Dear Editor,

We sincerely thank you for handling our manuscript, “Historical Climate and Future Projection in the North Atlantic and Arctic: Insights from EC-Earth3 High-Resolution Simulations” (EGUSPHERE-2025-2653). We are grateful to the reviewers for their evaluation. We appreciate Reviewer #1’s positive assessment, and thank Reviewer #2 for the constructive and detailed comments that have helped us improve the manuscript.

We have revised the manuscript accordingly and provide a point-by-point response below. We submit both a clean revised manuscript and a version with tracked changes; line numbers in this response refer to the clean revised manuscript.

The main revisions include:
(i) Clarifying the positioning of our diagnostic approach relative to existing methods. While our approach is not a direct implementation of previous methods, we acknowledge that conceptually related approaches have been used in earlier studies. We have therefore removed the term “novel” and revised the manuscript accordingly.
(ii) Revising and strengthening the DMV analysis in Section 4.4, including clearer methodological description and improved interpretation of the results.
(iii) Re-plotting Figure A5 and calculating the DMV using a shallower reference depth to evaluate sensitivity to the reference-depth choice.

Sincerely,
Pasha Karami
(on behalf of all co-authors)

Reply to Second review of ESD manuscript egusphere-2025-2653 **-** Historical Climate and Future Projection in the North Atlantic and Arctic: Insights from EC-Earth3 High-Resolution Simulations by KARAMI et al.

Since in my first review I mainly had remarks on section 4.4, I focused on that part of the paper [and the accompanying text in the abstract / introduction / discussion] in this second review

Line numbers refer to the revised manuscript

Overall impression and assessment
I am happy to see the authors have taken my criticism of this section to heart, and have revised it thoroughly. In particular, the section on their volume-conserving budget for deep water formation is now correct.

I only have minor remaining comments / questions on section 4.4, mostly requests for rephrasing, which should be relatively easy to address.
We are glad that the reviewer is satisfied with the first revised manuscript. We sincerely thank the reviewer for the constructive comments and valuable suggestions. Our point-by-point responses are provided below (our responses are shown in blue and numbers are related to the new revised manuscript).

A final point I want to make is that I disagree with the claim the authors (still) make that they have developed a novel method. From my 1st review it became clear that their approach has been used before. While the paper now actually cites some of these earlier papers, they should in my view more properly acknowledge this by (i) also adding citations to the older literature on this topic [i.e. adding the papers on which Katsman et al 2018 and Sayol et al 2019 build] and (ii) refrain from claiming novelty in this respect.
We agree that previous studies have examined the spatial distribution of vertical mass transport and decomposed the AMOC into contributions from different deep-water formation regions, and we have revised the manuscript accordingly, citing Katsman et al. (2018) and Sayol et al. (2019).

That said, we would like to clarify that the diagnostic approach used here is not a direct implementation of existing methods. Unlike previous approaches that rely on vertical velocity diagnostics, our framework quantifies regional contributions using gateway/cross-section transports derived from horizontal velocity components (u, v). This represents an alternative and independently developed methodology, with the practical advantage that it can be applied to model archives where vertical velocity output is unavailable (applicable to a broader range of model archives).

We have revised the manuscript to reflect this distinction more precisely, avoiding broad claims of novelty while emphasizing that the contribution lies in providing an alternative and practically more widely applicable diagnostic formulation. (Lines 523-530 which is copied below and lines 664-672 of the revised manuscript):

> "To complement the DMV, we introduce a volume-budget diagnostic of deep water formation (DWF) in the Labrador, Irminger, and Greenland Seas. This diagnostic is designed to quantify the regional contributions of these basins to the AMOC. Previous studies have characterized the spatial distribution of vertical mass transport and regional contributions to the AMOC using diagnostics based on vertical velocity fields (e.g., Katsman et al., 2018; Sayol et al., 2019). Here, we adopt an alternative approach, inferring these contributions

from horizontal transports across gateway sections within a volume-budget framework. This yields a framework that does not rely on vertical velocity fields and is therefore applicable even to model output in which this variable is not archived."

