In their manuscript “North Atlantic response to a quasi-realistic Greenland meltwater forcing in eddy-rich EC-Earth3P-VHR hosing simulations” Eneko Martin-Martinez and co-authors present results from a common Greenland hosing experiment though with a more realistic, moderate magnitude of 0.04 Sv and spatial distribution of the additional freshwater flux along the coast of Greenland. Most importantly, they use a climate model of very high resolution in ocean and atmosphere unprecedented for such experiment to my knowledge. Unfortunately, this limits the experiment length to only 21 year. However, they ran three ensemble members for estimating the imprint of internal variability on the response. The presentation of results is focused on the response by the AMOC and potential drivers of its decline.

The experiment is certainly timely as the discussion of Greenland ice sheet melting potentially impacting AMOC strength over the next decades is still ongoing. The experimental design is reasonable though slightly different to previous studies (of which numerous are referenced), which limits comparability and differences to previous studies cannot be attributed to model uncertainty unambiguously. On the one hand, the results align with previous findings, thus are supportive of the state of knowledge. Which is good. On the other hand, this means there is not much groundbreaking novelty inherent to the outcome. Maybe because of this, I have the impression that the authors do not really take full advantage of the unique part of their experimental setup, the model resolution. Overall the results section is largely descriptive, few connections are drawn, less are proven, and the analysis is mostly limited to surface fields. It would be helpful if the authors would connect the responses by indicating responsible processes, even if this is speculative (which should be noted then). And by putting more emphasis on the role of mesoscale features simulated by the high resolution of the model. Arguments related to explicitly resolved eddies, eddy mixing and sharper gradients at fronts are rather qualitative but could in principle be quantified. I wish there would be more specific analysis regarding, for instance, freshwater spreading and associated timescales, changes in surface fluxes (e.g. heat loss), and water mass transformation. How do these shape the initial decline of the AMOC?

> We thank the reviewer for the detailed and constructive comments. We have thoroughly addressed each point in the point-by-point response below (highlighted in blue), and have made a special effort to include new analyses that strengthen the physical interpretation of our results.
>
> Regarding the suggestions related to mesoscale features specifically, several of these were highly relevant and appealing to us. Unfortunately, many of the three-dimensional and high-frequency model outputs required for those analyses (including daily sea surface height) are currently inaccessible to us due to an ongoing issue with our tape storage system, and could not be performed. For instance, daily sea surface height fields would have allowed us to compute eddy kinetic energy (EKE) and track how mesoscale eddy activity along the boundary currents responded to the freshwater forcing and contributed to interior freshwater spreading. Ocean velocity fields at depth would have enabled a more quantitative assessment of eddy-driven lateral exchanges between the boundary currents and the Labrador Sea interior, including estimates of eddy diffusivity and mixing timescales.
>
> With the data currently available to us we have added a new surface-forced water mass transformation (SFWMT) analysis that directly addresses the reviewer’s request for a more process-oriented interpretation of the AMOC decline. This analysis quantifies the respective haline and thermal contributions to buoyancy forcing in the western and eastern subpolar regions, connects them to the changes in overturning transport across the OSNAP-West and OSNAP-East sections, and explains how the lightening of surface waters through freshwater injection reduces the availability of dense water for transformation. This ultimately provides a mechanistic link between the freshwater forcing, the suppression of deep convection in the Labrador Sea, and the warming of the DWBC discussed in the main text. The analysis has been added as a new subsection after the OSNAP cross-section analysis including the new Figures AC2.1 & AC2.2 and Table AC2.1, below.

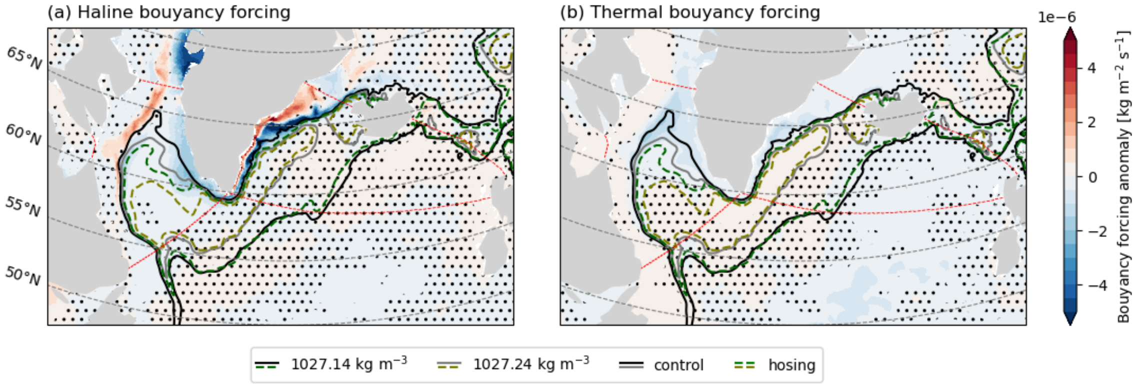


***Figure AC2.1**. Response to the hosing in the last 10 simulated years for the (a) freshwater flux and (b) heat flux contributions to surface forced water mass transformation; filled coloured contours represent the ensemble-mean anomalies of the hosing with respect to the reference control. Contour lines show the surface 1027.14 and 1027.24 kg m^{-3} isoline in control (black, grey) and hosing (green, olive). Non-significant values as identified by the bootstrap methodology are masked with dots to improve the visibility over the significant areas. Red dashed lines mark the limits for the analysis in Fig. AC2.2.*

