

Geosci. Model Dev. Discuss., referee comment RC4
<https://doi.org/10.5194/gmd-2022-27-RC4>, 2022
© Author(s) 2022. This work is distributed under
the Creative Commons Attribution 4.0 License.

Comment on gmd-2022-27

Andy Hogg (Referee)

Referee comment on "Cloud-based framework for inter-comparing submesoscale-permitting realistic ocean models" by Takaya Uchida et al., Geosci. Model Dev. Discuss., <https://doi.org/10.5194/gmd-2022-27-RC4>, 2022

This paper advocates for a cloud-based strategy to address the problems in sharing and analysing the large volumes of data that emerge from high-resolution ocean model simulations. I found the paper to be interesting and well-written, and concur with two previous reviewers that this manuscript is a worthwhile contribution to the literature. I have some minor comments, which the authors may like to take into account, listed below. I would be happy to recommend publication if these issues are addressed.

Line 8-9 - Consider deleting the sentence naming the 5 models from the abstract?

Section 2 — I found this description of the process of sharing data, and the ARCO format, to be particularly useful. But one thing I don't understand is whether the authors are arguing the Zarr files produced here are optimal for all operations. For example, if I wanted to filter with FFTs, average in time or average in space, would the Zarr chunking remain optimal for all of these operations? Or is there a trade-off between operations?

Line 115 — Improved GS separation is a nice feature, but the global ocean is bigger than just the North Atlantic and there are many more processes revealed by resolution than

WBC separation. I'm not convinced that separation is more "key" than other processes that are improved with resolution. Maybe just back away from this statement a little?

Line 122 - "this will..." is a little ambiguous.

Line 146 — my recollection is that the tides in LLC4320 had a bug in the tidal forcing which overestimates the tidal magnitude (but I apologise that I can't put my hands on the appropriate reference). I suggest the authors check on this issue as they revise the manuscript.

Line 180 — "... the two ..." - also ambiguous.

Line 191 - there is a case made about daily-averaged submesoscale fields, but it wasn't clear (to this reader) where these daily-averaged fields were used in this paper?

Line 222 — "This presents ..." ambiguous ...

Figure 6 — On a first read, I was amazed at the similarities between the parameterised submesoscale fluxes and measured buoyancy flux. Actually, it looked too good to be believable. But when I looked at D1 the comparison was underwhelming. I suspect the use

of the spatial median in Fig 6 is unfairly favouring the comparison. I would prefer the authors to show D1 as the main figure, or perhaps show both in the main text, for a warts-and-all comparison of the parameterisation.

Section 5 — The authors makes some good points here and I agree with most of them. But I found the approach to be slightly evangelical. Fundamentally, the argument seems to be “we have found the best approach, but if the scientists/funders don’t back us then it will fail”. I agree that the approach espoused here is good, and I would like to advocate for it myself. But a more dispassionate discussion of the pros and cons would probably be an advantage here. For example, a significant disadvantage here is the risk that the Google Cloud Platform is discontinued or unavailable to researchers in some nations, for whatever reason. That is not such an outlandish proposition, but could be catastrophic for an open platform like this. There are other risks of equal access, long term funding, etc. I am just asking here for a more objective analysis of the risks here — which would be a greater service to the reader than the advocative approach.