Specific comments

Abstract: given the limitations of the box model (appropriately discussed in the Discussion section, Ln. 321-351) and of the climate model simulation (not yet discussed), the Abstract should at least briefly acknowledge these caveats or perhaps tone down some of the confident language.

The abstract has been changed to:

The Atlantic Meridional Overturning Circulation (AMOC) is weakening in response to global warming, while the Nordic Seas Overturning Circulation (NOC) is projected to strengthen. So far, no causal link has been proposed between these two opposing trends. Using an idealized box model of the Atlantic Ocean and the Nordic Seas and a high-resolution climate model, we propose that a density reduction in the subpolar North Atlantic might weaken the AMOC by reducing the density difference with lighter waters further south, while at the same time strengthening the NOC by increasing the density difference with the heavier waters further north. The box model shows that the NOC initially increases moderately as the AMOC weakens in response to combined global warming and freshwater input, while a tipping point may be reached later if deep convection in the Nordic Seas shuts down and the NOC collapses together with the AMOC. These results are supported by GCM data.

Ln. 30: for completeness, this paragraph should mention that the historical evolution and potential weakening of the AMOC is currently still a subject of scientific debate (whereas a future AMOC decline is fully expected). In particular:

- Worthington, E. L. et al. A 30-year reconstruction of the Atlantic Meridional Overturning Circulation shows no decline. Ocean Sci. 17, 285–299 (2021). https://doi.org/10.5194/os-17-285-2021
- Latif, M., Sun, J., Visbeck, M. et al. Natural variability has dominated Atlantic Meridional Overturning Circulation since 1900. Nat. Clim. Chang. 12, 455–460 (2022). https://doi.org/10.1038/s41558-022-01342-4
- Terhaar, J., Vogt, L. & Foukal, N.P. Atlantic overturning inferred from air-sea heat fluxes indicates no decline since the 1960s. Nat Commun 16, 222 (2025). https://doi.org/10.1038/s41467-024-55297-5

We agree that any historical trend (weakening) of the AMOC is uncertain as a result of limited direct observations and large natural variability. We have added a sentence acknowledging this.

Similarly, the past evolution of the Nordic MOC should be briefly mentioned for context, e.g.

- Rossby, T., Chafik, L., & Houpert, L. (2020). What can hydrography tell us about the strength of the Nordic Seas MOC over the last 70 to 100 years?. Geophysical Research Letters, 47, e2020GL087456. https://doi.org/10.1029/2020GL087456

The sentence in line 48 has been changed to

The NOC is believed to have been stable over the past century (Rossby et al. 2020; Årthun et al. 2023; Larsen et al. 2024). For the future, global climate models predict a strengthening of the NOC (Årthun et al. 2023), in contrast to the AMOC's...

Ln. 36: "the AMOC has demonstrated the capacity to...": since this is referring to model results, care should be taken to use wording that does not imply any directly observed "capacity" of the real-world AMOC

"Capacity" has been changed to *"potential".*

Ln. 48-49: "This increased overturning...": this sentence is too vague, it is unclear where exactly those warm and salty waters appear, and what kind of changes in water mass transformation and circulation the authors have in mind

"This increased overturning takes place as anomalously warm and salty surface waters translate into changes in water mass transformation and horizontal circulation."

Has been changed to

"This increased overturning has been linked to large-scale redistribution of sea surface temperatures, modifying the surface density and thus impacting surface-forced water mass transformation, which increases the strength of the horizontal circulation as well as the zonal density gradient."

Ln. 88: "initial simulation": does this refer to a spin-up simulation to quasi-equilibrium before starting the preindustrial and 2xCO2 experiments? If so, that should be stated, since on Ln. 86 it is only said that "the model is initialized" with certain conditions, which does not sound like a proper spin-up simulation.

We clarified the description of the "initial simulation". The sentence now features a statement on the "100-year spin-up".

"*(…), followed by a spin-up simulation of 100 years into quasi-equilibrium."*

Equation (1): introduce all symbols and the values of the constants

Equation now features a detailed description of all symbols and constants.

"*(…) where $\rho 0$ is the density, which is assumed to be constant at 1025 $kg \cdot m{-3}$, Cp is the specific heat capacity of 4.007 $kJ \cdot kg{-1}K{-1}$, Θ is the potential temperature, v is the velocity and $\bar{}v$ being the mean velocity over the cross section for the full basin depth, (…)"*

Sections 2.2 - 2.3: it could be good to cite Msadek et al. (2013) which seems share a lot of the methodology used here:

- Msadek, R., W. E. Johns, S. G. Yeager, G. Danabasoglu, T. L. Delworth, and A. Rosati, 2013: The Atlantic Meridional Heat Transport at 26.5°N and Its Relationship with the

MOC in the RAPID Array and the GFDL and NCAR Coupled Models. J. Climate, 26, 4335–4356, https://doi.org/10.1175/JCLI-D-12-00081.1.

