Bayesian multi-proxy reconstruction of latitudinal temperature gradients for the Early Eocene

Kilian Eichenseer[[1]](#footnote-20) and Lewis A. Jones[[2]](#footnote-21)

# Abstract

Geochemical data such as oxygen isotopes (d18O), tetraether indices (TEX86) and magnesium calcium ratios (Mg/Ca) are routinely used to reconstruct ocean temperatures in deep time. However, limited spatial coverage and disagreements between geochemical proxies compromise the accuracy of these reconstructions. Geological data such as coral reefs, mangroves, and evaporites have the potential to supplement geochemical data by improving spatial coverage and constraining temperature estimates. Historically, these data have been used to qualitatively inform upon palaeoclimatic conditions at broad spatial scales, yet no transfer functions exist to convert geological data into temperature estimates. Quantitative temperature reconstructions have therefore not made use of these data. Likewise, the large spatial gaps in existing proxy records have often been bridged by simple averaging, without taking into consideration the spatial distribution of the samples. Here, we present a Bayesian hierarchical model to integrate geological data–with established temperature proxies–into a unified quantitative framework. We apply this approach to the early Eocene climatic optimum (EECO), the interval with the warmest sustained temperatures of the Cenozoic. Assuming the conservation of thermal tolerances of modern coral reefs and mangrove taxa, we establish broad sea surface temperature ranges for EECO coral reef and mangrove sites. We integrate these temperature estimates with the EECO geochemical shallow marine proxy record to model the latitudinal temperature gradient and its uncertainty. Our results confirm the presence of a flattened latitudinal temperature gradient and unusually high polar temperatures during the EECO. We show that the inclusion of ecological data can substantially reduce the uncertainty on temperature estimates in climate zones lacking geochemical data.

# Introduction

The increasing availability of large compilations and databases of deep-time climate data now offers the opportunity for empirical palaeoclimate reconstructions of unprecedented scope (e.g. Veizer & Prokoph 2015; Hollis *et al.* 2019; Scotese *et al.* 2021; **Grossman2022?**; **Judd2022?**). To exploit the full potential these data sets offer, two fundamental challenges need to be overcome: Disparate types of proxy data need to be integrated to inform a single climate reconstruction, and large-scale scale climate patterns need to be inferred from localised, often sparse observations. To address these challenges, we present a novel Bayesian method that combines quantitative proxies and palaeoecological constraints into a fully quantitative model of the latitudinal gradient of sea surface temperatures, bridging spatial gaps in sparsely sampled climate data.

Existing palaeoclimate reconstructions are generally based on one of three approaches (see **Burgener2023?**): 1) Quantitative, geochemical proxies such as oxygen isotopes (d18O), tetraether indices (TEX86) and magnesium calcium ratios (Mg/Ca) are used to derive palaeotemperature estimates using experimentally derived transfer functions (e.g. Liu *et al.* 2009). These data can be readily processed for regional or global-scale palaeoclimate reconstructions [@…]. 2) Palaeobiological, ecological and geological proxy data can provide broad climate constraints, but no experimentally determined transfer functions exist to directly infer quantitative climate parameters from these records (e.g. **Kvacek2007?**; **Cao2019?**; but see **Royer2012?**). Instead, palaeoclimate signals are inferred based on the climatic conditions of analogous, modern sediments, or based on the premise that the climatic requirements of ancient taxa, plant traits or ecological communities were similar to those of their nearest modern relatives (**Kvacek2007?**; **Greenwood2017?**). 3) Earth system models offer the most highly resolved and complete climate reconstructions, but ultimately rely on constraints provided by the geological and fossil record (**Lunt2021?**).

Even when using the full range of available records, proxy-based climate reconstructions beyond the Holocene are inevitably patchy (e.g. Hollis *et al.* 2019), and global temperature reconstructions based on those records alone may be biased (Jones & Eichenseer 2022). To infer regional or global-scale patterns from an incomplete record relies on the bridging of spatial gaps. A common approach that requires no additional computation is the spatial visualisation of proxy-derived temperatures against latitude, showing broad latitudinal temperature trends Vickers *et al.* (2021). Interpolation is sometimes used to bridge spatial gaps in palaeoclimate data (Taylor *et al.* 2004), taking advantage of the autoregressive nature of climatic data: much of the information on the climate of any given location is contained in the climate data of nearby locations (Reynolds & Smith 1994). Adding to this, some proxy-based reconstructions use statistical modelling to infer climatic patterns. For example, polynomial regression has been used to reconstruct latitudinal temperature gradients (Bijl *et al.* 2009), and 2D-reconstructions of surface temperatures and precipitation have been created with Gaussian process regression (Inglis *et al.* 2020; Chandra *et al.* 2021).

