Response to Reviewer 1:
We would like to thank the reviewer for the thoughtful comments and constructive feedback. Below, reviewer comments (RC) are followed by our responses (AC).

RC: In general the manuscript is nicely written and describes the nudging, however I could wish for a consideration of which data, experiments etc. should be in the manuscript. For instance in figure 4 and 5 is it necessary to show all different experiments with different relaxation? Not all are in focus of the manuscript. I think that it would make the points more clear.
AC: The purpose of presenting all relaxation experiments was to demonstrate how the system responds to varying relaxation time scales and, consequently, to verify the correct implementation of the relaxation algorithm in SIS2. A physically and numerically well-behaved relaxation scheme should exhibit a smooth and monotonic dependence of the solution on the relaxation time scale. As the relaxation becomes weaker (i.e., as the relaxation time scale, τ, increases), the solution should gradually approach the free-running model state without abrupt transitions or unexpected changes in sea ice conditions within the relaxation region. We believed that showing the sea ice thickness and concentration fields for all relaxation experiments helped illustrate this behavior.
Nevertheless, we agree with the reviewer's suggestion and have removed two intermediate relaxation experiments from Figures 4, 5, 9, and 10 to reduce redundancy. Results from all relaxation experiments are still retained in the remaining figures, where they provide a more compact quantitative assessment of model sensitivity to the relaxation strength. We have also revised the text to better describe the behavior of the omitted experiments, which was insufficiently discussed in the previous version.

RC: Redistribution. I am unsure why the model becomes unstable if all sea ice is put in the target category. Yes if the concentrations becomes very low as mentioned.
AC: SIS2 is technically able to start from a single-category ice thickness distribution (ITD), in which all ice area and volume are assigned to the primary thickness category, followed by redistribution within the model. In our tests, this approach generally performs adequately and does not lead to issues in most cases. However, we have occasionally encountered instances where a single-category initialization led to numerical failure during the early-stage redistribution step. We have not diagnosed the precise cause of these rare instabilities, but they appear to occur during the initial adjustment of the ITD when the model attempts to redistribute ice area and volume from a highly singular distribution. By contrast, initializing the ice state with a non-degenerate, category-resolved ITD (distributing ice area and thickness across multiple categories with small but non-zero values) has not exhibited such behavior in our experiments. For this reason, a category-resolved initialization was adopted as a practical robustness measure. The manuscript text has been revised to avoid implying that a single-category initialization is inherently unstable and instead reflects the observed behavior in our experiments.

RC: Secondly (step iv). How can one conserve volume and area when moving ice from the target category to thinner categories? This means that the ice thickness within each thickness category is reduced

AC: The goal of the redistribution algorithm is to conserve the aggregated ice partial area (eq. (7)) and the aggregated ice volume per unit grid-cell area (eq. (8)), i.e., to preserve the target ice concentration and mean ice thickness provided by the observational or reanalysis data (PIOMAS in this study). The redistribution is performed only within the ice thickness distribution and therefore does not alter these bulk quantities.

It is correct that the ice volume associated with individual thickness categories changes during the redistribution. When ice area and volume are transferred from the primary target category to neighboring categories, the category-wise ice concentrations and mean thicknesses are modified. However, the redistribution is constructed such that the sum of ice area over all categories and the sum of ice volume over all categories remain unchanged. Thus, conservation applies to the aggregated ice concentration and volume, not to the ice volume within each individual category. The manuscript text has been clarified.

RC: The introduction describes mismatch between ocean, sea ice and atmosphere and yet the authors have chosen to use two different products for the ocean and the sea ice. With the choice of the Glory's product for the ocean, ERA5 for the atmosphere and PIOMAS for the sea ice it seems as if the experiments have been unnecessary complicated as Glory's is forced by ERA5, and it contains a sea ice model which could have been used as the reference data instead of PIOMAS. Maybe the authors can elaborate on this? Rerunning the experiments seems to be beyond the scope.

AC: We agree that a more internally consistent experimental design could have been achieved by using sea ice fields from GLORYS together with ERA5 atmospheric forcing. The discussion in the Introduction primarily concerns challenges associated with lateral boundary conditions in regional sea ice models, including potential inconsistencies between prescribed boundary conditions and the evolving interior ice state. Similar considerations also apply to relaxation experiments, where inconsistencies between the target sea ice state and the atmospheric or oceanic forcing may affect the model response.

