Reviewer #1: Review of JQSR-D-22-00450
"Multi-centennial Holocene climate variability in proxy records and transient model Simulations" By Askjær et al.

This paper presents an investigation of multi-centennial temperature variability based on both proxy records and paleoclimate model simulations. The main results are based on spectral analysis on temperature. The authors find consistent statistical power density peaks on broad multi-centennial timescales, the most prominent at 100-200 years, in both proxy records and model simulations. Although there is an indication of high power density in the Arctic, the oscillation shows largely a global nature. This is an important paper, particularly from the perspective of modern-day climate change. The paper is generally very well written. I recommend accept with minor revision.

I appreciate that is really difficult to figure out the mechanisms behind such a complex phenomenon and the discussion section on potential drivers is very useful. The paper could benefit greatly from a bit of extended analysis on top of temperature as model simulations should have a full set of variables available. I believe these models are coupled atmosphere-ocean models. I am not asking for this to be done before the current paper is published, but an extension or a follow-up paper will be really useful. Is multi-centennial variability internal to the atmosphere-ocean system? I do remember it is prominent in the 3000yr HadCM3 control simulation although we didn't publish a paper. What are individual roles of the different forcings?

Reply:
We thank reviewer #1 for the evaluation of our manuscript. We agree that an extension of the analysis of model data beyond temperature as done in this study will be very interesting to study driving mechanisms and how these propagate through different parts and hence variables of the climate system. We consider this work a starting point i.e. through the analysis of all available proxies suitable for the task and the first comparison with available model runs. Follow-up studies should indeed also include the analysis of steady-state control runs like the mentioned 3000yr run. We have already started to generate different long control runs under different CO2-levels for this purpose and future projects.

Some specific points:
1) It would be useful to include some information about the external forcings and their power spectral properties for the model simulations.
Reply: In this paper we don't intend to discuss the mechanism, therefore we did not do the power spectral analysis on forcings. We plan to investigate how the spectrum varies in the Holocene transient simulations by using wavelet analysis in our following research. In that case, a thorough analysis of external forcings would be required to understand if and how the forcing might influence low-frequent variability.

2) In Fig.3, what are the magnitudes of centennial fluctuations in the models? How do they compare with the IPCC climate sensitivity estimates?
Reply: The magnitude of the fluctuations is different in different models, from visual observation in Fig3, some models show greater fluctuations (e.g. EC-Earth, MPI-ESM and HadCM3) than others (e.g. CESM1, LOVECLIM and AWI-ESM). How the magnitudes of low-frequency variations relate to climate sensitivity is not clear, e.g. both EC-Earth and CESM are the models that have high equilibrium climate sensitivity (ECS) (Haywood et al., 2020), but EC-Earth shows a large magnitude of centennial fluctuation and CESM shows small magnitude (Fig.3). To avoid any confusion, we choose not to discuss the relationship between the magnitude of centennial variability and climate sensitivity.

Reference:
Haywood, et al., 2020. The Pliocene Model Intercomparison Project Phase 2: large-scale climate features and climate sensitivity, Clim. Past, 16, 2095-2123, 10.5194/cp-16-2095-2020.

3) GMT or temperature, do you mean surface air temperature?
Reply: We have mentioned in the manuscript, for some models, it is near-surface air temperature, and for a few models, it is surface skin temperature (see lines 100-102 in the manuscript). The variability is the same.

4) Please add a few sentences at the end of the conclusions to say if there are any potential caveats of the current analysis.
Reply: The following has been added at the end of the conclusion; "In addition, both the model and proxy data vary in their age coverage despite all being located in the Holocene and the proxies are not synchronized, so further research into any potential changes in the variability pattern through time with e.g., wavelet, is a relevant aspect of both future research into potential drivers and testing/elaborating on the results of the analysis of this paper."

