Reviewer 2 Responses

General comments:

The authors conduct a numerical modeling study of secondary ice production and its possible effects on surface precipitation, based upon observations of the 3 March 2009 cold frontal passage across Southern England analyzed by Crosier et al. (2014). They address three kinds of secondary ice production: rime splintering, shattering of freezing raindrops, and collisional ice breakup, with several variations of coefficients that control the magnitude of each process. By enhancing the magnitudes of each process above that represented in past studies, or justified by laboratory experiments, they find (unsurprisingly) that the ice number concentrations can be greatly increased beyond that expected from primary nucleation. Through some (very rough) comparisons with the observations, they are still unable to replicate the maximum observed ice crystal number concentrations, and yet, estimates of the ice production rate actually exceed those based upon the observations. This disparity suggests that there might be some issues in comparing the observations and simulations, and/or that the simulated cloud dynamics are significantly different than in the observed clouds. The authors find an increase in surface precipitation of ~20% for the simulations that maximize secondary ice production, and advocate from that result that parameterizations of these processes should be included in large-scale models.

Thank you for your thorough reading of the work and suggestions to make it more rigorous. At both reviewers' behest, we have spent significant time to redo the simulations, adjust the structure of the study, and visualize new data from APPRAISE.

For the simulations, some of the parameter values that were chosen, particularly for rime splintering and droplet shattering were overly generous as you state. This was done, in part, to see if we were able to bring model and measured values into agreement with only modification to the secondary ice microphysics. Again as you state, we did not have agreement even at these extreme values, and so we have retained more conservative parameter values (sticking to, for example, 300 fragments (mg rime)$^{-1}$ for rime splintering) and modified the simulations as shown in the new Table 2 at the end of these responses.

Then we have added a section dedicated to reviewing the dynamical environment prior to any discussion of microphysical or precipitation (see responses to Comments 1, 6, and 8). This includes comparisons of radar reflectivity, updraft velocity, and surface wind speed.

Specific comments:

The authors are tackling a very difficult problem here, and studies like this are important and necessary. However, more care must be taken in what they can and cannot conclude from this study. I have some serious concerns with how the authors conducted some of the analyses, and/or their interpretations. In the manuscript, some very important details are omitted, that make it difficult to understand their interpretations.

Overall, I would like to see this study move forward, but feel that it would be of greater use to emphasize the temperature and dynamical regimes over which each of the secondary ice production processes is dominant, and how those differences assist (or do not) the formation of additional precipitation). That might be a more useful place to start when advocating that some of these processes be included in larger-scale models, as it would help focus case studies of the type of weather phenomena where they could have the most impact.

1. If observations and simulations are compared in this way, particularly when convective elements are contained within the weather phenomenon of interest, then it first must be demonstrated that the clouds and precipitation due to the cold frontal passage in the control case are consistent with that observed, and if not, to state clearly how they differ, and continue compensating for those differences when comparing the observed and simulated microphysical development. For example, a figure showing simulated radar echo can be produced, and shown alongside Crosier's Fig. 3 to understand how the general structure might differ. The timing is also important: if observations over a given time are averaged and compared with the simulations, any issues in doing so must be known. A small paragraph summarizing the dynamics of the observed clouds, based on the analysis of Crosier et al., would also be helpful in "setting the stage" for the reader, regarding the types of clouds (strengths of updraft speeds measured by aircraft, cloud top temperatures) being considered here.

Thank you for pointing out the need for these kinds of dynamical comparisons. We have compared modeled and observed radar reflectivity as you suggest. The comparison cannot be exact, as the field shown in Crosier et al. 2014 (and the CAMRa data in general) have much higher resolution along a radial whereas the model output exist only on the 2.8 km grid. In addition, the range-height indicator (RHI) of Crosier et al. 2014 comes from the CAMRa scan between 192207 and 192307 UTC, whereas we only have model output every 30 minutes.

Nevertheless, we calculate the latitude-longitude pairs along the 255° radial RHI in Fig 5b and identify the modeled lat-lon that minimize the Euclidean distance between these exact values and the model's spatial discretization. From here to generate an RHI, we iterate over a 1000 x 500 grid of distances and heights (relative to the Chilbolton Facility for Atmospheric and Radio Research) and do a bilinear interpolation over the nearest modeled lat-lons and 29 model levels to generate the radar reflectivities that are shown in the new Figure 4:

