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

Stephen D. Nicholls\textsuperscript{1,2}, Steven G. Decker\textsuperscript{3}, Wei-Kuo Tao\textsuperscript{1}, Stephen E. Lang\textsuperscript{1,4}, Jainn J. Shi\textsuperscript{1,5}, and Karen I. Mohr\textsuperscript{1}

\textsuperscript{1}NASA-Goddard Space Flight Center, Greenbelt, 20716, United States of America
\textsuperscript{2}Joint Center for Earth Systems Technology, Baltimore, NASA-Goddard Space Flight Center, Baltimore, 21250, United States of America
\textsuperscript{3}Department of Environmental Sciences, Rutgers, The State University of New Jersey, 08850, United States of America
\textsuperscript{4}Science Systems and Applications, Inc., Lanham, 20706, United States of America
\textsuperscript{5}Goddard Earth Sciences Technology and Research, Morgan State University, 21251, United States of America

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

1 Introduction

Bulk microphysical parameterization schemes (BMPSs) within numerical weather prediction models have become increasingly complex and computationally expensive. Modern prognostic weather models, such as the Weather Research and Forecasting (WRF) model (Skamarock et al., 2008), offer BMPS options ranging from simplistic, warm rain physics (Kessler, 1969) to complex, six-class, two-moment microphysics (Morrison et al., 2009). Microphysics and cumulus parameterizations drive cloud and precipitation processes within numerical weather prediction models and directly or indirectly impacts radiation, moisture, aerosols, and other simulated processes. Citing its importance, Tao et al. (2011) detailed more than 36 published, microphysics-focused studies focusing on idealized simulations, hurricanes, or mid-latitude convection. More recently, the observational studies of Stark (2012) and Ganetis and Colle (2015) investigated microphysical species variability within United States (U.S.) east coast winter-time cyclones (locally called “nor’easters”) and have called for further studies investigating how microphysical 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.
Recent nor’easter studies are scarce given the extensive research efforts in the 1980s. These historical studies addressed key nor’easter drivers including frontogenesis and baroclinicity (Bosart, 1981; Forbes et al., 1987; Stauffer and Warner, 1987), anticyclones (Uccellini and Kocin, 1987), latent heat release (Uccellini et al., 1987), and moisture transport by the low-level jet (Uccellini and Kocin, 1987; Mailhot and Chouinard, 1989). Despite extensive observational analyses, less attention has been provided to mid-latitude, winter cyclone simulations, especially those 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.

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

2 Methods

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. 2011). Model parameterizations include:

- Longwave radiation: New Goddard Scheme (Chou and Suarez, 1999; Chou and Suarez, 2001)
- Shortwave radiation: New Goddard Scheme (Chou and Suarez, 1999)
- Surface layer: Eta similarity (Monin and Obukhov, 1954; Janjic, 2002)
- Land surface: NOAH (Chen and Dudhia, 2001)
- 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.

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

To investigate BMPS influence upon W361 nor’eastern 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’eastern cases.

### 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
coverage. Stage IV is a six-hourly, 4-km resolution, gridded precipitation product covering the United States and is derived from rain gauge and radar data. Finally, MRMS is two minute, 1.3-km resolution, gridded 3D volume radar mosaic product derived from S- and C-band radars covering the United States and Southern Canada (Zhang et al. 2016). MRMS serves as an operational successor to the better known National Mosaic and Multi-Sensor QPE (NMQ; Zhang et al. 2011) radar mosaic products. Both Stage IV and MRMS, however are limited by the detection range of their surface-based assets. All cross comparisons between WRF and these validation data were conducted at identical 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:

$$PMR = \frac{1}{\rho g} \int P_{sfc} - P_{top} w dp$$

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

$$CS = 1 - \frac{\sum |PDF_m - PDF_o| h}{200}$$

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

3. Results

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, 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 (Fig. 5) from Domain 4. Composite profiles are averaged over the residence time of the nor’easter within Domain 4 (24-30 hours). To emphasize the fraction of supercooled water, two sets of dashed black lines are added to each panel in Figs. 4 and 5 to indicate the 0°C and -40°C heights from each model simulation. We exclude hail from our analysis because it is unique to GCE7 and its mixing ratio values are an order of magnitude smaller than other species.

