Point-by-point response to the reviewers' comments on the manuscript:

Freshwater input from glacier melt outside Greenland alters modeled northern high-latitude ocean circulation

Jan-Hendrik Malles, Ben Marzeion, and Paul G. Myers

We thank the two reviewers for the constructive and detailed feedback, which was very helpful to improve the manuscript! We have done our best to address the issues raised by the reviewers and provide a point-by-point reply to each of the reviewers' comments, including an explanation of subsequent changes to the manuscript. ("RC" for "reviewer comment", and "AR" for "authors' response")

1 Major/general issues

First, we would like to address issues that either came up repeatedly throughout the individual reviews and/or were raised in the reviewers' general comments before addressing the more specific line comments below.

1.1 Statistical significance

As both reviewers raised this issue, we give a detailed response on it here. Firstly, both reviewers posed the question why we used the Wilcoxon signed-rank test. We did so because it is not subject to the assumption that differences between samples are normally distributed, as is the paired Student's t-test. While the results are qualitatively similar applying the two tests, we think avoiding the normality assumption makes the results more robust.

To make this choice more clear to the reader we added the following to the manuscript at the end of section 3.1: "*In order to test the differences between our two model runs for statistical significance, we apply a Wilcoxon signed-rank test on the monthly means. We chose this test over the paired Student's t-test, as it is non-parametric and hence not subject to the assumption that differences between the tested samples are normally distributed. Although both tests yield qualitatively similar results, we chose the Wilcoxon signed-rank test* *for our analysis, since the normality assumption might be violated in some cases. If not stated otherwise, differences between the two NEMO simulations described in this section are statistically significant at p < 0.05.*"

As Reviewer 1 pointed out, there also is the issue of the ocean's chaotic behaviour (i.e., eddy activity). Naturally, in regions where the circulation is mainly chaotic (turbulent), as in the southeastern part of Fig. 6 , differences might be statistically significant, but might not give insight into whether the differences are systematic. Thus, it is true that one has to be careful when interpreting the results as to whether they reflect systematic or chaotic changes. In that sense it is probably useful to look at more persistent features than eddies, though internal variability still plays a role, as mentioned in the Discussion. Concerning SSH differences in the southeastern part of Fig. 7, the Carret et al. (2021) reference provided by the reviewer is useful to underline that differences there are predominantly caused by chaotic variability. We added the following to the manuscript in section 3.1.2 to address this issue: "*Though differences in the southeastern parts of Fig. 6 are statistically significant, they show a rather chaotic pattern; reflecting the high eddy activity in this area (Carret et al., 2021).*"

Reviewer 2 put forward the idea that bootstrapping could be used to "*to test for significance*". While it is an interesting idea, it is not clear to us at the moment how this would lead to a more robust significance testing of differences between our two model runs than the Wilcoxon signed-rank test without overcomplicating the issue.

Finally, Reviewer 1 mentioned that "*the numbers quoted in the text should be consistently equipped with p-values, and it should be clarified when you are only using the last 5 years or when you are using the full period (which is the default?)*". In the text we aimed to mainly discuss differences that are statistically significant. The significance of the statistical test is already given in the line plots. In order to not make the text more convoluted with additional p-values, we think it is better not to add them to the text, but to remind the reader where to find the p-values by adding the following to beginning of the Results section: "*Statistically significant results presented in this section also refer to the second half of the NEMO simulations, if not stated otherwise.*" and in the added paragraph on the statistical test: "*If not stated otherwise, differences between the two NEMO simulations described in this section are statistically significant at p < 0.05.*"

1.2 Description of the experimental setup

We amended the subsections 2.1 and 2.4 (now 2.3.1) with more details and more comprehensible wording. Also, we adjusted the subsection numbering in section 2 (Data and Methods), as it was erroneous in the previous manuscript version. The details are given in the answers to individual reviewer comments on the matter below.

1.3 Presentation of results

Reviewer 1 stated: "*Especially when discussing physical mechanisms, some of the results could be re-written with a focus on (assumed) causality, reminding the reader from time to time what is the link to the added freshwater, e.g. in L257-266, L274-277, L315-333 and elsewhere. In addition, please check with a critical eye whether the results section could be thinned out a little bit – there are many interesting results in the paper, but I feel that the most interesting parts sometimes get a little bit lost amid the overall volume.*"

We now rewrote the lines mentioned by the reviewer to make the proposed effect of the increased freshwater flux clearer there. We also removed following aspects from the Results: the statistically not significant increase in net northward volume flux between Iceland and Scotland, the proposed increase in strength of the gyres in the Norwegian and Greenland Seas, as well as the mentioning of potential changes in circulation at the junction of the Beaufort Gyre and the Transpolar Drift. Those were the parts of the results that we deemed dispensable.

1.4 Restructuring of figures

In order to address the comment of both reviewers on the amount and coherence of figures, we merged the difference plots of several variables into one figure for each examined region; with the following subplots: a) Potential temperature (0-200 m), b) salinity (0-200 m), c) potential temperature (200-600 m), d) sea surface height. This should make the relevant information more accessible to the reader, alleviating the need to go to the Appendix for certain plots. In the text we changed the reference to individual panels from, e.g., Fig. 2 panel a) to *Fig 2a* for better readability. Moreover, we added/moved plots, as described in the answers to individual reviewer comments below.

Reviewer 1 mentioned that the line plots could be moved to the Appendix without losing information in the main body of the manuscript. We decided to keep two of the line plots in the main body, as we deem them to carry enough information to justify this. The information they add compared to solely stating the significance levels is that they show the transient evolution of the differences between the simulations (and their statistical significance).

1.5 Restructuring of the Discussion

Following the idea of Reviewer 2, we now structured the Discussion in (sub-)sections and moved the previously discombobulated information into the respective (sub-)sections. We hope that this makes the Discussion more coherent and thus more intelligible.

1.6 Icebergs

Reviewer 1 stated the following: "*I commend the authors for incorporating an interactive iceberg model into their setup, but unfortunately no results from this model component are shown. For example, it would be very interesting to see a map of the meltwater distribution from icebergs released from the glaciers outside of Greenland, which should also help interpreting some ocean changes away from the coasts.*"

To address this comment we added a plot showing the difference in iceberg meltwater distribution. To accomodate for this plot, we changed the title of the subsection 3.1.4 to *Arctic Ocean, Sea Ice, and Icebergs* and added the following to the text: "*As expected from the difference in calving input between the halfsolid and the noOGGM simulations, Fig. A11 shows areas of statistically significant iceberg melt increase around Svalbard and the Russian Arctic islands. Moreover, we find more iceberg melt throughout the Arctic ocean, although this is not statistically significant. Since the presence of icebergs from Greenland in Baffin Bay and the CAA straits is already large, the addition of calving from the CAA in the halfsolid NEMO does not significantly alter the iceberg melt pattern in that region.*". Note that we used the the figure-numbering from the revised manuscript here, as the plot did not exist in the previous version.