Here, we present a Bayesian, hierarchical model for inferring sea surface temperatures (SSTs) that expands upon existing, spatial reconstructions of palaeoclimate by allowing for the integration of 1) prior information based on physical principles and on the observed, modern SST distribution, and of 2) geochemical and ecological climate proxies in a common, quantitative framework. nearest living relative…

We apply this approach to the record of the early Eocene climatic optimum (EECO), the interval with the warmest sustained temperatures of the Cenozoic (Pross *et al.* 2012). The early Eocene is the focus of a substantial body of palaeoclimate literature (**Berggren1998?**; **Thomas2000?**; **Hyland2017?**; **Lunt2021?**), as it potentially represents an analogue for extreme climate warming scenarios (Burke *et al.* 2018). For palaeoclimate modelers, the early Eocene presents a challenge: Geochemical proxy data (Hollis *et al.* 2019) and ecological data (Greenwood & Wing 1995; **Marwick1994?**) indicate tropical to subtropical conditions up to high latitudes, but early climate models have not been able to model warm polar temperatures and a shallow latitudinal temperature gradient under realistic conditions [(**Sloan1990?**); …]. Although some more recent climate models have been able to align their simulations with the proxy record [(**Huber2013?**);…], the nature of the early Eocene temperature gradient remains a matter of debate. A relatively recent, proxy-based studies inferred a SST gradient that was shallower than the modern by 32+-10 % (**Evans2018?**), which would correspond to a gradient of ~20+-3 degC. Later studies based on new proxy compilations only estimate temperature gradients up to mid-latitudes (Hollis *et al.* 2019), or combine SST proxies with other data (Inglis *et al.* 2020), making it difficult to infer a meaningful estimate of the SST gradient.

Assuming the conservation of thermal tolerances of modern coral reefs and mangrove taxa, we establish broad temperature ranges for EECO coral reef and mangrove sites. We integrate these temperature estimates with the EECO geochemical, shallow marine proxy record to model the latitudinal SST gradient and its uncertainty.

We apply this model to the early Eocene climate optimum,…

“nearest living relative” methods are used for deriving numerical climate estimates from taxonomic plant data for the Cenozoic (**Faucette2007?**; **Chevalier2022?**). These methods work on the the premise that . Proxy-based, deep-time palaeoclimate reconstructions relying solely on geochemical proxies (e.g. Veizer & Prokoph 2015; **Grossman2022?**) have significant spatial gaps, and global temperature reconstructions based on those records alone may be biased (Jones & Eichenseer 2022). Ecological and lithological data are sometimes used for quantitative climate reconstructions [(**Uhl2009?**);], but are rarely used in conjunction with geochemical proxies (**Burgener2023?**).

For example, a plethora of

Here, we combine sea surface temperatures from geochemical proxies with numerical temperature estimated from mangroves and coral reefs, derived with a nearest living relative

apply broadly analogous method to the marine record, deriving probability densities broadly reflecting the thermal preferences of modern coral reefs and mangrove taxa [@…], which we assign to fossil occurrences of coral reefs and mangroves.

, and sometimes biological proxies such as leaf physiognomy, . This can involve interpolation between data points, or extrapolation if entire climate zones remain unsampled. Examples of this approach include t Combining quantitative proxy data with not fully quantified climate constraints, e.g. the presence of tropical weathering products such as bauxite and laterite, has the potential to improve upon existing, purely proxy-based reconstructions (Scotese *et al.* 2021; **Burgener2023?**). Terrestrial and

approach is still rarely employed in deep-time palaeoclimatic studies, as it This requires transforming ordinal or interval-bound data into numerical estimates, which is common

Integrating this type of information into fully quantitative climate reconstructions requires

data provide a different method of reconstructing palaeotemperatures Palaeobiological or geological climate indicators that provide only broad climate constraints which are hard to quantify, and 3) (**Burgener2023?**) For the marine realm, Geochemical proxies such as oxygen isotopes (d18O), tetraether indices (TEX86) and magnesium calcium ratios (Mg/Ca) are used to derive palaeotemperature estimates for individual locations using experimentally derived transfer functions (e.g. Liu *et al.* 2009).

Climate data not based on geochemical, quantitative proxies generally provide broad categorical information, e.g. “tropical” or “dry”, or broad quantitative constraints, e.g. “> 20 degC mean annual temperature”. Integrating this information into fully quantitative climate reconstructions requires transforming it to a numerical scale. This approach is showcased by nearest living relative methods, commonly employed in terrestrial, Cenozoic climate reconstructions, which work on the premise that the climatic requirements of ancient plant taxa were similar to those of there nearest modern relatives (**Kvacek2007?**). A plethora of methods have been developed to derive numerical climate values from taxonomic data [(**Faucette2007?**); @…]. Approaches based on probability density functions have the advantage of naturally capturing the uncertainty associated with the climatic occupancy of modern nearest relatives (**Greenwood2017?**), and probability distributions can be readily integrated in a quantitative modelling framework. Here, we employ a similar approach, constructing probability distributions broadly reflecting the thermal preferences of modern coral reefs and mangrove taxa [@…], which we assign to fossil occurrences of coral reefs and mangroves.