We have also replaced the term "index" with "diagnostic" when referring to our deep water formation (DWF) analysis/calculation throughout the manuscript, as "diagnostic" more accurately conveys that the method represents a physically based quantity derived from a volume budget rather than an empirical index (e.g., Lines 523-530).

Section 4.4

1) l.454 Although I appreciate that the authors altered the title of section 4.4 in line with my comment, I find the new version ("Factors associated with the weakening of the AMOC") rather uninformative. Can the authors come up with a phrase that better represents this section's contents?

We have revised the Section 4.4 heading (Line 473) to "Changes in stratification, convection, and deep-water formation contributing to AMOC weakening" so that it more clearly reflects the specific content of the section while maintaining continuity with the preceding discussion of AMOC changes.

On the reference depth / DMV analysis:

2) L.492: the statement that the reference depth in the Greenland Sea [700 m] is motivated by the depth of the Denmark Strait sill does not seem in line with the statement directly above it [l.490] that the reference depth should reflect the renewal depth of water masses, as the former is a geographical feature and the latter is a physical feature. Please clarify this in the text

We have revised the sentence (Lines 487-490 of the revised manuscript):

> "To compute the DMV, a reference depth must be specified. We follow the definition of deep convection adopted by Brodeau and Koenigk (2016), building on Marshall and Schott (1999), ~~which emphasises that the depth threshold should reflect the renewal depth of water masses in each basin~~ which assumes that convection contributes to the AMOC only when it reaches either the depth of the southward-flowing branch (~1000 m) or the depth of the overflow sills (~700 m)."

3) L.493-496 I do not find the presented argumentation for choosing a reference depth of 1000m for the Irminger Sea logical. Following l.492, it would make sense to me to base the (regional) reference depths on the (present-day) stratification – for renewal of dense water volume, convection needs to be able "punch through" the stratified layer that is present in fall. The observations generally show shallower MLDs in the Irminger Sea than in the Labrador Sea, which suggests using 1000m underestimates DMV in the Irm Sea, and thus that a shallower reference depth is more appropriate. The authors continue with the statement that ECEarth reproduces this observation [l.494], i.e that the model's MLs are shallowr in the Irm than the Lab Sea, implying that also in the context of the model simulation h=1000m yields an underestimation of DMV and is thus a suboptimal choice. If EC Earth consistently simulates deeper mixed layers than the observations, using 1000m might be justified, but that is not what is written (and it does not look like that is the case - Response Fig 1). Please clarify the text.

In the revised manuscript, we clarify the rationale for the choice of reference depth and explicitly acknowledge its limitations. We have revised the text as following (Lines 492-498 of the revised manuscript):

> "For the Irminger Sea, given its location, we use the same reference depth as for the Labrador Sea. While mixed layer in the Irminger Sea is generally shallower than 1000 m, deep convection

events exceeding this depth have been reported (Våge et al., 2009; Piron et al., 2017) and are also simulated in EC-Earth3-HR. We acknowledge that the diagnosed DMV depends on the selected reference depth and using a 1000 m threshold in the Irminger Sea may underestimate the absolute DMV. For consistency across basins and to focus on the deepest convection events, we retain this threshold but also repeat the analysis using a shallower reference depth (500 m; Fig. A5)."

To address the reviewer's concern more directly, we now include a sensitivity analysis using a shallower threshold (500 m), presented in the revised manuscript as Fig. A5, which demonstrates that while absolute DMV values increase, the temporal is similar.

4) L.498 - the authorstate the change in DMV in the Lab Sea is "most pronounced", as it declines to zero by 2050. However, in that same sentence it is noted that DMV in the Greenland Sea reaches zero earlier in time, which raises the question what "most pronounced" means? The fastest decline? The largest change in m3 of DMV (which seems ambiguous as m3 depends strongly on the chosen reference depth)

We have revised the text to (Lines 499-502 of the revised manuscript):

"The DMV in EC-Earth3-HR shows notable internal variability across all three basins until the late twentieth century, followed by a marked and persistent decline (Figure 11). The Labrador Sea shows the largest DMV values, with a significant decline after 1980, approaching near-zero values by around 2050. The Greenland Sea also shows a significant decline after 1980, reaching near-zero values earlier, around 2020."

