Reviewer 1 response:

The authors would like to thank the reviewer, Dr. Julian Mak, for their constructive and creative feedback on this manuscript. Below, we address each point in order.

Technical comments:

- There is the paper of Rai et al. (2021) [Rai et al., Sci. Adv. 2021; 7 : eabf4920] that looks into eddy killing, providing an estimate of energy loss from the ocean arising from relative wind stress. The paper is close enough in spirit that appropriate citations should be made, probably in multiple places throughout the text, and in particular highlighting how the present methodology is different. Would be good to see a comparison on how the present results are similar/different to the results there.

  (There is also a Rai et al (submitted) but that's not published yet, and may be not as relevant as the 2021 paper.)

> *Response:*
> Thank you for pointing us in the direction of this study. To start, the intention of our study is to produce a simple dissipation rate and give a flavour of what it could look like. Rai et al. make some important conclusions regarding the length scale of eddy killing and its seasonality i.e. peaking in winter. They base this on the seasonality of EKE. In our work, the eddy dissipation parameterisation is based on the dissipation of APE, since this is typically the largest reservoir of eddy energy.
>
> We have cited Rai et al. (2021) in multiple places throughout where appropriate.

- A two-layer shallow water model is employed as a way to have a representation of the first baroclinic mode, however the choice of "mode" (or the basis) is the "standard" choice assuming flat bottom essentially. That choice of basis is not unique and increasingly there are studies using surface modes (e.g. LaCasce 2017; Groeskamp et al. 2020; Ni et al., accepted at GRL or 2023?) I am not suggesting you redo your analysis using surface modes, but given there is a choice of basis, minimally it would be good to speculate how the results are going to be the same and/or different (although of course if you redo the analysis, instead of speculating you could then quantify the similarity and differences, and strengthening the conclusions).

> *Response:*
> Thank you for bringing this point to our attention. We feel that the flat bottomed baroclinic mode suits our experiment well, particularly with regards to our two-layer baroclinic model and MITgcm setup.
>
> We include some sentences on surface modes on lines 148-149 and 483 with relevant

citations.

- Regarding Eq. (19), I'll be honest and didn't think about this as much as I probably should have. Two points here:

  1) The notation is a little un-rigourous, because it's not clear which quantities are scalars, vectors and/or tensors, missing some contraction operators and the like. For example, the final term on the right hand is written $(\nabla^2 u)^2$, but you really mean $\nabla^2 u \cdot \nabla^2 u = |\nabla^2 u|^2$. Is the second term a tensor ($|\nabla u|^2$) the hit by the Laplacian operator? This is mostly notation but probably could be cleaned up a bit.

  2) Normally in shallow water it is known that simple choices of the Laplacian as a diffusion leads to sign indefinite energy dissipation (e.g. Peter Gent in 1993, "The Energetically Consistent Shallow-Water Equations"; Gilbert et al., 2014, "On the form of the viscous term for two dimensional Navier–Stokes flows"; in my PhD thesis). One offending reason is that the primitive variables in shallow water are (h, u, v), but the conservative variables are (h, U = hu, V = hv). If for example in the prognostic equation we have $-\nabla^2 u$, then multiplying by hu then integrating by parts gives (abusing notation a bit)

     $\int hu \cdot \nabla^2 u = boundaryterms - \int \nabla(hu) \cdot \nabla u,$

     which gives the expected sign-definite term $-h|\nabla u|^2$, but there is a cross term involving ∇h floating around. The problem however doesn't exist in the conservative variables. It seems somehow you don't have this issue here? I am not seeing how your terms could be written as a flux given there is that annoying h term floating around, so a clarification would be useful.

     Ultimately I assume it's quantitatively not going to be important, because your "hyperdiffusion" $\nabla^4 u$ (I put quotation marks because I am arguing u is not the variable you should be hitting ∇^4 with) is presumably going to be a small effect. Clarification and maybe appropriate references here would be helpful (e.g. Gent 1993; note also he makes the point about energetic consistency, which I am not convinced you have here, but it's probably not too important).

Response:
Thanks for pointing this fact out. On reflection, the viscous term is not overly important and so we will leave it out of the paper entirely. The text in sections 2 and 4 have been modified accordingly and Fig. 3 has been changed too to account for zero damping in the absolute wind stress cases.

