Responses to Anonymous Referee #1 (Page 1)

My co-authors and I wish acknowledge and thank Reviewer #1 for the time, energy, and effort applied in the detailed review of this manuscript. We do feel that a more narrow focus on microphysics and removal of the energy norm has improved upon the original manuscript and also address most if not all of the highlighted concerns.

Responses to General Comments:

1) "…my main issue with the paper which is whether we can evaluate microphysics schemes against analyses such as these in a useful way."

Both your comments and those of Reviewer #2 highlight this point. While we do believe that GFS analysis data can be useful for broader themes of our analysis (e.g., large-scale water vapor fields), its coarseness proves problematic was addressing specific microphysical-related questions. The revised manuscript now includes a new analysis making use of the Multi-Radar Multi-Sensor (MRMS) 3D volume data. These observation data, we argue, permit a more thorough investigation of smaller-scale impacts from the microphysics.

2) "Errors in the forecast are dominated by other causes, such as the initial analysis error, considering that these are initialized 72 hours ahead of the precipitation events. Perhaps initializing closer to the event would have given more accurate representations that could be compared with analyses."

In light of your suggestion and a similar comment from Reviewer #2, we shifted the model initialization time forward until 24 hours prior to cyclogenesis off the Mid-Atlantic United States and re-ran all 35 WRF model simulations. We believe that initializing 24 hours prior to cyclogenesis is ideal because it ensures each model simulation is sufficiently spun-up prior to the main cyclogenesis period and yet there are only minimal deviations (< 50 km) between WRF simulations and the GFS model analysis storm tracks.

3) "I especially am not convinced that the energy norm metric has been demonstrated to be useful."

We concur and agree that the energy norm, although useful, is not the most effective vehicle by which to evaluate microphysical-related simulation errors. Thus the energy norm would be more apt in a more general, bulk analysis of nor'easters where a focus on large-scale players are key. Due to our shift in model initialization time (see #2 above) and our shift to focus on microphysics (see #1 above), the energy norm analysis has been redacted from the revised manuscript.

4) "There are also aspects of the model set-up that I would criticize. It seems that the central 1.67 km domain is at the same position for all storms, and this means that some storms pass through it while other would miss it and only be resolved in the 5 km domain"

The WRF model domain positions were fixed for all nor'easter cases. This lead to a situation WRF-simulated nor'easters in cases 1 and 4 either missed or never fully entered the 1.667 km model grid (Domain 4) as the reviewer hypothesized. We have since increased the sizes of the 5 km and 1.167 (Domains 3 and 4, respectively) by 50%, shifted domain 3 southward, and tailored the location of domain 4 for all seven nor'easter events. To physically demonstrate these changes, Figure 1 shows our original and new WRF model configuration. All 35 model simulations were re-run and reanalyzed accordingly. As can be seen below, each model analysis track moves through the center of each respective domain 4.

Responses to Anonymous Referee #1 (Page 2)

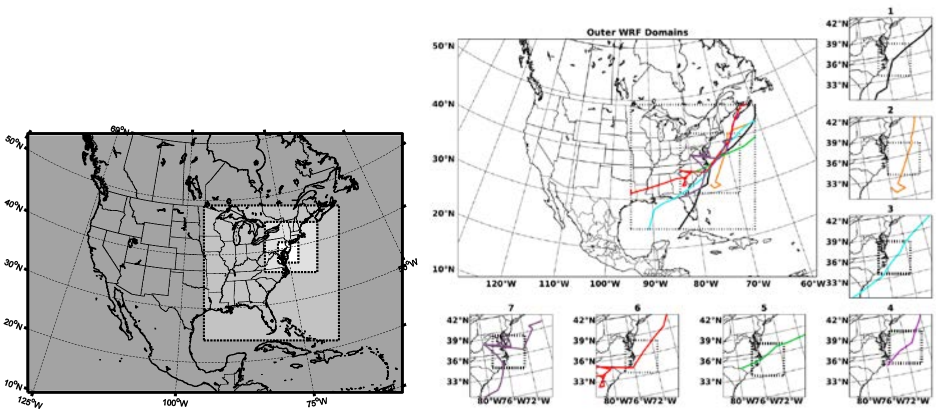


Fig. 1: Nested WRF configuration for the original manuscript (left) and the revised manuscript (right). The colored lines in the right panel show the GFS model analysis storm tracks for each of the seven cases.

Specific Comments:

1. line 141. What are the perturbations relative to, the GMA analysis? This is not stated.

All energy norm calculations are relative to the GFS model analysis. The energy norm section has been removed from the paper.

2. Section 3.2. It is not clear what area these results and Table 4 are for. It also seems that much of this would be in the 12 km domain where there is a cumulus scheme, and part is in domains 3 and 4 where there isn't.

Table 4 was originally based upon domain 2 (15 km domain). The revised manuscript keeps the same approach, but we use domain 3 (5 km grid spacing) instead because it is of similar resolution to the Stage IV precipitation product (4 km resolution), the cumulus parameterization is turned off, and we felt that domain 4 would be over too limited an area for comparison.

3. line 208. WRF's common heritage with GFS is implied. I don't think there is much common physics heritage except for some relationship in the land-surface scheme. What is meant here?

My assumption here was based upon that simulated storm tracks between GFS and WRF would be similar given WRF's common heritage in GFS. Similar tracks would, in theory, give a greater potential of similar forecasts. My comment about this heritage is no longer necessary and it has been removed from the revised manuscript.

Responses to Anonymous Referee #1 (Page 3)

4. Abstract does not mention that there are seven cases and five microphysics schemes and has nothing on the energy norm. It is not adequately describing the work carried out.

Given the significant changes to the manuscript in this revision, the abstract has been updated and overhauled to more aptly describe the work conducted.

5. line 234. What is meant by saturation heights?

Thank you for this asking this clarification. By saturation height, I am referring to the height at which each microphysical species reached its maximum value. This value however is part of the mixing ratio profile and I think distracts from the paper. I have elected to remove this term from the revised manuscript.

6. line 236. cloud water? This should probably be cloud droplet number concentration?

Thank you for finding this error. “Cloud water” has been changed to “cloud droplet number concentration” in the revised manuscript.

7. line 241-246. Without knowing where the freezing level is, it is difficult to follow this discussion. How much of the cloud water is supercooled?

Thank you for noting this challenge to understanding the microphysical species analysis section. To provide information on how much of the cloud water is super cooled, I have modified the composite mixing ratio diagrams with two dashed black lines which indicate both the 0°C and -40°C levels.

8. line 279. How does lack of a sedimentation term lead to low cloud ice? I thought sedimentation should reduce cloud ice extent and lifetime.

Thank you for the noting this logic error. A quick read into the literature found a cloud resolving model study addressing this very topic. Their findings do indeed show that the impact of the sedimentation in cloud ice is to increase its conversion rate to snow and graupel and thus decreasing the amount, extent, and lifetime of cloud ice hydrometeors. I have removed the erroneous comment from the revised manuscript.

Nomura, M., Tsuboki, K. and Shinoda, T., 2012. Impact of Sedimentation of Cloud Ice on Cloud-Top Height and Precipitation Intensity of Precipitation Systems Simulated by a Cloud-Resolving Model. *気象集誌. 第 2 輯*, *90*(5), pp.791-806.

9. line 282. ’assumed water saturation’. What assumption is made about water saturation in a purely ice process?

The original GCE6 scheme generated excess super cooled cloud water at temperature below -12°C where such droplets do not often occur. Therefore water saturation was extended down to much colder temperatures which allowed cloud ice to achieve supersaturation with respect to ice and made cloud ice to snow conversion rates

For further details please refer to page 2308 of the following reference:

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*.

Responses to Anonymous Referee #1 (Page 4)

10. Figure 7 (vapor) would have been better presented as a difference from analysis. Nothing can be seen with this plot as it is.

Thank you for the suggestion. In new manuscript, this diagram (now Figure 2) has been updated to show the difference in water vapor.

11. Section 3.4. It is hard to interpret what is meant by lowest energy norms and the metrics in Table 5 in general. Also make clearer what is meant by model-relative and GMA-relative norms.

"GMA-relative" denotes diagnosing the simulated environment within a 600-km wide box centered on the GMA-indicated cyclone center in both GMA and each WRF simulation. "Model-relative" uses the same box, but centers it on the cyclone center determined from each individual model simulation. The energy norm analysis is no longer part of the manuscript.

12. As mentioned in the general comments, I do not think the energy norm statistics are adding anything useful to the paper. It would be better and more focused without this. There are so many factors that could make one simulation look temporarily better than another, related to timing and structure developments, that using such a high-level bulk measure as this conflates too many things to be useful in such an intercomparison.

While we do see some value in the energy norm results with respect to diagnosing which dynamical fields are responsible for observed error, we agree that in context of a microphysics- focused paper this metric is not sensitive enough to be of use. Pending the suggestion of both reviewers, this section has been redacted from the revised manuscript.

13. line 334. Regarding the low-level jet which case is being referred to? Can it really be inferred from the v component of the energy norm that this jet is the cause? This looks highly speculative.

We agree with the reviewer's viewpoint that the energy norm by itself could be considered speculative for Case 7. Our decision to not include a figure of 850-hPa winds (See Figure 2 below) in the original manuscript was made on the assumption that presence of the cyclone center, the small size of the model domain, and a bump in the u and v energy norm components at 850-hPa would be sufficient circumstantial evidence to support our claim without the need for an additional figure. In the revised manuscript, the energy norm section has been removed from the paper.

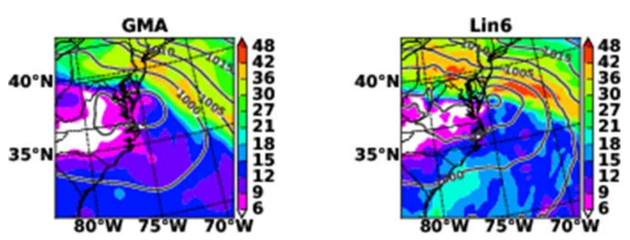


Fig. 2: 850-hPa wind speed (fills, m s-1) and sea-level pressure (contours, hPa) on 13 March 2010 at 18 UTC (Case 7).

Responses to Anonymous Referee #2 (Page 1)

My co-authors and I wish to thank Reviewer #2 for their time and consideration in reviewing this manuscript. Many comments are consistent with those of Reviewer #1 and have been incorporated into the revised manuscript.

General Comments

1) *I think that the spin-up time of 72 hours is too long for a simulation without any kind of assimilation. A test with a shorter spin up (12 hours) could be recommendable*

In light of your suggestion and a similar comment from Reviewer #1, we shifted the model initialization time forward until 24 hours prior to cyclogenesis off the Mid-Atlantic United States and re-ran all 35 WRF model simulations. We set our start time 24 hours beforehand because simulated radar reflectivity fields still appeared slightly "blooby" up through 9-10 hours. Starting the model simulations 24 hours before primary cyclogenesis allowed for full development of simulated radar reflectivity structures and WRF-GMA track differences tended to be modest (<50 km).

