Response to Reviewer #3

We thank the reviewer for his very constructive comments and suggestions as well as for the time spent reviewing the paper. In the following, the reviewer's comments are in black, and our responses are in red.

This paper presents a coordinated comparison of modelled and observed overturning transports across OSNAP-West and -East sections, using a large number of forced OMIP-type simulations spanning a wide range of ocean/sea-ice models and resolutions. The paper is explicitly framed as descriptive and benchmarking, rather than mechanistic explorations, and is largely successful in achieving that goal. The topic is highly relevant for the ocean and climate modelling community, and it will provide valuable insights into how well ocean models capture the OSNAP overturning circulation in both depth and density space. The paper is overall well-written, and I only have some minor/technical comments.

Thank you for your kind words, appreciating the relevance and value of the manuscript for the community.

- Near the end of the Introduction, the authors clearly state that the paper is "largely descriptive", which is welcome; however, before that, there is heavy text with extensive discussions of, e.g., the control of Labrador vs. Irminger Seas, which might make it read like a hypothesis-driven AMOC mechanism paper. I am not sure if/how this could be improved, though - how about clarifying this somewhere earlier in the text? I am not sure if the authors share a similar view, and I will leave it to them to decide.

  We think that it is important to give a relatively comprehensive overview of the AMOC and its impacts, variability, and drivers to provide a solid context to discuss the implications of the new findings from the OSNAP observations. This sets the stage for the importance of our benchmarking as well as our assessment of model fidelity. We now make the goals of the manuscript clearer at the beginning of this paragraph, including *assessment of model fidelity* as part of our effort.

- The differences in results between sigma_0 and sigma_2 space (such as the relative contributions from OW) are important findings. For a broader audience, it would be beneficial if the authors, where appropriate, could expand and articulate a bit why using sigma_2 alters the results.

  Such an explanation had been provided in the 4th paragraph of Section 2.5. In response, we added a reference to this section in the conclusions section (Section 9).

- L47, "…ocean - sea ice simulations…"

  Corrected.

- L206, "compensation"

  Corrected.

- L344-349, the authors claim that they "analyze the same forcing cycles for a given LR and HR set of simulations from the same group"; however, this does not seem to be the case in the following descriptions. For example, ANU10 (HR) is not in the same cycle as its LR counterparts of ANU1 and ANU25.

  This has been noted as an exception.

- L489, "FSU72"

  Corrected.

- Figure 7 and Figure 8 (and other pairs of figures), here the figures are not separated as HR vs. LR (as in previous text and figures), but rather separated alphabetically - is there any rationale behind the change? It would be helpful to explain a bit, or at least mention this change in the text.

  In these more extensive figures, we adopted an alphabetical ordering to expedite comparisons within a set of simulations from the same group. This is now mentioned at the beginning of Section 5.

- Fig. 18, if I understand correctly, the MOC here is the total, i.e., OW+OE; is there a reason not to separate it, as in Fig. 17? Also, would it be useful to add the OSNAP to the scatter plots, as in Fig. 17?

  Your understanding is correct: Original Fig. 18 uses the total overturning transports. In contrast to original Fig. 17 which shows the biases at the actual OSNAP sections, Fig. 18 is for the broader regions depicted in Fig. 1. For this reason, we think that it makes more sense to use the total transports in this figure. For the same reason, we do not plot the OSNAP values in Fig. 18.