



1 **Global mean surface temperature and climate sensitivity of the**
2 **EEOO, PETM and latest Paleocene**

3 Gordon N. Inglis^{1,2,†}, Fran Bragg³, Natalie Burls⁴, David Evans⁵, Gavin L. Foster⁶, Matt
4 Huber⁷, Daniel J. Lunt³, Nicholas Siler⁸, Sebastian Steinig³, Richard Wilkinson⁹, Eleni
5 Anagnostou¹⁰, Marlow J. Cramwinckel¹¹, Christopher J. Hollis¹², Richard D. Pancost¹ and
6 Jessica E. Tierney¹³

- 7 1. Organic Geochemistry Unit, School of Chemistry, School of Earth Science, University
8 of Bristol, UK
9 2. Cabot Institute for the Environment, University of Bristol, UK
10 3. School of Geographical Sciences, University of Bristol, UK
11 4. Department of Atmospheric, Oceanic and Earth Sciences, George Mason University,
12 USA
13 5. Institute of Geosciences, Goethe University Frankfurt, Frankfurt am Main, Germany
14 6. School of Ocean and Earth Science, University of Southampton, UK
15 7. Department of Earth, Atmospheric, and Planetary Sciences, Purdue University, USA
16 8. College of Earth, Ocean and Atmospheric Sciences, Oregon State University, USA
17 9. School of Mathematics and Statistics, University of Sheffield, UK
18 10. GEOMAR Helmholtz Centre for Ocean Research Kiel, Germany
19 11. Department of Earth Sciences, Utrecht University, Netherlands
20 12. GNS Science, Lower Hutt, New Zealand
21 13. Department of Geosciences, The University of Arizona, 1040 E 4th St Tucson AZ USA

22 [†] present address: School of Ocean and Earth Science, University of Southampton, UK

24 Corresponding author: Gordon N. Inglis

25 Email: gordon.inglis@soton.ac.uk. Telephone: +44 (0)117 954 6395



26 **Abstract:**

27 Accurate estimates of past global mean surface temperature (GMST) help to contextualise
28 future climate change and are required to estimate the sensitivity of the climate system to CO₂
29 forcing during the geological record. GMST estimates from the latest Paleocene and early
30 Eocene (~57 to 48 million years ago) span a wide range (~9 to 23°C higher than pre-industrial)
31 and prevent an accurate assessment of climate sensitivity during this extreme greenhouse
32 climate interval. Here, we develop a multi-method experimental framework to calculate GMST
33 during three target intervals: 1) the latest Paleocene (~57 Ma), 2) the Paleocene-Eocene
34 Thermal Maximum (56 Ma) and 3) the early Eocene Climatic Optimum (EECO; 49.4 to 53.3
35 Ma). Using six independent methodologies, we find that average GMST estimates during the
36 latest Paleocene and PETM are 11.7°C (± 0.6°C) and 18.7°C (± 0.8°C) higher than pre-
37 industrial, respectively. GMST estimates from the EECO are 13.3°C (±0.5°C) warmer than
38 pre-industrial and comparable to previous IPCC AR5 estimates (12.7°C higher than pre-
39 industrial). Leveraging the extremely large 'signal' associated with these extreme warm
40 climates, we combine estimates of GMST and CO₂ from the latest Paleocene, PETM and
41 EECO to calculate a gross estimate of the average climate sensitivity between the early
42 Paleogene and today. This yields gross climate sensitivity estimates for the latest Paleocene,
43 PETM and EECO which range between 2.8 to 4.8°C (66% confidence). These largely fall
44 within the range predicted by the IPCC (1.5 to 4.5°C per doubling CO₂), but appear
45 incompatible with low values (between 1.5 and 2.8°C per doubling CO₂).

46

47

48

49

50

51



52 1. Introduction

53 Under high growth and low mitigation scenarios, atmospheric carbon dioxide (CO₂) could
54 exceed 1000 parts per million (ppm) by the year 2100 (Stocker et al., 2013). The long-term
55 response of the Earth System under such elevated CO₂ concentrations remains uncertain
56 (Stevens et al., 2016; Knutti et al., 2017; Hegerl et al., 2007). One way to better constrain these
57 climate predictions is to examine intervals in the geological past during which greenhouse gas
58 levels were similar to those predicted under future scenarios. This is the rationale behind the
59 Deep-time Model Intercomparison Project (DeepMIP) which aims to investigate the behaviour
60 of the Earth System in three high CO₂ climate states in the latest Paleocene and early Eocene
61 (~ 57–48 Ma) (Lunt et al., 2017; Hollis et al., 2019).

62 Sea surface temperature (SST) and land air temperature (LAT) proxies indicate that
63 the latest Paleocene and early Eocene were characterised by global mean surface
64 temperatures (GMST) much warmer than those of today (Cramwinckel et al., 2018; Farnsworth
65 et al., 2019; Hansen et al., 2013; Zhu et al., 2019; Caballero and Huber, 2013). Having a robust
66 quantitative estimate of the magnitude of warming relative to modern is useful for two primary
67 reasons: (1) it allows us to contextualise future climate change predictions by comparing the
68 magnitude of future anthropogenic warming with the magnitude of past natural warming; (2)
69 combined with CO₂ proxy data, it allows us to estimate climate sensitivity, a key metric for
70 understanding how the climate system responds to CO₂ forcing. The Fifth IPCC Assessment
71 Report stated that GMST was 9°C to 14°C higher than for pre-industrial conditions during the
72 early Eocene (~52 to 50 Ma) (Masson-Delmotte et al., 2014). Subsequent studies indicate a
73 wider range of estimates, from 9 to 23°C warmer than pre-industrial (Cramwinckel et al.,
74 2018; Farnsworth et al., 2019; Hansen et al., 2013; Zhu et al., 2019; Caballero and Huber, 2013)
75 (Figure 1). It is an open question whether this range arises from inconsistencies between the
76 methods used to estimate GMST, such as selection of proxy datasets, treatment of
77 uncertainty, and/ analysis of different time intervals. This has thwarted a robust assessment
78 of GMST estimates for the latest Paleocene and early Eocene.



79 Here we calculate GMST estimates within a consistent experimental framework for the
80 target intervals outlined by the Deep-time Model Intercomparison Project (DeepMIP): i) the
81 Early Eocene Climatic Optimum (EECO; 53.3 to 49.4 Ma), ii) the Paleocene-Eocene Thermal
82 Maximum (PETM, ca. 56 Ma) and iii) the latest Paleocene (LP, ca. 57-56 Ma). We use six
83 independent methods to obtain new GMST estimates for these three time periods, employing
84 recently compiled datasets of SST and LAT estimates (Hollis et al., 2019) and BWT estimates
85 (Cramer et al., 2009; Westerhold et al., 2018; Barnet et al., 2019). We also undertake a suite
86 of additional sensitivity studies to explore the influence of particular proxies on each GMST
87 estimate. We then combine GMST estimates from all six methods to generate a “best
88 estimate” GMST for each time slice and use these, with existing estimates of CO₂ (Gutjahr et
89 al., 2017; Anagnostou et al., 2016) to develop new estimates of ECS during the latest
90 Paleocene, PETM and EECO.

91

92 2. Methods and Materials

93 Three different input datasets are used to calculate GMST. Dataset D_{surf} consists of surface
94 temperature estimates. Dataset D_{deep} consists of deep-water temperature estimates. Dataset
95 D_{comb} consists of a combination of surface- and deep-water temperature estimates. Six
96 different methodologies make use of these datasets to estimate GMST. Below we describe
97 these datasets and methods.

98

99 2.1. Dataset D_{surf}

100 Dataset D_{surf} is version 0.1 of the DeepMIP database, as described in Hollis et al (2019). It
101 consists of SSTs and LATs for the latest Paleocene, PETM and EECO. The SSTs are from
102 multiple proxies, including foraminiferal $\delta^{18}O$, foraminiferal Mg/Ca, clumped isotopes ($\Delta 47$),
103 and TEX₈₆. Foraminiferal $\delta^{18}O$ values are calibrated to SST following Bemis et al. (1998).
104 Foraminiferal Mg/Ca are calibrated to SST following Evans et al. (2018). TEX₈₆ values are



105 calibrated to SST using BAYSPAR (Tierney and Tingley, 2014). $\Delta 47$ values are reported using
 106 the parameters and calibrations of the original publications (Evans et al., 2018; Keating-Bitonti
 107 et al., 2011). LATs are derived from leaf fossils, pollen assemblages, mammal $\delta^{18}\text{O}$, paleosol
 108 $\delta^{18}\text{O}$ and branched GDGTs. LAT estimates are calculated using the parameters and
 109 calibrations of the original publication (see Hollis et al., 2019 and ref. therein). The location of
 110 the proxies is shown in Figure 2. For each site, we utilise the uncertainty range reported in
 111 Hollis et al. (2019). We do not explore calibration uncertainty, but instead focus on the
 112 methodologies used to calculate GMST.

113 Four methods (D_{surf-1} , D_{surf-2} , D_{surf-3} and D_{surf-4}) are employed to calculate GMST from
 114 dataset D_{surf} . These methods employ parametric (D_{surf-1} , D_{surf-2} , D_{surf-4}) or non-parametric
 115 (D_{surf-3}) functions to estimate temperature. Each method conducts a 'baseline' calculation
 116 which uses the SST and LAT data compiled in accordance with the DeepMIP protocols (i.e.
 117 Hollis et al., 2019). Our baseline calculation ($D_{surf-default}$) excludes $\delta^{18}\text{O}$ values from
 118 recrystallized planktonic foraminifera as these estimates are significantly cooler than
 119 estimates derived from the $\delta^{18}\text{O}$ value of well-preserved foraminifera, foraminiferal Mg/Ca
 120 ratios and clumped isotope values from larger benthic foraminifera (see Hollis et al., 2019 and
 121 ref. therein). For each method, we also conduct a series of sub-sampling calculations, based
 122 on varying assumptions about the robustness of different proxies (Table 1). The first sensitivity
 123 experiment ($D_{surf-Frosty}$) includes $\delta^{18}\text{O}$ values from recrystallized planktonic foraminifera. The
 124 second sensitivity experiment ($D_{surf-NoTEX}$) removes TEX_{86} values as these give slightly
 125 higher SSTs than other proxies, especially in the mid-to-high latitudes (Bijl et al., 2009; Hollis
 126 et al., 2012; Inglis et al., 2015). The third sensitivity experiment ($D_{surf-NoMBT}$) removes
 127 MBT(')/CBT values derived from marine sediment archives as they may suffer from a cool bias
 128 (Inglis et al., 2017; Hollis et al., 2019). The fourth sensitivity experiment ($D_{surf-NoMammal}$)
 129 removes mammal and paleosol $\delta^{18}\text{O}$ values as these proxies may suffer from a cool bias
 130 (Hollis et al., 2019). For each method, GMST is calculated for: i) the Early Eocene Climatic



131 Optimum (EECO; 53.3 to 49.4 Ma), ii) the Paleocene-Eocene Thermal Maximum (ca. 56 Ma)
132 and iii) the latest Paleocene (LP; ca. 57-56 Ma).