2) *"A microphysical comparison with observations could be useful because this topic is the main focus of the paper. Is it possible to retrieve data from radar or satellite platform"*

Thanks to your suggestion, we have given this revised paper more of a microphysics-style focus. I looked both into TRMM and CloudSat 2C-Ice products. TRMM offers a wide range radar observations but its orbital inclination is 35 degree (http://disc.sci.gsfc.nasa.gov/precipitation/additional/instruments/trmm_instr.html), which limits its usefulness when only half my analysis domains falls equatorward of 35°N. CloudSAT does provide profiles cloud ice, which my colleague used in a recent paper on global cloud species. It narrow swath range (see Figure 3) made getting a consistent "hit" on a nor'easter challenging.

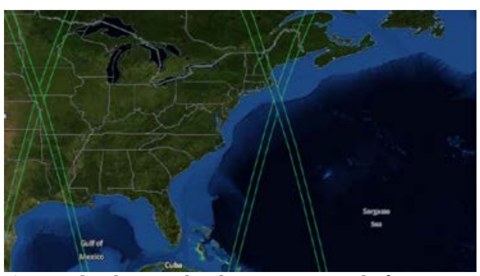


Fig. 3: CloudSAT orbital overpass sample from 2012.

I did find success with the Multi-Radar Multi-Sensor product from National Oceanagrahic and Atmospheric Association (NOAA), which provides hourly gridded 3D volume scans at 1-hour intervals (See Figure 4). Similar to StageIV, MRMS data only covers part of domain 4 in many of the seven cases, but the results thus far have been reasonable and useful.

Responses to Anonymous Referee #2 (Page 2)

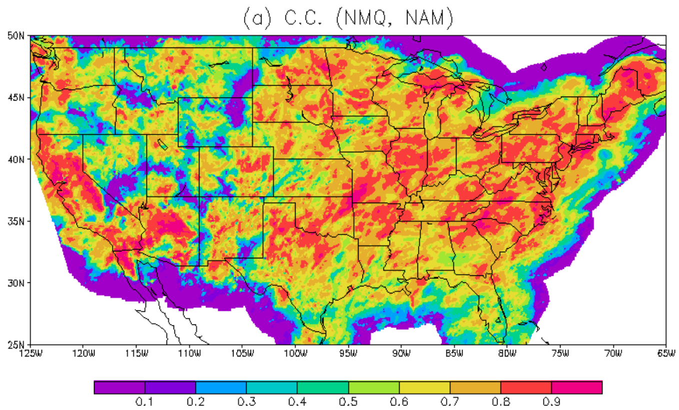


Fig. 4: MRMS coverage area (everywhere with colors).

Specific Comments:

1) Line 133: w is the mixing ratio of rain?

Although 'w' is often used in meteorology to denote mixing ratio, it represents vertical velocity in the energy norm equation. Instead, this formula uses 'q' to represent mixing ratio. With the removal of the energy norm from the paper's results this particular comment is no longer valid.

2) Line 203: Not Fig. 4 but Fig. 5

Thank you for catching the typo. I have corrected the manuscript to refer to Fig. 5.

3) Figs. 5-6-7: insert letters in the panel to easy the reading of section 3.

While I will not dispute that Figs. 5-7 do attempt to show much data. In an earlier form of this paper, I actually tried putting letters into the panels, but these letters were difficult to place without blocking or interfering with the displayed data. I thank you for the suggestion, but I have decided to keep my "Microsoft Excel-like" approach to plot labelling.

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

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

[1]*NASA-Goddard Space Flight Center, Greenbelt, 20716, United States of America*

[2]*Joint 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

[4]*Science Systems and Applications, Inc., Lanham, 20706, United States of America*

[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.

~~**Abstract.** This study evaluated the impact of five, single- or double- moment bulk microphysics schemes (BMPS) on Weather Research and Forecasting (WRF) model (version 3.6.1) winter storm simulations. Model simulations were integrated for 180 hours, starting 72 hours prior to the first measurable precipitation in the highly populated Mid-Atlantic U.S. Simulated precipitation fields were well-matched to precipitation products. However, total accumulations tended to be over biased (1.10–2.10) and exhibited low-to-moderate threat scores (0.27–0.59). Non-frozen hydrometeor species from single-moment BMPS produced similar mixing ratio profiles and maximum saturation levels due to a common parameterization heritage. Greater variability occurred with frozen microphysical species due to varying assumptions among BMPSs regarding ice supersaturation amounts, the dry collection of snow by graupel, various ice collection efficiencies, snow and graupel density and size mappings/intercept parameters, and hydrometeor terminal velocities. The addition of double-moment rain and cloud water resulted in minimal change to species spatial extent or maximum saturation level, however rain mixing ratios tended higher. Although hydrometeor differences varied by up to an order of magnitude among the BMPSs, similarly large variability was not upscaled to mesoscale and synoptic scales.~~

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 the simple, warm rain only Kessler scheme (Kessler, 1969) to the full, double-moment, six-class Morrison scheme (Morrison et al., 2009). Microphysics parameterizations (along with cumulus parameterizations) drive cloud and precipitation processes and have far reaching consequences within numerical weather simulations (radiation, moisture, aerosols, *etc.*). Given its importance for simulations, Tao et al. (2011) noted at least 36 major, published, microphysics-focused studies primarily in the context of idealized simulations, hurricanes, and mid-latitude convection. More recently, the observational studies of Stark (2012) and Ganetis and Colle (2015) investigated microphysical species variability within East Coast U.S. winter storms (locally called “nor’easters”) and have underscored the need to investigate how microphysical parameterizations alter simulations of these powerful cyclones, which is the objective of the present work.

A “nor’easter” is a large (~2000 km), mid-latitude cyclone occurring between October and 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). To illustrate their potential severity, ten strong December nor’easter events between 1980 and 2011 resulted in 29.3 billion U.S. dollars in associated damages (Smith and Katz, 2013). Such damages are possible given the high economic output (16 billion U.S. dollars per day) of the northeastern U.S. (Morath, 2016). Given their importance to prognostic weather and climate models, this study aims to evaluate how BMPSs within WRF impacts its simulations of nor’easter development, the storm environment, and precipitation.

Recent nor’easter studies are scarce in light of extensive research conducted on these cyclones, primarily during the 1980s, which addressed key drivers including frontogenesis and baroclinicity (Bosart, 1981; Forbes et al., 1987; Stauffer and Warner, 1987), anticyclones (Uccelini and Kocin, 1987), latent heat release (Uccelini et al., 1987), and moisture transport by the low-level jet (Uccellini and Kocin, 1987; Mailhot and Chouinard, 1989). Despite extensive observational analyses, there is much less work on nor’easter and winter storm simulations in general, particularly those related to 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 from single-moment schemes. However, single moment-based results vastly improved when snow-size distribution intercepts were derived from a diagnostic equation rather than set as 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, yet had notable differences in cloud top temperature and precipitation field errors. In this study, the WRF single-moment BMPS (Hong and Lim, 2006) produced marginally superior simulations of cloud and precipitation processes as compared to other schemes. For warmer, tropical cyclones, Tao et al. (2011) investigated how four, six-class BMPSs impacted WRF simulations of Hurricane Katrina and demonstrated that BMPS choice had a minimal impact upon storm track. However, variations in sea-level pressure (SLP) were considerably higher (up to 50 hPa).

Shi et al. (2010) evaluated several WRF single-moment BMPSs for a lake-effect snow and a 20-22 January 2007 synoptic event. Simulated radar reflectively and cloud top temperature validation revealed WRF accurately simulated event onset and termination times, cloud coverage, and lake-effect snow band extent. However, simulated station snowfall rates were less accurate due to error in predicting exact points within a mesoscale grid. WRF-simulated snow bands showed minimal BMPS-based differences because cold temperatures and weak vertical velocities prevented graupel generation in all simulations. A more recent lake-effect snow modeling study by Reeves and Dawson (2013) investigated WRF sensitivity to eight different BMPSs during a December 2009 event. Their study found precipitation rate and its coverage were highly sensitive to BMPS because in half of their simulations vertical velocities exceeded hydrometeor terminal fall speeds which prolonged hydrometeor residence times. Terminal fall speeds differences existed due to varying assumptions associated with frozen hydrometeor species (i.e., snow density values, temperature-dependent snow intercept values, and graupel generation terms).

In a similar spirit to previous studies, this work will test WRF nor'easter simulation sensitivity to six- and seven-class BMPSs and focus on storm and microphysical properties, precipitation, and the simulated storm environment. 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 fully-compressible, non-hydrostatic, Eulerian equations in terrain-following coordinates (Skamarock et al., 2008). There was a four-domain, convection-resolving WRF grid (Fig. 1) with two-way feedback. It had 45-, 15-, 5-, and 1.667-km grid spacing, 61 vertical levels, and a 50-hPa (~20 km) model top. Boundary conditions were derived from 1° × 1° resolution Global Forecasting System model operational analysis (GMA) data. Except for a fourth domain, this model configuration and the following parameterizations were successfully applied in a previous nor'easter study (Nicholls and Decker, 2015) and was 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)
- 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) which vary in both severity and time of year (Table 1). Nor'easter events in Table 1 list one case for each month in which nor'easters occur (October–March) to determine any seasonal dependence or biases, and they 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 is based 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.

Simulations are integrated for 180 hours, starting 72 hours prior to the first precipitation impacts in the highly populated Mid-Atlantic region. This lead time allows for sufficient model spin-up time, establishment of the coastal baroclinic zone, and surface latent heat flux generation which are crucial components for nor'easter development (Bosart, 1981; Uccelini 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). A smaller threshold is not used to avoid capturing isolated showers well ahead of the primary precipitation shield. A New Jersey-centric approach was chosen due to its high population density (461.6 / km^2), significant contribution ($473 billion) to the U.S. gross domestic product, and its central location in the Mid-Atlantic (United States Census Bureau, unpublished data, 2012).

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, all five BMPSs were each run for the nor'easter events listed in Table 1.

2.2 Verification and analysis techniques

Model output was evaluated against both GMA and 4-km resolution Stage IV precipitation data (Y. Lin and K.E. Mitchell, preprints, 2005). GMA data validated all model output (except precipitation) due to its extensive coverage, and lack of available in-situ data in data-sparse regions. Stage IV is a six-hourly, gridded precipitation product derived from rain gauge and radar data with 4-km spatial resolution. Prior to any validation, all data were interpolated to the coarsest grid spacing.

