

# Replies to Reviewer Comments

David Wade on behalf of the authors

We would like to extend our gratitude to Jim Kasting and the anonymous reviewers for the time and care they took in reviewing the paper. The comments will be dealt with in turn and changes to the updated manuscript will be described. In all cases, reviewer comments can be identified by the red text and the author reply in the black text. We've provided a tracked changes document for the editor where the new additions (bold black text here) are added in blue and text we have deleted is scored out in red for clarity.

---

## Anonymous Reviewer #2

Wade et al. quantify the climatic impact of changing atmospheric oxygen concentrations (pO<sub>2</sub>) using two ocean-atmosphere general circulation models in the Holocene, the Cretaceous and the Permian. They systematically conduct their simulations at 3 pO<sub>2</sub> levels (10, 21 and 35 ‰), which are shown to reasonably cover the pO<sub>2</sub> changes reported during the Phanerozoic. In their model, higher pO<sub>2</sub> values (and associated greater atmospheric mass) lead to two competing effects: an increase in Rayleigh scattering that induces an increase in albedo and surface cooling, and an increase in greenhouse effect that leads to surface warming. The authors first run the two models on the preindustrial Holocene configuration. Interestingly, the state-of-the-art IPCC-class model (HadGEM-AO) and the version of the model designed for deep-time studies (HadCM3-BL) provide climatic responses that agree at first-order, thus supporting the robustness of the subsequent deep-time HadCM3-BL integrations. In their Holocene simulations, the mean annual global climate response to an increase in pO<sub>2</sub> is a warming, with varying regional patterns. The warming is particularly strong in the northern high latitudes, especially during the cold month. A cooling is simulated at low latitudes, which is especially strong and extends to most continental areas during the warm month. Higher pO<sub>2</sub> values tend to flatten the equator-to-pole temperature gradient. They also lower the climate sensitivity to atmospheric carbon dioxide. Then the authors run the HadCM3-BL model in the Cretaceous and two Permian time slices. They show similar climatic behaviors and discuss two specific points related to each case study: the response of the terrestrial vegetation to changing O<sub>2</sub> levels during the Permian and the impact of changing O<sub>2</sub> levels on the capacity of their model to simulate the low latitudinal temperature gradients traditionally reconstructed for the Cretaceous based on proxy data. They notably show that changing oxygen concentration only slightly improves model-data agreement in the Maastrichtian.

Last but not least, they propose a quantification of the uncertainty in global temperature resulting from uncertainties in the pO<sub>2</sub> during the Phanerozoic. They show that the temperature bias associated with poorly constrained pO<sub>2</sub> levels is significantly lower than the uncertainty associated with the lack of constraints on the pCO<sub>2</sub>, with a notable contribution of pO<sub>2</sub> during the Permian though.

David Wade on behalf of the authors

It should be noted that Wade et al.'s implementation of O<sub>2</sub> forcing leads to results that agree at first order with most previous attempts, but differ in sign with the simulations of Poulsen et al. (2015; 10.1126/science.1260670). Analysis of the model runs led the authors to suggest that Poulsen et al.'s implementation may not be totally coherent.

We would like to thank the reviewer for their time in providing such an in depth review of our manuscript and for their comments and suggestions on how to improve it.

I think that Wade et al. provide a very interesting and innovative study that shades new light on the poorly explored question of the potential impact of changing pO<sub>2</sub> levels on deep-time climate. The results are based on numerous general circulation model simulations using two generations of climate models. The manuscript is relatively well organized (an exception if the methods section, see comments) and richly illustrated with high-quality figures and abundant information embedded in tables. The manuscript is lengthy (as testified by the length of my summary above). This is essentially due to the large amount of diagnostics provided by the authors but I also suggest below deleting the section of the manuscript relative to the impact of wind stress, which I think is not very useful and relatively badly integrated in the manuscript (see hereafter).

We thank the reviewer for this comment. This is also suggested by reviewer #2 and we agree that removing this from the paper will help make the overall messages clearer.

The discussion of the discrepancy with Poulsen et al.'s (2015) results is well conducted. Indeed, Wade et al. not only compared their results with the diagnostics provided by Poulsen et al. but also downloaded and analyzed the climatic simulations of the latter, by repeating key diagnostics that they previously provided for their own model runs. This effort deserves to be acknowledged. As a reviewer of this paper, I would be happy to have Poulsen et al.'s response, be it as another review or at least as a comment on the ClimPast Discussion forum. Therefore, I encourage the Editor to contact Poulsen et al. I also encourage the authors to make their implementation of O<sub>2</sub> forcing available online (as numerical – fortran – code or as equations) in order to allow other research groups to conduct similar experiments using other climate models, thus permitting to determine to what extent the discrepancy between Poulsen et al.'s results and theirs is model-dependent (and conversely, implementation-dependent; see major comment).

We thank the reviewer for the comments and suggestions for us to provide some numerical code or equations to help others. We did consider this at length but after consideration we feel that this would end up not being useful beyond users of the specific climate models we have used in this study. We will make clear in the manuscript that users of the versions of the climate models we have used are welcome to our code changes, upon request, but these changes are so specific to our models that they would not prove useful to other modelling groups. Indeed, implementing the changes in HadGEM3-AO was very different to implementing the changes in HadCM3-BL.

Added to the *Code and data availability* section:

“Readers who would like advice on how to implement alterations to pO<sub>2</sub> in their climate model are encouraged to contact the corresponding authors. UM users can obtain the code changes for these particular versions from the corresponding authors.”

Most of my comments are intended to help the authors sharpen and clarify their manuscript. The only (other) major (potentially critical) comment I have regards the robustness of the

analyzed climatic simulations. I suggest accepting this manuscript with moderate revisions, provided that Wade et al. can demonstrate that their climatic simulations are robust (i.e., sufficiently close to equilibrium).