5) L.500-501: this needs to be formulated more precisely: when using h=1000m, one can only make the statement that convection rarely exceeds that depth early in the timeseries and that this becomes even less frequent as time progresses m, mmn,m,m,m,nm,m,n m,m,nn, c; by comparing with panel c one can say the convection is weaker than in the Lab Sea rather than generally "weak"

In the revised manuscript, we now state (Lines 502-504 of the revised manuscript):

"The Irminger Sea has smaller DMV values than the Labrador and Greenland seas for the 1000 m criterion, and shows a more gradual reduction through the twenty-first century."

6) L.504 "confirm robustness" – since A5 shows a different variable I do not think this statement is justified; in my view Fig A5b / A5c show that the convection region is more cone-shaped in the Irm Sea and more cylindrical in the Lab Sea; interesting information (and in line with Response Fig 1) but not a confirmation that the conclusions drawn about the evolution of DMV (Fig 10) are not sensitive to the choice of the reference depth – such a statement should be based on repeating the analysis with only a different h

We have now included a sensitivity analysis using a shallower threshold (500 m), presented in the revised manuscript as Fig. A5. The results show that while absolute DMV values increase, the temporal evolution and long-term decline remain similar. We have updated the text accordingly (Lines 504-508 of the revised manuscript):

"Using a shallower threshold increases the absolute DMV values, but does not change the temporal evolution or the long-term declining tendency. This suggests that the qualitative behavior of DMV is not sensitive to the choice of reference depth (Fig. A5). The reduction in DMV across the three convection regions is consistent with the projected slowdown of AMOC, which also declines markedly after 1980 (Fig. 9)."

7) L.505-510: this comparison implicitly assumes that the chosen reference depths are also appropriate for EC-Earth-SR. Referring to remark 3), in my view it thus needs to be accompanied by some info on that aspect – discussing the evolution of DMV compared to HR only makes sense if this index correctly captures the convection to begin with [l.486]

The depth of maximum AMOC and the depth of the major overflow sills in EC-Earth3-HR and EC-Earth3-SR are very similar. Therefore, we consider the use of identical reference depths in the high- and standard-resolution configurations to be appropriate for comparison.

We agree that the choice of reference depth is important when comparing DMV across model configurations and acknowledge that an alternative approach would be to define regionally or temporally varying thresholds, for example based on stratification or mixed layer characteristics. However, establishing such a framework in a consistent manner across regions, model resolutions, and evolving climate states would require a dedicated methodological assessment and is beyond the scope of this study.

In the revised manuscript, we have clarified this rationale and revised the text (Lines 509-514 of the revised manuscript):

> "When compared to the DMV in EC-Earth3-SR, the DMV in EC-Earth3-HR is larger, in line with the findings of Koenigk et al. (2021), who reported an enhanced DMV with increasing resolution in HighResMIP models. This comparison uses identical reference depths in both model versions, as both configurations have similar depths of the major overflow sills and a comparable depth of maximum AMOC. In EC-Earth3-HR, convection in the Greenland Sea reaches near-zero values first, followed by the Labrador Sea after a delay of about two decades. In contrast, in EC-Earth3-SR, DMV declines earlier in the Labrador Sea and later in the Greenland Sea."

On the DWF analysis:

8) L.l.572 this formulation is not correct, the analysis does not assess **water mass transformation** = diapycnal mixing but **dense waterformation** as the authors define earlier[personal view: because the terms are so alike and thus easily confused, I much prefer sinking / vertical transport for the variable that is analyzed here]

We changed it to "dense water formation".

9) L.584 I think itis worth mentioning that this vertical transport may be partly attributable to waters descending over the Denmark Strait sill [so not necessarily driven by local processes]

We have added the following sentence (Lines 593-595):

> "However, the Irminger Sea DWF should not be interpreted solely as local deep-water formation, as the diagnosed downward transport partly reflects the downstream descent of dense overflow waters crossing the Denmark Strait sill."