Combining quantitative proxy data with other climate constraints has the potential to improve upon climate reconstructions based on a single type of proxy (**Burgener2023?**), but is rarely employed in deep-time palaeoclimatic studies.

Deep-time palaeoclimate reconstructions use geochemical proxy data

The first challenge requires that all proxy data is converted to a common scale.

A fundamental challenge in reconstructing Earth’s past climate is to infer large-scale climatic patterns from local observations, derived from disparate proxies. In order for the

Proxy-derived palaeoclimate reconstructions offer an important alternative to and a test of Earth system models, and as such are fundamental for reconstructing Earth’s past climate

The geological record holds evidence for episodes in Earth history with fundamentally different climates than our current (Scotese *et al.* 2021). Some episodes, for example the early Eocene, are thought to have been significantly warmer than the modern (Pross *et al.* 2012), potentially representing an analogue for extreme climate warming scenarios (Burke *et al.* 2018). Reconstructions of palaeoclimate commonly use geochemical proxy data, palaeobiological or lithological climate indicators, or earth system modelling.

Geochemical proxies such as oxygen isotopes (d18O), tetraether indices (TEX86) and magnesium calcium ratios (Mg/Ca) are used to derive palaeotemperature estimates for individual locations using experimentally derived transfer functions (e.g. Liu *et al.* 2009). Palaeobiological and lithological climate data provide a different method of reconstructing palaeotemperatures, using e.g. leaf physiognomy and floral assemblages to derive palaeotemperature proxies (Greenwood *et al.* 2003), or using lithological indicators to determine the range of palaeoclimatic belts (Scotese *et al.* 2021). To capture complex dependencies between proxies and climate data, machine learning approaches (Chandra *et al.* 2021) or articifical neural networks (**Malmgren1997?**; **Lauchstedt2017?**) are sometimes used. Earth system models take a fundamentally different approach by modelling the climate-generating processes directly, reconstructing spatially resolved climate estimates for the entire globe [@???]. However, this approach still relies on proxy data to constrain the large number of possible model setups and ground truth the model output [@???] It also requires substantial computing resources and technical expert knowledge, and is thus difficult to implement or modify for most palaeoclimate researchers.

To make inferences on a regional or global scale, proxy-based climate reconstructions need some way of extrapolating a generally sparse record of local findings. Taking the mean of multiple local estimates to summarise regional or global climate variables (e.g. Veizer & Prokoph 2015) is a straightforward approach, but is likely to result in biased climate reconstructions (Jones & Eichenseer 2022). Reconstructions that take into account the spatial distribution of climate data are… ### HERE### provide an alternative that can be employed more readily by a broad range of workers.

Here, we present a Bayesian, hierarchical model for inferring sea surface temperatures (SST) that expands upon existing, spatial reconstructions of palaeoclimate by allowing for the integration of 1) prior information based on physical principles and the observed, modern SST distribution, and of 2) geochemical and ecological climate proxies in a common, quantitative framework. We apply this approach to the early Eocene climatic optimum (EECO), the interval with the warmest sustained temperatures of the Cenozoic. Assuming the conservation of thermal tolerances of modern coral reefs and mangrove taxa, we establish broad temperature ranges for EECO coral reef and mangrove sites. We integrate these temperature estimates with the EECO geochemical, shallow marine proxy record to model the latitudinal SST gradient and its uncertainty.

# Materials & Methods

## Geochemical data

Geochemical climate proxy data were extracted from a latest Paleocene and early Eocene compilation (**hollis2019?**). This compilation provides four different geochemical proxies for reconstructing seawater temperature: *47*, Mg/Ca and TEX\_86\_. For our analyses, this dataset was restricted to the EECO and samples from the continental shelf. Recrystallised $18O samples were also excluded as secondary diagenetic calcite precipitated after deposition can bias isotope measurements and offset temperature values (**schrag1999?**). This resulted in most $18O samples being excluded from the dataset (x out of x). After data filtering, x geochemical proxy samples remained. For a detailed description of each proxy see (**hollis2019?**).

