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 reran 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.

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).