National Aeronautics and Space Administration
Goddard Space Flight Center
Goddard Institute for Space Studies
2880 Broadway
New York, NY 10025

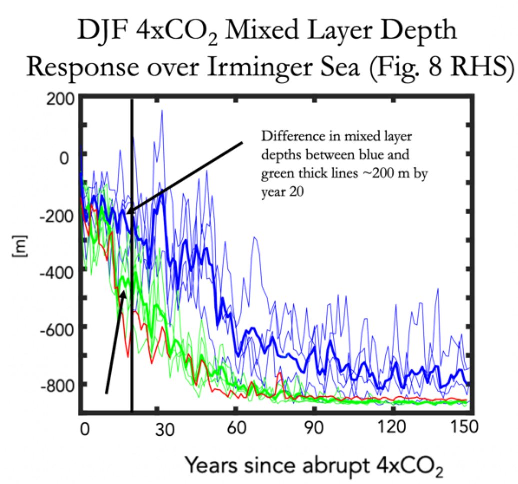
November 17, 2023

Dr. Yu Kosaka
Editor, Journal of Climate

Dear Dr. Kosaka:

Thank you for your evaluation of the reviewers' assessments. We have modified the manuscript after careful consideration of this feedback. A copy of the revised version of the manuscript indicating all changes in red text has also been included as reference material (the JCLI-D-23-0119_revision2_redchanges.pdf attachment) in order to with assist the review process.

First, in response to Reviewer 1 we have removed all references to "nonlinear response", which should address the reviewer's first major concern. While we are simply using the phrase in an identical fashion as in *Zhang et al. (2023)** (as we explain in detail to the reviewer) it is clear that this phrase produces more confusion than clarity and it is not essential to retain in the manuscript. In response Reviewer 1's second major comment, we continue to respectfully disagree with the reviewer's assertion that there is no meaningful mixed layer depth difference that emerges between the NINT and LINOZ ensembles before year 20 since CO_2 is abruptly quadrupled. We discuss this with the reviewer in the appropriate section in this response letter, but, for the sake of clarity to you, we reproduce one key figure below, which we are concerned that the reviewer is misinterpreting. In particular, we disagree with the reviewer's assertion that the blue (NINT) and green (LINOZ) thick lines below show no distinction before year 20 (indicated by the black vertical line):



As the arrows above highlight, there is a clear ensemble mean difference of ~200 m that establishes itself in the mixed layer field by year 20. We are not sure why the reviewer continues to assert that there is no difference between the green and blue lines, but we suspect that this may be because we are focusing on the difference that emerges in the *ensemble mean* (not individual ensemble members') responses. It is possible that our focus on the divergence between ensemble means (not individual ensemble members) may explain the reviewer's source of confusion. We hope that the reviewer now recognizes both of these points, but, in the event that any confusion persists, we have provided this figure and explanation to you directly.

Second, in response to Reviewer 2 we are in complete agreement with her/his push for us to clearly communicate "*throughout the paper that the mechanism only operates under an increased CO_2 background.*" In response, we have extended the new 1xCO_2 coupled experiments by an additional 20 years in order to illustrate the longer-term responses in these simulations, as requested by the reviewer and as is now illustrated in the new response figures R2_1 through R2_5 contained within this response file. We have also expanded our discussion of this result significantly in lines 615-628 and through addition of a new paragraph that we have added to the conclusions (lines 701-713). We hope that the combination of these extended runs, new figures and added discussions to the manuscript satisfy the reviewer's concerns.

Finally, in response to Reviewer 3, we have revised Figure A3 to reflect the 850 hPa wind response, which we agree with the reviewer is more intuitive to show, compared to the 300 hPa winds. This results in no major changes to the manuscript and should address the reviewer's major concerns. We are also extremely grateful to the reviewer for pointing out an error (flipped panels) in Figure A1, which has now been fixed.

Overall, we feel that we have addressed all of the reviewers concerns through our reworking of the manuscript and figures, which now focus more on the robustness and implications of our key findings. We hope that the reviewers feel the same and we look forward to receiving their reviews on the revised manuscript.

Kind regards,

Dr. Clara Orbe
Research Physical Scientist
NASA GISS
clara.orbe@nasa.gov

**Zhang, Xiyue, Darryn W. Waugh, and Clara Orbe. "Dependence of Northern Hemisphere Tropospheric Transport on the Midlatitude Jet under Abrupt CO_2 Increase." Journal of Geophysical Research: Atmospheres 128, no. 13 (2023): e2022JD038454.*

Response to Reviewer 1

Reviewer #1: The revised version of the manuscript is much improved compared to the previous version. Many inaccuracies were corrected, the introduction and discussion read better. I noticed some of my comments were not addressed, hence I decided to give the manuscript an entire new look without focusing too much on the points of disagreement. I still identified two major issues with the manuscript, which I had raised in my previous review already but I am elaborating better here: in particular, I believe they misuse (or use out of context) the definition of 'nonlinear response', and they confuse the surface forcings of AMOC with AMOC response in terms of timescales. More specifically, they make a point that ozone influences drivers of AMOC within a faster timescale (1-5 years) which I agree on, but the AMOC response is delayed by at least a few years to a decade – not occurring at the same time: hence, some of their statements and interpretation of figures need to be readdressed. I recommend the paper to be accepted after these remaining concerns are addressed.

We appreciate the thoughtful feedback from the reviewer and the time she/he has taken to provide a second round of constructive reviews.

Regarding the first remaining issue, we have removed all references to "nonlinear response", which should address the reviewer's concern. While we are simply using the phrase in an identical fashion as was used in the *Zhang et al. (2023)*, it is clear to us that retaining this phrase produces more confusion than clarity. Second, we have expanded our discussion of the AMOC response to account for the reviewer's observation that it seems to occur at a timescale that is slightly lagged, relative to the ozone-induced surface-wind (and mixed layer depth, surface heat flux, etc.) changes. Please see our responses to the relevant comments below.