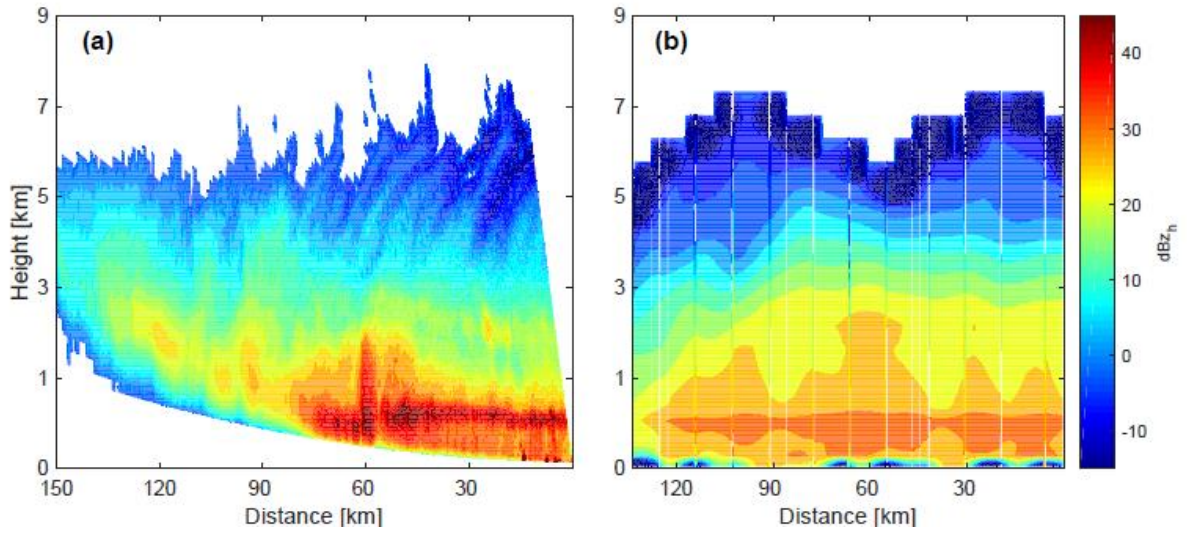

Figure 4 Model-measurement intercomparison of range-height indicator scans of radar reflectivity Z_{DH} along the 255 degree radial out from CFARR. CAMRa Doppler radar measurements are shown in panel a for the scan taken between 19:22:07 and 19:23:07 UTC, and modeled values are shown from the CTRL simulation at 19:00:00 UTC, both in dBZ_h.

We use the same methodology to calculate an RHI-type plot of the modeled updraft velocity; however this comparison includes an extra degree of inexactitude, as even the "observed field" is actually derived "using Doppler velocity measurements from RHIs by assuming mass-weighted flow continuity" [Crosier et al. 2014, Chapman and Browning 1998]. We have also compared modeled and observed (from the CFARR ground site) wind speeds and shown panels of modeled updraft velocity at two different altitudes during rain band passage in the new Figure 3:

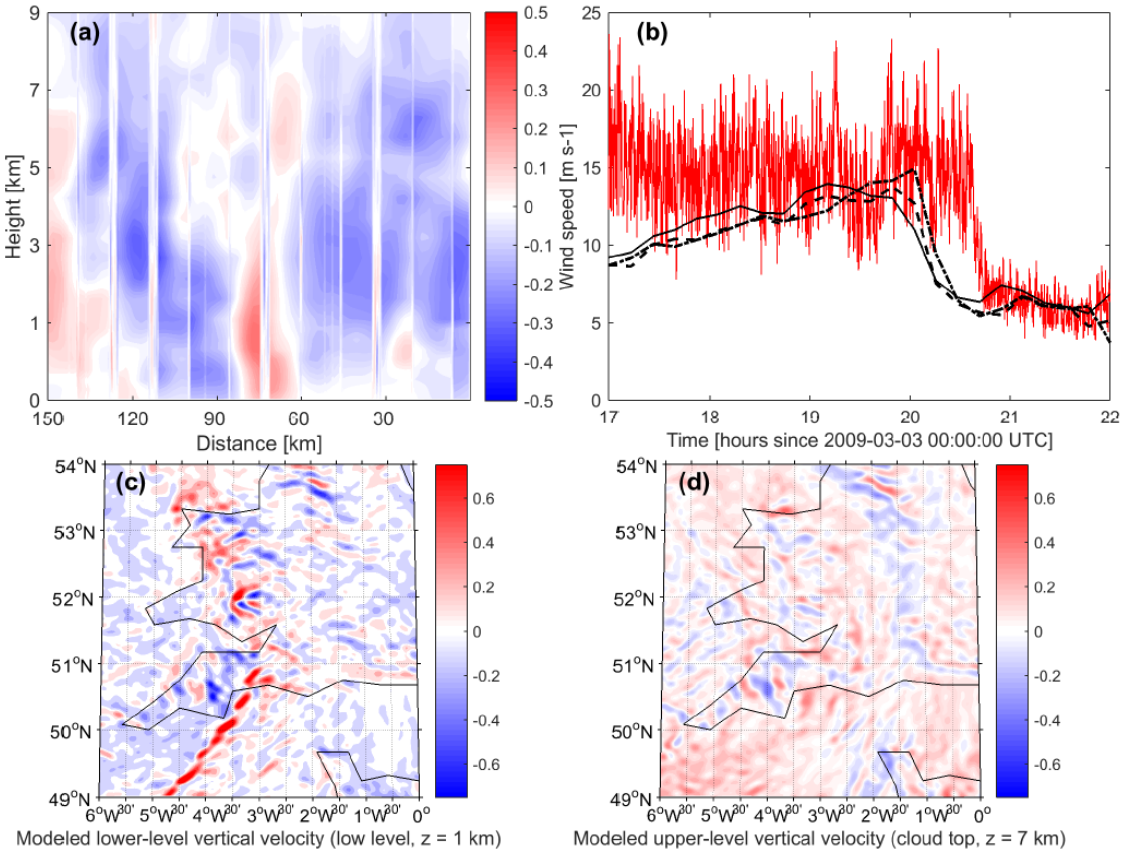

Figure 5 Different model dynamical fields. In panel a, we show the updraft velocity 150 km along the 255 degree radial from CFARR at 51.1N, 1.4W with values from the CTRL simulation at 19:00 UTC. Discontinuities are due to the minimization of Euclidean distance or interpolation aspects of an algorithm to approximate the radial from the model discretization. Surface wind speeds from 17:00 to 22:00 UTC from the CFARR three cup anemometer and our CTRL simulation are shown in panel b. Panels c and d show the modeled vertical velocities from the CTRL simulation at 1 and 7 km respectively at 18:00 UTC as the rainband began to pass over the UK.