In the experiments presented here, the atmospheric forcing was prescribed independently of the target sea ice fields. The NEP10k experiments were forced with ERA5 atmospheric fields, while the ARC10k experiments used JRA-55 forcing. The target sea ice state was obtained from PIOMAS, which is forced by NCEP/NCAR Reanalysis 1. For the NEP10k hindcast experiments, GLORYS ocean fields were used, whereas the ARC10k experiments did not employ GLORYS forcing.

The choice of PIOMAS was motivated primarily by its extensive validation and widespread use in studies of Arctic sea ice thickness and volume. In addition, these experiments were conducted as part of a broader system-development effort, and the experimental design was intended to remain consistent with existing hindcast and retrospective forecast configurations, particularly

with respect to atmospheric forcing. This limited the flexibility to redesign all components of the experimental setup.
We agree that using GLORYS sea ice fields would provide a more consistent combination of oceanic, atmospheric, and sea ice forcing for the NEP10k experiments. However, the primary objective of this study is to describe and evaluate the implementation of the sea ice relaxation algorithm in SIS2 rather than to assess the sensitivity of the results to a particular choice of forcing or reference product. The application of the relaxation procedure with independently derived target fields suggests that the method is sufficiently robust to accommodate realistic inconsistencies among data sources. Future hindcast applications of the NEP10k system will use GLORYS sea ice fields to ensure greater consistency between the prescribed oceanic and sea ice states.
Clarification on the choice of forcing fields in the experiments has been added to section 5 (Numerical simulations) and Section 9 (Discussion).

RC: Another topic that could be elaborated a bit more on is the limitations of using observations or forcing that do not contain all the fields that are required to constrain e.g the dynamics as this is often the case when applying nudging or other methods to assimilate/add lateral boundary conditions into the model.
AC: A discussion of the limitations associated with constraining only a subset of the sea ice state variables has been added to Section 9 (Discussion).

RC: Line 40: Duarte et al implemented this in CICE V5. Strictly speaking it is not available in the main branch of CICE6 however a nesting module should soon be included in this. Different implementations do exist but they are outside the main branch.
AC: Thank you for noting this. Corrected.

RC: Line 44 (and the section around) describes the challenges of nudging/boundaries if all parameters are properly described. I think that it is worthwhile to mention that often it is not possible to get for instance stresses for the nudging unless the large-scale model used for the forcing of the regional model originates from the same group (see also comment above).
AC: A sentence has been added to this paragraph noting that only a limited set of sea ice variables is typically available for prescribing boundary conditions, which can hinder the specification of a fully dynamically consistent sea ice state.

RC: Line 108: either list the barotropic and baroclinic timestep or skip the sentence with a baroclinic….
AC: A reference to Table 1 has been added, where the time-step values and definitions are listed.

RC: Line 257: How is salinity and enthalpy changed? Often these parameters are not known within the target values. If they are kept constant they should only be "unrealistic"if ice was not present before? A comment on this would be good.
AC: This is described in section 4.2.3. A clarification describing the "no-ice to ice" case has been added.

RC: Line 420: I would use a different word than error. In general most ice thickness products are associated with a relatively large bias/uncertainty. I would use bias instead.
AC: Here and in other similar places, the term has been corrected to "differences" for consistency with the figure captions.

RC: Line 485: The high impact is likely due to the slow ice thickness evolution. There should be a comment on why
AC: This is discussed in Discussion section.

RC: Line 535: Can you point to a reason. The ocean relaxation?
AC: The unconstrained ARC10k simulation is compared against PIOMAS in this section. The ice edge position in this experiment is determined by the response of the sea ice model to the simulated ocean state (MOM6) and the prescribed atmospheric forcing (JRA-55). The causes of the ice concentration bias along the East Greenland Current (EGC) and in the Barents Sea MIZ were not investigated in this study. Both regions are strongly influenced by ocean heat transport and air-sea interactions. The positive ice concentration bias may therefore suggest that the simulated upper ocean is too cold, either because oceanic heat transport into the ice-edge region is insufficient or because ocean heat loss to the atmosphere is too strong. In the EGC region, sea ice conditions are additionally influenced by ocean dynamical processes, including the Greenland Gyre, the EGC itself, and frontal eddies. The positive sea ice bias in this region may therefore also indicate excessive lateral export of Polar Water or sea ice across the EGC front in ARC10k into the Greenland Sea. A detailed investigation of these biases is beyond the scope of the present study. To maintain focus on the implementation and evaluation of the sea ice relaxation algorithm in SIS2, this speculative discussion has not been added to the manuscript.

