Review of "Impacts of Secondary Ice Production on the Microphysics and Dynamics of Deep Convective Clouds in Different Environments", by Waman and Colleagues

Verdict

I recommend the paper be published subject to major modifications to the text and analysis. Improved model validation and analysis is needed.

Major Comments

This paper is commendable in its overall aim to gauge holistically the effect on deep convection in contrasting cases, covering dynamics (latent heating, ascent), radiative effects and the microphysics.

Yet there are some issues:

- **Inconsistent model–observation comparisons (filtered vs total ice; medians vs means)**
- **Lack of key validation metrics (INP concentrations, vertical velocity statistics, precipitation time series)**
- **Missing process-level diagnostics (domain-integrated SIP budget)**
- **Overinterpretation of dynamical impacts (stratiform vs convective ascent)**

Specifically, the model validation against radar reflectivity is fine. But against the aircraft data, the comparison between modeled and observed ice concentrations is difficult to interpret because the model total ice concentration is compared with size-filtered observations. A consistent comparison (i.e., applying the same size threshold to model output) would be necessary to assess agreement. Also, none of the sample means of observations are shown, only medians, on the validation plots for aircraft data. It is the sample mean that is the estimator of the population mean.

Thus, the microphysical validation plots are not "apples-to-apples" comparisonsThus, what appears to be strong disagreement in ice concentration (e.g., Fig. 7d for DCMEX) may partly reflect inconsistencies in the quantities compared rather than true model bias. If the ice concentration were really orders of magnitude too high, then the supercooled liquid water content would be far too low, but that is not apparent for DCMEX, so the disagreement may have been overstated.

Also, there is no comparison of vertical velocity statistics (cumulative frequency histogram) between aircraft and model, so it is difficult to see if the aircraft has properly sampled the deep convection. If the DCMEX aircraft never sampled the convective cores, then that might explain why the measured ice concentrations are so low.

There seems no validation of INP active concentrations. In these field campaigns, it seems unlikely that there were no measurements available of INPs, either in the campaign itself or approximately from an adjacent campaign (e.g., similar month, different year). To get the ice concentrations correct and to infer the correct role of the SIP, the primary ice needs to be predicted correctly. And it is unclear how the aerosol chemistry and loadings were initiated for the cases and how aerosol conditions are related to the predicted INP activity.

There is also no validation of surface precipitation as a function of time for all cases, which makes it difficult to visualize the time evolution of each simulated storm.

Regarding the analysis of impact of SIP on microphysics, there needs to be a simple budget of ice initiation (a count of all ice particles initiated) in the entire domain for the 4 SIP mechanisms and primary ice (not including homogeneous freezing), displayed as a pie chart. This will allow the present simulation of MC3E to be compared with the pie chart shown by Waman et al. (2022) for the same SIP mechanisms. The same pie chart can be shown for the PHIL18 run too, to compare with SUL18.

Inspection of the plots shows that the dynamical impacts reported actually relate to the stratiform region (the only ascent shown to be sensitive is stratiform in magnitude), not the convective cores, yet the opposite is suggested in the text. This claim *"in the simulated DCCs, the exclusion of SIP significantly ...weakens the strength of convection through decreased vertical motion"* is not supported by the results shown.

On balance, there are some intriguing results about the thermodynamical and radiative effects from SIP. Yet there are some sweeping or misleading statements in the analysis of simulations and the model validation is unconvincing as regards the aircraft data. The manuscript attributes changes in convection strength to SIP; however, the vertical velocity magnitudes shown (~10 cm s^{-1}) are characteristic of stratiform ascent rather than convective updrafts. This distinction should be made explicit in the interpretation. This distinction is important, and the conclusions should be revised or additional diagnostics of convective updrafts (e.g., conditional statistics for $w > 1$ m s^{-1}) should be provided.

Detailed comments:

Line 36: The statement that "on a global scale, the majority of surface precipitation is from the ice-crystal ('cold rain') process" requires more careful qualification. The cited studies (Lau and Wu, 2003; Field and Heymsfield, 2015) rely on indirect inference methods that may overattribute precipitation to ice processes—for example, by assuming that the presence of ice aloft implies an ice-phase origin of surface precipitation. This assumption is not generally sufficient to establish causal microphysical pathways.

