Response to Referee #2

Manuscript number: EGUSPHERE-2026-3542

Manuscript: Ice–ice collisional breakup enhances secondary ice production but reduces surface hailfall in a simulated pulse hailstorm

Dear Referee,

Thank you for reviewing our manuscript and for your specific suggestions. In response, we have added comparisons of observed and simulated storm evolution, statistics of hydrometeor regions, and diagnostics of hail growth and hailfall. We respond to your comments point by point below and provide the corresponding figures and revised wording.

Major comments

Comment 1

One major issue is that the study does not have enough comparison or validation of the simulation with observation. In section 3.1, time series of the observed and simulated properties of the storm should be shown. For example, these properties can include the maximum radar reflectivity in the domain, the averaged radar reflectivity in the storm core area, precipitation at the surface, and so on. This way we can clearly see the evolution of the storm and to what extent the model simulations can represent the observed storm. Figure 1 only shows two snapshots. It is good to use the snapshots to show the vertical and horizontal structure of the storm. Section 3.1 provides some vague discussion on how the observed radar echo is reproduced in the CTRL experiment and how the SIP_ALL_2 experiment produces minor changes in the overall echo morphology. But it is very hard to use these snapshots to evaluate how well the model simulations represent the observed storm. I would suggest using time series to show how well the seven simulations reproduce the observations. The time series can also help identify which simulation (or including which secondary ice process in the model) can lead to simulated storm properties that are the closest to the observations.

Response

Thank you for your suggestion. The two times shown in the original manuscript mainly illustrate hydrometeor stratification and echo structure and are insufficient to assess the full evolution of the storm. We have added comparisons of the evolution of maximum reflectivity, core-mean reflectivity, and the strong-echo fraction between the observations and all seven simulations (Fig. R1). To compare periods of similar duration, we use 04:46–05:19 UTC for the observations and 97–130 min, covering the developing and mature stages, for the simulations. The core is defined as the region with composite reflectivity of at least 35 dBZ; the strong-echo fraction is the fraction of this core with reflectivity of at least 45 dBZ. Continuous surface precipitation observations are currently unavailable for comparison.

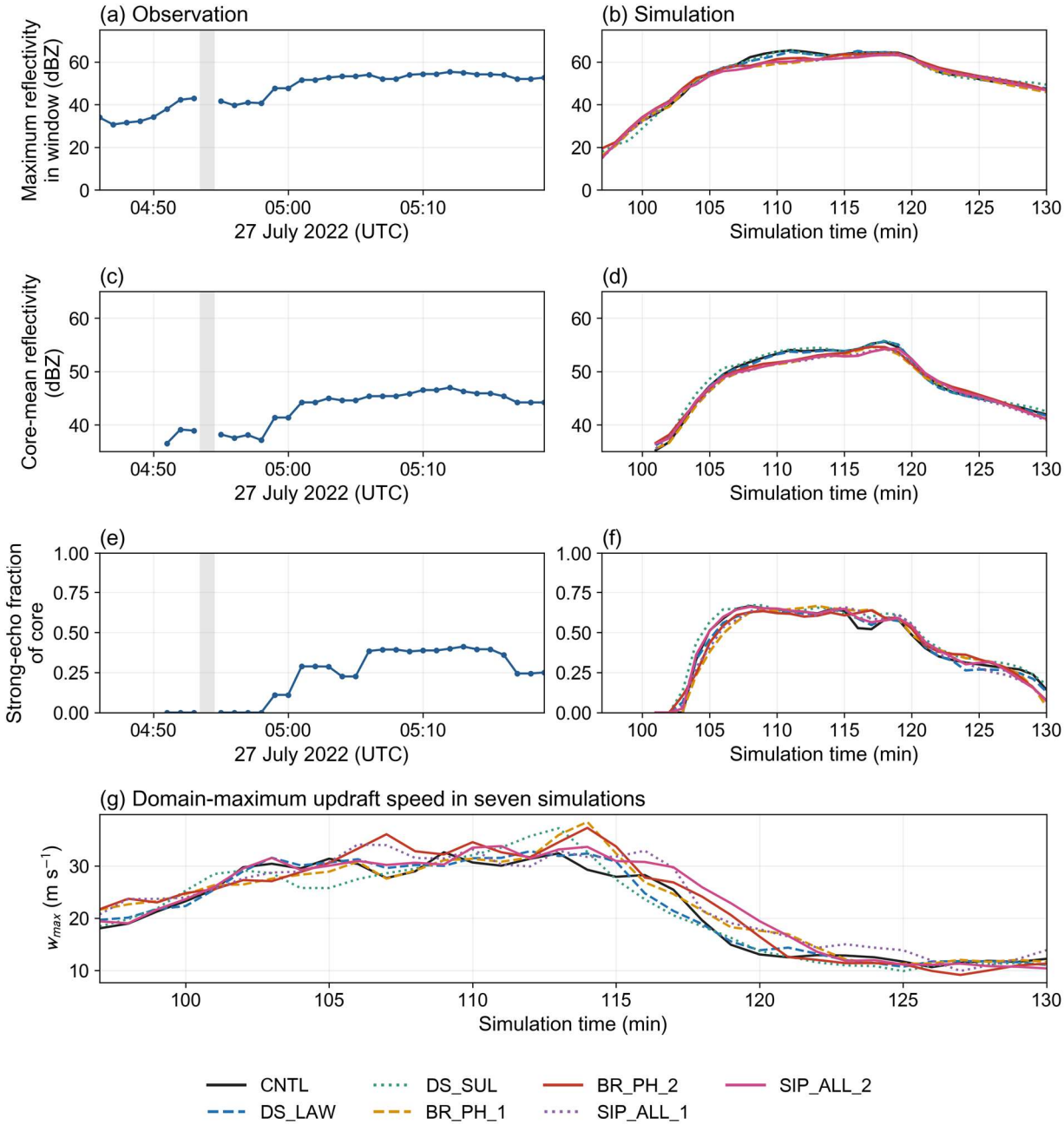


Figure R1. Observed and simulated storm evolution for the seven experiments. (a–f) Reflectivity, with observations in the left column and simulations in the right column. (g) Domain-maximum updraft speed in the seven experiments over the same simulation period.

