Storkey and colleagues run the climate model HadGEM3 using different nominal resolutions (1°, 1/4°, 1/12°) of the ocean component NEMO. They find that the model version with 1/4° resolution (referred to as eddy-permitting) has stronger Southern Ocean biases than the other versions. They hence proceed to attempting to reduce those biases by 1) adding an eddy-parameterisation to compensate for the resolution being insufficient to resolve all eddies, and 2) increasing bathymetric drag along the lateral boundaries. Both approaches act to dampen the overactive circulation in subpolar gyres and the Antarctic Slope Current (ASC), which leads to improvement in the biases. The authors link near-shelf biases to overly steep isopycnals and a too-strong ASC, preventing Circumpolar Deep Water (CDW) from reaching the shelf. They also show that the weak Antarctic Circumpolar Current (ACC) transport and near-surface warm biases in the open ocean can be linked to too flat isopycnals across the ACC.

The manuscript is generally well-written, with illustrative figures and a comprehensive discussion of the findings, though there is still room for some improvement. This work is of interest to other modelling groups running simulations at eddy-permitting ocean resolution, in particular those using NEMO configurations, and can help contribute to the understanding of origins of Southern Ocean biases in models, which is a common issue that is not easily resolved and occurs across a broad range of resolutions. This work does not solve all the issues in the eddy-permitting simulations (e.g., the ACC is still weak), and simulations are quite short, which I recognise as likely due to limited computational resources, but it is an insightful step on the way. Below, I outline my recommended revisions before publication of this manuscript.

General comments

- Upon reading the manuscript, I was immediately puzzled by the counterintuitive result that resolutions that are only partly eddy-resolving end up with too-flat isopycnals in the ACC (flatter than the resolution with parameterised eddies), despite supposedly not achieving the full extent of eddy compensation. To me, overly flat isopycnals in the ACC would suggest that there is likely an overcompensation by the eddy field to counterbalance the northward Ekman transport achieved by the prevailing wind field. I was asking myself how the authors explain the over-flattening of the isopycnals in the higher resolution model simulations, particularly when adding the scale-aware G&M in ORCA025 (i.e. adding more of an eddy effect) leads to steeper isopycnals across the ACC. A tentative explanation is later given in Conclusions, and noted as a topic for future work, but for the reader, it may be beneficial to acknowledge this counterintuitive characteristic of the simulation much earlier. I would prefer to see it discussed a bit more extensively, including a mention of how it could potentially be explained, already in the Discussion section, and then summarized in the Conclusions. Noting here that, in e.g., the Weddell Gyre, isopycnals are too steep in the eddy-permitting resolution, and adding G&M leads to flattening and a less active gyre, as expected.
- I can see how adding the scale-aware G&M and/or changing the slip condition would also be beneficial in the eddy-rich resolution. Has this been tested at all by the authors (I acknowledge that this is not the focus of the current manuscript)? If so, which of the two changes is expected to be most beneficial to alleviate biases at that resolution? A brief discussion on this could make the manuscript even more valuable to the broader ocean modelling community.

- In line with the previous review comment, I also note that the use of the EN4 1950-1954 climatology as an observational reference thus calls for some caution in phrasing.

Specific comments

Abstract

L. 1-2: "eddy-permitting ocean resolution" → "eddy-permitting ocean resolution without additional eddy parameterisation"

L. 8: "unresolved eddy processes or the representation of bathymetric drag" → and/or, since both can be (and are) causing issues at the same time.

L. 9: Already here in the abstract, the authors mention the shallower isopycnal slopes of the eddy-permitting resolution, and it immediately caught my attention as being counterintuitive. Hence, I would have liked to see this acknowledged earlier in the paper itself.

L. 30-32: It would be appropriate to also mention the emergence of models with unstructured grid configurations, as they are an approach to overcoming the issue with affording high-enough resolution to resolve high-latitude eddies in global models, e.g. *FESOM* (Wang et al., 2013, https://doi.org/10.5194/gmdd-6-3893-2013 ; Scholz et al., 2019, https://doi.org/10.5194/gmd-12-4875-2019), *ICON* (Jungclaus et al., 2022, https://doi.org/10.1029/2021MS002813;Korn et al., 2022, https://doi.org/10.1029/2021MS002952)

L. 61-62: "A free-slip boundary condition […]" - It would be valuable to also mention what viscosity scheme is used, and potentially also the parameter settings since they likely differ between resolutions, as these also have the potential to affect the flow and thus the biases.

L. 64-65.: "Diffusion of tracers along isopycnal surfaces, parameterising eddy mixing […]" - Is there any regional reduction of the parameterised eddy mixing around the equator where eddies are fully resolved, in particular in the higher resolutions?

L. 75: "with the N216 atmosphere" - Are the results consistent or at least similar if one of the other atmospheric resolutions are chosen? This would indicate that the results are more widely applicable to other coupled models, regardless of what atmospheric setup they use.

Page 3, footnote 1: Was the eddy-permitting model ever tested with no-slip? It would be useful to motivate the choice of partial-slip over no-slip in this case, and discuss how choosing a no-slip condition might have impacted the overly strong ASC. Useful references may be Penduff et al., 2007, https://doi.org/10.5194/os-3-509-2007 ; Deremble et al., 2011, https://doi.org/10.1016/j.ocemod.2011.05.002; Nasser et al, 2023, https://doi.org/10.1029/2022MS003594

L. 93-95: "As well as having more active gyres, the higher resolution models also have a stronger ASC […]" - In the next paragraph, you mention the consequences of this feature in the Drake Passage, but only after you discuss the over-flattened isopycnals. This makes this part of the text feel somewhat fractured. Maybe mention the Drake Passage briefly already here, or rearrange the next paragraph to discuss the ASC behaviour before the overly flat isopycnal slopes.