133

134 2.1.1. D_{surf-1}

135 Method D_{surf-1} was first employed by Caballero and Huber (2013) to estimate GMST from
136 early Eocene surface temperature proxies in the era after pervasive recrystallization of
137 foraminiferal $\delta^{18}O$ values was recognized (e.g. Pearson et al., 2001; Pearson et al., 2007). This
138 study used data compilations which were the predecessors to the DeepMIP compilation
139 (Huber and Caballero, 2011, Hollis et al., 2012).

140 Here, the anomalies of individual proxy temperature data points with respect to modern
141 values at the corresponding paleolocation are first calculated. The calculation involves binning
142 into low, mid, and high latitudes (30°N to 30°S, 30°N/S to 60°N/S, and 60°N/S to 90°N/S), and
143 calculating the unweighted mean anomaly within these bins between the median
144 reconstructed value at a given locality and the temperature at the same location today (from
145 reanalysis). The geographically binned means are then weighted according to relative
146 spherical area to calculate a globally weighted mean temperature anomaly between the paleo-
147 time slice and modern. All samples are treated equally and considered independent. The
148 associated errors are added in quadrature with the inter-sample standard deviation. These
149 two sources of error were combined and normalized by the square root of the number of
150 samples. This method is intended as an unsophisticated, brute force approach to estimating
151 GMST when dealing with many localities with poorly characterized errors in which there is a
152 large difference between the reconstructed temperature at a given location and the modern
153 equivalent. It is not intended to ferret out small differences in GMST nor is it expected to work
154 well under conditions in which temperature gradients are stronger than today, continents are
155 far removed from their current configuration, or in situations in which systematic errors are not
156 readily mitigated by large sample size (i.e. when there are correlations in systematic errors



157 between proxies). It is designed to be relatively straightforward to interpret and simple to
158 reproduce without relying overly on climate models or sophisticated statistical models.

159 Various sanity checks are performed along the way to determine if the method is likely
160 to produce useful results for a given sampling distribution and what corrections should be
161 applied to optimize it. For example, if we sampled the modern temperature field using a
162 geographic sampling distribution for a given time interval, what would the reconstructed
163 modern temperature be? If we sampled the modern global, annual average surface
164 temperature field in the reanalysis product ERA-5 (mean value: 15.1°C) with the geographic
165 distribution of samples we have in the past, we obtain values of 16.9°C ($\pm 1.5^\circ\text{C}$) in the latest
166 Paleocene, 14.2°C ($\pm 1.7^\circ\text{C}$) for the PETM, and 15.2°C ($\pm 1.1^\circ\text{C}$) for EECO at the distribution
167 of localities. For the sampling densities and spatial structure of the latest Paleocene and early
168 Eocene, this method can approach the true value within $\sim 1.5^\circ\text{C}$ and the error propagation
169 adequately characterizes the error, in this 'perfect knowledge' scenario. Seeking precision
170 beyond that range is probably unwarranted. However, estimating the latest Paleocene and
171 early Eocene GMST may be somewhat easier than estimating the modern GMST because
172 temperature gradients are roughly half modern values or less, thus spatial heterogeneity is
173 much reduced. Indeed, in the limit of a completely flat temperature gradient, only one perfect
174 sample would be required to estimate GSMT.

175 We can use paleoclimate model results to characterise how well the existing
176 palaeographic sampling network will impact results. For the latest Paleocene, the
177 reconstructed GMST is 24.6°C ($\pm 1.3^\circ\text{C}$), compared to the true paleoclimate model mean of
178 25.8°C. For the PETM, the reconstructed GMST is 27.2°C ($\pm 1.5^\circ\text{C}$), compared to the true
179 paleoclimate model mean of 29.3°C. For the EECO, the reconstructed GMST is 25.3°C
180 ($\pm 0.7^\circ\text{C}$), compared to the true paleoclimate model mean of 25.8°C. This method produces
181 estimates that are within random error given otherwise perfect knowledge. It is also clear that
182 systematic errors introduced by limited paleogeographic sampling can be alleviated by
183 incorporating the systematic offset in mean values between the true paleoclimate model
184 GMST and the sampled paleoclimate model GMST. This is the only component in which



185 paleoclimate model information is included and we utilise this offset to correct for systematic
 186 errors. While this approach could be applied uncritically, it is best applied only within the
 187 context of studying the random and systematic error structure as described above and caution
 188 should be taken in using systematic corrections that are significantly bigger than the estimated
 189 random error.

190 The calculations shown here utilize two CESM1 simulations, as described
 191 in Cramwinckel et al., (2018; EO3 and EO4). The two cases are chosen to minimize the
 192 magnitude of the correction to GMST and the final result is not sensitive to the choice of
 193 reference simulation among these two. The magnitude of the global correction could be
 194 sensitive to both using different models or boundary conditions.

195

196 2.1.2. D_{surf-2}

197 In this method, GMST estimates are calculated using the method described in Farnsworth et
 198 al. (2019) where a transfer-function is used to calculate global mean temperature from local
 199 proxy temperatures. The transfer function is generated from a pair of Eocene climate model
 200 simulations, carried out at two CO_2 concentrations. The first simulations are the same $2\times \text{CO}_2$
 201 and $4\times \text{CO}_2$ HadCM3L Eocene simulations from Farnsworth et al (2019). The second
 202 simulations are the $4\times \text{CO}_2$ and $8\times \text{CO}_2$ CCSM3 simulations of Huber and Caballero (2011),
 203 also discussed in Lunt et al (2012). We then provide a final estimate based on each of our two
 204 models. The two models are configured for the Eocene with different paleogeographies.

205 The principal assumption of this approach is that global temperatures scale linearly
 206 with local temperatures, and that a climate model can represent this scaling correctly. The
 207 resulting GMST estimate is independent of the climate sensitivity of the model but is
 208 dependent on the modelled spatial distribution of temperature. For a single given proxy
 209 location with a local temperature estimate (T^{proxy}) we estimate global GMST ($\langle T \rangle^{\text{inferred}}$) as:

210



$$\langle T \rangle^{inferred} = \langle T^{low} \rangle + (T^{proxy} - T^{low}) \frac{\langle T^{high} \rangle - \langle T^{low} \rangle}{T^{high} - T^{low}} \quad (1)$$

212

213 where $\langle T^{low} \rangle$ and $\langle T^{high} \rangle$ are the global means of a low- and high- CO_2 model simulation
 214 respectively, and T^{low} and T^{high} are the local temperatures (same location as the proxy) from
 215 the same simulations. T^{low} and T^{high} represent local modelled SSTs or local modelled near-
 216 surface LATs (in contrast to Farnsworth et al. 2019 who only used local modelled near-surface
 217 LATs to calculate T^{low} and T^{high} , even if T^{proxy} was SST). If the proxy temperature is greater
 218 than T^{high} or cooler than T^{low} , then the inferred global mean is found by extrapolation rather
 219 than by interpolation and is therefore more uncertain (Figure 3). We repeat this process for
 220 each proxy data point (Figure 4) and take an average (\pm standard error) as our best estimate
 221 of global mean temperature.

222

223 2.1.3. D_{surf-3}

224 For D_{surf-3} , GMST estimates are calculated using Gaussian process regression (Figure 5-6;
 225 Bragg et al., Submitted). In this method, temperature is treated as an unknown function of
 226 location, $f(x)$. There are many possible functions that can fit the available proxy dataset. By
 227 using a Gaussian process model of the unknown function, we assume that temperature is a
 228 continuous and smoothly varying function of location, and once fitted to the data, the posterior
 229 mean of the model gives the most likely function form for the temperature. We use a Gaussian
 230 process prior and update it using the proxy data to obtain the posterior model which we can
 231 then use to predict the surface temperatures on a global grid. Prior specification of the model
 232 is via a mean function $E(f(x)) = m(x)$, and a covariance function $\text{Cov}(f(x), f(x')) = k(x, x')$ (which
 233 tells us how correlated $f(x)$ is with $f(x')$). We also specify the standard deviation of the
 234 observation uncertainty about each data point (σ_i^2). If $\mathbf{f} = (f(x_1), \dots, f(x_n))^T$ is a vector of
 235 temperature observations at each location x_i , then the model is:

236



$$\mathbf{f} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma}) \quad (2)$$

238

239 where $\mu_i = m(x_i)$ and $\Sigma_{ij} = k(x_i, x_j) + \mathbb{I}_{i=j}\sigma_i^2$. The proxy temperatures are expressed as
 240 anomalies to the present-day zonal mean temperature at the respective paleolatitude. We
 241 subtract the mean temperature anomaly for each time period and core experiment prior to the
 242 analysis and therefore fit the model to the residuals, using a zero-mean prior function. This
 243 means the predicted field will relax towards the mean surface warming in areas of no data
 244 coverage. The covariance function – which considers the clustering of proxy locations –
 245 describes the correlation between $f(x_i)$ and $f(x_j)$ in relation to the distance of x_i and x_j . We use
 246 a squared-exponential covariance function with Haversine distances replacing Euclidean
 247 distances so that correlation is a function of distance on the sphere. A heteroscedastic noise
 248 model is used to weight the influence of individual proxy data by their associated uncertainty,
 249 i.e. the model will better fit reconstructions with a smaller reported error.

250 Proxy uncertainties are taken from Hollis et al., (2019) or are set to the average of the
 251 respective proxy method where no errors were reported. Standard deviations for TEX₈₆ and
 252 Mg/Ca records are derived from the reported 90% confidence intervals. A minimum value of
 253 2.5°C for the standard deviation is assumed for all other methods. The output variances of the
 254 covariance function are estimated using their maximum likelihood values, obtained with the
 255 GPy Python package (GPy, 2012). Note that the Gaussian process approach provides
 256 probabilistic predictions of temperature values, i.e., uncertainty estimates of the predicted
 257 field. We apply the method to the marine and terrestrial data separately and combine the
 258 masked fields afterwards in order to prevent mutual interference. The uncertainty reported for
 259 an individual GMST estimate is the standard deviation.

260 Model uncertainty (expressed as standard deviation fields) is typically highest in areas
 261 with sparse data coverage (e.g. the Pacific Ocean and Southern Hemisphere land masses;
 262 Figure S1) and the lower uncertainty for the latest Paleocene relative to the PETM and EECO
 263 is partly related to the smaller reported uncertainties in the training data rather than enhanced



264 data coverage. The large spread in reconstructed terrestrial temperatures for North America
 265 during the PETM (Figure S1d) and EECO (Figure Sf) also increases uncertainties for other
 266 continental areas during both time intervals.

