Responses to the reviewers

Significance of microphysical processes for uncertainties in ensemble forecasts of summertime convection over central Europe

by Christian Barthlott, Beata Czajka, Christoph Gebhardt, and Corinna Hoose March 17, 2026

We thank both reviewers for reading the manuscript and providing detailed comments. We have carefully considered all comments and changed the manuscript accordingly. Please find below our responses in blue.

Reviewer 1

Overall Evaluation

The manuscript, “Significance of microphysical processes for uncertainties in ensemble forecasts of summertime convection over central Europe,” investigates the impact of perturbing selected microphysical parameters in the two-moment bulk microphysics scheme of the ICON model, using a 108-member ensemble for 4 summertime convective cases. The methodology and statistical analyses are generally well presented. Figures 5, 10, and 11 are particularly creative and effectively illustrate the distinct effects of each ensemble perturbation, accompanied by clear discussions. Moreover, the authors convincingly demonstrate the importance of microphysical perturbations by generating an ensemble spread that compares with the operational ensemble results. Consequently, this study provides valuable insights into microphysics-related ensemble design and should be of interest to the broader forecasting and microphysics communities. However, some scientific clarifications and physical interpretations are needed. In its current form, the study feels incomplete, and I therefore recommend major revision, contingent upon addressing the comments below.

Major comments:

1. Line 48-54, Introduction, Although the methods differ, Thompson et al. (2021) could be mentioned here because they perturbed similar parameters, including the cloud droplet size distribution, CCN activation, and graupel size spectra within a single microphysics scheme. In addition, Barthlott et al. (2022a, b) are cited repeatedly, so the authors may consider introducing an abbreviation after the first occurrence.
The reference to Thompson et al. (2021) fits very well here and has been included. In places where references to the Barthlott et al. (2022a, b) papers are not essential, we have removed them and now believe that an abbreviation is no longer necessary.

2. Figure 7 shows that the precipitation rate reaches its maximum near the end of the simulation period for Cases 1, 3, and 4. In other words, the rainfall events have not yet completed. Thus, how can the perturbed microphysical parameters fully represent their total impact on the statistical analyses if the latter part of the event evolution is not captured?
We agree with the reviewer that in three cases the rainfall event did not come to an end. In two of these cases (cases 1 and 4), however, there is a morning/midday precipitation peak, which is covered in our 24-hour runs. In each case, the heavier precipitation occurs in the evening. An important point is that the maximum of the event is always captured in our 24-hour runs and the precipitation rates are already declining at the end of the integration time. One reason why we only simulated 24 hours is the fact that we did not want to have a very long lead time. In operational forecasts, new (ensemble) runs are started every one to three hours to keep the integration times relatively short. With 24 hours, we tried to find a compromise between

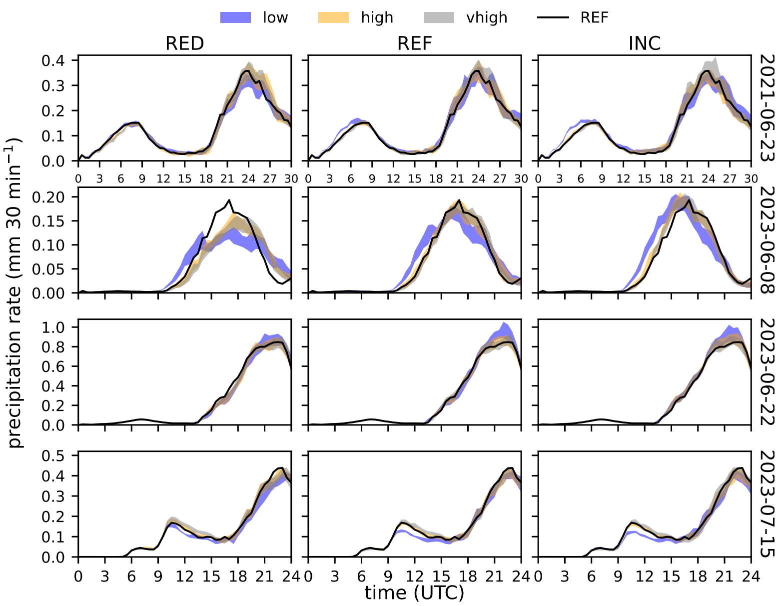

Figure R.1: Domain-averaged precipitation rates for reduced (left), reference (middle), and increased graupel sedimentation (right). Top row is for 30-h integration time.

