Dr Charles JR Williams, BA DPhil FRGS
Research Fellow
1.2n, School of Geographical Sciences
University Road, Bristol, BS8 1SS

Thursday 28 May 2020

To whom it may concern,

Re: Response letter to Reviewer 1, after submission of manuscript "*The UK contribution to CMIP6/PMIP4: mid-Holocene and Last Interglacial experiments with HadGEM3, and comparison to the pre-industrial era and proxy data*" by CJR Williams *et al*. to Climate of the Past (PMIP4 Special Edition).

Thank you very much for informing me that the discussion period for the above manuscript is now over. Thank you also for the opportunity to resubmit a revised manuscript, according to the reviewer's comments, and for the extension granted to me.

I extend my sincere appreciation to the Reviewer 1 for their thorough examination of my manuscript, and their detailed and highly constructive comments. I propose to address all of their concerns, both minor and major, so please see attached for a revised manuscript, still with the Track Changes included, to show my proposed modifications. I also attach a tidy version of this manuscript. Please note that the line numbers shown here relate to the Track Changes version.

Here, I address the reviewer's suggestions, comment-by-comment. In the following, the reviewer's comments are italicised and in a smaller font, and my corresponding response follows in a standard font.

I very much hope that my responses will satisfy the reviewer and meet your expectations, and therefore request you to consider our revised manuscript for publication in your Journal.

Yours faithfully,

Dr Charles JR Williams, and co-authors

REVIEWER 1

MAIN CONCERNS:

As mentioned above, it appears to me that the manuscript presents various data sets and experiments that in its current form lack relevance or seem somewhat out of place. This makes that the manuscript lacks a clear common thread, making the manuscript difficult to read. Clarify the relevance of all the presented results and consistently mention all of them in the abstract, introduction, results and conclusion sections. Clarify why you present: a comparison to previous versions of the model; a comparison to CMIP/PMIP results; a comparison with proxy-data. A change to the overall structure of the manuscript as detailed in the following could also improve the flow of the manuscript.

The manuscript has now been restructured. In short, the section on the spin-up results has been reduced and moved into the Methodology section (with the figures included in the Supplementary Material), and the Results section has been restructured to focus primarily on two measures: firstly a simulation comparison to assess changes relative to the same model's PI simulation, and secondly a model-data comparison in which the most recent version of the simulations are compared to both previous generations of the same model and newly-available proxy data. Importantly, all versions of the same model are now compared against all available proxy data. The section focusing on the Saharan greening has been reduced and is now discussed only briefly at the end, with the figure being moved into the Supplementary Material. Following the reviewer's comment above, this new structure is consistently mentioned in the abstract, introduction, results and conclusion sections.

The analysis of temperature on the one hand and precipitation on the other hand are very different. Temperatures are looked at globally, while the analysis of precipitation is solely on Africa. Clarify to the reader why this choice was made. The structure of the manuscript would potentially be improved if the analysis is firstly on global features (mostly temperature, but perhaps also precipitation considering the newly developed Last Interglacial precipitation reconstruction), and secondly zooms in on African precipitation changes.

This structure has now been modified, with an additional figure showing global precipitation changes (Figure 3) as well as appropriate discussion (lines 521-534) so that, when discussing the most recent simulations, both temperature and precipitation are considered at the global scale, before zooming into Africa as an example of monsoon changes.

An extensive description of the spin-up results is given in the manuscript. While in general I appreciate it that such potentially important modelling details are given, the relevance to the rest of the manuscript is not clear to me. Wouldn't it be sufficient to provide the numbers of the spin-up results in a table and refer to that table in the method section? Is a whole results section for the spin-up results needed, given also that they are not referred to anymore in the remainder of the analysis?

This section has now been significantly reduced, and moved to a subsection of the Methodology section (lines 394-429). The results are summarised in Table 2 and referred to here, and the figures (showing examples of timeseries for atmospheric and oceanic equilibrium) have been moved into the Supplementary Material (SM2 and SM4).

