Responses to Reviews of manuscript "A near-global eddy-resolving OGCM for climate studies" (gmd-2016-17) by X. Zhang, P. R. Oke, M. Feng, M. A. Chamberlain, J. A. Church, D. Monselesan, C. Sun, R. J. Matear, A. Schiller and R. Fiedler.
(The reviewer's comments are in back and our responses are in blue)

Referee Comment by anonymous referee #3:

I find little merit in this manuscript and feel it is unsuitable for publication. The paper describes a collection of ad hoc fixes and patches to the forcing of a high-resolution global model to compensate for shortcomings in the formulation of the problem (nonglobal, lack of sea-ice). The methods described are typical of those used in the 1980's and 1990's when available surface forcing products were far inferior. The apparent goal of the authors is to share "Experiences gained from our numerical experiments : : :" (pg 2 line 15). I cannot agree that this experience is something any other group should be following in the 21st century.

Please see our General Response #2 in the beginning about using "old" technique, hope you are convinced that there is still some room to use them wisely.

Also we want to point out that the OGCM for the Earth Simulator (OFES, Masumoto et al. 2014) developed by JAMSTEC used the same near-global domain (75S to 75N) as ours. The OFES model simulation is still one of the most popular eddy-resolving modelling to do climate-related studies Their model is neither fully global nor includes sea ice, but it is considered useful by the climate community.

There is no innovation in ocean model development described in this paper. There are some changes made relative to the model formulation described in the previous paper by Oke et al (2013), but none of these are new to the modeling community (e.g. the incorporation of the KPP mixing scheme). Further, the authors do not described the sensitivity of the solutions to these changes in formulation.

We want to clarify that this manuscript was submitted as a "Model experiment description paper", rather than a "Model description paper" or "Development and technical paper", according to the GMD guideline at http://www.geoscientific-model-development.net/about/manuscript_types.html. This paper is not about innovation in ocean model development, but about describing a model experiment based on an existing model. In particular, we developed new strategies to run the model so that it can be used for climate studies.

We modified the sentence to explain our usage of KPP mixing scheme.

The main argument for "innovation" in the manuscript is in the adjustment of the surface forcing to constrain global energy balance and addition of fictional forcing in the deep ocean to limit drift. The procedures are described as they were implemented without discussion of their merits relative to other alternatives, nor of why the particular choices were made. In the case of the surface flux adjustment, an obvious and simpler method would be to assess the global imbalance in the JRA-55 forcing using observed SST in the bulk formula prior to model integration and subtract the necessary correction a priori. What is the advantage of the iterative approach described in the paper? Why was the adjustment applied to long wave down flux? Why not reduce shortwave down? There is no rationale provided for the choices made.

Please see our General Response #1 in the beginning about our motivation of this model experiment. For both constant flux adjustment and non-adaptive restoring in the deep ocean, we have given enough details about their purposes and procedures.

For surface heat flux adjustment, we could adjust either downward longwave radiation or downward shortwave radiation. We chose to adjust longwave for its easy implementation. The results would not be sensitive if we chose to apply the small adjustment to the downward shortwave radiation.

For the iterative approach, please see our response below related to your comment about pg6 /line 27-bottom.

In the case of the deep restoring, the authors imply that this is common practice in high resolution modeling, but only reference their own work as an example. I cannot think of a single study since the pioneering work of Semtner and Chervin in the 1980s that has used deep restoring in a forward integration of a high-resolution model. The main “innovation” the authors claim is the introduction of “non-adaptive” restoring. This is not a new idea. The same basic ideas are described in for example Eden et al JPO, 34, 701-719, 2004 (and likely a number of other papers).

About non-adaptive restoring in the deep ocean, we give some detailed explanation in General Response #2 in the beginning, and also gave one example (Zhang et al. 1998) below related to your comment about pg4 /line 24. Eden et al. (2004) is also cited now.

