Answers to comments by reviewer #1

General comments:
This paper briefly describes the implementation of synthetic float capabilities in E3SMv2, then presents insights gained from the synthetic floats in several case studies. The manuscript is interesting and well written. However, even though I loathe gatekeeping based on journal scope, I do question whether the current form of the manuscript is suited for GMD as a development and technical paper. I would suggest more discussion about how the floats were implemented in the model and an evaluation of related development topics, like how much using the floats impacts the computational cost and whether it scales with the number of floats or the number of output variables saved, and also a more quantitative rather than qualitative evaluation of the model simulation accuracy.
We thank the reviewer for taking the time to provide comments on our manuscript. Inclusion of profiling floats increased the computational cost of the simulation by about 50% and scaled approximately linearly with numbers of processors, floats, and variables. However, as a proof-of-concept simulation, no attempt was made to optimize the new code's performance. In particular, interpolation weights from BGC tracer locations to particle locations were unnecessarily recalculated for every tracer which certainly caused significant slowdown. As suggested, we will provide more technical details on the new model developments including the computational cost in the revised version of the manuscript. Please also see our answers to all comments below.

Specific comments:

Lines 80-83: It should be clearer here that only the MPAS-O and -Seaice components are being used, and not the coupled land or atmosphere. This is stated in Section 2.2, but introducing the atmosphere and land components here gives the wrong impression that they are being used.
We have added a statement in the revised section 2.1 to make that clear:

"While we assess an ocean-sea ice-only simulation in this study, i.e., a simulation without coupled atmosphere and land model components (see section 2.2), we note that the new technical development described below can equally be used in the fully-coupled mode."

Section 2.2: For reproducibility and meeting interests of GMD readers, this section should provide more details about the model, particularly other components of the forcing like river runoff, atmospheric deposition, and atmospheric CO2 concentration if they are included in the forcing.
Freshwater from rivers is provided with the JRA forcing; nutrients inputs with river discharge are invariant in time and taken from Mayorga et al. (2010; GNEWS model; DOI 10.1016/j.envsoft.2010.01.007). The model uses the climatology of Luo et al. (2003; DOI: 10.1029/2003JD003483) for atmospheric deposition of dust and iron, and nitrogen deposition fields of Lamarque et al. (2010; DOI: 10.5194/acp-10-7017-2010). Atmospheric CO_2 is held constant at 405 ppm for the duration of the proof-of-concept simulation. We will provide this information in the revised method section.

First paragraph of Section 3.1: I think this paragraph buries the lede. It begins by discussing where 1000 m velocities are high and low in the model, and then showing that these Eulerian velocities are similar to those derived from the Lagrangian floats. Only then does it note that the model velocities are 2--3x lower than estimated from actual Argo floats. These seems to me to be a critical bias, and the impacts of it should be discussed and not glossed over.

We thank the reviewer for this remark. In the revised paper, we have revised the starting paragraph of this section to more clearly state the motivation behind the evaluation (also in response to comments by reviewer #2). Further, we have revised the text on the evaluation of current velocities to more clearly state the main take-away already in the topic sentence (instead of towards the end of the paragraph). The revised first two paragraphs of this section read:

"We evaluate the synthetic float capabilities in E3SMv2 in two ways: 1) by comparing the synthetic float data to the full Eulerian model output, we ensure the sampling by synthetic floats technically functions as intended and is sufficient in terms of spatio-temporal coverage, and 2) by comparing the synthetic float data to Core Argo data (Argo, 2023), we evaluate the extent to which the new synthetic observing network can be used for real-world applications. Specifically, we evaluate whether ocean currents at 1000 m, i.e., at the float parking depth, are adequately represented and whether environmental variables, such as temperature, salinity, and nitrate, are adequately simulated in E3SMv2 for realistic float-based sampling.

The simulated pattern of current velocities in E3SM agrees with an observation-based estimate, but current speeds are overall biased low in the model. In E3SM, current velocities at 1000 m are highest in the Antarctic Circumpolar Current and in the subpolar North Atlantic off the southeast coast of Greenland (locally >8 cm s-1), with velocities of less than 3 cm s-1 elsewhere (Fig. 3a). Fig. 3b shows a Lagrangian-based velocity estimate derived from all 10-day synthetic float positions which were averaged within 3°x3° boxes. The spatial patterns and magnitudes in current velocity produced by the Eulerian model output are largely captured by the synthetic floats (cf. Fig. 3a & b). The equatorial regions are the only exception, for which the Lagrangian E3SMv2 estimate suggests higher velocities (up to 4 cm s-1) than the Eulerian estimate (up to 2 cm s-1). This implies substantial variability of current speeds at 1000 m in E3SMv2 at sub-monthly time scales that are not captured by the Eulerian time-averaged output. In comparison to Argo-derived current speeds (Zilberman et al., 2022, 2023), velocities in E3SMv2 at 1000 m are a factor 2-3 too low (compare panels a-c in Fig. 3). This bias in ocean current speeds is a common feature in non-eddying ocean circulation models and is possibly related to how high-frequency dynamical processes are parameterized (e.g. internal mixing or tides; Su et al., 2023), in addition to limitations related to grid resolution. In spite of this bias, most high-velocity features present in the Argo-derived data set are reproduced in E3SMv2 (Fig. 3b & c). The only exception is the Gulf Stream, which is too shallow in E3SMv2 (not shown), resulting in much lower current speeds at 1000 m in E3SMv2 than in the Argo-based estimate in the northwest Atlantic (<2 cm s-1) compared to ~12 cm s-1)."