not deviating too much from reality and taking into account at least the maximum convective activity.

To be able to access the impacts of not covering the entire convective event throughout the night, we redid all 108 simulations with a longer integration time of 30 h for case 1 (23 June 2021, Fig. R.1). Unfortunately, the compilers used for these simulations are not available anymore on our HPC system. So the re-compilation of the model code with new compilers could also lead to slightly different outcomes. It turns out that although the same compilers are not available anymore, the sensitivities to the applied perturbations are not changed. In Fig. R.2, the general precipitation characteristics with respect to our microphysical perturbations remain the same even if individual points are not identical.

We also recalculated the ratio of the cold to warm rain processes with the 30-h simulation data and found an almost identical characteristic behaviour (Fig. R.3). As the sensitivities to our microphysical perturbations are robust and do not change for the longer integration time for case 1, we decided not to redo the other three cases, which would require another 130k CPU h for the full ensemble. In section 3.2 of the revised manuscript, we now mention that a longer simulation time, including the decaying stage of the convective events, did not change the sensitivity to our perturbations.

3. Line 302, Eq. 5, What does the symbol x represent? Also, why is the forecast error quantified relative to a reference simulation rather than to reanalysis or observations? Does the reference simulation objectively show the best performance? A justification is needed.
Thanks for pointing that out. $\mathbf{x}$ represents the position vector in a Cartesian coordinate system with components along the x, y, and z axes. This information has been added to the manuscript.

The difference total energy is computed between 2 (or more) simulations to illustrate the growth of the difference field over time. Although it is named tropospheric error growth, this variable is more related to errors/deviations induced by our perturbations and not to errors with respect to observations or reanalyses. In the original article by Zhang et al. (2003) or the work of Judt et al. (2018), DTE is computed using simulations only.

