Reply to Reviewer #1

We thank reviewer #1 for the positive and insightful comments. We have followed each of the comments, presented in black font below, to carefully revise our manuscript. Our point-by-point responses are given in blue.

General Comment:

This is a very well written manuscript about an improved and high resolution configuration of a coupled Ocean-Sea ice-Ice shelf model over the southern pole. The authors clearly explain the configuration of the model, and thoroughly evaluate several physical aspects against observational datasets.

We thank reviewer #1 for the recognition of our work!

Specific Comments:

1. Sentence at Lines 183-185. An appropriate reference supporting this statement would be helpful.

We thank the reviewer for this helpful suggestion.

Following the comment of the reviewer, we have now cited two relevant studies (Mack et al., 2017; Mack et al., 2019) that specifically discuss the need for topographic smoothing in the Regional Ocean Modeling System (ROMS) to mitigate pressure gradient errors. The revised sentence now includes these references (see lines 185 and 1633-1638 in the revised manuscript).

Mack, S. L., Dinniman, M. S., Klinck, J. M., McGillicuddy Jr., D. J., and Padman, L.: Modeling Ocean Eddies on Antarctica's Cold Water Continental Shelves and Their Effects on Ice Shelf Basal Melting, J. Geophys. Res. Oceans, 124, 5067-5084, https://doi.org/10.1029/2018jc014688, 2019.

Mack, S. L., Dinniman, M. S., McGillicuddy, D. J., Sedwick, P. N., and Klinck, J. M.: Dissolved iron transport pathways in the Ross Sea: Influence of tides and horizontal resolution in a regional ocean model, J. Mar. Syst., 166, 73-86, https://doi.org/10.1016/j.jmarsys.2016.10.008, 2017.

2. In Lines 262-264, the authors indicate that deeper than 5500m, they obtain the initial conditions by downward extrapolation, but the exact method is not described. Please clarify how this extrapolation is performed and what the effects of this choice are in the circulation of the model.

We thank the reviewer for raising this important point.

As the reviewer noted, the initial conditions for potential temperature (θ) and salinity (S) at depths exceeding 5500 m were obtained by a simple downward extrapolation. Specifically, for all wet grid cells located below 5500 m depth, the initial θ and S were simply set to the values of the deepest valid level from the WOA18 climatology for that same water column.

We acknowledge that this is a simplification. Yet, we think that it is still a reasonable and practical approach. First, as we know, the abyssal ocean is generally at rest, and the stratification in these layers is very weak. The initial conditions are not expected to strongly constrain the dynamics, as the model will quickly adjust based on the prescribed boundary conditions and forcing. Second, the number of bins with depths exceeding 5500 m is relatively small, confined to the deepest parts of the ocean basins. Third, observational data for such deep layers in the Southern Ocean are extremely sparse. The World Ocean Atlas 2018 (WOA18) itself does not provide climatological means below 5500 m, making a more sophisticated extrapolation scheme both difficult to validate and unlikely to yield a more accurate initial state.

Therefore, we think that this simple extrapolation does not introduce significant biases into the simulated large-scale circulation. We have now added the description of downward extrapolation as follows (see lines 265-266 in the revised manuscript):

'*Since the maximum depth of the climatological mean fields of WOA18 is 5500 m, the deeper layers (> 5500 m) of the initial conditions of θ and S in SOSIM are obtained by extrapolating downward, with all wet grid cells located below 5500 m depth set to the values of the deepest valid level from the WOA18 climatology for that same water* *column.*'

3. For Figure 5 and paragraph starting at Line 392. The map shows the three sectors, but the exact longitudes are not defined. Please specify the longitude ranges for each sector.

We thank the reviewer for this helpful suggestion, and we apologize for the omission in the original manuscript.

