Review of "Observing convective activities in complex convective organizations and their contributions to precipitation and anvil cloud amounts" by Zhenquan Wang and Jian Yuan.

Thank you to the authors for their responses to my comments. This revision is greatly improved relative to the initial submission. The text is much clearer, and the additional figures and text were very helpful, especially the revised section 3 and addition of Figure 2.

This paper is primarily one about methodology. The convective tracking algorithm is complex and sophisticated, and while the description of it was difficult to follow in the original submission, it is much clearer now, and the reader can understand the methodology being described. I appreciate the large amount of technical work that was done here, and think that this tracking system is potentially very useful to studies of tropical convection.

After describing the tracking algorithm, the authors use it to examine some properties of tracked convective segments in the West Pacific. While I agree that it is a good idea to show some results like this, I found this section to be much less compelling than the methodology itself, as the results are presented rather plainly without much context or discussion. In addition, this section still suffers from lack of clarity in some places, which makes it hard sometimes to understand the results. More details provided below.

General Comments:

1. **Clarity of writing.** While the writing in the paper is now clear enough to convey most points, it still does not read easily in many places, and the authors may wish to pursue more editing services if that is an option for them.

2. **Unclear definition of the OS life cycle.** Section 4 presents statistics about OS properties integrated over the OS life cycle. But since an OS is not the same as a whole convective system, it is not clear when the life cycle begins and ends. The lack of detail here prevented me from understanding and contextualizing the results. When exactly does the life cycle begin and end? How is it calculated when there are mergers and splits? For example, for OS #4 in the third row of Fig 5—when does the life cycle start and end in this case? The current explanation in Table 1 (last row) is not particularly helpful for the reader. A schematic may be helpful here.

   (Line 404 / Fig 7a) It is surprising, almost hard to believe, that so many OSs with cold-core-peak BT < 220 K would have a life cycle of just 1 hour. Convection with this BT is near the cold point, and I would expect the life cycle to far exceed 1 hour in almost all cases. Since the authors are using hourly BT images, a 1-hr lifecycle would mean that the OS appears in only a single image, which seems very strange for a convective plume with a BT as cold as 190 K. Is this real, or is this caused by something else (e.g., storms moving out of the study region, or a brief split in the OS before it merges back together). And for a 1-hr lifecycle consisting of just one BT image, how are the development and decay stages defined?

(Line 453 / Fig 9f) Because the life cycle is not clearly defined, the *N* parameter used in section 4 was confusing. Until *N* was introduced, I thought the results in section 4 were for individual OSs. I was then confused as to how a single OS could contain more than one OS. This must be because two OSs that merge together at a later time are actual considered as the same OS (as in Fig 5). Even so, I am not sure how *N* would be defined in many cases. E.g., if two cores merge together and later split back up, is *N* equal to 1, 2, 3, or 5?

3. **Motivation for section 4.** The results in section 4, if I am understanding correctly, are for individual OSs. It would be helpful to motivate this analysis more clearly. Why do we care about properties of individual OSs, as opposed to properties of entire convective systems?

Specific Comments:

- There is no comment or justification about the choice of the 50% threshold for the Dynamic Overlap Ratios (lines 237-245). Is this an arbitrary choice or based on previous work? Do the authors expect the results to be sensitive to this choice? Have there been any sensitivity tests conducted, or at least a tally of how often an OS match fails due to DOR below 50%?

- I understand why the situations in Fig 1c(iii) and Fid 1d(i, iii, and iv) are excluded based on DOR below 50%. These examples are obviously idealized cartoons used to explain the methodology, and I find them quite effective in doing so, but I am curious if situations like these would even be possible in the first place. Consider 1d(iii) as an example, which is the simplest because the structure of the OS is identical between the two time steps. In the first part of the tracking algorithm, cross-correlation is used to predict the location of the OS one hour into the future. This should predict a position that has maximum correlation with the actual BT11 observations from the next time step. So, why would the cross-correlation ever place the OS prediction in the spot that is shown in the figure? Shouldn't the maximum correlation occur when the two solid and dashed lines exactly overlap?

Nothing has to be changes here, since these are just cartoon examples and the real observations are more complex. But I'm curious if these is just an idealized example that would not actually occur, or if I am misunderstanding something about the tracking procedure.