We describe these methods and comment on their impact on microphysical comparisons in a new Section 4.1 on the *Dynamic environment*:

4.1 Dynamic environment

We begin by comparing the observed and modeled dynamics to understand how these differences may impact later microphysical ones. We show the updraft along a 255° radial out from the CFARR station (Figure 3a) as in Figure 5d of Crosier et al. 2014. Comparison to simulations cannot be exact, as the output from the model exists along a coarser spatiotemporal grid. We use those w values whose latitude-longitude pair minimizes the Euclidean distance to the precise values along the radial. Thereafter we interpolate both over distance and altitude on a 1000 x 500 grid to generate the range-height indicator (RHI)-type plot. This interpolation leads both to intermittent discontinuities and to weakening of extremes. The upright updraft region about 60 km from CFARR appears distinctly in the simulated field but with vertical velocity magnitude far smaller but extent far greater than those derived from measurements (maximum of about 1 m s^{-1} relative to about 6 m s^{-1}). Downdrafts of more similar magnitude and extent to those observed also form in adjacent regions. Values derived from Doppler velocities (v_u) also rely on an assumption that at low elevation, v_u approximates the horizontal wind and that any convergence or divergence of these horizontal winds within discrete columns must conserve mass with a compensating upward or downward velocity.

Comparison of the surface wind speed both in the CTRL simulation and from a three cup anemometer at CFARR is also shown in Figure 3b. Three series are shown from the simulation at latitude and longitudes closest to the center. Simulated wind speed peaks too early but to a value only slightly less than the average of the observations. Both series display a sudden drop in the strength of these winds with similar decay rates and plateau

values of about 5 m s^{-1}. Perhaps most important is the consistent underestimation of these surface winds prior to the rainband event, from 1700 to 1900 UTC. Given that the direction of low-level winds preceding the rainfall event was southwesterly (Crosier et al. 2014, their Figure 4b), underestimating their magnitude will diminish the oceanic moisture advection and moisture source ultimately available to form rain over the continent.

Figures 3c and d also show the vertical velocities in the CTRL simulation at altitudes of 1 and 7 km at 1800 UTC as the rainband reaches land. Its structure is apparent in the low-level updrafts of about 1 to 2 m s^{-1} (although these are again much weaker than those from observations) and their adjacent downdrafts with similar magnitudes of opposite signs. Elsewhere values are ± 0.2 m s^{-1} with slow descent presiding. For the upper-level field that corresponds to cloud top, the highest ascending motions also occur around the rainband region and slow ascent (≤ 0.4 m s^{-1}) dominates.

We next compare range-height indicator (RHI) scans of radar reflectivity (Z_{DR}) from the Chilbolton Advanced Meteorological Radar (CAMRa) and the CTRL simulation (Fig.4). The CAMRa is a 3 GHz Doppler instrument with a 0.28° beam, and its scan between 192207 and 192307 UTC along the 255° radial out from CFARR is shown, as in Figure 5a of Crosier et al. 2014. We use output from 190000 UTC in the CTRL simulation and again identify the modeled latitude-longitude pair that minimizes the Euclidean distance to the exact value from along the 255° radial. We then perform bilinear interpolation on the simulated values of Z_{DR} over a 1000 distance x 500 altitude grid.

The CAMRa scan shows the location of cloud top height and convective activity: the lowest Z_{DH} is around 6 to 8 km and fall streaks are present moving toward the CAMRa. These Z_{DH} fall streaks, as well as those in differential reflectivity (shown in Crosier et al. 2014 their Fig. 5c) indicate that ice crystal seeding may be occurring near cloud top. Intermediate values of Z_{DH} occur at altitudes of 2 to 5 km, and the highest ones occur around the melting layer at 1 to 2 km, as discussed by Crosier et al. 2014. General features are replicated in the simulated reflectivities. Very low Z_{DH} occur close to CFARR with cloud top around 7 km, but further out -- around 70 to 100 km along the radial -- these same very low reflecitivities occur more often than in the measurements. The gradient to higher Z_{DH} at lower altitudes is also apparent in simulations, but not as distinct fall streak structures. Z_{DH} has increased to about 10 dBZ_h by a height of 4 km and about 20 dBZ_h by a height of 2 km. The highest reflectivities also fall in the same altitudinal range, but importantly do not have the same maximum as in the observations. Z_{DH} in an updraft core 60 km from CFARR reaches a value of 45 dBZ_h in the CAMRa but only 30 dBZ_h in the CTRL simulation. This may be due to underestimation of graupel formation or too high CCN or INP concentrations that delay precipitation in the base COSMO model [e.g., Baldauf et al. 2011]. We keep this underestimation in mind in the proceeding discussion of microphysical adjustements.

2. The routes from ice to precipitation discussed in the introduction, shown in Fig. 1, and later discussed with respect to effects of the secondary ice upon precipitation, are not inclusive of a major route from ice crystals to precipitation: enhanced rimed ice / rimed snow / graupel / frozen raindrop formation that can melt to become surface precipitation. In Fig. 1, it is somewhat suggested by (3), but the arrow isn't drawn as leading to acceleration of precipitation like (2). Such an analysis of that route to precipitation is completely omitted in the manuscript. Why?

   Thank you for pointing out this missing mechanism. Figure 1 focused on showing how secondary ice production (solely) impacts precipitation via increased ice crystal number concentrations, so we had not included impacts of just riming (only rime splintering). But we agree that a more complete analysis cannot consider secondary ice processes in isolation, and we have added enhancement of precipitation by rimed hydrometeors to the schematic. To the discussion of cloud ice-precipitation linkages, we add the following:

   Efficient riming at mixed-phase temperatures may also simply generate larger hydrometeors that sediment more quickly, particularly in convective regions with a high degree of mixing.

   Even if no graupel were observed, the Crosier et al. paper discussed the importance of rimed snow, and noted that the aircraft did not sample the stronger convection where graupel might have resided. The authors only discuss in this study the possible effects on the Bergeron process leading to precipitation, but that would be more important in the stratiform precipitation regions, and not as

much in the convective regions of the cold front band, where the heaviest precipitation will fall. I would think that the precipitation enhancement see in Fig. 5 is due to rimed particles, not from an enhanced Bergeron process.

