Responses to reviews

"Secondary ice production affects tropical convective clouds under different aerosol conditions"

by Sun et al.

Thank you to the editor and reviewers for their continued time and thoughtful comments on our revised manuscript. We appreciate the further suggestions and have addressed them carefully in the revised manuscript and Supplement. Our point-by-point responses are provided in blue, with the corresponding revised text shown in grey italics.

Referee #1

The authors have done a thorough job in addressing each reviewer's comments. The way the figures are now introduced and presented is much clearer. I would like to give a few more suggestions regarding future perspectives, as well as some technical comments to consider before publication.

Technical comments:

• All figures: The labeling of figure panels (a, b, c, …) is currently placed inside each panel. It would be more visually consistent to put these labels outside each panel (above the top-left corner).

• Figure 1: The text (particularly the axis labels) in panel (b) is too small, especially compared to panel (a).

• Figure 2-3: The label text is also too small and should be increased for better readability.

• Figure 5: In panel (a), the panel title could be written after "a)" for consistency with the other panels.

• Figure 8: The names associated with panels (g), (h), and (i) could be added (as suggested for figure 5).

We have revised the figure formatting following these suggestions. The font sizes in Figures 1-3 have been increased, and the panel names in Figures 5 and 8 have been added for consistency. We have retained the panel labels inside the panels because we consider this placement clear and easy for readers to follow in multi-panel figures. However, if further formatting requirements are raised during the figure production stage, we will revise accordingly to comply with the journal guidelines.

• **Lines 429–430:** Do you mean that aggregation and sedimentation are much more dominant than the effect of SIP? This sentence could be reformulated for clarity. I remain somewhat skeptical about the low impact of SIP on snow, given the large increase in ice particle number. Anyway, these are the results produced by the model.

Our point was that the diagnosed snow category shows a weaker response in this case, and

that the SIP signal in snow may be partly moderated by subsequent aggregation and sedimentation processes in CASIM. We have reformulated the sentence as follows (lines 434-435):

"*By contrast, the diagnosed snow category shows a weaker response, suggesting that the SIP signal in snow may be partly moderated by aggregation and sedimentation in this case (Field et al., 2023).*"

• **Lines 663–664:** From my perspective, the dominance of each SIP process depends strongly on the microphysical scheme used, not only on the SIP parameterizations or the environmental conditions of each case study. This point partly addressed in next sentences with the example of drop immersion freezing but could be extended for microphysics schemes in general.

We agree that the relative importance of different SIP mechanisms can depend not only on the SIP parameterizations and environmental conditions, but also on the microphysics schemes. We have revised the sentence as follows (lines 667-670):

"*Part of this difference may arise from differences in the microphysics schemes, as well as from the environmental conditions in our simulations, where graupel sizes and collision energies are relatively modest (Sun et al., 2025), limiting the efficiency of BR. One specific example is the treatment of raindrop freezing in the two modeling systems. …*"

Recommendations for future work:

For future work, I strongly recommend that the authors conduct a more detailed validation against observational data to evaluate the ice/snow/graupel number concentrations (or even particle size distributions). Although Figure S1 in supplement provides some comparisons, it is limited to low-temperature regions near the cloud top, where SIP is less likely to significantly enhance ice particle concentrations. The simulated ice particle number appears to peak around 1 L^{-1} close to 6 km, which is relatively low compared to what in situ aircraft measurements reported for convective clouds.

We agree that a more detailed validation of ice-phase properties would provide a stronger constraint on the simulated microphysics.

In the present case, such validation is limited by the available aircraft observations: the Dornier mainly sampled lower-level cloud regions, while the Egrett sampled the upper-level anvil outflow at approximately 12–14 km. Therefore, the available measurements do not provide co-located sampling of the mixed-phase regions, where SIP-related enhancements are expected to be significant. Figure S1 is thus intended as a limited comparison of upper-level ice properties.

We note that the simulated ice particle number concentrations represent temporally and spatially averaged diagnostics in this study, which may not be directly comparable to in-situ aircraft measurements without applying a consistent sampling method. The Conclusions section also notes that constraining these responses through observations and targeted modeling will be essential for improving the representation of SIP processes across regimes.

