

SyncED Paper Revisions

Reviewer 1

We thank the reviewer for their thorough review, the detailed comments were really helpful.

This study is a very nice demonstration of a proof of concept of the use of glider targeted observations driven by a real time ocean forecasting system fed with those observations. I would recommend publishing this article with minor revisions. There are still concluding remarks about the improved realism of the simulations over the entire domain when glider observations are assimilated that are not supported by the analysis shown in the paper (1.402-403). Only local validation is shown with the glider observations in the small box where they are deployed. The lack of ocean interior BGC observation may prevent to conduct such validation on the extended geographical area. I would suggest to either add diagnostics against observations outside of the "box" domain to support the conclusion or to adjust the conclusion with the content of the paper. The later solution would be fine as the paper is already quite rich and the validation of the realism of the different simulations can be presented as a perspective.

The RMSD plots in figure 6 show the error between the forecast (at different lengths) and the analysis for a wider region. Whilst validating against independent observations is preferred when they are available, this approach is still valuable and shows how quickly the model diverges from future analysis states which take realistic observations into account. We have added a couple of lines to hopefully make this clearer

Figures 7 and 11 showing the number of days where the Chlorophyll threshold is crossed are very patchy. Can this be realistic patterns from our present knowledge?

If we look at a time series of the region, the assimilation of glider chlorophyll generally increases the amount of chlorophyll in the model system over the simulation period, which is what leads to an increase in threshold crossings. The patchiness is just a result of the fact we are plotting the crossings of an 'arbitrary' limit over a short mission time period, we would expect this trend to continue and smoothen out over a longer period. For oxygen, this area is not a hypoxic region but there is evidence that levels drop below the OSPAR threshold on occasions (see Ciavatta et al 2016)

In Section 3.4 on the Impact of gliders DA on path planning, which criteria was chosen for the path planning: transect sampling or targeting the largest uncertainty location? **Transect based – added this to the text**

Since the model fields that are feeding the path-planning are known, does the changes in the planning are coherent with the changes of patterns seen on figure 7 and 11 for the Oxygen and Chlorophyll?

We have changed figure 12 into something which should address this question a bit better, showing the difference in where the gliders spend most of their time under the different model outputs.

Specific comments

l.105: "optodes ." : A space to remove **Fixed**

l.125 to 130: It could be interesting to indicate if the DT processing described here has a very significant impact on data quality compared to the RT ones, and if the RT are still "useable" for assimilation, even with an increased observation error. This can have an impact when interpreting the difference between the DT and NRT AMM15 experiments.

In general, NRT data are typically usable for assimilation, particularly for temperature, where QC mainly removes outliers and spikes. However, more subtle drift and long-term degradation are harder to identify without delayed-mode processing, which can lead to underestimated observation errors.

The difference is more pronounced for salinity and especially biogeochemical variables, which are more sensitive to drift, fouling and hysteresis, and therefore benefit more strongly from delayed-mode correction and calibration.

l.139: two dots instead of one after m3. **Fixed**

l.141: data provided for the numerical model: for assimilation? **Clarified**

l.142: in the DT experiments, only the delayed data (gridded product) were used? **Yes**

l.196: Ford et al, 2012 should be inside parenthesis. **Fixed**

l.224: to be coherent with the rest of the text, chlorophyll should be used instead of chl-a. **Fixed**

l.253: "and" is repeated twice. **Fixed**

l.417: Assimilation of DT data alter or improve the results? **Fixed**

Reviewer 2

Thank you for comments, in particular they made me think about presenting the path planning results in a different way which helped make the paper clearer and stronger

In my original review, I requested a comparison of the actual glider trajectory with the DTO-recommended trajectory. Although the authors stated that this analysis was performed, it was not included in the manuscript. Even a brief comparison would enhance the study's operational relevance by showing how the DTO framework could influence sampling decisions. If available, I encourage the authors to add a short discussion or include an example in the supporting materials. We omitted the trajectory maps initially as they were messy and difficult to interpret, hence favouring the limited version in the original submission that focused on distance between potential planned routes. This has been changed to a heatmap of the glider locations, which sacrifices specific time/locations in favour of a clearer, broader picture of where the DTO and path planned gliders spend most of their time. Hopefully you agree that this demonstrates the impact of gliders and the DTO framework better than before.

The manuscript now includes additional validation against observations, a valuable improvement. However, in several sections, the discussion of forecast performance focuses primarily on comparisons among model configurations, such as AMM15-DT and AMM15-NRT. I recommend that the authors clearly distinguish between validation against independent observations and model-to-model comparisons throughout the manuscript. This clarification will help readers accurately interpret the improvements in predictive capability. **Added some additional text to make this clearer.**

In the discussion, it remains difficult to separate the effects of glider data assimilation from those due to differences in model resolution and configuration. Where appropriate, the authors should acknowledge this limitation and clarify that some reported improvements may result from both assimilation and model characteristics. **The impact of gliders section isolates the effect of glider data assimilation as that is the only difference between AMM15-NRT and AMM15-NoG. We agree it**

is difficult to separate the effects when comparing the different resolution models due to configuration differences, at the start of the impact of resolution section we say a full like for like comparison is not possible because of this.

Overall, the manuscript has improved significantly and will be a valuable contribution to Digital Twin Ocean systems and adaptive ocean observing strategies. The remaining comments are minor and primarily concern interpretation rather than additional analysis. I think this version is suitable for publication after a minor revision that does not necessarily require additional review.