Authors' Final Response to comments to <u>egusphere-2022-510</u>, entitled "*Ocean Modeling with Adaptive REsolution (OMARE, version 1.0) – Refactoring NEMO model (version 4.0.1) with the parallel computing framework of JASMIN. Part 1: adaptive grid refinement in an idealized double-gyre case*"

The authors would like to thank the editors and the two referees for the invaluable comments and suggestions, which helps us a lot in improving both our own understanding of the context and the manuscript. We have replied to the editors' comments and the referee's comments accordingly. In this final response we make a summary to these replies and attach the specific replies to the referee's comments (RC1 and RC2). Specifically, the reply to RC1 occupies the 3rd to the 8th page, and the reply to RC2 occupies the 9th to 13th page. The manuscript is revised accordingly, and the marked-up version provided.

In this final response, we would like to first re-emphasis our main goal of our work of OMARE, which is a refactorization of NEMO with the high-performance middleware software of JASMIN. Our main goal is to demonstrate that: **<u>advanced model functionalities, such as adaptive refinement and automatic parallelization can be achieved for current models with such a refactorization practice</u>**. In most models, the model parallelization is usually carried out by the model development team, which in effect results in a \`silo'-type software and undermines long-term development. But functionalities such as adaptive refinement (AMR) is a closely related model design to domain decomposition and parallelization, and it cannot be easily implemented and usually beyond the capacity of many model development teams. By utilizing JASMIN or similar general-purpose middleware software, the model developer is alleviated from these development efforts, and the model functionality can also be improved. **Furthermore, by adaptive refinement down to 0.02-deg resolution, OMARE simulates turbulent ocean processes and captures the dynamically changing mesoscale and submesoscale on the Western Boundary Current system**.

Second, we would like to summarize our reply to the status of JASMIN. **<u>We have provided several proofs and efforts alongside the manuscript to improve the usability of OMARE and JASMIN</u>**. (A) We have provided the full codebase of OMARE, consisting of over 40000 new lines of code, which complies fully with the GMD protocol. (B) We have provided several ways, including a step-by-step instruction, for the application of the proper version of JASMIN, which is currently released in binary form from its developing team. (C) We provide a brief guide for installing JASMIN, as well as compiling and running OMARE. These guides are provided as either Author Comment (https://editor.copernicus.org/#AC4) or the supplementary of the revised manuscript. JASMIN is independently developed and well managed at the Institute of Applied Physics and Computational Mathematic (IAPCM) and IAPCM retains all the copyright to it. We consider JASMIN an external, infrastructural software that we build OMARE with, but NOT the model itself. And we are fully aware of the various cases in the community that the infrastructural software is commercial (e.g., NVIDIA CUDA) and therefore not accessible in the open-source sense, and even the source of the model itself is NOT released (due to policy limitations of the governing body). Despite the practical limitations of using JASMIN, we hope that our practice could contribute a worthy try for the long-term development of geofluid models.

Third, we have finished the replies to both referees' general and specific comments, and attached them in this reply as well. The general comments from the referees mainly concerns the purpose of our work, the analysis of the experiments and the layout of the related text, as well as the usability of JASMIN (which we addressed above). We have finished item-to-item replies, and made revisions to the manuscript. The marked-up version of the revised manuscript is also provided, with specific notes to either of the referees' comments (RC1 or RC2).

We would like to express our sincere thanks again to the two editors and the two referees for their help to improve our manuscript. We would also like to send them our greetings for the upcoming holiday season and the new year of 2023!

Shiming Xu, on behalf of all authors

Reply to Referee Comment #1 (RC1):

The authors would like to thank the referee for the invaluable comments and suggestions. The following are the replies for each point of the comment, together with specific revisions that are made. The original comments are in *blue italic* font and listed in paragraphs, with our reply following each paragraph separately. The revisions are also highlighted in the revised manuscript in blue and marked by ***RC1***.

I find the approach proposed in the manuscript to be of interest for the journal. It follows the now discussed road of separation of concerns, whereby the code part dealing with numerical algorithms and the part dealing with the infrastructure (mesh, parallelization) are separated and treated in different ways. However, the manuscript in its present form misses the goal: one expects that the material of the manuscript describing a modeling approach will be sufficient for a reader to learn how to use the approach. I do not see that this goal is reached.

Reply: the authors thank the referee's comment on the relevance of our contribution to the journal. Regarding the comment on the main purpose of the manuscript, we would like to make the following clarifications and corresponding revisions to the manuscript.