RC: Line 560: It is likely that both the atmosphere and the Glorys nudging leads to similar cold biases as Glorys is forced with ERA5. The assimilated sea ice product within ERA5 and GLORYS are the same (OSISAF) whereas PIOMAS uses NSIDC (if I remember correct).
AC: We would like to clarify that the simulations discussed here are forced with JRA-55 atmospheric reanalysis, whereas PIOMAS is based on a different atmospheric forcing (NCEP/NCAR Reanalysis 1) and a different assimilation framework for sea ice and ocean conditions. No GLORYS forcing or relaxation is applied in this experiment. We agree that part of the differences between the ARC10k simulation and PIOMAS may reflect inconsistencies among reanalysis products rather than purely model deficiencies. In particular, differences in

atmospheric forcing characteristics between JRA-55 and NCEP/NCAR may lead to systematic differences in surface heat fluxes and, consequently, sea ice evolution. A colder atmospheric forcing in JRA-55 relative to the NCEP/NCAR Reanalysis 1 forced PIOMAS could contribute to the seasonal ice concentration differences presented in the manuscript. The text has been revised to clarify this point and to avoid over-attribution to a single mechanism.

RC: 645 Year-end. Not sure. Would end of the year be better? If year end is used. The version should be decided. It is year-end here and year end in line 658.
AC: Changed to "by the end of the year".

RC: Line 647 remove s from figures
AC: Corrected

RC: Line 694 The problem often arise when adding "new" ice is to water above the freezing point. I think that the thickness is less important for instability as long as it is not very small but it may or may not influence the freshwater balance.
AC: We agree that imposing ice over ocean water above the freezing point is a common source of thermodynamic inconsistency. More generally, inconsistencies may arise whenever the relaxed ice state is not in equilibrium with the underlying oceanic or atmospheric conditions, including cases where ice is removed under conditions favorable for ice growth. The discussion has been revised accordingly.

RC: Table 1+Table 2: These are nice to know but I would put them in an appendix as most of the information here is not directly necessary for the point of the manuscript.
AC: We appreciate the reviewer's suggestion. While moving Table 1, which summarizes the MOM6 configuration, to an appendix could be considered because the ocean model is not the primary focus of this study, we believe that Table 2 should remain in the main text. The sea ice relaxation algorithm is implemented and evaluated within SIS2, and the sea ice model configuration is therefore directly relevant to the methodology and interpretation of the results. The parameters listed in Table 2 may also serve as a useful reference for readers seeking to reproduce or adapt the experiments. Because the study presents simulations of a coupled MOM6–SIS2 system, we prefer to retain both tables in the main text for completeness and ease of reference. The tables are concise and provide key information about the model configurations used in the experiments without substantially increasing the length of the manuscript.

RC: Table 3: Are all these experiments addressed? I am not entirely sure where relaxation domain and time scale belongs to. Maybe it is the split of the table on two pages but it is a bit difficult to link the headers of each column with content. I think that this can be improved. What is COBALT?

AC: Yes, all these experiments are discussed in the manuscript. A detailed description of all simulations is provided in section 5. In the copy of the manuscript, the table heading got separated and ended up on different pages, which makes it confusing to relate the header with the table columns. This will be avoided in the revised version. COBALT is a biogeochemical component coupled to MOM6-SIS2 used in one of the multidecadal hindcast simulations for assessing the model drift (mentioned in section 2).

RC: Figure 3: For the numbers underneath the colorbar I would reference an equation from the figure text of figure 3c.
AC: We are not certain which equation the reviewer refers to. The colorbar values correspond directly to the relaxation rate, defined as $1/(3600 \times \tau)$, where τ is the prescribed maximum relaxation timescale in hours. Since the figure shows multiple experiments with different prescribed τ values, the colorbar is intentionally annotated with the corresponding numerical rates for clarity. The figure caption has been revised to improve the explanation of this representation.

RC: Figure 4. Last statement, where are the hours shown?
AC: We agree that this may not be immediately obvious from the figure. The relaxation timescale is indicated in the experiment names shown on the left-hand side of the panel. The figure and caption have been revised accordingly, following a previous reviewer's suggestion, to improve clarity.