Please note that the text refers in several places to supplementary figures. I did not find any SOM.

We thank reviewer #2 for pointing this out and this was an error at the submission stage we will rectify in the response.

#### A. Major comments:

- On the discrepancy with Poulsen et al.'s results. Since the current study casts doubts about the Poulsen et al. implementation of O<sub>2</sub> forcing, I suggest making the implementation of Wade et al. available online to allow other modelers to repeat such experiments using alternative climate models – using GENESIS in particular. I think that such common effort will

allow improving the implementation of oxygen forcing in a collaborative and efficient way.

As we have alluded to above we don't feel that this would be a practical action going forward. Instead, we would be very willing to help others with implementing the code changes required to repeat the calculations we have performed in their own models. Indeed, the formulation of equations used in different climate models may make it deeply difficult, if not nigh impossible, to perform these types of simulations. As we mentioned above, the code changes we made in the two different versions of the Hadley centre climate model were drastically different.

- On the robustness of the climatic simulations. I recently had the opportunity to attend a presentation by Dan Lunt showing that the climatic simulations published by Lunt et al. (doi:10.5194/cp-12-1181-2016), run for 1422 years, did not reach equilibrium. A longer duration in the order of 10 kyrs is necessary to reach deep-ocean equilibrium, with the global mean SST simulated at the end of the longer simulation significantly differing (several °C) from the SST simulated after 1422 years of model integration time. Therefore I logically wonder if the climatic results used in the present manuscript based on the 1422-year long model integrations (see page 6, line 15) can be trusted. To what extent are the model runs equilibrated? I encourage the authors to clarify this point. Otherwise, the subsequent publication of longer model runs may significantly question the robustness of this entire study.

We thank the reviewer for this comment, which touches on an important issue in climate modelling more widely. We reject the idea that the simulations we have shown are not sufficiently well spun up to allow us to make robust conclusions. The desirable length of a particular climate model integration depends intrinsically on the question being asked and on a number of factors including:

- Model components (e.g. inclusion of land ice would lead to a longer integration time due to the response time of that physical system)
- Magnitude of the imposed changes (with respect to the baseline state)
- Pragmatics (time and computational resources required to perform the integration)

The main question we ask here is how does the climate simulated by the models we've investigated, and by proxy the Earth, respond to changes in the amount of O<sub>2</sub> in the atmosphere. This is a form of a forcing-feedback study in which the forcing is being imposed

by changing  $pO_2$  in our climate models. Forcing-feedback studies are widely used for example to understand how climate change will evolve over the coming century. Ultimately we are trying to understand if the signal from the forcing is large or is within the internal noise present in the chaotic climate system. We believe we have robustly shown the amount of climate change from changing the amount of  $pO_2$  in the atmosphere is second order compared to changes in  $pCO_2$  but is not insignificant and we believe we have identified the major mechanisms behind the changes in climate. However, we wish to further elaborate on why we feel these simulations are robust. Hereafter we will refer to the doi:10.5194/cp-12-1181-2016 study as Lunt et al. 2016.

### *Model Components*

While the deep ocean does take a considerable amount of time to spin-up, the interest of our study is much more in the shallow ocean response to the imposed changes. Indeed, one simulation passing some ocean bifurcation and entering into a new pattern of circulation after several thousand model years would be an intriguing result but would not be insightful for our understanding of the impact of atmospheric  $pO_2$  changes. It is worth noting that on long enough time periods, orbital changes would need to be accounted for as these are not fixed on timescales of 10 kyr so would ideally need to be integrated over a number of Milankovitch cycles in order to be properly spun up. This would be too computationally expensive for all but Earth system models of intermediate complexity, which would not be suitable for this study due to their simplistic treatment of atmospheric radiation.

### *Magnitude*

The smaller the imposed change, the less time is required for integration. It is worth noting that the Lunt et al. (2016) experiment is initialised with an ocean at rest and an idealised temperature profile while the imposed atmospheric change is a quadrupling of  $CO_2$  from preindustrial conditions. With no initial ocean circulation, it will take a significant amount of time for the deep ocean to respond. This kind of spin-up process is required in the context of that study, where the new continental configuration has been imposed. In our proposed study we use existing, well spun-up simulations and perform perturbations to the  $pO_2$  content off these which, as shown by the model results, are substantially smaller than the impact of quadrupling the atmospheric  $pCO_2$  as performed in Lunt et al. 2016.

### *Pragmatics*

Given the range of time periods over which  $pO_2$  has varied, it is desirable to attempt to simulate a few different time periods. In addition, the desire to perform the idealised Holocene run with a more “state-of-the-art” model means that each simulation with HadGEM3-AO took at least 120 days of continuous simulation, compared to each HadCM3-BL which run around a month. Significantly longer simulations would not be possible given the available computational resource.

We therefore believe that the simulations presented are sound, given their application to understanding the impacts of  $pO_2$  variability.

ALSO: Page 8, line 6. “iterated for 100, 1000 and 100 years”. What’s the justification for the 100-year integration time used for two of the 3 experiments? I doubt that such duration is sufficient to reach equilibrium under a doubled  $CO_2$  level.

The 100 year simulations were designed explicitly for performing Gregory analysis to understand the forcing-feedback relationships and this requires fairly short runs. The 1000 year simulations for Ma-CM were run to enable a model-data comparison (section 3.5) which, based on the reviewers comments, we propose to remove. On this basis we will modify the text in the manuscript to read:

“The PI2x-CM\*, Ma2x-CM\* and As2x-CM\* experiments were spun off from the end of the PI-CM, Ma-CM and As-CM experiments and iterated for **100 years in order to perform a Gregory (2004) analysis.**”

#### B. Other comments:

- Title: I would suggest revising the title to clearly indicate that several case studies are considered – maybe something like: “Simulating the climate response to atmospheric oxygen

variability in the Phanerozoic – Holocene, Carboniferous and Permian case studies”. In my opinion, such title would be more instructive, notably permitting readers interested in these 3 key time slices to more easily find this paper.

