Reply to Referee Comment #2 (RC2):

The authors would like to express sincere thanks to the referee for the invaluable comments. The following are the replies for each point of the comment, together with specific revisions that are made. The original comments are in *green italic* font, and the reply to each of the comments are provided accordingly. Revisions are further highlighted in the revised manuscript in green and marked by ***RC2***.

The manuscript by Zhang et al describes an implementation of a nested regional ocean modeling capability using the parallel computing framework "J Parallel Adaptive Structure Mesh Applications Infrastructure" (JASMIN). They demonstrate the application of the implementation with simulations in an idealized basin double gyre configuration. The presentation is generally clear and the authors provide a realistic assessment of the remaining challenges with this particular implementation as well as for nested ocean modeling generally. My only significant concern with the study is that the entire work is based on a proprietary software stack, making the reproducibility and accessibility of the effort to the broader community questionable.

Reply: the authors thank the referee's comments on the fact that OMARE is based on JASMIN which is a closed-source middleware. During the early design stage of OMARE, we also faced the dilemma of the desired model functionalities (including AMR, automatic restart, parallel I/O, etc.) and the fact that there lacked (and still lacks) openly available high-performance computing middleware that satisfy these needs. We consider the choice of using JASMIN a worthwhile tryout to leverage the capabilities of traditional models. Furthermore, we do observe various cases in which commercial, closed source software does drive the geophysical modeling community forward, such as NVIDIA CUDA which is the main working horse for many fastest computers in the world.

Therefore, we have open-sourced the whole codebase of OMARE, complying with the rules of GMD, and provided comprehensive guides to applying and using JASMIN. In our opinion, JASMIN is a third-party, multi-purposed software that we utilize in constructing the model, but not the model itself. We definitely welcome the JASMIN developers and its governing bureau on any plan to open-source the platform. Besides, we are also actively searching for open-source alternatives to JASMIN for future developments.

I suggest below a few additional points and questions that I believe would be helpful to a reader. The grammatical construction (e.g. singular/plural usage) of the text is a little rough in places and could use some polishing by a technical editor. I have no substantial concerns with the manuscript as written and recommend publication after minor revision.

Reply: the authors have gone through the manuscript to correct small language issues, including those pointed out by the referee.

Line 31: "Other ensuing practice …" ?? Not right words here

Reply: we change "ensuing practice" into "follow-up modeling practice".

Line 37: "In the following up part of the paper …" -> In this paper ….

Reply: revised as suggested.

Line 58: suite -> suit

Reply: corrected.

Line 70 ASMIN -> JASMIN

Reply: corrected.

Line 114-119: It would be helpful to actually state the man-months of effort required to complete the refactorization.

Reply: as suggested by the referee, we have added the following sentence for the man-month information: *The whole code refactorization process took about 32 man-months to finish*.

*Line 128: 422 patches … Might be helpful to state where this number comes from .. (# of variables) * (# of grid hierarchy levels) * (???)*

Reply: we have reformulated the sentence into: *In total, during the refactorization we have formulated 155 components (i.e., FORTRAN subroutines) and 422 patches (i.e., model variables) in OMARE.*

Line 131, 133, 175 : Figure 3 should be Figure 1 or Listing 1?

Reply: for all three cases, *Figure 3* is corrected as *Figure 1*.

Line 180: does level here refer to depth level or grid hierarchy level?

Reply: we confirm that the depth level here refers to the resolution level (or the level of grid hierarchy). The text is modified to contain a more clear description of this information.

Line 295, Figure 5. Label the figures with the case name or include identifiers in the caption.

Reply: as suggested by the referee, we add the information of the experiment in the figure caption, as follows: *The background and the whole ocean basin is by default in 0.5-deg resolution and marked in cyan, and the region refined to 0.1-deg in each experiment marked in purple.*

Line 296: "Besides, L-M-I covering more area in the subtropical gyre " This seems backwards from the figure. Poor sentence construction.

Reply: we reconstruct the sentence as follows: *Besides, in the subpolar gyre, the refined region to 0.1-deg in L-M-I covers more area than that in L-M-II. On the contrary, in the subtropical gyre, the refined region is smaller in L-M-I than that in L-M-II*.