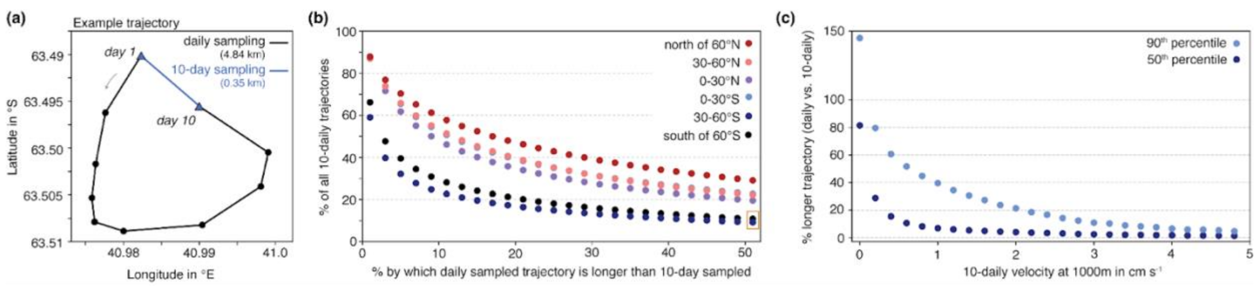
Section 3.1 in general: it would be good to have more quantitative comparisons between the model results and the Argo obs or between the model Eulerian/Lagrangian results. For Fig. 3, for example, it would be nice to know some combination of correlation, mean or mean bias, and RMSE. In addition to the speed it might also be good to know how accurate the current directions are, using something like the difference in direction.

In the revised manuscript, we will report some quantitative estimates of model-observation mismatch, e.g., by calculating the metrics the reviewer listed. While revising the manuscript, we will also explore the differences in current location. However, we note that while estimates from real-world floats are affected by lateral current shear during ascent, which will impact the diagnosed current direction (see e.g., Wang et al., 2022), our synthetic float-based estimates are not, complicating the comparison. We have thus not made a decision yet whether this is worth including in a revised manuscript.

Figure 5 and the associated text took a lot of time for me to understand, and I'm not entirely sure I still do. Maybe explicitly spelling out how the two values (daily, 10-daily) are being compared would be helpful (for example, maybe the x-axis of 5a would better be labeled as something like "% by which daily sampled trajectory is longer than 10-day sampled"). Also some clarity on how a longer trajectory means a lower velocity might help.
For the revised manuscript, we have changed the x-axis label as suggested by the reviewer. Additionally, we have added a new panel to illustrate differences in the trajectory length with one example (see new Figure 5 below). Lastly, to enhance clarity, we have added two new sentences at the beginning of the section to explain how trajectory lengths relate to velocities. The new text reads:

"Only knowing the position of typical Argo floats upon surfacing every ~10 days, Argo-float derived velocity estimates are subject to uncertainty stemming from the assumption of a linear trajectory between any two positions (see example from a synthetic float in Fig. 5a). With velocity calculated as the distance traveled per 10 days, a shorter trajectory length for 10-day sampling (blue line in Fig. 5a) than for daily sampling (black line) implies that the velocity derived from 10-day positions is underestimated relative to that derived from daily positions."

New Figure 5 for the revised manuscript:



Section 3.3.1: I don't understand why we're comparing monthly Eulerian model output with daily synthetic observations when, as summarized in the first paragraph of this section, it's more common to have daily Eulerian model output and sporadic observations.
The choice to compare monthly Eulerian model output to daily synthetic observations was dictated by what full Eulerian model output we had available when doing the initial analysis. We thank the reviewer for commenting on this inconsistency, which made us realize that this choice was probably more confusing than helpful. For the revised manuscript, we will homogenize the two data streams. We are currently re-assessing this case study using two approaches, i.e., 1) averaging the synthetic float observations to monthly means and 2) re-running the 6-year test simulation to produce daily mean Eulerian output, and will decide on one of the two for the revised manuscript.

Section 3.3.4: Is there any nuance to how the bloom timing is defined or is it just the day of year of the peak biomass?
Indeed, we define the bloom timing as the day of the year when biomass is highest. We have revised the text to make that clearer:

"In this case study, we assess the ability of float networks differing in float density to capture subsurface phytoplankton bloom characteristics, i.e., the timing (day of the year) and magnitude

of phytoplankton biomass peaks, in the subtropical Pacific, where subsurface maxima in phytoplankton biomass are commonly observed (Cornec et al., 2021; Yasunaka et al., 2022)."

Minor suggested edits:

Introduction: "e.g." is used very often
In the revised introduction, we have deleted some of the "e.g." from the original text.

Line 25: "have already permitted" -> "has already permitted"
Changed as suggested.

Line 54: "uncertain" or "unknown" may be a better choice than "less clear"
We rephrased this sentence to now state "uncertain".

Lines 120--121: This could explicitly state that "away from continental shelves and slopes" means "> 2000 m", at least according to the caption of Fig 2a.
The sentence now reads:

"Only floats seeded in the open-ocean away from continental shelves and slopes, i.e., at a water depth >2000 m, are retained for the analysis in this study […]."

Line 193: "ascertain knowledge of" -> "know" or "understand"
The revised sentence states *"[…] by the need to understand the present-day exposure […]"*.

Line 240: "any difference in the spatiotemporal variability"; difference compared to what?
In this sentence, we are referring to the differences in spatio-temporal variability of deep-ocean temperature and deep-ocean oxygen. To clarify, we have rephrased the sentence to now read:

"While any difference in the spatio-temporal variability of deep-ocean temperature and deep-ocean oxygen will directly impact the number of floats required to capture large-scale changes in each variable over time, this difference remains unquantified to date."