We thank the reviewer for this useful comment to help improve the visibility of the paper and suggest a subtle revision of the title (in bold) to:

“Simulating the climate response to atmospheric oxygen variability in the Phanerozoic: **A focus on the Holocene, Cretaceous and Permian.**”

- Page 1, Line 1. “10 %”: Fig. 1 suggests that it could have reached lower values.

Noted and changed in the abstract to reflect this.

- Page 1, line 5. “during different climate states” > “under different...”.

Noted and changed in the abstract to reflect this.

- Page 1, line 15. “increasing oxygen content leads to a slightly better agreement”.

As we are removing section 3.5 we have changed the abstract to read:

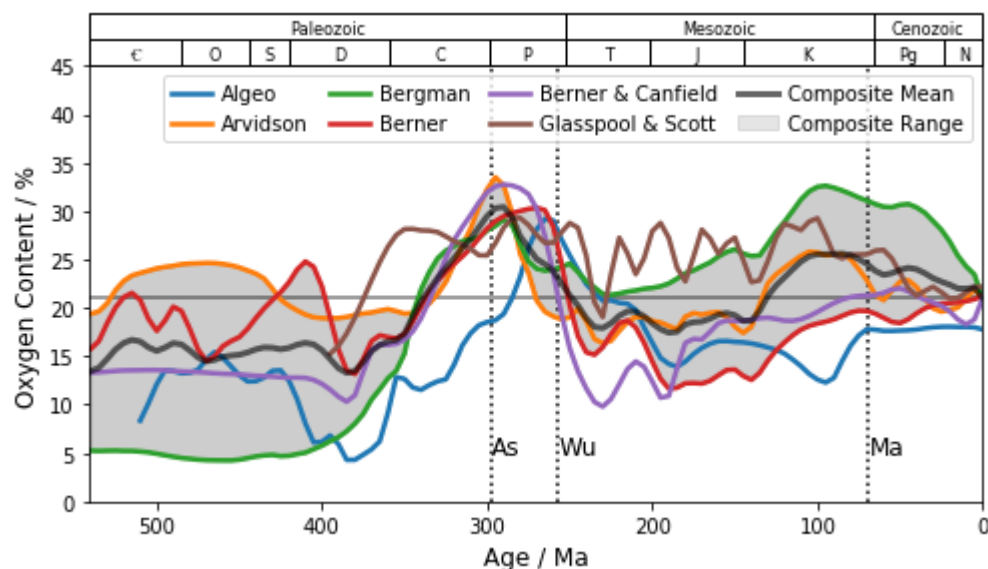
“Case studies from past climates are investigated using HadCM3-BL which show that in the warmest climate states **in the Maastrichtian (72.1-66.0 Ma)**, increasing oxygen may lead to a temperature decrease, as the equilibrium climate sensitivity is lower.”

- Fig. 1. Please show the different time slices used in each case study using for instance vertical lines.

This has been added to the adjusted figure (repeated below the next comment). New text has been added to the caption: “Timings of the palaeo case studies explored in this study are indicated by the vertical dotted lines (As: Asselian, Wu: Wuchiapingian, Ma: Maastrichtian).”

- Fig. 1 caption. “High and low limits on atmospheric oxygen are indicated by horizontal grey dashed lines”. What does that mean? Please clarify. I guess those horizontal lines indicate the 2 end-member O<sub>2</sub> levels considered in the deep-time case studies. In this case, the lower line is wrongly placed in the figure (this is not 10 %).

To reduce ambiguity we have removed the grey dashed lines from the revised figure which is provided below:



- Page 3, line 1. “to 20–35 % in the Permian and subsequently stabilized at levels around 15–30 % from the Mid Triassic onward” or similar.

Thanks for the suggestion which we have adopted directly in the text.

- Page 3, line 6. See studies by Dahl et al. (doi:10.1073/pnas.1011287107) and Lu et al. (doi:10.1126/science.aar5372) though, which provide very interesting insights into the evolution of pO<sub>2</sub> during the Phanerozoic. The authors may want to refer to these studies.

We’d like to thank the author for bringing these to our attention. We have adapted the sentence in question to “At the time of writing, there are no direct geochemical proxies for **atmospheric** pO<sub>2</sub> on the Phanerozoic timescale. However, there is isotopic evidence of oceanic oxygenation in steps at approximately 560 (Dahl et al 2010), 400 (Dahl et al 2010, Lu et al 2018) and 200 Mya (Lu et al 2018).”

- Page 3, line 12. “visible life”. OK, but I’m pretty sure it refers to the ocean realm, not to terrestrial life. Similarly, I think that the most prominent change between the Precambrian and the Phanerozoic is the advent of complex forms of life in the ocean during the Cambrian Explosion and subsequent Ordovician radiation.

We don’t mean to diminish the role of the ocean and so suggest modifying the manuscript to read:

“visible life and **one of** the marked.. ”.

- Page 3, line 15. “possibly led to the Ordovician glaciation” (there are a lot of alternative hypotheses and the spatial cover and thus climate impact of the primitive Ordovician vegetation remain poorly constrained).

We have added “**possibly**” as the reviewer suggests.

