Global mean surface temperature and climate sensitivity of the EECO, PETM and latest Paleocene

Gordon N. Inglis[1,2], Fran Bragg[3], Natalie Burls[4], Margot J. Cramwinckel[5], David Evans[6], Gavin L. Foster[1], Matt Huber[7], Daniel J. Lunt[3], Nicholas Siler[8], Sebastian Steinig[3], Jessica E. Tierney[9], Richard Wilkinson[10], Eleni Anagnostou[11], Agatha M. de Boer[12], Tom Dunkley Jones[13], Kirsty Edgar[13], Christopher J. Hollis[14], David K. Hutchinson[12] and Richard D. Pancost[2]

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

Corresponding author: Gordon N. Inglis

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

Abstract:

Accurate estimates of past global mean surface temperature (GMST) help to contextualise future climate change and are required to estimate the sensitivity of the climate system to CO_2 forcing through Earth history. Previous GMST estimates for the latest Paleocene and early Eocene (~57 to 48 million years ago) span a wide range (~9 to 23°C higher than pre-industrial) and prevent an accurate assessment of climate sensitivity during this extreme greenhouse climate interval. Using the most recent data compilations, we employ a multi-method experimental framework to calculate GMST during the three DeepMIP target intervals: 1) the latest Paleocene (~57 Ma), 2) the Paleocene-Eocene Thermal Maximum (PETM; 56 Ma) and 3) the early Eocene Climatic Optimum (EECO; 53.3 to 49.1 Ma). Using six different methodologies, we find that the average GMST estimate (66% confidence) during the latest Paleocene, PETM and EECO was 26.3°C (22.3 to 28.3°C), 31.6°C (27.2 to 34.5°C) and 27.0°C (23.2 to 29.7°C), respectively. GMST estimates from the EECO are ~10 to 16°C warmer than pre-industrial, higher than the estimate given by the IPCC 5th Assessment Report (9 to 14°C higher than pre-industrial). Leveraging the large 'signal' associated with these extreme warm climates, we combine estimates of GMST and CO_2 from the latest Paleocene, PETM and EECO to calculate gross estimates of the average climate sensitivity between the early Paleogene and today. We demonstrate that "bulk" equilibrium climate sensitivity (66% confidence) during the latest Paleocene, PETM and EECO is 4.5°C (2.4 to 6.8°C), 3.6°C (2.3 to 4.7°C) and 3.1°C (1.8 to 4.4°C) per doubling of CO_2. These values are generally similar to those assessed by the IPCC (1.5 to 4.5°C per doubling CO_2), but appear incompatible with low ECS values (< 1.5 per doubling CO_2).

1. Introduction

Under high growth and low mitigation scenarios, atmospheric carbon dioxide (CO_2) could exceed 1000 parts per million (ppm) by the year 2100 (Stocker et al., 2013). The long-term response of the Earth System under such elevated CO_2 concentrations remains uncertain (Stevens et al., 2016; Knutti et al., 2017; Hegerl et al., 2007). One way to better constrain these climate predictions is to examine intervals in the geological past during which greenhouse gas levels were similar to those predicted under future scenarios. This is the rationale behind the Deep-time Model Intercomparison Project (DeepMIP; www.deepmip.org) which aims to investigate the behaviour of the Earth System in three high CO_2 climate states in the latest Paleocene and early Eocene (~ 57–48 Ma) (Lunt et al., 2017; Hollis et al., 2019).

Sea surface temperature (SST) and land air temperature (LAT) proxies indicate that the latest Paleocene and early Eocene were characterised by global mean surface temperatures (GMST) much warmer than those of today (Cramwinckel et al., 2018; Farnsworth et al., 2019; Hansen et al., 2013; Zhu et al., 2019; Caballero and Huber, 2013). Having a robust quantitative estimate of the magnitude of warming at these times relative to modern is useful for two primary reasons: (1) it allows us to contextualise future climate change predictions by comparing the magnitude of future anthropogenic warming with the magnitude of past natural warming; (2) combined with knowledge of the climate forcing, it allows us to estimate climate sensitivity, a key metric for understanding how the climate system responds to CO_2 forcing. Using different proxy data compilations (Hollis et al., 2012; Lunt et al., 2012), the Fifth IPCC Assessment Report (AR5) stated that GMST was 9°C to 14°C higher than for pre-industrial conditions (*medium confidence*) during the early Eocene (~52 to 50 Ma) (Masson-Delmotte et al., 2014). However, subsequent studies indicate a wider range of estimates, from 9 to 23°C warmer than pre-industrial (Caballero and Huber, 2013; Cramwinckel et al., 2018; Farnsworth et al., 2019; Zhu et al., 2019; Figure 1 and Table 1). It is an open question whether this range arises from inconsistencies between the methods used to estimate GMST, such as selection of proxy datasets, treatment of uncertainty, and/or

analysis of different time intervals. This methodological variability has thwarted robust comparisons between GMST methodologies for key intervals through the latest Paleocene to early Eocene.

Here we calculate GMST estimates within a consistent experimental framework for the target intervals outlined by DeepMIP: i) the Early Eocene Climatic Optimum (EECO; 53.3 to 49.1 Ma), ii) the Paleocene-Eocene Thermal Maximum (PETM, ca. 56 Ma) and iii) the latest Paleocene (LP, ca. 57-56 Ma). We use six different methods to obtain new GMST estimates for these three time intervals, employing previously compiled SST and LAT estimates (Hollis et al., 2019) and bottom water temperature (BWT) estimates (Dunkley Jones et al., 2013; Cramer et al., 2009; Sexton et al., 2011; Littler et al., 2014; Laurentano et al., 2015; Westerhold et al., 2018; Barnet et al., 2019). We also undertake a suite of additional sensitivity studies to explore the influence of particular proxies on each GMST estimate. We then compile GMST estimates from all six methods to generate a 'combined' GMST estimate for each time slice and use these, with existing estimates of CO_2 (Gutjahr et al., 2017; Anagnostou et al., 2016) to develop new estimates of "bulk" equilibrium climate sensitivity (ECS) during the latest Paleocene, PETM and EECO.

2. Methods and Materials

Three different input datasets are used to calculate GMST: 1) dataset D_{surf} which consists of surface temperature estimates, both marine (sea surface temperature) and terrestrial, 2) dataset D_{deep} which consists of deep-water temperature estimates, and 3) dataset D_{comb} which consists of a combination of surface- and deep-water temperature estimates. Here we make use of six different methodologies, which are described in detail below, to estimate GMST from these datasets.

2.1. Dataset D_{surf}

106 Dataset D_{surf} is version 0.1 of the DeepMIP database, as described in Hollis et al (2019)

107 (Supplementary Information). It consists of SSTs and LATs for the latest Paleocene, PETM

108 and EECO. The SSTs are derived from foraminiferal $\delta^{18}O$ values, foraminiferal Mg/Ca ratios,

109 clumped isotopes ($\Delta 47$), and isoprenoid GDGTs (TEX_{86}). Foraminiferal $\delta^{18}O$ values and

110 Mg/Ca ratios are calibrated to SST following Hollis et al., 2019 and Evans et al. (2018),

111 respectively. TEX_{86} values are calibrated to SST using BAYSPAR (Tierney and Tingley, 2014).

112 $\Delta 47$ values are reported using the parameters and calibrations of the original publications

113 (Evans et al., 2018; Keating-Bitonti et al., 2011). LATs are derived from leaf fossils, pollen

114 assemblages, mammal $\delta^{18}O$ values, paleosol $\delta^{18}O$ values, paleosol climofunctions and

115 branched GDGTs. LAT estimates are calculated using the parameters and calibrations of the

116 original publications (see Hollis et al., 2019 and ref. therein). The locations of the proxy

117 datasets are shown in Figure S1 using the paleomagnetic-based reference frame (Hollis et

118 al., 2019). For each dataset, we utilise the uncertainty range of temperature estimates reported

119 in Hollis et al. (2019).

120 Four methods (D_{surf}-1, D_{surf}-2, D_{surf}-3 and D_{surf}-4) are employed to calculate GMST from

121 dataset D_{surf}. These methods employ parametric (D_{surf}-1, D_{surf}-2, D_{surf}-4) or non-parametric

122 (D_{surf}-3) functions to estimate temperature. We calculate GMST on the mantle-based

123 reference frame and employ the rotations provided in Hollis et al (2019). These differ very

124 slightly from those utilised in the DeepMIP model simulations (Lunt et al, 2020). Each method

125 conducts a 'baseline' calculation that uses the SST and LAT data compiled in accordance with

126 the DeepMIP protocols (i.e. Hollis et al., 2019). Our baseline calculation (D_{surf}-baseline; Table

127 2) excludes $\delta^{18}O$ values from recrystallized planktonic foraminifera because the resulting

128 temperature estimates are biased by diagenesis toward significantly cooler temperatures than

129 those derived from: i) the $\delta^{18}O$ value of similar aged and similarly located well-preserved

130 foraminifera, ii) foraminiferal Mg/Ca ratios and iii) $\Delta 47$ values from larger benthic foraminifera

131 (Pearson et al., 2001; Hollis et al., 2019 and ref. therein). For each method, we also conduct

132 a series of illustrative sub-sampling calculations relative to D_{surf}-baseline, based on varying

133   assumptions about the robustness of different proxies (Table 2). The first sensitivity

134   experiment (D_{surf}-Frosty; Table 2) includes $\delta^{18}O$ values from recrystallized planktonic

135   foraminifera. The second sensitivity experiment (D_{surf}-NoTEX; Table 2) removes TEX_{86} values

136   as these give slightly higher SSTs than other proxies, especially in the mid-to-high latitudes

137   (Bijl et al., 2009; Hollis et al., 2012; Inglis et al., 2015). The third sensitivity experiment (D_{surf}-

138   NoMBT; Table 2) removes MBT(')/CBT values derived from marine sediment archives as they

139   may suffer from a cool bias (Inglis et al., 2017; Hollis et al., 2019). The fourth sensitivity

140   experiment (D_{surf}-NoPaleosol; Table 2) removes mammal/paleosol $\delta^{18}O$ values and paleosol

141   climofunctions as these proxies may suffer from a cool bias (Hyland and Sheldon, 2013; Hollis

142   et al., 2019). For each method, GMST is calculated for: i) the Early Eocene Climatic Optimum

143   (EECO; 53.3 to 49.1 Ma), ii) the Paleocene-Eocene Thermal Maximum (ca. 56 Ma) and iii) the

144   latest Paleocene (LP; ca. 57-56 Ma).

145

146   *2.1.1. D_{surf}-1*

147   Method D_{surf}-1 was first employed by Caballero and Huber (2013) to estimate GMST from

148   early Eocene surface temperature proxies after it was recognised that pervasive

149   recrystallization of foraminiferal $\delta^{18}O$ could overprint the original SST signal (e.g. Pearson et

150   al., 2001; Pearson et al., 2007). That study used data compilations (Huber and Caballero,

151   2011, Hollis et al., 2012) which were the predecessors to the DeepMIP compilation (Hollis et

152   al., 2019).

153   Here, the anomalies of individual proxy temperature data points with respect to modern

154   values at the corresponding paleolocation are first calculated. The time period used is between

155   1979 and 2018 and we used a climatology of the full ERA-interim period (Dee et al., 2011).

