The authors did a great job in providing a global dataset with comprehensive information on the downwind areas of re-precipitation for evaporated water as well as for the upwind source areas of precipitation. Furthermore, it is the first global study of this sort using the recently published ERA 5 reanalysis data. The document is well written and understandable and provides interesting examples to exemplify possible uses.
I have a few comments as shown below:

Thank you for the encouraging words and constructive remarks.

1) The work does not provide a comparison of some of the results with previous work and a rationale for occurring differences. The recycling numbers (global land evaporation which re-precipitates over land & global precipitation over land originating from land sources) seem to be comparatively high. While reading the manuscript, the reader might interpret that differences to previous studies might solely be due to the usage of better data (most actual reanalysis data and finer resolutions). However, it would perhaps also be relevant to relate those relatively high numbers to errors in the precipitation and evaporation. Figure 8b indicates within this context significant differences (Deviation between UTrack and ERA5).

We agree that it is interesting to intercompare the recycling estimate, but the main goal of the current paper is the presentation of the dataset rather than this intercomparison. As the reviewer correctly points out, the literature on moisture recycling estimates uses different models and forcing datasets, and both these aspects contribute to the differences in moisture recycling estimates. Therefore, such a model intercomparison would require a more proper experiment, with different recycling models being forced with the same data. Nevertheless, we can speculate about the sources of the uncertainty in these intercomparisons. There are roughly two kinds of off-line moisture recycling models: those that focus on preserving the surface flux water balance (Eulerian schemes) and those that focus on preserving the moisture divergence balance as much as possible (Lagrangian schemes). In general, reducing the errors in one part of the atmospheric moisture path will increase these in the other part, but there may be a Pareto optimum. Unfortunately, the state of the art is that the moisture recycling studies do not report the errors in the moisture balance rigorously. (This also holds for our (OT and AS) previously published work.) By explicitly reporting the errors in our method (in Figure 8 in the current manuscript and by showing moisture tracking sensitivities in Tuinenburg and Staal, 2020 (HESS)), we hope to open the debate about such errors in the larger moisture tracking community.

2) The sample scripts seem not to work completely or there could be a bit more information on the necessary steps need to be done to get them running (e.g. how to derive the suitable netCDF file with monthly ERA5 data)

Thank you for this feedback. We added some more information in the sample scripts. Additionally, we added a script (ERA5_formatter.py) that can be used to sort the ERA5 data per month and to regrid the data to match the resolution of our dataset. These steps are also included in all other scripts that include ERA5 data.

3) Minor comment to caption of Figure 1:
Perhaps the last phrase could be formulated a bit more precise in order to avoid misunderstandings:
For instance from:
“The examples show the distribution of evaporation that precipitated (B) and the precipitation that evaporated (C) from Utrecht, the Netherlands, during 2008-2017, given as percentages of allocated moisture.”
To:

The examples show the distribution of re-precipitation for evaporated water from Utrecht (B) and the distribution of the city's sources of precipitation, given as percentages of allocated moisture.

Thank you. We changed this phrase to: "The examples show the distribution of re-precipitation for evaporated water from Utrecht (the Netherlands) during 2008−2017 (B) and the distribution of the city's source of precipitation (C), given as percentages of allocated moisture" (lines 441-442).

4) Regarding the PANGAEA dataset, it would perhaps be a nice add-on to have the results also as yearly averages. But users might of course simply build them by their own and this should not be considered as a "must-have".

We provide a script called yearly_average.py which calculates that. In this way we minimize the total size of the dataset(s).