More recent work (e.g., Gupta et al., 2023) uses tracer-based diagnostics in cloud-resolving simulations to explicitly track precipitation origin and demonstrates that the dominant pathway can vary substantially by case. In particular, Gupta et al. show that while ice-phase processes dominate in some mesoscale convective systems, warm-rain processes can dominate in other environments (e.g., tropical systems), even when ice is present aloft. These results suggest that stronger caution is needed when making general statements about the global dominance of ice-phase precipitation.

I recommend citing Gupta et al. (2023) and revising the statement to reflect this uncertainty and case dependence.

Additionally, the term “cold rain process” is potentially misleading. In many cases, the ice-crystal process produces solid precipitation (snow) that may or may not melt before reaching the surface, depending on thermodynamic conditions. Referring to this pathway as a “rain” process can obscure the underlying microphysics, particularly in cold environments where no liquid phase is involved at the surface. A more precise term such as “ice-phase precipitation process” or “ice-crystal process” would improve clarity.

Line 56: Where it is written “The HM process has been recognized as important for rapid ice enhancement in the early growth of DCCs” the language does not seem very scientific. Recognised by whom ? “Importance” can be subjective. Maybe the authors can try to be more precise in the citing of these studies: do you mean that these studies predicted that the HM process was active, and if so was it in the early stage of the growth of the cloud ? Which of these studies were validated with acuity (for ice concentration and liquid water content and surface precipitation or radar reflectivity) ? What was the evidence ?

The problem is that this paragraph poses a dilemma that any reader will be intrigued by: is the HM process real or not ? It would be helpful to see if there is any field campaign evidence of the HM process. Bower et al. (1996) published a vertical profile of observed ice concentration with a huge peak right in the centre of the HM generation region, in flights through deep stratiform cloud. Has there been anything like this more recently with modern technology for optical probes in stratiform clouds ?

Line 51: With raindrop-freezing fragmentation, it would be helpful to cite the early experimental studies, such as Takahashi and Yamashita (1977) or Johnson and Hallett earlier.

Line 62: It would be helpful to cite Phillips et al. (2017b, Part II) who were first to predict that breakup in ice-ice collisions prevails in the overall ice production of a simulated mesoscale convective system.

Line 75: It is written that "*While laboratory experiments (Oraltay and Hallett, 1989; Bacon et al., 1998; Dong et al., 1994) demonstrated fragmentation of sublimating dendrites and rimed particles under relative humidity with respect to ice (RHi) 50–90%, field observations (Korolev and Issac, 2004) and modeling studies (Waman et al., 2022; Yang et al., 2024) found minimal contribution of SBF to total ice formation in clouds*." But Figure 10 of Waman et al. (2022) shows that sublimational breakup in the convective descent and stratiform regions of a MC3E storm (MCS) is about the second-most important SIP mechanism in terms of ice concentrations from tagging tracers. The tagging tracer result is definitive as it shows the true concentration in the control simulation. For simulations of GOAmazon, Deshmukh in his PhD Thesis found in the ice initiation budget (after accounting for losses from sublimation of SB-fragments) the sublimational breakup is as important as ice-ice collisional breakup.

Did Yang et al. tag the concentrations rigorously ? It would be helpful to know.

The citation to Korolev and Isaac (2004) does not appear to support the statement regarding sublimational breakup, in the quote above. The paper focuses on shattering processes generally and does not discuss sublimation at all. A more appropriate reference should be used.

Line 71: It would be helpful to state that Lawson observed tropical convection… the very warm cloud-base creates a strong warm rain process. Also, they did not "show" that raindrop-freezing fragmentation dominates, since they did not model their case comprehensively. They inferred this with various assumptions, including the notion that the HM process is unimportant. Their aircraft chased very young ascending turrets, and parcel simulations of these cloud-top regions by Phillips et al. (2018) showed that both raindrop-freezing fragmentation and the HM process occurred. Over longer time-scales other SIP mechanisms would likely become prolific too.

Line 72: It is written that "*Furthermore, recent modelling studies of DCCs (Huang et al., 2021; Waman et al., 2022; Patade et al., 2025) reported that RDF dominates overall SIP in their early stages.*" But this is an exaggeration, at least for Waman et al. 2022. See

Figure 21 of Waman et al. (2022): HM process is much more prolific than RDF during the first 10 mins, which is very similar to sublimational breakup, in the budget of numbers of particles initiated. Also see Figure 18 of Waman et al., which shows this too.

Line 77: Patade et al. 2016 were not the first to discover that the HM process occurs in deep convection. There is a long literature before. It would be helpful to cite earlier studies such as Hallett et al (1978), Bower et al. (1996) and Latham and Blyth (1993) and Phillips et al. (2007).