Beyond the generally poorly motivated paper, the manuscript is quite superficial in its assessment of the solutions, and largely speculative when it comes time to attribute bias. Phrases like “may explain” or “could indicate” appear throughout when a more quantitative evaluation at a process level is called for. The aspects of the manuscript dealing with the BGC simulation seem to have been added as a complete afterthought.

We add more quantitative discussion now, and further discuss the motivation for including BGC in this experiment.

We are particularly interested in how the biophysical coupling is affected by mesoscale dynamics. As mentioned in the introduction, previous results indicate that mesoscale processes can modify the mean state. By including initial BGC results here we are preparing for further analysis, not just of the historical period, but also future climate projections.

DETAILED COMMENTS

pg 4 / lines 14-16. What is the implied global net heat flux for the JRA-55 product when using observed SST? This is a direct indication of the expected drift.

The usage of bulk formal to calculate heat flux on the fly makes the heat flux applied to the model ocean depends on ocean solution, including its global net heat flux. For example, in the Coordinated Ocean-ice Reference Experiments (COREs), a common atmospheric forcing dataset is applied to 7 global ocean-ice models, the global net heat flux is different among them (Griffies et al. 2009). In other words, the flux correction would be different for different models, even driven by the same atmospheric dataset. Therefore, in our experiment, we used this pragmatic method to derive the required flux correction, which was determined by the interaction between our model and JRA-55 forcing.

pg 4 / line 19: Why should the hydrologic cycle be balanced instantaneously? Cannot water be stored on land seasonally?

The balance of hydrologic cycle is not adopted because of model formalization. Our model can handle unbalanced freshwater flux.

We chose to balance freshwater flux at each time step, mainly because the quality of freshwater flux is insufficient to ensure a realistic temporal evolution of the global mean sea level. Since we are mainly interested in dynamic sea level (i.e., deviation of regional sea level from the global mean) rather than the global mean sea level itself, we chose to keep the global mean sea level as zero by balancing freshwater flux.

pg 4 / line 19: Global volume conservation is unrelated to the Boussinesq approximation

The sentence is modified to “The model adopts Boussinesq Approximation (Greatbatch 1994). The net global freshwater flux (i.e., sum of evaporation, precipitation and river run-offs) is balanced at each model time step, thus the global mean sea level is kept constant.”

pg 4 / line 24: if relaxation is "often used" provide some references of examples other than your own model

One example is Zhang J. et al. (1998), who specifically compared modelling results with and without climate restoring in their ocean-ice modelling experiment in the Arctic. They found that restoring of temperature and salinity has significant impact on prediction of the ice–ocean circulation in the Arctic. In particular, restoring salinity and temperature in the deeper ocean can reduce climate drift in the Arctic. This sentence is deleted because of restructure of Sections 1 and 2.

Reference:
Zhang, J., Hibler, W. D., Steele, M., and Rothrock, D. A.: Arctic ice–ocean modeling with and without climate restoring, J. Phys. Oceanogr., 28, 191–217, 1998.

pg 4 / line 33 : it is not a common technique (as above - provide some references)

See above

pg 5 / lines 8-22: This is a slew of no-conservative physics. There needs to be a fuller demonstration of its impact not just on global measures, but on the local structure of the solution. What is the spatial distribution of the restoring term? What spurious heat transports are implied by it? How does it impact the mesoscale?

Based on raised questions, I think you meant lines 8-12 (not 8-22) about restoring in the buffer zone near the northern boundary in the North Atlantic.

To minimize negative impacts of northern boundary, we designed a 5^o buffer zone which has linearly varying relaxation time scale, increasing from 30 days at the boundary to 365 days in the interior. "buffer zone" technique has been commonly adopted in many existing basin or near-global experiments, e.g., Smith et al. (2000) used a 3^o buffer zone at 72.6^oN in their $1/10^o$ North Atlantic basin model; Masumoto et al. also used a 3^o buffer zone in their $1/10^o$ near-global model domain from 75^oS to 75^oN. So our set-up is in alignment with many existing studies.

