Review of "A composite approach to produce reference datasets for extratropical cyclone tracks: Application to Mediterranean cyclones"

February 9, 2023

This manuscript presents a new approach which combines tracks from individual cyclone tracking algorithms to produce composite / best estimate tracks. Overall the manuscript is interesting and well written and the resultant data sets will be of use to the scientific community. I have three main concerns: (1) how applicable this method is / how necessarily this method is to extra-tropical cyclones in other geographic regions, (2) how physically consistent the resultant composite tracks are and (3) how the benchmark data set was created. These are discussed in more detail below along with some other minor comments:

1 Major Comments

1. Applicability / relation to extra-tropical cyclones in other geographic regions.

   (a) The manuscript discusses extra-tropical cyclones and Mediterranean cyclones almost inter-changeably without acknowledging that Mediterranean cyclones are a special subset of extra-tropical cyclones. Better discussion of how Mediterranean cyclones differ from e.g. North Atlantic cyclones should be included in the introduction. This is important because certain characteristics of Mediterranean cyclones make them much harder to track than extra-tropical cyclones in the major storm track regions which raises the question – is this method of composite tracks necessarily in all locations?

   (b) The manuscript does not clearly acknowledge that many (although not all) of the tracking algorithms were first developed primarily to identify synoptic-scale cyclones in the major oceanic storm tracks and as such, were not originally designed to be used to identify Mediterranean cyclones. This should be acknowledged and it should be made clearer in the main body of the text which tracking algorithms were specifically designed to identify Mediterranean cyclones.

   (c) Are bogus tracks a much bigger problem in the Mediterranean than in the North Atlantic / Pacific / Southern Hemisphere storm tracks? e.g. is this largely a Mediterranean problem? These are mentioned frequently throughout the manuscript yet no evidence of these are given. Could some examples be included or at least some evidence to show a reader that these really do exist and are a significant problem? In addition, it is my understanding that by default many algorithms tracks many cyclones but then filtering is performed to remove weak / stationary cyclones. Does this pre-existing filtering in many methods remove the bogus tracks?

2. Physical consistency. A major question I have regarding this method is, is the final produced composite track still physically consistent with the spatial pattern and time evolution of the input data (e.g. with the MSLP data from ERA5)? I would very much like to see some evidence of this in the manuscript e.g. plot the cyclone centre from the composite track and the mean sea level pressure at multiple time steps and see if it looks as physically plausible as one individual track.

3. Generation of the benchmark data set

   (a) Line 247. Selection of the 120 cyclones that manual subjective tracks were created for. There is not enough details of how these cyclones were selected nor of these cyclones' properties (life time, intensity) relative to all cyclones. It is stated that these cyclones were selected based on previous case studies. As case studies tend to only be performed on extreme / unusual storms, this suggests that these 120 selected cases are not representative of the whole data set.

   (b) Line 249 – 255. Information given to meteorologists to identify the tracks manually. I think this approach is somewhat flawed as the meteorologists have only been given the MSLP field which is either the same or less information than the tracking algorithms have. This approach wants the meteorologists to behave like the automated methods, rather than the intelligent meteorologists that they are. In my opinion, the meteorologist should have been allowed many more fields / information to create a true bench-mark data set that really identifies true cyclone tracks (i.e in the way a forecaster would), not only those cyclone tracks which are evident in MSLP.

2 Minor Comments

1. Table 1. Could any additional filtering applied to the tracks (e.g. length or duration of the track) be added to this table. It would better allow a reader to see how the methods differ.

2. Line 149, Appendix L618. In the text (line 149) it is written that only cyclone tracks which last for at least 24 hours are retained. When checking the Priestley reference in

line 618, it appears that M07 only keeps tracks which last 48 hours (and travel 1000 km). Please resolve this confusion (based on Figure 6 I think tracks that last at least 24 hours are retained). Adding more details to the appendix for M07 would resolve this. It also appears that M07 only keeps tracks if the maximum vorticity exceeds a given threshold. This information should also be added.

3. Line 161 – 162. The distance criteria for whether tracks overlap or not is given here as 300 km. This threshold value is justified based on the scale of Mediterranean cyclones. Therefore, would this value need to be different if this method was used for e.g. north Atlantic cyclones which are generally larger but potentially easier to track? Aspects of this new method of combining tracks which are specifically for Mediterranean cyclones need to be more clearly highlighted.

4. Line 200 / Figure 2, panels e and f. These panels are hard to read. Would it be possible to only show the composite track / black line and the individual tracks for each tracking algorithm? i.e. the black dots of varying sizes could be removed.

5. Line 307. In my opinion M06 also has a prominent seasonal cycle (albeit with a different phase to M04 and M09) and should be noted here.

6. Line 313 / track density. How was track density computed and is it the same for all 10 methods? Please add these details. Related to this is in Figure 5 the track density goes to the edge of the domain for some methods (e.g. 4, 7, 9) whereas for others (1,3,6,8) there is some halo / gap around the edge. Why is this?

7. Figure 5. There is some overlap with the longitude labels on the bottom right of each panel. Could this be fixed?

8. Line 333 – 334. This is a confusing and redundant sentence. It is not surprising that the lowest values occur at the time of minimum MSLP. In addition, "mature" stage is not clearly defined and could be misunderstood to mean occluded / after the point of minimum pressure.

9. Figure 6. To be consistent with the other figures, could method 01 be at the top and M10 at the bottom?

10. Line 508: Suggest replacing "extra-tropical cyclones" with "Mediterranean cyclones" here.