The Early Eocene equable climate problem: can perturbations of climate model parameters identify possible solutions?

Navjit Sagoo, Paul Valdes, Rachel Flecker and Lauren J. Gregoire†

School of Geographical Sciences, University of Bristol, University Road, Bristol BS8 1SS, UK

Geological data for the Early Eocene (56–47.8 Ma) indicate extensive global warming, with very warm temperatures at both poles. However, despite numerous attempts to simulate this warmth, there are remarkable data–model differences in the prediction of these polar surface temperatures, resulting in the so-called ‘equable climate problem’. In this paper, for the first time an ensemble with a perturbed climate-sensitive model parameters approach has been applied to modelling the Early Eocene climate. We performed more than 100 simulations with perturbed physics parameters, and identified two simulations that have an optimal fit with the proxy data. We have simulated the warmth of the Early Eocene at 560 ppmv CO$_2$, which is a much lower CO$_2$ level than many other models. We investigate the changes in atmospheric circulation, cloud properties and ocean circulation that are common to these simulations and how they differ from the remaining simulations in order to understand what mechanisms contribute to the polar warming. The parameter set from one of the optimal Early Eocene simulations also produces a favourable fit for the last glacial maximum boundary climate and outperforms the control parameter set for the present day. Although this does not ‘prove’ that this model is correct, it is very encouraging that there is a parameter set that creates a climate model able to simulate well very different palaeoclimates and the present-day climate. Interestingly, to achieve the great warmth of the Early Eocene this version of the model does not have a
strong future climate change Charney climate sensitivity. It produces a Charney climate sensitivity of 2.7°C, whereas the mean value of the 18 models in the IPCC Fourth Assessment Report (AR4) is 3.26°C ± 0.69°C. Thus, this value is within the range and below the mean of the models included in the AR4.

1. Introduction

The term equable climate has been used to describe the global extent of warmth in past climates, which have a reduced equator-to-pole temperature difference (EPTD), warm polar regions with a reduced seasonality and ice free conditions at both poles [1,2]. The extent of this warming is supported by a wide range of data. Recent syntheses of terrestrial [3] and marine [4] proxy climate data for the Early Eocene suggest that the polar temperatures were 15°C or more but that the tropics were only slightly warmer than modern. Moreover, palaeobotanical data also suggest that the high latitudes were above freezing throughout the year [5], which is a major change over present conditions despite the fact that the continents are not that different from the modern.

The Early Eocene equable climate problem relates to differences between climate model simulations and proxy reconstructions of the Early Eocene and the climate inferred from climate proxies. The modern generation of climate models has managed to capture much of this warmth from the proxy data in the low and mid-latitudes by forcing the climate with very high concentrations of CO₂, 16 times pre-industrial concentrations of CO₂ [3,6], but simulating above freezing temperatures at the poles all year round is difficult. The assumption of a strong seasonal bias in the proxy data must currently be assumed in order to reconcile proxy polar temperatures with climate model output [4].

Estimates of Early Eocene temperatures include annual sea surface temperatures (SSTs) of up to 27°C [7] and terrestrial mean annual temperatures (MATs) of up to 18°C [8] at palaeolatitudes greater than 80°N. In the Southern Hemisphere (SH), SSTs between 17°C and 32°C [9–11] have been reconstructed at palaeolatitudes greater than 60°S, while terrestrial MATs between 12°C and 18.8°C have been reconstructed at similar latitudes [12–14]. These high latitude temperatures are likely to have been sufficient to prevent any significant permanent ice cover. While there is reasonable data coverage for the mid- and high latitudes, data from the low latitudes are scarce. Tropical SST data are available from the Tanzania drilling project, which indicate that SSTs at a palaeolatitude of 18°S were approximately 33°C [15]. One of the features inferred from this distribution of temperatures is that the temperature difference between the pole and the equator was much reduced compared to the modern day. There is also evidence of an enhanced hydrological cycle in the high latitudes during the Early Eocene [16–18]. Water vapour has an impact on the radiation balance of the planet through the water vapour greenhouse effect, the cloud greenhouse effect and via reflection of shortwave radiation from clouds and ice [19]. Understanding what role an intensified hydrological cycle may play in developing and maintaining an equable climate is therefore also of interest.

The first paper on the Early Eocene equable climate problem was published over 30 years ago [20] and substantial modelling efforts have been undertaken since then in order to simulate the Early Eocene climate. Many advances in model development have also been made, and whereas the earliest Eocene models were limited to either energy balance models (EBMs) or early general circulation models (GCMs) with fixed seasons, the current generation of models simulate the dynamics of the atmospheres and oceans and in some cases vegetation, and are of higher resolution and have improved and revised physics. This has consequently improved our ability to simulate the Eocene climate. Meanwhile, advances in existing proxy methods, the development of new methods and the acquisition of additional proxy data have led to the warmer, revised temperatures for the tropical marine realm [21] and terrestrial realms at all latitudes [22–25]. Thus, the equable climate problem is still apparent in the proxy datasets. While modelling studies have improved in their simulation of the Early Eocene, the processes that contribute to the
amplification of polar temperatures during the Early Eocene are difficult to accurately model (e.g. clouds) and are not well understood. However, the model–data discrepancy persists in the high latitudes. The aim of this paper is to understand whether perturbing uncertain climate model parameters can offer insight into the climate processes involved in developing and maintaining the equable Early Eocene climate.

(a) Previous Eocene modelling studies

Many of the earliest model experiments of the Early Eocene, run with increased CO$_2$ concentrations compared to the modern, simulated high latitudes and continental interiors that were warmer than the modern, but not warm enough compared to the proxy climate evidence of the Early Eocene. Sloan & Barron [26,27] simulated high latitudes and continental interiors that are warmer than the modern, but still cooler than the proxy climate evidence. This model–data mismatch generated a range of possible explanations including missing components and processes in the models such as polar stratospheric clouds (PSCs) [27,28] and tropical cyclones [29–31] to approximations in the boundary conditions associated with coarse model resolution, for example, the presence of large lakes (e.g. [32,33]) altered orbits (e.g. [34]) and the role of heat transport [35,36].

A recurring theme in Early Eocene modelling studies is the contribution of clouds in equable climates. Sloan et al. [37], Kirk-Davidoff et al. [28] and Kirk-Davidoff & Lamarque [38] all investigated the role of PSCs in equable climates in response to elevated concentrations of CO$_2$ and CH$_4$. Sloan et al. [37] included idealized prescribed PSCs in the GCM GENESIS v. 2, which resulted in up to 20$^\circ$C warming in oceanic regions where sea-ice was reduced. This warming was still insufficient to account for warming seen in the proxy data available at the time, but compared to more recent proxy data these simulations were approximately 10$^\circ$C too cool at latitudes of around 60$^\circ$. Kirk-Davidoff et al. [28] and Kirk-Davidoff & Lamarque [38] investigated the mechanisms that led to the formation of PSCs and the response of the climate. PSCs were found to warm in response to higher CO$_2$ via changes in stratospheric circulation and water content, but the large radiative effects required to warm the polar regions were found to be related to ice crystal number density in the PSCs, and a lack of theoretical knowledge may have prevented these hypotheses from being developed further.

Abbot & Tziperman [39] identified a high latitude cloud radiative forcing feedback using a simple column model. They found that increased extra tropical surface temperatures led to the initiation of strong atmospheric convection and the convective clouds led to additional warming of the high latitudes. The radiative effect of the resulting convective clouds reduced the EPTD by 8–10$^\circ$C. Further work using this column model investigated the constraints atmospheric and oceanic heat transport and CO$_2$ concentration had on the convective cloud feedback [40]. This feedback was found to be present in modern model simulations forced with CO$_2 = 2240$ ppmv, and for the Eocene with CO$_2 = 560$ ppmv.

An alternative solution to the equable climate problem was suggested by Kump & Pollard [41] and Kiehl et al. [42]. Cloud condensation nuclei (CCN) play an important role in cloud properties such as cloud water content, cloud opacity and cloud lifetime. In the past, the distribution of CCN was probably different from today because the distribution and composition of atmospheric aerosols were different [43]. Based on this, Kump & Pollard [41] increased CCN radii in a Cretaceous climate simulation using the GENESIS (v. 3.0) GCM. This resulted in a decrease in cloud amount and cloud albedo leading to a dramatic warming, both globally and at the poles, and a decrease in the EPTD.

Lunt et al. [4] have recently published a review on Eocene modelling termed EoMIP (the Eocene modelling intercomparison project) in which they compare five recent modelling studies for the Early Eocene. The modelling studies have all been run with different objectives; different boundary conditions and multiple values of CO$_2$ have been used in some studies. The models used were: HadCM3L, the sister model of the FAMOUS model, which is used in this study [4];
ECHAM5/MPI-OM [44]; the GISS model [45]; and the two versions of the model CCSM—CCSM_H [3,46] that has no aerosol load following the approach of Andreae [43] and Kump & Pollard [41,43] and CCSM_W [6,47] that has a modern aerosol load. At a given CO₂ concentration CCSM_H and CCSM_W give different global means; for instance, there is a 3°C difference in mean surface air temperature (SAT) between CCSM_H and CCSM_W at 16 times pre-industrial CO₂ concentrations, the level at which the best match to Eocene proxy data was found for that model. The range of CO₂ concentrations resulting in the best Eocene simulation between the models varied between 2 times and 16 times pre-industrial concentrations, demonstrating the need for better constraints on actual CO₂ concentration during the Early Eocene.

A comprehensive comparison of model results with recent syntheses of proxy data was made [3,4] as part of the EoMIP and a one-dimensional EBM [44] was used to investigate, identify and understand the inter-model variability. The ACEX data points from the Arctic Ocean [7] indicate SSTs of approximately 13°C for the Ypresian (56.0–47.8 Ma) and an SST of approximately 22°C recorded during the Early Eocene Climatic Optimum (EECO, 53.1–49 Ma). Few of the models manage to simulate these temperatures. In the SH, SSTs greater than 25°C are measured from both EECO and Ypresian material from ODP 1172D in the Pacific Ocean [11] and Waipara River of the coast from New Zealand [48]. CCSM_H is the only model that managed to intersect the lower error bars of these temperature estimates.

In summary, there is considerable inter-model variability between the models in the EoMIP. The variability is considerably larger than present day (PD) inter-model differences, with very different CO₂ concentrations giving the best fit to data. The differences between models have been attributed to a combination of greenhouse effect and surface albedo feedbacks rather than differences in cloud feedbacks or heat transport [4]. Differences in the climates of CCSM_H and CCSM_W are related to differences in the assumed aerosol loads used. Despite the variation in boundary conditions between these five models, only CCSM_H manages to simulate temperatures within the lower boundaries of the estimates at the warmest high latitude locations (i.e. ACEX, ODP 1172D and Waipara River), thus demonstrating the need for alternative solutions to the equable climate problems.

(b) Parametrizations and ensembles

The EoMIP study highlights the inter-model variability between the models studied. Climate models are constructed by discretizing and then solving equations that represent the basic laws that govern the behaviour of the atmosphere, ocean and land surface [49] and many approximations are required in order to solve the nonlinear system of partial differential equations. Note that the solution of a partial differential equation depends on (i) the initial conditions, (ii) forcing boundary conditions (focus of the previous palaeoclimate studies) and (iii) approximations in the form of climate parametrizations (this study).

Parameter uncertainty stems from the fact that small-scale processes in all components of the climate system cannot be resolved explicitly in the climate system. This is the case in cloud processes for example [50,51]. Parametrization of sub-grid scale processes is a major source of uncertainty in climate prediction [52], and while in some parametrizations the processes, observational evidence or theoretical knowledge is well understood, where this information is scarce the values chosen for a parametrization may simply be because they appear to work [51]. Future climate change studies have recently focused on quantifying the uncertainty arising from these parameters using Monte Carlo-type techniques [53]. This type of work is referred to as perturbed physics ensembles (PPEs) because suites of simulations are generated by perturbing climate-sensitive model parameters. The resulting spread in predictions is quantified, leading to model-dependent probabilistic estimates of the distribution of future climate, warming and climate sensitivity. In a few cases, the ensembles are very large (i.e. a thousand member ensemble) [53,54] but in most cases the number of simulations is limited by the computational cost of complex climate models to a few tens or a hundred simulations as is the case in [50,55].
Ensembles with perturbed climate-sensitive model parameters have begun to be used in palaeoclimate research, primarily for the Late Quaternary and particularly on the issue of climate sensitivity and El Niño Southern Oscillation (e.g. [56–60]). Ensembles with perturbed climate-sensitive model parameters have also been used to ‘tune’ the climate model to proxy data for the last glacial maximum (LGM) [61]. However, few studies have investigated older time periods apart from a small set of simulations for the Pliocene [62].