RC: Figure 4+5: I think that it would be useful to show the PIOMAS as reference.
AC: PIOMAS seasonal ice concentration and ice thickness fields have been added to figures 4,5, 9, and 10.

RC: The Reference run for NEP10k (e.g row 5a) is more an artifact of a setup that is difficult to use for sea ice predictions without boundaries.
AC: We agree with the reviewer. The NEP10k control simulation employs closed sea ice boundaries and therefore is not expected to provide an optimal configuration for sea ice prediction in the Arctic region. It should be noted, however, that the primary forecast region of the NEP10k system lies south of the Bering Strait. Although the closed lateral sea ice boundaries introduce unrealistic ice conditions in the Arctic portion of the domain, their impact on sea ice conditions in the Bering Sea is comparatively limited. The purpose of including this simulation is to provide a baseline against which the boundary and domain relaxation approaches can be evaluated. One of the primary motivations of this study is to investigate whether sea ice relaxation can mitigate the limitations imposed by the closed sea ice boundaries in the NEP10k configuration.

Response to Reviewer 2:
We would like to thank the reviewer for the thoughtful comments and constructive feedback. Below, reviewer comments (RC) are followed by our responses (AC).

RC: The authors relax sea ice concentration and thickness to PIOMAS, which is a reanalysis product. That is understandable given the lack of observational thickness data. Yet, I'd still be interested in seeing the authors consider examining the sensitivity of their results to other products. Maybe another reanalysis like GLORYS? But more importantly, as well as assimilating to PIOMAS, the authors also evaluate their results in comparison with PIOMAS. I would definitely like to see some evaluation using independent products, both reanalysis, but also observational. There are some satellite sea ice thickness fields available, as well as lots for concentration.
AC: We agree that comparisons with independent observational and reanalysis products are important when assessing the realism of a sea ice simulation or the performance of a data-assimilation system. However, the objective of this manuscript is different. This study was submitted to GMD as a model development paper, with the primary goal of describing, implementing, and verifying the Newtonian relaxation capability in SIS2.

The numerical experiments are designed to demonstrate that the relaxation algorithm behaves as intended in a coupled ocean–sea ice model. Specifically, we assess whether the relaxation can robustly constrain the model state toward a prescribed target field, whether the solution approaches the target as the relaxation timescale decreases, whether the response is smooth and monotonic, and whether the procedure introduces undesirable effects such as numerical instability, model shocks, or substantial drift in the coupled ocean state. For these objectives, the choice of target dataset is not critical. During development, the method was tested using a variety of target fields, including fields derived from other products and idealized synthetic ice distributions, and the algorithm behaved robustly in all cases. We have added a statement to the revised manuscript noting that the methodology has been successfully applied using alternative target products (Section 3.1).

For the demonstration experiments presented here, we selected PIOMAS because it is a widely used Arctic sea ice reanalysis that has been extensively evaluated and validated in previous studies cited in the manuscript. In the context of the algorithm validation study, PIOMAS serves as the prescribed "truth" (or reference) dataset. The relevant measure of success is how closely the relaxed model solution approaches the prescribed PIOMAS state. Consequently, all evaluation metrics are computed relative to PIOMAS, regardless of potential differences between PIOMAS, observations, or other reanalysis products. A comprehensive assessment of the relative accuracy of PIOMAS and alternative datasets would constitute a separate scientific investigation and is beyond the scope of this model development study.

RC: The authors state there is an initialization step, where the grid points where nudging will occur are identified? Does that mean these grid points are fixed in time? They don't change between seasons? Or years, given there can be significant inter-annual variability in the Arctic sea-ice.

AC: We apologize if this point was not sufficiently clear in the manuscript. As described in Section 4.2.1, the initialization routine identifies the ocean grid points where nudging will be applied based on the user-specified relaxation mask and relaxation rates. Grid points with relaxation rates greater than zero are included in the nudging domain. The resulting set of grid points is fixed throughout the simulation and does not vary seasonally or interannually. The purpose of the relaxation mask is to define where relaxation is applied and to allow spatially varying relaxation strength (e.g., stronger relaxation in one region and weaker relaxation in another, or a spatially uniform relaxation rate as used in the ARC10k experiments). Temporal variability is introduced through the target sea-ice fields rather than through changes in the nudging domain itself. At each time step, the model solution at the selected grid points is relaxed toward the corresponding target values. Therefore, seasonal and interannual variability in the target sea-ice concentration and thickness fields is naturally reflected in the nudged solution, even though the nudging mask remains fixed in time. We have added a clarifying sentence in section 4.2.1 that states that the relaxation mask does not change in time.

