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

July 28, 2023

Dr. Yu Kosaka
Editor, Journal of Climate

Dear Dr. Kosaka:

Thank you very much for your evaluation of the reviewers' assessments. We have modified the manuscript significantly after 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_revision1_redchanges.pdf attachment) in order to assist the review process.

First, following the recommendation from Reviewers 1 and 2 we have taken care to more faithfully depict the midlatitude jet response in terms of its zonally varying characteristics. We believe that the revised text and new supplementary figure (Fig. A3) illustrating the two-timescale jet response over different regions addresses the concerns raised by the reviewers. We have also performed an entirely new 4-member ensemble of preindustrial CO_2 coupled ocean simulations (using the prescribed 4xCO_2 ozone forcing), per the request from Reviewer 2, and we now discuss the implications of these results in our response to the reviewer and in a new paragraph which we have added to the conclusions section. Finally, we share your concern (along with those raised by the reviewers) that the figures needed much improving. We have modified all colorbars, font sizes, etc., according to the recommendations provided.

Overall, we feel that we have addressed all 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

Response to Reviewer 1

Reviewer #1: This paper document a previously unpublished effect of stratospheric ozone changes on wind circulation, air temperature and surface fluxes, which in turn affect the AMOC, resulting in a stronger AMOC decline following an abrupt 4xCO2 forcing. The authors demonstrate this feedback in idealized model experiments in a hierarchy of climate models that are built to evaluate the impact of stratospheric ozone changes. I think the paper is overall of very good quality, but I had major issues in understanding the results and I believe more a mechanistic understanding of the impacts of ozone changes on air temperature and winds should be provided in the text. I have raised below several issues, which I think the authors should address before the paper can be accepted for publication.

We very much appreciate the thoughtful feedback from the reviewer. We hope that our incorporation of her/his feedback has significantly improved the manuscript.

L78-79: I think Delworth and Zeng (2016) would be a reference to add to those cited here:
****   Delworth, T. L., and F. Zeng, 2016: The Impact of the North Atlantic Oscillation on Climate through Its Influence on the Atlantic Meridional Overturning Circulation. J. Climate, 29, 941-962.***

Good point. We have added this reference. Please see the revised manuscript.

L86-104: These paragraph would better fit in a conclusion or discussion, there is too much text here in the introduction summarizing findings that are yet to be shown in the paper. It doesn't seem to be the right place at this point to anticipate all of the results.

Please see our response to the next comment.

L114-116: This also belongs to a discussion. In the introduction you need to highlight what is missing in the literature, and what you address. Not what you find and how it compares with previous work.

We will respectfully push back against the reviewer's suggestion to move all of this text to the discussion section. Given that the response that we are quantifying is complex (i.e., consisting of fast and slow timescales, spatially heterogenous, dependent on interactive chemistry, etc.) we feel that providing this context helps guide the reader through the development of our findings. In particular, the reviewer herself/himself was confused about the relationship between the stratospheric ozone changes and the surface wind response. We therefore feel that it is beneficial to help prime the reviewer about this (and the following) components of the stratospheric ozone feedback on the AMOC at the outset.

That said, we do agree that the *second* paragraph discussing the role of aerosols can be moved to the discussion section. We have moved this material to the Conclusions and merged it with some of the existing aerosol discussion in that section. Please see Section 4 in the revised manuscript.

L105-107: It should be mentioned that aerosols are also thought to possibly overly enhance AMOC variability in such a way that might be inconsistent with AMOC reconstructions. For example:

*   Zhang, R., Sutton, R., Danabasoglu, G., Kwon, Y.-O., Marsh, R., Yeager, S. G., et al. (2019). A review of the role of the Atlantic Meridional Overturning Circulation in Atlantic Multidecadal Variability and associated climate impacts. Reviews of Geophysics, 57, 316- 375.

*   Zhang, R., and Coauthors, 2013: Have Aerosols Caused the Observed Atlantic Multidecadal Variability?. J. Atmos. Sci., 70, 1135-1144.

* Robson, J., and Coauthors, 2022: The Role of Anthropogenic Aerosol Forcing in the 1850-1985 Strengthening of the AMOC in CMIP6 Historical Simulations. J. Climate, 35, 6843-6863.

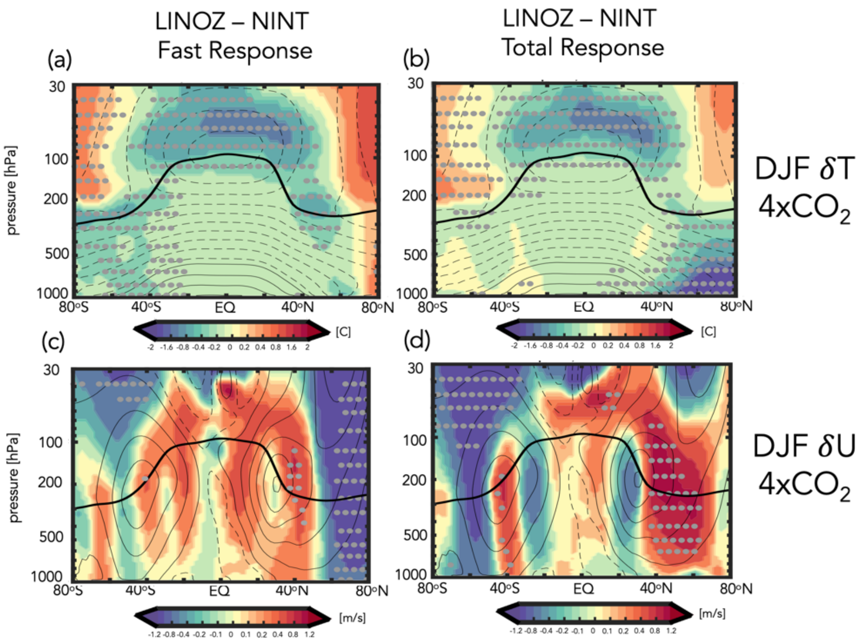
We thank the reviewer for providing these references. We have now included these references in the (newly expanded) aerosol discussion in the Conclusions. Please see the revised manuscript.

L190-197: The choice of using different time frames than CZP2018 is not justified in this work. I am left wondering what it would look like using the same 5-10 and 121-140 averages as in CZS2018. The larger ENSO amplitudes enhancing inter-annual variability that is used as a reason for this choice does not justify the choice of using the years 121-140 instead of 100-150, which are at much longer timescales than ENSO. Why did Zhang et al 2023 and CP2019 use a different approach? Is their approach the same as in the present study?

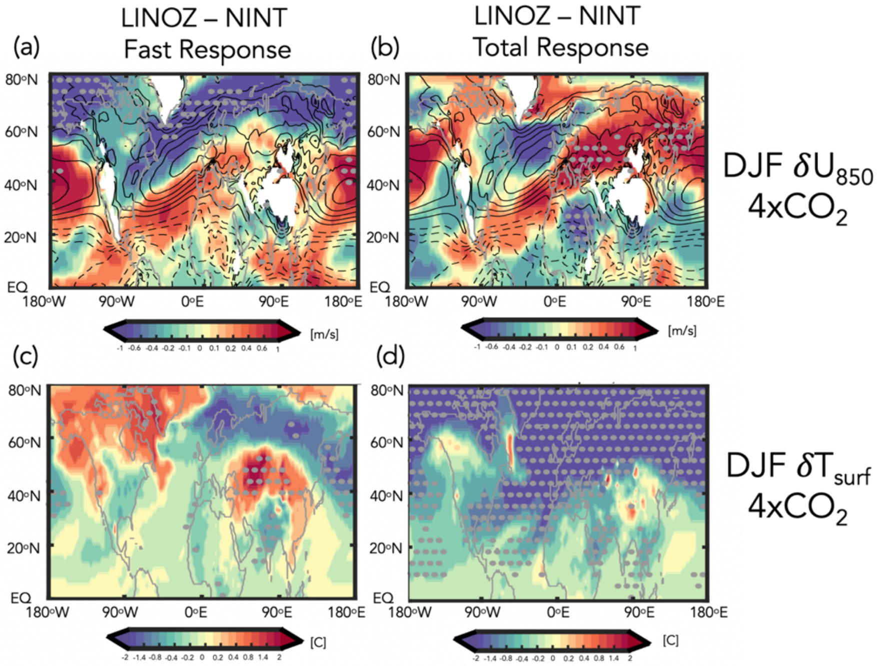
We thank the reviewer for her/his comment, which we address here:

1) "I am left wondering what it would look like using the same 5-10 and 121-140 averages as in CZS2018."

There is what Figures 5 and Figure 6 look like if we use the same time frames from CZP2018.



and



Comparing the above with the figures in the manuscript we find that there are no major differences in the responses. We now mention this explicitly in the text. Please see the revised manuscript.

"The larger ENSO amplitudes enhancing inter-annual variability that is used as a reason for this choice does not justify the choice of using the years 121-140 instead of 100-150, which are at much longer timescales than ENSO."

We use years 5-20 in lieu of 5-10 to account for large internal variability, rationale with which the reviewer seems to agree. The reviewer is more concerned, however, by our use of years 121-140. We argue, however, that the use of 121-140 is much is less conventional than using years 100-150 which is far more standard in the field and as has been applied in many studies (e.g., Grise et al. (2014), Grise and Polvani (2016), Menzel et al. (2019)). We now make this last point clearer in the text. In addition to being less conventional, we also have no physical reason to use their proposed years 121-140 (especially given the similar response shown above). Rather, we prefer using a longer averaging time period, especially when quantifying the (highly variable) NH DJF zonal wind.

"Why did Zhang et al 2023 and CP2019 use a different approach? Is their approach the same as in the present study?"

First, Zhang et al. (2023) and CP2019 do not examine the "fast" response. Second, they define the total (or slow) responses by averaging over the last 50 years of the abrupt (2)4xCO_2 simulations. As discussed earlier, this approach is much more standard in the field and so the question instead becomes "Why did CPZ2018 use years 121-140?" Unfortunately, that averaging choice is not explained in their study. Thus, in the absence of any physical (or other) justification for using years 121-140, we retain our 100-150 year averaging approach.

Fig. 1: the font of the scale on the color bars is too small to read (this problem arises also in all the other figures)

We thank for the reviewer for this comment and completely agree (another reviewer raised the same point). We have made extensive changes to the colorbars and labels in all figures, including increasing font sizes. Please see the revised manuscript.

L226-228: In the NINT northern high latitudes the sign of the change is negative (it's blue), so it's of the opposite sign… not only 'not statistically significant'

Good point. We now mention this in the text. Please see the revised manuscript.

Fig. 2: At which depth is the AMOC strength taken? Also this info is missing in fig. 7. It is also not clear why you have two figures (fig. 2 and 7) that essentially contain the same information. Is fig. 2 also reproduced in another paper (L413)? Why then reproducing the same figure here and not simply referring to the published one?

Our apologies for not including this information! We define AMOC strength as the maximum overturning streamfunction below 900m. We have checked the sensitivity of this calculation (i.e., searching below 500m) and this does not make a difference. This is now mentioned in the figure captions for Figure 2 and 7. We thank the reviewer for drawing attention to this oversight.

L251-252: I don't believe this is an AMOC non-linearity. It looks like feedbacks reactivate the AMOC circulation in this particular run.

By "AMOC non-linearity" we refer to the fact that the response of the AMOC to 4xCO_2 does not equal twice the response of the AMOC to 2xCO_2. The nonlinearity is then equal to the difference $1/2(\delta(4xCO_2))-\delta(2xCO_2)$ for any field of interest. This definition of linearity is consistent with its usage in the following studies:

Mitevski, Ivan, Clara Orbe, Rei Chemke, Larissa Nazarenko, and Lorenzo M. Polvani. "Non-monotonic response of the climate system to abrupt CO2 forcing." *Geophysical research letters* 48, no. 6 (2021): e2020GL090861.

Zhang, Xiyue, Darryn W. Waugh, and Clara Orbe. "Dependence of Northern Hemisphere Tropospheric Transport on the Midlatitude Jet under Abrupt CO2 Increase." *Journal of Geophysical Research: Atmospheres* (2023): e2022JD038454.

Orbe, Clara, David Rind, Jeffrey Jonas, Larissa Nazarenko, Greg Faluvegi, Lee T. Murray, Drew T. Shindell et al. "GISS model E2. 2: A climate model optimized for the middle atmosphere—2. Validation of large-scale transport and evaluation of climate response." *Journal of Geophysical Research: Atmospheres* 125, no. 24 (2020): e2020JD033151.

L253-254: Again, I fail to understand what is meant by AMOC non-linearity. 1.5Sv is regular AMOC swings, probably within one standard deviation.

Please see our comment above. Furthermore, we agree with the reviewer that 1.5 SV is within internal variability. Therefore, we have corrected our statement and now reference the fact that this nonlinearity is not statistically significant, relative to internal variability. Please see the revised manuscript.

L264-272: An enhanced and eastwardly extended North Atlantic jet stream is also seen in water hosing simulations, which further corroborates your hypothesis, see for example:

****	Bellomo, K., Meccia, V.L., D'Agostino, R. et al. Impacts of a weakened AMOC on precipitation over the Euro-Atlantic region in the EC-Earth3 climate model. Clim Dyn (2023).***

****	Jackson LC, Kahana R, Graham T et al (2015) Global and European climate impacts of a slowdown of the AMOC in a high resolution GCM. Clim Dyn 45:3299-3316.***

We thank the reviewer for providing these references, which we now cite in a new sentence that we have added to the paragraph. Please see the revised manuscript.

