Review of “An optimal transformation method for inferring ocean tracer sources and sinks” by Zika & Sohail for EGUsphere.

The paper presents a new approach (Optimal Transformation Method), rooted in water mass transformation methods, to infer changes in tracer distributions in the ocean interior as a result of ocean transport (circulation and mixing) and tracer sources/sinks. The novelty of this method is that it allows to separate the effect of air-sea fluxes, which often have biases, and mixing; this separation is not usually allowed by other inverse techniques. Also, the OTM method is not based on a steady state ocean circulation assumption, hence allowing to investigate changes in the ocean circulation.

The authors present an application of this new framework to a historical numerical model, after discussing the framework details with idealised case scenarios. This new framework is an interesting new approach, complimentary to other existing methods. The paper is overall very well written and some of the technical aspects of the methodology are clearly explained. I think this manuscript fits well in EGUsphere. Before publication, I think there are some aspects of the paper that need clarification. These are overall minor revisions, discussed below.

Comments:

- Line 48: More than in GF, it seems to me the method is rooted in transport matrix and water mass theory..?

- Line 118 (and following discussion at lines 123-127): Perhaps it might be worth to introduce a definition of a water mass? In the usual definition, which might not apply here, a water mass is defined as a “body of water with common formation history”, or a “body of water whose conservative properties are set by a single, identifiable process (and altered only by mixing)”. The conservative properties defining a water mass are most often set at the surface (some non-conservative properties can be acquired in the interior, e.g. an oxygen minimum, but most often that is not the case). Hence, why we usually describe properties in the interior as a linear combination of surface properties. My understanding is that in the OTM approach, the definition of a “water mass” is looser than the convention (e.g. line 118: using the definitions above, the mix of two known water masses is not a new, separate water mass), so it might be worth stating this difference from a conventional definition.

- Line 134: The reference to EMD is a bit confusing. Maybe I got it wrong, but my understanding is that $Q_{i,j}$ is the distance in tracer space between the early and late water masses due to sources/sinks. If that’s the case, it might be beneficial to write that explicitly in the definition of Qi,j at line 134, so that the following statement might become less confusing. Or rephrase/expand on the EMD reference (also because you are not using the EMD in the OTM, right?)

- Line 137: I think clarifying the point above about Qi,j definition would help to better understanding Eq.5. I was initially confused about gi,j acted on Qi,j.

- Line 148: What is the reasoning here? The previous statement says that the confidence in Qi,j is low, hence the confidence in the prior is low, correct? Why should the solution assume that Qadjust is small?

- Line 150: I might have missed it, but why is the cost function in eq. 7 called called “non-mixing cost”? Also, it was not until I read the Results section that it became clear that the steps are to (i) solve for gi,j in (7) and (ii) then calculate Qadjust in (8). I would suggest to state more clearly here.

- Fig 4: I am a bit confused by this figure. If I understand correctly, first the ocean is split in 16 T-S groups of equal global volume, and fig. 4 shows the volume of each of this groups in the 9 geographical regions considered. So, if we were to sum up the volumes across all nine regions per each water mass, we should retrieve the same volume? I might be reading this wrong, but I do not see this in Fig.4. Take for example the water mass defined by T[-2:5] and S[30:34.5] approx (bottom left box). This water mass has a relatively low volume compared to other water masses in almost all regions (but N. Pac, perhaps). It overall seems to be orders of magnitude lower than the volume of the water mass defined by T[-2:4] and S[34.5:35.5] approx. Again, I might be reading this wrong, so some clarification would be appreciated.
  Also: (1): it would be useful to add these boxes in Fig.3 and use the same x and y intervals and spacing, if possible; (2) only 14 of the water masses are visible in Fig.4, maybe change the axis to improve visualisation?

- Eq. 9: Ai is not the outcrop area at the early stage (right?), which is would one would most likely assume. Is it the outcrop area while transitioning between early and late periods? Some clarification would be useful. Also, should \Omega_i be \Omega_i(x,y,t) only (also in Eq. 19)?

- Line 311-312: Why don’t you attribute different adjustments, but Qadjust is the same for all i?

- Fig. 6: Perhaps change the colorbar for Qadjust, so that they are not just blank? Or remove the figure and just use the signal to noise reported (lines 328-329) to make the point that Qadjust << Qprior?

- Fig. 7: The number of points where transports can be inferred is limited by the number of regions selected, correct? Also, perhaps add the inferred transports for Case 2 and show that they are indistinguishable and change caption to mention both Case 1 and 2?

- Line 344: Can you add a justification of why you selected 5 W/m2 for the heat flux bias, and 5 mm/day for the fresh water flux bias? Why not larger/smaller (well, I guess larger would be more interesting) biases? And why not a percentage of the signal, rather than a fixed amount? And what if the biases were not uniformly positive/negative? Maybe I missed the point, but how could a fixed Qadjust reflect a mix positive/negative biases?

- Line 361-363: I think I am off here, comparing apple and oranges, but how does the result for the heat flux compare with the redistributed vs added heat? Can we interpret fig.9 (for the heat flux changes) as an indication that most of the ocean heat content changes are described by redistributed heat (gi,j explaining most of it), and only part of the changes are caused by added heat?

Minor comments:

- Line 18: Estimates (capital E)
- Line 19: Remove “However”?
- Line 38: Delete “[“ at the end of the line
- Line 70: “properties” misspelled
- Line 151: Add reference to section: “where wj is a relevant weighting (see section 2.5)”.
- Line 173: “early” and “late” in wrong order.
- Fig 3: colorbar label has kg spelled differently in the same label (Kg and kg)
- Line 291: Eq 15 (and not 7)?
- Line 339: we “find” (verb missing)
- Line 360: two “of”
- Line 361: add reference to fig. 9. Also, maybe change the colorbar for the adjusted heat flux?