Referee #2

A current limitation of the manuscript is that the microphysical realism of the model is not demonstrated as clearly as it could be, given the largely process-oriented conclusions drawn from the simulations. While the paper provides useful comparisons against radar reflectivity and radiative fluxes across the different sensitivity experiments, there is comparatively limited use of available in situ observations to evaluate the baseline model state.

In particular, aircraft observations from the TWP-ICE/ACTIVE campaigns appear to be under-utilised in the main text for model evaluation. This is important because these datasets provide valuable constraints on both aerosol and cloud microphysical properties, and comparable model validation has been demonstrated in earlier work (e.g. Kudzotsa et al., 2016 and associated PhD thesis) for similar simulation periods, including cases overlapping with the 6 February 2006 storm examined here.

It would strengthen the manuscript to re-name one of the existing simulations as a clearly defined "control simulation", with a corresponding evaluation section that compares its model output against key observational constraints from aircraft, satellite, and ground-based measurements. At present, the simulations are primarily presented as a set of sensitivity experiments, without first establishing a baseline level of agreement with observations for a standard version of the model.

We understand the reviewer's suggestion to define a single control simulation. However, in the present study, the model evaluation is performed across the experiment set. In the main manuscript, the simulations are evaluated against MTSAT-derived OLR spatial patterns, CPOL reflectivity CFADs, and CPOL-derived precipitation. This allows the model behavior to be assessed across Nd = 200, 400, and 800 cm−3. For this reason, we have not redefined one simulation as a "control".

In the Supplement, we provide additional targeted comparisons. Figure S1 compares upper-level aircraft-derived ice number concentration and IWC with the noSIP-400 and allSIP-400 simulations. This comparison is kept in the Supplement because the available aircraft observations do not sample the mixed-phase layer where SIP is expected to be most active. Figure S3 further compares the domain-averaged OLR time series from noSIP-400 and allSIP-400 with the MTSAT-derived equivalent OLR. We focus these additional comparisons on the Nd = 400 cm−3 simulations because this configuration shows a clear SIP-induced signal in the main analysis and therefore provides a representative case for these targeted observational comparisons.

Specific additional evaluation metrics that would be valuable to include are:

- **Ice particle number concentration:** Comparison between the control simulation and aircraft-derived estimates, as already demonstrated in prior studies of the TWP-ICE/ACTIVE case. Where possible, it would be helpful to consider conditionally averaged comparisons over cloudy regions (e.g. non-zero ice water content), even acknowledging the sampling limitations of upper tropospheric measurements.

We agree that aircraft-derived ice particle number concentration provides an important

constraint on the simulated ice-phase microphysics. We note that this point has already been addressed in the submitted revision in response to the previous comments.

Specifically:

(1) In the revised manuscript, we had added a description in Section 2.3 of the available aircraft observations and their sampling limitations. We stated that the Dornier sampled the lower troposphere, while the Egrett sampled the anvil outflow at 12–14 km. We also clarified that the available in-situ measurements of IWC and ice particle number concentration were confined to upper levels and did not adequately sample the mixed-phase region where SIP is expected to be most active.

(2) In the same section, we had also stated that a comparison of these upper-level ice properties is provided in Figure S1 of the Supplement.

(3) In Figure S1, we compare vertical profiles of ice number concentration and IWC from the noSIP-400 and allSIP-400 simulations with Egrett aircraft observations. The Nd = 400 cm−3 simulations are shown because this aerosol condition exhibits a clear SIP-induced response in the results. Following the reviewer's suggestion, both the simulated and observed profiles are sampled over regions where IWC > 0.01 g m−3 during 02:30–08:30 UTC on 6 February 2006.

(4) Figure S1 shows that the simulated upper-level ice number concentration and IWC are broadly consistent with the available aircraft observations.

The reason why we had kept this comparison in the Supplement rather than in the main text is that it provides a supporting evaluation of the simulated upper-level ice phase, but it does not directly sample the mixed-phase layer where the SIP mechanisms discussed in this study are most active. We therefore consider this point has already been addressed in the submitted revised manuscript and Supplement.