RC: If I am following section 4.2.2 correctly, with the multiple categories and figure 2, it seems like most of the ice goes in within one category, with small changes in the thinner ice categories. Is this optimal? Might a smoother distribution work? Or addition of some ice in all categories? I think some more discussion and potentially some sensitivity analysis would help convince readers the author's approach is optimal. And for the thickness assimilation, does it go in at the maximum thickness in a category, the mean, or just whatever value you take from PIOMAS? Given this is a key aspect of the work, I think more detail and evaluation is justified.

AC: We thank the reviewer for this comment. The objective of the redistribution procedure is not to reconstruct the true subgrid-scale ice thickness distribution (ITD) from the target ice concentration and thickness fields. Because only bulk ice concentration and volume are prescribed (Eqs. 7–8), the problem is inherently underdetermined, and many different category distributions can satisfy the same constraints.

For this reason, we do not view the redistribution problem in terms of finding a unique or optimal ITD. Instead, the procedure is designed to satisfy two practical requirements: (1) conservation of the prescribed bulk ice concentration and volume, and (2) generation of an ice state that can be robustly integrated by SIS2. The adopted approach satisfies both requirements and has proven stable in all experiments performed with the relaxation scheme.

The reviewer is correct that most of the ice is assigned to a primary thickness category, while smaller amounts are distributed among thinner categories. The extent of this redistribution is controlled by the parameter α_C (Eq. 14 and Fig. 2). The purpose of this redistribution is not to reproduce a realistic ITD, but rather to avoid highly singular states in which all ice resides in a single category. During development, we found that assigning all ice to a single category occasionally resulted in failures of the SIS2 category redistribution procedure, whereas the adopted multi-category initialization did not exhibit such behavior.

A smoother distribution, or one that populates all categories, could certainly be constructed. However, the available target fields provide no information about the underlying ITD shape, and therefore no objective basis exists for selecting one smooth distribution over another. We explored alternative approaches during development, including methods that produced smoother category distributions, but found no evidence that they improved the performance of the relaxation algorithm. Consequently, we did not pursue a systematic sensitivity analysis of alternative redistribution schemes.

Regarding the target thickness, the bulk ice thickness from PIOMAS is not assigned directly to an individual SIS2 category. Instead, the target ice concentration and thickness are redistributed among the SIS2 categories while conserving the prescribed bulk ice concentration and ice volume. The primary category is selected based on the mean ice thickness over the ice-covered fraction of the grid cell (Eq. 11), and the remaining categories are populated according to the redistribution procedure described in Section 4.2.2.

We have added clarifying text to Section 4.2.2 and refer the reviewer to our response to a related comment from Reviewer 1.

RC: Is the relaxation approach the same in the sea ice pack interior as within the marginal ice zone?

AC: The relaxation approach is the same, however relaxation rate may vary depending on the relaxation mask that the user designs for the simulation. In the ARC10k experiments, the relaxation rate is spatially uniform over the relaxation region (north of 60°N). Consequently, the same relaxation is applied within the marginal ice zone (MIZ) and the pack ice whenever both lie within the relaxation region. In contrast, the relaxation mask used in the NEP10k experiments is spatially varying (Fig. 3c). As a result, the relaxation rate depends on location rather than on the sea ice regime itself. For example, when the MIZ is located in the Bering Sea, the applied relaxation is weaker than over the Arctic pack ice. Later in the season, when the MIZ retreats northward into the Arctic Ocean, it may be subject to relaxation rates similar to those applied within the pack ice. Thus, the relaxation algorithm does not explicitly distinguish between pack ice and MIZ conditions. Any differences arise solely from the spatial structure of the relaxation mask.

RC: For ARC10K and looking at figure 3c shows the 60N nudging domain going through the middle of the Hudson Bay Complex, across the Labrador Current and into the Baltic, all areas of active sea-ice. Is this latitude the best? Wouldn't a slightly lower latitude work better?