156   The calculation involves binning into low, mid, and high latitudes (30°N to 30°S, 30°N/S to

157   60°N/S, and 60°N/S to 90°N/S), and calculating the unweighted mean anomaly within these

158   bins between the median reconstructed value at a given locality and the temperature in the

159   modern system (from reanalysis). The geographically binned means are then weighted

160 according to relative spherical area to calculate a globally weighted mean temperature

161 anomaly between the paleo-time slice and modern. All samples are treated equally and

162 considered independent. The associated errors are added in quadrature with the inter-sample

163 standard deviation. These two sources of error were combined and normalized by the square

164 root of the number of samples. This method is intended as an unsophisticated, brute force

165 approach to estimating GMST when dealing with many localities with poorly characterized

166 errors in which there is a large difference between the reconstructed temperature at a given

167 location and the modern equivalent. It is not intended to identify small changes in GMST; nor

168 is it expected to work well under conditions in which temperature gradients are stronger than

169 today, continents are far removed from their current configuration, or in situations in which

170 systematic errors are not readily mitigated by large sample size (i.e. when there are

171 correlations in systematic errors between proxies). It is designed to be relatively

172 straightforward to interpret and simple to reproduce without relying overly on climate models

173 or sophisticated statistical models.

174   Various sanity checks have been performed to determine if the method is likely to

175 produce useful results for a given sampling distribution and what corrections should be applied

176 to optimize it. For example, if the modern temperature field is sampled using a geographic

177 sampling distribution for a given time interval, what would the reconstructed modern

178 temperature be? Sampling the modern global annual average surface temperature field in the

179 reanalysis product ERA-5 yields a mean value of 15.1°C but when resampled at the equivalent

180 geographic distribution of our samples from the latest Paleocene, PETM and EECO yields

181 mean values for the modern of 16.9°C (±1.8°C), 14.2°C (±1.7°C), and 15.2°C (±1.1°C),

182 respectively. Thus, for the sampling densities and spatial structure of the early Paleogene, this

183 method can approach the true value within ~1.5°C and the error propagation adequately

184 characterizes the error, in this 'perfect knowledge' scenario. Seeking precision beyond that

185 range is unwarranted and as indicated above, systematic biases are a serious concern.

186 However, estimating the latest Paleocene and early Eocene GMST may be somewhat easier

187 than estimating the modern GMST because temperature gradients were much reduced from

188   modern. Huber and Caballero (2011) estimate a reduction to less than half the modern

189   temperature gradient whilst Evans et al (2018) constrain the low-to-high latitude SST gradient

190   to at least ~30% (+/- 10%) weaker than modern (Evans et al., 2018).

191   Alongside modern observations, we can also use paleoclimate model results to

192   characterise how well the existing palaeogeographic sampling network will impact results

193   (Figure 2). Here we utilize two CESM1 simulations, as described in Cramwinckel et al., (2018;

194   EO3 and EO4). The two cases are chosen to minimize the magnitude of the correction to

195   GMST and the final result is not sensitive to the choice of reference simulation between these

196   two (Supplementary Information). For each interval, the difference between reconstructed

197   global temperatures and the true paleoclimate model mean is <1 to 3°C. These comparisons

198   demonstrate that this method produces estimates that are within random error given otherwise

199   perfect knowledge.   The errors introduced by limited paleogeographic sampling can be

200   alleviated by incorporating the offset in mean values between the true paleoclimate model

201   GMST and the sampled paleoclimate model GMST outlined above (Figure 2). We utilise this

202   offset to correct for systematic errors, but this is the only component in which paleoclimate

203   model information is included in this GMST estimation methodology. This approach is best

204   applied within the context of studying the random and systematic error structure as described

205   above and caution should be taken in using systematic corrections that are significantly bigger

206   than the estimated random error. The underlying assumption is that the bias in the global

207   mean estimate that exists due to uneven sampling is the same in the 'proxy' Eocene world as

208   in the 'model' Eocene world, i.e. that the zonal and meridional gradients are well characterised

209   by the model, even if the absolute temperatures are not.

210   We note that the magnitude of the global correction could be sensitive to different

211   models and/or boundary conditions. To explore this further, we performed the same analysis

212   using Community Earth System Model version 1.2 (CESM1.2) at 6x CO_2. This model

213   simulation offers a major improvement over earlier models (Zhu et al., 2019) due to the

214   improved treatment of cloud microphysics and is able to reproduce key features of the early

215   Paleogene (e.g. the meridional SST gradient; Zhu et al., 2019; Lunt et al., 2020). We find that

216   CESM1 (8x and 16x CO_2) and CESM1.2 (6x CO_2) yield similar GMST estimates during the

217   PETM, EECO and latest Paleocene. For example, GMST values (obtained using D_{surf}-

218   baseline) during the EECO average 24.5°C, 24.6°C and 25.2°C for CESM1 (x8 CO_2), CESM1

219   (x16 CO_2) and CESM1.2 (6x CO_2), respectively. This indicates that the final result is not overly

220   sensitive to the choice of reference simulation, at least within the CESM model family. In the

221   following sections, we only discuss CESM1 simulations to avoid circularity if the results from

222   this paper are used to evaluate more recent simulations (e.g. CESM1.2; Lunt et al., 2020).

223

224   *2.1.2. D_{surf}-2*

225   GMST estimates are calculated using the method described in Farnsworth et al. (2019), in

226   which a transfer-function is used to calculate global mean temperature from local proxy

227   temperatures. The transfer function is generated from a pair of early Eocene climate model

228   simulations, carried out at two CO_2 concentrations. The first simulations are the same 2x CO_2

229   and 4x CO_2 HadCM3L Eocene simulations from Farnsworth et al (2019). The second

230   simulations are the x 4CO_2 and 8x CO_2 CCSM3 simulations of Huber and Caballero (2011),

231   also discussed in Lunt et al (2012). The two models are configured for the Eocene with

232   different paleogeographies (Supplementary Table S1). We provide a final estimate based on

233   the mean of our two models.

234   The principal assumption of this approach is that global temperatures scale linearly

235   with local temperatures, and that a climate model can represent this scaling correctly (see

236   below). The resulting GMST estimate is therefore independent of the climate sensitivity of the

237   model but dependent on the modelled spatial distribution of temperature. For a single given

238   proxy location with a local temperature estimate (T^{proxy}), Farnsworth et al. (2019) estimate

239   global GMST ($<T>^{inferred}$) as:

240

241     $$<\text{T}>^{inferred} = <T^{low}> + (T^{proxy} - T^{low}) \frac{<T^{high}> - <T^{low}>}{T^{high} - T^{low}}$$ (1)

242

243   where $<T^{low}>$ *and* $<T^{high}>$ are the global means of a low- and high-CO_2 model simulation

244   respectively, and T^{low} *and* T^{high} are the local temperatures (same location as the proxy) from

245   the same simulations. T^{low} and T^{high} represent local modelled SSTs or local modelled near-

246   surface LATs (in contrast to Farnsworth et al. 2019, who only used local modelled near-surface

247   LATs to calculate T^{low} and T^{high}, even if T^{proxy} was SST). If the proxy temperature is greater

248   than T^{high} or cooler than T^{low}, then the inferred global mean is found by extrapolation rather

249   than by interpolation and is therefore more uncertain (Figure 3). This will be sensitive to the

250   choice of model simulation; models that simulate less polar amplification (e.g. HadCM3L) are

251   more likely to obtain $<T>^{inferred}$ (i.e. GMST) via extrapolation. We repeat this process for each

252   proxy data location (Figure 4) and take an average over all proxy locations as our best estimate

253   of global mean temperature.

254   Recent work has demonstrated that CESM1.2 and GFDL model simulations offer a

255   major improvement over earlier models (Zhu et al., 2019; Lunt et al., 2020). As such, we also

256   calculated GMST using CESM1.2 (3x and 6x CO_2; Zhu et al., 2019; Table S1) and GFDL (3x

257   and 6x CO_2; Hutchinson et al., 2018; Lunt et al., 2020; Table S1). We find that all four

258   simulations (i.e. HadCM3L, CCSM3, CESM1.2 and GFDL) yield similar GMST estimates. For

259   example, GMST during the PETM ranges between 32.3 and 34.5°C (Supplementary

260   Information). This demonstrates that D_{surf}-2 is not overly sensitive to the climate model

261   simulation. However, as CESM1.2 and GFDL have greater polar amplification than other

262   models (e.g. HadCM3L), GMST is more likely to be found by interpolation (c.f. extrapolation).

263   To explore whether GMST scales linearly with local temperatures, we used CESM1.2 to re-

264   calculate GMST using the same method as above but using the 9x CO_2 simulation in place of

265   the 6x CO_2 simulation. We find that GMST estimates are very similar (±0.4°C). This is

266   because, although the relationship between GMST and CO_2 is non-linear (Zhu et al, 2019),

267   the relationship between local and global temperature is relatively constant. In the following

268   sections, we employ CCSM3 and HadCM3 simulations to avoid circularity if the results from

269   this paper are used to evaluate more recent simulations (e.g. CESM1.2, GFDL; Lunt et al.,

270   2020).

271

272   *2.1.3. D_{surf}-3*

273   For *D_{surf}-3*, GMST estimates are calculated using Gaussian process regression (Figure 5;

274   Bragg et al., in prep). In this method, temperature is treated as an unknown function of location,

275   *f(x)*. Many possible functions can fit the available proxy dataset. By using a Gaussian process

276   model of the unknown function, we assume that temperature is a continuous and smoothly

277   varying function of location, and once fitted to the data, the posterior mean of the model gives

278   the most likely function form for the temperature. We use a Gaussian process prior and update

279   it using the proxy data to obtain the posterior model which we can then use to predict the

280   surface temperatures on a global grid. Prior specification of the model is via a mean function

281   *E(f(x)) = m(x),* and a covariance function *Cov(f(x), f(x')) =k(x,x')* (which tells us how correlated

282   *f(x)* is with *f(x')*). We also specify the standard deviation of the observation uncertainty about

283   each data point (σ_i^2). If $\boldsymbol{f} = \left(f(x_1), \dots f(x_n)\right)^T$ is a vector of temperature observations at each

284   location x_i, then the model is:

285

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

287

288   where $\mu_i = m(x_i)$ and $\Sigma_{ij} = k\left(x_i, x_j\right) + \mathbb{I}_{i=j}\sigma_i^2$. The proxy temperatures are expressed as

289   anomalies to either the marine or terrestrial present-day zonal mean temperature at the

290   respective paleolatitude. We subtract the mean temperature anomaly (weighted by the

291   paleolatitude) for each time period and core experiment prior to the analysis and therefore fit

292   the model to the residuals. This means the predicted field will relax towards the mean surface

293   warming in areas of no data coverage. The covariance function – which considers the

294   clustering of proxy locations – describes the correlation between *f(x$_i$)* and *f(x$_j$)* in relation to the

295   distance of x_i and x_j. We use a squared-exponential covariance function with Haversine

296   distances replacing Euclidean distances so that correlation is a function of distance on the

297   sphere.

