Influence of Bulk Microphysics Schemes upon Weather Research 1 and Forecasting (WRF) Version 3.6.1 Nor'easter Simulations 2

- Stephen D. Nicholls^{1,2}, Steven G. Decker³, Wei-Kuo Tao¹, Stephen E. Lang^{1,4}, Jainn J. Shi^{1,5}, and 3 Karen I. Mohr¹ 4
- ¹NASA-Goddard Space Flight Center, Greenbelt, 20716, United States of America
- 5 6 7 8 9 ²Joint Center for Earth Systems Technology, Baltimore, NASA-Goddard Space Flight Center, Baltimore, 21250, United States of America
- 3Department of Environmental Sciences, Rutgers, The State University of New Jersey, 08850, United States of America
- 10 ⁴Science Systems and Applications, Inc., Lanham, 20706, United States of America
- 11 ⁵Goddard Earth Sciences Technology and Research, Morgan State University, 21251, United States of America
- 12 *Correspondence to: Stephen D. Nicholls (stephen.d.nicholls@nasa.gov)*

13 Abstract. This study evaluated the impact of five, single- or double- moment bulk microphysics schemes (BMPS) on 14 Weather Research and Forecasting (WRF, version 3.6.1) model simulations of seven, intense winter time cyclone 15 events impacting the Mid-Atlantic United States. Five-day long WRF simulations were initialized roughly 24 hours 16 prior to the onset of coastal cyclogenesis off the coast of North Carolina. Validation efforts focus on microphysics-17 related storm properties including hydrometer mixing ratios, precipitation, and radar reflectivity by comparing model 18 output to model analysis and available gridded radar and rainfall products across 35 WRF model simulations (5 19 BMPSs and seven cases). Comparisons of column integrated mixing ratios and mixing ratio profiles revealed little 20 variability in non-frozen hydrometeor species due to their common programming heritage, yet assumptions about 21 snow and graupel intercepts, ice supersaturation, snow and graupel density maps, and terminal velocities lead to 22 considerable variability in frozen hydrometeor species and in turn radar reflectivities. WRF model simulations were 23 found to produce similar precipitation coverage, but simulations favored excessively high precipitation amounts 24 compared to observations and low to moderate (0.217–0.414) threat scores. Finally, comparison of contoured 25 frequency with altitude (CFAD) plots between WRF and gridded observed radar reflectivity fields yielded notable 26 variations between BMPSs with schemes favoring lower graupel mixing ratios and better aggregation assumptions 27 compared more favorably to observations.

28 1 Introduction

29 Bulk microphysical parameterization schemes (BMPSs) within numerical weather prediction models have 30 become increasingly complex and computationally expensive. Modern prognostic weather models, such as the 31 Weather Research and Forecasting (WRF) model (Skamarock et al., 2008), offer BMPS options ranging from 32 simplistic, warm rain physics (Kessler, 1969) to complex, six-class, two-moment microphysics (Morrison et al., 2009). 33 Microphysics and cumulus parameterizations drive cloud and precipitation processes within numerical weather 34 prediction models and directly or indirectly impacts radiation, moisture, aerosols, and other simulated processes. 35 Citing its importance, Tao et al. (2011) detailed more than 36 published, microphysics-focused studies focusing on 36 idealized simulations, hurricanes, or mid-latitude convection. More recently, the observational studies of Stark (2012) 37 and Ganetis and Colle (2015) investigated microphysical species variability within United States (U.S.) east coast 38 winter-time cyclones (locally called "nor'easters") and have called for further studies investigating how microphysical 39 parameterizations impact simulations of these powerful cyclones.

A "nor'easter" is a large (~2000 km), mid-latitude cyclone occurring from October to April and is capable of bringing punishing winds, copious precipitation, and potential coastal flooding to the Northeastern U.S. (Kocin and Uccellini 2004; Jacobs et al., 2005; Ashton et al., 2008). This region is home to over 65 million people and produces 16 billion U.S. dollars of daily economic output (Morath, 2016). Given its high output, nor'easter-related damages and disruptions can be extreme. Just ten strong, December nor'easters, between 1980 and 2011, produced 29.3 billion U.S. dollars in associated damages (Smith and Katz, 2013). BMPSs are key to accurate simulations of a nor'easter's precipitation and microphysical properties and will be the focus of this study. 47 Recent nor'easter studies are scarce given the extensive research efforts in the 1980s. These historical studies 48 addressed key nor'easter drivers including frontogenesis and baroclinicity (Bosart, 1981; Forbes et al., 1987; Stauffer 49 and Warner, 1987), anticyclones (Uccelini and Kocin, 1987), latent heat release (Uccelini et al., 1987), and moisture 50 transport by the low-level jet (Uccellini and Kocin, 1987; Mailhot and Chouinard, 1989). Despite extensive 51 observational analyses, less attention has been provided to mid-latitude, winter cyclone simulations, especially those 52 focused on BMPSs.

Reisner et al. (1998) ran several single and double-moment BMPS Mesoscale Model Version 5 simulations of winter storms impacting the Colorado Front Range for the Winter Icing and Storms Project. Double moment-based simulations produced more accurate simulations of supercooled water and ice mixing ratios than those originating from single-moment schemes. However, single-moment simulations vastly improved when the snow-size distribution intercepts were derived from a diagnostic equation rather than from a fixed value.

Wu and Pretty (2010) investigated how five, six-class BMPSs affected WRF simulations of four polar-low events (two over Japan, two over the Nordic Sea). Their simulations yielded nearly identical storm tracks, but notable cloud top temperature and precipitation errors. Overall, WRF single-moment BMPS (Hong and Lim, 2006) produced marginally better cloud and precipitation process simulations compared to other BMPSs. For warmer, tropical cyclones, Tao et al. (2011) investigated how four, six-class BMPSs impacted WRF simulations of Hurricane Katrina. They found BMPS choice minimally impacted storm track, yet sea-level pressure (SLP) varied up to 50 hPa.

64 Shi et al. (2010) evaluated several WRF single-moment BMPSs during a lake-effect snow event. Simulated radar 65 reflectively and cloud top temperature validation revealed that WRF accurately simulated the onset, termination, cloud 66 cover, and band extent of a lake-effect snow event, however snowfall totals at fixed points were less accurate due to 67 interpolation of the mesoscale grid. They found BMPSs produced only minimal simulation differences because cold 68 temperatures and weak vertical velocities prevented graupel generation. Reeves and Dawson (2013) investigated WRF 69 sensitivity to eight BMPSs during a December 2009 lake-effect snow event. Their study found precipitation rates and 70 snow coverage were sensitive to BMPSs because vertical velocities exceeded hydrometeor terminal fall speeds in half 71 of their simulations. Vertical velocity differences were attributed to varying BMPS frozen hydrometeor assumptions 72 concerning snow density values, temperature-dependent snow-intercepts, and graupel generation terms.

Similar to previous studies, we will evaluate WRF winter storm simulations and their sensitivity to six- and sevenclass BMPSs, but our primary focus will be microphysical properties and precipitation. The remainder of this paper is divided into three sections. Section 2 explains the methodology and analysis methods. Section 3 shows the results.

Finally section 4 describes the conclusions, its implications, and prospects for future research.