AC: We agree that the 60°N relaxation mask used in the ARC10k experiments is not necessarily optimal for all regions of the domain. In particular, the relaxation boundary intersects several seasonally ice-covered areas, including Hudson Bay, the Labrador Sea, and the Baltic Sea, where a different relaxation mask might be more appropriate for operational or processoriented applications. However, the primary objective of the ARC10k experiments was to evaluate the implementation and behavior of the relaxation algorithm rather than to optimize the relaxation mask for a particular application. The spatially uniform relaxation north of 60°N was intentionally chosen as a simple test configuration that allows the effects of the relaxation timescale to be assessed without introducing additional complexity associated with a spatially varying mask.

We agree that a different mask, potentially including spatially varying relaxation rates similar to those used in the NEP10k experiments, would likely be more appropriate for realistic simulations in the subpolar seas. The design and optimization of such masks, however, is beyond the scope of the present study. We have added clarification to the manuscript to make this distinction more explicit (section 4.3).

RC: No multi-year hindcasts with ARC10K? In any case, for the single-year hindcasts, why those specific years? Might the relaxation work differently in high vs low sea-ice years?

AC: Unlike the NEP10k configuration, no multi-year ARC10k hindcast simulations were available for this study. The ARC10k experiments were designed primarily to test and evaluate the implementation of the relaxation algorithm, and we considered one-year integrations sufficient for this purpose. Long-term NEP10k hindcast simulations were used instead to assess whether relaxation introduces slow-evolving drift in the coupled ocean–ice system. No evidence of such drift was found.

The specific year selected for the ARC10k experiments was not intended to represent a particular sea ice state. In principle, the effectiveness of the relaxation algorithm should not depend on whether a year is characterized by relatively high or low sea ice extent, provided that the relaxation mask encompasses the regions of interest. However, the impact of relaxation can depend on the spatial overlap between the sea ice cover and the relaxation mask. For example, in the NEP10k configuration the relaxation strength decreases southward. During years with extensive ice cover, portions of the marginal ice zone may extend into regions where the relaxation is weaker, whereas during low-ice years the marginal ice zone may be located farther north where relaxation is stronger. This illustrates the importance of carefully designing the relaxation mask for a particular application.

A systematic assessment of the sensitivity of the relaxation algorithm to interannual variations in sea ice conditions was beyond the scope of the present study. The primary objective here was to demonstrate and evaluate the implementation of the relaxation capability in SIS2 rather than to optimize relaxation parameters for specific sea ice regimes.

RC: In terms of evaluation, would like to see how the relaxation impacts the sea-ice velocity. And for the upper ocean, how the stratification and mixed layer depths respond.

AC: We agree that the response of sea ice velocity, upper-ocean stratification, and mixed-layer depth to the relaxation procedure would provide useful additional diagnostics. However, the primary objective of this study is to describe, implement, and validate the sea ice relaxation capability in SIS2, with emphasis on its ability to constrain sea ice concentration and thickness while maintaining stable and physically consistent coupled model behavior. Accordingly, our evaluation focuses on variables that are directly affected by the relaxation algorithm, namely sea ice concentration and thickness. Sea ice velocity is not directly modified by the relaxation procedure and therefore falls outside the scope of the present validation. Similarly, while upper-ocean stratification and mixed-layer depth may respond to changes in sea ice conditions induced by relaxation, these diagnostics are more closely related to the physical response of the coupled ocean–ice system than to validation of the relaxation algorithm itself.

The manuscript already includes an assessment of Arctic surface temperature and salinity in the multi-decadal NEP10k simulations to evaluate whether the relaxation introduces spurious ocean drift. We have also examined impacts on additional oceanic properties and fluxes, which demonstrate that the relaxation influences the coupled system through its effects on sea ice distribution and associated air–sea interactions. However, presenting a comprehensive analysis of the dynamical response of the sea ice and upper ocean would substantially broaden the scope of the manuscript beyond its intended focus on development and validation of the relaxation capability. We agree that a detailed investigation of the response of sea ice dynamics, upper-ocean stratification, and mixed-layer processes to sea ice relaxation would be valuable and is a natural direction for future work.

RC: For the discussion, would the authors expect this to work for other models? What about for sea-ice in the southern hemisphere? In a changing/warming climate?
AC: The relaxation framework is not model-specific and can, in principle, be applied to any sea ice model that allows external forcing or state adjustment of ice variables. The algorithm itself can be used in both hemispheres as well as in global or regional configurations, provided the relevant sea ice state variables are available for adjustment. Regarding application in a changing or warming climate, the effectiveness of the method depends primarily on the fidelity of the target fields (e.g., observations, reanalyses, or externally generated model fields) used in the relaxation. The relaxation procedure itself simply constrains the model toward these prescribed states and does not introduce additional climate information. Therefore, its performance under climate change conditions will reflect the ability of the target dataset to represent evolving sea ice conditions, rather than any inherent limitation of the relaxation approach.