Although it's natural to think that global model should do a better job than regional model (with boundaries), it's not always the case. For example, Maltrud and McClean (2005) found that the performance of their $1/10^o$ global model is worse than a North Atlantic regional model (same model grid as the global model, but truncated to 20^oS to 72^oN) in representing the Gulf Stream/North Atlantic Current System. Although they didn't give clear conclusion about the contrast between global and regional models, but they suggested that better water mass conditioning through restoring at boundaries (20^oS and 72^oN) could play some role.

pg 5 / line 8 : what is the "correct" spin-up. The ocean state in 1979 was not in equilibrium with the forcing in 1979.

I think you meant pg 6 (not 5).

How to spin up a model is a challenging question. It may require tailored treatments for different model experiments. Here, we intended to strengthen the importance of "correct spin-up". The methodology we used in this paper may be one of feasible solutions to a similar set-up. The sentence and terminology are modified now.

The idea behind normal-year forcing experiment as proposed in CORE-I experiment (Griffies et al. 2009) is to run the ocean model with repeated normal year forcing (1979 in our experiment) long enough so that the ocean can finally reach equilibrium with forcing. Our model experiment basically follows the design principle of CORE-I.

pg 6 / line 15-20: How is the year-end discontinuity handled?

We didn't do special treatment on this. Following the protocol of normal-year forcing experiment by CORE-I (Griffies et al. 2009), we simply repeated the forcing year after year. We didn't notice any obvious discontinuity in our model fields (e.g., see Figs. 2 & 3).

pg 6 / lines 27-bottom: There is no justification provided for this ad hoc procedure. (see above)

This procedure is ad-hoc, but driven by the intention to make the heat imbalance smaller and stabilized as quickly as possible. We are interested in examining heat uptake and distribution in recent decades simulated by this model. The observation indicates a net heat imbalance on the order of ~0.5 W/m2. If we didn't adjust, the ocean absorbs heat at the rate of ~5 W/m2 (ten times of realistic magnitude) over at least the first several years, which leads to a much warm-biased ocean for us to start our historical experiment.

So the main motivation was to reduce excessive heat uptake during the spin-up, though the way how it's implemented isn't so critical. For example, alternatively, we could adjust during Year 3 to 5, then maintain the correction from Year 6 (rather than adjust during Year 3 to 7 as done in our experiment)

pg 7 / line 12 "does not necessarily imply a net heat flux correction" Of course it does. That is what you have constructed it to do!

If the model is forced with flux form forcing, any heat flux correction will be directly passed to the ocean. But if the model is forced with bulk formula (heat flux is calculated on the fly and thus depends on model state, rather than being provided), as implemented in our model experiment, some air-sea feedback can take place, which will affect heat flux. About this point, refer to Table 1, since Year 7, the model is repeatedly driven by the same atmospheric forcing through bulk formula until Year 20, but the annual net heat flux keeps changing (gradually decreases).

pg 7 / line 32-35: Why does it agree better later in the run? The SST has diverged further from the observed initial conditions. You should be comparing to OAflux for 1979, not its climatology.

The initialized temperature field (including SST) is from the CARS climatology (Ridgway and Dunn 2003) based on data collected over the recent decades, not from the 1979 ocean state. So the SST evolves from its initial climatological value, partially because the model is adjusting to reach equilibrium with the forcing in 1979. In the first several years, the ocean is still in the initial shock of spinup and a large and unrealistic heat imbalance is not a surprise.

OAFlux product (1983 to current) doesn't have data in 1979, thus climatology is plotted instead.

pg 9 / line 12 : "model may be too efficient" This is not a meaningful diagnosis and is purely speculative. Could be equally well attributed to any other process.