298   A heteroscedastic noise model is used to weight the influence of individual proxy data

299   by their associated uncertainty, i.e. the model will better fit reconstructions with a smaller

300   reported error. Proxy uncertainties are taken from Hollis et al., (2019). Standard deviations for

301   TEX_{86}, Mg/Ca and $\delta^{18}O$ records are derived from the reported 90% confidence intervals (Hollis

302   et al., 2019). A minimum value of 2.5°C for the standard deviation is assumed for all other

303   methods. The output variances and length scale of the covariance function are estimated

304   using their maximum likelihood values, obtained with the GPy Python package (GPy, 2012).

305   We apply the method to the marine and terrestrial data separately and combine the masked

306   fields afterwards to prevent mutual interference. We further constrain the lower bound of the

307   lengthscale parameter to 2000 km to always fit a reasonably smooth surface, even in some

308   continental areas with noisy proxy data (e.g. western North America). We note that our choice

309   of the minimum lengthscale and the separation of land and ocean temperatures influence the

310   predicted regional surface temperature patterns but do not significantly change our GMST

311   estimates.

312   The Gaussian process approach provides probabilistic predictions of temperature

313   values, i.e., uncertainty estimates of the predicted field. The uncertainty reported for an

314   individual GMST estimate is calculated via random sampling. We generate 10,000 surfaces

315   from a multivariate normal distribution based on the predicted mean and full covariance matrix

316   and calculate the GMST for each sample. Uncertainty of the mean estimate is then defined as

317   the standard deviation of the 10,000 random samples. Regional model uncertainty (expressed

318   as standard deviation fields) is typically highest in areas with sparse data coverage (e.g. the

319   Pacific Ocean and Southern Hemisphere landmasses; Figure S2). The lower uncertainty for

320   the latest Paleocene relative to the PETM and EECO is related to the smaller reported

321   uncertainties in the proxy dataset rather than enhanced data coverage. The large spread in

322   reconstructed terrestrial temperatures for North America during the PETM and EECO (Figure

323   S2) propagates through into relatively large uncertainties in the GMSTs estimates for these

324   intervals.

325

326   *2.1.4. D_{surf}-4*

327   For D_{surf}-4, GMST estimates are calculated using a simple function of latitude (θ), tuned to

328   best fit the proxy data:

329

330     $$T(\theta) \approx a + b\theta + c \cos \theta \qquad (3)$$

331

332   where $T(\theta)$ is the Eocene zonal-mean temperature, and the coefficients a, b, and c are chosen

333   to minimize the sum of the squared residuals relative to D_{surf} (i.e. the SST and LAT data from

334   Hollis et al. 2019). This new model represents $T(\theta)$ well in the modern climate (Figure S3)

335   when supplied with similar number of data points as are in the Hollis et al (2019) dataset, and

336   it ensures a global solution that is consistent with the physical expectation that temperature

337   should decrease - and the meridional gradient in temperature should increase - from the

338   tropics toward the poles (Figure S3).

339   For each data point, we account for three types of uncertainty (i.e. temperature,

340   elevation, latitude). For temperature, we assume a skew-normal probability distribution based

341   on the stated 90% confidence intervals. Where uncertainty estimates are not given, we

342   assume a (symmetric) normal distribution with a 90% confidence interval of ±5K. For elevation,

343   we assume a skew-normal distribution with a 90% confidence interval equal to the lowest and

344   highest elevations of adjacent grid points in the paleotopography data set of Herold et al.

345   (2014), with a lower bound of zero.

346        $T(\theta)$ was estimated by sampling temperature, elevation, and latitude from their

347   respective distributions at each location (Figure S4) and a lapse-rate adjustment of 6°K/km

348   was applied. Then, using a standard Monte Carlo bootstrapping method, the same number of

349   data points were resampled via replacement, and the coefficients in Equation 3 were found

350   that best fit the sub-sampled data. This procedure was repeated 10,000 times to find a

351   probability distribution of $T(\theta)$. The uncertainty associated with an individual GMST estimate

352   is the standard deviation.

353

354   **2.2. Dataset D_{deep}**

355   Dataset D_{deep} consists of benthic foraminiferal $\delta^{18}O$-derived bottom water temperatures

356   (BWTs) for the latest Paleocene, PETM and EECO. The benthic foraminiferal $\delta^{18}O$ dataset is

357   based on previous compilations (Dunkley Jones et al., 2013; Cramer et al., 2009), updated to

358   include more recently published datasets (Sexton et al., 2011; Littler et al., 2014; Laurentano

359   et al., 2015; Westerhold et al., 2018; Barnet et al., 2019). The EECO dataset is sourced from

360   eleven sites, providing spatial coverage of both the Pacific, Atlantic and Indian Oceans

361   (DSDP/ODP Sites 401, 550, 577, 690, 702, 738, 865, 1209, 1258, 1262, & 1263). The PETM

362   and latest Paleocene datasets are sourced from a compilation of nine and seven sites,

363   respectively, differing from Dunkley-Jones et al. (2013) in that: i) more recent datasets were

364   added, and ii) PETM sites with a muted CIE magnitude (< 1.5 ‰) were excluded as these

365   datasets may be missing the core PETM interval (Table S2). Benthic foraminifera $\delta^{18}O$ values

366   are adjusted to *Cibicidoides* following established methods (Cramer *et al.*, 2009), allowing

367   temperature to be calculated using Eq. 9 of Marchitto et al (2014):

368

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

370

371   where t is bottom water temperature in Celsius, δ_{cp} is $\delta^{18}O$ of $CaCO_3$ on the Vienna-

372   Pee Dee Belemnite (VPDB) scale, and δ_{sw} is $\delta^{18}O$ of seawater on the Standard Mean Ocean

373   Water (SMOW). δ_{sw} is defined in accordance with the DeepMIP protocols (−1.00 ‰; see Hollis

374   et al., 2019).

375

376   *2.2.1. D_{deep}-1*

377   For *D_{deep}-1*, GMST estimates are calculated following the method of Hansen et al. (2013),

378   which utilises only the deep ocean benthic foraminifera $\delta^{18}O$ dataset, and we refer the reader

379   to that study for a detailed justification of the approach. Briefly, for time periods prior to the

380   Pliocene, GMST is scaled directly to deep ocean temperature. Specifically, ΔGMST = ΔBWT

381   prior to ~5.3 Ma, where early Pliocene BWT and GMST was calculated following Eq. 3.5, 3.6,

382   and 4.2 of Hansen et al. (2013). As such, the calculations presented here differ from those of

383   Hansen et al. (2013) only in that: i) we use the revised benthic $\delta^{18}O$ compilation described

384   above rather than that of Zachos et al. (2008), and ii) a different equation (Eq. 4) to convert

385    $\delta^{18}O$ to temperature.

386

387   **2.3. Dataset D_{comb}**

388   Dataset *D_{comb}* uses a combination of (tropical) surface- and deep-water temperature

389   estimates. The deep ocean dataset (D_{deep}) is identical to that described in Section 2.2. The

390   tropical SST dataset utilises all relevant surface ocean proxy data from the DeepMIP

391   database, i.e. those with a palaeolatitude in the magnetic reference frame within 30° of the

392   equator. An expanded (relative to modern) definition of the tropics is used because tropical

393   SST reconstructions are relatively sparse; 30° was chosen because it retains tropical SST

394   data from several proxies for all three intervals whilst SST seasonality remains relatively low

395   within these latitudinal bounds.

396

397   2.3.1. D_{comb}-1

398   For D_{comb}-1, GMST estimates are calculated for each time interval based on the difference

399   between tropical SSTs and deep-ocean BWTs (Evans et al., 2018), such that:

400

401 $$GMST = 0.5(\overline{tropical\ SST} + \overline{BWT}) \tag{5}$$

402

403   The fundamental assumptions of this approach are that: 1) GMST can be approximated by

404   global mean SST, 2) global mean SST is equivalent to the mean of the tropical and high

405   latitude regions, 3) benthic temperatures are representative of high latitude surface

406   temperatures and 4) that the temperature gradient between the abyss and high latitude SST

407   is fixed through time (c.f. Sijp et al., 2011). To test these assumptions from a theoretical

408   perspective, we modelled the shape of the latitudinal temperature gradient using a simple

409   algebraic function (Figure S5). These results suggest that D_{comb}-1 may underestimate GMST

410   by 0.75 to 1.25 °C in the modern. We also compared GMST from the EO3 and EO4 model

411   simulations of Cramwinckel et al. (2018) to that calculated using D_{comb}-1 (Figure S5) and find

412   a similar cold bias during the Eocene (~1 to 3°C). However, we note that these findings depend

413   on the accuracy of the modelled deep ocean temperatures.

414   Probability distributions for each time interval were computed as follows. In the case

415   of the tropical SST data, 1000 subsamples were taken, following which a random normally

416   distributed error was added to each data point in the DeepMIP compilation, including both

417   calibration uncertainty and variance in the data where multiple reconstructions are available

418   for a given site and time interval. Mean tropical SST was calculated for each of these

419   subsamples. To provide a BWT dataset of the same size as the subsampled tropical SST

420   data, 1000 normally distributed values were calculated for each time interval, based on the

421   mean ±1SD variation of the pooled benthic $\delta^{18}O$ data from all sites including calibration

422   uncertainty.

423

3. Results and Discussion

3.1. Comparison of surface- and bottom water temperature-derived GMST estimates

424

425

426 The following section discusses our 'baseline' GMST estimates calculated on the mantle-

427 based reference frame only. During the latest Paleocene and PETM, GMST estimates derived

428 from D_{surf}-*baseline* average ~27 and 33°C, respectively (Table 3; Figure 6). These values are

429 consistent with previous studies analysing the latest Paleocene (~27°C; Zhu et al., 2019) and

430 PETM (~32°C; Zhu et al., 2019). During the EECO, GMST estimates calculated using D_{surf}

431 average ~27°C (Figure 6). These values are up to 3°C lower compared to previous estimates

432 from similar time intervals (ca. 29 to 30°C; Huber and Caballero, 2011; Caballero and Huber,

433 2013; Zhu et al., 2019). This is likely because we use an expanded LAT dataset (n = 80)

434 compared to previous studies (n = 51; Huber and Caballero, 2011). Several of these proxies

435 saturate between ~25 and 29 °C (e.g. leaf fossils, pollen assemblages and brGDGTs; see

436 Hollis et al., 2019 and ref. therein) and/or are impacted by non-temperature controls (e.g.

437 paleosol climofunctions; see below) and could skew GMST estimates towards lower values.

438 To confirm this, we calculated GMST values using LAT proxies only (Supplementary

439 Information). We show that LAT-only GMST estimates are up to 6°C lower than our 'baseline'

440 (SST + LAT) calculations, suggesting that EECO GMST estimates (D_{surf}-*baseline*) may

441 represent a minimum temperature constraint.

442 GMST estimates for the latest Paleocene, PETM and EECO, calculated using D_{deep},

443 are 25.8°C (±1.4°C), 31.1 (± 2.9°C) and 28.0°C (±1.3°C) respectively (Table 3; Figure 6).

444 These estimates are comparable to those derived from surface temperature proxies alone

445 (Table 3). GMST estimates from the EECO are also comparable to previous estimates based

446 on globally distributed benthic foraminifera data (~28°C; Hansen et al., 2013). As benthic

447 foraminifera are less susceptible to diagenetic alteration than planktonic foraminifera (e.g.

