Review of ‘‘Ice–ice collisional breakup enhances secondary ice production but reduces surface hailfall in a simulated pulse hailstorm’’

The manuscript describes sensitivity tests of hail production in the NSSL 2-moment bulk microphysics scheme. This topic is particularly interesting to me as the developer of this scheme. Although I think it would be more interesting if it were applied to a severe storm with larger hail production, I can accept that this is done in the context of the previous work with these kinds of storms. Severe storms (supercells) could be a follow-on (and using more recent microphysics code that uses a new graupel-to-hail conversion method). The provision of the CM1 code and analysis files on Zenodo were very helpful for reviewing the manuscript by providing more details than the figures alone. There are a number of issues that I have listed in the comments below, but I think generally the method and implementations are good, and the paper should be acceptable with major revisions. (I have not examined the breakup code, but I trust that the authors have tested it sufficiently, and it seems to yield reasonable results.) The authors are welcome to contact me directly with any questions.

best regards, Ted Mansell (ted.mansell@noaa.gov)

Primary Comments:

1. Fig. 3d: The rain+cloud mass near z=0 looks wrong. Something is strange there to cause a spike of 3g/kg. From panels f and g, it cannot be melting. The files in `processed_data/fig3/nc` show that it is rain that has the large value, and only at the lowest grid point. This looks like an error. Hail mass and number also have an increase at the lowest level, which seems unrealistic.

   Likewise in Fig. 6a, the sudden reduction in diameter at k=1 is similarly curious and possibly an error. The diameter increase below the melting level is expected from melting-driven size sorting. What does it look like if you plot, for example, the 90th percentile?

   I did some tests with the code provided on Zenodo, and I found a problem with the added variables in writeout.F. The additional output code was accidentally inserted within another output conditional. For my compiler (gfortran 15) and netcdf library, this resulted in a corrupted time variable: each file had 96 time levels, but only one valid one. Effects may vary with different compilers. I ran a coarse (dx=500m) test and saw no indications of large qr at k=1 (with corrected writeout code). There could otherwise be an issue with the analysis code.

2. Fig. 5: The mean rate here uses mixing ratio (g/kg), which is not physically correct for vertical averages. This is because air density varies with height, so that the same mixing ratio at different altitudes represents different mass content. It is the mass content (g/m3) that is the correct variable to use here. To average correctly, the rates must first be multiplied by air density to convert to g/m3/s. Since dz also varies with height, the rates should further be multiplied (weighted) by the grid cell volume (dx*dy*dz) to get kg/s (i.e., the total mass rate over the domain region). This is because the same mass rate in a larger grid cell implies a larger contribution to the total storm mass. So I recommend plotting the total mass rate over

time, which has a simple interpretation and comparison. A mean value would be the total mass divided by the total grid cell volume, which seems less meaningful to me.

This is likewise true for number concentration (c) and change in mass (b), which means calculating the total number and total masses.

I do not know if this will change any conclusions, but it is better to use the correct physical quantities and not have any questions about it.

3. Fig. 4: At $T < 0$C, 'shedding' is not a net mass loss but rather a reduction in the gain. There is always at least some mass gain with wet growth, so to plot this as a mass loss is not very meaningful because the hail is not actually losing net mass. The net riming gain could be plotted instead to show a difference between the cases.

4. The BR parameterization was used with either 5 or 20 diameter bins, but there is no discussion about the differences in results. The BR1 plot in Fig. 2 is notably different from the others by having only one relative maximum. How do they compare with 10 bins? Were there any simple tests to compare the different resolutions?

Minor Comments:

1. At first I was concerned that there was no mention of surface observations of hail. Otherwise, hail aloft that completely melts before reaching the ground has no severe impacts is not so interesting. But I found that the preceding paper, Dong et al. 2024, does list observed hail (8mm max diameter) for this date. It would be good to repeat that information here.

2. Line 116: X-band experiences strong attenuation in heavy precipitation. How was this detected and corrected? The netcdf files included in the Zenodo archive show fairly obvious attenuation (Z and Zdr), but I don't see anything for attenuation correction.

