Reply to Referee Comment #3 (RC3):

The authors would like to thank the referee for the follow-up comments to the revised version of the manuscript. We would like to make the following replies, specific to each comment. The original comments are in *purple italic* font and listed in paragraphs, with our reply following each paragraph separately. The revisions are also highlighted in the revised manuscript in purple and marked by ***RC3***.

The manuscript has improved, and I am almost satisfied with it except for several small issues.

1. The main point of section 1.1 is that other existing solutions (AGRIF or unstructured meshes) do not have the capability adaptive grid refinement. I do not think this is the main point. There are unstructured-grid models that use adaptive mesh refinement, especially as concerns tsunami applications. However, in practical terms it would be very difficult to propose a situation when adaptive mesh refinement is needed in realistic ocean simulations. One generally knows where eddies are present, and static refinement is sufficient in most cases. Besides, remeshing destroys geostropic and other balances, which is not necessarily a welcome addition. This is the main reason why adaptive meshes are not used in ocean modeling, different from computational fluid dynamics.
If one is following the development of a shock wave, mesh adaptivity is very helpful. The discussion of missing adaptivity is therefore misleading.In my opinion, the interesting part of the manuscript is that the authors present a use case of an approach based on the separation of concerns, where the parallel infrastructure is separated from the numerical part, and this should serve as the main motivation.

Reply: the authors fully agree with the referee that there exist ocean modeling activities that utilize adaptive refinement. Therefore, we would like to revise the text to state clearly that OMARE is targeted at the adaptive refinement for the modeling of the ocean's general circulation. Hereby we would also like to make clarification of our understanding of the certain scenarios in ocean modeling that do require adaptive refinement. One particular scenario is the tracking of certain features that are of interest by dynamically changing. For example, the mesoscale eddies shed from the Agulhas Retroflection propagate into the southern Atlantic. By dynamically tracking these eddies with refined resolution (e.g., 10km and finer) would help better understand the salinity transport and climatic impact of these eddies, as well as their dissipation and interaction with bathymetry.

Another clarification is about potentially breaking the geostrophic balance during the simulation. Even with dynamic refinement, the model still carries out full simulation on the coarse resolution. Therefore, the geostrophic balance on the coarse resolution level is not violated. Indeed, across the resolution boundaries, there exist potential inconsistencies during long-term runs, and we are working on improving the model to feedback (or, in a sense, improve) the status on the coarse resolution. And this would alleviate the aforementioned inconsistencies, and it will be an add-on functionality of OMARE's next version.

The authors thank the referee for expressing the interest in the separation of model development concerns. We do consider this practice of value to the community, since it alleviates the model developers from the details of parallel computing, while maintaining good overall computational performance. According to the referee's comment, we revise the Section 1 accordingly. (1) We add the specific statement of the adaptive refinement that better clarify our motivation, as well as the role of Double-Gyre case which is to demonstrate this functionality. (2) We better introduce our motivation of refactoring/constructing models with JASMIN-like middlewares that better separate computation from model physics/numerics and alleviate the modeler's efforts.

2. I have already written previously that the extensive analysis of the double-gyre case is of very limited sense. The metric of KE is not very much telling because the author do not control energy input in their model. I would recommend to further shorten the description. As concerns diagnostics, it would be advisable to look at mean eddy kinetic energy distribution in space (for 0.5 - 0.1 degree experiments). In the adaptive refinement case one can hardly learn anything from multiple panels after 5 20 or 50 days. These time intervals are too short for studying real ocean. The authors raise a question on matching dynamics between coarse and fine subdomains, but only in passing. Do one needs, for example, to adjust dissipation at the boundary?

Reply: the authors agree with the referee that the energy budget of the Double-Gyre should be better accounted for. We do confirm with the referee that the energy input into the system, both dynamic and thermodynamic, is indeed constrained. As evident from the manuscript, the energy cycle of the 0.5-deg run and that of the 0.1-deg run are very different, especially given that no mesoscale turbulence and energy reservoir is present in the 0.5-deg run.

We confirm that we did investigate the eddy kinetic energy of the 0.1-deg run. The WBC and its extension is the 'hot spot' of kinetic energy. We now provide the figure as an extra supplementary to the revised manuscript, and provided below as well.

