A: Referee #1 asks for more details (including quantitative measures) in the validation of the physical circulation. Further he is interested in details regarding the biogeochemistry in the model's reference run and short sensitivity runs. In addition, he wants to see a more comprehensive discussion of the "overturning concept" in the introduction.

The specific comments of Referee #1 are:

Major Comments
R: page 2, lines 4-9 : This is not an accurate representation of current thinking on the overturning circulation of the Southern Ocean. There is no debate as to whether the bolus overturning due to eddy fluxes of thickness opposes the so-called Deacon cell. Itis extremely well established that it does so, to the degree that the Southern Ocean literature does not discuss the Deacon cell and instead talks about Eulerian overturning and bolus overturning. There is an on-going debate regarding the degree to which the bolus overturning cancels out the Eulerian, and that to which their residual responds to forcing changes, particularly with respect to wind stress changes, and model differences. However, that this cancellations takes place is widely accepted. See, e.g.,Marshall & Radko (2003), Viebahn & Eden (2010), and Abernathey et al. (2011), etc.

A: We will include a more comprehensive discussion of the overturning in the introduction of our model description paper and cite the papers proposed by the reviewer in the revised version of the manuscript.

R: page 2, line 24 onwards : There is at least one example of a fully-equilibrated, both thermodynamically and biogeochemically, model study that the authors could cite here; Munday et al. (2014). Whilst the model in Munday et al. is too coarse to be truly eddy resolving, it does have substantial internal variability and large-scale vortices, which leads to changes in sensitivity to wind stress of the physical circulation consistent with higher resolution models

A: We will add the respective reference to in the revised version of the manuscript.

R: Section 2.1 : Also of importance is that, since z* is essentially an extension of the nonlinear free surface method of Campin et al. (2004) to all model levels, it also gives very accurate conservation of tracers.

A: Thanks for pointing this out. We will highlight this advantage of our choice of numerics accordingly in the revised version of the manuscript.

R: page 5, lines 4-5 : The Smagorinsky scheme has a single coefficient, usually taken to be in the range 2-4, that allows some control over the viscosity. Please add the coefficient value you chose here. In addition, the issues arising near steep topography suggests that this coefficient was too small. Is there a reason

why the authors chose to introduce additional viscosity instead of just increasing the viscosity value? Additionally, some of the issues could result from the use of purely Laplacian viscosity. Given the grid scale variance at the resolution of MOMSO, it is probably appropriate to start applying some biharmonic viscosity/diffusion to the model.

A: We will add the information on the coefficient in the revised version of the manuscript. Further, we will elaborate on our attempts to fix the problem which indeed included biharmonic diffusion and/or viscosity, overall higher Smagorinski coefficient, decreased winds, changed bottom topography and changes to the vertical diffusion/friction. We really went at it - but did not list all of our failed attempts to fix the problem in the original version of the manuscript.

R: page 5, lines 5-7 : The PPM scheme is a good choice for the advection of temperature/salinity. The choice of advection scheme for the biogeochemistry is also veryimportant (see Levy et al. 2001). Having different schemes for these two choices could have repercussions. Have the authors investigated this?

A: No, we have not investigated this. But we agree that advection numerics adds uncertainty. We will state that in the revised version of the manuscript. We also agree that a comprehensive study on the effects of advection numerics would be very interesting. We expect repercussions even if the same scheme is used for all (passive and active) tracers because spurious numerical effects are dependent on tracer gradients and those differ between temperature and salinity which affect the saturation state of carbon dioxide and DIC on the other hand. We will add a paragraph on this issue in the revised version of the manuscript.

R: Section 2.5 : Initialising from a previous stratification/tracer distribution can be a very good way to shorten the spinup of high resolution models. Is there a reason why the temperature and salinity from the same run that produced the biogeochemical tracers wasn't also used?