77 2 Methods

78 2.1 Study design

We utilized WRF version 3.6.1 (hereafter W361) which solves a set of fully-compressible, non-hydrostatic,
Eulerian equations in terrain-following coordinates (Skamarock et al., 2008). Figure 1 shows the four-domain WRF
model grid configuration with 45-, 15-, 5-, and 1.667-km grid spacing used for this study. This grid also has 61 vertical

- levels, a 50-hPa (~20 km) model top, two-way feedback, and turns off cumulus parametrization in Domains 3 and 4.
 The fourth domain is convection-resolving and moves for each simulation set (Fig. 1). Global Forecasting System
 model operational analysis (GMA) data was used for WRF boundary conditions. This model configuration (except
 the 4th domain) and the below parameterizations are identical to those in Nicholls and Decker (2015) and are consistent
 with past and present WRF model studies at NASA-Goddard Space Flight Center (i.e., Shi et al., 2010; Tao et al.
- 87 2011). Model parameterizations include:
- Longwave radiation: New Goddard Scheme (Chou and Suarez, 1999; Chou and Suarez, 2001)
- 89 Shortwave radiation: New Goddard Scheme (Chou and Suarez, 1999)
- 90 Surface layer: Eta similarity (Monin and Obukhov, 1954; Janjic, 2002)
- Land surface: NOAH (Chen and Dudhia, 2001)
- 92 Boundary layer: Mellor-Yamada-Janjic (Mellor and Yamada 1982; Janjic 2002)
- Cumulus parameterization: Kain-Fritsch (Kain, 2004) (Not applied to domains 3 and 4)

This study investigates the same, diverse, selectively chosen sample of seven nor'easter cases from Nicholls and Decker (2015) detailed in Table 1 and storm tracks are shown in Fig. 1. The seven, nor'easter cases in Table 1 include at least one event per month (October–March) and are sorted by month rather than chronological order. In Table 1, the Northeast Snowfall Impact Scale (NESIS) value serves as proxy for storm severity (1 is notable and 5 extreme) and its value depends upon the population impacted, area affected, and snowfall severity (Kocin and Uccellini, 2004). Early and late season storms (Cases 1, 2, and 7) did not have snow and thus do not have a NESIS rating.

100 Five-day, WRF model simulations were initialized 24 hours prior to the first precipitation impacts in the highly 101 populated Mid-Atlantic region and prior to the onset of rapid, coastal cyclogenesis. A 24 hour lead time provides 102 sufficient time for WRF to fully-develop mesoscale circulations and atmospheric vertical structure (Kleczek et al., 103 2014) and also to establish key surface baroclinic zones and sensible and latent heat fluxes (Bosart, 1981; Uccelini 104 and Kocin, 1987; Kuo et al., 1991; Mote et al., 1997; Kocin and Uccellini, 2004; Yao et al., 2008). We define the first 105 precipitation impact time as the first 0.5 mm (~0.02 inch) precipitation reading from the New Jersey Weather and 106 Climate Network (D. A. Robinson, pre-print, 2005) associated with a nor'easter event. A smaller threshold is not used 107 to avoid capturing isolated showers occurring well ahead of the primary precipitation shield.

To investigate BMPS influence upon W361 nor'easter simulations, five BMPS are used (Table 2). As shown in Table 2, the selected schemes include three, six-class, three-ice, single-moment schemes Lin (Lin6; Lin et al., 1983; Rutledge and Hobbs, 1984), Goddard Cumulus Ensemble (GCE6; Tao et al., 1989; Lang et al., 2007, and WRF single moment (WSM6; Hong and Lim 2006), a seven-class, four-ice, single-moment scheme (GCE7; Lang et al. 2014), and finally, a six-class, three-ice, double-moment scheme (WRF double-moment, six class (WDM6; Lim and Hong 2010)). For this study, we ran 35 W361 simulations covering five BMPS and seven nor'easter cases.

114 2.2 Verification and analysis techniques

Model validation and analysis efforts focused on comparisons of WRF to GMA, Stage IV precipitation (Fulton et al. 1998; Y. Lin and K.E. Mitchell, preprints, 2005), and Multi-Radar, Multi-Sensor (MRMS) 3D volume radar reflectivity (Zhang et al. 2016). GMA offers six-hourly, gridded dynamical fields, including water vapor, with global 118 coverage. Stage IV is a six-hourly, 4-km resolution, gridded precipitation product covering the United States and is

119 derived from rain gauge and radar data. Finally, MRMS is two minute, 1.3-km resolution, gridded 3D volume radar

120 mosaic product derived from S- and C-band radars covering the United States and Southern Canada (Zhang et al.

121 2016). MRMS serves as an operational successor to the better known National Mosaic and Multi-Sensor QPE (NMQ;

122 Zhang et al. 2011) radar mosaic products. Both Stage IV and MRMS, however are limited by the detection range of

123 their surface-based assets. All cross comparisons between WRF and these validation data were conducted at identical

124 grid resolution.

Analysis of WRF model microphysical, precipitation, and simulated radar output was comprised of three main parts: precipitable mixing ratios and domain-averaged mixing ratio profiles, simulated precipitation, and simulated radar reflectivity. Precipitable mixing ratio are calculated for all six microphysical species (vapor, cloud ice, cloud water, snow, rain, and graupel) using the equation for precipitable water:

129 $PMR = \frac{1}{\rho g} \int_{P_{top}}^{P_{sfc}} w \, dp \tag{1}$

130 In Eq. (1), PMR is the precipitable mixing ratio in mm, ρ is the density of water (1000 kg m⁻³); g is the gravitational 131 constant (9.8 m s⁻²); p_{sfc} is the surface pressure (Pa), p_{top} is the model top pressure (Pa); w is the mixing ratio (kg kg⁻ 132 ¹); dp is the change in atmospheric pressure between model levels (Pa). Only water vapor can be validated because 133 the other species are nonexistent in GMA and ground and space validation microphysical data are lacking, especially 134 over the data-poor North Atlantic (Li et al., 2008; Lebsock and Su, 2014). Similarly, mixing ratio profiles will only 135 be inter-compared amongst BMPSs because satellite-derived cloud ice profile products (e.g., CloudSat 2C-ICE; Deng 136 et al. 2013), have a narrow scan width (1.3–1.7 km) and do not have direct overpass of Domain 4 during coastal 137 cyclogenesis. WRF-simulated precipitation fields and their distribution were qualitatively compared to Stage IV data 138 and then evaluated with bias and threat score (critical success index; Wilks, 2011). Finally, contoured frequency with 139 altitude diagrams (CFADs) will validate WRF against observed MRMS data as in similar radar validation efforts of 140 Yuter and Houze (1995), Lang et al. (2011) and Lang et al. (2014). A CFAD offers the advantage of preserving 141 frequency distribution information, yet is insensitive to both spatial and temporal mismatches. Additionally, CFAD 142 scores will also be calculated at each height level and evaluated with time using Eq (2).

143

$$CS = 1 - \frac{\sum |PDF_m - PDF_0|_h}{200}$$
(2)

144 In (2), CS is the CFAD score and PDF_m and PDF_o (%) are the probability density functions (PDF) at constant 145 height for the model-simulated and observed radar reflectivity, respectively. The CFAD score ranges between 0 (no 146 PDF overlap) to 1 (identical PDFs).