Major points

Misuse of 'nonlinearity':

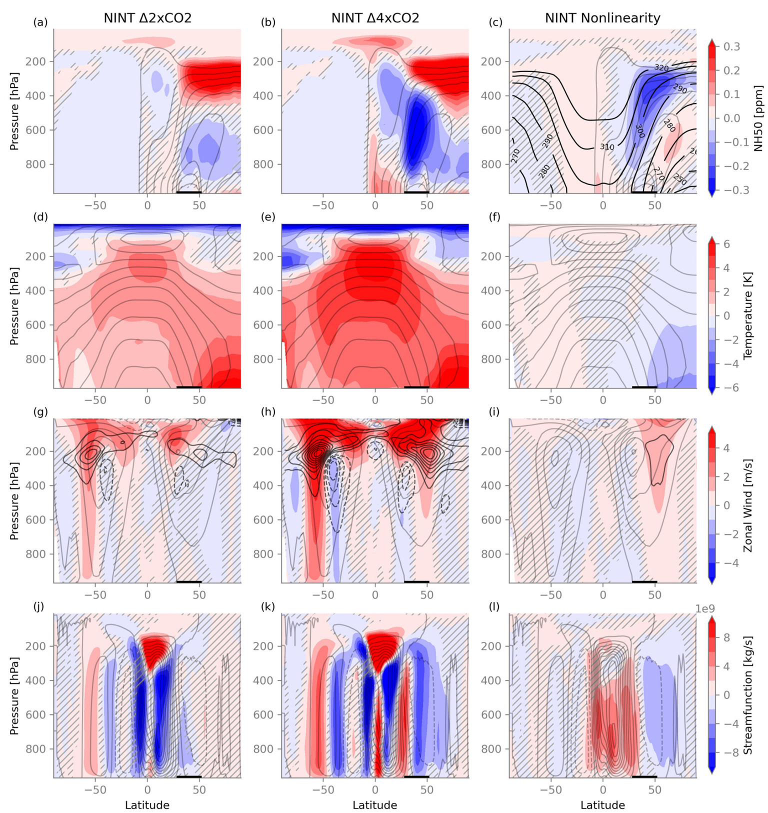
L248-250: I still believe that the authors misuse the term 'nonlinear'. When they say "This difference between OMA and NINT occurs at both 2- and at 4xCO2, resulting in a nonlinearity in the jet (and tracer transport) response in NINT that is not present in the OMA simulations." It implies that that the jet response in OMA is linear then? Hard to say without doing a linear fit, or at least an attempt at a calculation. What the figure shows is: "The jet response in NINT is negative, while in OMA is enhanced and of the same sign".

Later they say: Here "nonlinearity" is defined as the difference 1/2(δ4xCO2)-δ2xCO2, consistent with its use in previous studies (e.g., Chadwick and Good (2013); Mitevski et al. (2021)). This is fine, but you need to quantify it, the reader shouldn't try to do the math from a figure.

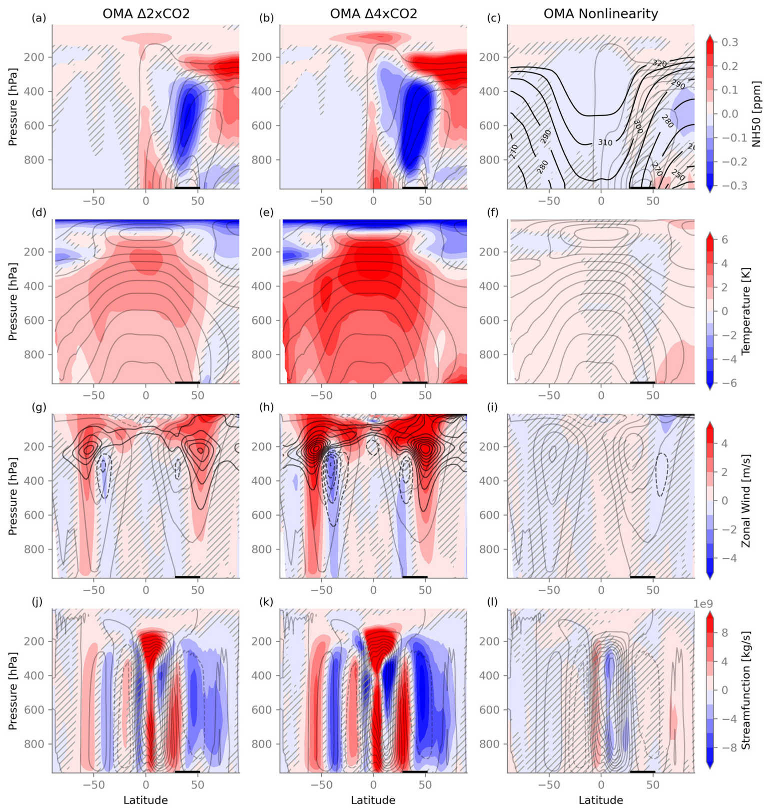
At L254-263 they attempt a quantitative estimate of "AMOC nonlinearity" stating that "the AMOC weakens by ~7 and ~17 SV in the 2- and 4xCO2 simulations, respectively, representing only a very weak (and not statistically significant) nonlinearity in the long-term response of the AMOC". I am not even sure why the authors insist with the nonlinear/linear description of their results. I don't think anyone would think the AMOC would respond linearly to a forcing. In fact, the AMOC is a tipping element in the climate system. The AMOC response to a 4x is not expected to be double than the 2x forcing: oceanic response is very slow and likely if these experiments were extended you would see a recovery of the AMOC in both 2x and 4x in all configurations. If it happens that the AMOC decline at a particular latitude in the 4x is double than 2x, it's just a coincidence in this model.

The reviewer is correct that Figure 1 in our manuscript does not explicitly show the nonlinearity in the wind pattern. To do this, one would need to explicitly compare the bottom panel ($4xCO_2$ response) with $2*2xCO_2$ (the top panel). Thus, we suspect that the primary source of the reviewer's confusion here is that she/he has not seen Figures 3 and 6 in *Zhang et al. (2023)*. While it is true that we could perhaps have pasted these figures directly in our response to the first round of reviews (and we apologize for not doing this), we do note that we already reference these figures in the manuscript.

In particular, Figure 3 from *Zhang et al. (2023)*, reproduced below, shows that there is a large nonlinearity in the tracer (third column, top row), temperature (third column, second row), zonal wind (third column, third row) and Eulerian stream function (third column, fourth row) responses in the NINT simulations. Note that "nonlinearity" in the third column is defined as $4xCO_2-2*(2xCO_2)$.



This nonlinearity in the NINT simulations contrasts with the OMA simulations, for which Figure 6 from that study shows that the nonlinearity is next to negligible:

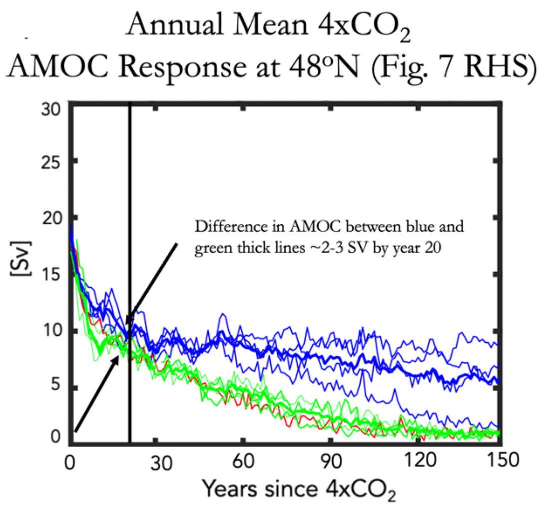


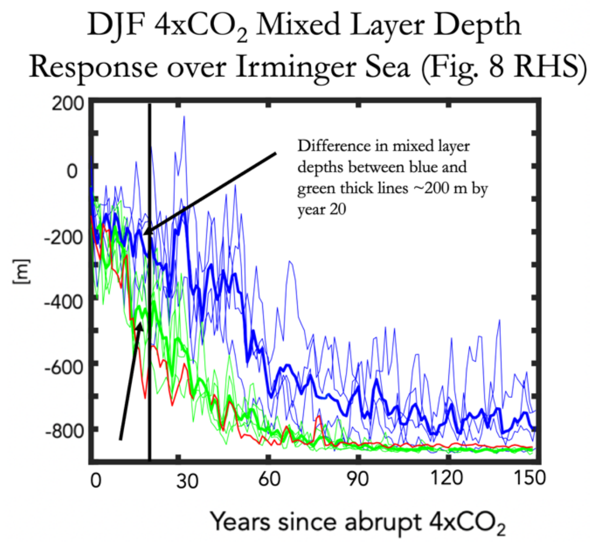
Nonetheless, in order to avoid any further confusion, we have simply chosen to remove all references to "nonlinearity" in the manuscript and have rephrased accordingly. We refer the reviewer to lines 248-254 in the red marked version of the manuscript. This should satisfy the reviewer's concerns.

Impact of ozone on AMOC within the first 20 years:

L457-460: "By comparison, the results in Figure 7 identify an impact of ozone on the AMOC that occurs within the first 20 years of the initial CO2 forcing – that is, over the period during which ozone is also rapidly evolving (Chiodo et al. 2018) and stratospheric temperature gradients are most Impacted by changes in ozone (not aerosols)." In my previous review I already had stated that fig. 7 shows no difference for the first 15 years of the simulation. I reinstate that in therevised manuscript the figure still shows no difference in the first 15 years and this statement is thus inaccurate. The authors in the reply showed this figure below and comment (copied and pasted here). Even from here if you take the first 15 years in this figure, the differences between the lines are indistinguishable. Differences start appearing after 15 years. I think timescale is a crucial point in this paper, and the conclusions the authors draw from this figure are not robust and do not corroborate their following statements at L460-462. Also, quite interestingly the authors do put a vertical bar at year 15 in a following figure (fig. 11), which proves my point.

We respectively push back against the reviewer's assertion that Figure 7 shows no indication of a separation among the ensembles within the initial 20 years of the forcing. Furthermore, we note that the reviewer previously questioned this finding in relation to *Figure 8*, not Figure 7 (*we have reproduced the reviewer's initial comment below in italics to remind the reviewer). Regardless, we argue that both Figures 7 and 8 do show indications of a separation between the ***ensemble means*** before year 20. While the reviewer is certainly correct that the clearest distinction between the ensembles occurs after year 20, there are still indications of a separation between the LINOZ and NINT ensemble means beforehand. In particular, we ask that the reviewer please examine the arrows that we have added to the modified versions of Figures 7 and 8 (right panels from both) shown below, where a vertical line has been added at year 20. Even in the left panel we see that by year 20 there is a difference between the thick blue (NINT) and green (LINOZ) lines of ~2-3 SV, a difference that has already begun to emerge by years ~5-10. We have added a sentence to the manuscript clarifying and quantifying this point (please see new lines 439-440). The distinction between the ensemble means is still clearer in the mixed layer depths (right panel). We have added another sentence to the manuscript reporting on this apparent ~200 m mixed layer depth difference that has emerged between the NINT and LINOZ ensemble means by year 20 (please see new lines 467-468).





We suspect that the source of confusion surrounding these figures is that we are focusing on the ***ensemble means corresponding to LINOZ and NINT***. If one focuses, rather, on the individual ensemble members, as the reviewer might be, there is certainly room for questioning the robustness of this early timescale of divergence. To make this point crystal clear, we have added more emphasis throughout the manuscript that we are focusing on the ensemble mean response. We are grateful to the reviewer for pushing us on this issue, as it emphasizes an important key point – that the signal emerges only after we consider the *ensemble mean* and that single runs would not be sufficient for capturing this response. We refer the reader to lines 438, 439, 457, and 468, where we have added new emphasis to the ensemble mean response.

*We repeat here the original question raised by the reviewer during the first round of revisions: "*Again, the rate of change in mixed layer depths is remarkably similar among all simulations within the first 15-20 years as seen from **fig. 8**".*

Later on, they state: "In particular, the weaker AMOC in the LINOZ and OMA runs is found to be accompanied by a rapid reduction in mixed layer depths, which occur primarily in the Irminger Sea region (55◦N-65◦N, 40◦W-20◦W) (Figure 8). The mixed layer depth differences among the configurations in the Labrador Sea are, by comparison, negligible." By eye, they both look negligible in the first 15 years in fig. 8. I don't think the authors demonstrate this point with sufficiently enough evidence that they can draw any robust conclusions.

As discussed above, we respectfully disagree with the reviewer's claim that there is no sign of a difference between the LINOZ and NINT ensemble means in Figure 8 before year 20. Please see the arrows that have been overlayed in the right panel above. By year 20, there is an ensemble mean difference of ~200 m that has become established. We have now added a sentence to the manuscript commenting on this point. Please see new lines 467-468.

The discussion of fig. 9 seems to be a bit more on point. Here they show surface conditions that would lead to a reduction in AMOC strength. These conditions are stronger in the LINOZ compared to NINT. However, the AMOC response (as shown in fig. 7) estimated by the AMOC strength index, seems to be delayed by about 5 to 10 years, not occurring at the same time, which is expected.

Here we perhaps reach some point of agreement with the reviewer. We agree that the AMOC response shown in Figure 7 may appear somewhat lagged with respect to the more rapid surface variable response (which occur within 5 years). Any apparent lag, however, is subtle, as we emphasize to the reviewer that the AMOC *ensemble mean lines already begin to diverge by year 5-10*. Again, here we refer to the figure shown above, where we have overlaid arrows on the RHS on Figure 7. Note that there are already differences between the AMOC ensemble means occurring by year ~5.

That said, as the reviewer strongly notes her/his prefer to communicate to the reader the potential for a lagged response, we have added a sentence to the manuscript mentioning this possibility. We refer the reviewer to lines 527-530 in the revised manuscript, reproduced below:

"Note that the emergence of these surface changes happens somewhat earlier than the response in the AMOC maxima, which show clearer differences by year ~10 (Fig. 7). As a thorough examination of any lags in the response of the AMOC, relative to the surface, are beyond the scope of this study, this will be examined in future work."

Minor points

L339-340: "warming" at high latitudes seems a bit overstated. Can you quantify the warming and indicate the latitudes at which it occurs, as you do for the cooling? In text, it says cooling in the tropics, looks like cooling at all latitudes except some very small warming at high latitudes in the SH mostly. Also high latitudes are masked actually. Plot is between 60S-60N.
Also: "warmer temperatures over high latitudes (> 40◦N)": seems to me there's still cooling here up to 60°N at the surface.

