Rewiever 2

1. The manuscript concerns ensembles of trajectories of tracers generated by ocean models. The authors wish to generate the "variability" of multiple ensembles by manipulating a single trajectory. The manipulations are primarily perturbing an initial condition in space, or by starting the trajectory at a different time. In order for such a study to be useful, pinning down exactly what "variability" means is crucial. The word "variability" is used extensively in the first several pages of the manuscript without stating what it actually is; I think perhaps on page 7, in relation to "connectivity", the reader begins to see what might be meant. As far as I could tell, according to this manuscript, variability means either "connectivity", or various types of entropy of coarse-grained future distributions of trajectories. A case is not clearly made for why these quantities are useful for ocean dynamicists or oceanographers. That is, why are these quantities the gold standard by which oceanographers should assess "sameness" of (collections of) trajectories.

We thank the reviewer for this important comment; they are absolutely right that the term “variability” needs a clear definition. In the revised manuscript, we have added the paragraphs below (lines 34-47 in the track-changed manuscript):

“*However, similar to above, the trajectories obtained from the particle tracking in a single OGCM ensemble member may not be representative of the full probability density of the system's state. Because pure advection is deterministic, there will be only one trajectory resulting from a virtual particle that starts at a certain place and time.*

This deterministic nature limits what we define as ‘trajectory variability’ – the range of possible pathways and end locations that particles could follow given uncertainties in ocean conditions. We define trajectory variability as the spread in particle positions, pathways, and connectivity patterns that emerges when accounting for uncertainties in initial conditions or modelled ocean states.

Capturing the trajectory variability is crucial for practical oceanography applications. For example, search and rescue professionals may want to compute a full probability density function of possible object locations – even when the starting location and time of an object lost at sea is known exactly – due to uncertainties in the ocean model. Similarly, marine pollution studies need to assess the range of possible contamination pathways, while connectivity studies in marine ecology require understanding the full spectrum of larval dispersal routes between habitats. In each case, a single deterministic trajectory provides insufficient information, limiting the generalisability of the results, as it cannot represent the inherent uncertainty in ocean dynamics and model predictions.”

2. The manuscript does not mention models where subgrid-scale dynamics is simulated e.g. stochastically. This would appear to be a very relevant set of comparators. It is also well known that trajectories are influenced by the resolution of the model grid, and that very different dynamics

can arise from the same model with different resolutions. This aspect is also not addressed; as far as I understand, only a single 1/4 degree model is used.

This point was also mentioned by reviewer 1, and we have now added a third set of experiments to the manuscript where we simulate particles with some Brownian diffusion added (lines 203-217 in the track-changed manuscript). We as well updated the results and discussion section to include these results (lines 397-399 and lines 514-519, respectively, in the track-changed manuscript)

Line 90: "The first strategy varies the release locations". Isn't this exactly part of what one does with ensemble generation? What is the difference?

We think there may be some confusion here between model trajectories in ensemble modelling and particle trajectories in Lagrangian analysis. We have thus now changed this to "*varies the release locations of the virtual particles spatially*" (lines 117-118 in the track-changed manuscript).

Section 2.2: I believe it is a poor choice to put the first results in an appendix. In fact, the manuscript reads as though it was recently chopped up and rearranged because it is impossible to read from start to finish via the Appendices. The Appendices contain definitions and details that the reader has not yet come across when reading from start to finish. I would strongly recommend removing the appendices and putting the material in the body of the paper to help with the narrative flow.

We have now moved the definition of the decorrelation time and length scales to the methods section of the manuscript (lines 152-166 and lines 186-200 of the track-changed manuscript). We also moved the Figure B5 (Figure 5 in the track-changed and updated manuscript) out of the appendix, showing the probability distributions of one ensemble member at different ages with different strategies, to the results section and adding a description of the figure in lines 428-443. This figure helps illustrate the probability distributions and the subsequent quantitative analyses for comparing the distributions. These modifications improved the readability of the manuscript.

Section 2.4: I could not understand how the probability distributions were being formed. The description is wordy, vague, and a bit sloppy. It needs precision and some formulas wouldn't hurt. Is a hexagon a bin? I could not find it stated.