Comparing Figs. 2 and 4 to Fig. 3, reveals a strong correspondence between radar reflectivity signatures and particular precipitable hydrometeor species structures, especially graupel and snow and to a lesser extent cloud water. Analysis of Fig. 4 reveals that cloud water at 4,000 m is super-cooled and graupel mixing ratios values are near their peak and given the corresponding precipitation mixing ratio values in Fig. 2, these two species are well correlated with the strongest, convective reflectivity signatures (> 35 dBZ). Fig. 4 also reveals snow mixing ratio, except for Lin6 are also comparatively high at this level, yet precipitable snowfall values better correlate best with moderate reflectivity (20-35 dBZ) regions within the broader, more stratiform, precipitation shield. Notably, for Lin6, reduced snow mixing ratios are partially offset by an increase of graupel mixing ratio values within the precipitation shield. Inter-BMPS mixing ratio variability amongst BMPSs, both at this level and throughout the troposphere, is due to identifiable trends within the underlying assumptions made by BMPSs and will explained in more detail below.

All evaluated BMPSs share a common heritage in the Lin6 scheme. With the exception of the two-moment cloud water and rain and CCN-cloud droplet feedbacks in WDM6, the BMPSs differ primarily in how each addresses frozen hydrometeor species (cloud ice, graupel, and snow). Their common programming heritage is evident from the nearly identical (exception: WDM6) rain mixing ratio profiles (Figs. 4 and 5) and precipitable water vapor (Fig. 2) and is consistent with Wu and Petty (2010). WDM6, unlike single-moment BMPSs, explicitly forecasts CCN, rain and cloud droplet number concentrations and does not apply derivative equations (Hong et al., 2010). The forecasts result produce minimal changes to maximum mixing ratio height (Figs. 4 and 5) and precipitable rain coverage (Fig. 2), yet 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 water mixing ratios. Although in WDM6 cloud water is double-moment, the maximum mixing ratios are only decreased slightly relative to WSM6. This result suggests that WDM6-forecasted cloud water number concentrations are likely close to prescribed 300 cm$^{-3}$ number concentration assumed in WSM6 (Hong et al. 2010) and/or the larger-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
size distribution intercept (decreased surface area), and its auto-conversion of snow to either graupel or hail at high
mixing ratios. GCE6 turns off dry collection of snow and ice by graupel, greatly increasing the snow mixing ratios at
the expense of graupel and reducing snow riming efficiency (Lang et al. 2007). Snow growth in GCE6 is further
augmented by its assumption of water saturation for the vapor growth of cloud ice to snow (Reeves and Dawson,
2013; Lang et al. 2014). GCE7 addressed the vapor growth issue of GCE6 and applied numerous other changes
including the introduction and of a snow size and density mapping, snow breakup interactions, a relative humidity
(RH) correction factor, and a new vertical-velocity-dependent ice super saturation assumption (Lang et al., 2007; Lang
et al., 2011; Lang et al., 2014; Chern et al., 2016; Tao et al., 2016). Despite the reduced efficiency of vapor growth of
cloud ice to snow stemming from the both the new RH correction factor and the ice super saturation adjustment, the
new snow mapping and enhanced cloud ice to snow auto-conversion in GCE7 offset this potential reduction and keep
GCE snowfall mixing ratio higher than in non-GCE BMPSs. Unlike Lin6, WSM6 and WDM6 assume grid cell graupel
and snow fall speeds are identical (Dudhia et al., 2008) and that ice nuclei concentration is a function of temperature
(Hong et al., 2008). These two aspects, effectively eliminate the accretion of snow by graupel and increase snow
mixing ratios at colder temperatures (Dudhia et al., 2008; Hong et al., 2008). Figure 4 and 5 show the height of
maximum snow mixing ratio is roughly conserved in all non-Lin6 BMPSs. Lin6’s assumption of non-uniform graupel
and snow fall speeds and dry collection of snow by graupel reduce snow mixing ratios in the middle troposphere and
raise its maximum snow mixing ratio height.