We agree with the reviewer that the dominant signal is cooling at most latitudes. We have revised the manuscript by removing the reference to warming at high latitudes. Please see lines 337-338 in the revised manuscript.

Response to Reviewer 2

I thank the authors for the careful consideration of my concerns and for performing the extra experiments. I am delighted to see the strong similarities between the new and already performed experiments.

However I think it is important to show the zonal wind response of these new experiments not only for the fast response, which is shown here, but for the total response as well. This would demonstrate to what extend the changes in ozone by itself contribute to the response on the longer timescale. As the authors already noted the changes in the AMOC might not be as large without the accompanying 4xCO_2 changes, therefore the long term response might be significantly different? This would be important to know. While it would not devaluate the proposed mechanism it would be important to clearly communicate throughout the paper that the mechanism only operates under an increased CO_2 background (and possibly to understand why this is the case). At at the moment the reasoning presented is that the ozone changes have a direct fast impact and an indirect long term impact via changes to the AMOC, independent from the CO_2 increase (although the dependence is briefly acknowledged in l.619-621).

We thank the reviewer for her/his thoughtful feedback. We are in complete agreement that it is critical that we clearly communicate "***throughout the paper that the mechanism only operates under an increased CO_2 background.***" In response to the reviewer's first round of comments, we did add a caveat/note of explanation along these lines, but we realize that this was not sufficient and have made this point much clearer in the current version of the manuscript.

In particular, for the reviewer we have now extended the new 1xCO_2 ensemble simulations by 20 years and have accordingly redone Figures R2_1 through R2_5, originally provided for the reviewer as part of the first round of revisions, except now accounting for the difference between years 5-20 (the "fast" response) and years 50-70 of the simulations. We now clarify to the reviewer that the previous version of these figures was showing an *average over years 1-50* as we did not anticipate a two-timescale response for that experiment, given the different (1xCO_2) forcing employed.

We expect that the reviewer may find the new versions of Figures R2_1 through R2_5 quite interesting, because they clearly show that the "fast" response is captured in the runs (absent background 4xCO_2 forcing), but that the longer response is significantly muted in all of the fields and is not statistically significant in several regions. This point is key as it underscores that the ozone feedback on the AMOC has a lasting impact, provided that it is operating in the *presence of background 4xCO_2 forcing*. We refer the reviewer back to the new paragraph that we added in response to the previous revision (the paragraph concluding Section III). We have now modified that paragraph to emphasize this strong dependence on the 4xCO_2 forcing. Please see lines 615-628 in the revised manuscript.

Regarding *why* this is the case, we feel that this is beyond the scope of the study and will be examined in future work, as part of planned analysis to explore not only the sensitivity to the background CO_2 forcing, but also to evaluate the robustness of our findings using the CESM2 model. However, in the meantime, we can speculatively appeal to the oceanographic literature. In particular, the concept of hysteresis may be applicable here as it refers to how a change in forcing (in this case, a stronger surface buoyancy forcing driven by an ozone-induced change in the NAO) beyond a critical threshold can cause the AMOC to weaken to a weak or reversed state beyond which it cannot recover to its original strength, even upon reversal of the forcing. In the context of this manuscript, although the AMOC recovers in neither the NINT nor LINOZ simulations, the differences in their rates of decline may be casually interpreted as reflective of sensitivities around an underlying hysteresis in the model. The key point here is that hysteresis occurs only beyond a certain threshold (i.e., upon the AMOC weakening sufficiently). In our experiments, this can only occur in the presence of increased background CO_2 forcing and may help to explain why our simulations show a long-term ozone feedback only when the AMOC has sufficiently weakened under $4xCO_2$ forcing.

Regarding the occurrence of hysteresis in other climate models, although this behavior has typically been observed only in intermediate complexity or very low resolution climate models, *Jackson et al. (2022)* show that a broad range of CMIP6 (including higher resolution) climate models are capable of exhibiting this behavior (see their Figure 3) in response to a prescribed freshwater forcing. More to the point, we know from two recently published studies (*Orbe et al. (2023), Romanou et al. (2023)*) that the low-top version of the GISS climate model (E2-1-G) does exhibit a hysteresis-type behavior under SSP 2-4.5 forcing. While the analogous behavior has not been reported for the high-top version of the climate model employed in this study, we note that both E2-1-G and E2-2-G use the same ocean model. Therefore, it is highly likely that E2-2-G also exhibits hysteresis.

We have now added a new paragraph in the Conclusions that a) stresses the dependence on background CO_2 forcing and b) briefly speculates on the role that hysteresis plays in this dependency. We refer the reviewer to lines 701-713 in the revised manuscript. As our musings are highly speculative, however, we reserve the right to keep this section as brief and concise as possible. Only rigorous follow-up in future work will adequately address our hypothesis. Overall, however, we are very grateful to the reviewer for suggesting the new experiments and critically questioning the implications of those results. We feel that this feedback has greatly improved the manuscript.

Jackson, Laura Claire, Eduardo Alastrué de Asenjo, Katinka Bellomo, Gokhan Danabasoglu, Helmuth Haak, Aixue Hu, Johann Jungclaus et al. "Understanding AMOC stability: the North Atlantic hosing model intercomparison project." Geoscientific Model Development Discussions 2022 (2022): 1-32.

Orbe, Clara, David Rind, Ron L. Miller, Larissa S. Nazarenko, Anastasia Romanou, Jeffrey Jonas, Gary L. Russell, Maxwell Kelley, and Gavin A. Schmidt. "Atmospheric Response to a Collapse of the North Atlantic Circulation Under A Mid-Range Future Climate Scenario: A Regime Shift in Northern Hemisphere Dynamics." Journal of Climate 36, no. 19 (2023): 6669-6693.

Romanou, Anastasia, David Rind, Jeff Jonas, Ron Miller, Maxwell Kelley, Gary Russell, Clara Orbe, Larissa Nazarenko, Rebecca Latto, and Gavin A. Schmidt. "Stochastic Bifurcation of the North Atlantic Circulation Under A Mid-Range Future Climate Scenario With The NASA-GISS ModelE." Journal of Climate (2023): 1-49.

