riming from multiple storms and multiple seasons. Four BMPs are selected based upon their diverse means of characterizing snow size distributions, fall speed relationships, means for graupel production, and simulation of riming characteristics (Tables 1 and 2 of Molthan and Colle 2012). Schemes were selected based upon their frequent use in operational numerical weather prediction and for continuity with the previous study to suggest continued improvements in the simulation of winter weather. Since detailed in situ aircraft observations are not available for this multi-season sampling of storms, evaluations are performed against available ground observations and radar remote sensing of reflectivity and particle fall speed. Model simulations of these quantities and comparison to observations will clarify whether these schemes capture variability in size distribution and fall speed during periods of varying riming degree, necessary to improve simulations of winter precipitation. Comparisons will also identify future opportunities for improvement in the simulation of riming processes.

This study is motivated by two key research questions:

- How realistic are selected single- and double-moment WRF BMPs for simulating snow size distributions, fall speeds, and radar reflectivity for observed periods of light and moderate riming during winter storms over Long Island, New York?
- How does the WRF BMP performance change for these categories of observed riming, and when large aggregates are present?

This paper will be organized as follows: Section 2 will discuss the data and methods used in this study. Section 3 will discuss the model verification results, and the conclusions are in section 4.

2. Data and Methods

This study uses several computed and observed variables related to snow particle fall speed, which we define in Table 1 for clarity. Differences in these variables need to be taken into account when comparing among them. Since the instruments available did not directly observe small-scale turbulence (*E)* and vertical air motion (*w*), we can only directly compare among the measured and computed values when we can assume *E* and/or *w* are zero. In low horizontal wind conditions, it is often assumed that w=0 for surface-based in situ instruments such as disdrometers. Surface observations for events described herein were limited to periods of horizontal wind speeds of 5 m s^{-1} or less. In these environments, small-scale turbulence (*E)* will be smaller than typical snowfall events, following Schreur and Geertsema (2008), who estimated E as related to half the squared difference of wind gust and average wind speeds.

a. Observations

Microphysical and radar observations for this study were taken during the 2009-2012 winter seasons at Stony Brook, NY (SBNY, see Colle et al. 2014; their Fig. 1), which is on the north shore of Long Island (LI), approximately 93 km east of New York City, New York (NYC). Stark et al. (2013) and Colle et al. (2014) provide details on the experimental setup and location. A vertically-pointing METEK Ku-band micro rain radar (MRR; Peters et al. 2002) was used at SBNY to observe the profile of reflectivity and Doppler velocities to 7750 m above sea level every minute. The MRR has been used to study winter snowstorms in several locations (Cha et al. 2009; Keighton et al. 2009; Prat and Barros 2010; Kneifel et al. 2011a,b; Xie et al. 2012; Maahn and Kollias 2012, Stark et al. 2013; Colle et al. 2014; Maahn et al. 2014; Pokharel et al. 2014a,b; Garrett et al. 2015). The radar reflectivities from the short wavelength of the radar (1.25 cm) are subject to attenuation in heavier precipitation (Löffler-Mang et al. 1999) and in conditions when wet snow builds up on the antenna (Stark et al. 2013). The latter is the more relevant for conditions during snowstorms. The MRR data were post-processed to improve sensitivity and data quality using the method of Maahn and Kollias (2012).

Observations from the MRR were supplemented by the WSR-88D radar at Upton, NY (KOKX). Vertical profiles of interpolated WSR-88D reflectivity were computed from Level II KOKX data for the vertical column nearest the verification point of each model simulation. The WSR-88D radar data have coarser native sensor spatial resolution, about 500 m in the vertical and horizontal at the 30 km range over the SBU site as compared to the MRR resolution volume size of 250 m in the vertical and ~100 m in the horizontal. For convenience in generating comparisons, the WSR-88D data were interpolated to a Cartesian grid with vertical and horizontal spacing of 250 m and 100 m, respectively.

A Particle Size and Velocity (PARSIVEL; Löffler-Mang and Joss 2000; Löffler-Mang and Blahak 2001; Yuter et al. 2006) disdrometer was placed about 1 m above the 1-storey roof surface to collect hydrometeor size and fall speed distributions. Battaglia et al. (2010) note that the PARSIVEL measures a "PARSIVEL diameter" based on the maximum shadowed area of the particle as it passes through the disdrometer laser beam. In a limited set of conditions where the snow flake is horizontally aligned, this measurement is equivalent to the widest horizontal dimension of the snow particle, otherwise, the PARSIVEL diameter represents an estimate of widest horizontal diameter with an error less than or equal to 20%. In calm conditions, the PARSIVEL measurement of particle fall speed is equivalent to settling speed (Table 1). Battaglia et al. (2010) determined that the PARSIVEL measured fall speed has a variance less than 20% for individual particles and tends to underestimate the mean fall speeds of smaller particles. The larger errors in their reported fall speeds are less relevant here as their instruments were typically operated in windy conditions. Analysis herein was restricted to winter storms with winds less than 5 m s^{-1} to avoid the potential for blowing snow from the surface impacting the results (D. Kingsmill, personal communication, 2011), and to emphasize periods of reduced small-scale turbulence.

In order to further characterize precipitation during these events, a stereo microscope and camera were used to observe the snow habit and riming intensity at the

SBU site as described in Colle et al. (2014) and Stark et al. (2013). The ice habits were categorized into several main types (needles and columns, dendrites, plates, side planes and bullets), and riming was categorized as light, moderate, or heavy. Heavy riming did occur during short intervals within three sampled storms, but the sample size of heavy riming was insufficient for a comprehensive analysis.

b. WRF simulations

The Weather Research and Forecast Model (WRF; Skamarock et al. 2008) version 3.3 was utilized for simulations of several of the observed winter storms. The North American Mesoscale (NAM) model analysis data at 12 km grid spacing (NAM 218 hereafter) and six-hourly time increments were used as initial and boundary conditions in the majority of the simulations, though in a limited number of events, the Global Forecasting System (GFS) analysis data at 0.5^{o} grid spacing were used because simulations with the NAM 218 data were too dry, or precipitation placement was not in agreement with observations. Sea surface temperature and snow cover data were included in these initial and boundary condition datasets at model initialization. The WRF was run using an outermost 36 km resolution domain with one-way nesting for three inner domains at 12 km, 4 km, and 1.33 km grid spacing as illustrated in Figure 1. The 1.33 km domain was used in the analysis for this paper. Thirty-nine vertical levels were used, and the top of the model was 100 hPa. Model physics included the Betts-Miller-Janjic cumulus scheme (Betts and Miller 1993; Janjic 1994) on the 36 km and 12 km domains, Yonsei University (YSU, Hong et al. 2006) planetary boundary layer scheme, and the Unified Noah land surface physics (Ek et al. 2003). Within the 4 and 1.33 km resolution domains, a convective parameterization was not used, and all cloud or precipitation processes were simulated with the WSM6 (Hong et al. 2006), THOM2 (Thompson et al. 2008), SBU-YLIN (Lin and Colle 2011), or MORR (Morrison et al. 2009) bulk microphysics schemes. Molthan and Colle (2012; their Table 1) provide a detailed overview of the characteristics of snowfall within the WRF v3.3 schemes used in this study. The WRF model and BMPs were specially configured to output the particle size distribution intercept (N_{os}) and slope parameter (λ_s) of snowfall size distributions, along with parameters necessary to obtain the radar reflectivity and radar reflectivity-weighted fall speeds of precipitating species in each scheme. Molthan and Colle (2012) provide details on the derivation of model reflectivity, size distribution parameters, and fall speeds.

A list of cases simulated, their respective initialization times, and initial conditions are given in Table 2, which represents a subset of a larger number of storms evaluated by Colle et al. (2014). The verification point in the WRF model was obtained through a bilinear interpolation of 1.33 km resolution grid boxes nearest to SBNY in each simulation. For the simulations of 19-20 December 2009, the simulated heavy snow band was approximately 58 km southwest of the actual location. In this case, a representative point for SBNY was selected relative to the simulated snow band. The verification points for each simulated case and BMP are shown in Table 3. With the exception of the 7 January 2011 event (~4 hours), each simulation included at least six hours of spin-up time to generate precipitation prior to verification. Other simulations of the 7 January 2011 event with a longer start-up time did not capture the precipitation that occurred over SBNY.

c. Comparisons of volumetric characteristics

Derivations of the model-derived reflectivities and fall velocities are straightforward and computed using assumptions consistent with each of the BMPs (Molthan and Colle 2012). Model-simulated properties were calculated for WRF grid boxes with at least 0.001 g kg^{-1} of hydrometeor mixing ratio, thus, reflectivity and fall speed distributions correspond to model volumes with at least a trace of snow, graupel, or rain. Comparison of model output to observations is more complex as there are several limitations of the observations that preclude direct comparison. As noted previously, MRR observed reflectivity is subject to attenuation when snow accumulates on the MRR antenna. The differences in sensor spatial resolution between the MRR and WSR-88D will manifest most strongly when the storm structure is more spatially heterogeneous and non-uniform beam filling is present (Rinehart 1991). Though many schemes represent sub-grid variability in clouds through a cloud fraction defined in both microphysics and radiation schemes, their representation is not sufficient to account for the same effects of a non-uniformly filled or partially-filled radar resolution volume. In addition, the model-derived reflectivities are not subject to instrument sensitivity constraints and can be computed for lower precipitation ice concentrations than can be detected by either of the two radars. The cm-wavelength MRR and WSR-88D radars do not have sufficient sensitivity to observe the non-precipitating portions of cloud.

We compare the radar reflectivity in the vertical column from the WRF model directly over the SBU measurement site to the radar reflectivity observed by the MRR and to the vertical column of WSR-88D data from KOKX taken over the site. Simulations from the innermost, 1.33 km domain are separated into whether there was light, moderate or heavy riming observed at SBNY (Colle et al. 2014). The set of these profiles are accumulated into an asynchronous volume of data from which joint frequency distributions of reflectivity and height using contoured frequency by altitude diagrams (CFADs; Yuter and Houze 1995). We truncate the observed CFADs for the MRR and WSR-88D at the altitude where the number of samples is less than 20% of the maximum number of samples at one level in the volume (Yuter and Houze 1995). Some differences existed in the specific timing between the simulated and observed precipitation as shown in Figure 2 .

Similarly, we compare joint frequency distributions of measures of snow fall speed with height between the model column over the measurement site and the MRR. The MRR Doppler velocity is not directly equivalent to the model's computed mean fall speed. Errors in any combination of model vertical air motion, size distribution, size-fall speed relation, and/or particle density would yield errors in the model-computed mean reflectivity-weighted fall speed (V_{cf}, Table 1).

As a net result of these differences, we do not expect close quantitative matches between the model and observed Z or fall speed variables. Rather, we focus on the degree of agreement in the trends of the modes of the distributions with height and changes in width of the distribution with height. We also note large differences in trends in maximum Z values with height. The interplay among the model kinematics, microphysics and latent heat release is such that it is not possible to attribute differences solely to individual components within the microphysics parameterizations such as size distributions and particle densities.