Following the suggestion of the reviewer, we have now specified the exact longitude ranges for each sector in both the main text and the caption of Fig. 5 as follows (see lines 396-397 and 406-408 in the revised manuscript):

'*To quantitatively evaluate the model results in different basins, we introduced a basin-scale regional division in the deep Southern Ocean, including the Southern Atlantic sector (67 °W-20 °E), the Southern Pacific sector (146 °E-67 °W), and the Southern Indian sector (20 °E-146 °E), respectively (Fig. 5a).*'

'*Figure 5. The regional division for model evaluation. (a) The basin-scale regional division and the shelf sea around Antarctica. The red semi-transparent region denotes the Southern Atlantic sector (67 °W-20 °E), the yellow semi-transparent region denotes the Southern Indian sector (20 °E-146 °E), and the cyan semi-transparent region denotes the Southern Pacific sector (146 °E-67 °W). (b) The alternatively red and yellow regions denote the division of the Antarctic shelf seas (the south of the 700 m isobath) by following Stewart et al. (2018). The blue lines denote three transects across the continental slope, as shown in Figs. 19-21.*'

4. Paragraph starting at Line 849. A possible explanation for the weak positive correlation with AVISO observations for surface MKE might be the large discrepancy between the resolution of SOSIM and AVISO, a byproduct of the interpolation from the low resolution of AVISO to match the higher resolution of SOSIM. One would expect that the features would be broader in AVISO than in SOSIM, leading to areas with higher values in AVISO than SOSIM (the area below the 1:1 line in Figure 16c). And the intrinsically higher resolution of SOSIM will lead to larger values of MKE in areas with low and moderate values of MKE in AVISO (area above the 1:1 line in the same

figure).

We thank the reviewer for this insightful suggestion!

We agree that the discrepancy in horizontal resolution between SOSIM (~5 km) and the AVISO product (0.25 °) may be a key factor contributing to the statistical differences shown in Fig. 16c. The interpolation from a coarser grid smooths spatial gradients, leading to the AVISO underestimating the MKE in high-variability regions and overestimating it in low-variability regions, resulting in the pattern of deviations from the 1:1 line.

Following the suggestion of the reviewer, we have now incorporated this reasoning into the manuscript to better interpret the SOSIM-AVISO comparison as follows (see lines 860-864 in the revised manuscript):

'*The simulated surface MKE shows a weak positive correlation with observations from AVISO (r = 0.40) and an RMSE of 1.5 × 10^{-2} m^2 s^{-2} (Fig. 16c). There is a significant downward offset of the linear regression line from the 1:1 line. This deviation may be attributable to the mismatch in horizontal resolution between SOSIM (~5 km) and the original AVISO product (0.25°). The coarser resolution of AVISO leads to broader representations of spatial features, such as frontal jets, which can result in higher MKE values than SOSIM in some regions. Conversely, the higher resolution of SOSIM allows it to resolve finer-scale flow structures that are smoothed out in AVISO, contributing to higher MKE values in such regions. However, since the PDFs for both SOSIM and observations peak at ~1 × 10^{-4} m^2 s^{-2} and are concentrated near this value, the overall dynamic range of SOSIM still appears constrained within the range of observations.*'

5. Section 3.7. Sea Ice Production (SIP) (for March-October) in the Atlantic sector (at 45W and between 60-65S) is smaller than observation (HU) (around -1 m/yr, Figure 25b). At the same time, the Sea Ice Thickness (SIT) is considerably larger in the same area in September (+0.2-0.4 m, Figure 24e) (albeit a bit smaller in February at between 0 and -0.2 m). How are these things reconciled?

We thank the reviewer for raising this important point.

The reviewer is correct in noting that the simulated Sea Ice Thickness (SIT) in this region is higher than observations, while the local Sea Ice Production (SIP) is lower. This can be reconciled by considering the distinction between thermodynamic and dynamic processes in sea ice modeling. SIT is not only determined by local thermodynamic growth (i.e., SIP) but also strongly influenced by sea ice dynamics, particularly the advection and convergence/divergence. An area with low local SIP can still develop thick ice if it acts as a convergence zone where ice produced elsewhere is transported into the region and accumulates. Conversely, a region of high local SIP may not exhibit thick ice if the newly formed ice is rapidly exported by winds or currents.

6. Paragraph starting at Line 1245 and Figure 26d. How exactly were the statistical metrics (e.g., correlation and RMSE) calculated in this case of the area-integrated ISMRF within each sector over the continental shelf?

We thank the reviewer for raising this important point.

The statistical metrics (r, p, RMSE, and mean bias) for the area-integrated Ice Shelf Melt/Freezing Rate (ISMFR) were calculated as follows: we treated each of the eight shelf sectors shown in Figure 5b as an individual data point. For each sector, we computed the total ISMFR (in Gt yr^{-1}) by integrating the basal melting/freezing rate over all ice shelf grid cells within that sector. This yielded two sets of eight values: one from SOSIM and one from the A2020 observational product. The correlation, RMSE, and mean bias were then calculated on these paired eight-point series.

