**Author response to reviews of manuscript "Model spread in multidecadal NAO variability connected to stratosphere-troposphere coupling"**

We thank the Editor for sourcing two thorough reviews of our manuscript. We are pleased both reviewers find merit in our study and support its publication in WCD subject to minor revisions. We have responded to the points raised below. We hope the Editor will find the revised manuscript improved and ready for publication.

**Response to Reviewer 2**

**The manuscript assesses multidecadal variability of the winter North Atlantic Oscillation (NAO) using historical simulations from 15 different climate models, each with at least 10 members. It is found that NAO variance is underestimated by models and potential causes are investigated. It is shown that NAO variance correlates significantly across the models with both the variance of the stratospheric polar vortex (SPV) and the coupling between the SPV and the NAO. Though causality cannot be identified, together these two factors explain 70% of the inter model spread in NAO variance. Furthermore, the coupling between the SPV and the NAO appears to be related to a measure of atmospheric eddy feedback. The authors also investigate other relationships finding no link to the Pacific, and a weaker though significant link to Atlantic Multidecadal Variability but only by processing in a way that precludes comparison with observations.**

**The manuscript is well written and the results are interesting. I recommend publication after addressing the minor comments below.**

First, we would like to thank the reviewer for her/his careful reading, interest in our study and the insightful comments that helped us to improve the manuscript. Figure S2 has been modified, the two missing models were added and we've changed the color code to match the figure on the article. Please find below the answers to the comments point-by-point. For clarity, all reviewer comments are in **bold** and responses in **blue**.

**Table 1: A minor point but I believe there are 50 ensemble members for MIROC6, or are the data you need not available?**

Indeed, not all members of the large ensembles provided daily wind data on the ESGF at the time of analysis. We added a sentence to clarify this in the revised manuscript:

"Therefore, we analyse CMIP6 models providing at least 10 ensemble members for the DECK historical experiment that provide daily zonal wind (ua) and meridional wind (va) variables needed to calculate the eddy feedback parameter described in Section 1 (see Section 2.2.5). Note that some models have additional ensemble members in ESGF which are not included here because they did not provide daily wind data at the time of analysis."

**Lines 143-144 Eddy feedback parameter: looks like something is missing here?**

Indeed, this was a problem with the .pdf file. We removed the "obj" that were in the PDF and we also added the unit hPa unit after 500 in the revised manuscript.

**Lines 158-159: please explain how serial correlation in the timeseries is accounted for**

As we do not test the significance of the correlation between variables (e.g. the AMV and NAO) within a single member, we do not need to take account for autocorrelation in the significance testing. All members are assumed to be independent and so the standard error on the ensemble mean calculated from the interensemble standard deviation follows the equation on L157. Each model is assumed to be independent for the purposes of the significance tests to assess intermodel relationships (L158-159).

**Fig 2: the model labels don't seem to match up with the dots - at least in my version**

Thanks. The labels are now fixed in the revised manuscript.

**Fig S2: looks like not all of the models are included here?**

Indeed, two models were missing and have now been added to the revised manuscript. The colors used have also been modified to match the rest of the manuscript.

**Fig 5a: is the ensemble mean the average of the regressions for the individual ensemble members, or is it the regression between the ensemble means of the NAO and AMV? I think it is the former (this needs clarifying) but the latter would be highlight forced responses and might also be interesting (though not so easy to compare with observations).**

In Figure 5a the ensemble mean is indeed the average of the regressions calculated for each individual member and not the regression between the ensemble means. We did not put the ensemble mean as it would not be directly comparable with the observations. It is now clarified in the legend of the Figure 5a:

"Figure 5: (a) Regression slope between the 20-year running mean NAO and AMV indices for DJFM over 1900-2010 for each model ensemble member (dot) and the ensemble mean of these regressions (diamond)."

**Fig 5b: if I am reading it correctly it looks like the strongest observed regressions occur with AMV leading NAO? If so, I wonder how that relates to previous work suggesting AMV can be explained by the integrated NAO e.g., https://journals.ametsoc.org/view/journals/clim/32/22/jcli-d-19-0177.1.xml**

This study and previous work by Li et al. (2013) relate low-pass filtered (7-year for O'Reilly et al., 2019 and 11-year for Li et al., 2013) AMV to the integrated NAO over all previous timesteps. This is different from our approach which uses 20-year running means for both AMV and NAO, so this could explain the apparent difference. Furthermore, O'Reilly et al. (2019) find around 60% of the AMV variance is not explained by the integrated NAO. In the manuscript we show the regression slope as a function of lead-lag which may not describe the statistical strength of the relationship. In Figure R9 we show the correlation coefficient of the NAO-AMV regression, which does reveal the strongest correlation with AMV leading the NAO by 10 years and weaker correlation with the NAO leading AMV consistent with the magnitude of the regression slopes in Figure 5b of the manuscript.

Meanwhile, both observational and modeling studies have shown that a positive AMV (associated with a stronger AMOC) can also induce a negative winter NAO response in the atmosphere, as seen in Figure R9. Peing and Magnusdottir (2016), for example, showed that the positive phase of the AMV promotes negative NAO in winter using three different configurations of the Community Atmospheric Model version 5. Based on the long-term NOAA 20CRv3 (Slivinski et al., 2019) and ERA20C (Poli et al., 2016) reanalyses, Kwon et al. (2019) also showed the influence of the AMV on the NAO at multidecadal timescales.

