**Reply to review of "The Lagrangian moisture source and transport diagnostic WaterSip V3.2", submitted to GMD by H. Sodemann**

**Reviewer #2**

*This paper describes the software WaterSip, a widely used Lagrangian moisture source diagnostic that is based on the accounting algorithm developed by Sodemann et al. (2008). Apart from diagnosing moisture sources for precipitation or vapor, WaterSip can provide information on moisture source conditions, transport, and arrival quantities. The paper gives an overview over the diagnostic method, describes how to configure and run WaterSip, and explains the different output files and variables using a test case in Scandinavia as an example. It also briefly discusses potential errors and uncertainties in the results.*

*WaterSip is a very powerful tool, but so far has been difficult to use due to restricted access and limited documentation. This paper will greatly enhance the accessibility and usability of WaterSip. It is well-written and nicely structured, providing useful guidelines for potential users. I also very much like the idea of providing a test case with all the necessary input and configuration files for running WaterSip, which is a good starting point for new WaterSip users. There are some problems related to the test case, but they can hopefully be fixed. Apart from that my comments are mostly minor, and I recommend publication after addressing those.*

**Reply:** Thank you for the thorough and constructive review. All comments are addressed below.

***General comments***

*1) I tried to run WaterSip with the provided test case, but it did not really work. During compilation it first did not find the netcdfcpp.h file, and I think this is because there is a mistake in the makefile. The NETCDFINC path should be added to the compilation step instead of the linking step:*

```
COMPILE = $(CC) $(CFLAGS) -c $(NETCDFINC)
FLINK = $(NETCDFLIB) -lnetcdf_c++4 -lnetcdf_c++ -lnetcdf -lsz -lz -ldl -lm
-lpthread -lcurl -lstdc++ -fopenmp
```

*During runtime there were some other problems:*
- *The startDate in the input file is after the last shortposit file provided on Zenodo. It should be changed to 20220811-000000 (or alternatively more shortposit files should be provided).*
- *The particle number maxPart is too low, I got the error "***ERROR: could not assign particle, maximum particle number exceeded!". With maxPart = 100000001 it worked.*
- *The reading of the input file ended up in an infinite loop in Parser::skipBlanks. I had to add filestr.get() on line 688 in Parser.cpp for it to work.*
- *After completing 92.2%, WaterSip crashed with a segmentation fault. I did not figure out why.*

**Reply:** I'm sorry about the trouble getting the software to compile and run, and thank you for testing out some solutions. Indeed, the makefile is currently not sufficiently system independent. In fact, the solution is to add the library path to the netCDF-C++ library include files to the makefile, then the netcdfcpp.h file can be found. An updated makefile will be made available with the resubmission.

Regarding the runtime problems, it appears that the wrong INPUT file was included in the archive, where the settings don't match the test data. In addition, the Parser.cpp routine was not sufficiently robust to handle additional blank characters that had been added to the input file for nicer formatting. An updated version of the Parser.cpp file will be made available with the source code.

Furthermore, to facilitate running the test case, a bash script will be made available with WaterSip that will download the test case, build the file structure implied in the INPUT file, and thus allow users to more easily run the test case.

*2) A new version of FLEXPART has been released recently (FLEXPART11, Bakels et al., 2024). Does WaterSip work with this new version as well, or only the 10.4 version? Since the new version writes trajectory output to NetCDF files, it would probably require rewriting the routines reading the input data. However, it might be worth it because FLEXPART11 has several advantages compared to FLEXPART10.4, for example (relevant for Waterip) improved trajectory accuracy and the option to write out average instead of instantaneous values along trajectories.*

**Reply:** This is an interesting question. Currently, WaterSip is not able to read NetCDF files from FLEXPART11. The simplest way to make this work is to create a post-processor that will convert netCDF FLEXPART11 particle files to the binary dump format used with FLEXPART10.4 and before. This information will be added to the revised manuscript.