First, we would like to clarify that we have multiple purposes for the manuscript: (1) we demonstrate that by refactoring existing model (here NEMO) with JASMIN we can attain extra functionality, while breaking the `silo'-type model development in which modelers make (almost) everything by themselves; (2) the resulting model OMARE is capable to simulate realistic ocean processes with the AMR function working as intended; (3) the introduction of the actual refactorization process, using NEMO and JASMIN as an example. We consider OMARE a worthy try in (re-)constructing the model with a software middleware and satisfactory results are achieved. We agree with the referee that the technical details of how to refactorize the model is an important aspect, but we consider it one of these goals. In response to the referee's suggestion, we provide a technical guide as supplementary material for using JASMIN and compiling/running OMARE (see also the reply to the next paragraph). The text referring to the new technical guide is also added to the revised version of the manuscript.

In order to better explain the overall design idea of OMARE, we intend to add the following figure of the model structure of OMARE and differentiate it with NEMO. The figure, together with the added paragraph to the manuscript is now included in the beginning of Sec. 2 of the revised version of the manuscript.

NEMO, as a typical ocean model, consists of a layered structure. Like many models, it relies on an intermediate layer that contains both a self-developed parallelization software solution, and third-party add-on's (AGRIF, XIOS). As a consequence, certain limitations exist, including limited flexibility in parallelization and adaptivity, especially given that these issues are inherently intertwined. OMARE, on the contrary, relies on JASMIN for managing the parallelization, adaptive refinement, and parallel I/O. Therefore, no extra effort is spent on designing/building the intermediate layer specifically for the model, hence no software `silo' is constructed. More importantly, the model does not suffer from the aforementioned limitations.

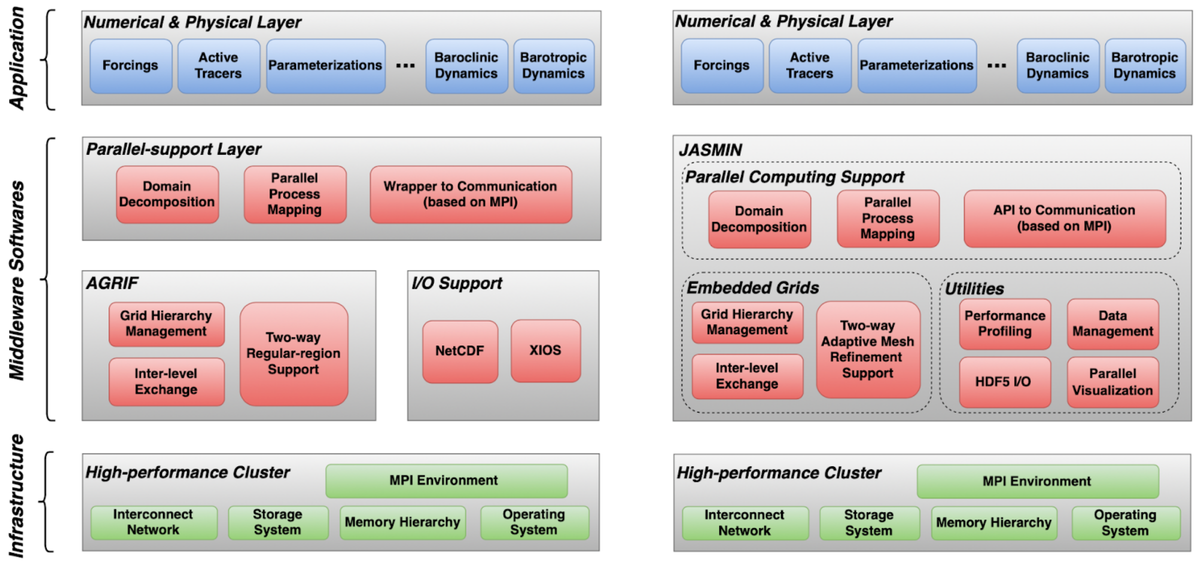

Figure 1. Model structure and abstraction layers of NEMO (left) and OMARE (right).

Of course, there are also limitations associated with OMARE, such as the lack of direct support for accelerator architectures (e.g., GPUs) in JASMIN. Besides, there are potential usability issues due to the current binary release form of JASMIN, which is an issue also raised by other editors. We are aware that these detailed issues should be addressed during future development of OMARE.

The text on lines 165 - 190 shortly describes how the JASMIN is involved, but I doubt a reader can get any understanding of what and why is done. Moreover, it is not at all clear how to use JASMIN in conjunction with the updated code. How the JASMIN environment can be installed, how code is compiled, etc. The description should be essentially extended and be such that those who are willing to follow author's approach can do it.