• **CCN and aerosol constraints:** ACTIVE provides CCN and aerosol measurements that could be used to evaluate whether the simulated cloud droplet number concentration is consistent with observed aerosol–cloud relationships, particularly given the central role of prescribed Nd in the model formulation. Yes active CCN concentrations are not used to predict droplet number in the present simulation, but it is still possible to relate in convective cores the prescribed droplet concentration to an implied CCN concentration (e.g. the Twomey formula for a certain cloud-base ascent). Yes, there are no INP observations in ACTIVE/TWP-ICE but it is still possible to use the mineral dust loadings observed to infer with a dust-aware INP scheme the active INP concentration for comparison with the Cooper scheme used (which was not based on measurements anywhere near Darwin).

We agree that CCN, aerosol, and INP observations can provide additional constraints. In the present study, however, the main objective is to examine how different CCN conditions, represented by prescribed in-cloud droplet number concentration (Nd), modulate the impacts of SIP on cloud microphysics, radiation, and precipitation. For this purpose, Nd is prescribed directly in CASIM, so that the CCN-related perturbation can be controlled across the sensitivity experiments.

Following the reviewer's comment, the revised text now clarifies the observational basis for

the prescribed Nd values. Specifically, it states that these values are broadly consistent with ACTIVE observations reported by Connolly et al. (2013), which suggest cloud droplet number concentrations of approximately 150, 500, and 900 cm−3 for the monsoon break, pre-monsoon, and burning periods, respectively. This supports the use of Nd = 200, 400, and 800 cm−3 as representative clean, moderate, and polluted CCN conditions in this study.

We also agree that a Twomey estimate or a dust-aware INP calculation could provide additional information. However, such calculations would require further assumptions about cloud-base ascent, CCN activation, dust properties, and the choice of INP parameterization that are not part of the present experimental design. The conclusions of this study are based on relative sensitivity experiments using prescribed Nd and a fixed primary-ice treatment. We have therefore not included these additional calculations. The simulations retain the prescribed-Nd approach and the Cooper (1986) heterogeneous ice nucleation scheme, with INP concentrations not varied or linked to Nd, as stated in the model description.

Revised text (lines 187-189):
"*These values are broadly consistent with ACTIVE observations, which suggest cloud droplet number concentrations of approximately 150, 500, and 900 cm−3 for the monsoon break, pre-monsoon, and burning periods, respectively (Connolly et al., 2013).*"

• **Cloud droplet number concentration:** Dornier observations in ACTIVE include in situ cloud measurements (e.g. FSSP-based droplet concentrations in relevant cloud-intersecting segments, such as in weak cloudy stratiform ascent). These could be used to evaluate the model's representation of N_d in shallow cloud or stratiform boundary-layer conditions.

We agree that Dornier FSSP measurements can provide a useful constraint on cloud droplet number concentration in low-level cloud conditions. Following the reviewer's suggestion, we examined the available Dornier FSSP data for the 6 February 2006 case and matched them with the corresponding AIMMS position and height data.

However, after applying a threshold of LWC > 0.001 g m−3, the number of valid samples was very limited (less than 10 valid samples). In this case, only a few samples remained, confined to a narrow low-level altitude range around 0.4 km.

We also note that these measurements do not sample the mixed-phase layer where the SIP-related processes discussed in this study are most active. Therefore, these data are not sufficient for a robust quantitative comparison with the modeled cloud droplet number concentration. For the above reasons, we have not added an additional comparison based on these limited observational data.

• Cloud fraction and vertical structure: Following previous TWP-ICE/ACTIVE analyses (e.g. Kudzotsa et al.), comparisons of vertical cloud fraction profiles would provide useful context for assessing the simulated cloud field.

We agree that vertical cloud fraction profiles can provide useful context for assessing the simulated cloud field. However, the available observational cloud fraction estimates in previous TWP-ICE/ACTIVE analyses appear to represent broader campaign estimates, rather

than case-specific ACTIVE observations for 6 February 2006. Therefore, they are not directly matched to our simulation period in this study, and we have not added a direct comparison. Instead, the simulated vertical cloud structure has already been evaluated using CPOL reflectivity CFADs, while satellite OLR and CPOL precipitation have been used to assess the storm-scale cloud and precipitation evolution in the manuscript.

• Radiative fluxes: While domain-mean radiative quantities are shown, additional direct comparison with observational estimates would further strengthen confidence in the simulated energy budget.

