DeepMIP: experimental design for model simulations of the EECO, PETM, and pre-PETM.

Daniel J. Lunt[1], Matthew Huber[2], Michiel L.J. Baatsen[3], Rodrigo Caballero[4], Rob DeConto[5], Yannick Donnadieu[6], David Evans[7], Ran Feng[8], Gavin Foster[9], Ed Gasson[5], Anna S. von der Heydt[3], Chris J. Hollis[10], Sandy Kirtland Turner[11], Robert L. Korty[12], Reinhardt Kozdon[13], Srinath Krishnan[7], Jean-Baptiste Ladant[6], Petra Langebroek[14], Caroline H. Lear[15], Allegra N. LeGrande[16], Kate Littler[17], Paul Markwick[18], Bette Otto-Bliesner[8], Paul Pearson[15], Chris Poulsen[19], Ulrich Salzmann[20], Christine Shields[8], Kathryn Snell[21], Michael Starz[22], James Super[7], Clay Tabour[8], Jess Tierney[23], Gregory J.L. Tourte[1], Gary R. Upchurch[24], Bridget Wade[25], Scott L. Wing[26], Arne M.E. Winguth[27], Nicky Wright[28], James C. Zachos[29], and Richard Zeebe[30]

[1]School of Geographical Sciences, University of Bristol, UK
[2]Department of Earth Sciences, University of New Hampshire, USA
[3]Institute for Marine and Atmospheric Research, Utrecht University, Netherlands
[4]Department of Meteorology (MISU), Stockholm University, Sweden
[5]Department of Geosciences, University of Massachusetts-Amherst ,USA
[6]Laboratoire des Sciences du Climat et de l'Environnement, CNRS/CEA, France
[7]Department of Geology and Geophysics, Yale University, USA
[8]National Centre for Atmospheric Research, USA
[9]Ocean and Earth Science, National Oceanography Centre Southampton, University of Southampton, UK
[10]GNS Science, New Zealand
[11]Department of Earth Sciences, University of California - Riverside, USA
[12]Department of Atmospheric Sciences, Texas A&M University, USA
[13]Lamont-Doherty Earth Observatory of Columbia University, USA
[14]Bjerknes Centre for Climate Research, University of Bergen, Norway
[15]School of Earth and Ocean Sciences, Cardiff University
[16]NASA-GISS, USA
[17]Camborne School of Mines, University of Exeter, UK
[18]Getech PlC, UK
[19]Department of Earth and Environmental Sciences, University of Michigan, USA
[20]Department of Geography, Northumbria University, UK
[21]Department of Geological Sciences, University of Colorado, USA
[22]Alfred Wegener Institute, Germany
[23]Department of Geosciences, University of Arizona, USA
[24]Department of Biology, Texas State Univesity, USA
[25]Department of Earth Sciences, University College, London, UK
[26]Department of Paleobiology, Smithsonian Institution, USA
[27]Earth and Environmental Science, University of Texas - Arlington, USA
[28]School of Geosciences, University of Sydney, Australia
[29]PBSci-Earth & Planetary Sciences Department, Institute of Marine Sciences, University of California - Santa Cruz, USA
[30]Department of Oceanography, University of Hawaii at Manoa, USA

Correspondence to: Dan Lunt (d.j.lunt@bristol.ac.uk)

Abstract. Past warm periods provide an opportunity to evaluate climate models under extreme forcing scenarios, in particular high ($>800\,\mathrm{ppmv}$) atmospheric CO_2 concentrations. Although a post-hoc intercomparison of Eocene ($\sim$50 million years ago, Ma) climate model simulations and geological data has been carried out previously, models of past high-CO_2 periods have never been evaluated in a consistent framework. Here, we present an experimental design for climate model simulations of three warm periods within the latest Paleocene and the early Eocene. Together these form the first phase of DeepMIP – the deep-time model intercomparison project, itself a group within the wider Paleoclimate Modelling Intercomparison Project (PMIP). The experimental design consists of three core paleo simulations and a set of optional sensitivity studies. The experimental design specifies and provides guidance on boundary conditions associated with palaeogeography, greenhouse gases, orbital configuration, solar constant, land surface parameters, and aerosols. Initial conditions, simulation length, and output variables are also specified. Finally, we explain how the geological datasets, which will be used to evaluate the simulations, will be developed.