Model output analysis consisted of several parts. Nor'easter storm tracks were derived via an objective, self-coded algorithm similar to that used at the Climate Prediction Center (Serreze, 1995; Serreze et al., 1997). At each storm position, minimum SLP (MSLP), maximum wind speed, and track error were stored and compared to model analysis. Precipitation values and their distribution were evaluated against Stage IV data and validated using bias and threat score (critical success index) calculations (Wilks, 2011). The simulated hydrometeor species analysis was comprised of two parts: precipitable mixing ratios, and composite mixing ratio profiles. Precipitable mixing ratio is derived from the equation for precipitable water and is defined as the following:

$$PMR = \frac{1}{\rho g}\int_{P_{top}}^{P_{sfc}} w\, dp \qquad (1)$$

In Eq. (1), PMR is the precipitable mixing ratio in m, ρ 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}); dp is the change in atmospheric pressure between model levels (Pa). Composite mixing ratio profiles were calculated within a 600-km wide cubic volume centered at both model- and GMA-relative surface cyclone locations (hereafter, model-relative and GMA-relative storm environments, respectively). For illustrative purposes, the red, dashed box in Figure 2, panel 1 denotes the GMA-relative storm environment extent at 12 UTC 15 October 2009. Finally, the accuracy of model- and GMA-relative storm environment WRF simulations will be validated using the non-hydrostatic, moist, total energy norm (Kim and Jung, 2009). Energy norm integrations were capped at ~100 hPa to limit large temperature errors near the model top and calculated using Eq. (2).

$$E_m = \iiint_{\sigma,x,y} \frac{1}{2}\left[u'^2 + v'^2 + w'^2 + \left(\frac{g}{N_r\theta_r}\right)^2 \theta'^2 + \left(\frac{1}{\rho_r c_s}\right)^2 p'^2 + \omega_q \frac{L^2}{c_p T_r} q'^2\right] dy\, dx\, d\sigma \qquad (2)$$

In Eq. (2), E_m is the moist total energy norm (J m^2 kg^{-1}); u', v', and w' are the zonal, meridional, and vertical wind perturbations (m s^{-1}), respectively; p' is the pressure perturbation (Pa); θ' is the potential temperature perturbation (K); q' is the mixing ratio perturbation (kg kg^{-1}). N_r, θ_r, ρ_r, T_r, and c_s are the reference Brunt Väisälä frequency (0.0124 s^{-1}), reference potential temperature (270 K), reference air density (1.27 kg m^{-3}), reference air temperature (270 K), and speed of sound (329.31 m s^{-1}), respectively. Finally, c_p is the specific heat at constant pressure (1005 J kg^{-1} K^{-1}) and ω_q is a scaling factor (0.1). Finally, y, x, and σ, denote the zonal, meridional, and sigma (terrain following) directional components, respectively. Our analysis focus on the energy norm was influenced by Buizza et al. (2005), who made a compelling case for its usage at ECMWF for model validation given its total model volume integration, lack of single-layer sensitivity, and inclusion of temperature, wind, pressure, and moisture errors. Similar to root mean square error, smaller values denote less error.

Commented [NSD(RAU(1]: Perrturbations are relative to GMA. Section not in new manuscript, no change to text is made.

Commented [NSD(RAU(2]: Refers to vertical velocity and not mixing ratio.

3. Results

3.1 Nor'easter track and property analysis

Figure 2 displays storm tracks from W361 BMPS simulations (colors) and GMA (black), and Fig. 3 shows GMA-relative track errors for all seven cases. In Fig. 3, smaller, colored symbols denote six-hourly track error, whereas the larger, black symbols denote the model mean. Similar to Wu and Petty (2010) and Tao et al. (2011), BMPS choice yields modest storm track changes (Δ BMPS average; 84 km) and no apparent directional biases among the schemes. As compared to GMA, six-hourly storm track errors vary greatly ranging from 30 km (GCE6, Case 6) to 1,594 km (GCE7, Case 2). Nor'easters with less track error (Case 3, 4, and 6) formed within a regions of stronger differential cyclonic vorticity advection (CVA) aloft, whereas for higher track error cases (Cases 2 and 7) CVA was far weaker (not shown). To quantify case-to-case track errors, Table 3 lists average track errors for each case, using bold type for large errors (> 400 km). Both Table 3 and Fig. 3 indicate that the GCE6-based simulations have the least average track error in four out of seven cases (Cases 1, 3, 4, and 6) and overall (406 km). However, this conclusion is not definitive, given a 187 km maximum track error spread (Case 1, WSM6-GCE6) among BMPSs.

In addition to average track errors, Table 3 also contains other key nor'easter properties including MSLP, maximum MSLP deepening rate, and maximum wind speed within the model-relative storm environment. To supplement Table 3, Fig. 4 displays six-hourly MSLP and maximum 10 m wind speeds from all W361 runs and GMA for Cases 2, 3, 4, and 5. These cases have the least and greatest average track errors (See Table 3). In Table 3, large deviations from GMA are in bold type (ΔMSLP > 5 hPa, Δ deepening rate > 5 hPa / 6 hours, and Δ10 m winds > 5 m s^{-1}). Consistent with the storm track analysis, Case 2 has notable deviations in both MSLP (up to 8.6 hPa) and 10 m winds (up to 7.1 m s^{-1}). Large track errors however are not required for MSLP and wind speed errors to be large. The highest MSLP errors originate from Cases 3 (10.5 hPa; Lin6) and 4 (9.3 hPa; Lin6) and are statistically significant in the former (maximum p-value 0.032, GCE6). Although sizable, these MSLP differences fall well short of the 50-hPa MSLP differences cited in Tao et al. (2011) possibly due to the less extreme MSLP values associated with nor'easters as compared to hurricanes. Consistency between BMPSs simulations in Fig. 1, Fig. 4, and Table 3 suggests that nor'easter MSLP and wind errors are more associated with differences in steering flow and cyclonic vorticity advection aloft rather than BMPS selection. Case 3 best illustrates this hypothesis as MSLP lags notably behind GMA starting when all simulations diverged from GMA on December 19 (See Figs. 1 and 4), yet once the secondary low developed further north along the Gulf Stream, latent heat fluxes increase greatly (> 1000 W m^{-2}) and the MSLP gap in Fig. 4 closes considerably. A similar situation occurs in Case 2, where 10 m maximum winds became far stronger (> 10 m s^{-1}) in GMA than in W361 simulations. Stronger winds exist in GMA than W361 simulations because its cyclone remains over the strong baroclinic zone associated with the Gulf Stream, rather than the more energy-poor inland track exhibited by all W361 simulations track (See Fig. 2, panel 2).

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. ~~One of the most crippling potential impacts associated with nor'easters comes from precipitation, which is partially driven in simulations by BMPSs. To demonstrate any potential BMPS sensitivity, Fig. 5 displays 72-hour precipitation accumulations (forecast hours 48–120) from Stage IV and Lin6 (top panels), differences between the remaining BMPSs and Lin6 (middle panels), and finally precipitation probability density and cumulative distribution functions (PDF and CDF, respectively) from Cases 4 and 6.~~ These two cases have the lowest track errors in Table 3 which facilitated easier comparisons to Stage IV precipitation data. Table 4 contains bias and threat scores values from all seven cases assuming a 12.5 mm to quantify simulated precipitation field accuracy and tendency.

Threat score and bias values in Table 4 indicate Cases 2 and 3 to be clear outliers given bias scores exceeding 4 and less than 1, respectively. These outlier values result from the spatial limitations of the Stage IV product due to its reliance upon radar and rain gauge data. In Cases 2 and 3, either the GMA or W361 simulated cyclone crossed the data cut-off region prematurely resulting in a severe over-bias (4.50–4.72) and an under-bias (0.71–0.85), respectively. For the remaining five nor'easter cases, Table 4 indicates low (0.29, GCE7, Case 7) to moderate (0.59, WDM6, Case

Commented [NSD(RAU(3]: Corrected to mention to case number. Seen here is a excerpt from the revised manuscript relevant to your request about knowing to which model domain Table 4 (now Table 3) refers.

6) threat scores and over-biased precipitation totals (bias range: 1.10–2.10). Although case-to-case threat score and bias vary up to 0.27 and 0.98, inter-BMPS threat scores and biases (except Case 4) are an order of magnitude smaller. 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, despite being the simplest BPMS tested, Lin6 did manage marginally better threat scores in three of the five nor'easter events and has the lowest overall average bias.

As Fig. 5 illustrates, Case 4 W361 simulations produce a precipitation extent similar to Stage IV (except off Georgia), yet exact precipitation totals along the coast are too high. Case 6 exhibits similar behavior and has well-matched extent, but excessive precipitation totals. Precipitation PDF and CDFs show three distinctive bin categories: 5–10 mm, 10–55 mm, and 55 mm+. The strong-convection modeling studies of 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. ~~Given WRF's common heritage with GFS, similar precipitation biases would be expected. However, two nor'easter cases (Cases 6 and 7) deviate from this expectation and generated too little light precipitation (5–10 mm) and too much heavier precipitation (10–55 mm).~~ Once above 55 mm, all cases produce too much precipitation. These findings likely stem from two sources: different Stage IV domain exit times and the focus in previous studies on convective rather than stratiform events, which may lead to differences in simulated precipitation generation. Marginal changes in QPF (< 15 mm) and threat scores between the BMPS W361 runs are consistent with Fritsch and Carbone (2004) and Wang and Clark (2010) who evaluated the accuracy of simulated precipitation in warm-season events and quasi-stationary fronts, respectively.

Commented [NSD(RAU(4]: Corrected to Fig. 5

Commented [NSD(RAU(5]: Line removed from revised manuscript.

3.3 Hydrometeor species analysis

Figure 6 displays precipitable mixing ratios for six microphysics species (water vapor, cloud water, graupel, cloud ice, rain, and snow) at 18 UTC 26 January 2015 over the entirety of Domain 3. This time is selected for its exceptionally small track error (< 50 km) and because all simulated cyclones are located within the 5-km Domain 3 and 1.667-km Domain 4. Figure 6 depicts precipitable mixing ratios rather than column-integrated mixing ratios as it is easier to express these data as a height (mm) than as a weight (kg m^{-2}). Hail is excluded as it is specific to GCE7 and is an order of magnitude less (on average) than the other hydrometeor species. Figure 6 shows most precipitable mixing ratio species (especially cloud ice and snow) vary considerably among BMPSs though there are identifiable trends due to the underlying assumptions made within the BMPS as explained in more detail below. Figure 7 shows Case 4, domain 3, composite hydrometeor mixing ratio values averaged from the model-relative storm environments of each W361 BMPS simulation. The first five panels exclude water vapor (two orders of magnitude larger), but do include composite vertical velocity as a black, solid line. Composite water vapor mixing ratios are shown for all W361 simulations in the last panel of Fig. 7. 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).

All 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 (Fig. 7), and precipitable rain fields (Fig. 6) and is consistent with Wu and Petty (2010). WDM6 varies from the other single-moment BMPSs because CCN, rain and ~~cloud water~~ cloud droplet number concentration are forecasted rather than diagnosed from derivative equations (Hong et al., 2010). While such changes have minimal impact upon ~~maximum saturation heights~~ or the precipitable rain coverage area, maximum rain mixing ratio values are noticeably higher aloft and decrease sharply towards the surface.

Commented [NSD(RAU(6]: Maximum saturation height is maximum mining ratio height. This term has been corrected throughout.

Similar to rain mixing ratios, cloud water mixing ratios exhibit little variability in either the precipitable cloud water extent (Fig. 6) or the maximum saturation level (Fig. 7), but maximum mixing ratio values vary even between single-moment schemes. 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, which allows the number concentrations to vary, in this instance, the maximum mixing ratios are only decreased slightly relative to WSM6. Small variations in cloud water between WSM6 and WDM6 suggest cloud water number concentrations in WDM6 are potentially close to the assumed 300 cm^{-3} number concentration in WSM6 (Hong et al. 2010) and/or the larger-scale environment/forcing is a dominant factor as water supersaturation is negligible.