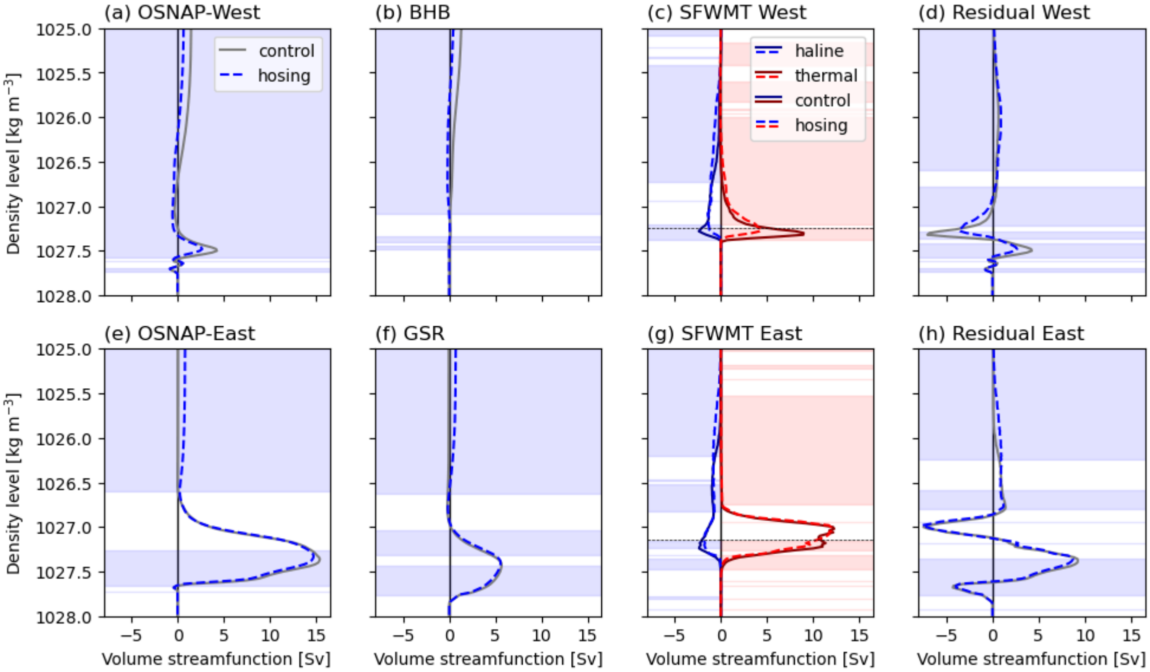


***Figure AC2.2**. Response to the hosing in the last 10 simulated years for the overturning streamfunction in density space in (a) OSNAP-West, (b) BHB, (e) OSNAP East, (f) GSR; values in the hosing (blue) and control (grey) ensembles. Thermal (blues) and haline (reds) contributions to the SFWMT in the (c) West and (g) East regions. Residual WMT computed as the net transport out of the region minus the total SFWMT for (d) West and (h) East). The ensemble means are shown by lines, the background is coloured when the anomalies are deemed statistically significant by the bootstrap methodology. The lines in (c) and (g) represent the 1027.14 and 1027.24 kg m^{-3} density levels, respectively, which isolines are plotted in Fig. AC2.1.*

***Table AC2.1**. Volume streamfunction maximum in density space for OSNAP transect for the last 10 years of the hosing.*

	OSNAP West	*OSNAP East*	*OSNAP Full*
control	*4.3 Sv*	*15.4 Sv*	*17.2 Sv*
hosing	*2.7 Sv*	*14.7Sv*	*15.4 Sv*
anomaly	*-1.6 Sv*	*-0.7 Sv*	*-1.8 Sv*

The other main criticism I have is that there is little validation of the model provided. The authors use the OSNAP section in Figure 5 to show the depth structure of changes in the subpolar North Atlantic. This would be an opportunity to show the mean state of the reference run in comparison to OSNAP observation, at least in the supplementary material. Some discussion of ocean biases is found in Moreno-Chamarro et al. (2025), which is referenced in line 77. Please emphasize this. Nevertheless, a specific discussion of biases with respect to the subpolar North Atlantic would be a great addition to the paper since the model seems to perform rather well (judging from global plots in above reference).

> Apart from Moreno-Chamarro et al. (2025), Frigola et al. (2025) also explores the biases in the historical run with the same model, and few extra discussion for the vertical profiles in the Labrador Sea and AMOC mean states are found in Martin-Martinez et al. (2025). These are now reflected in the revised manuscript which also includes the corresponding references in the discussion.
>
> Regarding further validations, we have computed the OSNAP East and West transports in the control simulation and compared them with the observed ones. We note, however, that both datasets are not fully compatible, as the control run uses fixed 1950 radiative conditions and observations cover the period 2014-2022 with time-varying forcing. For EC-Earth3-VHR we get 16.6 ± 2.7 Sv (mean ± std of monthly data) for OSNAP East and 5.6 ± 2.3 Sv, which lay within the uncertainty range of observations (https://www.o-snap.org/data-access/), corresponding to 16.4 ± 2.9 Sv and 2.7 ± 1.3 Sv, respectively. Interestingly, EC-Earth3-VHR reproduces the stark asymmetry between the observed OSNAP eastern and western components, with much stronger transports associated to the eastern side. Other studies using NEMO-based climate models have shown a positive bias in the transport in the OSNAP-West transport, e.g., Petit et al. (2023). Note that our transport biases are smaller than those presented by Petit et al. (2023). In general, the contributions to the SFWMT are comparable, although the maximum transport and SFWMT do not occur at exactly the same density levels.

Considering a major revision and **sharpening of the arguments**, maybe some **additional analysis emphasizing more the role of the explicitly resolved mesoscale dynamics** (beyond reducing biases), this paper could become a valuable contribution and suitable for publication in OS.

Minor comments by line

93 Separating lines of the six sub-drainage basins could be added to Figure 1b

We have included the sub-drainage basins in Fig. B1b, modified the caption accordingly and added a reference in the text.

