Response Reviewer #2

The authors have accounted for my comments and have done a very good job to produce an almost publishable manuscript.

Reply: We thank the reviewer for the positive feedback and the additional comments.

I only have a few minor comments that the authors might find of interest.

#1. Line 35: this last statement has not been proved properly in Caesar et al. (2021). To be able to attribute a signal to anthropogenic forcing, you need to compare it with forced signal from model simulations. This has been done in Latif et al. (2022) and Bonnet et al. (2021) who both concluded that internal variability might be more likely to explain the long term trend found in Caesar et al. (2018), not the external forcing. This might be useful to state this. Observed Greenland melting is unlikely to also explain this (e.g. Devilliers et al. 2021). The next paragraph is exploring the role of internal variability, but I think it is misleading since the focus will be mainly on interannual variability, up to decadal variations, while the statement from the paragraph before clearly discusses multi-decadal trend, which might be driven by different processes. I think it might be useful to clarify the different time scale involved and the relevance of this work in this respect.

Reply: We thank the reviewer for the feedback. We now cite all the articles suggested, to give a more complete discussion of the recent AMOC weakening. We also clarify in the following paragraph that the discussion refers to interannual to decadal internal variability.

#2. Line 324 : “sea-ice” should be replaced by “sea ice” as used in the rest of the manuscript when it is referring to a noun (e.g. line 345)

Reply: Done

#3. Line 354: As stated later on, the forced signal from HIST cannot be compared per se with observations, as those ones include both signals (forced and internal). Maybe it's worth clarifying this already here. Also, the motivation of this last part is not very clear to me in the general narrative of the paper. I wonder if a better transition with the former part, or a slight introduction to the motivation of this one might not be useful.

Reply: These are two very good points. We have included a sentence to motivate this final section and also expanded the paragraph to clarify why the forced signal from HIST cannot be directly evaluated from observations.

#4. Line 358 and caption of Fig. 11: It might be worth discussing the FT here. We can see on Fig. 5 that this is almost constant for HIST. Is it entirely or are there some edge effects (maybe not since the time of analysis has been chosen to avoid this). Please clarify. Also the use of ACC for HIST is surprising and might need to be discussed since it is not that straightforward to understand (we usually look into trends when analysing HIST and observations).

Reply: We confirm that the ACC values for the HIST experiments remain constant across all forecast years because the evaluation period used for their computation is fixed. This information is now explicitly indicated in the figure caption. We now also mention the long-term trends in the discussion of Figure 11. We also clarify that we chose to evaluate the forced signals in the models using the ACC metric, as the response to external forcings is not accurately captured by a simple linear trend, particularly due to the episodic nature of several volcanic eruptions during the evaluation period. However, we acknowledge that the main results shown in Figure 11 would likely hold if the metric were replaced by the linear trend, as Figure 4 demonstrates a strong relationship between trend-based skill and ACC.

#5. Line 416: this last point might be better substantiated if observations were also indicated as a grey overlap for instance in Fig. 11, to highlight that models with best ACC in OHC in HIST have very poor representation of stratification and heat fluxes which is actually surprising results when compared to Hegerl et al. (2021)

Reply: Observations are indicated as a dashed black line on the Figure, and support the message that we make in that line.

#6. Line 421: “(Quin et al., 2020)” should be “Quin et al. (2020)”

Reply: Corrected

#7. Line 431: I have the feeling that Fig. 5 does not show this, with the skill of NorCPM1 being not significant from FY 4. Can you please clarify or correct this statement?

Reply: The reviewer is right. Figure 5 clearly shows that this statement does not apply to the whole Labrador Sea region. We have rephrased the text, which now specifically mentions the westernmost side of the Labrador Sea, for which the added value of initialization persists up to forecast year 10 (Figure 1).

Response Reviewer #3

The authors have addressed the reviewers' suggestions thoroughly, and the manuscript is in good shape for acceptance. However, I recommend a minor revision before final acceptance to improve clarity and accuracy.

Reply: We thank the reviewer for the positive feedback and the additional comments.

Specific comments:

#1. For the benefit of future readers who may not be well-acquainted with the decadal prediction system, I suggest adding a few sentences explaining how the DCPP simulations are performed, as well as the methodology behind the Historical simulations in CMIP6. This information is not explicitly stated in the manuscript and would help improve understanding for future readers.

Reply: This is a fair point. We have reorganised and extended the paragraph in the introduction where both DCPP and HIST simulations are explained.

#2. Line 180: The parentheses in "(figure A3)" is missing and should be corrected.

Reply: Corrected

#3. Line 203: The reference to the "western Subpolar Gyre" seems inappropriate, given that the described region, including the Irminger Sea, is significantly larger than what is typically considered the western Subpolar Gyre. If this designation is based on a specific reference, further justification is needed. Otherwise, I recommend deleting this statement and avoiding the term "western Subpolar Gyre" throughout the manuscript. The use of "LS region" is acceptable.

Reply: We have removed the sentence

#4. Figure 3 Caption: There is a typo in the caption where "for" appears incorrectly. Please revise accordingly.

Reply: Corrected