Commented [NSD(RAU(7]: Added lines showing 0C and -40C temperature heights in new Figs. 4 and 5.

Among the BMPSs, Figs. 6 and 7 show that precipitable snow and snow mixing ratios vary considerably with Lin6 having the smallest and GCE6 the largest amounts. Dudhia et al. (2008) and Tao et al. (2011) associate the dearth of snow in Lin6 to its high rates of dry collection by graupel, low snow size distribution intercept (decreased surface area), and auto-conversion of snow to either graupel or hail at high mixing ratios. In GCE6, dry collection of snow and ice by graupel is turned off and results in a large increase in snow at the expense of graupel (Lang et al. 2007). Although the snow riming efficiency was reduced, the omission of dry collection along with and the continued assumption of water saturation for the vapor growth of cloud ice to snow contributes to its high snow-mixing ratios (Reeves and Dawson, 2013; Lang et al. 2014). In GCE7, this latter issue has been addressed and along with numerous other changes, including the introduction and of a snow size and density mapping, snow breakup interactions, and a new vertical-velocity-dependent ice super saturation assumption (Lang el al., 2007; Lang et al., 2011; Lang et al., 2014; Chern et al., 2016; Tao et al., 2016). Figures 6 and 7 show that although the combination of an RH correction factor (Lang et al., 2011) in conjunction with the new ice super saturation adjustment (Tao et al., 2016) reduce the efficiency of vapor growth of cloud ice to snow, the new snow mapping and enhanced cloud ice to snow auto-conversion in GCE7 help to keep snow mixing ratios higher than in non-GCE BMPSs. Unlike Lin6, WSM6 and WDM6 graupel and snow fall speeds are assumed to be identical within a grid cell (Dudhia et al., 2008) and the ice nuclei concentration is a function of temperature (Hong et al., 2008). These two changes effectively eliminated the accretion of snow by graupel and increased snow mixing ratios at colder temperatures (Dudhia et al., 2008; Hong et al., 2008). Figure 7 shows that the level of maximum snow content is largely conserved across the BMPSs, except for Lin6, which is 100 hPa lower as differential snow and graupel fall speeds allow graupel to collect snow.

Maximum mean graupel mixing ratios in the column are generally much less than for snow except for Lin6 where dry collection aloft is dominated by graupel and is unrealistic (Stith et al., 2002). In contrast, GCE7 produces the most snow and the least amount of graupel. GCE7 includes a graupel size mapping, but the combination of the snow size mapping, which generally decreases snow sizes aloft (thus increasing their surface area and vapor growth), the addition of deposition conversion processes wherein graupel/hail particles experiencing deposition growth at colder temperatures are converted to snow, and changes to the cloud ice that lead to more cloud ice and less super-cooled cloud water (see below) and thus reduced riming, favor snow over graupel even more (Lang et al. 2014; Chern et al., 2016; Tao et al., 2016). Consistent with Reeves and Dawson (2013), graupel mixing ratios are around 30-50 % that of snow for WSM6 and WDM6. Despite having a smaller peak mean graupel mixing ratio in the column (Fig. 7), WDM6 produces locally enhanced precipitable graupel values in Fig. 6 relative to WSM6.

Although up to ninety percent smaller in magnitude than snow (GCE6), cloud ice mixing ratios vary greatly amongst the BMPSs in Figs. 6 and 7. They are highest in GCE7 and lowest in Lin6. Wu and Petty (2010) similarly found low cloud ice mixing ratios from their Lin6 simulations and ascribed it to dry collection by graupel, ~~lack of an ice sedimentation term~~, and fixed cloud-ice size distribution. Similar to Lin6, in GCE6 the cloud-ice size distribution is monodispersed, but as noted in Lang et al. (2011) and Tao et al. (2016), 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 ~~the vapor growth of cloud ice to snow in GCE6 was still based upon an assumed water saturation, which made this term too efficient and helped keep cloud ice mixing ratios lower~~. This term includes an RH correction factor in GCE7, which depends upon the amount of ice supersaturation, which in turn is dependent on the vertical velocity in GCE7. These factors effectively blunt this term's over-efficiency. Additionally, in GCE7, contact and immersion freezing terms are included (Lang et al., 2011), cloud ice collection by snow efficiency is a function of snow size (Lang et al., 2011; Lang et al., 2014), there is a maximum limit on cloud ice particle size (Tao et al., 2016), the ice nuclei concentration follows the Cooper curve (Cooper, 1986; Tao et al., 2016), and cloud ice can persist even in ice subsaturated conditions (i.e., when RH values for ice are greater than or equal to 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), these changes combine to produce almost 100 % more cloud ice in GCE7 than in GCE6 (See Fig. 7). 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 also include ice sedimentation terms (Hong et al., 2008). Despite the differences in the cloud ice mixing ratio amounts, the level of maximum mean cloud ice mixing ratio is around 300 hPa for all of the BMPSs.

Neither precipitable mixing ratio nor vertical velocity exhibit notable sensitivity to the BMPSs despite the above hydrometeor results. Close inspection of Fig. 7 reveals that GMA water vapor mixing ratios are slightly higher below 800 hPa on average than those from the W361 BMPS simulations and slightly lower above that level, while Fig. 6 hints at a potential small dry bias in WRF. Although one order of magnitude or more smaller than water vapor mixing ratios, slight differences in the other hydrometeor species (notably cloud ice and snow) act to drain the available

Commented [NSD(RAU(8]: Reviewer is correct. Not having a sedimentation term would increase cloud ice concentrations.

moisture (GCE7 versus Lin6) at slightly different rates. In contrast to Reeves and Dawson (2013), model-relative vertical velocities in nor'easters extend through the depth of the troposphere, whereas for lake-effect snow, positive vertical velocities may only extend to 700 hPa. Enhanced vertical velocities above 770 hPa are driven primarily by isentropic lift associated with the warm-conveyor belt (Kocin and Uccelini, 2004)

3.4 Energy norm-based analysis of model- and GMA-relative storm environments

~~Figure 8 displays the model-relative storm environment fully-integrated Lin6 energy norm with time (black) and the percent difference between the Lin6 energy norm and all other BMPSs for all seven cases. Lin6 energy norm values provide a fixed reference to inter-compare WRF simulation accuracy because both a WRF and GMA data are used to calculate energy norm values~~. Figure 9 shows the similar information to Fig. 8, except the energy norm is integrated at each model level and averaged in time. To complement these two figures, Fig. 10 depicts the model-relative time-averaged total energy norm (black) and its six component parts integrated for each level for cases 1, 2, 4, and 7 from Lin6, GCE7, and WDM6. Table 5 summarizes the energy norm results for both the model- and GMA-relative storm environments. Given the similar appearance between the GMA- and model-relative storm environment plots (similar shape, slightly different magnitude), we elected to only show model-relative energy norm plots in this section.

Commented [NSD(RAU(9]: Section no longer relevant. Model- relative = evaluation at cyclone location as forecasted by each model simulation. GMA-relative = evaluation for all models as at GMA-indicated cyclone location.

Closer observation of Figs. 3, 8, and 9 reveal energy norm variability has strong links to both storm track uncertainty (e.g., Fig. 8, Case 7, GCE6) and the energy norm magnitude (e.g., Fig. 9, Case 1, GCE7), yet track errors need not be large to have higher energy norms (i.e., Case 3). Energy norm differences in Fig. 8 vary from 95 % (Case 3, GCE7) to -39 % (Case 4, WDM6) where positive percentage values denote higher energy norms than Lin6. Similarly, time-averaged energy norms in Fig. 9 show a slightly smaller range between -24 % (Case 1, WDM6) and 79 % (Case 1, GCE7)). Overall, Figs. 8 and 9 show that no one BMPS scheme consistently outperforms the other four schemes, a result quantified in Table 5. In Table 5, the Lin6 scheme has the highest tendency for the lowest energy norm values, but its energy norms are lowest only in 18 out of 62 times (29 %) and 24 out of 67 times (35.8 %) and for 3 out of 7 cases in the model- and GMA-relative storm environments, respectively. There was no statistically significant differences between Lin6 and other BMPSs in two-tailed T-Tests (min p-value: 0.206 (GCE7, Case 1)) with the exception of the GCE schemes from Case 7. For this case and these BMPSs, statistical significance is only achieved due to highly variable storm track errors at the last three analysis times when differential CVA aloft was fairly weak. Complicating the energy norm results, WDM6 has the least average error in the GMA-relative storm environment which only makes drawing a decisive conclusion more difficult.

Although we could not detect a clearly preferable BMPS for WRF nor'easter simulations, the Figs. 9 and 10 can help diagnose key sources of error. For Cases 1, 2, 4, and 7 (also true for the remaining 3 cases), model-relative storm environment total energy norms are highest near the surface and decrease until the tropopause. Figure 10 shows the total energy norm to be dominated by its temperature and horizontal wind components. By comparing the magnitude of these errors between BMPSs, it is possible to diagnose that GCE7 has a less accurate depiction of the low-level jet given its higher horizontal wind energy norm values at 858 hPa than as represented by Lin6. Alternatively higher meridional wind errors at and above 500 hPa for GCE7, Case 7, indicate errors in the speed or location of the warm-conveyor belt.

Commented [NSD(RAU(10]: This claim is 'sketchy" without further evidence, even if it is a true statement.

Formatted: Indent: First line: 0"

4 Conclusions

The role and impact of five BMPSs upon seven, W361 nor'easter simulations is investigated and validated against GMA and the Stage IV precipitation product. Tested BMPSs include four single-moment (Lin6, GCE6, GCE7, and WSM6) and one double-moment BMPSs (WDM6). Consistent with previous studies, storm track, MSLP, and maximum 10 m winds exhibits only a minor dependence upon BMPS with up to 187 km, 7.0 hPa, and 7.6 m s^{-1} of error variability between BMPSs, respectively. Relative to GMA, model track errors average 406 km and MSLP and maximum 10 m winds vary up to 10.5 hPa, and 11.2 m s^{-1} and are only statistically significant when storm track errors involve the Gulf Stream (e.g., Case 3).

Simulated precipitation fields exhibit low-to-moderate (0.27–0.59) threat score skill and varying degrees of over-bias (1.10–2.10) when compared to the Stage IV precipitation product. Although most cases generate too much light precipitation and too little heavy precipitation (up to 55 mm) as in previous studies (Ridout et al., 2005; Dravitzki and McGregor, 2011), two cases (6 and 7) reverse this trend. At notably high precipitation accumulation (55 mm+) all BMPSs generate excessive precipitation (relative to Stage IV). These digressions from previous studies are potentially related to the general lack of strong convection in nor'easters, whereas in previous studies their focii lie on strong-convective events (e.g., hurricanes and squall lines), but validating this claim would require investigation beyond the scope of the present work.