Elsewhere in the introduction, we note that *in cases of ice-initiated precipitation, the requisite crystal growth can occur **via riming** or the Bergeron process.* In the discussion of ice-precipitation linkages, we note that *this pathway [of small crystal formation depleting supersaturation until the Bergeron process initiates] should be more important for stratiform precipitation, given the narrow range of requisite ambient vapor pressures: indeed for an integral ice radius of 100 um cm^{-3}, the updraft must be less than about a 1 m s^{-1} for the Bergeron process to occur [Korolev 2007].* To the discussion of impacts on precipitation, we add that *this amplification may be due in part to more riming in the ascending regions, which feeds into precipitation both directly as rimed particles sediment and melt (Fig. 1(3)) and indirectly as they splinter and generate more rimable particles (Fig. 1(4)). Crosier et al. 2014 note that higher values of differential reflectivity around cloud top could be due to high concentrations of rime particles.*

3. The implementation of the secondary ice parameterizations in the two-moment Seifert and Beheng scheme are confusing.

a. Why is the second moment not taken advantage of here? Everything seems to depend upon mass.

We are indeed taking advantage of the second moment outside of the fragment number parameters. Within the ice-ice collisional breakup formulation for example, the D_j and D_k in Equation 6 are the particle diameters associated with particle mass through a power law as detailed in Appendix B. Within the droplet shattering formulation, the tendency of freezing droplets in time is a function of the mean mass per raindrop as detailed in Appendix A.

The importance of the second moment extends beyond a collision or freezing tendency to the fragment number parameters, and you are right that this dependency has not yet been incorporated. There is not enough consensus of laboratory and in-situ measurements to support one fragment number function in our opinion. We expand the section on the frozen droplet shattering parameterization (Section 2.1) and add to our comment in the earlier version that *future studies should add dependency on droplet size to the ejected fragment number*:

Recent droplet levitation experiments and high speed video are elucidating the exact physics behind the shattering of droplets as they freeze [Leisner et al. 2014, Wildeman et al. 2017]. Droplet shattering has been previously parameterized statistically in a bin microphysics scheme with the fragment number as a function of drop diameter to the fourth power, using data from the Ice in Clouds Experiment - Tropical (ICE-T) campaign [Lawson et al. 2015, Lawson et al. 2017]. But while measurements continue to confirm a strong dependence of fragment number on droplet size, even recent studies could not confirm this fourth-power dependence [e.g., Lauber et al. 2018]. The laboratory studies of Lauber et al. 2018 in particular add important quantitative results to existing secondary ice measurements but are taken at two droplet sizes (83 and 310 um) so that it remains difficult to rigorously formulate fragment number.

To the section on the ice-ice collisional breakup parameterization (Section 2.2), we also add discussion of other parameterizations:

Vardiman first parameterized ice-ice collisional breakup using fragment generation functions based on the momentum exchange between two particles upon impact and leading coefficients dependent upon crystal type [Vardiman78]. More recently, Yano and Phillips 2011 and Yano et al. 2016 constructed a dynamical system-type models that tracks only ice crystal, small graupel, and large graupel number densities and illustrated the ability of ice-ice collisions to generate huge ice crystal enhancements in the absence of vapor limitation. Recently a more complete parameterization has used an exponential formulation with the initial kinetic energy of two particles, their temperature- and humidity-dependent collision type, and asperity fragility coefficients [Phillips 2017a, Phillips 2017b].

We choose to focus on temperature dependence in a more straightforward, if less physically rigorous, product of fragment number and hydrometeor collision tendency.

For example, rime-splintering appears to have a constraint of rimed mass, but for a two-moment scheme, the Cotton et al. (1986) second formulation that uses the number of fragments per number of 25 um diameter drops accreted would be a better prediction. The lab studies have shown that if the rimed drops don't achieve this size they won't splinter. As implemented here, there is no drop size dependence, so splintering might be greatly overestimated, and some commentary needs to be given in regards to that limitation.

In connection with the previous comment, this statement is a bit confusing, as it advocates for use of the first moment *rather* than the second moment. Nevertheless, we had made the rime splintering parameterization active only for rain drops. At the end of Section 2.3, we had stated: *We limit rime splintering to occur only after collisions between raindrops and ice crystals, graupel, or hail.* However, raindrops in the Seifert and Beheng scheme have a radius of 40 um, and as you state, the radius of onset for rime splintering is lower, around 12 um. So we have gone back and implemented rime splintering with cloud droplets as well, defining a 12 um threshold radius for these cases. We adjust the Section 2.3 description:

We also limit rime splintering to occur only after collisions between cloud droplets of diameter greater than 25 um or raindrops (r > 40 um) and ice crystals, graupel, hail, or snow [e.g., Phillips et al. 2001, Connolly et al. 2006b].

b. Along similar lines, what is the justification for the experiments using broader temperature ranges and/or increased fragment numbers for rime splintering? That process has been studied much more in the laboratory than others (Hallet and Mossop and Saunders & Hosseini, AR, 2001). The results here seem to rely on the expansion of this process to a broader area of temperature than appears justified by the laboratory work. The latter study also looks at the importance of fall speed, where graupel is more favored for greater splinter production. Since the simulation has little / no graupel here, then it would seem to imply using smaller splintering rates is appropriate.