Minor points:
Line 45: add "data" after paleoclimate
Line 196: remove "bu" before AMOC

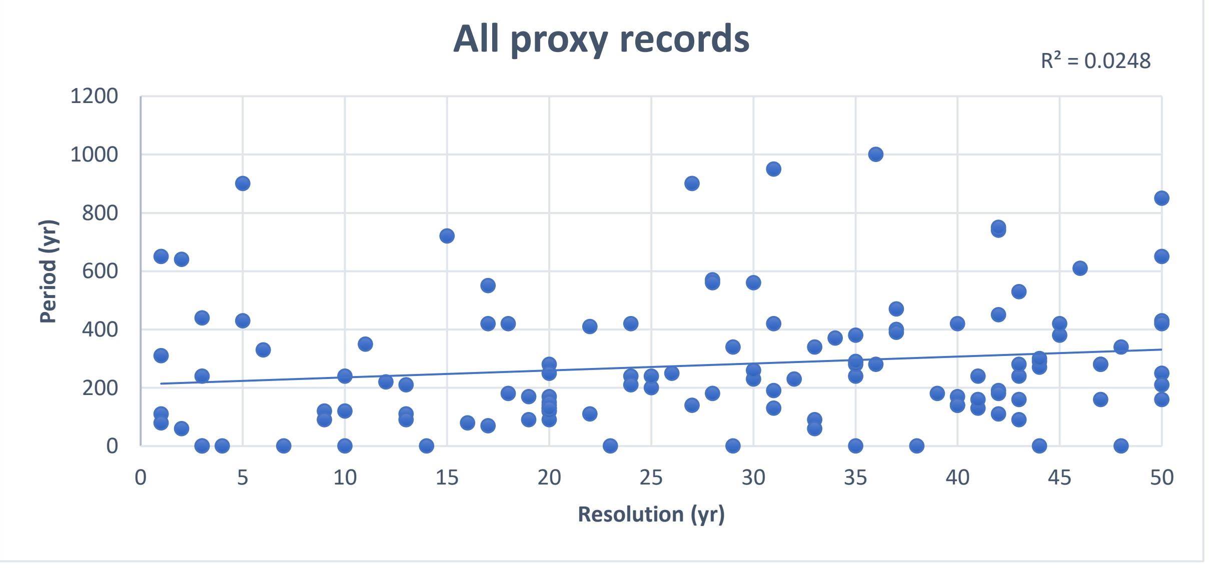
Line 280: do you mean "only have become available"?
Lines 327-330: the first sentence in the conclusions does not read well. Please break it into short sentences and write it properly.
Reply: the minor point has been fixed in the manuscript

Reviewer #2: The manuscript presents a very interesting analysis of the Holocene climate focusing on multidecadal-centennial modes of variability from both model and proxy-data perspectives. The main conclusion is a centennial pacing in climate (both on land and at sea, both in proxy data and models) that mostly relates to internal mechanisms. Hence, the results of the analyses deserve to be published.
Reply: We thank reviewer #2 for the assessment of our study.

While the study focuses on spectral analyses of time series, little information is provided about the time series from proxy-data and some clarification is needed:

- The authors referred to the Holocene climate database "Temperature 12K" published by Kaufman et al. (2020) after quality control. One of the criteria retained to build the 12K data set is a temporal resolution better than one data point for 400 years. Clearly, such a criterium is too much inclusive for the analysis of the multidecadal variability. Hence, some information about the temporal resolution of proxy-based time series is indispensable to demonstrate the adequation of the sub-data set analysed.
Reply: We only selected records with a temporal resolution of at least 50 years for our study (see line 77) and ended up using 120 records with an average resolution of 28 years (line 78). As we do not include the (multi-)decadal frequency range in our analysis (which could indeed be compromised by resolution issues), we have avoided any problems resulting from too coarse resolution. The overall criteria for the temp12k (and Sundqvist et al, 2014) are too weak for our analysis, which is why we added the 50-year resolution criteria and only used records from the databases that could live up to that. Otherwise, the analysis would have included 1000+ records instead of 120.

- Could time series with lower resolution explain the spread of oscillation periods illustrated Fig. 7 ?
Reply: Good question, it could have been the case. Looking at fig. 3 d), ice cores for example seem to be less right skewed than some of the others. However, looking at the scatterplot with resolution vs period, they do not seem to be any correlation.



- Many proxy time series are characterized by a change of frequency during the late Holocene (e.g., Allan et al. Paleoceanography and paleoclimatology, 2018; Bae et al. Scientic Report, 2022). Could this explain, at least in part, the spread of oscillation periods ?
Reply: Good question. The investigation of changes in frequency would require doing the analysis for different sub-periods. This is beyond the scope of this study, where we attempted to show the overall spectrum (or dominant ones) in proxy data and model data. The same is true regarding the question of state-dependency whether a specific variability is bound to a specific climate state (or forcing). In the revised manuscript we have added the suggested references and mentioned that it would be useful to use wavelet analysis to examine the varied spectrum and to understand how the different forcing affect the spectrum in future work.