During the period examined, the observed peak maximum reflectivity is 55.3 dBZ, compared with 63.1–65.4 dBZ in the seven simulations. The observed peak core-mean reflectivity is approximately 47.0 dBZ, compared with 54.2–55.8 dBZ in the simulations. Although the experiments including BR have lower peak maximum reflectivity, closer to the observed value, the strong-echo fraction shows no marked improvement. Peak maximum updraft speeds range from 32.6 to 38.4 m s^{-1} across the seven simulations, with substantial differences and no clear pattern. The seven experiments roughly capture the overall evolution of the echoes, but strictly speaking, the current results are insufficient to identify the experiment closest to the observations.

However, the simulations are idealized and are designed to examine the effects of SIP under the environmental conditions of this case, rather than to reproduce the observed storm in detail. The observational comparison assesses the basic characteristics of the simulated storm to establish whether it has the characteristics of a pulse hailstorm. We consider that differences between the simulated and observed storms are therefore expected. The differences among the seven experiments also show that the storm-scale effects of BR are not pronounced, but the earlier wording does need to be narrowed further.

In our previous study (Dong et al., 2024), Fig. 4 (shown here as Fig. R2) compared the observed and simulated storm characteristics and structures. Together with Fig. 5 (shown here as Fig. R3), it showed a high convection initiation height, rapid development, concentrated hail production, and rapid decay in CNTL, consistent with the basic characteristics of a pulse storm.

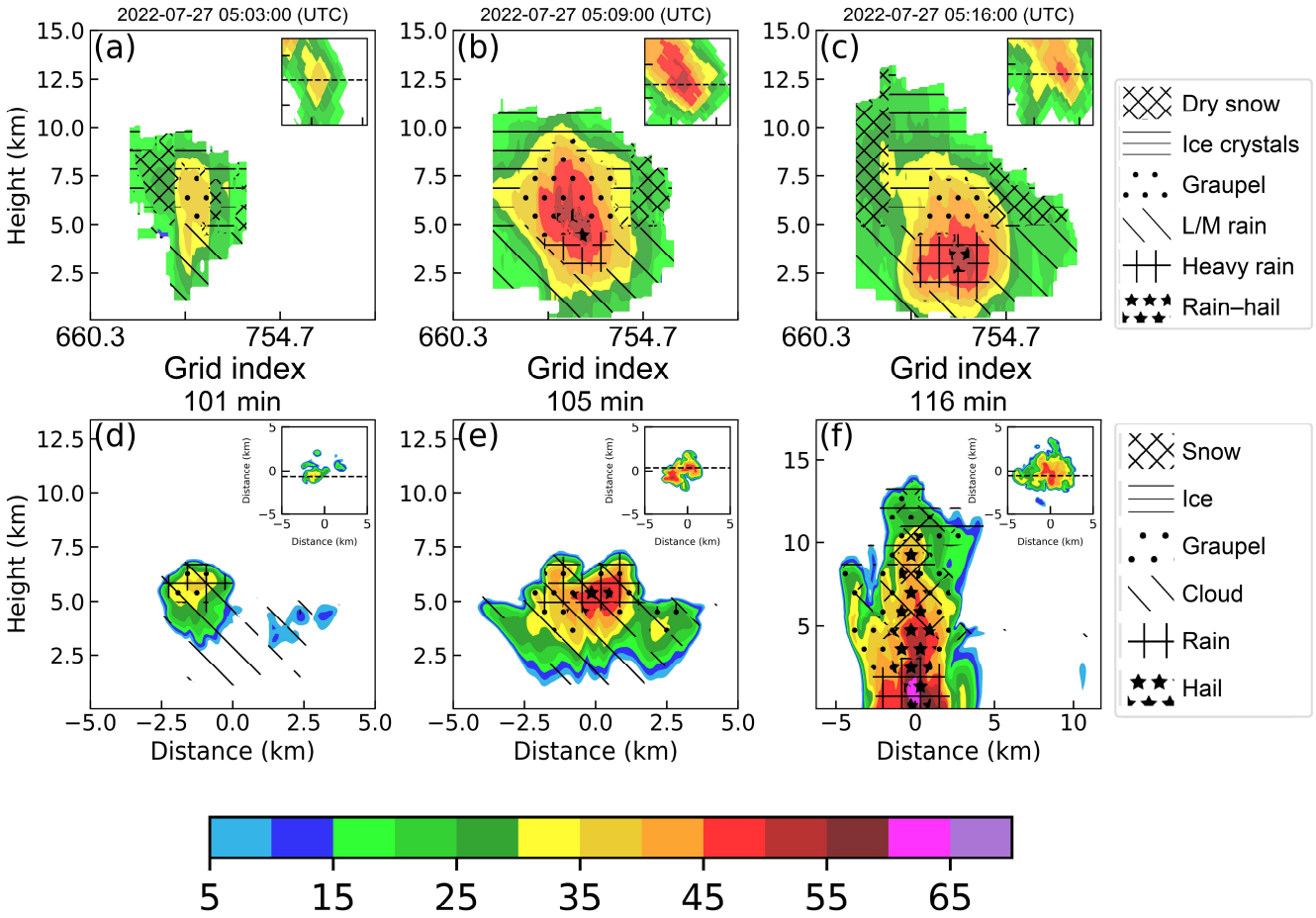


Figure R2. Storm structures at representative stages in the observations and CNTL. (a–c) Observations. (d–f) Simulations. Colors show reflectivity (dBZ), patterns indicate the hydrometeor classes listed in the legends, and dashed lines mark the cross-section locations.