147 **3. Results**

148 **3.1 Hydrometeor species analysis**

Figure 2 displays precipitable mixing ratios (mm) for six microphysics species (water vapor, cloud water, graupel,
 cloud ice, rain, and snow) from Case 5, Domain 4 at 06 UTC February 2010. Corresponding simulated radar

reflectivity (dBZ) at 4,000 m is shown as Fig. 3. This case and time was selected for its negligible storm track error,

- 152 centralized location in Domain 4, and expansive radar reflectivity coverage at 4,000 m where hydrometeor mixing
- ratios are high. Notably, MRMS are currently not available for this date. To supplement these data, Figs. 4 and 5
- depict composite mixing ratios, temperature, and vertical velocity profiles for Case 5 (Fig. 4) and over all seven cases
- 155 (Fig. 5) from Domain 4. Composite profiles are averaged over the residence time of the nor'easter within Domain 4
- 156 (24-30 hours). To emphasize the fraction of supercooled water, two sets of dashed black lines are added to each panel
- 157 in Figs. 4 and 5 to indicate the 0°C and -40°C heights from each model simulation. We exclude hail from our analysis
- 158 because it is unique to GCE7 and it mixing ratio values are an order of magnitude smaller than other species.
- 159 Comparing Figs. 2 and 4 to Fig. 3, reveals a strong correspondence between radar reflectivity signatures and 160 particular precipitable hydrometeor species structures, especially graupel and snow and to a lesser extent cloud water. 161 Analysis of Fig. 4 reveals that cloud water at 4,000 m is super-cooled and graupel mixing ratios values are near their 162 peak and given the corresponding precipitation mixing ration values in Fig. 2, these two species are well correlated 163 with the strongest, convective reflectivity signatures (> 35 dBZ). Fig. 4 also reveals snow mixing ratio, except for 164 Lin6 are also comparatively high at this level, yet precipitable snowfall values better correlate best with moderate 165 reflectivity (20-35 dBZ) regions within the broader, more stratiform, precipitation shield. Notably, for Lin6, reduced 166 snow mixing ratios are partially offset by an increase of graupel mixing ratio values within the precipitation shield. 167 Inter-BMPS mixing ratio variability amongst BMPSs, both at this level and throughout the troposphere, is due to 168 identifiable trends within the underlying assumptions made by BMPSs and will explained in more detail below.
- 169 All evaluated BMPSs share a common heritage in the Lin6 scheme. With the exception of the two-moment cloud 170 water and rain and CCN-cloud droplet feedbacks in WDM6, the BMPSs differ primarily in how each addresses frozen 171 hydrometeor species (cloud ice, graupel, and snow). Their common programming heritage is evident from the nearly 172 identical (exception: WDM6) rain mixing ratio profiles (Figs. 4 and 5) and precipitable water vapor (Fig. 2) and is 173 consistent with Wu and Petty (2010). WDM6, unlike single-moment BMPSs, explicitly forecasts CCN, rain and cloud 174 droplet number concentrations and does not apply derivative equations (Hong et al., 2010). The forecasts result 175 produce minimal changes to maximum mixing ratio height (Figs. 4 and 5) and precipitable rain coverage (Fig. 2), yet 176 rain mixing ratios remain higher aloft and decrease sharply towards the surface unlike in single-moment simulations.
- Similar to rain mixing ratios, cloud water mixing ratios exhibit little variability in either the precipitable cloud water extent (Fig. 6) or the maximum mixing ratio height and freezing level (Fig. 7), but maximum mixing ratio values vary even between single-moment BMPSs. Differing allowances in the amount of ice supersaturation between GCE7 (Chern et al. 2016) and WSM6 (Hong et al. 2010) are likely to account for the differences in the maximum cloud
- 181 water mixing ratios. Although in WDM6 cloud water is double-moment, the maximum mixing ratios are only 182 decreased slightly relative to WSM6. This result suggests that WDM6-forecasted cloud water number concentrations
- 183 are likely close to prescribed 300 cm⁻³ number concentration assumed in WSM6 (Hong et al. 2010) and/or the larger-
- 184 scale environment/forcing is a dominant factor as water supersaturation are negligible.
- Amongst the BMPSs, Figs. 2, 4, and 5 show that precipitable snow and snow mixing ratios vary considerably with Lin6 and GCE6 having the smallest and highest amounts of snow, respectively. Dudhia et al. (2008) and Tao et al. (2011) attribute low snow mixing ratios in Lin6 to its high rates of dry collection of snow by graupel, its low snow

188 size distribution intercept (decreased surface area), and its auto-conversion of snow to either graupel or hail at high 189 mixing ratios. GCE6 turns off dry collection of snow and ice by graupel, greatly increasing the snow mixing ratios at 190 the expense of graupel and reducing snow riming efficiency (Lang et al. 2007). Snow growth in GCE6 is further 191 augmented by its assumption of water saturation for the vapor growth of cloud ice to snow (Reeves and Dawson, 192 2013; Lang et al. 2014). GCE7 addressed the vapor growth issue of GCE6 and applied numerous other changes 193 including the introduction and of a snow size and density mapping, snow breakup interactions, a relative humidity 194 (RH) correction factor, and a new vertical-velocity-dependent ice super saturation assumption (Lang el al., 2007; Lang 195 et al., 2011; Lang et al., 2014; Chern et al., 2016; Tao et al., 2016). Despite the reduced efficiency of vapor growth of 196 cloud ice to snow stemming from the both the new RH correction factor and the ice super saturation adjustment, the 197 new snow mapping and enhanced cloud ice to snow auto-conversion in GCE7 offset this potential reduction and keep 198 GCE snowfall mixing ratio higher than in non-GCE BMPSs. Unlike Lin6, WSM6 and WDM6 assume grid cell graupel 199 and snow fall speeds are identical (Dudhia et al., 2008) and that ice nuclei concentration is a function of temperature 200 (Hong et al., 2008). These two aspects, effectively eliminate the accretion of snow by graupel and increase snow 201 mixing ratios at colder temperatures (Dudhia et al., 2008; Hong et al., 2008). Figure 4 and 5 show the height of 202 maximum snow mixing ratio is roughly conserved in all non-Lin6 BMPSs. Lin6's assumption of non-uniform graupel 203 and snow fall speeds and dry collection of snow by graupel reduce snow mixing ratios in the middle troposphere and 204 raise its maximum snow mixing ratio height.

205 Compared to snow, graupel mining ratios are generally smaller for non-Lin6 schemes due to Lin6's assumption 206 of dry collection by snow dominates species growth which was proven unrealistic by Stith et al. (2002). GCE7 is in 207 many ways at opposition to Lin6, where it simulations generate the most snow, yet the least graupel. GCE7 includes 208 graupel size mapping, but the combination of the snow size mapping (decrease snow size aloft, increases snow surface 209 area, and enhances vapor growth), the addition of deposition conversion processes (graupel/hail particles experiencing 210 deposition growth at colder temperatures are converted to snow), and a reduction in super cooled droplets available 211 for riming (cloud ice generation is augmented, see below) all favor snow growth at the expense of graupel (Lang et 212 al. 2014; Chern et al., 2016; Tao et al., 2016). Consistent with Reeves and Dawson (2013), graupel mixing ratios value 213 are typically 30-50 % of their snow counterparts for WSM6 and WDM6.