1.7 Boundary currents

Reviewer 2 stated: "*I am curious as to how boundary currents are represented in the simulations. As near surface, shelf currents transport the majority of freshwater input from the coast, there is a certain degree that must be resolved. Is 1/4th degree sufficient? What is the role that shelf currents may play on freshwater pathways, particularly in the Canadian Arctic Archipelago which is made up of a network of straits and basins?*"

Since our configuration of NEMO uses a tri-polar grid, resolution is finer closer to the pole in the CAA than might be expected at 1/4 degree (see Fig. A1). As shown in papers such as Garcia-Quintana et al. (2019, doi: 10.1029/2018JC014459), this means resolutions of 10-14 km in the CAA region. Although such a resolution can still not represent the inshore coastal components of the boundary currents (which need resolutions of 1/20 degree or higher - e.g. Gou et al. (2022, doi: 10.1029/2022JC018404), simulations at this resolution do a solid job of representing the main shelf break boundary currents (e.g., Gillard et al. (2022, doi: 10.1016/j.ocemod.2022.101974). Hu et al. (2019, doi: 10.1029/2019JC015111) also showed that twin 1/4 and 1/12 degree simulations had very similar transports through the CAA. In summary, we agree with the reviewer that a higher resolution will do a better job of representing the details of the current structure in the CAA region. But as we argue above, 1/4 degree provides an acceptable representation of the circulation in the region. And unfortunately it is completely impractical to run sensitivity experiments at resolution higher than a 1/4 degree - our last ANHA12 experiment effectively took a year of real time to run 20 years, and that was a single realization based on our learnings from the 1/4 degree ANHA4 configuration.

On the CAA transports we added the following to the Discussion: "*Concerning transports across the CAA, Hu et al. (2019) showed that twin 1/4 and 1/12 degree simulations were similar in this regard.*"

1.8 Model evaluation

While we agree that more model evaluation would certainly be valuable, it is out of the scope of this study. Our aim was to establish a first estimate of the ice-ocean coupling effects outside Greenland, and the manuscript is already lengthy. Though systematically wrong model behaviour might impact our results, we deem it unlikely that the model setup we chose is fundamentally so flawed that it would nullify our findings. We added the following to the Discussion: "*Finally, a thorough evaluation of the model results with observations could reveal whether the inclusion of glacial melt runoff outside Greenland actually enhances the model's fidelity.*"

We rewrote the end of section 2.1 to also add more references on previous NEMO-ANHA4 evaluation efforts : "Apart from our newly added freshwater flux, NEMO-ANHA4 setups akin to the one described here have been used before to study ocean circulation processes in the northern high-latitudes *(Marson et al., 2021; Hu et al., 2019; Castro de la Guardia et al., 2015). Furthermore, NEMO-ANHA4 has been evaluated in previous studies on aspects as, for example, circulation in the northern Baffin Bay (Ballinger et al., 2022), eastern CAA (Izett et al., 2022), and Labrador Sea (Gillard et al., 2022; Pennelly and Myers, 2021; Garcia-Quintana et al., 2019; Holdsworth and Myers, 2015), as well as on eddy (Müller et al., 2017) and sea-ice features (Bouchat et al., 2022; Hutter et al., 2022; Ballinger et al., 2022). The model proved to generally agree well with observations, and further evaluation is not in the scope of this work.*"

2 Reviewer 1: line comments

RC: *L34 and following: Some more recent references such as Swingedouw et al. (2022), Martin and Biastoch (2023), Devilliers et al. (2024), could be added here.*

AR: We incorporated the mentioned references and did some changes to the paragraph; see answer to comment below.

RC: *L35: As far as I can see, these studies all cover the historical period. Have there been any studies on the impact of glacier meltwater on ocean circulation in future projections? Or in more idealized setups?*

AR: Since we study a historical period, we previously only included studies doing this as well. On the suggestion of the reviewer, we now also added references on future periods. The part of the manuscript now reads as follows:

"*While there has been previous research on the impact of Greenland melt on modeled ocean properties, mostly focusing on the AMOC, they have either added an idealized high 'worst-case scenario' (∼0.1 Sv; e.g., Jackson et al., 2023; Weijer et al., 2012; Castro de la Guardia et al., 2015; Swingedouw et al., 2013) or realistic historical (∼0.01 Sv; e.g., Martin et al., 2022; Martin and Biastoch, 2023; Schiller-Weiss et al., 2024) freshwater flux from Greenland only, or did not disentangle the impact of the freshwater flux from Greenland and from the glaciers in regions surrounding it (e.g., Devilliers et al., 2021; Swingedouw et al., 2022; Devilliers et al., 2024). Since climate models used for decadal or centennial projections mostly do not include future GrIS melt (Swingedouw et al., 2022), the influence of GrIS melt on future climate model projections has also been studied (Jungclaus et al., 2006; Swingedouw et al., 2015; Saenko et al., 2017).*"

RC: *L41: "roughly half that of the GrIS": in which period? Some values would also be insightful*

AR: We added: "[...] roughly half that of the GrIS *over 2010 - 2019 (∼125 Gt* a^{-1}*, see e.g.,* Hugonnet et al., 2021; Zemp et al., 2019; Slater et al., 2021) [...]"

RC: *L42: increased compared to what?*

AR: For clarity we added: "[...] whether increased freshwater input at the coasts of the aforementioned regions *due to glacial melt* does [...]"

RC: *L56: Is there no observational study that measured submarine melt (and could be cited here)?*

AR: We added Sutherland et al., 2019 (doi: 10.1126/science.aax3528).