- The use of a low-pass filter of 20 years for the proxy data has to be better explained : to which series does it apply ? ice cores and speleothems ? probably not lake and marine sediment cores.
Reply: The pre-processing for the proxies is in two steps: step 1, interpolate the records to 10-year resolution. Step 2, filter with the 20-year butterworth filter. This is the case for all the proxy records, regardless of their type, location etc. Same procedure for all the data. Effectively that means that some are upsampled and some are downsampled by the interpolation. The filter mainly affects the records with an original resolution of more than 20 years. The records that original have a resolution of less than 20 years, still won't have fluctuations faster than the original resolution after the interpolation. We have detailed the procedure in section 2.3 line 126-133 in the revised manuscript.

The paper makes clear the point of centennial to multi-centennial variability of mean annual temperature in both model and proxy data. This is by itself very interesting, but the mean annual temperature is not necessarily the best climate parameter for high latitudes, where seasonal gradients of temperature are large. So, the use of mean annual temperature may result in sensitivity loss for high latitude regions. There is also a possible difference between terrestrial and marine areas. This is shown in figure 2 for proxy records but is not documented for model data. Distinguishing the respective modes of variability on land and at sea would be useful.
Reply: This effect is indirectly studied in climate model simulations as part of differentiating between latitude bands for both hemispheres that strongly differ regarding the amount of land mass contained. Because low-frequency variability generally originates from the ocean (unless it would be a direct forcing by low-frequent solar forcing), the variability observed in the ocean can be expected to be dominant in GMT and southern latitudes. However, we do not observe any "sensitivity loss" for northern latitude bands that contain more land mass and there is no "sensitivity loss" at all for high latitude regions. As mentioned in the study, high northern latitudes appear to be even more sensitive (e.g., line 333) most likely due to strong positive feedback linked to interactions with the cryosphere. A more detailed study of mechanisms is required here which is beyond the scope of our study.

A lack of latitudinal dependency in the spectral distribution is mentioned for both model and proxy data. The distribution of proxy data points is concentrated in the Northern Hemisphere, with low numbers of records in the Southern Hemisphere, which is a limitation. In the Northern Hemisphere, the number of proxy data points is sufficient for making the assessment about latitudinal distribution. However, at high latitude the summer temperature is probably more sensitive than the mean annual temperature. It would have been interesting to verify the seasonal weight on the mode of oscillation.
Reply: According to many lines of proxy evidence, as well as instrumental data, the annual mean (and winter-half year) temperature variability at the high latitudes is larger than the summer variations. From the evidence we are aware of it is not necessarily correct that the summer season is more sensitive. However, we fully agree with the reviewer that investigations using seasonal weight on the mode of oscillations would be of interest but will feel we have to leave this to future follow-up studies.

The transient models provide interesting results with different absolute values (ranging 11 to 15 degree C) and somewhat different trends. It would be interesting to make a comparison with reanalysis data of the last century to see which model is closer to data and which are warm-biased or cold-biased.
Reply: Such an analysis and additional questions linked to the performance of individual models are beyond the scope of this study. We are surprised to see the large spread in global mean temperature simulated by the different models. We agree, that this would be quite interesting to understand if the absolute value of global mean temperature in different models would affect their low-frequency variability. Model bias could be one aspect among many others like model resolution or different complexity etc. A potential dependency on the model bias would also imply a potential climate state-dependency of low-frequent variability. All these important questions go far beyond the scope of this study, though.

Should we add some comments to raise the question of model differences in the revision?

- *It is already described a little in section 4.2, but if you want to add more, feel free to do so. I'm also okay with the way it is briefly mentioned in the manuscript now.*

In brief, the paper is very interesting and also challenging because the comparison of the modes of variability on a global scale (model) with local-regional heterogenous proxy records (different resolutions, different proxies) is a difficult task. From this viewpoint, the paper is a very good contribution although some clarifications, notably about the proxy data set, are needed.