214 Although cloud ice mixing ratios are up to ninety percent smaller than those for snow (GCE6), cloud ice mixing 215 ratios still vary greatly amongst the BMPSs as illustrated in Figs. 2, 4, and 5. Cloud ice mixing ratios are highest in 216 GCE7 and lowest in Lin6. Wu and Petty (2010) similarly found low cloud ice mixing ratios in Lin6 simulations and 217 ascribe it to dry collection by cloud ice by graupel and its fixed cloud-ice size distribution. Similar to Lin6, GCE6 218 uses a monodispersed cloud-ice size distribution (20 µm diameter), but assumes vapor growth of cloud ice to snow 219 under an assumption of water saturation conditions (yet supersaturated with respect ice) leading to higher cloud ice 220 amounts, but also increased cloud ice to snow conversion rates (Lang et al., 2011; Tao et al., 2016). GCE7 blunt this 221 cloud ice to snow conversion term using a RH correction factor which is dependent upon ice supersaturation which is 222 itself dependent up vertical velocity. Additionally, GCE7, also includes contact and immersion freezing terms (Lang 223 et al., 2011), makes the cloud ice collection by snow efficiency a function of snow size (Lang et al., 2011; Lang et al., 224 2014), sets a maximum limit on cloud-ice particle size (Tao et al., 2016), makes ice nuclei concentrations follows the

- 225 Cooper curve (Cooper, 1986; Tao et al., 2016), and it allows cloud ice to persist in ice subsaturated conditions (i.e.,
- 226 RH for ice \geq 70%) (Lang et al, 2011; Lang et al., 2014). Despite the increased cloud ice-to-snow auto conversion
- (Lang et al. 2014; Tao et al. 2016), all the above changes nearly doubled cloud ice amounts in GCE7 than in GCE6
- 228 (See Fig. 2). Similar to GCE7, WSM6 runs generate larger cloud ice mixing ratios than Lin6, which Wu and Petty
- 229 (2010) attribute to excess cloud glaciation at temperatures between 0°C and -20°C and its usage of fixed cloud ice size
- 230 intercepts. Additionally, both WSM6 and WDM6 include ice sedimentation terms which promote smaller cloud ice
- amounts (Hong et al., 2008). Despite their varying assumptions, the maximum cloud ice amounts for both Case 5 and
- 232 overall (Figs. 4 and 5) are consistent between BMPSs.

233 **3.2 Stage IV precipitation analysis**

234 Excess precipitation, whether frozen or not, is one of the most potentially crippling impacts from a nor'easter. 235 WRF precipitation is generated from its microphysics and cumulus parameterization; the latter is turned for Domains 236 3 (5 km grid spacing) and 4 (1.667-km grid spacing). Figures 6 and 7 show Domain 3, 24-hour accumulated 237 precipitation, their difference from Stage IV, and the associated probability and cumulative distribution functions 238 (PDF and CDF, respectively) of precipitation for Cases 5 and 7. As for our composite microphysics plots, the data 239 accumulation period only covers the nor'easter's residence time in Domain 4. We focus on Domain 3 rather than 240 Domain 4 because the latter is located near the boundary of the Stage IV dataset where its radar-based data tends to 241 fade. Cases 5 and 7 are shown here because these cases have near-shore tracks (Fig. 1) good coverage of their 242 associated precipitation by Stage IV. Table 3 includes threat score and bias information for all seven cases their 243 associated standard deviation statistics. Both threat score and bias assume a 10 mm precipitation accumulation 244 threshold value, which as seen in Figs. 6 and 7 is approximately the 25th percentile of accumulated precipitation.

245 Table 3 shows Case 4 as a clear outlier where its low threat score and bias values deviate more than two standard 246 deviation from the composite mean due to its non-coastal track (Fig. 1) and thus it will be excluded from this section 247 of the analysis. For the remaining six cases, Table 4 indicates low (0.217; Lin6, Case 2) to moderate (0.414; Lin6, 248 Case 5) threat scores and a 10 mm precipitation contour spatial covers an area far exceeding Stage IV (bias range: 249 1.47 [Lin6, Case 7] – 4.05 [GCE7, Case 3]). Inter-BMPS barely varied with threat score and biases varying only up 250 to an order of magnitude less than the threat and bias scores themselves. Consistent with Hong et al. (2010), threat 251 score and bias values for WSM6 are equal to or improved upon by WDM6 due to its inclusion of a cloud condensation 252 nuclei (CCN) feedback. Overall, WDM6 generated marginally better simulated precipitation fields and has the lowest 253 threat score in four out of six cases and it also has the lowest model mean (0.322), yet Lin6 was found to be the least 254 bias in four out of six cases and it also has the lowest model mean (2.55).

As illustrated Figs. 6 and 7, all WRF simulations tended to generate similar coverage to Stage IV, but its precipitation values tended to be smaller than for corresponding grid points in WRF resulting in low to moderate forecast skill and excessively heavy precipitation totals as illustrates in the PDF and CDF diagrams. Previous modelling studies of strong-convection by Ridout et al. (2005) and Dravitzki and McGregor (2011) found both GFS and Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS) produced too much light precipitation and too much heavy precipitation, which stands in contrast to our results, which show the opposite tendency. Unlike these

- two studies, nor'easters often track over the data spare North Atlantic, a region with no rain gauge data and is at the
- 262 operation range limits of S-band radars. These issues could lead to an under bias in Stage IV precipitation data,
- especially near the data edges, which likely suggests that threat scores and biases are likely closer to observations than
- shown. Marginal changes in accumulated precipitation (<10 mm) between BMPS simulations and threat scores is
- 265 consistent with investigation of simulation precipitation during warm-season events and quasi-stationary front (Fritsch
- and Carbone, 2004; Wang and Clark 2010).

267 **3.3 MRMS and radar reflectivity analysis**

268 Figure 8 show statistical CFADs for Case 4, Domain 4 constructed over a 24 hour period (12 UTC 26–27 January 269 2015) with 0° C and -40°C heights at approximately 3,000 and 9,000 m above mean sea level (not shown). Similar to 270 the previous section, all CFAD and CFAD products are based only upon the 24-30 hour period a nor'easter resided 271 within Domain 4. We selected Case 7 because its radar volume data from NMQ has been reprocessed with the latest 272 algorithms associated with MRMS. To supplement Fig. 8, MRMS and WRF simulated radar reflectivities are shown 273 at 4,000 and 9,500 m above mean sea level on 18 UTC 26 January 2015 are shown as Figs. 9 and 10, respectively. 274 These two heights were selected because they pass through the two MRMS dBZ frequency maxima shown in Fig. 8. 275 Finally, Fig. 11 shows CFAD scores with height and time and their differences over the same time period as Fig. 8.

276 Figure 8 show a wider ranges of dBZ values (up to 40 dBZ) from WRF simulations than from MRMS (up to 27 277 dBZ) below the melting layer. Qualitatively, all model simulations below the melting layer have dBZ frequency ranges 278 exceeding that of MRMS, yet only Lin6 and especially GCE7 correctly capture the core of maximum frequencies 279 between 5-10 dBZ. All other schemes produce this same core, but at values over 10 dBZ. Figure 9 illustrates these 280 radar reflectivity differences at the 4,000 m above sea level where radar reflectivity values from GCE6, WSM6, and 281 WDM6 simulations are often 15 dBZ or more greater than MRMS. Between 3,000 and 6,000 m, only GCE7 produces 282 a narrow core of maximum frequency values below 10 dBZ consistent with MRMS. Lang et al. (2014) attribute the 283 narrow core to changes in aggregation which made it both temperature and mixing ratio dependent and to the new 284 snow map. Together these changes favored the production of small hydrometeors at colder temperature and larger 285 hydrometeors at warmer temperatures. Eventually above 6,500 m, all WRF CFADs collapse to very small radar 286 reflectivities values (< 5 dBZ) whereas the core of dBZ frequencies increases in MRMS up through 11 km. As Fig. 287 10 shows, at 9,500 m in altitude radar reflectivity coverage has become spotty and quite sensitive to even small radar 288 signatures.