RC: *L92: Since it is referred to several times later, what is the physical or numerical motivation behind increasing the mixing at discharge locations in NEMO?*

AR: We added the following sentence subsequently: "*This is to mimic vertical mixing due to inertial shear at locations where runoff enters the ocean (Horner-Devine et al., 2015; doi: annurev-fluid-010313-141408) and thus to prevent that freshwater accumulates too strongly in the top grid cell.*"

RC: *L102: The causality between lack of data and computational cost of a model is not clear to me.*

AR: We changed the sentence to: "Because observational data on glaciers, needed to constrain more complex representations of glaciological processes (e.g., ice thickness, spatial distribution of mass balance, albedo, basal velocity) are scarce, *such processes are not included in the model* and its computational cost is hence relatively low."

RC: *L105: Reference Fig. A2 here (and consider moving that figure to the main text). Why use exactly this set of glaciers and not, for example, also the Scandinavian glaciers?*

AR: We referenced Fig. A2 (as previously named) and moved it to the main text (now Fig. 2). We also added to the text: "*Scandinavian glaciers are not included, since their melt rate is roughly one order of magnitude lower than that in the regions we included, and thus unlikely to alter our results meaningfully. Moreover, there are no marine-terminating glaciers in this region.*"

RC: *L141: "halfsolid" is a rather unintuitive name, it sounds more like a sensitivity run to test different liquid-to-solid ratios. Please consider renaming it.*

AR: We see the point in the reviewer's comment, but chose to stay with that name for the moment, lacking a convincing alternative. *OGGM* vs. *noOGGM* are too similar in our opinion. Namings like, e.g., *reference* vs. *sensitivity* can be confusing, as they do not clearly indicate what they stand for. If we found a convincing alternative naming, we would be open to adopting it.

RC: *L146: Remove or clarify the parenthesis here, as "baseline" could be ambiguous (Dai and/or Bamber runoff?) to the reader that doesn't remember Section 3.1 in detail.*

AR: Added: "[...] (excluding the *Dai et al. (2009) baseline runoff) [...]*"

RC: *L146 and following: Why not just quote the numbers for Greenland (which are prescribed and should be the same across both runs) and the contribution from OGGM separately? This would also help reduce the confusion mentioned in the general comments.*

AR: We changed the whole paragraph/subsection to: "*In one of our NEMO experiments, we use the OGGM output of glaciers' surface mass loss in addition to half of the frontal ablation as additional liquid freshwater forcing. The other half of the frontal ablation is added to the iceberg module, as is done for the Greenland solid ice discharge (this experiment hereafter is named halfsolid). We neglect the OGGM-freshwater and -iceberg fluxes in the other NEMO experiment (hereafter called noOGGM). Note that the liquid freshwater and iceberg input along Greenland's coast is derived from Bamber et al. (2018). This data set contains total runoff and solid ice discharge, including from peripheral glaciers, and is the same in both NEMO runs. While this dataset also contains runoff, but no calving, data for glacierized regions outside Greenland (e.g. Svalbard),* *we do not use it for these regions, but add the OGGM-derived glacier mass loss estimates in the halfsolid run. The distribution of the resulting liquid freshwater forcing (excluding the Dai et al., 2009 baseline runoff described in the previous section) is displayed in Fig. 2. Liquid freshwater input along Greenland's coast (derived from Bamber et al., 2018), averaged over 2010 to 2019, amounts to approximately 28.6 mSv (≈ 903 Gt a^{-1}) in the noOGGM run, while OGGM adds roughly 3.4 mSv (≈ 108 Gt a^{-1}) outside Greenland in the halfsolid run. The calving input distribution is displayed in Fig. A2 and amounts to an average of approximately 8.7 mSv (≈ 248 Gt a^{-1}) along the coast of Greenland in the noOGGM experiment (derived from Bamber et al., 2018). Outside Greenland, OGGM adds roughly 1.0 mSv (≈ 28 Gt a^{-1}) of solid freshwater input. This means that OGGM contributes a total of ca. 4.4 mSv (≈ 136 Gt a^{-1}) additional freshwater in the halfsolid run; close to the roughly 4.8 mSv (≈ 150 Gt a^{-1}) Bamber et al (2018) display in their Fig. 3. Note that for Greenland the Bamber et al. (2018) data does not only account for ice mass loss, but gives total runoff values and hence replaces the Dai et al. (2009) baseline runoff along the coast of Greenland. The liquid freshwater from surface melt and the calving of individual glaciers deducted from OGGM output are put into the NEMO-ANHA4 grid cell with the lowest haversine distance to the respective glacier terminus location recorded in the RGI.*"

We hope this makes the different freshwater contributions clearer to the reader.

RC: *L149: As a check, it would be good to compare these 4 mSv to observational estimates (e.g. the "glacier and ice cap runoff outside of Greenland" from Bamber et al. 2018)*

AR: In Bamber et al. (2018; Fig. 3) the glacier and ice cap runoff outside of Greenland appears to be ∼4.8 mSv (≈ 150 Gt a^{-1}), although it is not clear to us whether this is solely glacier mass loss, or also includes snow melt/tundra runoff. Note that the Bamber et al. (2018) also does not include calving data outside Greenland. We added the following to the manuscript: "This means that OGGM contributes a total of ca. 4.4 mSv additional freshwater in the *halfsolid* run*, close to the roughly 4.8 mSv (≈ 150 Gt a^{-1}) Bamber et al (2018) display in their Fig. 3.*"

RC: *L150: "approximately half the freshwater amount released to the ocean due to GrIS mass loss": please provide a reference. Or remove this part, since for the ocean circulation, it does not really matter whether the freshwater is due to mass loss or not.*

AR: Removed that part, since we already give this information (with references) in the introduction.

RC: *L214-220: I do not find that the choice of integrating a relatively large volume (area*depth range) is justified well enough here. It is true that we* *don't know whether picking a single cell "actually reflects water properties at the glacier front best", but certainly this would be the simplest choice, and therefore the authors should argue why averaging over a 50km+ radius may reflect the properties at the glacier front more accurately. Similarly, I do not know how the bathymetry of this NEMO configuration looks like exactly, but it might vary between different (near-)coastal grid points, and averaging over different depths may introduce some unintended artifacts. It would be good to check (perhaps for some selected glaciers) how sensitive the melt rates are to these choices. Or is this already covered by one of the parameters in the latin hypercube ensemble?*

AR: There are two arguments that can be made in favor of not taking solely the nearest ocean grid point, but a radius. The first argument would be that in the case of complex coastal topography (i.e. in the CAA), the situation can arise that the nearest ocean model grid cell actually is not the one nearest to the opening of the fjord the glacier sits in. The second argument is that with a varying grid size, the (horizontal) size of the reservoir the thermal forcing is sourced from also varies. It is imaginable that in a setup, probably with higher resolution than in our case, the width of a fjord is larger than the ocean model grid size. In that case taking the data from more than one ocean model grid cell would be necessary to estimate the thermal forcing the fjord receives from the ocean. (Note that we use a distance-weighted averaging. Hence, the nearest grid cell still has the largest influence on the thermal forcing.)