The model-data comparison is limited to temperature proxies while recently a new compilation of precipitation reconstructions for the Last Interglacial has become available (Scussolini et al., 2019) which should be used here as well.

This compilation has now been included in the manuscript, with a new figure (Figure 10) as well as new text (lines 769-790), which we feel further adds to the novelty of the manuscript.

I find the whole analysis of African precipitation changes in the different simulations incomplete and too simplistic, leading to figures and results that are misleading.
To improve this situation I suggest the following:
-> Remove the ocean grid cells from the domain over which the analysis is performed. Presenting zonal-mean figures is not appropriate if the results are clearly not zonally homogeneous as is the case around the equator in this analysis.
-> Why is the focus on JJA precipitation over Africa? The results in figure 12 are again annual mean. Be consistent and clarify your reasoning.
-> The authors seem to use the words 'monsoon' and 'ITCZ' interchangeably. While indeed they cannot easily be separated based solely on the analysis of precipitation, they are driven by fundamentally different processes and an attempt should be made to separate the two. An interesting read on this topic is by Nicholson (2009).
-> The results for African precipitation are rather different for the different periods that are considered and for the various versions of the climate model. Provide some analysis as to what drives these differences.

All of the above points have now been addressed:
-> The zonal mean figure has now had all ocean grid cells removed

-> An explanation is given for why JJA precipitation is considered over Africa (i.e. because it is the primary wet season) versus annual precipitation in figure 12 (which has now been moved to the Supplementary Material) i.e. to be consistent with the proxy data-based threshold (lines 546-551) and (lines 926-927)
-> The confusion between the words "monsoon" and "ITCZ" has been clarified, by sticking to terms such as "rainfall maxima" or "rainbelt" (lines 555-565)
-> The reason for the difference in African precipitation between the different periods has been made more clear (lines 488-603), and a possible reason for one of the differences between versions of the model has been added (lines 817-820).

In many figures no measure of the robustness of the results is given. Provide measures to determine if your results are significant or if we are looking at internal variability of the climate system. A possibility would be to use long-term variability within the different simulations (PI, 6k, 127k) to deduce whether the depicted anomalies are outside of the range of this variability.

This has now been addressed, with the global temperature and precipitation figures of the most recent simulations (Figure 2 and Figure 3) now showing the results of a Students t-test (at the 99% confidence level). The same measure of statistical significance was not included when comparing the most recent simulations with previous versions of the same model and proxy data because stippling would make the proxy data locations harder to visualise.

MINOR COMMENTS:

Line 48: are surface (0 meter) temperatures discussed in the manuscript? Or do the authors mean SSTs here?

This line has now been addressed, to differentiate between 1.5m air temperature and SST, both of which are used in different places (lines 49-69)

Line 82: these periods are not 'warm' everywhere and all the time. Please clarify.

This has now been clarified, emphasising that they are two interglacial simulations representing the most recent warm periods (particularly in the Northern Hemisphere) (lines 106)

Line 83: 129-116 is not the period that is discussed in this manuscript. Why not simply 127ka?

This has now been clarified (lines 107)

Lines 108-112: Clarify that this warming is mainly located at high latitudes.

This has now been corrected (lines 169)

Line 126: convection in the atmosphere or ocean or both?

This has now been clarified, to confirm that it should be atmospheric convection (lines 190)

Line 127: are the ocean and sea ice models completely new or have parts been updated?

This has now been corrected, to confirm that these models are not completely new, but rather updated (lines 191)

Line 140-141: This division in two subsections (3.1 and 3.2) suggest to me that the two topics are of similar importance while in reality this is certainly not the case, with the results on the spin-up phase being only a small side topic. Consider changing this structure to better represent the importance of the different topics.

This is now been corrected as per the comment above, with significantly less emphasis being given to the spin up section (which is now in the Methodology section and summarised in Table 2, when the figures in the Supplementary Material (see above comment) (lines 394-429)

Line 141: For me the term productions runs is a little strange, perhaps it is CMIP kind of language, but in the context of a manuscript is doesn't mean much to me.