Reply: we agree with the referee on the importance of the technical aspects (of how JASMIN and the application work together). We do consider the detailed content will be too much to be included in the main part of the paper, considering the overall purpose of the work, as well as the limited relevance of these details only to model developers. Therefore, we supply a concise manual for the installation of JASMIN and the compilation and running of OMARE in the supplementary. The text here is also revised with the addition of the reference to this manual.

The manuscript devotes more than a half of its volume to the description of simple test configuration, going into too much detail, which is hardly optimal. The test case remains the test case, and one can only learn that the approach proposed by the authors is working, yet not without drawbacks related to one-way nesting (the development of errors on fine-coarse boundary). I do not think that this test case is well suited to demonstrate the need of adaptivity. Figure 11 shows clearly that small-scale turbulence occupies 3/4 domain on full mesh, and it occupies only pieces where the resolution is refined in b and c. It is different from the initial phases in Fig. 9 and 10, but 5 or 20 days is a too short time for turbulence to equilibrate, and this transient phase is of no interest (it depends on coarse initial conditions, and does not model any reality). So the conclusion here is that dynamic adaptivity is an interesting, but perhaps not very needed possibility as concerns eddying flows. Static refinement might be doing the work, and one will take a decision where to resolve based on one interest. I foresee, however, one direction, where dynamic refinement still might be of interest -- the simulations of seasonal course of variability. Submesoscale eddies might be suppressed in warm seasons, and a coarser mesh will be sufficient for mesoscale. My recommendation are to make the experimental part more compact. One can hardly learn anything from detailed description of particular eddy features or the comparison of transects (Fig. 13, 14), and there is very little sense in Fig. 9 and 10.

Reply: we would like to thank the referee for the comment on the test case. We would like to make the following replies and revisions to the manuscript.

First, on the choice of Double-Gyre testcase. The Double-Gyre testcase is chosen as a testbed for the major functionalities of OMARE. The case is of intermediate complexity: (1) the model physics is complete, making it capable to produce realistic three-dimensional large-scale ocean processes; (2) it produces an idealized Western Boundary Current system and in particular, the seasonality of submesoscale processes; (3) the testcase omits realistic (or any) bathymetry, therefore avoiding the complex issues such as inconsistent bathymetry between the non-refined and the refined regions. We agree with the referee that the Double-Gyre is a choice among many possible ones. And we chose it mainly because that the aforementioned 3 characteristics. Similar Double-Gyre cases or even more complex idealized cases such as NeverWorld2 (Marques et al., 2022) are also commonly used in the community for model benchmarking or the study of certain processes.

Second, on the demonstration of the need for adaptivity. We agree with the referee that small-scale turbulence is prevalent in the basin for full-basin 0.02-deg experiments, which is actually expected especially during winter. Here we would like to emphasize our perspective on adaptive refinement: where & when we need adaptivity and AMR is an open issue, and the purpose of OMARE is to provide a framework that supports various possible AMR scenarios.

In the Double-Gyre case, the major simulated features include the WBC, the lateral boundaries of the basin, and the associated mesoscale-submesoscale processes. <u>The refinement criteria are designed to capture the kinetically active and submesoscale</u>

processes, and hence based on surface velocity and surface relative vorticity. The two AMR cases have already been shown to capture a majority of the basin's (surface) kinetic energy (Fig. 8), which directly indicates the validity of AMR.

A further proof is provided below in Fig. 2 for the vertical motion induced by submesoscale flow. The enhanced vertical motion and the associated heat (and other tracer) transport is a key characteristics of the submesoscale processes (Taylor & Thompson, 2022). The following figure shows that: (1) the strong vertical motion is strongly concentrated at WBC and some other regions (i.e., ocean fronts, basin's boundary), and (2) the two AMR cases capture these key regions for strong vertical motion. A further investigation of the PDF of vertical speed in Fig. 3 confirms that the strong vertical motion in AMR experiments matches closely with the full-field 0.02-deg experiment.