## Ecological data

**Coral reefs.** Today, shallow warm-water coral reefs are limited to tropical and subtropical latitudes (~34°N–32 degS), with minimum sea surface temperature (~18 degC) tolerances being the primary constraint on this distribution (Johannes *et al.* 1983; Kleypas *et al.* 1999; Yamano *et al.* 2001). As coral reefs reside at the upper thermal limit of the oceans today, their maximum sea surface temperature tolerance is less well-constrained with some studies suggesting up to 35.6 degC in the geological past (Jones *et al.* 2022). Nevertheless, the distribution of coral reefs have frequently been recognised as tracers of past (sub-)tropical conditions (Ziegler *et al.* 1984; Kiessling 2001). During the Eocene, coral communities and reefs expanded across tropical and temperate latitudes, with some communities found up to palaeolatitudes of 43 degC N (Zamagni *et al.* 2012). Using a compilation of Paleocene–early Eocene coral reefs and community localities (Zamagni *et al.* 2012), we generate quantitative sea surface temperature estimates for the ECCO. To do so, we extract localities inferred to be Ilerdian (early Eocene) coral reefs from the compilation, and that can be confidently assigned to the EECO. We exclude coral knobs and coral-bearing mounds which might have broader climatic limits than coral reef ecosystems. This filtering resulted in four unique coral reef localities remaining for the EECO, all of which conform to the modern latitudinal range of coral reefs (<34 deg N). Subsequently, we use statistically derived temperature limits (minimum = 21 degC, average = 27.6 degC, maximum = 29.5 degC) from the published literature (Kleypas *et al.* 1999) to define a normal probability distribution of potential temperature values for coral reef localities. This normal probability distribution was defined with a mean of 27.6 and a standard deviation of 2.125, placing 97.5% of the probability density above the minimum. As the distribution of modern corals is skewed towards warmer temperatures, this approach results in 16.5% of the probability being placed on temperatures > 29.5 degC, allowing for the possibility that Eocene coral reefs were adapted to warmer conditions than modern-day coral reefs.

**Mangroves.** Mangroves are distributed throughout the tropics and subtropics today. While factors beside SST influence the distribution of mangroves, empirical, lower temperature limits have been established for the genera *Avicennia* (15.6 degC) and *Rhizophora* (20.7 degC) (Quisthoudt *et al.* 2012). Both *Avicennia* and members of the Rhizophoraceae family were widespread in the early Eocene, but only *Avicennia* occurred at polar latitudes (Suan *et al.* 2017; Popescu *et al.* 2021). Assuming that Eocene members of these mangrove taxa conform to the same climatic requirements as their modern relatives, the presence and absence of *Avicennia* and Rhizophoraceae pollen can be used as a palaeotemperature indicator. For this analysis, published mangrove occurrence data were taken from Popescu *et al.* (2021), and converted to quantitative temperature estimates. From this data, we identify two types of pollen assemblages which we ascribe different temperature distributions:

1. *Avicennia*-only assemblages (): the absence of Rhizophoraceae is indicative of temperatures being between 15.6 degC (lower temperature limt of *Avicennia*) and 20.7 (lower temperature limit of *Rhizophora*). However, a value of 22.5 degC is ascribed as the upper temperature limit here as *Rhizophora* is rare below this temperature. We define the *Avicennia*-only temperature distribution as a normal distribution with a mean of 19.05 degC and a standard deviation of 1.725, resulting in 95% of the probability density being placed within the temperature limits.
2. *Avicennia* and Rhizophoraceae assemblages (): the presence of both groups suggests that the locality should have a minimum temperature of 20.7 degC (lower temperature limit of *Rhizophora*). As the upper thermal limits of *Aviciennia* and *Rhizophora* are not well established in Quisthoudt *et al.* (2012), we assign the same maximum temperature limits (29.5 degC) as coral reef localities as mangroves are also widely distributed throughout tropical regions. Consequently, we define the temperature distribution for this locality as a normal distribution with a mean of 25.1 and a standard deviation of 2.2, with 95% probability density within the temperature limits.

[Fig 1: Map and distribution of proxy data and ecological data]

## Palaeogeographic reconstruction

The palaeogeographic distribution of geochemical and ecological data was reconstructed using the Merdith *et al.* (2021) plate rotation model via the palaeoverse R package ver. XXX (REF). The midpoint age of the EECO (51.2 Ma), along with the present-day coordinates of geochemical and ecological data, were used for palaeorotation.

## Bayesian framework

**Model structure (Fig 2).** We model the mean temperature () at location as a function of absolute latitude () with a logistic regression (growth curve or Richard's curve) of the form:

where and denote the lower and upper asymptote, respectively, specifies the latitude of maximal growth, the most quickly with latitude, denotes the growth rate (Fig. 2), denotes the residual standard deviation, and denotes the number of locations.

We infer from individual temperature observations , derived from geochemical data, at location as

where is the number of observations at each location, and is the estimated standard deviation of the temperatures at location .

Similarly, is inferred for locations with ecological proxies from the associated normal temperature distributions with a given mean and standard deviation, and , as

This structure implies that is not fixed at the mean proxy temperature at location , , but is drawn towards the overall logistic regression curve, i.e. towards . The pull towards tends to be high when is low, when the observations are scattered, i.e.  is high, and/or when the overall standard deviation is low. In practice, this has the desirable consequence that locations with few observations and large temperature differences between observations have less influence on the overall regression than well-sampled locations with consistent temperature observations.