- So intuitively the action of the wind is at the surface, but that is being argued to be felt over the structure of the baroclinic mode, so there is implicitly some assumption about the time-scale of communication. Do we actually know what that is, and is that small enough for the conclusions here to internally consistent? I think I wouldn't raise this question so much if surface modes were used for example. Clarifications on this point would be welcome, maybe commenting on the evolution of the MITgcm model results.

> *Response:*
> Thank you for this point. We felt adding a comment directly in the results of the MITgcm results (Section 4) didn't feel right. So, we have added a few sentences at the beginning of section 5, line 335-338. We essentially add a further justification (relative wind stress timescale of communication) for using potential energy, rather than just a scaling and that potential energy is largest.

- The toy model uses an eddy as a noun (quasi-circular object), but results are then applied generically to a velocity with no specific mention to the eddy, so I assume there is no eddy detection that is being done here, and the application is then considering eddies as the verb (the fluctuation, a special case being the quasi-circular object). Could the authors comment on this distinction, and which one would be the more appropriate thing to do? One could for example argue that we might want to identify the coherent eddies (so the noun, via Eulerian or Lagrangian approaches), then consider the impact of relative wind-stress on those identified eddies for consistency of the theory and application.

  (I don't personally believe you should do what I just suggested. I am just raising the point that the word "eddy" is sometimes used to mean different things by different people, and sometimes the intention is not as precise as it should be.)

> *Response:*
> You are right. We initially consider the eddy to be a noun in the toy model, and this helps us derive the parameterisation. In our application of the parameterisation, we are considering eddies to be deviations from the mean, of which coherent eddies make up a small part of this. We have added the following two sentences on line 347-348 "In addition, we now consider *eddies* to be deviations from the time-mean, rather than just being a singular coherent eddy. The interpretation of the dissipation rate can also be thought of as one for these eddy time-mean deviations."
>
> We have also added a sentence on line 469 that suggests the use of a eddy detection method in future work.

- Maybe a hypocrite for asking this (because I didn't do it either in the 2022a paper), but a dissipation rate is estimated here but no estimation of a power? Could you estimate a power, and how might that compare with the results of Rai et al (2021) say? Or if you refrain from doing so, give a reason on why you don't?

> *Response:*
> Computing an estimation for power is not something we considered doing. We could try and attempt an estimation, but this would require eddy detection e.g. eddy amplitude A and eddy radius R, similar to Chelton et al. (2011). We would put the detected values of A and R into our equation for P_{rel} and compute globally. This is something we could do, but is beyond the scope of the papers goal. Again, perhaps an idea for a future piece of work.

Presentation comments

- (line 6): unless you can definitively say and show the "nonlinear baroclinic processes" causality, I would lessen the strength of the wording and say "there is divergence from the analytical model at around day 150, likely due to the presence of nonlinear baroclinic processes" (because it's really more an observation at the moment).

> *Response:*
> Agree. Changed.

- (line 12 and later): the 10^{-7} is mentioned but its significance (or lack of) is never given explicitly. Easiest to say up front why that that reference value is chosen.

> *Response:*
> Removed 10^{-7} for now in abstract.

- (two sentences spanning line 10-12): reads a bit clunky, could do with a re-write.

> *Response:*
> Sentence changed/combined.

- (line 34-35): I would argue that's not a good comparison, because the low explicit eddy energy is to do with the coarse resolution model and much less on the GM parameterisation itself. I would personally just remove that sentence.

> *Response:*
> Removed.

- (line 45-47): Jumpy sentence, consider rewrite (eddy saturation <-> GEOMETRIC while "other" <-> turbulent energy cascade, and as written the is ambiguity in how the sub-clauses are related to each other)..

> *Response:*
> Removed the Jansen citation and focused on GEOMETRIC.

- (line 59): formatting of reference, brackets.

> *Response:*
> Fixed.

- (paragraph of line 60): Rai et al. (2021) should be cited and results compared accordingly in this paragraph.

> *Response:*
> Included on line 69.

- (line 95): g should be the gravitational acceleration constant

> *Response:*
> Done.