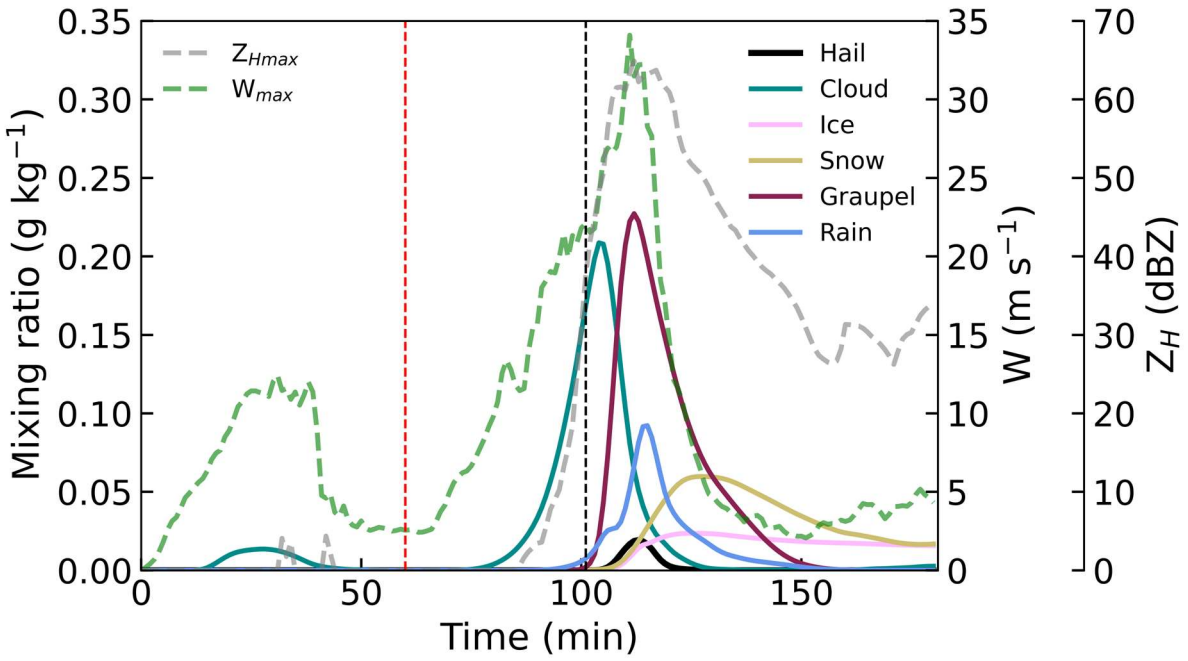


Figure R3. Evolution of hydrometeor mixing ratios, maximum vertical velocity, and maximum reflectivity in CNTL. Hydrometeor mixing ratios are spatial averages over a 10 km × 10 km × 12 km region centered on the simulation domain center. Solid lines of different colors indicate different hydrometeor classes; the gray and green dashed lines show maximum reflectivity and maximum vertical velocity, respectively. The red and black vertical dashed lines mark the increase in surface heat flux at 60 min and convective initiation at 101 min, respectively.

SIP_ALL_2 also shows this evolution (Fig. R4): maximum updraft speed peaks at 111 min, regional-mean hail mixing ratio at 113 min, and lowest-layer hail mixing ratio at 116 min, followed by a marked decrease. Thus, with BR included, the simulation retains rapid convective development and hail production and can be used to compare the effects of secondary ice processes on ice-phase evolution and hailfall in this storm.

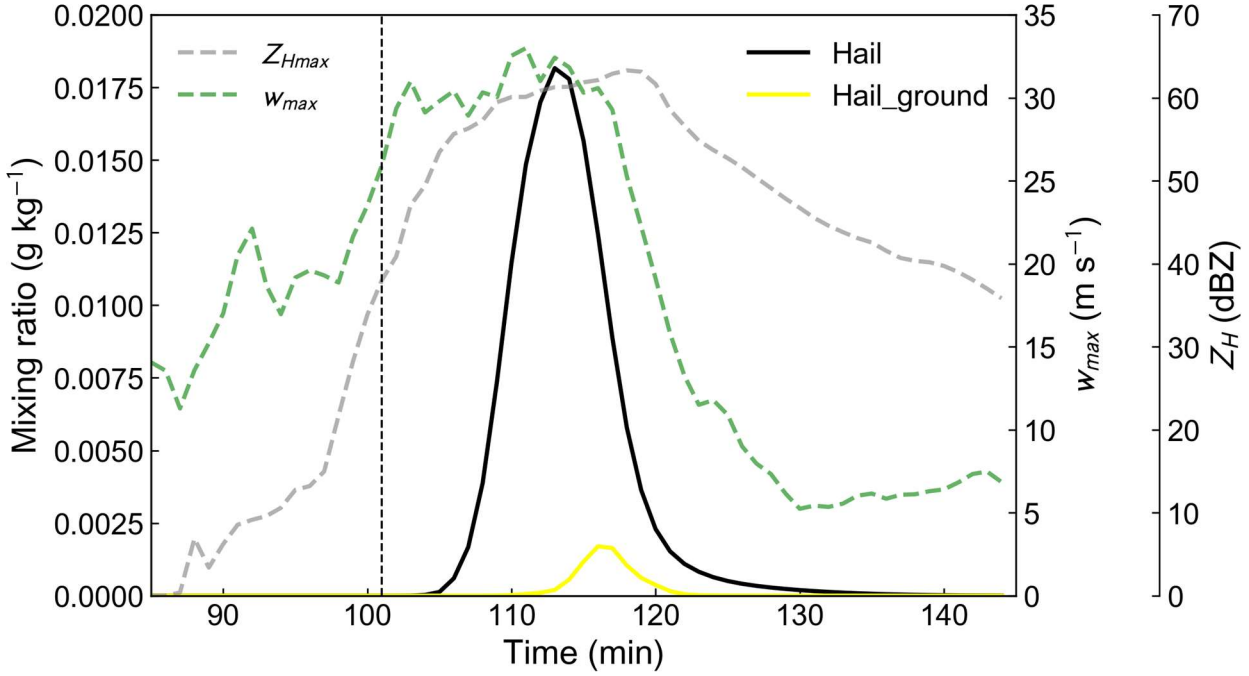


Figure R4. Evolution of hail mixing ratio, maximum updraft speed, and maximum reflectivity in SIP_ALL_2. The black and yellow solid lines show mean hail mixing ratio over the central region of approximately 10 km × 10 km and 0–12 km in height and mean hail mixing ratio in its lowest layer, respectively. The green and gray dashed lines show maximum updraft speed and maximum reflectivity, respectively.

