Reply to reviewers

We thank the editor, reviewers, and community commenters for the time and effort invested in evaluating our manuscript.

Before responding to the individual comments, we would like to emphasize what we consider to be the central issue raised during the review process. This manuscript was submitted as an invited *Opinion article*. Its purpose is not to present new observational evidence, but to critically evaluate and interpret the existing body of published evidence. Such an interpretation should be judged on whether it is scientifically defensible, whether it accurately represents the published literature, and whether it transparently acknowledges uncertainties and limitations.

We fully accept that scientists may reasonably disagree with our interpretation of the current observational evidence. However, after carefully considering all comments and revisiting the underlying literature, we do not believe that the reviews demonstrate that our interpretation is inconsistent with the available evidence or scientifically indefensible. Rather, many of the comments concern the relative weight that should be assigned to different observational records, the interpretation of statistically non-significant trends, or differing views on how multiple lines of evidence should be synthesized. These are precisely the kinds of scientific questions that Opinion articles are intended to discuss.

A central point of disagreement concerns the interpretation of estimated trends versus statistical significance. Several comments appear to conflate the absence of statistical significance with the absence of a trend. However, these are fundamentally different concepts. Figure 3 consistently displays standard linear trend estimates obtained by applying the same transparent calculation to all time series. The statistical significance merely assesses whether the estimated trend is sufficiently large relative to the variability in the data to reject the null hypothesis of no trend. This is fundamentally different from estimating the trend itself. This distinction is central to our Opinion article, which aims to assess the balance of evidence across multiple observational records rather than draw conclusions from any single dataset in isolation.

We are concerned by comments suggesting that the trends shown in Figure 3 were somehow “created by us” rather than objectively derived from the published observational records. All analyses used to produce Figure 3 were fully documented, and the scripts used to calculate the displayed trends were made publicly available alongside the manuscript to allow readers and reviewers to inspect and reproduce every step of the analysis. The reviewers have not pointed out any errors in our calculations.

The Worthington reconstruction illustrates this point. The published paper concluded that the reconstruction showed no overall trend but did not report a linear trend estimate. The associated PhD thesis clarifies that this conclusion was based on a Mann–Kendall test, which found no statistically significant monotonic trend at the 95% level. Before performing our analysis, we also confirmed with the lead author that we were using the correct dataset corresponding to the published reconstruction. Applying the same linear regression procedure used consistently for all datasets in Figure 3 yields a negative trend that likewise does not reach the conventional 95% significance threshold. We therefore consider our result entirely consistent with the original analysis.

For brevity, to structure our detailed reply we use the summary of the reviewer concerns provided by the editor, cited in italics below.

1. There are concerns about a bias in the used literature towards ones that fit in the narrative. The list of cited literature showing an AMOC time series is incomplete and biased towards decline. Papers that are showing an increase or no trend are discarded for invalid reasons.

We strongly disagree with this statement. Reviewer 1 did not mention any studies we have missed, and we have carefully examined the review papers and additional references suggested by Reviewer 2. These papers cover a broad range of AMOC-related observations and analyses, but they do not all provide directly comparable long-term observations-based reconstructions of basin-wide AMOC strength south of 50°N, which were the stated selection criteria.

On those that do:

- (Sanchez-Franks et al., 2021): This reconstruction was originally included in our figure, but we removed it following correspondence with the lead author, who pointed out that it is detrended by construction and thus does not contain trend information.

- (Willis, 2010) as extended in (Jackson et al., 2022) based on sea level (altimeter) data since 1993: How and whether sea level can be used to reconstruct AMOC is an open question, e.g. Eresania et al. (preprint DOI:10.21203/rs.3.rs-10131592/v1) also use altimeter data to find a strong AMOC decline since 1993. Given the scope of this Opinion article, we chose not to include this additional methodological debate, noting that studies using similar data have reached contrasting conclusions.

- Wett et al. (2023) does satisfy these criteria, providing an observational AMOC reconstruction at 47°N, and we will include it in a revised manuscript.

Some of the other studies cited by the reviewer concern shorter observational records or individual components of the circulation rather than multidecadal basin-wide AMOC reconstructions. For example, Meinen et al. (2013) presents a short (20 months with a later extension to 6 years) direct AMOC record at 34.5°S, while Hummels et al. (2015) examines western-boundary circulation at 11°S rather than a continuous basin-wide reconstruction.

2. The Bayesian statistical approach is compelling, but the underlying assumption - that all lines of evidence are independent from each other – is not met.

We do not use or mention a Bayesian statistical approach in the manuscript. What we do state is that a non-significant weakening trend is not an argument against weakening: “If we evaluate multiple pieces of evidence and three largely independent data sets each find evidence for an 80% significant weakening, the conclusion is *not* that “three studies show there is no weakening”. Rather, these studies taken together reinforce the evidence for an AMOC decline, increasing the confidence level.” That is a fact – and we nowhere stated or assumed “that all lines of evidence are independent”. Rather, we agree that the observational products are not completely statistically independent. But they represent different observational approaches that provide partially independent information and all contribute to our interpretation of the balance of evidence.

