Response to interactive comment on "Use of an observation-based aerosol profile in simulations of a mid-latitude squall line during MC3E: Similarity of stratiform ice microphysics to tropical conditions" by Ann M. Fridlind et al. by Anonymous Referee #1

General comments

This article constructs hygroscopic aerosol size distribution profiles from MC3E aircraft and ground-based data over six days. These profiles are used to run 4 NU-WRF simulations of a squall line case study. Observed and simulated cloud ice microphysical properties in the stratiform outflow region are then compared. The work is very comprehensive and cites the existing literature thoroughly. The results about similarity between continental and tropical ice microphysics are quite interesting.

We very much appreciate the helpful questions and comments. Point-by-point responses below have greatly improved the manuscript by reducing figures, adding section numbers, and making corrections and clarifications throughout.

Although factors like "fall speeds, aggregation and vapor growth rates, [etc]" are listed in the results, I would have appreciated more discussion on how the modeled ice microphysics might be improved to bring something like the number and mass size distributions into better agreement with observations.

Clarification added to Section 5: "The NU-WRF biases relative to observations shown here are consistent with the hypothesis that microphysics schemes are missing a key aspect of an updraft microphysics pathway that can largely determine outflow size, most likely associated with warm-temperature ice multiplication (e.g., Ackerman et al., 2015; Lawson et al., 2015; Ladino et al., 2017). Here we show that NU-WRF biases in stratiform ice mass size distribution are worsened when warm-temperature contributions to ice formation are decreased; Ackerman et al. (2015) find the same in parcel simulations and also demonstrate how biases can be decreased when warm-temperature contributions are substantially increased. In the simulations shown here, we also speculate that gravitational collection of stratiform ice may be too efficient, at least in the mid-troposphere, as evidenced by reflectivity increasing and number concentration decreasing substantially more rapidly than observed between 8 and 6 km (cf. Figs. 10 and 17)."

I missed also a discussion of the one-hour offset between the simulated and observed rain event initiation. Is there a hypothesis for this?

Clarification added to section 4.1: "The simulated squall line passes roughly an hour earlier than observed, which could be attributable to two general causes: (i) uncertainties in the initial and boundary conditions, including those influential to surface heat fluxes, and (ii) errors in model parameterization components, including microphysics scheme elements, which can independently influence the rainfall structure in NU-WRF simulations in this case (cf. Tao et al., 2016, their Fig. 11)."

There were two other points on which I would have appreciated clarification. I was surprised by the result that modeling only homogeneous freezing (HOMF) results in "substantially larger and

fewer outflow ice crystals". Normally homogeneous freezing yields many more and smaller crystals (e.g. DeMott et al. 1998 GRL). Why does the opposite occur here?

Clarification added to section 4.2.2: "Whereas favoring homogeneous freezing of droplets generally yields more ice particles in an updraft parcel (e.g., DeMott et al., 1998), here we find the opposite in aged stratiform outflow, where snow is the dominant hydrometeor class. Snow number concentration maxima intermittently reach ~500 L^{-1} in all simulations except HOMF, where they reach only ~30 L^{-1}. Since 500 L^{-1} is the limit imposed on the Cooper (1986) parameterization contributions to total ice number concentration (see Section 4.1), we conclude that removing that source is likely chiefly responsible for larger ice in HOMF outflow. We note that ice number concentrations are not conserved by design in order to enforce limits on size distribution slope parameters (Morrison et al., 2009), which complicates drawing firm conclusions about the contributions of specific processes."

Then I found the results for the size distributions in Figures 14 to 17 and radar reflectivity in Figure 22 incongruous: the distribution comparisons indicate that the simulated ice crystals are far too big, while it is suggested from the reflectivity comparison that the simulated ice is too small. Am I missing something? Clarification in both cases would be helpful.

Clarification added to section 4.3.1: "Thus, specifically at the elevations where the aircraft sampled (Fig. 16, white bars in observed reflectivity), simulated reflectivity is substantially greater than observed, consistent with ice particles substantially larger than observed (Figs. 11–13), but that is not the case at all elevations."

Otherwise my comments are related to readability. I find the article rather figure-heavy, and I think the results would be made be more accessible if the figures were condensed in some places and simplified in others. For example, Figure 2 is only referred to once, and since only the 20 May panel is particularly relevant, this panel could be combined with Figure 12. In Figure 13, only the rain gauge-corrected QPE measurements and BASE simulation are discussed, so panels a and c could be removed. Or Figures 17 and 18 could be moved to Supplemental Information, since the altitudinal dependence of Ni and mass distributions is already seen between Figures 15 and 16 and the discussion of 2DC images is quite brief.

We combined Figs. 4 and 5 and removed 6, 10–11, and 19–20. We retained 2 (for reader to quickly assess other case study conditions), 13 (emphasizes substantial uncertainty in rainfall products), 17 (15-17 are main focus), and 18 (for modelers to know what ice looks like).