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

Although cloud ice mixing ratios are up to ninety percent smaller than those for snow (GCE6), cloud ice mixing
ratios still vary greatly amongst the BMPSs as illustrated in Figs. 2, 4, and 5. Cloud ice mixing ratios are highest in
GCE7 and lowest in Lin6. Wu and Petty (2010) similarly found low cloud ice mixing ratios in Lin6 simulations and
ascribe it to dry collection by cloud ice by graupel and its fixed cloud-ice size distribution. Similar to Lin6, GCE6
uses a monodispersed cloud-ice size distribution (20 μm diameter), but assumes vapor growth of cloud ice to snow
under an assumption of water saturation conditions (yet supersaturated with respect ice) leading to higher cloud ice
amounts, but also increased cloud ice to snow conversion rates (Lang et al., 2011; Tao et al., 2016). GCE7 blunt this
cloud ice to snow conversion term using a RH correction factor which is dependent upon ice supersaturation which is
itself dependent up vertical velocity. Additionally, GCE7, also includes contact and immersion freezing terms (Lang
et al., 2011), makes the cloud ice collection by snow efficiency a function of snow size (Lang et al., 2011; Lang et al.,
2014), sets a maximum limit on cloud-ice particle size (Tao et al., 2016), makes ice nuclei concentrations follows the
Cooper curve (Cooper, 1986; Tao et al., 2016), and it allows cloud ice to persist in ice subsaturated conditions (i.e., RH for ice ≥ 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 (See Fig. 2). Similar to GCE7, WSM6 runs generate larger cloud ice mixing ratios than Lin6, which Wu and Petty (2010) attribute to excess cloud glaciation at temperatures between 0°C and -20°C and its usage of fixed cloud ice size 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 overall (Figs. 4 and 5) are consistent between BMPSs.

3.2 Stage IV precipitation analysis

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

Table 3 shows Case 4 as a clear outlier where its low threat score and bias values deviate more than two standard deviation from the composite mean due to its non-coastal track (Fig. 1) and thus it will be excluded from this section of the analysis. For the remaining six cases, Table 4 indicates low (0.217; Lin6, Case 2) to moderate (0.414; Lin6, Case 5) threat scores and a 10 mm precipitation contour spatial covers an area far exceeding Stage IV (bias range: 1.47 [Lin6, Case 7] – 4.05 [GCE7, Case 3]). Inter-BMPS barely varied with threat score and biases varying only up to an order of magnitude less than the threat and bias scores themselves. Consistent with Hong et al. (2010), threat score and bias values for WSM6 are equal to or improved upon by WDM6 due to its inclusion of a cloud condensation nuclei (CCN) feedback. Overall, WDM6 generated marginally better simulated precipitation fields and has the lowest 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 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 sparse North Atlantic, a region with no rain gauge data and is at the 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 consistent with investigation of simulation precipitation during warm-season events and quasi-stationary front (Fritsch and Carbone, 2004; Wang and Clark 2010).

3.3 MRMS and radar reflectivity analysis

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

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

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

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

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

GFS model analysis data boundary condition data can be obtained from the NASA’s open access, NOMADS data server (ftp://nomads.ncdc.noaa.gov/GFS/Grid3/). Stage IV precipitation data is publically available from the
National Data and Software Facility at the University Center for Atmospheric Research (http://data.eol.ucar.edu/cgi-bin/codiac/fgr_form/id=21.093).

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.

8 Acknowledgements

This research was supported by the Joint Center for Earth Systems Technology (JCET), the University of Maryland Baltimore County (UMBC), and in part by the New Jersey Agricultural Experiment Station. Resources supporting this work were provided by the NASA High-End Computing (HEC) Program through the NASA Center for Climate Simulation (NCCS) at Goddard Space Flight Center.

References


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.


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.

<table>
<thead>
<tr>
<th>Case Number</th>
<th>NESIS</th>
<th>MSLP (hPa)</th>
<th>Event Dates</th>
<th>Model Run Start Date</th>
<th>Model Run End Date</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>N/A</td>
<td>991.5</td>
<td>15–16 Oct 2009</td>
<td>10/15 00UTC</td>
<td>10/20 00UTC</td>
</tr>
<tr>
<td>2</td>
<td>N/A</td>
<td>989.5</td>
<td>07–09 Nov 2012</td>
<td>11/06 18UTC</td>
<td>11/11 18UTC</td>
</tr>
<tr>
<td>3</td>
<td>4.03</td>
<td>972.6</td>
<td>19–20 Dec 2009</td>
<td>12/18 18UTC</td>
<td>12/23 18UTC</td>
</tr>
<tr>
<td>4</td>
<td>2.62</td>
<td>980.5</td>
<td>26–28 Jan 2015</td>
<td>01/25 12UTC</td>
<td>01/30 12 UTC</td>
</tr>
<tr>
<td>5</td>
<td>4.38</td>
<td>979.7</td>
<td>05–07 Feb 2010</td>
<td>02/05 06UTC</td>
<td>02/10 06UTC</td>
</tr>
<tr>
<td>6</td>
<td>1.65</td>
<td>1005.5</td>
<td>02–03 Mar 2009</td>
<td>03/01 00UTC</td>
<td>03/06 00UTC</td>
</tr>
<tr>
<td>7</td>
<td>N/A</td>
<td>993.5</td>
<td>12–14 Mar 2010</td>
<td>03/11 18UTC</td>
<td>03/16 18UTC</td>
</tr>
</tbody>
</table>
Table 2. Applied bulk microphysics schemes and their characteristics. The below table indicates simulated mixing ratio species and number of moments. Mixing ratio species include: QV = water vapor, QC = cloud water, QH = hail, QI = cloud ice, QG = graupel, QR = rain, QS = snow.