Accordingly, the revised wording about validation in Sect. 3.1 is:

> *"Previous work showed that CNTL exhibits a high convection initiation height, rapid development, concentrated hail production, and rapid decay, characteristic of a pulse storm (Dong et al., 2024). At the two representative times shown in Fig. 1, the main altitude ranges of graupel and hail in CNTL broadly agree with the observations, but the early upper-level ice crystal region is too small. Relative to CNTL, SIP_ALL_2 has a more extensive ice crystal region at middle and upper levels, with small changes in the overall echo morphology. These comparisons show that the simulations have the main ice-phase stratification of this case. On this basis, we analyze the effects of secondary ice processes on hail growth and surface hailfall in this pulse-storm setting."*

The detailed time-series plots for all seven experiments do not provide clear validation results and do not affect the conclusions of this study. We therefore do not include them directly in the main text at this stage. If needed, we will add them to the main text or an appendix later.

Comment 2

> *Data from X-band radars can provide information on the hydrometeor types and also the spatial distribution of these hydrometeors. I think the manuscript should also compare the observed and simulated hydrometeor distributions. For example, in Figure 2b, cloud ice volume from the seven simulations are shown. (I suppose that cloud ice volume is the air volume that contains cloud ice?) I think the cloud ice volume from the observation should also be shown in Figure 2b. I addition, there is also hail volume (which is the air volume that contains hail, as can be seen in the blue contours in Figure 1). Hail volume from the observation and the seven simulations should also be shown. It would be interesting to see which simulation has the hail volume that is the closest to the observation. It would also be interesting to see if the hail volume from the seven simulations changes in the same manner as the peak surface hail mass as seen in Figure 2c. As the manuscript found that ice-ice collisional breakup can expand the cloud ice region and graupel-to-hail conversion is suppressed, it would be interesting to see if the hail region is reduced by ice-ice collisional breakup.*

Response

Thank you for your suggestion. Figure 1 in the current main text compares the observed and simulated hydrometeor distributions. Your interpretation of the volume is right and we will add its definition at first mention in the main text. Figure 2b shows cloud-ice volume, defined as the air volume with a cloud-ice number mixing ratio $N_i > 10^5$ kg^{-1}. The added comparison uses $N_i > 10^4$ kg^{-1} and $q_h > 0.3$ g kg^{-1} in the simulations to show cloud-ice and hail-containing volumes (Fig. R5), with the same thresholds applied to all seven simulations. We also compare results for $N_i > 10^3$, 10^4, and 10^5 kg^{-1} and $q_h > 0.1$, 0.3, and 1.0 g kg^{-1} (Tables R1 and R2).

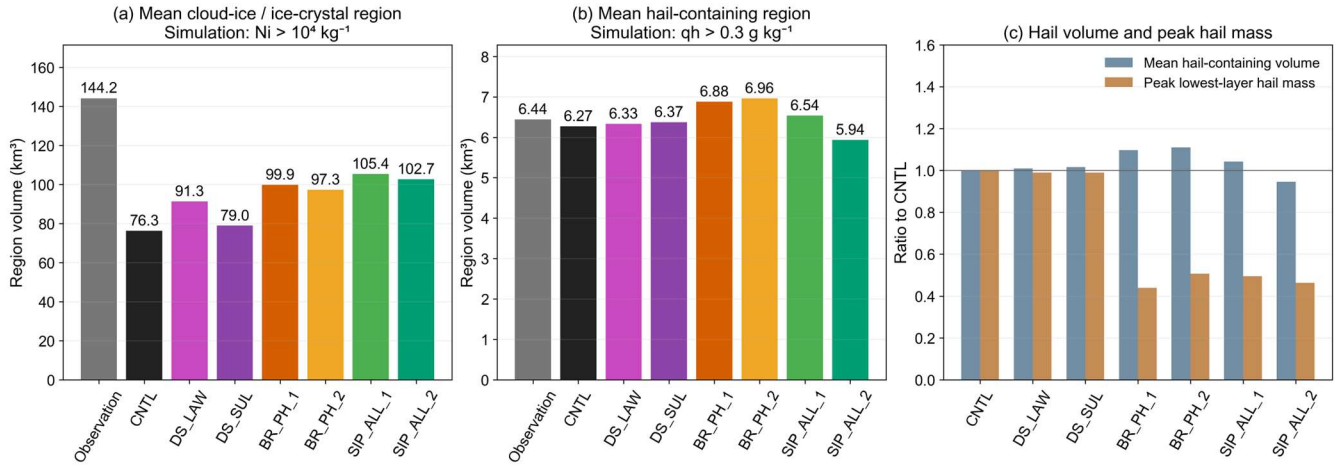


Figure R5. Ice-crystal and hail-containing volumes and near-surface hail mass. (a, b) Mean values during 05:09–05:16 UTC in the observations and 107–111 min in the simulations. Observed volumes are estimated from the ice-crystal and rain–hail mixture classes; the simulations use $N_i > 10^4$ kg^{-1} and $q_h > 0.3$ g kg^{-1}, respectively. Both observations and simulations use 10 km × 10 km horizontal regions. (c) Simulated mean hail-containing volume and the lowest-layer hail mass peak for 79–149 min.