289 Consistent with the above discussion, CFAD scores with height and time (Fig. 11) show Lin6 to qualitatively 290 perform best overall, however, GCE7 simulations below 5,000 m typically attained even higher CFAD scores. Other 291 BMPSs as shown in Fig. 8 typically favor unrealistically higher reflectivity values and the exhibit lower CFAD scores 292 in the melting layer which is likely associated with higher graupel and cloud ice concentrations. Further aloft, 293 aggregation of hydrometeors toward smaller sizes and entrainment likely cut off cloud tops in GCE7 and results in its 294 lower CFAD scores above 6,000 m. The other six cases produce similar tendencies in their CFAD and CFAD scores 295 as noted above for Case 7, except cloud heights become higher and CFADs become wider with the introduction of 296 stronger convection with early and late season events.

297 4 Conclusions

298 The role and impact of five BMPSs upon seven, W361 nor'easter simulations is investigated and validated against 299 GMA, Stage IV precipitation, and MRMS 3D volume reflectivity. Tested BMPSs include four single-moment (Lin6, 300 GCE6, GCE7, and WSM6) and one double-moment BMPSs (WDM6). Simulated hydrometer mixing ratios show 301 general similarities for non-frozen hydrometeor species (cloud water and rain) due to their common Lin6 heritage. 302 However, frozen hydrometeor species (snow, graupel, cloud ice) demonstrate considerably larger variability between 303 BMPSs. Larger changes exist for frozen species due to different assumptions about snow and graupel intercepts, 304 degree of allowable ice supersaturation, snow and graupel density maps, and terminal velocities made by each BMPS. 305 WRF-Stage IV accumulated precipitation comparisons reveal WRF demonstrate that although WRF generates 306 precipitation fields of similar coverage to Stage IV precipitation intensities tended to be higher than observations and 307 resulting in low to moderate (0.217–0.414) threat scores with WDM6 demonstrating marginally better forecast skill 308 than its single-moment counterparts. Finally, MRMS-based CFAD and CFAD scores show Lin6 and GCE7 to be 309 notably better than GCE6, WSM6 and WDM6 in the lower troposphere, with GCE7 being the only BMPS scheme to 310 produce the narrow core of maximum frequencies below10 dBZ due to its temperature and mixing ratio dependent 311 aggregation and new snow map. Above 5,000 m GCE7 however becomes less skilled the combination of smaller 312 hydrometers and entrainment reduced it cloud top height relative to other BMPSs.

313 The study has shown that although subtle in the large-scale environment, cloud microphysics do make small, but 314 noticeable impacts in the microphysical and precipitation properties of a nor'easter. While no BMPS leads to 315 consistently improved precipitation forecast skill, the underlying assumptions do make notable change in the 316 composition of radar reflectivity structure which itself can vary notably from observed radar reflectivity structures. 317 Follow-on studies could investigate additional nor'easter cases or simulate other weather phenomena (polar lows, 318 monsoon rainfall, drizzle, etc.). Results covering multiple phenomena may provide guidance to model users in their 319 selection of BMPS for a given computational cost. Additionally, potential studies could specifically address key 320 aspects of a nor'easter's structure (such as the low-level jet) or validation of model output against current and recently 321 available satellite-based datasets from MODIS (Justice et al., 2008), CloudSat (Stephens et al., 2008), CERES, and 322 GPM (Hou et al. 2014). Finally, other validation methods including object-oriented (Marzban and Sandgathe, 2006) 323 or fuzzy verification (Ebert 2008) could be utilized.

5 Code availability

WRF version 3.6.1 is publically available for download from the WRF Users' Page (http://www2.mmm.ucar.edu/
 wrf/users/download/get_sources.html).

6 Data availability

328 GFS model analysis data boundary condition data can be obtained from the NASA's open access, NOMADS 329 data server (ftp://nomads.ncdc.noaa.gov/GFS/Grid3/). Stage IV precipitation data is publically available from the 330 National Data and Software Facility at the University Center for Atmospheric Research (http://data.eol.ucar.edu/cgi-

331 bin/codiac/fgr_form/id=21.093).

332 **7** Author contributions

S. D. Nicholls designed and ran all experimental model simulations and prepared the manuscript. S. G. Decker
supervised S. D. Nicholls' research efforts, funded the research, and revised the manuscript. W. -K. Tao, S. E. Lang,
and J. J. Shi brought their extensive knowledge and expertise on model microphysics which helped shape the project
methodology and rationalize the results. Finally, K. I. Mohr helped to facilitate connections between the research
team, supervised S. Nicholls' research, and was pivotal in revising the manuscript.

338 8 Acknowledgements

339 This research was supported by the Joint Center for Earth Systems Technology (JCET), the University of

340 Maryland Baltimore County (UMBC), and in part by the New Jersey Agricultural Experiment Station. Resources

341 supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center

342 for Climate Simulation (NCCS) at Goddard Space Flight Center.

343 References

- Ashton, A. D., Donnely, J. P., and Evans, R. L.: A discussion of the potential impacts of climate change on the
 shorelines of the Northeastern U.S.A. Mitig. Adapt. Strat. Glob. Change, 13, 719–743, 2008.
- Bosart, L. F.: The Presidents' Day Snowstorm of 18–19 February 1979: A subsynoptic-scale event, Mon. Wea. Rev.,
 109, 1542–1566, 1981.
- Chen, F., and Dudhia, J.: Coupling an advanced land-surface/ hydrology model with the Penn State/ NCAR MM5
 modeling system. Part I: Model description and implementation, Mon. Wea. Rev., 129, 569–585, 2001.
- 350 Chern, J. -D., Tao, W. -K., Lang, S. E., Matsui, T., J. -L. F. Li, J. -L. F., Mohr, K. I., Skofronick-Jackson, G. M.,
- and Peters-Lidard, C. D.: Performance of the Goddard multiscale modeling framework with Goddard ice
 microphysical schemes, J. Adv. Model. Earth Syst., 7, doi:10.1002/2015MS000469, 2016.
- Chou, M. -D. and Suarez, M. J.: A solar radiation parameterization for atmospheric research studies. NASA Tech,
 Memo NASA/TM-1999-104606, 40 pp., 1999.
- Chou, M. -D., and Suarez, M. J.: A thermal infrared radiation parameterization for atmospheric studies, NASA Tech.
 Rep. NASA/TM-1999-10466, vol. 19, 55 pp., 2001.
- 357 Deng M, G. G. Mace, Z. Wang, and R. P. Lawson: Evaluation of several A-Train ice cloud retrieval products with in
- situ measurements collected during the SPARTICUS campaign, J. Appl. Meteor. Climatol., 52, 1014–1030,
 2013.