3. Model Microphysical Evaluation

a. Evaluation of simulations during observed light riming events

We first examine combined statistics from 21 occurrences of light riming within nine events (Fig. 2). For these time periods with the light riming designation, surface stereo microscope observations indicate that less than 1% of particles were graupel (Colle et al. 2014; their Figure 6). Colle et al. (2014) showed that cold type crystals (side planes and bullets, plates, and needles) were dominant (~80%) during observed, light riming periods (~80%), whereas the schemes examined assume slower-falling dendritic habits. Mean profiles and CFADs of simulated hydrometeor categories are shown in Figures 3 and 4, respectively, for simulations sampled during observed periods of light riming. Mean profiles of cloud ice for the SBU-YLIN and MORR schemes are similar (Figs. 3a, 3d), and frequently less than 0.05 g kg^{-1} within an altitude range of 4-9 km (Figs. 4b, 4n). The WSM6 produces cloud ice throughout the column (Fig. 3b), in sharp contrast to the THOM2 scheme, which produces the smallest amount of cloud ice, confined to 6-10 km (Fig. 3c). In the THOM2 simulations, cloud ice mixing ratios were smaller than 0.05 g kg^{-1}, or a single CFAD joint histogram cell size (Fig. 4j). Though the four schemes differ in their partitioning of ice mass into cloud ice, snow, or precipitating ice, they produce a similar vertical distribution of total snow and ice mixing ratios. Partitioning of mixing ratios among these categories exaggerates some of their differences. Rather than simulating small crystals through production of cloud ice mixing ratio, the THOM2 applies a bi-modal size distribution within the simulated snow category. The MORR, WSM6, and THOM2 schemes produce mean cloud water profiles of 0.02 g kg^{-1} or less within the lowest 4 km (Figs. 3a-c), where mean temperatures range from -15 to 0^{o}C (Fig. 5a). The SBU-YLIN scheme increases mean cloud water throughout the column, to 0.06 g kg^{-1} at 4 km. Increased cloud water mixing ratio continues through 8 km, inconsistent with surface observations of lightly rimed particles (Fig. 3d). Cloud water CFADs capture infrequent amounts of cloud water content greater than mean values for each scheme, predominately in the lowest 4 km (Fig. 4). Each scheme shows the greatest increases in mean snow (or precipitating ice) mixing ratios between echo top and 4 km where the rate of increase slows and then decreases toward the surface. The MORR and WSM6 schemes produce graupel with mean profile amounts greater than 0.01 g kg^{-1} confined to the lowest 4 km (Figs. 3a-b), with infrequent occurrence of amounts exceeding 0.25 g kg^{-1} (Figs. 4d, 4h). The THOM2 simulations produced very small amounts of graupel with mean values less than 0.001 g kg^{-1}. Graupel in produced by several and different processes within the MORR, WSM6, and THOM2 schemes, however, detailed microphysical process budgets for each simulation are beyond the scope of this study. The SBU-YLIN scheme represents snow and graupel through a rimed precipitating ice category but produces excessively high riming intensities as a result of excessive cloud water (Fig. 6a). The THOM2 scheme produces a larger mean profile of snow and occasional, larger amounts of snow mixing ratio than the MORR and WSM6 schemes that produce graupel. Median, liquid-equivalent precipitation from these schemes ranges from 0.1-0.2 mm h^{-1} with the highest amounts resulting from the WSM6 scheme (Fig. 7).

Figure 8 shows the frequency distributions of observed MRR and WSR-88D reflectivity (Z) along with values derived from WRF simulations during light riming periods. In the lowest 2 km, the most frequently occurring (modal) values of MRR reflectivity are around 16-20 dBZ (Fig. 8a), comparable to the WSR-88D reflectivity within the same altitude range (Fig. 8b). As compared to the MRR, the WSR-88D has a broader range of Z values at each height, with small occurrences of values that exceeded those observed by the MRR. The cause for the lack of Z values greater than 24-28 dBZ in the MRR data is not clear.

Simulated reflectivity is highly variable and lacks the distinct modes observed by the MRR and WSR-88D, suggesting that observed precipitation structures are more uniform at various altitudes than the corresponding model simulations during light riming events (Fig. 8c-f). Both the modal and maximum Z values observed by the WSR-88D increased between 6 km and 1 km altitude consistent with increase in particle sizes via depositional growth and aggregation. The increase was from 8 to 18 dBZ for the modal values and 18 to 34 dBZ for the maximum values. The THOM2 scheme has a similar trend of maximum values with height while the other three schemes have maximum reflectivities that are too high from 4-6 km. The THOM2 and SBU-YLIN schemes produce reflectivity distributions comparable to observations in the lowest 4 km while the WSM6 and MORR schemes exceed the observed reflectivity distribution from the WSR-88D. The higher reflectivity values in the WSM6 and MORR simulations than observed likely result from the prediction of graupel, representing moderate to heavily rimed particles in contrast to observed, light riming.

Figure 9 shows the distributions of observed Doppler velocity and simulated fall speed variables. Throughout the vertical column, there is a fairly consistent range in Doppler velocities observed by the MRR, from 0.3-2.0 m s^{-1} and above 4 km altitude, the mode in MRR Doppler velocity is less distinct than lower levels (Fig. 9a). The most frequently observed Doppler velocities of around 1.0 m s^{-1} are consistent through an altitude of 4 km. There is broadening of the observed fall speed distribution to nearly

1.75 m s^{-1} below 2 km associated with the increased particle growth and reflectivity increase in this layer. The MORR and THOM2 fall speed distributions are narrower than observed below 2 km altitude (Figs. 9b, 9d), while the SBU-YLIN (Fig. 9e) has a second mode below 5 km altitude at high fall speed values, inconsistent with observations. Schemes incorporating a temperature-dependent size distribution, such as the WSM6, THOM2, and SBU-YLIN have trends of increasing fall speed with decreasing height between 3 and 7 km. The MORR scheme has less of a change in fall speed with height and is more consistent with observations. Unfortunately, lack of data from the MRR in the lowest 1 km precludes validation of the ~0.2 m s^{-1} increase in fall speed peak frequency for the MORR. At 1-2 km, where the MRR provides observations, the MORR and WSM6 simulations provide the best match to MRR fall speeds (~1.0 m s^{-1}) while the THOM2 and SBU-YLIN slightly underestimated fall speeds by around 0.25 m s^{-1}.

b. Evaluation of simulations during observed moderate riming events

During moderate riming periods (21 occurrences within seven events; Fig. 2) the observed snow contained about 50% dendrites and plates, 20-25% needles, less than 10% cold type crystals, and small amounts (< 4%) of graupel (Colle et al. 2014). All schemes increase their predicted mean snow and combined cloud ice and snow mass (Fig. 10), shown as higher frequency of larger mixing ratios, particularly in the lowest 3-4 km (Fig. 11). Cloud water also increases in all schemes, particularly in the lowest 2 km, along with increases in the mean profile through 6 km. Increases occur at a range of temperatures from -15°C to 0°C (Fig. 5a), where the simulations for these moderate riming events average as much as 1.5°C warmer than light riming cases (Fig. 5b). The MORR scheme exhibits an increase in mean cloud water and frequency from 1-4 km where mean temperatures range from -15 to -5°C, up to 1.5°C warmer than light riming cases. Overall, the most frequent mixing ratio of cloud water remains less than 0.1 g kg^{-1}, and the mean value is 0.05 g kg^{-1} or less for all schemes except the SBU-YLIN. The simulated cloud water amount is less than expected for periods of moderate riming. For example, Lin and Colle (2011) and Lin et al. (2011) showed for two cases over the Washington Cascades that observed and simulated cloud water was 0.1 to 0.3 g kg^{-1} for moderate riming periods. Some of the cloud water in the SBU-YLIN scheme appears to be erroneously high (> 0.3 g kg^{-1}), resulting in heavily-rimed precipitating ice, increased precipitation, and decrease in snow mass by fallout in the lowest 3 km. Excess production of cloud water may be related to issues with the saturation adjustment process in the SBU-YLIN scheme (Molthan and Colle 2012). Although the aforementioned schemes produce some additional snow and cloud water, the MORR and WSM6 simulations produce amounts and frequencies of graupel comparable to light riming simulations with similar maximum values and frequencies (Figs. 11d, 11h). Mean profiles of graupel in moderate riming events are similar to light riming events for the MORR, WSM6, and THOM2 profiles (Fig. 10a-c).

The CFAD of MRR reflectivity has a distribution mode below 2 km altitude that is sharper and greater than light riming periods (20 dBZ, Fig. 12a). The modal value of WSR-88D reflectivity in the lowest 2 km is similar at 20-24 dBZ (Fig. 12b). Both MRR and WSR-88D distributions of Z indicate a steady increase in the modal value of reflectivity with decreasing altitude from 6-2 km. Such "diagonalization" of the reflectivity CFAD indicates growth of particles as they descend (Yuter and Houze 1995). As with light riming cases, the CFAD of WSR-88D reflectivity included small frequencies of higher reflectivity near the surface, as high as 32-34 dBZ. Similarly, the MRR data included small frequencies of reflectivity from 24-28 dBZ. All four models exhibit clear modes in the reflectivity distribution that increased with decreasing height, similar to the diagonalization seen in observations, though their modal values and ranges differ from the MRR and WSR-88D data. The MORR and WSM6 schemes (Figs. 12c-d) produce modal and maximum reflectivity exceeding WSR-88D and MRR observations near the surface (Fig. 12c-d). The THOM2 and YLIN schemes are a better match to observed reflectivity values and trends with altitude in terms of modal and maximum values (Figs. 12e-f).

The MRR Doppler velocity distribution produces a mode that increases with decreasing altitude from to 1.0 m s^{-1} at 4 km altitude to 1.25 m s^{-1} at the surface (Fig. 13a). The mode in the MRR near-surface Doppler velocity increases by 0.25 m s^{-1} versus the light riming periods. In the WSM6 simulations, an increase in predicted snow mass and larger particles inferred from radar reflectivity contributes to an overall increase in mean fall speeds. Excessive fall speeds above 5 km in WSM6 likely result from erroneously large particles associated with simulated reflectivity greater than observed by the MRR and WSR-88D.

The MORR and THOM2 schemes produce vertical profiles of modal fall speeds for the moderate riming periods similar to their performance during light riming periods despite increases in snow and graupel content (Figs. 9 and 13). In WSM6, the modal fall speeds increase between 6 km and 2 km but do not change much above or below that layer. The inference is that increases in snow content from additional riming did not translate to increases in fall speeds through changes in their diameter-fall speed relationships. In addition, cloud water is likely underpredicted during moderate riming events, contributing to an underprediction of fall speeds at lower levels even if the schemes accounted for varied riming conditions. In SBU-YLIN, riming effects are allowed to influence fall speeds, but excessive cloud water contributed to high riming factors and exacerbated fall speed errors previously observed in light riming events (Fig. 6b). These errors resulted in isolated occurrences of fall speeds of 1.5-3 m s^{-1}, comprising as much as 15% of fall speeds in the lowest 1-2 km (Fig. 13e).

c. Surface size distribution and fall speed

In Figures 14-16, we compare observed and simulated surface size distributions and fall speed measures for four short cases from 1 to 4 hours in duration. These cases highlight some details of the representation of particles as a function of degree of riming and whether aggregation is present. The BMP scheme size distributions are average values for the set of 15-minute intervals in each case and are compared to the observed distribution of PARSIVEL diameter (Section 2a). For context, we also show average mixing ratio profiles for snow, cloud water and graupel in the lowest 3 km. We compare distributions of PARSIVEL fall speed (or settling speed, w=0 and E≠0, Table 1), to a computed mean mass-weighted fall speed (w=0, E=0, Table 1) for each 15 min model-simulated period that includes contributions from snow, graupel, and rain.

i. No riming

A period of mainly cold type crystals (51% side planes and 20% bullets) occurred from 0145 to 0500 UTC on 16 February 2010, with little or no riming observed. A small amount of plates (~10%) and columns (~10%) were also observed with light riming. The observed and simulated size distributions for this period are illustrated in Figure 14a. All four BMPs slightly underestimate the number concentrations of aggregated snow particles with D < 8 mm, with the MORR and SBU-YLIN closest to the observed for diameters from 6 to 8 mm. Each scheme simulates between 0.1 and 0.15 g kg^{-1} of snow above 1 km, but decreases the amount to around 0.05 g kg^{-1} at the surface (Fig. 14b). This reduction results from sublimation, since these cold type habits are mainly observed near the outer edges of the comma head (Colle et al. 2014), where the low-levels are still moistening. The reduction in snow mixing ratio near the surface may contribute to an overall reduction in particle number concentrations as shown by BMPs that predict lower number concentrations of particles across all observed size bins.