We have included the citation to our manuscript.

"Similar to Msadek et al. (2013), we compute MHT as follows (…)"

Fig. 1 caption: "winter" in what hemisphere (what months specifically)? And why winter?

We added *"The winter period was chosen because convection is active during that time."*, before "All stacked boxes…". We further added the precise months *(January-March)* to the figure description.

Equation (5): What are the values of the constants in this equation?

We added:

"where γ is the thermal coupling constant (10 kg K^-1 s^-3), C_s is the specific heat capacity (4000 J kg^-1 K^-1), ρ the density of water (1025 kg m^-3) ,and d_i the depth of box i."

Ln. 141: "the light blue arrow flips": specify which arrow exactly, since several of them are light blue

"the light blue arrow" has been changed to "*the arrow pointing from box 2 to box 5"*.

Box model equations (Eqs. 6-21): this collection of equations is a bit hard to parse. I would suggest putting it in an Appendix to save some space in the main text. Also, since each term in each equation belongs to one of several categories (i.e., restoring terms with lambda, overturning terms with m, gyre terms with K, convection terms with c, diffusion terms with eta), it could be beneficial to align these terms vertically so that e.g. all terms involving lambda are vertically aligned, then all overturning terms with m, etc. Lines with e.g. no restoring term (for boxes below the surface) would then just have empty space in the restoring column, if that makes sense.

Maybe this would take up too much space but it would certainly increase the readability of the model equations.

We relocated the equations to the appendix and added

"The full model equations for positive m_0 and m_1, as well as $m_0 > m_1$, can be found in Appendix A."

Ln. 195: "The weights for the salinity and temperature differences are the values of α and β in the linear EOS." This sounds intuitive, but is there really a physical motivation to choose the loss function weights in this way? I would assume that the weights are arbitrary hyperparameters that could in principle be tuned, even if the chosen values (1.7 and 7.9) may happen to give good results.

The section has been changed to

The weights for the salinity and temperature differences are the values of α and β (in units of 10^{-4}*) in the linear EOS, because a difference in salinity impacts the density scaling with β, while the impact of a temperature difference scales with α. While these hyperparameters could potentially be tuned, this choice yields satisfactory results.*

Uncertainties (section 3.1): all values reported in tables 1-3 come without uncertainty estimates/confidence intervals. This makes statements about particular decompositions and “dominant” processes etc. less convincing. If more than one ensemble member has been run for this configuration, the ensemble spread should be stated. If similar runs from other climate models are available (e.g., from the CMIP archive), these could be used to get some sense of the variance in these results. Otherwise, the Discussion section should fully discuss these caveats, not only the limitations of the box model.

We agree that the original manuscript was lacking a more detailed uncertainty discussion of the results presented in table 1-3. Hence, we have added a paragraph to the discussion section of the manuscript:

“Nonetheless, the CM2.6 results are based on a single realization, such that no formal uncertainty estimates or confidence intervals can be provided for the values reported in Tables~1--3. Internal climate variability, which is known to be substantial in the subpolar North Atlantic and the Nordic Seas, may therefore influence both the magnitude of the simulated heat transport changes and their decomposition into dynamic, thermodynamic, and nonlinear components. As a result, statements regarding the relative importance of individual processes should be interpreted as representative of the CM2.6 response rather than as statistically robust estimates of the forced signal. Nevertheless, the consistent dominance of circulation-driven contributions across regions and decomposition frameworks suggests that the underlying physical mechanism identified here is robust within this model configuration. Quantifying the spread of these responses and assessing their generality across models would require multi member ensembles or inter-model comparisons, which are left for future work.”

Ln. 228-229: “Hence, the findings imply an anticorrelation between AMOC and NOC-driven MHT changes under increased CO2 forcing in CM2.6.” Since there are really only single data points presented in these tables (e.g. a decreasing AMOC and an increasing NOC), “anticorrelation” might not be the most accurate term. It would be interesting to test whether the evolution of the AMOC and the NOC are correlated in time on decadal (or even inter-annual) time scales in CM2.6, as well as across hosing/warming experiments in the box model in later sections.

We have revised that statement to replace “anticorrelation” with more precise language describing the contrasting or opposing response in AMOC- and NOC-driven heat transport to increased CO2 forcing in CM2.6.