109 "…by matching the years between …" I am not sure I understand this sentence. Do you mean that you compute differences by subtracting the experiment from control year-wise, i.e. not comparing the experiment to a long-term mean of control? Please rephrase. The interactive atmosphere would lead to independently evolving internal variability. The shortness of the simulations (21 years) suggests however, that low-frequent variability is likely still preserved. This could be noted here.

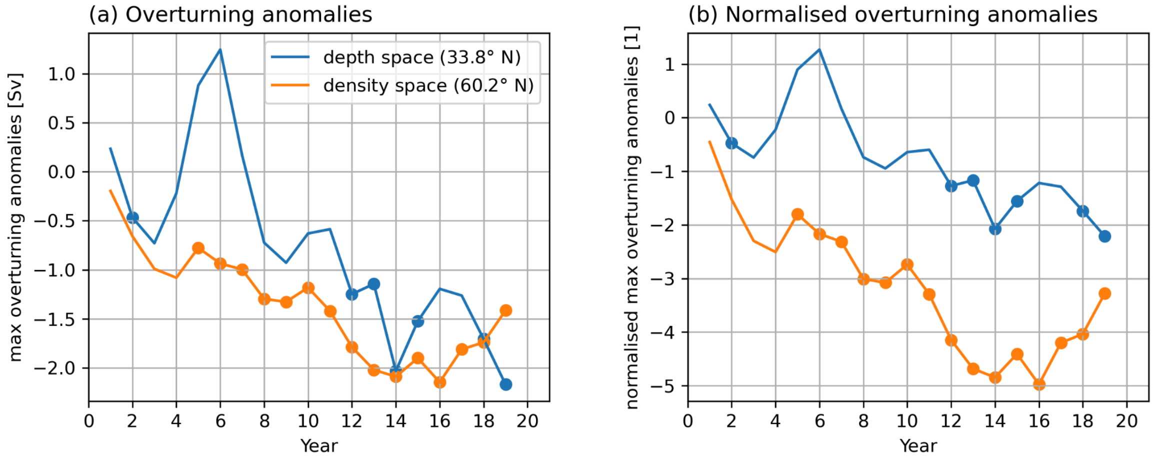
Yes, we compared with the control year-wise. We have extended the sentence to now say: "The anomalies between the hosing and control experiments are computed by matching the years between both experiments after the initialisation, , i.e., the hosing member initialised in 1987 is matched with its respective 21-year period in the control (1987–2007) year-by-year, and the same is done for the other members with their respective control years. This is intended to remove any remaining background model drift and align the phase of internal climate variability. Although the atmosphere could evolve into different states of internal variability, the low-frequency variability is likely to be preserved due to the relatively short length of the simulations."

124 and 129 The latitudes given (33.8N and 60.2N), are these locations of max AMOC in control or of the maximum change in experiment minus control? This is somewhat unclear here. Please specify.

We have modified the sentence as follows: "that is the latitude where the AMOC exhibits the strongest response to the forcing over the last 10 years (Fig. 3a), i.e., the magnitude of the decadal anomaly is maximum".

132 "The ensemble-mean difference …" Please provide numbers here. Judging from Figure 2 I'd say there is a significant difference with an AMOC weakening of 1.5 and 2 Sv respectively. Also, simply comparing the last year is likely misleading and an average weakening of at least the last 5 year should be given.

We agree that the sentence would have benefited from some numerical comparison, as we are doing a deeper comparison later with the full streamfunction in Fig. 3. We have decided to remove this sentence. Also even though signals are relatively comparable in magnitude their signal-to-noise ratio is very different, see Fig. AC2.3 below (included as a supplementary figure now).


(a) Overturning anomalies

(b) Normalised overturning anomalies

Figure AC2.3. *Maximum volume overturning streamfunction and in the Atlantic basin for (a) annual data with 3-year moving average and (b) annual data with 3-year moving average divided by the control 3-year moving average's standard deviation. The values have been computed with annual averages and filtered with a 3-year moving average. The latitudes were selected where the corresponding maximum change in the last 10* *years happens, see Fig. 3. The gray dots show the annual average maximum volume of the overturning stream function in the control during the first year.*

134 not sure which “ratio” is meant hear. Isn’t this a simple difference?

> We have rewritten the sentence “However, the signal in the hosing ensemble deviates from the control from the beginning. When normalizing by the temporal standard deviation of the control, the signal-to-noise ratio in density space is 1 to 3 times higher than at 33.8° N in depth space (Fig. B2), indicating that density coordinates provide a clearer separation between the forced response and internal variability”.

135 replace “happens” with “is located”

> Changed.

139 remove last part of sentence: “…, that is when the AMOC…” Due to the shortness of the runs, it remains unclear whether the AMOC response is already fully developed.

> We have replaced it by “to remove the initial years where the signal of the forcing in the AMOC is too small compared to the noise from internal variability”.

143 add “… by 1.4 Sv on decadal average” at the end of this sentence.

> Added.

145 replace “lowest latitudes” with “subtropical region”

> Replaced.

146 there appears to be something missing at the end of the sentence: “…, representing a 7%.” of what?

> We have added “of reduction computed with respect to the average value in the control for the same area” after “7 %”.

148 remove “the” from “…AMOC in the density space” and “… than in the depth space”.

> Removed.

151 Which water mass is simulated in this density range?

> We computed the average densities shown in Fig. AC2.4. This was done for both the hosing and control experiments, since the injected freshwater could modify the position of the density levels, even though the changes between the experiments are relatively small, and the density range 1027.4-1027.6 kg m^{-3} seems to represent the same water masses in both experiments. For that density range the water mass in the 8-30° N range is located approximately between 1000–1600 m. Between 30° N and 50° N, the range increases to 500-1800 m. The lower limit increases slightly until reaching 1500 m at around 63° N, after which there is a jump to 250-600 m. In high latitudes, the water mass corresponds to the Labrador Sea Water (LSW). We selected this range because it encompasses the maximum change, as the LSW in our model can extend to 1027.2 kg m^{-3}.