**Priors (Fig2x).** In the Bayesian framework, priors need to be placed on the unknown parameters of a model. We placed weakly informative, conjugate inverse-gamma priors on and :

We set , allowing these priors to be quickly overwhelmed by the data as and increase, as we have little *a priori* knowledge of these parameters.

In contrast, we put informative priors on the regression coefficients , , and , based on physical principles, and vaguely based on the modern climate system:

**A.** Predicted seawater surface temperatures are not allowed to be , the freezing point of sea water. The highest prior density of is placed around , and it slowly tapers off towards higher temperatures. This shape is achieved by placing a skew-normal prior on the lower asymptote, specified as

where , , and are the location, scale and shape parameters.

**K.** Input of solar energy decreases from the tropics to the poles. Hence, the latitudinal temperature gradient is broadly negative, i.e. temperature decreases with absolute latitude. This is achieved by setting . The prior on the upper asymptote is a truncated normal distribution with the mean set to of the modern SST gradient, with a broad standard deviation:

The distribution is truncated to the left at , but not truncated to the right ().

**M.** The steepness of the gradient is presumed to be highest in mid-latitudes; this is expressed with a normal prior on with the mean set to of the modern SST gradient:

**B.** The steepness or growth rate of the gradient is constrained to be and to not be exceedingly high, as oceanic and atmospheric heat transfer is bound to prevent very abrupt SST changes across latitudes. A gamma-distributed prior of the form

was placed on . The shape and rate parameters and were chosen such that the highest prior density is at of the modern SST gradient, .

[Fig. 2: Model structure and visualisation of parameters and priors]

## Model validation

To test whether our logistic regression model can adequately describe latitudinal temperature gradients, we applied a simplified version of the model to modern sea surface temperatures from Bio-Oracle [@…]. As these modern, annual sea surface temperatures estimates are associated with little uncertainty, there is no need for a hierarchical model structure, and we substituted temperature () for in eq. ? and ?. To verify that the modern gradient can be approximated with limited sampling, we resampled modern sea surface temperatures at modern latitudes corresponding the palaeolatitudes of the Eocene samples. This process was repeated 100 times, randomly chosing a longitude for each latitude in each repetition. The parameters of the non-hierarchical temperature model were estimated for each of the 100 samples, and the iterations after burn-in from the posterior of the parameters were pooled to generate the resulting median temperature gradient with 95 % credible intervals ( to percentile).

## Parameter estimation

We estimated the posterior distributions of the model parameters using a Markov chain Monte Carlo (MCMC) algorithm, written in R. Specifically, we sampled the unknown parameters , , and with Metropolis-Hastings, and used Gibbs sampling to estimate all other unknown parameters. Posterior inference on the modern gradient is based on four chains with 60,000 iterations each, 10,000 of which were discarded as burn-in, and keeping every 10th iteration, resulting in a total of 20,000 iterations. The modern temperature gradient with the Eocene sampling distribution was estimated in one chain with 25,000 iterations for each of the 100 temperature samples. 5,000 iterations each were discarded as burn-in, and every 10th iteration was kept, resulting in a total of 200,000 iterations across all 100 model runs. For the Eocene model, we ran four chains with 600,000 iterations each, discarding 100,000 as burn-in and keeping every 100th iteration, as the hierarchical model structure results in higher autocorrelation of the chains. The Eocene posterior inference is thus based on a total of 20,000 iterations with low autocorrelation (effective multivariate sample size for , , and is > 18,000). Trace plots of the MCMC chains indicate convergence and good mixing of the chains (Supplementary Figure S1).

## Processing of model results

Modelled sea surface temperature estimates were generated with eq. ?, calculating the sea surface temperatures at any latitude with the parameter estimates of each iteration from the posterior. The median, and percentile of temperatures where then taken from all temperature estimates obtained at the latitude of interest.

Differences between Eocene and modern temperatures at a certain latitude were calculated by randomly pairing all iterations of the posterior from the Eocene and modern temperature gradient model, calculating the Eocene and modern temperature using the respective iterations, taking the difference, and then calculating the median (, percentile) from all pairs of iterations.

Global average temperatures were calculated by taking the weighted mean of the median (, percentile) temperature estimates in 1 degree latitudinal bins. The weights were set to the proportion of global surface area in each latitudinal bin, i.e. decreasing with increasing latitude as:

where is the upper, and is the lower latitudinal boundary of bin .

# Results

## Model validation

[Fig 3 - Modern gradient] The average modern temperature gradient can be closely approximated with our model when using the full modern SST dataset (Fig. 3a). Reducing the amount and the spatial coverage of data points to match the early Eocene sampling distribution () still results in a close match of the modeled gradient with the empirical gradient (Fig. 3b). The 95% credible interval (CI) of reconstructed gradients (Fig. 3b) is noticeably narrower than the spread of empirical temperature values (Fig. 3a). The modern, latitudinal gradient, modeled with all data, spans 29.6 degC from the equator to the poles. When modeled with only the early Eocene sampling locations, the median temeprature gradient is 29.3 (95% CI: 27.0 - 31.6) degC. A random sampling of modern temperatures at latitudes corresponding to the sampled Eocene palaeolatitudes would thus allow the reconstruction of the modern latitudinal temperature gradient with good accuracy, with expected deviations from the actual gradient as low as -2.6 to 2.1 degC (95 % CI).