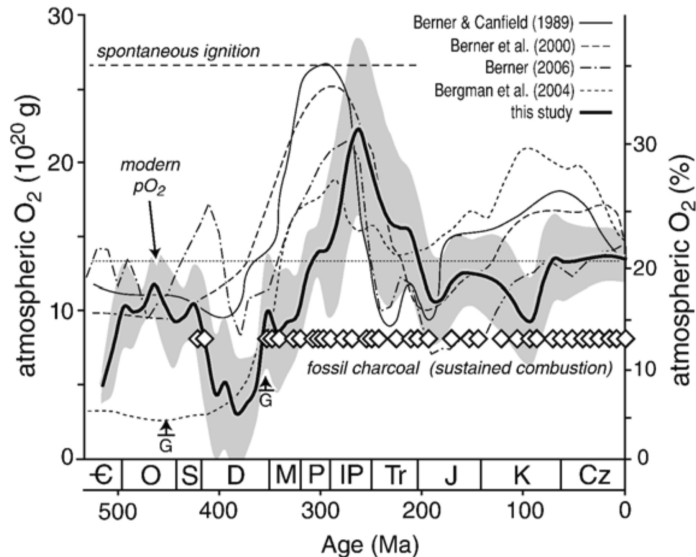
- Page 3, lines 17–18. “which is consistent with a long-term sensitivity of the Earth system to CO<sub>2</sub>”. I do not understand what the authors want to convey here, please rephrase.

We have removed this last part of the sentence starting on line 16 page 3.



• Page 3, line 21. “continuously since the late Silurian”. Fig. 5 of Algeo and Ingall (doi:10.1016/j.palaeo.2007.02.029) (below) suggests that the charcoal record is more or less continuous since the latest Devonian or so.

We thank the reviewer for the comment and have modified the text to reflect the point: “Charcoal appears in the fossil record continuously since the **late Devonian (~ 360 Ma, Algeo et al 2007, Scott and Glasspool 2006).**”



• Page 3, lines 24–34. In my opinion, this paragraph is off topic or, at least, should not be included here.

We acknowledge the reviewers comment but would rather keep this paragraph in for completeness as it helps provide some motivation for discussions on Earth system feedbacks which many readers may be interested in.

• Page 4, line 12. “Cenomanian (mid Cretaceous, ~95 Ma)”.

A typo, thanks for spotting.

• Page 4, line 31. “Changes to the incoming solar radiation [reference?]”.

Reference to Lunt et al (2016) added.

• Section 2 “Methods & Simulations”. This section should be better organized. I suggest using subsections. Here are suggestions:

o Page 5, line 4. “2.1. Models”

o Page 6, line 16. “2.2. Experiences” or “2.2. Boundary conditions”

o Page 9, line 1. “2.3. Data”

o Page 9, line 8. “2.4. 1D energy balance model”

o Page 9, line 26. “2.5. Climate sensitivity”

We thank the reviewer for their suggestions in improving the layout of this section and have added in the subsections and headings as they have suggested.

• Page 5, lines 13–14. “A fixed vegetation distribution of plant functional types is employed”. Which one? A present-day one?

We have changed the text to make it clearer:

“A fixed **present-day** vegetation distribution of plant functional types is employed.”

- Page 5, line 33. “increases in thickness”

We have changed the text to:

“In the vertical, 31 model levels are used which increase **in thickness** steadily between”

- Fig. 2. Temperature unit?

We are not sure what the reviewers comment is? The figures clearly show the surface temperature in degrees Celsius -- the most commonly used unit for the variable -- indicated in both the figure and the figure caption.

- Page 6, line 5. “limited to 4 m thick”?

Changed.

- Page 6, lines 18–19. “as it is possible to alter the model topography and bathymetry”. Please delete.

Deleted.

- Table 1.

o Please explain how the experiments name is built. As it is, the reader has to figure it out himself. The use of “2x” and “4x” in particular, is not obvious. This is placed at the beginning or in the middle of the experiment name and does not refer to any CO<sub>2</sub> level but rather seems to multiply the CO<sub>2</sub> value used in the baseline runs. Please, explain all this, for instance in the caption of Table 1. Also, what’s the “\*”?

o What’s the horizontal bar delimiting the 2 parts of Table 1? I guess that “baseline runs” and “sensitivity tests” may be included to refer to each part.

The simulation names have been harmonised throughout such that XX-YYY refers to time period XX and model YYY, while a prefix of Nx indicates that the CO<sub>2</sub> content has been multiplied by a factor of N with respect to experiment XX-YYY, i.e. 4xPI reads “Four-times preindustrial” which is more intuitive. The horizontal lines now separate the equilibrium experiments from the transient Gregory experiments. The table caption has been expanded to read:

“Experiment names AA-BBB include the continental configuration (AA) and model used (BBB). Experiment names NxAA-BBB indicate a multiplier of CO<sub>2</sub> with respect to AA-BBB. A star (\*) indicates that the CO<sub>2</sub> multiplier was applied instantaneously and the transient adjustment to climate was analysed for the purpose of a Gregory 2004 analysis.”

o Here and throughout (Table 2, Fig. 6, Fig. 9, Fig. 12, Fig. 16 etc.), I would prefer to see the unit in parentheses rather than with a “/”: “CO<sub>2</sub> / Pa” > “CO<sub>2</sub> (Pa)”. The use of “/” is confusing when it does not represent a ratio.

<https://www.bipm.org/en/publications/si-brochure/section5-3.html> suggests that SI recommendations are to use “/” to express units so this convention has been retained in the revised manuscript.

- Fig. 3. Precipitation unit? + “Continental outline is represented with the thick black line” or similar.



We will add “**Continental outline is represented with the thick black line.**” to both Figure 2 and Figure 3 for consistency.

- Page 7, lines 2–4. “The annual average ... Figs. 2 and 3.” Please move these lines and figures into the results section.

These are not really results per se, they reflect the standard output of the model for the base simulations and so we argue they belong where they are in the method section as a reference for the reader.

- Page 8, line 4. I guess this is “O<sub>2</sub> content”.

We have changed “O<sub>2</sub>” to “pO<sub>2</sub>” to reflect this.