In Fig. R5, simulated cloud-ice volume is smaller than observed. At all three thresholds, cloud-ice volume in the four experiments including BR is greater than in CNTL, showing that BR does increase cloud-ice volume (Table R1).

Table R1. Sensitivity of cloud-ice volume to the N_i threshold.

N_i threshold (kg^{-1})	Volume range across seven simulations (km^3)	BR_PH_1	BR_PH_2	SIP_ALL_1	SIP_ALL_2
10^3	120.8–136.8	+9.2%	+8.4%	+13.2%	+11.0%
10^4	76.3–105.4	+30.9%	+27.5%	+38.1%	+34.6%
10^5	43.1–66.7	+44.5%	+41.2%	+54.9%	+48.4%

With $q_h > 0.3$ g kg^{-1}, mean hail-containing volume increases by 9.7%, 11.0%, and 4.3% relative to CNTL in BR_PH_1, BR_PH_2, and SIP_ALL_1, respectively, but decreases by 5.4% in SIP_ALL_2. This pattern remains the same at $q_h > 0.1$ and 1.0 g kg^{-1} (Table R2). Peak lowest-layer hail mass during 79–149 min decreases by approximately 49%–56% in all four experiments. Thus, hail-containing volume does not correspond to peak near-surface hail mass, and it does not decrease in every experiment that includes BR.

Table R2. Sensitivity of hail-containing volume to the q_h threshold.

q_h threshold (g kg^{-1})	Volume range across seven simulations (km^3)	BR_PH_1	BR_PH_2	SIP_ALL_1	SIP_ALL_2
0.1	10.57–12.05	+6.4%	+8.4%	+2.6%	-4.9%
0.3	5.94–6.96	+9.7%	+11.0%	+4.3%	-5.4%
1.0	1.81–2.53	+14.3%	+21.6%	+2.1%	-13.1%

However, the observed ice-crystal and hail classes and the simulated hydrometeors use different criteria, and the results are strongly affected by the model thresholds. The simulated volume with reflectivity of at least 20 dBZ is also smaller than observed. This comparison can therefore only show changes in cloud-ice and hail-containing volumes in the simulations and cannot directly validate simulation accuracy.

Comment 3

Regarding BR_PH_1 and SIP_aLL_1, results are only discussed in Figure 2. The whole manuscript does not have much discussion on these two simulations. It seems that 5 bins and 20 bins do not make much difference in the simulated results. Please consider moving all the discussions related to BR_PH_1 and SIP_aLL_1 to Appendix. I think it is very good that Appendix provides some discussion on additional sensitivity tests. It might be good to put a paragraph here for BR_PH_1 and SIP_aLL_1. What is the motivation of doing 5 bins and 20 bins?

Response

Thank you. We accept your suggestion and will move the detailed discussion of BR_PH_1 and SIP_ALL_1 to the Appendix, where we will also briefly explain why the bin experiments were conducted and compare the sensitivity results.

The text for the Appendix is:

"The bin-sensitivity experiments examine how the BR calculation is affected by the discretization of the particle size spectrum. The mass and number information predicted by NSSL is used to reconstruct the size spectrum, from which collisions and fragment production are calculated between bins. 5 and 20 bins provide relatively coarse and fine representations of the size spectrum, respectively. We use the 20-bin configuration for the main process analysis and the 5-bin configuration to test the sensitivity of the results. Both configurations give the same conclusion that BR reduces hailfall."

Comment 4

Regarding DS_SUL, it seems that this simulation is not very useful in this manuscript. The study focuses on the effect of ice-ice collisional breakup. Droplet shattering is not an important secondary ice process. So it does not seem very straight forward to discuss two drop shattering configurations. Again, my suggestion is moving DS_SUL to Appendix. But I'll leave it to the authors to decide whether or not to move. Lines 198-200 mentioned about "influences of raindrop size and thermodynamic control", please explain more on this.

Response

Thank you for your suggestion. This study focuses on ice–ice collisional breakup, while the contribution of droplet shattering is relatively small in the simulated storm. This can be shown by retaining DS_LAW in the main text. We therefore agree to move the DS_SUL results to the Appendix while keeping descriptions of both DS configurations in the main text.

The phrase "raindrop size and thermodynamic control" was somewhat unclear. It was intended to express the effects of raindrop size and temperature on the two DS processes. Thank you for pointing this out. We have revised the wording as follows:

"The DS_LAW and DS_SUL experiments compare two parameterizations of droplet shattering during freezing. In DS_LAW, the number of fragments produced by a single freezing raindrop increases with the fourth power of its diameter. In DS_SUL, the shattering probability is controlled by a temperature function, reaching a maximum of 0.2 at 258 K, with a temperature distribution width of 3 K. The two configurations therefore emphasize the effects of raindrop size and freezing temperature on fragment production, respectively."

Comment 5

For the modeled storm, it is easier to define the development stage and the mature stage. But for the observed storm, how to define the stages?

Response

In our previous study, Fig. 4 (see Fig. R2) shows 05:03, 05:09, and 05:16 UTC. As in the simulation, these times correspond to an early structure in which graupel is clearly present (convective initiation was at 05:01 UTC), a distinct hail-containing structure, and the mature stage, respectively. In this study, the observed developing and mature stages are likewise distinguished by whether hail has formed. A small part of the displayed cross-section is already classified as rain–hail mixture at 05:07 UTC, whereas a distinct hail-containing structure is not evident until 05:09 UTC. We therefore retain this division into stages. We will clarify this in Sect. 2.4 and the caption of Fig. 1 and make the stage names in the observation figures consistent.

Comment 6

In section 2.3.2, it seems that equation 12 should be put before equation 11, because equation 10 is used in equation 12. Also please check every equation and make sure every symbol in the equations is explained.

Response

Thank you for your suggestion. We will reorder the equations and their cross-references in Sect. 2.3.2 according to the relationships among the physical quantities. The sequence of the original equations will be (10), (12), (11), and (13).