The observed fall speed distribution is generally between 1 and 1.5 m s^{-1} (Fig. 14c). The MORR and SBU-YLIN schemes have particle fall speeds clustered around 1 m s^{-1}, with a few values at ~1.25 m s^{-1} (Fig. 14d), while the WSM6 and the THOM2 produces fall speeds slower than the peak in the observations, ranging from 0.5 to 0.75 m s^{-1}. These schemes simulate a small amount of cloud water (< 0.05 g kg^{-1}) above 1 km (Fig. 14b). There was little riming observed during this event as well, so the observed fall speeds (tail > 1.5 m s^{-1}) are likely related to faster falling cold type crystals compared to the conventional plates and dendrites used in these schemes. Underestimation of surface fall speeds in this sample of observed crystals is comparable to the underestimate of fall speeds in the broader sampling of light riming simulations (Fig. 9), therefore, schemes may not be accounting for faster fall speeds for cold type crystals. The SBU-YLIN includes a temperature dependence term for fall speeds, but it is based on the local temperature and not necessarily where the snow particles are formed, and the observed side planes and bullets are likely formed in the middle and upper levels of the cloud.

ii. Light to moderate riming

A mix of 70% plates and 16% side planes was observed from 1000 to 1200 UTC on 21 February 2011 with observed riming intensities that range from none to moderate, with the peak riming intensity occurring at 1115 UTC. The WSM6 scheme slightly underestimates number concentrations of particles across all sizes, and the SBU-YLIN simulates a greater number concentration of particles than observed for all sizes, with the exception of particles around 2 mm (Fig. 15a). The MORR and THOM2 schemes produce number concentrations of particles similar to observations for diameters 1 mm or greater (Fig. 15a).

Each scheme simulates snow mixing ratios of 0.15 to 0.25 g kg^{-1} near the surface (Fig. 15b) and the SBU-YLIN scheme has a small amount (< 0.05 g kg^{-1}) of cloud water between 1.5 and 2 km. However, in general, all schemes have very little cloud water, which likely contributes to the lack of fall speeds greater than 1.5 m s^{-1} (Fig. 15d). Most of the observed fall speeds in this two-hour period were between 1 and 1.5 m s^{-1} (Fig. 15c). All BMPs concentrate their fall speeds around 1 m s^{-1} (Fig. 15d), while the observed peak was slightly greater at ~1.1 m s^{-1} (Fig. 15c). The small amount of cloud water within the SBU-YLIN scheme resulted in an increase in the diagnosed riming factor for the precipitating ice class and some of the greater fall speeds. This also likely contributes to larger standard deviations in the particle size distribution when the scheme transitioned from between periods of rimed and unrimed precipitation.

iii. Light to moderate riming and many aggregates

Figure 16a shows the observed and simulated size distributions for a time period with 65% dendrites and 20% plates observed from 1530 UTC 26 February 2010 to 0000 UTC 27 February 2010. During this period the observed degree of riming is light to moderate (0.5 to 2.5) and many aggregates of dendrites are also observed (not shown). The particle size distributions from the examined BMPs are similar to the observed size distributions for particles smaller than 4 mm, but the BMPs underestimate the number concentrations of particles larger than 4 mm. The particle size distribution from the MORR scheme produces slightly larger particles than the other BMPs, and thus a better fit to observations. One hypothesis for the underestimation of the number concentrations of larger ($D > 4$ mm) particles is a poor representation of snow aggregation, or the shift to particle size distributions comprised of larger particles at the expense of smaller crystals (Fig. 16b). Each scheme produces 0.05 to 0.1 g kg^{-1} of snow below 1 km but varied in their production of cloud water, ranging from 0.05 to 0.15 g kg^{-1} of cloud water between 1 and 2 km. Riming of snow is implied by the colocation of snow and cloud water in the models. An increase in snow content rather than cloud water might have contributed to larger numbers of larger particle sizes in modeled size distributions.

Unlike the above cases and the combined CFAD results, the THOM2, MORR, and SBU-YLIN schemes are able to produce more cloud water (0.05 to 0.15 g kg^{-1}). The distribution of observed fall speeds peak at ~0.75 m s^{-1} and ~1.1 m s^{-1} (Fig. 16c), with a tail to fall speeds exceeding 2 m s^{-1}. Model-simulated fall speeds are clustered between 0.8 and 1.1 m s^{-1} (Fig. 16d). The better fall speed prediction in the model, including some fall speeds greater than 1 m s^{-1} is likely the result of better simulation of the cloud water. The scheme with the least amount of cloud water (WSM6) has the worst fall speed prediction near the surface (peaking around 0.8 m s^{-1}). The THOM2 has several periods with fall speeds from ~2 to 3 m s^{-1} for mixing ratios below 0.1 g kg^{-1} (Fig. 16d). These faster fall speeds result from trace amounts of faster-falling graupel, or brief production of drizzle with 0.005 to 0.01 g kg^{-1} of rain water simulated at the surface between 2215 and 0000 UTC (not shown). Aggregates and lightly rimed snow likely comprised the smaller peak (~0.75 m s^{-1}), while the moderately rimed habits likely fell within the second peak (~1.1 m s^{-1}) and the tail of higher fall speeds. As observed in the CFAD analysis, larger riming factors in the SBU-YLIN scheme contribute to faster fall speeds of 1.25 to 1.75 m s^{-1} for predicted snow mixing ratios less than 0.1 g kg^{-1} (Fig. 16d).

4. Conclusions

Reflectivity and fall speed from four BMPs (MORR, WSM6, THOM2, and SBU-YLIN) run down to 1.33 km grid spacing within the WRF model were compared to vertically-pointing radar observations at Stony Brook, NY for nine snow events that were partitioned into periods of observed riming intensity. Comparisons of observed and modeled particle size distributions and fall speeds at the surface were made for selected periods with distinct sets of crystal habits. Motivating research questions sought to examine whether the selected schemes were able to reproduce key characteristics of the observed distributions of reflectivity and fall speed within various categories of observed riming. In light riming periods, the WSM6 and MORR schemes produced larger reflectivities (Z) than observed, particularly in the lowest 4 km where they produced higher-density graupel particles inconsistent with the light degree of riming observed at the surface. The THOM2 scheme only produced trace, insignificant amounts of graupel and the SBU-YLIN scheme limited the occurrence of higher riming factors, with a better representation of observed WSR-88D reflectivity in the surface to 4 km altitude range. These results encourage a more detailed examination of graupel sources within the WSM6 and MORR simulations to reduce the generation of higher-density graupel particles in periods of observed, light riming. For fall speed variables, MRR observations captured a consistent mode around 1 m s^{-1} in the lowest 1-2 km. The THOM2 and SBU-YLIN schemes produced a consistent mode in the same range, but underestimated fall speeds by around 0.25 m s^{-1}. An analysis of the 16 February 2010 event suggests that this may be related to the underpredicted fall speeds of cold-type snow habits.

During moderate riming, the THOM2 and SBU-YLIN were both closer to Z observations in terms of the most frequent and maximum values at varying altitudes. The MORR and WSM6 simulations produced modal and maximum values of Z that exceeded observations. As in light riming cases, the MORR and WSM6 schemes produced higher-density graupel within the lowest 4 km, which may contribute to their much higher than observed values of radar reflectivity. The MORR, WSM6, and THOM2 schemes increased the amount of cloud water between 1 and 4 km as the observed riming intensity increased from light to moderate categories, but the amounts were less than expected for moderate riming, based on previous field studies. The SBU-YLIN scheme produced the largest mean cloud water profile and infrequent occurrences of higher cloud water amounts from 4-6 km, contributing to an increase in particle density and slight overestimation of radar reflectivity.

There was relatively little change in the near-surface fall speeds with increasing riming category among the MORR, WSM6, and THOM2 simulations, and as a result modeled fall speeds underestimated MRR-obtained surface fall speeds by 0.25 to 0.5 m s^{-1}. These schemes may not produce enough cloud water during these winter storms to capture the observed riming processes (e.g., there is little cloud water in the THOM2 scheme), or the snow fall speeds do not account for increases in particle fall speeds during observed moderate riming periods, likely because they assume dry snow fall speeds unless the scheme contributes mass to the graupel category. The THOM2 and WSM6 schemes incorporate temperature-dependent particle size distribution characteristics for snowfall that provided an increase in fall speeds approaching the surface but were ~0.25 to 0.50 m s^{-1} slower than surface observations. The MORR and THOM2 simulations produce similar fall speed distributions with height regardless of riming category.

Comparisons between surface and simulated particle size distributions suggest that while BMPs simulate reasonable number concentrations of particles smaller than 4 mm, there was an underestimate of the number concentrations particles larger than 4 mm. It is hypothesized that underestimation of the number of larger particles results from BMPs not adequately simulating the aggregation process, or by allowing for particle size distribution parameters to evolve to smaller intercept and slope parameters as aggregation occurs. Fixed parameters of density and spherical shape may be a limiting factor, though schemes such as THOM2 and SBU-YLIN that incorporate variable bulk density did not markedly improve performance. Lower density aggregates cannot be well represented in a BMP that assumes a fixed density for snow, such as used in the WSM6 and MORR schemes. However, the double moment MORR scheme seemed to provide a better comparison to observations than other schemes during periods of aggregation, perhaps benefitting from greater flexibility in the determination of size distribution parameters by predicting both mass and total number concentration. Schemes that predicted higher-density and more heavily rimed graupel particles during light and moderate riming periods resulted in excessive radar reflectivity contrary to radar observations and the lack of these heavily rimed particles at the surface. However, schemes that produced more unrimed snow were not able to capture increases in fall speed during observed moderate riming periods, suggesting that they were unable to predict the observed changes in riming degree.

For schemes that favor the production of higher-density graupel rather than unrimed snow, future work should examine opportunities for a smoother transition between the dry and heavily rimed ice categories to improve the representation of a broader range of riming categories. Additional vertical levels should be included to better capture convective-scale processes contributing to the development of cloud water or small-scale microphysical variability. Such an examination would be best achieved with detailed in situ observations from multiple riming regimes to guide improvements, supplemented by additional radar remote sensing comparisons where in situ data are unavailable.

Acknowledgements

The authors thank Spencer Rhodes for assistance with the processing of MRR data. Financial support for this research project was given by the National Science Foundation

(ATM-1347491 for Colle and ATM-1347499 for Yuter), as well as the NASA Precipitation Measurement Missions Program (NNX13AF88G for Colle and Molthan). A special thanks to Dr. Gordon Taylor for allowing us to use his microscope and camera throughout the observational portion of this study. The authors also thank three anonymous reviewers who provided suggestions that improved the clarity of discussions and improved the related figures.

References

Battaglia, A., E. Rustemeier, A. Tokay, U. Blahak, and C. Simmer, 2010: PARSIVEL snow observations: A critical assessment. *J. Atmos. Oceanic Technol*., **27**, 333–344.

Betts, A. K., and M. J. Miller, 1993: The Betts-Miller scheme. The representation of cumulus convection in numerical models, *Meteor. Monogr*. **24**, 107-121.

Cha, J., K. Chang, S. Yum, and Y. Choi, 2009: Comparison of the bright band characteristics measured by Micro Rain Radar (MRR) at a mountain and a coastal site in South Korea. *Adv. Atmos. Sci.*, 26, 211–221, doi:10.1007/s00376-009-0211-0.

Colle, B. A., D. Stark, and S. E. Yuter, 2014: Surface microphysical observations within east coast winter storms on Long Island, New York. *Mon. Wea. Rev.*, **142**, 3126-3146.

Doviak, R. J., and D. S. Zrnic, 1993: Doppler radar and weather observations, 2nd edition, Academic Press, New York, 562 pages.

Ek, M., K. E. Mitchell, Y. Lin, E. Rogers, P. Grunmann, V. Koren, G. Gayno, and J. D. Tarpley, 2003: Implementation of Noah land surface model advances in the National Centers for Environmental Prediction operational mesoscale Eta Model. *J. Geophys. Res.*, **108,** 8851, doi:10.1029/2002JD003296.

Garrett, T. J., S. E. Yuter, C. Fallgatter, K. Shkurko, S. R. Rhodes and J. L. Endries , 2015: Orientations and aspect ratios of falling snow. *Geophys. Res. Lett.*, doi:10.1002/2015GL064040.

Garvert, M. F., B. A. Colle, and C. F. Mass, 2005: The 13–14 December 2001 IMPROVE-2 Event. Part I: Synoptic and mesoscale evolution and comparison with a mesoscale model simulation. *J. Atmos. Sci.*, **62**, 3474–3492.

Han, M., S. A. Braun, T. Matsui, and C. R. Williams, 2013: Evaluation of cloud microphysics schemes in simulations of a winter storm using radar and radiometer measurements. *Journal of Geophysical Research: Atmospheres*, **118**, 1401-1419, doi:10.1002/jgrd.50115

Hong, S.-Y., J. Dudhia, and J.-O. J. Lim, 2006: The WRF single-moment 6-class microphysics scheme (WSM6). *Journal of the Korean Meteorological Society*, **42**, 129–151.