Lines 238-239 now read: *“Hence, the findings imply opposing responses of AMOC- and NOC-driven MHT changes under increased CO2 forcing in CM2.6.”*

We agree with the reviewer that it would be interesting to investigate the correlation between AMOC and NOC on various time scales (inter-annual to decadal). Nonetheless, we focused our analysis on 20-year periods from both simulations due to computational costs and therefore limited data. To prevent confusion, we added the following statement to the revised manuscript:

Lines 96-98 now read *"Owing to data storage limitations, we had to restrict the analysis to 20-year periods drawn from the last 20 years of each simulation. Nonetheless, this experimental design allows us to diagnose the responses to increased atmospheric CO2 forcing."*

Table 2: The heat budget terms for the subpolar North Atlantic contain a large compensating effect of the dynamic vs. nonlinear term (-0.36PW vs. +0.35PW). Is it thus really possible to interpret that circulation changes "dominate" (Ln. 221 and 228)? It could also be said that nonlinear effects dominate, or that dynamic and nonlinear effects cancel and the thermodynamic effect dominates.

We have revised the paragraph accordingly. It now features a clearer interpretation of the respective terms.

"For the subpolar North Atlantic, the decomposition reveals large but compensating circulation-related contributions: dynamic changes in ocean circulation contribute −0.36 PW, while the nonlinear term contributes +0.35 PW (see table 2). This strong compensation results in a comparatively small net MHT anomaly. The thermodynamic contribution is considerably smaller in magnitude (−0.08 PW). While no single term dominates the net MHT change, the largest individual contributions are associated with circulation-related processes, indicating that changes in ocean circulation and its interaction with temperature anomalies primarily control the response."

Ln. 238: "...the weakening of the overturning circulation is the dominant factor...": I would argue that a 57.5% contribution is not very "dominant", especially since no uncertainty bounds are stated (see comment about uncertainties further above)

We have revised the wording to more accurately reflect the relative contributions and to avoid overinterpretation.

"Investigating changes to MHT entering the subpolar North Atlantic under increased CO2 forcing, the MOC component accounts for 57.5% of the total MHT anomalies, indicating that the weakening of the overturning circulation accounts for the largest individual contribution to the diagnosed reduction in MHT, while the gyre component provides a substantial additional contribution."

Ln. 243: It would be helpful to have a paragraph briefly summarizing the climate model results before starting the box model section.

We have added a concise summary of the climate model results at the end of the respective section to the revised manuscript.

Lines 228-237 now read: *"In summary, the CM2.6 simulations indicate a contrasting response of heat transport in the subpolar North Atlantic and the Nordic Seas under increased atmospheric CO2 forcing. While meridional heat transport into the subpolar North Atlantic decreases, heat transport into the Nordic Seas increases. Decomposition of the MHT anomalies shows that circulation-related processes (dynamic and nonlinear terms) provide the largest individual contributions in both regions, although substantial compensation between these terms occurs, particularly in the subpolar North Atlantic, resulting in a comparatively small net anomaly. Thermodynamic contributions are smaller in magnitude. An overturning–gyre decomposition further indicates that changes in the overturning circulation account for the largest individual contribution to the diagnosed MHT anomalies, with gyre-related changes providing a substantial additional contribution. Together, these results suggest that changes in large-scale circulation, rather than temperature anomalies alone, control the opposing mean responses of heat transport in the two regions under sustained CO2 forcing in CM2.6."*

Ln. 252: "...the imposed warming accelerates convection shutdown due to decreasing surface buoyancy": this is not apparent from Fig. 3a/c/e, it seems that the high- and low-end warming scenarios (red and light blue lines) bracket the response of the intermediate warming scenarios. Additionally, the red high-end warming line seems to behave qualitatively differently from the other scenarios. In any case, the differences due to different warming rates seem to be relatively small.

We agree that the wording is confusing and changed "accelerates convection shutdown" to *"Initiates earlier convection shutdown"*. However, we do not see exactly why the red line behaves qualitatively differently from the others, as it shows the same progression, merely slowed down. The dominating effect of the freshwater hosing is addressed in the text.

Ln. 314: "warming in the Nordic Seas": it appears that Li and Liu (2025) are concerned more with the subpolar gyre warming hole than with Nordic Seas warming? See their Fig. 3a

The part has been changed to

"Li and Liu 2025 compare CMIP ensembles with and without a weakened AMOC and find that the warming hole in the subpolar North Atlantic is directly correlated with a weaker AMOC state; however, their Fig. 3b also reveals statistically significant warming in the Nordic Seas in models with a weakened AMOC."