3. Figure 3 lacks to show and quantify the uncertainty estimates of the used data in combination with the used smoothing. Some of the original studies show no statistically significant trend in their original work, but are here presented as if they do, and likely would not when the uncertainty would be added. This has been considered misleading.

We do not agree that Figure 3 presents statistically non-significant trends as though they were significant. However, the fact that this impression arose during the review process indicates that we can make this distinction much more explicit in the revised manuscript. We will therefore present the unsmoothed data (which were always used to calculate the trends) and include a table listing all estimated trends together with their corresponding p-values. Our main argument is different (as stated above): namely that even trends that are not statistically significant do add to the evidence of a decline, even if each considered in isolation is not compelling. The dominant uncertainty in many cases stems not from uncertainty in the data but from how representative a certain data series (like South Atlantic salinity) is for AMOC changes, and this uncertainty cannot readily be quantified objectively.

4. The authors are making a quantitative argument without doing the necessary analysis and then framing it as an opinion. An opinion piece is an assessment based on existing literature. I agree with the reviewers that here an extra analysis is applied, and that this analysis is not properly reviewed, but is fundamental to the opinion. This mixes up science and opinion in an unclear manner.

We respectfully disagree with the characterization that we are presenting a new quantitative scientific analysis, apart from standard linear trend calculation. The principal quantitative statement in our conclusions – "Where these indicators give quantitative estimates, they consistently indicate a weakening by roughly 2 ± 1 Sv since the mid-twentieth century" is not derived from our trend calculation but summarizes the published estimates reported in the cited literature.

5. This opinion attributes citations to claims that are not in the papers that are cited.

Regarding the examples given by reviewer 2: (Zhang, 2008) was cited for nicely illustrating that "AMOC slowing leads to a northward shift of the Gulf Stream" in her Figure 2, whereas the mechanism (vorticity conservation) indeed is detailed in their previous (Zhang & Vallis, 2007) paper. We'll be happy to cite both.

We did not state that (Majumder et al., 2016) found "an accelerated weakening of the Atlantic overturning circulation since the 1980s." That conclusion was explicitly attributed to (Zhu et al., 2023) while the subsequent citation of Majumder et al. refers to Zhu et al. stating that "this post-1980s AMOC weakening trend appears consistent with the AMOC observations and reanalyses after the 1990s", for which they cite Majumder. Thanks for pointing out a phrasing that could be misunderstood, which is easily fixed in the revision.

Several references in the Figure 3 caption were cited specifically as sources for the applied time lags, rather than because their time series were plotted. For example, Wett et al. (2023) explicitly conclude that "the 26°N AMOC lags the 47°N AMOC by about one year," which is why the 47°N record was shown with a one-year lead relative to RAPID. Similarly, Stepanov and Haines (2014) and Kanzow et al. (2010) were cited in connection with the respective published lag estimates and corrections. We have revised the caption to make the purpose of these citations explicit.

The citation of Dotto et al. (2025) in relation to a productivity lag was imprecise: the paper discusses propagation timescales of upper-ocean properties into deeper water masses, not delayed productivity responses. We will therefore remove this citation from that statement. The applied lag was intended as an approximate alignment based on the expected propagation of AMOC-related anomalies from the subtropical to the subpolar North Atlantic and the delayed response of the ecological proxy. As no published study explicitly derives a 10-year lag for this specific comparison, we have revised the caption accordingly.

6. The community comments (particularly CC5-CC9) discuss the level of uncertainty in different air-sea flux products and ocean reanalysis products. To me, the arguments used in this discussion are not well reflected in the opinion piece and could change the conclusion.

Indeed, we have discussed these reanalysis products with the authors of these community comments at the EGU Assembly this year, after submitting our Opinion article, and are planning to revise our section of reanalysis data accordingly. The latest reanalysis data as analysed in the *Marine Environment Reanalyses Evaluation Project MER-EP* – both atmospheric and oceanic – strongly support the AMOC weakening we discuss in our opinion piece.

Additional points not included in the editor's summary:

To reviewer 2: We do not "discard" studies. The Fraser and Cunningham reconstruction is shown in our figure, and we merely state that for this type of computation "the available data before ~1950 are too sparse and uncertain to allow reasonably robust conclusions." That is a completely valid assessment in line with what Fraser and Cunningham themselves write in their discussion section: "However, from the 1930s onwards we see qualitative agreement with Caesar et al. (2018), with

mostly a high AMOC until the 1950s, followed by a weakening throughout the 1960s and then a lesser peak around 2000. The disagreement lies between 1900 and 1930, when we found the AMOC increased by around 4 Sv (Figure 3a) whereas Caesar et al. (2018) see a high AMOC throughout. Our uncertainty estimate exceeds 3.6 Sv over most of this period (Figure S6). Thus, although our results do not resolve AMOC weakening over the last century, **they should not be interpreted as evidence to the contrary**." (Emphasis added by us.)