In practice, there are several hundreds of parameters that are poorly constrained in climate models and it is impossible to vary all of them. Gregoire et al. [61] identified a total of 10 parameters to be varied in FAMOUS, of which six parameters had been tuned in a previous study [63] and recognized as having a high impact on the climate of HadCM3 [50]. The study by Murphy et al. [50] identified key parameters that had a major impact on Charney climate sensitivity (the global average temperature increase associated with a doubling in CO₂ and including a specific set of feedbacks).

This paper investigates the effect of parametric uncertainty on the Early Eocene equable climate problem using the model FAMOUS. The motivation of this study is to attempt to detect ensemble simulations that match the proxy data available for the Early Eocene and to understand how processes in these simulations vary from the rest of the ensemble. We deliberately do not limit the parameter set perturbations to only those sets that perform well for modern conditions because we wish to explore if any combination of parameters are able to simulate the Early Eocene equable climates.

2. Material and methods

(a) Model description

FAMOUS (Fast Met Office/UK Universities Simulator) is an atmosphere and ocean GCM (AOGCM) that is based on HadCM3 (Hadley Centre Model v. 3) [64]. While its parametrizations of physical and dynamical processes are almost identical to those of HadCM3, FAMOUS has a reduced resolution in both the atmosphere and ocean, and a longer time-step that reduces the computational resources required to run FAMOUS to 10% of that required by HadCM3 [65]. This favours the use of FAMOUS in experiments where large amounts of computational resources are required.

The atmosphere component of FAMOUS is based on the Hadley Centre atmosphere model (HadAM3) (see [66] for full details). The atmosphere resolution in FAMOUS is 7.5° longitude × 5° latitude grid, with 11 levels in the vertical. The ocean model in FAMOUS is the Hadley Centre ocean model (HadOM3) (see [64] for full details), which is a rigid lid model. The ocean resolution in FAMOUS is 3.75° longitude × 2.5° latitude grid with 20 levels and a 12 h time step (using a distorted momentum equation), which is the same resolution as the model HadOM3L, and is lower than the resolution of HadOM3 (1.25° longitude × 1.25° latitude grid). Since the resolution of the ocean model is greater than the atmosphere, FAMOUS uses a coastal tiling scheme that combines the properties of land and sea in coastal grid boxes. The ocean model can then use the more detailed coastline allowed by its higher resolution grid while conserving coupled quantities [65]. FAMOUS does not use flux adjustments. Land processes are modelled with the UK Meteorological Office’s land surface scheme, MOSES 1 [67]. Smith et al. [65] give a detailed description of FAMOUS and highlight the major differences between FAMOUS and HadCM3. The version of FAMOUS used in this work is identical to that of Gregoire et al. [61] and slightly differs from Smith et al. [65] as described in Gregoire et al. [61].

The resolution of FAMOUS is not as high as the models used to investigate future climate change; horizontal resolution of the order of 1° to 2° is now commonly used in the ocean component of most climate models [68]. However, FAMOUS fills an important niche in the current generation of models sitting between the higher resolution AOGCMs and the lower resolution, highly parametrized Earth system models of intermediate complexity. The reduced resolution allows us to fully spin-up the ocean, with all of the simulations presented in this work extended...
to at least 8000 years. This would be impossible with the higher resolution models but is essential since the time scale for ocean equilibration is measured in thousands of years.

(b) Present-day simulation

In the original tuning of FAMOUS, Jones et al. [63] systematically tuned the model to reproduce both the equilibrium climate and climate sensitivity of HadCM3. Smith et al. [65] then undertook manual tuning to reduce a cold bias in the northern high latitudes, which led to the removal of Iceland. Gregoire et al. [61] conducted ensembles with perturbed climate-sensitive model parameters for the PD and LGM climates. Building on this work, we use the PD control parameter values in Gregoire et al.’s [61] configuration as our control PD simulation.

The PD version of FAMOUS uses the following concentrations of greenhouse gases: CO₂, −280 ppmv; CH₄, 760 ppmv; N₂O, 270 ppmv. The orography is derived from the US Navy 10 min resolution dataset, with some small additional smoothing at latitudes poleward of 60° (see [65] for full details). The ocean resolution of FAMOUS does not allow for flow between the Atlantic and the Mediterranean. Instead a simple mixing has been parametrized for this region in an area that extends from the surface to a depth of 1300 m. An artificial island is used at the North Pole to avoid the problem of converging meridians [65].

(c) Early Eocene model configuration and uncertainties

(i) Palaeogeography and orography

Palaeogeographic reconstruction is a critical boundary condition in palaeoclimate modelling, and reconstructing continental interiors, dimensions of palaeo-orography, palaeo-shorelines of ancient lakes and the widths of epicontinental seaways is challenging as the geological evidence left by these features can be minimal [69–72]. Modelling experiments have been used to explore the impact on climate for some of these poorly constrained variables. For example, experiments have investigated the impact of the inclusion of a large lake in western North America [32], the opening and closing of the Arctic seaways during the Early Palaeogene [45] and the impact of uncertain orography [27,73–76] on Eocene climate. Results suggest that uncertain palaeogeography tends to increase regional uncertainty in modelled climate, with some potential for climatic tele-connections and modification of global climate.

The Early Eocene simulations presented here use a palaeogeography created using similar methods to Markwick & Valdes [72]. The palaeogeography is similar to the HadCM3 Early Eocene simulations conducted by Tindall et al. [77] but at the resolution of FAMOUS. There is no flow between the global oceans and the Arctic Ocean in these simulations although opening these gateways could impact climate [45] and FAMOUS does not explicitly represent lakes.

(ii) Greenhouse gases and orbit

Early Eocene atmospheric CO₂ concentration is an important but uncertain boundary condition. Proxy measurements indicate that CO₂ in the Early Eocene was higher than present and estimates range from as low as 300 ppmv to more than 4400 ppmv [7,15,78–82]. Climate modelling studies of the Early Eocene have used different CO₂ values that span the entire proxy range. For these Early Eocene simulations, CO₂ was set at 560 ppmv (twice pre-industrial concentrations). While this is at the lower end of the range of predicted CO₂ values for the Early Eocene, it has been used because Early Eocene sensitivity simulations (P. J. Valdes 2010, unpublished data) showed that the Eocene configuration of FAMOUS is relatively sensitive to CO₂. All other greenhouse gases (CH₄, N₂O) were set to pre-industrial values. Indirect evidence indicates that during the Early Cenozoic methane concentrations of these other greenhouse gases could have been much higher due to the expansion of peat lands and the consequent increase in methanogenesis for instance [83–86]. However, in the absence of suitable proxy data to quantify this increase we use PD values.
Orbital changes have been calculated for the past 250 Myr [87] and studies have identified a strong eccentricity and precession signal from Early Eocene sediments [85–87]. We attempt to simulate a very long multi-million year interval in which many orbital configurations would have occurred. While the role of orbital forcing may be a driver for short-term hyperthermal events [84,88], we are interested in simulating the overall warmth of this period and thus have used a modern orbital configuration. Modelling studies that have investigated the impact of orbital forcing on the Early Eocene climate have improved the model–data fit if specific orbits are chosen. Sloan & Morrill [34] showed that extreme orbital values from the calculated Pleistocene range could reduce temperatures in the Northern Hemisphere (NH) continental interiors compared to the orbital configuration for the PD. Sloan & Huber [33] showed that between precessional end members for an Eocene greenhouse world widespread regional variation occurred, including: SST variation of 5°C in the high northern latitudes; up to a twofold variation in upwelling strength in tropical regions; and changes in net surface moisture balance (precipitation–evaporation) of up to 3 mm d−1 in the tropics. Uncertainty in orbital forcing has a limited impact on global mean climate values and a larger impact on regional and seasonal climate, in particular at high latitudes. In the studies referenced here [33,34], uncertainty was more pronounced in high latitude terrestrial realms and in the low latitude marine realm.

(iii) Vegetation

There is very little data available for vegetation reconstruction of past climates and the data that do exist may not be fully representative of the diversity of the area they come from. Numerous modelling studies have investigated the impact of vegetation on palaeoclimate [89–92] and several studies have looked specifically at Early Eocene modelled vegetation [73,93,94]. While the impact on global climate has been noted to be small, changes to regional climate can be distinct [73,93,94].

Vegetation in model simulations can either remain static and unchanging through time or dynamic and responding to the changing climatic conditions. Both approaches have advantages and disadvantages, as reviewed in Peng [95]; for example, dynamic vegetation may increase precipitation and reduce temperature extremes [96]. The work presented here used a static and uniform vegetation configuration of shrub-like plants everywhere as we consider the effect of vegetation feedbacks to be secondary compared to the parameter perturbations. Future work will examine the impact of vegetation change.

(d) Perturbed parameter ensemble

Table 1 gives a description of each parameter perturbed in this work. We perturb 10 parameters within their upper and lower bounds. The uncertainty bounds were based on previous studies [50]. The uncertainty arises because of the large spatial and temporal variation of many of these processes.

We have run two sets of perturbed physics simulations. In the first set all 10 groups of uncertain parameters are perturbed simultaneously and at 10 equal intervals between the lower and upper boundaries of their uncertain range; we refer to these simulations as the multiple parameter perturbations (MPPs). In order to facilitate the best use of computing time and the greatest coverage of different parameter sets a statistical method of Latin hypercube sampling (LHS) is used to define the parameter values for the MPP simulations [97]. Using LHS with 10 parameters requires of the order of one hundred simulations to obtain a reasonable coverage of the parameter space [98]. We therefore generated one hundred unique parameter sets, maximizing the parameter space that is sampled for a finite number of simulations in a statistically robust way. Full details of the LHS methodology are available in Gregoire et al. [61] who originally ran PD simulations with the same MPP sets. The MPP simulations were initially set up to run for 6000 years, though runs of particular interest were integrated for 10 000 years. This length of the runs is required in order to achieve full equilibrium in both the surface and deep ocean in the Early Eocene climate.
Table 1. Name and description of the 10 parameters or groups of parameters that are perturbed in this study. The minimum, maximum and intermediate values for each parameter are also given with the standard value highlighted in bold. The parameters are derived from the uncertainty study by Murphy et al. [50] and from known climate sensitive parameters in FAMOUS as described in [61]. RHCrit, VF1, CT, CW_LAND and CW_SEA are all involved in cloud processes. ATM_DIFF, OCN_DIFF_H and OCN_DIFF_V are associated with diffusion processes. The elements of CW are varied as a pair in the MPPs but are perturbed separately in the SPPs.