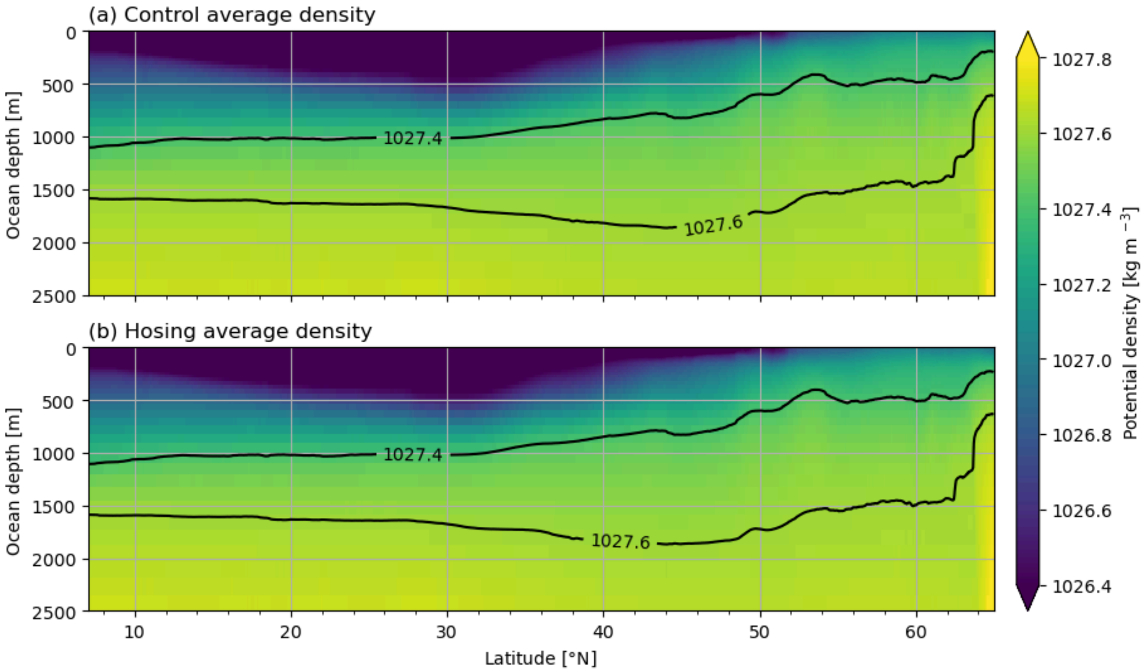


***Figure AC2.4**. Average potential density in the Atlantic basin for (a) control experiment and (b) hosing experiment.*

154 impact “on” not “in”

> We have removed that sentence following a comment from RC1.

154 replace first occurrence of “reduction” with “weakening” to avoid repeat wording in same sentence

> We have removed that sentence following a comment from RC1.

165 add “-“ to “boundary-interior”

> Added.

167 The buoyancy loss could also happen over the path of the NAC in the eastern North Atlantic where the initial densification of the Atlantic water begins.

> We agree that buoyancy loss could happen in other regions. As we want to focus on the role of the Labrador Sea in the reduction of the AMOC, we have replaced the sentence by “This also suggests that the Labrador Sea may play an important role in the AMOC response”, following a recommendation of RC1. Later we show the contributions of OSNAP East and West to the overturning changes, which support a role of western part in the AMOC reduction.

175 rephrase last part of sentence: “… Greenland coast, suggesting an absence of any local impact of the meltwater injection.”

> Done.

178 suggest to rephrase: “This cooling is unlikely driven by export with the Labrador Current, …” and add “s” to “shows” then.

> Changed.

178 SST is not an ideal indicator of the heat content of the Labrador Current. However, I agree that the SST change in the NAC is likely indication of an AMOC weakening.

> We agree that measuring the total heat content of both currents would be more accurate for the discussion. However, the small SST change in the Labrador Current relative to the larger change in the NAC suggests that something more than the Labrador Current is driving the temperature change in the NAC. To be more accurate and acknowledge a potential contribution of the Labrador Current, we have rephrased it adding the word “only”, i.e. “This cooling is unlikely to be driven only by export with the Labrador Current, which shows a weaker SST response than in the NAC region…”.

186 remove “depth” from “mixed layer depth (Fog. 4c).” and remove “,which implies that the impact is relatively high.” Leave such judgement to the reader.

> Done.

188 “, where the mixing also weakens” should be dropped for readability of the sentence.

> Dropped.

189 From the thin contour lines in Figure 4c it is very difficult to see whether the controls run has deep convection in the Irminger and Nordic Seas. Please state this more clearly.

> We have made the contours thicker in Fig. 4 and added a legend to more easily identify the strength of the convection in each area. We also added the sentence “In addition, the Nordic Seas has very shallow climatological mixing due to a positive bias in sea ice (Martin-Martinez et al. 2025, Moreno-Chamarro et al. 2025), which limits the potential impact of the freshwater anomalies”.

190f The deceleration of the gyre by missing deep convection could be countered by a stronger density gradient to the boundary current, which freshened (but also cooled).

> We have added that to the discussion in the next sentence: “The expected deceleration of the SPG in response to the reduction in deep mixing may be counteracted by the acceleration of boundary currents caused by the strong freshwater-driven density gradient”.

193 Please rephrase: “This pattern does not … of the Gulf Stream, as no significant changes in the mean latitudinal position of the current is found when estimating its position from SST and SSH gradients (not shown).”

> Done.

196 “with a slowing of the Gulf Stream” and “temperature and salinity in[!] the NAC”

> Changed.

197 “change of[!] the wind stress” and replace “discards” with “excludes”

> Changed.