<table>
<thead>
<tr>
<th>Microphysics Scheme</th>
<th>QV</th>
<th>QC</th>
<th>QH</th>
<th>QI</th>
<th>QG</th>
<th>QR</th>
<th>QS</th>
<th>Moments</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin6</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>1</td>
<td>Lin et al. (1983); Rutledge and Hobbs (1984)</td>
</tr>
<tr>
<td>GCE6</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>1</td>
<td>Tao et al. (1989); Lang et al. (2007)</td>
</tr>
<tr>
<td>GCE7</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>1</td>
<td>Lang et al. (2014)</td>
</tr>
<tr>
<td>WSM6</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>1</td>
<td>Hong and Lim (2006)</td>
</tr>
<tr>
<td>WDM6</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>2 (QC, QR)</td>
<td>Lim and Hong (2010)</td>
</tr>
</tbody>
</table>
Table 3. Stage IV-relative, accumulated precipitation threat scores and biases assuming a threshold value of 10 mm (25th percentile of 24 hour accumulated precipitation). Bolded value denote the model simulation with the threat score closest to 1 (perfect forecast) or a bias values closest to 1 (number of forecasted cells matches observations). The lower two panels indicate the number of standards deviations (stdev) each threat score and bias value deviates from the composite (all models + all cases) mean.

<table>
<thead>
<tr>
<th>Threat Score</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Mean</th>
<th>Mean w/o 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin6</td>
<td>0.289</td>
<td>0.217</td>
<td>0.291</td>
<td>0.091</td>
<td><strong>0.414</strong></td>
<td><strong>0.304</strong></td>
<td>0.332</td>
<td>0.277</td>
<td>0.308</td>
</tr>
<tr>
<td>GCE6</td>
<td>0.286</td>
<td><strong>0.243</strong></td>
<td>0.320</td>
<td>0.091</td>
<td>0.406</td>
<td>0.291</td>
<td>0.356</td>
<td>0.285</td>
<td>0.317</td>
</tr>
<tr>
<td>GCE7</td>
<td>0.288</td>
<td>0.235</td>
<td>0.319</td>
<td><strong>0.096</strong></td>
<td>0.405</td>
<td>0.300</td>
<td>0.337</td>
<td>0.283</td>
<td>0.314</td>
</tr>
<tr>
<td>WSM6</td>
<td><strong>0.293</strong></td>
<td>0.237</td>
<td>0.315</td>
<td>0.093</td>
<td>0.404</td>
<td>0.292</td>
<td>0.356</td>
<td>0.284</td>
<td>0.316</td>
</tr>
<tr>
<td>WDM6</td>
<td>0.290</td>
<td><strong>0.243</strong></td>
<td><strong>0.329</strong></td>
<td>0.094</td>
<td>0.411</td>
<td>0.299</td>
<td><strong>0.357</strong></td>
<td><strong>0.289</strong></td>
<td><strong>0.322</strong></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Bias</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>Mean</th>
<th>Mean w/o 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin6</td>
<td>2.47</td>
<td><strong>3.53</strong></td>
<td><strong>2.72</strong></td>
<td>7.82</td>
<td><strong>2.22</strong></td>
<td>2.90</td>
<td><strong>1.47</strong></td>
<td><strong>3.30</strong></td>
<td>2.55</td>
</tr>
<tr>
<td>GCE6</td>
<td><strong>2.37</strong></td>
<td>3.88</td>
<td>2.85</td>
<td>8.09</td>
<td>2.26</td>
<td>2.93</td>
<td>1.64</td>
<td>3.43</td>
<td>2.66</td>
</tr>
<tr>
<td>GCE7</td>
<td>2.52</td>
<td>4.05</td>
<td>2.85</td>
<td><strong>7.76</strong></td>
<td>2.23</td>
<td>2.82</td>
<td>1.57</td>
<td>3.40</td>
<td>2.67</td>
</tr>
<tr>
<td>WSM6</td>
<td>2.47</td>
<td>3.75</td>
<td>2.86</td>
<td>8.13</td>
<td>2.26</td>
<td>2.93</td>
<td>1.62</td>
<td>3.43</td>
<td>2.65</td>
</tr>
<tr>
<td>WDM6</td>
<td><strong>2.37</strong></td>
<td>3.80</td>
<td>2.76</td>
<td>8.09</td>
<td>2.23</td>
<td><strong>2.82</strong></td>
<td>1.57</td>
<td>3.38</td>
<td>2.59</td>
</tr>
</tbody>
</table>