Determining the underlying processes needs some further investigation, such as the mixed layer heat budget. For this paper, we are trying to suggest (rather than diagnose) possible reasons by considering ocean dynamics and existing studies.

pg 9 / line 18: "This systematic difference : : :" The biases in the Gulf Stream and Kuroshio appear well to the north of separation points, an indication of poor boundary current separation. This is consistent with the biases int he SSH as well.

Thank you for your suggestion. The poleward position of the separation point in the model, relative to observations, is a factor. We now add some discussion on this.

pg 9 / line 29: "Their global means are almost the same : : :" By construction – you formulated to model to keep global mean sea level constant!

This sentence is deleted now.

pg 10 / line 28 "we do not repeat the detailed comparison" Then you don't need the detailed Table.

The sentence is revised. We also added observation-based transports for comparison, upon suggestion by Referee #2.

pg 11 / line 10: A more useful and critical comparison would be against the vertical structure of the RAPID overturning.

As the reviewer suggested, we compare the overturning stream function at 26^{o}N in the Atlantic between the model and RAPID Array observation (McCarthy et al. 2015), over the overlapping period 2004-2014 (Fig. B below). The maximum of the stream function, often referred to the AMOC, is 15.9+-2.4 SV from the model, close to 16.9+-3.5 from the RAPID Array. The depth of maximum stream function is around 1000 m in both model and observation. The temporal correlation of AMOC between model and observation is 0.66 over 2004-2014, but jumps to 0.81 over the later period (2009-2014). Fig. B is now included in the manuscript.

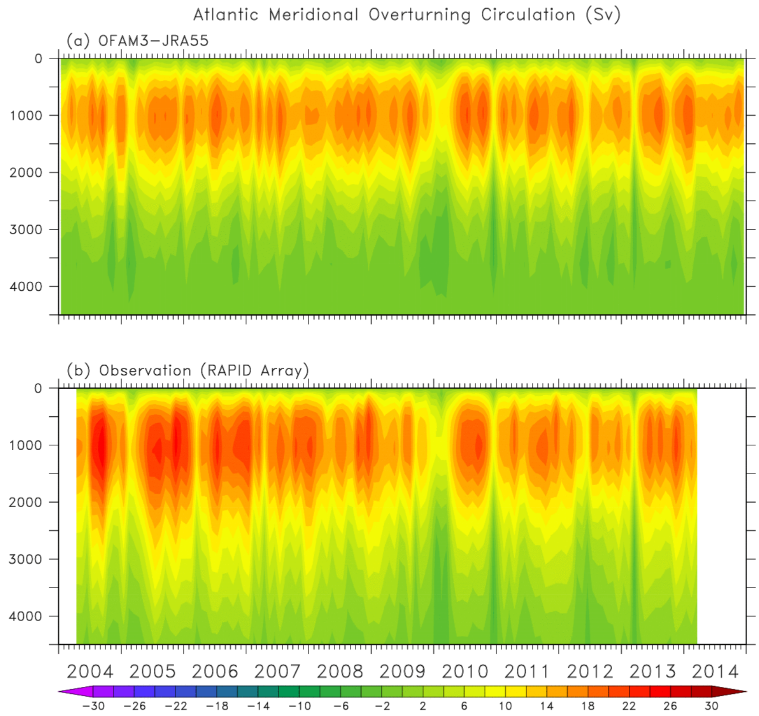

Figure B. Overturning stream function (Sv) at 26^{o}N in the Atlantic from (a) the model and (b) the RAPID Array observations over 2004-2014.

Reference:

McCarthy, G., and Coauthors, 2015: Measuring the Atlantic meridional overturning circulation at 26°N. Prog. Oceanogr., 130, 91–111, doi:10.1016/j.pocean.2014.10.006.

pg 11 / line 27 : They appear more dissimilar than similar to me. The model is completely missing most of the upwelling productivity and vastly overestimates the equatorial productivity.

