

We thank the two reviewers for their very helpful comments on the original submission, which have greatly improved the paper. Here we provide a point-by-point response to those comments. The original comment is in 8-point font and our response is in **10-point bold font**. The line numbers refer to the 'track-changes' version of the manuscript.

### **Reviewer 1:**

Paleogeography - I am generally happy with both the field and your explanation of it. However, I note that there appears both an isolated very high peak as well as a 'lake' at the join between Greenland and (what I think is) Scandavia. This feels unrealistic to me. Can you either smooth them out if that's the case, or explicitly discuss the geological evidence for them?

**This very high peak is actually in the Straume et al reconstruction, but we agree with the reviewer that this looks “odd”, being so isolated and significantly higher than other topography globally, and may cause model instabilities, so we have smoothed it. The lake is actually an inlet in the original reconstruction, but it is too narrow to be resolved at the 0.25 degree resolution of the DeepMIP file, so we filled it in. Both these changes are summarised in the text [Lines 219-220].**

Figure 4. Can you please harmonise the shape and color between the panels? It is the ice cover that switches with savannah that is most confusing.

**It is not possible to completely harmonise the colour scales between the BIOME4 and LPJ-GUESS vegetation types because they are fundamentally different classifications. However, we have made them more consistent by changing the colour of ice and savanna, and have modified the forest greens [Figure 4].**

ESGF. On L386 you state that participants should upload their simulation to the DeepMIP archive. Is there a reason you do not also allow the possibility of uploading data created using CMIP7 models onto the ESGF? [Now is the time to establish the relevant controlled vocabulary, and you are within the CMIP7 special issue]. This would have the advantage of combining the Eocene with the other PMIP experiments.

**We have applied for the *deepmip-p2-stand-5xCO* simulation to be a formal “Tier 2” CMIP7 experiment, named “*eocene-5x*”. We have encouraged groups to submit this simulation to the Earth System Grid Federation (ESGF). [Lines 470-471]**

Experiment name: Are you sure that you have selected the most helpful experiment names?

- The precedent from the wider PMIP and CMIP efforts does not include phase numbering. Bear in mind that all phase 2 data is stored within a specific directory (L391).

**As our boundary conditions are considerably different from Phase 1, we do think it is important in this case to make that distinction in the “local” experiment name. However our CMIP7 name is simply *eocene-5x* [Table 1].**

- No period is included in the name. This is especially pertinent, as I understood the Miocene is also included in DeepMIP. Since you focus only on the EECO, why not choose either 'eocene' or 'EECO' instead of *deepmip*

**We have modified “*deepmip*” to “*deepmip-eocene*” in the experiment name, for consistency with Steinig et al. [Table 1].**

- I find the inclusion of 'stand' to be unhelpful. I can understand defining subsidiary experiments, with a 'sens' flag. Surely this implicitly assumes that other than the named feature being changed, everything else is as set as standard - inconsistent with the requirement of a 'stand' flag.

**We have removed “stand” and “sens” from all simulation names. [Table 1].**

- If you do choose to allow ESGF inclusion, your main experiment name will be the longest around (and will not make sense to people not involved in *deepmip*).

**See above - we have applied for the *deepmip-eocene-p2-5xCO* simulation to be a formal CMIP7 experiment, named “*eocene-5x*”. [Lines 470-471]. We also shortened our standard simulation names (Table 1).**

Please be aware that equilibrium-4xCO<sub>2</sub> could be exactly the simulation (but at a different point) as abrupt-4xCO<sub>2</sub>. How is CMIP7 DECK treating this possibility?

**We do not foresee this being a problem, as we will not register the equilibrium-4xCO<sub>2</sub> simulation on ESGF.**

Sect 3. You state this section includes 'plans for analysis', but it doesn't really. Either expand on this a little, or perhaps just removed them from the section title.

**Removed. [Line 472].**

Sect 2.3.4: Can you please shout out to the relevant sensitivity experiments listed in Table 1

**Clarified that this section describes additional sensitivity studies not listed in Table 1. [Line 382].**

L148: can you please move 'red' to before 'line'? This would fit the same structure as description of the other lines in the figure.  
**Done [Line 155].**

L154: can you clarify the word 'records'? I believe that you mean that two compilations ingest multiple of the same individual readings. But it could mean that the error bars on the two timeseries overlap.

**Clarified. [Line 161].**

e.g. L178, L255: there are a couple of instances of the wrong \cite command, leaving the bracket in the after, rather than before the authors.

**Done.**

L239: Consider removing `field of`

**Re-worded [Line 250].**

L241: I appreciate you stating that Herold et al provide river routing. Can you comment of this fields relevance, given the different topography?

