General comments:

R: This paper presents a new ocean model configuration, which couples an ecosystem model to an active ocean-sea ice model, with a refined horizontal resolution in the Southern Ocean. The stated aim of this new configuration is to examine the Southern Ocean carbon budget and its sensitivity to wind changes. Although this new model configuration is promising, this paper fails to provide quantitative evidence of the added value of this configuration, and lacks adequate discussion of the applicability and limitations of the technical choices made or the resulting biases. Most results included are side-by-side plots of surface metrics in the simulation and observations. The model evaluation needs to include quantitative metrics and/or plots of differences, and would benefit from comparisons with existing reanalysis products or state estimates. Biases are only briefly (if at all) mentioned and there is generally no attempt to identify (and provide evidence for) the possible causes of these biases. In addition, the manuscript mentions a 'twin' configuration with coarser resolution, but does not include any comparison with that existing configuration (which would significantly improve the manuscript). The objectives of the wind sensitivity experiment are not clearly stated in the introduction, and the limitations of this experiment need to be discussed any time its results are included. This manuscript also needs to include a more detailed discussion about the applicability of the results and the limits of this configuration.

A: To summarize, the reviewer proposes to add more quantitative estimates of model-data misfits and suggests to identify the possible causes of these biases.
In addition, he asks for a more comprehensive comparison between the coarse resolution simulation (which is not part of this GMD model documentation paper) and the high-resolution model simulation that is presented in this paper and suggests to include the objectives of the wind sensitivity experiment into the introduction.

We thank the reviewer for his constructive comments and we will add the respective information about the wind sensitivity experiment to the introduction (including its limitations). Following his suggestions we will, further, add more quantitative estimates of model-data misfits to the revised version of the manuscript. Finally, we will put additional effort into identifying possible causes for model biases in the revised version of the manuscript.

As concerns a comprehensive comparison with the coarse resolution model: it is work in progress and will be presented in a forthcoming paper which references the high-resolution configuration presented in this GMD paper. We find that a comprehensive comparison is beyond the scope of this manuscript which merely should document the high-resolution model configuration and map out potential applications of it. We realize, however, that the manuscript - in its current form - raises wrong expectations. We will refocus accordingly in the revised version of the manuscript.

Specific comments:

R: Abstract: The abstract makes the case that the model 1) combines a carbon inventory and eddy-permitting resolution in a way that outperforms existing models, 2) is suited to sensitivity experiments at decadal timescales, 3) can provide boundary condition to ice-sheet models. This manuscript would benefit from focusing the analysis and figures to demonstrate these 3 points, and from including a discussion section on the limitations of the model with regards to those 3 points.

A: We agree with the reviewer that that the abstract is misleading in its present form and will revise it. The major point of the configuration is not that it outperforms existing models. The point is that, although eddy-resolving in the Southern Ocean and near-global in spatial extend, the remaining drift after the spinup in simulated dissolved inorganic carbon content is sufficiently small to study decadal carbon inventory variability. To our knowledge this is the first configuration/spinup of a "free" (i.e. not data-assimilated) model.

R: Introduction: Overall, this section would benefit from a more focused approach, targeting specific dynamics or metrics that will be covered in the body of the paper.

A: Agreed. Yes, this is confusing. We will make a major revision.

R: Page 1, Line 20: There is no justification as to why 40S was chosen, or even simplyhow it relates to the above discussion about the location of the STF.

A: Agreed. The only reason was to make comparisons with other papers using 40S easier. We will change this text. The other reviewer did not like it either ...

R: Page 2, 29-31: Please consider how this fits within other work of Southern Ocean response to winds,such as (Spence et al, 2010).

A: Thanks for the reference. We will use it in the respectively revised consideration.

R: Page 3, 14: Where is this comparison between coarse and finer resolution model included? The manuscript would benefit from one.

A: In our opinion a full coarse resolution / high resolution comparison in this manuscript will make the manuscript too extensive and is beyond our aim. Our aim here is to document the high-resolution model configuration settings, link them to open-accessable model output and provide a reference point for forthcoming studies (including those comparing different resolutions). The forthcoming studies will need additional model evaluation because each scientific questions posed at a model typically asks for more "fit-for-the-purpose evaluation".

R: Page 3,15-22: the discussion of the role of meltwater on ocean circulation or biogeochemistryis out of place here. Given that the current configuration does not include sub-ice-shelf cavities or parameterized meltwater fluxes, it will not be able to account for meltwater feedbacks. At best, this could be included in a discussion section to discuss options for further development.

A: Agreed. We will put this more into the background.

R: Page 4, line 15: Some justification on why topography needs to be filtered 3 times (even just a qualitative target) is needed here.