Some symbol definitions were omitted when following the Phillips parameterization. We now provide the physical meanings of the variables in Eq. (12):

"M denotes the collision type and the morphology and physical state of the more fragile particle. A(M) is the number of breakable surface features per unit area on that particle, in m^{-2}. C is an energy-scale coefficient, in J^{-1}; γ is a dimensionless empirical exponent; and N_{max} is the upper limit on the number of fragments produced in a single collision. The parameters are assigned according to collision type, temperature, particle size, and degree of riming (Phillips et al., 2017, Table 1)."

Comment 7

In section 3.2, Figure 3 and Figure 4, some properties are not plotted consistently in the two figures. For example, in Figure 3, snow number and mass mixing ratio are not plotted. Please add these to the figure, or explain why snow is not plotted. In Figure 4, graupel and hail deposition are not plotted. Please explain why. Another important thing is that the mass growth rates for cloud ice, snow, graupel, and hail through aggregation of these ice particles are not shown in Figure 4. It would be good to see if these aggregation processes are slowed down or enhanced when compare the SIP_ALL_2 case with the CTRL case. These mass growth rates should also be consistent with the mass mixing ratio of ice phase particles in Figure 5c.

Response

Thank you for your suggestion. Snow number and mass mixing ratios have been added to the original Fig. 3, as shown in Fig. R6b and g. The original Fig. 4 omitted processes with relatively small values; graupel and hail deposition and ice-phase aggregation have now been added, as shown in Fig. R7d–f.

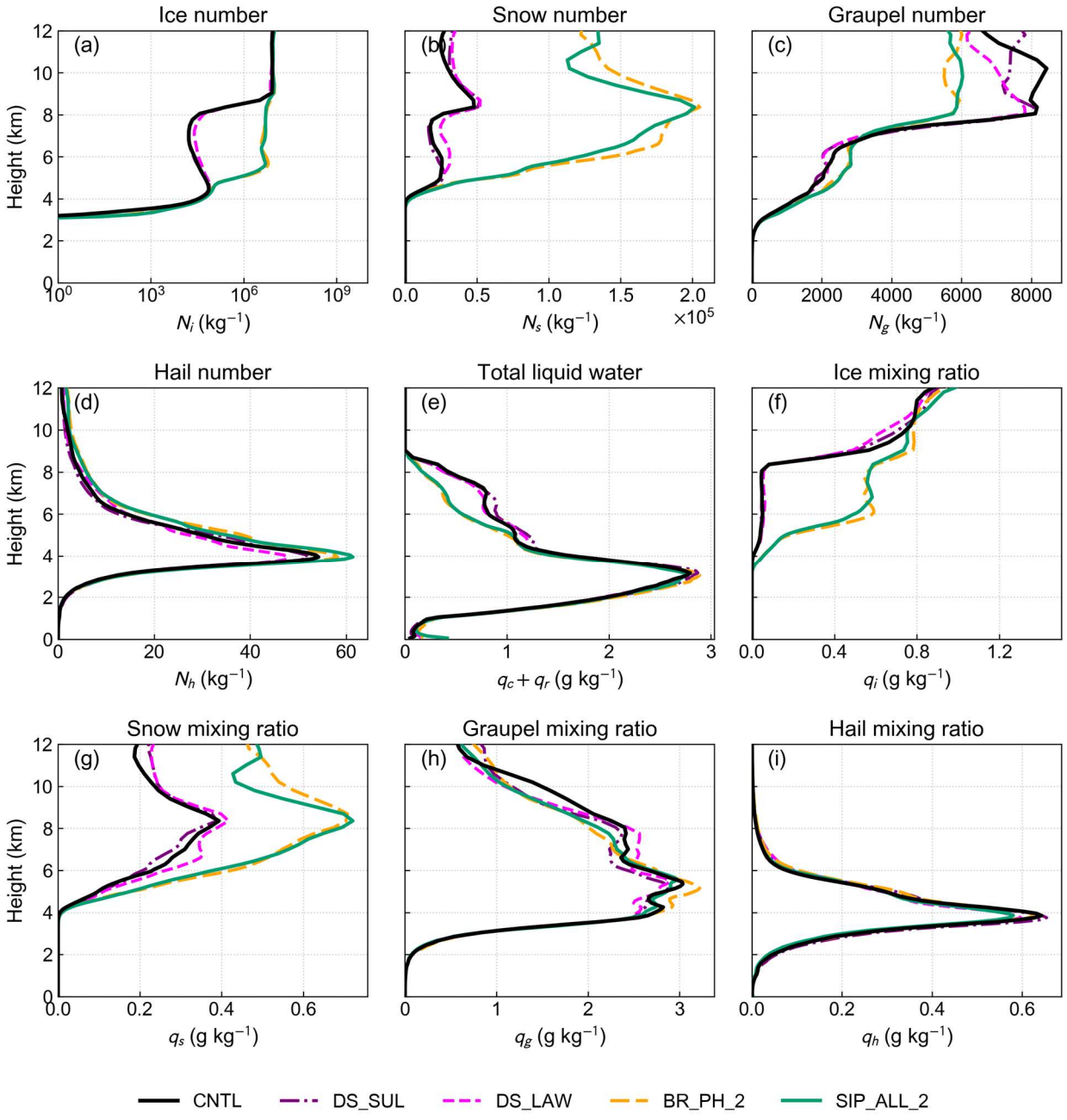


Figure R6. Mean profiles for each experiment during 108–120 min over grid cells with w > 2 m s^{-1} in the central analysis region. (a–d) Number mixing ratios of cloud ice, snow, graupel, and hail. (e) Total liquid water mixing ratio. (f–i) Mass mixing ratios of cloud ice, snow, graupel, and hail.

In Fig. R6, snow mass and number mixing ratios are both higher in SIP_ALL_2 than in CNTL. Figure R7a shows secondary ice mass production rates in SIP_ALL_2, while panels b–f show rate differences, SIP_ALL_2 minus CNTL. Depositional growth of cloud ice and snow is enhanced in the main ice-phase layer, while deposition onto graupel and hail is reduced at approximately 6–9 km.