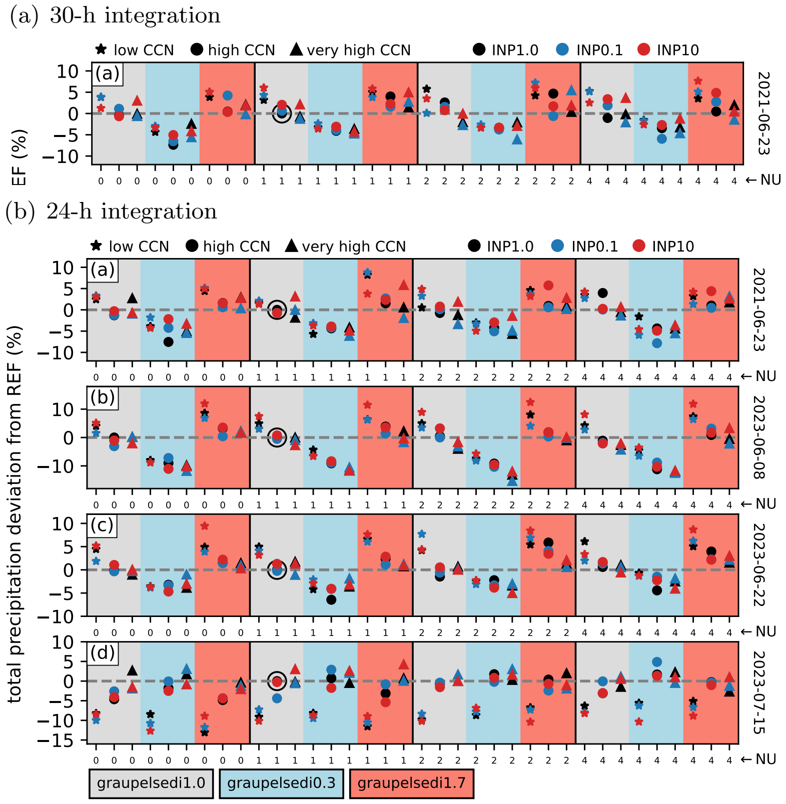

Figure R.2: (a) Precipitation deviation from the respective reference run (marked with a black circle) for case 1 with 30-h integration time and (b) all cases with 24-h integration time.

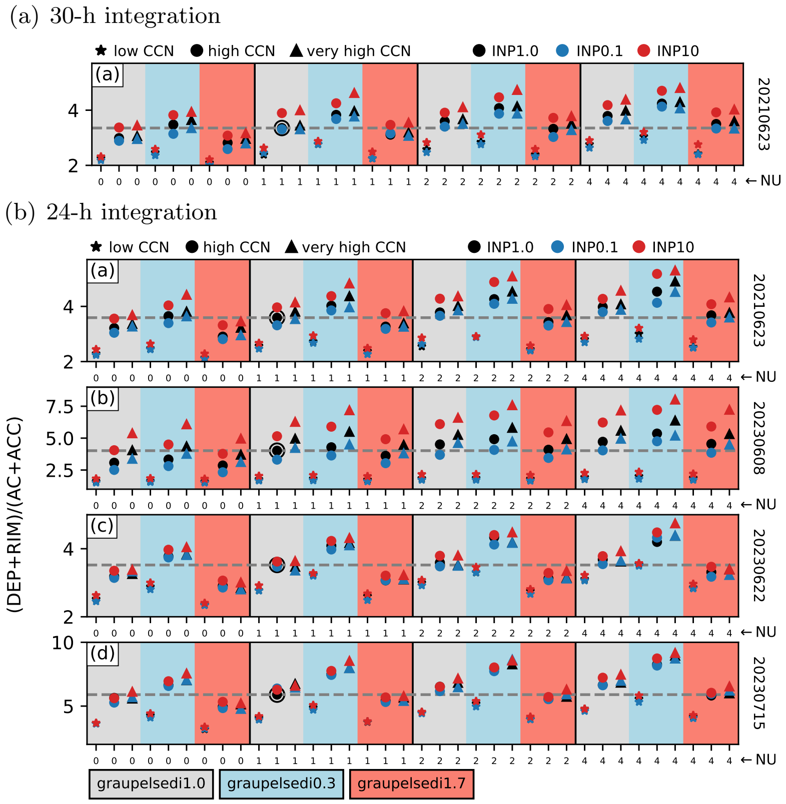

Figure R.3: Ratio of cold-rain formation (deposition DEP and riming RIM) to warm-rain formation (autoconversion AC and accretion ACC).