Section 3.3: A key conclusion of this section is that the mis-representation of APE->KE transfer in the eddy-parameterized parent-grid leads to boundary forcing biases in the child grid. Later in the paper, the authors allude to the need for a GM-type parameterizations in this regime. Does the implementation support different parameterization methods at different levels of the hierarchy (they show that different parameter values can be used)? An additional case with GM turned on in the 0.5 degree grid would be informative of the suitability of this approach with a scale aware parameterization strategy.

Reply: we confirm that different parameterization schemes could be applied to the different resolution levels in the OMARE system. The mixing schemes already differ among the 3 resolution levels in the numerical experiments of the paper. The experiments based on 0.5-deg relied on a 1st-order Laplacian mixing scheme for lateral mixing of the momentum equation, and the others (with either 0.1-deg or 0.02-deg) adopted a 2nd-order Bi-Laplacian mixing scheme. We plan to further implement other more sophisticated parameterization schemes in the future, in order for a more systematic analysis of the energy cycle and balance of the Double-Gyre system at different resolutions.

Line 346: “proxies” Not sure this is the correct word here. I believe you mean something more like “metrics of grid quality”

Reply: we change “proxies” to “indicating parameters”.

Line 398: bu -> by

Reply: corrected.

Line 423: LaTex typo for circle symbol

Reply: corrected.

Line 424 : “following up part” -> following part

Reply: corrected.

Line 442: We pick the a -> We pick a

Reply: corrected.

Line 474: This is indicated by that both … -> This is indicated by the fact that both …

Reply: revised as indicated by the referee.

Line 505-514: Could alternatively provide an estimate of the man-months of effort here.

Reply: the information of man-month is added as follows: *The refactorization process in total involves about 32 man-months to finish*.

Line 529: (Section on refinement criteria) : Some description of how the current implenmexation aggregates points into individual patches would be helpful.

Reply: we hereby confirm with the referee that the aggregation of model points into patches are mainly carried out by the middleware software of JASMIN. For the refined region, the user can specify the granularity of the refined region, which indirectly controls

the patch size for the refined region. For example, we use the granularity parameter of 3, and given the refinement ratio of 5, the refined region is then consisted of regions of 15x15 in size on the refined level (or 3x3 grid points on the coarse level), and the resulting patch size is 15x15 as well. This granularity parameter affects the computational performance and should be treated as a trade-off for model tuning. We consider it a technical detail and hence omitted it from the manuscript. As suggested by the referee, we add the following concise description in this section: *Specifically, a prescribed parameter for refinement granularity can be used to control the patch size on the refined region, ensuring both the full coverage to these marked grid cells and patch sizes which affects computational performance.*

Line 597 (Section on realistic cases): Some discussion of challenges expected with realistic topography and coastlines would be helpful. Would topography or the coastline be refined along with the grid? How would mismatches in land-ocean boundaries across hierarchy levels at the lateral boundary be handled?

Reply: as suggested by the referee, we have included some discussion of potential issues and challenges with realistic bathymetry, including the choice of refinement, as well as the cross-boundary consistency of model bathymetry. The added sentences are: "*Specifically, spatial refinement can be carried out in key regions with bathymetric features, such as land-sea boundaries, continental shelves, sea mounts. Due to the different bathymetry across the resolutions in OMARE, the model status on the coarse grid contains inherent inconsistencies for the refined region. Therefore, after spatial and temporal interpolation, the lateral boundary conditions to the refined region need to be modified accordingly, in order to reduce any potential physical and numerical issues*".

Other: I presume some data on scalability and performance will be presented in the companion manuscript, but a very brief statement about this would be helpful to make the present manuscript self-contained.

Reply: we have revised the last sentence of the first paragraph in Summary, in order to include the brief information for the planned accompanying paper. The revision is as follows: "*Another planned paper (part 2) will further introduce the computational aspects of OMARE, including the scalability and computability of OMARE, with a particular focus on AMR and its role in improving the computational efficiency of high-resolution simulations*."