**Done [Line 255-256].**

L263: add CO2 after 6xPI

**We have now clarified the nomenclature in an earlier Section [Line 125].**

L294: "in, that" -> "in and that"

**Reworded [Line 317].**

Sect 2.4: I am happy with the explanation of all the various methods. Can you please reiterate that whatever approach is selected should be documented in the simulation publication?

**Done [Line 395].**

L402: Can you please spell out that 'std' stands for 'standard deviation'

**Done [Line 454].**

L402 and L404: 'timeseries' seems to be doubly defined, and I don't understand the distinction.

**Not changed. One occurrence is in the directory and one in the filename.**

L407: You ask for temperatures at 3500m depth. Can you clarify if you just want the layer containing 3500m, or to vertically interpolating the full profile to get the value at 3500m for intercomparison?

**Clarified – the layer containing 3500 m [Line 460].**

## **Reviewer 2:**

L20: The Mallory reference is perhaps belittling the importance of learning about processes and mechanisms that led to things being the way they were/are even if they are not so directly relevant for future climate change. I suggest removing.

**We have added more detail here emphasise the important improved understanding that this 'curiosity-driven' research brings [Lines 20-21].**

L76-78: Do we have any idea why there are such differences in CO2 estimate for the best agreement with the data?

**Added that the reason for the discrepancy is yet to be explored [Line 82].**

L97-98: Can you perhaps infer that this may be the cause of some disagreement between models?

**Yes, done [Line 103].**

L98: Over what period were the simulations run but still did not reach equilibrium?

**Reworded – none of the simulations reached full equilibrium [Line 102].**

L63-92: I wonder if this section could be reorganised into consistencies between models and data, consistencies between models but disagreement with data, inconsistencies between models (some of which may agree with the data), and inconsistencies between models and data so that is clearer what the main certainties and uncertainties are?

**We have considered this, but prefer the current structure of going through the papers in an order related to the processes that they focussed on, as this is more intuitive, in our view.**

Figure 1: I presume the divergence in the colour scheme around the EOT is to highlight the shift from the greenhouse world to the icehouse world - might be worth adding those labels to the figure to explain why that was chosen as the divergence point. **We now note that the centre of the colour scale is broadly the value associated with the EOT [Figure 1 caption].**

L122: Include Eocene/EECO in the simulation naming to avoid confusion with the other DeepMIP simulations. I believe MioMIP is using Deepmip-Mio, perhaps here therefore they should be Deepmip-Eo or Deepmip-EECO?

**We have modified “deepmip” to “deepmip-eocene” in the experiment name [Table 1].**

L254-244: Ambiguous as to whether you are suggesting that this runoff routing field could be used for Phase 2 or not  
**Clarified [Line 255-256].**

L270: What criteria is used to determine if regions have "sufficient" data coverage? A comment on the vegetation data availability for this period would also be useful

**We have highlighted that this is a manual qualitative process, established within the PlioMIP framework, and documented in the Appendix (Figure A2c). We also commented on the variable geographical data coverage. [Lines 286-289; Lines 291-293].**

L330-331: Is the best-fit GCM CESM1.2 for Thomsson in BIOME4 the same as the best-fit GCM CESM1.2 for Brugger in LPJ-GUESS? If so, worth commenting on that.

**Done [Line 363].**

L336-338: Are the discrepancies between the BIOME4 version of the vegetation and the data because these are where the model has been corrected towards the data in the LPJ-GUESS version of the vegetation? i.e. because the BIOME4 version has not undergone the hybrid proxy reconstruction correction step? Or are these different regions?

**Clarified that this is not due to the ‘nudging’ process, and is a more widespread effect in BIOME4 that is not present in the underlying LPJ simulations [Lines 369-370].**

L362-363: You suggest that groups can use initial conditions from the Phase 1 model database, but in L98 you state that many models did not reach full equilibrium but the end of their simulations. Are there some models in the Phase 1 data that you would therefore recommend are NOT used for the initial conditions for Phase 2? Similarly, are there any in the database that suffered from crashing early or running away that should NOT be used? Or are these not in the database?

**We have added text recommending that groups take account of the TOA imbalance in the previous simulations, as given in the Supp info of Lunt et al (2021). Simulations that crashed are not in the database [Lines 400-404].**

Figure 4: It is very difficult to see the details of the legends for (a) and (b), and the distinction between some of the greens is so small it is nearly impossible to discern which is which on the map. It also appears that the same colour has been used for Ice in (a) as Savannah in (b) and (c). Can you try to use a more distinctive colour scheme?

Figure 5: Similar comments on the colour scheme to Figure 4.