This has been removed at this line, and has been clarified when used elsewhere by a new Terminology subsection (section 2.1.1) (lines 216-226)

Line 214: Isn't precipitation impacted by ENSO?

This reference to ENSO has been removed

Lines 248-252: I don't think such details (number of output variables) are relevant for a manuscript.

These have been removed

Line 273: what is your definition of 'summer'?

This has been clarified, to confirm 'summer' here = July September (lines 362)

Line 276: You constructed this 127 ka time-slice of the Hoffman et al data? Do you provide this data for future work?

Yes, that is correct. A sentence has been added into the Data availability section, detailing access (lines 1014-1016)

Line 306: Is a trend of 0.16 degrees per century small? Sounds significant to me. Please clarify.

The reviewer is correct, so the ambiguous word "small" has been removed

Line 335: "the current two warm climate", what does that mean or refer to?

This term has now been clarified in the new Terminology subsection (section 2.1.1), see above comment

Line 335: Which newly-available proxy data are you referring to? Did you gather new data? Or do you mean the 127 ka time slice based on the Hoffman et al data?

This has now been clarified (lines 477)

Line 342: HadGEM3 warm climate simulations?

This has now been clarified (lines 488)

Line 364: 30 degrees east doesn't sound like west African to me. Please clarify why this domain was chosen, also in light of my main concern on this topic.

The domain used to calculate this zonal mean has now been changed, to become 20°W-15°E, with appropriate clarification text (lines 573)

Line 369: The wind patterns to me show an increase at nearly all latitudes, is that typical for an ITCZ shift?

Reference to the ITCZ has now been removed, instead referring to the regions of rainfall maxima, and an appropriate sentence has been added to clarify that the enhanced wind patterns do indeed occur at all latitudes, but especially over regional rainfall maxima (lines 558)

Lines 374-377: Do we see the same kind of pattern to the south of the equator, so the South African region?

This sentence has now been removed, so is no longer problematic

Line 407: Do proxies suggest a global annual mean warming during the MH?

This has now been addressed, to clarify that we don't see a global annual mean warming from proxies, but rather do see warming in many locations (lines 628)

Line 439: "within the average uncertainty range'? Please clarify this statement.

As part of other changes, this sentence has now been removed, as it was ambiguous

Line 466: The model is seasonally dependent? What does that mean? Do you mean the comparison of models and data?

This sentence has now been clarified, to confirm that it should read "the accuracy of the model is seasonally dependent" (lines 799)

Lines 488-492: Why would you compare your results to results from previous model version to see if you get sufficient precipitation over the Sahara to promote vegetation growth?

This entire paragraph has now been removed as part of other changes, so is no longer problematic

Lines 515-529: These kind of detailed (small) differences make me wonder whether we are really discussing forced differences or if we are discussing internal variability of the system. Please show statistics to argue either way.

This entire paragraph has now been removed as part of other changes, so is no longer problematic. However, as detailed above, significance testing has now been carried out on the most recent warm climate simulations

Lines 546-553: When you are talking about an 'improvement' this suggest that we know what 'good' means. What kind of data or proxies do you use to determine 'good' and what is the uncertainty of these estimates?

This entire paragraph has now been removed as part of other changes, so is no longer problematic. However, as detailed above, all model versions have now be compared against all available proxy data, allowing a quantitative and qualitative determination of 'good'

Line 555: Aren't the paragraphs before already discussing "rainfall across the Sahara"?

Yes, they are. This is therefore now been changed to a more appropriate title (lines 909)

Lines 570-583: What is the relationship between vegetation in the Sahara (the topic of this paragraph) and the state of the equatorial Atlantic ('drying'?)? Please clarify.