L280-289: Here the authors argue that corresponding to different AMOC decline rates in OMA and NINT, there is an impact of interactive chemistry on the zonal wind changes. While the long term total response is more convincing because is consistent with water hosing experiments (see refs cited above), I don't think the 5-20 years average (16 years total) bears any statistical significance. The authors mention this result is consistent with CP2019, but I don't see any explanation of physical mechanisms that would explain this jet response in either the Atlantic or Pacific Oceans. I strongly encourage the authors to seek or at least provide evidence for a sound physical process that would explain the jet response at the fast time scale.

We thank for the reviewer for her/his comment. We respond to each point separately:

"I don't think the 5-20 years average (16 years total) bears any statistical significance."

The 5-20 year OMA-NINT response in the 850 hPa zonal winds is statically significant, as indicated in the grey dotted regions (Fig. 3c). Having only had one OMA ensemble member, however, we generated a 4-member LINOZ ensemble, which also shows that this surface wind response is statistically significant (Figure 6a).

"The authors mention this result is consistent with CP2019, but I don't see any explanation of physical mechanisms that would explain this jet response in either the Atlantic or Pacific Oceans."

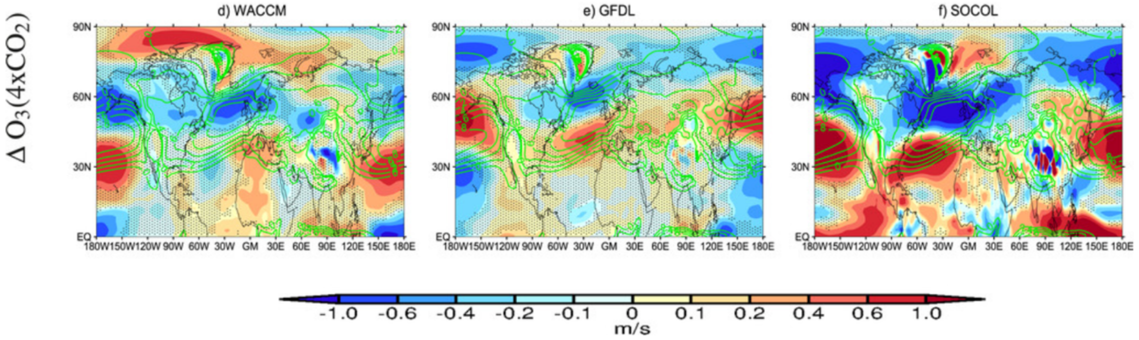
We do not discuss plausible mechanisms in the text at this point since we first want to confirm a) that we see this response in a larger (4-member) ensemble and 2) that stratospheric ozone changes (not aerosols) are driving this near-surface wind response. To this end, it is most appropriate to discuss mechanisms only *after* we have confirmed these findings in Section 3b; otherwise, we would be explaining either an insignificant result and/or one related to tropospheric composition changes.

"***I strongly encourage the authors to seek or at least provide evidence for a sound physical process that would explain the jet response at the fast time scale"***

We agree with the reviewer that a physical mechanism needs to be proposed (after first confirming the response is present in the LINOZ ensemble, per our comment above). This occurs in Section 3c, paragraph 1 where we ascribe the acceleration of the NH midlatitude jet above and along the jet core (and weakening on poleward flank) to the "reduction in the meridional temperature gradient near the tropopause" associated with the anomalous cooling due to ozone transport out of the tropical lower stratosphere. The fact that we see a similar response in the winds at 300 hPa (Supplementary Figure A2) supports our hypothesis that local changes in lower stratospheric meridional temperatures gradients imply, through thermal wind, an equatorward shift of the NH jet in the upper troposphere. We find that this signature is barotropic over the North Atlantic and, hence, evident at 850 hPa.

The idea that lower stratospheric ozone-induced temperature gradients are associated with an equatorward jet shift through thermal wind was already explored in CP2019 and Li and Newman (2022). In particular, in CP2019 Section 4a they report a very similar temperature response to what we find (see their Figure 4) and they argue that "The temperature response to $\delta(O_3)(4xCO_2)$ implies a reduction in the meridional temperature gradient near the tropopause, which has major consequences for the atmospheric circulation…Hence, $\delta(O_3)(4xCO_2)$ substantially reduces the poleward shift of the Atlantic jet due to CO_2…These results are consistent with the ozone-induced temperature perturbation near the tropopause, and the resulting change in the meridional temperature gradient at these levels."

To see the parallels between their results and ours consider, for example, their Figure 6, which shows the zonal wind changes in one model (WACCM-SC) induced by the $4xCO_2$ ozone responses from three different models (WACCM, GFDL, SOCOL). First, you can see that this figure shows very similar negative NAO-like responses over the North Atlantic (similar to our result). At the same time, it also confirms that the ozone feedback over the Pacific is not robust. For that reason, we do not expect that any ozone feedback over the Pacific that we find in our model will be robust.



The attribution of the North Atlantic jet change to tropical lower stratospheric meridional temperature gradient was also invoked in the Li and Newman (2022) study, who also showed a very similar (negative NAO-like) response of the NH boreal winter jet to ozone feedbacks (their Supplementary Material Figure S2):

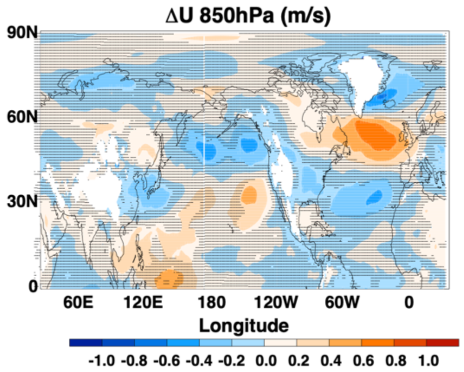


Fig. S2 Different response in the annual-mean zonal wind to 4×CO_2 between the PC and IC experiment (PC minus IC). Stippling indicates that the difference is not statistically significant at the two-tailed 5% level

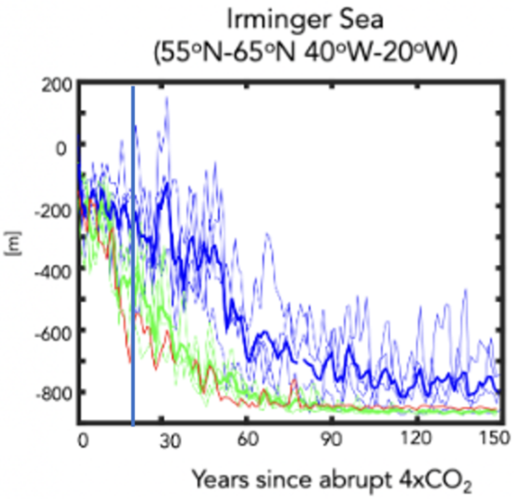
Please note that the above figure plots the difference between prescribed vs. interactive chemistry runs (not LINOZ– NINT), which is why the sign of their response is opposite to what we and CP2019 show.

To summarize: It is clear that the ozone feedback on the North Atlantic jet is something that is increasingly becoming a more robust result. Furthermore, while the North Atlantic response identified in Li and Newman (2022) is incidental and not the focus of that SH-centered study, they too attribute the jet responses in both hemispheres to the ozone influence on lower stratospheric tropical temperature gradients. In particular, the authors argue that by examining seasonal variations in the ozone-induced temperature changes that "the seasonality of meridional temperature gradient change in the tropopause region is important for determining the seasonality of the downward coupling of stratospheric circulation anomalies." They argue that the season of maximum influence in the SH of stratospheric ozone changes is when the strongest meridional temperature gradient increases are found.

Note that we already show in Supplementary Figure A2 that the wind response over the Atlantic is manifest in all interactive chemistry configurations at 300 hPa, which supports the idea that the increased meridional temperature gradients in the UTLS region are descending into the lower stratosphere. We therefore do not think that there are new figures that need to be added. Nonetheless, we have tried to make clearer the proposed connection relating the reduced meridional temperature gradients in the upper troposphere to the equatorward shift of the jet, along the lines of what previous studies have already argued.

L438-440: This statement does not make sense to me. The NINT, LINOZ, OMA all show the same decline rate in the first 15 years, which is clearly due to 4xCO2 forcing (c.f. Bellomo et al. 2021). I don't see any role for stratospheric forcing of the AMOC. 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. There is some gap between the red and blue line the first 20 years in the Irminger Sea, but the blue and black lines appear indistinguishable, likely falling within internal variability.

We do not agree. Please look again at the figure, which we have reproduced here and added a line at year 20. There is a clear gap between the blue (NINT) vs. red (OMA)/green (LINOZ) thick (ensemble mean) lines before year 20. Also, there is no black line.



L462-464 should be modified to accurately report the findings of the cited study: I suggest to add 'artificially' before 'adding (extracting)' at L 462, and 'in a climate model idealized experiment' before (Delworth and Zeng 2016) at L463.

Good point – we have modified this sentence as requested. Please see the revised manuscript.

L469: Add after years1-5: 'Averages for years 5-20 are shown in fig. 10'. Otherwise the reader is left wondering why choosing years 1-5, why everywhere else they were looking at the period 5-20.

Very good point – we have added the suggested text. Please see the revised manuscript.

L480-481: mechanistically from the text it is not clear to me why the surface winds and frictional speed is reduced more in LINOZ rather than NINT. Can you clarify?

The reduced surface winds (and surface friction speed) reflect the same 850 hPa zonal wind weakening shown in Figure 6a. This is already described in the text. Note this is a barotropic wind response that is coupled to the winds at 300 hPa (Supplementary Figure A2), which bolsters our argument that the stratospheric ozone changes are impacting the AMOC via changes in lower stratospheric meridional temperature gradients and winds that couple to the surface.

Section g (L558) finally clarifies some of the issues I raised above. Up to this point I had no idea how the ozone would drive the wind response the feed-back onto the ocean circulation, but still this is not detailed here. I suggest expanding this section and perhaps move it up in the text. At L577 I get there is a striking similarity with the earlier figure, and CP2019, but still I have no idea why the winds and air temperature respond this way to ozone changes.

Please see our response to the reviewer's previous comment asking for a physical explanation relating the ozone changes to the surface wind response.

L605-608: Why is this happening? What are the physical processes involved? Figs 9-10 provide convincing evidence of the processes involved in the coupling with the AMOC, but the first part - how the ozone drives air temperature and wind changes - is not demonstrated. There often is in the paper a reference to CP2019, but aside from the fact that results are similar it is not even clear to me why I am reading this paper and not CP2019, if there is not an independent story in here to read.

Again, please see our previous response to a similar comment.

Our manuscript presents a very independent story from the results in CP2019. The independent story here is that in our model the ozone feedback is modulated by coupling with the ocean, resulting in a long-term stratospheric ozone feedback on the midlatitude jet that is *opposite* to what they find. In particular, like CP2019, we find that stratospheric ozone changes initially result in an equatorward shift of the midlatitude jet which is concentrated over the North Atlantic. However, in our model this subsequently sets of a response in the AMOC that is accompanied by a longer-term poleward shift of the jet over the Pacific and an acceleration and eastward extension of the Atlantic jet over Europe.

Response to Reviewer 2

Reviewer #2: In this manuscript the authors try to isolate the midlatitude jet response to the ozone change caused by abrupt CO2 forcings. To these ends they perform several simulations using the NASA GISS model, with various combinations of fixed, evolving or prescribed CO2, ozone, aerosols and SST's. The authors find that on short timescales (5-20 years) the ozone leads to an anomalous equatorward shift of the midlatitude jet, while on longer timescales (100-150 years) the jet shifts poleward. They explain this by an anomalous weakening of the Atlantic meridional overturning circulation (AMOC), which shifts the jet poleward on these timescales. The anomalous weakening of the AMOC in turn is explained by changes in wind stress due to the anomalous fast response.

I think the manuscript is well written and provides a novel contribution, although the presentation of the figures has to be improved on. However I disagree with the general statement that the total response to ozone is a poleward shift, which I don't think can be concluded from the presented evidence, see below for more details. Also, I am suggesting to perform another experiment to strengthen the presented arguments. For these reasons I am recommending major revisions, but I believe that these concerns can be addressed.

We very much appreciate the thoughtful feedback from the reviewer. We hope that our incorporation of her/his feedback has significantly improved the manuscript.