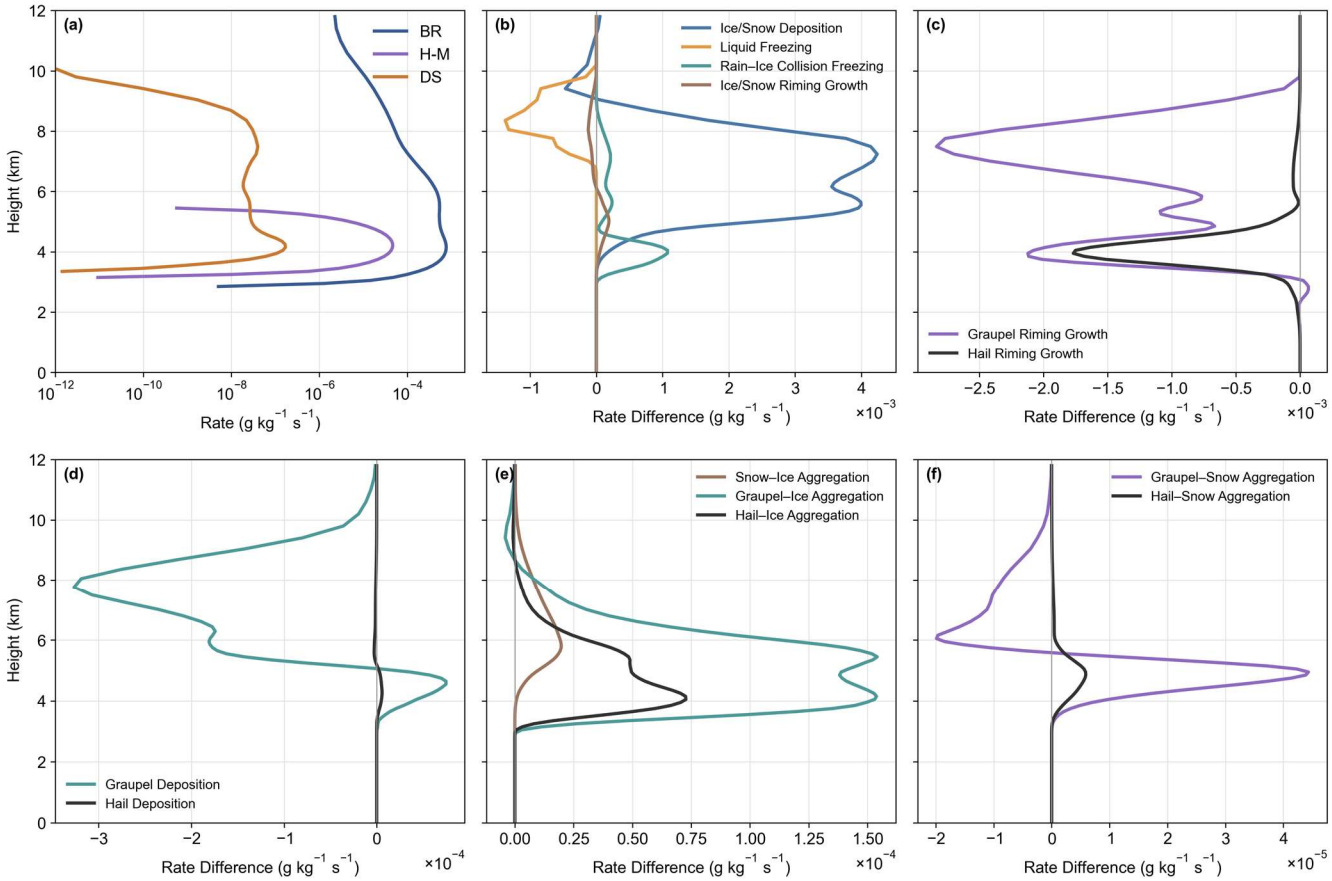


Figure R7. Mean profiles during 108–120 min over grid cells with w > 2 m s⁻¹ in the central analysis region. (a) BR, H–M, and DS mass production rates in SIP_ALL_2. (b–f) Rate differences, SIP_ALL_2 minus CNTL; positive values indicate enhancement and negative values indicate reduction. (b) Cloud-ice/snow deposition, droplet freezing, freezing through rain–ice collisions, and cloud-ice/snow riming growth. (c) Graupel and hail riming growth. (d) Graupel and hail deposition. (e) Growth of snow, graupel, and hail through aggregation with cloud ice. (f) Growth of graupel and hail through aggregation with snow.

At approximately 5–8 km, growth through aggregation with cloud ice is enhanced for snow, graupel, and hail. Graupel growth through aggregation with snow is reduced at approximately 6–9 km, while hail growth through aggregation with snow is enhanced at 4–5.5 km. Enhanced snow deposition and growth through aggregation with cloud ice are consistent with the increase in snow mass mixing ratio. However, these newly shown processes are smaller in magnitude than hail and graupel riming growth and do not dominate the differences.

Figure R7c (the original Fig. 4c) differs from the original manuscript mainly near 4 km, where the graupel riming growth rate difference relative to CNTL is now negative. This difference arose because adaptive time stepping was inadvertently enabled, although the earlier manuscript stated that a fixed time step was used. Figure R7c should be used as the corrected result. This change does not affect the conclusions.

Comment 8

The CTRL case and the SIP_ALL_2 case have quite different hail mass at the surface (as can be seen in Figure 6d and Figure 2c). However, the conversion rate of graupel to hail is very similar for the two cases (as seen in Figure 6c1. So what process leads to the hail mass difference between the CTRL and SIP_ALL_2 cases?