Janjic, Z. I., 1994: The step-mountain eta coordinate model: Further development of the convection, viscous sublayer, and turbulence closer schemes. *Mon. Wea. Rev.*, **122**, 927-945.

Keighton, S., and Coauthors, 2009: A collaborative approach to study northwest flow snow in the Southern Appalachians. *Bull. Amer. Meteor. Soc.*, **90**, 979–991, doi:10.1175/2009BAMS2591.1.

Kneifel, S., M. S. Kulie, and R. Bennartz, 2011a: A triple-frequency approach to retrieve microphysical snowfall parameters. *J. Geophys. Res.*, **116**, D11203, doi:10.1029/2010JD015430.

Kneifel, S., M. Maahn, G. Peters, and C. Simmer, 2011b: Observation of snowfall with a low-power FM-CW K-band radar (Micro Rain Radar). *Meteor. Atmos. Phys.*, **113**, 75–87, doi:10.1007/ s00703-011-0142-z.

Lang, S.,W.-K. Tao, R. Cifelli, W. Olson, J. Halverson, S. Rutledge, and J. Simpson, 2007: Improving simulations of convective systems from TRMM LBA: Easterly and westerly regimes. *J. Atmos. Sci.*, **64**, 1141–1164.

Lin, Y., and B. A. Colle, 2009: The 4–5 December 2001 IMPROVE-2 Event: Observed microphysics and comparisons with the Weather Research and Forecasting Model. *Mon.Wea. Rev.*, **137**, 1372–1392.

Lin, Y., and B. A. Colle, 2011: A new bulk microphysical scheme that includes riming intensity and temperature dependent ice characteristics. *Mon. Wea. Rev.*, **139**, 1013–1035.

Löffler-Mang, M., M. Kunz, and W. Schmid (1999), On the performance of a low-cost

K-band doppler radar for quantitative rain measurements, *J. Atmos. and Ocean. Tech.*, **16**, 379–387, doi:10.1175/1520-0426(1999)0162.0.CO;2.

Löffler-Mang, M., and J. Joss, 2000: An optical disdrometer for measuring size and velocity of hydrometeors. *J. Atmos. and Ocean. Tech.*, **17**, 130-139.

Löffler-Mang, M. and U. Blahak, 2001: Estimation of the Equivalent Radar Reflectivity Factor from Measured Snow Size Spectra. *J. Appl. Meteorol*., **40**, 843-849.

Maahn, M., and P. Kollias, 2012: Improved micro rain radar snow measurements using Doppler spectra post processing. *Atmos. Meas. Tech.*, **5**, 2661–2673, doi:10.5194/amt-5-2661-2012.

Maahn, M., C. Burgard, S. Crewell, I. V. Gorodetskaya, S. Kneifel, S. Lhermitte, K. Van Tricht, and N. P. M. Lipzig, 2014: How does the spaceborne radar blind zone affect derived surface snowfall statistics in polar regions? *J. Geophys. Research: Atmos.*, **119,** 24, 13,604-13,620, doi: 10.1002/2014JD022079.

Molthan, A. L., W. A. Petersen, S. W. Nesbitt, and D. Hudak, 2010: Evluating the snow crystal size distribution and density assumptions within a single-moment microphysics scheme. Monthly Weather Review, 138, 4254-4267.

Molthan, A. L. and B. A. Colle, 2012: Comparisons of single- and double-moment microphysics schemes in the simulation of a synoptic-scale snowfall event. Monthly Weather Review, 140, 2982-3002.

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

Mosimann, L., E. Weingartner, and A. Waldvogel, 1994: An analysis of accreted drop sizes and mass on rimed snow crystals. *J. Atmos. Sci.*, **51**, 1548–1558

Nielsen, P., 1993: Turbluence effects on the settling of suspended particles. *J. Sed. Petrology*, **63**, 835-838.

Peters, G., B. Fischer, and T. Andersson. 2002. Rain observations with a vertically looking Micro Rain Radar (MRR). *Bor. Environ. Res*. **7**, 353–362.

Petersen, W. A., and Coauthors, 2007: NASA GPM/PMM participation in the Canadian CloudSat/CALIPSO validation project C3VP: Physical process studies in snow. Preprints, 33rd Int. Conf. on Radar Meteorology, Cairns, Australia, Amer. Meteor. Soc., P12.8. [Available online at https://ams.confex.com/ams/33Radar/webprogram/Paper123652.html.]

Pokharel, B., B. Geerts, J. Xiaoqin, K. Friedrich, J. Aikins, D. Breed, R. Rasmussen, and

A. Huggins, 2014: The impact of ground-based glaciogenic seeding on clouds and precipitation over mountains: A multi-sensor case study of shallow precipitating orographic cumuli. *Atmospheric Research*, **147**, 162-182.

Pokharel, B., B. Geerts, and X. Jing, 2014: The impact of ground-based glaciogenic seeding on orographic clouds and precipitation: A multisensory case study. *J. Appl. Meteor. Climatol.*, **53**, 890-909.

Prat, O. P., and A. P. Barros, 2010: Ground observations to characterize the spatial gradients and vertical structure of orographic precipitation—Experiments in the inner region of the Great Smoky Mountains. *J. Hydrol.*, **391**, 141–156, doi:10.1016/j.jhydrol.2010.07.013.

Rinehart, R.E., 1991: *Radar for Meteorologists*, 2d Ed. University of North Dakota Press, 334 pp.

Schreur, B. W., and G. Geertsema, 2008: Theory for a TKE based parameterization of wind gusts. HIRLAM Newsletter, **54**, 177-188. [Available online at http://hirlam.org/index.php/component/docman/doc_download/150-hirlam-newsletter-no-54-paper-25-wichers-schreur?Itemid=70]

Shi, J. J., W-K. Tao, T. Matsui, R. Cifelli, A. Hou, S. Lang, A. Tokay, N-Y. Wang, C. Peters-Lidard, G. Skofronick-Jackson, S. Rutledge, and W. Petersen, 2010: WRF simulations of the 20-22 January 2007 snow events over eastern Canada: Comparison with in situ and satellite observations, *J. Appl. Met. Clim.*, 49, 2246-2266.

Skamarock, W. C., and Coauthors, 2008: A description of the advanced research WRF version 3. NCAR Tech. Note NCAR/TN-4751STR, 125 pp.

Skofronick-Jackson, G., and Coauthors, 2015: Global Precipitation Measurement cold season precipitation experiment (GCPEX): For measurement's sake, let it snow. *Bull. Amer. Meteor. Soc.*, **96**, 1719-1741.

Stark, D., B. A. Colle, and S. E. Yuter, 2013: Observed microphysical evolution for two east coast winter storms and the associated snow bands. *Mon. Wea. Rev.*, **141**, 2037-2057.

Tao, W.-K., and Coauthors, 2003: Microphysics, radiation and surface processes in the Goddard Cumulus Ensemble (GCE) model. *Meteor. Atmos. Phys.*, **82**, 97–137.

Thompson, S., R. M. Rasmussen, K. Manning, 2004: Explicit Forecasts of Winter Precipitation Using an Improved Bulk Microphysics Scheme. Part I: Description and Sensitivity Analysis. *Mon. Wea. Rev.*, **132**, 519-542.

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

Wang, L.-P., and M. R. Maxey, 1993: Settling velocity and concentration distribution of heavy particles in homogeneous isotropic turbulence. *Journal of Fluid Mechanics*, **256**, 27-68, doi: 10.1017/S0022112093002708.

Xie, X., U. Löhnert, S. Kneifel, and S. Crewell, 2012: Snow particle orientation observed by ground-based microwave radiometry. *J. Geophys. Res.*, **117**, D02206, doi:10.1029/2011JD016369.

Yuter, S. E., D. Kingsmill, L. B. Nance, and M. Löffler-Mang, 2006: Observations of precipitation size and fall speed characteristics within coexisting rain and wet snow. *J. Appl. Meteor. Climo.*, **45**, 1450-1464.

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

List of Tables:

Table 1. Snow particle fall speed variables. Vertical air motion is w and turbulence is E. In the third and fourth equations below, w is assumed to be constant within a radar resolution volume and a WRF model grid box.

Name		Description	Reference
Terminal velocity	V_t	Velocity of hydrometeor in still air (w=0, E=0) usually measured for individual particles	Locatelli and Hobbs (1974)
Settling speed	V_s	$V_s = V_t + E$ (w=0). Usually measured for individual particles	Wang and Maxey 1993; Nielsen 1993
Mean Doppler velocity	V_r	$V_r = \frac{\int_D (V_t(D) + E) Z(D) dD}{\int_D Z(D) dD} + w$, within a vertically-pointing radar resolution volume.	Doviak and Zrnic (1993)
Computed mean reflectivity-weighted fall speed	V_{cf}	$V_{cf} = \frac{\int_D V_t(D) Z(D) dD}{\int_D Z(D) dD} + w$, within a WRF model grid box. (E=0)	Calculated from scheme assumptions
Computed mean mass-weighted fall speed	V_m	$V_m = \frac{\int_D V_t(D) m(D) dD}{\int_D m(D) dD}$, within a WRF model surface grid box (w=0, E=0).	Calculated from scheme assumptions

Table 2. Initialization time, total run time, and initial and boundary conditions used in the cases simulated with the WRF V3.3 model.

Case	Initialization Time	Total Run Time (Hours)	Initial and Boundary Conditions
19-20 December 2009	19 December 2009 1200 UTC	24	NAM 218
8 January 2010	8 January 2010 0000 UTC	18	NAM 218
28 January 2010	28 January 2010 0000 UTC	18	NAM 218
16 February 2010	15 February 2010 1800 UTC	30	NAM 218
26 February 2010	26 February 2010 0000 UTC	24	NAM 218
7 January 2011	7 January 2011 1200 UTC	12	NAM 218
21 February 2011	21 February 2011 0000 UTC	18	0.5° GFS
21 January 2012	21 January 2012 0000 UTC	24	NAM 218
11 February 2012	11 February 2012 0000 UTC	18	0.5° GFS

Table 3. Verification points used for the WRF model validation results.

19-20 December 2009		
BMP	**Verification Point**	**Location**
WSM6	40.7220, -73.7655	Queens, NY
THOM2	40.7720, -73.8754	La Guardia, NY
SBU-YLIN	40.6910, -73.9757	Brooklyn, NY
MORR	40.7428, -73.9908	Manhattan, NY
Remaining Simulated Events		
All	40.9044, -73.1184	SBNY

List of Figures

7 Box and whisker plots of liquid equivalent precipitation from various microphysics schemes accumulated for light (left, light shading) and moderate (right, heavier shading) precipitation events. Shaded regions bound the first and third quartiles with median values inset. Extended, dashed lines represent the 10^{th} and 90^{th} percentiles.

8 CFADs of observed reflectivity (dBZ) for light riming periods shown in Figure 2. (a) from the MRR and (b) from the WSR-88D at OKX. (c) Simulated reflectivity (dBZ) from the MORR scheme, (d) from the WSM6 scheme, (e), from the THOM2 scheme, and (f) from the SBU-YLIN scheme. The dashed line at 5.25 km in panel a) and 6 km in panel b) indicates the altitude at which point the observed CFADs were truncated aloft due to the limited number of observations above these altitudes, as described in Section 2b.

9 CFADs of fall speed variables (positive downward, m s^{-1}) for light riming periods shown in Figure 2. (a) Doppler velocity observed from the MRR, and (b) computed mean reflectivity-weighted fall speed simulated from the MORR scheme, (c) from the WSM6 scheme, (d), from the THOM2 scheme, and (e) from the SBU-YLIN scheme. The dashed line at 5.25 km in panel a) indicates the altitude at which point the observed CFADs were truncated aloft due to the limited number of observations above these altitudes, as described in Section 2b.

10 As in Figure 3, but for moderate riming periods shown in Figure 2.

11 As in Figure 4 but for moderate riming periods shown in Figure 2. Note that the x-axis scale for the snow or precipitating ice column differs from the remaining panels.