198 “, which therefore may only be tied to a reduction…”

> Changed.

230 What is the origin or cause of this warming of the DWBC?

We hypothesise that the warming is linked to a local reduction in the vertical mixing and thus isolation of the DWBC from colder waters. We expose the following in the next sentence: "This may be linked to a local reduction in the vertical mixing (Fig. 5a) which feeds the DWBC". The role of the Labrador Sea may be clear in developing this signal as it does not appear at the Greenland coast. In order to show the reduction of the volume transport in the Labrador Sea, we have computed the OSNAP West and East transport, see Fig. AC2.2. Although the maximum transport is simulated in the OSNAP East transect, as in the observations, the reduction is bigger in OSNAP West (Table AC2.1), which is linked to a reduction of the thermal contribution to the buoyancy forcing (Fig. AC2.2). We have added this table to the main part and the figure to the appendix, and enriched the discussion about Fig. 5):

> "We hypothesize that reduced deep-water mixing in the Labrador Sea could lead to reduced DWBC feeding. This would not only reduce the SPNA AMOC, but also warm the current. This hypothesis is explored further in the water mass transformation analysis in Section 3.2".

32 replace "since" with "from"

Done.

239 How are these numbers (0.4 psu and 1˚C) diagnosed, based on which years?

The quantities represent an estimation of the average difference the whole period. To be more precise we have computed the difference between the mean in both ensembles for the whole period with two decimal, which are now mentioned as follows: "The average salinity reduction for the whole period is of about 0.40 psu, while the temperature reduction is of around 0.94˚ C"

241 replace "since" with "from"

Changed.

246 I agree with the current acceleration but this seems to have a different timescale than the response in temperature and salinity. T and S show a rather rapid adjustment during the first 6 years whereas the Labrador Current exhibits a gradual, cumulative increase in speed. Please explain, eventually speculate, and emphasize this difference.

The two timescales reflect distinct physical mechanisms. The current speed also has a fast initial response, which is consistent with a fast geostrophic adjustment: the freshwater-induced increased cross-current density gradient strengthens the thermal wind balance. The subsequent cumulative increase likely reflects slower ocean-atmosphere feedbacks operating on the subpolar gyre. One plausible pathway follows the adjustment mechanism described by Oltmanns et al. (2020): a progressive freshening of the subpolar region gives rise to a surface cooling, promoting enhanced storminess and wind stress curl, which ultimately reinforces the subpolar gyre circulation in a positive feedback loop. We have added a sentence to the text acknowledging this distinction and its likely cause:

> "It is evident that both currents have also accelerated since the beginning of the freshwater hosing (Fig. 7c). Also, in both currents, while the temperature and salinity response show a rapid adjustment during the first years of hosing, the current acceleration exhibits a more gradual increase. The initial speed-up is consistent with a fast geostrophic response to the enhanced cross-current density gradient, while the longer-term trend likely reflects progressive subpolar gyre adjustments driven by ocean–atmosphere feedbacks (e.g. Oltmanns et al., 2020)"

250 "… 5 years of the simulation (marked by dots)."

> Added.

253 and 258 Why is the NAC responding later than the Norwegian Current? Isn't the Norwegian Current a continuation of the NAC and would thus lag the NAC response? Please explain otherwise this appears to be inconsistent.

> We thank the reviewer for raising this apparent inconsistency. The earlier response of the Norwegian Current relative to the NAC can be explained by the freshwater forcing applied directly in the Nordic Seas, part of which can escape the Greenland Current and reach the Norwegian Current without requiring long-range advection. The NAC response, by contrast, requires the freshwater anomalies to be integrated into the subpolar gyre circulation and transported along a longer interior pathway to the NAC region, explaining the delayed emergence eleven years.

266 replace "gets consistently reduced over time" with "is permanently reduced after year 6"

> Replaced.

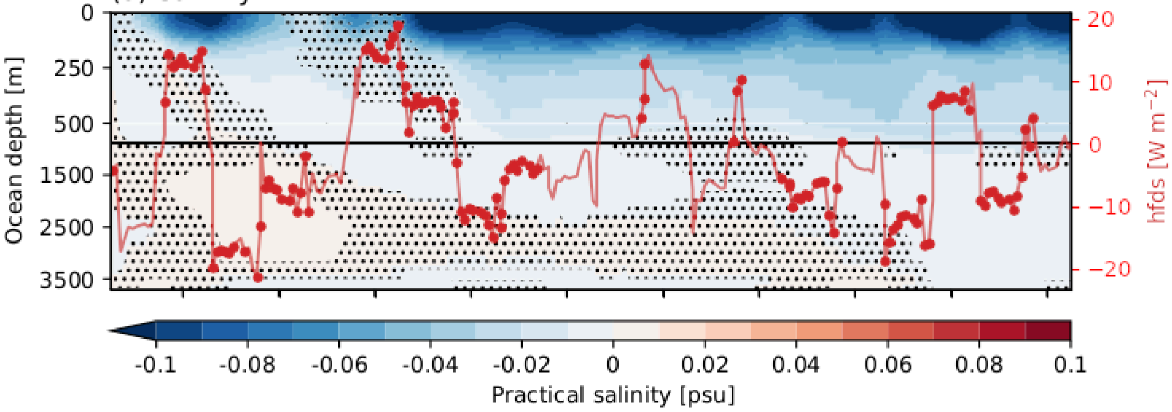
267 Shallow mixed layers in the Nordic and Irminger Seas are not worth mentioning since there is no deep convection there in control either – except for there would be an expectation for deep convection to occur under hosing. This would need to be explained though.

> The average mixing in the control for the interior of the seas defined in Fig. 6 is 478 m for the Labrador, 312 m for the Irminger, and 123 m for the Nordic Seas. We agree that in particular for the Nordic Seas there is no motivation to discuss a potential shallowing of the mix layer. However, the Irminger Sea may expect some reduction, even if the mixing is not as strong as in the Labrador Sea. We have included the climatological values in the main text.