448 Edgar et al., 2013), this implies that benthic foraminiferal $\delta^{18}O$ values could be used to provide

449 the 'fine temporal structure' of Cenozoic temperature change (e.g. Lunt et al., 2016; Hansen

450 et al., 2013). However, we also urge caution as this approach scales GMST directly to BWT

451 prior to the Pliocene and assumes that the characteristics of polar amplification are constant

452 through time (c.f. Evans et al., 2018; Cramwinckel et al., 2018). Changes in ice volume may

453 also influence the benthic foraminiferal $\delta^{18}O$ signal (see Hansen et al., 2013) and additional

454 corrections are required before applying this method to other time intervals (e.g. the Eocene-

455 Oligocene transition). D_{deep} also implies that vertical ocean stratification is fixed, even though

456 vertical ocean stratification has been proposed to change dramatically in the past (e.g. Sijp et

457 al., 2013; Goldner et al., 2014) and may shift the slope and/or intercept of the relationship

458 between BWT and GMST.

459   GMST estimates for the latest Paleocene, PETM and EECO, calculated using D_{comb},

460 are 21.6°C (±1.2°C), 26.6 (± 2.1°C) and 22.8°C (± 1.0°C), respectively (Figure 6). These

461 estimates are consistently lower (up to 5°C) than GMST estimates derived using D_{surf} and

462 D_{deep}. Although D_{comb}-1 can estimate modern GMST within ~1 to 2 °C of measured values,

463 whether this approach can be applied in greenhouse climates remains to be confirmed. As

464 described above, we used CESM1 simulations (EO3 and EO4 from Cramwinckel et al., 2018)

465 to compare the "true" model simulation GMST to that calculated using D_{comb}-1 (Supplementary

466 Information). We find that D_{comb}-1 underestimates GMST by 1°C during the Eocene when the

467 model high latitude SST is used a proxy for the deep-ocean, and 2-3°C when the model deep

468 ocean temperature is used. As such, we suggest that D_{comb}-1 may reflect a minimum GMST

469 constraint. We suggest that variable weighting of the deep ocean and tropics could improve

470 the D_{comb} method in future studies (Eq. 5 gives an equal weighting to each).

471

472 **3.2. Influence of different proxy datasets upon D_{surf}-derived GMST estimates**

473 To explore the importance of the proxies themselves upon D_{surf}-derived GMST estimates, we

474 conducted a series of illustrative subsampling experiments relative to D_{surf}-baseline (Table 2).

475 This was performed for methods D_{surf}-1, -2, -3 and -4. In the first subsampling experiment

476 (D_{surf}-Frosty; Table 2), we include $\delta^{18}O$ SST estimates from recrystallized planktonic

477   foraminifera. This yields lower GMST estimates (<1 to 4°C; e.g. Figure S6-8) and is consistent

478   amongst all four methods. This agrees with previous studies which indicate that $\delta^{18}O$ values

479   from recrystallized planktonic foraminifera are significantly colder than estimates derived from

480   the $\delta^{18}O$ value of well-preserved foraminifera (Pearson et al., 2001; Sexton et al., 2006; Edgar

481   et al., 2015), foraminiferal Mg/Ca ratios (Creech et al., 2010; Hollis et al., 2012) and clumped

482   isotope values from larger benthic foraminifera (Evans et al., 2018).

483   The removal of TEX_{86} results in lower GMST estimates (~1 to 4 °C; e.g. Figure S6-8)

484   across all methodologies (D_{surf}-NoTEX; Table 2). This is consistent with previous studies which

485   indicate that TEX_{86} gives slightly higher SSTs than other proxies, especially in the mid-to-high

486   latitudes (e.g. Hollis et al., 2012; Inglis et al. 2015). The functional response of TEX_{86} at higher-

487   than-modern SSTs remains relatively uncertain, which may explain why TEX_{86} gives slightly

488   higher SSTs than other proxies (see discussion in Hollis et al., 2019). New indices or

489   calibrations could help to reduce the uncertainty associated with TEX_{86}-derived SST estimates

490   beyond the modern calibration range. TEX_{86} values can also be complicated by the input of

491   isoGDGTs from other sources (see discussion in Hollis et al., 2019). The DeepMIP database

492   excludes samples with anomalous GDGT distributions (Hollis et al., 2019). However, a

493   Gaussian process regression (GPR) model may help to better identify anomalous GDGT

494   distributions in the sedimentary record using a nearest neighbour distance metric (Eley et al.,

495   2019). This methodology could be employed in future studies to further refine GDGT-based

496   SST datasets, but this methodology is currently under review and is not considered here.

497   Despite the caveats and concerns raised in previous work, the exclusion of TEX_{86} data shifts

498   GMST by a relatively small amount.

499   The input of brGDGTs from archives other than mineral soils or peat can bias LAT

500   estimates towards lower values (Inglis et al., 2017; Hollis et al., 2019) and the exclusion of

501   MBT(')/CBT-derived LAT estimates could yield higher GMST values. Excluding MBT(')/CBT

502   in marine sediments does yield slightly warmer GMST estimates (0.5 to 1.0°C). However, the

503   impact of excluding MBT(')/CBT values is relatively minor because there are other proxies

504 (e.g. pollen assemblages, leaf floral) which yield comparable LAT estimates in the regions

505 where MBT(')/CBT values are removed (e.g. the SW Pacific).

506   The removal of $\delta^{18}O$ values from paleosols/mammals and paleosol climofunctions

507 (D_{surf}-NoPaleosol; Table 2) also leads to slightly warmer GMST estimates (~0.5°C). This may

508 be related to additional controls on paleosol and mammal $\delta^{18}O$ values. This includes variations

509 in the isotopic composition of rainfall (i.e. meteoric $\delta^{18}O$; Hyland and Sheldon, 2013),

510 variations in soil water $\delta^{18}O$ values (Hyland and Sheldon, 2013) and/or $\delta^{18}O$ heterogeneity

511 within nodules (e.g. Dworkin et al. 2005). Temperature estimates from paleosol climofunctions

512 may also be prone to underestimation (e.g. Sheldon et al., 2009) and Hyland and Sheldon

513 (2013) suggest that paleosol climofunctions are only applied as an indicator of relative

514 temperature change. Intriguingly, D_{surf}-1 method yields much higher GMST estimates during

515 the EECO when $\delta^{18}O$ values from paleosols/mammals and paleosol climofunctions are

516 excluded (~3°C higher than D_{surf}-baseline). This is attributed to the inclusion of two "cold" LAT

517 estimates from the Salta Basin, NW Argentina (Hyland et al., 2017) which overly influence

518 GMST (e.g. Figure 2). For D_{surf}-1, a direct comparison of new and prior estimates (Caballero

519 and Huber, 2013) can be made in which the only change has been the use of a newer data

520 compilation. For our new estimate, the EECO is ~4.5°C colder than previous estimates

521 (29.75°C; Caballero and Huber, 2013). Given that the floristic LAT estimates are identical

522 between the DeepMIP compilation and the older compilation, the lower GMST estimates are

523 largely due to the incorporation of additional LAT datasets (e.g. paleosol climofunctions).

524

525 **3.3. A combined estimate of GMST during the DeepMIP target intervals**

526 To derive a combined estimate of GMST during the latest Paleocene, PETM and EECO, we

527 employ a probabilistic approach, using Monte Carlo resampling with full propagation of errors.

528 Our combined estimates employs GMST estimates from each 'baseline' experiment (except

529 D_{surf}-1 for the EECO for which we use D_{surf}-NoPaleosol; see discussion above). We generated

530 1,000,000 iterations for each of the six methods, for each time interval (latest Paleocene,

531 PETM and EECO). In these iterations, the GMST estimates were randomly sampled with

532 replacement within their full uncertainty envelopes, assuming Gaussian distribution of errors.

533 As the different GMST estimates ultimately derive from the same proxy dataset, we do not

534 consider them to be independent. The resulting 6,000,000 GMST iterations for each time

535 period are thus simply added into a single probability density function, in order to fully

536 represent uncertainty (Figure 7). From this probability distribution, the median value and the

537 upper and lower limits corresponding to 66 and 90% confidence limits were identified (Table

538 4).

539   Sequential removal of one GMST method at a time (jackknife resampling) was

540 performed to examine the influence of a single method upon the average GMST estimate.

541 Jackknifing reveals that that no single method overly influences the mean GMST or 66%

542 confidence intervals during the latest Paleocene, PETM or EECO (±1.5°C; Supplementary

543 Information and Figure S9). However, the removal of D_{surf}-2 (which has relative large error

544 bars; Figure 6) reduces the 90% confidence interval (Supplementary Information). We also

545 show that removing D_{comb}-1 removes the bimodality of the temperature distribution (Figure

546 S9). This is because D_{comb}-1 is associated with consistently lower GMST estimates compared

547 to other methods (see Section 3.1).

548   During the latest Paleocene, the average GMST estimate is 26.3°C and ranges

549 between 22.3 and 28.3°C (66% confidence interval; Table 4; Figure 7). During the PETM, the

550 average GMST is higher (31.6°C) and ranges between 27.2 and 34.5°C (66% confidence

551 interval; Table 4; Figure 7). Assuming a preindustrial GMST of 14°C, our average GMST

552 estimates indicate that the latest Paleocene, and PETM are 12.3°C and 17.6°C warmer than

553 pre-industrial, respectively. Our results indicate that GMST likely increased by ~4 to 6°C

554 between the latest Paleocene and PETM (66% confidence), in keeping with previous

555 estimates (Frieling et al., 2019; Dunkley Jones, 2013). During the EECO, the average GMST

556 estimate is 27.0°C and likely ranges between 23.2 and 29.7°C (66% confidence interval; Table

557 4; Figure 7). Assuming a preindustrial GMST of 14°C, our average GMST estimate indicates

558 that the EECO is 13.0°C warmer than pre-industrial. The GMST anomaly for the EECO is

559 ~2°C lower than previous studies (~15°C warmer than pre-industrial; Caballero and Huber,

560 2013; Zhu et al., 2019) but the median falls within the range quoted previously in the IPCC

561 AR5 (9 to 14°C warmer than pre-industrial). The EECO is approximately 4 to 5°C colder than

562 the PETM (66% confidence). This is larger than previously suggested (~3°C; Zhu et al., 2019)

563 and may related to a cold bias in EECO GMST estimates (see Section 3.1).

564

565 **3.4. Equilibrium climate sensitivity during the latest Palaeocene, PETM and EECO**

566 Equilibrium climate sensitivity (ECS) can be defined as the equilibrium change in global near

567 surface air temperature, resulting from a doubling in atmospheric CO_2. Various "flavours" of

568 ECS exist, some of which specifically exclude various feedback processes not always included

569 in climate models, such as those associated with ice sheets, vegetation, or aerosols (Rohling

570 et al., 2012). ECS may also be state-dependent (Caballero and Huber, 2013) and there is no

571 reason to expect that it has not changed with time or as a function of climate state (Farnsworth

572 et al, 2019; Zhu et al., 2020). Therefore, direct comparison of ECS in the past to modern

573 conditions is a fraught enterprise. For our purposes we define a "bulk" ECS (ECS_{bulk}) as being

574 a gross estimate of ECS, between our three intervals and preindustrial. i.e.