12 As in Figure 8, but for moderate riming periods shown in Figure 2.

13 As in Figure 9 but for moderate riming periods shown in Figure 2.

14 Period of 51% side planes and 20% bullets observed from 0145 UTC to 0500 UTC on 16 February 2010. (a) Observed and simulated surface size distribution, (b) mean mixing ratio for snow, cloud water, and graupel (g kg^{-1}), (c) the distribution of PARSIVEL settling speeds, V_s (m s^{-1}), normalized to the number of particles every 15 minutes, and (d) mean mass-weighted fall speed, V_m (m s^{-1}) for total precipitation mixing ratio (snow, rain, and graupel, g kg^{-1}). Error bars represent one standard deviation above and below the simulated size distribution. The diameter* for panel a) notes that the panel compares the "PARSIVEL diameter" for observations discussed in Section 2, and the diameter of assumed, spherical and frozen hydrometeors within the model, where schemes assume a single crystal habit.

15 As in Fig. 8 except during a period of 70% plates and 16% side planes observed from 1000 to 1200 UTC 21 February 2011.

16 As in Fig. 8 except during a period of 65% dendrites and 20% plates observed from 1530 UTC on 26 February 2010 to 0000 UTC on 27 February 2010.

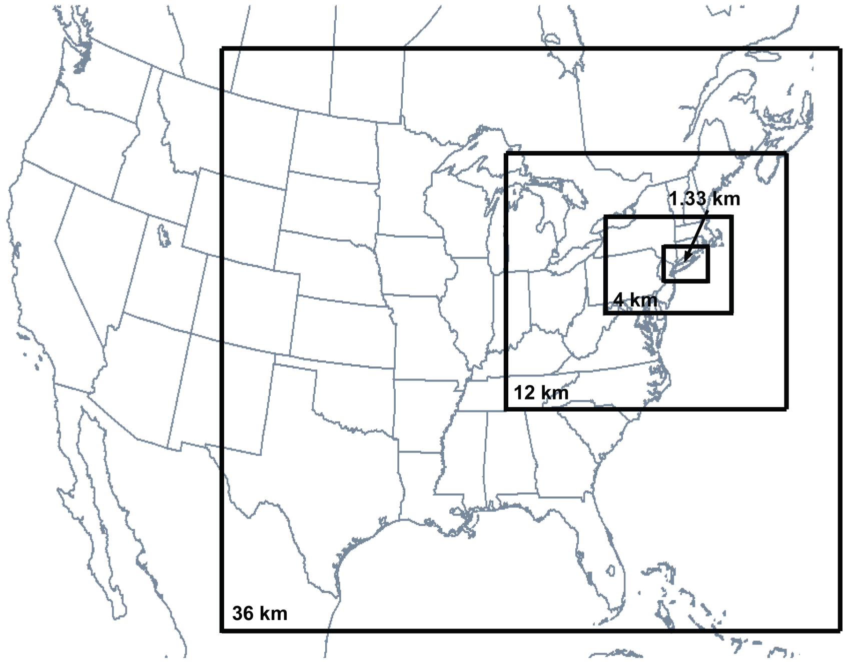
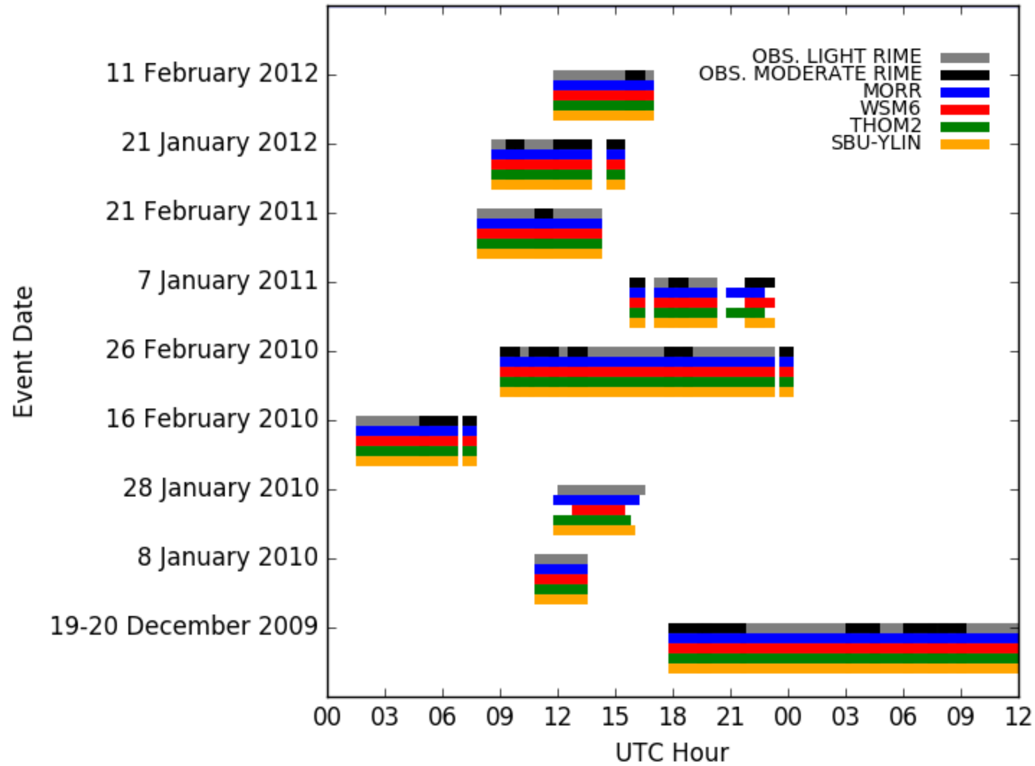
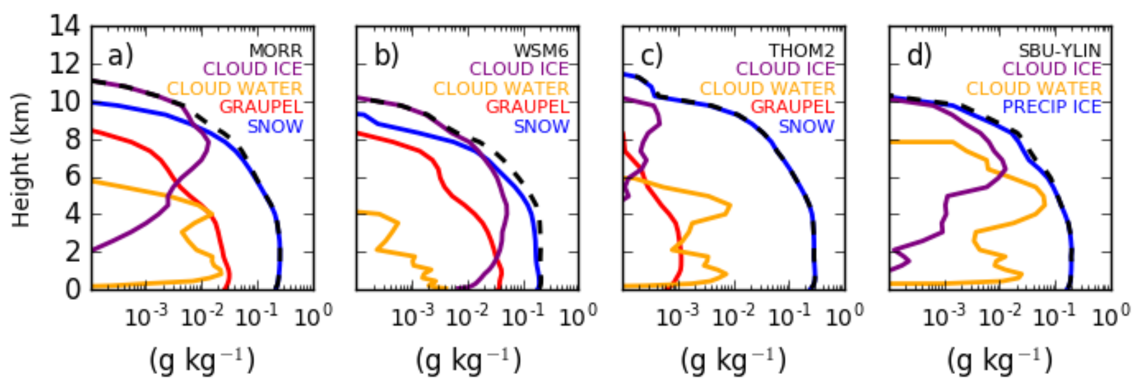


Figure 1. WRF model domains used in this study.



916

917 Figure 2. Observed light or moderate riming periods and corresponding simulation times
918 when model profiles were extracted for performing comparisons.



919

920 Figure 3. Mean profiles of hydrometeor content for selected microphysics schemes and
921 simulations sampled during observed periods of light riming shown in Figure 2.
922 Combined mixing ratios of cloud ice and snow are shown as a black dashed line.

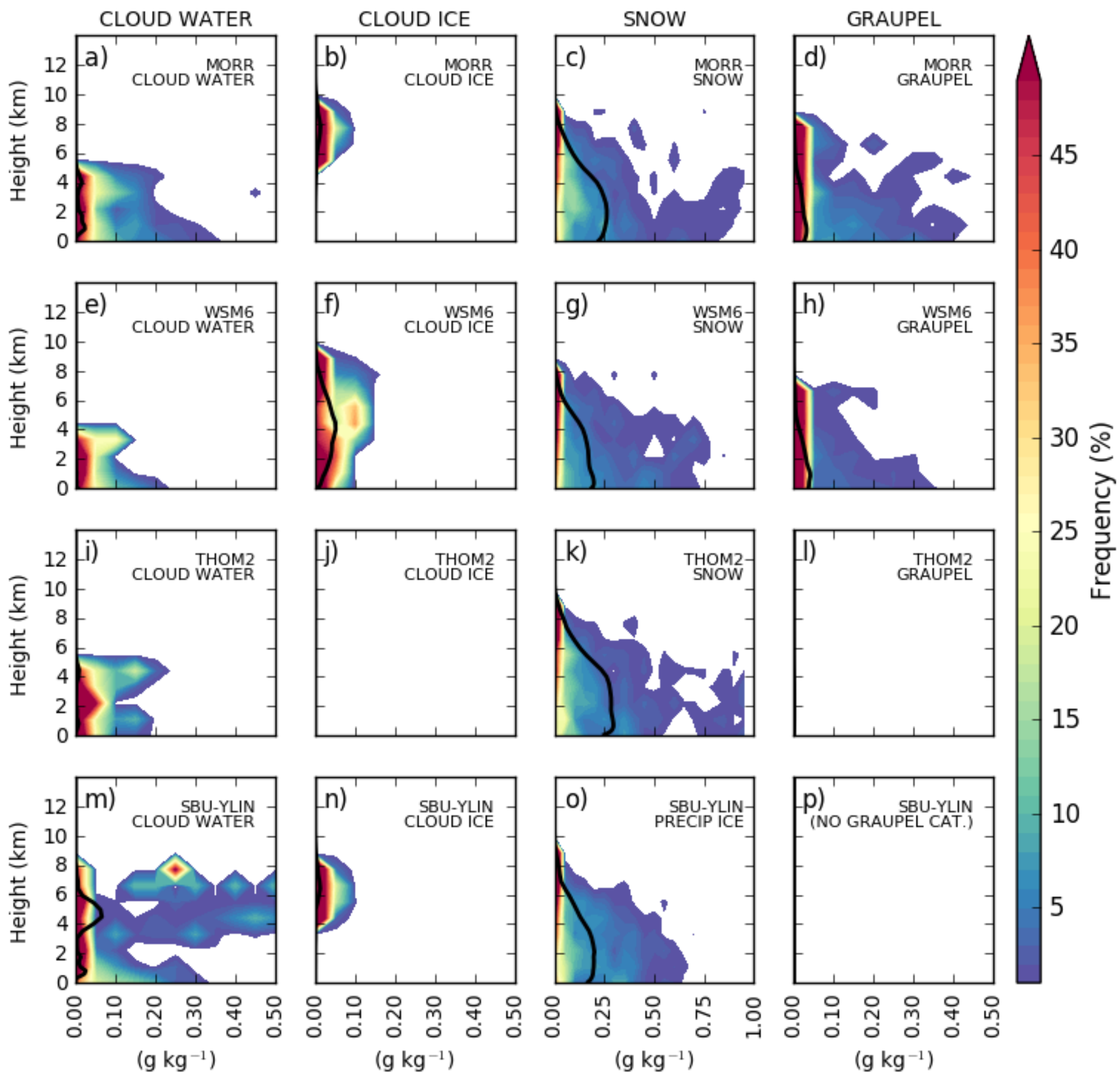


Figure 4. Contoured frequency with altitude diagrams (CFADs) for hydrometeor species obtained from surveyed microphysics schemes during the light riming periods shown in Figure 2. Mean profiles of each hydrometeor type are provided as a solid black line. Trace amounts of cloud ice and graupel in the THOM2 scheme are shown as mean profiles in Figure 3. Note that the x-axis scale for the snow or precipitating ice column differs from the remaining panels.