I think breaking down the "Evaluation of hydrometeor size distributions in 20 May case study simulations" section into subsections, e.g. "Precipitation intensity", "Mass and number concentration distributions", and "Radar retrievals", would also ease readability.

We now use two levels of subsections in Sections 3 and 4.

Specific comments

Page 4, Line 6 – Please be consistent in the instrument acronyms. What is called the "DMA" here is later called the "HTDMA" in Figure 3 and introduced as the "TDMA" in Section 2. Again for the CPC, it is not always clear whether the measurements to which you refer are form the ground-

based or aircraft CPC; it is inferred from the other instruments you mention. You could make this more explicit.

HTDMA now used throughout. CPC now always preceded by "ground-based" or "airborne."

Page 6, Line 31 – The statement “unknown aerosol source terms are neglected” is unclear to me. The airport and power plants are mentioned in the section of Aerosol input data, but there is no discussion of back trajectories or systematic confirmation of hypothesized sources.

By unknown we meant that aerosol source terms cannot be readily observed and specified. Simplification and clarification made: "Aerosol source terms beyond advection across outer domain boundaries are neglected (e.g., primary emission and gas-to-particle conversion)."

Page 7, Lines 1-2 – Is there also a quantitative basis (other than “similarity to April case studies”) for the 8000 cm-3 and 0.005 um values chosen in the NUCL simulation? If so, this should be mentioned.

Clarification added also in response to referee 2: "Based on the April and 1 May nucleation-mode fits listed in Fig. 6, this represents the most commonly fit mode diameter and mode standard deviation, and a modest number concentration (maximum on 1 May) that is lower than typically observed in the 10–30-nm diameter range during intense new particle formation events (e.g., Crippa and Pryor, 2013)."

Page 7, Line 3 – The statement “simulations use a preliminary version of the 20 May aerosol input data” is unclear to me. The Aerosol input data section does not mention multiple processings or versions.

Clarification added: "During the course of this study, minor changes were made to aerosol observation processing concurrently with the simulations being run; simulations therefore use a preliminary version of the 20 May aerosol input data, which is negligibly different from the final version for our purposes. AERO and NUCL aerosol input files are included in Supplement 1 for completeness."

Page 8, Line 1 and Page 47, Table 2 – Could you please include the standard deviation in the “top three elevations”, e.g. 7.6 ± x m, and associated temperatures?

We prefer not to complicate the table because the elevations and temperatures are in a narrow range based on level legs within horizontally homogeneous conditions and the table is already complicated by showing a range of minimum and maximum values from two observational data sets.

Page 11, Line 16 – It is not clear what “similarly coherent” means here. Could you word this more substantively?

Reworded to "both predicted and observed stratiform ice size distributions exhibit relatively well-defined properties that do not vary rapidly in time."

Page 28, Figure 7 – The caption indicates that the CPC profiles on the left and UHSAS profiles on the right are in red and blue respectively as in Figures 4 and 5, but this is not the case. The thick black line for layer-wise median ratio is not so easily distinguished from the thinner black lines; perhaps the UHSAS/CPC traces can also be changed from black in the rightmost subplot. Finally

it is not clear what the “layer-wise” ratio means; are these values also calculated for km-deep layers?

Clarification added to figure and caption: "The median of airborne CPC and UHSAS aerosol number concentrations within 1-km-deep layers for each MC3E flight, and the ratio of those median values for the seven flights with both instruments (black lines). The median of profile values at each elevation (red lines) are archived as Supplement 2."

Page 29, Figure 8 – The numbers in the subpanels of this figure need to be moved to a table. This will significantly ease comparing the values between days and allow the y-axis to be readjusted for better comparison of the different traces. It is also unclear to me what the various colors (red, green, blue, purple, black) represent. The caption refers to “measurement time”, but this should be clarified. A brief discussion of why the 2-mode fit is better than the 3-mode and vice versa at certain times might also be included in the third paragraph of page 5.

We used a fixed vertical axis to emphasize case study differences. The black values are archived with Supplement 1 and we disagree that the underlying values deserve a dedicated table. Clarifications added to caption also in response to referee 2: "Aerosol dry number size distributions ($dN_a/dlogD_a$) reported from HTDMA during the two-hour pre-rain period (colored solid lines; legend indicates Julian date in UTC), lognormal fits to HTDMA (colored dashed lines; text indicates fitted number concentrations in cm^{-3}, geometric mean dry diameter in μm and standard deviation), and the final case study distribution derived from the mode-wise linear mean of contributing parameters and its hygroscopicity parameter (κ) derived as the number-weighted mean of contributing HTDMA values (black dashed lines and black text; archived with Supplement 1). In the 20 May case, zero and 8000 cm^{-3} particles in the nucleation mode illustrate BASE and NUCL simulation inputs (dotted black lines)."

Reworded "It is found that two to three modes provide the best fit" to "The Vogelmann et al. (2015) algorithm optimizes a fit of two or three modes" to emphasize that we relied entirely on that algorithm since results appeared consistently satisfactory.