575

576 $$ECS_{bulk} = (\Delta T_{CO2\text{-}vs\text{-}PI}) / (\Delta F_{CO2\text{-}vs\text{-}PI}) \qquad [6]$$

577

578 where $\Delta T_{CO2\text{-}vs\text{-}PI}$ is the temperature difference between pre-industrial and the time period of

579 interest that can be attributed to CO_2 forcing, and $\Delta F_{CO2\text{-}vs\text{-}PI}$ is the CO_2 forcing relative to

580 preindustrial. The result is then normalised to a CO_2 forcing equal to a doubling of CO_2. Such

581 calculations have been performed previously (e.g. Anagnostou et al., 2016) and they provide

582 some constraint on the range of climate sensitivity values that are relevant for near-modern

583 prediction (Rohling et al., 2012). For example, Anagnostou et al. (2016) indicated that early

584 Eocene ECS (excluding ice sheet feedbacks) falls within the range 2.1–4.6 °C per CO_2

585 doubling with maximum probability for the EECO of 3.8 °C. These values (2.1–4.6 °C per CO_2

586   doubling) are similar to the IPCC ECS range (1.5–4.5 °C at 66% confidence). Here we

587   calculate bulk ECS estimates using the change in GMST and CO_2 in the latest Paleocene,

588   PETM and EECO intervals with reference to the pre-industrial. Following the approach of

589   Anagnostou et al. (2016) and using the forcing equation of Byrne and Goldblatt (2014), we

590   first determine the relative change in climate forcing relative to pre-industrial ($\Delta F_{CO2\text{-vs-PI}}$):

591

592                    $$\Delta F_{CO2\text{-vs-PI}} = 5.32\ln(C_t/C_{PI}) + (0.39[\ln(C_t/C_{PI})])^2 \qquad [7]$$

593

594   where C_{PI} is the atmospheric CO_2 concentration during pre-industrial (278 ppm) and C_t refers

595   to the CO_2 reconstruction at a particular time in the Eocene. The mean proxy estimate of

596   CO_2 for the PETM is ~2200 ppmv (+1904/-699 ppmv; 95% confidence) (Gutjahr et al., 2017).

597   The mean proxy estimate of CO_2 for the LP is ~870 ppmv (Gutjahr et al., 2017). The

598   uncertainty of latest Paleocene CO_2 represents two standard deviations of pre-PETM CO2

599   (Gutjahr et al. 2017), equal to ±400 ppm. The mean proxy estimate of CO_2 for the EECO is

600   ~1625 ppmv (±750 ppmv; 95% confidence) (Anagnostou et al., 2016; Hollis et al., 2019). To

601   calculate bulk ECS, we then use radiative forcing from a doubling of CO_2 from Byrne and

602   Goldblatt (2014) to translate CO_2 into forcing relative to preindustrial (ΔF_{CO2}):

603

604                         $$ECS = (\Delta T_{CO2\text{-vs-PI}}) / \Delta F_{CO2\text{-vs-PI}} * 3.875 \qquad [8]$$

605

606   , where GMST (ΔT) distributions are based on output generated via our Monte Carlo

607   simulations (see Section 3.3). Some of the temperature anomaly of the latest Paleocene,

608   PETM, and EECO is caused not by CO_2 but by the different paleotopography,

609   paleobathymetry, and solar constant compared with preindustrial. Furthermore, we choose

610   here to calculate an ECS that explicitly excludes feedbacks associated with vegetation, ice

611   sheets, and aerosols, i.e. $S_{[CO2,LI,VG,AE]}$ in the nomenclature of Rohling et al (2012). To account

612   for these effects, we subtract a value of 4.5°C (Caballero and Huber, 2013; Zhu et al. 2019)

613   from GMST; i.e.

614

615    $$\Delta T_{CO2\text{-}vs\text{-}PI} = \Delta GMST - 4.5°C \qquad\qquad [9]$$

616

617   Following Anagnostou et al. (2016), the uncertainty on the slow-feedback correction on

618    $\Delta GMST$ follows a uniform 'flat' probability ($\pm 1.5°C$). This value of 4.5°C is based upon a

619   comparison of preindustrial and Eocene simulations (both 1x CO_2) conducted with CESM1.2

620   (Zhu et al., 2019), which incorporates the paleogeographic, solar constant, ice sheet,

621   vegetation, aerosol, and ice sheet changes from preindustrial to Eocene. Our value is similar

622   to previous studies which attribute ~4 to 6°C to the non-CO_2 and non-aerosol forcings and

623   feedbacks (Anagnostou et al., 2016; Caballero and Huber, 2013, Lunt et al., 2012). However,

624   the sensitivity to these Eocene boundary conditions is likely model-dependant and this value

625   may differ between model simulations. The uncertainties in our estimated ECS are the

626   products of 10,000 realizations of the latest Paleocene, PETM and EECO CO_2 values and the

627   respective $\Delta GMST$ estimate (the mean estimate and propagated uncertainty) based on

628   randomly sampling each variable within its 66% and 90% confidence interval uncertainty

629   envelope

630    $S_{[CO2,LI,VG,AE]}$ values (66% confidence) for the EECO and PETM average 0.80 (0.46 to

631   1.15) and 0.92 (0.60 to 1.20), respectively. This yields ECS estimates (66% confidence) for

632   the EECO and PETM compared to modern which average 3.1°C (1.8 to 4.4°C) and 3.6°C (2.3

633   to 4.7°C), respectively (Figure 8). These are broadly comparable to previous estimates from

634   the early Eocene which account for paleogeography and other feedbacks (~2.1 to 4.6°C;

635   Anagnostou et al., 2016) They are also similar to those predicted by the IPCC (1.5 to 4.5°C

636   per doubling CO_2). $S_{[CO2,LI,VG,AE]}$ values (66% confidence) during the latest Paleocene average

637   1.16 (0.61 to 1.75), which is somewhat higher than the other DeepMIP intervals. This yields

638　ECS estimates (66% confidence) for the latest Paleocene which average 4.5°C (2.4 to 6.8°C)

639　(Figure 8). Higher ECS values are attributed to relatively high GMST estimates (~26°C) and

640　relatively low CO_2 values (~870ppm) during the latest Paleocene. As latest Paleocene CO_2

641　estimates remain highly uncertain (Gutjahr et al., 2017; see above), new high-fidelity CO_2

642　records are required to accurately constrain ECS during this time.

643　　　ECS may be strongly state-dependant and model simulations indicate a non-linear

644　increase in ECS at higher temperatures (Caballero and Huber, 2013; Zhu et al., 2019) due to

645　changes in cloud feedbacks (Abbot et al., 2009; Caballero and Huber, 2010; Arnold et al.,

646　2012; Zhu et al., 2019). This implies caution when relating geological estimates to modern

647　climate predictions (e.g. Rohling et al., 2012; Zhu et al., 2020) and it may be more appropriate

648　to calculate ECS between different time intervals (e.g. latest Paleocene to PETM; Shaffer et

649　al., 2016). To this end, we also calculate ECS between the transition from the latest

650　Palaeocene to the PETM, assuming that non-CO_2 forcings and feedbacks are negligible. This

651　yields an ECS estimate of 3.6°C. However, we note that early Paleogene CO_2 estimates

652　remain uncertain (Gutjahr et al., 2017) and well-synchronised, continuous and high-resolution

653　CO_2 records are required to accurately constrain ECS during the DeepMIP intervals.

654

655　**4. Conclusions**

656　Using six different methods, we have quantified global mean surface temperatures (GMST)

657　during the latest Paleocene, PETM and EECO. GMST was calculated within a coordinated,

658　experimental framework and utilised six methodologies including three different input

659　datasets. After evaluating the impact of different proxy datasets upon GMST estimates, we

660　combined all six methodologies to derive an average GMST value during the latest Paleocene,

661　PETM and EECO. We show that the 'average' GMST estimate (66% confidence) during the

662　latest Paleocene, PETM and EECO is 26.3°C (22.3 to 28.3°C), 31.6°C (27.2 to 34.5°C) and

663　27.0°C (23.2 to 29.7°C), respectively. Assuming a preindustrial GMST of 14°C, the latest

664　Paleocene, PETM and EECO are 12.3°C, 17.6°C and 13.0°C warmer than modern,

665 respectively. Using our 'combined' GMST estimate, we demonstrate that "bulk" ECS (66%

666 confidence) during the latest Paleocene, PETM and EECO is 4.5°C (2.4 to 6.8°C), 3.6°C (2.3

667 to 4.7°C) and 3.1°C (1.8 to 4.4°C) per doubling of CO_2. Taken together, our study improves

668 our characterisation of the global mean temperature of these key time intervals, allowing future

669 climate change to be put into the context of past changes, and allowing us to provide a refined

670 estimate of ECS.

671

Data availability

673 Data can be accessed via the online supporting information, via www.pangaea.de/, or from

674 the author (email: gordon.inglis@soton.ac.uk).

675

Authorship tiers and contributions

677 Authorship of this manuscript is organized into three tiers according to the contributions of

678 each individual author. Inglis (Tier I) organized the structure and writing of the manuscript,

679 contributed to all sections of the text and designed the figures. Tier II authors (listed

680 alphabetically following Inglis) assumed a leading role by contributing methodologies used in

681 the text. Tier III authors (listed alphabetically following Wilkinson) contributed intellectually to

682 the text and figure design.

683

Declaration of competing interest

685 The authors declare that they have no known competing financial interests or personal

686 relationships that could have appeared to influence the work reported in this paper.

687

Acknowledgements

700

References

702   Abbot, D.S., Huber, M., Bousquet, G. and Walker, C.C.: High-CO_2 cloud radiative forcing

703   feedback over both land and ocean in a global climate model, Geophysical Research

704   Letters, 36, 2009.

705   Arnold, N.P., Tziperman, E. and Farrell, B.: Abrupt transition to strong superrotation driven by

706   equatorial wave resonance in an idealized GCM, Journal of the Atmospheric

707   Sciences, 69, 626-640, 2012.

708   Anagnostou, E., John, E. H., Edgar, K. M., Foster, G. L., Ridgwell, A., Inglis, G. N., Pancost,

709   R. D., Lunt, D. J., and Pearson, P. N.: Changing atmospheric CO_2 concentration was

710   the primary driver of early Cenozoic climate, Nature, 533, 380-384,

711   10.1038/nature17423, 2016.

712   Barnet, J. S., Littler, K., Westerhold, T., Kroon, D., Leng, M. J., Bailey, I., Röhl, U., and Zachos,

713   J. C.: A high-Fidelity benthic stable isotope record of late Cretaceous–early Eocene

714   climate change and carbon-cycling, Paleoceanography & Paleoclimatology, 34, 672-

715   691, 2019.

716   Bemis, B. E., Spero, H. J., Bijma, J., and Lea, D. W.: Reevaluation of the oxygen isotopic

717   composition of planktonic foraminifera: Experimental results and revised

718   paleotemperature equations, Paleoceanography & Paleoclimatology, 13, 150-160,

719   10.1029/98pa00070, 1998.

720 Bijl, P. K., Schouten, S., Sluijs, A., Reichart, G.-J., Zachos, J. C., and Brinkhuis, H.: Early

721   Palaeogene temperature evolution of the southwest Pacific Ocean, Nature, 461, 776-

722   779, 2009.