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. Despite the increased complexity of WDM6, it did not produce vastly different results from the single-moment BMPSs. The Lin6 hydrometeor species vary the most relative to other schemes, especially graupel and snow, due to its low snow size intercept and its snow-to-graupel conversion rates. Validations of hydrometeor species (except water vapor) were not performed due to lack of either sufficient radar coverage off the U.S. East Coast or a high-quality, satellite-based hydrometeor product covering all major species (excluding hail).

Model and GMA-relative storm environment energy norms indicate that with the exception of Case 7 (due to track error at three times), combined temperature, wind, pressure, and moisture errors failed to yield statistically significant differences (min p-value: 0.206) attributable to BMPS option. These differences, although not statistically significant do show the Lin6 simulations produce the lowest energy norm in 29 % and 35.8 % of all evaluated model-and GMA-evaluated storm track positions. Energy norms from the remaining BMPSs did not frequently stray more than 20 % from the Lin6 scheme and demonstrated that the greatest contributions to the energy norm were horizontal winds and temperature in the lower troposphere (especially between 850 and 500 hPa). Energy norm results also show that although hydrometeor species mixing ratios varied up to an order of magnitude (snow, Lin6 vs all others), these large changes were not upscaled to mesoscale and synoptic scales.

Although none of these results proved definitive, they do strike a cautionary note where higher computational costs associated with double-moment or even sophisticated single-moment BMPSs do not guarantee better results. Furthermore, microphysics-focused studies tend to focus on strong convective events (i.e., squall lines, hurricanes, etc.), yet provide little attention to strongly precipitating, stratiform-dominated events (such as nor'easters). Although not conclusive, this study has shown that assumed precipitation tendencies may vary in light of the dominant precipitation mode. 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 2011 Graduate Student Summer Program (GSSP), 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

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.

Buizza, R., Houtekamer, P. L., Pellerin, G., Toth, Z., Zhu, Y., and Wei, M.: A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems, Mon. Wea. Rev., 133, 1076–1097, 2005

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.

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.

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.

Ebert, 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.

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.

Kain, J. S. (2004), The Kain–Fritsch Convective Parameterization: An Update, J. Appl. Meteor., 43, 170–181, 2004.

Kessler, E. (1969), On the distribution and continuity of water substance in atmospheric circulation, Meteor. Monogr., 32, Amer. Meteor. Soc., 84 pp, 1969.

Kim, H. M., and Jung, B. -J.: Influence of moist physics and norms on singular vectors for a tropical cyclone. Mon. Wea. Rev., 137, 525–543, 2009.

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*.

Lang, S. E., Tao, W. -K., Chern, J. -D., Wu, D., and Li, X.: Benefits of a fourth ice class in the simulated radar reflectivities of convective systems using a bulk microphysics scheme, J. Atmos. Sci., 71, 3583–3612, 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.

Li, J. -L. F., Waliser, D., Woods, C., Teixeira, J., Bacmeister, J., Chern, J. -D.,, Shen, B. -W., Tompkins, A., Tao, W. -K.,, and Kohler, M.: Comparisons of satellites liquid water estimates to ECMWF and GMAO analyses, 20th century IPCC AR4 climate simulations, and GCM simulations, Geophys. Res. Lett., 35, L19710, doi:10.1029/2008GL035427, 2008.

Lim, K.-S. and Hong, S. -Y.: Development of an effective double-moment cloud microphysics scheme with prognostic cloud condensation nuclei (CCN) for weather and climate models, Mon. Wea. Rev., 138, 1587–1612, 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 meso-scale model, Mon Wea. Rev., 117, 1311–1343, 1989.

Marzban C., and Sandgathe, S.: Cluster analysis for verification of precipitation fields, Wea. Forecasting, 21, 824–838, 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.

Morath, E. (2016), Will a blizzard freeze U.S. economic growth for the third straight year, Wall Street Journal, 20 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.

Nicholls, S. D. and Decker, S. G.: Impact of coupling an ocean model to WRF nor'easter simulations, Mon. Wea. 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.

Rutledge, S. A., and Hobbs, P. V.: The mesoscale and microscale structure and organization of clouds and precipitation in midlatitude cyclones. XII: A diagnostic modeling study of precipitation development in narrow cloud-frontal rainbands. J. Atmos. Sci., 20, 2949–2972, 1984.

Serreze, M. C.: Climatological aspects of cyclone development and decay in the Arctic, Atmos.-Ocean, 33, 1–23, 1995.

Serreze, M. C., Carse, F., Barry, R. G., and Rogers, J. C.: Icelandic low cyclone activity: Climatological features, linkages with the NAO and relationships with recent changes in the Northern Hemisphere circulation, J. Climate, 10, 453–464, 1997.

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.

Stauffer, D. R., and Warner, T. T.: A numerical study of Appalachian cold-air damming and coastal frontogenesis, Mon. Wea. Rev., 115, 799–821, 1987.

Stephens, G. L., *et al.:* CloudSat mission: Performance and early science after the first year of operation, J. Geophys. Res., 113, *D00A18, doi:10.1029/2008JD009982, 2008.*

Stith, J. L., Dye, J. E., Bansemer, A., Heymsfield, A. J., Grainger, C. A., Petersen, W. A, and Clfelli, R.: 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 NU-WRF 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.*

Uccellini, L. W. and Kocin, P. J.: The Interaction of jet streak circulations during heavy snow events along the east 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.

Wilks, D. S.: Statistical methods in the atmospheric sciences, third edition, Academic Press, Oxford, in press., 2011.

Wu, L., and Petty, G. W.: Intercomparison of bulk microphysics schemes in model simulations of polar lows, Mon. 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.

Table 1. Nor'easter case list. The NESIS number is included for storm severity reference. The last two columns denote the first and last times for each model run. Tracks are plotted in Fig. 2.

Case Number	NESIS	Event Dates	Model Run Start Date	Model Run End Date
1	N/A	15–16 Oct 2009	10/12 12UTC	10/20 00UTC
2	N/A	07–09 Nov 2012	11/04 06UTC	11/11 18UTC
3	4.03	19–20 Dec 2009	12/16 06UTC	12/23 18UTC
4	2.62	26–28 Jan 2015	01/23 00UTC	01/30 12 UTC
5	4.38	04–07 Feb 2010	02/02 18UTC	02/10 06UTC
6	1.65	01–02 Mar 2009	02/26 12UTC	03/06 00UTC
7	N/A	12–14 Mar 2010	03/09 06UTC	03/16 18UTC

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.

Microphysics Scheme	QV	QC	QH	QI	QG	QR	QS	Moments	Citation
Lin6	X	X		X	X	X	X	1	Lin et al. (1983); Rutledge and Hobbs (1984)
GCE6	X	X		X	X	X	X	1	Tao et al. (1989); Lang et al. (2007)
GCE7	X	X	X	X	X	X	X	1	Lang et al. (2014)
WSM6	X	X		X	X	X	X	1	Hong and Lim (2006)
WDM6	X	X		X	X	X	X	2 (QC, QR)	Lim and Hong (2010)

Table 3. Various simulated nor'easter characteristics. Bolded values indicate sea-level pressure values or rate errors > 5 hPa (/6 hours), wind errors > 5 m s^{-1}, and average track errors > 400 km.

GMA	1	2	3	4	5	6	7
Min SLP (hPa)	991.5	989.5	972.6	980.5	979.7	1000.5	993.5
Max SLP decrease (hPa/6hrs)	-6.0	-5.9	-6.4	-10.8	-7.9	-3.2	-2.7
Max 10 m Wind (m s^{-1})	24.4	24.8	23.4	22.9	23.1	16.4	15.2

Lin6	1	2	3	4	5	6	7
Min SLP (hPa)	995.2	**982.8**	**983.1**	**989.8**	978.2	1001.9	998.1
Max SLP decrease (hPa/6hrs)	-4.1	-6.0	-6.9	**-5.5**	-6.4	-3.3	-2.7
Max 10 m Wind (m s^{-1})	24.1	20.0	**30.6**	26.2	23.3	14.2	**26.4**
Avg Track Error (km)	**505**	**767**	356	131	**490**	219	**404**

GCE6	1	2	3	4	5	6	7
Min SLP (hPa)	990.0	**982.2**	976.7	**988.0**	981.7	1002.2	996.4
Max SLP decrease (hPa/6hrs)	-8.5	-6.7	-9.0	-6.0	-6.2	-3.5	-3.9
Max 10 m Wind (m s^{-1})	28.7	**18.1**	**33.0**	22.1	23.5	15.5	**23.5**
Avg Track Error (km)	366	**789**	311	140	**465**	197	**576**

GCE7	1	2	3	4	5	6	7
Min SLP (hPa)	989.0	**983.1**	976.9	**987.3**	976.2	1002.1	996.3
Max SLP decrease (hPa/6hrs)	-4.3	-7.2	-9.8	-6.0	-6.4	-3.2	-3.7
Max 10 m Wind (m s^{-1})	24.3	**19.1**	**30.2**	20.6	23.0	16.1	**24.6**
Avg Track Error (km)	**445**	**792**	317	129	**479**	225	**541**

WSM6	1	2	3	4	5	6	7
Min SLP (hPa)	996.0	**982.3**	**978.6**	**989.3**	976.2	1002.5	996.3
Max SLP decrease (hPa/6hrs)	-3.9	-5.9	-8.9	**-5.3**	-5.2	-3.2	-6.1

Max 10 m Wind (m s^{-1})	22.1	**17.7**	25.6	24.4	21.5	21.1	**21.5**
Avg Track Error (km)	**553**	**789**	327	140	**518**	233	**544**
WDM6	1	2	3	4	5	6	7
Min SLP (hPa)	992.7	**980.9**	977.1	**988.6**	978.5	1001.4	995.0
Max SLP decrease (hPa/6hrs)	-4.9	-6.5	-8.7	**-5.5**	-8.7	-2.7	-5.8
Max 10 m Wind (m s^{-1})	23.1	**19.6**	**33.2**	20.4	23.2	15.9	**23.4**
Avg Track Error (km)	**543**	**804**	333	138	**567**	219	**452**

Table 4. Stage IV-relative, storm-total precipitation threat scores and biases assuming a threshold value of 12.5 mm (0.5”). Bolded value denote the model simulation with the threat score closest to 1 (perfect forecast) and bias values closest to 1 (no precipitation bias).

Threat Score	1	2	3	4	5	6	7	Mean
Lin6	0.40	0.16	**0.25**	**0.40**	**0.58**	0.57	**0.31**	**0.38**
GCE6	**0.41**	**0.17**	0.23	0.34	0.54	0.57	**0.31**	0.37
GCE7	0.40	**0.17**	0.23	0.35	0.56	0.56	0.29	0.37
WSM6	0.39	0.16	0.23	0.35	0.55	0.57	0.30	0.36
WDM6	0.39	0.16	0.23	0.36	**0.58**	**0.59**	**0.31**	0.37

Bias	1	2	3	4	5	6	7	Mean
Lin6	1.38	4.62	0.71	**1.79**	**1.34**	1.33	1.12	**1.76**
GCE6	**1.34**	4.52	0.81	2.10	1.45	1.33	1.12	1.81
GCE7	1.40	**4.50**	**0.85**	2.04	1.40	1.35	1.20	1.82
WSM6	1.45	4.72	0.81	2.07	1.44	1.33	1.14	1.85
WDM6	1.45	4.68	0.82	2.01	1.36	**1.30**	**1.10**	1.82

Table 5. Energy norm analysis for model- and GMA-relative cyclone locations. Energy norm values are derived from domain 2 data and only within a 600-km diameter box centered on the model-indicated cyclone location. "Per case rank order" ranks the models based upon number of instances of lowest model error for each of the seven cases and allows for ties.