Modeling the latitudinal temperature gradient facilitates accurate reconstruction of global average temperatures. The modern, global mean sea surface temperature estimate is 17.6 degC when using the full, modern data set. When reduced to the early Eocene sampling distribution, the modern global mean sea surface temperature is estimated at 17.8 (16 - 19.7) degC. The deviations from the full modern estimate are again very low, with a 95 % credible interval of -0.8 to 1.3 degC.

## EECO reconstruction

The modeled Eocene temperature gradient is starkly different from the modern (Fig 4). Modelled, median equatorial temperatures are 4.2 (0.2 - 8.3) degC higher for the Eocene, and polar temperatures are higher by 25.0 (17.0 - 29.1) degC. This results in a flattened latitudinal temperature gradient of 9.0 ( 2.5 - 17.8) degC for the Eocene, as opposed to 29.6 degC for the modern.

The high variability of early Eocene palaeotemperature proxies, particularly in the mid-latitudes, and the scarcity of high-latitude data, result in substantial uncertainties in the modeled temperature gradient. This is reflected in the residual standard deviation () of the early Eocene gradient, 4.9 (3.8 - 6.5) degC, more than twice as high than of the modern gradient, 2.2. This signifies that the early Eocene data fit less well to the logistic, latitudinal gradient model, which can also be seen from the drastic departure of some of the proxy data from the gradient estimates (Fig. 4).

The global early Eocene mean sea surface temperature is estimated at 28.8 (25.7 - 31.7) degC, significantly higher than the modern (17.6 degC).

A model run excluding the ecological proxies results in a similar median latitudinal gradient, 9.2 degC, but with a 20% wider 95% CI (Supplementary materials). This indicates that the ecological proxy data are broadly in agreement with the geochemical proxies, while providing additional constraints on the shape of the early Eocene temperature gradient.

Due to the limited spatial coverage of the early Eocene proxy record, and due to the added model complexity of simultaneously estimating a model across both hemispheres, we pooled the proxy data across both hemispheres. Applying the model separately to each hemispheres results in substantial differences in hemispherical, average temperatures, with the southern hemisphere being warmer by 6.5 (3.5 - 9.4) degC. The inferred latitudinal gradient is somewhat steeper ( 4.8 degC, although the 95% CI spans -6.6 to 14.3 degC) in the northern hemisphere, but the large uncertainties associated with both gradients, and the lack of polar proxy data in the southern hemisphere preclude a more precise statement (Fig. S3).)

## Fig 4 - compare Eocene gradient with modern gradient, also with previous reconstructions / ES models?

* x1 Draw samples from the posterior gradient to get a latitudinal gradient of deg C / deg lat with uncertainty (intercept gradient at 0 and 90 deg lat)
* x2 generate global average temperature, and tropical, temperate, polar temperatures in a similar way, accounting for area

# Discussion

**Eocene temperature gradients.**

**Methodological advances.** Latitudinal temperature gradients have been estimated with a variety of parametric and non-parametric statistical techniques. Parametric methods include polynomials (Greenwood & Wing 1995; Bijl *et al.* 2009) and a cosine function (Inglis *et al.* 2020). Whilst some of those functions are flexible enough to model a range of empirical temperature gradients, the logistic function used herein has the added advantage of an easily understandable relationship between each of the model parameters and the shape of the latitudinal gradient (Fig. 2). Non-parametric models can very flexibly follow the shape of an unknown temperature gradient Zhang *et al.* (2019), but have the disadvantage of having no model parameters that can be compared between data sets. Parametric methods should therefore be preferred to non-parametric temperature gradient models, as long as very broad prior information on the shape of the gradients exists. The ease with which such prior information can be integrated is perhaps the major advantage of our method: The Bayesian framework allows to specify arbitrary priors on the four parameters that control the upper and lower temperature limits, the steepness of the gradient, and the latitudinal position of the steepest point in the gradient (see Methods). A further improvement to existing methods is that the hierarchical structure of our model explicitly accounts for variation of temperature estimates at individual localities, and for uncertainty associated with the ecological temperature proxies. An extension of the model additionally accounts for measurement uncertainty of the geochemical proxies (Supplementary Materials).

-GP regression and interpolation can give rise to inverse gradients ect. –> our method nicely prevents this.