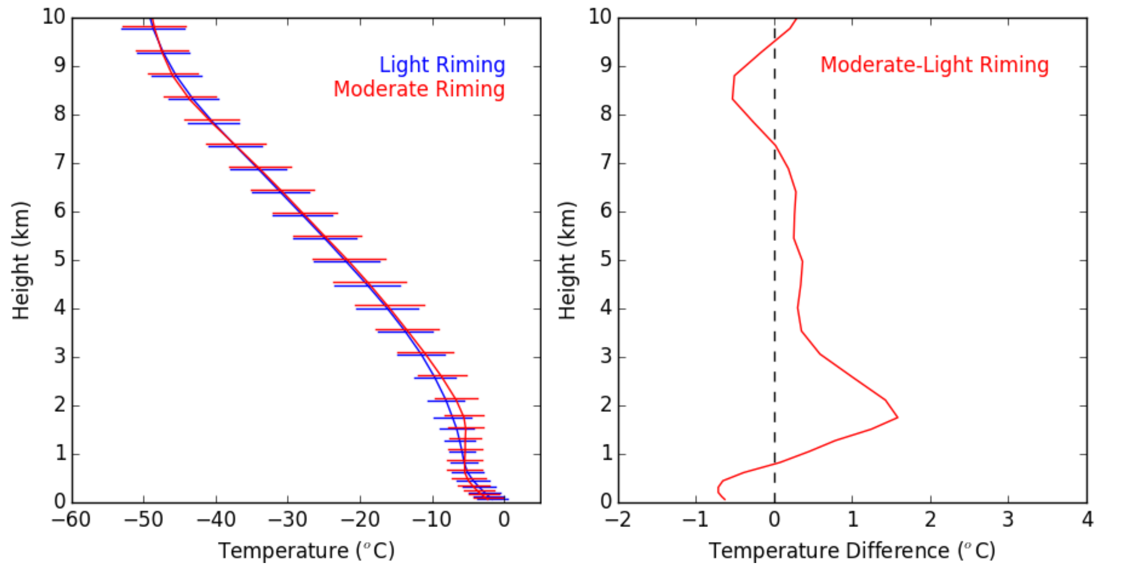


Figure 5. (left) Mean temperature profiles and standard deviations of temperature during periods of light and moderate riming shown in Figure 2, and (right) difference of the mean temperature profile between moderate and light riming periods.

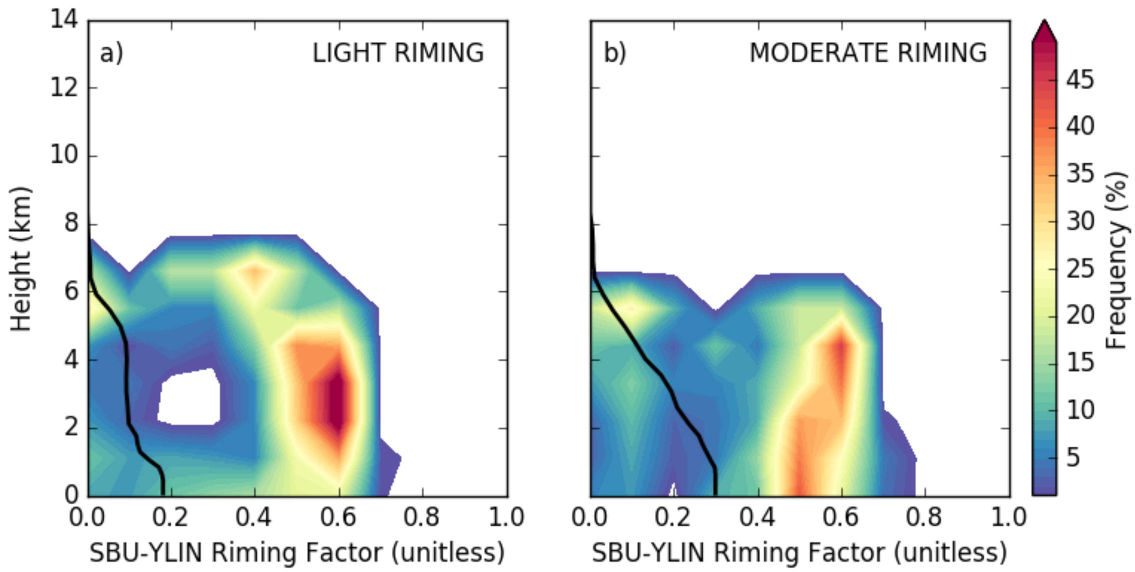


Figure 6. CFADs (shaded) and mean profiles (black line) for the unitless riming factor used to parameterize the mass-diameter and diameter-fall speed relationships for the precipitating ice category within the SBU-YLIN scheme, partitioned into model simulations of observed a) light and b) moderate riming periods shown in Figure 2.

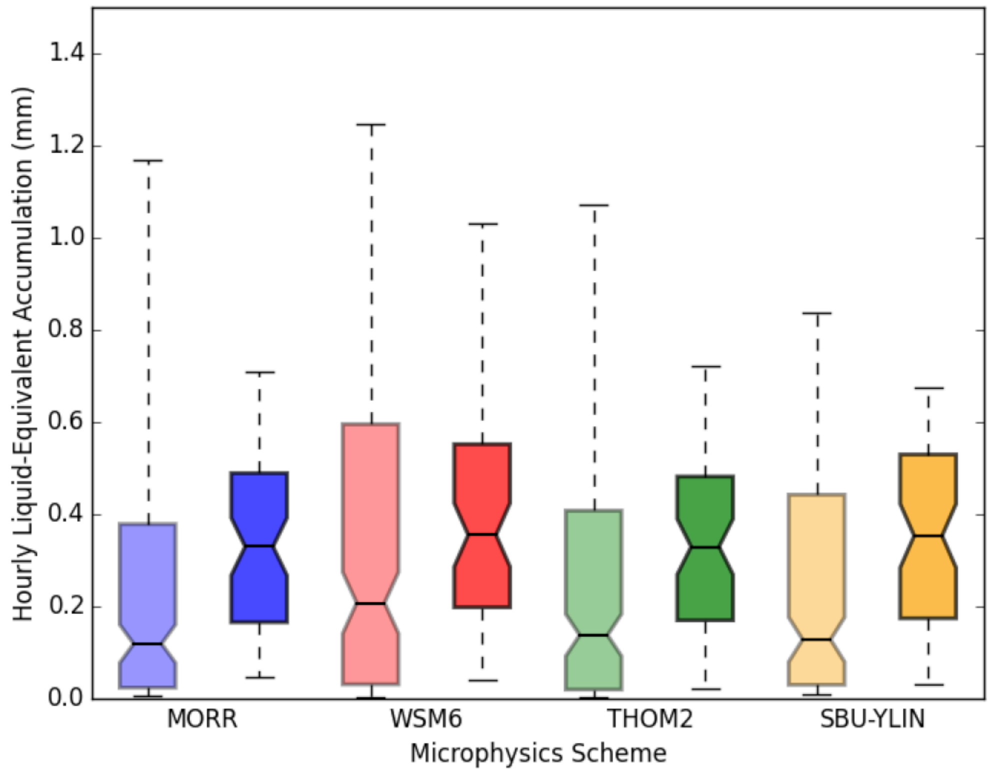


Figure 7. Box and whisker plots of liquid equivalent precipitation from various microphysics schemes accumulated for light (left, light shading) and moderate (right, heavier shading) precipitation events. Shaded regions bound the first and third quartiles with median values inset. Extended, dashed lines represent the 10^{th} and 90^{th} percentiles.

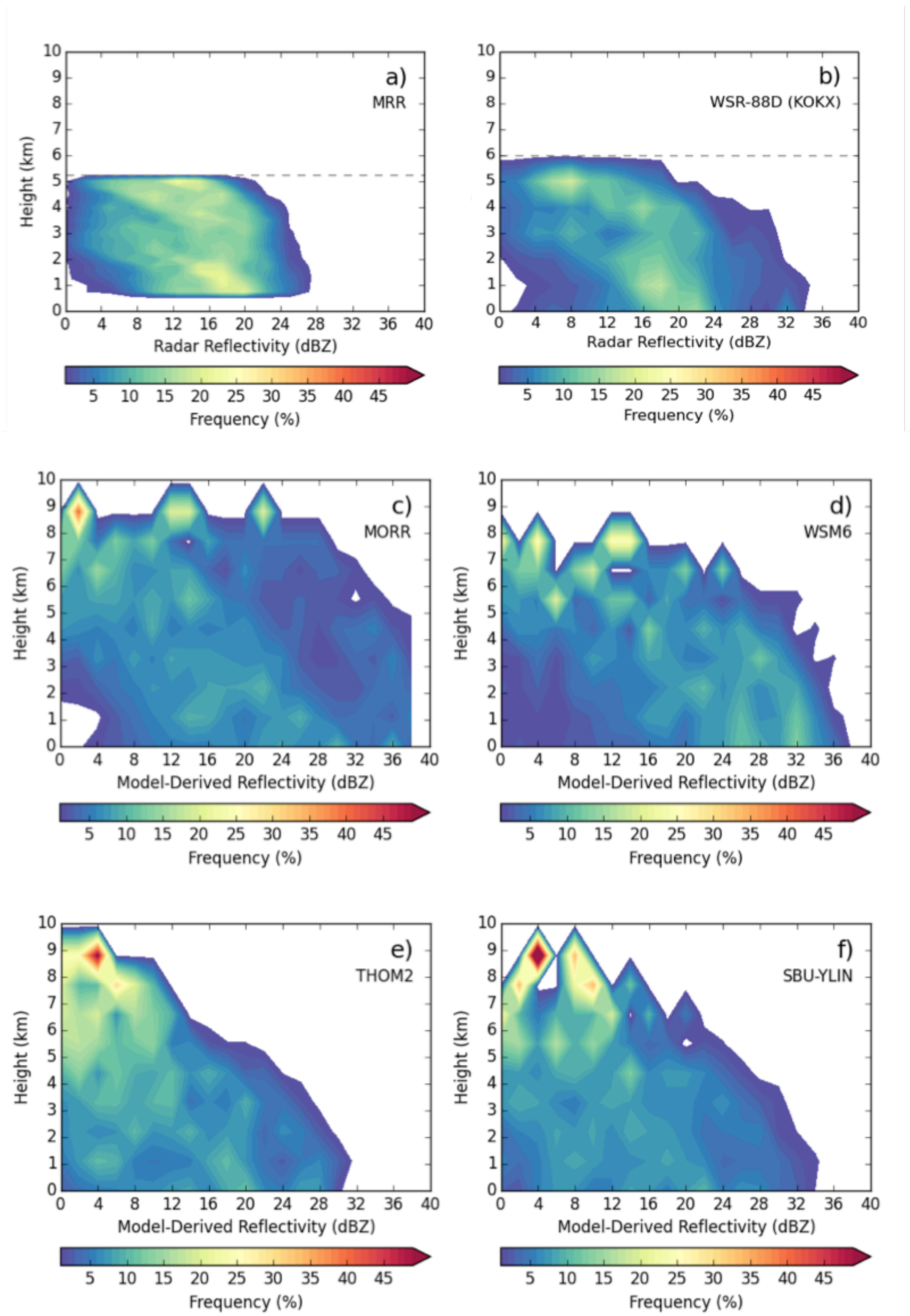


Figure 8. CFADs of observed reflectivity (dBZ) for light riming periods shown in Figure 2. (a) from the MRR and (b) from the WSR-88D at OKX. (c) Simulated reflectivity (dBZ) from the MORR scheme, (d) from the WSM6 scheme, (e), from the THOM2 scheme, and (f) from the SBU-YLIN scheme. The dashed line at 5.25 km in panel a) and 6 km in panel b) indicates the altitude at which point the observed CFADs were truncated aloft due to the limited number of observations above these altitudes, as described in Section 2b.

