Review of ***Offline Fennel: A High-Performance and Computationally Efficient Biogeochemical Model within the Regional Ocean Modeling System (ROMS)*** by J. Crespín, J. Solé and M. Canals

This manuscript describes an offline implementation of a well-stablished biogeochemical model, the Fennel model, and offers an evaluation of its performance against the same biogeochemical model coupled online to an ocean model configured for the Gulf of Mexico, from which physical variables are extracted to run the offline case. The offline model performance is evaluated considering different time steps. In addition, the role of three vertical mixing scheme closures is explored.

This offline model is a valuable tool that reduces the computational cost of running biogeochemical models. It is also appreciated that it is publicly available to the community. After evaluating the revised manuscript version as well as the replies to initial reviewers, I find it well-structured, relevant and fits well within the scope of the journal, for which I recommend it for publication. I have, however, a few comments that aim to enhance the clarity of the manuscript.

(Line numbering in comments below refers to the latest version of May 17 with tracked changes)

MAJOR COMMENTS

i) I am confused about how biogeochemical tracer diffusion is implemented in the offline model regarding the three points below, which should only require some clarifications in the text.
   - Authors mention the usage of three different mixing schemes, but these are vertical mixing schemes while nothing is said about horizontal mixing at subgrid scale. Please add a line or two about it.
   - In L126 authors refer to the GLS scheme but this is a generic scheme. Authors should specify if it was implemented as k-kl (a specific case of which is the MY25 scheme), as k-ε or k-ω as indicated in Warner et al. (2005), as well as the parameters used (or refer to a publication that uses the same parameters) for reproducibility of results.
   - L161-L168. Authors mention that AKs, AKt, and AKv are used to force the offline model in GLS and MY25 mixing schemes but not the LMD one. I find this very confusing. The aim of using a vertical mixing scheme is to obtain a vertical diffusion coefficient to use for subgrid-scale vertical mixing in the biogeochemical tracer equations (e.g., eq. 9 in Fennel et al. 2022). Which of the three (AKs, AKt, and AKv) is being used to diffuse vertically biogeochemical tracers in GLS and MY25 mixing schemes? Why do you need the TKE and the generic length scale output in the offline model if the vertical diffusivity is already provided? Additionally, the text reads as no vertical diffusivity is used in the LMD scheme case, is this correct?

ii) Reviewer 2 had a concern about the additional value of showing the RMSE besides the SS. The concern arises because the SS is the RMSE normalized by the online values, therefore both properties quantify differences in amplitude between the online and offline. Authors could replace Table 1 by a Taylor diagram (Taylor 2001), offering a visual and faster comparison between simulations (DT and vertical mixing scheme) to the reader. This diagram shows RMSE but also the correlation coefficient, a metric on how the online and offline values covary in time and/or space, which is currently lacking in the evaluation. In this case, statistics for both p and RMSE should be weighted by volume given that it varies among grid cells (I believe depth levels are unevenly spaced in the model) and tracer concentrations are per unit volume. Volume-weighted averages should be also used for the computation of the SS time series in Fig. 2.

I also agree with reviewer 2 that the results shown in Fig. 3, and its related discussion (L286-L295), are applicable to Fig. 6, which also seems more relevant to me. The only differences are the discrepancies in PO_4 and NO_3 offshore, which are not very relevant to the manuscript and would likely be unveiled in Fig. 6 if considering the full water column.

Finally, related to model evaluation of results shown in Fig. 4, how is equation (1) applied when two offline simulations are compared? That is, which values are used in the denominator to normalize the RMSE? Could this be related to the fact that the heatmap is not exactly symmetric?

MINOR COMMENTS

L15. Comma after fields.

L20. Express the time reduction in percentage since it is more meaningful as a general message.

Authors could also mention in the abstract the fact that NH_4 is a challenging tracer to simulate accurately offline since its timescale of change is faster than other tracers. This is a nice general result applicable to other offline biogeochemical model implementations.

L90. Which type of shift? A time shift?

L121. Add “vertical” before “mixing”.

L131-L132. Instead of listing all variables, authors could say that the Fennel model variables are included except those involved in carbon pools (which imply using ALK). This is a more direct way for the reader to understand what is exactly included and avoids redundancy since all Fennel model variables are listed in L100-102.

L147. Climatology over which time period?

L147. Explain what zeta, ubar- and vbar- velocity properties are.

L174-L175. Delete last sentence to avoid too much redundancy. The value of the barotropic time step is already stated in L107.

L124. Delete “subgrid-scale turbulent mixing closure scheme” since all three schemes are.

L309. Remove “(GLS, LMD, or MY25)” to avoid redundancy.

L330. Specify the time length of the averaged output.

Figure 8. I think panels C are redundant differences from B are shown in A. Also, in caption remove “(x5 DT)” after “offline” to avoid repetition. Finally, consider a colorblind friendly colorbar in panels B such as the cmocean ones (Thyng 2016).

REFERENCES

Fennel, K., Mattern, J.P., Doney, S.C. et al. (2022). Ocean biogeochemical modelling. *Nat Rev Methods Primers* 2, 76. https://doi.org/10.1038/s43586-022-00154-2.

Taylor, K. E. (2001). Summarizing multiple aspects of model performance in a single diagram, *J. Geophys. Res.*, 106(D7), 7183–7192, doi:10.1029/2000JD900719.

Thyng, K. M., Greene, C. A., Hetland, R. D., Zimmerle, H. M., & DiMarco, S. F. (2016). True colors of oceanography. *Oceanography*, 29(3), 10.

Warner, J.C. et al. (2005). Performance of four turbulence closure models implemented using a generic length scale method, *Ocean Modelling*, Volume 8, Issues 1–2, 2005, Pages 81-113.