Reviewer 2

In this study, a Lagrangian moisture tracking model driven by ERA5 reanalysis data is presented, and recommendations concerning both input data resolution and model set-up in light of accuracy and efficiency are given based on sensitivity experiments. This method is designed for tracking the fate, or as the authors phrase it, the ‘footprint’ of evaporative moisture released at point locations around the globe. Such a sensitivity analysis comes at the right time, as ERA5 drastically improved both the temporal and spatial resolution as compared to the widely used ERA-Interim in moisture tracking studies, yet this wealth of data is also accompanied with data storage limitations, and considerable simulation time increases. This well-written, nicely presented study may therefore boost scientific progress by providing guidelines on how to use ERA5 data most efficiently, and additionally, demonstrates the power of Lagrangian modelling for tracing moisture. Although generally easy to understand, some descriptions lack clarity, and I believe the manuscript could be ameliorated if the authors considered the following points listed below (not sorted), and thus become suitable for publication in HESS.

Thank you for these encouraging words.

- **Number of particles:** If I understand it correctly, then even the setting with the lowest number, 100 particles per mm of evaporation, is still high compared to some previous studies; in case of e.g. Manaus, for a range of say 3 to 6 mm of daily evaporation, this would correspond to 300 to 600 particles being released per day.
Studies such as Gimeno et al. (2012) and many, many others follow the method introduced in Stohl & James (2004) and are based on air parcels (that, altogether, represent the entire atmosphere), and not a ‘collection of water particles’, as is the case here, but I believe the number of particles may still be compared. Even though some studies based on the Stohl & James (2004) approach do not even mention the number of parcels (or ‘particles’), a lot of publications are based on simulations with 6-hourly timesteps for which 2 million air parcels for the entire globe are used (e.g. Sorí et al., 2017; García-Herrera et al., 2019, …). Läderach & Sodemann (2016) used 5 million parcels globally, which corresponds to an average of about 70 parcels per column residing over each grid cell at any given timestep, so that about 280 parcels are available per grid cell and day, compared to about 120 in many other papers (2 million parcels globally). Therefore, the setting used in this study with ‘only’ 100 particles per mm of evaporation, at least in the case of high evaporation rates as in Manaus, still results in far more particles than were employed in many peer-reviewed moisture tracking studies. Indeed, as the authors point out in the discussion, larger study areas and longer analysis periods likely decrease the number of particles needed. Still, it would be interesting to see an extended Fig. 4 for e.g. 50, 25 or even 10 particles per mm of evaporation, and this might also reveal a lower limit, since even ‘only’ 100 particles perform nearly as well as 10’000. Considering the near-linear runtime dependency on the number of particles, this information might be valuable.

Thank you for this suggestion and we agree. Therefore, we extended our sensitivity analysis for the number of parcels with 50 and 10 parcels per mm evaporation. We added two new panels to the respective figures in the main text (Fig. 4) and the supplement (Figs. S8-S13), and updated the respective sections in the Methods (2.2.2) and Results (3.2). Lowering the amount of parcels in this way had a small effect on the statistics and made the footprints less smooth. Therefore, we decided to keep the setting of our optimal model at 500 parcels mm^{-1} evaporation. Moreover, we mentioned in the methods section that the number of parcels used here is high compared to other moisture tracking studies (lines 194-196 in the version with tracked changes).

- **Particle vs parcel:** Related to the comment above, I am not sure if the usage of 'particle' throughout most of the manuscript is ideal. Particle by definition implies a small size, but depending on the number of particles (per unit of evaporation) used, each particle represents a considerable amount of water, and obviously not a tiny droplet or even a single water molecule, as one might first think. As of now, both 'particle' and 'parcel' are used in the manuscript, so I suggest to remove one of these to ensure consistency, preferably the former.

We now consistently refer to parcels throughout the manuscript.

- **Omega:** It is unclear to me why the authors decided to omit omega in their default configuration. In this context, it is important to point out that the statement on l. 423 is incorrect: the approach of Stohl et al. (2005) does use omega, yet it is complimented by random vertical displacements of air parcels to represent convective (vertical) redistribution.

Thanks, we changed the statement to reflect that the study by Stohl et al. (2005) does include omega and uses a scheme based on turbulent mixing to create additional vertical displacements (lines 487-489): "Therefore, moisture tracking models may complement the omega-based vertical displacement by using a mixing scheme based on turbulent mixing (Stohl et al., 2005), or …".

Also, the argument starting on l. 436, "We assume that this speed of mixing is rapid enough to supersede larger-scale vertical flows so as to simplify the model and exclude omega", seems to contradict a previous statement, in which the large effect of omega (vs no omega) is highlighted; moreover, evaporation footprints with and without omega are only similar for (probably unrealistically) fast mixing times of 1 hour, which strongly indicates that this aforementioned 'superseding' does not really take place for mixing at timescales of 24 hours. The lack of a proper reference, since the baseline model is essentially identical to the 3D Lagrangian model (a major limitation that is not really emphasized in the manuscript), clearly makes it difficult to justify any choice here; however, I am not convinced by the explanations provided so far, other than omitting omega 'for simplicity' and to achieve faster simulation times.