267

268 2.1.4. D_{surf-4}

269 For D_{surf-4} , GMST estimates are calculated using a simple mathematical model, tuned to best
 270 fit the proxy data:

271

$$272 \quad T(\theta) \approx a + b\theta + c \cos \theta \quad (3)$$

273

274 where $T(\theta)$ is the Eocene zonal-mean temperature, and the coefficients a , b , and c are chosen
 275 to minimize the sum of the squared residuals relative to D_{surf} (i.e. the SST and LAT data from
 276 Hollis et al. 2019). This model accurately represents $T(\theta)$ in the modern climate (Figure S2)
 277 when supplied with similar number of data points as are in the Hollis et al (2019) dataset, and
 278 it ensures a global solution that is consistent with the physical expectation that temperature
 279 should decrease - and the meridional gradient in temperature should increase - from the
 280 tropics toward the poles (Figure S2).

281 For each data point, we account for three types of uncertainty (i.e. temperature,
 282 elevation, latitude). For temperature, we assume a skew-normal probability distribution based
 283 on the stated 90% confidence intervals. Where uncertainty estimates are not given, we
 284 assume a (symmetric) normal distribution with a 90% confidence interval of $\pm 5K$. For elevation,
 285 we assume a skew-normal distribution with a 90% confidence interval equal to the lowest and
 286 highest elevations of adjacent grid points in the paleotopography data set of Herold et al.
 287 (2014), with a lower bound of zero. For latitude, we assume a uniform distribution spanning
 288 the stated paleomagnetic and mantle estimates.



289 To estimate $T(\theta)$, we randomly sample temperature, elevation, and latitude from their
 290 respective distributions at each location (Figure S3), and apply a lapse-rate adjustment of
 291 6°K/km . Then, using a standard Monte Carlo bootstrapping method, we resample the same
 292 number of data points with replacement, and find the coefficients in Equation 3 that best fit the
 293 sub-sampled data. We repeat this procedure 10,000 times to find a probability distribution of
 294 $T(\theta)$. The uncertainty associated with an individual GMST estimate is the standard deviation.

295

296 2.2. Dataset D_{deep}

297 Dataset D_{deep} consists of bottom water temperatures (BWTs) for the latest Paleocene, PETM
 298 and EECO. Benthic foraminifera $\delta^{18}\text{O}$ values for the latest Paleocene, PETM and EECO come
 299 from previous compilations (Westerhold et al., 2018; Barnett et al., 2019; Cramer et al., 2009),
 300 adjusted to *Cibicidoides* following established methods (Cramer et al., 2009), allowing
 301 temperature to be calculated using Eq. 9 of Marchitto et al (2014):

302

$$303 \quad (\delta_{\text{cp}} - \delta_{\text{sw}} + 0.27) = -0.245 \pm 0.005t + 0.0011 \pm 0.0002t^2 + 3.58 \pm 0.02 \quad (4)$$

304

305 where t is bottom water temperature in Celsius, δ_{cp} is $\delta^{18}\text{O}$ of CaCO_3 on the PeeDee
 306 Belemnite (PDB) scale, and δ_{sw} is $\delta^{18}\text{O}$ of seawater on the Standard Mean Ocean Water
 307 (SMOW). δ_{sw} is defined in accordance with the DeepMIP protocols (-1.00‰ ; see Hollis et al.,
 308 2019). A single method ($D_{\text{deep-1}}$) is used to calculate GMST from D_{deep} following the
 309 methodology outlined in Hansen et al. (2013). For this method, GMST is calculated for: i) the
 310 Early Eocene Climatic Optimum (EECO; 53.3 to 49.4 Ma), ii) the Paleocene-Eocene Thermal
 311 Maximum (ca. 56 Ma) and iii) the latest Paleocene (LP; ca. 57-56 Ma).

312

313 2.2.1. $D_{\text{deep-1}}$



For D_{deep-1} , GMST estimates are calculated following the method of Hansen et al. (2013), which utilises only the deep ocean benthic foraminifera $\delta^{18}O$ dataset, and we refer the reader to that study for a detailed justification of the approach. Briefly, GMST is scaled directly to deep ocean temperature before the Pliocene. Specifically, $\Delta GMST = \Delta BWT$ prior to ~ 5.3 Ma, where early Pliocene BWT and GMST was calculated following Eq. 3.5, 3.6, and 4.2 of Hansen et al. (2013). As such, the calculations presented here differ from those of Hansen et al. (2013) only in that we use a more recent benthic $\delta^{18}O$ compilation and a different equation to convert $\delta^{18}O$ to temperature in the ice-free Paleogene. For each time-slice, the reported uncertainty incorporates the mean calibration uncertainty and standard deviation (1σ) in calculated BWTs.

2.3. Dataset D_{comb}

Dataset D_{comb} uses a combination of (tropical) surface- and deep-water temperature estimates. The deep ocean dataset (D_{deep}) is identical to that described in Section 2.2. The tropical SST dataset utilises all relevant surface ocean proxy data from the DeepMIP database, i.e. those with a palaeolatitude in the magnetic reference frame within 30° of the equator. An expanded definition of the tropics is used as tropical SST reconstructions are relatively sparse; 30° was chosen as it retains tropical SST data from several proxies for all three intervals whilst SST seasonality remains relatively low within these latitudinal bounds.

2.3.1. D_{comb-1}

For D_{comb-1} , GMST estimates are calculated for each time interval based on the difference between tropical SSTs and deep-ocean BWTs (Evans et al., 2018), such that:

$$GMST = 0.5(\overline{tropical\ SST} + \overline{BWT}) \quad (5)$$



339 The fundamental assumptions of this approach are that: 1) GMST can be approximated by
 340 global mean SST, 2) global mean SST is equivalent to the mean of the tropical and high
 341 latitude regions, and 3) benthic temperatures are representative of high latitude surface
 342 temperatures. Applying these assumptions to the modern ocean would generate a GMST
 343 estimate within $\sim 1^\circ\text{C}$ of measured and a modern latitudinal SST gradient within $\sim 1^\circ\text{C}$ of the
 344 surface ocean dataset (as discussed in Evans *et al.*, 2018).

345 Probability distributions for each time interval were computed as follows. In the case
 346 of the tropical SST data, 1000 subsamples were taken, following which a random normally
 347 distributed error was added to each data point in the DeepMIP compilation, including both
 348 calibration uncertainty and variance in the data where multiple reconstructions are available
 349 for a given site and time interval. Mean tropical SST was calculated for each of these
 350 subsamples. To provide a BWT dataset of the same size as the subsampled tropical SST
 351 data, 1000 normally distributed values were calculated for each time interval, based on the
 352 mean $\pm 1\text{SD}$ variation of the pooled benthic $\delta^{18}\text{O}$ data from all sites including calibration
 353 uncertainty.

354

355 3. Results

356 3.1. $D_{\text{surf-1 to -4}}$

357 GMST estimates ($D_{\text{surf-default}}$) during the latest Paleocene ($n = 4$) range between 25.7 and
 358 26.8°C (Table 3). GMST estimates ($D_{\text{surf-default}}$) during the PETM ($n = 4$) range between 31.1
 359 and 33.6°C (Table 3). GMST estimates ($D_{\text{surf-default}}$) during the EECO ($n = 4$) range between
 360 25.4 and 29.0°C (Table 3). All four methods indicate that: 1) the PETM is warmer than the
 361 latest Paleocene (by ~ 4 to 9°C) and: 2) the PETM is warmer than the EECO (by ~ 3 to 7°C).
 362 GMST estimates derived using $D_{\text{surf-Frosty}}$ (i.e. which include planktonic foraminifera $\delta^{18}\text{O}$
 363 values) are consistently lower (up to 3.5°C) than those derived using $D_{\text{surf-default}}$. GMST
 364 estimates derived using $D_{\text{surf-NoTEX}}$ (i.e. which exclude TEX_{86} estimates) are also consistently



lower (up to $\sim 2^{\circ}\text{C}$) than those derived using $D_{\text{surf-}default}$. GMST estimates derived using $D_{\text{surf-}NoMBT}$ (i.e. which exclude MBT/CBT values from marine sediments) are higher than GMST estimates derived using $D_{\text{surf-}default}$ (up to 1°C). GMST estimates derived using $D_{\text{surf-}NoMammal}$ (i.e. which exclude $\delta^{18}\text{O}$ mammal or paleosol estimates) are similar to GMST estimates derived using $D_{\text{surf-}default}$ ($\pm 0.5^{\circ}\text{C}$), with the exception of $D_{\text{surf-}1}$ during the EECO which is $\sim 3^{\circ}\text{C}$ higher when $\delta^{18}\text{O}$ mammal or paleosol values are excluded.

371

3.2. $D_{\text{deep-}1}$

GMST estimates (D_{deep}) during the latest Paleocene, PETM and EECO average 24.3°C ($\pm 1.8^{\circ}\text{C}$), 30.2°C ($\pm 9.2^{\circ}\text{C}$) and 28.0°C ($\pm 2.6^{\circ}\text{C}$), respectively (Table 3). This method indicates that: 1) the PETM is warmer than the latest Paleocene (by $\sim 6^{\circ}\text{C}$) and, 2) the PETM is warmer than the EECO (by $\sim 2^{\circ}\text{C}$).

377

3.3. $D_{\text{comb-}1}$

GMST estimates (D_{comb}) during the latest Paleocene, PETM and EECO average 21.0°C ($\pm 1.7^{\circ}\text{C}$), 26.0°C ($\pm 5.0^{\circ}\text{C}$) and 22.7°C ($\pm 2.3^{\circ}\text{C}$), respectively (Table 3). This method indicates that: 1) the PETM is warmer than the latest Paleocene (by $\sim 5^{\circ}\text{C}$) and, 2) the PETM is warmer than the EECO (by $\sim 3^{\circ}\text{C}$).

383

4. Discussion

4.1. Influence of different proxy datasets upon GMST estimates

To explore the importance of other datasets upon our reconstructed latest Paleocene, PETM and EECO GMST estimates, we conducted a series of subsampling experiments. This was performed for methods $D_{\text{surf-}1}$, -2 , -3 and -4 . In the first subsampling experiment, the inclusion of $\delta^{18}\text{O}$ SST estimates from recrystallized planktonic foraminifera yields lower GMST



estimates (ca. ~ 1 to 3°C ; e.g. Figure 6b). This is consistent amongst all four methods and agrees with previous studies which indicate that $\delta^{18}\text{O}$ values from recrystallized planktonic foraminifera are significantly colder than estimates derived from the $\delta^{18}\text{O}$ value of well-preserved foraminifera, foraminiferal Mg/Ca ratios and clumped isotope values from larger benthic foraminifera (Hollis et al., 2019). The removal of TEX_{86} also results in lower GMST estimates (ca $2\text{--}4^{\circ}\text{C}$; e.g. Figure 6c) across all methodologies. This is consistent with previous studies which indicate that TEX_{86} gives slightly higher SSTs than other proxies, especially in the mid-to-high latitudes (e.g. Hollis et al., 2012; Inglis et al. 2015). This implies that the inclusion of TEX_{86} may lead to a slight warm bias in GMST estimates.