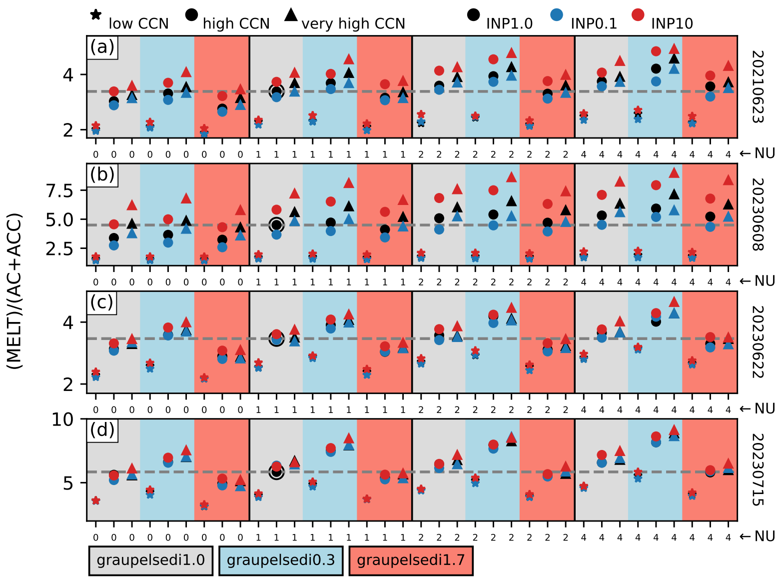

Figure R.4: Ratio of cold-rain formation (melting MELT) to warm-rain formation (autoconversion AC and accretion ACC).

Judt, F., 2018: Insights into Atmospheric Predictability through Global Convection-Permitting Model Simulations. J. Atmos. Sci., 75, 1477–1497, https://doi.org/10.1175/JAS-D-17-0343.1.

4. Figure 10 is very useful for illustrating the relative importance of the different microphysical processes in terms of magnitude and frequency. Because graupel fall speed directly influences surface precipitation through sedimentation, would it be possible to include the hydrometeor sedimentation rate in the analysis additionally?
Thanks for this suggestion. Unfortunately, this is not possible as the sedimentation velocity is not a model output variable. Moreover, as Figure 10 contains phase changes or shifts between hydrometeor classes, the inclusion of sedimentation velocity with the unit of speed would not be comparable in this figure.

5. Figure 11, The ratio of cold-rain to warm-rain production is an important diagnostic. Since melting of frozen hydrometeors is the most direct pathway to cold-rain formation, it would be more appropriate to use melting, rather than DEP + RIM, to compute this ratio.
We agree with the reviewer that melting could be used instead of the sum of deposition and riming. However, the results are almost identical, and we see the same systematic behaviour with only slightly different values (Fig. R.4). We already used this metric using deposition and riming in several recent papers:

   (a) Barthlott, C. and C. Hoose (2018): Aerosol effects on clouds and precipitation over central Europe in different weather regimes, J. Atmos. Sci. 75, 4247-4264, DOI:10.1175/JAS-D-18-0110.1

   (b) Schneider, L., C. Barthlott, C. Hoose, and A.I. Barrett (2019): Relative impact of aerosol, soil moisture and orography perturbations on deep convection, Atmos. Chem. Phys., 19, 12343-12359, DOI:10.5194/acp-19-12343-2019

   (c) Baur, F., C. Keil, and C. Barthlott (2022): Combined effects of soil moisture and microphysical perturbations on convective clouds and precipitation for a locally forced case over Central Europe, Q. J. R. Meteorol. Soc. 148, 2132-2146, DOI:10.1002/qj.4295

   (d) Barthlott, C., A. Zarboo, T. Matsunobu and C. Keil (2022a): Importance of aerosols and