We have to admit that, with such a small sample size (n=8), the p-value is indeed expected to be relatively high, and the correlation should be interpreted with caution. Therefore, we have explicitly noted the p-value (0.30) to highlight this statistical limitation (see line 1260 in the revised manuscript).

7. Paragraphs starting at Line 1267 and Line 812: Regarding the significant underestimation of ACC transport. Some possible explanations are given in the manuscript. Are there additional diagnostics that can be used to narrow down even more the possible causes?

We thank the reviewer for raising this important point.

We agree that the underestimation of ACC transport in SOSIM is a notable bias, and we appreciate the opportunity to discuss its possible causes in more depth. At the current stage, we hypothesize that this bias may stem from the proximity of the northern open boundary to the ACC core, which could artificially constrain the full development of the current within the model domain. Additionally, the prescribed open boundary conditions from ECCO2 and WOA18 may not fully capture the dynamical complexity of the ACC at the northern edge.

To further narrow down the causes, we think that additional sensitivity experiments would be valuable. These could include: (1) testing different northern open boundary locations, (2) exploring alternative advection schemes, (3) varying horizontal resolution, and (4) comparing different northern open boundary forcing datasets. Yet, such a systematic investigation would require a dedicated set of experiments that are beyond the scope of the present study, which focuses on the overall configuration and evaluation of SOSIM, particularly the subpolar gyres and shelf regions south of the ACC. Therefore, while the ACC bias is acknowledged, we hope we can still focus on the region south of the ACC. Nevertheless, we have added more descriptions to highlight that future development could include sensitivity experiments to better understand and potentially reduce this bias as follows (see lines 1289-1291 in the revised manuscript):

‘*To further understand the underestimated ACC transport, future sensitivity experiments (e.g., testing different northern open boundary configurations, horizontal resolutions, or advection schemes) would be valuable for diagnosing and potentially mitigating this bias.*’

We thank the reviewer for prompting this important point!

Technical Comments:

1. Lines 614-615: This sentence is somewhat difficult to follow due to the “and while”. Please consider revising the sentence to improve the clarity.

We thank the reviewer for pointing out the lack of clarity in this sentence.

We agree that the original phrasing 'and while' was redundant. Following the suggestion, we have revised the sentence for improved readability as follows (see lines 624-625 in the revised manuscript):

'*In the Southern Indian sector, strong warm biases dominate the upper mixed layer at lower latitudes, while weak cold anomalies are present below the thermocline (Fig. 9h)*.'

Reply to Reviewer #2

We thank reviewer #2 for the positive and insightful comments. We have followed each of the comments, presented in black font below, to carefully revise our manuscript. Our point-by-point responses are given in blue.

This manuscript presents a valuable new circum-Antarctic coupled ocean–sea ice–ice shelf configuration based on MITgcm and documents it through an extensive evaluation against a wide range of observational and reanalysis products. The configuration itself is a substantial technical achievement, combining relatively fine horizontal resolution, explicit ice-shelf cavities, and a long hindcast forced by ERA5 over 1979–2022. I also appreciate the authors' transparency in discussing remaining biases and limitations. Overall, I see this work as a strong contribution that will likely be of broad interest to the Southern Ocean and Antarctic modeling communities.

In my view, the manuscript is already in good shape scientifically, but it would benefit from a to help clarify the model's novelty, better contextualize some of the remaining biases, and strengthen the discussion of where the configuration is especially robust versus where caution is still warranted.

I want to emphasize that one of the best features of the manuscript is its transparency. The authors do not oversell the model and openly discuss important remaining issues such as the warm abyssal drift, sea-ice seasonal bias, and underestimated total ice-shelf melt. That gives the paper credibility.

We thank reviewer #2 for the recognition of our work!

We fully agree with the reviewer that we should strengthen the discussion.

Rather than changing the underlying results, I think the manuscript would benefit mainly from reorganizing some of this discussion so that readers can more easily distinguish:

1) where the model performs particularly well,

2) where biases are moderate but acceptable for many applications,

3) where caution is needed.

A summary figure or table listing strengths, limitations, likely causes, and implications for model use would be extremely helpful.