<table>
<thead>
<tr>
<th>no.</th>
<th>parameter name</th>
<th>parameter description</th>
<th>max.</th>
<th>int.</th>
<th>min.</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>RHCrit</td>
<td>threshold of relative humidity for cloud formation</td>
<td>0.9</td>
<td>0.687</td>
<td>0.6</td>
</tr>
<tr>
<td>2</td>
<td>VF1</td>
<td>precipitating ice fall-out speed</td>
<td>2</td>
<td>1.75</td>
<td>0.5</td>
</tr>
<tr>
<td>3</td>
<td>CT</td>
<td>conversion rate of cloud liquid water droplets to precipitation</td>
<td>$4 \times 10^{-4}$</td>
<td>$9.41 \times 10^{-5}$</td>
<td>$5 \times 10^{-5}$</td>
</tr>
<tr>
<td>4</td>
<td>CW*</td>
<td>threshold value of cloud liquid water for formation of precipitation over the sea and over land</td>
<td>$5 \times 10^{-4}$</td>
<td>$3.82 \times 10^{-5}$</td>
<td>$2 \times 10^{-5}$</td>
</tr>
<tr>
<td></td>
<td>CW_SEA</td>
<td></td>
<td>$2 \times 10^{-3}$</td>
<td>$1.61 \times 10^{-4}$</td>
<td>$1 \times 10^{-4}$</td>
</tr>
<tr>
<td>5</td>
<td>G_WAVE</td>
<td>gravity wave parameters (two parameters)</td>
<td>$2 \times 10^{4}$</td>
<td>$1.50 \times 10^{4}$</td>
<td>$1 \times 10^{4}$</td>
</tr>
<tr>
<td></td>
<td>K_GWAVE</td>
<td></td>
<td>$3 \times 10^{5}$</td>
<td>$2.20 \times 10^{5}$</td>
<td>$1.50 \times 10^{5}$</td>
</tr>
<tr>
<td></td>
<td>KAY_LEE_GWAVE</td>
<td></td>
<td>$3 \times 10^{-8}$</td>
<td>$2.20 \times 10^{-8}$</td>
<td>$2 \times 10^{-8}$</td>
</tr>
<tr>
<td>6</td>
<td>Z0FSEA</td>
<td>the free convective roughness length over the sea for boundary layer processes</td>
<td>0.005</td>
<td>0.00111</td>
<td>0.0002</td>
</tr>
<tr>
<td>7</td>
<td>ALPHAM</td>
<td>albedo (reflectivity) of sea-ice variability with temperature</td>
<td>0.65</td>
<td>0.5</td>
<td>0.2</td>
</tr>
<tr>
<td>8</td>
<td>ATM_DIFF</td>
<td>the horizontal atmospheric diffusion parameters varied together (two parameters)</td>
<td>$4.19 \times 10^{9}$</td>
<td>$3.85 \times 10^{9}$</td>
<td>$3.50 \times 10^{9}$</td>
</tr>
<tr>
<td></td>
<td>DIFF_COEFF</td>
<td></td>
<td>$2.40 \times 10^{8}$</td>
<td>$2.20 \times 10^{8}$</td>
<td>$2 \times 10^{8}$</td>
</tr>
<tr>
<td></td>
<td>DIFF_COEFF_Q</td>
<td></td>
<td>$1.10 \times 10^{-5}$</td>
<td>$1 \times 10^{-5}$</td>
<td>$8 \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>KAPPA0_SI</td>
<td></td>
<td>$1.10 \times 10^{-5}$</td>
<td>$1 \times 10^{-5}$</td>
<td>$8 \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>DKAPPA_DZ_SI</td>
<td></td>
<td>$3.08 \times 10^{-8}$</td>
<td>$2.80 \times 10^{-8}$</td>
<td>$2.20 \times 10^{-8}$</td>
</tr>
<tr>
<td>9</td>
<td>OCN_DIFF_H***</td>
<td>oceanic horizontal diffusion parameters varied together (four parameters)</td>
<td>$0.00605$</td>
<td>$0.0055$</td>
<td>0.005</td>
</tr>
<tr>
<td></td>
<td>FNUB_SI</td>
<td></td>
<td>$1.10 \times 10^{-5}$</td>
<td>$1 \times 10^{-5}$</td>
<td>$8 \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>KAPPA0_SI</td>
<td></td>
<td>$1.10 \times 10^{-5}$</td>
<td>$1 \times 10^{-5}$</td>
<td>$8 \times 10^{-6}$</td>
</tr>
<tr>
<td></td>
<td>DKAPPA_DZ_SI</td>
<td></td>
<td>$3.08 \times 10^{-8}$</td>
<td>$2.80 \times 10^{-8}$</td>
<td>$2.20 \times 10^{-8}$</td>
</tr>
<tr>
<td></td>
<td>FNUO_SI</td>
<td></td>
<td>$0.00605$</td>
<td>$0.0055$</td>
<td>0.005</td>
</tr>
</tbody>
</table>

(Continued)
In order to understand some of the causes of the changes in climate, we selected a simulation with a promising Early Eocene climate based on the 6000 year results (from herein referred to as E6000). The climate in E6000 exhibited global warmth (SAT more than 30°C) and polar regions with SAT more than 10°C. We used the 10 groups of perturbed parameter values in E6000 to set up a further set of simulations in order to investigate the response of the climate to changes in one parameter at a time. This second group of experiments was termed the single parameter perturbations (SPPs). We ran 15 SPP simulations in total from the original 10 parameter groups by separating the parameters in CW (threshold value of cloud liquid water at which precipitation commences) into land and sea components; the four parameters in the OCN_DIFF_H group, horizontal ocean diffusivity, were also split into three separate experiments. Finally, OCN_DIFF_V, vertical ocean diffusivity and ATM_DIFF, horizontal atmosphere diffusion parameter groups were sampled twice: once using the values in E6000 and then a second set of simulations were conducted reducing the values even further than in E6000. These simulations are run for up to 9000 years. A summary of the different sets of simulations and criteria used to assess them is shown in table 2. Although E6000 does not make it into the final Eocene simulations, the parameter values of E6000 are shown in table 3 for reference.

(e) Present-day simulations with twice pre-industrial CO₂

PD simulations and with twice pre-industrial CO₂ concentrations (560 ppmv) were used to calculate Charney climate sensitivity values for the same MPP sets that were used in the Early Eocene simulations (E_MPP). The PD configuration is identical to that described in §2b with the exception that CO₂ concentrations of 560 ppmv were used. These simulations were run for 200 years.

(f) Model–data comparison

Model output is compared to published multi-proxy datasets that have undergone comprehensive selection and standardization. We use the terrestrial dataset first compiled in Huber & Calallero [3] and also used in Lunt et al. [4]. Our marine dataset is also from Lunt et al. [4]. An outline of the proxy data and consideration of the uncertainty associated with this data is given below.

The terrestrial proxy dataset compiled by Huber & Caballero [3] contains fifty Early Eocene data of Ypresian (56.0–47.8 Ma) and Lower Lutetian age. The Lutetian occurred between 47.8 and 41.3 Ma; however Lu1, the first global section of the Lutetian, is dated at 47.47 Ma thus we take the age span of the terrestrial data to be between 56.0 and 47.47 Ma. PETM (Palaeocene–Eocene Thermal Maximum) and other hyperthermal events were excluded in the compilation of the dataset by Huber & Caballero [3]. One Middle Eocene data point approximately 45 Ma from
Table 2. Summary of the three groups of experiments discussed in this paper and the criteria used to assess and rank these simulations. The three groups of experiments are: PD_MPP, a present day 100 member multiple perturbed parameter ensemble; E_MPP, an Eocene 100 member multiple perturbed parameter ensemble; and E_SPP, an Eocene 14 ensemble single perturbed parameter ensemble. In the PL_MPP only 14 simulations outperform the control parameter set. Only 15 E_MPP and 2 E_SPP simulations are deemed successful, which does not include the control parameter set.

<table>
<thead>
<tr>
<th>ID</th>
<th>description</th>
<th>details</th>
<th>assessing the simulations</th>
<th>final</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD_MPP</td>
<td>100 pre-industrial MPP (multiple perturbed parameter) ensemble run initially for 200 years, with a subset continued for an additional 300 years (500 years in total)</td>
<td>10 parameters perturbed as identified from [50] and from climate sensitive parameters in the model FAMOUS [61]. Perturbed parameter sets were generated using Latin hypercube sampling (see [61] for full details)</td>
<td>an Arcsin Mielke [99] score was calculated for all 100 simulations and the control simulation (see [61]). Simulations with a higher Arcsin Mielke score than the control (14) and the control simulation were continued for an additional 300 years. Thus 86 simulations were not continued</td>
<td>14 simulations and control simulation with standard parameter set were run for 500 years</td>
</tr>
<tr>
<td>E_MPP</td>
<td>100 Eocene MPP (multiple perturbed parameters) ensemble run for up to 10,000 years.</td>
<td>parameter sets are identical to those in PL_MPP</td>
<td>successful simulations ran for the allotted time (10,000 years) and had stable topa (top of atmosphere) net energy balance and surface air temperatures. Fifty-nine simulations failed within 100 years; a further four failed to complete 1000 years, and 19 failed to complete 4000 years. Eighteen simulations complete 10,000 years of which three are identified as unstable. Simulation E6000 is one of the three unstable simulations</td>
<td>15 simulations run for up to 10,000 years. Three parameter sets overlap with the final PL_MPP simulations</td>
</tr>
<tr>
<td>E_SPP</td>
<td>14 Eocene SPP (single perturbed parameter) ensemble run for up to 9000 years</td>
<td>based on the climatologies of E_MPP at 6000 years, a simulation with a promising Eocene climate (E6000) was selected and used as the basis of the SPP. Eight parameter sets were varied as in E6000. Parameters in CW and OCN_DIFF_H were varied separately to create five further simulations. DIFF_COEF and DIFF_V were reduced further than in E6000 to give an additional split to give two further simulations; and DIFF_COEFF and DIFF_V were reduced further than in E6000 to create two further simulations</td>
<td>simulations were deemed successful if they ran to their allotted time and were stable as assessed by top of energy net energy balance and surface air temperature drift. Only three simulations completed their time, of which one was unstable</td>
<td>two simulations run for up to 9000 years</td>
</tr>
</tbody>
</table>
Table 3. Parameter values of the final 17 Eocene simulations as a percentage of the original standard parameter value (for standard parameter value see Table 1). Simulations are ranked in the order of lowest to highest mean annual temperatures (SAT, also shown), i.e. simulation E1 has the coolest SAT and E17 has the warmest SAT. The parameter values of simulation E6000 on which the single parameter perturbations were based are also included for reference, although note that this simulation is not part of the final 17 Eocene simulations.

