Response to Reviewer Comments

Reviewer #2

The manuscript is very well written, it introduces in detail a very complete tracker. This reviewer appreciates the very thorough and inclusive literature review of other existing algorithms as well as the authors decision to include the respective websites to advertise the other trackers. This reviewer also appreciates that the authors dedicate a whole section to provide the potential user with post-processing tools and visualization. This reviewer has never seen a tracker description manuscript with that many details and believes that this manuscript could serve as a model for future papers describing new trackers or new versions of existing trackers. The authors have also addressed the concerns from previous reviewers.

We thank the reviewer for recognizing our efforts on the manuscript and the software.

It is suitable for publication once the following comments are addressed:

Because the main goal of the paper includes demonstrating the capabilities of the PyFLEXTRKR, this reviewer thinks that some additional prove on how the algorithm manages splits and mergers should be included to support the skilled performance of the tracker (this reviewer acknowledges the conceptual schematic in Fig. 8). Adding statistics on how the number of split and mergers vary using different Tb thresholds sensitivity tests can be a way to go about this for a convective cell and/or MCS case. Statistics can be included in table form (an example can be following Núñez Ocasio et. al 2020 Table 2). This will be especially interesting to include given this tracker has the capacity to provide a comprehensive list of splits and mergers. Below a few minor comments.

We thank the reviewer for the excellent suggestion. We have performed a sensitivity test on a variety of Tb thresholds and the impacts on merge/split and MCS statistics. We added Table 1 and the following paragraph on line 370:

"While merging and splitting of clouds can occur naturally during the lifetime of MCSs, it is often more important to identify whether an MCS initiates or decays naturally (i.e., a complete lifecycle) as opposed to starting as a split from an existing system or ending as a merger to another system (i.e., a partial lifecycle). PyFLEXTRKR provides two variables (*start_split_cloudnumber*, *end_merge_cloudnumber*) to easily identify the status of each MCS at the track start and end time (a positive value indicates a split or merge). Table 1 shows the sensitivity of the fraction of MCS merging/splitting to T_b thresholds used to define CCS for tracking. As T_b thresholds increase, more MCS are identified, MCS maximum cloud area and lifetime both increases, and the fraction of MCSs that start as split or end as merge also increase from ~16% at the lowest T_b thresholds to ~45% to the highest T_b thresholds. This is expected because as larger part of the anvil clouds surrounding the cold cores or clouds with warmer cloud-tops are included in the tracking with higher T_b thresholds, the probability of merging/splitting would increase due to more complex morphologies as anvils spread and interact with nearby convection. The default T_b thresholds result in ~28% splits/mergers during the test period, suggesting that majority of the MCS tracks obtained have complete lifecycles."

Minor comments:

Introduction: Please include in the literature review GTG by Kim Whitehall: Whitehall, K., and Coauthors, 2015: Exploring a graph theory based algorithm for automated identification and characterization of large mesoscale convective systems in satellite datasets. Earth Sci. Info., 8, 663–675, https://doi.org/10.1007/S12145-014-0181-3.

We thank the reviewer for the reference. We have added citing this work in the introduction on lines 53-54:
"The "Grab 'em, Tag 'em, Graph 'em" (GTG) algorithm (Whitehall et al., 2015) also uses area overlap method combined with graph theory to track satellite IR brightness temperature (Tb) defined objects and identify MCSs."

Line 208: Why 5-km radius? Did you test other radii? Please include discussion on the reasoning.

We have added the following statement to explain this choice:
"The dilation radii in this example are tuned to work well for 15-min temporal resolution data (discussed further in Section 3.2). The dilation radii are user-defined parameters, which can be easily adjusted to adapt for different datasets and research applications."

Line 240: Also, TAMS

Thank you. We have added the reference.

Reviewer #3

The manuscript "PyFLEXTRKR: a Flexible Feature Tracking Python Software for Convective Cloud Analysis" by Feng et al. presents the Python based implementation of the FLEXTRKR and its application to a large variety of datasets including radar observations, LES model output, and global climate model data. I am impressed by the functionality and technical ability of this software package. Also, the parallelization capabilities are noteworthy. Additionally, the paper is well written, and the images are of high quality. I am convinced that this package will be very well received by the research community. I only have some minor comments and recommend publishing the article after those are addressed.

We thank the reviewer for the positive comments. Below please find our responses in blue.

Minor comments:

L28: "...the growth of weather and climate model dataset..." I assume that you are talking about the data volume here. Growth of datasets is easy to misunderstand.

Yes. We revised the statement as: "… due to the growth of weather and climate model *data volume*, …"

L109: Please add a brief statement about what you consider a contiguous area. Are these grid cells that cross a threshold and are adjacent in the x, y, or diagonal direction?

We added an example for a contiguous region:
"(e.g., grid cells exceeding a threshold and are adjected in the x, y direction)".

L330: I assume you mean gridded surface precipitation with the same time resolution as the Tb data here?

Yes. We revised the sentence as:
"When collocated *gridded mean* surface precipitation data is available *with the same temporal resolution as the* T_b *data*, …"

L413-6: I assume that you could run longer if you would not be constrained by the wallclock time on your HPC system. Discussing RAM memory and wallclock time constraints more explicitly might be beneficial.

We added the following two sentences in the revision:
"It is possible to run continuous global MCS tracking for even longer periods (e.g., multiple years), but system memory and HPC wall-clock constrain may negate such benefit, as tracking of individual years can be run in simultaneous jobs on an HPC system and completed in much shorter wall-clock time."

L465-70: While I agree with your interpretation of these results, I suggest to tune down the emphasis on model biases and equally emphasize observational deficiencies.

We revised the last sentence of this paragraph as follows (the *italicized* words are revised):
"…, while the stratiform cloud area and associated precipitation that dominate PF area are *weaker*, though *satellite* observational estimates of precipitation also likely contain biases (Cui et al., 2020; Li et al., 2022), *therefore caution is needed when interpreting the tracking results*.

L513-8: How much of this difference is due to the coarse model resolution (25 km)? Would these differences go away at km-scales?

We added the following sentences to address the reviewer's comments:

"This is likely due to the simulated precipitation being much weaker than observations, a typical bias associated with cumulus parameterizations in global models (e.g., Caldwell et al., 2019, Fig. 12). In state-of-the-art global convection-permitting models with kilometer-scale grid spacing, MCS precipitation is generally much better simulated, although challenges remain in faithfully representing the observed spectrum of deep convective systems (Feng et al., 2023)."