We thank the reviewer for this helpful suggestion.

Following the advice of the reviewer, we have added a table (Table R1) to categorize the strengths, biases, likely causes, and implications of SOSIM across its main components, including the open ocean, continental shelf, continental slope, sea ice, and ice shelves as follows (see lines 1381-1386 in the revised manuscript):

Table R1. A summary of the performance of SOSIM.

Component/ Process	Key strengths	Main biases	Likely causes	Implications for usage
Open oceans	1) Large-scale θ and S structures 2) Frontal positions 3) Spatial patterns of surface MKE and EKE	1) Underestimated ACC transport 2) Warm and fresh drift in abyssal waters 3) Weakened AABW	1) Northern boundary proximity to ACC 2) Insufficient dense water formation/mixing 3) Open boundary conditions	**Robust:** subpolar gyre structure; frontal dynamics. **Caution:** ACC transport studies; abyssal trend analyses.
Continental shelf	1) Bottom water mass properties 2) Meridionally averaged θ and S structures 3) large-scale CDW/MCDW intrusion patterns	Regional biases in CDW intrusion	1) Biases in open-ocean source waters 2) Overestimated ASC strength	**Robust:** regional process studies; water mass transformation. **Cautious:** quantitative heat transport in specific sectors.
Continental slope	1) Three typical ASC/ASF	1) Overestimated ASC strength	1) Local wind biases 2) No tides	**Robust:** Three typical

	regimes 2) Good θ and S correlations with observations	2) Absent Dense SW overflow		ASC dynamics. **Caution:** quantitative along-slope velocities; AABW production
Sea ice	1) Winter SIC and SIE distribution 2) Interannual variability 3) SIP spatial pattern	1) Exaggerated seasonal cycle 2) Overestimated coastal SIP 3) Regional SIT biases	1) Zero-layer thermodynamics 2) Overestimated air-ice stress 3) Neglect of ice tongue barriers	**Robust:** winter sea ice state; interannual variability. **Caution:** summer sea ice projections; quantitative SIP in polynyas.
Ice shelves	Spatial pattern of high melting rates in West Antarctica	1) Underestimated total mass loss 2) Absence of basal freezing in large ice shelves	1) No tides 2) Coarse vertical resolution near ice–ocean interface 3) No frazil ice	**Robust:** regional patterns of basal melting. **Caution:** quantitative basal mass balance of ice shelves.

‘*To provide a concise summary of the performance of SOSIM, we synthesize our evaluation in Table 1. By outlining the primary strengths, major biases, the likely causes of these biases, and the practical implications for future applications, this summary is intended to clearly show the regimes and processes that are robust in SOSIM, as well as those where caution is warranted or further development is still necessary.*’

Moreover, the use of repeated 1979 forcing for spin-up is a reasonable practical choice, especially given the desire to begin the hindcast consistently in 1979. At the same time, the manuscript shows that some domain-integrated drifts remain, especially in deeper waters and in the early evolution of kinetic energy. I do not see this as undermining the value of the model, but I do think the paper would benefit from a slightly fuller discussion of what this means for interpretation. In particular, it would help readers if the authors could state more explicitly that:

1) the deep ocean remains less equilibrated than the shelf regions,

2) some large-scale integral diagnostics should be interpreted cautiously,

3) shelf-focused applications may be more robust than basin-scale abyssal trend analyses.

That would strengthen the manuscript considerably.

We thank the reviewer for this constructive comment. Following the suggestion of the reviewer, we have added two paragraphs to state these points more explicitly.

First, at the end of subsection 3.2 (Model spin-up and drift), we state that the deep ocean remains less equilibrated than the shelf and slope regions, and that large-scale integral diagnostics should be interpreted cautiously as follows (see lines 558-562 in the revised manuscript):

'*Overall, the model equilibration varies substantially across different regions of the model domain (Figs. 6-7). The continental shelf and slope regions reach a quasi-equilibrium state much more rapidly than the deep ocean, with no statistically significant trends over the last two decades of the simulation. In contrast, there are persistent, albeit slow, drifts in the deep ocean throughout the integration period. Therefore, large-scale integral diagnostics should be interpreted with caution, particularly for the abyssal circulation or long-term climate variability.*'

Second, in subsection 4.1 (The simulated oceanic biases), we note that shelf-focused applications are likely to be more robust than basin-scale abyssal trend analyses as follows (see lines 1298-1299 in the revised manuscript):