We acknowledge similarities of the simulated BGC fields to observations are not as good as similarities of physical metrics (such as SST and SSH). Correlations of simulated BGC fields with observations are challenging due the fields being sensitive to both physical and biological processes and their uncertainties in the simulation, and also due to limitations and uncertainties in the observations.

As a way of comparison, we have calculated correlations of the mean fields shown to the available observations, and compared these to the range of correlations from a recent ocean BGC model intercomparison paper (Kwiatkowski et al 2014). Our correlation of primary productivity is 0.31, compared to a reported range of -0.08 to 0.64, and for carbon flux the correlation is 0.69, compared to a range of 0.01 to 0.68. These are now reported, in the manuscript.

Reference:

Kwiatkowski, L., Yool, A., Allen, J. I., Anderson, T. R., Barciela, R., Buitenhuis, E. T., Butenschön, M., Enright, C., Halloran, P. R., Le Quéré, C., de Mora, L., Racault, M.-F., Sinha, B., Totterdell, I. J., and Cox, P. M.: iMarNet: an ocean biogeochemistry model intercomparison project within a common physical ocean modelling framework, Biogeosciences, 11, 7291-7304, doi:10.5194/bg-11-7291-2014, 2014.

pg 13 / line 21 : as above - they are more dissimilar than similar. The largest observed variability is in the upwelling areas, not the central Pacific.

The position of BGC variability is related to gradients in the mean fields, which have biases that are larger than physical biases, but are reasonable relative to other BGC simulations, as discussed above.

pg 14 / line 14: "its possible that : : :" There is no need for speculation here. Smooth the model solution to the length scales used in the OI procedure for the observations and compare the smoothed fields.

Smoothing model SST (from 1/10^{o} to 1/4^{o}) doesn't change the result too much, still higher than the observation. In response to suggestion by Referee #2, we have added a comment to indicate that the model variability may be related to forcing, or due to unresolved, sub-mesoscale processes which are found to lower variability (Lévy et al. 2012).

pg 14 / line 22 : How can the mean upwelling be simulated poorly, but the variability be well reproduced?

Representation of mean and variability in models isn't necessarily tied with each other. Good representation of mean doesn't imply good representation of variability, and vice versa (e.g., Bates et al. 2012). Later in the paragraph, we did point some locations where model is different from observation. This sentence is revised.

Reference:

Bates, S. C., B. Fox-Kemper, S. R. Jayne, W. G. Large, S. Stevenson, and S. G. Yeager,: Mean biases, variability, and trends in air–sea fluxes and sea surface temperature in the CCSM4. J. Climate, 25, 7781–7801, 2012.

pg 15 / line 15: What are we supposed to be seeing in the single-month plots? The eddies will all be in different places for a forward model run.

The reviewer is right. Eddies from model and observation at any time aren't identical due the "chaotic" nature of eddies. We can only compare them by statistical measures, like long-term mean in Fig. 21a, b. Here we give the monthly snapshots to show the model can simulate a similar eddy distribution in those eddy-active regions (like western boundary currents) as observed by satellite altimeter.

pg 15 / line 31 - : The full depth OHC is completely determined by the surface heat flux, so it is not surprise these are similar. Why are there no comparisons with the vertical structure from the reanalysis products?

The reviewer is right - change of full-depth OHC is determined by the surface heat flux (because the model or the real ocean conserves energy). But the surface heat flux forcing is not necessarily the same for all ocean reanalysis products and our model experiment. In particular, heat flux is calculated on the fly through bulk formula in our experiment, rather than prescribed from atmospheric reanalysis, so the heat flux forcing depends on the model state, and we have no direct control over it.

Because the manuscript is already quite lengthy with 26 figures, we are planning a separate scientific paper describing the heat uptake and redistribution, and underlying physical processes based on this historical modelling experiment.

Only EKE is derived indirectly from altimetry sea level data. We repeated model EKE calculation using sea level rather than surface velocity, and got very similar results in most areas, except some regions (like western boundary currents) where ocean EKE can't be fully derived from sea level.