<u>Abstract / Introduction / Discussion</u>

The parts of the text referring to Section 4.4 have hardly been altered while the contents of the section it summarizes are very different in the new manuscript version. This needs to be better aligned, in particular rephrasing the following:

- l. 31 the authors they claim to introduce a novel method, I do not think this is appropriate – the DMV analysis builds heavily on the Brodeau and Koenigk (2016) paper, and the DWF analysis is the same as used in Katsman et al (2018) and Sayol et al (2019) - papers that they now cite. The new part is that they apply the latter approach to a climate scenario simulation;

We have removed the word "novel." As noted in our response on page 1, our diagnostic approach is not a direct implementation of existing methods but instead provides a practical alternative framework.

- L.33 is easily misunderstood as "processes in basin X are responsible for… " – this is not a good representation of the analyses in the paper. What is shown / analyzed is how the vertical transports at h=1000m change in various regions in response to climate warming.

We have revised the text (line 31-35 of the revised manuscript):

> "We also introduce a framework to diagnose deep-water formation (DWF) in the Labrador, Irminger, and Greenland Seas and to quantify their regional contributions to the AMOC. Applying this framework, we find that projected DWF weakens across all regions, with the largest reduction in the Labrador Sea, making it the dominant contributor to long-term AMOC weakening. By 2100, diagnosed DWF ceases in the Labrador Sea, compared with declines of 62% in the Greenland Sea and 13% in the Irminger Sea."

- l.103 repeats the claim of a new method

The sentence was revised:

> "Additionally, we introduce a volume-budget framework for diagnosing deep-water formation (DWF) in the North Atlantic Ocean."

- L.645 "However, while the DMV index effectively captures… " – please re-evaluate based on remarks below if this statemen can be made like this

The sentence was revised (Lines 656-657):

> "However, while the DMV index effectively captures the presence and breakdown of deep convection, it does not directly quantify ~~deep water formation rates or~~ the relative importance of individual convection ~~formation~~ sites."

- L.646 "… it does not directly quantify…. " – this sentence stillsuggests convection and dense water formation / sinking are the same thing

The sentence was revised (please see the response above)

- L.649 "indicating that DMV and MLD.." – this is not at all new finding as the authors seem to suggest here; it needs to be rephrased as more certain ["confirming once more … "] + references added

The sentence was revised, a new sentence was added with references (Lines 657-663 of the revised manuscript):

> "Moreover, our analysis suggests that despite DMV's relevance for AMOC variability and decline, approximately two-thirds of the AMOC strength persists even after a complete cessation of deep mixing in the North Atlantic. This confirms that DMV and MLD-based diagnostics should not be interpreted as direct proxies for AMOC strength, consistent with previous work showing that deep convection, sinking of dense waters, and the overturning circulation are dynamically related but not interchangeable diagnostics (Katsman et al 2018, Brüggemann and Katsman 2019)."

- L.651-653: again the claim the method is developed by the authors and novel – which I do not think is correct, see above

The term "novel" was removed

We also updated the whole paragraph implementing the comments above (Lines 664-672).

> "To overcome this limitation and further investigate the drivers of AMOC weakening, we developed a volume-budget diagnostic for deep water formation (DWF). Our method offers an alternative and accessible framework for estimating DWF and quantifying the regional contributions of key DWF sites to the AMOC by using horizontal transports across gateway and cross-section boundaries. Earlier studies infer regional contributions to the AMOC from vertical velocity fields (e.g., Katsman et al., 2018; Sayol et al., 2019), but our approach provides an alternative framework that does not rely on vertical velocity output and is therefore applicable to a broader range of model archives. When applied to future projections, our DWF diagnostic shows that DWF declines across all three basins (Labrador Sea, Greenland Sea, and Irminger Sea) over the twenty-first century. By 2100, DWF in the Labrador Sea undergoes…"

Typos / text suggestions

- l. 478-479 caption refers to fig 9b twice

Corrected

- l.529 Denmark Strait ~~overflow~~ sill

Corresponding sentence is changed.

- L.503 can be more concisely rephrased as "sensitivity analyses using a shallower criterion"

Corresponding sentence is changed.

- L.505 grammatical error - DMV [a volume] cannot be stronger

Corrected

- L.530 please give proper credit to earlier papers emphasizing the difference between convection and siking [see my first review]

We added Spall and Pickart (2001) and Brüggemann and Katsman (2019).

- L.562 a depth-specific

corrected