*3) The section on errors and uncertainties is very short, but I think it is an important section and should be extended a bit. I would suggest to add a few figures showing the sensitivity of the results for the Scandinavia case to the settings, specifically the timeStep, the uptakeThreshold and precipThreshold, and arrivalRHMin. This would be very useful for new users to understand the influence of these settings. Also some potential error sources are currently not mentioned in this section, e.g. the fact that WaterSip always assumes either only moisture uptakes or only moisture losses during one timestep but not both. This could lead to an overestimation of remote and an underestimation of local moisture sources. Or, when using WaterSip as a diagnostic for surface evaporation, the assumption that water evaporated from the location (lat,lon) where the moisture is taken up by the air parcel might not always be true.*

**Reply:** A similar comment has been made by the other reviewer. While my general approach has been to refer to systematic sensitivity tests in earlier studies, I agree that it will be useful to include some sensitivity results for the test case in this manuscript. The additional error sources mentioned by the reviewer will also added in the revised manuscript.

*4) I did not fully understand the (difference between the) moisture source and transport quantities. For the moisture source quantities, the values are multiplied by $\Delta q * f * m$, for the transport quantities, they are multiplied only by $f * m$. I think the problem is that I don't really know what f and $\Delta q$ are in this case. Are these the values after the accounting? If so, isn't f exactly $\Delta q/q_0$? So why is then f multiplied by $\Delta q$ again? I am sure this is all done correctly in the code, but it could be explained a bit better (see also my specific comments on Equations 19 & 20).*

**Reply:** The difference between source and transport quantities is that the weighting of the source quantities is only done once for the final evaluated contribution from each source region, whereas the transport quantities are also gridded at each time step during the airmass transport. Regarding the question why $\Delta q * f * m$ is calculated rather than $f * m$ only, in the first case, the resulting quantity is the accounted mass of water (since m is a constant, giving the air mass of the particle).

**Specific comments**

*L9-11: This is a repetition of L2-4*

*L22: Bracket around Stohl and James (2004)*

*L74: chose -> choose*

*L108: Figure reference broken*

*L133: doe -> does*

*L151: What do you mean by „in case trajectories are used rather than air parcels"?*

*L154: I think this first sentence is not needed here.*

**Reply:** These items will be corrected in the revised manuscript

*Equation 12: Shouldn't this be the sum over i?*

*L168: This is interesting. Why does this (f_tot > 1) happen? This would be a case where the algorithm from Dütsch et al. (2018) is not equivalent, because there f_tot is by definition always <= 1.*

**Reply:** Yes, the summation should be over i. The f_tot > 1 can happen due to the delta_q_c threshold. This will be clarified in the revised manuscript.

*L233: Why „differences in"? Not just the source contributions themselves?*

*L266: Do you mean atmospheric properties? Because the positions come from the trajectory calculation, so that would be the second part of the sentence.*

**Reply:** These comments will be addressed by rephrasing the respective sentences.

*L314: What would be a reason for setting a maximum relative humidity?*

**Reply:** A hypothetical example would be in the tracing of vapour rather than precipitation properties, where one could be interested in cloud-free arrival locations only.

*L323: partcels -> parcels*

*L360-369: I didn't understand this part with the gridRadius.*

*L395: Could you briefly explain what these methods do, specifically Gustafsson et al. (2010) and Dirmeyer and Brubaker (1999)?*

*L403: Remove „are"*

*L436: Closing bracket missing*

*L467: chose -> choose*

*L483: Maybe start with a brief description of the meteorological situation for the event?*

*L498 (and others): Sometimes day is d, dy, or day. Please be consistent.*

*L515: Stohl et al. (2005) or Stohl and James (2004)?*

**Reply:** These comments will be addressed by rephrasing the respective sentences.

*L539-L545: I don't understand how the quantities moisture transport, air mass mixing, and rainout are obtained. Could you explain this better? For example, what is meant by gridded product of the specific humidity?*

**Reply:** The explanation of transport quantities will be rephrased and extended in the revised manuscript.

*L551: I don't understand this first sentence. Do you mean differences _in_ moisture sources and transport? But moisture sources are not shown here… (?)*

*L554: quantity -> quantities*

**Reply:** These comments will be addressed by rephrasing the respective sentences.

*L554-L557: How does the precipitation estimate by WaterSip compare to ERA5 precipitation? This would be a good validity check.*

**Reply:** See reply to major comment #3

*Equation 19: Shouldn't the denominator be P_tot (without k)? And the enumerator would correspond to T^k_0 * P^k_tot?*