Ln. 371: "warming and freshening of the northern Atlantic caused by an AMOC slow-down" Since both warming and freshening are actually imposed in the box model, to what extent can one really conclude that the mechanism involves the AMOC itself driving further warming and freshening? Can this feedback contribution be quantified?

We added a plot with the AMOC and hosing contributions to the freshening of the North Atlantic in the appendix, and added

“This feedback is stronger for low hosing rates. For high hosing rates (4 mSv/year), the hosing contribution to the freshening of the North Atlantic is significantly stronger than that of the AMOC (fig. B1).”

After “Due to the weakening of the AMOC, more freshwater accumulates in the northern Atlantic, decreasing the surface density and continuing to steepen the density gradient between the Nordic Seas and the northern Atlantic.” in section 3.2.

Technical corrections

Typesetting: the LaTeX could be improved in a number of places:

- Uses of MHT' and similar "math terms" in the text should be placed inside an inline equation ($...$)
- 2xCO2 should use \times and non-italic letters
- Some citations should not be in parentheses: Ln. 133, 189, 198, 205

All points above have been addressed accordingly.

Ln. 1: "while Nordic..." -> while the Nordic...

We changed the phrasing as written above.

Ln. 2: "increase" -> strengthen

We changed the phrasing as written above.

Ln. 23: "winter temperatures": specify that this refers to (surface) air temperature

"winter temperatures" has been changed to "winter surface air temperatures"

Ln. 33: "above 90% significant": the wording should be improved

The sentence has been changed to "Observations reveal a statistically significant decline in density in the Irminger Sea, exceeding the 90% confidence level and approaching 95%, despite pronounced internal variability."

Ln. 35: "hosing" -> freshwater hosing

We changed the phrasing as suggested above.

Ln. 41: "southern shift" -> southward shift

We changed the phrasing as suggested above.

Ln. 45: "to name a few": this could be removed

We changed the phrasing as suggested above.

Ln. 79-80: "much greater fidelity, ... more realistic": compared to what? Other models of the same generation/resolution?

We added *"(...) compared to coarser resolved climate models, (...)"* to clarify the statement.

Ln. 106-107: "In the following, ...": this sentence is redundant with the previous sentence and could be removed

The sentence has been removed as suggested.

Ln. 127: refer to Fig. 2 after "... eight boxes"

We changed the phrasing as suggested above.

Ln. 185: “The parameter c”: there are actually several parameters c_i

We changed “The parameter c is” to *“the parameters c_0 and c_1 are…”*

Ln. 185: “The parameter c...”: although I can understand the meaning of this sentence, it would be easier to understand when put in equation form (with a large brace containing the value of c_i in the two cases)

The suggestion was implemented precisely as written above.

Ln. 191: “eta is the heat diffusion constant”: it is also responsible for salt diffusion (Eq. 21)

We changed the phrasing to *“eta determines the rate of diffusion.”*

Ln. 214: “surface heat flux”: specify the sign convention

The sign conventions have been specified.

Fig. 3 caption: “Grey lines”: since they are all grey, it is not possible to tell which line belongs to which warming rate.

The grey has been changed to the color of the respective warming rate.

Ln. 261: “because of the nonlinearity in the equation of state”: expand on this a bit, I would assume this refers to the temperature dependence of alpha and the different temperatures in the two regions.

The referred sentence has been changed to

“Warming impacts the density in the northern Atlantic more than in the Nordic Seas because of the temperature-dependent thermal expansion coefficient in the equation of state, causing a nonlinear response of density to temperature changes. This increases the density gradient between the two ocean basins.”

Ln. 264: “salt advection feedback”: the accumulation of freshwater in the northern Atlantic after an AMOC decline is not yet a feedback, it is only one of the two “arrows” in the feedback loop. I would thus remove this parenthesis.

The parentheses have been removed.

Ln. 290: “warming prevents the decline of the NOC”: it only “prevents” it in the warming-only experiments. The NOC does end up declining in the hosing+warming experiments, so a more accurate term could be “delays” or similar.

“prevents” has been changed to *“delays”*

Ln. 304: “Labrador Sea”: it looks like the box covers the Irminger Sea more than it does the Labrador Sea?

“Labrador Sea” has been changed to *“subpolar North Atlantic”*

Ln. 340: “dsicussed”: typo

We changed the phrasing as suggested above.

Ln. 355: “in northern Atlantic” -> in the northern Atlantic

We changed the phrasing as suggested above.

Ln. 358: “5 mSv/year”: should this be 4 mSv/year instead?

We corrected this and thank the reviewer for spotting the mistake!