As part of other changes, this paragraph has been rewritten and shortened (and indeed the figure has been moved into the Supplementary Material), and therefore this sentence no longer exists

Lines 570-583: Not only is a vegetation model missing to directly determine whether or not vegetation would grow with the simulated amount of precipitation, but also all vegetation related feedbacks on the climate are missing. Discuss the possible impact of these missing feedbacks on you results.

As part of other changes, this paragraph has been rewritten and shortened. However, an additional sentence has been added at the end of this paragraph, briefly discussing the current lack of vegetation-related feedbacks (lines 933-935)

Lines 616-619: meltwater does not only yield a warming, it usually results in a spatially varying pattern with regions of warming and regions of cooling. Please clarify.

This has been clarified, to reflect the accurate comments of the reviewer (lines 986-987)

Lines 619-621: Is the length of the spin-up really a potentially important caveat? Do you have evidence to support this?

As part of other changes, this sentence has now been removed

Line 628: Only MH or both MH and 127ka?

This has been clarified (lines 997)

Table 2: If some values are for the full ocean depth and others for the top 1054 meter, can we still compare them? Isn't it comparing apples and oranges?

This has been corrected, such that all the values are for the full ocean depth.

Table 4: I appreciate the attempt to provide a lot of information, but I find this table very confusing. Perhaps it can be split or rearranged?

This has now been corrected, with the table being split into Table 4a and b

Figure 1: Have calendar effects been taken into account when making this figure? Please apply corrections, following for instance the methodology outlined by Bartlein et al. (2019).

Calendar adjustments, both this figure and all subsequent figures involving monthly or seasonal data, have now been applied, following the methodology of Pollard & Reusch (2002) and Marzocchi et al. (2015). This is briefly discussed in the introduction (lines 118-121), with examples of the data on the modern calendar (for comparative purposes) included in the Supplementary Material (SM1).

Figure 2: There seems to be a gap between the control data and the start of the 127k simulation, is this a real data gap or an error in the figure?

This has now been moved to the Supplementary Material (SM2). However, to answer the reviewer: yes, there is a purposeful gap between the end of the control data and the start of the LIG simulation, because a number of model crashes caused the first ~50 years of the spin-up to be unstable giving highly varied global mean temperatures. This is briefly noted in the figure caption (SM2)

Figure 2: Are the temperatures in the left-hand figure surface or 1.5 meter temperatures?

This has now been moved to the Supplementary Material (SM2). However, to answer the reviewer: these are 1.5m air temperatures, and this has now been clarified in the figure

Figure 2: Consider: 'b) TOA radiation balance'

This has now been moved to the Supplementary Material (SM2). However, to answer the reviewer: yes, this has now been corrected

Figure 2: This figure gives a good idea of the amount of internal variability in the system, which seems considerable in both the MH and lig127k simulations. Use this information to define which of your results are robust with respect to this internal variability. Is it true that variability is larger in the 'warm climates' than it is in de control?

We now use a Student's test (at the 99% level) as a matter of significant or robustness, which accounts for the interannual variability. Yes, it is true that variability is larger in the warm climate simulations than the PI, and this has been briefly noted in the text (lines 402)

Figure 3: For the control simulation the full depth is used instead of the top 1054 meters according to the main text, please clarify.

This has now been corrected (see comment above)

Figure 4: These figures show some well-know climate change features, including polar amplification. The mechanisms of such spatial temperature anomaly patterns are not discussed. Provide a discussion or refer to previous work on the topic.

This now been addressed, with a short discussion on one of the mechanisms of polar amplification, namely sea-ice interactions, has now been added (lines 515-519), along with an accompanying figure in the Supplementary Material (SM6)

Figure 6: Rainfall anomalies on y-axis?

As part of other changes, this figure has now been removed. However, in subsequent zonal mean figures, the y-axis label has been changed from simply "Rainfall" to "Rainfall anomalies" (relative to the PI), which is what we understand the reviewer to mean here

Figure 6: x-axis values are not easy to read in this format.

As part of the changes, this figure has been removed

Figure 6: consider showing absolute precipitation values because I think those give a much better idea of the width of the wet and dry regions as discussed in the main text.