- Dravitzki, S., and McGregor, J.: Predictability of heavy precipitation in the Waikato River Basin of New
 Zealand, Mon. Wea. Rev., 139, 2184–2197, 2011.
- Dudhia, J., Hong, S. -Y., and Lim, K. -S.: A new method for representing mixed-phase particle fall speeds in bulk
 microphysics parameterizations, J. Meteor. Soc. Japan, 86A, 33–44, 2008.
- Beert, E. E.: Fuzzy verification of high-resolution gridded forecasts: A review and a proposed framework, Meteor.
 Applic., 15, 51-64, 2008.
- Forbes, G. S., Thomson, D. W., and Anthes, R. A.: Synoptic and mesoscale aspects of an Appalachian ice storm
 associated with cold-air damming, Mon. Wea. Rev., 115, 564–591, 1987.
- Fulton, R. A., J. P. Breidenbach, D.-J. Seo, D. A. Miller, and T. O'Bannon: The WSR-88D rainfall algorithm. Wea.
 Forecasting, 13, 377–395. 1998.
- Fritsch, J. M., and Carbone, R. E.: Improving quantitative precipitation forecasts in the warm season: A USWRP
 research and development strategy, Bull. Amer. Meteor. Soc., 85, 955–965, 2004.
- Ganetis, S. A. and Colle, B. A.: The thermodynamic and microphysical evolution of an intense snowband during the
 Northeast U.S. blizzard of 8–9 February 2013. Mon. Wea. Rev., 143, 4104-4125, 2015.
- Hong, S -Y., and Lim, J. -O. J.: The WRF single-moment 6-class microphysics scheme (WSM6), J. Korean Meteor.
 Soc., 42, 129-151, 2006.
- Hong, S. -Y., Lim, K. -S. S., Lee, Y. -H., Ha, J. -C., Kim, H. -W., Ham, S. -J., and Dudhia, J.: Evaluation of the
 WRF double-moment 6-class microphysics scheme for precipitating convection, Adv. Meteor., 2010,
 doi:10.1155/2010/707253, 2010.
- Hou, A. Y., Kakar, R. K., Neeck, S., Azarbarzin, A. A., Kummerow, C. D., Kojima, M., Oki, R., Nakamura, K., and
 Iguchi, T.: The Global Precipitation Measurement Mission, Bull. Amer. Meteor. Soc., 95, 701–722, 2014.
- Jacobs, N. A., Lackmann, G. M., and Raman, S.: The combined effects of Gulf Stream-induced baroclinicity and
 upper-level vorticity on U.S. East Coast extratropical cyclogenesis, Mon. Wea. Rev., 133, 2494–2501, 2005.
- Janjic, Z. I.: Nonsingular implementation of the Mellor–Yamada level 2.5 scheme in the NCEP meso model, NCEP
 Office Note 437, 61 pp., 2002.
- Justice, C. O. et al. (1998), The Moderate Resolution Imaging Spectroradiometer (MODIS): land remote sensing for
 global change research, IEEE Transactions on Geoscience and Remote Sensing, 36, 1228–1249, 1998.
- 387 Kain, J. S.: The Kain–Fritsch Convective Parameterization: An Update, J. Appl. Meteor., 43, 170–181, 2004.
- Kessler, E.: On the distribution and continuity of water substance in atmospheric circulation, Meteor. Monogr., 32,
 Amer. Meteor. Soc., 84 pp, 1969.
- Kleczek, M. A., G.-J. Steenveld, and A. A. M. Holtslag: Evaluation of the Weather Research and Forecasting
 Mesoscale Model for GABLS3: Impact of boundary-layer schemes, boundary conditions and spin-up,
 Boundary-Layer Meteorol, 152, 213–243, 2014.
- Kocin, P. J. and Uccellini, L. W.: Northeast snowstorms. Vols. 1 and 2, Meteor. Monogr., No. 54., Amer. Met. Soc.,
 818 pp., 2004.
- Kuo, Y. H., Low-Nam, S., and Reed, R. J.: Effects of surface energy fluxes during the early development and rapid
 intensification stages of seven explosive cyclones in the Western Atlantic. Mon. Wea. Rev., 119, 457–476, 1991.

- Lang, S., Tao, W. -K., Cifelli, R., Olson, W., Halverson, J., Rutledge, S., and Simpson, J.: Improving simulations of
 convective system from TRMM LBA: Easterly and westerly regimes, J. Atmos. Sci., 64, 1141–1164, 2007.
- Lang, S. E., Tao, W. -K., Zeng, X., and Li, Y.: Reducing the biases in simulated radar reflectivities from a bulk
 microphysics scheme: Tropical convective systems, J. Atmos. Sci., 68, 2306–2320, 2011.
- 401 Lang, S. E., Tao, W. -K., Chern, J. -D., Wu, D., and Li, X.: Benefits of a fourth ice class in the simulated radar
- 402 reflectivities of convective systems using a bulk microphysics scheme, J. Atmos. Sci., 71, 3583–3612,
- 403 doi:10.1175/JAS-D-13-0330.1, 2014.
- Lebsock, M., and Su, H: Application of active spaceborne remote sensing for understanding biases between passive
 cloud water path retrievals, J. Geophys. Res. Atmos., 119, 8962–8979, doi:10.1002/2014JD021568, 2014.
- 406 Li, J.-L. F., Waliser, D., Woods, C., Teixeira, J., Bacmeister, J., Chern, J.-D., Shen, B.-W., Tompkins, A., Tao,
- 407 W. -K., and Kohler, M.: Comparisons of satellites liquid water estimates to ECMWF and GMAO analyses,
- 408 20th century IPCC AR4 climate simulations, and GCM simulations, Geophys. Res. Lett., 35, L19710,
- 409 doi:10.1029/2008GL035427, 2008.
- 410 Lim, K.-S. and Hong, S. -Y.: Development of an effective double-moment cloud microphysics scheme with
- 411 prognostic cloud condensation nuclei (CCN) for weather and climate models, Mon. Wea. Rev., 138, 1587–
 412 1612, 2010.
- Lin, Y. -L., Farley, R. D., and Orville, H. D.: Bulk parameterization of the snow field in a cloud model, J. Climate
 Appl. Meteor., 22, 1065–1092, 1983.
- Mailhot, J. and Chouinard, C.: Numerical forecasts of explosive winter storms: Sensitivity experiments with a mesoscale model, Mon Wea. Rev., 117, 1311–1343, 1989.
- 417 Marzban C., and Sandgathe, S.: Cluster analysis for verification of precipitation fields, Wea. Forecasting, 21, 824–
 418 838, 2006.
- Mellor, G. L., and Yamada, T.: Development of a turbulence closure model for geophysical fluid problems, Rev.
 Geophys. Space Phys., 20, 851–875, 1982.
- Monin, A. S., and Obukhov, A. M.: Basic laws of turbulent mixing in the surface layer of the atmosphere. Tr. Akad.
 Nauk SSSR Geophiz. Inst., 24, 163–187, 1954.
- 423 Morath, E. (2016), Will a blizzard freeze U.S. economic growth for the third straight year, Wall Street Journal, 20
 424 Jan. 2016.
- Morrison, H., Thompson, G., and Tatarskii, V.: 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, 2009.
- Mote, T. L., Gamble, D. W., Underwood, S. J., Bentley, M. L.: Synoptic-scale features common to heavy snowstorms
 in the Southeast United States, Wea. Forecasting, 12, 5–23, 1997.
- 430 Nicholls, S. D. and Decker, S. G.: Impact of coupling an ocean model to WRF nor'easter simulations, Mon. Wea.
 431 Rev., 143, 4997–5016, 2015.
- Reeves, H. D. and Dawson II, D. T.: The dependence of QPF on the choice of microphysical parameterization for
 lake-effect snowstorms, J. Appl. Meteor. Climatol., 52, 363–377, 2013.