- Page 8, line 11. This sounds unlikely, see for instance Fig. 1 of Royer et al. (doi:10.1130/1052-5173(2004)014<4:CAAPDO>2.0.CO;2).

We stand by the fact that absolute constraints on pre-Quaternary CO<sub>2</sub> levels are not as good as those in the Quaternary. However, we suggest toning down/rewording the sentence to make it clearer that we mean that the upper limits of CO<sub>2</sub> are on the order of ~ 100s of Pa, rather than specifically about 100 Pa. The text now reads:

“however there is growing evidence that CO<sub>2</sub> is unlikely to have been significantly higher than the **order of hundreds of Pa** since the radiation of land plants”

- Page 9, line 3. “heuristically”. Well, this is obviously “by hand”.

Agreed, and as we suggested above this section will be removed.

- Page 9, line 24. What’s  $\tau_{s,ebm}$  referring to?

The tau was a typo and should be T. We have corrected this in the modified version.

- Page 10, line 5. Please define “CS” and “CRE”.

They are defined on line 3.

- Page 10, lines 14–21. Here and throughout: the text is sometimes difficult to follow because

the authors do not refer to figure panels. Please explicitly include “(Fig. 4b)” etc. when appropriate. Also page 13.

References to Figure 4a-f have been added at relevant points in the first and second paragraphs of the Surface Climate subsection and references to Figure 6a-f in the final paragraph of the Surface Climate subsection.

- Page 10, line 23. “(Figure 4 centre)” > “(Fig. 4 middle column)”

Changed

- Page 10, lines 25–27. “These could be ... reduction with height”. Is this effect really significant? This could be tested with a flat Earth simulation.

While the result of this sort of idealised study would be of interest, the intention was to note the possible mechanism rather than to quantify the impacts.

- Fig. 4a. Is Panama really open?

The plots show the land sea mask employed by the model. In the case of HadGEM3-AO the ocean is not allowed to mix across the isthmus of Panama but the atmosphere component does not resolve the land bridge hence appears open while the ocean is separated. This is a peculiarity of the differing atmosphere and ocean grids employed by the atmosphere and ocean components.

- Fig. 4. Please define the “cold month” and the “warm month”.

This has been added as “change in the mean gridbox temperature of the coldest month in the monthly mean climatology”

- Table 2.

- o Missing data for 4xPI-GEM.

While preparing the manuscript it became evident that the 4x-PI-GEM21 had not been performed so these results are unfortunately not available.

- o The authors may want to include data for their EXP21 experiments in brackets next to their EXP35 results to permit the comparison with Poulsen et al.’s results.

We’d like to thank the reviewer for the suggestion and did implement this, however the resulting table was so crowded as to render it unclear so this has not been updated in the revised manuscript.

- Page 12, line 9. “Comparing the surface temperature (Fig. 4a,b) and precipitation response”. Precipitation is showed in the next paragraph.

- Page 12, line 12. “air temperature and precipitation anomalies”. Precipitation is showed in the next paragraph.

We’d like to thank the reviewer for noticing these discrepancy. We have moved “Comparing the surface temperature and precipitation response between HadCM3-BL and HadGEM3-AO suggests that the model responses are broadly consistent.” and “A gridbox-by-gridbox comparison of annual mean surface air temperature and precipitation anomalies for PI-GEM<sup>35</sup><sub>10</sub> vs PI-CM<sup>35</sup><sub>10</sub> is presented in Fig. S1. The largest discrepancy in surface air temperature response between the two models occurs for the largest temperature changes simulated by HadGEM, which is strongest in Northern Hemisphere polar regions. This could be linked to differences in the representation of polar climate processes between the two models. There is broad consistency in cold and warm-month means (Figure 4a and b) with stronger warming in the cold month mean and terrestrial cooling in the warm month mean.” to a new paragraph at the end of the Surface Climate subsection.

- Page 12, line 12. “Fig. S1”. Missing supplementary figures? Please check throughout, including on page 15, line 8 + page 18, lines 7–11.

We would like to thank the reviewer for drawing this to our attention. We will ensure that the supplementary figures will be submitted with the revised manuscript - based on the reviewers suggestions we will be including many of these in the main text having removed the Model-data comparison and wind stress sections.

- Page 13, line 1. “representation of polar climate processes between the two models” + amplification by polar ice feedbacks.

The sentence has been adapted to read “This could be linked to differences in the representation of polar climate processes **and amplification by polar ice feedbacks** between the two models”

- Page 13, line 12. **Bjerkness compensation.**

This has been added to the sentence:

“A northward shift in the ITCZ would be consistent with stronger warming in the Northern Hemisphere **due to Bjerknes compensation (Bjerknes 1964, Broccoli et al. 2006).**”

- Page 13, line 15. “suggests that  $pO_2$  could mediate monsoon climate”. Please check in the model output.

Have removed “which suggests that  $pO_2$  could mediate monsoon circulations” as there is no supporting analysis.

- Fig. 6 caption. “Global mean values (mm/day) are offset”. Please rephrase. As it is, this suggests that values are really offset, which would be annoying. The text label is offset.

The global mean values are offset from the plot to the top-right. We have rephrased this from “are offset” to “are offset **to the top-right of each plot**” as for figure 4.

- Fig. 7. and Fig. 8:

- o Bottom left: What are the dashed lines?