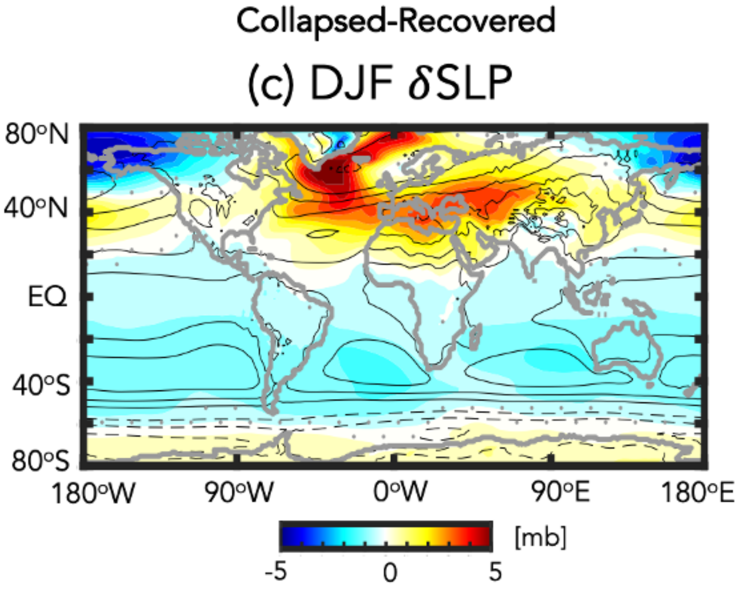
Mayor comments:

General: I think it would be valuable to perform an experiment like the last one in table 1, but with a coupled ocean. That means prescribing the ozone response to 4xCO2, while keeping aerosols and CO2 fixed. You observe an equatorward shift to decreasing ozone, as shown in fig. 12. By your arguments in section 3f, this should lead to a weaker AMOC, and thus to an altered total response. Performing this experiment would significantly strengthen the causality of your argument and provide direct experimental evidence of your mechanism. If these hypothesised responses are not observed, it would be necessary to investigate why they only appear with an increased CO2 background state, as observed in LINOZ.

We thank the reviewer for the comment. At her/his recommendation, we have performed a new 4-member ensemble of coupled atmosphere-ocean 1xCO_2 simulations, in which we prescribe the time evolution of the 4xCO_2 O_3 response from the interactive LINOZ simulations. We have run these simulations out 50 years each in order to see if the evolution of the response is as expected from the coupled 4xCO_2 interactive ozone simulations. To facilitate comparison between all of the different runs, we here present figures showing the following comparisons: a) LINOZ-NINT 4xCO_2 CPLD (top right) b) LINOZ-NINT 1xCO2 CPLD (top left) and c) LINOZ-NINT 1xCO_2 FIXED PICONTROL SST (bottom). The results in a) and c) are already shown in the manuscript – the new results are what are shown in b). Note that we also ran an additional experiment that was not requested, but which tests the impact that reading in the LINOZ ozone field vs. calculating them online makes on the 4xCO_2 response of the circulation. That is, we ran one coupled atmosphere-ocean 4xCO_2 experiment in which we read in the time-evolving 4xCO_2 O_3 response from LINOZ. Those results (not shown) demonstrate that our main findings are not sensitive to the interactivity of the compositional forcing (i.e., using online vs. read-in LINOZ ozone fields).

Examining Figures R2_1 through R2_5 shown below, we find that the LINOZ-NINT 1xCO_2 coupled ocean response does a good job of capturing the LINOZ-NINT 4xCO_2 coupled ocean differences for most variables. These variables include annual mean temperature (Fig. R2_1 – compare with figures 5a and 12b in the manuscript), zonal mean DJF wind (Fig. R2_2 – compare with figures 5c and 12a) and 850 hPa zonal wind (Fig. R2_3 – compare with figures 6a and 12c in the manuscript). The responses in surface winds (Fig. R2_4 – compare with figure 10a (second row) in manuscript), surface friction speed (Fig. R2_5 – compare with figure 10b (second row) in manuscript), mixed layer depth (Fig. R2_6 – compare with figure 10c (second row) in manuscript), net heat flux into the ocean (Fig. R2_7 – compare with figure 10d (second row) in manuscript), and sea level pressure (Fig. R2_8 – compare with supplementary figure A4 in manuscript) and are also captured, but are somewhat weaker in the $1xCO_2$ experiment.

These figures support our hypothesis that stratospheric ozone changes produce reduced meridional temperature gradients in the lower stratosphere -> equatorward zonal mean jet shift in the NH -> NAO-like response in the lower troposphere -> reduced surface winds and surface friction speed -> increased heat fluxes into the ocean and -> positive SLP over Greenland. At the same time, however, the wind anomaly is not translated as efficiently into the enhanced (into the ocean) heat flux and mixed layer depth anomalies in the $1xCO_2$ experiment, compared to $4xCO_2$. We hypothesis that this is because the AMOC response to $4xCO_2$ is itself associated with reduced surface zonal winds over the North Atlantic, increased sea level pressure, etc., a result that has been documented in several previous studies. For example, the results from our recent study (Orbe et al. (2023)), show increased sea level pressure over the North Atlantic in identically forcing SSP 2-4.5 ensemble members in which the AMOC collapses, compared to ensemble members in which the AMOC recovers:



That is, an AMOC decline under increasing GHG forcing is associated with the development of the warming hole and enhanced sea level pressure. This prompts an atmospheric-temperature feedback associated with colder air (and sea surface) temperatures which hold less moisture. Together with reduced SSTs this results in reduced evaporation and freshening of the of the subpolar gyre, acting as a positive feedback to the AMOC decline (reported in Rind et al. (2018)). This feedback is simply not present in the $1xCO_2$ runs. We now have added a new paragraph briefly summarizing these results to the section discussing the fixed SST and SIC results. Please see the revised manuscript.

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* (2023): 1-52.

Rind, David, Gavin A. Schmidt, Jeff Jonas, Ron Miller, Larissa Nazarenko, Max Kelley, and Joy Romanski. "Multicentury instability of the Atlantic meridional circulation in rapid warming simulations with GISS ModelE2." *Journal of Geophysical Research: Atmospheres* 123, no. 12 (2018): 6331-6355.

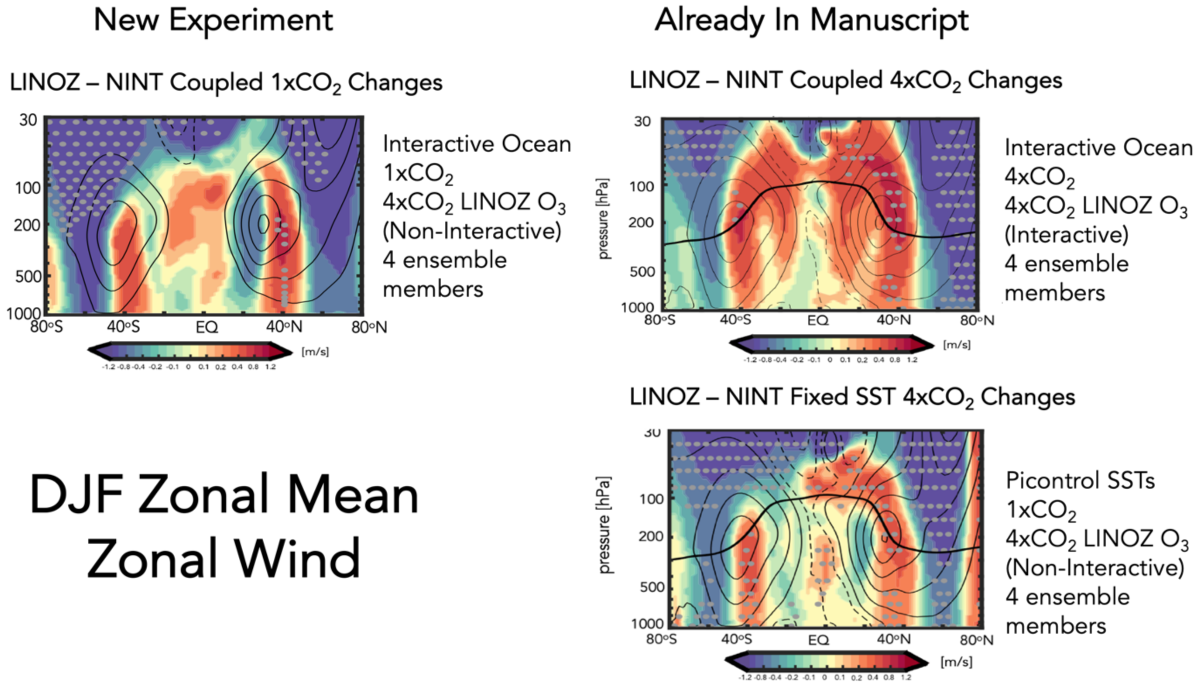


Figure R2_1: Comparison of the DJF LINOZ-NINT zonal mean zonal wind differences between the 4xCO_2 coupled atmosphere-ocean (top right) and 4xCO_2 fixed picontrol SST (bottom right) ensembles. The results from a new 4-member ensemble of LINOZ-NINT 1xCO_2 coupled atmosphere-ocean runs are shown in the left panel.

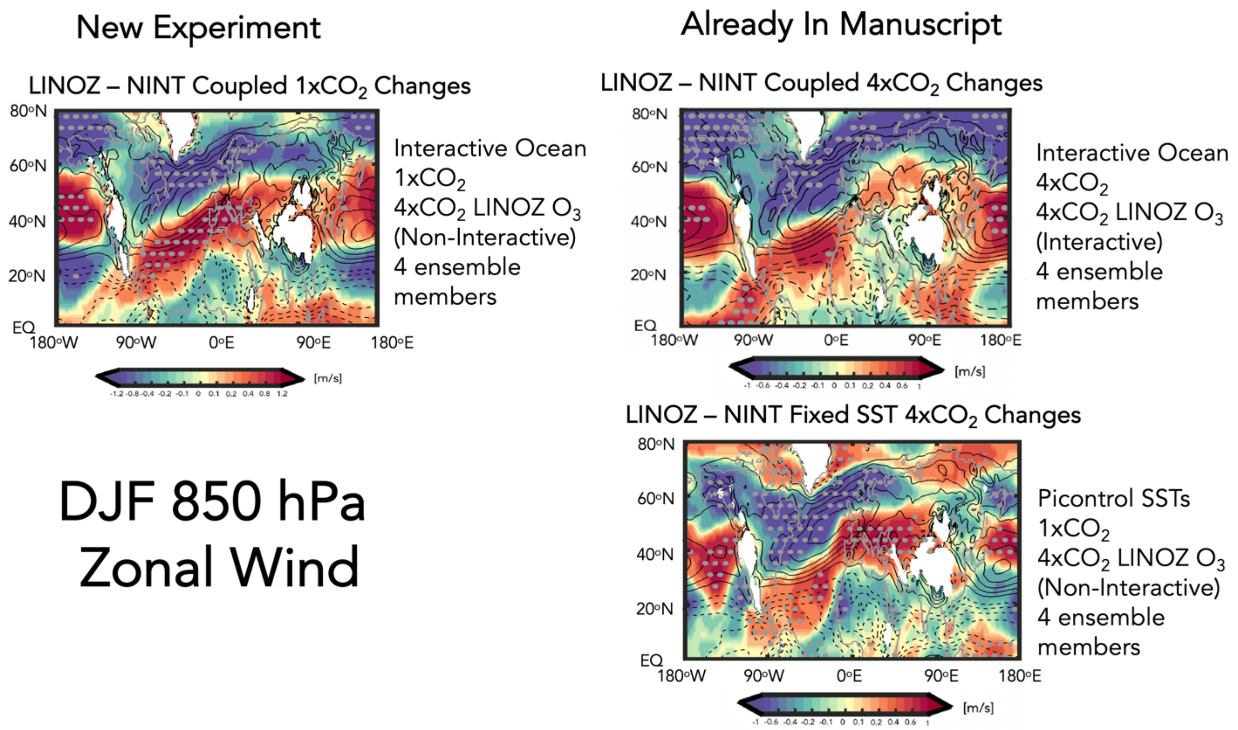


Figure R2_2: Same as Figure R2_1, except showing the 850 hPa zonal winds.

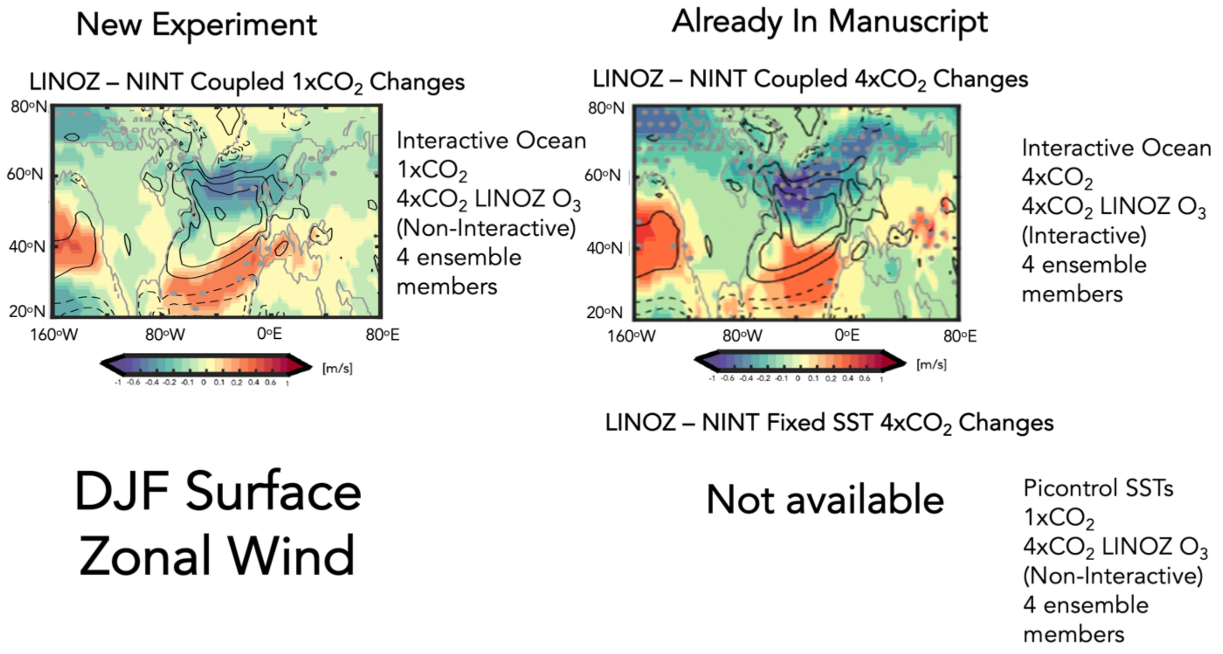


Figure R2_3: Same as Figure R2_1, except showing the surface zonal winds.