According to the referee's suggestion on the 5-day, 20-day and 50-day results of the 0.02-deg experiments, we further shorten and simplify the text. We intend to keep the figures for the content to be indicative of the Adaptive Refinement working as we expected. For the inter-boundary exchange, we first plan to better match the coarse-fine grid status by updating the coarse grid status by that of the fine grid for future work of OMARE (which we mentioned in the final section of the manuscript). Meanwhile, numerical dissipation may be needed after synchronizing the model status, which we plan to carry out further.

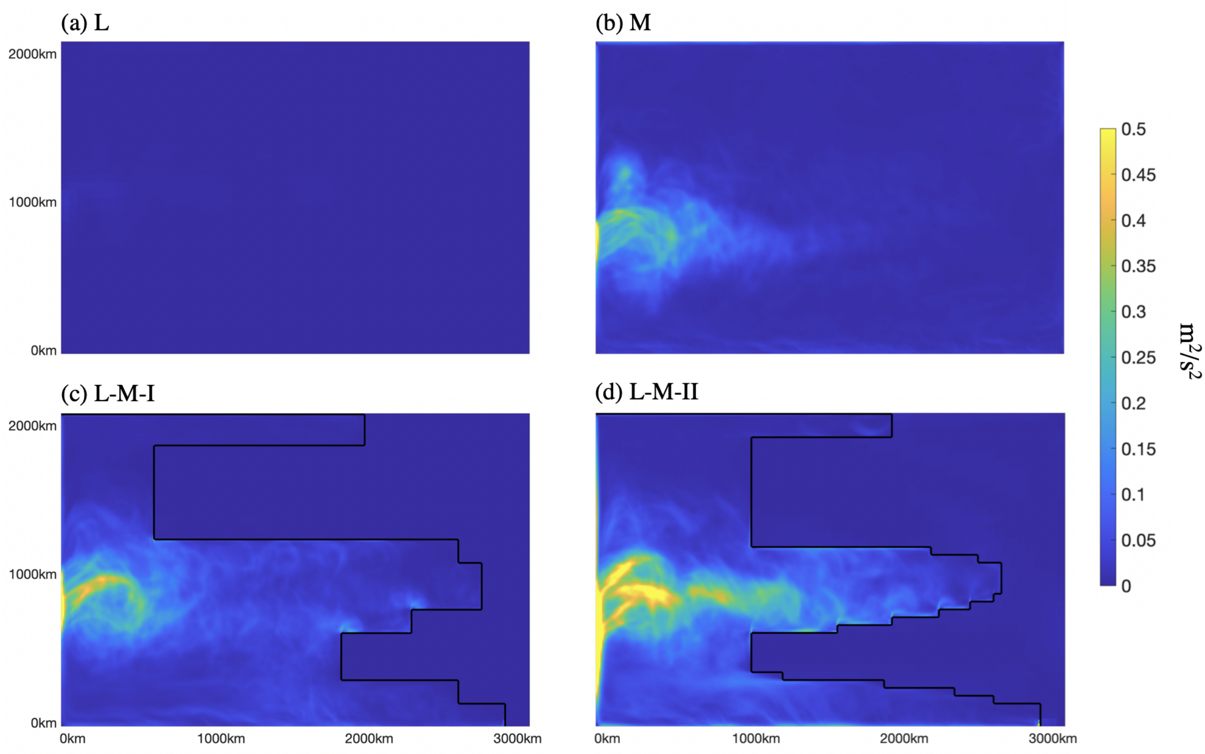

Figure. Annual mean eddy kinetic energy (m^2/s^2) for 0.1-deg simulation based on daily instantaneous model status.

3. There are still too many grammatical errors (singular-plural is the most frequent one), and some phrases have to be edited. I have not compiled the list, but please read the text very carefully, improvements are needed in many places.

Reply: we have gone thoroughly through the manuscript to correct for the errors. These revisions are also highlighted. We apologize for any inconveniences that are caused to the referee.

Below I specifically mention some places.

line 47 "First ..." --- and where is "Second..."?