**It is not possible to completely harmonise the colour scales between the BIOME4 and LPJ-GUESS vegetation types because they are fundamentally different classifications. However, we have made them more consistent by changing the colour of ice and savanna, and have modified the forest greens [Figure 4, Figure 6, Figure A2].**

L416: This section is very short and the title is misleading. Perhaps Section 2.5 might fit better here. Maybe Summary and Next Steps or similar might be a better descriptor?

**Section title renamed, see also comment from Reviewer 1 [Line 472].**

L424: Are there plans to version control the proxy database so it is clear which dataset is being used for model-data comparisons and to ensure consistency between the analyses by different groups?

**Yes, we will ensure that future versions of the database are version controlled.**

Figure A1. The colour scheme used may prove difficult for some colour blindness conditions.

**We have modified this colour scale [Figure A1].**

L1-2: No commas required

**Our feeling is that the sentence is clearer with commas.**

L2: Define when the early Eocene is

**Done [Line 2].**

L8: Word missing, a new vegetation... scheme/distribution/boundary condition

**Done [Line 9].**

L10-12: Pluralisation of concentrations seems a little odd in the context of the model simulations where preindustrial CO<sub>2</sub> is a constant and 4 x preindustrial CO<sub>2</sub> is therefore also a constant.

**We are now consistent throughout the paper with the use of “concentrations” (proxy**

**estimates which vary with time) versus “concentration” (model values, or preindustrial control value) [e.g. Lines 11-12].**

L38: Word missing, we/DeepMIP extend...

**We think this is fine as-is. It is DeepMIP-Eocene and DeepMIP-Miocene that are extending.**

L38: Comma after Quaternary not required

**As this is a long sentence we prefer the commas.**

L46: Word missing, also indicate

**Done [Line 48].**

L47: No comma required before 'and a vegetation consistent'

**As this is a long sentence we prefer the commas.**

L47: Duplication of wording, perhaps change 'a generally warmer and wetter climate than modern' to something like 'those warmer and wetter conditions'?

**We prefer to keep as-is because hydrological proxies indicate wetter high latitudes, and separately from that the vegetation indicates generally warmer and wetter conditions.**

L53: Is Herold the only reference for all these differences? Or do you mean to say Herold and references therein?

**Done [Line 56].**

L55: No comma required before 'and has been used'

**As this is a long sentence we prefer the commas.**

L61: No comma required before 'and have been discussed'

**Done [Line 65].**

L70: Pluralisation of concentrations again

**We are now consistent throughout the paper with the use of “concentrations” (proxy estimates which vary with time) versus “concentration” (model values, or preindustrial control value). [e.g. Line 73].**

L73: Unnecessary 'and' before 'drier conditions'

**Done [Line 76].**

L120: Maybe introduce PI=preindustrial when talking about CO2 concentrations earlier in the manuscript

**Done [Line 125].**

L157 & 161: Inconsistent wording, runaway versus 'runs away'. Maybe also explain what this term means.

**'Runs away' is a verb, whereas 'runaway' is a noun. We added an explanation [Line 165-166].**

L173: Typo, Appemdix

**Corrected [Line 181].**

L200: Wrong wording?, should be 'or between' rather than 'as well as between'

**Changed “as well as” to “or” [Line 209].**

L205-206: Incorrect wording, 'As the paleogeographic reconstructions used here all have modified' should be 'As all the paleogeographic reconstructions used here have modified'

**Done [Line 214].**

L222: Commas needed around the word 'particularly'

**Reworded [Line 232].**

L227: No comma needed after the word 'resolved'

**As this is a long sentence we prefer the commas.**

L286-289: The long sentence is a little confusing. Perhaps reword to something like 'The recommendations for all other boundary conditions remains unchanged from Phase 1 (refer to Lunt et al. 2017). Using the section numbering from Lunt et al. (2017), detailed recommendations are given for soils and lakes in Section 4.2.2,' etc

**We reworded this sentence [Line 308-309].**

**Other changes:**

In addition to the changes listed above, we also made the following additional changes:

- **Following offline comments, we have corrected the caption to Figure 1, and to the text that references it, including the correct citation, and the fact that this is a composite. We also added clumped isotope deep ocean temperatures [Figure 1, caption].**
- **We have updated the zenodo dois, and think that these are now correct.**
- **We have added some more quantitative guidance on model initialisation from deep ocean temperature proxies [Lines 405-416].**
- **We have added uncertainties in the Hönisch et al CO<sub>2</sub> curve in Figure 2 [Figure 2, caption].**
- **We have added a link to a github repository that allows people to carry out their own paleo rotations [Code Availability section].**
- **We have provided some additional files on zenodo for the paleogeography sensitivity studies, and shown these in a new Figure 5; namely a PETM geography at 56 Ma, and a Herold 55Ma paleogeography rotated to our new plate model and reference frame.**