Responses to the comments of Reviewer 1

Below, comments from the referee are reproduced in *italics*, while our detailed responses to each comment are in **bold**. We have produced a revised version addressing the reviewer comments, which can be found at https://github.com/KEichenseer/PalaeoClimateGradient/blob/main/manuscript/main_revision_track_changes.docx

RC1:

The authors present a novel approach in an attempt to improve latitudinal temperature gradients from spatially scarce proxy estimates.

While their efforts can clearly help substantially towards better assessing and interpreting the available proxies, the authors do not provide a convincing case to support their claims regarding the early Eocene. The technicalities and implementation of the method is solid, but the results allow only for a limited assessment of its applicability. Many of the limitations considering these proxy estimates lie in the methods behind the proxies themselves, something that is not adequately addressed in my opinion. I would therefore like to see some more tests including the potential effects of e.g. seasonality, lacking upper/lower bounds, or differently-sourced temperatures (ML ocean, SST, SAT etc.).

Regardless, I believe that this study can substantially benefit the field and therefore suggest publication after some adjustments/additions.

We thank the referee for their feedback and for endorsing the publication of our manuscript after some adjustments. We agree that the limitations regarding the proxy data primarily lie in the methods behind the proxy data. Our model is developed to help address issues with limited and uneven spatial sampling, not seasonality and other shortcomings with which individual samples may be afflicted.

There is indeed potential evidence that seasonal effects may play a role in the higher-latitude geochemical proxy data: We have now added a figure to the Supplementary Material (Fig. S6) showing that the standard deviation of the temperature estimates provided by proxy data from the same location increases with absolute palaeolatitude. There are, however, other effects that could explain this pattern. For example, mid- to high-latitude regions are more sensitive to climate change through time, and time averaging may thus contribute to the larger variability of higher-latitude temperature data.

The structure of our model inherently accounts for such effects by not using the proxy temperatures directly, but instead estimating a local temperature that preserves the uncertainty associated with the variation in proxy estimates. Seasonal effects would thus only be problematic if they systematically bias the temperature proxy record in one direction. The published data compilation we use does not provide information that would allow us to account for this.

As an additional test, we have now expanded our analysis to also include the uncertainties associated with each individual proxy measurement, and present these new results in the Supplementary Material (Fig. S5). As the reviewer notes, the issue primarily lies in the geochemical proxies. This is not a limitation of our model, but of the input data, and should be addressed therein. Currently, the resolution required to be able to account for seasonality in geochemical proxy records is not available. While

it might be tempting to apply some standardised distribution of seasonality with respect to latitude, seasonality is known to–though not well constrained–vary with climate state.

General remarks:

Introduction

- *spatial patterns and distribution are discussed in detail, but how is temporal variation covered?*

  When considering proxies within a certain interval, they may represent entirely different subsets of this interval and therefore not be compatible.

Time averaging is a common problem at all scales, whether it be decadal, millennial, or so forth. We restrict our analyses to the Early Eocene Climatic Optimum (ECCO), an episode of sustained global warming. We treat the ECCO as a distinct interval, and while subsets of this interval might represent warmer or cooler sub-intervals, we assume the long-term average of this interval is represented by the data. We would of course be interested in conducting more temporally-resolved analyses, however, data at temporal scales <50,000 years is often unavailable for pre-Pleistocene intervals due to limitations in dating.

- *(L73) The introduction could use some background and references on Bayesian modelling and why one could expect this method to be a suitable tool to the problem raised.*

Thank you. We have now expanded the justification of the Bayesian method, and added references in the introduction (see lines 78–89). We also now provide further details in the methods and discussion sections.

Methods

- *I assume the authors consider near-surface, surface or mixed layer ocean temperatures? I am missing some clarification and motivation here.*

We use the sea surface temperature (SST) data compilation from Hollis et al. 2019. We minimise depth differences in the origin of the SST temperature proxy data by excluding any data labelled as thermocline or sub-thermocline, and otherwise rely on the screening of Hollis et al. 2019 for excluding data not representative of SSTs. Realistically, the proxy compilation from Hollis et al. 2019 used herein may be more accurately characterised as including data from the mixed layer, but we use the term SST as is conventional for palaeoclimate studies, including the Hollis et al. 2019 compilation. We have clarified this in the Methods section (lines 107–110). We would also like to clarify that our manuscript is intended to introduce a new approach to reconstructing latitudinal temperature gradients from spatially incomplete and uneven datasets, which is why we used an already published dataset.

