Convective controls on anvil cloud evolution in the ICON km-scale global climate model

This study investigates the relationship between convective updrafts simulated by the global storm-resolving model ICON and their corresponding anvil properties. The authors first introduce a couple of new methodologies to improve the three-dimensional tracking of convective cores and their linkage to associated anvils. Specifically, they propose a combined tracking framework in which different components of deep convective systems are initially identified and tracked separately, and subsequently merged to analyze the storm as a coupled dynamical system. These developments build upon the existing tracking library *tobac*.

Using the new tracking methodology, the authors identify and analyze deep convective cores and their associated anvils and updrafts over the Amazon region. The relationships between updraft properties and anvil characteristics are quantified. The results demonstrate that both the updraft area and updraft intensity influence anvil area, with updraft area emerging as the more dominant control. A key finding is that systems with the largest and most intense updrafts are associated with anvils that are up to four times larger in area compared to the smaller and less intensive convective system of the tracked distribution. This result suggests that changes in anvil extent are not linearly proportional to convective core properties, but instead reflect nonlinear and potentially amplifying dynamical processes that can substantially enhance anvil growth.

The paper reads very well and is of clear relevance to both the global km-scale modeling community and the storm-tracking community. The analysis is thorough, carefully executed, and clearly presented. I have a few minor comments and suggestions for clarification (also a few points that would be worth taking up in the discussion), but overall I believe the paper is suitable for publication once these points are addressed. No additional data analysis is required.

General minor comments:

- **Importance of vertical resolution:** Based on the presented analysis and development of the enhanced 3D tracking capabilities, do you have a sense of how important the vertical resolution is? Many km-scale models with similar grid spacing may have only ~ 50 vertical levels or so, depending on the vertical coordinates used. I am thinking in particular of Fig. 7e) which shows that you are not missing too many deep convective systems that have strong updrafts but are spatially disconnected from the anvil cloud. I wonder if that could potentially change with the vertical resolution.
- **Tropics vs. mid-latitudes:** Similarly, it would be helpful to discuss potential caveats if the same method was applied outside the tropics. Do you expect the tracking enhancements to work equally well or were they so tuned for a specific cloud regime that they might only be useful for the tropics?
- **Microphysics:** The presented analysis is based on cloud ice which heavily depends on the liquid vs. ice partitioning in the applied microphysics scheme. In addition, the (somewhat arbitrary) partitioning between different hydrometeor types as well as the particle fall speed that determines the removal of ice are factors that affect anvil growth. Can you just add a brief discussion which of the studied aspects might look different in simulations with different microphysical parametrizations?
- **Feasibility/performance of 3D tracking:** Can you provide some information about the performance of the 3D tracking (in terms of computational costs and time)? It would be great to know if the same analysis would be feasible to do on a global scale or if the analysis of storm-resolving models always still needs to be confined to a smaller region.

- **Open science:** Lastly, I wonder if there is a plan to publish the code enhancements as they would be hugely beneficial for the larger cloud tracking and high-resolution climate modeling communities.

Detailed comments:

Fig. 1. I suggest plotting the liquid and ice water path separately, because one could gain a little insight into the phase partitioning of deep convection. Since the entire analysis is based on cloud ice, it would be great to see if there is a significant area in the tropics that has quite high LWP values but no high IWP values.

Table 1: Why did the authors choose the same thresholds for the centroid detection threshold and the boundary threshold for each variable? Given that cloud ice shows huge variations in magnitude within a single convective cloud, would it not make more sense to have a higher threshold for the centroid detection and a lower threshold to define the boundary of the cloud?

Fig. 3: Since the authors decided on the ice water content threshold based on its distribution but chose a more arbitrary value for vertical velocities, I would suggest showing the ice water content distribution in this figure instead. This would contain more information that supports the choices made for the presented analysis. If you still want to keep this figure too, maybe attach it as supplementary material.

L. 9: I suggest replacing “convective intensity” with “updraft intensity” in the abstract, since this describes more accurately which variables were analyzed as the controlling factors for anvil clouds.

L. 52 - 55: Why is the “scatteredness” or lack of convective organization a result of the overestimated vertical motions and precipitation intensity? The causal relationship between those two sentences is unclear to me.

L. 81: Since the enhancements of the existing tracking library focuses on the linking method of features with complex geometries and combining tracks identified on two different atmospheric fields, I would suggest to be specific and clarify what “advancing the capability” refers to here.

L. 134: would replace “data-driven” by “data-informed” to not confuse with actual ML techniques

L. 194: What does “coincident in time and space” mean here exactly? It does not seem straightforward how to combine lifecycle tracks of two components such as updrafts vs. ice clouds. Is the same overlapping technique used that was used for these variables separately? Is the total lifetime of the DCC truncated by the co-occurrence of updrafts and cloud ice or is it possible to have updrafts first which are then followed by a connecting anvil cloud later on (as would be common because I believe there is a time lag between cloud ice formation and the initial updrafts that drive liquid condensation). A few more details here would be helpful.

L. 211: Refer to the nice visualization of this concept in Fig. 2. (panel 5)?

L. 217: This line is a bit confusing. Does this mean that DCCs that have been determined to be “isolated” could actually be part of a larger complex? If so, I do not understand the criteria made to distinguish between “isolated” and “complex”. The text suggests to me that disconnected updrafts can oc-occur in one system connected horizontally by cloud ice, but would that not exclude that systems with isolated updraft and cloud ice are still part of the larger organized cloud scene?

L. 391-392: Can you clarify whether any of the results of your study on the relationship between anvil thickness and convective cores, suggests a revision of previous assessments wrt anvil feedbacks? Or in other words, since radiative effects are the main motivation behind this investigation, can you conclude what the implications of increases/decreases in deep convective core intensity/area would be for ice cloud feedback assuming the relationship stays the same as in this study? Additionally, maybe you could also mention how NextGEMS specifically can be utilized to test whether the relationship will hold in a warmer climate or not.