**Comparison to Eocene proxy and modelling studies**.

A major theme of the discussions on the early Eocene greenhouse climate has been the discrepancy between proxy records, particularly ecological climate indicators such as macrofloral assemblage data, have

* x1 compare lat gradient to literature estimates, e.g. Zhang2019
* compare lat gradient to climate models and the “early EECO problem”
* discuss insufficiency of proxies (high variability, …)
* discuss utility of the model and future applications

# Conclusions

# Acknowledgements

# Author contributions

# Data accessibility

# References

BIJL, P. K., SCHOUTEN, S., SLUIJS, A., REICHART, G.-J., ZACHOS, J. C. and BRINKHUIS, H. 2009. [Early Palaeogene temperature evolution of the southwest Pacific Ocean](https://doi.org/10.1038/nature08399). *Nature*, **461**, 776–779.

BURKE, K. D., WILLIAMS, J. W., CHANDLER, M. A., HAYWOOD, A. M., LUNT, D. J. and OTTO-BLIESNER, B. L. 2018. [Pliocene and Eocene provide best analogs for near-future climates](https://doi.org/10.1073/pnas.1809600115). *Proceedings of the National Academy of Sciences*, **115**, 13288–13293.

CHANDRA, R., CRIPPS, S., BUTTERWORTH, N. and MULLER, R. D. 2021. [Precipitation reconstruction from climate-sensitive lithologies using Bayesian machine learning](https://doi.org/10.1016/j.envsoft.2021.105002). *Environmental Modelling & Software*, **139**, 105002.

GREENWOOD, D. R. and WING, S. L. 1995. [Eocene continental climates and latitudinal temperature gradients](https://doi.org/10.1130/0091-7613(1995)023<1044:ECCALT>2.3.CO;2). *Geology*, **23**, 1044.

———, MOSS, P. T., ROWETT, A. I., VADALA, A. J. and KEEFE, R. L. 2003. [Plant communities and climate change in southeastern Australia during the early Paleogene](https://doi.org/10.1130/0-8137-2369-8.365). *In* *Causes and Consequences of Globally Warm Climates in the Early Paleogene*, Geological Society of America.

HOLLIS, C. J., DUNKLEY JONES, T., ANAGNOSTOU, E., BIJL, P. K., CRAMWINCKEL, M. J., CUI, Y., DICKENS, G. R., EDGAR, K. M., ELEY, Y., EVANS, D. and OTHERS. 2019. The DeepMIP contribution to PMIP4: Methodologies for selection, compilation and analysis of latest paleocene and early eocene climate proxy data, incorporating version 0.1 of the DeepMIP database. *Geoscientific Model Development*, **12**, 3149–3206.

INGLIS, G. N., BRAGG, F., BURLS, N. J., CRAMWINCKEL, M. J., EVANS, D., FOSTER, G. L., HUBER, M., LUNT, D. J., SILER, N., STEINIG, S., TIERNEY, J. E., WILKINSON, R., ANAGNOSTOU, E., DE BOER, A. M., DUNKLEY JONES, T., EDGAR, K. M., HOLLIS, C. J., HUTCHINSON, D. K. and PANCOST, R. D. 2020. [Global mean surface temperature and climate sensitivity of the early Eocene Climatic Optimum (EECO), Paleocene (PETM), and latest Paleocene](https://doi.org/10.5194/cp-16-1953-2020). *Climate of the Past*, **16**, 1953–1968.

JOHANNES, R., WIEBE, W., CROSSLAND, C., RIMMER, D. and SMITH, S. 1983. Latitudinal limits of coral reef growth. *Marine ecology progress series. Oldendorf*, **11**, 105–111.

JONES, L. A. and EICHENSEER, K. 2022. [Uneven spatial sampling distorts reconstructions of Phanerozoic seawater temperature](https://doi.org/10.1130/G49132.1). *Geology*, **50**, 238–242.

JONES, L. A., MANNION, P. D., FARNSWORTH, A., BRAGG, F. and LUNT, D. J. 2022. Climatic and tectonic drivers shaped the tropical distribution of coral reefs. *Nature communications*, **13**, 1–10.

KIESSLING, W. 2001. Paleoclimatic significance of phanerozoic reefs. *Geology*, **29**, 751–754.

KLEYPAS, J. A., MCMANUS, J. W. and MEÑEZ, L. A. 1999. Environmental limits to coral reef development: Where do we draw the line? *American zoologist*, **39**, 146–159.

LIU, Z., PAGANI, M., ZINNIKER, D., DECONTO, R., HUBER, M., BRINKHUIS, H., SHAH, S. R., LECKIE, R. M. and PEARSON, A. 2009. [Global Cooling During the Eocene-Oligocene Climate Transition](https://doi.org/10.1126/science.1166368). *Science*, **323**, 1187–1190.

MERDITH, A. S., WILLIAMS, S. E., COLLINS, A. S., TETLEY, M. G., MULDER, J. A., BLADES, M. L., YOUNG, A., ARMISTEAD, S. E., CANNON, J., ZAHIROVIC, S. and OTHERS. 2021. Extending full-plate tectonic models into deep time: Linking the neoproterozoic and the phanerozoic. *Earth-Science Reviews*, **214**, 103477.