- *Are the temperature limits mentioned considering yearly, seasonally, monthly or extreme values?*

  Especially towards higher latitudes, seasonal temperatures may be much more restrictive than yearly averages.

We used mean annual temperatures to define our ecological proxies, which we deem most appropriate as we are not modelling seasonal temperatures; instead the geochemical proxy compilation likely consists of temperature signals averaged across seasons, and/or from different seasons. We acknowledge that using seasonal temperatures would be preferable for a seasonally resolved analysis, which we cannot provide herein. However, these limitations lie in the geochemical data and the methods used to generate temperature estimates, not our modelling approach. When robust data capturing seasonality is made available, we will of course expand our model to account for seasonality effects.

Model validation

- *As the authors mention in the introduction, the spatial distribution of temperature estimates can greatly limit the skill of derived latitudinal gradients.*
- *In the model validation, is there a way to not only consider the shear amount of samples, but also their spatial clustering?*

  Looking at many random spatial distributions of temperature estimates, one may get a too optimistic view on how well they could capture the considered gradient.

We agree with the reviewer that it is important to also consider spatial clustering. We provide an empirical example of data clustering and how well our model performs across different climate states by using the spatial distribution of our Eocene dataset. There are of course a vast amount of potentially different spatial distributions for geochemical data. However, our approach demonstrates a model validation which would be preferable for all datasets: test whether the empirical spatial distribution of the data would allow different gradients to be reconstructed (i.e. under different climate states).

Results

- *Are hemispheric asymmetries considered? These may differ substantially between the current and Eocene climate.*

We chose a simplified approach, combining both hemispheres, due to the scarcity of EECO data, and particularly the absence of high-latitude southern hemispheric data. We show separate, hemispheric analyses of the EECO data in the supplementary materials (Fig. S4), and have now added hemispheric, empirical temperature gradients of the modern for comparison.

- *Somewhat philosophical question: considering the idealised profiles in Figure 1, one would hardly see any difference between the extreme icehouse, icehouse and present climate using the adopted method to determine the latitudinal temperature gradient. Were other measures explored in this sense?*

The idealised icehouse gradient was indeed closely resembling the present gradient. We have now replaced this idealised gradient with the actual modern gradient. The extreme icehouse gradient is loosely based on reconstructions of Pleistocene ice ages, showing a much steeper gradient than the modern.

- *Looking at the main result in figure 4, I find it hard to see the added value of the method presented in this work.*

The temperature estimates based on coral reefs in the tropics seem highly doubtful (this is briefly touched upon in the discussion), while information at higher latitudes is still extremely scarce.

Likely, the potential influence of seasonal biases in some high latitude proxies are potentially problematic for the method, this is again only briefly mentioned in the discussion.

In that sense, I am not convinced about the authors' claim that this method succeeds in providing an unbiased estimate of the latitudinal temperature gradient of the Early Eocene climate.

Our results are novel in that they are offering an estimate of EECO latitudinal SST gradients that quantitatively combines geochemical and ecological proxy data, while acknowledging the uncertainties inherent in these data with the best of our ability. We understand the concern that the ecological proxies come with considerable uncertainty, and show an alternative analysis, leaving out the ecological proxy data, in the supplementary materials (Fig. S2). We have now also added an analysis with wider uncertainties in the ecological proxy data in the supplementary materials, following the suggestions of both referees. However, we do not believe that ecological proxies are inherently less reliable than geochemical proxies; both are subject to considerable uncertainties.

Specific Comments:

- *L43: and strongly limited by the possibilities to determine temperature estimates from different proxies.*

  This is associated with many assumptions, the most important probably being that any relations found in present experiments still hold in the distant past.

We have now acknowledged this in the following paragraph (lines 46–48).

- *L55: this is indeed the case in a general sense, but there are clear exceptions e.g. near fronts (e.g. the ACC or a WBC). How can we tell whether a certain proxy is representative of the surrounding region?*

We acknowledge this limitation in the discussion (l. 431); our model predicts the average temperature gradient, to account for regional phenomena such as the WBC, a 2D model would be needed.