**Reply:** Correct, thank you for pointing this out. This will be corrected and clarified in the revision.

*L579: Earlier λ and φ were used for lon and lat. Please use consistent notation throughout the manuscript. How is the mean over longitudes calculated? Are the coordinates converted to a Cartesian grid first? Otherwise there would be problems for e.g. lon1=-179 and lon2=179.*

*Equation 20: What is M? I assume the time steps of the trajectories? You could use this also in Equation 12 for consistency. Shouldn't Δq also appear in the denominator?*

**Reply:** Yes, the longitude averaging is done using conversion to a Cartesian grid. The notation will be adjusted to be consistent throughout as suggested.

*L583: zeta -> \zeta*

*L595: This is a detail, but wouldn't it be better to use the land fraction itself in the averaging (instead of 0 or 1)?*

*L600: and Eq. 19*

*L604 and L617: „the" too much.*

*L607: Maybe briefly explain what d-excess is or define it?*

*Equation 22: Is N the same as M in Equation 20? If so, use M instead?*

*L650: where -> were*

*L698: Would it be possible to provide a script for reading the .traj files? Otherwise this is difficult because of the binary format.*

*L763: arrival _in the_ target region?*

*L780: The link to the WaterSip source code does not work. The one given on Zenodo is correct.*

**Reply:** These comments will be addressed by editing the respective sentences and equations.

**Figure and table comments**

*Figures 5,6,7: I would use a different colormap. This one looks like topography and is not very intuitive for sequential values (e.g. white means high).*

**Reply:** I will try to find an alternative, suitable colormap for the revised manuscript.

*Figure 9: The locations of the colors don't correspond to the labels of the colorbar. Which one is correct?*

**Reply:** The labels have been flipped in the figure. This will be corrected, as well as potentially using a projection that shows locations north of 80ºN.

*Figure 10: Are these quantities weighted by f? From Figure 10 it seems so, but the text doesn't mention it. In (d) where can we see the time before arrival mentioned in the caption?*

**Reply:** Yes, the text will be updated accordingly. The caption to (d) gives the PDF as a function of distance, which will be corrected.

*Table 3: Why do the uptakes have negative values?*

**Reply:** The negative quantity is actually time before arrival in hours, rather than uptake, which is the gridded quantity. The table will be adjusted to make this clear in the revised manuscript.

*Figure 11b: I am not sure what is shown here. The figure says specific humidity, but the text says total accounted fraction. Where do we see that „moisture uptakes of more than 10 days before arrival are mostly overwritten by later uptake events“? I would suggest to use a different colormap, because red for moist and blue for wet is a bit counterintuitive. Also jet has many other problems (irregular lightness, not perceptually uniform, not colorblind-friendly).*

**Reply:** The figure gives the correct information, the text will be corrected in the revised manuscript. I will consider using a different colormap.

***References***

*Bakels, L., Tatsii, D., Tipka, A., Thompson, R., Dütsch, M., Blaschek, M., Seibert, P., Baier, K., Bucci, S., Cassiani, M., Eckhardt, S., Groot Zwaaftink, C., Henne, S., Kaufmann, P., Lechner, V., Maurer, C., Mulder, M. D., Pisso, I., Plach, A., Subramanian, R., Vojta, M., and Stohl, A.: FLEXPART version 11: improved accuracy, efficiency, and flexibility, Geoscientific Model Development, 17, 7595–7627, https://doi.org/10.5194/gmd-17-7595-2024, 2024.*

*Dütsch, M., Pfahl, S., Meyer, M., and Wernli, H.: Lagrangian process attribution of isotopic variations in near-surface water vapour in a 30-year regional climate simulation over Europe, Atmospheric Chemistry and Physics, 18, 1653–1669, https://doi.org/10.5194/acp-18-1653-2018, 2018.*

*Sodemann, H., Schwierz, C., and Wernli, H.: Interannual variability of Greenland winter precipitation sources: Lagrangian moisture diagnostic and North Atlantic Oscillation influence, J. Geophys. Res., 113, D03 107, http://dx.doi.org/10.1029/2007JD008503, 2008.*