A: Ideally, we would not have to filter at all. Numerical stability is guaranteed only for topographies so smooth that they barely represent actual conditions. A typical approach is to try with what little smoothing one can get away with. Turns out that 3 times is almost enough except for in a small confined area in the ACC where we had to apply additional viscosity. We will add a more formal explanation in the revised version of the manuscript.

R: Page 5, line 8-11:I have some concerns about the increase of viscosity over the Drake Passage. Some justification and possible root causes of this problem over Drake Passage need to be included.

A: This relates back to your last issue with topographic filtering. We will provide a more comprehensive explanation on these issues in the revised version of the manuscript.

R: Section 2.3 This section is too short and vague. If a change was made to the standard code and led to different results, please show the impact that this change has made (inan Appendix if necessary).

A: We will add more information on this issue in the revised version of the manuscript. In a nutshell: sea ice attracted more sea ice until the model crashed. The state before the crash shows a LOT of sea ice in one spot and none elsewhere while the revised version shows a sensible distribution of sea ice. We tried dozens of different settings for the sea ice model - all to no avail. The problem persisted. We really got desperate and send our approach to the MOM-discussion forum. The response was that levitating sea ice is a viable option. We found also that this is used in ne NEMO community.

R: Section 2.4 Design: please specify "complex-enough" for what? Or the model fidelity inrepresenting what? At least include the target variables and their scientific relevance.

A: Agreed. We will revise the text accordingly.

R: Why is the comparison with the coarse-resolution model not included in this paper?This manuscript needs to be self-sufficient.

A: The aim of the manuscript is to document the high-resolution configuration. The coarse resolution configuration already published (https://www.biogeosciences.net/14/1561/2017/bg-14-1561-2017.html). We are sorry for the confusion. A thorough comparison has not been intended. We will state more clearly that we only intend to document our high resolution model configuration rather than raising the expectation that the paper contains a full high-res coarse-res model comparison.

R: Section 2.6 Please include the frequency (or time scale) of boundary restoring. Also, the boundary restoring is necessary because the domain does not include dense water formation regions (it is not simply lack of processes).

A: Agreed. We will add the respective information to the revised version of the manuscript.

R: It is unclear what the wind experiment is supposed to address. If it is supposed to mirror the historical trend, please address why the physics are set to a historical trend while the carbon is maintained at a pre-industrial level. The objective of this experiment needs to be set up here, to provide context for the interpretation of results and discussion of their applicability.

A: Agreed. We will add the respective information to the revised version of the manuscript

R: Section 3. Why does this section not include any comparison with reanalysis data or state estimates? These datasets are commonly used in the literature, and providemore comprehensive data (sub-surface, time-varying...), which would help make a more compelling case for the fidelity of this new model configuration. As it is, thissection appears extremely light and unconvincing.

A: The problem we have with reanalysis data and state estimates in this context is that they rely on very similar models than the one we present here. Here is an over-exaggeration of my problem: when comparing our model with a state estimate/reanalysis product in a region/time-interval where there are no observations, we essentially compare one model with another. Now, they could be similar without actually representing reality. So, in this specific case we find a comparison between our model and state estimate/reanalysis misleading.

R: Section 3.1.1. This section lack qualitative evidence of the model's ability to represent the ocean circulation metrics mentioned.

A: We will add more quantitative information in the revised version of the manuscript.

R: It is necessary to define the meridionaloverturning as calculated here. Without a clear definition, the quantitative results are difficult to interpret or compare to other studies.

A: Agreed, sorry, we will fix it (Reviewer 1 stumbled over the same issue).

R: Section 3.1.2 This section lack qualitative evidence of the model's ability to produce 'realistic' levels of EKE. Given that this is a main point (included in the paper's title), putting plots side-by-side (Fig 8) is far from enough. Please include a plot showing the difference between the two datasets, and a quantification of mean EKE per region to provide some level of quantitative assessment.

A: Thanks for the constructive comment. Will be done in the revised version of the manuscript.

R: In addition, some discussion of the results is needed: what are some of the biases present? What may cause these biases? Finally, the conclusion that a 'realistic' level of EKE necessarily equates a good representationof eddy-driven processes is simplistic. Here, the only metric is surface EKE, where as eddy-driven processes, (including eddy-driven upwelling of nutrient-rich waters) occur over a range of depths. There is a vast body of literature investigating eddy processes in the Southern Ocean and what resolution may be necessary/sufficient to represent them adequately. Some discussion of this literature, and of how this particular model configuration fits in the context of other modeling studies is necessary.

A: We will elaborate on this in the revised version of the manuscript.