- *L121: What is the motivation to use a normal distribution when it is known that the actual distribution is skewed?*

Given the possibility of warmer-than-modern coral reef habitats in the early Eocene, a skewed distribution based on the modern coral reef environmental parameters would exacerbate the potential underestimation of temperatures at Eocene coral reef locations. We prefer a broad, normal distribution which acknowledges the large uncertainty inherent in this proxy.

Furthermore, the statistically derived temperature range falls well short of the potential maximum of 35.6C mentioned earlier, how is this consistent?

- *L146: What motivates the maximum value of 29.5C for mangroves? This may push tropical temperature ranges down considerably, and thus needs to be justified.*

We have added a sensitivity test using a broader temperature range of 16–35.6ºC (combining this suggestion and that of reviewer 2) for both the corals and the Avicennia-Rhizophoraceae mangrove assemblage. The result is very similar to the original analysis (Fig. S3). The value of 29.5ºC is not a hard maximum, as the tail of the probability distribution extends beyond that value.

- *L210: It is unclear to me how the emulated climatic states represent a realistic simulation. If these are highly idealised, they may not be suitable to purposefully test the bayesian model.*

We have now replaced one of the idealised gradients with the empirical, modern gradient. The goal of this validation is only to show how the model setup performs with differently shaped gradients, and that the model is able to infer gradient shapes that are different from the modern.

- *L290: The us of the word 'modelled' is a bit ambiguous here, as the study still considers estimates from proxy-based data rather than numerical climate models.*

  The autors may instead consider using e.g. 'estimates from our Bayesian model', or 'proxy-based model estimates'.

Thank you. We have now followed this recommendation throughout the text.

- *L308: Considering hemispheres separately is new at this point and should therefore not be considered solely in the results section.*

Thank you. We have now mentioned it in the Methods section. (lines 182–184)

- *L359: A big caveat here is that the inclusion of ecological constraints is highly dependent on the underlying assumption, something I feel is not adequately addressed here.*

  There is yet, however, sufficient discussion on several related aspects further down.

Thank you. We think that the discussion on this aspect belongs in the next paragraph (l.401).

Figures:

- *Figure 1: This figure needs some grid lines and coordinate labels.*

Thank you. We have modified Figure 1 accordingly.

- *Figure 3: As shown clearly in this figure, equatorial temperatures are on average cooler than tropical temperatures.*

  This may be too detailed for the scope of the study, but this would suggest *using average tropical temperatures rather than equatorial ones would be better suitable to estimate the latitudinal gradient.*

In our study, we simplified the shape of the gradient to assume highest temperatures at the equator. We consider this simplification justified, as the difference in averages between equatorial and the highest tropical sea surface temperatures in the modern is only ~0.5 C, which is much lower than the uncertainties in our reconstructed gradient.

In general, using average tropical temperatures to calculate the magnitude of the latitudinal gradient has its own problems, as the width of the tropical climate belt has varied through time, and using tropical averages may thus be influenced not only by the magnitude, but also by the shape of the gradient within the tropics.

- Figure 2/4: It is described in the methods that the latitudinal gradient of the prior lies at middle latitudes, this is however very different from the emulated greenhouse climates.

  In figure 4, we again see the maximum gradient shown at middle latitudes. This seems inconsistent with the emulator cases and also poses the question what determines this position?

  A profile much like the one shown in the bottom panels of Figure 2 would likely result in highly different polar temperatures and thus gradients.

We agree that the prior on M we used in the initial analysis was too restrictive to fully emulate greenhouse climates. We have decided to broaden the prior on M, which now allows the latitudinal SST gradient to follow the data more closely, and to better depict greenhouse climate gradients.

Tables

- *Table 1: why not use the recently published DeepMIP model results for this estimate?*

  In this table, it should be explained better what the gradient means and at least have units (I assume degree C?). Is this a regression, a difference between points/regions?

We have added the DeepMIP model results from Lunt et al. (2021) to the table.

Small remarks:

- *L121: values are missing units (likewise in other distribution values further down).*

Thank you, this has been corrected.

- *L291: usage of degC after the brackets is a bit awkward.*

This has now been corrected.