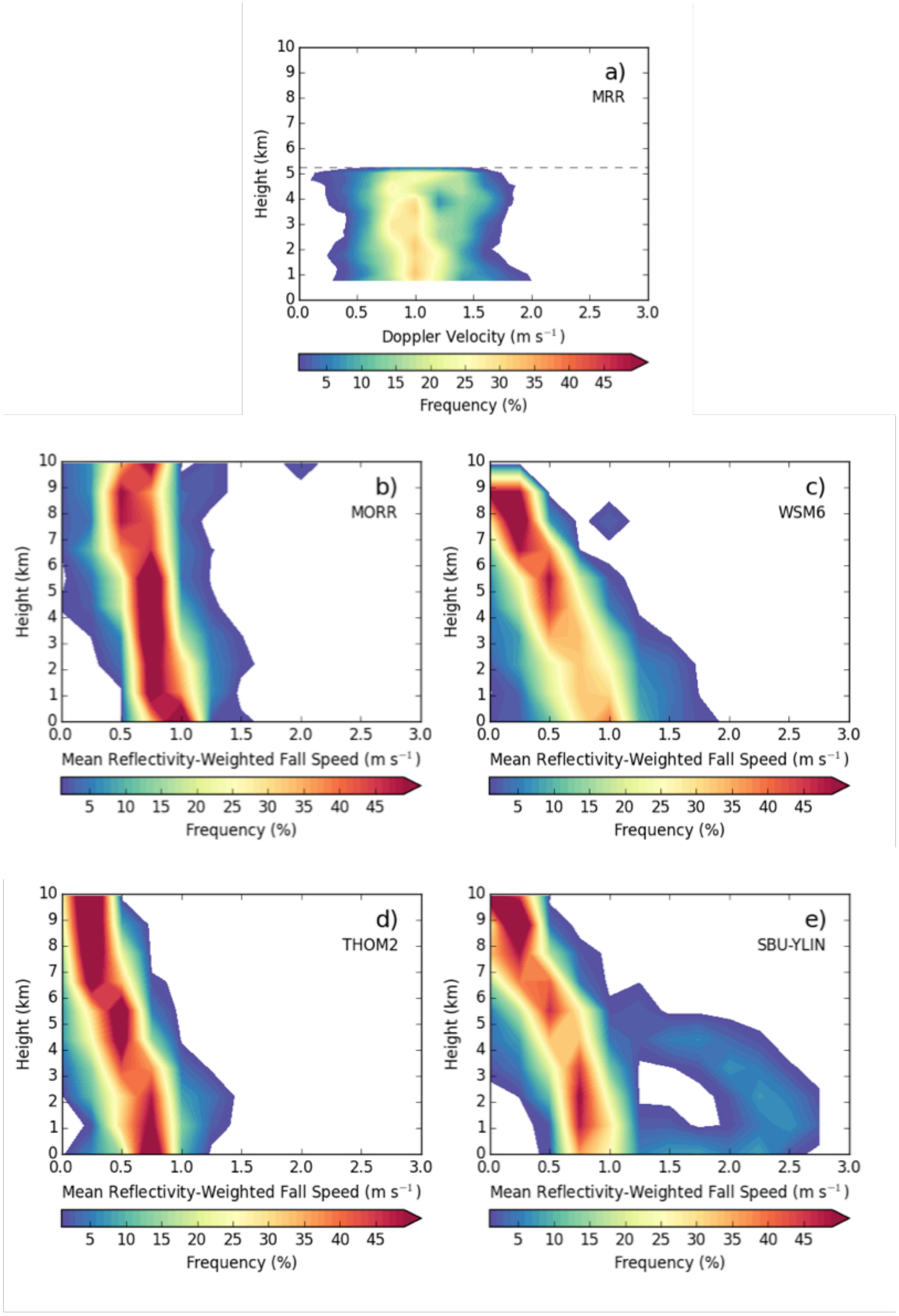


Figure 9. CFADs of fall speed variables (positive downward, m s^{-1}) for light riming periods shown in Figure 2. (a) Doppler velocity observed from the MRR, and (b) computed mean reflectivity-weighted fall speed simulated from the MORR scheme, (c) from the WSM6 scheme, (d), from the THOM2 scheme, and (e) from the SBU-YLIN scheme. The dashed line at 5.25 km in panel a) indicates the altitude at which point the observed CFADs were truncated aloft due to the limited number of observations above these altitudes, as described in Section 2b.

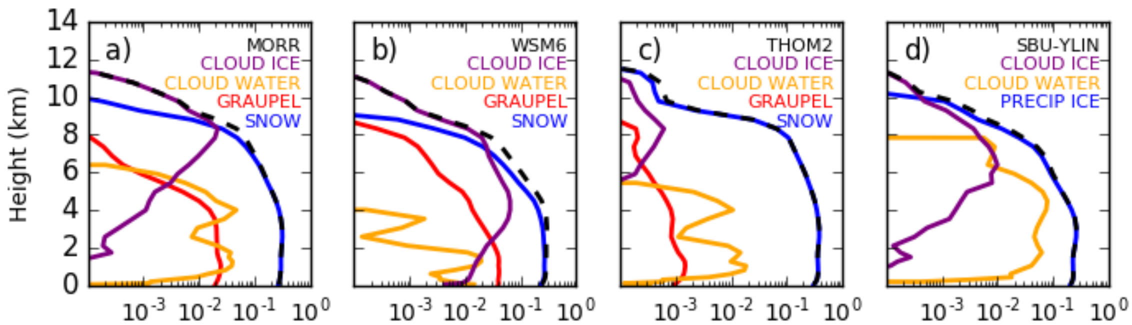


Figure 10. As in Figure 3, but for moderate riming periods shown in Figure 2.

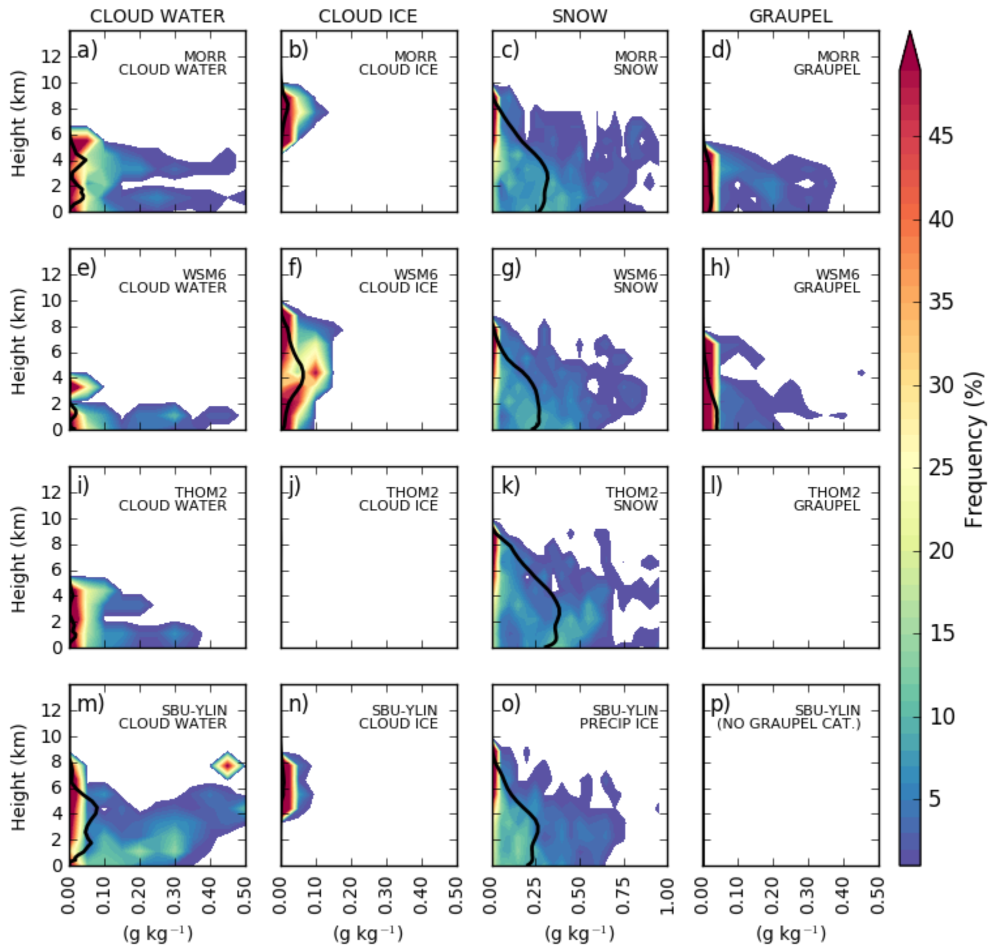


Figure 11. As in Figure 4 but for moderate riming periods shown in Figure 2. Note that the x-axis scale for the snow or precipitating ice column differs from the remaining panels.

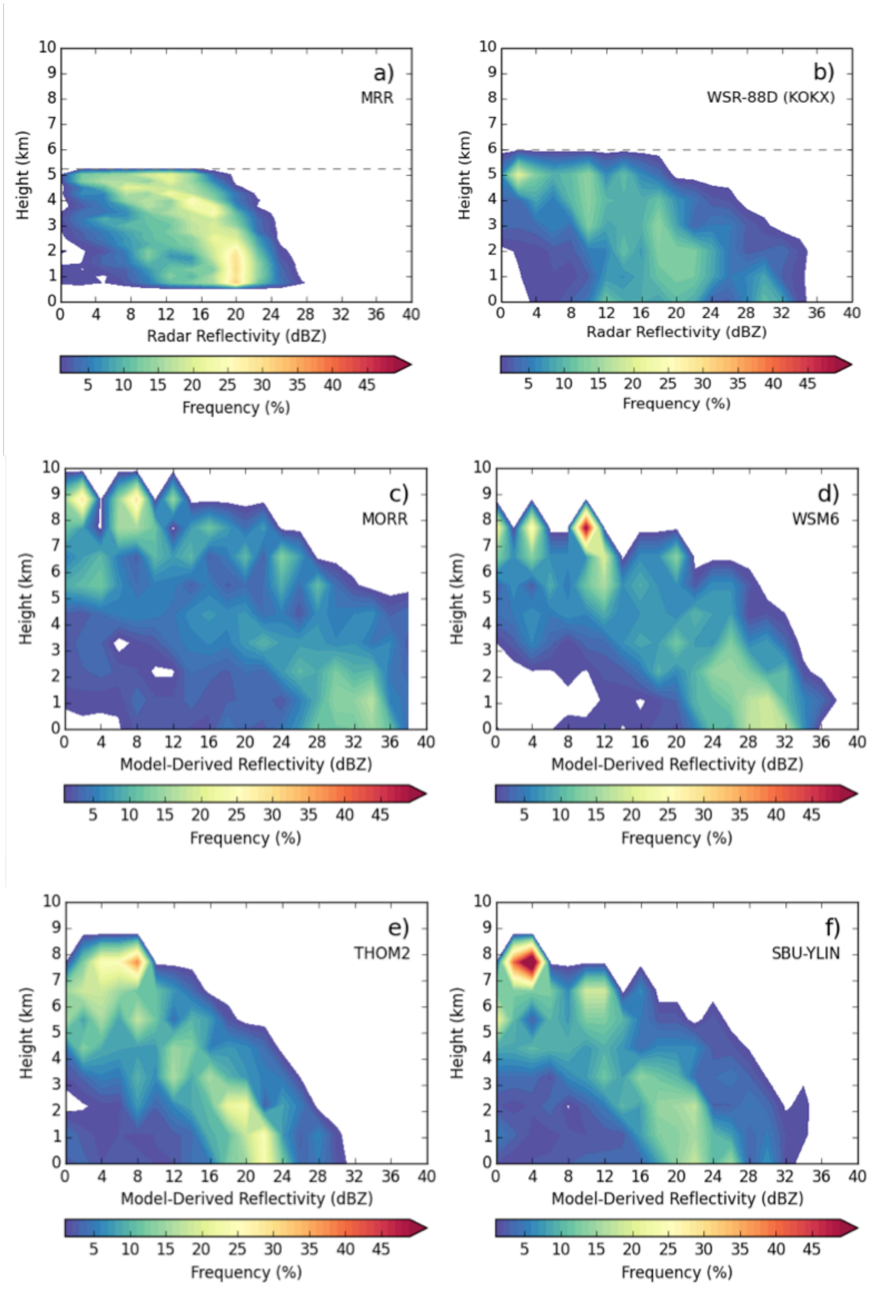


Figure 12. As in Figure 8, but for moderate riming periods shown in Figure 2.