L. 98-99: Klatt et al. (2005) is one of few observationally-based estimates of the Weddell Gyre strength published in recent decades. It is, however, not the only one. Reeve et al. (2019, https://doi.org/10.1016/j.pocean.2019.04.006) estimate it to 32 +/- 5 Sv based on ARGO data. Older observational estimates by Farbach et al. (1991) and Yaremchuk et al. (1998, https://doi.org/10.1007/s00585-998-1024-7) are also lower c.f. Klatt et al. Meanwhile, the transport in models ranges between at least 10-80 Sv (see e.g., Neme et al., 2021, https://doi.org/10.1029/2021JC017662; Wang, 2013, https://doi.org/10.1002/jgrc.20111)

L. 105-106: "The weaker ACC transport in the higher resolution models is associated with a flattening of the time-mean isopycnal slopes in the Drake Passage" - Given that particularly the ORCA025 resolution is not fully eddy-resolving in the ACC region, it is somewhat counterintuitive that there is an over-flattening of the isopycnals, which suggests too much eddy compensation. From reading the text here, it makes one wonder how that can be. As mentioned in the general comments, this counterintuitive behaviour is mentioned by the authors later in the manuscript, but should be acknowledged sooner. However, looking at the figures, it looks like the isopycnals are very steep (steeper than ORCA1) in the northern part of DP, where the core of the ACC is, and then flattened in the centre, where there is weaker(ORCA12)/counter(ORCA025) flow. In ORCA12, steeper isopycnals than in ORCA1 are also observed between -63 and -62 degN. These details should be mentioned, as they might help elucidate what is actually happening to the ACC transport.

L. 107: "counterflowing currents […] associated with the Shackelton fracture zone" – The authors mention that the modelled counterflow along the southern shelf break is unrealistic c.f. Meijers et al. (2016), but they give no indication of whether counter flows at the Shackelton fracture zone correspond to observations or not.

L. 122-128: On resolution-dependent temperature biases - What are the differences in global mean surface temperature in these simulations, and compared to the observational dataset? As some of the biases in SST can stem from the atmospheric model, or results of the difference in resolution in other regions, it might be useful to make a supplementary figure where the biases are normalised to the observational global mean surface temperature.

L. 132: "observational estimates of Klatt et al. (2005)" – as this is not the only observational estimate of the Weddell Gyre strength (see L. 98-99), this may need to be modified, or at least motivated why this particular estimate is used for comparison.

L. 165: "in the eddy-permitting model […] deep water formation has been suppressed" - Does this lead to (more) unrealistic open-ocean convection than in the other two model versions?

L. 168-169: Cold, fresh water advecting around the Antarctic peninsula in the eddy-rich model suggests the ASC and is too strong through the Drake Passage also in this model version.

L. 185-189: On introduction of the scale-aware G&M: What are the implications of running a resolution that to some degree allows eddy formation, but then also parametrising eddies on top of it using G&M, which can lead to “smoothing out” of the actual eddies? This should be clarified in the text.

L. 190-192: On the introduction of the partial-slip condition – As mentioned above, it would be useful to motivate why the choice was to go with partial-slip and not no-slip for this resolution. Also, based on the description of how partial slip affects the biases in ORCA025, it seems that introducing it south of 50S in ORCA12 could potentially fix the remaining biases with counter flows in the Drake Passage, and thus with cold waters being advected around the Antarctic Peninsula in this resolution as well.

L. 223-224: “the Fresh Shelf, the Dense Shelf and the Warm Shelf. In Figure 7” - It would help guide the reader’s eye if it were also indicated in the figure and/or the figure caption what column exemplifies which one of these three regimes.

L. 238-239: About the V-shaped pattern of isopycnals in the higher-resolution models - I struggle to find this V-shape in the drawn isopycnals in the eddy-permitting resolution. If so, I can only see it in the two isopycnals labelled 27.5 (this might be a mistake in the labelling, as it occurs twice).

L. 262-265: The EN4. 1 climatology does not show the same steep ispoycnal slopes near the continent as the model in this region but, as mentioned, the observations included in the climatology are sparse. It could be useful to cite other data (not included in the EN4.1 climatology) that give indications about the isopycnal structure in the area even if those are from a later time period.

L. 310-320: On open-ocean polynyas - Are these events stronger/more frequent in the eddy-permitting c.f. the eddy-rich resolution, given that the latter appears to have some more capability of forming dense water on the shelves? (see also L. 165)

L. 332-336: About biases along the continental slope/shelf – Here, it would be helpful to clarify which factor is the more important in reducing these biases: adding G&M or introducing the partial-slip condition.

Figure 1 (caption): See comments to L. 98-99 and L. 132 regarding Klatt et al. (2005)

Figure 5: mean SST-lines - It is unclear to me which mean SST this refers to, and as the lines are completely unlabelled, there is no indication of what temperatures the different lines represent.

Figure 9: In this figure, it might be more illustrative to show the model results as anomalies from the reanalysis data. In the other figures, observational datasets are shown in the bottom-right subpanel. It would be helpful to keep the same structure throughout the manuscript.

Figures overall: It could be helpful with supplementary figures showing the differences between the standard N216-ORCA025 and the other two ORCA025 versions (GS and PM) as anomalies from the standard (GS-standard, and PM-standard)