Some justification comes from the potential dependence of the process on rimer surface temperature rather than ambient temperature [Heymsfield and Mossop, 1985], but not to the extent that our initial rime splintering temperature weightings (w_{RS}) reached. We have amended w_{RS} to the typical triangular weighting between -3 and -8°C and a uniform one between 0 and 10°C. The latter is still somewhat generous in order to pick up on any potential "cascade effect" between a droplet shattering or ice-ice collisional breakup "trigger" and rime splintering. We note this specifically in Section 2.3:

We add a second, uniform temperature weighting (UNI) between 263 and 273 K to investigate the possibility of a droplet shattering or ice-ice collisional breakup `trigger' that feeds into a rime splintering `cascade'. The rimer surface temperature may in fact be the more important factor and can remain between 265 and 270 K, even for cloud temperatures a few degrees colder [Heymsfield 1984].

c. Also, there is no mention of the recent work on ice-ice collisions, and its parameterization by Phillips et al. (Phillips et al., JAS, 2017 and 2017) or for shattering of freezing drops ((Phillips et al., JAS, 2018). How do their parameterizations compare to those used here, and how might that influence differences in the effects upon precipitation?

Thank you for pointing out this oversight. As noted above (in response to point 3a), we have noted the existing breakup parameterizations in an expanded Section 2.2. To the end of this section, we note the differences in the importance of temperature, as our formulation depends directly and solely on temperature: *We expect a strong influence of temperature from our breakup tendency* $(\delta N_{ice}/\delta t)_{BR}$ *than was discussed in Phillips et al. 2017b, given the direct and sole dependence in Equation 5.*

4. It is stated that Crosier et al. noted fall streaks at cloud top in the radar measurements. This would seem to imply a seeding mechanism of ice from above that could also have fallen to the observation level of the aircraft, unless this has somehow been ruled out?

Thank you for this suggestion. Generally, an $N_{i,pri}$ peak at an altitude and an N_{ice} peak at a slightly lower altitudinal band should be a signature of seeding. The new Figure 7 (old Figure 4) could address this somewhat, but to be more thorough we have included two supplemental figures that show the vertical profiles of N_{ice} and $N_{i,pri}$ over time. For example, three locations near CFARR are shown in panels a, b, and c here with the darker colors corresponding to earlier times and lighter ones to later times:

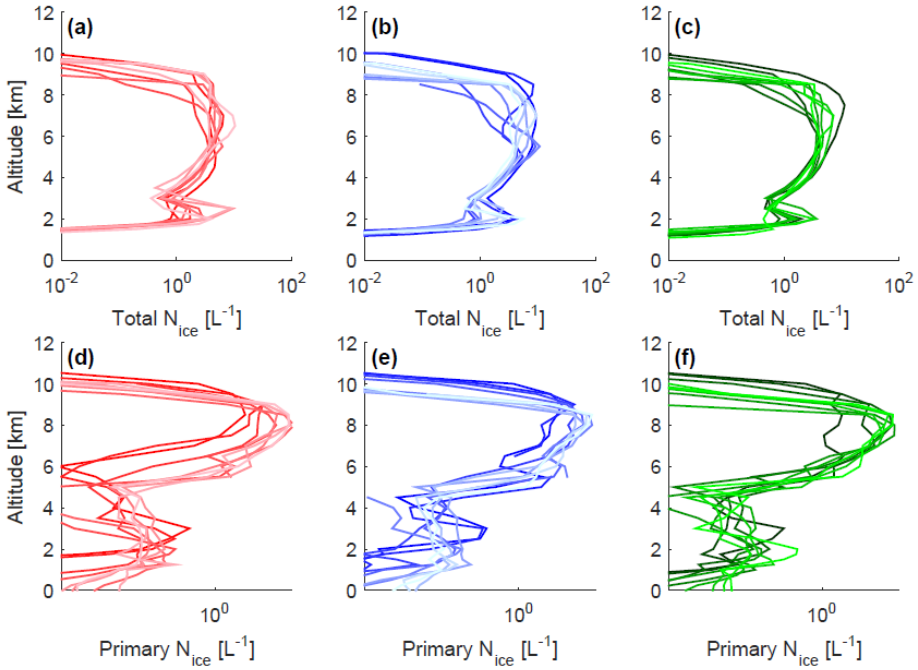

Figure S3 Temporal evolution of N_{ice} (panels a, b, c) and $N_{i,pri}$ (panels d, e, f) profiles in the RS1 simulation from three, randomly-sampled latitude / longitude locations in the vicinty of CFARR. Eight profiles are shown for each location, one for each half hour from 18:00 UTC to 21:30 UTC with the darker colors representing earlier times and the lighter ones later times.

Around cloud top, we do see that the peak in N_{ice} extends below that in $N_{i,pri}$, but these temperatures are too cold for droplets to exist, so this should just be the signature of ice sedimentation. There are secondary peaks in both N_{ice} and $N_{i,pri}$ at lower altitude, but both also exhibit a "pinch point" above this, i.e. there is not consistently a nucleation source about the low-level N_{ice} peak. This low-level peak should mostly be due to secondary ice production.

We summarize this by stating the following in the new section on the Dynamic environment: "*these Z_{DH} fall streaks, as well as those in differential reflectivity (shown in Crosier et al. 2014, their Fig. 5c) are signatures of ice crystal sedimentation and aggregation near cloud top. Ice crystal seeding may also be occurring with lower-level sedimentation, but the altitudinal peak in $N_{i,pri}$ does not fall consistently above that in N_{ice} (Figs. S3 and S4) so that secondary ice production must generate a portion of this low-level ice.*"

5. I would contend that most "larger-scale models" do not have two-moment microphysical schemes, so that the suggestion at lines 21-24 on page 10 are not practical.