723 Bragg, F. J., Paine, P., Saul, A., Lunt, D. J., Wilkinson, R., and Zammit-Mangion, A.: A

724   Statistical Algorithm for Evaluating Palaeoclimate Simulations Against Geological

725   Observations, Geoscientific Model Development, In preperation.

726 Byrne, B., and Goldblatt, C.: Radiative forcing at high concentrations of well-mixed

727   greenhouse gases, Geophysical Research Letters, 41, 152-160, 2014.

728 Caballero, R. and Huber, M.: Spontaneous transition to superrotation in warm climates

729   simulated by CAM3, Geophysical Research Letters, 37, 2010

730 Caballero, R., and Huber, M.: State-dependent climate sensitivity in past warm climates and

731   its implications for future climate projections, Proceedings of the National Academy of

732   Sciences, 110, 14162-14167, 2013.

733 Cramer, B. S., Toggweiler, J. R., Wright, J. D., Katz, M. E., and Miller, K. G.: Ocean overturning

734   since the Late Cretaceous: Inferences from a new benthic foraminiferal isotope

735   compilation, Paleoceanography & Paleoclimatology, 24, 10.1029/2008pa001683,

736   2009.

737 Cramwinckel, M. J., Huber, M., Kocken, I. J., Agnini, C., Bijl, P. K., Bohaty, S. M., Frieling, J.,

738   Goldner, A., Hilgen, F. J., Kip, E. L., Peterse, F., van der Ploeg, R., Rohl, U., Schouten,

739   S., and Sluijs, A.: Synchronous tropical and polar temperature evolution in the Eocene,

740   Nature, 559, 382, 2018.

741 Creech, J.B., Baker, J.A., Hollis, C.J., Morgans, H.E. and Smith, E.G.: Eocene sea

742   temperatures for the mid-latitude southwest Pacific from Mg/Ca ratios in planktonic

743   and benthic foraminifera, Earth and Planetary Science Letters, 299, 483-495, 2010.

744 Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U.,

745   Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de

746   Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J.,

747   Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, L., Kållberg, P.,

748   Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park,

749   B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N. and Vitart, F.: The

750   ERA-Interim reanalysis: configuration and performance of the data assimilation

751   system, Quarterly Journal of the Royal Meteorological Society 137, 553-597, 2011.

752 Dunkley Jones, T., Lunt, D.J., Schmidt, D.N., Ridgwell, A., Sluijs, A., Valdes, P.J. and Maslin,

753   M.: Climate model and proxy data constraints on ocean warming across the

754   Paleocene–Eocene Thermal Maximum, Earth Science Reviews, 125, 123-145, 2013.

755 Dworkin, S.I., Nordt, L. and Atchley, S.: Determining terrestrial paleotemperatures using the

756   oxygen isotopic composition of pedogenic carbonate, Earth and Planetary Science

757   Letters, 237,56-68, 2005.

758 Edgar, K.M., Anagnostou, E., Pearson, P.N. and Foster, G.L: Assessing the impact of

759   diagenesis on $\delta^{11}B$, $\delta^{13}C$, $\delta^{18}O$, Sr/Ca and B/Ca values in fossil planktic foraminiferal

760   calcite, Geochimica et Cosmochimica Acta, 166, 89-209, 2015.

761 Edgar, K.M., Pälike, H. and Wilson, P.A.: Testing the impact of diagenesis on the $\delta^{18}O$ and

762        $\delta^{13}C$ of benthic foraminiferal calcite from a sediment burial depth transect in the

763   equatorial Pacific, Paleoceanography, 28, 468-480. 2013.

764 Eley, Y., Thompson, W., Greene, S.E., Mandel, I., Edgar, K., Bendle, J.A and Dunkley Jones,

765   T.: OPTiMAL: A new machine learning approach for GDGT-based palaeothermometry,

766   Climate of the Past Discussions, 2019.

767 Evans, D., Sagoo, N., Renema, W., Cotton, L. J., Müller, W., Todd, J. A., Saraswati, P. K.,

768   Stassen, P., Ziegler, M., Pearson, P. N., Valdes, P. J., and Affek, H. P.: Eocene

769   greenhouse climate revealed by coupled clumped isotope-Mg/Ca thermometry,

770   Proceedings of the National Academy of Sciences, 115, 1174-1179,

771   10.1073/pnas.1714744115 %J Proceedings of the National Academy of Sciences,

772   2018.

773   Farnsworth, A., Lunt, D., O'Brien, C., Foster, G., Inglis, G., Markwick, P., Pancost, R., and
774   Robinson, S.: Climate sensitivity on geological timescales controlled by non-linear
775   feedbacks and ocean circulation, Geophysical Research Letters, 2019.

776   Gallagher, T. M., and Sheldon, N. D.: A new paleothermometer for forest paleosols and its
777   implications for Cenozoic climate, Geology, 41, 647-650, 10.1130/G34074.1, 2013.

778   Goldner, A., Herold, N. and Huber, M.: Antarctic glaciation caused ocean circulation changes
779   at the Eocene–Oligocene transition, Nature, 511,574-577, 2014.

780   Gutjahr, M., Ridgwell, A., Sexton, P. F., Anagnostou, E., Pearson, P. N., Pälike, H., Norris, R.
781   D., Thomas, E., and Foster, G. L.: Very large release of mostly volcanic carbon during
782   the Palaeocene–Eocene Thermal Maximum, Nature, 548, 573-577,
783   10.1038/nature23646, 2017.

784   Hansen, J., Sato, M., Russell, G., and Kharecha, P.: Climate sensitivity, sea level and
785   atmospheric carbon dioxide, Philosophical Transactions of the Royal Society A:
786   Mathematical, Physical Engineering Sciences, 371, 20120294, 2013.

787   Hegerl, G. C., Zwiers, F. W., Braconnot, P., Gillett, N. P., Luo, Y., Marengo Orsini, J., Nicholls,
788   N., Penner, J. E., and Stott, P. A.: Understanding and attributing climate change, IPCC,
789   2007: Climate Change 2007: the physical science basis. contribution of Working Group
790   I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change,
791   2007.

792   Herold, N., Buzan, J., Seton, M., Goldner, A., Green, J. A. M., Müller, R. D., Markwick, P., and
793   Huber, M.: A suite of early Eocene (~ 55 Ma) climate model boundary conditions,
794   Geoscientific Model Development, 7, 2077-2090, 10.5194/gmd-7-2077-2014, 2014.

795   Hollis, C. J., Taylor, K. W. R., Handley, L., Pancost, R. D., Huber, M., Creech, J. B., Hines, B.
796   R., Crouch, E. M., Morgans, H. E. G., Crampton, J. S., Gibbs, S., Pearson, P. N., and
797   Zachos, J. C.: Early Paleogene temperature history of the Southwest Pacific Ocean:
798   Reconciling proxies and models, Earth and Planetary Science Letters, 349–350, 53-
799   66, http://dx.doi.org/10.1016/j.epsl.2012.06.024, 2012.

800 Hollis, C. J., Dunkley Jones, T., Anagnostou, E., Bijl, P. K., Cramwinckel, M. J., Cui, Y.,

801   Dickens, G. R., Edgar, K. M., Eley, Y., Evans, D., Foster, G. L., Frieling, J., Inglis, G.

802   N., Kennedy, E. M., Kozdon, R., Lauretano, V., Lear, C. H., Littler, K., Lourens, L.,

803   Meckler, A. N., Naafs, B. D. A., Pälike, H., Pancost, R. D., Pearson, P. N., Röhl, U.,

804   Royer, D. L., Salzmann, U., Schubert, B. A., Seebeck, H., Sluijs, A., Speijer, R. P.,

805   Stassen, P., Tierney, J., Tripati, A., Wade, B., Westerhold, T., Witkowski, C., Zachos,

806   J. C., Zhang, Y. G., Huber, M., and Lunt, D. J.: The DeepMIP contribution to PMIP4:

807   methodologies for selection, compilation and analysis of latest Paleocene and early

808   Eocene climate proxy data, incorporating version 0.1 of the DeepMIP database,

809   Geoscientific Model Development, 12, 3149-3206, 10.5194/gmd-12-3149-2019, 2019.

810 Huber, M., and Caballero, R.: The early Eocene equable climate problem revisited, Clim. Past,

811   7, 603-633, 10.5194/cp-7-603-2011, 2011.

812 Hutchinson, D. K., de Boer, A. M., Coxall, H. K., Caballero, R., Nilsson, J. and Baatsen, M.:

813   Climate sensitivity and meridional overturning circulation in the late Eocene using

814   GFDL CM2.1, *Climate of the Past,* 14**,** 789-810, 2018.

815 Hyland, E. G., Sheldon, N. D., and Cotton, J. M.: Constraining the early Eocene climatic

816   optimum: A terrestrial interhemispheric comparison, GSA Bulletin, 129, 244-252,

817   10.1130/B31493.1, 2017.

818 Hyland, E.G. and Sheldon, N.D.: Coupled CO2-climate response during the early Eocene

819   climatic optimum, *P*alaeogeography, Palaeoclimatology, Palaeoecology, 369, 125-

820   135, 2013.

821 Inglis, G. N., Farnsworth, A., Lunt, D., Foster, G. L., Hollis, C. J., Pagani, M., Jardine, P. E.,

822   Pearson, P. N., Markwick, P., Galsworthy, A. M. J., Raynham, L., Taylor, K. W. R., and

823   Pancost, R. D.: Descent toward the Icehouse: Eocene sea surface cooling inferred

824   from GDGT distributions, Paleoceanography, 30, 1000-1020, 10.1002/2014pa002723,

825   2015.

826 Inglis, G. N., Collinson, M. E., Riegel, W., Wilde, V., Farnsworth, A., Lunt, D. J., Valdes, P.,
827   Robson, B. E., Scott, A. C., Lenz, O. K., Naafs, B. D. A., and Pancost, R. D.: Mid-
828   latitude continental temperatures through the early Eocene in western Europe, Earth
829   and Planetary Science Letters, 460, 86-96, https://doi.org/10.1016/j.epsl.2016.12.009,
830   2017.

831 Keating-Bitonti, C. R., Ivany, L. C., Affek, H. P., Douglas, P., and Samson, S. D.: Warm, not
832   super-hot, temperatures in the early Eocene subtropics, Geology, 39, 771-774, 2011.

833 Knutti, R., Rugenstein, M. A., and Hegerl, G. C.: Beyond equilibrium climate sensitivity, Nature
834   Geoscience, 10, 727-736, 2017.

835 Lauretano, V., K. Littler, M. Polling, J. C. Zachos, and L. J. Lourens.: Frequency, magnitude
836   and character of hyperthermal events at the onset of the Early Eocene Climatic
837   Optimum, Climate of the Past, 11, 1313-132, 2015.

838 Littler, K., Röhl, U., Westerhold, T. and Zachos, J.C.: A high-resolution benthic stable-isotope
839   record for the South Atlantic: Implications for orbital-scale changes in Late Paleocene–
840   Early Eocene climate and carbon cycling, Earth and Planetary Science Letters, 401,
841   18-30, 2014.

842 Lunt, D. J., Jones, T. D., Heinemann, M., Huber, M., LeGrande, A., Winguth, A., Loptson, C.,
843   Marotzke, J., Roberts, C., and Tindall, J.: A model-data comparison for a multi-model
844   ensemble of early Eocene atmosphere-ocean simulations: EoMIP, Climate of the Past,
845   8, 2012.

