Reviewer 2 response:

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

Section 5:

1. Line 367: There is a formula given here for the reduced gravity, g'. The first problem I have with this formula is the term involving the eigenmode, phi_1. The eigenmodes can always be scaled by a constant multiplying factor and since we are not told how the eigenmodes are normalized, the formula for g', as written, is not sensible. It follows that we need to be told how the eigenmodes are normalized. Nevertheless, I cannot make sense of where this expression for g' comes from. The fact the value of the eigenmode at z=0 appears is mysterious; the fact it appears squared even more so. Some explanation is required.

> *Response:*
> The equation for g' is correct, although we accept the description could be made clearer. In Xu et al. (2011), to get Eq (37) on page 539, they rearranged (from table 2 page 352 in Flierl (1978)) $F_1(0)$ for H_1, which is where the square of the eigenmode and its definition at $z = 0$ comes from. From the same table in Flierl (1978), the eigenvalue λ_1 is rearranged for reduced gravity g', where $g' = \varepsilon g$ (see Flierl, page 347). Then substituting (37) into (38) from Xu et al. (2011), the equation for g' follows. In the updated manuscript we keep g' in terms of λ_1, and not c. So, $g' = \frac{f^2(1+\phi_1^2(0))}{\lambda_1 H_2}$, where ϕ_1 is normalized using Eq (36) in Xu et al. (2011).

2. In the 2-layer model, the wave speed, c, for the first baroclinic mode is related to the reduced gravity by c^2 = g' H_1 H_2/H where H = H_1 + H_2 and H_1, H_2 are the undistiurbed depths of the upper and lower layers respectively (see Gill(1982), equation (6.3.7)). Presumably the term involving the eigenmode in the expression for g' above corresponds to the factor H/H_1?

> *Response:*
> Yes, it does. In the previous response, when rearranging to find g', we actually have the equation $g' = \frac{c^2 H}{H_1 H_2}$. But in our equation, we replace H with $H_1(1+\phi_1^2(0))$, and do not use c.

3. Line 366: Here, it is written that H_1 is taken to be the depth of the zero crossing of the eigenmode. Why write this in terms of the minimum of the absolute value of the eigenmode when this is obviously zero at the zero crossing? I also wonder if this is the appropriate choice for H_1? It seems reasonable but might also turn out to be on the deep side? On the other hand, it is clear that the choice of H_1 is of great importance

since once H_1 is known, so is H_2 and hence g', given that c and H are known (see above).

> *Response:*
> This is a mistake. We will write "We take the depth of zero crossing of the first baroclinic mode to be H_1" on line 375-376. We did also attempt to find H_1 by computing the depth of the main thermocline, but discovered this to be a non-trivial task. We therefore chose the depth of zero crossing of the first baroclinic mode as H_1.

4. Line 369: How is mu computed? We are not told anything about this. On line 186, mu is given by mu = -g'H_2/gH. How does this connect to the mu being used here?

> *Response:*
> μ is computed using the expression on line 185. We have reiterated this fact in the text on line 378.

5. I found myself wondering if a simpler model for an eddy than the 2 layer model used here (i.e. allowing the lower layer to be in motion and selecting only the baroclinic model) might be to assume that the lower layer is at rest, corresponding to a projection onto both the baroclinic and barotropic modes. Such a set-up would be equivalent to the 1 ½ layer model. The 1 ½ layer model is recovered from the 2 layer model in the limit that the depth of the lower layer, H_2, goes to infinity. The mathematics for the 1 ½ layer model is simpler than for the 2-layer model and I feel sure the expression in equation (29) would be the same as given in the manuscript with mu = g'/g (since H_2/H tends to one as H_2 tends to infinity). This shows very clearly the importance g' for determining the dissipation rate. The authors might want to consider briefly discussing the 1 ½ layer model as an alternative to their single baroclinic mode setup.

> *Response:*
> The 1.5-layer model is certainly interesting and offers an alternative approach to the 2-layer model. We have briefly mentioned how the 1.5-layer model alters the dissipation rate in Eq. (28) (old Eq. (29)), showing the importance of the reduced gravity g'. The choice of the 2-layer model in this work essentially follows on from some earlier work we carried out on the 1-layer model (see this thesis https://ueaeprints.uea.ac.uk/id/eprint/92511/).

6. An obvious question that is not addressed is what the authors do about the fact the ocean actually has variable bottom topography. I assume that the eigenvalue problem stated at the beginning of Section 5.1.1 uses H = 4000m? This should be made clear. How do the authors deal with the temperature and salinity data if the local depth is less than 4000m? Again, this needs to be explained.