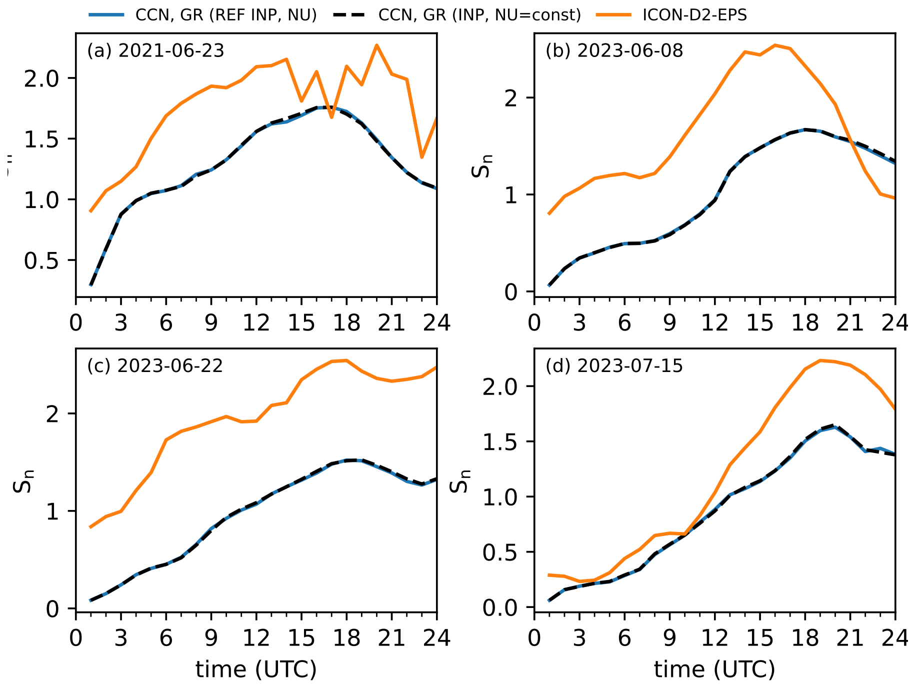

Figure R.5: Domain-averaged normalized ensemble spread S_n based on hourly precipitation amounts for a 9-member ensemble only perturbing CCN concentrations and graupel sedimentation (blue) for reference INP and NU, averaged 9-member ensemble only perturbing CCN concentrations and graupel sedimentation and constant INP and NU (black dashed), and the operational single-moment microphysics ensemble (ICON-D2-EPS, orange).

shape of the cloud droplet size distribution for convective clouds and precipitation, Atmos. Chem. Phys., 22, 2153-2172, DOI:10.5194/acp-22-2153-2022

(e) Barthlott, C., A. Zarboo, T. Matsunobu and C. Keil (2022b): Impacts of combined microphysical and land-surface uncertainties on convective clouds and precipitation in different weather regimes, Atmos. Chem. Phys., 22, 10841-10860, DOI:10.5194/acp-22-10841-2022

(f) Lucas, L., C. Barthlott, C. Hoose, P. Knippertz (2025): Aerosol effects on convective storms under pseudo-global warming conditions: insights from case studies in Germany, Atmos. Chem. Phys., 25, 18527-18548, DOI:10.5194/acp-25-18527-2025

Thus, we would like to keep it in that form. However, we added a statement in the text that the use of melting provides similar results and the same statistics.

6. Figure 12, CCN concentration and graupel fall speed appear to be the two most influential parameters. Therefore, it would be valuable to compute the ensemble spread using only the corresponding subsets of ensemble members. This would allow an assessment of whether a high-impact parameter represented by a small subset contributes more to the overall spread than a low-impact parameter represented by a larger subset, thereby supporting a clearer conclusion.
Thanks for this suggestion. We computed the ensemble spread in two additional ways: (i) all members where CCN and graupel sedimentation are perturbed with reference INP and NU (9 members) and (ii) averages of 12 subensembles with 9 members each in which CCN and GR are perturbed with constant combinations of INP and NU (Fig. R.5).

Even though CCN and graupel sedimentation have the strongest impact on total precipitation amounts, the spread is mostly smaller than in the operational ICON-D2-EPS due to the small ensemble size. It makes no difference whether only the reference values of INP and NU are used or whether the average of all combinations of constant INP and NU is taken. We therefore conclude that just perturbing CCN and GR with only 3 different perturbed values each are not sufficient to produce enough spread, and that either combinations with INP and shape parameter are necessary, or more perturbed values are necessary to reach a larger ensemble size. We added these statements in section 3.6 of the manuscript:

"It is also of interest to evaluate the ensemble spread based only on the parameters that have the strongest influence on precipitation amounts. In most of the cases analyzed, these parameters are the CCN concentration and the graupel sedimentation velocity (see Section 3.2). Using these parameters, we additionally calculated the ensemble spread using reference values for INP concentrations and shape parameters (9 members) and also averaged over 12 sub-ensembles of 9 members each, in which CCN and GR are perturbed with constant combinations of INP concentrations and shape parameters. Although CCN and graupel sedimentation have the strongest influence on total precipitation amounts, the spread is mostly smaller than in the operational ICON-D2-EPS (not shown) due to the smaller ensemble size. It makes no difference whether only the reference values of INP and NU are used or whether the average of all combinations of constant INP and NU is taken. We therefore conclude that the high-impact parameters represented by a small subset contribute less to the overall spread than a combination with lower-impact parameters using a larger subset."

Minor comments:

1. Line 101, INP uncertainty, what is Tmin in Eq. (1)? Is it necessary to list the full formulations of Eqs. (1)–(3)? Please make a check whether an upper-limit constraint is imposed by the prescribed ice-nucleating particle number concentration (i.e., using up available ice nuclei) to reduce the potential overproduction of ice crystals and thereby limit the impact of INP perturbations.
   T_{min} is 237.15 K, but there is also an upper limit as this parameterisation must only be applied within a specific temperature range to prevent unrealistically high concentrations at very cold temperatures (Hande et al. 2015). However, we agree with the reviewer that equations 1-3 are not necessary here as they are provided in the referenced literature. We removed them and adapted the text accordingly.

   The heterogeneous ice nucleation is active between 237.15 K and 261.15 K with an upper limit of potential INP concentrations at 10^5 m^{-3}. In a very narrow temperature range between the lower limit and 240.3 K (range of 3.15 K), does the increased INP concentration exceed the INP limit. We therefore believe that this does not affect the results in a significant way.

2. Line 128, Eq. 4, A and λ denote the intercept and slope parameters, respectively. Also, there is one typo of v' = 2, 5, 8, 14 in line 140.
   We added the information about the intercept and slope parameters, and the typo was corrected as well.

3. Figure 3, a minor typo of "radar-derived".
   We modified the caption to meet ACP's guidelines. Now Radar stands at the beginning of a sentence and should be written with a capital first letter.

4. Figure 4, consider adding the spatial correlation between the ensemble median and the observations shown in Fig. 3 to demonstrate how well the precipitation systems are captured.
   Thanks for this suggestion, but we believe that a spatial correlation between two fields only is not suitable, as a classic Pearson correlation is not defined pointwise for a 2D map. Instead, we calculated difference plots (RADAR - ensemble median) in Fig. R.6. Please note that the RADOLAN product is also erroneous, as Radar is not an instrument for surface precipitation in

a quantitative sense. Moreover, the 1-km RADOLAN data were interpolated on the 2-km ICON grid. As can be seen from the red and blue areas, some regions have more and other regions have