Model-Relative Energy Norm Analysis

Total 62 Periods	Lin6	GCE6	GCE7	WSM6	WDM6
Lowest Energy Norm (% of total)	18 (29.0 %)	8 (12.9 %)	8 (12.9 %)	15 (24.2 %)	13 (21.0 %)
Avg ΔENorm vs. Lin6 (% of Lin Enorm)	N/A	3.23E+5 (5.73 %)	8.75E+4 (1.55 %)	1.85E+4 (0.33 %)	3.72E+5 (6.59 %)
2-Tailed P-Value (vs Lin6)	N/A	0.406	0.11	0.941	0.652
Per Case Rank Order (of 5)	2113312	4223334	2423154	1233223	4521211

GMA-Relative Energy Norm Analysis

Total: 67 Periods	Lin6	GCE6	GCE7	WSM6	WDM6
Lowest Energy Norm	24 (35.8 %)	5 (7.5 %)	6 (9.0 %)	17 (25.4 %)	15 (22.4 %)
Avg ΔENorm vs. Lin6 (% of Lin Enorm)	N/A	2.69E+5 (6.16 %)	2.61E+5 (5.97 %)	1.54E+4 (0.35 %)	-1.14E+5 (-2.58 %)
2-Tailed P-Value (vs Lin6)	N/A	0.414	0.24	0.882	0.589
Per Case Rank Order (of 5)	2221121	3454242	3414545	1141224	3233212

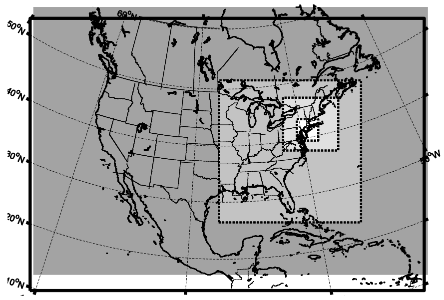


Figure 1. Nested WRF configuration used in simulations. Horizontal resolution for domains 1, 2, 3, and 4 are 45, 15, 5, and 1.667 km, respectively.

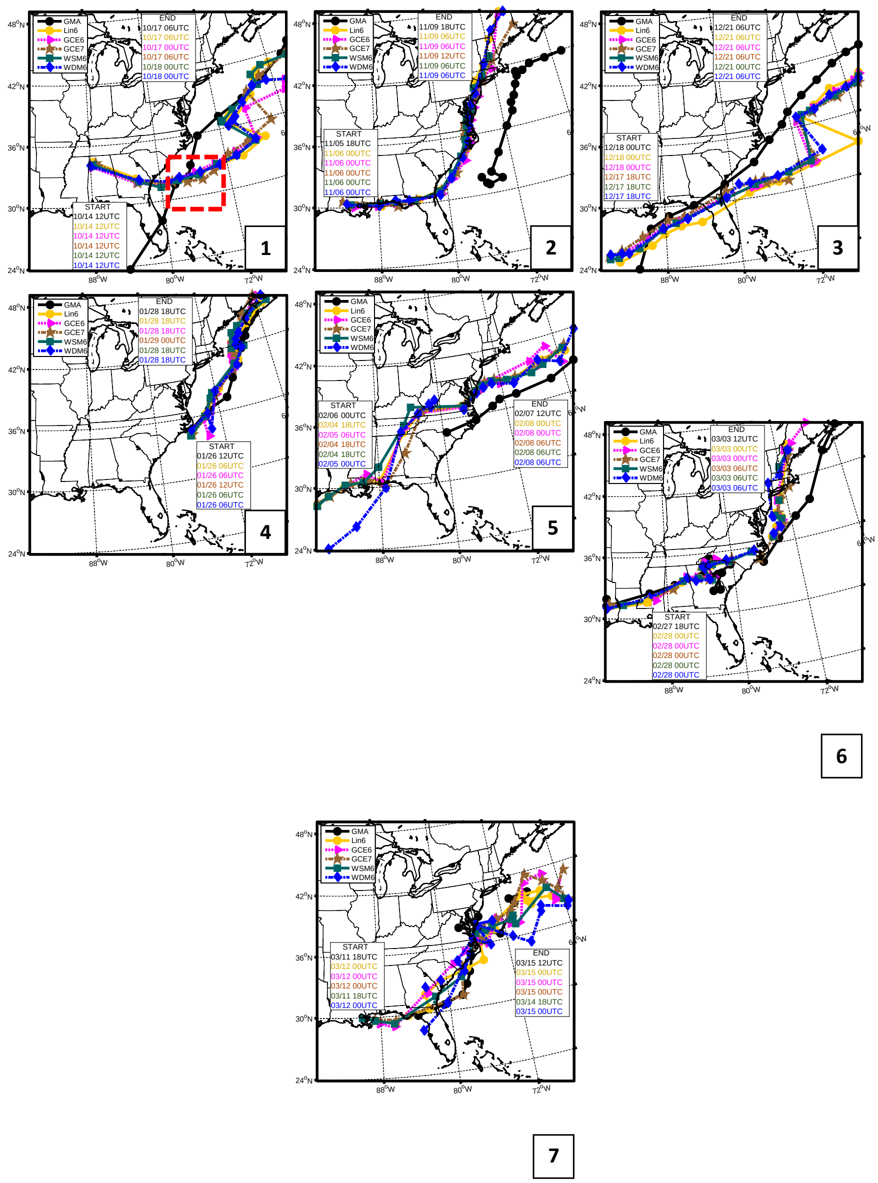

GMA
Lin6
GCE6
GCE7
WSM6
WDM6
START
END
1
2
3
4
5
6
7

Figure 2. Storm tracks from GMA and the model runs. Line legend is shown on the upper-left of each plot. Shown symbols indicate simulated storm position every six hours. Black numbers indicate case number. The red, dashed box in case 1, shows the size of a 600-km diameter box.

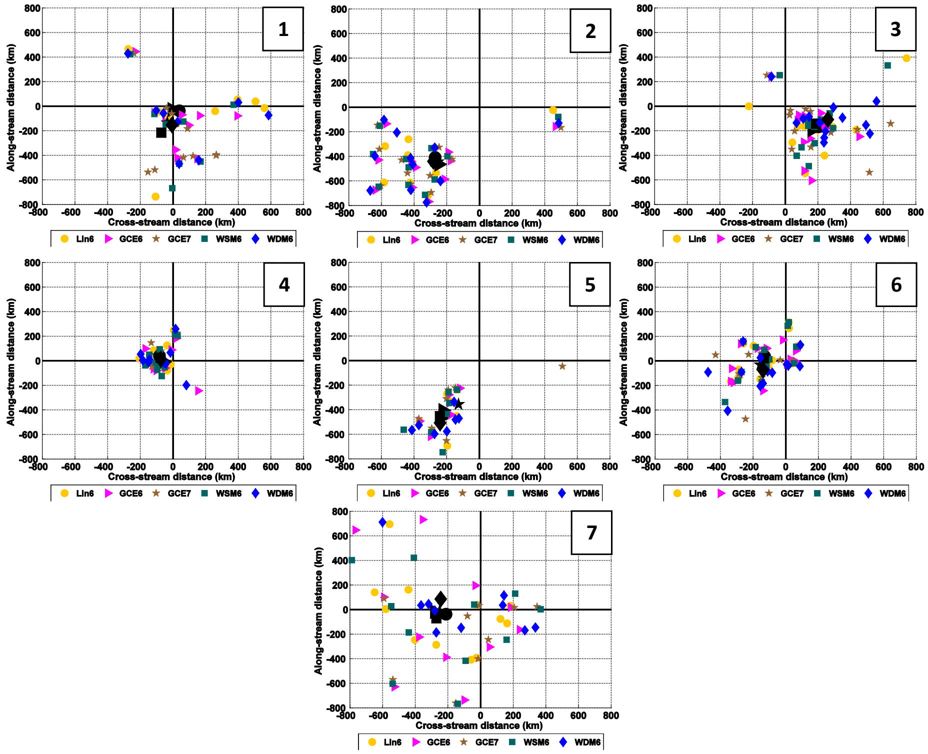


Figure 3. GMA-relative storm track error (km). Smaller, colored symbols denote storm track error every six hours and the large, black symbols denote the model mean error. The positive y-axis is aligned to six-hourly, GMA-relative storm track propagation direction. Black numbers indicate case number.

Minimum Sea-level Pressure (hPa) Maximum surface wind (m s^{-1})

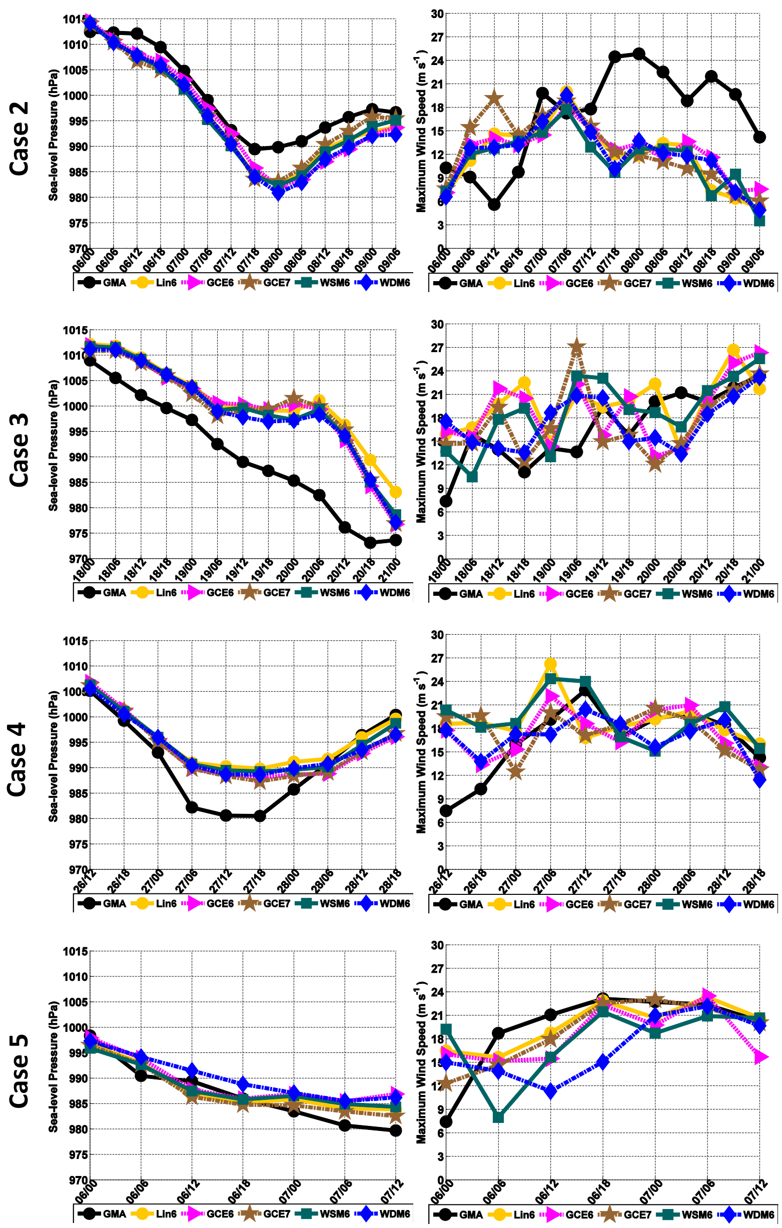


Figure 4. Plots of storm minimum sea-level pressure (hPa, left-hand panels) and maximum surface wind speed (m s^{-1}) within 600 km of the cyclone center from cases 2, 3, 4, and 5.