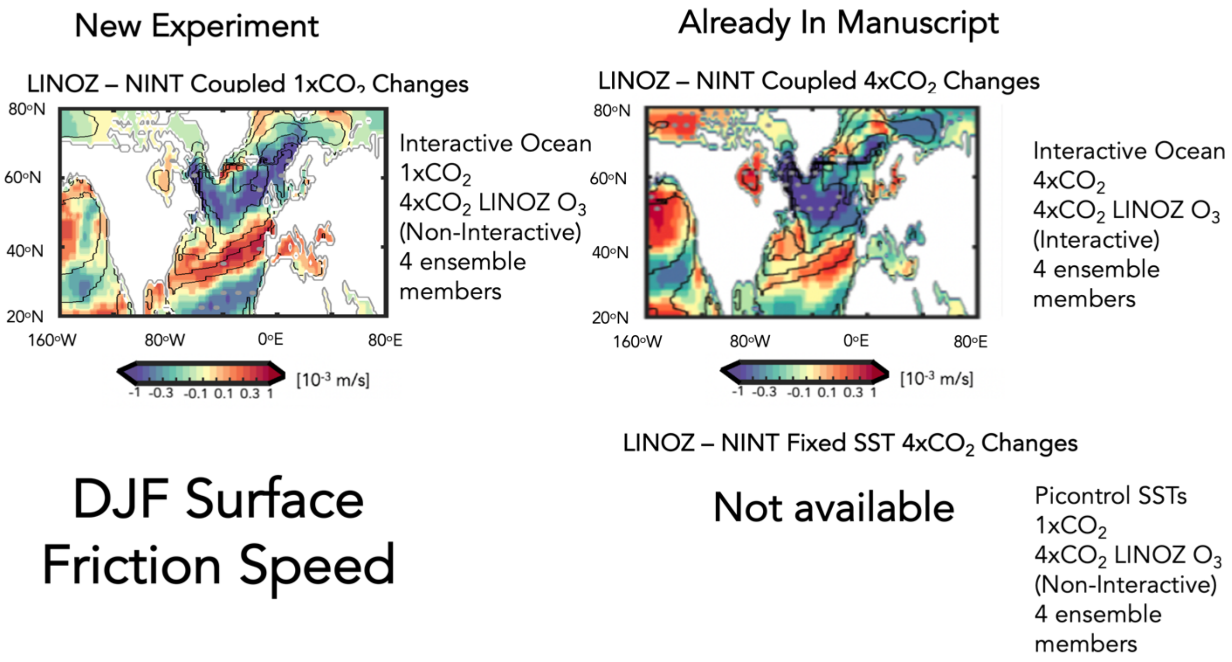


Figure R2_4: Same as Figure R2_1, except showing the surface friction speed.

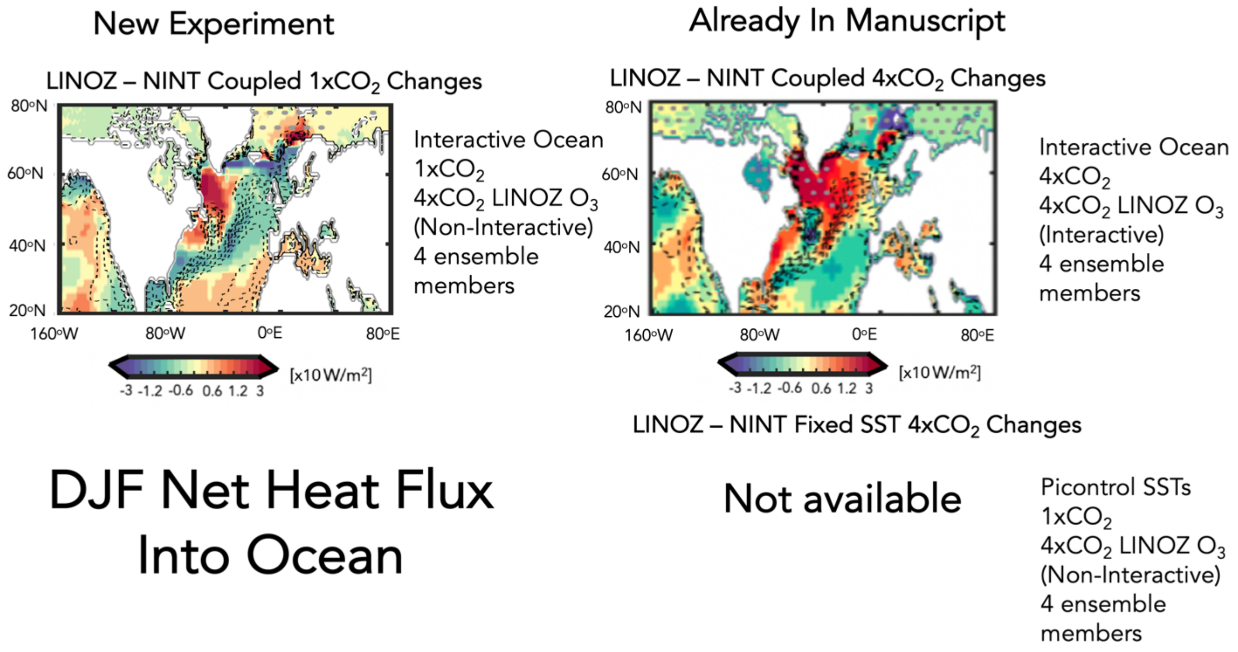


Figure R2_5: Same as Figure R2_1, except showing the net heat flux into the ocean.

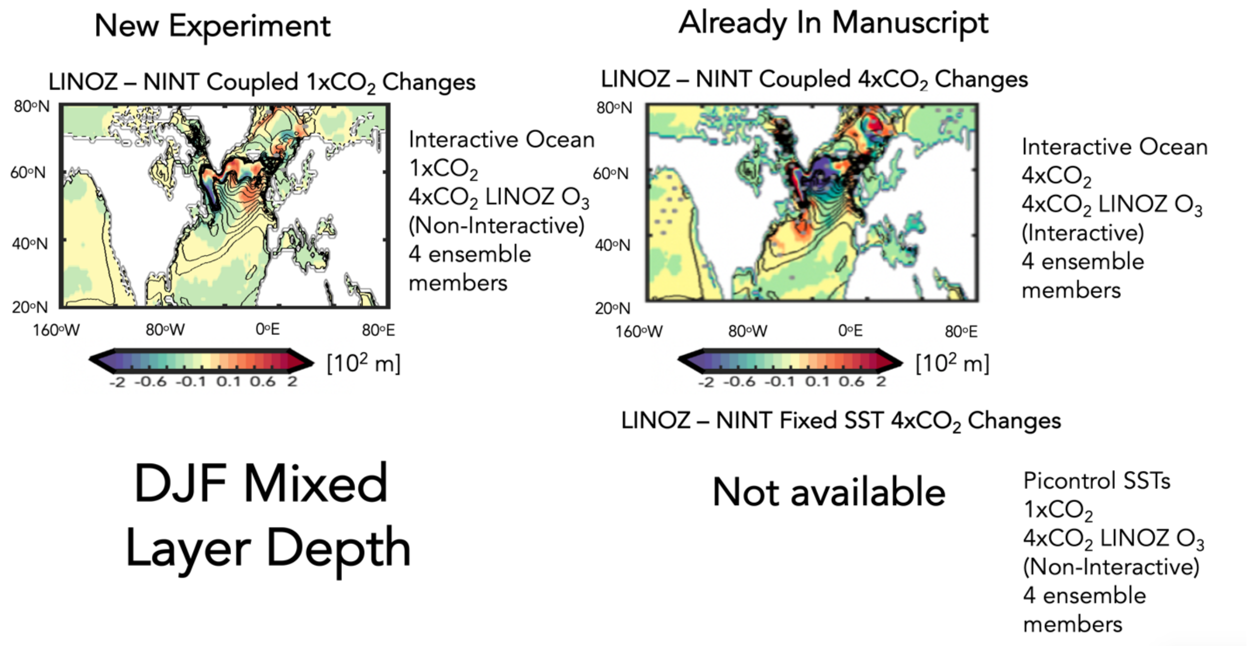


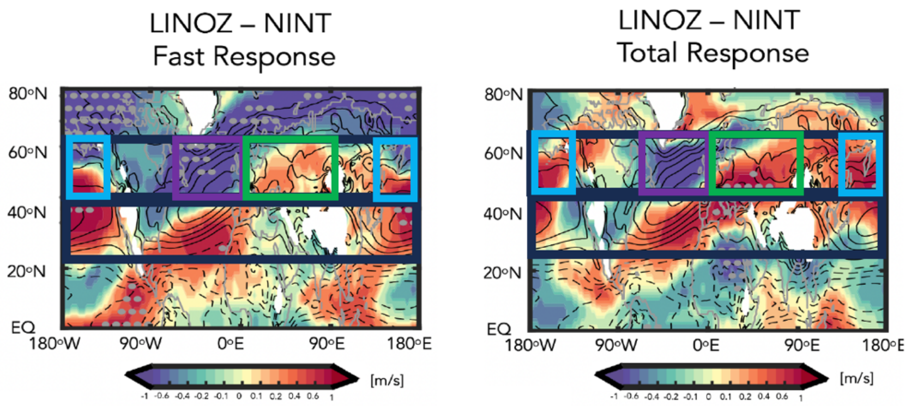
Figure R2_6: Same as Figure R2_1, except showing the ocean mixed layer depths.

l.390-407: I am concerned about the interpretation of the results in section 3d. Just to confirm, that I fully understood them: In fig. 5 and fig. 6 you show the differences in the fast and total response between LINOZ and NINT, which gives a measure of the influence of Ozone on the responses. You show the zonal mean winds in fig. 5 c/d and point out that the total response (d) shows a poleward shift. However, fig. 6b shows strong asymmetries in both the climatology and the jet response, which questions how to meaningfully interpret the zonal mean picture. Furthermore, we can see from 6b that the jet response to ozone is actually an equatorward shift over the Atlantic and European region, which is in contrast to l.400 where you state that a poleward shift is found at all longitudes. Over the Pacific we do see a poleward shift accompanied by a strengthening. Please enlighten me if I misunderstood your findings or claims. However, since a core claim of your study (a poleward shift of the jet due to ozone on long timescales) rests on these results, I feel like they have to be explored with more scrutiny. Especially the zonal mean picture seems misleading and as far as I can see, your claim of a general poleward shift in the total response due to ozone does not hold. An analysis based on physically motivated ocean and land basins might bring more insight.

We completely agree with the reviewer that we have oversimplified our description of the jet response in the manuscript. Our focus on the zonal mean picture in Figure 5 is mainly to connect back to the CP2019 study and provide the reader with the opportunity to compare directly between our results and theirs (using the same colorbars). However, we realize now that, in the absence of more caveats about interpretation in the text, this misleadingly places too much emphasis on the zonal mean jet shift.

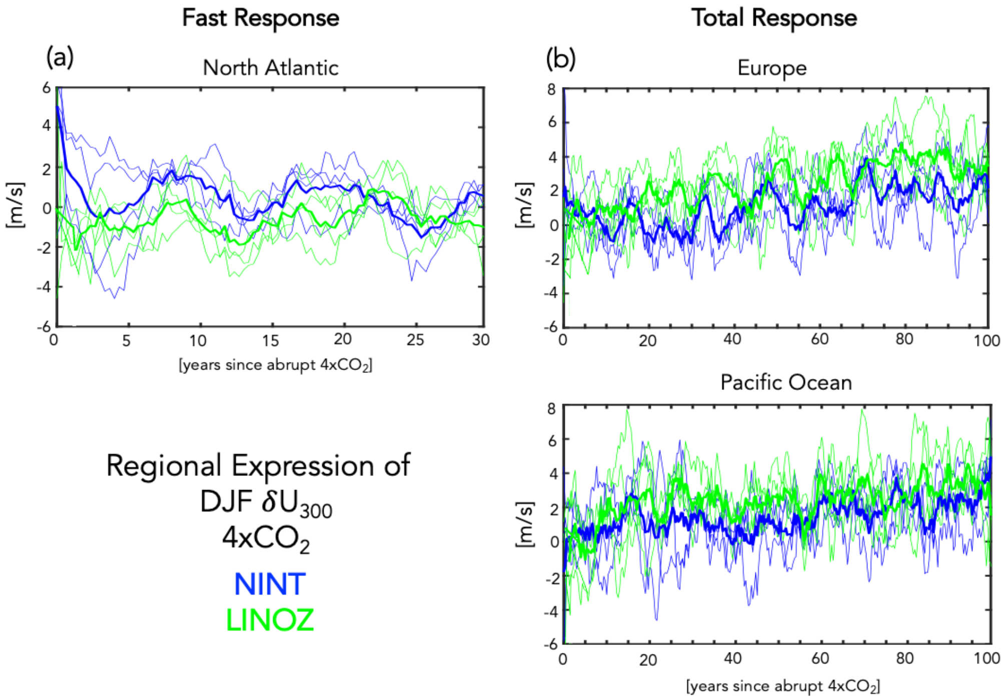
Another reviewer articulated a similar concern, additionally asking to see the jet shift manifest in the timeseries. This request to see the timeseries also addresses this reviewer's concern as it highlights that the zonal mean jet shift evident in Figure 5 is expressed regionally in different ways. In particular, as the reviewer here points out, Figure 6 highlights how the zonal mean changes reflect a complex set of responses occurring over different basins (colored boxes below) during the "fast" vs. "total" timescales.