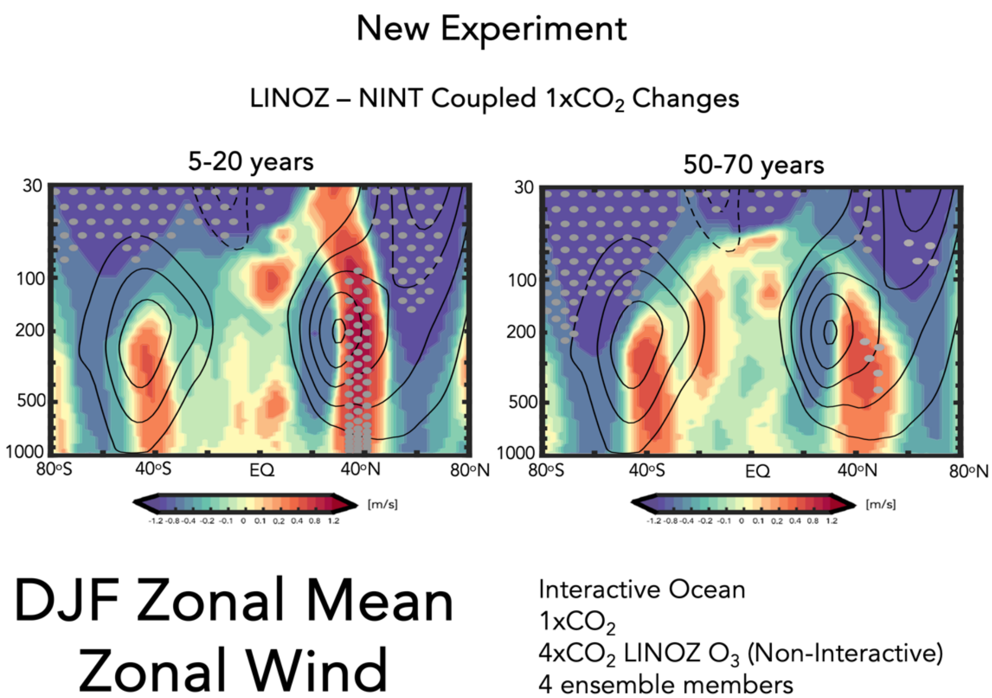


Figure R2_1_ Revison2: The DJF 4-member ensemble LINOZ-NINT differences between the 1xCO_2 coupled atmosphere-ocean runs over years 5-20 (left) and years 50-70 (right). The zonal mean zonal wind response is shown.

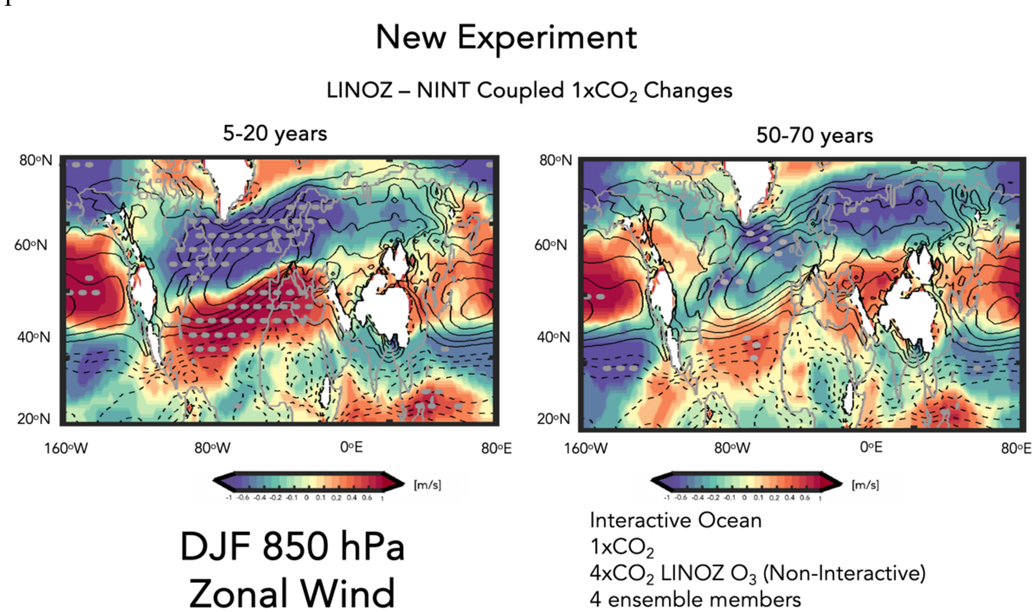


Figure R2_2_Revison2: Same as in Figure R1_1_Revision2, except showing the 850 hPa zonal winds.

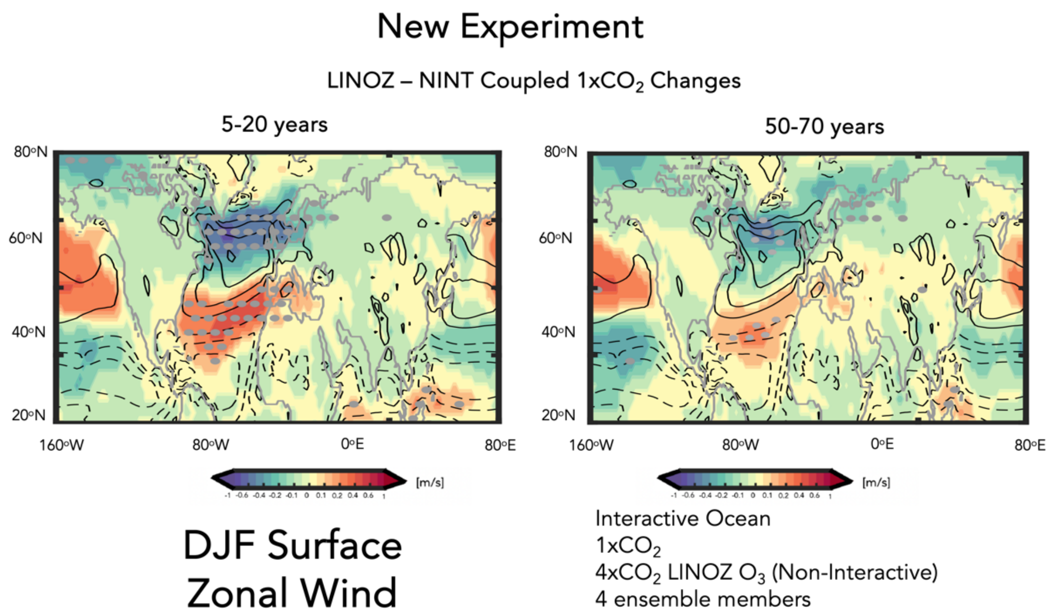


Figure R2_3_Revison2: Same as in Figure R1_1_Revision2, except showing the surface zonal winds.