Yes, you are right; within the CMIP5 models, the majority still use single-moment cloud microphysics schemes. But two-moment schemes generally lead to better representation of factors like greenhouse gas warming trends, snowfall intensity, and stratiform precipitation extent [Ekman 2014, Molthan and

Colle 2012, Morrison et al. 2009]. The IPCC has summarized its Fifth Assessment Report by saying that "climate models are incorporating more of the relevant aerosol-cloud interaction processes than at the time of AR4, but confidence in the representation of these processes remains weak." We take these results and trends to indicate that future model development will favor the incorporation of two-moment schemes. Indeed, given the uncertainties still associated with secondary ice production parameterizations, their model inclusion should generally come after the shift to a two-moment scheme.

Nevertheless we qualify our recommendation and note, as you mention above, that it may be more helpful initially to identify thermodynamic regimes under which these processes need to be included:

As meso- and large-scale models transition toward two-moment cloud schemes, secondary ice production could be included in parameterizations with criteria for the number concentration of large hydrometeors: the droplet shattering and collisional breakup parameterizations are activated only for those cloudy grid cells with more than a threshold concentration of large droplets and graupel, hail, or snow respectively. In one-moment schemes, parameterizations on the basis of favorable thermodynamic regimes will be more useful for the time being.

6. Comparison of observations and modeling: unless the authors can justify that the simulated movement of the rainband, and its dynamical nature, was very similar to that observed, it would make the comparisons shown in Fig. 4 not very useful. (If the comparison is not "fair", it could even be the case that the model IS producing sufficient secondary ice, for example, if the dynamical / thermodynamical conditions of the simulated clouds in that region and over that time are different than observed!) The description of the analysis for panels b and c on page 12 is helpful, but the reader cannot see what is being compared (which is why some comparison of the observed radar evolution to the simulated one is needed early on in the paper.) The temperature ranges also need to be specified in Fig. 4, too.

This is an important and difficult point that you raise again. We have provided a description of the simulated dynamics in our response to your first comment and see that the embedded updrafts, the preceding surface winds, and the reflectivities within the rainband are all weaker than in the observations. But it is also the case that the old Figure 4 (now Figure 6) intended to show how important the secondary production is relative to primary nucleation in the model. We were considering the simulated fields only because there is no way to make an analogous comparison with the available data. However, we still note in the rewritten discussion that "*underestimation of the updrafts within the rain band core (Sec. 4.1) will lead to errors that offset each other somewhat: too few raindrops will form when the vertical velocity and supersaturation are too low, but these will also be lofted more quickly through the altitudinal band where rime splintering is favoarble, leaving less time for collisions to occur.*"

In general, this kind of "dynamical buffering" for secondary ice production (i.e. that they require large hydrometeors but that also sufficient time in an appropriate temperature zone) should give some "dynamical resilience" to the simulations. In discussion of the altitudinal profiles of N_i, we also point out that "*The underestimated updrafts and radar reflectivities noted above in Section 4.1 may also help explain the too low N_i around 2 km: larger vertical velocities could loft graupel to high altitudes and boost the contribution from collisional breakup.*

To aid with visualization of what is shown in the new Figure 7 (altitudinal profile, histogram, and time series of N_{ice}), we have made the outline of the subdomain surrounding CFARR more prominent in Figures 5 and 6 and mentioned explicitly that we are drawing from this subdomain in the figure caption.

We have also switched Figure 6 to an altitudinal slice rather than a pressure level. We indicate in the caption now that "*The median temperature at this altitude is 255 K with a minimum value of 245 K and a maximum value of 267 K*."

7. Ice production rates: it needs to be stated more clearly how these derived from both the observations and the modeling results. Right now, it is unconvincing that this comparison is valid. Also, it should be stated somewhere that the CIP-15 observations were corrected using an algorithm designed to remove shattering artifacts, but some still likely remain because anti-shattering tips were not used on the instrument, as stated by Crosier et al.

   Yes, thank you for pointing out that the details of this calculation were missing. We had used a centered finite difference:

$$\left.\frac{dN_{ice}}{dt}\right|_{t_i} = \frac{N_{ice}\,(t_{i+1}) - N_i(t_{i-1})}{t_{i+1} - t_{i-1}}$$

   But we agree that this analysis is not as informative as focusing on the temperature-dynamic regimes where ice production is largest. For that reason, we remove the ice production rate figure.

   Within what is now Section 4.2 on Ice production rates, we also describe the CIP-15 measurements more thoroughly:

   IAT algorithms, with those particles below a threshold IAT of 10^{-4} s classified as artifacts, were used to correct the ICNCs [Field et al. 2006b]. No shatter-resistant tips were used on the probe, but given the strict IAT threshold, the possibility of artifacts is limited.

8. Qualifiers / limitations: need to be stated clearly throughout the paper. For example:

a. The model uses the primary ice nucleation parameterization of Phillips et al. (2008). Since INP measurements were not collected, it is unknown if this is an accurate representation, or not, and this might greatly affect the ratios of secondary to primary nucleated ice, including the possible importance of secondary ice to precipitation.

   Yes, this is a good point to include. In Section 3 on the Simulation setup, we add

   Previous studies have noted that limited nucleating efficiencies in the PDA08 may lead to underestimation of ICNC at mixed-phase conditions [Barahona et al. 2010, Curry and Khvorostyanov 2012, Morales-Betancourt et al. 2012]. No ice nucleating particle (INP) measurements were made during this case study, but from other observational datasets, PDA08 still yields better agreement with in-situ ICNCs than purely lab-based or theoretical parameterizations [Sullivan et al. 2015].

   Here it is also important to point out that Crosier et al. have noted the potential for large contributions of homogeneous nucleation in the convective regions. We point out right after the comment above that

   Crosier et al. also note that the low cloud top temperatures and strong updrafts in convective regions generate supersaturations that could favor large ice production from homogeneous nucleation. While not observationally confirmed, these conditions could buffer the ice nucleation tendency to our choice of parameterization.