Table 1: It seems unlikely that all four cases would have the same cloud-top, even if it near the tropopause. Perhaps call it "maximum cloud-top". If the latitudes are completely different among the cases, then the tropopause temperature will differ too among them.

Line 228: In Equation 2, this is the wrong temperature range for the HM process. It should be -3 to -8 degC. Perhaps there is a typo ? If not, then one cannot call this process the "HM process".

Line 275: Better to track the dendritic snow with a tagging tracer and not limit the sublimational breakup of it to the dendritic temperature regime. Waman et al. (2022) did that.

Line 279: *d* is the diameter of the sublimating parent ice particle, and it is not generally a crystal. Rather it is either graupel or snow here.

Figures 6–9: The validation plots for ice concentration are difficult to interpret due to inconsistencies in the quantities being compared. The filtered ice concentration from the observations appears to be compared against the total ice concentration from the model. For a consistent ("apples-to-apples") comparison, the model output should be processed using the same filtering criteria as applied to the observations.

If the objective is to evaluate mean cloud properties, the comparison should be based on consistent statistics—i.e., model mean versus observed sample mean. In that case, the observational mean should be accompanied by an estimate of sampling uncertainty (e.g., standard error or confidence intervals), particularly given the limited and correlated nature of aircraft measurements.

The current box-and-whisker representation primarily conveys the distribution (median and percentiles) of the observations and does not directly quantify agreement in the mean. This is useful for assessing variability, but it does not by itself support conclusions about agreement in mean values.

If sample means are presented, it would also be helpful to indicate the effective sample size or sampling duration (e.g., time or distance of aircraft segments), given the incomplete and potentially autocorrelated sampling of clouds.

Finally, the inclusion of model extrema (minima/maxima) is of limited interpretive value, as these depend strongly on model resolution and sampling and are not directly comparable to observational extrema.

Also it is not always helpful to plot all the raw data on the validation plots, as it is impossible to interpret when innumerable points just create a wall of opaque shading.

Line 520: This statement seems to suggest an impact on the convective updrafts from SIP: "*Excluding SIP in the simulated DCCs causes noticeable changes in the 90th-percentile vertical velocities (Fig. 18), thereby influencing the overall convective strength*." But when one looks at Figure 18, these are not convective-scale rates of ascent. Rather, the order of magnitude is 10 cm/sec, which is stratiform.

It would be helpful to show a cumulative frequency histogram of convective ascent and define the zero-th percentile of ascent to be 1 m/sec. Then the 90th percentile of convective ascent will be a few metres per second at least.

It would be helpful to ascertain if there is any appreciable sensitivity of convective ascent at all in any of the control simulations.

Without any such re-working of the data, it would be helpful to be open about the fact that the sensitivity of dynamics reported is actually stratiform weak ascent and not the convective cores.

Line 549: This sentence has a typo "*both SUL18 and PHIL18 schemes predict about 1.5 times higher mean INC than SUL18.*" In this section, it would be helpful to show the process rate profile for PHIL18 in a similar fashion to Fig. 10.

One suspects that the relative difference in raindrop-freezing fragmentation rates between both schemes must be large (e.g. an order of magnitude ?) in most of the cases, because for CAIPEX, the total ice concentration between 0 and -10 degC is changed by up to half an order of magnitude at levels where there are likely far more HM (or IIC) splinters than raindrop-freezing splinters.

Line 590: Point 5 suggests an alternative explanation to that given in point 4: if the cloud-top is much lower without SIP (due to latent heating changes), then the rate of LW emission to space is very much more than with SIP by Stefan's Law. So that explains the reduction in LW heating rate when SIP is excluded.

Line 593: It would be helpful to give the range of percentage change of vertical velocity. 10% seems to be an exaggeration for DCMEX. For one or two cases the response is more like 1%.

Line 597: It is misleading to state that "*In all simulated DCCs, both SUL18 and PHIL18 schemes of SIP through RDF predict similar INC and dynamical properties*." In fact, only the total ice concentration in simulations using both RDF schemes is shown. Nowhere is the actual component of ice concentration from each scheme shown or their budgets of ice initiated.   It seems almost inevitable that simulations with both schemes give similar total ice concentrations in long-term averaged profiles, since breakup in ice-ice collisions is by far the leading SIP mechanism at most levels in general.