

Geosci. Model Dev. Discuss., referee comment RC2  
<https://doi.org/10.5194/gmd-2022-204-RC2>, 2022  
© Author(s) 2022. This work is distributed under  
the Creative Commons Attribution 4.0 License.

## Reply on RC1

Anonymous Referee #2

---

Referee comment on "How does 4DVar data assimilation affect the vertical representation of mesoscale eddies? A case study with observing system simulation experiments (OSSEs) using ROMS v3.9" by David E. Gwyther et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2022-204-RC2>, 2022

---

The goal of this paper is to investigate impact of data assimilaion on ocean sub-surface fields. A set of "fraternal" twin OSSEs have been conducted. Please see below for: 1. Major and 2. Minor comments.

1. 1 L101- 102 "Further details of the free-running and DA configuration used in these OSSEs are given in Gwyther et al. (2022)." However, not much information has been provided in this present paper. It is simple to cite past work! However, such details, even in a summary format are very relevant to this paper, a table of differences between `ref` and `DA` model configurations would be most illustrative. Also a discussion of why those differences actually yield a "good" set-up to address the sub-surface impacts of data assimilation (discussed in the Introduction) is needed.

1.2 Table 1 "Along-track satellite-observed sea surface height al- timetry and sea surface temperature." A plot is needed here to show the data coverage.

1.3 Figures such as 2 and 3 are time-averaged. But given the eddying flow, the errors are expected to vary in time, therefore a time-series plot or Hovmoller plot of error (with respect to `ref`) standard deviations would be best.

1.4 L 223 "The XBT-N+S OSSE has a slightly higher EKE difference than XBT-N or XBT-S but performs better than Surf (cf. Fig.3j and Fig.3g)" Why so? The authors do not explain this counterintuitive result though more observations have been added in XBT-N+S than XBT-N or XBT-S.

1.5 L377 "DA simulations will potentially struggle to generate representative baroclinic

mode structure." The authors generalize their conclusions without demonstrating how does the model used for DA performs without any assimilation (see comment above 1.1). They haven't shown the nature of errors that this model has without any observations - unless those are described, how can one draw conclusions whether assimilating observations helps or not? It is important to design the experiments properly, by making sure the model used for DA is different from that to generate the 'ref' trajectory.

1.6 I enjoyed reading Sec 2.3.4 and 3.5; other sections/results identified problems with the data assimilation experiments, but never explained their causes. In the end, I am not what exactly sure what is the take home message? Every data assimilation scheme/implementation has a specific treatment of background errors. In this paper, there is no concrete evidence that it is the background errors that are to be blamed; mere speculations have been raised (L391, L418). It would be much better if there was a set of experiments with different B (formulation or changed values) that proved these speculations- even partly. Otherwise, what exactly is the contribution of this work? A diagnostic tool presented in Sec 2.3.4 and its use case in Sec 3.5? I hope some of this criticism helps improve this work on a very important topic.

2. 1 L19 "they" deliver

2.2 L68 "focus on two" ??

2.3 L340 "sim" should be "\$\sim\$"

2.4 Most figures do not have axes labeled (for e.g., "Latitude (deg N)"), same remark for colorbar - Fig.2

2.5 Most figures would be more readable if XBT-N, -S lines are superimposed.

2.6 L87 "surface forcing conditions from BARRA-R" Are they also daily?

2.7 a. Fig.6(e) The transect line is further from XBT-S than in other panels. Is it same section?

2.7 b. Fig 6(k), Fig. 7(k) What is the colorscale?