Article: AMOC weakening across latitude and time in CMIP6 future scenarios

March 2026

The article, "AMOC weakening across latitude and time in CMIP6 future scenarios" examines the evolution of AMOC weakening across latitude and time in a set of CMIP6 models under four different future emission scenarios. The authors analyze the propagation speed of the weakening in both depth and density space, concluding that the latter may provide a more suitable early warning signal. They further find that the propagation speed of a given weakening threshold decreases as the threshold increases, with this decrease being more pronounced in low-emission scenarios than in high-emission scenarios.

The study presents interesting results with clear relevance for the current AMOC observing system. However, I have major concerns regarding the interpretation of the results and their physical and statistical robustness. Only if these concerns are adequately addressed would I consider the manuscript suitable for publication in Ocean Science.

We thank the reviewer for taking the time to assess our manuscript and provide thoughtful and thorough feedback. Their comments have helped us clarify and substantially improve the manuscript.

Major Comments

1. The result that the weakening gradient declines with increasing percentage threshold is interesting. My concerns regarding this result are twofold. First, Fig. 9 illustrates that this behavior is not a general feature across all models. As noted by the authors, this may partly be explained by natural variability. However, since this appears to be one of the main conclusions of the manuscript, it is essential to demonstrate that the result is statistically robust. I therefore suggest repeating the analysis using the ensemble mean for each model in order to reduce the influence of internal variability.

   Thank you for highlighting this. We agree the use of multiple ensemble members is important for improving the robustness of our results as well as informative of broader model behaviour. We have added ensemble members for 5 models, where additional ensemble members for all future scenarios were available. This allows us to analyse the ensemble mean of these models. Analysis in Section 3.3 was adjusted to focus on these 5 models. For three of these models (CanESM5, MIROC6, UKESM1-0-LL) the ensemble mean is similar to the individual ensemble members, indicating internal variability is not such a large factor. For CNRM-CM6-1 there is a wide range of behaviour in the ensemble members associated with internal variability and uses the ensemble mean has yielded clearer results. Three models show a decreasing trend in weakening gradient with increased threshold, while two of the models show either a variable or slightly increasing trend.

   Second, if such a statistically robust relationship indeed exists, the physical mechanism underlying the decline in weakening gradient with increasing percentage threshold should be discussed. At present, the manuscript provides no plausible explanation for this behavior, nor any suggestion for how it might be investigated. While I agree that a comprehensive analysis may be beyond the scope of the present study, the authors should at least outline a possible physical interpretation and suggest directions for future analysis, supported by clear arguments.

   We argue that it is advection of negative density anomalies along the DWBC and interior pathways that determines the southward propagation of AMOC weakening by reducing the west-east density and pressure gradient in the southward flowing deep branch of the AMOC. We support this reasoning by comparing the speed of the weakening gradients (~0.6 cm/s) with wave speeds (~1 m/s). It also follows that as the AMOC weakens, then so too will the propagation of the weakening signal slow. This argument is presented in the Discussion (Lines 510-526)

2. While the mechanism underlying the propagation is stated to be beyond the scope of the present study, the Discussion section should at least provide a more detailed quantitative comparison between the propagation speeds diagnosed in these models and those expected from theoretically proposed mechanisms. Such a comparison would help place the results in a clearer physical context. Lines 453–460 briefly mention this point, but no quantitative interpretation is provided.

   This has been added to the discussion, adding the point made above.

3. The first two to three paragraphs of the Discussion and Conclusion sections substantially overlap. I suggest combining these sections into a single Summary and Discussion section to avoid repetition and improve the overall structure of the manuscript.

   Thank you, we agree there was a large amount of repetition in these sections. We have condensed the conclusions into a single paragraph and combined it into the discussion section.

4. The abstract presents the results in a very formal and quantitative manner. I suggest revising it to focus less on specific numerical values and instead emphasize the main qualitative findings, their interpretation, and their implications for the AMOC observing system.

   We have removed the quantitative elements of the abstract to highlight more the broad findings of the paper.