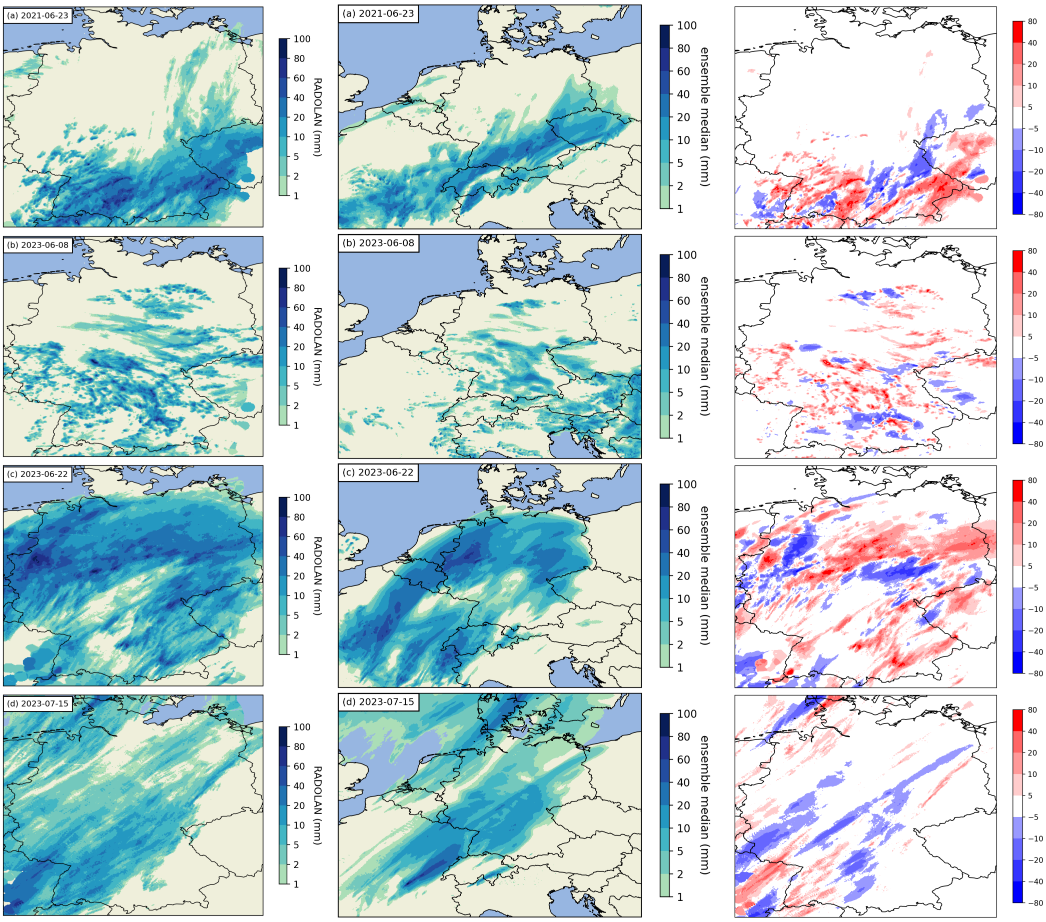

Figure R.6: Radar-derived total precipitation amounts (left), ensemble median (middle), and difference field (Radar-median, right).

less precipitation than in the Radar-derived product. However, we believe that the characteristics of the different convective scenarios on these days are generally well simulated by the ICON model, even if not all cells are captured at the right place or the right intensity. The location of precipitation and the sizes of convective areas generally agree with the observed precipitation, which means that there is a good basis for the sensitivity analyses using the microphysical perturbations. However, a systematic model evaluation or quantitative forecast improvement was not the goal of our study. As we already mentioned in the manuscript, the convective scenarios are well captured by the model, we believe that no further changes to the text or inclusion of these difference plots are necessary.

5. Line 253, the sentence "This may be overcompensated by the reduced riming efficiency associated with the smaller cloud droplets" is unclear. How are smaller cloud droplets related to enhanced graupel fall speed? Further clarification is needed.
We did not state that smaller cloud droplets lead to an enhanced graupel fall speed. This part is about the effects of our perturbations on the riming process: the increase of the graupel fall velocity is expected to enhance the riming growth of ice hydrometeors. When CCN or the shape parameter are perturbed, and cloud droplets become smaller, the riming efficiency decreases. So the riming increase due to higher graupel sedimentation can be compensated by a reduced riming efficiency if the CCN concentration or the shape parameter is increased. So we did not imply that smaller cloud droplets lead to an enhanced graupel fall speed. We slightly rephrased

that sentence in the manuscript to make that clearer, it now reads:

"A higher graupel fall velocity is expected to enhance the riming growth of ice hydrometeors, affecting total precipitation either in solid or melted form. If the size of cloud droplets is reduced either by higher CCN concentrations or larger shape parameters, the increased riming growth due to higher graupel fall velocities may be overcompensated by the reduced riming efficiency associated with smaller cloud droplets."