The input of brGDGTs from archives other than mineral soils or peat can bias LAT estimates towards lower values (Inglis et al., 2017; Hollis et al., 2019) and the removal of MBT'/CBT-derived LAT estimates leads to a warm bias in GMST. However, excluding these proxies has a relatively minor impact on GMST ($\sim 0.5^{\circ}\text{C}$). This is because in regions where MBT'/CBT values are discarded (e.g. the SW Pacific), there are other proxies (e.g. pollen assemblages, leaf floral) which yield comparable LAT estimates. The removal of $\delta^{18}\text{O}$ values from paleosols or mammals also leads to a small warm bias in GMST estimates ($\sim 0.5^{\circ}\text{C}$). Intriguingly, $D_{\text{surf}}-1$ yields much higher GMST estimates ($\sim 3^{\circ}\text{C}$ higher than $D_{\text{surf}}\text{-default}$) when $\delta^{18}\text{O}$ values from paleosols or mammals are excluded. This is attributed to the inclusion of two "cold" LAT estimates from the Salta Basin, NW Argentina (Hyland et al., 2017) which overly influence GMST (Figure 6e; Figure 7b-c;). These estimates are derived from the salinization index (SAL) (Sheldon et al., 2002) and the paleosol weathering index (PWI) (Gallagher and Sheldon, 2013), both of which yield a cold bias in the original DeepMIP database (Hollis et al. 2019).

413

4.2. Intercomparison of methods for calculating GMST

For consistency, the following section discusses 'baseline' GMST estimates only. During the latest Paleocene and PETM, GMST estimates derived from D_{surf} average ~ 27 and 32°C ,



respectively (Figure 8). These values agree with previous studies analysing the latest Paleocene (~27°C; Zhu et al., 2019) and PETM (~32°C; Zhu et al., 2019). During the EECO, GMST estimates calculated using D_{surf} range between ~25 and 28°C (Figure 8). These values are comparable to previous estimates from similar time intervals (ca. 29 to 30°C; Huber and Caballero, 2011; Caballero and Huber, 2013; Zhu et al., 2019), but are up to 4°C lower. This cooling can be attributed to two factors. Firstly, our EECO dataset is largely comprised of land air temperature proxy data ($n = 80$ LAT estimates; $n = 27$ SST estimates) which can suffer from a cold bias (Hollis et al., 2019). To explore whether LAT estimates skew GMST estimates towards lower values, we derived GMST using only SST or only LAT data. This analysis was performed using $D_{surf}-1$, -2 and $D_{surf}-4$ and indicates that the GMST estimate are ~2 to 4°C lower when calculated using LAT proxies only. This may be less pronounced in previous studies (i.e. Zhu et al. 2019) because they utilise a different compilation with fewer LAT estimates ($n = 51$; Huber and Caballero, 2011). Secondly, the inclusion of $\delta^{18}\text{O}$ values from paleosols or mammals leads to a cold bias in GMST estimates. For $D_{surf}-1$, a direct comparison of new and prior estimates (Caballero and Huber, 2013) can be made in which the only change has been the use of a newer data compilation. For this new method (Figure 7), the EECO is ~3.5°C colder than previous estimates (29.75°C; Caballero and Huber, 2013). Given that the floristic LAT estimates are identical between the DeepMIP compilation and the older compilation, this strongly suggests that the cooling with respect to older estimates is largely due to the incorporation of paleosol temperature estimates. This suggests that more investigation of the systematic cold bias introduced by paleosols is warranted.

During the latest Paleocene, PETM and EECO, GMST estimates calculated using D_{deep} average ~24°C ($\pm 1.8^\circ\text{C}$), ~30 ($\pm 9.2^\circ\text{C}$) and ~28°C ($\pm 2.6^\circ\text{C}$), respectively (Figure 8). These estimates are comparable to those derived via surface temperature proxies (Table 3). GMST estimates from the EECO are also comparable to previous estimates based on globally distributed benthic foraminifera data (~28°C; Hansen et al., 2013). This implies that benthic foraminiferal $\delta^{18}\text{O}$ values could be used to provide the ‘fine temporal structure’ of Cenozoic



temperature change (Lunt et al., 2016; Hansen et al., 2013). However, we also urge caution as this approach scales GMST directly to BWT prior to the Pliocene, and therefore assumes that the characteristics of polar amplification are constant through time or balanced by other processes. We also note that GMST estimates for the PETM are associated with a large uncertainty. This is due to differences in $\delta^{18}\text{O}$ values between sites and an overall lack of PETM benthic data ($n = 22$ from 3 sites) rather than an inherent uncertainty in the proxy or method of calculating GMST.

During the latest Paleocene, PETM and EECO, GMST estimates calculated using D_{comb} average $\sim 21^\circ\text{C}$ ($\pm 1.7^\circ\text{C}$), ~ 26 ($\pm 5.0^\circ\text{C}$) and $\sim 23^\circ\text{C}$ ($\pm 2.3^\circ\text{C}$), respectively (Figure 8). These estimates are consistently lower (by ~ 2 to 5°C) than GMST estimates derived using D_{surf} ($n = 4$) and D_{deep} ($n = 1$). We suggest this mismatch could be related to two factors. First, if deep water formation preferentially takes place during the winter months, GMST estimates will be biased towards lower values. Secondly, there are relatively few tropical SST estimates during the EECO ($n = 10$ sites), such that D_{comb} may not be fully representative of actual tropical warmth.

4.3. A 'best estimate' of GMST during the latest Paleocene, PETM and EECO

To derive the 'best estimate' of GMST during the latest Paleocene, PETM and EECO, we combine GMST estimates from each 'baseline' experiment (except D_{surf-1} for the EECO which uses $D_{surf-NoMammal}$) and calculate a weighted average (Figure 8). This approach is useful because it assigns lower confidence to GMST estimates associated with larger uncertainties (e.g. D_{deep-1} during the PETM). The reported uncertainty is the reciprocal square root of the sum of all the individual weights. Sequential removal of one time series at a time (jackknife resampling) was performed to examine the influence of a single method upon the average GMST estimate. Jackknifing reveals that that no single method overly influences the mean GMST during the latest Paleocene, PETM or EECO (ca. $\pm 1.0^\circ\text{C}$).



We find that the average GMST estimate for the latest Paleocene, PETM and EECO are 25.7°C ($\pm 0.6^{\circ}\text{C}$), 32.7°C ($\pm 0.8^{\circ}\text{C}$) and 27.3°C ($\pm 0.5^{\circ}\text{C}$), respectively (Figure 8). Assuming a preindustrial GMST of 14°C , our average GMST estimates indicate that the latest Paleocene, PETM and EECO are $+11.7^{\circ}\text{C}$, $+18.7^{\circ}\text{C}$ and $+13.3^{\circ}\text{C}$ warmer than pre-industrial, respectively. The GMST anomaly for the EECO is skewed to cooler values than previous work ($\sim 15^{\circ}\text{C}$ warmer than pre-industrial; Caballero and Huber, 2013; Zhu et al., 2019) but lies within the range quoted previously in the IPCC AR5 (12.7°C warmer than pre-industrial). On average, GMST increases by ~ 6 to 7°C between the latest Paleocene and PETM, in keeping with previous estimates (Frieling et al., 2019; Dunkley Jones, 2013). The PETM is approximately 5°C warmer than the EECO. This is higher than previously suggested ($\sim 3^{\circ}\text{C}$; Zhu et al., 2019) and may related to a cold bias in EECO GMST estimates (see Section 4.2).

4.4. Equilibrium climate sensitivity during the latest Palaeocene, PETM and EECO

Equilibrium climate sensitivity (ECS) can be defined as the equilibrium change in global near surface air temperature, resulting from a doubling in atmospheric CO_2 . Various “flavours” of ECS exist, some of which specifically exclude various feedback processes not always included in climate models, such as those associated with ice sheets, vegetation, or aerosols (Rohling et al., 2012). ECS may also be state-dependent (Caballero and Huber, 2013) and there is no reason to expect it has not changed with time. Therefore, direct comparison of ECS in the past to modern conditions is a fraught enterprise. For our purposes we define a ‘bulk ECS’ as being a gross estimate of ECS across time between our three intervals and preindustrial. Such calculations have been performed previously (Shaffer et al., 2016; Anagnostou et al., 2016) and they provide some constraint on the range of climate sensitivity values that are relevant for near-modern prediction (Rohling et al., 2012). For example, Anagnostou et al. (2016) indicated that early Eocene ECS (excluding ice sheet feedbacks) falls within the range 2.1 – 4.6°C per CO_2 doubling with maximum probability for the EECO of 3.8°C . These values (2.1 – 4.6°C per CO_2 doubling) are similar to the IPCC ECS range (1.5 – 4.5°C at 66% confidence).



Here we calculate bulk ECS estimates using the change in GMST and CO₂ in the latest Paleocene, PETM and EECO intervals with reference to the pre-industrial. Following the approach of Anagnostou et al. (2016) and using the forcing equation of Byrne and Goldblatt (2014), we first determine the relative change in climate forcing relative to pre-industrial ($\Delta F_{\text{CO}_2-\text{vs-PI}}$):

502

$$\Delta F_{\text{CO}_2-\text{vs-PI}} = 5.32 \ln(C_i/C_{\text{PI}}) + (0.39 [\ln(C_i/C_{\text{PI}})]^2) \quad [6]$$

504

where C_{PI} is the atmospheric CO₂ concentration during pre-industrial (278 ppm) and C_i refers to the CO₂ reconstruction at a particular time in the Eocene. The mean proxy estimate of CO₂ for the PETM is ~2200 ppmv (+1904/-699 ppmv; Gutjahr et al., 2017). The mean proxy estimate of CO₂ for the LP is ~870 ppmv (Gutjahr et al., 2017; n.b. no published uncertainty available; here we assign an uncertainty of ±400ppm). The mean proxy estimate of CO₂ for the EECO is ~1625 ppmv (±750 ppmv) (Anagnostou et al., 2016; Hollis et al., 2019). To calculate bulk ECS, we then use radiative forcing from a doubling of CO₂ from Byrne and Goldblatt (2014) to translate CO₂ into forcing relative to preindustrial (ΔF_{CO_2}):

513

$$\text{ECS} = (\Delta \text{GMST}) / \Delta F_{\text{CO}_2-\text{vs-PI}} * 3.875 \quad [7]$$