In total, incorporating these elements would likely require only a small number of additional figures (on the order of ~2–3 multi-panel figures), but would significantly improve confidence in the model's baseline realism and therefore in the interpretation of the sensitivity experiments.

We agree. Following this suggestion, we have added a comparison between the simulated OLR and MTSAT-derived equivalent OLR in the Supplement (**new Figure S3**). Because the strongest SIP-induced OLR reduction occurs at Nd = 400 cm−3, Figure S3 compares the noSIP-400 and allSIP-400 simulations with the MTSAT-derived estimate. The comparison shows that allSIP-400 better captures the magnitude of the MTSAT-derived OLR minimum during 06:00–08:00 UTC, whereas noSIP-400 underestimates the OLR reduction. We focus this observational comparison on OLR because the available MTSAT infrared brightness temperature data provide an equivalent longwave emission estimate, whereas deriving outgoing shortwave radiation (OSR) would require additional calibration and conversion, which is beyond the scope of this analysis.

Revised text (lines 273-276): "*A direct comparison between the MTSAT-derived equivalent OLR and the corresponding simulations is therefore provided in Figure S3 for the Nd = 400 cm−3 pair. In this comparison, allSIP-400 better captures the magnitude of the MTSAT-derived OLR minimum during 06:00–08:00 UTC, whereas noSIP-400 underestimates the OLR reduction.*"

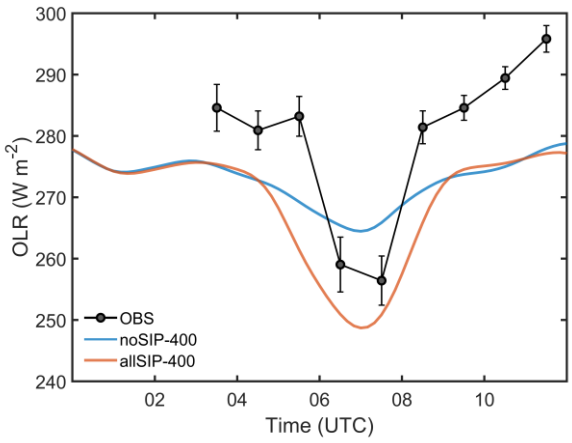


Figure S3. *Time series of domain-averaged outgoing longwave radiation (OLR; W m–2) from*

noSIP-400 and allSIP-400 simulations, compared with MTSAT-derived equivalent OLR. Black markers denote the MTSAT-derived mean OLR, with error bars indicating ±1 standard error. Observation times are shifted by +1 h for comparison with the earlier simulated storm development. The allSIP-400 simulation better captures the magnitude of the observed OLR minimum during 06:00–08:00 UTC on 6 February 2006, although both simulations produce lower OLR than observed before and after this period.

Finally, a more explicit discussion of the treatment of heterogeneous raindrop freezing (e.g. Bigg-type parameterisation) would be helpful. In particular, clarifying how this formulation relates to the assumed INP activity elsewhere in the model would improve physical consistency. Some recent approaches (e.g. Waman et al.) explicitly link raindrop freezing to INP concentrations, and it may be useful to discuss how the present approach compares to these frameworks. The current discussion of SIP processes (around line 660) would also benefit from a more quantitative comparison with these studies.

We have revised the discussion and now state that the formulation based on Bigg (1953) is represented separately from the heterogeneous ice nucleation scheme and is not explicitly linked to INP activity. We also clarify that this differs from Waman et al. (2022), in which raindrop freezing is related to INP concentrations.

Regarding a more quantitative comparison: we agree that such comparisons are useful, but a quantitative comparison is not straightforward because our study is based on sensitivity experiments with different SIP configurations, whereas Waman et al. (2022) used tagging-tracer approach to estimate the relative contributions of different SIP mechanisms to ice concentrations. Instead, we have revised the text to provide a clearer process-based comparison.

The revised text reads (lines 670-675):

"*UM-CASIM employs a formulation based on Bigg (1953) with modified parameter values, which is represented separately from the heterogeneous ice nucleation scheme and is not explicitly linked to INP activity. This treatment affects the number of freezing raindrops and graupel particles formed in the mixed-phase region where droplet shattering can occur. This differs from recent approaches in which raindrop freezing is linked to INP concentrations, such as Waman et al. (2022).*"