<table>
<thead>
<tr>
<th>ID</th>
<th>SAT (°C)</th>
<th>RHCRIT</th>
<th>VF1</th>
<th>CT</th>
<th>CW_LAND</th>
<th>CW_SEA</th>
<th>Z0FSEA</th>
<th>KAY_GWAVE</th>
<th>ALPHAM</th>
<th>ATM_DIFF</th>
<th>OCN_DIFF_H</th>
<th>OCN_DIFF_V</th>
</tr>
</thead>
<tbody>
<tr>
<td>E1</td>
<td>12.3</td>
<td>100</td>
<td>73</td>
<td>67</td>
<td>754</td>
<td>787</td>
<td>429</td>
<td>86</td>
<td>312</td>
<td>99</td>
<td>103</td>
<td>98</td>
</tr>
<tr>
<td>E2</td>
<td>14.9</td>
<td>107</td>
<td>75</td>
<td>75</td>
<td>807</td>
<td>843</td>
<td>353</td>
<td>88</td>
<td>221</td>
<td>89</td>
<td>98</td>
<td>83</td>
</tr>
<tr>
<td>E3</td>
<td>15.2</td>
<td>107</td>
<td>52</td>
<td>77</td>
<td>1113</td>
<td>1169</td>
<td>167</td>
<td>69</td>
<td>132</td>
<td>97</td>
<td>105</td>
<td>88</td>
</tr>
<tr>
<td>E4</td>
<td>15.9</td>
<td>91</td>
<td>40</td>
<td>60</td>
<td>562</td>
<td>583</td>
<td>161</td>
<td>78</td>
<td>189</td>
<td>99</td>
<td>102</td>
<td>85</td>
</tr>
<tr>
<td>E5</td>
<td>16.9</td>
<td>117</td>
<td>88</td>
<td>71</td>
<td>630</td>
<td>655</td>
<td>376</td>
<td>91</td>
<td>182</td>
<td>85</td>
<td>97</td>
<td>99</td>
</tr>
<tr>
<td>E6</td>
<td>17.6</td>
<td>106</td>
<td>67</td>
<td>53</td>
<td>118</td>
<td>112</td>
<td>198</td>
<td>82</td>
<td>205</td>
<td>95</td>
<td>97</td>
<td>83</td>
</tr>
<tr>
<td>E7</td>
<td>20.8</td>
<td>101</td>
<td>107</td>
<td>195</td>
<td>1098</td>
<td>1152</td>
<td>100</td>
<td>88</td>
<td>286</td>
<td>87</td>
<td>89</td>
<td>86</td>
</tr>
<tr>
<td>E8</td>
<td>21.1</td>
<td>113</td>
<td>102</td>
<td>124</td>
<td>971</td>
<td>1018</td>
<td>328</td>
<td>94</td>
<td>239</td>
<td>90</td>
<td>90</td>
<td>82</td>
</tr>
<tr>
<td>E9</td>
<td>21.1</td>
<td>109</td>
<td>91</td>
<td>117</td>
<td>992</td>
<td>1040</td>
<td>135</td>
<td>67</td>
<td>306</td>
<td>94</td>
<td>89</td>
<td>99</td>
</tr>
<tr>
<td>E10</td>
<td>21.2</td>
<td>117</td>
<td>104</td>
<td>99</td>
<td>771</td>
<td>805</td>
<td>439</td>
<td>55</td>
<td>309</td>
<td>88</td>
<td>89</td>
<td>86</td>
</tr>
<tr>
<td>E11</td>
<td>22.1</td>
<td>91</td>
<td>111</td>
<td>313</td>
<td>1201</td>
<td>1262</td>
<td>209</td>
<td>61</td>
<td>153</td>
<td>86</td>
<td>104</td>
<td>91</td>
</tr>
<tr>
<td>E12</td>
<td>24.6</td>
<td>89</td>
<td>42</td>
<td>135</td>
<td>909</td>
<td>951</td>
<td>401</td>
<td>81</td>
<td>202</td>
<td>99</td>
<td>85</td>
<td>90</td>
</tr>
<tr>
<td>E13</td>
<td>25.0</td>
<td>89</td>
<td>65</td>
<td>248</td>
<td>1017</td>
<td>1067</td>
<td>305</td>
<td>68</td>
<td>265</td>
<td>96</td>
<td>104</td>
<td>98</td>
</tr>
<tr>
<td>E14</td>
<td>25.3</td>
<td>89</td>
<td>100</td>
<td>100</td>
<td>10</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>E15</td>
<td>26.3</td>
<td>92</td>
<td>49</td>
<td>212</td>
<td>1210</td>
<td>1272</td>
<td>349</td>
<td>87</td>
<td>165</td>
<td>93</td>
<td>102</td>
<td>99</td>
</tr>
<tr>
<td>E16</td>
<td>29.7</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>10</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>E17</td>
<td>31.8</td>
<td>95</td>
<td>108</td>
<td>321</td>
<td>1050</td>
<td>1102</td>
<td>108</td>
<td>100</td>
<td>108</td>
<td>97</td>
<td>99</td>
<td>92</td>
</tr>
<tr>
<td>E6000</td>
<td>n.a.</td>
<td>89</td>
<td>90</td>
<td>309</td>
<td>379</td>
<td>389</td>
<td>25</td>
<td>93</td>
<td>214</td>
<td>86</td>
<td>102</td>
<td>82</td>
</tr>
</tbody>
</table>
the tropics is included in the absence of any tropical data from the Early Eocene [100,101]. There is no data coverage at latitudes greater than 65°S and coverage is highest in the NH particularly over North America.

In order to account properly for systematic bias and spatio-temporal sampling uncertainty, the authors have reconstructed MAT based on leaf-margin analysis (LMA) where possible. CLAMP (physiognomic analysis of leaf fossils) is used for MAT reconstruction when LMA is not available. MATs are calculated using the Kowalski & Dilcher [24] calibration when feasible as this offsets the well-established cool biases that may have been incorporated in the original calibrations (see [3] and references therein). Error bars are included in the terrestrial dataset to encompass the uncertainty introduced from the age of the material, topographic uncertainty and from the calibration method. All palaeolatitudes are adjusted to 55 Myr plate configuration using the GPlATES software (www.gplates.org) and the plate model of Muller et al. [102]. Palaeo-elevation uncertainty is quantified by calculating the standard deviation of PD topography at elevations greater than 1500 m, and then applying this to Eocene data to calculate the uncertainty in temperature as a result of lapse rate (±2.4°C), based on the work by Hren et al. [103].

The marine dataset used in this work was compiled by Lunt et al. [4] and includes data from 13 locations. The age range of the data spans the ages of approximately 55.0–49.0 Ma. Data are grouped into three categories by Lunt et al. [4] and include (i) data aged approximately 55 Ma that is termed Late Palaeocene data but excluding the PETM in [4]; (ii) well-constrained EECO data between 53.1 and 49 Myr; and (iii) Early Eocene data which is constrained to the Ypresian, but not thought to be a representative of the EECO. This final dataset is referred to as the background Ypresian. Given the recent new boundaries of the Ypresian (56–47.8 Myr) (www.stratigraphy.org) the Late Palaeocene data referred to in Lunt et al. [4] now are categorized as the earliest Eocene; thus, we term this dataset the earliest Eocene. Multiple data are available at several locations where either two proxy methods have been used or data of different ages are available and our final marine dataset contains 15 data points in total. Data are generally well constrained with the exception of the data from Seymour Island in the Antarctic Ocean [9], which is provisionally classed as background Early Eocene, although this may potentially be Middle Eocene in age.

Climate data are included from a range of proxies; δ18O (planktic foraminifera), δ18O (benthic dwelling molluscs), Mg/Ca (planktic foraminifera), clumped isotopes and Tex86. The authors have calculated three temperatures for the δ18O data [77,104,105] in order to capture the upper and lower bounds of temperature estimates. Similarly three assumed values of Mg/Ca values (3, 4 and 5 mol mol⁻¹) are used to calculate Mg/Ca temperatures. There are now several published calibrations available for Tex86 and the ‘high’, ‘low’ and ‘inverse’ calibrations are all used. In addition, samples with a branched versus isoprenoidal tetraether (BIT) index greater than 0.3 are excluded where possible as this is now accepted as good practice (see [106] for further details). However, Ypresian samples from Tanzania [15] and Hatchetigbee Bluff, coastal North America [107] were included by Lunt et al. [4] despite higher BIT indices (0.3–0.5), in order to include more Early Eocene data points.

We have averaged proxy temperatures calculated with different methods at the same location but we have not averaged data of different ages. As a result our dataset contains 15 points that we use to compare to model output. The temperature ranges of these data points are summarized in table 4. Minimum and maximum temperature estimates from the multiple proxy methods and calibration errors are plotted in all our estimates. The terrestrial dataset spans the ages of 56.0–47.47 Ma, and no divisions are specified. The marine dataset spans a slightly narrower age range (55.0–49.0 Ma), which is encompassed by the Ypresian, but has been subdivided into three categories: earliest Eocene, EECO and background Ypresian. Non-EECO data (i.e. earliest Eocene and Ypresian) is referred to as the background Early Eocene as it does not include the peak EECO temperatures.

A wide range of proxy data using different methods have been used in these datasets, which introduces uncertainty from numerous sources. For example, uncertainties are associated with reconstructing palaeolocation and depositional environments [73], age control and diagenesis and alteration [115]. The geochemical effects on biological material are another source of
Table 4. Summary of 15 marine proxy data points used for marine model–data comparison. The original dataset (19 data points) was compiled in Lunt et al. [4] from 13 locations. We have taken the mean temperature value from different proxy methods at each location but have not calculated means for data of different ages.

<table>
<thead>
<tr>
<th>LOC_ID</th>
<th>ID</th>
<th>palaeolatitude</th>
<th>palaeolongitude</th>
<th>median MAT (°C)</th>
<th>max. MAT (°C)</th>
<th>min. MAT (°C)</th>
<th>calibration (+/− °C)</th>
<th>proxy method</th>
<th>age</th>
<th>SITE_ID</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>-65.3</td>
<td>-1.2</td>
<td>13.2</td>
<td>16.9</td>
<td>7.2</td>
<td>0.7</td>
<td>foram. δ¹⁸O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>-65.7</td>
<td>-59.5</td>
<td>11.7</td>
<td>18.4</td>
<td>2.0</td>
<td>1.4</td>
<td>mollusc δ¹⁸O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>3</td>
<td>-61.1</td>
<td>79.3</td>
<td>13.8</td>
<td>17.1</td>
<td>10.3</td>
<td>0.7</td>
<td>foram. δ¹⁸O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>4</td>
<td>-63.9</td>
<td>156.8</td>
<td>24.9</td>
<td>28.9</td>
<td>23.1</td>
<td>4.0</td>
<td>Tex86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>-63.9</td>
<td>156.8</td>
<td>29.0</td>
<td>30.0</td>
<td>24.5</td>
<td>4.0</td>
<td>Tex86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>-54.2</td>
<td>-163.7</td>
<td>27.4</td>
<td>29.0</td>
<td>25.4</td>
<td>2.5</td>
<td>Tex86 and Mg/Ca</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>-18.0</td>
<td>34.6</td>
<td>30.7</td>
<td>32.1</td>
<td>27.3</td>
<td>2.3</td>
<td>Tex86 and δ¹⁸O</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>8</td>
<td>30.8</td>
<td>-71.6</td>
<td>27.7</td>
<td>29.4</td>
<td>25.2</td>
<td>3.0</td>
<td>clumped isotopes and Tex86</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Continued.*
<table>
<thead>
<tr>
<th>LOC_ID</th>
<th>ID</th>
<th>palaeolatitude</th>
<th>palaeolongitude</th>
<th>median MAT (°C)</th>
<th>max. MAT (°C)</th>
<th>min. MAT (°C)</th>
<th>calibration (+/− °C)</th>
<th>proxy method</th>
<th>age</th>
<th>SITE_ID</th>
<th>reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>9</td>
<td>83.2</td>
<td>28.9</td>
<td>22.4</td>
<td>14.5</td>
<td>11.1</td>
<td>4.0</td>
<td>Tex86</td>
<td>background Ypresian</td>
<td>302-4A</td>
<td>Sluijs et al. [7]</td>
</tr>
<tr>
<td>8</td>
<td>10</td>
<td>83.2</td>
<td>28.85</td>
<td>13.1</td>
<td>24.7</td>
<td>18.0</td>
<td>4.0</td>
<td>Tex86</td>
<td>EECO</td>
<td>302-4A</td>
<td>Sluijs et al. [7]</td>
</tr>
<tr>
<td>9</td>
<td>11</td>
<td>38.0</td>
<td>−56.12</td>
<td>25.6</td>
<td>27.4</td>
<td>23.7</td>
<td>2.3</td>
<td>Tex86 and foram. δ¹⁸O</td>
<td>earliest Eocene Bass River, USA</td>
<td>Sluijs et al. [111]; Watterson [112]</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>12</td>
<td>38.2</td>
<td>−56.69</td>
<td>25.6</td>
<td>26.7</td>
<td>24.1</td>
<td>4.0</td>
<td>Tex86</td>
<td>earliest Eocene Wilson Lake, USA</td>
<td>Zachos et al. [113]</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>13</td>
<td>5.5</td>
<td>−143.9</td>
<td>29.4</td>
<td>30.1</td>
<td>28.5</td>
<td>1.1</td>
<td>Mg/Ca</td>
<td>earliest Eocene ODP 865</td>
<td>Tripati &amp; Elderfield [18]</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>14</td>
<td>−31.1</td>
<td>−7.38</td>
<td>27.4</td>
<td>28.1</td>
<td>26.7</td>
<td>1.1</td>
<td>Mg/Ca</td>
<td>earliest Eocene DSDP 527</td>
<td>Tripati &amp; Elderfield [18]</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>15</td>
<td>22.0</td>
<td>−162</td>
<td>28.1</td>
<td>28.4</td>
<td>27.8</td>
<td>1.1</td>
<td>Mg/Ca</td>
<td>earliest Eocene ODP 1209</td>
<td>Zachos et al. [114]</td>
<td></td>
</tr>
</tbody>
</table>
considerable uncertainty, for example, while the effects of temperature and seawater $\delta^{18}$O on foraminiferal $\delta^{18}$O have been recognized for a long time [116], the effects of seawater CO$_2$ chemistry on foraminiferal $\delta^{18}$O were only recognized through culturing experiments in the late 1990s [117,118]. This led to the realization that foraminifera $\delta^{18}$O-based temperature estimates may be too low for periods of the past where atmospheric CO$_2$ was high, such as the Early Eocene [78,119–121]. Better constraints on Early Eocene CO$_2$ will also help improve temperature estimates from foraminiferal $\delta^{18}$O; however, other ‘unknown’ or currently unquantified factors which affect foraminiferal $\delta^{18}$O may not have been recognized yet.