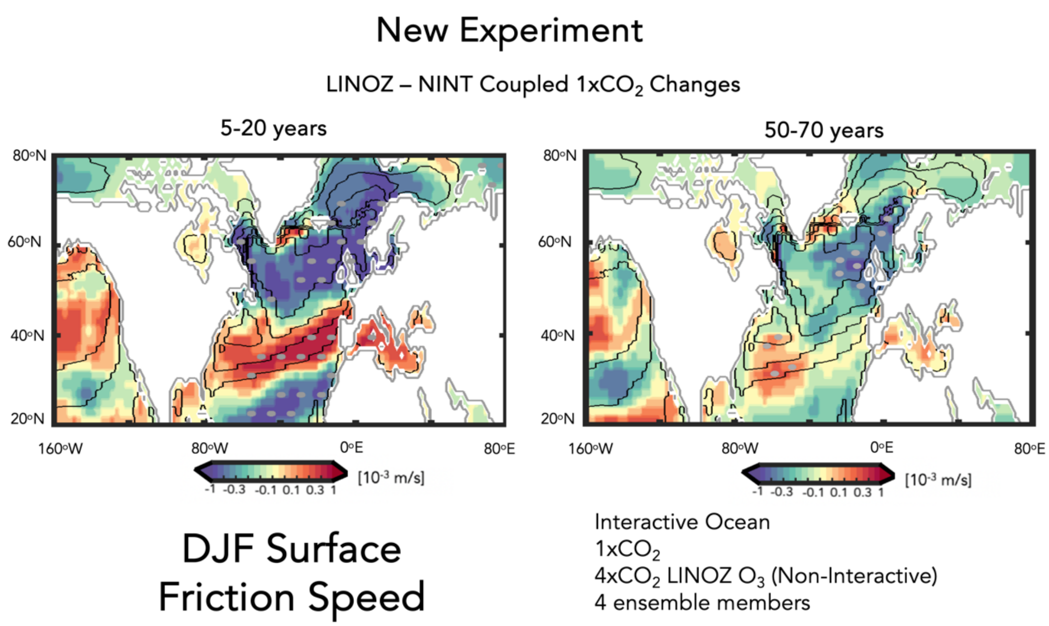


Figure R2_4_Revison2: Same as in Figure R1_1_Revision2, except showing the surface friction speeds.

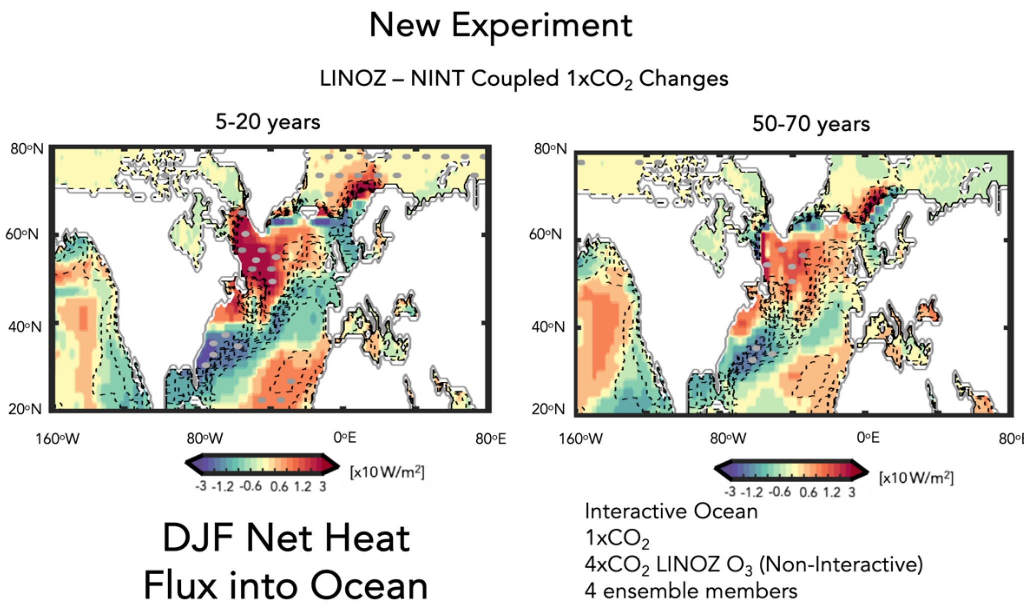


Figure R2_5_Revison2: Same as in Figure R1_1_Revision2, except showing the net heat flux into the ocean.

Lastly, I'd like to point out that not all revisions were highlighted in red, as seen in line 419. This added a layer of complexity to the evaluation process.

We apologize for not accurately highlighting all of our changes and we are sorry that this has unintentionally caused more work for the reviewer. We have taken more care during this round to ensure that all changes have been appropriately highlighted. We very much value the reviewer's thoughtful feedback and are grateful for the time she/he has spent evaluating the manuscript.

Response to Reviewer 3

Reviewer #3: I thank the authors for their thoughtful consideration and attention to the reviewer comments. The authors' revised approach in addressing the regional character of the jet changes is much more accurate and addresses many of my previous comments. I respect the complications that arise from examining this issue in other models, and I'm glad to hear that the authors are pursuing similar experiments with a different model for a future study.

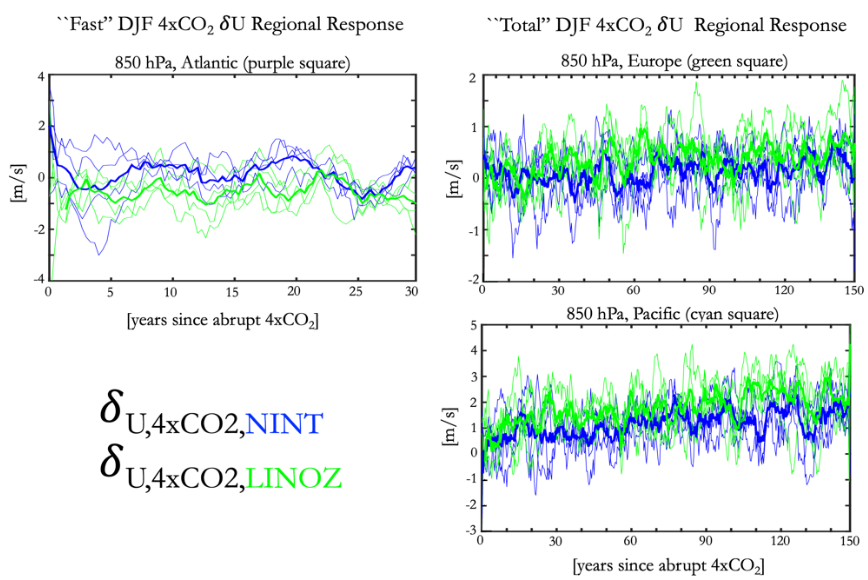
I think the authors need to clarify several additional points in the manuscript (mostly about the new Fig. A3), but once these issues are addressed, I'm happy to recommend publication.

We thank the reviewer for her/his generally positive feedback.