846 Lunt, D. J., Farnsworth, A., Loptson, C., Foster, G. L., Markwick, P., O'Brien, C. L., Pancost,
847   R. D., Robinson, S. A., and Wrobel, N.: Palaeogeographic controls on climate and
848   proxy interpretation, Climate of the Past, 12, 1181-1198, 2016.

849 Lunt, D. J., Huber, M., Anagnostou, E., Baatsen, M. L. J., Caballero, R., DeConto, R., Dijkstra,
850   H. A., Donnadieu, Y., Evans, D., Feng, R., Foster, G. L., Gasson, E., von der Heydt,
851   A. S., Hollis, C. J., Inglis, G. N., Jones, S. M., Kiehl, J., Kirtland Turner, S., Korty, R.
852   L., Kozdon, R., Krishnan, S., Ladant, J. B., Langebroek, P., Lear, C. H., LeGrande, A.

853   N., Littler, K., Markwick, P., Otto-Bliesner, B., Pearson, P., Poulsen, C. J., Salzmann,

854   U., Shields, C., Snell, K., Stärz, M., Super, J., Tabor, C., Tierney, J. E., Tourte, G. J.

855   L., Tripati, A., Upchurch, G. R., Wade, B. S., Wing, S. L., Winguth, A. M. E., Wright, N.

856   M., Zachos, J. C., and Zeebe, R. E.: The DeepMIP contribution to PMIP4: experimental

857   design for model simulations of the EECO, PETM, and pre-PETM (version 1.0),

858   Geoscientific Model Development, 10, 889-901, 10.5194/gmd-10-889-2017, 2017.

859   Lunt, D.J, Bragg, F., Chan, W, Hutchinson, D., Ladant, J-B., Morozova, P., Niezgodzki, I.,

860   Steinig, S., Zhang, Z., Zhu, J., Abe-Ouchi, A., Anagnostou, E., de Boer, A.M.,

861   Coxall, H.K., Donnadieu, Y., Foster, G., Inglis, G.N., Knorr, G., Langebroek, P.,

862   Lear, C.H., Lohmann, G., Poulsen, C.J., Sepulchre, P., Tierney, J., Valdes, P.J.,

863   Dunkley Jones, T., Hollis, C., Huber, M and Otto-Bliesner, B.L.: DeepMIP: Model

864   intercomparison of early Eocene climatic optimum (EECO) large-scale climate

865   features and comparison with proxy data, Climate of the Past Discussion, 1-20,

866   2020.

867   Marchitto, T., Curry, W., Lynch-Stieglitz, J., Bryan, S., Cobb, K., and Lund, D.: Improved

868   oxygen isotope temperature calibrations for cosmopolitan benthic foraminifera,

869   Geochimica et Cosmochimica Acta, 130, 1-11, 2014.

870   Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., Gonzalez Rouco,

871   J. F., Jansen, E., Lambeck, K., Luterbacher, J., Naish, T., Osborn, T., Otto-Bliesner,

872   B., Quinn, T., Ramesh, R., Rojas, M., Shao, X., and Timmermann, A.: Information from

873   Paleoclimate Archives, in: Climate Change 2013 – The Physical Science Basis:

874   Working Group I Contribution to the Fifth Assessment Report of the Intergovernmental

875   Panel on Climate Change, Cambridge University Press, Cambridge, 383-464, 2014.

876   Müller, R.D., Sdrolias, M., Gaina, C., Steinberger, B. and Heine, C.: Long-term sea-level

877   fluctuations driven by ocean basin dynamics, Science, 319, 1357-1362, 2008.

878   Pearson, P. N., Ditchfield, P. W., Singano, J., Harcourt-Brown, K. G., Nicholas, C. J., Olsson,

879   R. K., Shackleton, N. J., and Hall, M. A.: Warm tropical sea surface temperatures in

880   the Late Cretaceous and Eocene epochs, Nature, 413, 481-487, 2001.

881   Pearson, P. N., van Dongen, B. E., Nicholas, C. J., Pancost, R. D., Schouten, S., Singano, J.

882   M., and Wade, B. S.: Stable warm tropical climate through the Eocene Epoch,

883   Geology, 35, 211-214, 10.1130/g23175a.1, 2007.

884   Rohling, E. J., Sluijs, A., Dijkstra, H. A., Köhler, P., van de Wal, R. S. W., von der Heydt, A.

885   S., Beerling, D. J., Berger, A., Bijl, P. K., Crucifix, M., DeConto, R., Drijfhout, S. S.,

886   Fedorov, A., Foster, G. L., Ganopolski, A., Hansen, J., Hönisch, B., Hooghiemstra, H.,

887   Huber, M., Huybers, P., Knutti, R., Lea, D. W., Lourens, L. J., Lunt, D., Masson-

888   Delmotte, V., Medina-Elizalde, M., Otto-Bliesner, B., Pagani, M., Pälike, H., Renssen,

889   H., Royer, D. L., Siddall, M., Valdes, P., Zachos, J. C., Zeebe, R. E., and Members, P.

890   P.: Making sense of palaeoclimate sensitivity, Nature, 491, 683-691,

891   10.1038/nature11574, 2012.

892   Sexton, P.F., Norris, R.D., Wilson, P.A., Pälike, H., Westerhold, T., Röhl, U., Bolton, C.T. and

893   Gibbs, S.: Eocene global warming events driven by ventilation of oceanic dissolved

894   organic carbon, Nature, 471, 349-352, 2011.

895   Sexton, P.F., Wilson, P.A. and Pearson, P.N.: Microstructural and geochemical perspectives

896   on planktic foraminiferal preservation: "Glassy" versus "Frosty", Geochemistry,

897   Geophysics, Geosystems, 7, 2006.

898   Sijp, W.P., England, M.H. and Huber, M.: Effect of the deepening of the Tasman Gateway on

899   the global ocean, Paleoceanography, 26, 2011.

900   Shaffer, G., Huber, M., Rondanelli, R., and Pedersen, J. O. P. J. G. R. L.: Deep time evidence

901   for climate sensitivity increase with warming, 43, 6538-6545, 2016.

902   Sheldon, Nathan D., Retallack, Gregory J., and Tanaka, S.: Geochemical Climofunctions from

903   North American Soils and Application to Paleosols across the Eocene-Oligocene

904   Boundary in Oregon, The Journal of Geology, 110, 687-696, 10.1086/342865, 2002.

905   Sheldon, N.D.: Non-marine records of climatic change across the Eocene-Oligocene

906   transition, The Late Eocene Earth—Hothouse, Icehouse, and Impacts: Geological

907   Society of America Special Paper, 452, 241-248, 2009.

908 Stevens, B., Sherwood, S. C., Bony, S., and Webb, M. J.: Prospects for narrowing bounds on
909   Earth's equilibrium climate sensitivity, Earth's Future, 4, 512-522, 2016.

910 Tierney, J. E., and Tingley, M. P.: A Bayesian, spatially-varying calibration model for the
911   TEX86 proxy, Geochimica et Cosmochimica Acta, 127, 83-106, 2014.

912 Westerhold, T., Röhl, U., Donner, B., and Zachos, J. C.: Global extent of early Eocene
913   hyperthermal events: A new Pacific benthic foraminiferal isotope record from Shatsky
914   Rise (ODP Site 1209), Paleoceanography & Paleoclimatology, 33, 626-642, 2018.

915 von der Heydt, A.S., Dijkstra, H.A., van de Wal, R.S., Caballero, R., Crucifix, M., Foster, G.L.,
916   Huber, M., Köhler, P., Rohling, E., Valdes, P.J. and Ashwin, P.: Lessons on climate
917   sensitivity from past climate changes, Current Climate Change Reports, 2, 148-158,
918   2016.

919 Zachos, J.C., Dickens, G.R. and Zeebe, R.E.; An early Cenozoic perspective on greenhouse
920   warming and carbon-cycle dynamics, Nature, 451, 279-283, 2008.

921 Zhu, J., Poulsen, C. J., and Tierney, J. E.: Simulation of Eocene extreme warmth and high
922   climate sensitivity through cloud feedbacks, Science Advances, 5, 2019.

923 Zhu, J., Poulsen, C.J. and Otto-Bliesner, B.L.: High climate sensitivity in CMIP6 model not
924   supported by paleoclimate, Nature Climate Change, 10, 378-379, 2020,

925

926

927

928

929

930

931

932

Label in Fig. 1	Source	Time	GMST (°C)	Uncertainty (°C)	Proxy system
1a	Farnsworth et al. (2019)	EE	23.4	±3.2	$\delta^{18}O$ planktonic
1b	Farnsworth et al. (2019)	EE	37.1	±1.4	$\delta^{18}O$ planktonic + TEX_{86}
2a	Zhu et al. (2019)	LP	27	n/a	Multiple
2b	Zhu et al. (2019)	EECO	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	$\delta^{18}O$ benthic
5	Cramwinckel et al. (2018)	EE	29.3	n/a	Multiple

933

934 **Table 1:** Previous studies that have determined GMST for the early Eocene (EE), EECO,

935 PETM or latest Paleocene (LP). n/a indicates that no error bars were reported in the original

936 publications.

937

938

939

940

941

942

943

944

945

946

Experiment	Description
D_{surf}-*Baseline*	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}-*baseline* but including recrystallized planktonic foraminifera $\delta^{18}O$ values
D_{surf}-*NoTEX*	D_{surf}-*baseline* but excluding TEX_{86} values
D_{surf}-*NoMBT*	D_{surf}-*baseline* but excluding MBT(')/CBT values from marine sediments
D_{surf}-*NoPaleosol*	D_{surf}-*baseline* but excluding mammal/paleosol $\delta^{18}O$ values and paleosol climofunctions

947   **Table 2**: Baseline and optional subsampling experiments applied to D_{surf}

948

949

950

951

952

953

954

955

956

957

958

959

960

961

GMST (°C)						
	D_{surf}-1	D_{surf}-2	D_{surf}-3	D_{surf}-4	D_{deep}-1	D_{comb}-1
LP	26.6 (±1.3)	26.8 (±6.9)	27.6 (±1.5)	26.8 (±1.3)	25.8 (±1.4)	21.6 (±1.2)
PETM	33.9 (±1.4)	33.4 (±10.3)	32.6 (±1.5)	30.7 (±1.6)	31.1 (±2.9)	26.6 (±2.1)
EECO	27.2 (±0.7)	26.7 (±8.9)	29.8 (±1.5)	25.7 (±1.1)	28.0 (±1.3)	22.8 (±1.0)

962

963 **Table 3:** Individual GMST estimates for latest Paleocene (LP), PETM and EECO. Reported

964 GMST estimates utilise 'baseline' experiments except D_{surf}-1 during the EECO which uses

965 D_{surf}-*NoPaleosol.* GMST estimates are based on the mantle-based reference frame. Error bars

966 on each individual method are the standard deviation (1σ), except D_{surf-1} and D_{surf-2} which use

967 the standard error (1$\sigma_{\bar{x}}$).