- (line 95): "∇_h IS THE horizontal gradient operator" or simila

> *Response:*
> Done.

- (sentence of line 114-115): Could be read as there is current feedback onto the wind profile, which I assume is not what was intended.

> *Response:*
> Not what I mean. Have removed that last sentence and included some description of what the figure shows in the text.

- (equation 9): consider using `\begin{align} \end{align}` with some & according to break the lines, probably

```
\begin{align}

    W_{rel} &= \tau_rel \cdot u_g \\

            &= etc. \\

            &=

    \end{align}
```

> *Response:*
> Done.

- (line 143-145): citation to Rai et al. (2021) here also probably.

> *Response:*
> Done.

- (section 2.2 opening paragraph): probably comment about surface modes here, or say the discussion will be given in the conclusion section

> *Response:*
> Included LaCasce citation but discuss in conclusion.

- (line 163): "…where $\cdot_{1,2}$ denotes the upper and lower layer variables,…"

> *Response:*
> Done.

- (line 202, 203): might consider swapping u and η ordering so you can have equation reference ordering as (5) and (6)

> *Response:*
> Done.

- (line 223): not sure why you wouldn't just RK4 the whole thing, comment to clarify? (Because piggy-backing on the MITgcm AB3 time-stepper?)

> *Response:*
> Clarified.

- equation (23): `\left( and \right)` for brackets

> *Response:*
> Done.

- (liner 262): so you have sponge layers or diffusion to soak it up? clarification would be useful (if discussed in previous work, citation here would also be appropriate).

> *Response:*
> This information is provided in Wilder et al. (2022) and first cited in the first paragraph of section 3.1.

- (line 273): I might have opted to define $n_0(z)$ as the NEGATIVE vertical gradient to soak up that negative sign in PE partly because it is never used again anyway (I'm not remotely attached to this suggestion).

> *Response:*
> Thanks. Will leave as is though.

- (line 318): remove comma maybe, "…but with vertical diffusion did result…"

> *Response:*
> Done.

- (line 343): "…zonal wind velocity, because wind patterns…"

> *Response:*
> Done.

- (line 345): "…is on a 2 degree horizontal grid."

> *Response:*
> Done.

- (line 355): comment on differences / similarities with Rai et al. (2021)

> *Response:*
> Added a sentence on line 392-393 when describing figure 7.

- (text below equation 31): comment on how surface modes might change results here, or say this will be talked about in conclusions section (I would minimally speculate what is expected to change with surface modes here though).

> *Response:*
> Comments on surface modes will go in the discussion. We feel that they will flow better in an 'alternative approach' paragraph at the end. See line 483.

- (line 370-372): "…while there is a slow down…between seasons, with the largest absolute values…"

> *Response:*
> Done.

- (line 388): against, may want to be explicit about significance of 10ˆ-7

> *Response:*
> Done.

- (line 390): extra space after 0°

> *Response:*
> Done. Added a dash after.

- (line 395): for completeness it is also in Mak et al (2022a), Fig 7 I suppose.

> *Response:*
> Added other citation here.

- (line 405): is negative dissipation a problem?

> *Response:*
> There is no negative dissipation? We take the log of $\Lambda_{rel}/10^{-7}$.

- (line 407): would recommend $\widehat{lon}$, or even better $\widehat{\mathrm{lon}}$

> *Response:*
> Done.

- (line 428): how significant are the uncertainties? a quantification would be useful

> *Response:*
> It's difficult to quantify really without diving into more experiments. Have modified the sentence on line 445.

- (line 438-439): is the issue of "time-scale of response" an issue?

> *Response:*
> It shouldn't be. See our earlier point in technical comments.

- (line 451): Mak et al (2022a) the one actually intended? (2022b is the prognostic calculation, while the 2022a is the inverse calculation). How are the results here compared to that say?

> *Response:*
> Removed sentence. Didn't really follow from the previous sentences.

- (line 467-468): Rewrite? I assume you want “A further advantage of this work is having a simple analytical expression for this dissipation rate that can be applied to ocean datasets” or something similar. It doesn’t flow very well at the moment.

> *Response:*
> Thanks.