Similarly, Tex$^{86}$ is a relatively new palaeothermometer [122] and understanding the environmental signal recorded by Tex$^{86}$ for the Early Eocene is exacerbated by use of this proxy outside of its calibration conditions. High latitude areas from which very warm Early Eocene temperatures have been recorded by Tex$^{86}$ (for example, the Arctic Ocean) would have undergone several months of darkness due to the boreal winter; the lack of these organisms in the modern high latitude oceans makes use of this proxy method problematic in polar regions [123]. Incubation experiments are required to calibrate the Tex$^{86}$ palaeothermometer for tropical SSTs as the PD ocean is simply not hot enough [124].

Proxy data are compared point by point with model output at grid box resolution and with zonal mean values. Where the same land surface type is not present in the model as in the proxy data the nearest matching land surface location is used along a band of longitude. Terrestrial data are compared with the surface air temperature at 1.5 m in the model over terrestrial surfaces while marine data are compared to the ocean temperature at a depth of 5 m.

3. Results

(a) Successful runs

Some combinations of model parameters generated by our sampling technique result in climates which are far from realistic, for either a modern climate or a palaeoclimate [61]. Moreover, in the extreme conditions of the Early Eocene, 82 out of 100 MPP simulations fail to complete due to the model generating very extreme climates (e.g. tropical temperature in excess of 50$^\circ$C) resulting in numerical instabilities in the model. Eocene MPP simulations were required to run for in excess of 10 000 years and Eocene SPPs ran for up to 9000 years. A summary of the initial number of simulations, the selection criteria and the final number of simulations for each set of experiments conducted is given in table 2. It should be noted that in all Eocene simulations, we needed to perform multi-millennial runs in order to reach near equilibrium in both the surface and deep ocean. In some cases, initial results from the first 1000 years of the Eocene MPP simulations gave significantly different results. For instance, some simulations showed an 8$^\circ$C change of global SAT between the end of 1000 years runtime and the end of 10 000 years. The latitudinal gradients were also impacted such that in some simulations the EPTD changed by more than 15$^\circ$C from 1000 to 10 000 years. Even between 4000 and 10 000 years, the gradient changed in some simulations by up to 5$^\circ$C. The changes seemed to be strongly linked with the effects of ocean overturning and the time scales are consistent with this. These results highlight the potential for misinterpretation of the climatic effects of model changes (either parameter or boundary conditions) if the simulations are less than a few thousand years in duration and justify the use of a relatively fast but comprehensive model such as FAMOUS.

In order to verify the stability of the Eocene simulations that completed 10 000 year runs, the time series of the global mean top-of-atmosphere (toa) net energy balance and global surface air temperatures were plotted against each other [125]. In three simulations global surface air temperature appeared to be in equilibrium but the toa net energy was not tending to zero and so these simulations were discarded. In the remaining 17 Eocene simulations (15 MPP and two SPP) the global mean net toa energy balance is less than 0.3 W m$^{-2}$ (and in most cases less than 0.1 W m$^{-2}$) indicating that the simulations were in radiative balance. Trends in time-series plots
of global mean annual surface air temperature are small in the final simulation set with most simulations varying less than 2°C over the final 1000 years of simulation.

(b) Climate of the final simulations

In initial condition ensembles, model parameters and forcings are identical throughout the ensemble but each simulation has a different starting state. In these ensembles, the natural variability in the system is of interest and thus an ensemble mean value is a useful measure. In PPEs such as the work described here, model parameters and forcings have been changed while the initial conditions are identical. The value of PPEs is in understanding where and how the climate converges and diverges within the ensemble. We therefore describe the range of climates simulated but do not present the ensemble mean.

The parameter values of the final simulations and of the control parameter set are given in table 5. The simulations are ranked in the order of ascending global SAT and this ranking is used to identify the different simulations, i.e. the simulation with the lowest SAT is termed E1 and simulation with the highest SAT is referred to as E17. We performed simple regression analysis of each parameter against a number of global annual climate values (i.e. SAT, mean annual precipitation (MAP), tropical SSTs, polar SSTs, EPTD planetary albedo, low cloud and high cloud global values) but the resulting $R^2$ correlation coefficients were all below 0.5 indicating that direct correlation between these variables is not strong and that it is the combination of changes which is key.

(i) Temperature and precipitation

Table 3 summarizes some climate variables for the final 17 simulations. SAT in our ‘final’ simulations ranges from 12°C to 32°C, MAP ranges from 2.7 to 4.1 mm d$^{-1}$ and there is a strong positive correlation between SAT and MAP, with an $R^2$ of 0.97 and a slope equivalent to a 0.76 mm d$^{-1}$ (approx. 25%) increase per 10°C. This strong relationship also holds for the land and ocean precipitation, i.e. the fraction that falls over land versus ocean (approx. 30% of total precipitation falls over land) remains approximately constant across the range of simulations.

Figure 1 shows the SAT averaged from years 9900 to 10 000 for two example runs: simulation E1, which has the coldest global mean temperature of 12.3°C; and the warmest model, E17, with a much higher global mean temperature of 31.8°C. Not surprisingly, the basic spatial patterns are quite similar between the two simulations but with a large offset of approximately 15°C. In E1 (figure 1a), the SATs are significantly below zero at high latitudes in both the north and south. These cold temperatures are even more pronounced seasonally (not shown) as temperatures decrease below $-20°C$ in large parts of the high latitude continents. By contrast, annual mean temperatures in the warmest models remain above freezing for almost the whole globe. Simulations E16 and E17 have no annual mean temperatures below zero and E15 has a small area of sub-freezing temperatures (reaching $-10°C$) in the very heart of Antarctica, although the coastal regions of Antarctica remain above freezing. Seasonally, there are still some sub-freezing temperatures but in the warmest models, these are confined to very small regions in the heart of the continents polewards of about 60° and in regions where there are no proxy data to evaluate such values.

The spatial patterns of precipitation for simulations E1 and E17 are shown in figure 2 for both summer and winter seasons. The patterns over land are broadly consistent between warm and cold models but with some of the most marked differences in precipitation occurring at high latitudes. In E1, the North Pole is a ‘polar desert’ (shown clearly in figure 2a), whereas in E17 the poles are relatively moist. This is unsurprising given the much warmer and sea-ice free polar regions in the warm model. In the tropics, there are some important differences particularly over the ocean where the cold model shows a distinct split intertropical convergence zone (also clear in figure 2a) whereas the warmer model has a much broader feature and centred on the equator. However, over land there are somewhat smaller differences in the patterns of precipitation.
Table 5. Global mean values for the final 17 Eocene simulations. Tropical mean temperatures are calculated from the mean temperatures between 10°S and 10°N. Polar temperatures are defined between 60° and 90° in each hemisphere (NH = Northern Hemisphere and SH = Southern Hemisphere). The EPTD for each hemisphere is calculated by subtracting the polar temperatures from the tropical temperatures in each hemisphere.

<table>
<thead>
<tr>
<th>ID</th>
<th>tropical SAT (°C)</th>
<th>tropical SSTs (°C)</th>
<th>NH polar SAT (°C)</th>
<th>NH polar SSTs (°C)</th>
<th>SH polar SAT (°C)</th>
<th>SH polar SSTs (°C)</th>
<th>NH terr. temp. (°C)</th>
<th>SH terr. temp. (°C)</th>
<th>MAP (mm d⁻¹)</th>
<th>surface albedo (%)</th>
<th>planetary albedo (%)</th>
<th>low cloud (%)</th>
<th>high cloud (%)</th>
<th>total cloud (%)</th>
<th>net solar radiation toa (W m⁻²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>14.5</td>
<td>27.9</td>
<td>0.8</td>
<td>0.1</td>
<td>27.0</td>
<td>0.1</td>
<td>27.7</td>
<td>27.7</td>
<td>33.6</td>
<td>33.1</td>
<td>24.7</td>
<td>24.7</td>
<td>25.4</td>
<td>23.4</td>
<td>22.3</td>
</tr>
<tr>
<td>E1</td>
<td>14.9</td>
<td>26.0</td>
<td>0.6</td>
<td>0.4</td>
<td>27.6</td>
<td>0.3</td>
<td>28.5</td>
<td>27.2</td>
<td>34.8</td>
<td>34.3</td>
<td>24.8</td>
<td>24.5</td>
<td>25.5</td>
<td>24.3</td>
<td>23.3</td>
</tr>
<tr>
<td>E2</td>
<td>15.2</td>
<td>26.0</td>
<td>0.6</td>
<td>0.4</td>
<td>38.2</td>
<td>0.0</td>
<td>39.6</td>
<td>39.6</td>
<td>40.6</td>
<td>37.4</td>
<td>25.7</td>
<td>25.4</td>
<td>26.3</td>
<td>25.1</td>
<td>24.6</td>
</tr>
<tr>
<td>E3</td>
<td>15.9</td>
<td>26.6</td>
<td>0.2</td>
<td>0.9</td>
<td>27.9</td>
<td>11.5</td>
<td>27.0</td>
<td>27.0</td>
<td>35.8</td>
<td>36.3</td>
<td>26.1</td>
<td>26.5</td>
<td>28.0</td>
<td>27.0</td>
<td>26.0</td>
</tr>
<tr>
<td>E4</td>
<td>16.9</td>
<td>27.7</td>
<td>0.2</td>
<td>1.3</td>
<td>29.8</td>
<td>10.1</td>
<td>30.6</td>
<td>30.6</td>
<td>32.1</td>
<td>31.7</td>
<td>27.7</td>
<td>27.5</td>
<td>28.3</td>
<td>27.2</td>
<td>26.5</td>
</tr>
<tr>
<td>E5</td>
<td>17.6</td>
<td>28.0</td>
<td>0.6</td>
<td>1.7</td>
<td>30.0</td>
<td>8.4</td>
<td>31.6</td>
<td>31.6</td>
<td>34.0</td>
<td>33.6</td>
<td>28.6</td>
<td>28.4</td>
<td>30.4</td>
<td>29.3</td>
<td>28.3</td>
</tr>
<tr>
<td>E6</td>
<td>20.8</td>
<td>32.1</td>
<td>0.9</td>
<td>3.1</td>
<td>33.7</td>
<td>7.4</td>
<td>35.4</td>
<td>35.4</td>
<td>37.4</td>
<td>35.3</td>
<td>26.5</td>
<td>26.3</td>
<td>28.7</td>
<td>27.6</td>
<td>26.7</td>
</tr>
<tr>
<td>E7</td>
<td>21.1</td>
<td>32.2</td>
<td>0.9</td>
<td>3.5</td>
<td>34.6</td>
<td>6.7</td>
<td>37.1</td>
<td>37.1</td>
<td>39.5</td>
<td>37.9</td>
<td>29.5</td>
<td>29.4</td>
<td>31.5</td>
<td>30.4</td>
<td>29.6</td>
</tr>
<tr>
<td>E8</td>
<td>21.1</td>
<td>32.3</td>
<td>0.0</td>
<td>6.2</td>
<td>34.0</td>
<td>9.5</td>
<td>36.5</td>
<td>36.5</td>
<td>39.1</td>
<td>39.0</td>
<td>30.9</td>
<td>30.8</td>
<td>32.8</td>
<td>32.0</td>
<td>31.8</td>
</tr>
<tr>
<td>E9</td>
<td>21.1</td>
<td>32.3</td>
<td>0.9</td>
<td>3.4</td>
<td>34.8</td>
<td>7.0</td>
<td>37.4</td>
<td>37.4</td>
<td>39.8</td>
<td>38.7</td>
<td>31.0</td>
<td>31.0</td>
<td>33.4</td>
<td>32.7</td>
<td>32.7</td>
</tr>
<tr>
<td>E10</td>
<td>21.2</td>
<td>32.3</td>
<td>0.9</td>
<td>3.4</td>
<td>34.8</td>
<td>7.0</td>
<td>37.4</td>
<td>37.4</td>
<td>39.8</td>
<td>38.7</td>
<td>31.0</td>
<td>31.0</td>
<td>33.4</td>
<td>32.7</td>
<td>32.7</td>
</tr>
<tr>
<td>E11</td>
<td>22.1</td>
<td>32.5</td>
<td>2.3</td>
<td>5.0</td>
<td>34.5</td>
<td>3.8</td>
<td>37.5</td>
<td>37.5</td>
<td>39.5</td>
<td>38.5</td>
<td>31.0</td>
<td>31.0</td>
<td>33.4</td>
<td>32.7</td>
<td>32.7</td>
</tr>
<tr>
<td>E12</td>
<td>24.6</td>
<td>34.9</td>
<td>2.9</td>
<td>6.4</td>
<td>38.2</td>
<td>1.0</td>
<td>40.2</td>
<td>40.2</td>
<td>42.8</td>
<td>41.8</td>
<td>33.1</td>
<td>33.1</td>
<td>35.8</td>
<td>35.0</td>
<td>34.6</td>
</tr>
<tr>
<td>E13</td>
<td>25.3</td>
<td>34.8</td>
<td>3.9</td>
<td>7.4</td>
<td>37.9</td>
<td>0.3</td>
<td>40.3</td>
<td>40.3</td>
<td>42.6</td>
<td>41.6</td>
<td>34.5</td>
<td>34.5</td>
<td>37.9</td>
<td>37.9</td>
<td>37.9</td>
</tr>
<tr>
<td>E14</td>
<td>25.3</td>
<td>34.1</td>
<td>4.3</td>
<td>13.3</td>
<td>36.7</td>
<td>1.7</td>
<td>40.9</td>
<td>40.9</td>
<td>43.7</td>
<td>42.7</td>
<td>36.2</td>
<td>36.2</td>
<td>39.9</td>
<td>39.9</td>
<td>39.9</td>
</tr>
<tr>
<td>E15</td>
<td>26.3</td>
<td>36.4</td>
<td>4.8</td>
<td>8.0</td>
<td>40.2</td>
<td>2.1</td>
<td>43.7</td>
<td>43.7</td>
<td>47.1</td>
<td>46.1</td>
<td>38.8</td>
<td>38.8</td>
<td>42.9</td>
<td>42.9</td>
<td>42.9</td>
</tr>
<tr>
<td>E16</td>
<td>27.7</td>
<td>38.2</td>
<td>11.4</td>
<td>15.2</td>
<td>42.3</td>
<td>9.7</td>
<td>46.0</td>
<td>46.0</td>
<td>49.3</td>
<td>48.3</td>
<td>41.8</td>
<td>41.8</td>
<td>46.1</td>
<td>46.1</td>
<td>46.1</td>
</tr>
<tr>
<td>E17</td>
<td>31.8</td>
<td>39.4</td>
<td>14.3</td>
<td>19.3</td>
<td>44.5</td>
<td>12.4</td>
<td>50.9</td>
<td>50.9</td>
<td>55.1</td>
<td>54.1</td>
<td>47.9</td>
<td>47.9</td>
<td>53.8</td>
<td>53.8</td>
<td>53.8</td>
</tr>
</tbody>
</table>
In both simulations, the sub-tropics are seasonally dry but annual averages reveal only very small areas which are dry throughout the year.