Thanks for raising this issue. We agree that we may have been too quick to disregard inclusion of omega for the weaker mixing assumptions. It is true that the baseline model is 'just' the 3D Lagrangian model, because it is quite hard to get reliable mixing, or source-sink relations of atmospheric moisture, especially in the context of the current study in which many sensitivity test are performed.

Nevertheless, we changed the manuscript in the method section to state that there are some sensitivity analyses for which we have a true value of moisture recycling (the numerical ones; time step, degraded resolutions, number of particles, etc.) and there are some for which we do not have a true estimate (vertical mixing, release height) (lines 173-176): “For some of the sensitivity tests, these criteria are evaluated against the simulation with the most detailed settings (most parcels, highest resolution, etc.), in which case there is a numerical true estimate. However, for some tests, there is no information to derive a true value. For these tests, the uncertainty remains higher and we derive the sensitivity of moisture recycling to the assumptions.”

We have changed the discussion (lines 502-505) and conclusion (lines 547-548) to reflect that the vertical transport is the main uncertainty regarding moisture tracking.

As a small note, given the data volume of the ERA5 data, the reason for excluding the omega data (or reducing the forcing data in general) has more to do with computer memory (RAM) use of the simulation than with the CPU time needed. Once all data is loaded into RAM, looking up specific values is done in about constant time.

- **Horizontal resolution:** Especially because of the authors’ valuable recommendation not to degrade the vertical resolution, it would be helpful for the scientific community to know whether the same is true for degrading the horizontal resolution. As stated in the manuscript, many previous Lagrangian studies are based on input data at 0.75x0.75°, or even 1.0x1.0° horizontal resolution, so that upgrading to 0.25x0.25° represents a massive improvement. However, this comes at the cost of extensive input data size, which makes studies covering multiple decades rather than a few days extremely challenging. I suggest adding this analogously to the degradation of vertical information; e.g. for increments of 0.25° from 0.25x0.25° up to 1.0x1.0°. For the same reasons, investigating the sensitivity to temporal resolution would be helpful too, but I am aware that the authors cannot include everything in their manuscript.

Thank you for this suggestion and we agree that testing for degrading horizontal resolution is a worthwhile addition to our paper. Therefore, apart from 0.25°, we performed our analyses also for 0.5°, 1.0°, and 1.5°. We added a new Fig. 8, supplementary Figs. S32-37 and new sections 2.2.6 and 3.6.

The horizontal degradation had quite considerable effects on the statistics, so we decided not to change the settings of our optimal model based on these new experiments.

- **Tracking time:** 99% of moisture allocated, or 30 days: Dirmeyer & Brubaker (2007) used 90% & 15 days, and nearly all tracking studies do not exceed 15 (or even 10!) days either, as the trajectory accuracy is known to decrease with increasing length (Stohl & Seibert, 1998). Due to how the model is set up, at least if I understand it correctly (see also below), this choice might not really affect the results, but I still suggest to check if the conclusions hold for considerably shorter trajectory lengths, such as 15 or 10 days.

We appreciate the suggestion, but we would advise the model users against using short tracking times. It happens often that after ten days not all tracked moisture has been allocated yet, in which case continuing the tracking will be better than terminating it. This happens especially over drier areas. Although we agree that the uncertainty increases with time, the quality of the forcing data is not reduced with longer tracking, as they are observation-based and not model-based.

- **Model description:** is insufficient. According to my understanding, tracked moisture remains static until precipitation occurs ('over' the grid cell underneath), and then decreases accordingly to the ratio of precipitation over precipitable water (of the entire column, I presume?). If so, this invokes another assumption, namely that each water (vapor) molecule within a tropospheric column has the same odds of condensing and precipitating (Dirmeyer & Brubaker, 2007). Please clarify.

Yes, this is correct. The allocation happens according to the ratio of P/PW, in which PW is the precipitable water of the entire column. This indeed assumes that every unit of water in the entire column has an equal probability of precipitating out of the column. We have added this fact and assumption explicitly in the model description section (lines 104-107): "However, once there is precipitation, a fraction of the moisture (precipitation over precipitable water of the entire atmospheric column, $\frac{P}{PW}$) that is still present in the atmosphere is allocated to rain out in that location. This assumes that all moisture in the atmospheric column has the same probability of raining out."

Minor comments and suggestions

In addition to the major comments above, a few additional comments and suggestions are listed here.

- l. 25: This sentence describes *continental* moisture recycling, not the more general concept of moisture recycling as it is most commonly defined (e.g. Brubaker et al., 1993; Dirmeyer et al., 2009). I suggest to include 'Continental' at the beginning of this sentence, and replace 'continental evaporation' e.g. by 'terrestrial evaporation'. This would also be consistent with the results section, where the continental recycling ratio (CRR) is used already.

Thank you, we revised it as suggested.

- l. 113: Not all Lagrangian models are initialized with a collection of water particles, but indeed, this is true for the approach presented in the manuscript. I suggest rephrasing this.