515

Some of the temperature anomaly of the latest Paleocene, PETM, and EECO is caused not by CO₂ but by the different paleotopography, paleobathymetry, and solar constant compared with preindustrial. Furthermore, we choose here to calculate an ECS that explicitly excludes feedbacks associated with vegetation, ice sheets, and aerosols, i.e. $S_{[\text{CO}_2, \text{LI}, \text{VG}, \text{AE}]}$ in the nomenclature of Rohling et al (2012). To account for these effects, we subtract a value of 4.5°C (Caballero and Huber, 2013; Zhu et al. 2019) from the GMST in Table 3. This value of 4.5°C is based upon a comparison of preindustrial and Eocene simulations (both 1x CO₂) conducted



with CESM1.2 (Zhu et al., 2019), which incorporates the paleogeographic, solar constant, ice sheet, vegetation, aerosol, and ice sheet changes from preindustrial to Eocene. This value is similar to previous studies which attribute ~4 to 6°C to the non-CO₂ forcings and feedbacks (Anagnostou et al., 2016; Caballero and Huber, 2013, Lunt et al., 2012). However, we note that the sensitivity to these Eocene boundary conditions is likely model-dependant and this value will likely differ between model simulations. The uncertainties in our estimated ECS are the products of 10,000 realizations of the latest Paleocene, PETM and EECO CO₂ values and the respective ΔGMST estimate (the mean estimate and uncertainty in Table 3) based on randomly sampling each variable within its 95% confidence interval uncertainty envelope

We estimate $S_{[CO_2, LI, VG, AE]}$ for the latest Paleocene, EECO and PETM to range between 0.73 and 1.12 (66% confidence; Figure 9). This yields bulk ECS estimates for the latest Paleocene, EECO and PETM compared to modern which range between 2.8 to 4.8 °C per doubling CO₂ (66% confidence). These values are comparable to previous estimates from the early Eocene which also account for paleogeography and other feedbacks (~2.1 to 4.6°C; Anagnostou et al., 2016) and fall within the modern ECS range predicted by the IPCC (1.5 to 4.5°C per doubling CO₂). However, care must be exercised when relating geological estimates to modern climate predictions (e.g. Rohling et al., 2012). In addition, published CO₂ estimates remain uncertain (especially during the latest Paleocene and PETM) and new high-fidelity records are required to accurately constrain ECS during these super warm climates.

542

543 5. Conclusions

Using six different methods, we have quantified global mean surface temperatures (GMST) during the latest Paleocene, PETM and EECO. GMST was calculated within a coordinated, experimental framework and utilised three different input datasets. After evaluating the impact of different proxy datasets upon GMST estimates, we combined all six methodologies to derive an average GMST value during the latest Paleocene, PETM and EECO. Our results indicate high GMSTs during the latest Paleocene (25.7°C ± 0.6°C), PETM (32.7°C ± 0.8°C) and EECO



550 (27.3°C ± 0.5°C). Assuming a preindustrial GMST of 14°C, our average GMST estimates for
551 the latest Paleocene, PETM and EECO are 11.7°C, 18.7°C and 13.3°C higher than pre-
552 industrial, respectively. Using our 'combined' GMST estimates, we then estimated a bulk ECS
553 during the latest Paleocene, PETM and EECO. Our results range between 2.8 to 4.8°C (at
554 66% confidence) per doubling of atmospheric CO₂ when feedbacks associated with ice
555 sheets, vegetation, and aerosols are accounted for. Taken together, our study improves our
556 characterisation of the global mean temperature of these key time periods, allowing future
557 climate change to be put into the context of past changes, and allowing us to provide a refined
558 estimate of ECS.

559

560 **Data availability**

561 Data can be accessed via the online supporting information, via www.pangaea.de/, or from
562 the author (email: gordon.inglis@soton.ac.uk).

563

564 **Authorship tiers and contributions**

565 Authorship of this manuscript is organized into three tiers according to the contributions of
566 each individual author. Inglis (Tier I) organized the structure and writing of the manuscript,
567 contributed to all sections of the text and designed the figures. Tier II authors (listed
568 alphabetically following Inglis) assumed a leading role by contributing methodologies used in
569 the text. Tier III authors (listed alphabetically following Wilkinson) contributed intellectually to
570 the text and figure design.

571

572 **Declaration of competing interest**

573 The authors declare that they have no known competing financial interests or personal
574 relationships that could have appeared to influence the work reported in this paper.

575



576 Acknowledgements

577 This research was funded from NERC through NE/P01903X/1 and NE/N006828/1, both of
 578 which supported GNI, DL, SS and RDP. GNI was also supported by a Royal Society Dorothy
 579 Hodgkin Fellowship. N.J.B. is supported by NSF AGS-1844380. FB, DL, and RDW were
 580 funded by the EPSRC 'Past Earth Network'. MH was funded by NSF OPP 1842059.

581

582 References

- 583 Anagnostou, E., John, E. H., Edgar, K. M., Foster, G. L., Ridgwell, A., Inglis, G. N., Pancost,
 584 R. D., Lunt, D. J., and Pearson, P. N.: Changing atmospheric CO₂ concentration was
 585 the primary driver of early Cenozoic climate, *Nature*, 533, 380-384,
 586 10.1038/nature17423, 2016.
- 587 Barnet, J. S., Littler, K., Westerhold, T., Kroon, D., Leng, M. J., Bailey, I., Röhl, U., and Zachos,
 588 J. C.: A high-Fidelity benthic stable isotope record of late Cretaceous–early Eocene
 589 climate change and carbon-cycling, *Paleoceanography & Paleoclimatology*, 34, 672-
 590 691, 2019.
- 591 Bemis, B. E., Spero, H. J., Bijma, J., and Lea, D. W.: Reevaluation of the oxygen isotopic
 592 composition of planktonic foraminifera: Experimental results and revised
 593 paleotemperature equations, *Paleoceanography & Paleoclimatology*, 13, 150-160,
 594 10.1029/98pa00070, 1998.
- 595 Bijl, P. K., Schouten, S., Sluijs, A., Reichert, G.-J., Zachos, J. C., and Brinkhuis, H.: Early
 596 Palaeogene temperature evolution of the southwest Pacific Ocean, *Nature*, 461, 776-
 597 779,
 598 http://www.nature.com/nature/journal/v461/n7265/supinfo/nature08399_S1.html,
 599 2009.
- 600 Bragg, F. J., Paine, P., Saul, A., Lunt, D. J., Wilkinson, R., and Zammit-Mangion, A.: A
 601 Statistical Algorithm for Evaluating Palaeoclimate Simulations Against Geological
 602 Observations, *Geoscientific Model Development*, Submitted.



- 603 Byrne, B., and Goldblatt, C.: Radiative forcing at high concentrations of well-mixed
 604 greenhouse gases, *Geophysical Research Letters*, 41, 152-160, 2014.
- 605 Caballero, R., and Huber, M.: State-dependent climate sensitivity in past warm climates and
 606 its implications for future climate projections, *Proceedings of the National Academy of*
 607 *Sciences*, 110, 14162-14167, 2013.
- 608 Cramer, B. S., Toggweiler, J. R., Wright, J. D., Katz, M. E., and Miller, K. G.: Ocean overturning
 609 since the Late Cretaceous: Inferences from a new benthic foraminiferal isotope
 610 compilation, *Paleoceanography & Paleoclimatology*, 24, 10.1029/2008pa001683,
 611 2009.
- 612 Cramwinckel, M. J., Huber, M., Kocken, I. J., Agnini, C., Bijl, P. K., Bohaty, S. M., Frieling, J.,
 613 Goldner, A., Hilgen, F. J., Kip, E. L., Peterse, F., van der Ploeg, R., Rohl, U., Schouten,
 614 S., and Sluijs, A.: Synchronous tropical and polar temperature evolution in the Eocene,
 615 *Nature*, 559, 382, 2018.
- 616 Evans, D., Sagoo, N., Renema, W., Cotton, L. J., Müller, W., Todd, J. A., Saraswati, P. K.,
 617 Stassen, P., Ziegler, M., Pearson, P. N., Valdes, P. J., and Affek, H. P.: Eocene
 618 greenhouse climate revealed by coupled clumped isotope-Mg/Ca thermometry,
 619 *Proceedings of the National Academy of Sciences*, 115, 1174-1179,
 620 10.1073/pnas.1714744115 %J *Proceedings of the National Academy of Sciences*,
 621 2018.
- 622 Farnsworth, A., Lunt, D., O'Brien, C., Foster, G., Inglis, G., Markwick, P., Pancost, R., and
 623 Robinson, S.: Climate sensitivity on geological timescales controlled by non-linear
 624 feedbacks and ocean circulation, *Geophysical Research Letters*, 2019.
- 625 Gallagher, T. M., and Sheldon, N. D.: A new paleothermometer for forest paleosols and its
 626 implications for Cenozoic climate, *Geology*, 41, 647-650, 10.1130/G34074.1, 2013.
- 627 Gutjahr, M., Ridgwell, A., Sexton, P. F., Anagnostou, E., Pearson, P. N., Pälike, H., Norris, R.
 628 D., Thomas, E., and Foster, G. L.: Very large release of mostly volcanic carbon during
 629 the Palaeocene–Eocene Thermal Maximum, *Nature*, 548, 573-577,
 630 10.1038/nature23646, 2017.



- 631 Hansen, J., Sato, M., Russell, G., and Kharecha, P.: Climate sensitivity, sea level and
 632 atmospheric carbon dioxide, *Philosophical Transactions of the Royal Society A:*
 633 *Mathematical, Physical Engineering Sciences*, 371, 20120294, 2013.
- 634 Hegerl, G. C., Zwiers, F. W., Braconnot, P., Gillett, N. P., Luo, Y., Marengo Orsini, J., Nicholls,
 635 N., Penner, J. E., and Stott, P. A.: Understanding and attributing climate change, IPCC,
 636 2007: *Climate Change 2007: the physical science basis. contribution of Working Group*
 637 *I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*,
 638 2007.
- 639 Herold, N., Buzan, J., Seton, M., Goldner, A., Green, J. A. M., Müller, R. D., Markwick, P., and
 640 Huber, M.: A suite of early Eocene (~ 55 Ma) climate model boundary conditions,
 641 *Geoscientific Model Development*, 7, 2077-2090, 10.5194/gmd-7-2077-2014, 2014.
- 642 Hollis, C. J., Taylor, K. W. R., Handley, L., Pancost, R. D., Huber, M., Creech, J. B., Hines, B.
 643 R., Crouch, E. M., Morgans, H. E. G., Crampton, J. S., Gibbs, S., Pearson, P. N., and
 644 Zachos, J. C.: Early Paleogene temperature history of the Southwest Pacific Ocean:
 645 Reconciling proxies and models, *Earth and Planetary Science Letters*, 349–350, 53–
 646 66, <http://dx.doi.org/10.1016/j.epsl.2012.06.024>, 2012.
- 647 Hollis, C. J., Dunkley Jones, T., Anagnostou, E., Bijl, P. K., Cramwinckel, M. J., Cui, Y.,
 648 Dickens, G. R., Edgar, K. M., Eley, Y., Evans, D., Foster, G. L., Frieling, J., Inglis, G.
 649 N., Kennedy, E. M., Kozdon, R., Lauretano, V., Lear, C. H., Littler, K., Lourens, L.,
 650 Meckler, A. N., Naafs, B. D. A., Pälike, H., Pancost, R. D., Pearson, P. N., Röhl, U.,
 651 Royer, D. L., Salzmann, U., Schubert, B. A., Seebeck, H., Sluijs, A., Speijer, R. P.,
 652 Stassen, P., Tierney, J., Tripathi, A., Wade, B., Westerhold, T., Witkowski, C., Zachos,
 653 J. C., Zhang, Y. G., Huber, M., and Lunt, D. J.: The DeepMIP contribution to PMIP4:
 654 methodologies for selection, compilation and analysis of latest Paleocene and early
 655 Eocene climate proxy data, incorporating version 0.1 of the DeepMIP database,
 656 *Geoscientific Model Development*, 12, 3149-3206, 10.5194/gmd-12-3149-2019, 2019.
- 657 Huber, M., and Caballero, R.: The early Eocene equable climate problem revisited, *Clim. Past*,
 658 7, 603-633, 10.5194/cp-7-603-2011, 2011.