We added the following to the manuscript: "*In the case of complex coastal topography (for instance in the CAA), the situation can arise where the nearest ocean model grid cell is not actually the one nearest to the opening of the glacier's fjord. While the value of the radius could be adjusted in future work, it also ensures that the thermal forcing's source area is similar among glaciers, as the horizontal resolution is a function of the horizontal position in the modeling domain (see. Fig. A1).*"

It is not clear to us what unintended artifacts could appear by averaging over different depths. We us a depth-weighted averaging for each grid cell, hoping this information can attenuate the concerns. To the respective sentence we added: "[...] compute a depth-averaged *(weighted by vertical level thickness)* value [...]".

As already alluded to by the reviewer, two parameters in the equation we used to estimate submarine melt rates are related to the ocean-glacier heat transfer (B and β). Thus, a change in the thermal forcing by using a different approach for sourcing the thermal forcing should (at least to certain extent) be attenuated by latin hypercube sampling.

Ultimately, it is impossible to tell at the moment whether a change in the thermal forcing source region would improve or deteriorate the accuracy of the results, since submarine melt is only constrained by observed total frontal ablation and uncertainties in this quantity are thus large. Hence, we argue that our approach, while it could certainly be improved, is suitable for a first estimate. And while checking the influence of the thermal forcing source region would certainly be valuable, it might not add crucial information to an already quite extensive manuscript. We already stated the following in the manuscript: "Another aspect that could be further investigated concerning the submarine melt parameterization is which part of the ocean in the marine-terminating glaciers' vicinity should be used to source the thermal forcing from before inserting it in Eq. 2. Refining the distance from the glacier termini as well as the ocean depth range that should be taken into account could help to better constrain submarine melt estimates."

RC: *L230-234: I suggest removing this note about water masses. It sounds overly negative and does not add anything to the analysis.*

AR: Removed that part.

RC: *L247: For stratification, it would be informative to show a comparison of the potential density (and maybe T-S) profiles averaged over a suitable region.*

AR: While we agree that much more analysis could be done, this would go beyond the scope of this manuscript, which is a first overview. We added the following to the Discussion: "*An examination of stratification and vertical heat losses could further refine the processes producing the simulated changes.*"

RC: *L253: Why does the east-west SSH gradient change qualitatively in the same way when adding freshwater at either the eastern or western side? This does not seem intuitive.*

AR: We now added a SSH difference plot (now Fig. 3d) to the main text and moved the SSH gradient/volume flux plot to the Appendix (now Fig. A3). We also slightly changed the method of selecting the points for calculating the gradient (using only the last 5 years instead of all years, which is more consistent). Moreover, to make things clearer to the reader we changed the text to the following: "[...] *We find an increase in sea surface height (SSH) gradient from the eastern and western shelves of Baffin Bay towards its center (see Figs. 3d and A3). As the additional freshwater input in the halfsolid run takes place along the western coast of Baffin Bay, the increase in SSH gradient we find from the west towards the center of the gyre is roughly double the gradient we find from the east. This increase in SSH gradient from the (west) coast towards the center of Baffin Bay leads to a stronger cyclonic circulation* [...]" (Upon a new look at Fig. 2 in Castro de la Guardia et al. (2015), we found that the order of magnitude statement was not accurate, because their figure shows the actual gradient and not differences to a reference. Hence, we deleted that statement.)

RC: *L279: Does Fig. 6 show annual means? If the focus is on deep convection, would it be more useful to show/discuss March or winter means (at least for mixed layer depth) instead?*

AR: Changed to JFM differences, which are qualitatively similar, but naturally have higher values.

RC: *L281: But surely the interannual variability is not independent between the two model runs, since some/most of it is determined by the surface forcing?*

AR: Yes, it certainly is not independent. Changed to only: "[...] *are within the observed interannual/natural variability (Kieke and Yashayaev, 2015)."*

RC: *L304: Here it could be mentioned that the lack of a signal in the deep ocean (e.g., the AMOC lower limb) could also be due to the short simulation time, especially since the freshwater signal will take some time to transit from the main input regions to the subpolar North Atlantic.*

AR: Changed the sentence to: "[...] *the differences between and/or the length of our model runs might just not be large enough to have an effect on that large-scale circulation feature (Böning et al., 2016; Garcia-Quintana et al., 2019).*"

RC: *L312: Percentage changes of FW flux are not really useful, since they depend on the reference salinity (e.g., Schauer & Losch 2019), especially if, like here, the baseline FW flux is not too far from zero. If the aim is to decompose FW transport anomalies, it would be more appropriate to apply a decomposition into a velocity anomaly and a salinity anomaly (and a nonlinear residual) term.*

AR: We agree that the freshwater content/transport concept is problematic and we changed that part to the following to avoid any misleading/confusing statements: "[...] *The volume flux out of the Barents Sea increases (Fig. 9c), leading to a net volume flux decrease of 0.11 Sv (≈ 4 %). However, some of the additional freshwater input from Svalbard and the Russian Arctic remains in the western Barents Sea and in the Kara Sea (see Fig. A12).*" Note that Fig. 9 was formerly Fig. 10.

RC: *L314: "freshwater input leaving through the BSO is salinified": The two opposing pairs of words make this statement very confusing.*

AR: Resolved by the changes made based on the comment above.

RC: *L317 and Fig. 10: How is "positive volume flux" defined here?*

AR: We already wrote "northward (positive)", but now added "[...] *through*

Fram Strait" to the end of the sentence.

RC: *L346: Especially for the western Greenland Sea, I would also expect sea ice advection from the Arctic Ocean to play an important role in setting sea ice thickness. Could you check the differences in sea ice volume transport across Fram Strait and the BSO between your two simulations?*

AR: We added the following sentence to the section: "*We also find a decrease of 1 % in the southward (negative) sea ice velocity's absolute value across Fram Strait ($p < 0.05$; not shown).*". Please also see the response to the comment after the next.