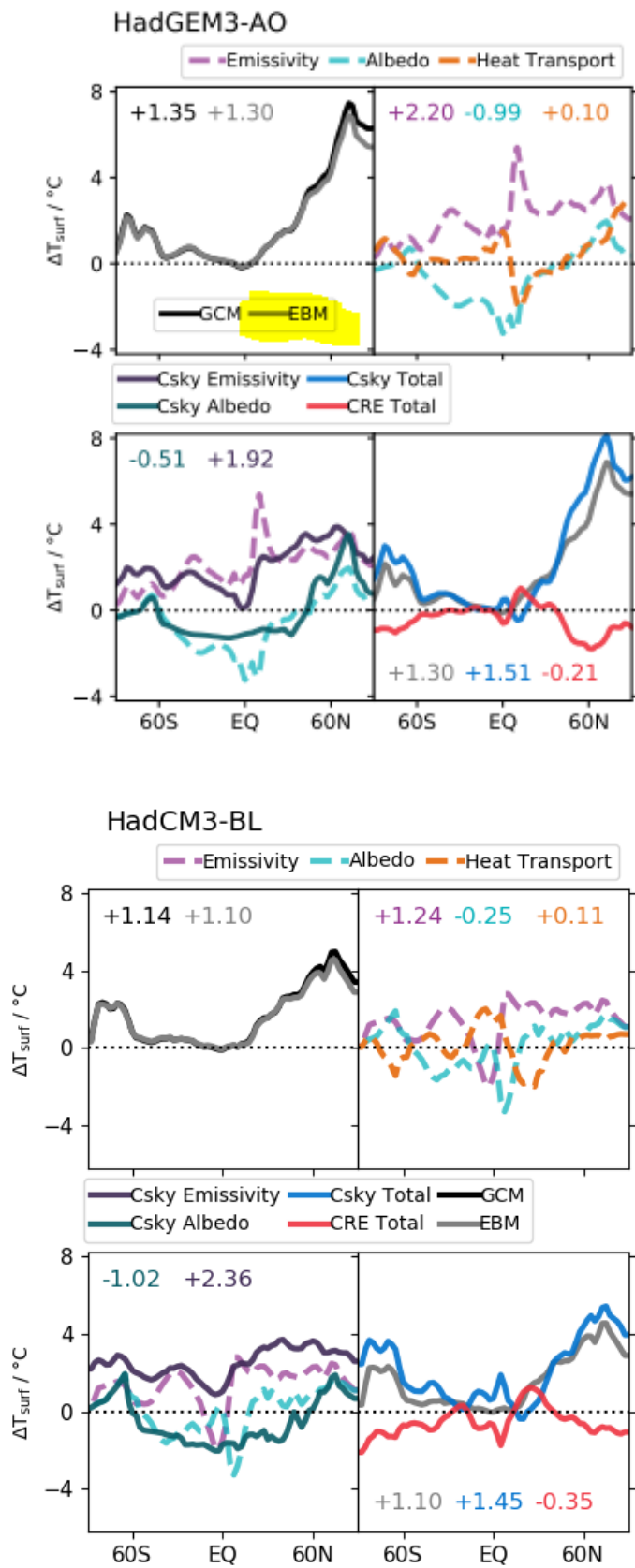
- o Bottom right: Grey line is missing in the legend.

The emissivity is indicated by the dashed purple line and the clearsky emissivity is indicated by the solid purple line while the albedo is indicated by the dashed green line and the clearsky albedo is indicated by the solid green line. The dashed green and dashed purple lines are included again in the bottom-left panel so that the reader can contrast the clear-sky vs all-sky components (the remainder being the cloudy-sky component). We infer from the questions that ambiguity arises from the figure caption, hence this is updated to ensure clarity.

The grey line is included in the legend, however is inset in the top-left plot. To ensure that the reader can more easily interpret that the legend entries are common across the plot Figures 7 and 8 have been slightly adjusted:

“Bottom left: Clear-sky emissivity (dark purple) and **clear-sky** albedo (dark green) components of the EBM. **The all-sky components are included for comparison.** Bottom right: Decomposition of EBM into the total clear-sky (blue), cloudy-sky (red) and all-sky (grey) components.”

Old Figure



• Fig. 9, caption. "top-of-atmosphere radiative imbalance".

Thanks - changed “top-of-atmosphere radiative imbalance”

- Page 18, line 11. “numerically unstable”. Any idea why?

An inspection of the model dump files before the crash revealed very high temperatures in the tropics, however a root cause of the instability could not be found. “Runaway” temperatures\* have been found in the MPI model at high CO<sub>2</sub> (see Heinmann M., PhD Thesis [https://pure.mpg.de/rest/items/item\\_993927\\_4/component/file\\_2388525/content](https://pure.mpg.de/rest/items/item_993927_4/component/file_2388525/content) Chapter 3 p40). This limit may have been reached in HadCM3-BL in the context of the Maastrichtian climate.

\*note this is distinct from a runaway greenhouse effect, the physics of which are typically not properly accounted for in climate models as water vapour is usually assumed to be a minor gas while in a runaway greenhouse will become a major gas and therefore come across similar issues with atmospheric pressure and heat capacity changes. In addition, models would need to better account for water vapour continuum lines than current radiation schemes allow. The rapid runaway temperatures are therefore a model feature rather than an indication that a runaway greenhouse has been reached.

- Section 3.4.

- o I suggest changing the title for something more specific like “Response of Permian vegetation to changing O<sub>2</sub> levels” since this section really deals with the Permian case study.

At the suggestion of the reviewer we have renamed the section “Response of Permian vegetation to pO<sub>2</sub>”

- o The temperature and precipitation dependence of the dominant PFT simulated in the Permian should also be considered. To what extent are the changes in vegetation cover and type due to changes in temperature and precipitation? Changes in precipitation in the Wu-CM runs (Fig. 6e), in particular, seem to spatially correspond to the expansion of the BLT PFT (Fig. 10). I would like to see a short analysis of the environmental affinities of the main PFTs shown on the maps in Fig. 11. I suspect that temperature and precipitation threshold values may play a more important role than changing O<sub>2</sub>/CO<sub>2</sub> ratios.

Disentangling the myriad of impacts on vegetation cover would be a challenge. The goal of this section is to highlight that pO<sub>2</sub> will have implications beyond those of photorespiration, including the impacts of precipitation and temperature changes.