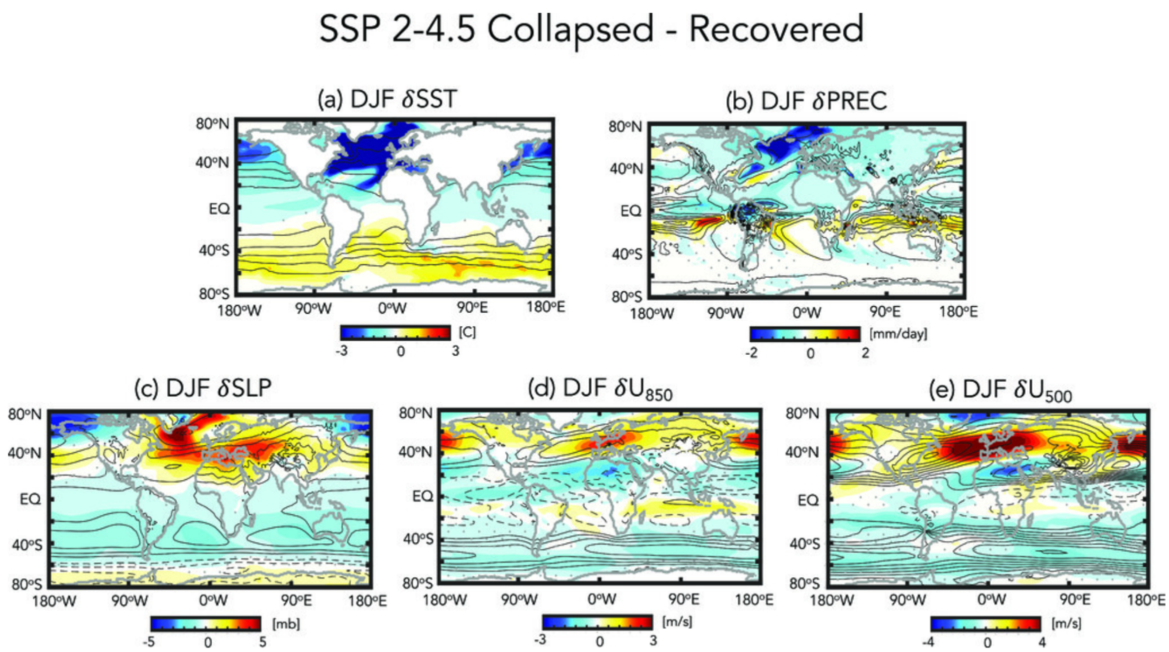
Figure A3: I thank the authors for adding this figure, but I have a few questions about it, mostly in terms of the methodology:

1. Why is 300 hPa U used in this figure, as opposed to 850 hPa U in the figures in the main text that examine the regional jet response?

Our apologies for causing any confusion by showing 300 hPa, as opposed to 850 hPa. The reason why we were showing 300 hPa was in order to highlight the barotropic nature of the zonal wind changes. We felt that it is more convincing to the reader to see that the changes occurring at 850 hPa are linked to changes at 300 hPa, as the latter are more obviously connected to ozone-induced changes in lower stratospheric temperature gradients. We realize now, however, that this is confusing, and so we now replaced Appendix Figure 3a with the new version shown below. Please note that we have also extended the x-axes in panels (b) and (c) to include years > 100.



Comparing the above with the original version of Fig. A3, we see that the regional responses are very similar between 300 hPa and 850 hPa. Perhaps the acceleration over Europe is a bit weaker, compared to 300 hPa, which is consistent with our previous work quantifying the impact of a collapsed AMOC on the zonal winds. In particular, *Orbe et al. (2023)* showed that the eastward extension of the jet over Europe and the poleward shift over the Pacific were both larger at 500 hPa, compared to 850 hPa, especially the response over Europe. This can be seen by comparing panels (d) and (e) in the following figure (Figure 3 from that study):



The weaker response at 850 hPa is also consistent with *Bellomo et al. (2020)*, who did "not find a similarly robust relationship between the eddy-driven jet at 850 hPa and AMOC change [compared to 250 hPa], which could be related to the fact that the North Atlantic Warming Hole has been found to have a relatively weak impact on the eddy-driven jet, compared to other drivers."

To summarize: We have now replaced Fig. A3 with the version above. We have also added a new sentence to the manuscript noting that the responses between the lower and upper troposphere are similar, although the response over Europe is somewhat muted at 850 hPa, which is consistent with *Orbe et al. (2023)*. We refer the reviewer to lines 378-383 in the revised manuscript.

Bellomo, Katinka, Michela Angeloni, Susanna Corti, and Jost von Hardenberg. "Future climate change shaped by inter-model differences in Atlantic meridional overturning circulation response." Nature Communications 12, no. 1 (2021): 3659.

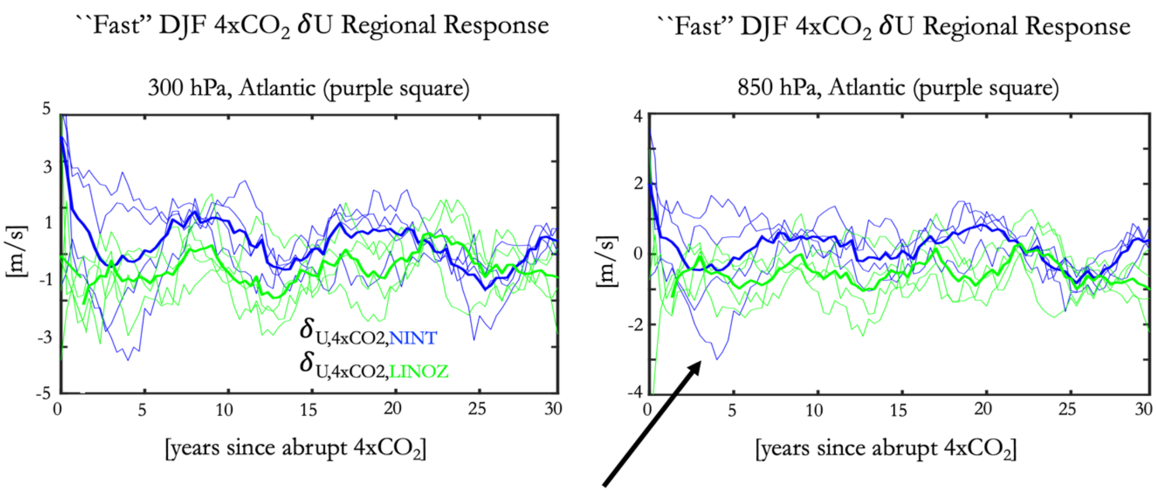
Orbe, Clara, David Rind, Ron L. Miller, Larissa S. Nazarenko, Anastasia Romanou, Jeffrey Jonas, Gary L. Russell, Maxwell Kelley, and Gavin A. Schmidt. "Atmospheric Response to a Collapse of the North Atlantic Circulation Under A Mid-Range Future Climate Scenario: A Regime Shift in Northern Hemisphere Dynamics." Journal of Climate 36, no. 19 (2023): 6669-6693.

2. Why does the x-axis in the right column only extend to Year 100, as the total response is defined from Years 100-150? In the existing figure, you can clearly start to see the divergence of the NINT and LINOZ ensembles by Year 100 for the European sector, but you don't really see this in the Pacific sector.

Again, our apologies for any confusion caused. The timeseries plots are challenging to show because we want to clearly highlight the differences among the ensembles occurring ***both*** on the fast and long timescales (our main motivation for cutting off the axes at 100 years). We agree, though, that cutting it off at 100 years can be confusing, so we've revised this figure to extend to year 150 (see comment and new figure in response to comment #1 above). While it is true that there is a large amount of internal variability, the differences between the ensemble means still show that the zonal winds accelerate over the Pacific and over Europe. The point here is that ensembles of four members (or more) are needed. To communicate this last point, we have added several references to our focus on "ensemble means" throughout the manuscript.