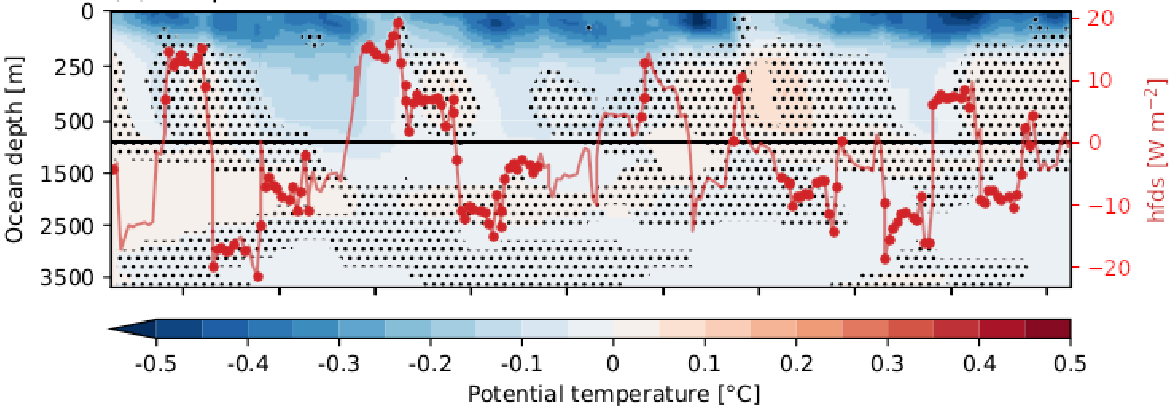
282 Another reason could be internal variability of surface heat fluxes forced by the coupled atmosphere.

> The referee is right. We have produced a time series with the same characteristics (a 12-month moving mean of the spatial average in the Labrador Sea interior) as the downward heat flux at the sea surface (hfds; Fig. AC2.5), overlaying the depth vs. time plots from Fig. 8. We can see that the less cold anomalies around year 6 may be linked to the positive anomalies in hfds. However, we cannot conclude from this that internal variability in local hfds is causing the faster cooling at the beginning. The ensemble-mean signal is negative, but not coherent in sign during the early years, which indicates that at least one member has a positive hfds value that should induce positive temperature anomalies, making temperature signal not coherent across members, which is not what happens. Nevertheless, part of the signal may be explained by this, with other processes counteracting the warming tendency in the member(s) with positive hfds values. We have included a sentence before the mentioned sentence: "Although heat flux loss to the atmosphere could contribute, it does not seem to dominate the faster response of surface temperature in the Labrador Sea interior (not shown). We hypothesize that the fast temperature drop arises because lateral advection of heat (and thus temperature anomalies) is more efficient than the advection of salt/freshwater anomalies."

(a) Salinity



(b) Temperature



(c) Density

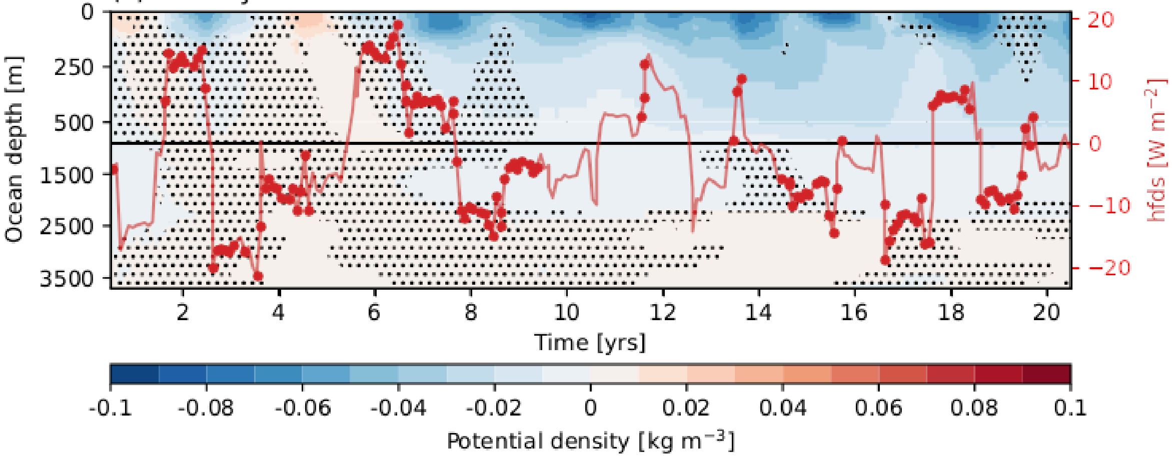


***Figure AC2.5**. Depth vs time Labrador Sea interior ensemble-mean anomalies in (a) salinity, (b) temperature and (c) density as computed between the hosing and the reference control. Anomalies that are not fully consistent in sign between the three ensemble members are masked with dots to highlight the visibility of the fully consistent ones. The Labrador Sea interior ensemble-mean downward heat flux at sea water surface anomalies time series is overlapping in red, positive values indicate warming of the ocean coming from the atmosphere and negative values indicate cooling of the ocean towards the atmosphere. The anomalies are computed as 12-month moving averages.*

314 Suggest to summarize the main process of this teleconnection as explained in Diamond et al (2025). The wave train is mentioned in line 336. This is too late.

We thank the reviewer for this suggestion. Rather than restructuring the narrative, which is organised by hemisphere and variable, we have added a brief sentence immediately after the sentence in line 314 to hint at the wave train mechanism and direct the reader to the later discussion where it is described in more detail.