- 659 Hyland, E. G., Sheldon, N. D., and Cotton, J. M.: Constraining the early Eocene climatic
 660 optimum: A terrestrial interhemispheric comparison, *GSA Bulletin*, 129, 244-252,
 661 10.1130/B31493.1, 2017.
- 662 Inglis, G. N., Farnsworth, A., Lunt, D., Foster, G. L., Hollis, C. J., Pagani, M., Jardine, P. E.,
 663 Pearson, P. N., Markwick, P., Galsworthy, A. M. J., Raynham, L., Taylor, K. W. R., and
 664 Pancost, R. D.: Descent toward the Icehouse: Eocene sea surface cooling inferred
 665 from GDGT distributions, *Paleoceanography*, 30, 1000-1020, 10.1002/2014pa002723,
 666 2015.
- 667 Inglis, G. N., Collinson, M. E., Riegel, W., Wilde, V., Farnsworth, A., Lunt, D. J., Valdes, P.,
 668 Robson, B. E., Scott, A. C., Lenz, O. K., Naafs, B. D. A., and Pancost, R. D.: Mid-
 669 latitude continental temperatures through the early Eocene in western Europe, *Earth*
 670 *and Planetary Science Letters*, 460, 86-96, <https://doi.org/10.1016/j.epsl.2016.12.009>,
 671 2017.
- 672 Keating-Bitonti, C. R., Ivany, L. C., Affek, H. P., Douglas, P., and Samson, S. D.: Warm, not
 673 super-hot, temperatures in the early Eocene subtropics, *Geology*, 39, 771-774, 2011.
- 674 Knutti, R., Rugenstein, M. A., and Hegerl, G. C.: Beyond equilibrium climate sensitivity, *Nature*
 675 *Geoscience*, 10, 727-736, 2017.
- 676 Lunt, D. J., Jones, T. D., Heinemann, M., Huber, M., LeGrande, A., Winguth, A., Loptson, C.,
 677 Marotzke, J., Roberts, C., and Tindall, J.: A model-data comparison for a multi-model
 678 ensemble of early Eocene atmosphere-ocean simulations: EoMIP, *Climate of the Past*,
 679 8, 2012.
- 680 Lunt, D. J., Farnsworth, A., Loptson, C., Foster, G. L., Markwick, P., O'Brien, C. L., Pancost,
 681 R. D., Robinson, S. A., and Wrobel, N.: Palaeogeographic controls on climate and
 682 proxy interpretation, *Climate of the Past*, 12, 1181-1198, 2016.
- 683 Lunt, D. J., Huber, M., Anagnostou, E., Baatsen, M. L. J., Caballero, R., DeConto, R., Dijkstra,
 684 H. A., Donnadieu, Y., Evans, D., Feng, R., Foster, G. L., Gasson, E., von der Heydt,
 685 A. S., Hollis, C. J., Inglis, G. N., Jones, S. M., Kiehl, J., Kirtland Turner, S., Korty, R.
 686 L., Kozdon, R., Krishnan, S., Ladant, J. B., Langebroek, P., Lear, C. H., LeGrande, A.



- 687 N., Littler, K., Markwick, P., Otto-Bliesner, B., Pearson, P., Poulsen, C. J., Salzmann,
 688 U., Shields, C., Snell, K., Stärz, M., Super, J., Tabor, C., Tierney, J. E., Tourte, G. J.
 689 L., Tripathi, A., Upchurch, G. R., Wade, B. S., Wing, S. L., Winguth, A. M. E., Wright, N.
 690 M., Zachos, J. C., and Zeebe, R. E.: The DeepMIP contribution to PMIP4: experimental
 691 design for model simulations of the EECO, PETM, and pre-PETM (version 1.0),
 692 Geoscientific Model Development, 10, 889-901, 10.5194/gmd-10-889-2017, 2017.
- 693 Marchitto, T., Curry, W., Lynch-Stieglitz, J., Bryan, S., Cobb, K., and Lund, D.: Improved
 694 oxygen isotope temperature calibrations for cosmopolitan benthic foraminifera,
 695 Geochimica et Cosmochimica Acta, 130, 1-11, 2014.
- 696 Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., Gonzalez Rouco,
 697 J. F., Jansen, E., Lambeck, K., Luterbacher, J., Naish, T., Osborn, T., Otto-Bliesner,
 698 B., Quinn, T., Ramesh, R., Rojas, M., Shao, X., and Timmermann, A.: Information from
 699 Paleoclimate Archives, in: Climate Change 2013 – The Physical Science Basis:
 700 Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental
 701 Panel on Climate Change, Cambridge University Press, Cambridge, 383-464, 2014.
- 702 Pearson, P. N., Ditchfield, P. W., Singano, J., Harcourt-Brown, K. G., Nicholas, C. J., Olsson,
 703 R. K., Shackleton, N. J., and Hall, M. A.: Warm tropical sea surface temperatures in
 704 the Late Cretaceous and Eocene epochs, Nature, 413, 481-487, 2001.
- 705 Pearson, P. N., van Dongen, B. E., Nicholas, C. J., Pancost, R. D., Schouten, S., Singano, J.
 706 M., and Wade, B. S.: Stable warm tropical climate through the Eocene Epoch,
 707 Geology, 35, 211-214, 10.1130/g23175a.1, 2007.
- 708 Rohling, E. J., Sluijs, A., Dijkstra, H. A., Köhler, P., van de Wal, R. S. W., von der Heydt, A.
 709 S., Beerling, D. J., Berger, A., Bijl, P. K., Crucifix, M., DeConto, R., Drijfhout, S. S.,
 710 Fedorov, A., Foster, G. L., Ganopolski, A., Hansen, J., Hönlisch, B., Hooghiemstra, H.,
 711 Huber, M., Huybers, P., Knutti, R., Lea, D. W., Lourens, L. J., Lunt, D., Masson-
 712 Delmotte, V., Medina-Elizalde, M., Otto-Bliesner, B., Pagani, M., Pälike, H., Renssen,
 713 H., Royer, D. L., Siddall, M., Valdes, P., Zachos, J. C., Zeebe, R. E., and Members, P.



- 714 P.: Making sense of palaeoclimate sensitivity, *Nature*, 491, 683-691,
 715 10.1038/nature11574, 2012.
- 716 Shaffer, G., Huber, M., Rondanelli, R., and Pedersen, J. O. P. J. G. R. L.: Deep time evidence
 717 for climate sensitivity increase with warming, 43, 6538-6545, 2016.
- 718 Sheldon, Nathan D., Retallack, Gregory J., and Tanaka, S.: Geochemical Climofunctions from
 719 North American Soils and Application to Paleosols across the Eocene-Oligocene
 720 Boundary in Oregon, *The Journal of Geology*, 110, 687-696, 10.1086/342865, 2002.
- 721 Stevens, B., Sherwood, S. C., Bony, S., and Webb, M. J.: Prospects for narrowing bounds on
 722 Earth's equilibrium climate sensitivity, *Earth's Future*, 4, 512-522, 2016.
- 723 Tierney, J. E., and Tingley, M. P.: A Bayesian, spatially-varying calibration model for the
 724 TEX86 proxy, *Geochimica et Cosmochimica Acta*, 127, 83-106, 2014.
- 725 Westerhold, T., Röhl, U., Donner, B., and Zachos, J. C.: Global extent of early Eocene
 726 hyperthermal events: A new Pacific benthic foraminiferal isotope record from Shatsky
 727 Rise (ODP Site 1209), *Paleoceanography & Paleoclimatology*, 33, 626-642, 2018.
- 728 Zhu, J., Poulsen, C. J., and Tierney, J. E.: Simulation of Eocene extreme warmth and high
 729 climate sensitivity through cloud feedbacks, *Science Advances*, 5, eaax1874, 2019.
- 730
- 731
- 732
- 733
- 734
- 735
- 736
- 737
- 738



Label in Fig. 1	Source	Time window	GMST (°C)	Uncertainty	Proxy system
1a	Farnsworth et al. (2019)	EE	23.4	±3.2	δ ¹⁸ O planktonic
1b	Farnsworth et al. (2019)	EE	37.1	±1.4	δ ¹⁸ O planktonic + TEX ₈₆
2a	Zhu et al. (2019)	LP	27	n/a	Multiple
2b	Zhu et al. (2019)	EEO	29	±3	Multiple
2c	Zhu et al. (2019)	PETM	32	n/a	Multiple
3	Caballero and Huber (2013)	EE	29.5	±2.6	Multiple
4	Hansen et al (2013)	EE	28	n/a	δ ¹⁸ O benthic
5	Cramwinckel et al. (2018)	EE	29.3	n/a	Multiple

739

740 **Table 1:** Previous studies that have determined GMST for the early Eocene (EE), EEO,
 741 PETM or latest Paleocene (LP). n/a indicates that no error bars were reported in the original
 742 publications.