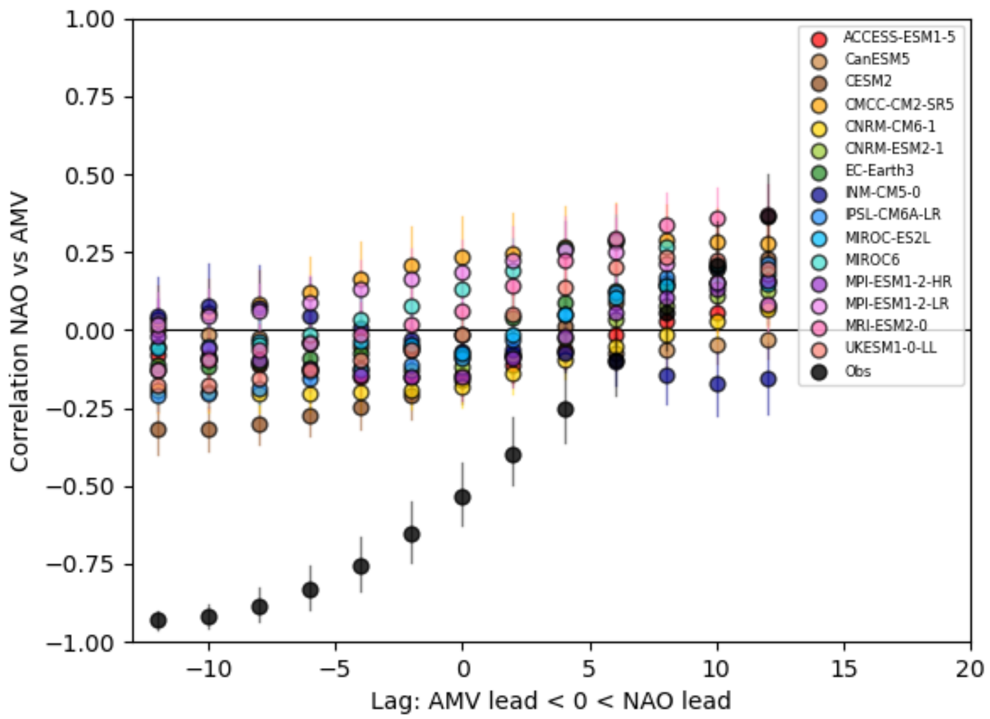


**Figure R9:** *Lead-lag DJFM correlations between the 20-yr running mean AMV and the 20-yr running mean NAO for each model and the observations (dot) and their relative uncertainties (bars) calculated over the 1900-2014 period.*

**Fig S3: the caption is a bit confusing. Presumably this is the average of the values for each ensemble member, since the ensemble mean has been removed?**

Indeed, it is the average of the member values for each ensemble. We clarified the caption in the revised manuscript to avoid confusion:

"Figure S3: Scatter plot of the (a) average low frequency NAO variance (hPa²) and versus the regression slope calculated between the AMV and the NAO from the members for each of the ensemble of climate model simulations for the DJFM months over the 1850-2014 period. The ensemble mean is removed for both variables. The black line represents the least square regression with Pearson correlation r and p-value (see Section 2.3). (b) Same as (a) using a 10-yr running mean."

**Line 359: the values given in fig 9b appear to be different, r=0.52 p=0.05**

Sorry for this mistake, it is changed in the revised manuscript

**Discussion: I presume that constraining the NAO variance (lines 393-394) is beyond the scope of this study, though it would enhance it if it could be done.**

We considered using an emergent constraint methodology similar to Simpson et al. (2021), however, since the constrained NAO variance would still be systematically biased low wrt observations, this seems of limited value. We did not wish to mislead readers that a constrained distribution of NAO variability is more accurate given the evidence for other systematic biases.

**Discussion: Fig 5b appears to show further evidence of errors in all models. Perhaps a brief discussion of potential causes and implications could be included.**

Thanks for the suggestion. We have added to the sentence "This overall weak role for the AMV on the NAO may be related to weak atmosphere-ocean coupling in models as has been suggested in other studies (e.g., Simpson et al., 2018)." with "Indeed, the analysis in Figure 5b shows all models analysed fail to capture the lead-lag relationships between the NAO and AMV compared to observations. The amplitude of the NAO-AMV regression slope is too weak in the models with AMV leading the NAO, and has the opposite sign to observations when the NAO leads the AMV. Consistent with earlier studies (e.g. Kim et al. 2018; Simpson et al., 2018; Bracegirdle, 2022) this points to systematic errors in the modelled representation of atmosphere-ocean coupling in the North Atlantic on multidecadal timescales."

**References:**


Kwon, Y. O., Seo, H., Ummenhofer, C. C., & Joyce, T. M. (2020). Impact of multidecadal variability in Atlantic SST on winter atmospheric blocking. *Journal of Climate*, *33*(3), 867-892.

Li, J., Sun, C., & Jin, F. F. (2013). NAO implicated as a predictor of Northern Hemisphere mean temperature multidecadal variability. *Geophysical research letters*, *40*(20), 5497-5502.

Peings, Y., & Magnusdottir, G. (2016). Wintertime atmospheric response to Atlantic multidecadal variability: Effect of stratospheric representation and ocean–atmosphere coupling. *Climate dynamics*, *47*, 1029-1047.

Poli, P., Hersbach, H., Dee, D. P., Berrisford, P., Simmons, A. J., Vitart, F., ... & Fisher, M. (2016). ERA-20C: An atmospheric reanalysis of the twentieth century. *Journal of Climate*, *29*(11), 4083-4097.

Slivinski, L. C., Compo, G. P., Whitaker, J. S., Sardeshmukh, P. D., Giese, B. S., McColl, C., ... & Wyszyński, P. (2019). Towards a more reliable historical reanalysis: Improvements for version 3 of the Twentieth Century Reanalysis system. *Quarterly Journal of the Royal Meteorological Society*, *145*(724), 2876-2908.