We thank the reviewer for recognizing the important issue underlying our work. We have below carefully addressed all the comments.

These are all good questions, that we now clarify in the revised manuscript. Specifically:

1. We have now clarified in the main text that we computed the decorrelation length scale "*of the initial particle velocities at the release location on the release day (2 January 2010)*" (lines 144-146 of the track-changed pdf). We used this to inform the size of the spatial release patches because it gives an estimate of the flow scales *at the moment of release*, so we think it is a relevant metric.

2. We moved the description of the correlation time and length scales to the methods section to improve readability (as suggested by reviewer 2; lines 155-166 and lines 186-200 of the track-changed pdf). We moved Figure A1 to main text, as Figure 2. We also added the equation that we used for calculating the decorrelation length and time scales (Eq.(1) and (2) of the track changes document).

3. We clarified that these metrics were computed as particle-pair correlations to inform the correlation in the spatial and temporal varying initial conditions used to generate variability (line 152 and 186, respectively, of the track-changed pdf). However, we explicitly state, lines 169-172: "*While decorrelation scales likely evolve over the 6-year simulation period due to particle spreading and varying flow conditions, computing their evolution at every particle age is computationally intensive and impractical, compared to other metrics.*"

4. We have now clarified that the virtual particles were advected for six years (from 2 January 2010 to the end of the hydrodynamic data availability on 31 December 2015) – or slightly less in the case of the temporal varying release, where some particles were released up to 20 weeks later (lines 101 of the track-changed pdf). The time-of-flow is the same as the particle age that we use in e.g. Figures 3 and 4.

5. We also clarified in lines 139-141: "*The choice of spatial release radii (9-180 km) spans the range from sub-mesoscale to mesoscale oceanographic features, allowing us to test how initial condition uncertainties at different scales affect long-term particle dispersion.*"

2. It seems that all calculations were done with 2 January as the release date. Given the unsteadiness (time-dependence) of the Eulerian velocity field, one will of course get different results if using a different release date. From an oceanographer's viewpoint, why is 2 January so important? Alternatively, there needs to be convincing results to show that "similar" results are obtained for different choices of release dates. Also, one needs some understanding of the sensitivity of the results of the time-of-flow chosen.

The reviewer has a good point that there is nothing 'special' with 2 January 2010. We chose it for the practical reason that it allowed us the longest particle advection time given the length of the CJMCYC3 model fields. However, since our goal was to introduce an analysis methodology, we feel that a full analysis of how the results change for different release dates is beyond scope (and computational resources available).

We have now clarified this in the limitations paragraph of the discussion section of the revised manuscript: "*We performed simulations for only one release period, on 2 January 2010, because that allowed us the longest particle advection time; but realise that there is nothing 'special' about that 2 January 2010 and that the results presented here might depend on the release time*" (lines 653-656 of the track-changed pdf).

Our advection strategy was apparently not clear in the first version of our manuscript. All particles travel until the end of 2015. We have now clarified in the methods section that "*all particles were advected until the end of 2015, so that in this simulation some particles reached a maximum 'age' (time-of-flow) of 6 years and others only 5.6 years. This is a minor effect though, as most of our analysis will focus on the first few months of advection only.*" (Lines 181-183 of the track-changed pdf).

As also mentioned in our previous response, we seem not to have been clear enough about our temporal release experiment strategy. The point of a temporally varying release is to seed particles that will have different trajectories, as the temporal decorrelation functions suggest. This is in fact the release strategy that is currently most widely used by Lagrangian ocean modellers to simulate different flow regimes, and the motivation of our study is to investigate how to assess whether this release strategy is equivalent to sampling different ensemble members. We have now further clarified this in the Introduction: "*Missing variability in particle trajectories is typically created by releasing particles at different locations (spatial variation; e.g., Rossi et al 2013) at different times (temporal variation; e.g., Qin et al 2014, van Sebille et al 2015) and/or with a small amount of random walk diffusion added to the advection (e.g., Hart-Davis and Backeberg 2021)*" (lines 64-70 of the track-changed pdf)

Yes, the reviewer is correct that indeed we added a small amount of spatial variability in the temporal-variability runs. However, as we also explain in that paragraph "*we kept this noise amplitude small to (as much as possible) isolate the effects of the temporal release strategy alone*" (lines 179-180 of the track-changed manuscript).

This is an interesting perspective by the reviewer. We have now added this relation to streaklines in the introduction section, where we also cite the papers suggested by the reviewer: "*In fluid mechanics, this is related to the concept of streaklines, transport barriers and coherent structures (e.g., Haller 2004, Zhang 2013, Karrasch 2016, Balasuriya 2017)*" (lines 51-52 of the track-changed manuscript).