A: 3-D model initialisation is indeed an issue and we will add a paragraph on this subject in the revised version of the manuscript. Ideally, we would have liked to start all prognostic variables of the high-resolution model from observations (and we would have liked the high-resolution model to stay close to these initial conditions). Unfortunately, observational data coverage is very sparse for the biogeoschemical tracers (especially for iron). Thus, we decided to opt for a model-based product. Since we planned a comparison with a coarse resolution model anyways we decided to use the coarse resolution model output as input. In a way, we explore how a coarse-resolution biogeochemical model state "performs" in a model with "enhanced" (higher resolution) physics.

R: Section 3.1 : A general criticism of this section is that it is largely qualitative, even when quantitative comparisons could be carried out, e.g. the spatial pattern and values of EKE/SSH variance are available from the altimetry the authors use. A properly quantitative comparison would allow more detail to be drawn out. In addition, the authors highlight the surface velocities and EKE as having "remarkable" comparison to the altimetry. This is hardly remarkable; the surface velocities and EKE look largely as one would expect for a model of~1/5o grid spacing in the Southern Ocean (see, e.g. Delworth et al., 2012, or Barnier et al., 2006, for examples of how similar resolution ocean models, or ocean components of coupled models, look). In addition, the deficiencies in particular comparisons are thoroughly glossed over. Something that stood out to me is that the flow is too zonal and there is too much EKE south of Australia. The East Australia Current is also poorly represented and the Agulhas retroflection has too small a region of high EKE. I'd expect the mean flow and eddies south of Australia to be related to the high temperatures south of this region shown in Figure 9.

A: (1) concerning "remarkable EKE comparison to altimetry": This is a misunderstanding that we will fix in the revised version of the manuscript. It is true, the configuration presented here is featuring a performance comparable to other high-resolution configurations. The selling point here is that the configuration has a sufficiently-spun-up biogeochemical cycle to address some carbon-related questions.

I started my career with 1/3 degree models and I am still amazed by the remarkable realism of (almost) ALL high-resolution models.

(2) concerning "too much EKE south of Australia/East Australia Current/ Aghulas retroflection": We will include a more detailed discussion in the revised version of the manuscript. E.g. the East Australia Current is already affected by the coarser resolution which presides outside our zone of interest (i.e. outside the Southern Ocean).

R: Most of the verification of the model is carried out on the surface quantities. Whilst this is appropriate given the length of the model run, which means that below the mixed layer there might not have been a great deal of adjustment, it is still useful to consider it briefly as it would place the later biogeochemical validation in context. At the very least, they should look at some temperature/salinity/density/biogeochemical transects across the Southern Ocean and the pycnocline depth. This will likely tie into the noted low value of the ACC transport. There is no discussion of the mixed layer depth. This is fairly well observed in recent years thanks to Argo floats and would be a key parameter for the exchange of carbon with the atmosphere and for the subduction of biogeochemical tracers from the surface.

A: We will add a meridional section of T/S and PO4. Further we will show a comparison with surface mixed layer depth.

R: Section 3.2 : Unfortunately, this section is completely deficient. There are many ways to calculate overturning (see discussion above about Eulerian vs. residual vs. bolusoverturning) and there is no information given here as to what variety of overturning is presented in Figure 23. It should be some measure of the residual circulation, although given the increase in the WIND simulation it looks like it might be the Eulerian over-turning. A better comparison would be to calculate the residual overturning in density coordinates and compare the residual overturning, as well as the Eulerian and bolus overturning. This would give a much better sense of how the overturning is changing between the two experiments and how much the remaining drift in REF is affecting their results. If the subsurface stratification is changing, then it will make interpretation more difficult, since the there could be significant diapycnal transformations in the streamfunction. This is another reason to look at the subsurface stratification earlier in the paper.

A: We will clarify the computation of overturning in the revised version of this manuscript. A comprehensive study between the experiments WIND and REF is in preparation but beyond the scope of this manuscript. This GMD manuscript is a first step. It is dedicated to describing the model setting and hinting at potential projects which could be carried out within such a model framework. We will clarify this in the revised version of the manuscript.

Minor Comments

R: page 1, line 20 : The choice of 40oS is not well justified, why not 35oS or 45oS? Thereare physical aspects of the circulation that could be cut off by this choice, for example the region in which the Agulhas current interacts with the ACC is very close to 40oS.