(ii) Equator-to-pole temperature difference

Although some simulations achieve very warm polar temperatures, they also have very warm tropical temperatures so that the resulting EPTDs are generally very similar to present. EPTDs are calculated for all simulations by subtracting mean polar temperatures (70°N to 90°N and 70°S to 90°S) from the mean equatorial temperatures (10°N to 10°S). We have calculated the marine EPTD for the PD control simulation (table 5), and the NH marine EPTD is 27.2°C. Marine EPTDs in the coldest and warmest Eocene simulations E1 and E17 are both 25.1°C. Intermediate models (E7–E15) have a greater NH marine EPTD of up to 31.9°C. In these intermediate simulations, sea-ice acts as a buffer, keeping the marine temperatures at high latitudes around 0°C. Once this reduces, polar oceans begin to warm which then reduces the gradient. Examination of the equivalent terrestrial gradients helps confirm this as it shows a
simpler gradual reduction in temperature gradient between the coldest and warmest models. The NH terrestrial EPTD for the PD is 39.2°C. The NH terrestrial EPTD for simulations E1–E15 ranges between 38.2°C and 43.5°C, whereas E16 and E17 have a NH terrestrial EPTD of approximately 32°C, a 6–7°C reduction compared to the PD. The SH terrestrial EPTD for the PD is very large (58.4°C) due to the ice covered Antarctic. During the Early Eocene, with no ice and no circumpolar current, the largest SH terrestrial EPTD for the Eocene is 42.7°C. However, the warmest Eocene models have a terrestrial EPTD of approximately 30.5°C which is a notable reduction. The SH marine EPTD in the PD simulation is 27.8°C, and the majority of the Eocene simulations have an SH marine EPTD between 24.6°C and 28.9°C. Three simulations have a smaller SH marine EPTD: these are E14 (20.9°C), E16 (22.9°C) and E17 (20.1°C). Thus, in our two warmest Eocene simulations (E16 and E17) terrestrial EPTD in both hemispheres and SH marine EPTD do show a small reduction in temperature gradients in both hemispheres, which are compatible with the reduced EPTD suggested by the sparse data available for the Eocene.

The 6 times pre-industrial CO2 HadCM3L simulation in the EoMIP [123] had the least polar amplification of temperature from the five models compared. HadCM3L is an intermediate model (resolution) between HadCM3 and FAMOUS. As it is part of the same family of models as FAMOUS we compare the EPTD as calculated using the method above with these simulations. The NH SST EPTD is 29.7°C and the SH SST EPTD is 26.3°C. These EPTDs are larger than those in our PD control but are well within the EPTD range simulated by the Eocene ensemble, maximum values of which are 32.2°C for the NH and 29.0°C for the SH.

(iii) Other climate characteristics

The sensitivity of the Early Eocene, proto-Atlantic Ocean meridional overturning circulation to changes in the concentration of CO2 (which changed the warmth and the presence of sea-ice) was described in Lunt et al. [4], for HadCM3L. We find similar results for the Atlantic overturning circulation in our suite of simulations. The warmer the simulation, the stronger the Atlantic intermediate-water formation, with a jump in the strength between simulations E6 and E7 related to a loss of year round sea-ice in the North Atlantic. Further increase in Atlantic intermediate-water formation in the very warm simulations (E14–E17) is associated with the almost complete loss of seasonal sea-ice. However, the location of oceanic convection, as indicated by the mixed layer depth, remains quite similar in all models. An intermediate to deep anticlockwise flow develops in the models where sea-ice disappears in the South Pacific (e.g. simulations E9, E14, E16 and E17). The centre of the cell is between 1000 and 2000 m with the bottom of the cell extending to 4000 m in E16 and up to 3000 m in E9, E14, E16 and E17. This replaces a deeper, small bottom water cell in the cooler models that have year round South Pacific sea-ice. In addition to the high latitude sources, there is also a source of intermediate water within the Tethys seaway. The relatively enclosed basin is very warm and experiences high evaporation. As a consequence, the surface waters are sufficiently saline to sink and these then spread out at about 2 km depth (figure 3).

A substantial increase in tropospheric mid-latitude westerlies which increase by more than 25% between coldest and warmest simulation is observed. Moreover, the strength of the tropical easterlies weakens considerably but they do not transition to westerlies as seen in [124]. It is probable that some of this difference is related to the resolution of FAMOUS which does not represent the atmospheric wave dynamics (particularly, the Madden Julian Oscillation) reported in [126]. The strengthening of the westerlies in our simulations seems to be strongly linked to a much intensified Hadley cell.

(c) Model–data comparison

Figures 4 and 5 show, respectively, the zonal means for the mean annual SSTs and terrestrial SATs for all 17 final simulations. The marine and terrestrial proxy datasets [3,4] are overlaid in these
Figure 3. Atlantic meridional streamfunction ($Sv$) for the present day (PD) and final ensemble of 17 Eocene models. Positive values indicate clockwise motion and negative values indicate anticlockwise motion. (Online version in colour.)

plots along with the lower and upper temperature bounds and the calibration errors for each data point. No simulation has a SAT or SST zonal mean that intersects all the proxy points (including the error bars).

In order to assess rigorously how well each simulation matches the proxy temperature estimates, we calculate the root mean square error (r.m.s.e.) for the difference between the simulation temperature predictions and the proxy data temperature estimates for all marine and terrestrial data (table 6). The r.m.s.e.’s for the different time periods included in the marine dataset have also been calculated (i.e. Ypresian, earliest Eocene and EECO) as have r.m.s.e.’s for mid- and high latitudes in the terrestrial dataset. Low r.m.s.e. values indicate that there is a better model–data fit than large r.m.s.e.’s. Simulations E1–E10 have greater terrestrial r.m.s.e.’s than marine r.m.s.e.’s, whereas simulations E12–E17 have greater marine r.m.s.e.’s than terrestrial r.m.s.e.’s indicating that above 22.1°C (SAT of E11) an improved fit with terrestrial proxy data is at the expense of the fit with the marine dataset.
Figure 4. Early Eocene SSTs as compiled in Lunt et al. [4] shown as solid black circles. Upper and lower temperature error bars are shown in black. Model simulated zonal SSTs are plotted over the top. The four warmest simulations, E14 (dotted line), E15 (dashed double-dotted line), E16 (dashed single-dotted line) and E17 (dashed line), are highlighted with thicker lines for clarity. (Online version in colour.)

Figure 5. Early Eocene terrestrial SATs as compiled in Huber & Cabellero [3] shown as solid black circles. Upper and lower temperature error bars are shown in black and calibration errors are plotted in grey. Model simulated terrestrial zonal SATs are plotted over the top. The four warmest simulations, E14 (dotted line), E15 (dashed double-dotted line), E16 (dashed single-dotted line) and E17 (dashed line), are highlighted with thicker lines for clarity. (Online version in colour.)
Table 6. Root mean square error (r.m.s.e.) calculations for differences between simulation temperature predictions and proxy data temperature estimates. The r.m.s.e. has been calculated for the entire combined terrestrial and marine proxy dataset and the separated marine and terrestrial datasets (highlighted in italics). The r.m.s.e. are also calculated for each subdivision of age in the marine data: earliest Eocene (approx. 55 Ma), Early Eocene Climatic Optimum (EECO) and Ypresian. The r.m.s.e. for different geographical subsets of terrestrial data have also been calculated. The minimum r.m.s.e. for each group and subgroup is highlighted in bold. All simulation estimates are calculated from a grid box mean centred over the proxy data points’ palaeolocation.