DJF δU_{850}
4xCO_2



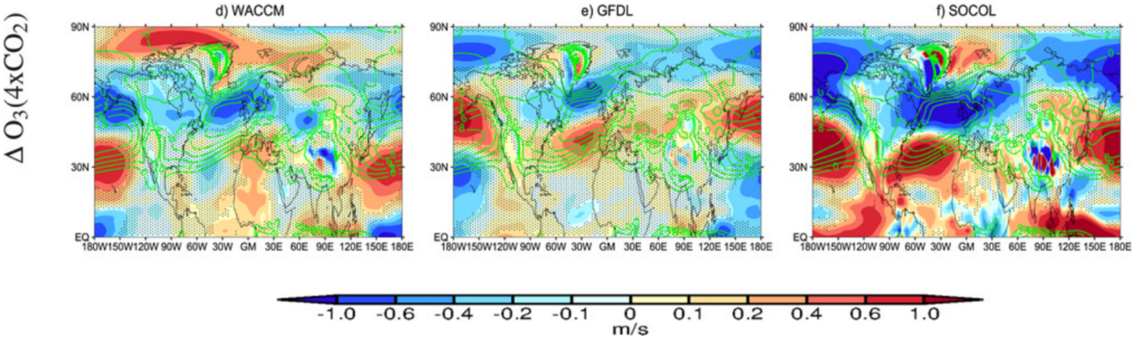
Analyzing the timeseries (figure below) over the boxed regions indicated above reveals that:

1) The "fast" equatorward shift evident in the zonal mean primarily reflects a "fast" equatorward shift over the Atlantic at 850 hPa (purple boxes).
2) The "total" zonal wind response, by comparison, consists of an acceleration of the jet eastward over Europe (green boxes, top right timeseries below) and a poleward shift over the Pacific (cyan boxes, bottom right timeseries below).

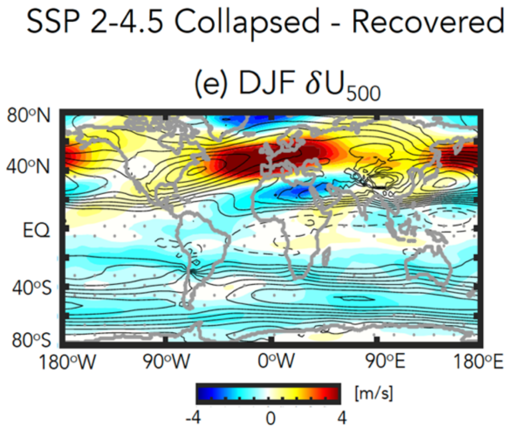


We now have included this figure in the appendix (see new Fig. A3). We also incorporate this in our more nuanced discussion of the zonally asymmetric jet response.

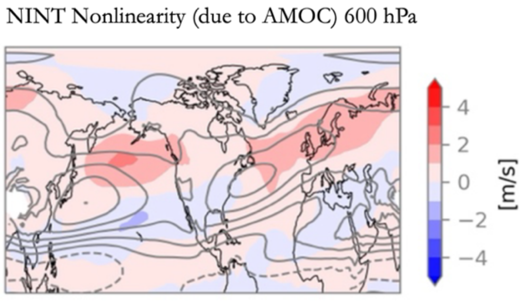
We also want to take this opportunity to remind the reviewer that we focus on the ``fast" response over the North Atlantic not only because it is key in instigating the subsequent changes that occur in the AMOC, but also because we do not expect to find any robust ozone signature in the Pacific, based on the results in CP2019. Consider, for example, their Figure 6, which shows the zonal wind changes in one model (WACCM-SC) induced by the 4xCO_2 ozone responses from three different models (WACCM, GFDL, SOCOL). This figure shows very similar negative NAO-like responses over the North Atlantic (similar to our result), but very different results over the Pacific. For that reason, we do not expect that any ozone feedback over the Pacific that we identify with our model will be robust.



With respect to the acceleration of the jet occurring over Europe and Pacific in response to a stronger AMOC decline, we also want to highlight that this has been reported in several previous studies as well. In particular, we highlight the results from a recent study in which we recently compared the jet response between identically forced SSP 2-4.5 simulations between ensemble members in which the AMOC collapsed vs. recovered (left panel below, Orbe et al. (2023)). In another recent study we also showed this long-term response of the AMOC collapse on the NH zonal winds (right panel, below; Zhang et al. (2023)):



Orbe et. al. (2023)



Zhang et. al. (2023)

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* (2023): 1-52.

Zhang, Xiyue, Darryn W. Waugh, and Clara Orbe. "Dependence of Northern Hemisphere Tropospheric Transport on the Midlatitude Jet under Abrupt CO2 Increase." *Journal of Geophysical Research: Atmospheres* (2023): e2022JD038454.

To summarize: We recognize that the zonal mean response does not faithfully represent the jet changes that are occurring regionally and which, taken together, give a clearer picture of what we mean by "two-timescale" response. We have significantly revised our discussion of the jet response to reflect this more nuanced view and we now show the timeseries in a new Supplementary Figure (Fig. A3). We hope that the combination of this new figure and clarified discussion in the text addresses the reviewer's concerns. Please see the revised manuscript.

l.201-203: You are not specifying a confidence level. Also how was the data sampled? Simply using all data is not valid I believe, because the data points inside a sample will not be independent.

Our apologies for not clarifying confidence intervals (5% level), which we do now in the manuscript. Furthermore, we identified a bug (confidence level set to 2% level by accident), which affected our LINOZ-NINT significance stippling in the relevant panels in Fig. 5,6,9,10 and 12. Fixing this produces only subtle differences in the stippling (see revised figures). Statistical significance of the four-member ensemble mean LINOZ-NINT and single member OMA-NINT abrupt CO_2 differences is assessed relative to the interannual variability in the corresponding preindustrial control simulation. This is consistent with the approach used in Zhang et al. (2023) and is now mentioned in the text.

l.328: You say that the cooling is about 3K, but from figure 4 it looks more like 1-2K.

Apologies – we thank the reviewer for catching this mistake. We have now corrected this in the manuscript.

l.339: You state that the range of temperature change is 2-4K, again this is not supported by fig.4

We thank the reviewer for catching this mistake. We have now corrected this in the manuscript.

l.343: Related to the two comments above, you state that the 2K cooling in NINT is 50% of the cooling in OMA/LINOZ. This is incompatible with fig.4.

The cooling in OMA/LINOZ relative to NINT is 1.5-2 K. Therefore, if anything, this statement underestimates the cooling due to the ozone feedback. We have therefore revised "~50%" to "same order of magnitude".

Minor comments:

Abstract: Consider making clear that those are findings based on one model

We have added "Using simulations produced with the NASA Goddard Institute for Space Studies (GISS) high-top climate model (E2.2)…" to the abstract. Please see the revised manuscript.

l.18: Consider removing 'midlatitude eddy-driven' as it might be confusing and lead people to think it refers to a different structure than the 'jet' in l.17

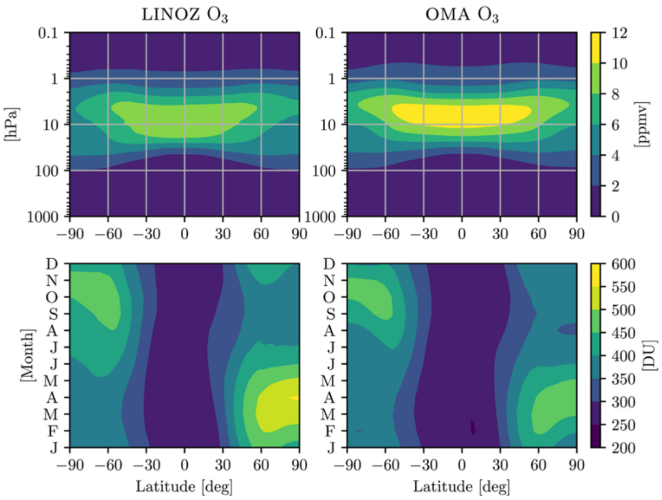
Done – this has been removed. Please see the revised manuscript.

l.142-148: You are giving two different description of how ozone is calculated and state that the second one is used for tropospheric ozone. The first one is thus used for stratospheric ozone? Providing an additional sentence or two of explanation on how the LINOZ experiments are set up would be valuable I think.

We have added "stratospheric" to the first sentence for clarification. We are not clear, however, which additional details on experimental setup the reviewer is seeking. These runs are identical to the NINT 4xCO_2 coupled ocean runs, except that ozone is calculated online via use of the LINOZ parameterization (whereas in NINT it is prescribed to a preindustrial climatology).

l.145: How well does a first order Taylor expansion hold during the end of the simulation period? This is somewhat answered in l.317-325. Maybe you can add a sentence here?

We already discuss the fidelity of the LINOZ parameterization in the last sentence of that paragraph: "DallaSanta et al. (2021) previously showed that the LINOZ ozone parameterization reproduces well the vertical structure and seasonal cycle of stratospheric ozone obtained from the fully interactive OMA configuration (see their Figure 1)." This is the figure to which we are referring:



Comparisons of ozone climatologies for LINOZ (linearized ozone) and OMA (comprehensive chemistry) schemes (row 1) annually averaged zonal-mean volume ratio, as a function of latitude and pressure; and (row 2) monthly averaged zonal-mean total column ozone, as a function of month and latitude. In LINOZ, enhanced ozone concentration at lower levels leads to greater total column ozone, despite its weaker stratospheric concentration.

l.187-188: What are the opposites? Ocean basins?

Our apologies – we should have been clearer. Yes, we mean "oppositely signed shifts between the Pacific and Atlantic basins." We have now clarified this in the manuscript.

l.242-244: The logic about the non-linearity is not clear to me. Why is a stronger response in OMA that is observed in both 2x and 4xCO2 forcing indicative of a non-linearity in NINT?

By "AMOC non-linearity" we refer to the fact that the response of the AMOC to 4xCO2 does not equal twice the response of the AMOC to 2xCO_2. The nonlinearity is then equal to the difference $1/2(\delta(4xCO_2))$- $\delta(2xCO_2)$ for any field of interest. We have now added a line in the manuscript noting this definition and appropriate references. Note that this definition of linearity is consistent with its usage in the following studies:

Chadwick, Robin, and Peter Good. "Understanding nonlinear tropical precipitation responses to CO2 forcing." *Geophysical research letters* 40, no. 18 (2013): 4911-4915.

Mitevski, Ivan, Clara Orbe, Rei Chemke, Larissa Nazarenko, and Lorenzo M. Polvani. "Non-monotonic response of the climate system to abrupt CO2 forcing." *Geophysical research letters* 48, no. 6 (2021): e2020GL090861.

Zhang, Xiyue, Darryn W. Waugh, and Clara Orbe. "Dependence of Northern Hemisphere Tropospheric Transport on the Midlatitude Jet under Abrupt CO2 Increase." *Journal of Geophysical Research: Atmospheres* (2023): e2022JD038454.

Orbe, Clara, David Rind, Jeffrey Jonas, Larissa Nazarenko, Greg Faluvegi, Lee T. Murray, Drew T. Shindell et al. "GISS model E2. 2: A climate model optimized for the middle atmosphere—2. Validation of large-scale transport and evaluation of climate response." *Journal of Geophysical Research: Atmospheres* 125, no. 24 (2020): e2020JD033151.

l.251: How is this -5SV calculated? What does it mean? I feel like l.251-254 has to either be expanded on or discarded.

We have now added our definition of nonlinearity to the manuscript (see response to the previous comment) which should address the reviewer's concern.

l.368-374: You could consider showing OMA-LINOZ in the appendix, as it is a bit complicated to gauge the difference between them from the two plots.

We do not believe that showing this figure would be very meaningful since it would the difference between one ensemble member (OMA) and a four-member ensemble (NINT). As we already discus in the text, there is large variability in the "fast" response, so we do not think it is necessary to show this figure directly.

l.590: regarding fig.12c/d. Why are you not showing the all longitudes? This would also help to compare to fig.6a,c as advised in l.582

We completely agree with the reviewer. We now show all longitudes in Figure 12 c and d. Please see the revised manuscript.