   In what is now Section 4.2 on ice production, we also add the following caveat:

   The magnitude of these values is subject to uncertainty from the nucleation parameterization, which, as noted above, has underestimated INP numbers in previous studies.

b. To show an appreciable effect on surface precipitation (20% increase), rime splintering had to be increase over that typically depicted in models based on the laboratory measurements (e.g. Cotton et al. 1986 parameterization).

We have only left one test with an extended rime splintering where rime splintering is permitted to happen for 2 K below the typical threshold of -8°C. Earlier we had also made the error of showing the precipitation accumulation at 18:00 UTC ("as the rainband begins to pass over the UK" stated in the caption). We now look at the accumulation from the final model output file (23:30 UTC) after complete rainband passage and see accumulation differences on the order of 10 mm. Precipitation intensity differences remain about ±5 mm h^{-1} for a maximum simulated intensity of 30 mm h^{-1} (the 20% you cite) for the adjusted simulations.

c. Reasons for why the other two secondary ice processes might be less important here: (i) minimal graupel, which is important for ice-ice collisional breakup; (ii) limited number of raindrops? (Not sure what else would have limited that process here, but it would be good to know.) Also, it should be explained that the Crosier et al. study noted limited graupel in their observations, and I don't think they found any evidence of shattered frozen raindrops, but they clearly state that the former could have been limited by the inability of the aircraft to fly in the more convective regions.
Thank you for your thoughts. Building off these, we have included the spatial fields of graupel and snow mixing ratios and rain drop number concentrations in a new supplemental Figure:

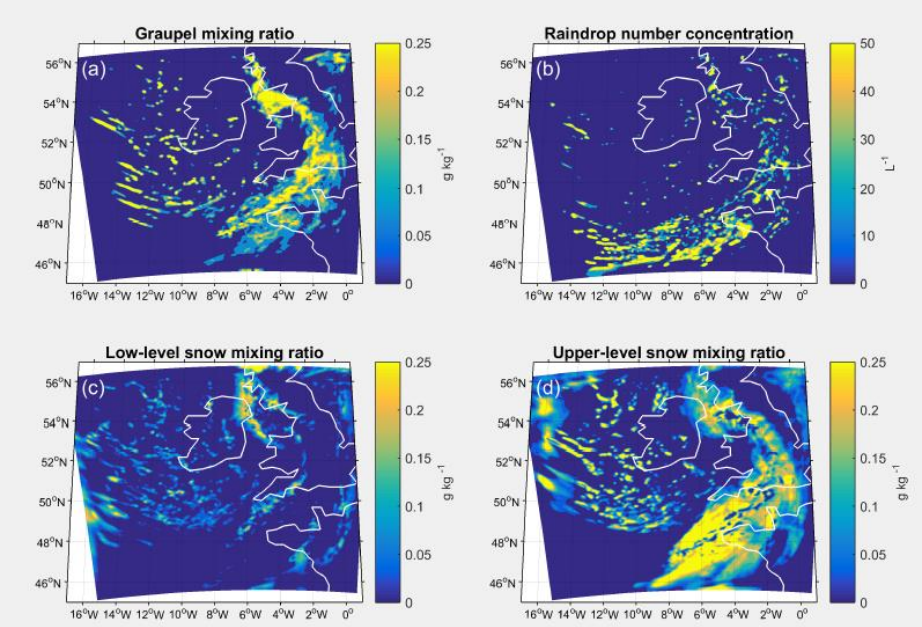

Figure S5 Graupel mixing ratio (a), snow mixing ratio (b), large-scale graupel quantity (c), and rain drop number concentration (d) in the simulation domain at 18:00 UTC for the RS2 simulation.

From these, we agree that low graupel and raindrop concentrations at the appropriate altitudes are limiting the $N_{i,sec}$ from collisional breakup and droplet freezing. To Section 4.2 we add the following:
If graupel were present at higher altitudes, $N_{i,sec}$ from breakup could increase significantly, as both the snow mixing ratio and fragment number parameter increase at colder temperatures. Similarly very limited raindrop number concentrations exist at the altitude where shattering probability is non-negligible. This importance of large hydrometeor number concentration for $N_{i,sec}$ suggests a means of parameterizing secondary ice production as meso- and large-scale models transition toward two-moment cloud schemes: the droplet shattering and collisional breakup parameterizations could be activated only for those cloudy grid cells with more than a threshold concentration of large droplets and graupel, hail, or snow respectively. In one-moment schemes, parameterizations on the basis of favorable thermodynamic regimes will be more useful for the time being.

d. The enhancement of the updrafts and precipitation, and downdrafts, mentioned on page 14 needs to be backed up with some evidence.

We have deleted this statement because we are not showing a feedback of the ice production on dynamics yet. With the updraft field in the new Figure 3 we can only state that the regions of highest precipitation intensity (and precipitation intensity deviation) occur in the regions of highest vertical velocity.

e. Some discussion of the model resolution effects should be included, for both properly simulating the microphysics as well as the dynamics.
Within the discussion section, we have added discussion on the potential impact of resolution and the particular microphysics scheme on our results:

The choices of spatiotemporal resolution and microphysics scheme are particularly important for convective clouds because of the fine structure of precipitation and strong feedbacks between hydrometeor formation, latent heating, and cloud dynamics. Previous studies have generally found a spatial resolution of 4 to 6 km to be sufficient to reproduce precipitation extremes [e.g., Prein et al. 2013, Pieri et al. 2015]. This resolution dependence results from changes in the vertical moisture advection, in turn due to adjustments of vertical velocity with resolution [Yang et al. 2014]. For simulations whose resolutions border on the ``gray zone" scales (around a tenth of a degree), over-representation of convective activity is possible by both the parameterization and explicit resolution [Pieri 2015]; however, our simulations are at a fully convection-permitting scale and use only a reduced form of the Tiedtke mass-flux scheme for shallow convection [Tiedtke 1989] and this should not be a concern.