<table>
<thead>
<tr>
<th>ID</th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
<th>E4</th>
<th>E5</th>
<th>E6</th>
<th>E7</th>
<th>E8</th>
<th>E9</th>
<th>E10</th>
<th>E11</th>
<th>E12</th>
<th>E13</th>
<th>E14</th>
<th>E15</th>
<th>E16</th>
<th>E17</th>
</tr>
</thead>
<tbody>
<tr>
<td>entire marine dataset</td>
<td>16.1</td>
<td>14.9</td>
<td>14.8</td>
<td>14.7</td>
<td>14.2</td>
<td>14.0</td>
<td>12.2</td>
<td>12.0</td>
<td>11.4</td>
<td>12.0</td>
<td>10.9</td>
<td>10.6</td>
<td>10.3</td>
<td>8.9</td>
<td>10.0</td>
<td>8.1</td>
<td>7.9</td>
</tr>
<tr>
<td>earliest Eocene data</td>
<td>8.0</td>
<td>6.0</td>
<td>6.1</td>
<td>5.3</td>
<td>4.8</td>
<td>4.1</td>
<td>2.2</td>
<td>2.5</td>
<td>2.8</td>
<td>2.7</td>
<td>1.9</td>
<td>2.7</td>
<td>2.7</td>
<td>2.2</td>
<td>3.5</td>
<td>5.1</td>
<td>6.4</td>
</tr>
<tr>
<td>background Early Eocene data</td>
<td>15.9</td>
<td>15.0</td>
<td>14.8</td>
<td>14.6</td>
<td>14.3</td>
<td>14.0</td>
<td>12.7</td>
<td>12.5</td>
<td>11.9</td>
<td>12.5</td>
<td>11.4</td>
<td>11.2</td>
<td>10.9</td>
<td>9.7</td>
<td>10.6</td>
<td>8.8</td>
<td>9.4</td>
</tr>
<tr>
<td>EECO data</td>
<td>24.5</td>
<td>23.0</td>
<td>23.0</td>
<td>23.2</td>
<td>22.2</td>
<td>22.0</td>
<td>19.0</td>
<td>18.5</td>
<td>17.6</td>
<td>18.4</td>
<td>16.9</td>
<td>16.0</td>
<td>15.4</td>
<td>12.9</td>
<td>14.7</td>
<td>9.9</td>
<td>6.2</td>
</tr>
<tr>
<td>Antarctic Ocean data</td>
<td>17.0</td>
<td>15.0</td>
<td>15.3</td>
<td>15.0</td>
<td>14.0</td>
<td>13.5</td>
<td>9.6</td>
<td>9.2</td>
<td>7.8</td>
<td>9.0</td>
<td>7.7</td>
<td>6.0</td>
<td>5.5</td>
<td>5.2</td>
<td>5.2</td>
<td>7.7</td>
<td>9.2</td>
</tr>
<tr>
<td>Pacific Ocean data</td>
<td>18.2</td>
<td>17.3</td>
<td>17.1</td>
<td>17.0</td>
<td>16.5</td>
<td>16.4</td>
<td>14.6</td>
<td>14.2</td>
<td>13.4</td>
<td>14.2</td>
<td>12.9</td>
<td>12.6</td>
<td>12.1</td>
<td>9.5</td>
<td>11.9</td>
<td>9.3</td>
<td>8.2</td>
</tr>
<tr>
<td>Atlantic Ocean data</td>
<td>8.1</td>
<td>6.2</td>
<td>6.0</td>
<td>5.5</td>
<td>5.1</td>
<td>4.3</td>
<td>2.1</td>
<td>2.5</td>
<td>3.0</td>
<td>2.7</td>
<td>1.7</td>
<td>2.0</td>
<td>2.0</td>
<td>1.3</td>
<td>2.1</td>
<td>3.5</td>
<td>5.4</td>
</tr>
<tr>
<td>Arctic Ocean data</td>
<td>20.0</td>
<td>19.8</td>
<td>19.6</td>
<td>19.7</td>
<td>19.7</td>
<td>19.6</td>
<td>19.8</td>
<td>19.7</td>
<td>19.9</td>
<td>19.8</td>
<td>18.9</td>
<td>19.0</td>
<td>18.8</td>
<td>17.5</td>
<td>18.2</td>
<td>11.3</td>
<td>8.5</td>
</tr>
<tr>
<td>Late Palaeocene and background Early Eocene data</td>
<td>13.2</td>
<td>12.1</td>
<td>12.0</td>
<td>11.7</td>
<td>11.4</td>
<td>11.0</td>
<td>9.8</td>
<td>9.7</td>
<td>9.3</td>
<td>9.7</td>
<td>8.8</td>
<td>8.7</td>
<td>8.5</td>
<td>7.5</td>
<td>8.4</td>
<td>7.5</td>
<td>8.3</td>
</tr>
<tr>
<td>entire terrestrial dataset</td>
<td>21.6</td>
<td>18.8</td>
<td>18.1</td>
<td>17.1</td>
<td>16.3</td>
<td>14.9</td>
<td>12.9</td>
<td>12.9</td>
<td>14.0</td>
<td>12.9</td>
<td>11.0</td>
<td>8.8</td>
<td>8.0</td>
<td>7.7</td>
<td>7.0</td>
<td>4.8</td>
<td>5.1</td>
</tr>
<tr>
<td>terrestrial SH polar region</td>
<td>22.6</td>
<td>19.2</td>
<td>18.5</td>
<td>17.9</td>
<td>17.0</td>
<td>16.1</td>
<td>12.4</td>
<td>12.0</td>
<td>10.8</td>
<td>12.0</td>
<td>10.7</td>
<td>7.9</td>
<td>7.2</td>
<td>5.3</td>
<td>6.7</td>
<td>3.7</td>
<td>5.5</td>
</tr>
<tr>
<td>terrestrial SH mid-latitudes</td>
<td>18.6</td>
<td>15.2</td>
<td>15.2</td>
<td>14.7</td>
<td>13.7</td>
<td>12.8</td>
<td>9.8</td>
<td>9.3</td>
<td>8.8</td>
<td>9.2</td>
<td>8.5</td>
<td>6.3</td>
<td>5.8</td>
<td>4.9</td>
<td>5.4</td>
<td>3.7</td>
<td>5.5</td>
</tr>
<tr>
<td>terrestrial NH polar region</td>
<td>18.5</td>
<td>16.8</td>
<td>16.3</td>
<td>15.0</td>
<td>14.4</td>
<td>12.9</td>
<td>10.6</td>
<td>10.9</td>
<td>11.9</td>
<td>10.8</td>
<td>9.1</td>
<td>7.0</td>
<td>6.3</td>
<td>6.8</td>
<td>5.6</td>
<td>4.8</td>
<td>5.2</td>
</tr>
<tr>
<td>terrestrial NH mid-latitudes</td>
<td>33.3</td>
<td>27.9</td>
<td>26.1</td>
<td>25.8</td>
<td>24.1</td>
<td>22.5</td>
<td>22.1</td>
<td>21.5</td>
<td>24.4</td>
<td>21.8</td>
<td>18.5</td>
<td>15.7</td>
<td>14.8</td>
<td>13.1</td>
<td>12.3</td>
<td>5.8</td>
<td>4.1</td>
</tr>
<tr>
<td>entire marine and terrestrial dataset</td>
<td>20.5</td>
<td>18.0</td>
<td>17.4</td>
<td>16.6</td>
<td>15.8</td>
<td>14.7</td>
<td>12.7</td>
<td>12.7</td>
<td>13.4</td>
<td>12.7</td>
<td>11.0</td>
<td>9.2</td>
<td>8.6</td>
<td>8.0</td>
<td>7.8</td>
<td>5.7</td>
<td>5.9</td>
</tr>
</tbody>
</table>
Differences between the proxy temperature and simulation temperature estimates have been calculated for the marine and terrestrial datasets and are shown in figures 6 and 7. These are used to assess how well the simulations match the proxy data and to visualize any bias in the simulations. The simulation errors in the terrestrial data (figure 6) have an ‘approximately’ normal distribution. Simulations E1–E15 consistently underpredict terrestrial temperatures (e.g. the distribution is centred below zero). Simulations E16 and E17 over- and underpredict an equal number of terrestrial data points by up to ±10°C (e.g. the distribution is centred about zero).

Figure 7 shows the differences for the marine dataset. There are not enough marine data points to assess the distribution of the data. Many of the simulations are skewed to the right indicating an overprediction of SSTs. Simulations E14 and E15 are centred near zero and overpredict SSTs in half the data points by up to 5°C but underpredict the remaining SSTs by between 10°C and 20°C. Simulations E16 and E17 are also centred near to zero; both simulations overpredict SSTs in half the dataset by up to 10°C and underpredict SSTs by the same amount (E17) or slightly more, up 15°C (E16).

The four warmest simulations (E14–E17) all consistently overpredict SSTs at four locations. These are the Ypresian age data recorded at Tanzania and Hatchetigbee Bluff and the earliest Eocene data from ODP 865 and ODP 1209 and Seymour Island. E14 has the smallest error and E17 has the largest error in all these locations, with the error varying between 1°C and 11°C at these locations. Three of these locations have specific uncertainties associated with them, uncommon with the other marine data points. The data at Tanzania and Hatchetigbee Bluff were included in the marine data compilation despite having BIT indices between 0.3 and 0.5, indicating a large terrestrial organic matter component in the data signal. This has increased uncertainty in the SST estimate [106,127] but to what degree is not stated. Similarly, the data from Seymour Island have provisionally been aged as earliest Eocene; however, the possibility of these data being Middle Eocene age has been raised [128]. If these data are re-assigned to a Middle Eocene age it may be assumed that Early Eocene temperatures at this location would be higher. The overestimate
at Seymour Island is the greatest for simulations E14–E17 from the marine dataset and this is possibly the marine data point with the largest age uncertainty. Better age constraints at Seymour Island will allow this uncertainty to be resolved in the future. In contrast the Middle Eocene age tropical data point included in the terrestrial dataset actually compared reasonably well with the warmest simulations: temperatures are approximately 1.2°C and 2.9°C warmer than the proxy temperatures in simulations E16 and E17, which are well within the published error bars for this data point.

Three marine locations are consistently too cold in the Eocene simulations, all of which are high latitude EECO aged data points. The data are from the ACEX core in the Arctic Ocean at a latitude of approximately 83°N [7], Waipara river off New Zealand at a latitude of approximately 54°S [48,129] and at ODP 1172D in the southwest Pacific at a latitude of approximately 64°S [11]. The palæo-reconstruction of all three of these locations is for a shallow marine environment, with Waipara river and ODP 1172D being coastal and from restricted environments. The ACEX data point, although in a restricted basin, is in the most open setting. There are two factors that may contribute to the underestimation of temperatures at these locations, these are the bathymetry in the model and the use of a static orbital configuration. The original references (table 4) for these proxy data identify these locations as shallow water or restricted environments with water depths of up to approximately 2000 m; however, the bathymetric reconstruction in our model is between 2000 and 3000 m at all these locations. There is evidence of orbital forcing pacing the EECO climate [89,130,131]. Previous modelling studies that have investigated orbital forcing during the Early Eocene identified the climate of the high latitude terrestrial realm as being sensitive [33,34]. Further work with a dynamic orbital configuration may reduce the model–data discrepancy with the EECO proxy data.

Overall, differences between the terrestrial dataset and the simulations are much smaller than with the marine dataset, particularly in simulations E16 and E17. Temperatures at three
locations in North America are consistently overpredicted by approximately 10°C by simulations E16 and E17 compared to proxy temperatures. These locations are all along the south or west coast of North America, which was mountainous during the Early Eocene. The uncertainty in orographic reconstruction in these particular locations is high and close to ±1000 m [3]. Huber & Cabellero [3] calculate the temperature uncertainty associated with orographic uncertainty in the terrestrial dataset as ±2.4°C for an uncertainty of ±450 m based on the environmental lapse rate of 5.2°C km⁻¹ [103]. Given the larger orographic uncertainty at these locations and the coarser resolution of the land surface in our model than CCSM_H, the model this dataset was prepared for comparison with, a larger temperature error of at least ±5.2°C may be more representative here, and which provides a much improved fit between simulations and proxy data.

Taking the marine and terrestrial comparisons together, of the 17 final simulations, two simulations have a more optimal fit with the Early Eocene proxy data; these simulations are E16 and E17 which are the simulations with the highest SAT. The SAT of the best performing simulations for each model in the EoMIP study range between 24.0°C and 29.5°C, with the ECHAM model (2× CO₂) and HadCM3L (6× CO₂) at the bottom end of this range and CCSM_H (16× CO₂) at the upper end. Our two best simulations have higher SATs than the EoMIP models. The two optimal Eocene simulations are described below.

