To begin with, we wish to thank the critics from the Anonymous Referee #1, which we considered in revising the manuscript. For the sake of clarity, we reproduce the comments from the referee and then we present our answers. We also made several small corrections in the text. All implemented changes may be tracked down in the attached files with the revised manuscript. However, as far as we understood, these files may be submitted only after the Discussion is closed. Therefore, the main purpose here is to briefly explain how we addressed the comments from the referee.

**Reviewer comments and our answers**

**RC1:** 'Comment on gmd-2022-23', Anonymous Referee #1, 09 Mar 2022  reply

**Duarte et al.: Implementation and evaluation of open boundary conditions for sea ice in a regional coupled ocean (ROMS 3.7) and sea ice (CICE 5.1.2) modelling system**

The authors implement a new coupled configuration of the ROMS regional ocean model and CICE sea ice models for two regional modeling domains around the Barents Sea and Svalbard for operational (Barents-2.5km) and research (S4K) purposes. The authors present results showing that using time varying sea ice boundary conditions for the CICE sea ice model improve regional model results. It does appear that there are improvements to the model results with the time varying boundary conditions. However, the manuscript does not fully explain some of the model set up that would be necessary to replicate results, nor do they fully explain results and in some cases seem to make conclusions not supported by the figures. I think the authors need to expand their descriptions and discussions in the paper before the manuscript should be accepted.

**Answer:** We note that some of the general comments made above by the referee are
detailed below, under “Major concerns”. Therefore, we also detail below how we addressed the issues raised by the referee. Here we briefly state that we added more detailed explanations about the modeling set up and also about the description and discussion of our results and related figures.

**Major concerns**

- The authors do not provide some crucial details about the model assumptions:
  - Line 174: What slightly negative value did you use? -1.8C?
    - **Answer:** We used -0.00001 °C. This is now specified in the text. Please note that this is assumed at the snow or ice-atmosphere interface when air temperatures are > 0.
  - Line 175: What sort of interpolation? How did this work if there was snow?
    - **Answer:** Linear interpolation. The same temperature trend was assumed for snow and ice. Therefore, when snow was present its height was considered at the thickness of each ice layer. These clarifications were added to the revised manuscript.
  - Line 260: Did you test the mixing of initial conditions to see the impact on the solution?
    - **Answer:** The initial conditions came all from TOPAZ and we let the S4K model spin up before comparing the results with observations collected in 2015.
  - Line 401: I disagree with the statement that the lower resolution model needs the doesn’t necessarily need the same number of categories and layers, but you do need information to interpolate from the original to the new model.
    - **Answer:** Please note that this sentence refers to what we understand as “ideal conditions”. Actually, the sentence starts with “The ideal output...”. So, we are not talking about the necessary conditions where interpolations are certainly possible, as we did in this study, where boundaries were far from what we consider “ideal”.

**Impact of time varying Boundary Conditions:**

- Line 287/Figure 4: You state that the time varying boundary conditions lead to lower RMSE. However, in Fig. 4a it looks like the drop in RMSE happens BEFORE the change in boundary conditions. What is going on here? It is also impossible in Fig 4b to see the difference in M1 and M2. Can you use different colors or shapes for the markers?
  - **Answer:** Yes, the drop in RMSE indeed occurs before the change in boundary conditions. That is due to the re-initialization from TOPAZ (based on 1-month spin up) as described in line 283 of the “old” manuscript version. Moreover, please note that the average RMSE is much lower in the period between the beginning of usage of time-varying boundaries (red vertical line in Figure 4a) and the beginning of data assimilation (green vertical line in Figure 4a) than in the previous period, despite the drop in RMSE mentioned above and related to model re-initialization. We added a few comments to the last paragraph of 3.1.2 to clarify these issues. We also changed Fig. 4b and now the two red circles are visible despite their great proximity.

