Responses to the comments of Reviewer 2

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

RC2:

The authors provide a novel quantitative method for reconstructing Eocene temperature gradients from sparse proxy data. The paper is well written, and the implementation is mostly solid, but I am not yet convinced of their claims for the early Eocene. In particular, I'm concerned that the authors may have overestimated their model's skill by neglecting to include proxy biases and noise in their model's validation. I would also appreciate a bit more context and/or justification for several methodological choices – such as the choice of a logistic model, and the sensitivity to the ecological parameters. That said, I think this paper has excellent potential to improve the paleoclimate field, and I recommend its acceptance following various revisions.

We thank the referee for their feedback and for endorsing the publication of our manuscript after some adjustments. We expanded the justification of the method in the introduction (lines 85-89), and have added additional details on the choice of the logistic functions in the Methods (lines 177-184).

Comments:

Introduction – I would like some more context on the use of Bayesian models for paleoclimate reconstruction here. How have BHM's been used before, and why are they a good choice for this reconstruction? Same for the choice of the logistic model. I thought there were some nice points in lines 320-330 that could be useful here.

We have expanded the introduction (lines 79-81) to add some examples. Bayesian hierarchical models are increasingly commonly used in (palaeo-) climatic studies, and a de-novo explanation of the Bayesian approach and philosophy is not deemed necessary here. We have added some general references on Bayesian modelling in the introduction (l.76). We now justify the choice of the logistic function in more detail in the introduction (lines 85-89).

Equations 2, 5, 6 - Similarly, I'd like more context for the choice of this model and its design. Is this logistic model's design a common setup for paleoclimate? If so, some citations would be nice. If this is a completely novel approach, then I'd appreciate more discussion as to why the authors made these choices. By contrast, I appreciated the discussions around equations 7-10 and thought these were well justified.

Logistic functions are ideally suited for reconstructing latitudinal temperature gradients, as they are flexible, yet require few shape parameters. Despite this, logistic functions are not yet commonly used to model latitudinal temperature gradients for palaeoclimate reconstructions. We have added a justification to the methods (lines 177-182).

119 – Have the authors done any sensitivity testing of the model's ecological constraints?

One immediate example: The minimum and maximum temperatures used to define the coral distributions (21 – 29.5 C) seem to be drawn from the mean values listed in Table 3 of Kleypas et al., 1999. However, the range of extreme values listed in that table (16 – 34.4 C) is considerably broader and could also be a reasonable choice. Does using the broader range noticeably change the results of the analysis? If so, this should be noted.

We do not agree that the extreme range (based on coldest and hottest weekly temperature) is a good alternative to describe the modern coral habitat in terms of annual mean temperature: A mean annual temperature corresponding to the extreme end of temperature tolerance would imply even more extreme temperatures during part of the year, due to seasonality. However, we have added a sensitivity test using the temperature range of 16 – 35.6ºC for both the corals and the Avicennia-Rhizophoraceae mangrove assemblage to follow the suggestion of referee 1. The result is very similar to the original analysis (Fig. S3).

211 - I would appreciate a slightly more detailed description of the gradients and how they were constructed.

We have added additional description to the Methods (lines 177-182).

216 – I'm concerned that this validation is neglecting the effects of bias and noise in the proxy data. The idealized gradients and limited spatial sampling are a great start, but the current setup seems to assume that proxy data is a perfect record of past temperature. In reality, this is not the case, and I would like to see the validation take this into account. Incorporating the effects of proxy seasonality and auto-regressive noise would be my two foremost concerns.

We have now adjusted the validation to allow for uncertainty in the proxy data by including the average uncertainty of the geochemical proxy values reported by Hollis et al. 2019, corresponding to a standard deviation of 3.8. This may serve to simulate some seasonal effects, but as we do not know the magnitudes and variations of seasonal effects at the EECO sampling locations, it is hard to devise a comprehensive evaluation of potential seasonal effects. Similarly, the proxy data used herein do not contain information on auto-regressive noise.

280 – I'm curious why the authors have limited the prior to the modern empirical gradient. Is there a reason for not using priors derived from Eocene climate model simulations?

We wanted to avoid letting a prior based on Eocene climate model simulations push our model results in the direction of already existing, Eocene climate model estimates. A prior broadly based on the modern is conservative in that it ensures that a latitudinal temperature gradient very different from the modern is inferred based mainly on the data, not on the prior. Following a comment from reviewer 1, we have now made the priors on the parameter M wider, to allow for a larger range of gradient shapes, which allows the modelled gradient to more closely fit the data.

301 – "the early Eocene data does not fit as well to the logistic latitudinal gradient model" - This begs the question of whether the logistic model is a reasonable choice here. Again, I'd suggest adding more context for the selection of this model.

We have broadened the prior on M, which allows for shapes that less resemble the logistic function (see reply to previous comment). We have added context on the model choice in the Methods (lines 177-182).

This might be beyond the scope of the paper – but are there other models that might fit this data better?

Models with more parameters or non-parametric models such as GAMs or splines will be able to fit the data better, at the cost of allowing inverse gradients and other unrealistic shapes. The high residual standard deviation (compared to the modern) is not a problem of the specific function we chose, but of the scattered data. These data are potentially, as the referee remarked earlier, influenced by seasonal and temporal signals, and possibly by miscalibrated proxies.

Figure 3 **–** *Please add R^2, and sample size (N) to either the figure or the caption. Also, I suspect many of the grey dots are obscuring data points behind them. If this is the case, consider using a heatmap-style shading for the grey dots.*

Thank you. We have added R^2 and sample size, and added transparency to the grey dots to indicate the density of points. We experimented with heatmap-style shading, but the result was difficult to interpret.

Minor notes:

83 - This sentence runs on a bit. Consider splitting.

93 - I think it would be best to reference Figure 1 in the text of this section.

121, 141, 149 – If I'm understanding this correctly, the standard deviations were selected specifically with the 97.5/95% distributions in mind. I'd suggest rewording slightly to clarify this point.

127 - Extra comma after "empirical"

138 - Remove "being"

139 - Is "ascribed" the right verb here? Perhaps "used" or "assumed" instead?

252 - Missing capitalization

254 - "where then" -> "were then"

256 - This paragraph changes tenses several times

260 - I believe this references Figure 4 before figures 2 and 3

Thank you. We now have corrected all these issues.