743

744

745

746

747

748

749

750

751

752

753

754

755



756

Experiment	Description
$D_{surf}^{default}$	All SST and LAT data compiled in Hollis et al. (2019) but excluding recrystallized planktonic foraminifera $\delta^{18}O$ values
D_{surf}^{Frosty}	$D_{surf}^{default}$ but including recrystallized planktonic foraminifera $\delta^{18}O$ values
D_{surf}^{NoTEX}	$D_{surf}^{default}$ but excluding TEX_{86} values
D_{surf}^{NoMBT}	$D_{surf}^{default}$ but excluding MBT(')/CBT values from marine sediments
$D_{surf}^{NoMammal}$	$D_{surf}^{default}$ but excluding mammal and paleosol $\delta^{18}O$ values

757 **Table 2:** Default and optional subsampling experiments applied to D_{surf}

758

759

760

761

762

763

764

765

766

767

768

769

770

771

772



773

GMST (°C)							
	D _{surf-1}	D _{surf-2}	D _{surf-3}	D _{surf-4}	D _{deep-1}	D _{comb-1}	Average
LP	25.9 (±1.0)	26.8 (±1.2)	25.7 (±6.0)	27.6 (±1.3)	24.3 (±1.1)	21.0 (±1.7)	25.7 (±0.6)
PETM	33.6 (±1.2)	33.4 (±1.6)	31.2 (±7.6)	31.3 (±1.6)	30.2 (±9.2)	26.0 (±5.0)	32.7 (±0.8)
EECO	26.3 (±0.7)	26.7 (±0.9)	27.9 (±7.0)	25.4 (±1.1)	28.0 (±2.6)	22.7 (±2.3)	27.3 (±0.5)

774

775 **Table 3:** GMST for latest Paleocene (LP), PETM and EECO. Reported GMST estimates utilise
 776 ‘baseline’ experiments except D_{surf-1} during the EECO which uses $D_{surf-NoMammal}$.

777

778

779

780

781

782

783

784

785

786

787

788

789

790



	ECS (°C) (66% confidence)	ECS (°C) (95% confidence)
Latest Paleocene	3.9 – 4.8	3.6 – 5.5
PETM	3.5 – 4.4	3.2 – 5.5
EECO	2.8 – 3.8	2.6 – 5.2

791

792 **Table 4:** Estimates of ECS (66% and 95% confidence) during the latest Paleocene, PETM
 793 and EECO.

794

795

796

797

798

799

800

801

802

803

804

805

806

807

808

809

810



811 **Figure captions:**

812 **Figure 1:** Published GMST estimates during the early Paleogene (57 to 48 Ma). Dots
 813 represent average values. The horizontal limits on the individual dots represent the reported
 814 error. y-Axis labels refer to previous estimates (see Table 1).

815

816 **Figure 2:** Location of proxies within the surface temperature dataset (D_{surf}). A) SST proxies
 817 with time intervals indicated as followed: black circles, all three-time intervals represented.
 818 Red circles: PETM \pm latest Paleocene intervals; orange circles, EECO interval (b) Terrestrial
 819 sites with time intervals indicated as in (a) and green circles, LP only.

820

821 **Figure 3:** An illustration of Method D_{surf-2} for 2 sites: (a) Tanzania in the EECO as diagnosed
 822 using HadCM3L, and (b) Mid Waipara in the PETM as diagnosed using CCSM3. The vertical
 823 dashed line shows $\langle T \rangle^{inferred}$ and the horizontal dashed line shows T^{proxy} , which intercept at
 824 the orange dot. The dark blue dots show the intercept of T^{low} with $\langle T^{low} \rangle$, and the red dots
 825 show the intercept of T^{high} with $\langle T^{high} \rangle$.

826

827 **Figure 4:** Inferred global mean temperature ($\langle T \rangle^{inferred}$) for each EECO-aged proxy in the
 828 DeepMIP database using D_{surf-2} , as diagnosed using CCSM3. The final estimate of global
 829 mean temperature is the average of all the individual sites.

830

831 **Figure 5:** Predicted surface warming by Gaussian process regression using D_{surf-3} for the (a)
 832 latest Paleocene, (b) PETM and (c) EECO. Anomalies are relative to the present-day zonal
 833 mean surface temperature. Circles indicate all available SST and LAT proxy data for the
 834 respective time slice that were used to train the model. Circles for locations where multiple
 835 proxy reconstructions are available are slightly shifted in latitude for improved visibility.



836

837 **Figure 6:** Predicted surface warming by Gaussian process regression using D_{surf-3} for the
 838 EECO for the five core experiments (see Table 2). Anomalies are relative to the present-day
 839 zonal mean surface temperature. Circles indicate all available SST and LAT proxy data for the
 840 respective time slice and experiment that were used to train the model. Circles for locations
 841 where multiple proxy reconstructions are available are slightly shifted in latitude for improved
 842 visibility.

843

844 **Figure 7:** An illustration of Method D_{surf-1} during the EECO. (a) Modelled early Eocene (2240
 845 ppm) temperatures utilising CCSM3 (b) Interpolated absolute SST reconstructions, (c) Data-
 846 model difference between (a) and (b).

847

848 **Figure 8:** Summary of GMST estimates for the (a) latest Paleocene, (b) Paleocene-Eocene
 849 Thermal Maximum and (c) early Eocene Climatic Optimum. Error bars on each individual
 850 method are the standard deviation, except D_{surf-1} and D_{surf-2} which use the standard error.
 851 Error bar on weighted average is the reciprocal square root of the sum of all the individual
 852 weights.

853

854 **Figure 9:** Probability density function of bulk ECS during the latest Paleocene, PETM and
 855 EECO that explicitly accounts for non- CO_2 forcings of palaeogeography and solar constant, and
 856 feedbacks associated with land ice, vegetation, and aerosols (Zhu et al., 2019), i.e.
 857 $S_{[\text{CO}_2, \text{LI}, \text{VG}, \text{AE}]}$ in the nomenclature of Rohling et al (2012).