Response

The difference in surface hailfall between the two experiments mainly results from weaker graupel-to-hail conversion and hail riming growth. We apologize for incorrectly labeling the cumulative quantities in the previous Fig. 6c as rates. Panels a and b of Fig. R8 now compare the rates of these two processes. Domain-wide integration over 79–149 min shows that graupel-to-hail conversion and hail riming growth in SIP_ALL_2 are approximately 34% and 23% lower than in CNTL, respectively. Thus, although the mean curves look relatively similar, the actual cumulative growth differs substantially.

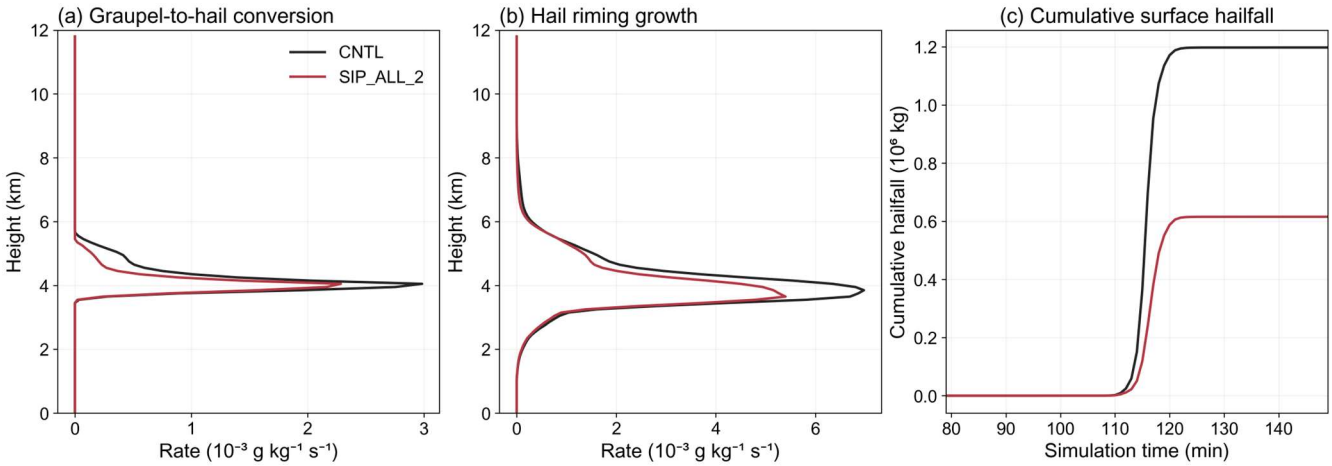


Figure R8. Hail growth and surface hailfall in CNTL and SIP_ALL_2. (a, b) Mean graupel-to-hail conversion and hail riming growth rates in the updraft region during 108–120 min. (c) Actual cumulative hail mass reaching the surface.

During this period, total losses in SIP_ALL_2 decrease by approximately 22%. Combining growth and losses gives a net microphysical supply approximately 50% lower than in CNTL. Reduced melting and liquid water shedding offset part of the reduction in growth but do not compensate for weaker graupel-to-hail conversion and riming growth. Actual cumulative hail mass reaching the surface decreases by 48.6% (Fig. R8c).

We further checked the mass budget using growth, losses, changes in airborne hail mass, and actual cumulative hail mass reaching the surface. The estimated difference in surface hailfall between the two experiments agrees with the actual output.

Minor comments

Comment 9

The ice phase collision pairs, such as H-H, H-G, etc., should be written as complete words when they first appear in the manuscript.

Response

Thank you for your suggestion. At the end of the paragraph on collision efficiency in the original Sect. 2.3.2, the full names for H, G, S, and I are already given. The relevant text is:

"Here, H, G, S, and I denote hail, graupel, snow, and cloud ice, respectively. For example, H–G represents hail–graupel collisions, and G–I represents graupel–cloud ice collisions."

Comment 10

In section 2.1, it is mentioned that data from two X-band radars are used. In Figure 1, is the radar reflectivity from one of the radars or from the combination of the two radars?

Response

We apologize. The observed reflectivity and classification results in Fig. 1 come from the single X-band dual-polarization phased-array radar at Xinming Village, ZWN02. The original Sect. 2.1 described the field instrument deployment together with the data actually used in this study, which caused the misunderstanding.

The data source in Sect. 2.1 and the caption of Fig. 1 is clarified as follows:

"This study uses reflectivity and hydrometeor classification data from the X-band dual-polarization phased-array radar at Xinming Village (ZWN02; 103.94° E, 26.98° N). All observations shown in Fig. 1 come from this single radar."

Comment 11

In section 2.1, the hydrometeor classification using X-band radar should be described in more details. It should at least point out that how many types of hydrometeors are classified.

Response

Thank you for your suggestion. Hydrometeor classification uses a fuzzy-logic method. Polarimetric variables, including reflectivity factor Z_H, differential reflectivity Z_{DR}, specific differential phase K_{DP}, and the copolar correlation coefficient ρ_{hv}, are used to calculate membership degrees. Weighted scores then determine the dominant echo class.

The additional text in Sect. 2.1 is:

"X-band radar hydrometeor classification uses a fuzzy-logic method based on membership functions of polarimetric variables. It identifies eight hydrometeor echo classes: dry snow, wet snow, ice crystals, graupel, big drops, light-to-moderate rain, heavy rain, and rain–hail mixtures. The algorithm also includes ground clutter/anomalous propagation and biological scatterers as two non-hydrometeor classes, giving ten echo classes in total. We use the ice-crystal, graupel, and rain–hail mixture classes to describe the main ice-phase structures, with the rain–hail mixture class representing the observed hail-containing region."

Comment 12

Title of 3.1 should be "observed and simulated storm structure ……"

Response

We accept your suggestion. We have changed the title of Sect. 3.1 to "Observed and simulated storm structure and evolution" to match the discussion of the structure comparison in this section.

Comment 13

The fonts in almost all the figures are too small.

Response

Thank you for your suggestion. We have increased the font sizes in the figures and adjusted parts of the layout.

Yi Dong, Shaofeng Hua, Baojun Chen, and Tuanjie Hou