On slab ocean models: The limitations of these models are well established and directly relevant to the interpretation of these results. A slab ocean model has a fixed prescribed transport which thus cannot respond to changes in surface forcing, hence the ocean response to changes in surface flux is purely thermodynamic, which leads to an over-sensitive sea surface temperature response. Also, (He et al., 2022) already start with an unrealistic coupled model run with two 'cold blobs', the usual Irminger Sea one and a second wind-cooled region in the Labrador Sea which contradicts observations (their Fig. 1). In the slab ocean driven by this atmosphere the Irminger Sea cold blob disappears, but the wind-driven cooling from the Labrador Sea extends further, reaching the Irminger Sea given the stronger sensitivity of the slab ocean. If that cooling looks at first sight similar to the AMOC-driven cold blob in the coupled model, that is because a different colour scale is used in the slab ocean figure which amplifies the effect by a factor of 2.5 relative to the scale of the coupled model. We chose to not include such detailed critique of individual papers, since our goal of this Opinion article is not to criticise the work of colleagues but to promote a constructive discussion about the ongoing AMOC slowing.

Re the questionable version of *reanalysis surface fluxes* used in (Terhaar et al., 2025): this indeed was also used by one of us in a recent study (Rahmstorf et al., 2026), since at the time of finalising that paper we were not aware of the problems with these fluxes. We regret that, although for this study that has very limited impact since in the region we analysed the trend of both versions of the fluxes are very similar (Mayer et al., 2023) and the reliable energy-budget based fluxes would only have strengthened our conclusion, while for (Terhaar et al., 2025) the latter fluxes imply an AMOC weakening, in contrast to the conclusion of that paper.

In summary, we see a number of aspects brought up by the reviewers as a useful basis for improving our manuscript in the revisions and we are grateful for those suggestions. However, after carefully revisiting every criticism and the underlying literature, we do not believe that the reviews demonstrate that the central interpretation presented in this Opinion article is unsupported by the available evidence. Rather, there appear to be differences of opinion, the discussion of which is exactly the purpose of opinion articles.

He, C., Clement, A. C., Cane, M. A., Murphy, L. N., Klavans, J. M., & Fenske, T. M. (2022). A North Atlantic Warming Hole Without Ocean Circulation. *Geophysical Research Letters*, *49*(19). https://doi.org/10.1029/2022gl100420

Jackson, L. C., Biastoch, A., Buckley, M. W., Desbruyères, D. G., Frajka-Williams, E., Moat, B., & Robson, J. (2022). The evolution of the North Atlantic Meridional Overturning Circulation since 1980. *Nature Reviews Earth & Environment*. https://doi.org/10.1038/s43017-022-00263-2

Majumder, S., Schmid, C., & Halliwell, G. (2016). An observations and model-based analysis of meridional transports in the South Atlantic. *Journal of Geophysical Research: Oceans*, *121*(8), 5622-5638. https://doi.org/10.1002/2016jc011693

Mayer, J., Haimberger, L., & Mayer, M. (2023). A quantitative assessment of air–sea heat flux trends from ERA5 since 1950 in the North Atlantic basin. *Earth System Dynamics*, *14*(5), 1085-1105. https://doi.org/10.5194/esd-14-1085-2023

Rahmstorf, S., Jendrkowiak, J., Gou, R., Cheng, L., Ruiz-Angulo, A., & Björnsson, H. (2026). Multidecadal Atlantic “Warming Hole” Heat Content Variations Are Caused by Ocean Heat Transport, Not by Surface Fluxes. *Geophysical Research Letters*, *53*(11). https://doi.org/10.1029/2025gl118383

Sanchez-Franks, A., Frajka-Williams, E., Moat, B. I., & Smeed, D. A. (2021). A dynamically based method for estimating the Atlantic meridional overturning circulation at 26° N from satellite altimetry. *Ocean Science*, *17*(5), 1321-1340. https://doi.org/10.5194/os-17-1321-2021

Terhaar, J., Vogt, L., & Foukal, N. P. (2025). Atlantic overturning inferred from air-sea heat fluxes indicates no decline since the 1960s. *Nat Commun*, *16*(1), 222. https://doi.org/10.1038/s41467-024-55297-5

Willis, J. K. (2010). Can in situ floats and satellite altimeters detect long-term changes in Atlantic Ocean overturning? *Geophysical Research Letters*, *37*(6), n/a-n/a. https://doi.org/10.1029/2010gl042372

Zhang, R. (2008). Coherent surface-subsurface fingerprint of the Atlantic meridional overturning circulation. *Geophys. Res. Lett.*, *35*, L20705. https://doi.org/10.1029/2008GL035463

Zhang, R., & Vallis, G. K. (2007). The Role of Bottom Vortex Stretching on the Path of the North Atlantic Western Boundary Current and on the Northern Recirculation Gyre. *Journal of Physical Oceanography*, *37*(8), 2053-2080. https://doi.org/10.1175/jpo3102.1

Zhu, C., Liu, Z., Zhang, S., & Wu, L. (2023). Likely accelerated weakening of Atlantic overturning circulation emerges in optimal salinity fingerprint. *Nat Commun*, *14*(1), 1245. https://doi.org/10.1038/s41467-023-36288-4