

Dear editor and authors,

The Vogt et al. manuscript offers a step-change in our understanding of the relationship between the preindustrial control state of current generation models and their ocean heat uptake under anthropogenic forcing. The manuscript is really well written. Apart from some minor clarifications, there are two important issues that I would like to see addressed.

First, I would like the authors to probe more deeply into their first principal component of intermodel variance in the preindustrial annual mean stratification. It seems to me that EOF1 is very much related to the strength of the ocean gyres and may point to intermodel differences and biases in the surface wind-stress curl.

Second, I would like the authors to clarify whether some individual CMIP models exhibit a bias towards convection in the Labrador Sea while underestimating processes in the Irminger Sea. How do such intermodel differences correlate with the AMOC in the preindustrial and forced experiments?

I therefore recommend **minor revisions**.

Major comments:

Lines 269-271 and the Synthesis section: Are you sure that "broadly uniform" is the best way to describe the first EOF of preindustrial annual mean stratification? I would argue that EOF1 is very much related to the strength of ocean gyres, with the pattern being weighted towards the stronger subtropical gyres compared to the weaker ones.

Is EOF1 related in any way to the near-surface component of the large-scale atmospheric circulation? Does it represent systematic biases and intermodel spread in the surface wind stress curl?

Line 210: Are there some models that exhibit convection only in the Labrador Sea and others that have convection in both the Labrador and the Irminger basins in their control state? Is there also a connection between the spatial pattern of the preindustrial North Atlantic MLD and the AMOC weakening in response to forcing?

Minor comments:

Line 42: add "in a historical context" to distinguish from future scenarios

Line 148: Specify what time period of the preindustrial control run do you include in the averaging? Is it only the period preceding the branching, the period starting at the same time as the branching, or the full preindustrial control simulation?

Lines 150-151: Does this process leave you with 23 models, or you get 28 models left after the elimination of outliers?

Line 158: 28 models before or after the removal of outliers?

Lines 518-520, this manuscript is now published, and the updated reference is:
Gregory, J.M., Bloch-Johnson, J., Couldrey, M.P. et al. A new conceptual model of global ocean heat uptake. *Clim Dyn* 62, 1669–1713 (2024). <https://doi.org/10.1007/s00382-023-06989-z>