A: We will clarify that the choice of 40S has been motivated by other studies (to some of which we compare our results) using this exact threshold. We agree that this choice is arbitrary and not suitable for all purposes.

R: page 6, line 5-8 : Do the climatological estimates include the seasonal cycle?

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

R: Section 3.1.4: A missing quantitative comparison here is between the flux due to the surface restoring of salinity and the other surface fluxes that contribute to salinity changes.

A: Agreed, we will add the respective information in the revised version of the manuscript.

A: Figures such as 14 & 15, where the reader is invited to compare the REF and WIND experiments variation with the observed quantity would work much better as a single figure. Preferably as a single panel, with the observations as a third line.

A: Agreed we will change that in the revised version of the manuscript.

R: Section 3.1.6 : Based on Figures 19 & 20, the average concentration of both iron and phosphate appears low. Is this because the initial conditions for these fields are biased low? And if so, could the issues raised in this section be improved by simply adding more phosphate & iron?

A: I guess so, and yes, I guess this could be improved by changing the initial conditions - however - at a price: an increased drift towards just these low biased conditions. We will look into this and add a discussion in the revised version of the manuscript.

R: page 12, line 26 : In what sense to the compiler optimisations break reproducibility? Will the same model year, run on the same machine, be different if the model was rerun? Or will it only change when run on a different machine with different compilers?The first is rather alarming, the second would be very common.

A: Yes, the former (depending on compile options). This is well known for some (very performant) compiler settings and we have tested the effect of this in a similar setup. We will include more information on this issue (including some links to intel's compiler documentation) in the revised version of the manuscript.

R: page 13, line 12 : Quantum leap is an overstatement, given the length of the physical model spinup, the still low resolution of the physical model and the documenteddeficiencies of both physical and biogeochemical system.

A: Sorry, for the offensive phrasing - we will change it in the revised version of the manuscript. All "free" (i.e. not data assimilating) eddy-resolving ocean-circulation biogeochemical Southern Ocean model configurations we are aware of are of limited use in terms of simulating decadal changes of Southern Ocean carbon inventories because of either (1) they are regional configurations that are affected by prescribed spatial boundary conditions or (2) they could not be integrated long enough (spun-up) such that the remaining model drift is significantly less than the signals under considerations or (3) both the latter and the former.

Technical Issues

R: page 2, lines 13-14 & 27 : Work in progress for whom? The authors or the community in general, it could be either!

A: We will change the phrasing.

R: page 2, line 14 : Put comma after progress and delete comma from after extent.

A: Thanks!

R: page 2, line 18 : “As for now we know”?

A: We will rephrase this expression in the revised version of the manuscript.

R: page 3, line 11 “by chance”, isn’t it more by design?

A: We did not expect that the high-resolution model configuration is so similar to the coarse resolution simulations we carried out so far. We expected substantial differences. These potential differences were the motivation to switch form coarse to high resolution. We will delete "by chance" in the revised version of the manuscript because it is confusing.

R: page 3, line 13 : “allows to test” -> “allows us to test”.

A: Thanks!

R: page 4, line 9, page 7, line 10 : Orphan lines like this pop up quite a bit throughout the paper.

A: You are right! We will comb through and delete annoyances like these in the revised version of the manuscript.

R: page 4, line 19 : 42.429.759 -> 42 429 759

A: Yes.

R: page 5, line 15 : I think the authors mean vicious cycle! (Even though viscous seems appropriate for numerical models)

A: Ups, yes!

R: page 5, line 24 : Galbreith -> Galbraith

A: Sorry!

R: page 8, line 5 : Spacial is an accepted alternative to the more common spatial spelling. However, the paper uses both spellings at different places.

A: We will make it consistent.

R: page 10, lines 21-24 : Is the reference to "module" meaning a specific part of the numerical code? or an inaccuracy in the overall representation of part of the system?

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

R: page 12, line 16 : There's something wrong with the brackets here.

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

R: Figure 23 : noetig? in the caption.

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