858



Figure 1

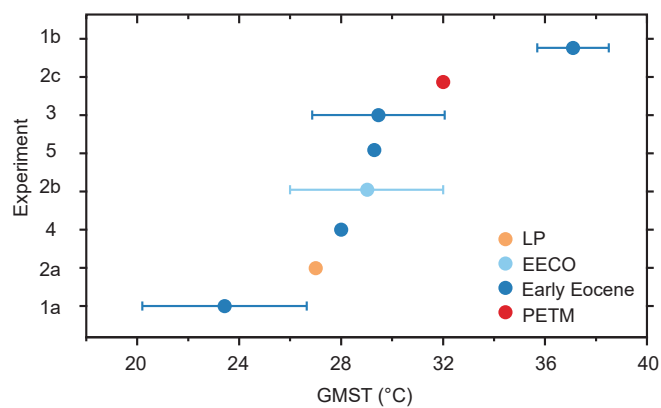




Figure 2

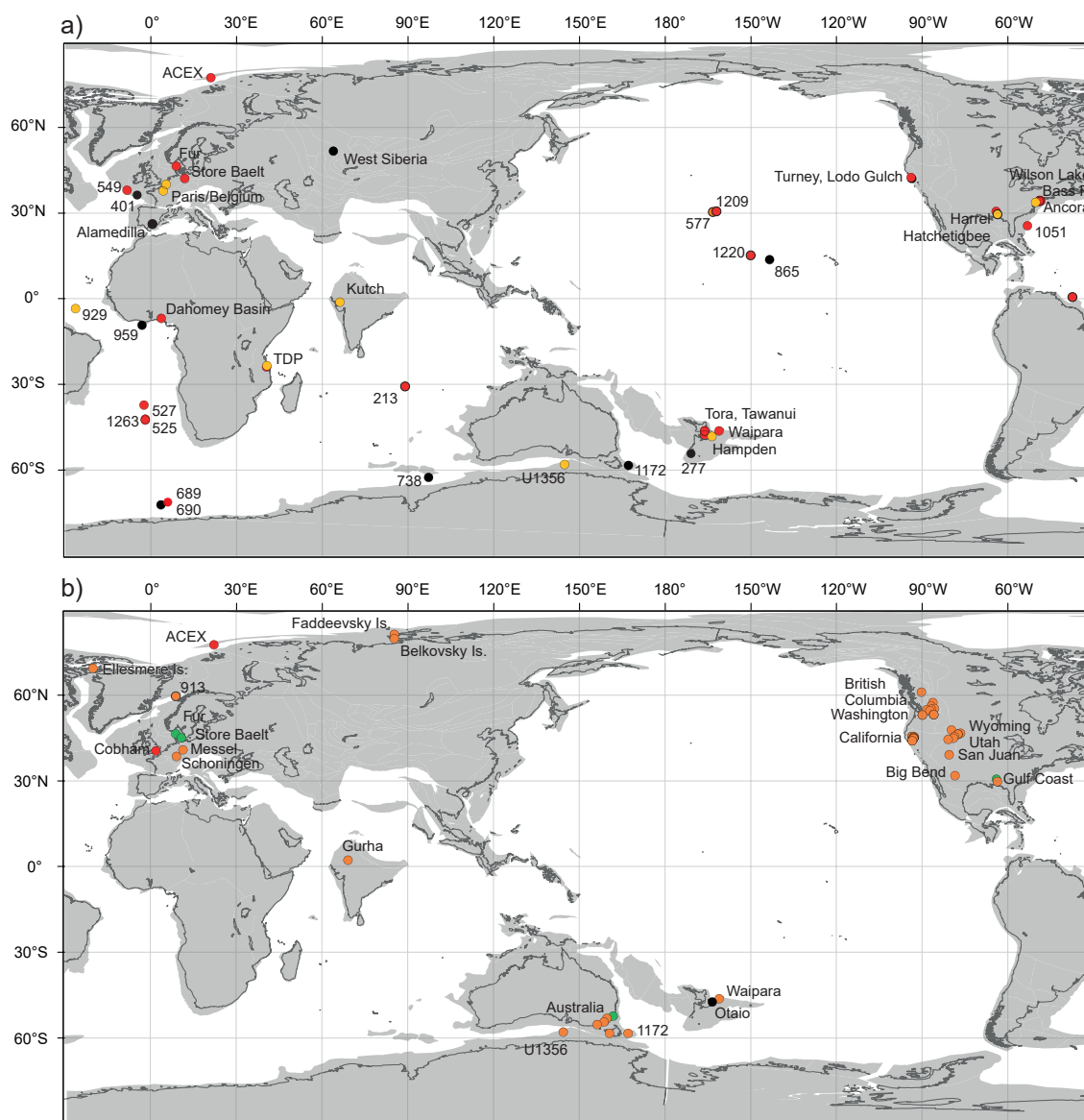




Figure 3

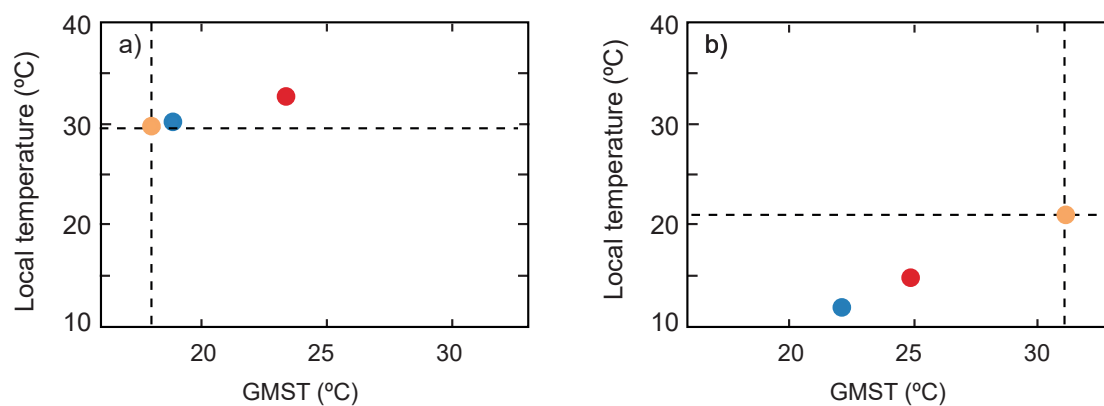




Figure 4

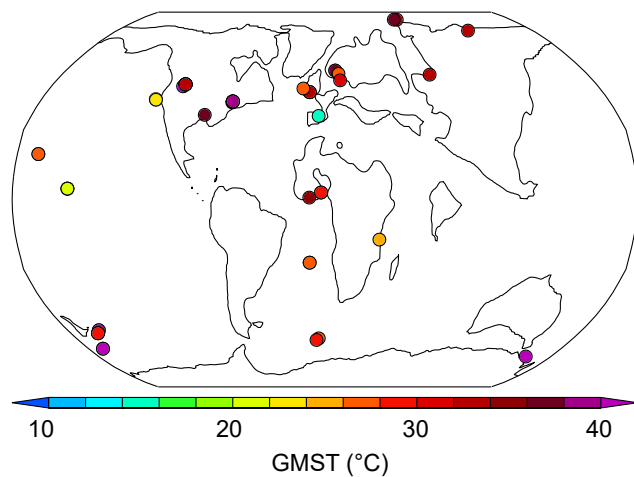




Figure 5

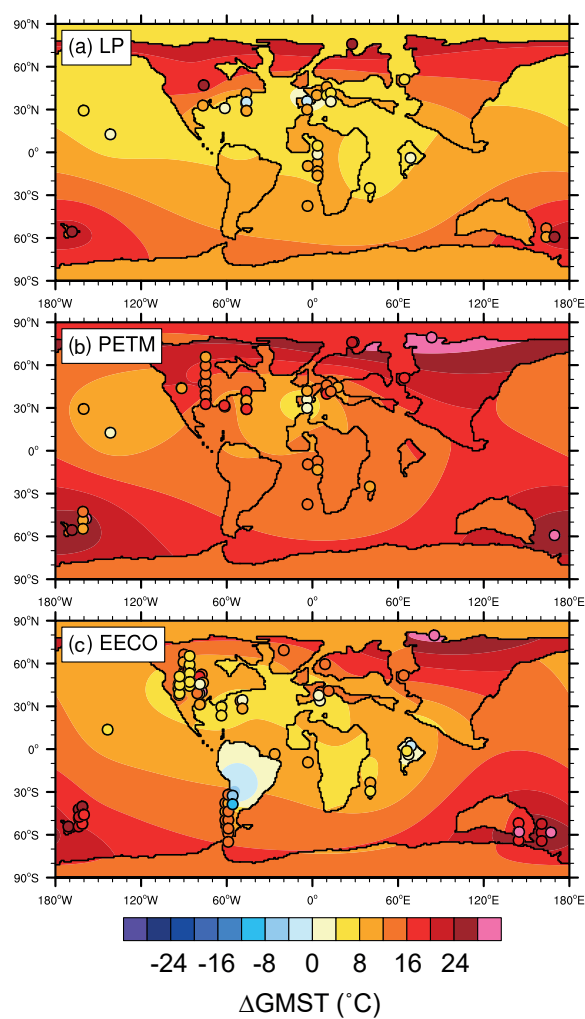




Figure 6

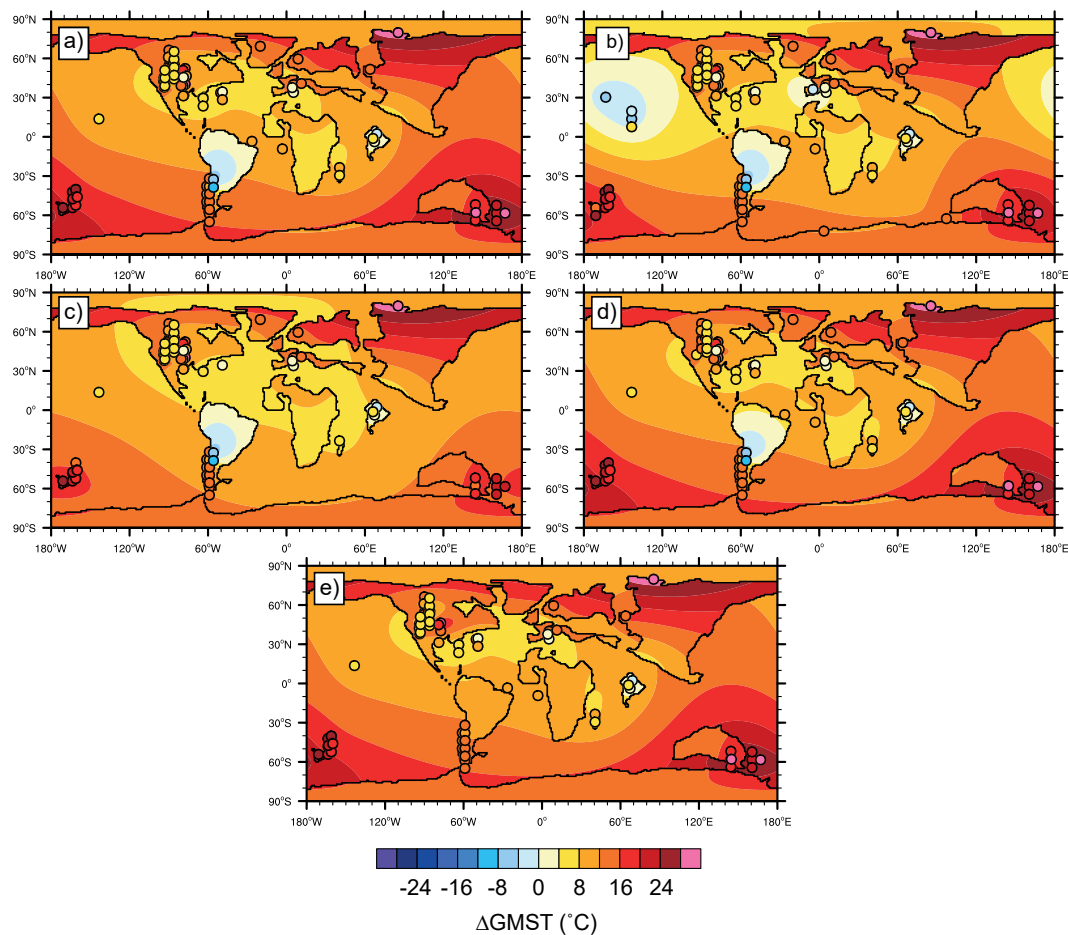




Figure 7

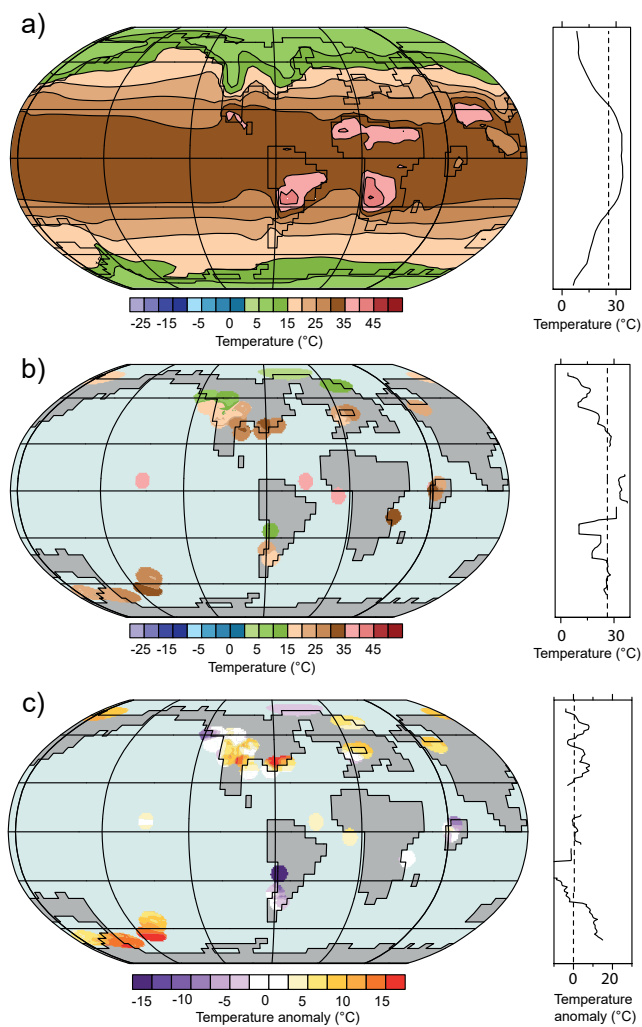




Figure 8

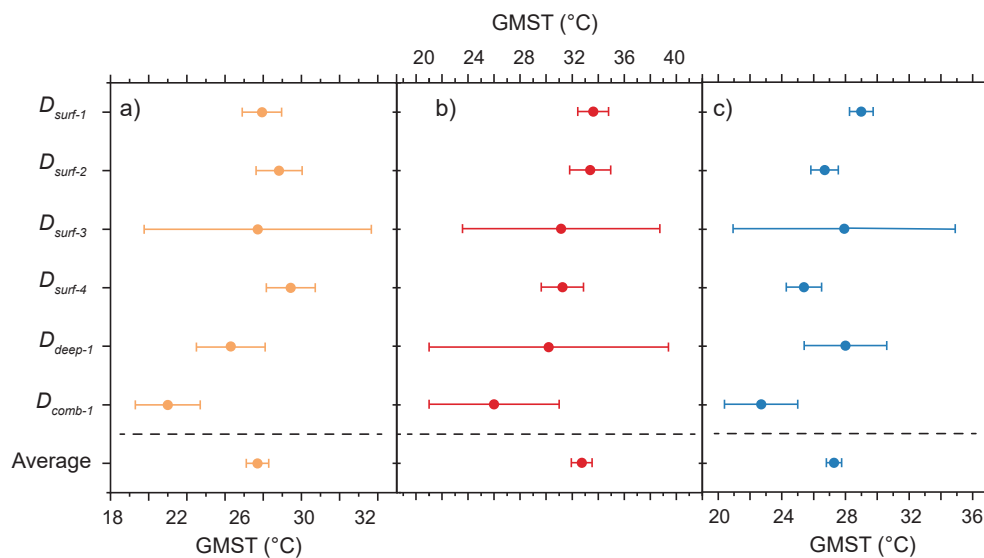




Figure 9