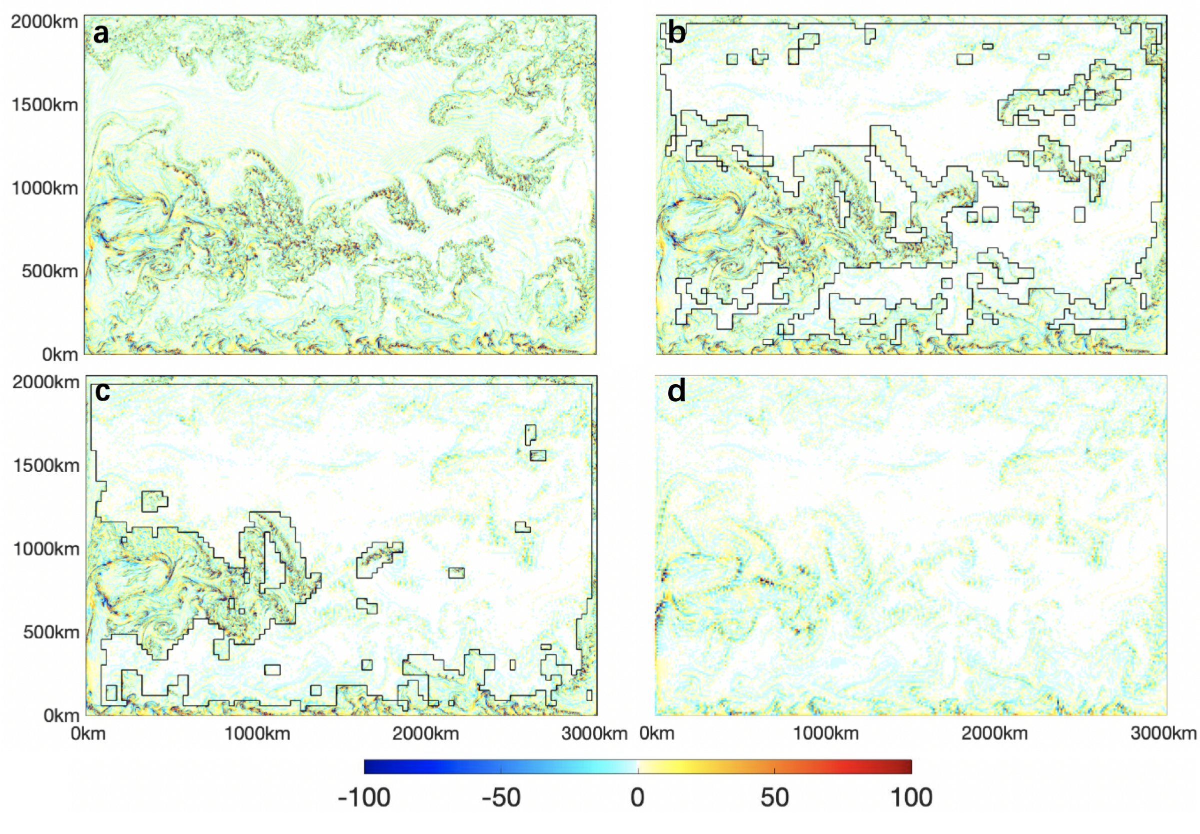

Figure 2. Vertical velocity (w, in m/d) at the 5-th layer after 20 days since Feb.-1st for all experiments. The depth is about 40m. Panel a, b and c shows the result for full-field 0.02˚ experiment (i.e., **S**), that of AMR setting 1 (i.e., **M-S-I**), and that of AMR setting 2 (i.e., **M-S-II**), respectively. In both AMR experiments (panel b and c), the boundary between 0.02˚ and 0.1˚ is marked by black lines. Panel d shows the reference simulation (i.e., **M**) result at 0.1˚ on the same day.