3. From Fig. A3a, it appears that abrupt North Atlantic jet weakening (similar to that in the LINOZ ensemble) actually occurs in two out of the four NINT ensemble members. From Fig. 7, it does appear that one of the NINT members does experience an AMOC weakening similar to LINOZ, but not two. I think the authors need to address this apparent discrepancy. Perhaps this is because they are using 300 hPa wind instead of lower tropospheric wind, which is more relevant for the surface wind stress.

In response to the reviewer's previous comment we have now replaced the old version Figure A3 with a version showing the response at 850 hPa. The new version (850 hPa) of the Atlantic panel is reproduced below (right panel) and compared with the older version (300 hPa) (right panel). Indeed, it does perhaps shows a clearer separation between the single NINT (blue) ensemble member (black arrow), compared to at 300 hPa (left).

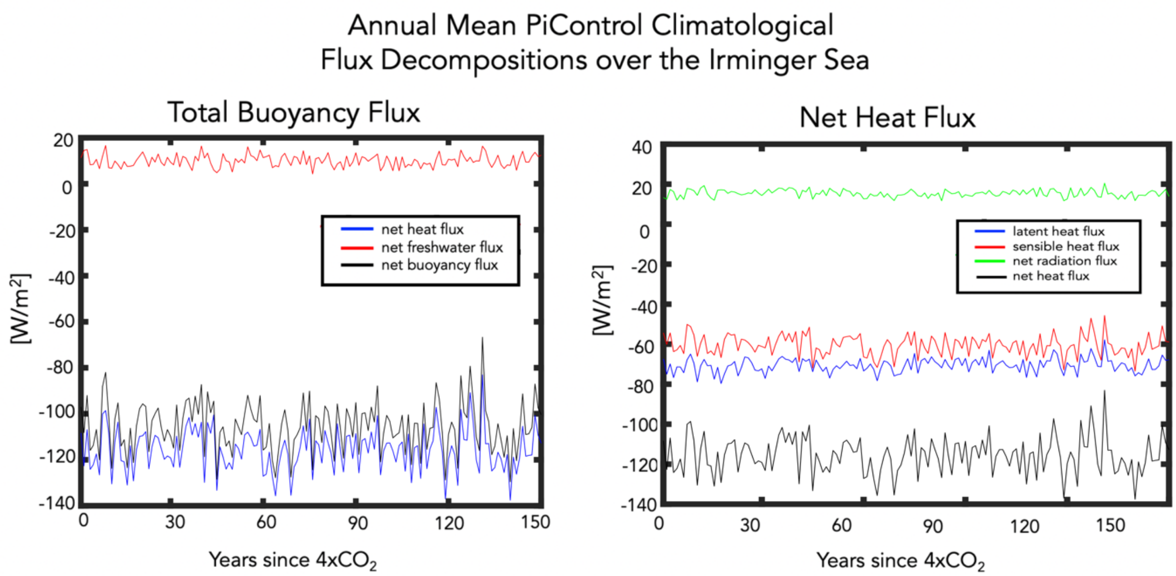


That said, we wish to be extremely clear that our statements apply to the ***ensemble means*** and, hence, we are very hesitant to comment on any causal relationship linking surface wind forcing to the AMOC behavior for a given ensemble member, due to the larger internal variability of the (DJF) system. It is also extremely clear that ensembles are essential for addressing this problem, as the results from certain single abrupt 4xCO$_2$ experiments would not show a significant LINOZ vs. NINT difference. Another review raised a similar concern, and we have modified the language in the manuscript to emphasize that our focus is on the behavior of the different ensemble means. We refer the reviewer to the main manuscript and ask that she/he note the additional references to "ensemble means" (lines 438, 439, 457, and 468) that have been made in the red text.

Finally, another important point to note is that the AMOC is evaluated as the maximum stream function value at 48N and below 900 m. Therefore, there will be processes that may mitigate the surface wind forcing as it is communicated below the ocean's surface. Disentangling those processes needs to be examined in the future work, but is beyond the scope of this present manuscript. We refer the reader to new lines 527-530 in the manuscript.

Lines 217, 220, Fig. A1 caption: References to right and left panels of Fig. A1 are switched both in main text and in caption. Please make sure that the layout of the figure is consistent with references in the main text and caption.

Wow! Thank you for catching this mistake! We are extremely grateful to the reviewer that she/he spotted this error both in the manuscript text and the caption. We have switched the panels in the figure to be consistent with the figure subtitles, figure caption and description in the main text (see below). The discussion of this figure should now be consistent with this revised figure. Please see the revised manuscript.



Lines 325-326: It might be clearer to say black dashed contours.

OK – this has been changed in the caption. Please see the revised manuscript.

Lines 493-496, Fig. 9: These changes in panels a-c are weak and not statistically significant. An alternative viewpoint would be that the warming climate associated with the abrupt change in CO2 is simply causing the changes seen in panels d and e. The Pacific sector is not shown in this figure, but are similar heat flux anomalies seen over the Kuroshio current? If so, then I think the signal in panels d-e is just that associated with a warming climate and is not really associated with the wind anomalies.

The reviewer makes a good point (so long as the comment applies to the top panels in Figure 9, not the bottom ones). While the *anomalous* winds in the LINOZ simulations (bottom panels) are being associated with a stronger reduction in the heat fluxes into the ocean and mixed layer depths, it is certainly possible that increased CO_2 alone can drive the changes in SSTs and heat fluxes observed in the NINT $4xCO_2$ response. We have revised this description in the manuscript to remove the suggestion that the NINT $4xCO_2$ response is solely wind-driven. Specifically, the new sentences now read: "Over the subpolar North Atlantic the weakening of the surface winds leads to a significant reduction in surface friction speed (Fig. 9b, top). At the same time, there is a reduction in mixed layer depths…". Please see lines 496-499 in the revised manuscript.

Lines 567-569: Isn't this shown in Fig. 11f? It looks like the signals begin to diverge around year 40. Maybe the authors are referring to something else here since the text states "not shown", but it would be good to clarify this as it's confusing looking at Fig. 11f.

The sentence in question refers to the specific contribution to freshwater forcing from **sea ice**. Figure 11f shows the total freshwater forcing term. No changes to the manuscript.

Line 588: I don't think Fig. A4 (right) is the correct figure to refer to here, if the authors are trying to show the impact of AMOC on the atmospheric circulation. This figure panel is from the run with fixed SSTs.

Thank you for identifying the mistake. The reference should be to Fig. A4 (**left**). We have corrected this in the manuscript.

Line 605: Fig. A2e

We have revised this section in order to account for the new version of Appendix Figure A3. Please see the revised manuscript.