- Line 393: Is it possible that some errors are caused by the discontinuity at the boundary? Many models - e.g. WRF for the atmosphere (https://doi.org/10.1175/1520-0477(1997)078%3C2599:ATOLBC%3E2.0.CO%3B2) or RASM in the ocean (https://doi.org/10.3189/2015AoG69A760) – have some sort of buffer zone to help with spurious boundary issues. Did you investigate this at all? Is it worth investigating?
  - **Answer:** Thanks for pointing this out. We have recently investigated this problem. It is true that such spurious boundary issues are mainly due to the imperfect ocean open boundary conditions, particularly when tide is activated (as the TOPAZ4 does not include tide). Such imperfect ocean boundary conditions will result in artificial boundary currents, transfer erroneous energy along the open boundary, thus affecting the sea ice there. We have tested some cases by changing the sponge zone settings, which mitigate the effect of this error. This work is underway and will be reported later in a
following study. For the sea ice boundary conditions applied here, our experience with a higher resolution model (500 m horizontal resolution) implemented with CICE, nested in the Barents-2.5km model, and using exactly the same sea ice data of the larger model, showed a near perfect transition between both domains. This suggests that model will perform robustly when the ocean boundary conditions are correctly treated, and the sea ice boundary conditions match the exact needs of the model. This was already referred in the manuscript following line 393 (in the old version) and we kept it in the revised version.

- Line 320-322: Your description about the biases is missing a lot of the detail from Figure 5 and in some cases contradicts it. For example in Figure 5 shows that S4k biases are largest from the surface to 100m whereas your description states the opposite. You also need to explain how the agreement shown in Figure 7 is better – larger range, warmer temperatures, fresher waters?

  **Answer:** The temperature biases between the surface and 100 m is much smaller for S4K than for TOPAZ. The overall salinity biases within this depth range is smaller for TOPAZ but at the surface is smaller for S4K (Figures 5 and 6). The agreement in Figure 7 is better for S4K because both the salinity and temperature ranges compare well with those of the observations (upper panel versus lower panel). In the case of TOPAZ both ranges are much narrower than those of the observations (upper panel versus middle panel). We added text to 3.2.1 to detail better these results.

- Figures 8, 9, 10: Can you zoom in on the relevant domain rather than show pan-Arctic figures. It’s hard to see anything meaningful in the full Arctic scale when the important details are at the transition/boundary of the regional model. Also, making difference figures might help as well to highlight the differences you describe because at present the differences are challenging to see. Adding arrows or other markers may help too.

  **Answer:** Figures were zoomed as suggested.

- Figure 9: the boundary looks smoother, but why is there so much less ice in S4K (9d) than Topaz (9c)? This seems important and worth explaining. It seems the S4K model is performing worse with so little ice.

  **Answer:** Please note that Figure 9 corresponds to 2014 results when the model is spinning-up. We emphasize this in the revised manuscript in section 3.2.2. However, comparisons with measurements done in 2015 in Figures 11 show a better correspondence between S4K and observations as pointed out in the paper.

- Figure 11: c/d. Why do the frequencies for HEM/EM31/Magna vary between the columns? Are they on the same domain, because if so they should be identical frequencies. Is it possible to put all three PDFs on one figure for the same domain?

  **Answer:** This is a result of the different resolution of TOPAZ and S4K. Observational data were averaged for TOPAZ4 or S4K model cells located in the same areas, resulting in slightly different observed frequency distributions, given the different spatial resolution of the models (12.5 and 4 km, respectively). This explanation was added to the revised manuscript in Figure 11 caption.

**Minor concerns**

- Figures 2 and 3: Please provide more descriptive, standalone captions.

  **Answer:** Done as suggested.

- Introduction: You should probably address some of the points from this paper on whether sea ice models from climate models should be used for forecasting (https://doi.org/1007/s40641-020-00162-y). Also at line 386-388 you mention relevant operational reasons this model improvement has been useful. It seems this should be expanded on more in the introduction.

  **Answer:** We added a paragraph to the end of the Introduction citing the paper mentioned by the referee and addressing some of its points, as suggested. We also expanded a bit about the importance of the improvement in the operational forecasts in the first paragraph of the Discussion.