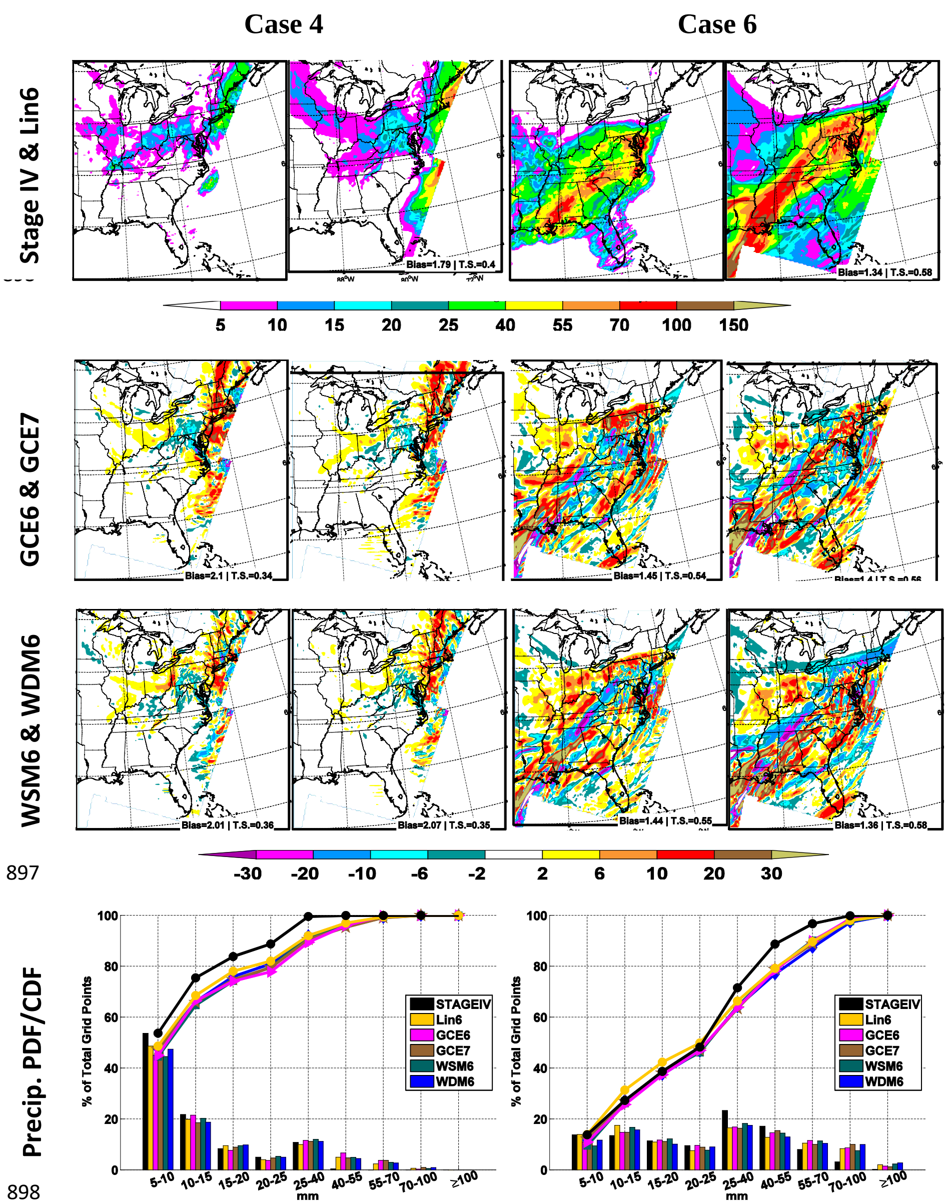
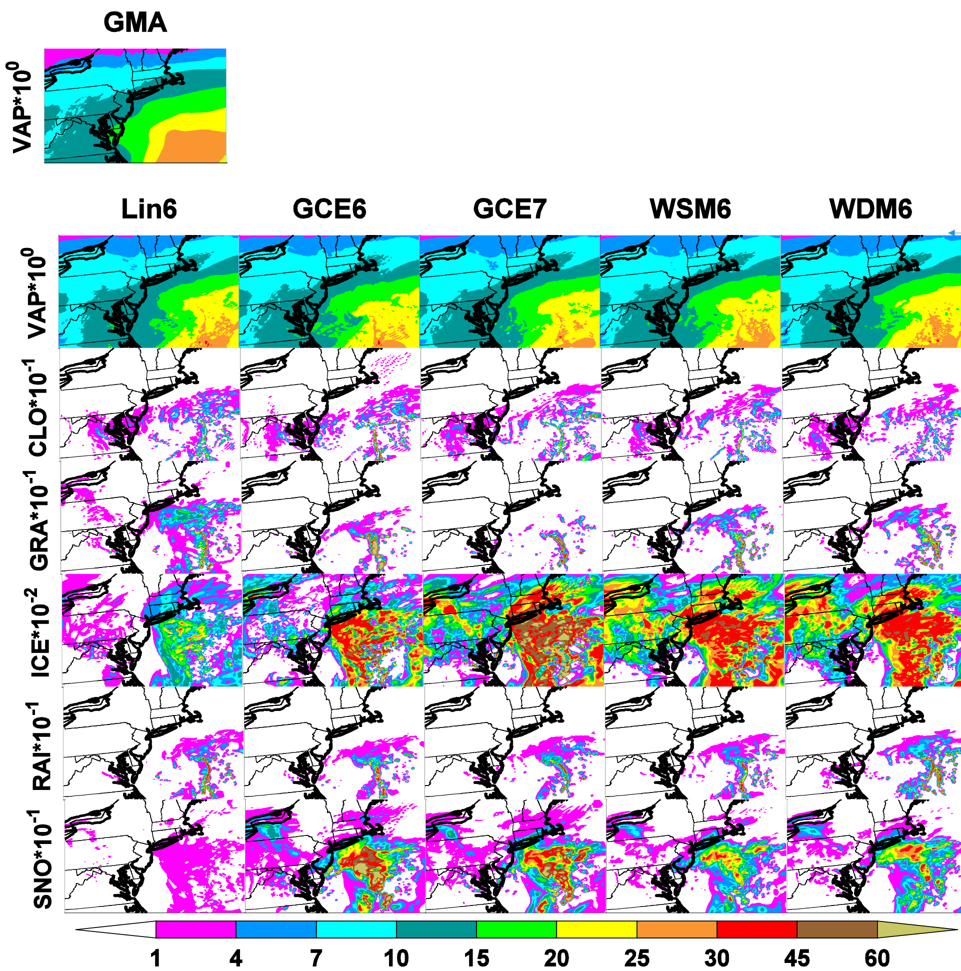
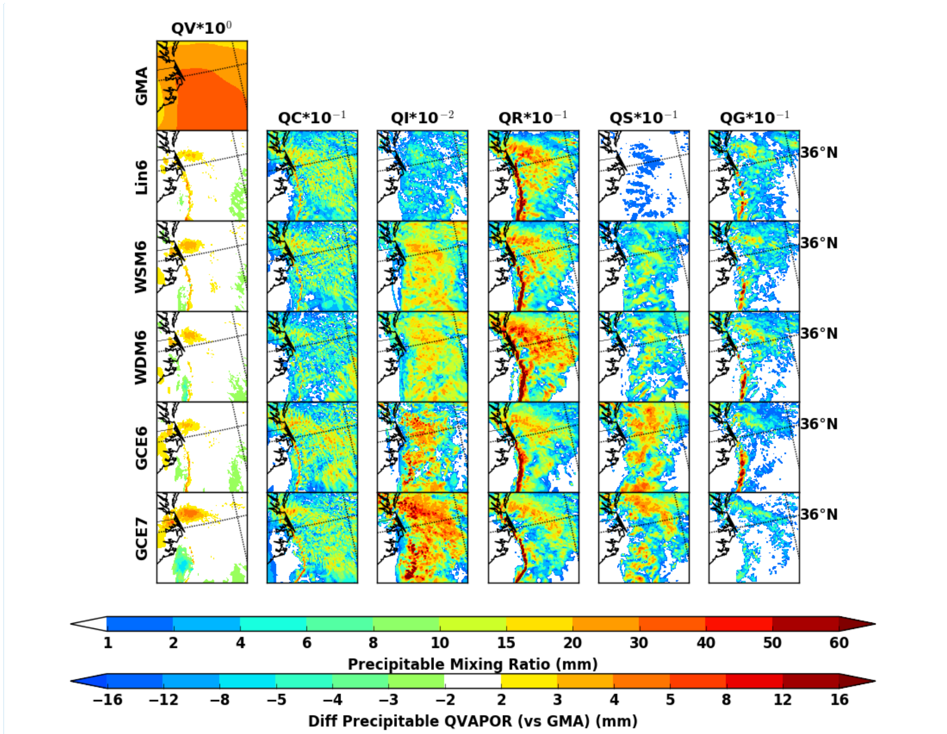


Figure 5. (top) 72-hour total precipitation accumulation (mm; forecast hours 48–120) from Stage IV and Lin6. (middle) Difference between other models and Lin6 (mm, model-Lin6). (bottom) Probability density and cumulative distribution functions of 72-hour accumulated precipitation for Stage IV and all models. Left-hand panels are for Case 4 and ri~~ght-hand panels are for Case 6.~~

Commented [NSD(RAU(11]: Lettering would make things more confusing. No change made for Figures 4, 5, and 6.



Formatted: Justified, Space After: 8 pt



Commented [NSD(RAU(12]: Modified to show water vapor mixing ratio differences.

Figure 6. Domain 3, precipitable mixing ratios (mm) at 18 UTC 26 Jan 2015. Shown abbreviations for mixing ratios include: VAP = water vapor, CLO = cloud water, GRA = graupel, ICE = cloud ice, RAI = rain, SNO = snow.