R: Section 3.1.3 Again, some quantitative comparison between the observation-baseddatasets and the modeled values is needed (e.g. plot of the difference), as well as a discussion of the biases (especially between the bottom temperature values, which has no discussion of biases at all). Combining Fig 10 and 11 would help the comparison.

A: Thanks for the constructive comment. Will be done in the revised version of the manuscript.

R: Section 3.1.5 A more comprehensive assessment of sea ice would make a more compelling case (e.g. sea ice concentration, sea ice thickness, annual cycle of sea ice area).

A: We agree and will see what we can do (i.e. what observational products we can get).

R: Section 3.1.6 This section needs a more comprehensive assessment of the model'sperformance in representing observed patterns of biogeochemical properties. A comparison to the Biogeochemical Southern Ocean State Estimate (B-SOSE) (Verdy andMazloff, 2017) would be a good step forward.

A: I can understand the reviewer's push to promote B-SOSE because it is a good product. I do also agree that output from physical-biological data assimilation models are super valuable for a lot of purposes. In this specific case, however, I disagree: B-SOSE is based on a 1/3 degree (i.e. non-eddy resolving) model. Wherever there are no observation I would essentially compare my "free" eddy-resolving physical-biological data with a non eddy-resolving model. The breach in logic here is that: We do not know yet if non eddy-resolving models are sufficiently realistic. They may well be, but we do not know yet. The eddy-resolving configuration presented here will be used to work on this question. Please note that our only aim, for now, is to document the model settings of this configuration. We will make this more clear in the revised version of the manuscript and apologize for confusing phrasing in the original version of the manuscript.

R: In addition, the biogeochemical performance of this model configuration should be shown (not just said) to be comparable to the one from the existing coarse resolution model. Likewise, there is little discussion of the possible causes of the biases, and lack of evidence to support the possible causes mentioned.

A: O.K.

R: Section 3.2 It is unclear what this sensitivity experiment is for, and why only the winds were changed.

A: We will add the respective information in the revised version of the manuscript. (This section has been motivated by Lovenduski, N. S., Long, M. C., Gent, P. R., and Lindsay, K.: Multi-decadal trends in the advection and mixing of natural carbon in the Southern Ocean, Geophys. Res. Lett., 40, 139–142, doi:10.1029/2012GL054483, 2013.)

R: It lacks discussion of the mechanisms leading to the change in over-turning circulation, or to the change of bottom water temperature. For the carbon results, it should be specified that the change in the experiment is showing only the

impact of the physical adjustment to winds (given that the carbon concentration is main-tained to pre-industrial levels).

A: Agreed - we appear to promise more than we deliver. All we actually wanted to do is to sketch out potential applications of our configuration rather than presenting new science. We will adjust the text accordingly.

R: Section 4: summary and conclusions This section makes qualitative statements about the model's fidelity, which have not been adequately supported by the body of the paper(similar to the abstract). As the model performance with respect to biogeochemistry is said to be similar to the coarser configuration, it is essential to demonstrate the benefits of this configuration with respect to eddy processes. Describing the configurationas a 'quantum leap' in modeling sounds over-reaching, given the lack of quantitative evidence included in the current manuscript to demonstrate the improved performanceof this model configuration compared to existing configurations. This section needs to include an in-depth discussion of the relevance and applicability of this model configuration. For example, the time scales examined here are the decadal timescales, while modeling of ice-shelf melt or Antarctic Bottom Water are more relevant to longer timescales.

A: We apologize for the confusion. The "leap" refers to the fact that the "free" (i.e. not data assimilated in the Southern Ocean) eddy-resolving coupled ocean circulation biogeochemical model presented here has been successfully spun up such that the (spurious) trend in simulated Southern Ocean carbon content is small enough to study decadal variability and underlying mechanisms.

The way we presented the configuration seems to be misleading in that it implies that the fidelity of our model in terms of reproducing observations is a quantum leap. This is certainly not the case – the model performance is just what is to be expected from the current generation of these "free" models. Model configurations which apply data assimilation are - naturally - much closer to observations (even the coarser, non-eddy resolving).

The benefit of our "non-assimilated" configuration presented here is that it provides an additional tool to explore the interplay between atmospheric drivers, circulation and oceanic biogeochemistry. This interplay is of relevance because it affects processes of societal concern such as oceanic heat transport to ice-shelfs and oceanic carbon uptake. We thank the reviewer for highlighting this aspect and will clarify this issue.

Technical Corrections:

A: We thank the reviewer **for the time he put into identifying and listing the technical issues and his/her** constructive suggestions. We do not list them here - but, naturally, will take them all into account.