The use of a two-moment scheme is also important for simulation of extreme precipitation [Otkin 2006]. Certain one-moment schemes tend to generate overly large drops and too high precipitation rates [Thompson 2004], but SB06 tends to produce especially large quantities of graupel [Otkin et al. 2006]. The Bigg parameterization, as a precursor to our droplet shattering additions, has been shown in previous studies to predict very low numbers of frozen drops [e.g., Morrison et al. 2005, Fan et al. 2009], which may contribute to underestimate of secondary ice here. The more rigorous alternative would be to account for immersed surface area and scavenging of ice nucleating particles as in [Paukert and Hoose 2014, Paukert et al. 2017], and future work should implement both an updated immersion freezing and secondary ice parameterizations.

Technical Corrections:

1. Table 2 and Fig 2 are confusing, since most of these runs are not discussed in the paper. I would suggest only showing those that are discussed here, and just noting somewhere (if important) that other variations did not show much change in the results. Also, if the text could explain the naming convention for the different simulations used, that would be much easier for the reader to interpret them.
In part because no results were shown from several simulations and in part because some parameter values were too extreme, we have adjusted the simulations. These changes are listed in a new Table 2:

Rime splintering				Ice-ice collisional breakup			
RS1:	$\aleph_{RS} = 300,$	$w_{RS} = \text{TR}$		**BR1ig**:	*graupel_breakup_ice*		
					$F_{BR} = 180,$	$T_{min} = 256,$	$\gamma = 3$
RS2:	$\aleph_{RS} = 300,$	$w_{RS} = \text{UNI}$		**BR2ig**:	*graupel_breakup_ice*		
					$F_{BR} = 360,$	$T_{min} = 249,$	$\gamma = 5$
				BR2sg:	*graupel_breakup_snow*		
					$F_{BR} = 360,$	$T_{min} = 249,$	$\gamma = 5$
Droplet shattering				**Combinations**			
DS1:	$\aleph_{DS} = 2,$	$p_{max} = 5\%,$	$\sigma = 3$	**RS2 + BR2ig**:	$\aleph_{RS} = 300,$	$w_{RS} = \text{UNI}$	
					graupel_breakup_ice		
					$F_{BR} = 360,$	$T_{min} = 249,$	$\gamma = 5$
DS2:	$\aleph_{DS} = 10,$	$p_{max} = 10\%,$	$\sigma = 5$	**DS2 + BR2ig**:	$\aleph_{DS} = 10,$	$p_{max} = 10\%,$	$\sigma = 5$
					graupel_breakup_ice		
					$F_{BR} = 360,$	$T_{min} = 249,$	$\gamma = 5$
				RS2 + DS2:	$\aleph_{RS} = 300,$	$w_{RS} = \text{UNI}$	
Control					$\aleph_{DS} = 10,$	$p_{max} = 10\%,$	$\sigma = 5$
CTRL:	$\aleph_* = 0$	$F_{BR} = 0$		**ALL**:	$\aleph_{RS} = 300,$	$w_{RS} = \text{UNI}$	
					$F_{BR} = 360,$	$T_{min} = 249,$	$\gamma = 5$
					*graupel_breakup_**		
					$\aleph_{DS} = 10,$	$p_{max} = 10\%,$	$\sigma = 7$

The previous naming schemes were based on the idea of four levels of experiments – **C**onservative, **M**oderate, **N**arrow, and **G**enerous – but this was not clear since there was no description in the manuscript. The idea had been to have a fractional factorial experimental design that probed the whole parameter space without an expensive, one-at-a-time approach. In the new scheme shown above, we have only defined two "levels" for each process to avoid the most extreme values we had used earlier, particularly for rime splintering and droplet shattering.

We have also condensed Figure 2 for the new simulation layout:

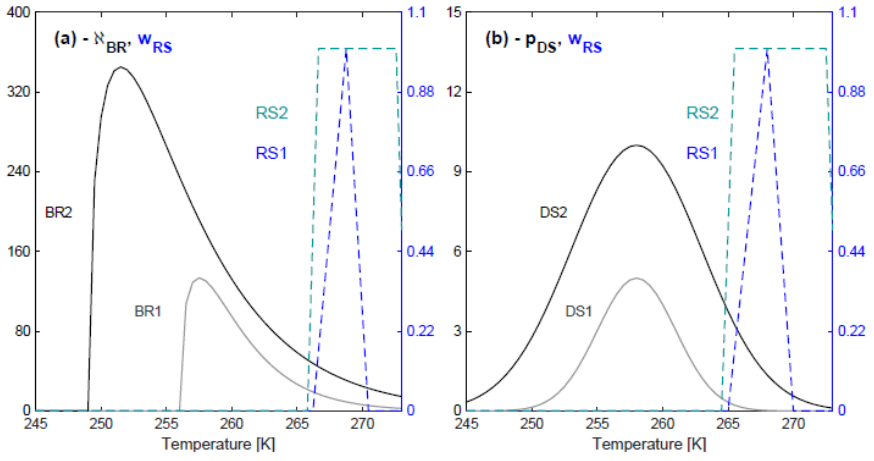

Figure 2 Fragment numbers, weightings, and probabilities from the secondary ice production parameterizations. In panel a, we show N_{BR} from both ice-ice collisional breakup simulations (BR1 and BR2) as well as the triangular and uniform $w_{RS}(T)$. In panel b, we show p_{DS} from both droplet shattering simulations (DS1 and DS2) and w_{RS} once again. The overlapping temperature regions of these are particularly important to understand any feedback between the processes.

If the major issues in this review are addressed, I suspect that much of the wording will be revised, and thus I will refrain from noting any suggestions at this time.