> *Response:*
> Thanks for pointing this out. The eigenvalue problem uses the depth from the WOA dataset, so we have made this clear on line 371. We also neglect data shallower than 300 m, so we write this on line 373. Reviewer 1 (Julian Mak) suggested the use of surface modes which account for rough bathymetry. This may be something that we should look at in future work.

7. Another issue throughout section 5 and in the figures is how the authors deal with the Rossby radius of deformation near the equator where the Coriolis parameter goes to zero? I assume a band around the equator must be left out. But we do not see this in Figure 7 and neither is anything ever said about it in the text.

> *Response:*
> Thanks for bringing this up. We did not originally deal with the Rossby radius of deformation at the equator. Considering this fact, we therefore include masks over the equator regions in the Rossby radius (Fig. 7), and the dissipation rate (Fig. 8). The density plot (Fig. 9) has also been recomputed by neglecting the equator band, and the figure is not too different to the previous one. The mask is pointed out in the figure captions, and in text on lines 393-394 and 414-415.

8. Line 334: We need to be told what expression is being used for the energy in (29). In particular, we need to be told which term in (17) comes from the thermocline displacement? I also feel confident that it is easy to show that this term dominates the expression in (17) by substituting appropriate values for the parameters. Showing this would justify the use of the "key assumption" made here and make it clear it is not an ad hoc assumption.

> *Response:*
> This has been made much more clear in the text. We have included a scaling and have cited Gill et al. (1974) for the dominance of potential energy. A simplified eddy energy E has now been shown [see new Eq. (27)]. We have removed the phrase "key assumption" because it is rather a fact.

9. It seems that the wind speed used in (29) is the seasonal mean wind speed. This also needs to be stated clearly. What about the fact that the winds are not steady? How could this affect the result? At least in (29), wind speed appears linearly (although the speed itself is not linear). In reality, however, the drag coefficient also depends on the wind speed. Indeed, I am reminded of the paper by Thompson, Marsden and Wright (1983, JPO) from 40 years ago…

> *Response:*
> In Section 5.1, we state that the data is averaged into seasons (line 356). On line 468-469 we briefly mention the work by Zhai et al. (2012) on high-frequency wind power input, where high-frequency wind may also impact the dissipation rate.

10. Line 385: Is L_e being used or what is written on line 355?

> *Response:*
> Thanks. Will make clear that we are using either R_d or an eddy length scale L_e scaled by $R_d/\hat{R_d}$.

11. Lines 389-390: To talk about the influence of bottom topography is a bit glib since what is plotted comes from equation (29) and hence must arise from the terms in (29). Maybe R, maybe the wind speed? Of course, it is true that this is where the Antarctic Circumpolar Current takes a turn to the north, consistent with the form drag effect across the Drake Passage sill, so bottom topography may play an indirect role.

> *Response:*
> In this text we are trying to connect our results with something tangible that readers can understand more easily. Indeed, the results are a result of R or wind speed, but R is also directly a result of temperature and salinity, which must be steered to some extent by features in the Southern Ocean.

12. Line 394: Likewise, here, the authors could (should) be able to say which terms in (29) are making the important contribution.

> *Response:*
> Yes. Have added more detail to line 407-409.

13. Line 409: The dissipation rates plotted in Figure 9 have to be consistent with Figure 8 by construction!

> *Response:*
> Removed the line.

Other comments

1. Lines 20-21: There are lots of examples of how eddies modulate volume transport, going back to the early quasi-geostrophic models of Holland et al., e.g. Holland, Rhines and Keffer (1984), or even Holland (1978), to more recent papers such as Wang et al. (2017, GRL.

> *Response:*
> Thanks. We have included Holland (1978) and Wang et al (2017) on line 21.

2. Lines 33-35: Tandon and Garrett (1996, JPO) were perhaps the first authors to ask what happens to the energy released by the GM scheme. The discussion here reminds me of the backscatter scheme. . .

> *Response:*
> Thanks. We have included this citation on line 33

3. Line 37: Although Jansen et al. gets referenced, there is no explicit mention of the backscatter scheme in the text?

> *Response:*
> That's right. The point we want to make with that reference is that they use an energy equation to inform a GM transfer coefficient, and thus constraining the dissipation of energy is important for these types of parameterisations.

4. Line 63: The drag coefficient is known to be a function of wind speed – there is no "could be" about it! Think of Large and Pond (1981, JPO) and all the subsequent updates.

> *Response:*
> Corrected.