We appreciate the observation. Based on it, we rewrote the sections describing the probability distributions and mixture probability distributions, by adding equations describing how we computed this distributions (Eq.(4) and Eq.(5) and lines 219-275, of the track-changed manuscript). We also added that these bins are indeed the same as the hexagons, by referring specifically to "*(hexagonal) bin*s" (e.g lines 299 and 355 in the track-changed manuscript).

Section 2.5: I could not understand what a mixture probability distribution is. This seems to be a crucial object in the manuscript, but the description was brief and ambiguous. Again, some formulas may help. There is a discussion about the optimal number of particles. In what sense optimal? Again the reader is referred to an Appendix that is too brief and does not provide any insight.

As stated previously, we have added some extra mathematical formulations to better explain the concept of the mixture probability distribution (lines 247-275 of the track-changed manuscript). We also made sure to stay consistent with the mathematical notation throughout the manuscript to avoid confusion.

Section 2.6: Connectivity is not defined, and it is not explained in the Appendix. What is it?

We have now clarified in the Introduction section that we define connectivity as "*a metric that maps the origin of substances (water, nutrients, plankton, plastic objects) to their destinations*" (lines 30-31 of the track-changed pdf).

Section 2.7: Similar to section 2.4, the section is written verbosely and somewhat sloppily, to the extent that I could not understand what the various definitions were.

We improved the section 2.7 "*Marginal Entropy and Relative Entropy Calculation*" (section 2.4.4 in the track-changed manuscript, lines 296-371) by homogenizing mathematical notation and variables used in the previous sections describing the probability distributions, to stay consistent and avoid confusion. This removes some ambiguities in the explanations of the concepts marginal and relative entropy. In general, we followed the notation and explanation style of Shannon's 1958 article, which is a is a mix of verbose and mathematical expressions.

Line 202: "ensemble of bins" what does this mean?

We now realise that using the word "ensemble" for anything other than the NATL025-CJMCYC3 ensemble could lead to confusion. We have thus replaced these uses of ensemble by "*set*" (e.g. lines 141, 152, 309 of the track-changed pdf).

Line 203: "t is the particle age of the distribution". This may make sense if the distribution was unambiguously defined at some point.

In the revised manuscript, we now mathematically define the probability distribution (lines 241 of the track-changed pdf).

Line 210: The authors write "P_A(X) = (1/2, 1/4, 1/8, 1/8)". As far as I understood from e.g. line 201, P(x_i) should be the probability of event x_i occuring, i.e. a number. Therefore since X=(x_1,x_2,x_3,x_4) -- see line 211 -- P_A(X) should equal 1, not a string of probabilities. This is just one example of the vague writing. If the authors really want P_A(X) to be a string of probabilities, that is fine, define some suitable object, and ensure that the writing is clear and consistent.

We thank the reviewer for highlighting this issue. We stated that the set of hexagonal bins is defined as $X = (x_1, .., x_B)$, where x_i is a hexagonal bin, and B is the number of bins in the domain (see lines 229-230 of the track changes pdf). Following this notation $P(X)$ represents the probability distribution over all hexagonal bins and $P(x_i)$ is the probability of finding particles in one bin. This was added in lines 243-245 of the track changed manuscript.

In summary, on the basis of both the writing and the scientific impact, I do not recommend publication.

We thank the reviewer for their comments, and hope that our responses have clarified some confusion so that the reviewer can be more supportive of this revised version.

The manuscript could be improved by going back to the drawing board and asking what exactly are the properties by which it is *most meaningful* to oceanographers to compare trajectories or ensembles of trajectories. Strong justifications and illustrations would need to be provided. Comparing with ocean models of different types, including stochastic components, and across different grid resolutions would add to the robustness and generalization of the subsequent results. A linear narrative and a much more precise presentation would also be required.

This is a good point by the reviewer, that we have taken on board. We think that in this revised manuscript, we have balanced further clarification of what many oceanographers find meaningful (connectivity!) and that the added comparison to adding Brownian motion is meaningful.