- Fig. 11. The color map is reversed from a to b, which makes it difficult to read. Please revise.

The color maps used are different in all three plots as they represent different variables. If the reader has deuteranopia the scales should still be discernible but will appear to be reversed from a to b.

- Page 19, line 15. “expansive”. What does than mean?

Greater in extent - reworded to “broadleaf trees **cover more area** in...”

- Page 20, line 8. I guess it means that the simulated changes in carbon storage on land do not impact the pCO<sub>2</sub> level? It would be good to clearly state what the authors mean by “not interactive”.

David Wade on behalf of the authors

This is correct. The sentence now reads “the carbon cycle is not interactive (**atmospheric CO<sub>2</sub> is fixed**)” to clarify this point.

- Page 20, lines 20–22. “however, it is likely ... equator-to-pole temperature gradient”.

Please

provide references to support this statement.

This section has been removed from the revised manuscript.

- Section 3.6 “Importance of Wind Stress”. I get that atmospheric mass impacts wind stress, which in turn impacts the ocean circulation and the heat transport (see page 4, line 20). Unless I get it wrong, those effects are included in the coupled ocean-atmosphere simulations conducted by the authors, which is a good point. However, I do not understand why the authors test the impact of removing wind stress. In my opinion, this section is off topic and should be deleted and possibly kept for another contribution, which would simultaneously shorten the present manuscript and leave the possibility to conduct a robust analysis of the climate response (including the response of ocean dynamics, the analysis of which is essentially lacking so far). (For this reason, I did not include the minor comments relative to this section in this review).

We have removed the section, as advised by the review in the major comments above.

- Fig. 12, caption. “Proxy data locations (Upchurch et al., 2015) are indicated”.

This section has been removed, as advised by the reviewer in the major comments above.

- Fig. 14, caption. What’s the unit of precipitation in panel c? Is this an annual mean?

The unit mm day<sup>-1</sup> has been added to the caption (is already included in the plot). It is annual mean, this has also been added. Now reads:

“**annual mean** precipitation (**mm day<sup>-1</sup>**)”

NB this is now Figure 15 in the revised paper.

- Page 24, line 1. “mainly due to”.

We have added the word mainly:

“HadCM3-BL simulates a reduced equilibrium climate sensitivity **mainly** due to changes in longwave cloud feedbacks.”

- Page 24, lines 2–3. “The pre-industrial Holocene ... of the Archean”. Please support this statement with appropriate references.

The references Payne et al (2016), Chemke et al (2016) and Charnay et al (2013) have been added to support the statements made in this sentence.

- Page 24, lines 13–15. So, the implementation of pressure broadening is not the same as in Poulsen et al.? Page 7 line 2 suggests that O<sub>2</sub> forcing is analogous to Poulsen et al. Please clearly state what’s common between both studies and what’s different. Also, I encourage the authors to make their numerical code available for future work (see major comment).

The O<sub>2</sub> forcing *ought* to have been made in an analogous way, however without a deep understanding of the structure of the climate model used and the relevant subcomponents it would not be possible to verify the implementation or identify the commonalities between the implementations. Even if the GENESIS model code and settings were made available, it would be beyond the expertise of the authors to be able to provide a detailed verification. It



is for this reason that we do not believe that releasing model code changes would be useful particularly as the HadCM3-BL and HadGEM3-AO model codebases are under UK crown copyright and the required model changes may vary considerably between models. It is worth noting that while the general approach to changing the  $pO_2$  was the same in HadCM3-BL and HadGEM3-AO the specific details are considerably different. Hence we propose that the similarities between the model results and similarities with the more idealised studies elsewhere are what give confidence in the way that the general approach was applied in the specific model versions.

- Page 24, lines 18–21. Since sub-daily model output was not written on disk in Wade et al. model runs and thus not made available for analysis, I suggest deleting this comparison that is not that instructive.

These two sentences have been removed in the revised manuscript.

- Page 24, lines 22–24. Please be cautious: even in a slab model, the continental configuration can impact the ocean heat transport due to the varying ocean area and global climate and thus temperature (which is also impacted by the continental configuration). Another, maybe more robust argument to support the comparison, is that (i) both reconstructions are not so different at first-order and (ii) both simulations provide a relatively close global climate state (compare the mean annual SAT in Poulsen et al. 21% and this study 21% – ca. 18°C vs. 22°C).

We thank for reviewer for this suggestion and have included in this sentence, which we have slightly adjusting to clarify which Cretaceous stage was used.

“Note that the Poulsen et al. 2015 simulations were for an earlier Cretaceous period (**Cenomanian**) than those performed in HadCM3-BL (**Maastrichtian**), **however the continental configurations and the global mean temperatures are reasonably similar (22.2 °C in HadCM3-BL vs 20.5 °C in Poulsen et al 2015).**”

- Page 24, lines 25–31. The contribution of changing ocean dynamics / deep circulation to the simulated climate changes is addressed for the first time here. I suggest either deleting this unsupported statements or providing clear diagnostics of the changes in ocean dynamics.

We have removed lines 25-31 and lines 32- p25 l1-2 as both paragraphs discuss the offending material.

- Page 25, lines 1–2. The authors may want to refer to Pohl et al. (doi:10.5194/cp-10-2053-2014), who demonstrated the importance of ocean dynamics to simulate Ordovician climate changes.

We'd like to thank the reviewer for this suggestion. We have removed this paragraph as a result of the above comment, so have not included it in the revised manuscript.

- Page 25, lines 5–6. “Increases ... high latitudes”. Please provide a reference. Kiehl and Shields 2013 reference has been added.