   Was the hydrometeor identification done with the attenuated data? If so, that raises questions about the results unless there is some kind of correction built in to the algorithm. The grayscale reflectivity in Fig. 1 is a bit hard to follow quantitatively, but it looks like a maximum of 40-45 (maybe 45-50?) dBZ, which is pretty marginal for hail (not accounting for attenuation).

3. Line 134: 'initialized with an intensive radiosonde sounding'

   Please add a plot of the model sounding. The archived sounding (input/input_sounding) has had the temperature profile replaced with an analytical function, which is not the same as plotted in Dong et al. 2024 and has a much larger surface-based CAPE.

4. Line 182: 'DS_LAW uses the empirical expression of Lawson et al. (2015)'

   Note that more recent versions of NSSL-MP apply the Sullivan et al. (2018, parcel model) temperature-dependent profile (Pds) to the Lawson formula. That could be done here, too, which would reduce the upper maximum in Fig.4a. Or leave it as-is for the maximum effect and comparison with BR.

5. Line 185: 'd is the equivalent-volume diameter of the raindrop'

   It is not a single raindrop: this is just what is already in the code, which is the mean volume diameter of the converted rain drops, derived from qrfrz/crfrz (or qiacr/ciacr).

6. Line 218: 'For each ice-phase category, the particle size distribution was assumed to follow a gamma distribution in diameter space.'

   Note that in the NSSL scheme, ice crystals and snow use a PSD of volume (mass) rather than diameter. So this assumption for ice and snow is inconsistent. It might not make a large difference, but it can be acknowledged.

7. Line 285: 'corresponding fragment mass'

8. Line 324: 'CNTL is the control experiment, in which only the default H–M process in the NSSL scheme is retained.'

   In the Zenodo archive, namlist_CNTL has

   ```
   ienable_sip_ds = 0 ! ydong_sip
   ienable_sip_br = 1
   num_bins =  10
   nsplinter = 10
   sip_scaling_factor = 0
   sip_max_mass_loss_rato = 1
   ```

   With ienable_sip_ds=0, the logic goes to the branch for (nsplinter .ge. 0), which results in 10 splinters per frozen drop (of any size) because nsplinter=10 and there is no factor of P-DS (i.e., the original code). The BR-Only namelists also have nsplinter=10. This would not be a large contribution, but it is not zero, either, and they probably should be rerun with nsplinter=0.

   Also, the CNTL namelist shows `ienable_sip_br=1` which may be having some effect? It is also possible that the archived namelist is accidentally different from what was actually used.

9. Line 341: 'The mature-stage mean cloud-ice volume is calculated within the same horizontal analysis box and is defined as the volume of grid cells satisfying $N_i > 105\text{kg}^{-1}$.

   Note that using number per kilogram leads to inconsistency in the physical quantity of number per cubic meter. This threshold might be too high to capture some regions of SIP at higher temperature.

10. Figure 2d (and elsewhere for 'surface hail'): I suggest calling this 'near-surface' hail mass to avoid confusion because there is no output of actual surface-accumulated ice. The instantaneous value of the sum of the lowest model level is not the same thing as surface (accumulated) hail. Note that the most recent version (22.0) of CM1 now has a separate hail (total ice) accumulation. (The 'rain' variable is still total precipitation.)

11. Fig. 2d: It is a little curious that DS-SUL is more different from CNTL than DS-LAW, since DS-LAW ought to have higher crystal production than DS-SUL. Can you comment on this?

12. Fig. 4: Here (and in Fig. 3) it would be preferable to use the physical variable of mass content (g/m3) instead of mixing ratio g/kg (advection variable). For example, the difference in air density between z=4km and 10km is about a factor of 2. That would show mainly a slight shift on the log scale (panel a), but much more apparent on the linear scales (b,c).

13. Line 471: By ‘direct liquid freezing’ does this mean homogeneous/hetergeneous nucleation (i.e., qrfrz, without ice crystals)? Does it include cloud droplet freezing?

14. Fig. 4: The hail conversion mass rates seem large. Is this a sum over the domain?

15. Fig. 4d: How is the surface hail mass spectrum averaged?

Editorial comments:

1. Line 580: Suggest deleting ‘stronger’ (or replace with ‘changing’)