- Line 315-327 / Fig 4a: I've read this paragraph several times and think I mostly understand the point the authors are making here, but I think some of it is still not getting though. I think the authors are just saying that fixed BT thresholds do not capture the structural complexity that still exists in the region where BT is colder than the threshold. But I am not sure how this related to Fig 4a, which I am struggling to understand. Why would most of OSs have cold-center BTs equal to their cold-core BT? Shouldn't the cold-center BTs always be warmer? And if they are indeed equal, wouldn't that mean that most cores are only ~1K colder than the rest of the convective complex? I am pretty sure I am misunderstanding something here, so it would be helpful to clarify this section.

- Line 343-344: "The results in Figs. 4c-e might imply that the OSs of colder cores have increased precipitation efficiency, which contributes to both more precipitation and anvil clouds." I do not see how the authors can claim that greater precip efficiency leads to greater anvil cloud area. What would be the proposed mechanism for this? Lindzen et al (2001) suggested the exact opposite, although I am not presently aware of any evidence for their claim that does not rely on model microphysics parameterizations. The authors find that storms with lower BT have greater precip efficiency, greater precip area, and greater anvil area. But this might simply mean that storms with lower BT are larger storms. To assess the relationship between precip efficiency and anvil area, one would have to control for BT. I suggest revision of this sentence. Another conclusion could be that the observed relationship between BT and precip efficiency might be expected – storms with higher precip efficiency generally have less dry air entrainment, which may allow updrafts to reach higher altitudes and lower BTs.

- How exactly is the lifetime-accumulated anvil fraction defined? Are you simply adding up the anvil areas from each BT image? The units are km^2, but if you are measuring area over a period of time, shouldn't the units be hours*km^2?

- Fig 9 / line 446-450. I imagine that much of the differences in anvil area and precip between four these subgroups can be explained simply by the differences in life cycle duration shown in Fig 9c. The fractional changes in anvil/precip seem to roughly line up with the fractional changes in duration. I would not expect this to be exact of course, but maybe this could explain most of the difference.

- Line 452: Is the difference in life cycle duration not another mechanism that could explain these differences?

- Line 467: if A and P are hourly anvil and precip, and N is the total accumulated number of OSs, I do not understand how AN and PN are the lifecycle accumulated A and P. Doesn't there need to be a life cycle duration term in here to achieve that result? E.g., PND, where P is mean hourly precip for a single OS, N is the number of OS, and D is the life cycle duration of each OS?

I do not doubt that the author's analysis and units are correct, but I think there is a miscommunication or mislabeling here.

- It would help contextualize the results in section 4 if the frequency of the four life cycle categories are provided somewhere. The authors state that simple life cycle events are rare, but I don't believe the numbers are not actually provided.

- In section 4, it would be appropriate to remind the readers that "anvil" as defined here still requires BT<260 K. In reality, much the area of detrained cirrus has BT warmer than 260. Berry & Mace 2014 and Sokol & Hartmann 2020 show that anvils with optical depth of 1-2 are extremely common, and Gasparini et al 2022 (DOI: 10.1175/JCLI-D-21-0211.1) showed that these anvils can have BTs warmer than 260.

- An interesting validation experiment for the tracking algorithm could be done using a cloud-resolving model with high-frequency output and a BT11 simulator. "Observations" could be taken from the simulation at every hour, and the tracking could be applied to those “observations”. The tracking results could then be compared to the higher-frequency model output to see if the segments are correctly tracked. This is a big undertaking and is *not* a suggestion for the current paper, simply an idea if the authors ever wished to further validate the algorithm while avoiding the uncertainties associated with the wind observations.

Minor/Line Comments:

- Line 311-314: I suggest revising this section, as I am not sure what it is saying after reading it a few times: “The complexity of convective organizations can be inferred from the cold-center BT11 of OSs. Only when the cold-center BT11 is 260 K is the OS of the isolated convective body. Under the fixed BT11 threshold, the OS of the cold-core BT11 that is warmer than the selected threshold cannot be identified. The OS of the cold-center BT11 that is colder than the selected threshold cannot be isolated from CCOs.”

- Fig 1c,d: the terminology in the labels is a bit confusing here, and it took me a while to figure out what was being shown. One possible revision is to label the solid lines as “OS position predicted by cross correlation” and the dashed lines as “observed OS position”.

- Line 247-248: “The variation in the cold-core BT11 is prior to the variation…and decay.” The use of the word “prior” here was a bit confusing – perhaps “considered first” instead?

- Line 329: “the colder OS” -> “a colder OS”

- Line 330: “the warmer OS” -> “a warmer OS”

- Line 434-435: “is more distributed”. Is there a word missing here between “more” and “distributed”?- Line 480: “anvil production is enhanced” – is the evidence for this just that the $\overline{N}A'$ term is positive? I am just trying to understand.