968

969

970

971

972

973

974

975

976

977

978

979

980

	GMST (°C) (Average)	GMST (°C) (66% CI)	GMST (°C) (90% CI)
LP	26.3	22.3 – 28.3	21.3 – 29.1
PETM	31.6	27.3 - 34.5	25.9 – 35.6
EECO	27.0	23.2 – 29.6	22.2 – 30.7

981

982 **Table 4:** 'Combined' GMST estimates (66% and 90% confidence intervals) during the: i) latest

983 Paleocene (LP), ii) PETM, and iii) EECO.

984

985

986

987

988

989

990

991

992

993

994

995

996

997

	ECS (°C) (Average)	ECS (°C) (66% CI)	ECS (°C) (90% CI)
LP	4.5	2.4 – 6.8	1.6 – 8.0
PETM	3.6	2.3 – 4.7	1.9 – 5.2
EECO	3.1	1.8 – 4.4	1.3 – 5.0

998

999 **Table 5:** Estimates of ECS (66% and 90% confidence) during the: i) latest Paleocene (LP), ii)

1000 PETM and iii) EECO.

1001

1002

1003

1004

1005

1006

1007

1008

1009

1010

1011

1012

1013

1014

1015

Figure captions:

Figure 1: Published GMST estimates during the early Paleogene (57 to 48 Ma). Dots

represent average values. The horizontal limits on the individual dots represent the reported

error. *y*-Axis labels refer to previous estimates (see Table 1).

Figure 2: An illustration of Method D_{surf}-1 during the EECO. (a) Modelled early Eocene

temperatures utilising CESM1.2 at 6x pre-industrial CO_2, (b) Interpolated absolute SST

reconstructions, (c) Data-model difference between (a) and (b).

Figure 3: An illustration of Method D_{surf}-2 for 2 sites: (a) Big Bend LAT in the EECO as

diagnosed using HadCM3L, and (b) DSDP Site 401 SST in the PETM as diagnosed using

CCSM3. The vertical dashed line shows $< T >^{inferred}$ and the horizontal dashed line shows

T^{proxy}, which intercept at the orange dot. The dark blue dots show the intercept of T^{low} with $<$

$T^{low} >$, and the red dots show the intercept of T^{high} with $< T^{high} >$.

Figure 4: Inferred global mean temperature ($< T >^{inferred}$) using D_{surf}-2, for (a) each EECO-aged

LAT proxy as diagnosed using HadCM3L, and (b) each PETM-aged SST proxy as diagnosed

using CCSM3. For (a) and (b), the final estimate of global mean temperature is the average

of all the individual sites. The solid line shows the continental outline in each model, and the

dashed line shows the continental outline.

Figure 5: Predicted surface warming by Gaussian process regression using D_{surf}-3 for the (a)

latest Paleocene, (b) PETM and (c) EECO. Anomalies are relative to the present-day zonal

mean surface temperature. Circles (triangles) indicate all available SST (LAT) proxy data for

1040 the respective time slice that were used to train the model. Symbols for locations where

1041 multiple proxy reconstructions are available are slightly shifted in latitude for improved visibility.

1042

1043 **Figure 6:** GMST estimates during the (a) PETM, (b) EECO and (c) latest Paleocene for each

1044 methodology. GMST estimates utilise 'baseline' experiments except D_{surf}-1 during the EECO

1045 which uses D_{surf}-*NoPaleosol.* GMST estimates are based on the mantle-based reference

1046 frame. Error bars on each individual method are the standard deviation (1σ), except D_{surf-1} and

1047 D_{surf-2} which use the standard error (1σ).

1048

1049 **Figure 7:** Probability density function of 'combined' GMST during the DeepMIP intervals with

1050 full propagation of errors. GMST estimates are calculated on the mantle-based reference

1051 frame.

1052

1053 **Figure 8:** Probability density function of 'bulk' ECS during the latest Paleocene, PETM and

1054 EECO that explicitly accounts for non-CO_2 forcings of palaeography and solar constant, and

1055 feedbacks associated with land ice, vegetation, and aerosols (Zhu et al., 2019), i.e.

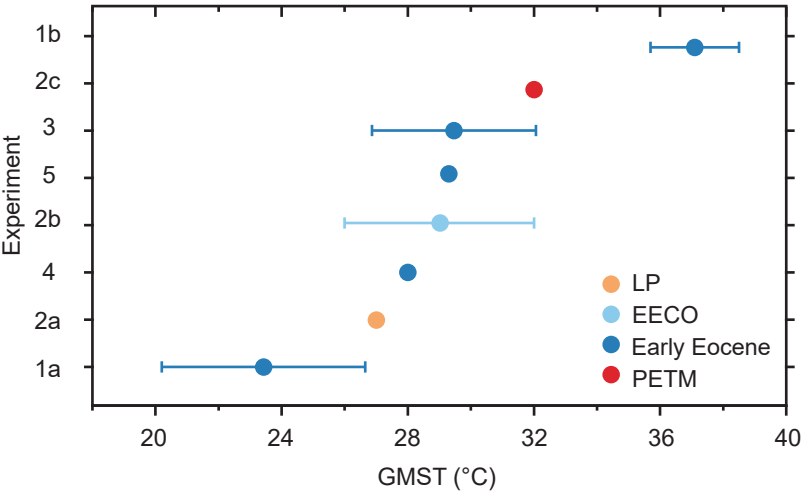1056 $S_{[CO2,LI,VG,AE]}$ in the nomenclature of Rohling et al (2012).

Figure 1

Figure 2

a)

Mean GMST: 30.5°C
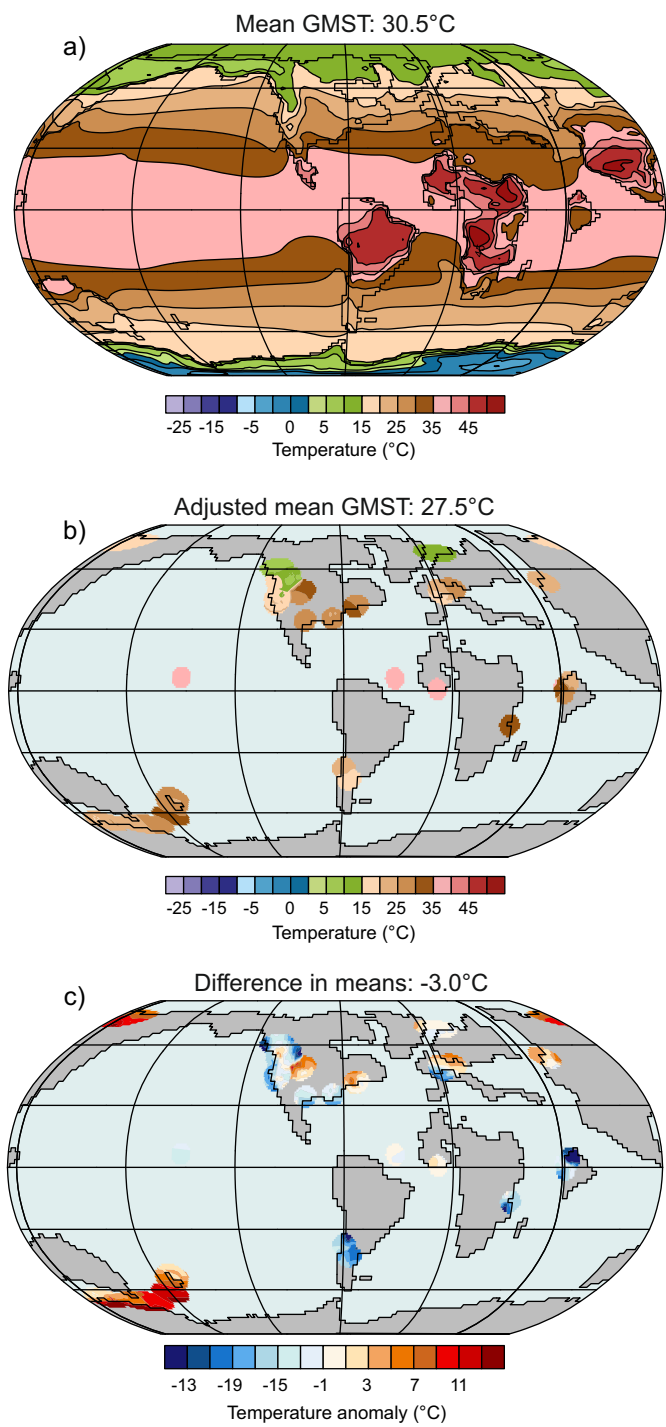

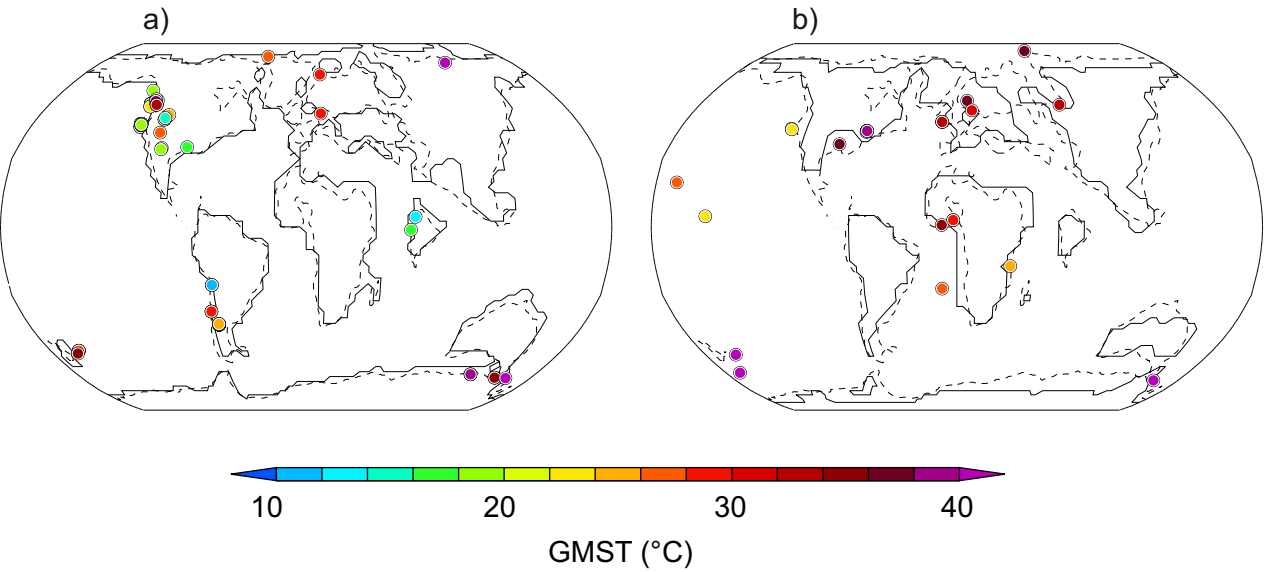-25 -15 -5 0 5 15 25 35 45
Temperature (°C)

b)

Adjusted mean GMST: 27.5°C

-25 -15 -5 0 5 15 25 35 45
Temperature (°C)

c)

Difference in means: -3.0°C

-13 -19 -15 -1 3 7 11
Temperature anomaly (°C)

Figure 3

Figure 4

a) b)

GMST (°C)

10   20   30   40

Figure 5

Figure 6

a) PETM

b) EECO

c) latest Paleocene

Figure 7

Figure 8