General comments:

This paper analyses the variations of the ocean heat content averaged over the first 700 m (OHT700) of the ocean into an area surrounding the Labrador Sea region for the period 1970-2014. For this purpose, it uses observational data available and two multi-model sets of climate simulations, one with only the external forcing (historical simulations) and the other with decadal hindcasts starting from observation-based estimate. The analysis shows a very wide range of response in the models, especially for the historical simulations. The authors try to estimate the skill of the different systems to reproduce the OHT700 in decadal prediction and historical simulations, and found an interesting link between this skill and the capability of the models to reproduce observed mean state of stratification and ocean heat fluxes in the Labrador area.

This is an interesting and well-written paper. The analysis led is impressive given how difficult it is to deal with so many climate model data. The interpretation of the results is wise and useful, even though no definitive conclusions can be drawn from this type of multi-model analysis. At least, this is presenting an interesting intercomparison of present-day models to reproduce heat storage in the Labrador Sea area and a few interesting predictors that might be of use for observations constraints approaches. I therefore think this paper is suitable for publication. I have mainly some comments that might allow to strengthen the demonstrations and possibly improve the interpretation of the results.

Reply: We thank the reviewer, Didier Swingedouw, for his constructive comments on our manuscript. We greatly appreciate the time and effort he has invested in reviewing our work. His feedback has been invaluable in improving the quality and clarity of our paper, and we believe the revisions have significantly strengthened our manuscript.

We have carefully addressed each of the suggestions and have provided a detailed, point-by-point response to all the comments.

1. Line 35-40: here the authors are mixing discussions about the subpolar gyre and the wider North Atlantic and ocean heat content and SST. It might be worth to be a bit more specific in the description of those papers.

   Reply: We have rewritten this part of the introduction to make a more clear distinction between the regions and variables.

2. Line 61: a reference after forecast range might be useful to support this claim.

   Reply: This part of the introduction has been removed following a suggestion of the other reviewer.

3. Line 202: The LS, as represented in Figure 2, does not entirely correspond to the Labrador Sea but is going far the east, including the Irminger Sea for instance. In this respect the agreements between observation-based datasets are not that clear to the east (cf. Figure A.1), while the good agreement is taken as a reason to focus on this region in line 204. Please clarify. Have the authors tried a more tied region?

Reply: The reviewer is correct. We had not noticed that the large LS box considered includes areas where the reanalyses and EN4 show significant discrepancies in terms of OHC variability. We no longer refer to the agreement between those datasets as one of the reasons for focusing on this region. However, we do not believe that the local areas of disagreement between the observation-based datasets affect our results, as we are using regional averages and the datasets largely agree over most of the selected area. Additionally, the main reason for using this broader region, which includes both the Labrador Sea and the Irminger Sea, still holds. This area exhibits large inter-model differences in OHC skill in the DCPP ensemble across all forecast ranges, as well as in the HIST ensemble. It is also a characteristic region of deep vertical ocean mixing, with common preconditioners and drivers, whose representation varies greatly across models, contributing to the inter-model spread (as shown later in Figure 11). For these reasons, we have decided to retain the entire region in our analysis.

4. Line 249-253: ocean stratification and heat fluxes are two variables clearly linked in the convection region. If the halocline is too strong, convection is not allowed and heat fluxes can lead to sea ice formation. It might be worth to state this coupling between these two variables (maybe in the discussion).

   Reply: We now comment this in the text.

5. Line 266: it is said line 155 that density is computed with reference 1000 m (sigma_1) while in Figure 6, the caption talks about reference to the surface. Given that the numbers in Figure 6 are larger than 28, I assume this is actually sigma_1. This choice is surprising given that then authors are focusing on the very upper layer. I think it might be better to consider sigma_0 as stated in the caption (while it is not what is shown).

   Reply: The reviewer is correct: it is the caption of Figure 6 that was incorrect. Indeed, we have only used sigma_1 in the study. We understand the suggestion of using sigma_0 instead, but still believe that sigma_1 is more suitable, because we use it to understand and characterise the mixed layer in the Labrador Sea, whose influence goes beyond the upper ocean levels. We have corrected the caption in the revised manuscript.

6. Line 291-296: why are the observations not shown on Figure 7?

   Reply: We prefer to keep everything consistent within the figure, to ease its interpretation. If we included an additional panel with observations it would not have crosses as for the HIST simulations, since there is only one instance of observations as opposed to the different realizations in HIST. For that reason, we show observations elsewhere (i.e. new Figure A6).

7. Line 395-400: The use of residual ACC (Scaife & Smith 2018) might be interesting as well. I'm wondering if this might work for this type of complex quantity like OHT700, especially given the complexity of its forced response. A discussion on this aspect might be interesting here I think.

Reply: We have added a figure with the residual correlations (new Figure A3) and a discussion addressing the validity of that metric for the OHC700. To complement the residual correlations, we have also included another figure with the differences in ACC between the DCPP and HIST experiments (new Figure A2). The main takeaway from the new analysis is that residual correlations help to identify more clearly the added value of initialization in the Labrador Sea and the Eastern North Atlantic areas, which is less evident when directly comparing the DCPP and HIST ACC maps. We also note that the results of the residual ACC need to be interpreted with caution, as some areas show large inter-model uncertainties in terms of OHC skill for the HIST ensemble. This implies that some models do not correctly represent the observed forced signal, a requirement for the residuals correlation to be meaningful.

8. Line 412-416: I have the feeling that this aspect has not been much depicted in the result section, so that this discussion seems a bit coming out of the blue. Maybe useful to add a few points on this in the results section.

   Reply: We now address the differences in OHC skill between full-field and anomaly-initialized prediction systems in the discussion around Figure 1. Also, the specific benefits of each initialization approach for the representation of mean stratification in the Labrador Sea and the local NAO forcing at different forecast times are discussed in their respective results sections.

9. Line 431: Yes, the omission of advective processes is clearly missing in this paper, but I can understand that it is far from easy to have those quantities from such a large ensemble of simulations. What about citing Ortega et al. (2015) that was also discussing this type of processes in details? + typo at “mechanisms”.

   Reply: Thanks for spotting the typo and for suggesting the reference. The typo has been corrected, and the reference is now cited.

References

Scaife A.A., Smith D., 2018. A signal-to-noise paradox in climate science. npj Clim. Atmos. Sci., 1, 28. doi: 10.1038/s41612-018-0038-4

Ortega P., et al. (2015) Reconciling two alternative mechanisms behind bidecadal AMOC variability. Progress in Oceanography 137, pp. 237-249.