> "Similar positive temperature anomalies in the Amundsen Sea are described by Diamond et al. (2025) in an analysis focused on NAHosMIP runs. These remote Southern Hemisphere anomalies are consistent with poleward-propagating Rossby wave train forced by anomalous tropical heating associated with the AMOC weakening, a mechanism that is discussed further down in the context of the sea level pressure response."

327ff It is not clear to me how this teleconnection works. As just mentioned, please provide some more insight here likely to be summarized from Diamond et al. (2025).

The teleconnection described by Diamond et al. (2025) may explain the regional response over the Amundsen Sea, which is prominent in JJA but not in DJF. In the austral winter, the SLP response over the Southern Hemisphere is more zonally uniform and is consistent with a negative SAM, which tends to weaken the westerlies and partially counterbalance the broad Southern Ocean cooling. We have extended the discussion of the wave-train mechanism from Diamond et al. (2025), but in the relevant paragraph of the JJA section.
Additionally, the regional anomaly over the Amundsen Sea may reflect an early signature of the tropical-Antarctic Rossby wave train described by Diamond et al. (2025), which originates from Central America and propagates eastward into the Southern Ocean, generating regionally asymmetric pressure and temperature anomalies. We note that our analysis focuses on the second decade of hosing, so this signal is at an early stage of development; confirming the wave train would require upper-tropospheric diagnostics for which we do not have the required 3D outputs.

> "*This last change could explain the positive temperature anomaly previously seen over the same region and season, as the warming is located to the west of the sea level pressure anomaly, where warm southward advection is expected. These changes could be induced by a Rosby-wave train originating in Central America and propagating eastward into the Southern Ocean, as described in Diamond et al. (2025), generating sea level pressure and temperature anomalies with regional asymmetry.*

334 replace "previously seen over" with "highlighted above for"

Done.

364 "…significant impacts on the short 20-year time scale of the experiments, like …"

Added.

368 I am missing an item summarizing the key findings using a VHR model. What is specifically enabled by having an eddy-rich ocean and correspondingly well resolved atmosphere? This is the unique selling point of the paper!

We thank the reviewer for this important suggestion. We have expanded the first paragraph after the itemized conclusions to better highlight the specific advantages enabled by the eddy-rich model configuration and their connection to the key results of the paper.

"The fidelity of the spatio-temporal distribution of the injected meltwater fluxes and the model capacity to resolve mesoscale ocean processes are two key aspects of our study, both jointly expected to enhance the realism of the simulated responses. Compared to eddy-parameterised models, the higher ocean resolution has enabled a more realistic representation of the narrow boundary currents that quickly carry the freshwaters southward around Greenland, and of the mesoscale eddies that govern the lateral exchanges with the interior of the Labrador Sea (Georgiou et al., 2020; Schiller-Weiss et al., 2024) while keeping the Irminger Sea interior isolated. In the hosing experiments, the water mass transformation analysis has shown that the AMOC weakening is driven by a thermally-mediated reduction in dense water formation confined to the Labrador Sea, with the Irminger and Nordic Seas showing no significant convective response, a regional selectivity that reflects the model's capacity to resolve the eddy-driven boundary-interior exchanges that control how the injected freshwater reaches the deep convection region. Further south, the eddy-rich configuration enables a more accurate Gulf Stream separation, as shown in other studies (Marzocchi et al., 2015; Moreno-Chamarro et al., 2021; Frigola et al., 2025), and sharper frontal gradients, with important implications for the responses in the northward heat and salinity transports and the atmospheric circulation aloft ."

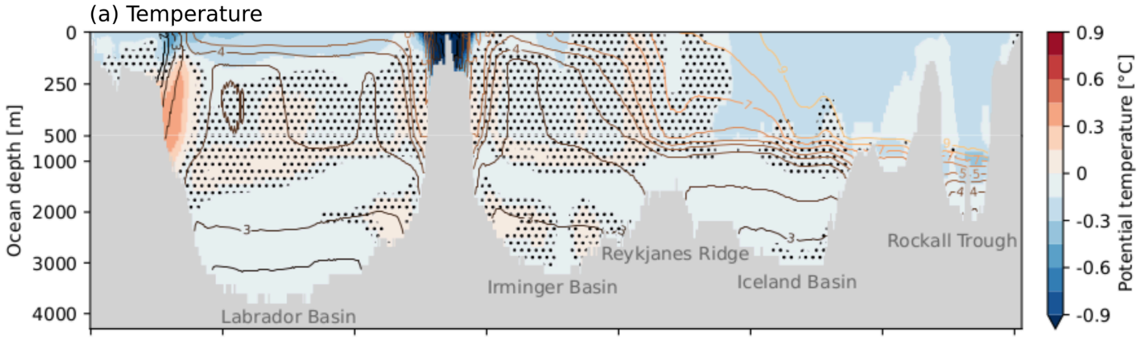
397 (Appendix A) "…, many global 3D temperature (but not SST) data files …" I assume here that SST was available and that this is the reason why your analysis is mostly surface focused, right?

Yes. We have SST data available for the entire globe, as well as 3D temperature data for a small region in the SPNA, as 3D data for this region was preprocessed and saved separately before the loss. Therefore, the OSNAP section in Fig. 5c and the vertical profiles of the Labrador Sea in Fig. 8b are based on the model output temperatures, rather than estimations. However, this limited our potential for analysis in other areas. We do not want to rely heavily on the estimated temperatures, as they may introduce errors that could lead to incorrect conclusions. For example, we would have liked to compute the heat transport by the AMOC, but we cannot proceed due to this important data constraint.