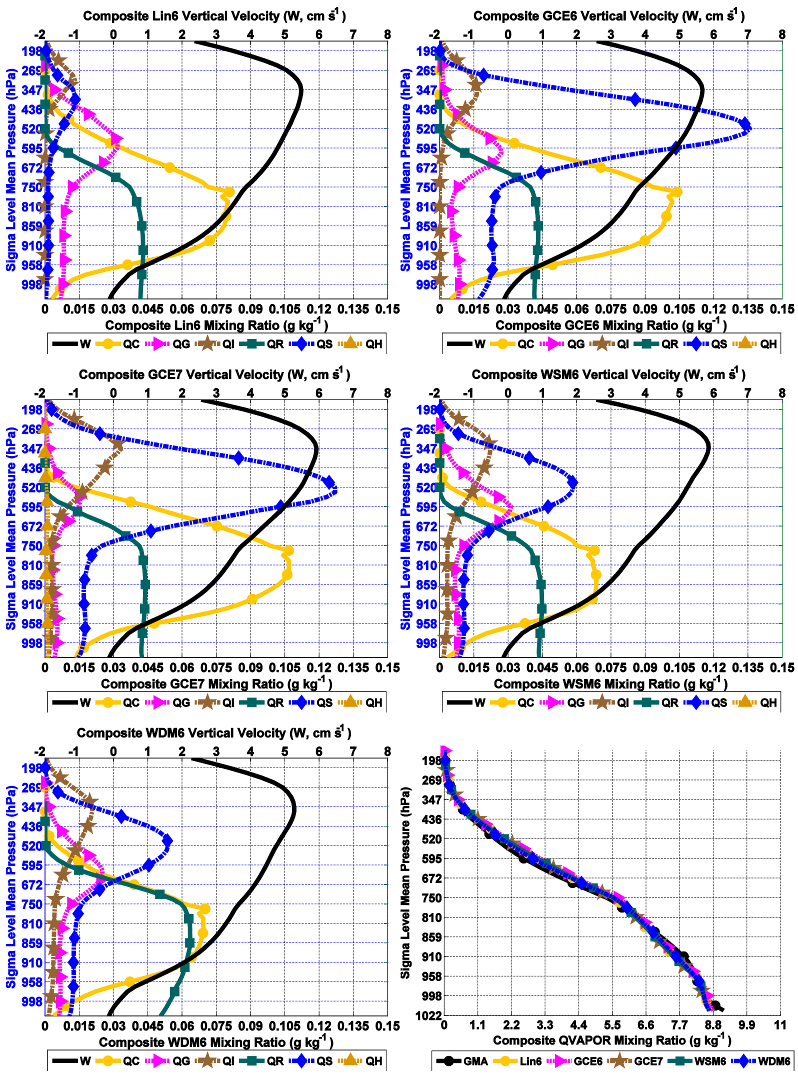


Figure 7. Composite mixing ratios (g kg^{-1}) and vertical velocities (cm s^{-1}) averaged over at all model-relative storm track locations (within 600 km diameter box) and all seven nor'easter cases. Mixing ratio species abbreviations are QC (cloud water), QG (graupel), QI (cloud ice), QR (rain), QS (snow) and QH (hail), and QVAPOR (water vapor, lower-right panel only).

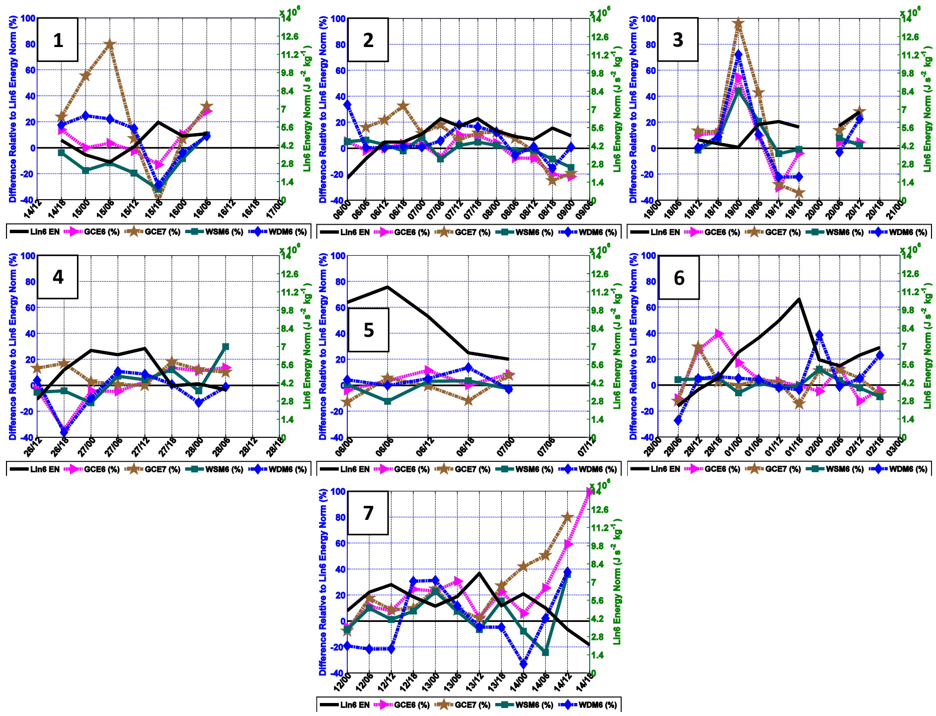


Figure 8. Model-relative total energy norm every six hours for each storm from Lin6 (black line, right y-axis) and difference (in percent) between energy norm from all other runs and Lin6 (colored lines, left y-axis). All energy norms were integrated only within a 600-km diameter box centered at the model indicated surface cyclone location. Postive precentage values indicate higher energy norm values than Lin6.

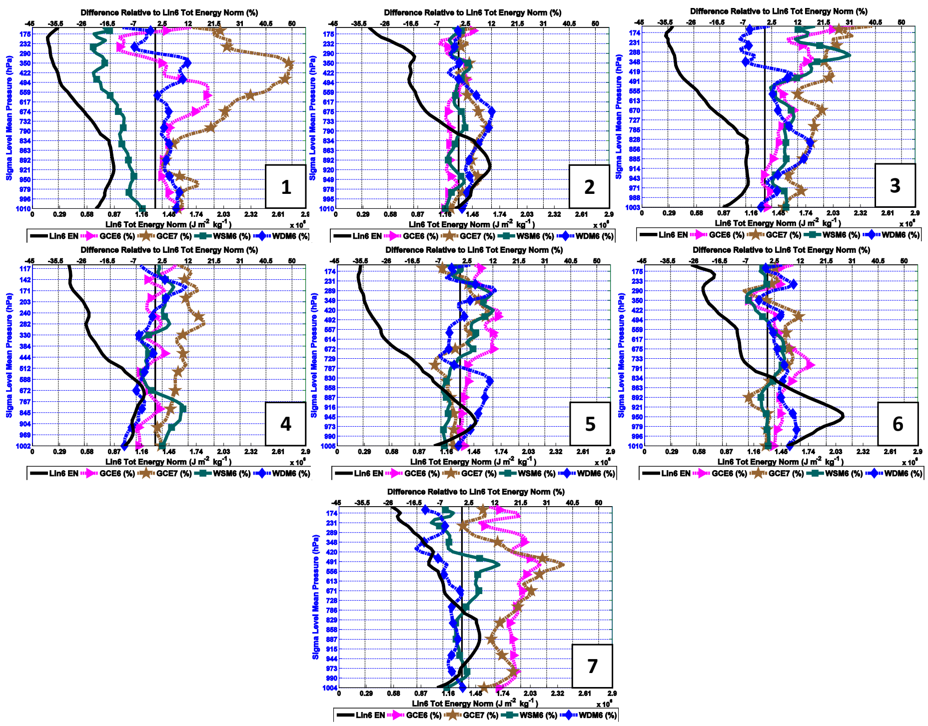


Figure 9. Model-relative total energy norm integrated on each model level and averaged over all times from Lin6 (black line, bottom x-axis) and difference (in percent) between energy norm from all other runs and Lin6 (colored lines, top x-axis). All energy norms were integrated only within a 600-km diameter box centered at the model indicated surface cyclone location. Postive precentage values indicate higher energy norm values than Lin6.

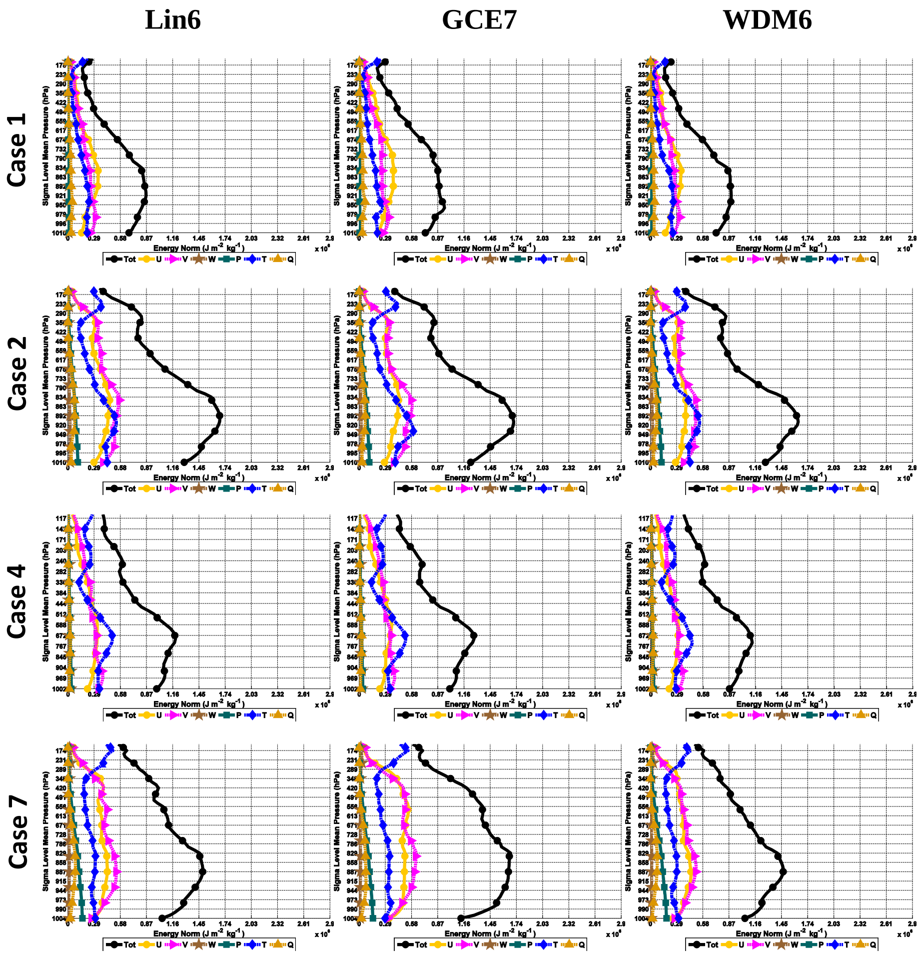


Figure 10. Time-averaged, model-relative storm environment energy norm components for cases 1, 2, 4, and 7 form the Lin6, GCE7, and WDM6 simulations. Shown lines include total energy norm (Tot; black) and its six-components (colors) including zonal wind (U; yellow), meridional wind (V; pink), vertical velocity (W; brown), atmospheric pressure (P; green), temperature (T, blue), and mixing ratio (Q; gold).