We rephrased this sentence to "In Lagrangian models, the internal model state is not a model grid, but generally a collection of water parcels." (line 128).

- l. 166: Is this the only instance where the (very short) study period is referred to? I think it would be justified to add a sentence in the discussion to remind the reader of this limitation.

In lines 533-534 in the Discussion, we added "… that we tracked only moisture for one month that evaporated during five days mean[s] that one should be cautious with generalizing the implications of these outcomes."

- l. 171: Is there any reason behind the choice of both Utrecht and Stockholm, which are rather 'close' both geographically and climatologically compared to all other point sources, other than the authors' affiliations? This is merely a question/comment, not a suggestion, since the conclusions drawn by the authors do not depend on this choice.

As you suspected, the choice of these two European study locations is no coincidence. Indeed, the two are relatively close, but we found it worthwhile to include a high-latitude evaporation source in addition to a more centrally located European location, as moisture recycling models have not always been performing well at higher latitudes. But apart from that, choosing both Utrecht and Stockholm helps us to communicate (the performance of) our model at our institutions.

- l. 265: Is the usage of time units (#particle mm -1 **h -1**) correct here? Besides, elsewhere in the text, the number of particles simply relates to some evaporation amount/volume (without referring to time at all), which is already a sufficient description to me.

Thank you for spotting this mistake. We deleted the reference to time in this sentence.

- l. 306: Do the authors have an explanation as to why the simulation time increases with fewer vertical levels? I do not doubt that this is correct, I simply find it counter-intuitive.

We agree that this is counter intuitive, as all simulations were run on the same computer. However, the simulations for the degraded vertical were run at a different moment than the other ones, which may explain the small difference of 1% in their CPU time.

- l. 423 contains an incorrect statement, as Stohl et al., 2005 do not disregard omega in their approach (also mentioned above)

We apologize for this incorrect reference and modified the example it was based on.

- Fig. 9: Why does B resemble Fig. 3D more than C? According to the text (l. 204), the default choice for the 3D Lagrangian model is 2'000 particles per mm of evaporation.

Figure 9B (now Figure 10B) equals Figure 3D, because both use the default settings of the model: no omega, 6h mixing, and 2000 parcels. All panels in Figure 10 use 2000 parcels per mm.

Further comments

Below are a few additional, language-related comments.

- l. 36: "Universal approach **es** and principle **s** moisture tracking models **is** that they

apply…", this reads a bit weird to me.
- l. 105: "courser" => coarser?
- l. 127: "weighed" => weighted?
3
- l. 185: "on full resolution" => at full resolution?
- l. 345: "CRR decreased much more rapidly ***with (increasing)*** mixing time than ...", maybe a word (or two) went missing here?
- l. 442: " ***The*** Continental recycling ratio", not sure if this is correct without any article

Thank you very much for these detailed comments. We fixed these mistakes. In the last case we changed it to "Continental recycling ratios …".

Concluding remarks
The paper investigates a multitude of choices and assumptions related to setting up a (Lagrangian) moisture tracking framework based on ERA5 data, and even though the main figures are only for a single location (Manaus) among a total of seven, and the analysis time is very short, I believe this choice is justified. Except for the discussion on vertical mixing and omega, as well as the interpretation on the sensitivity to the number of particles, the conclusions stated in the study make sense to me. Additional experiments (e.g. horizontal & temporal information degradation) and some more explanations might complete the package, hence I would gladly serve as a reviewer again and look forward to reading the revised paper.

Thank you very much.

References
Brubaker, K. L., Entekhabi, D. & Eagleson, P. S. Estimation of continental precipitation recycling. *J. Clim.* **6,** 1077–1089 (1993).
Dirmeyer, P. A., Schlosser, C. A. & Brubaker, K. L. Precipitation, recycling, and land memory: An integrated analysis. *J. Hydrometeorol.* **10,** 278–288 (2009).
García-Herrera, R. et al. The European 2016/17 drought. *J. Clim.* **32,** 3169–3187 (2019).
Läderach, A. & Sodemann, H. A revised picture of the atmospheric moisture residence time. *Geophys. Res. Lett.* **43,** 924–933 (2016).
Sodemann, H., Schwierz, C. & Wernli, H. Interannual variability of Greenland winter precipitation sources: Lagrangian moisture diagnostic and North Atlantic Oscillation influence. *J. Geophys. Res.* **113,** 1–17 (2008).
Sorí, R., Nieto, R., Vicente-Serrano, S. M., Drumond, A. & Gimeno, L. A Lagrangian perspective
of the hydrological cycle in the Congo River basin. *Earth Syst. Dyn.* **8,** 653–675 (2017).
Stohl, A. & Seibert, P. Accuracy of trajectories as determined from the conservation of meteorological tracers. *Q. J. R. Meteorol. Soc.* **124,** 1465–1484 (1998).
Stohl, A., Forster, C., Frank, A., Seibert, P. & Wotawa, G. Technical note: The Lagrangian particle dispersion model FLEXPART version 6.2. *Atmos. Chem. Phys.* **5,** 2461–2474 (2005).