- Line 35: You should provide a DOI for whatever version of CICE you are using. See the
Answer: This release table does not seem to include CICE 5.1.2, but we cite the manual of CICE5.1 (Hunke, E. C., Lipscomb, W. H., Turner, A. K., Jeffery, N., Elliot, S.: CICE: the Los Alamos Sea Ice Model. Documentation 474 and User's Manual Version 5.1. Los Alamos National Laboratory, USA. LA-CC-06-012, 2015.). Moreover, we include the DOIs to the exact versions we used in the paper, that include our own modifications. In fact, we appreciate this comment very much because we realized we made a mistake here by forgetting to specify that the Barents 2.5 km model is indeed based on CICE5.1.2 but the S4K model is based on a “columnar” version of CICE that the first author received from the CICE consortium in 2017. Thus, it is not possible to identify it in the table cited above. Because of this, we removed 5.1.2 from the title since more than one version of CICE is used in this study. The alternative is to specify both versions in the title but since the columnar does not correspond to any of those listed in the link provided by the referee, we can refer to it only as “columnar” which is a bit vague. Most likely it was a developing version between releases. However, this makes no big difference with regard to the implementation of time-varying boundary conditions since, as explained in the last paragraph of the Discussion, the CICE files changed are common across different CICE versions including CICE + Icepack. We added some comments to 2.1.1 and 2.1.2 to clarify that different CICE versions were used in the Barents 2.5km and the S4K models. We also added some text to the Acknowledgements regarding the CICE columnar version used for the latter model and following the agreement with the CICE consortium. By the way, we also checked (https://github.com/CICE-Consortium/CICE-svn-trunk/wiki/CICE-Versions-Index-(older)) for CICE versions released prior to CICE6 but we did not find an exact match to the versions used herein.

- Line 37: You may want to clarify that CICE6 and later have Icepack separate from the dynamics but CICE5 and any earlier versions they were combined. Use the specific version numbers here.
  
  Answer: Done as suggested. Please see also our answer to your previous comment and the 1st paragraph of the Introduction.

- Line 40: You should provide more information about other modeling systems that use CICE, even older versions, beyond the CMIP6. E.g. RASM (https://doi.org/10.3189/2015AoG69A760), CESM (https://doi.org/1029/2019MS001916), Canadian Operational Center (https://doi.org/10.5194/gmd-2020-255), Danish Meteorological Institute (https://doi.org/10.1002/2017JC013481), etc.
  
  Answer: Done as suggested. More citations and corresponding references added to provide examples of modeling systems that use CICE.

- Line 68: It would be helpful to have 1 sentence about TOPAZ4 and why it’s optimal for this set up since not all readers will be familiar with it.
  
  Answer: Done as suggested. A sentence added about TOPAZ.

- Line 71: AMSR2 link doesn't work. Also maybe mentioning here that in section 2.4.1 you will describe how the satellite product at 6km is downscaled.
  
  Answer: We have just tested the link and it worked. We wonder what the problem might be. We added a small sentence following the suggestion made by the referee about downscaling details.

- Line 87: Is this run continuously or with restarts (as reanalyses are usually run)?
  
  Answer: Please note that this is not a reanalysis product. It is run without restarts unless something wrong happens, in which case restart files are used for a “perfect” restart. For the sake of clarity, we added “run continuously” to the sentence. Please refer also the Introduction where we characterize both modeling systems and explain that while Barents 2.5km is an operational system, the S4K model is a research tool.

- Line 133-135: Do you have an estimate about the extra computational cost for using more memory? 10% or 90% increase?
  
  Answer: We do not have such estimate, but it should be marginal (< 10%) as compared to other files (especially ROMS boundary files, which are huge due to the
many time slots) because the ice forcing routines read only two time slices of the boundary conditions for proper time interpolation. Therefore, we are talking about a few arrays (as many as the boundary variables) with dimensions equal to those of the grid X 2 time slots and, for some variables, X the number of ice layers.

- Line 161: there are two periods at the end of the sentence.
  - Answer: Corrected.
- Line 166: use “internal” instead of “inner”.
  - Answer: Done as suggested.
- Line 294: Remove “likely” as you have not yet shown this to be the case. The sentences below do show this, but at this point it seems unproven.
  - Answer: There is no “likely” in line 294.
- Line 406: There is no figure 10e.
  - Answer: We deleted this wrong reference to a figure.