Minor Comments

- Line 29: What about the uncertainty associated with AMOC tipping? The possibility of AMOC tipping introduces a substantial model spread and should be acknowledged.
    - We agree that the possibility of the AMOC experiencing a tipping point adds to uncertainty of AMOC weakening. We have added an acknowledgement that CMIP6 models may be biased to be too stable and reference recent studies suggesting AMOC tipping is more likely than the CMIP6 dataset suggests (Lines 35-39)
- Line 36: Such a mechanism is not exclusive to climate change forcing. See, for example, *Diagnosing the causes of AMOC slowdown in a coupled model: a cautionary tale*.
    - Added a comment to acknowledge this (Line 48)
- Line 69: The study by Marotzke and Klinger (2000) should be mentioned here.
    - Added reference this this study regarding the advective transport of density anomalies along the DWBC (Line 78). This study is also referenced in the discussion.
- Line 106: I do not understand why the streamfunction is estimated from the mass transport. Models that do not provide ψ typically output vo. Why not calculate the volume transport directly from vo instead of using vmo, which may introduce errors in the presence of a strongly variable density field?
    - We originally chose to use vmo as it includes the bolus velocity and much better compares to the online streamfunctions. Due to concerns Reviewer 2 had about the hybrid use of online and offline streamfunction products, especially when comparing depth-space and density-space streamfunctions, we have adopted the use of vo for calculating the depth-space streamfunctions. For completeness, we show the comparisons between online streamfunctions and the vmo/vo streamfunctions in Appendix A.
- Section 2.3: How does the mean weakening found in this subset of models compare to results from earlier studies using larger ensembles (e.g. Weijer et al., 2020; Fox-Kemper et al., 2021)?

- Performing a quick comparison using the method described in Weijer et al, 2020, we find our model set has a mean weakening of 45%, whereas Weijer et al, 2020 find a mean weakening of 39%, indicating our model set skews a bit towards models with greater AMOC weakening. We have added this to Section 2.3 (Lines 168-172).
- Line 177: Does the standard deviation refer to the spread within each model ensemble?
  - Yes. Adjusted the line to read "The standard deviations of the differences ***in the model ensemble*** reveal that the largest ***model*** spread is not aligned with maximum weakening"
- Line 184–188: Redundant.
  - Removed.
- Line 195: What does "slow" mean in this context? Please quantify.
  - Rephrased to avoid the ambiguous slow and instead mention the timescale to cross the Atlantic for the 5% and 25% threshold lines (Lines 220-223).
- Line 207: Does the spread in years represent inter-model spread?
  - The spread in years represents the increasing difference of timings as threshold increases. This has been specified in the text (Line 235).
- Figure 4: Why are fewer models available for the weak percentage thresholds compared to, for example, the 10% threshold?
  - EC-Earth3 begins the experimental period (starting year 2000) between 5% and 7.5% weaker than in the reference period. As such a weakening gradient cannot be calculated. This has been added to the text (Lines 270-272).
- Line 270: This may be a transient feature. At 30°S the system may still be far from adjustment.
  - We agree it might be. We are not suggesting that the state has finished responding to climate forcing, only that the high northern latitudes response under climate forcing both occurs (where the weakening in the line plots becomes positive) sooner and has greater magnitude by the end of the century than at 30S. We have added a comment about timing to clarify this (Lines 299-301).
- Line 277: Is this behaviour consistent across most models, or is it dominated by one or a few outliers?
  - This is a behaviour consistent across most models, the main exception being ACCESS-ESM1-5. Added a comment stating this (Lines 306-306).
- Figure 6: Why use a different colour scale compared to Fig. 3? Also, why is SSP585 shown twice?
  - The yellow-red-black colourmap was used to more clearly show weakening, rather than show weakening and strengthening as is done with the red-blue colourmap. We agree this is visually confusing so we revert to the red-blue colourmap on all Hovmöller diagrams.
  - Results for ssp585 is reproduced for ease of comparison between future scenarios. This is clarified in the figure caption for Fig 6 and Fig 8.
- Line 333: Why not use the abbreviations EMMM and SMMM?
  - Corrected to use EMMM and SMMM.
- Section 3.3: The section is organised somewhat confusingly. Lines 344–356 first discuss results, and only afterwards describe the models from which these results are obtained. It would be clearer to first introduce the models used, then present the figures, and finally discuss the results they imply.
  - Section 3.3 has been reworked to focus on the ensemble mean behaviour. We have significantly reduced the emphasis on model categorisation. The robustness of categorisation was questioned by Reviewer 2 and considering the results from the ensemble analysis, particularly that the spread of weakening gradients seen in the CNRM-CM6-1 ensemble is of similar magnitude to the spread in weakening gradients for the model set, we agreed it was not robust.

References