Reply: we revised the second sentence to add the missing "*Second*".

line 57 "Similarly ..." -- it is not an English sentence

Reply: we revised the sentence from "*Similarly, with MPAS, grid generators …" to "There is similar practice for MPAS, which utilize grid generators …*"

Line 58 "Although existing models (which are based on orthogonal grids) ... " -- What are you willing to say? Do you mean structured-mesh models?

Reply: the sentence is revised as: "*Although existing structured-grid based models can no longer …*"

Line 60 "(1) the model grids ...." --- there is reason to avoid this

Reply: we have revised the sentence by segmenting it into 4 individual ones, as follows: "*However, there are certain limitations. First, the model grids cannot change arbitrarily with time, hence limited 'adaptivity' and 'flexibility'. Second, scale-aware parameterization schemes should be developed to accommodate gradual change of model grid resolution. Third, due to CFL limitations, the time step is usually controlled by the smallest grid cell size, resulting in extra computational cost.*"

line 64 "Furthermore, there is no ocean model ..." --- because nobody ever shown that moving grids are needed for ocean simulations. In neXtSim they are trying to preserve the structure of linear kinematic features, but why it should be used in the large-scale ocean?

Reply: we would like to clarify that to our understanding, we have the following motivation for using adaptive refinement for modeling the ocean's general circulation. First, the tracking of dynamic processes requires adaptive refinement, such as mesoscale eddies, ocean fronts, etc. Second, large-scale features such as polar sea ice and submesoscale processes are of clear seasonal characteristics, and they serve as potential candidates for adaptive refinement. Third, certain non-scientific ocean applications such as the tracking of arbitrarily moving objects/Lagrangian points. We would like to re-emphasize that in this work, we use Double-Gyre test case to demonstrate the AMR capability of OMARE, which is shown to simulate reasonable ocean processes. The possible use of OMARE and its AMR function is still subjected to specific simulation scenarios.

line 80 "tripolar boundaries" --- what is tripolar boundaries?

Reply: tripolar grids are commonly used for many high-resolution global ocean models as model grids. The tripolar boundary is the grid's northest boundary, which connects the 'eastern' hemisphere and the 'western' hemisphere of the model grid. We have revised the sentence and include the necessary reference to Murray (1996).

line 95 "including first-order and second-order viscosity/diffusion" --- What do you mean? Do you mean harmonic and biharmonic viscosity and diffusion? They can be of arbitrary order of accuracy.

Reply: we have revised the manuscript accordingly. Indeed we mean harmonic and biharmonic viscosity/diffusion schemes.

Fig. 1 Please provide an extended caption. I can hardly see what is the intention of this figure.

Reply: we have revised the figure caption to include relevant information of the figure. The main purpose of this added figure (in revision 1) is to contrast the model structure of NEMO and OMARE (i.e., after refactorization).

Fig. 3 Same comment. I do not see what I should conclude from this figure.

Reply: we have revised the figure caption to include relevant information of the figure. The main purpose is to demonstrate the typical time integration in OMARE, including the adaptive refinement cycle, which is a sequence of baroclinic steps of the coarsest resolution during simulation.

line 296 What is 'kinetic'? Do you mean kinetic energy? This issue is present in several places.

Reply: we revised “the surface ocean kinetics” into “the mean surface kinetic energy”. Other places of misuse are also corrected.

Section 3.3 and further -- use roman font for units

Reply: we have revised the manuscript accordingly, including the text and Figure 7.

line 318 PE is seemingly the potential energy, which is not stated. I do not see the definition of PE, and from what I can reconstruct from the text, the authors identify the PE with the difference in sea surface height which is an error. One is interested in equilibrated regimes and not in transient phenomena in toy models. Why all this discussion is needed?

Reply: we have revised to include the precise statement, which is the “mean potential energy (PE)”. We have now added the statement at its first use. At 0.5-deg, the model simulates systematically higher mean PE than that at 0.1-deg. As a future study, a holistic analysis of the energy reservoir and the transfer and comparison between 0.5-deg and 0.1-deg experiments are planned. We would also like to clarify that the comparison of 0.5-deg and 0.1-deg experiments, including the transitional phase, is to contrast the systematic difference of these full-field and regionally-refined simulations.