RC: Figure 1 caption – For ARC10K, since the caption mentions relaxation north of 60N, it would be good to indicate that latitude with a line on the figure.
AC: The figure and the figure caption have been edited.

RC: Table 1 – Is the background kinematic viscosity really zero?
AC: This is correct; in the present MOM6 configuration the background kinematic viscosity is set to zero. The MOM6 numerical formulation reduces the need for a prescribed constant

viscosity lower bound, as stability is primarily ensured through flow-dependent parameterizations and conservative numerics. Importantly, the effective viscosity is not zero. In this configuration, a biharmonic Smagorinsky-type nonlinear scheme is used to compute the lateral viscosity, which provides spatially and temporally varying dissipation and remains nonzero in regions of resolved shear. The effective viscosity is determined as the maximum of the background viscosity, the Smagorinsky viscosity, and other contributing terms, ensuring nonzero dissipation where shear exists.

RC: Section 3.2 title – Model validation metrics or Model evaluation metrics? I'm assuming the model code is valid, so this should more likely be the latter option.
AC: We have revised the section title accordingly to avoid ambiguity.

RC: Line 449 – "This supports the assumption that the Bering Sea ice state is weakly sensitive to Arctic LBCs." Is this surprising given the flow through Bering Strait is generally to the north. What about during a northerly wind event when ice is exported to the south (which has been shown to occur in the literature).
AC: We agree that the result is not unexpected given that the mean ocean and sea ice transport through Bering Strait is northward into the Arctic Ocean. Nevertheless, this behavior cannot be assumed a priori, particularly in the NEP10k configuration where the northern lateral boundary is located relatively close to the Bering Sea. As the reviewer notes, observations have shown episodic southward ice transport through Bering Strait during northerly wind events, which can influence sea ice and ocean conditions in the northern Bering Sea. This possibility was one of the motivations for testing relaxation along the northern lateral boundary. In unconstrained simulations, unrealistically thick ice is often produced near Bering Strait and in the northern Bering Sea during heavy ice years. Although 2001 was not an extreme case, positive ice thickness anomalies are still evident in this region (Fig. 5b revised version). Our results indicate that relaxation applied only along the lateral boundary is insufficient to effectively constrain the northern Bering Sea ice state. In contrast, domain-wide relaxation, even with a weaker relaxation rate over the Bering Sea, substantially improves the simulated ice conditions. These findings support the conclusion that the Bering Sea ice state is only weakly sensitive to Arctic lateral boundary conditions in the context of the present experiments.

RC: Line 571 – Why might the relaxation being having no effect during the freeze-up?
AC: There are several possible reasons why the relaxation has a weaker impact on sea ice concentration during the freeze-up season. As discussed in the manuscript, sea ice concentration is inherently more difficult to constrain than sea ice thickness because it responds to oceanic and atmospheric forcing on much shorter timescales. During freeze-up, when sea surface temperatures are near the freezing point, relatively small variations in surface heat fluxes can rapidly promote or suppress ice formation, leading to changes in ice concentration that may partially counteract the relaxation tendency. In addition, the ARC10k simulations are forced with

JRA-55 atmospheric fields, whereas the PIOMAS target fields are generated using NCEP/NCAR Reanalysis 1 forcing. Differences between these atmospheric datasets may produce local thermodynamic conditions that are not fully consistent with the prescribed target ice state, reducing the effectiveness of the relaxation. This effect is expected to be more pronounced for ice concentration than for ice thickness, which evolves on longer timescales and is therefore more readily constrained. Finally, the relaxation experiments presented here use monthly target fields. Higher-frequency target fields, together with stronger relaxation, may provide improved control of ice concentration during the freeze-up period. Additional discussion of these factors has been added to Sections 7.2.1 and 9 of the revised manuscript.

RC: Figure 13 – Looks like the different experiments have different run length. I.e. one goes to 2025. Unless there is something relevant to be discussed in that extended period, just truncate in in the plots and show all results for the same time frame.
AC: Figure 13 has been revised to show all experiments over a common time period.