- Reisner, J. R., Rasmussen, R. M., and Bruintjes, R. T.: Explicit forecasting of supercooled liquid water in winter
 storms using the MM5 mesoscale model. Quar. J. Roy. Met. Soc., 124, 1071-1107, 1998.
- Ridout, J. A., Y. Jin, and Liou, C. -S.: A cloud-base quasi-balance constraint for parameterized convection:
 Application to the Kain–Fritsch cumulus scheme, Mon. Wea. Rev., 133, 3315–3334, 2005.
- 438 Rutledge, S. A., and Hobbs, P. V.: The mesoscale and microscale structure and organization of clouds and precipitation
- in mid-latitude cyclones. XII: A diagnostic modeling study of precipitation development in narrow cloud-frontal
 rainbands. J. Atmos. Sci., 20, 2949–2972, 1984.
- Shi, J. J. et al.: WRF simulations of the 20-22 January 2007 snow events of Eastern Canada: Comparison with in situ
 and satellite observations, J. Appl. Meteor. Climatol., 49, 2246–2266, 2010.
- Skamarock, W.C., Klemp, J. P., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X. -Y., Wang, W., and
 Powers, J. G.: A description of the advanced research WRF version 3, NCAR Tech. Note NCAR/TN-475+STR,
 125 pp., 2008.
- Smith, A. B., and Katz, R. W.: US billion-dollar weather and climate disasters: Data sources, trends, accuracy and
 biases, Natural Hazards, 67, 387–410, 2013.
- Stark, D.: Field observations and modeling of the microphysics within winter storms over Long Island, NY. M.S.
 thesis, School of Marine and Atmospheric Sciences, Stony Brook University, 132 pp., 2012.
- 450 Stauffer, D. R., and Warner, T. T.: A numerical study of Appalachian cold-air damming and coastal frontogenesis,
 451 Mon. Wea. Rev., 115, 799–821, 1987.
- 452 Stephens, G. L., et al.: CloudSat mission: Performance and early science after the first year of operation, J. Geophys.
 453 Res., 113, D00A18, doi:10.1029/2008JD009982, 2008.
- 454 Stith, J. L., Dye, J. E., Bansemer, A., Heymsfield, A. J., Grainger, C. A., Petersen, W. A, and Clfelli, R.:
 455 Microphysical observations of tropical clouds, J. Appl. Meteor., 41, 97–117, 2002.
- Tao, W. -K., Simpson, J. and McCumber, M.: An ice-water saturation adjustment, Mon. Wea. Rev., 117, 231–235,
 1989.
- Tao, W. -K., Shi, J. J., Chen, S. S., Lang, S., Lin, P. -L., Hong, S. -Y., Peters-Lidard, C., and Hou, A.: The impact of
 microphysical schemes on hurricane intensity and track, Asia-Pacific J. Atmos. Sci., 47, 1–16, 2011.
- Tao, W. -K., Wu, D., Lang, S., Chern, J. -D., Peters-Lidard, C., Fridlind, A., and Matsui, T.: High-resolution NUWRF simulations of a deep convective-precipitation system during MC3E: Further improvements and
 comparisons between Goddard microphysics schemes and observations, J. Geophys. Res. Atmos., 121, 1278–
 1305, doi:10.1002/2015JD023986, 2016.
- 464 Uccellini, L. W. and Kocin, P. J.: The Interaction of jet streak circulations during heavy snow events along the east
 465 coast of the United States, Wea. Forecasting, 2, 289–308, 1987.
- Wang, S.-Y., and Clark, A. J.: NAM Model forecasts of warm-season quasi-stationary frontal environments in the
 Central United States, Wea. Forecasting, 25, 1281–1292, 2010.
- 468 Wilks, D. S.: Statistical methods in the atmospheric sciences, third edition, Academic Press, Oxford, in press., 2011.
- 469 Wu, L., and Petty, G. W.: Intercomparison of bulk microphysics schemes in model simulations of polar lows, Mon.
- 470 Wea. Rev., 138, 2211–2228, 2010.

- Yao, Y., Pierre, W., Zhang, W., and Jiang, J.: Characteristics of atmosphere-ocean interactions along North Atlantic
 extratropical storm tracks, J. Geophys. Res., 113, doi:10.1029/2007JD008854, 2008.
- 473 Yuter, S. E., and R. A. Houze: Three-dimensional kinematic and microphysical evolution of Florida cumulonimbus
- 474 part II: frequency distributions of vertical velocity, reflectivity, and differential reflectivity, Mon. Wea. Rev., 123,
 475 1941–1963.
- 476 Zhang, J., K. Howard, C. Langston, S. Vasiloff, B. Kaney, A. Arthur, et al.: National mosaic and multi-sensor QPE
- 477 (NMQ) system: description, results, and future plans. Bulletin of the American Meteorological Society, 92,
- 478 1321-1338, 2011.
- Zhang, J., K. Howard, C. Langston, B. Kaney, B., Y. Qi, L. Tang, H. Grams, Y. Wang, S. Cocks, S. Martinaitis, and
 A. Arthur: Multi-radar multi-sensor (MRMS) quantitative precipitation estimation: Initial operating capabilities.
- 481 Bull. Amer. Meteor. Soc., 97, 621–638, 2016.
- 482
- 483

Table 1. Nor'easter case list. The NESIS number is included for storm severity reference. Mean sea-level pressure
(MSLP) indicates maximum cyclone intensity in GMA. The last two columns denote the first and last times for each
model run. GMA storm tracks are displayed in Fig. 1.

Case	NEGIC	MSLP	Event Detec	Model Run Start	Model Run End	
Number	NE313	(hPa)	Event Dates	Date	Date	
1	N/A	991.5	15-16 Oct 2009	10/15 00UTC	10/20 00UTC	
2	N/A	989.5	07–09 Nov 2012	11/06 18UTC	11/11 18UTC	
3	4.03	972.6	19–20 Dec 2009	12/18 18UTC	12/23 18UTC	
4	2.62	980.5	26–28 Jan 2015	01/25 12UTC	01/30 12 UTC	
5	4.38	979.7	05–07 Feb 2010	02/05 06UTC	02/10 06UTC	
6	1.65	1005.5	02–03 Mar 2009	03/01 00UTC	03/06 00UTC	
7	N/A	993.5	12–14 Mar 2010	03/11 18UTC	03/16 18UTC	

Table 2. Applied bulk microphysics schemes and their characteristics. The below table indicates simulated mixing

491	ratio species and number of moments. Mixing ratio species include: QV = water vapor, QC = cloud water, QH = hail,
402	

Microphysics Scheme	QV	QC	QH	QI	QG	QR	QS	Moments	Citation
Lin6	v	v		v	v	v	v	1	Lin et al. (1983);
LIIIO	Λ	Λ		Λ	Λ	Λ	Λ	1	Rutledge and Hobbs (1984)
CCE6	v	v		v	v	v	v	1	Tao et al. (1989);
GCE0	Λ	Λ		Λ	Λ	Λ	Λ	1	Lang et al. (2007)
GCE7	Х	Х	Х	Х	Х	Х	Х	1	Lang et al. (2014)
WSM6	Х	Х		Х	Х	Х	Х	1	Hong and Lim (2006)
WDM6	Х	Х		Х	Х	Х	Х	2 (QC, QR)	Lim and Hong (2010)

QI = cloud ice, QG = graupel, QR = rain, QS = snow.