Threat Score Stats: All Stdev 0.094 All Mean 0.284

<table>
<thead>
<tr>
<th>Threat Score</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin6</td>
<td>0.06</td>
<td>-0.71</td>
<td>0.08</td>
<td>-2.05</td>
<td>1.39</td>
<td>0.22</td>
<td>0.52</td>
</tr>
<tr>
<td>GCE6</td>
<td>0.03</td>
<td>-0.43</td>
<td>0.39</td>
<td>-2.05</td>
<td>1.31</td>
<td>0.08</td>
<td>0.77</td>
</tr>
<tr>
<td>GCE7</td>
<td>0.05</td>
<td>-0.52</td>
<td>0.38</td>
<td>-2.00</td>
<td>1.29</td>
<td>0.18</td>
<td>0.57</td>
</tr>
<tr>
<td>WSM6</td>
<td>0.10</td>
<td>-0.50</td>
<td>0.34</td>
<td>-2.03</td>
<td>1.28</td>
<td>0.09</td>
<td>0.77</td>
</tr>
<tr>
<td>WDM6</td>
<td>0.07</td>
<td>-0.43</td>
<td>0.48</td>
<td>-2.02</td>
<td>1.36</td>
<td>0.16</td>
<td>0.78</td>
</tr>
</tbody>
</table>

Bias Stats: All Stdev 2.008 All Mean 3.389

<table>
<thead>
<tr>
<th>Bias</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lin6</td>
<td>-0.46</td>
<td>0.07</td>
<td>-0.33</td>
<td>2.21</td>
<td>-0.58</td>
<td>-0.24</td>
<td>-0.96</td>
</tr>
<tr>
<td>GCE6</td>
<td>-0.51</td>
<td>0.24</td>
<td>-0.27</td>
<td>2.34</td>
<td>-0.56</td>
<td>-0.23</td>
<td>-0.87</td>
</tr>
<tr>
<td>GCE7</td>
<td>-0.43</td>
<td>0.33</td>
<td>-0.27</td>
<td>2.18</td>
<td>-0.58</td>
<td>-0.28</td>
<td>-0.91</td>
</tr>
<tr>
<td>WSM6</td>
<td>-0.46</td>
<td>0.18</td>
<td>-0.26</td>
<td>2.36</td>
<td>-0.56</td>
<td>-0.23</td>
<td>-0.88</td>
</tr>
<tr>
<td>WDM6</td>
<td>-0.51</td>
<td>0.20</td>
<td>-0.31</td>
<td>2.34</td>
<td>-0.58</td>
<td>-0.28</td>
<td>-0.91</td>
</tr>
</tbody>
</table>
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.
Figure 2. Domain 4 (1.667 km grid spacing), precipitable mixing ratios (mm) at 06 UTC 06 February 2010. Shown abbreviations for mixing ratios include: QV = water vapor, QC = cloud water, QG = graupel, QI = cloud ice, QR = rain, QS = snow.
Figure 3. Simulated radar reflectivity (dBZ) at 4,000 m above mean sea level and their difference at the same time as Fig. 2.
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).
Figure 6. Case 5, 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 WRF model output. Accumulation period is from 00 UTC 06 February 2010 – 00 UTC 07 February 2010. Shown differences are model - Stage IV (StIV).
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 WRF model output. Accumulation period is from 18 UTC 12 March 2010 – 18 UTC 13 March 2010. Shown differences are model - Stage IV (StIV).
Figure 8. Contoured frequency with altitude diagram (CFAD) of radar reflectivity and indicated differences from Case 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 (HMSL).
Figure 9. MRMS radar reflectivity and WRF simulated radar reflectivity (dBZ) at 4,000 m above sea level at 18 UTC 26 January 2015. Show radar reflectivity differences are as indicated.
Figure 10. MRMS observed radar and WRF simulated radar reflectivity (dBZ) at 9,500 m above sea level at 18 UTC 26 January 2015. Show radar reflectivity differences are as indicated.
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 corresponds to the same time period as in Figure 5. The y-axis shows height above mean sea