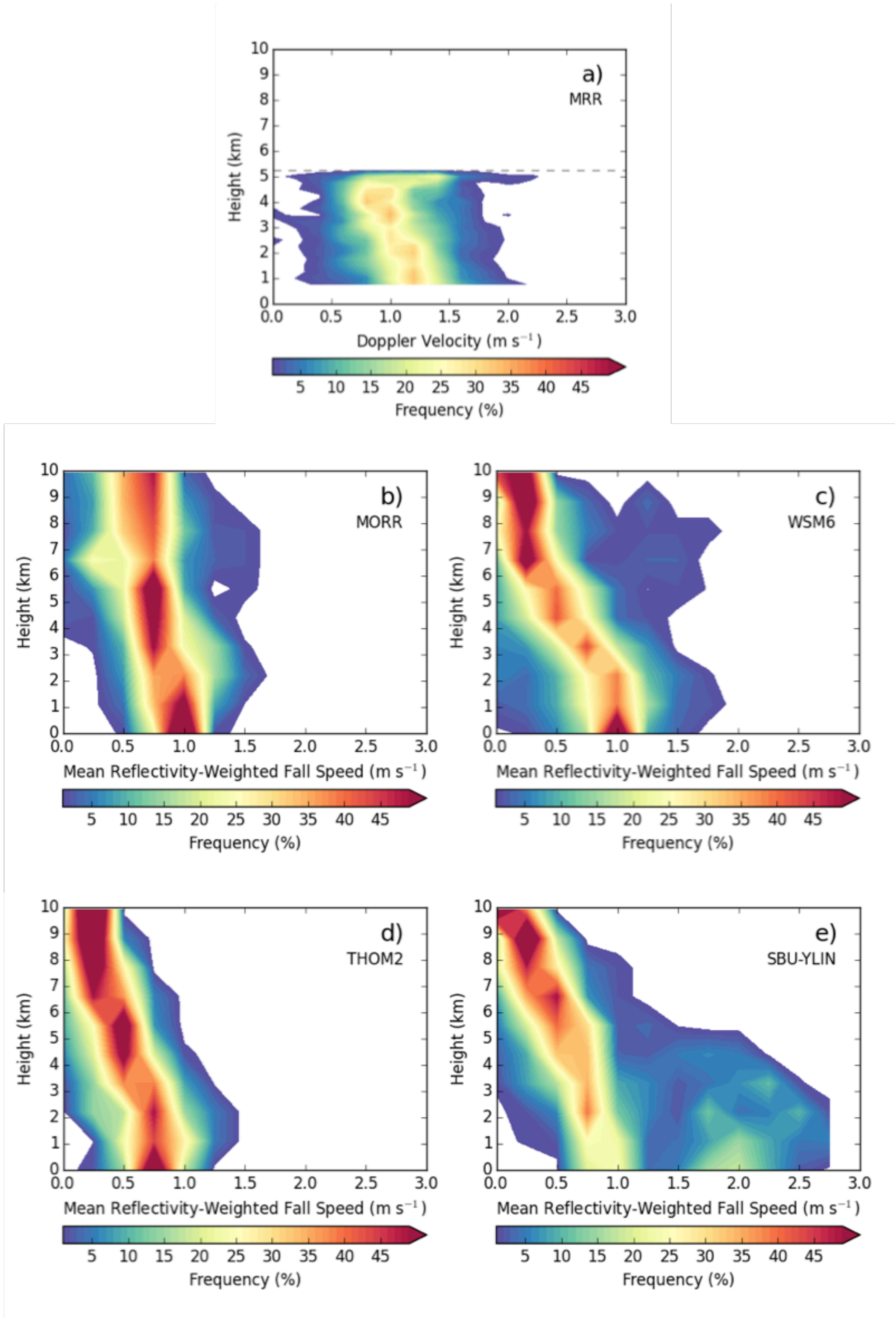


Figure 13. As in Figure 9 but for moderate riming periods shown in Figure 2.

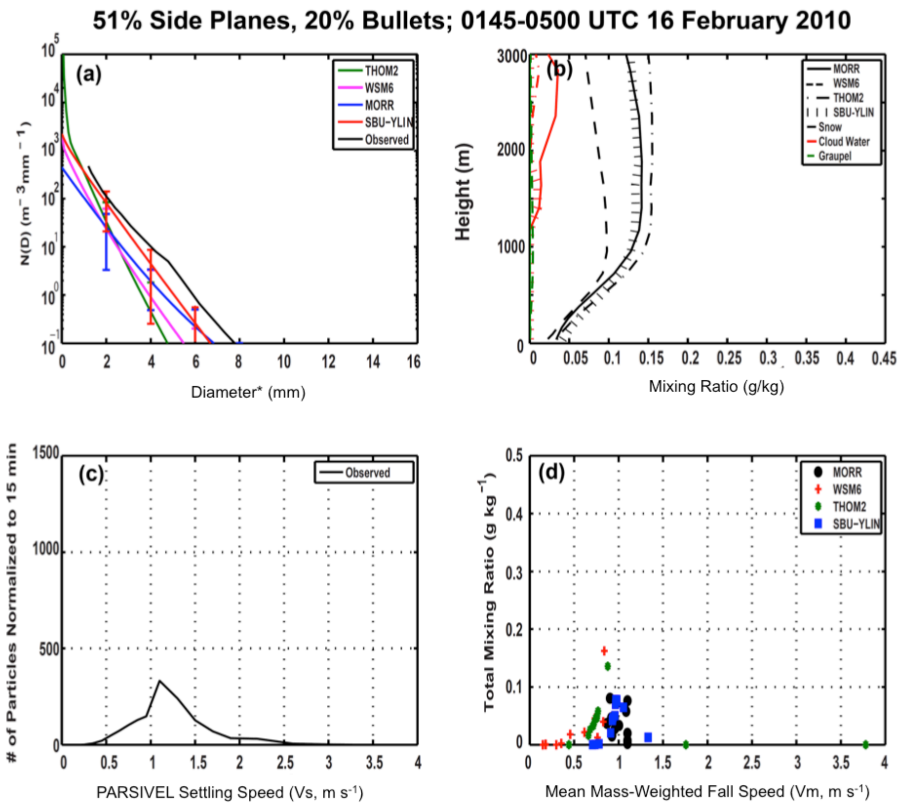


Figure 14. Period of 51% side planes and 20% bullets observed from 0145 UTC to 0500 UTC on 16 February 2010. (a) Observed and simulated surface size distribution, (b) mean mixing ratio for snow, cloud water, and graupel (g kg^{-1}), (c) the distribution of PARSIVEL settling speeds, V_s (m s^{-1}), normalized to the number of particles every 15 minutes, and (d) mean mass-weighted fall speed, V_m (m s^{-1}) for total precipitation mixing ratio (snow, rain, and graupel, g kg^{-1}). Error bars represent one standard deviation above and below the simulated size distribution. The diameter* for panel a) notes that the panel compares the "PARSIVEL diameter" for observations discussed in Section 2, and the diameter of assumed, spherical and frozen hydrometeors within the model, where schemes assume a single crystal habit.

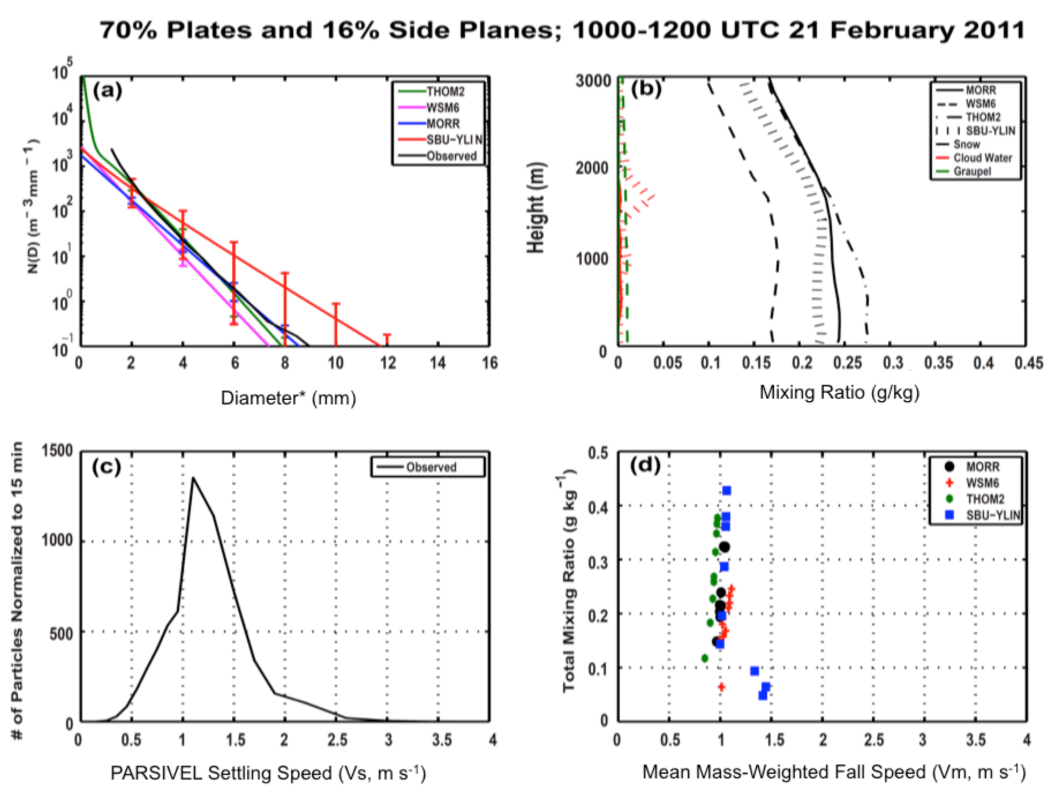


Figure 15. As in Fig. 8 except during a period of 70% plates and 16% side planes observed from 1000 to 1200 UTC 21 February 2011.

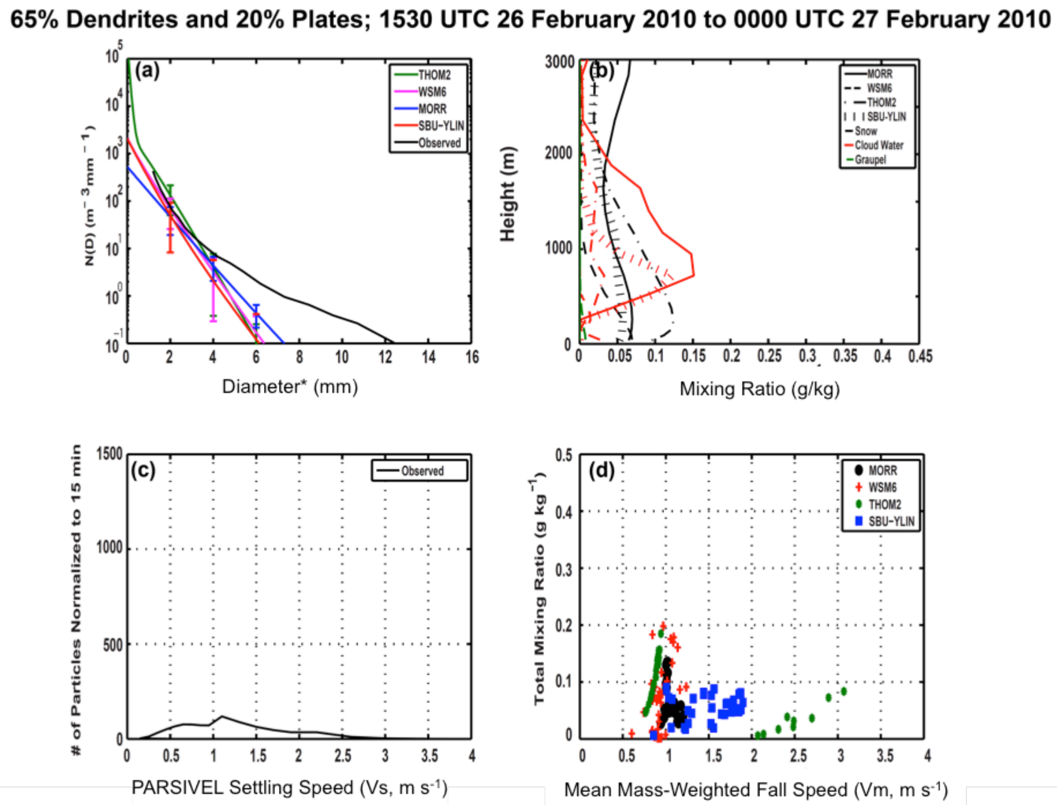


Figure 16. As in Fig. 8 except during a period of 65% dendrites and 20% plates observed from 1530 UTC on 26 February 2010 to 0000 UTC on 27 February 2010.