RC: *Fig. 10a+d: The p-values (not significant) and line styles (significant) are contradicting each other.*

AR: As stated in the figure caption(s): "Values in the lower left corners show the p-values of Wilcoxon signed-rank tests of differences between the differences of the first and last five modeled years.". This means it is the statistical significance of the difference between the data along the blue and the red line.

RC: *L358: I also find it surprising that there is less sea ice in the "halfsolid" run, I would have expected the opposite for two reasons: a) the meltwater should strengthen the halocline, and b) the latent heat flux from melting icebergs should cool the ocean surface locally. Could you comment on this a bit more? Plotting the seasonal cycle of sea ice thickness and a comparison with Marson et al. (2021) might provide some insights.*

AR: While we agree that the mechanisms pointed out by the reviewer would counteract an increase in heat supply to the sea ice, the dynamic/kinematic change would be unchanged. We now added a plot of differences in sea ice production to have a look closer look at where which effect (thermodynamic vs. dynamic) dominates where. It shows that there indeed are areas of increased sea ice production (CAA, Barents/Kara Sea), but also areas of decreased sea ice production (Svalbard and western Greenland Sea), which leads to a (slight) overall decrease in sea ice thickness. We changed the paragraph on sea ice in the manuscript incorporating the new figure: "[...] *The smaller increase in upper layer temperature in the CAA compared to the Fram Strait and eastern Greenland areas suggests that other factors than increased ocean heat content play a role there. The decrease in ice thickness in the CAA is driven by less sea ice advection, since the increase in SSH across the region leads to a divergent flow out of the area (see Figs. 5d and A9). This is also reflected in Fig. A10 showing more sea ice production in the CAA area in the halfsolid run, likely due to the decreasing salinity increasing the freezing point. As expected from the higher temperatures in Baffin Bay in the halfsolid run, the sea ice is slightly thinner in this area as well, although differences in sea ice production are heterogeneous (see Fig. A10). This indicates that dynamical factors, i.e. more southward sea* *ice transport through Davis Strait, also play a role here. The only area where we find a slightly increased sea ice thickness is between the Barents and Kara Seas, which is most probably related to the decreased heat transport into Barents Sea due to the rerouting of Atlantic water described above, as well as a local increase in the freezing point. The net difference in sea ice thickness in the northern hemisphere between the two NEMO experiments is intriguing, since we only add freshwater to the ocean, which should not increase its heat content. This suggests that increased high-latitude freshwater input due to glacial melt can decrease sea ice thickness by the changes in ocean circulation it induces.*". Note that Figs. 5d and A9 were formerly Figs. 13 and A16, and that Fig. A10 is the newly created sea ice production figure.

RC: *L359: It is unclear what you mean by "structural changes"*

AR: Changed the sentence to: *This suggests that increased high-latitude freshwater input due to glacial melt can decrease sea ice thickness by the changes in ocean circulation it induces.*

RC: *Fig. 14/Table 1: Layperson question: Is there no contribution from surface runoff or basal melting from these glaciers?*

AR: The quantity we are interested in here is frontal ablation, as this is relevant for the ice-ocean coupling. Frontal ablation is usually defined as calving plus submarine melt. Since we are only interested in frontal ablation we only show these two quantities. Differences in surface melt might be caused by differences in surface thinning close to the glacier front, but will be marginal. Basal melt is not included in the model, as stated in section 2.2.

RC: *L403: It would be good to remind the reader what the Castro de la Guardia et al. paper was about.*

AR: Changed the sentence to "[...,] *who investigated the impact of increased freshwater input along the (west) coast of Greenland on the Baffin Bay circulation and its exchanges with the Arctic through the CAA*, [...]"

RC: *L403 and following: Why compare only with Castro de la Guardia et al.? How about other papers from the introduction? Of course, the setup is not identical, but especially for regional features Devilliers et al. (2021, 2024) might provide a useful comparison with a lower-resolution ocean model.*

AR: We added the following paragraph to the Discussion: "*While there have been numerous studies on the effect of Greenland melt on modeled Labrador Sea convection and subsequently the AMOC, it is not straightforward to compare their results to ours. That is because in this study we focus on the impact of freshwater added in different locations, which naturally leads to different impacts on the ocean circulation. Still, one location that is comparable is the*

SPG/Labrador Sea region, as at least some fraction of the additional freshwater we included in the halfsolid run will impact that area. We see that the patterns in SSH differences between our two runs are qualitatively consistent with previous studies (Saenko et al., 2017; Stammer et al., 2011). This pattern consist of a larger SSH in the Labrador Sea/western SPG area and a lower SSH in the eastern SPG (see Fig. 7). The differences in mixed layer depth we find, are also consistent with previous studies (Schiller-Weiss et al. 2024; Devilliers et al., 2021). The mixed layer depth is decreased in the Labrador Sea and increased in the eastern SPG, though this increase is not statistically significant in our case (see Fig. 6). Areas of increased mixed layer depth that are intertwined with the areas of negative differences in the Labrador Sea are also presented in Schiller-Weiss et al. (2024), while they are not in Devilliers et al. (2021). Using a freshwater forcing around Greenland similar to the one used in Devilliers et al. (2021), but a different climate model, Devilliers et al. (2024) find the opposite mixed layer depth response (increase in the Labrador Sea and decrease in the eastern SPG). Overall, these comparisons with previous studies suggest that the additional freshwater from glacial melt outside Greenland in the halfsolid run might exacerbate the impact of increasing Greenland freshwater input in the Labrador Sea/SPG area.

RC: *L417: I do not find the comparison with the PIOMAS trends from Labe et al. very convincing, since from my reading their magnitude (over the entire 1979-2015 period) is at least one order of magnitude larger than your values.*

AR: It is not necessarily one order of magnitude, since our color scale is cut off at the extreme values. And while it is to be expected that our values are smaller, since additional freshwater input from glaciers outside Greenland is presumably only a minor contributor to total changes in sea ice thickness, the spatial patterns are still consistent with Labe et al. (2018) in some areas. We changed the wording to a less strong statement though: "[...], it is intriguing that Labe et al. (2018) find *negative trends of sea ice thickness between 1979 to 2015 in some similar areas.*

RC: *L421-434: The discussion around individual figures is probably a bit too detailed here – one would need to flip back and forth between this paragraph and the figures (in the results section) quite a bit.*

AR: We now start this paragraph with the following sentence: "*Placing our results for the SPG circulation (see Figs. 6 and A7) further in the context of existing literature, we find* [...]", and removed the following references to the figures in the paragraph. Note that Fig. A7 was previously Fig. A10.