5. Equation (6) should not include "A" explicitly since the amplitude is already contained in eta as given by (5).

> *Response:*
> Corrected.

6. Equation (8): It should be noted that this approximation assumes that that the absolute value of u_a is much bigger than the absolute value of u_g.

> *Response:*
> Have added a sentence to line 109

7. Equation (10): I assume that the drag coefficient is assumed to be a uniform constant here, independent of u_a?

> *Response:*
> Yes. Have added a sentence in section 2.1.2 after Eq. (7) on line 105 saying C_d is kept

> constant.

8. Equation (10): Furthermore, if I understand correctly, it is not W_rel that is being integrated but only the last 4 terms in the expression for W_rel in (9c)?

> *Response:*
> Yes, but for consistency sake in text we write the integral of W_{rel}.

9. Line 142: It is already clear from the last four terms in (9c) that P_rel is negative. This is because the only term that is not negative-definite is the (u_g)^3 term and this integrates out for a circular eddy. I am also noting here that the absolute value of u_a, plus u_a itself, is always positive or zero.

> *Response:*
> Yes you are right. It's interesting to look at the terms individually, but we will leave the discussion as is.

10. Lines 151-152: Better to write "zero net vertically-integrated flow".

> *Response:*
> Done.

11. Line 163: Should say that eta_2 is measured positive upward. This is because interface displacement in the 2 layer model is often measured positive downward.

> *Response:*
> Thanks. Added.

12. As I mentioned earlier, the authors could consider also discussing the simpler 1 ½ layer model in addition to their 2 layer model.

> *Response:*
> Thanks. We have included a brief mention of the 1.5-layer model on line 343-346.

13. Line 186: It should be noted that these solutions for lambda and mu are only valid in the limit g'/g goes to zero – see Gill (1982), Section 6.2. From (13), for the barotropic mode this is obvious because lamda = 1 implies g' = 0. But it is also true for the baroclinic mode.

> *Response:*
> Done.

14. Lines 198-199: It is not correct to say that the terms in the middle represent the redistribution of energy by nonlinear advection. The pressure work term is also contained in these terms! Think of the energy equation for the equations linearized about a state of rest, e.g. Gill (1982), Section 5.7.

> *Response:*
> Added a line on 196 to add ‘pressure work terms’.

15. Line 215: Surely the higher order boundary conditions on biharmonic viscosity must be used for the first two terms on the right hand side of (19) to drop out on integration? The no normal flow condition is not enough on its own.

> *Response:*
> Have removed the viscous term. See comments/response by/to reviewer 1.

16. Section 3.1: I am assuming that the imposed atmospheric wind is uniform, that the drag coefficient is independent of wind speed and that the model domain has closed boundaries? Please be clear about these things! Also, what depth is used for the model ocean, 4000m?

> *Response:*
> As we mentioned at the beginning of the section, more details are given in Wilder et al. (2022). Have included ocean depth in the text. The default drag coefficient in MITgcm depends on wind speed and this is what we use. At moderate wind speeds we don’t expect there to be too many differences in wind power input (see this thesis on page 55 https://ueaeprints.uea.ac.uk/id/eprint/92511/). Have added a small paragraph from line 252 detailing the wind field.

17. Line 260 and thereabouts: It might be helpful to show a plot of the vertical profile of the eddy used for initialization?

> *Response:*
> We don’t show that here since we showed that in Wilder et al. (2022).

18. Section 4: I do not like the title. How about “Verifying the analytic model”?

> *Response:*
> That’s a better title, thanks.

19. Section 4, first paragraph: It should be stated in the text that this paragraph refers only to the first 150 days of integration, otherwise the impression one gets from Figure 3 is quite different.

> *Response:*
> Have added ‘Now, focusing on the first 150 days…’ on line 291.

20. Lines 318-319: What is the vertical mixing scheme used in the model? Vertical mixing by itself will always increase potential energy – vertical mixing does work against gravity by mixing dense water upwards. Perhaps the greater increase in potential energy due to vertical mixing in the cyclonic case is because the stratification is weaker than in that case?

> *Response:*
> The model uses only an explicit vertical diffusion and viscosity (see Wilder et al. (2022)). Examining the impact of vertical mixing on eddy energy is something that we could look to examine further at a later date.

21. Figure 4: I assume what is plotted is geostrophic relative vorticity, not total relative vorticity? Please make clear.

> *Response:*
> Added to caption.

22. Line 454: The 2 layer model used by the authors is not “linearized”, as written in the text, although the nonlinear terms do not play a role in the analysis.

> *Response:*
> Thanks. Have changed the start of this paragraph slightly.