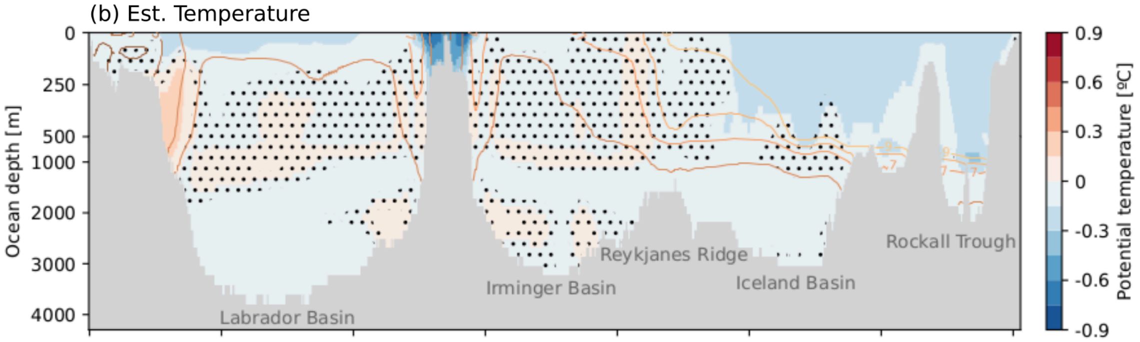
407 "… that the intensity of the anomalies … anomalies are greater." I do not understand this sentence. Please rephrase. Are these two different anomalies? What is smaller, what is greater?

We have rephrased it to:
"However, a test in the OSNAP section (not shown) shows that the estimated temperature anomalies tend to be weaker than the true model outputs. Nevertheless, the sign and significance of the anomalies remain consistent in both cases".

See the values of the comparison of model output (Fig. 4c) and estimated temperature below in Fig. AC2.6.



(a) Temperature



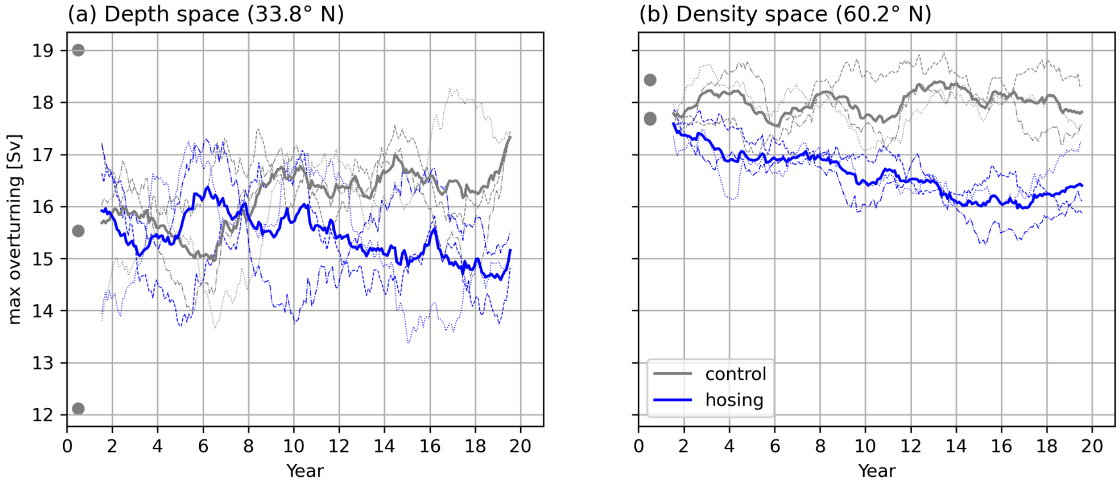
(b) Est. Temperature

***Figure AC2.6**. Response to the hosing in the last 10 simulated years for the (a) model output potential temperature and (b) estimated potential temperature across the OSNAP section (see red lines in Fig. 4); filled coloured contours represent the ensemble-mean anomalies of the hosing with respect to the reference control and the contour lines the climatology of the control. Non-significant values as identified by the bootstrap methodology are masked with dots to improve the visibility over the significant areas.*

Figures

Figure 2: Please add bold dots for the start values of each of the three experiments in panel (a) to highlight the three different AMOC states chosen.

We have now included the values for AMOC at 45° N in depth space in the description of the methods when mentioning the three states, which are the ones that we took into account for selecting the three members. We think that adding the starting values (first year means) to the figure does not help, as the strong spread at 33.8° N substantially expands the range of the y-axis (See Fig. AC2.7), which hinders the comparability of the timeseries:



(a) Depth space (33.8° N)

(b) Density space (60.2° N)

Figure AC2.7*. Monthly volume overturning streamfunction in (a) depth space at 33.8° N and (b) potential density space at 60.2° N in the Atlantic basin. The solid lines are the ensemble averages, each member and its corresponding control year are plotted with a unique linestyle. The values have been computed with monthly averages and filtered with a 36-month (3-year) moving average. The grey dots denote the starting year average in the control for each of the three members. The latitudes were selected where the corresponding maximum change in the last 10 years happens, see Fig. 3.*

Figure 4 The thin contour lines indicating the control climatology are color coded. Please provide a color legend for this. Alternatively use black lines and provide few labels and an increment (in the caption).

We have updated all the figures with contours to include the legends on one side for a better visualization.

Figure 4d Is the sea ice edge in the hosing experiment really that far south as the stippling indicates? This seems unrealistic.

The sea-ice edge, commonly taken as 15 % of sea ice concentration can be identified by the orange contour line, better seen in Fig. B3 (Fig. B2 in the first version). The far south stippling is produced by small anomalies which produce the difference not being exactly 0.

Figure 7d Showing Irminger and Nordic seas here only makes sense if you also discuss a potential east- and northward migration of the deep convection.

We prefer to keep the values of all the deep convection areas for completeness, irrespective of the signal and significance of the response.

Figure B2 Please increase line width of the contour lines indicating the control state for improved visibility. Similarly for Figure 4 only that in the latter there is less space for this.

We have increased the linewidth of all the contours to be the same in all the figures.