Technical comments:

l.114: North Atlantic Deepwater (NADW)

No change to the manuscript as the official term appears to be North Atlantic Deep Water (NADW). For example, see the relevant discussion in:

Buckley, Martha W., and John Marshall. "Observations, inferences, and mechanisms of the Atlantic Meridional Overturning Circulation: A review." *Reviews of Geophysics* 54, no. 1 (2016): 5-63.

l.243: resulting in -> indicating

Zhang et al. (2023) do actually ascribe the transport changes to the enhanced isentropic mixing, so we feel that this statement is an accurate depiction of that study. No change to the manuscript.

l.349: enhancement -> decrease (?)

We thank the reviewer for catching this mistake! This has been fixed.

l.358: reference back to fig.1

Good point – we have added this to the manuscript.

l.366: zonal mean response -> zonal wind response at 850hPa

We thank the reviewer for catching this mistake. This has been fixed in the revised manuscript.

l.636: are -> our

We thank the reviewer for catching this mistake. This has been fixed in the revised manuscript.

Figures:

Fig.1: Decrease fontsize of titles and CO2 specification. Increase fontsize of axis label and especially colorbar tick labels (almost unreadable in print).

This has been fixed.

Fig.3: Decrease fontsize of title, increase fontsize of colorbar tick labels. Could consider increasing figure size. Why not show c/d on the same color scale as a/b?

This has been fixed. Note that we show c/d on a different color scale as a/b because that is what CP2019 did and we want to enable as much comparison with that study as possible.

Fig.4: Increase fontsize of axis label and especially colorbar tick labels. Could consider increasing figure size. In c/d unit should be °C or K.

This has been fixed.

Fig.5: Same as fig. 1,3 and 4

This has been fixed.

Fig.6: As fig. 5

This has been fixed.

Fig.7: There seam to be gaps in the curves in 8b

We thank the reviewer for noticing this mistake! This has been fixed.

Fig.8: Reducing the title and subtitle texts and increasing the size of the plots would help. The stippling is not the same size in all plots and very hard to see in some.

This has been fixed.

Fig.9: Same as fig.8. Additionally the subfigures have different sizes, are not aligned and gaps between the subfigures have different widths.

This has been fixed.

Fig.11: Same as fig.7

This has been fixed.

Fig.12: Increase fontsize of colorbar tick labels. In b/d unit should be °C or K.

This has been fixed.

A1: Move the legend so it does not obstruct data

This has been fixed.

A2: Increase fontsize of colorbar tick labels

This has been fixed.

A3: As A2

This has been fixed.

Response to Reviewer 3

Reviewer #3: This study focuses on the role of stratospheric ozone in the Northern Hemisphere eddy-driven jet's response to abrupt 4xCO2 forcing. The authors show that changes in the ozone layer in response to abrupt 4xCO2 forcing initially cause the wintertime North Atlantic jet to shift equatorward. This initial equatorward jet shift, through changes in surface wind stress, weakens the Atlantic Meridional Overturning Circulation of the ocean, which is associated with a poleward shift of the zonal-mean Northern Hemisphere midlatitude jet.

Overall, this study presents an interesting analysis and highlights the importance of ozone changes in an increased CO2 scenario in modifying the atmospheric and oceanic general circulation, building on the results of past studies.

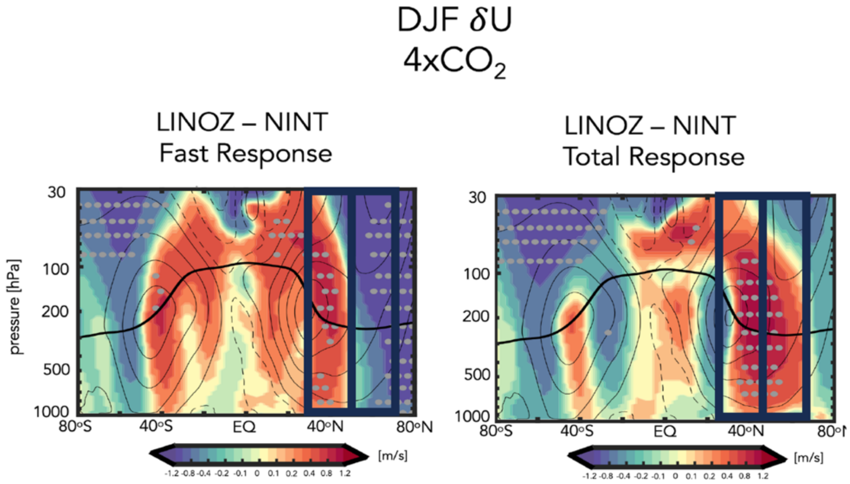
We very much appreciate the thoughtful feedback from the reviewer and we believe her/his contribution has significantly improved the manuscript.

I see two major weaknesses of the study in its present form:

1. What's really missing from this study is a time series of the jet response (zonal-mean jet latitude and/or North Atlantic jet latitude) to the abrupt 4xCO2 forcing in the ensemble of simulations used in this study (i.e., similar to Fig. 7, but for the jet latitude). The paper argues extensively that, in runs with time-varying ozone, the jet response has two timescales and that it has a large enough magnitude to be discernible from the forced CO2 changes. However, this is not explicitly shown anywhere in the paper in its present form. From Fig. 6, it appears the two-timescale jet response is more present in the Pacific sector than in the Atlantic sector, which contradicts the authors' hypothesis. This is why it is critical to show a plot of jet latitude versus time to explicitly demonstrate the two-timescale jet response that is the focus of this study.

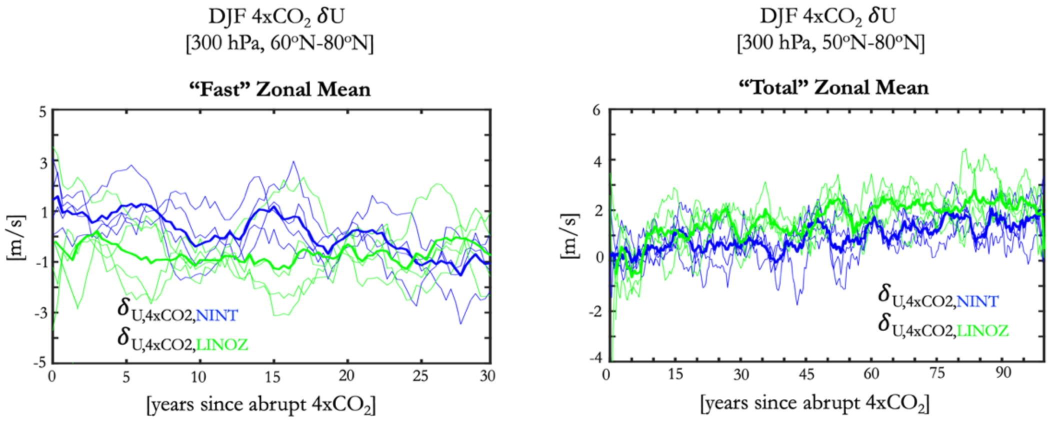
We agree with the reviewer that it is important to confirm the existence of a two-timescale jet response by showing the timeseries.

First, we remind the reviewer of the spatial patterns of the "fast" vs. "total" responses (evident in Fig. 5c and 5d):



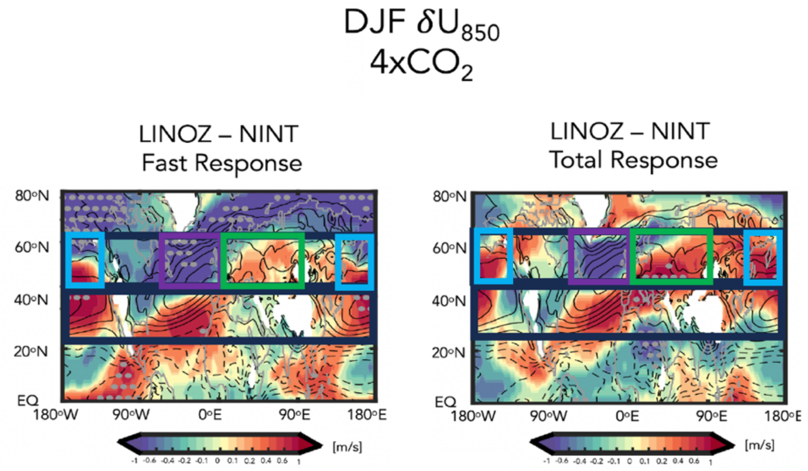
Here we have added the black boxes to highlight how 1) during the “fast” response there are stronger winds <50ºN and weaker winds >50ºN, especially poleward of 60ºN (very similar to Figure 5 in CP2019), whereas 2) during the “total” response there are stronger winds <50ºN *and* stronger winds >50ºN up to ~70ºN. The question, therefore, is how clearly these signals show up in the timeseries.

The “fast” zonal mean response -- consisting of weakened winds north of 60ºN -- is captured in the timeseries (bottom, left). So is the “total” response on longer timescales – consisting of stronger winds at and poleward of 50ºN (bottom, right):



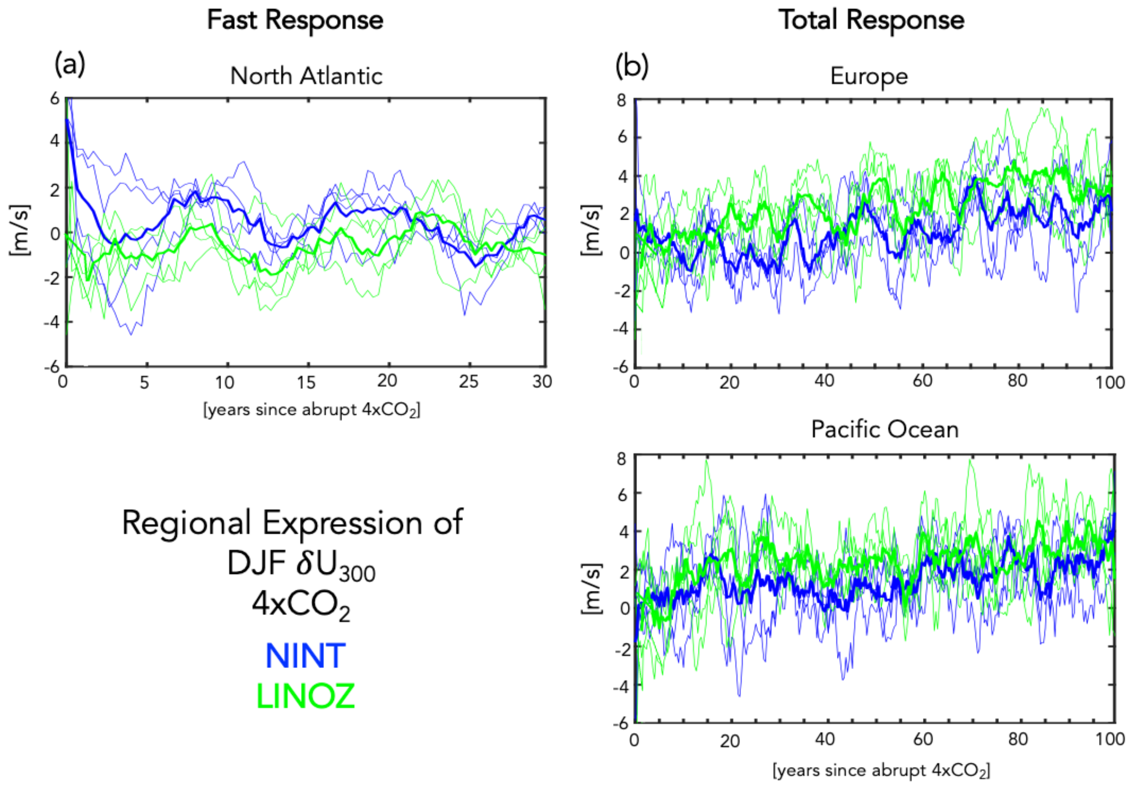
While these zonal mean timeseries do support the idea of a two-timescale response, upon further reflection of the reviewer’s comments we realize that this view is much too simplistic and not representative of the regional and basin-specific changes in the jet (which are key to influencing the stability of the AMOC and to capturing the impacts of a weakened AMOC). Indeed, we hope to clarify to the reviewer that our main purpose in showing the zonal mean figure (Figure 5) is to highlight the close correspondence between our results and previous literature (CP2019), where we’ve gone so far as to even use the same colorbars as in CP2019 to help contextualize our results.

As we already show in Figure 6 (reproduced below), but now emphasize much more clearly in the revised manuscript, the zonal mean changes in the winds reflect a complex set of responses occurring over different basins (colored boxes below) during the “fast” vs. “total” timescales:



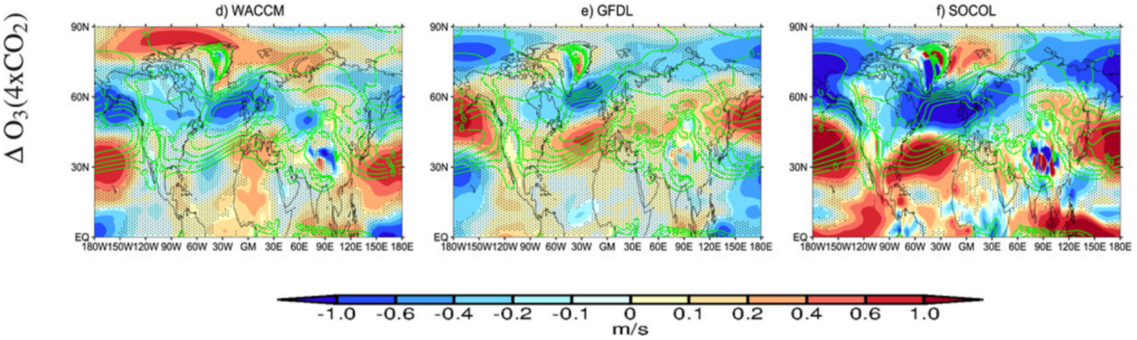
Analyzing the timeseries (figure below) over the boxed regions indicated above reveals that:

3) The “fast” equatorward shift evident in the zonal mean primarily reflects a “fast” equatorward shift over the Atlantic at 850 hPa (purple boxes).
4) The “total” zonal wind response, by comparison, consists of an acceleration of the jet eastward over Europe (green boxes, top right timeseries below) and a poleward shift over the Pacific (cyan boxes, bottom right timeseries below).