RC: *L473-485: Which of these many possibilities do you find most interesting/relevant? It might be worth focusing this part a bit.*

AR: Due to the restructuring of the Discussion described above (in the response to the general comments), we now have a subsection ”Simulation length”. Here, we go a bit more into detail on this issue.

RC: *L490: Do slower quantities (temperature, salinity, sea ice) not suffer from a (potentially longer) initialization shock either?*

AR: Any drift due to initialization will be similar in our two setups. Thus, it is not clear to us how this would hinder a meaningful (first) exploration of the added glacier melt's impact on near-surface conditions. (We made a similar statement in the Discussion.)

RC: *Fig. A17: Are you sure the units of this figure are correct? 150 Sv is on the order of the ACC transport...*

AR: Yes, this is the northward (not net) transport across 47°N, mostly driven by the North Atlantic Current, which is consistent with, e.g., Mertens et al. (2014, doi: 10.1002/2014JC010019).

RC: *L549: “even” Why is this surprising?*

AR: Removed ”even”.

RC: *How is the baseline (Dai et al.) runoff treated over the OGGM model area? Surely some river runoff in this area is sourced from (seasonal) glacier melt in the first place, so how is double-counting avoided?*

AR: Upon inspection of Fig. 1 in Dai et al. (2009) we found that there is no river runoff measurement station in the regions where we added the glacial melt runoff (except for Iceland). In Dai & Trenberth (2002, doi: 10.1175/1525-7541(2002)003¡0660:EOFDFC¿2.0.CO;2), which Dai et al. (2009) is based on, they state that the approach they use to close such data gaps does not explicitly account for changes in land water storage (which includes glaciers). Moreover, glacier area in Svalbard, the Russian Arctic, and the CAA is around 50 % marine-terminating; the runoff of such glaciers will probably not end up in rivers, decreasing the danger of double-counting. We changed the following sentence: ”The Dai et al. (2009) data does not cover our model period from 2010 to 2019 *and does not explicitly account for runoff caused by (marine-terminating) glacier mass loss.*”

RC: *Why was this specific modelling period chosen? 2010–2019 is quite short (probably too short for the AMOC, see above) and it is unfortunate that it does not fully overlap with the Bamber et al. coverage.*

AR: There is no specific reason for this period other than that we aimed at modeling the most recent decade we had atmospheric reanalysis and glacier model data for when we started the work. While we agree that it is unfortunate that the chosen period does not fully overlap with the Bamber et al. data, we do not think that this challenges the main findings from our generally somewhat idealized model setup, as the Greenland runoff is not the freshwater forcing we are actually interested in. Moreover, we aimed at a first estimate of the effect of increased freshwater input from (high-latitude) glacial melt outside Greenland and not at an optimized hindcast. We also mentioned this issue in the Discussion.

RC: *Is any salinity restoring used?*

AR: We added the following sentence to section 2.1.: ”*No salinity restoring was employed, as that would tend to dampen the freshwater signal and hence suppress the response to the perturbation in the forcing we are interested in.*”

Technical corrections

RC: *L84: Put a line break here*

AR: Done.

RC: *L96: Enderlin 2016 say “up to” 50%*

AR: Changed ”roughly” to ”*up to*” in the sentence.

RC: *L346 simulation -¿ simulations*

AR: Done.

RC: *L404 units are missing*

AR: Added a ”*K*”.

RC: *L541 and following: The use of past and present tense is inconsistent here, please decide on one of the two.*

AR: We decided for the present tense now.

3 Reviewer 2: line comments

RC: *L10: Regarding the abstract, there could be a few more words on the reasoning behind the increase in Baffin Bay’s heat content. Particularly some more specifics could be given, rather than just stating “changes in the subpolar gyre’s structure”.*

AR: Changed the sentence to: "[...] *increase in heat content of Baffin Bay due to an enhanced gyre circulation that leads to an increased heat transport through Davis Strait. We also find changes in the subpolar gyre's structure; an increase/decrease in density/SSH in the eastern part and vice versa in the western part.* [...]"

RC: *L35-37: Can some magnitudes of Greenland melt be offered? For example, artificial hosing experiments typically consist of .1 Sv (Jackson et al. 2018) while current magnitudes of Greenland FWFs are at .01 Sv (Martin et al. 2022)*

AR: Added the orders of magnitude to the text. (Also see answer to Reviewer 1's comment on "L35")

RC: *L56: Would replace the word intricate with complex*

AR: Replaced.

RC: *L70: Rather a long sentence. Would recommend shortening it or breaking it into two parts*

AR: Changed to: "*We then explore the differences in results obtained from the two different NEMO and OGGM experiments. By this we aim to obtain a first-order estimate of the effect ice-ocean coupling outside Greenland has on ocean properties as well as on marine-terminating glacier mass loss.*"

RC: *L85/L89: Does Dai et al. 2009 provide only the continental runoff i.e. river discharge or does it also provide glacial runoff? Bamber et al. 2018 provides freshwater fluxes from Greenland and the Arctic, so do you include both freshwater from continental runoff (Dai et al. 2009) as a baseline and also add runoff from Bamber simultaneously? Can you clarify which freshwater forcing you use and when per data set, as this part is not clear for me*

AR: This part of subsection 2.1 now reads as follows and we hope this resolves the lack of clarity: "*The Dai et al. (2009) data does not cover our model period from 2010 to 2019 and does not explicitly account for runoff caused by (marine-terminating) glacier mass loss. We therefore applied the 1997 to 2007 monthly average baseline runoff. Freshwater input from Greenland is derived by remapping the data published by Bamber et al. (2018) to the NEMO-ANHA4 grid. This data gives the total runoff, including from the ice sheet and peripheral glaciers, thus replacing the Dai et al. (2009) baseline runoff in this region. As this data set only ranges to the end of 2016, we use the 2010 to 2016 average for the three missing years. Note that the Bamber et al. (2018) data also provides runoff, but no calving, estimates for other high-latitude glacierized regions in the northern hemisphere (e.g., Svalbard), but we only use the estimates for Greenland. The handling of additional freshwater from other glacierized regions* *is described in section 2.3.1.*". Note that section 2.3.1 was formerly section 2.4.