— E16 (SAT of 29.7°C) is an SPP where the horizontal atmospheric diffusion parameter (atm-diff) was reduced to 72% of control value, the parameter choice in this simulation being based on the parameter values in a promising Eocene MPP simulation at approximately 4000 years (the original simulation this was based on did not make it into the final selection).
— E17 (SAT of 31.8°C) is a multi-parameter perturbation where all 10 uncertain parameters were varied together in order to maximize the parameter space sampled in these experiments.

These two simulations also have marine and terrestrial EPTDs which are at the lowest end of the simulations. These simulations are much better at reproducing high latitude SH warmth than NH warmth. While neither simulation manages to replicate the high temperatures recorded in the marine EECO proxy data, the global warmth of the Early Eocene is captured. E16 and E17 do have limitations and neither fit the proxy data perfectly; however, investigating how climate processes and heat transport in these simulations may give us insights into understanding low polar seasonality and continental warmth during the Eocene.

(d) Are the models too hot?

It should perhaps also be noted that the tropical SSTs in the warmest models are very warm. The zonal mean SST is almost 40°C and in places within the tropics it even exceeds 42°C. Such high temperatures exceed the optimum for many modern day species of ocean biological processes [132,133] such as growth. Thresholds in foraminifera with symbiotic algae have also been linked to enzyme inactivation at temperatures more than 35°C [132]. However, it should be noted that these temperatures decrease away from the equator so that by about 15°N and 15°S they are nearer 35°C. Similarly, at a depth of 50 m the temperatures have decreased to 36°C. Temperature is a strong biogeographic control on ocean biota and reduced zonation of foraminifera and poleward migration of foraminifera has been shown during the Early Cenozoic [134,135]. Similarly, the selective extinction of warm water ocean taxa during subsequent climatic cooling events such as that at the Eocene–Oligocene transition [135] indicates that modern foraminifera are not representative of greenhouse climates such as the Eocene and the possibility that species can adapt to the extreme conditions these temperatures indicate cannot be ruled out [136]. Conversely, for the EECO marine data points, the data may not be hot enough. The uncertainty associated with biological proxy data from past warm periods continues to be problematic and the omission of strong orbital forcing in our model may preclude these temperatures from being simulated in this ensemble.
(e) Causes of the warmth

While it is relatively easy to analyse the reasons for the warmth in these simulations relative to the PD control climate it is more difficult to analyse the causes of the warmth between the two warmest models. If we compare simulations E16 and E17 to the PD control simulation we see that there are a number of drivers of change beyond the increase in CO₂. Firstly, the relative humidity within the simulations remains relatively constant (albeit with some small decreases at high latitudes in the mid-troposphere) so that the specific humidity increases at all levels and latitudes in the warmer simulations (E16 and E17) compared with the colder simulations (E1–E15) and the PD simulation (PI), resulting in a strong positive feedback from water vapour.

Secondly, the removal of land ice greatly decreases the surface albedo. However, this is not a straightforward feedback. In the colder runs, the land ice is largely replaced by heavy snowfall so that the global mean surface albedo does not change appreciably (table 5). However, in the warmer climate simulations there is a major decrease in snow cover and hence we have a strong positive feedback. Sea-ice also experiences major decreases in the warmest simulations.

The planetary albedo follows a similar relationship as surface albedo, with decreased albedo with warmer temperatures. As SAT increases in the ensemble, there is also an increase in net solar radiation at the top of the atmosphere (toa) indicating increased radiative forcing. However, there are some more complicated variations from this simple pattern. Specifically simulations E12 and E15 increase their planetary albedo compared to the overall downward trend and subsequently reduce the net solar radiation toa relative to the remaining simulations. This appears to be strongly linked to changes in cloud cover. Overall, the warmer models generally have less total cloud cover which is consistent with the idea that clouds are acting as a positive feedback in these simulations. Moreover, the total cloud cover is strongly correlated with the planetary albedo (table 5). However, the patterns are quite complicated. Low clouds have a tendency to cool the climate system (through their impact on albedo) and hence the large reductions in this type of cloud in our simulations are producing a positive feedback. However, high clouds also decrease which moderates this somewhat. At higher latitudes, all types of clouds act to warm the climate system and in most of the simulations we have an increase in high latitude cloudiness. The ratio of low clouds to high clouds decreases as SAT throughout the ensemble; in E1 this ratio is 1.1 and in E16 and E17 this ratio is 0.8 and 0.9, respectively. Further complicating matters, the parameters perturbed in these simulations impact cloud physical properties such as cloud water content, cloud ice content and subsequently cloud albedo. These variables were not output in these simulations and would need to be assessed alongside any changes in cloud amounts before any definitive conclusions on the radiative balance can be drawn, particularly in relation to the processes suggested in previous studies such as PSCs [28,37,38,137] and high latitude convective cloud feedbacks [39,40,138,139].

In terms of changes in EPTD, it is also useful to examine the poleward heat transport in the simulations. Peak values of heat transport (HT) occur at approximately 40° latitude in the Eocene ensemble and in the PD control. In the PD simulation peak values of HT are 5.3 PW in the NH and 4.9 PW in the SH. In the Eocene ensemble, peak HT ranges between 5.1 and 6.0 PW in the NH and between 5.0 and 5.7 PW in the SH. At the latitude of peak HT, atmospheric heat transport (AHT) accounts for between 89 and 94% of heat transport in the NH and between 85 and 93% of heat transport in the SH. For the modern, peak values of HT are approximately 5 PW at 35° latitude, with AHT comprising 78% and 92% of the total heat transport in the NH and SH in good agreement with [140]. Ocean heat transport (OHT) peaks much closer to the equator and can be important at those latitudes but is relatively unimportant further polewards. Figure 8 shows the distribution of HT for the PD and for the warmest Eocene simulations. The ranges of OHT in the SH and of AHT in the NH are particularly large but the total variation is always dominated by the atmosphere.

The OHT acts to transfer heat from the tropics to higher latitudes and to weaken the latitudinal temperature gradient. The correlation between tropical ocean temperatures and OHT is clearly
Figure 8. Ocean, atmosphere and total heat transport in four warmest simulations (E14–E17) and PD control simulation for each hemisphere plotted against the latitude. (Online version in colour.)

Figure 9. OHT calculated as a per cent of total heat transport in each hemisphere for the Eocene simulations and plotted against (a) tropical SSTs and (b) EPTD. NH data plotted as diamonds and SH data plotted as squares. \( R^2 \) correlation coefficients are also shown. (Online version in colour.)

shown in figure 9a. However, the link between OHT and EPTD is less clear (figure 9b). This is because of two reasons. Firstly, the OHT is not strong, and is almost negligible beyond about 45°N and S, and hence has its strongest effect on mid-latitudes. Most of the heat transport further polewards is performed by the atmosphere. Secondly, the link between total heat transport and EPTD is also complicated because the albedo varies between the simulations. This implies that the total heat transport required to maintain the gradient will also vary [141].

(f) Equivalent modern simulations

Of the 15 MPP simulations in the final 17 Eocene simulations, only one parameter set (from E17) surpasses the control parameter set for a PD simulation in the PD (see table 2 for criteria used to assess PD simulations). The PD equivalent of the Eocene MPP simulation E11 is ranked only one place behind the control simulation in the PD ensemble. The control parameter set, however, does not make it into the final ensemble of Eocene simulations.

For all of the 15 final Eocene MPP simulations, we have an equivalent PD control and 2× CO\(_2\) concentration simulation. Thus, it is possible to calculate the Charney climate sensitivity for
these parameter sets. The Charney climate sensitivity is broadly defined as the equilibrium global mean surface temperature change following a doubling of CO₂ concentration. The mean ± 1 s.d. value for Charney sensitivity for the 18 models assessed in the Fourth Assessment Report (AR4) of the IPCC [142] was 3.26°C ± 0.69°C. Of the subset of our 15 simulations that are run at 2 × CO₂ concentration the mean ± 1 s.d. value for the Charney sensitivity is 3.25°C ± 0.58°C, which is very similar to the AR4 mean value. The Charney sensitivity for our best model, E17, is calculated to be 2.7°C, which is below this mean estimate. CCSM3, which was used for the Eocene simulations by [3,46], has a PD climate sensitivity of 2.7°C, while HadCM3 the sister model of FAMOUS has a Charney sensitivity of 3.3°C. Thus, our best performing parameter set for the Eocene, and which was able to simulate the extreme warmth of the Eocene, actually has a reduced climate sensitivity compared to the control parameter set, and a very similar climate sensitivity to CCSM.

Moreover, Gregoire et al. [61] use the same 100 MPP parameter sets in their tuning study which focused on the PD and the LGM. Simulation S4 which is highlighted in their study as having a favourable fit has identical parameters to our Eocene simulation E17.

4. Conclusions

Our work is the first attempt at a comprehensive ensemble with perturbed climate-sensitive model parameters for the Early Eocene. The results show that we can get a large diversity in response, with global mean temperature changes which vary considerably, from temperatures that are slightly cooler than the modern to temperatures that are extremely warm. We have managed to simulate levels of warmth comparable to that of the Early Eocene at only twice pre-industrial CO₂ which is a much lower concentration than used by many other models.

Although many aspects contribute to this warmth, a strong sensitivity to albedo changes associated with cloud cover was apparent. Clouds remain one of the most uncertain aspects of climate modelling with little consensus over the sign of the cloud feedback. In this work, the choice of perturbed parameters affected the physical properties of the clouds. The physical properties of the clouds and the effect on radiative balance will be examined in future work.

Within the ensemble, as SAT increases OHT decreases in both the NH and SH. In the SH as tropical SSTs increase and polar SSTs increase this also correlates to a reduction in OHT. However, this relationship is not apparent when OHT and the EPTD are compared across the ensemble. This implies that OHT is not a major control on the EPTD. If OHT is not a major part of the EPTD, AHT and local radiative effects are likely to be involved in driving changes in the EPTD.

Proxy–model discrepancies are larger in the marine dataset than the terrestrial dataset. Simulation of the marine EECO temperatures is the most problematic, with the warmest simulation still 12°C too cool compared to the proxy Tex86 temperature estimated. Some of this temperature difference may be attributable to the use of a modern orbital forcing in these simulations. There is evidence for a strong precessional and eccentricity signal pacing the EECO; all the EECO data used in this study are from the high latitudes and previous studies indicate that the high latitudes are most sensitive to orbital forcing during the Eocene [33,34] and other periods [143,144].

It has been known for some time that perturbing the parameters of models can result in a wide spread of results. However, one of the most exciting aspects of our results is that the ‘best’ climate simulation for the Early Eocene was also one of the best simulations for the PD and LGM. For the Early Eocene, our results have to be partly tempered by the uncertainty in boundary conditions, particularly the lack of a precise indicator of past greenhouse gas concentrations. Therefore, we may be obtaining a good comparison to data for the wrong reasons.

When we apply this parameter set to a future climate change simulation, we find that the resulting temperature increase due to an instantaneous doubling of CO₂ (the so-called Charney climate sensitivity) is 2.7°C. This value is slightly below the mean estimates of Charney sensitivity from the AR4 [142]. This is perhaps surprising since there have been indications [145] that palaeoclimate data would imply that models were under sensitive. Our new results show that
it is possible that a model can respond strongly to past changes without it necessarily resulting in a high sensitivity to future changes.

Palaeoclimate research focused on comparing proxy data to models will never be able to ‘prove’ that climate models work. However, it does provide a unique test of models’ ability to simulate climates different to present. It is worth bearing in mind that even with an optimal choice of parameters there will be irreducible structural deficiencies in the model that cannot be mitigated. However, it is still very encouraging that a single model parameter set exists that results in a model that simulates well the PD, LGM, and Early Eocene.

Acknowledgements. The authors would like to thank the two anonymous reviewers for particularly helpful suggestions and comments. This work was carried out using the computational facilities of the Advanced Computing Research Centre, University of Bristol, http://www.bris.ac.uk/acrc. The Eocene palaeogeography was created by Fugro-Roberston.

Data accessibility. The simulations described in this work are available at the following website: http://www.bridge.bris.ac.uk/resources/simulations.

References


142. Solomon S. 2007 *Climate change 2007: contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change*. Cambridge, UK: Cambridge University Press.