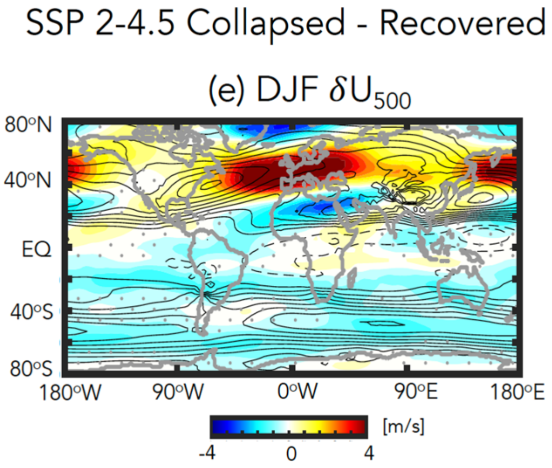


We now have included this figure in the appendix (see new Fig. A3). We also incorporate this in our more nuanced discussion of the zonally asymmetric jet response.

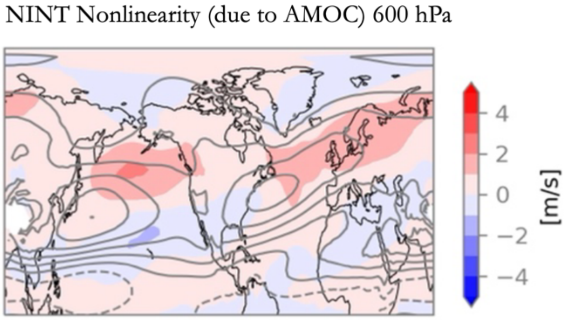
We also want to take this opportunity to remind the reviewer that we focus on the ``fast” response over the North Atlantic not only because it is key in instigating the subsequent changes that occur in the AMOC, but also because we do not expect to find any robust ozone signature in the Pacific, based on the results in CP2019. Consider, for example, their Figure 6, which shows the zonal wind changes in one model (WACCM-SC) induced by the $4xCO_2$ ozone responses from three different models (WACCM, GFDL, SOCOL). This figure shows very similar negative NAO-like responses over the North Atlantic (similar to our result), but very different results over the Pacific. For that reason, we do not expect that any ozone feedback over the Pacific that we identify with our model will be robust.



With respect to the acceleration of the jet occurring over Europe and Pacific in response to a stronger AMOC decline, we also want to highlight that this has been reported in several previous studies. In particular, we highlight the results from a recent study in which we recently compared the jet response between identically forced SSP 2-4.5 simulations between ensemble members in which the AMOC collapsed vs. recovered (left panel below, Orbe et al. (2023)). In another recent study we also showed that the long-term response of the AMOC collapse in E2.2. (right panel, below; Zhang et al. (2023)):



Orbe et. al. (2023)



Zhang et. al. (2023)

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* (2023): 1-52.

Zhang, Xiyue, Darryn W. Waugh, and Clara Orbe. "Dependence of Northern Hemisphere Tropospheric Transport on the Midlatitude Jet under Abrupt CO2 Increase." *Journal of Geophysical Research: Atmospheres* (2023): e2022JD038454.

To summarize: We recognize that the zonal mean response does not faithfully depict the jet changes that are occurring regionally and which, taken together, give a clearer picture of what we mean by "two-timescale" response. We have significantly revised our discussion of the jet response to reflect this more nuanced view and we now show the timeseries in a new Supplementary Figure (Fig. A3). We hope that the combination of this new figure and clarified discussion in the text addresses the reviewer's concerns. Please see the revised manuscript.

2. Another weakness of this study is that it only focuses on this behavior in a single model, and the lingering question in my mind while reading the paper was whether this behavior is just unique to this particular model. While conducting new experiments with other models is beyond the scope of this study, I think the authors could make a better effort to see if this behavior is present in other CMIP6 models by using existing data from the CMIP6 archive:

a. It should be straightforward to check if there are multiple timescale jet responses in the North Atlantic jet across other CMIP6 models. In other words, make a figure similar to the one suggested in my previous comment, but for the abrupt 4xCO2 scenario of various CMIP6 models. Ceppi et al. (2018) suggested that the Atlantic jet shift response is almost entirely dominated by the response within the first decade, but if this is not true for all models, it would be important to note this and would provide support for what the authors are arguing in this study.

We appreciate the concerns raised by the reviewer. However, we should be clear here that this "two-timescale response" is defined **relative to the NINT version of the model**. Therefore, in the absence of having both LINOZ (or interactive ozone) and NINT configurations of other individual CMIP6 models, it is not clear how analyzing the CMIP6 models will be helpful here. We should note that we are planning on setting up identical experiments using the CESM2 model. This will provide an important opportunity to assess the robustness of our findings.

b. The varying methods by which CMIP6 models deal with ozone (Table 1 of Keeble et al. 2021; may also provide a pathway to check whether this behavior exists within the CMIP6 ensemble. This is probably more difficult to do in practice (as many processes vary among models), but perhaps it's something worth looking into. In particular, the MRI-ESM2.0 model significantly diverges from the others in capturing the evolution of ozone over the historical record, suggesting that its jet response may be very biased compared to other models.

We agree with the reviewer that the diverging ozone evolution of the MRI-ESM2.0 model over the historical period is very interesting. However, we remind the reviewer that the stratospheric ozone response to 4xCO2 is structurally very different from historical ozone changes (dominated by SH polar ozone depletion). Thus, while it could be interesting to compare, for example, the single-forcing historical runs (with and without ozone depleting substances), we believe that analysis of the historical runs is somewhat tangential to the focus of our study. Rather, analysis of the SSP runs or the 4xCO_2 experiments would be much more relevant. However, even for these forcing scenarios, the absence of a clean NINT vs. LINOZ comparison would make it difficult to confidently attribute any jet (or AMOC) differences to ozone. Again, the CESM2 experiments will be illuminating.

For these reasons, I'm suggesting major revisions. Other minor (line-by-line) comments/suggestions are listed below.

Minor (Line-by-Line) Comments

Introduction: A lot of this discussion in the Introduction refers vaguely to the NH midlatitude jet. As the authors make apparent in their results, the NH jet often reflects competing behavior between the

Atlantic and Pacific. It would be good to clarify in the discussion of these previous studies, whether they are referring to the zonal-mean jet or explicitly one of the jets over the two ocean basins.

We thank the reviewer for this comment and for their more general emphasis on the need to clearly report the zonally asymmetric response of the jet (see our previous response to the first major comment). In the Introduction we have added a statement clarifying that the NH midlatitude jet equatoward shift reported in CP2019 and Li and Newman (2022) was not demonstrated to be robust over the Pacific (we already indicated that it was primarily a North Atlantic response). We have also clarified that Zhang et al. (2023) identified a poleward shift in the zonal mean that was expressed as an eastward extension of the Atlantic jet over Europe and as a poleward shift of the Pacific jet, very similar to our "total" response captured in Figure 6b. Please see the revised manuscript.

Introduction: It would also be good to clarify in the literature review, for studies that relied on a single model to investigate ozone influence on the jet (CP19, Li and Newman 2022, Zhang et al. 2023), which model was used. In other words, is this behavior mostly been noted before in the GISS model? Or other models as well?

Good point – while the GISS model was used in Zhang et al. (2023), CP2019 used WACCM and Li and Newman (2022) used GEOS. Note that these models have very different historical lineages and do not derive from either other. We have added a sentence mentioning this.

Lines 29-34: The jet shifts, especially in the Northern Hemisphere, are seasonally dependent and vary by basin (Atlantic vs. Pacific). See, for example, Fig. 12 of Barnes and Polvani (2013).

We completely agree. As part of this round of revisions, we hope to convey to the reviewer that we very much appreciate the zonally varying component of our story.

Lines 44-46: It would be good to clarify the relationship of these ozone changes with the associated changes in the Brewer-Dobson circulation.

Excellent point. Indeed, CP2018 show a strong relationship between the strength of tropical lower stratospheric residual mean upwelling and the ozone changes in four different models (their Figure 6, reproduced below):

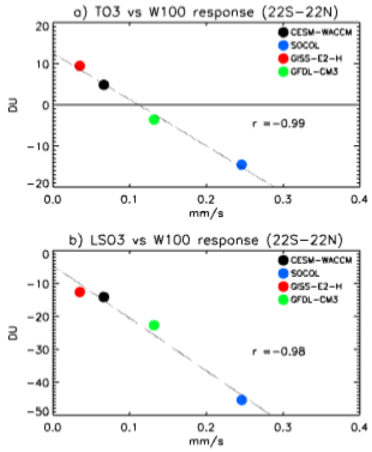
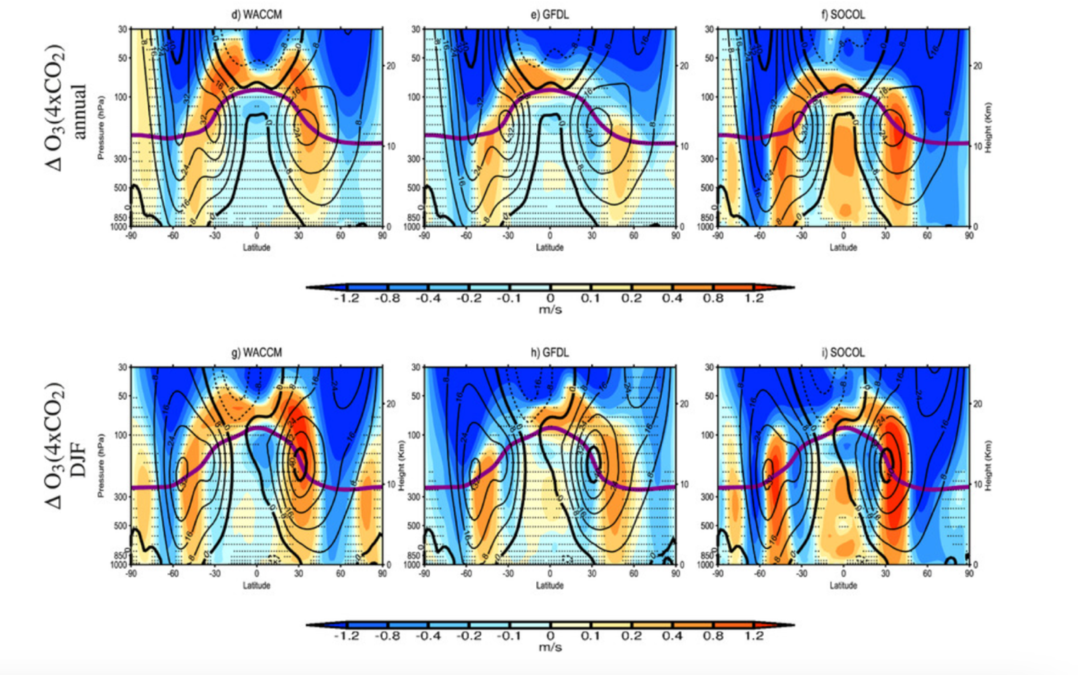


FIG. 6. Scatterplot of upward velocity ($\overline{w}$) change at 100 hPa in response to 4×CO_2 and (a) total column ozone, and (b) lower-stratospheric ozone column-averaged in the tropical region (22°S–22°N).

This figure shows that models simulating a stronger BDC response tend to simulate a greater reduction in lower stratospheric ozone (panel b)). We have now added a sentence discussing this relationship between BDC strength and ozone changes. Please see the revised manuscript.

Lines 53-57, 217-219: It would be good to clarify whether these influences are year-round, or only occur during certain seasons. One would expect stratosphere-troposphere coupling to be most active during seasons when the stratospheric polar vortex is westerly.