RC: *L86-L90: Since both the Dai et al. 2009 and Bamber et al. 2019 freshwater forcings do not cover the entire time period (2010 - 2019), can a sentence be added on what may be assumed using some climatological means? Particularly as 2019 was an anomalous year of Greenland melt and using a mean of 2010 - 2016 will likely smooth out that interannual variability. What are the potential implications of using a 2010 - 2016 mean?*

AR: Later in the manuscript (Discussion) we stated the following: "Additionally, the baseline runoff and the Bamber et al. (2018) data not covering the whole modeling period, might induce some uncertainty in our results, since the impact of the additional freshwater we examined could be altered. If, for instance, the ratio of the additional freshwater in the halfsolid run to the baseline plus Greenland runoff was larger (smaller), the impact would presumably be larger (smaller) as well.". We think that the Discussion is the suitable place to critically reflect on the model setup choices we made, while the Data and Methods section should be mainly descriptive.

RC: *L92: Why does the additional runoff entail an increase in the diffusivity? Can this be elaborated on further? Would this affect local sea ice formation along the shelves?*

AR: We added the following sentence: "*This is to mimic vertical mixing due to inertial shear at locations where runoff enters the ocean (Horner-Devine et al., 2015; doi: annurev-fluid-010313-141408), and thus to prevent that freshwater accumulates too strongly in the top grid cell.*" As this affects vertical mixing in the upper 30m, it might affect sea ice formation, but we presume it would be of second order, since it is confined to the grid cells where runoff enters the ocean.

RC: *L121: What is the influence of basal melt and how is it represented in the glacier model? Are basal melt rates typically smaller than submarine melting rates?*

AR: Basal melt is not included in OGGM. If by basal melt the removal of ice by melting at the (grounded) base is meant: This is one of the mentioned "intricate processes", since it is hard to quantify/parameterize. Thus, it is implicitly included in the surface mass balance. If by "basal melt" the submarine melt of ice shelves is meant: We do not permit the formation of ice shelves in OGGM, since this would require a special numerical representation of those, which is complex, and the large majority of marine-terminating glaciers that are not connected to the ice sheets do not possess a floating tongue anymore. To avoid confusion about this, we changed one sentence to the following: "[...] *and neglects more intricate processes such as refreezing, basal melt, or the surface energy balance.*". We also added the following sentence to the subsection: "*For simplicity, and since the large majority of northern hemisphere marine-*

terminating glaciers outside the Greenland ice sheet do not possess a floating tongue anymore, we neglect the formation of ice shelves in OGGM.

RC: *Fig A2: It's a bit unclear to me if the freshwater forcing (liquid and solid) is distributed uniformly around the Canadian Arctic? Bamber et al. 2018 non-uniformly distributes freshwater along the coastlines. Can this be reiterated or stated in section 2.4 OGGM to NEMO?*

AR: We amended the following sentence like this: "The liquid freshwater from surface melt and the calving of individual glaciers deducted from OGGM output are put *non-uniformly* into the NEMO-ANHA4 grid cell with the lowest haversine distance."

RC: *L149-L150: I am confused about this sentence, the OCGM has 4 mSv of more freshwater than in the halfsolid run, which is about half of the freshwater released from Greenland? Can this be clarified?*

AR: Based on a comment by Reviewer 1 above ("L146 and following"), we partly rewrote the subsection to hopefully avoid confusion now. We initially wanted to express that the glacial melt in the regions regarded outside Greenland is half that of the glacial melt taking place across Greenland. But since this statement was already made in the introduction, we removed it from this subsection, as it seemingly caused confusion in the way we put it.

RC: *L214: Can you clarify why you use a 50km radius upon obtaining the thermal forcing values?*

AR: We gave a detailed response to a similar remark by Reviewer 1 ("L214-220"). As a result, we added the following to the manuscript: "*In the case of complex coastal topography (for instance in the CAA), the situation can arise where the nearest ocean model grid cell is not actually the one nearest to the opening of the glacier's fjord. While the value of the radius could be adjusted in future work, it also ensures that the thermal forcing's source area is similar among glaciers, as the horizontal resolution is a function of the horizontal position in the modeling domain (see. Fig. A1).*"

RC: *L219: Over which depths do you average? The full depth range near the terminus of the glacier? Is this a relatively shallow depth range?*

AR: Yes, the full depth range of the included cells. Table 2 gives the average depth of these, which is relatively shallow (125 - 250 m, depending on the region).

RC: *L229: This is a long sentence. I would recommend rephrasing something along the lines of: "we use the term Atlantic for water (warmer and saltier)* *moving from the Atlantic towards the Arctic Ocean, and the term Arctic for water (cooler and fresher) moving in the opposite direction."*

AR: Changed to: "[...] *we use the term Atlantic for (warmer and saltier) water moving from the Atlantic towards the Arctic Ocean, and the term Arctic for (cooler and fresher) water moving in the opposite direction.*"

RC: *L235: Can a reference or two be added confirming most of the oceanic changes occurring in these depth ranges (0 - 200m and 200 - 600m)?*

AR: We changed the wording to the following: "*In the following sections we will focus on the depth ranges 0-200 m, and 200-600 m, where we found most significant changes. We chose these two ranges to represent the upper layer, and the interface between the upper and the intermediate layer; in these two layers most Atlantic water masses are present and/or formed (Liu and Tanhua, 2021). The upper 600 m are also most relevant to potential feedbacks with marine-terminating glacier mass loss induced by submarine melt, as marine-terminating glaciers outside Greenland rarely exceed 500 m water depth.*"

RC: *L239: It appears the cooler area along the western coast is not statistically significant, based on the significance testing*

AR: Added: "[...] slight *(non-significant)* cooling, [...]

RC: *L247: Regarding stratification, when looking at difference maps of stratification say averaged over the top 200m, do you see an overall positive difference between the two runs?*

AR: We had an initial look at stratification, but it did not appear to change significantly. We also added the following to the Discussion: "*An examination of stratification and vertical heat losses could further refine the processes producing the simulated changes.*"

RC: *L254: Why might there be such a discrepancy in the increase in SSH gradient at the east compared to Castro de la Guardia et al. 2015? Is it due to the freshwater forcing magnitudes?*

AR: Upon a new look at Fig. 2 in Castro de la Guardia et al. (2015), we found that the order of magnitude statement was not accurate, because their figure shows the actual gradient and not differences to a reference. Hence, we deleted that statement. We also made changes to this subsection based on a comment by Reviewer 1 ("L253") above, and hope this makes it clearer to the reader.