‘*Given the persistent drifts in the abyssal ocean (Fig. 7), shelf-focused applications are likely to be more robust than basin-scale abyssal trend analyses.*’

A third point is the topography workflow. The geometry preparation appears to have required careful expert judgment, especially in handling remapping artifacts and problematic isolated cells. That is completely understandable in a configuration of this complexity. Still, because this paper will likely serve as a reference for future users, it would be valuable to document these steps a bit more fully, perhaps in the supplement. A brief workflow description, a mask of edited cells, or a short note on the scale of the edits would likely be enough. This would make an already useful paper even more reusable by others.

We thank the reviewer for this constructive comment.

Following the suggestion of the reviewer, we have added a brief description of the topography construction workflow in the Supplement as follows (see lines 66-91 in the revised Supplement):

‘*Text S1. Topography construction workflow*

The topography datasets for SOSIM were derived from the Refined Topography dataset version 2 (RTopo-2), which provides consistent topography of bedrock, elevation, and ice draft on a spherical grid. The workflow for constructing the model topography is as follows.

1) The RTopo-2 data south of 35°S (43201 × 6601 cells, 30 arcsec resolution) is interpolated onto the SOSIM orthogonal curvilinear grid (1800 × 1800 cells, ~5 km resolution) using a nearest neighbour scheme. This approach preserves the original values without smoothing but can introduce artifacts in regions with complex coastlines or steep topographic gradients.

2) After remapping, a thorough inspection of the resulting topography fields is performed to identify problematic isolated wet cells that were disconnected from the main ocean domain (e.g., enclosed by dry land or ice shelves with no flow pathway). Since the model applies additional constraints during initialization, such as minimum *water column thickness requirements and partial-cell adjustments, the model modifies the actual distribution of wet cells relative to the raw input topography. Therefore, the identification of isolated cells should be based on the mask file generated by the model (e.g., the 'hFac' file) rather than on the theoretical water column thickness (ice draft minus bedrock depth). The mask file can be generated by running the model for one time step without additional components (e.g., sea ice, atmospheric forcing, and open boundary forcing).*

3) Problematic cells are corrected through a manual editing process. The edits of the bedrock depth and ice draft mainly focus on removing isolated wet cells by converting them to dry cells. The manual edits are concentrated in regions with complex coastal geometry, particularly around the Antarctic Peninsula, the sub-Antarctic islands (e.g., South Shetland Islands and South Orkney Islands), and the margins of large ice shelves (e.g., the Ross Ice Shelf and Filchner-Ronne Ice Shelf). The revised topography is then used to generate an updated mask file by running the model for a single time step.

4) The updated mask file is visually inspected again to ensure that the manual edits have not introduced unrealistic features or new isolated wet cells. If problematic cells are identified again, further modifications to the bedrock depth or ice draft are necessary, and the process of model re-runs and re-inspection is repeated. This iterative loop continues until all problematic cells are resolved in the final generated mask file. In total, ~0.15% of grid cells are manually adjusted in SOSIM v1.0.'

We have also added a reference to this supplement description in the revised manuscript (lines 257-258).

A final comment concerns the underestimated ice-shelf melt. The manuscript's treatment of ice-shelf melt is appropriately cautious, and the underestimation relative to observational products is clearly shown, with the omission of tides identified as one likely explanation. I would encourage the authors to frame this more explicitly as a priority avenue for future model development. This would leave readers with a clearer sense not only of the current model performance, but also of the most promising next steps.

We thank the reviewer for this insightful suggestion.

We have added the description to more explicitly identify the inclusion of tidal forcing as a primary and priority direction for future development as follows (see lines 1410-1411 in the revised manuscript):

'*We therefore highlight the inclusion of tidal forcing as a priority avenue for future development of SOSIM.*'

Overall, I consider this manuscript to be a strong and worthwhile contribution. The configuration is impressive, the evaluation is extensive, and the manuscript is strengthened by the authors' transparent discussion of remaining biases. The revisions I suggest are primarily aimed at sharpening the interpretation and improving the utility of the paper as a reference for the wider community.

We fully agree with the reviewer that these revisions can substantially strengthen the manuscript. We thank the reviewer again for these insightful and supportive comments, which have helped us improve the clarity and impact of our work!