As part of the changes, this figure has been removed. However, when zonal mean rainfall is shown, both anomalies and absolute values are now shown (Figure 6)

Figure 6: Can't this figure be combined with figure 9?

This has now been done

Figure 8: Remove the ice core data points if the corresponding modeled surface temperature anomalies are not shown.

This has now been done

Figure 9: What does this figure add that is not already depicted in figures 10 and 11? Can't it be removed?

It was decided to remove existing figures 10 and 11 instead, because the same information is shown at the global scale in the new Figure 8 and Figure 10

Figure 12: why are the grey dashes that show required rainfall for grassland growth only start from 16 degrees north?

This has now been moved to the Supplementary Material (SM7), and it has been modified so that the model latitudes begin approximately where the rainfall threshold needed for grassland (the grey dashes) begins. To directly answer the reviewer, the grey dashes in this figure were taken directly from Figure 3a in Joussaume *et al*. (1999), which only has data beginning at 15.5°N. This reference has been added to the figure

Figure 12: Rainfall anomalies on y-axis?

This has now been corrected (see above comment)

Figure 12: Why are anomalies shown? Doesn't the threshold to support grassland depend on the absolute amount of precipitation?

Anomalies, rather than absolute values, are shown because this is following the methodology of Joussaume *et al*. (1999), who also considered annual mean anomalies in their study. Likewise, the region of averaging is larger here than Figure 6 (up to 30°E, as opposed to 15°E), again to be consistent with Joussaume *et al*. (1999). This has been made more clear in the text (lines 926)

TECHNICAL COMMENTS:

Line 41: are similar

This has been corrected

Line 41: period, but

This has been corrected

Line 53: generations of the same

As part of other changes, this sentence has now been removed

Line 121: therefore in the

This has been corrected

Line 146: consider removing "indeed"

This has been corrected

Line 149: double space before "Full"?

This has been corrected

Line 186: tuning of

This has been corrected

Line 201: including a reduction of the temperature bias in many regions

This has been corrected

Line 221: remove comma after 'design'

This has been corrected

Line 244: Too many brackets

This has been corrected

Line 272: annual mean surface

This has been corrected

Line 298: radiation balance?

As part of other changes, this sentence has now been removed

Line 384: in the core monsoon region?

As part of other changes, this has now been removed

Line 395: 'recent', what do you mean?

This has now been corrected

Line 399: what kind of uncertainty in simulated anomalies are you referring to, please clarify.

This has now been corrected

Line 400: remove 'often'

This has now been corrected

Line 437: small number of reconstructions?

As part of other changes, this has now been removed

Line 449: remove double comma

As part of other changes, this has now been removed

Line 499: refer to figure 10?

As part of other changes, this has now been removed

Line 505: smaller northward displacement?

As part of other changes, this has now been removed

Line 590: 'auspices', not sure if that is the right word for it.

This has now been corrected

Line 590: replace comma by a dot.

This has now been corrected

Line 591: remove 'time'?

This has now been corrected

Line 592: were assessed?

This has now been corrected

Line 603: 'time', are we talking seasons or different geological intervals?

This has now been corrected

Line 626: 'necessity' is perhaps a bit too strong in this context.

This has now been corrected

Line 1007: better not to use the & symbol.

This has now been corrected

Line 1008: for each

As part of other changes, this has now been removed

Line 1018: in this caption and some others the words 'simulated gridded anomalies' are used. This sounds a little double to me since nearly all climate models work on spatial grids so the output is per definition also gridded.

This has now been corrected

Lines 1024-1027: Is there no overlap between these two data sets? No single core was used in both of them?

As discussed in the Methodology section (section 2.3), the two datasets use different reference chronologies and methodologies to infer temporal surface temperature changes. Whilst they may use the same core, the methodologies are very different, and therefore they should not be combined

Line 1033: erroneous bracket?

As part of other changes, this has now been removed