Excellent point. We agree with the reviewer that one should expect a strong seasonality to this coupling. We have now added “boreal winter” and appropriate references to these sentences. In addition, we highlight to the reviewer this figure from CP2019 which shows that a consistent ozone feedback on the NH jet is only evident in DJF and not in the annual mean:



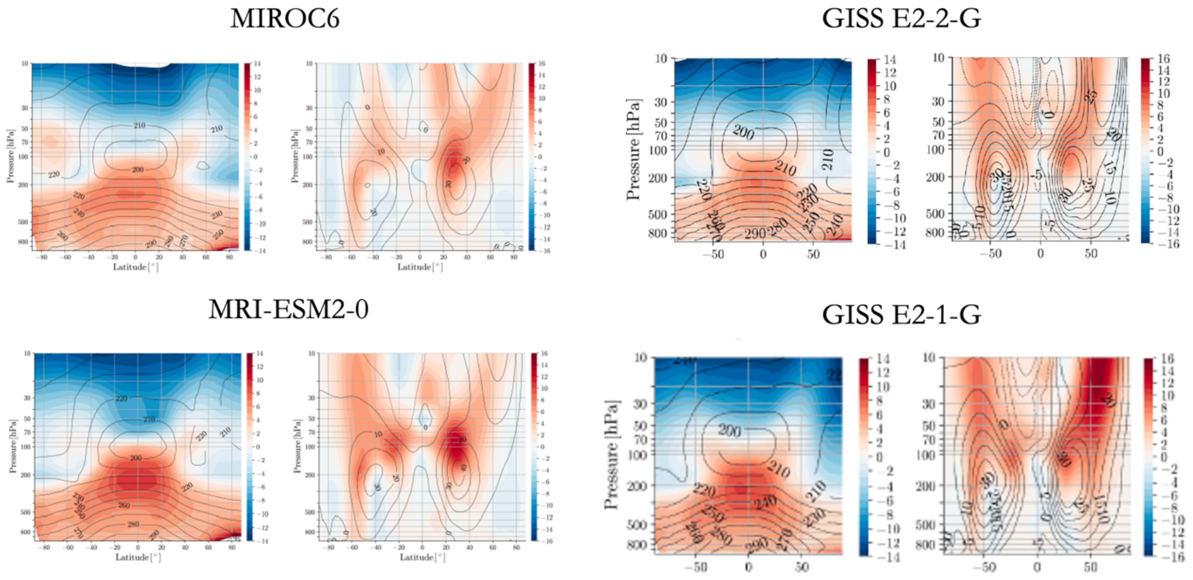
Lines 193, 258, 281: The fast response shown in Fig. 2 does not appear to correspond to the year 5-20 period definition of the fast response listed on line 193.

We thank the reviewer for catching this plotting mistake. Indeed, the light grey boxes should start at year 5. We have now fixed this. Please see the revised manuscript.

Section 2a: It would be good to discuss how representative the zonal wind responses to abrupt CO2 forcing in the GISS model are of CMIP6 models in general.

We are not aware of any exhaustive analysis of the zonal wind 4xCO$_2$ response which comments on the behavior of the GISS model, but we have pulled some data from the ESGF archive to perform a preliminary comparison. This figure compares the DJF temperature and zonal wind responses between the GISS E2-2-G model used in this study (top right) and three other CMIP6 models, i.e., MIROC6, MRI-ESM2-0 and GISS-E2-1-G. Comparison among these models shows that GISS E2-2-G captures a similar response as the other models and there is nothing suggesting that it is an outlier. Indeed, the MIROC6, GISS E2-1-G and MRI-ESM2-0 results seem more dissimilar than the GISS E2-2-G results from any of these models.

4xCO$_2$ DJF Temperature (Left Panels) and Zonal Winds (Right Panels)



Note that we have included the GISS E2-1-G results also to highlight how we intentionally did not use the latter in our study because we wanted to use a higher vertical resolution version of the model (102 vs. 40 levels). We have shown in previous studies that GISS E2-2-G has a much more credible stratospheric circulation, compared to E2-1-G (Orbe et al. (2020)), which is important for studying stratospheric ozone feedbacks.

Orbe, Clara, David Rind, Jeffrey Jonas, Larissa Nazarenko, Greg Faluvegi, Lee T. Murray, Drew T. Shindell et al. "GISS model E2. 2: A climate model optimized for the middle atmosphere—2. Validation of large-scale transport and evaluation of climate response." *Journal of Geophysical Research: Atmospheres* 125, no. 24 (2020): e2020JD033151.

Line 251: Please clarify how these values of "AMOC nonlinearity" are calculated.

By “AMOC non-linearity” we refer to the fact that the response of the AMOC to 4xCO$_2$ does not equal twice the response of the AMOC to 2xCO$_2$. The non-linearity is then equal to the difference $1/2(\delta(4xCO_2))- \delta(2xCO_2)$ for any field of interest. We have now added a line in the manuscript noting this definition and appropriate references. Note that this definition of linearity is consistent with its usage in the following studies:

Chadwick, Robin, and Peter Good. "Understanding nonlinear tropical precipitation responses to CO$_2$ forcing." *Geophysical research letters* 40, no. 18 (2013): 4911-4915.

Mitevski, Ivan, Clara Orbe, Rei Chemke, Larissa Nazarenko, and Lorenzo M. Polvani. "Non-monotonic response of the climate system to abrupt CO2 forcing." *Geophysical research letters* 48, no. 6 (2021): e2020GL090861.

Zhang, Xiyue, Darryn W. Waugh, and Clara Orbe. "Dependence of Northern Hemisphere Tropospheric Transport on the Midlatitude Jet under Abrupt CO2 Increase." *Journal of Geophysical Research: Atmospheres* (2023): e2022JD038454.

Orbe, Clara, David Rind, Jeffrey Jonas, Larissa Nazarenko, Greg Faluvegi, Lee T. Murray, Drew T. Shindell et al. "GISS model E2. 2: A climate model optimized for the middle atmosphere—2. Validation of large-scale transport and evaluation of climate response." *Journal of Geophysical Research: Atmospheres* 125, no. 24 (2020): e2020JD033151.

Lines 328, 339: This value of 3K does not appear to correspond to what is shown in Fig. 4, where the maximum value on the color bar is 2K.

Apologies – we thank the reviewer for catching this mistake. We have now corrected this in the manuscript.

Lines 330-331: How are cooler temperatures at higher latitudes associated with increased radiative heating?

We again thank the reviewer for catching this mistake. The sentence should have read "…in our model the cooler temperatures in the tropics (20S-20N) and the warmer temperatures over high latitudes (>40N) are respectively associated with reduced and increased radiative heating…". We have corrected this in the manuscript.

Line 349-352: Is the meridional temperature gradient enhanced or reduced? There is inconsistency between these two sentences.

We thank the reviewer for catching this mistake. The meridional temperature gradient is *reduced*. We have corrected this is in the manuscript, noting that there was another point where we erroneously referenced to enhanced, not reduced, gradients.

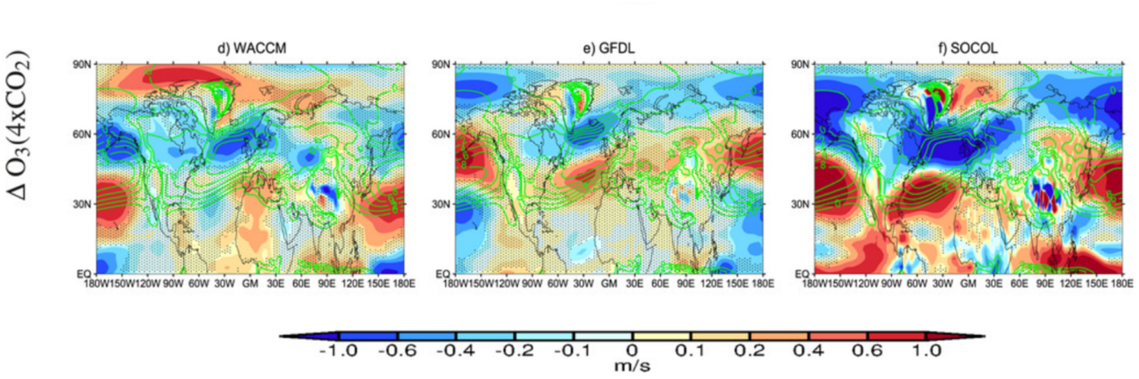
Lines 366-367: I'm not sure I would characterize this as a weakening of the jet at all longitudes. The weakening really is focused over the Atlantic. If anything, the Pacific jet strengthens in Fig. 6a.

Our reference to "all longitudes" applies to latitudes poleward of 60^{o}N. We now clarify this point and the fact that the weakening of the midlatitude jet is concentrated only over the Atlantic ocean. Please see the revised manuscript.

Lines 373-379: While I generally agree with the authors' interpretation in this paragraph, this is not actually what is shown in Fig. 6a. The signal is actually more statistically significant in the Pacific, compared to the Atlantic.

The point we are making is that the OMA and NINT responses over the Pacific are different, regardless of whether they are statistically significant for each respective chemistry configuration. This can be seen by comparing Fig. 5a with Figure 3c.

Note that this is not surprising as it was already discussed in CP2019. Consider, for example, their Figure 6 (shown also in response to a previous comment). This figure shows very similar negative NAO-like responses over the North Atlantic, but very different results over the Pacific.

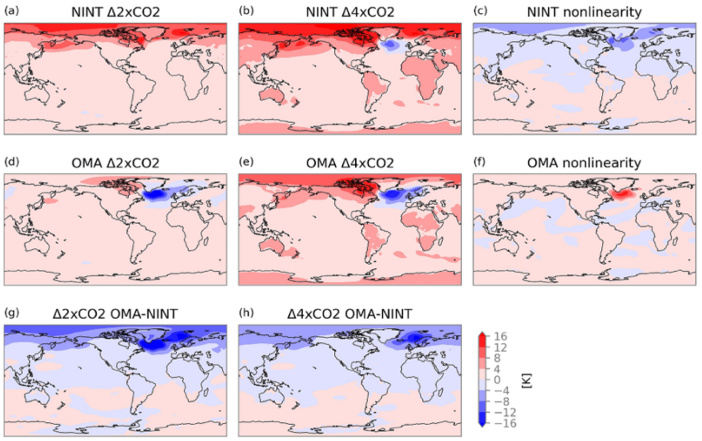


Line 381: I would specify northern North America here, as there is actually cooling over much of the United States and southern Canada.

Right – thanks for pointing this out. We have corrected this in the manuscript.

Line 399: How do you know this? At least with the color scheme on Fig. 6d, there is no way of knowing that the cooling is strongest over the North Atlantic warming hole.

Fair point. We know this based on our examination of this figure without using such a restrictive colorbar. This was also reported in Zhang et al. (2023) for the OMA-NINT comparison – see panels G and H in their Figure 1 (reproduced below). We now reference that study at this point in the manuscript. Note that our use of such a limited colorbar scheme is to maintain consistency with the colorbar used to depict the ``fast" response (which is the same one used in CP2019).



December–January–February surface air temperature response for non-interactive (NINT) (a) 2xCO$_2$, (b) 4xCO$_2$, (c) nonlinearity, and one-moment aerosol (OMA) (d) 2xCO$_2$, (e) 4xCO$_2$, (f) nonlinearity. Differences between OMA and NINT responses are shown on (g) for 2xCO$_2$ and (h) for 4xCO$_2$.

Lines 399-401: Based on Fig. 6b, I would not agree with this assessment. If anything, there still appears to be an equatorward jet shift over the Atlantic. The biggest differences between Fig. 6a and Fig. 6b appear over east Asia and the west Pacific, not the North Atlantic. This appears to be the origin of the poleward jet shift seen in the zonal mean, not the North Atlantic.

We agree and thank the reviewer for their careful reading of the manuscript. As mentioned in response to an earlier comment we now have combed through the text to make sure that our depiction of the jet changes, on both fast and long timescales, is more accurately reflective of the changes over individual ocean basins and over land.

Figure 4, bottom row: Why is there troposphere cooling in these figures? Is this the total temperature response to 4xCO2, in which case there should be warming at tropospheric levels? Or, are these the differences from the NINT response? Please clarify this in the figure caption.

Our apologies. Yes, the bottom panels show LINOZ/OMA-NINT. We have changed the subtitles in the figure. We thank the reviewer for catching this mistake.

Figures 12 c,d: Why is the map domain in these figures different from Fig. 6? It makes it difficult to compare the results in the two figures (Fig. 6 vs. Fig. 12).

We agree with the reviewer. We now show all longitudes in Figure 12 c and d. Please see the revised manuscript.

Color bars: I hesitate to make this comment, as I know it will be a pain to fix. But, the number labels on the color bars have an extremely small font size, which would make them almost impossible to read for someone who printed a PDF of the final paper.

We completely agree with the reviewer! We have increased the font sizes of all colorbars in all figures. Please see the revised manuscript.

Typos

Line 157: row 9, not row 8

Thanks – this has been fixed.

Line 228: weakened zonal winds in NINT

Thanks – this has been fixed.

Line 482: Fig. 6a, not Fig. 6c (top)

Thanks – this has been fixed.

References

Barnes, E. A., and L. Polvani, 2013: Response of the Midlatitude Jets, and of Their Variability, to Increased Greenhouse Gases in the CMIP5 Models. J. Climate, 26, 7117-7135.

Keeble, J., Hassler, B., Banerjee, A., Checa-Garcia, R., Chiodo, G., Davis, S., Eyring, V., Griffiths, P. T., Morgenstern, O., Nowack, P., Zeng, G., Zhang, J., Bodeker, G., Burrows, S., Cameron-Smith, P., Cugnet, D., Danek, C., Deushi, M., Horowitz, L. W., Kubin, A., Li, L., Lohmann, G., Michou, M., Mills, M. J., Nabat, P., Olivié, D., Park, S., Seland, Ø., Stoll, J., Wieners, K.-H., and Wu, T.: Evaluating stratospheric ozone and water vapour changes in CMIP6 models from 1850 to 2100, Atmos. Chem. Phys., 21, 5015-5061.