- Page 25, lines 3–14. The authors may also want to refer to the climatic mechanism demonstrated by Rose and Ferreira (doi: 10.1175/JCLI-D-11-00547.1), which was subsequently invoked by Ladant and Donnadieu (doi:10.1038/ncomms12771) to explain the climate changes observed in their Cretaceous model runs.

We'd like to thank the reviewer for this insight, while not completely analogous to Rose and Ferreira the enhanced convection at low  $pO_2$  is consistent with an atmospheric moistening. Mechanistically, this is more sensitive in a warmer climate due to Clausius-Clapeyron. The subsequent analysis and addition to the manuscript has been included in a response for Anonymous Reviewer #3.

- Page 25, line 16. "While subsequent experiments have put this in doubt". Please support this statement with a reference.

Wildman et al. 2004 reference added

- Page 25, lines 24–25. "although there is evidence of vegetation which causes C4-like fractionation". During which geological period?

Lower Carboniferous - added to the text as "fractionation in the Mississippian, suggesting"

- Page 25, line 31. "Other approaches such as trait based methods". The authors may want to cite Porada et al. (doi: 10.1038/ncomms12113) who applied a trait-based model to simulate the impact of the Ordovician primitive terrestrial vegetation on weathering.

We'd like to thank the reviewer for drawing our attention to this interesting and relevant article and have added to the reference in this sentence: "Other approaches such as trait based methods (Van Bodegom et al., 2012; **Porada et al., 2016**)"

- Page 25, line 32 to page 26, line 2. I suggest deleting this paragraph.

We have removed "We also have not accounted for changes... may be sensitive to atmospheric  $pO_2$  (Clarkson et al., 2018)" inclusive.

- Figure 16. The authors previously demonstrated that Poulsen et al.'s implementation of  $O_2$  forcing may not be robust. I think this is thus relatively unexpected that they here use those results in their Phanerozoic calculations, even if this may constitute a conservative estimate. I suggest using the results of the current study instead.

The motivation for doing so is that *even if* the climate state was that sensitive to  $pO_2$  (which seems unlikely based on previous studies and is also not supported by the proposed study) the differences this causes to global mean temperatures is minor compared to that due to the uncertainty in  $CO_2$  content.

- Page 27, lines 17–18. "If  $pCO_2$  and  $pO_2$  ... in the Phanerozoic". Why? Why would cooler climates be associated with higher  $pO_2$  levels? I cannot imagine any clear and straightforward explanation to this.

$CO_2$  and  $O_2$  are linked by photosynthetic productivity and organic carbon burial - increased carbon burial reduces  $CO_2$  and increases  $O_2$  ( $CO_2 \rightarrow C_{org} + O_2$ ), see e.g. Montañez 2016

(<https://doi.org/10.1073/pnas.1600236113>) in the context of  $pCO_2$  and  $pO_2$  in the Carboniferous-Permian. This mechanism was invoked by Poulsen et al 2015 also.

#### C. Minor points:

- Page 3, line 2. "(grey shading in Fig. 1)".

Changed

- Page 5, line 13. "which simulates".

Changed

- Page 6, line 12. Reference formatting: “(Valdes et al., 2017)”.

Changed

- Page 6, lines 20–21. Reference formatting.

Changed

- Page 7, line 3. Please use correct experiment names instead of 4 x PI-CM21.

This has been updated and also now includes the sub-letters as the reviewer kindly suggested in an earlier comment:

**“Fig. 1 shows the annual average surface temperatures and Fig. 2 shows the annual average precipitation for the (a) PI-CM<sup>21</sup>, (b) Ma-CM<sup>21</sup>, (c) Wu-CM<sup>21</sup> and (d) As-CM<sup>21</sup> simulations.”**

- Page 8, line 17. “monotonically increasing ozone column”. Please rephrase.

The ozone column increases monotonically with  $pO_2$  in the Harfoot et al 2007 study, hence was described as such in the manuscript.

- Page 8, line 29. Reference formatting.

Changed

- Page 10, line 19. “Wu-CM” (lower-case).

Changed

- Page 10, line 21. “This suggests that ... but is non-linear”. Please revise.

We have reworded to “This suggests that the climate response to  $pO_2$  variability depends on the background climate state.

- Page 12, line 14. “, which are strongest”.

Changed

- Fig. 5, caption. “4XPI-GEM(red)”: missing space. See also multiple occurrences on page 18, lines 1–5.

Changed (fig 5 caption) and updated 5 occurrences page 18.

- Fig. 6 color bar labels. Font size issue leading to overlapping text.

An adjusted figure has been provided in the revised manuscript

- Page 19, line 6. “Fig. 11a”.

Changed

- Page 19, line 16. “reduces” > “reduce” (2 occurrences on the same line).

Changed

- Page 20, line 13. Please delete question mark.

Section removed in the revised manuscript

- Page 20, lines 25–26. “These show that across both CO2 contents that increasing”. Please rephrase.

Section removed in the revised manuscript

- Page 22, lines 4–5. “has the capacity to alter the radiative budget of the atmosphere and therefore on Earth’s climate”.

This is verbatim from the original manuscript. Rereading we are suggested to rephrase this text to: “therefore **has implications for** Earth’s climate.” We hope this aids in the clarity of the sentence.

- Page 22, line 6. “with increasing  $pO_2$ ”.

Colon added

- Fig. 15, caption. Missing space.

Space added

- Page 25, line 3. “also contribute”.

This sentence has been removed from the revised manuscript.

David Wade on behalf of the authors

- Page 25, line 9. “compared”.

This sentence has been removed from the revised manuscript.

- Page 25, line 13. “which may increase lead to”. Please revise.

This sentence has been removed from the revised manuscript.

- Page 27, line 7. “PAL”. Please write in full or provide meaning.

Changed: PAL → “present atmospheric levels of”

- Page 27, line 16. “When pO<sub>2</sub> was higher”.

Changed: were → was