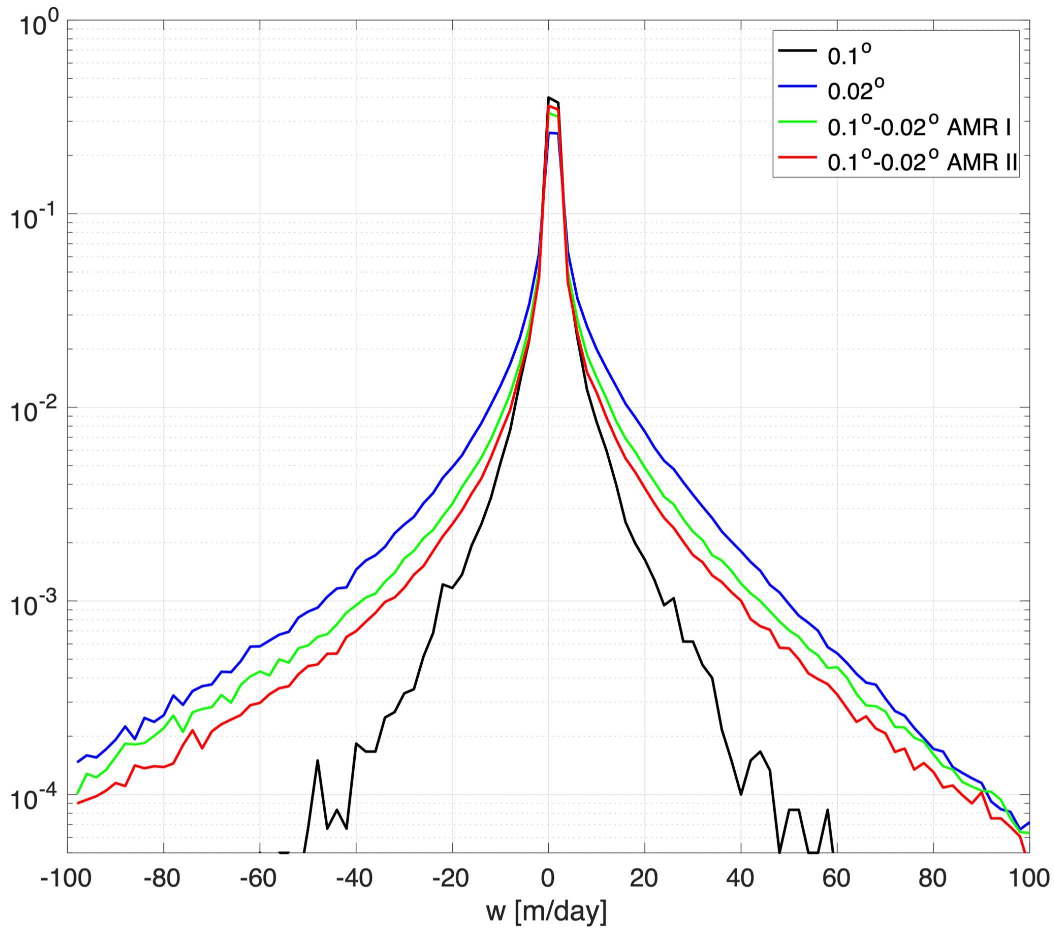

Figure 3. Distribution of vertical velocity (w) of the four experiments in Fig. 2.

Regarding the referee's suggestion on utilizing the seasonal variability to demonstrate adaptivity, we totally agree and consider it a great suggestion to our work. Actually we observe certain seasonality of the overall kinetic energy in our AMR experiments (Fig. 8) and that in the ratio of the refinement region. For 0.1-deg run, the mean KE peaks in summer (Jul-Aug), while the full-field 0.02-deg run definitely shows a different time of the KE peak during the seasonal cycle. Also evident is the seasonally enhanced submesoscale features with 0.02-deg. Therefore we argue that the Double-Gyre could serve as a typical case for demonstrating OMARE's adaptivity capabilities.

To summarize, the criteria of adaptive refinement, in our opinion, should be process dependent as well as study dependent. For example, mesoscale/submesoscale feature identification and tracking is another possible choice for our planned use of the model. <u>In this study, the criteria based on velocity and vorticity, in our opinion, fully served the purpose of demonstrating the capability of AMR and capturing the dynamically changing WBC system</u>. With the suggestions of the referee, we plan to fully explore the design of adaptive refinement for both one-way and two-way (i.e., with feedback) in the future.

Third, on the 5th-, 20-th and 50-th day results of the wintertime AMR run. We agree with the referee that the model status immediately after the refinement to 0.02-deg does not contain valid physics at 0.02-deg, since the spin-up has not fully finished. We want to emphasize that the purpose of showing results on the Feb. 5-th (5 days after refinement) is to: (1) demonstrate that the AMR criteria does capture the instantaneous WBC pattern, and (2) illustrate the spin-up process within the refined regions for all three refined experiments, and the consistency among them. Also it (Fig. 9) compares nicely with that after 20 days' refinement and that after 50 days, showing specific temporal changes in the refined regions. Therefore the comparison across the whole development process further demonstrates that the model captures the dynamically changing WBC with the AMR capability.

According to the referee's comment to make it more concise, we now shorten the result description on Feb. 5th and 20th, but keep the figure (i.e., Fig. 9). We would also like to notice that for the summer case, we only include the 50th-day results.

Please also check the text: there are numerous cases when plural/singular and some terms (e.g. kinematics) are used inappropriately.

Reply: we have gone through the text for corrections of grammatical errors, including the 3 cases of misuse of "kinematics".

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: <u>*The refactorization process in total involves about 32 man-months to finish*</u>.

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*."