RC: *Fig 3: I'm a bit confused as to the difference between the dashed and solid blue lines. The horizontal lines represent the differences between the two runs which are statistically significant? I would suggest adding to the legend* *indicating the black, blue (dashed/solid), and red lines. This may help to clarify the figure.*

AR: This figure was now moved to the Appendix. The caption of the line plot figures already stated the following: "Differences between the two NEMO runs that are statistically significant, according to Wilcoxon signed-rank tests ($p < 0.05$), are drawn as solid lines and dashed otherwise.". We slightly shortened the, admittedly rather long, caption and moved that sentence further up to make the figure clearer to the reader.

RC: *L273: Can you elaborate on the vertical mixing coefficient here as a way to remind the readers, i.e. more mixing resulting in a potential decrease in temperature in the upper 200m layers?*

AR: Changed to: "Again, the *increased* vertical mixing coefficient*, exposing more water to the cold atmosphere,* might play a role here as well."

RC: *L280: Is the MLD averaged over the whole year, or is it during winter months? Please clarify. I would recommend winter months (JFM) or at least March as the MLD is likely deepest then*

AR: Changed to JFM differences, which are qualitatively similar, but naturally have higher values.

RC: *L299: Why do you pick 47N latitude to look at the AMOC and transport, can this be justified more? There may not be huge differences but if you take say the OSNAP West section to compute the volume transport, do you get a similar result?*

AR: We now also checked the AR7W section, which should be quite similar to the OSNAP West section. Based on this we changed the respective sentence in the manuscript to the following: "Concerning the AMOC, we neither find a statistically significant difference in north-/southward or total volume flux across the 47°N latitude in the Atlantic *or the AR7W section across the Labrador Sea (Yashayaev, 2007)*, nor a significant change in the meridional overturning streamfunction."

RC: *L309: This may be the first time introducing the Barents Sea Opening and thus the acronym, please add (BSO)*

AR: Added.

RC: *L320: Can more be clarified, potentially in the methods and model setup description (see general comment), on the NEMO integration time? Or refer the readers to check the model setup discussion?*

AR: Added: "[...] roughly half of the NEMO integration time *(i.e., 5 years)*, [...]"

RC: *L338: "Relatively strongly increased salinity" appears to have opposing meanings, is it a strong or weaker increase in salinity? Can this be rephrased?*

AR: Removed "relatively strongly".

RC: *L344: Can these plots A14 and A15 be placed side by side or consolidated in some way? It may allow the reader to more clearly see the effect that changes in salinity has on SSH in the interior Arctic (see general comment as well)*

AR: As described in the answers to the general comments above, we merged these two figures (amongst others).

RC: *L346: On the part about sea ice thickness, assuming that much of sea ice is advected from the Arctic and thus can influence the thickness, particularly along East Greenland via Fram Strait. Can you comment more on sea ice advection? Or one could check the sea ice thickness change over say an averaged area (i.e. western Greenland sea) or cross section (Fram Strait) and compare that with the change in sea ice production*

AR: In response to comments of Reviewer 1 on sea ice thickness ("L346" and "L358"), we added the following to the subsection: "*We also find a decrease of 1 % in the southward (negative) sea ice velocity's absolute value across Fram Strait ($p < 0.05$).*". Moreover, we added a figure of sea ice production and rewrote the subsection slightly based on this.

RC: *L357: Rework the sentence starting with "That there is a net sea ice thickness...". I would recommend saying instead: "The decrease in net sea ice thickness in the northern hemisphere between the two NEMO experiments is intriguing..."*

AR: Changed to: "*The net difference in sea ice thickness in the northern hemisphere between the two NEMO experiments is intriguing*, [...]"

RC: *L403: Can you explain what the major results from Castro de la Guardia et al. 2015 are to remind the reader? Also highlight the differences between yours and their experiments, as you say the model setup and scope is very similar*

AR: Changed the beginning of the Discussion to: "*Concerning Baffin Bay, we compare our results to those of Castro de la Guardia et al. (2015), who investigated the impact of increased freshwater input along the (west) coast of Greenland on the Baffin Bay circulation and its exchanges with the Arctic through the CAA. They found an increase in heat transport into Baffin Bay through Davis*

Strait and a reduction in volume fluxes through the CAA into Baffin Bay; both related to an increase in the Baffin Bay gyre's strength."

RC: *L404: "lower increase" sound a bit opposing to me, perhaps reword to "smaller increase"*

AR: Done.

RC: *L413: Rather than say the gradients did not change sufficiently, I would refer to the statistical significance of the SSH gradient differences between the two model experiments*

AR: We changed the wording to a less strong statement: "Volume flux through the CAA into Baffin Bay *was found to be* mainly controlled by the SSH gradients across the straits connecting Baffin Bay to the Arctic Ocean (McGeehan and Maslowski, 2012; Hu and Myers, 2014)*, suggesting* that these gradients did not change sufficiently to alter the total volume flux between the Arctic and Baffin Bay in a notable manner, comparing our two NEMO experiments."

RC: *L545: I would cite that previous study i.e. Castro de la Guardia et al. 2015 to remind the readers which study you are referring to*

AR: Done.

RC: *L554: "also goes along with" sounds a bit colloquial, I would suggest alternative wording*

AR: Changed "goes along with" to "*results in*".

RC: *L561: Given this small increase in submarine melt, can a research outlook or statement be made on the contribution of submarine melt when including additional glacier-sourced freshwater input? For example, should future model sensitivity studies on freshwater input consider including submarine melting and associated diffusivity scheme at depth? What would the authors advise for future modeling studies?*

AR: The main problem here is that also the surface runoff of marine-terminating glaciers is usually, at least partly, injected into the fjords at the grounding line depth. This poses the question of how to deal with this in a general manner (i.e., not an individual treatment of each glacier in the runoff scheme). We added the following to sentence to the Discussion: "*As surface and submarine melt of marine-terminating glaciers enter the fjords at depth and is subsequently mixed, the modeling setup could be enhanced by a more accurate representation of meltwater injection at depth, especially when individual fjords are not resolved in the ocean model.* "

RC: *L573: What is meant by decadal snapshot simulations? Decadally averaged or filtered simulations? Can this be elaborated?*

AR: Snapshot simulation means to run the model with the respective boundary conditions for different time periods of a certain length (e.g., decade), such that they could be run in parallel. We added the following sentence: "*Such a snapshot approach would avoid the computational cost of long transient simulations, but might not suffice to capture longer transient processes, as for instance changes in the AMOC.*"