POPESCU, S.-M., SUC, J.-P., FAUQUETTE, S., BESSEDIK, M., JIMÉNEZ-MORENO, G., ROBIN, C. and LABROUSSE, L. 2021. [Mangrove distribution and diversity during three Cenozoic thermal maxima in the Northern Hemisphere (pollen records from the Arctic regions)](https://doi.org/10.1111/jbi.14238). *Journal of Biogeography*, **48**, 2771–2784.

PROSS, J., CONTRERAS, L., BIJL, P. K., GREENWOOD, D. R., BOHATY, S. M., SCHOUTEN, S., BENDLE, J. A., RÖHL, U., TAUXE, L., RAINE, J. I., HUCK, C. E., VAN DE FLIERDT, T., JAMIESON, S. S. R., STICKLEY, C. E., VAN DE SCHOOTBRUGGE, B., ESCUTIA, C. and BRINKHUIS, H. 2012. [Persistent near-tropical warmth on the Antarctic continent during the early Eocene epoch](https://doi.org/10.1038/nature11300). *Nature*, **488**, 73–77.

QUISTHOUDT, K., SCHMITZ, N., RANDIN, C. F., DAHDOUH-GUEBAS, F., ROBERT, E. M. R. and KOEDAM, N. 2012. [Temperature variation among mangrove latitudinal range limits worldwide](https://doi.org/10.1007/s00468-012-0760-1). *Trees*, **26**, 1919–1931.

REYNOLDS, R. W. and SMITH, T. M. 1994. Improved global sea surface temperature analyses using optimum interpolation. *Journal of climate*, **7**, 929–948.

SCOTESE, C. R., SONG, H., MILLS, B. J. W. and VAN DER MEER, D. G. 2021. [Phanerozoic paleotemperatures: The earth’s changing climate during the last 540 million years](https://doi.org/10.1016/j.earscirev.2021.103503). *Earth-Science Reviews*, **215**, 103503.

SUAN, G., POPESCU, S.-M., SUC, J.-P., SCHNYDER, J., FAUQUETTE, S., BAUDIN, F., YOON, D., PIEPJOHN, K., SOBOLEV, N. N. and LABROUSSE, L. 2017. [Subtropical climate conditions and mangrove growth in Arctic Siberia during the early Eocene](https://doi.org/10.1130/G38547.1). *Geology*, **45**, 539–542.

TAYLOR, S. P., HAYWOOD, A. M., VALDES, P. J. and SELLWOOD, B. W. 2004. [An evaluation of two spatial interpolation techniques in global sea-surface temperature reconstructions: Last Glacial Maximum and Pliocene case studies](https://doi.org/10.1016/j.quascirev.2003.12.003). *Quaternary Science Reviews*, **23**, 1041–1051.

VEIZER, J. and PROKOPH, A. 2015. [Temperatures and oxygen isotopic composition of Phanerozoic oceans](https://doi.org/10.1016/j.earscirev.2015.03.008). *Earth-Science Reviews*, **146**, 92–104.

VICKERS, M. L., BERNASCONI, S. M., ULLMANN, C. V., LODE, S., LOOSER, N., MORALES, L. G., PRICE, G. D., WILBY, P. R., HOUGÅRD, I. W., HESSELBO, S. P. and OTHERS. 2021. Marine temperatures underestimated for past greenhouse climate. *Scientific reports*, **11**, 1–9.

YAMANO, H., HORI, K., YAMAUCHI, M., YAMAGAWA, O. and OHMURA, A. 2001. Highest-latitude coral reef at iki island, japan. *Coral Reefs*, **20**, 9–12.

ZAMAGNI, J., MUTTI, M. and KOŠIR, A. 2012. The evolution of mid paleocene-early eocene coral communities: How to survive during rapid global warming. *Palaeogeography, palaeoclimatology, palaeoecology*, **317**, 48–65.

ZHANG, L., HAY, W. W., WANG, C. and GU, X. 2019. [The evolution of latitudinal temperature gradients from the latest Cretaceous through the Present](https://doi.org/10.1016/j.earscirev.2019.01.025). *Earth-Science Reviews*, **189**, 147–158.

ZIEGLER, A., HULVER, M., LOTTES, A. and SCHMACHTENBERG, W. 1984. Uniformitarianism and palaeoclimates: Inferences from the distribution of carbonate rocks. *Geological journal. Special issue*, 3–25.

1. Department of Earth Sciences, Durham University, South Road, DH1 3LE, Durham, United Kingdom [↑](#footnote-ref-20)
2. Centro de Investigación Mariña, Grupo de Ecoloxía Animal, Universidade de Vigo, 36310 Vigo, Spain [↑](#footnote-ref-21)