Page 35, Figure 14 – In my opinion, this figure could be removed, and the simulated values added to Table 2.

We have retained it because this figure conveys information that is difficult to fully capture in a table and we removed six other figures.

Page 41, Figure 20 – I am confused by the black BASE trace for number concentration. Doesn’t this simulation have a fixed droplet concentration of 250 mg-1, as stated on page 6, line 26?

Correction made to text: "250 cm^{-3}."

Page 42, Figure 21 – It is unclear whether only the top left panel is an integrated reflectivity; it seems so given its different scale, but this should be clarified in the caption. A definition of ZHH (as the horizontally-polarized radar reflectivity, right?), along with definitions for the pink, white, and red circles in various subpanels, would help in the interpretation of this figure.

Clarifications added to caption: "Horizontally polarized radar reflectivity (Z_{HH} in dBZ) from KVNX radar (left, dotted red circle): (top) example updraft object at ~12 UTC (solid red) among others identified in units of dBZ km (red-enclosed, see text), (middle) movement of example updraft

from initial location (solid red) towards intersection with the aircraft sampling location (white-enclosed, see text) projected onto 2-km Z_{HH} at ~14 UTC, and (bottom) Z_{HH} curtain obtained from column-wise averages over tracked regions from ~12–15 UTC with Citation ascent legs in time and height (white bars) and averaging time used in Fig. 22 (white lines). From the AERO simulation (right): (top) identification of a typical updraft object projected onto simulated Z_{HH} at ~11 UTC (solid red) among others identified (red enclosed, see text), (middle) its movement from the identified location (solid red) to intersection with the aircraft sampling location (white-enclosed, see text) projected onto simulated 2-km Z_{HH} at ~13 UTC, and (bottom) Z_{HH} curtain obtained from column-wise averages over tracked regions from ~11–14 UTC with mid-point of hour-long averages used in Fig. 22 (white lines)."

Page 43, Figure 22 – "Time 1", "Time 2", etc. have not been defined for the simulations. It would be clearer to label the gray traces 'AERO, Time1' etc. so that the reader knows these are only from that simulation.

Clarification added to caption: "AERO simulation times 1, 2, 3 and 4 indicated in Fig. 21 (light to dark grey lines)."

Technical comments / suggestions Page 3, Line 25 – A term like "droplet activation" or "ice nucleation" or "new particle formation" would more clearly indicate the process(es) meant by "aerosol consumption" here.

Clarification added: "via droplet activation".

Page 3, Line 26 – Remove the second "be".

Removed, thank you.

Page 6, Lines 9 – 10 – Add a 'to': "appears to be variably biased relative to the groundbased measurements".

Added, thank you.

Page 8, Lines 16-22 – Reword through here for clarity, e.g. "Consistent with underestimated Ni, the Dmax at which BASE mass distributions peak is roughly 3-5 times larger than that at which the observed distribution peaks. The Dmax at which the BASE mass distributions peak increases monotonically with increasing mass concentration, whereas the observed mass distributions tend to..

Reworded, thank you.

Page 8, Line 28 – There is an unfinished sentence beginning with "At 6.7 and 7.6 km".

"At 6.7 and 7.6 km. However," corrected to "At 6.7 and 7.6 km, however,"

Page 12, Line 18 – "Updrfts" to "updrafts"

Corrected, thank you.

Page 14, Line 4 – "have" to "has"

Corrected, thank you.

Page 14, Line 6 – “are” to “is”

Corrected, thank you.

Page 22, Figure 1 – It would ease readability if the ARM central facility were marked with a color other than yellow, since the pentagon, bull’s eyes, and thumbtacks are all yellow as well.

Agreed. Since this is a stock figure that we did not generate, we did not attempt to adjust it.

Page 24, Figure 3, panel d – Is there a red trace for 0.013 um here? If so, it is not visible.

It is strongly intermittent. Clarification added to caption: "(intermittent at smallest cut)."

Page 33, Figure 12 – It would ease readability if Q2 were expanded to National Mosaic and Multi-Sensor QPE system in this caption, as well as in the text, and again if QPE were expanded here and in the text.

We have now spelled out "National Mosaic and Multi-Sensor Quantitative Precipitation Estimate" in the caption to Figure 12 and in the text.

Pages 36-38, Figures 15-17 – The red and blue traces should be labeled PMS 2DC and HVPS rather than obs1 and obs2.

Both are merged PSDs from the same raw data, adopted here as an estimate of poorly established uncertainty. Clarification added to caption: "Size distributions of ice mass (left) and number (right) in four ranges of ice water content (IWC, ranges in parentheses in g m^{-3}) derived from merger of 2DC and HVPS raw data independently by Wang et al. (2015a, 'obs1' in red) and Wu and McFarquhar (2016, 'obs2' in blue). Both are shown as an estimate of poorly established uncertainty."

Page 41, Figure 20 – The y-axis should be ‘[km]’ not ‘[m]’.

Figure removed.