494 Table 3. Stage IV-relative, accumulated precipitation threat scores and biases assuming a threshold value of 10 mm 495 (25th percentile of 24 hour accumulated precipitation). Bolded value denote the model simulation with the threat score 496 closest to 1 (perfect forecast) or a bias values closest to 1 (number of forecasted cells matches observations). The 497 lower two panels indicate the number of standards deviations (stdev) each threat score and bias value deviates from 498 the composite (all models + all cases) mean.

499

Threat Score	1	2	3	4	5	6	7	Mean	Mean w/o 4
Lin6	0.289	0.217	0.291	0.091	0.414	0.304	0.332	0.277	0.308
GCE6	0.286	0.243	0.320	0.091	0.406	0.291	0.356	0.285	0.317
GCE7	0.288	0.235	0.319	0.096	0.405	0.300	0.337	0.283	0.314
WSM6	0.293	0.237	0.315	0.093	0.404	0.292	0.356	0.284	0.316
WDM6	0.290	0.243	0.329	0.094	0.411	0.299	0.357	0.289	0.322
Bias	1	2	3	4	5	6	7	Mean	Mean w/o 4
Lin6	2.47	3.53	2.72	7.82	2.22	2.90	1.47	3.30	2.55
GCE6	2.37	3.88	2.85	8.09	2.26	2.93	1.64	3.43	2.66
GCE7	2.52	4.05	2.85	7.76	2.23	2.82	1.57	3.40	2.67
WSM6	2.47	3.75	2.86	8.13	2.26	2.93	1.62	3.43	2.65
WDM6	2.37	3.80	2.76	8.09	2.23	2.82	1.57	3.38	2.59
Threat Score Stats:	All Stdev	0.094	All Mean	0.284				_	
Threat Score	1	2	3	4	5	6	7	-	
Lin6	0.06	-0.71	0.08	-2.05	1.39	0.22	0.52	-	
GCE6	0.03	-0.43	0.39	-2.05	1.31	0.08	0.77		
GCE7	0.05	-0.52	0.38	-2.00	1.29	0.18	0.57		
WSM6	0.10	-0.50	0.34	-2.03	1.28	0.09	0.77		
WDM6	0.07	-0.43	0.48	-2.02	1.36	0.16	0.78	_	
Bias Stats	All Stdev	2.008	All Mean	3.389				-	
Bias	1	2	3	4	5	6	7	-	
Lin6	-0.46	0.07	-0.33	2.21	-0.58	-0.24	-0.96		
GCE6	-0.51	0.24	-0.27	2.34	-0.56	-0.23	-0.87		
GCE7	-0.43	0.33	-0.27	2.18	-0.58	-0.28	-0.91		
WSM6	-0.46	0.18	-0.26	2.36	-0.56	-0.23	-0.88		
WDM6	-0.51	0.20	-0.31	2.34	-0.58	-0.28	-0.91		





Figure 1. Nested WRF configuration used in simulations. The large panel shows the first 3 model domains (45-, 15 , 5- km grid spacing, respectively). The smaller panels show the location of domain 4 (1.667-km resolution) for each
 of the seven cases. The colored lines show the cyclone track as indicated by GMA for each nor'easter case.



505

506 Figure 2. Domain 4 (1.667 km grid spacing), precipitable mixing ratios (mm) at 06 UTC 06 Feburary 2010. Shown

507 abbreviations for mixing ratios include: QV = water vapor, QC = cloud water, QG = graupel, QI = cloud ice, QR = 508 rain, QS = snow.



509

510 Figure 3. Simulated radar reflectivity (dBZ) at 4,000 m above mean sea level and their difference at the same time as

511 Fig. 2.



512

Figure 4. Domain 4 (1.167-km grid spacing), composite mixing ratios (kg kg⁻¹), temperature (K), and vertical velocities (cm s⁻¹) composited over Case 5 (00 UTC 6–7 January 2010). The black dashed lines denote the height above mean sea level (MSL) where the air temperature is 0°C or -40°C. The upper-left panel shows composited and model-averaged profiles of temperature (red line) and vertical velocity (blue). Mixing ratio species abbreviations are QCLOUD (cloud water), QGRAUP (graupel), QICE (cloud ice), QRAIN (rain), QSNOW (snow) and QHAIL (hail).



Figure 5. Domain 4 (1.167-km grid spacing), composite mixing ratios (kg kg⁻¹), temperature (K), and vertical velocities
(cm s⁻¹) composited over all seven nor'easter events. The black dashed lines denote the height above mean sea level
(MSL) where the air temperature is 0°C or -40°C. The upper-left panel shows composited and model-averaged profiles
of temperature (red line) and vertical velocity (blue). Mixing ratio species abbreviations are QCLOUD (cloud water),
QGRAUP (graupel), QICE (cloud ice), QRAIN (rain), QSNOW (snow) and QHAIL (hail).



525 526 Figure 6. Case 5, 24-hour precipitation accumulation and their differences (mm, small panels) and corresponding 527 probability density and cumulative distribution functions (big panel) of these same data derived from Stage IV and 528 WRF model output. Accumulation period is from 00 UTC 06 February 2010 - 00 UTC 07 February 2010. Shown

529 differences are model - Stage IV (StIV).



530 531

Figure 7. Case 7, 24-hour precipitation accumulation and their differences (mm, small panels) and corresponding
 probability density and cumulative distribution functions (big panel) of these same data derived from Stage IV and

- 534 WRF model output. Accumulation period is from 18 UTC 12 March 2010 18 UTC 13 March 2010. Shown
- 535 differences are model Stage IV (StIV).



536

537 **Figure 8.** Contoured frequency with altitude diagram (CFAD) of radar reflectivity and indicated differences from Case

538 4 (January 2015). Data accumulation period spans 12 UTC 26 January 2015 – 12 UTC 27 January 2015 during the transit

of the nor'easter through WRF model domain 4 (1.167 km grid spacing). The y-axis shows height above mean sea level

540 (HMSL).



Figure 9. MRMS radar reflectivity and WRF simulated radar reflectivity (dBZ) at 4,000 m above sea level at 18 UTC

543 26 January 2015. Show radar reflectivity differences are as indicated.



544

545 Figure 10. MRMS observed radar and WRF simulated radar reflectivity (dBZ) at 9,500 m above sea level at 18 UTC

546 26 January 2015. Show radar reflectivity differences are as indicated.



548

549 **Figure 11.** Domain 4 (1.667 km grid spacing), hourly CFAD scores (See Eq. 2) of radar reflectivity and indicated

```
differences from Case 4 starting 12 UTC 26 January 2015 and ending on 12 UTC 27 January 2015. The time period
```

551 corresponds to the same time period as in Figure 5. The y-axis shows height above mean sea