Review of paper

LARA: a Lagrangian Reanalysis based on ERA5 spanning from 1940 to 2023

by L. Bakels et al.

submitted to ESSD

This is an excellent paper about a very interesting open-access dataset (LARA). The authors offer a state-of-the-art global Lagrangian climatology to the research community, which they calculated using FLEXPART and more than 80 years of ERA5 reanalyses. For the first time, such a comprehensive and valuable dataset can be used by researchers worldwide – and potential applications are numerous. The authors provide a selection of interesting example applications that nicely illustrate the potential of LARA. The paper is generally well written, but a few clarifications can further improve the paper. My main concern is that the first application is not at all easy to understand (the way it is presented at the moment) and this might give a wrong impression about the usability of Lagrangian diagnostics. Overall, I recommend accepting this excellent paper with minor revisions.

Comments (most of them are minor)

- 1) L17: "by time" sounds strange, maybe "as the time" works better.
- 2) L54: we typically write "airstreams" as one word
- 3) L69: should read "ERA-Interim"
- 4) Table 1: units should be "m s⁻¹" instead of "m/s"; symbols should be in math mode, e.g., *u*, *T*, *q*; not sure why you use rather unusual variable names like "sp", "2t" etc.; *p_s*, *T_{2m}* would be more common → consider using more standard variable names
- 5) L123: it would be interesting to have a bit more information about the 6 million particles: how did you decide for this number? What is the resulting mass represented by each particle? How many particles are on average in a 1x1 deg column?
- 6) L126: a bit strange "the full period ... would take ... to complete", maybe better "the calculations for the full period ..."
- 7) L128/129: confusing, first the overlap period is one year, then 3 months only?
- L132: FYI: the 300 s (= 5 min) time step correspond to 1/12 of the data input interval (1 h), which is the default approach used in LAGRANTO
- 9) L141: the periods are 8-y long

- 10) L148: I assume that BL height is directly from ERA5; what about the tropopause height? Which definition is used to calculate the tropopause height?
- 11)L149-151: I don't understand these sentences, why are BL and TP height not simply interpolated to the horizontal position of the air parcels?
- 12) L160: I don't understand "a month's worth of files normalised per variable"
- 13) L172: difficult to understand what these Spearman coefficients refer to (particle positions or tracer concentrations?)
- 14) L210: in line with my comment 5: please provide more information about the density of parcels typically available in a box of, e.g., 100 km x 100 km x 100 hPa
- 15) L220: "or are fully explored on themselves" sounds strange, I suggest something like ", nor are they investigated in full detail here"
- 16) L227: "periods that use different assimilation data" maybe this aspect could be discussed briefly in Sect. 2.1. The ERA5 dataset is as consistent as it can be, but there are still issues with changes in the datasets available for assimulation.
- 17) L238: "The left panel of figure 2" \rightarrow please use panel lables and write "Figure 2a ..."
- 18) L241: "within the lower parts of the lower atmosphere" is not very clear, can you give a pressure or height range?
- 19) L243: "in panel (a2)" should maybe read "in Fig. 2a₂" (however quite unusual), or you better change panel labels to (a, b, ... j)
- 20) L243-266: honestly, I am a bit lost with the text and the figure ... all rather complicated. Do you need so many panels to make the main point? Do you need the correlations with AMO, PDO, ...? If yes, then the reader requires a more careful explanation of what is shown and why and how to interpret the results. It would be a pity if the first application of LARA was so complicated that readers get the impression that Lagrangian investigations are hard to understand.
- 21) L285: 500 hPa should read 600 hPa
- 22) L290: I don't understand the criterion "air mass within a WCB per square metre to be larger than 1 kg", is this a criterion to guarantee a certain density of particles that fulfill the WCB criterion of ascent?
- 23) L293: maybe worth noting that your WCB climatology uses different units than, e.g., Madonna et al. if I interpret your unit correctly, then it refers to a vertical mass flux in WCBs (kg / m^2

/ s). The comparison with WCB frequencies (as in Madonna et al.) is therefore qualitative (which is fine, just maybe helping the readers to get along with the different units and values)

- 24) L321: I very much liked this application; it is easy to understand and clearly demonstrates something that could not been obtained by Eulerian analyses
- 25) L344: this could be better mentioned already in Sect. 3.1
- 26) L388: it is a very interesting results that "... ATCE_c values for specific humidity increase throughout the ERA5 period" do you have a hypothesis why this is the case?
- 27) L390: "Similar results to above" sounds strange, maybe "to the ones discussed above"