The reviewer has a good point here. In the revised manuscript, we have also added the results of a simulation where the particle spread was created by adding a small amount of diffusion on top of the advection. We described the method to add diffusion in section 2.3.3 in the revised manuscript, in lines 203-217 in the track-changed manuscript. As well, we updated Figures 4, 5, 6, 7 and 8 in the revised manuscript, to include the results from the added diffusion release strategy. As well, we updated the results section to include diffusivity (lines 397-399 and lines 514-519, of track-changed manuscript).

In the discussion section (lines 601-614 of the track-changed manuscript), we added: "*The added diffusion simulations use a stochastic differential equation approach with Brownian motion terms, following established methods in Lagrangian oceanography (Griffa, 1996). However, our results demonstrate that this approach has limitations in reproducing the full ensemble variability compared to spatial and temporal releases, particularly over particle ages < 100 days. We chose a physically reasonable diffusion coefficient based on literature values for subgrid-scale parameterization at our model resolution. Larger diffusivity coefficients could potentially increase trajectory variability, they would become unphysical as the artificial noise would exceed the magnitude of typical unresolved subgrid processes in the model, effectively forcing the particles to behave as if experiencing constant strong turbulent mixing. The underperformance of the diffusion strategy likely reflects that it represents primarily small-scale turbulent mixing processes, and the diffusivity cannot create noticeable variability in trajectories, specially for particle ages 100 days, missing the larger-scale flow uncertainties and mesoscale variability captured through other strategies.*"

The reviewer raises an interesting point here; these approaches are indeed nicely complementary. Hence, we have now added a few sentences to the introduction of our revised manuscript: "*Our approach is complementary to other new approaches for computing stochastic sensitivity of Lagrangian trajectories in the ocean, such as those by Balasuriya (2020), Badza et al. (2023) and Branicki and Uda (2023). However, our approach is particularly also useful for particles with added `behaviour', such as in the case of plastic particles (e.g., Denes and Van Sebille, 2024)*" (lines 73-76 of the track-changed pdf).

9. The authors use several quantifiers for dispersal based on their Lagrangian simulations: mixture probability distributions, connectivity, entropy, and KL-divergence. Computing each of these requires the final Lagrangian distributions. So, interpreting any of these depends strongly on how those Lagrangian distributions were obtained in a physically reasonable fashion (see my earlier points on this). Because of this, it is hard for me to interpret any of these information-theoretic results.

We agree with the reviewer that the fidelity of our results depends crucially on the details of the Lagrangian integration scheme. We expect that the changes we made in response to other comments already help to clarify our methodology, and to even further highlight where details on the trajectory computations can be found, we have restructured the Methods section so that there is now a separate section (2.2) "*Lagrangian simulations*" (line 101 of the track-changed pdf).

10. It appears that a major point that is being made is that the full-ensemble variability is being approximated by single-member simulations. This is stated in many places, but I am a little uncertain as to whether my interpretation of this statement is correct. My understanding is that "single-member" means one particular initial condition is chosen. By choosing "50 members" (line 158), Lagrangian simulations associated with 50 different initial conditions are chosen. Then the probability distributions associated with each of these 50 are combined in a mixture model to get the "full" distribution. Is this understanding correct? If so, it would appear that the "full ensemble" can be thought of as the 50 simulations collectively, and so this is collectively the "full ensemble"? If this is so, the claim that the "full-ensemble" is obtained from "single-member" simulations, presumably demonstrating some advantage, isn't too different from simply saying that one is choosing a full-ensemble comprising 50 members, presumably chosen to cover a region of interest at the initial time sufficiently well? If this is not true, things have not been expressed clearly. Basically, I think some clarification on this claim is necessary, explaining exactly what is meant and how it is achieved (and exactly what a "single-member simulation" means).

The reviewer is right that the term full-ensemble variability is somewhat unclear. In the revised manuscript, we have replaced it with "*variability of the full 50-member ensemble*" (lines 6 in the abstract) and "*Our goal is to reproduce this dispersion of what we refer to as the `full ensemble' using just one ensemble member*" (lines 115 of the track-changed manuscript)

We agree with the reviewer that the fidelity of the results depend on the quality of the Lagrangian integration scheme and choice to include pure advection as well as stochastic approaches. We hope that our responses above have convinced the reviewer that this is the case. However, we would also like to highlight that our proposed analysis approach would have value by itself to other numerical oceanographers wanting to analyse their own ensemble simulations, for their own region of interest.

Based on my comments above, I feel that a major revision would be necessary for this article to be acceptable for publication in Nonlinear Processes in Geophysics.

We hope that our changes help convince the reviewer of the quality of our work, and thank them again for the very useful comments and ideas that have significantly improved our manuscript.