1 Introduction

There is a large community of Earth scientists who focus on 'deep-time' palaeoclimates, here defined as climates of the pre-Pliocene (i.e., prior to $\sim$5 Ma). Recently, a growing community of modelling groups focussing on these periods is also beginning to emerge. DeepMIP – the deep-time model intercomparison project – brings together these modellers, and the data community, into a multidisciplinary international effort dedicated to conceiving, designing, carrying out, analysing, and disseminating an improved understanding of these time periods. It also aims to assess their relevance for our understanding of future climate change. DeepMIP is a working group in the wider Paleoclimate Modelling Intercomparison Project (PMIP), which itself is a part of the sixth incarnation of the Coupled Model Intercomparison Project (CMIP6). In DeepMIP, we will focus on three time periods in the latest Paleocene and early Eocene ($\sim$55–50 Ma), and for the first time, carry out a formal co-ordinated model–model–data intercomparison. In addition to the experimental design presented here, DeepMIP will synthesise existing proxy records, and develop new ones if appropriate, and carry out the model–model–data comparison. The aim will be to best characterise our understanding of the palaeoclimate of the chosen interval through the synthesis of proxy records, to compare this with the model simulations, and to understand the reasons for the intra and inter model and data differences. The ultimate aim is to encourage development of models in response to any robust model deficiencies which emerge from the model–data comparison. This is of particular relevance given the relative warmth and high CO_2 which characterises many intervals of deep-time.

2 Previous Work

An informal, post-hoc model–model–data intercomparison has previously been carried out for the early Eocene (Lunt et al., 2012). This compared the results of four models from five modelling groups with marine and terrestrial data syntheses, and explored the reasons for the model–model differences using energy balance diagnostics. That study contributed to the recent

IPCC AR5 report (Box 5.1, Fig. 1), but it also revealed challenging differences between model simulations of this period, intriguing model–data mismatches, as well as inconsistencies between proxies. Further work resulting from this intercomparison included Gasson et al. (2014) which investigated the CO_2 thresholds for Antarctic ice sheet inception, Lunt et al. (2013) which compared the ensemble and data to further Eocene simulations, and Carmichael et al. (2016) which investigated the hydrological cycle across the ensemble and compared model results with proxies for precipitation.

The exercise pointed to the need for a more coordinated experimental design (different modelling groups had carried out simulations with different boundary conditions, and different initial conditions etc.), and a greater understanding for the reasons behind proxy–proxy differences. Those challenges provide the motivation for DeepMIP.

3   The chosen intervals – the Early Eocene Climatic Optimum (EECO) the Palaeocene–Eocene Thermal Maximum (PETM), and the pre-PETM.

The choice of time interval on which to focus is a balance between (i) the magnitude of the anticipated climate signal (larger signals have a higher signal-to-uncertainty ratio, and larger signals provide a greater challenge to models), (ii) the uncertainties in boundary conditions which characterise the interval (small uncertainties result in more robust conclusions as to the models' abilities, and minimise the model sensitivity studies required to explore the uncertainties), and (iii) the amount and geographic distribution of palaeoclimate data available with which to evaluate the model simulations.

We have chosen to focus on the latest Paleocene and early Eocene – $\sim$55 to $\sim$50 Ma (Ypresian stage), as it is the most recent geological interval characterised by high ($>$800 ppmv) atmospheric CO_2 concentrations. Within the latest Paleocene and early Eocene, DeepMIP will focus on three periods:

1. The Early Eocene Climatic Optimum (EECO, $\sim$53–51 Ma)

2. The Palaeocene-Eocene Thermal Maximum (PETM, $\sim$55 Ma)

3. The period just before the PETM (pre-PETM, or latest Paleocene)

These three intervals have been the focus of numerous studies in the geological literature, and some syntheses of proxies from these intervals already exist (e.g. Huber and Caballero, 2011; Lunt et al., 2012; Dunkley Jones et al., 2013). The pre-PETM provides a reference point for both the PETM and the EECO. In addition, all three time periods can be referenced to modern or pre-industrial. This is in recognition that both modelling and proxies are strongest when considering relative changes.

Compared to earlier warm periods, such as the mid-Cretaceous, the palaeogeography during the early Eocene is reasonably well constrained, and freely available digital palaeogeographic datasets exist; however, there are wide uncertainties in estimates of atmospheric CO_2 at this time. Furthermore, due to the recent interest in the Eocene and PETM for providing information of relevance to future warming, there is a relative wealth of palaeoclimate proxy data with which the model results can be compared.

4 Experimental design

The DeepMIP experimental protocol consists of four main simulations (pre-industrial, two early Eocene, and one latest Pale-ocene), plus a number of optional sensitivity studies (see Section 4.3).

4.1 Pre-industrial simulation

The pre-industrial simulation should be as close to the CMIP6 standard as possible. Many groups will already have carried out this simulation as part of CMIP6. Some groups may need to make changes to their CMIP6 model configuration for the DeepMIP paleo simulations (for example changes to ocean diffusivity). If this is the case, we encourage groups to also carry out a non-CMIP6 preindustrial simulation with the model configuration used for DeepMIP paleo simulations.

4.2 Latest Paleocene and early Eocene simulations

This section describes the DeepMIP paleo simulations. There are three standard simulations, which differ only in their atmospheric CO_2 concentration, plus a number of optional sensitivity studies.

4.2.1 Palaeogeography and land-sea mask

Herold et al. (2014) (henceforth H14) is a peer-reviewed, traceable, freely-available digital reconstruction of the early Eocene interval. It includes land-sea mask, topography and sub-gridscale topography, bathymetry, tidal dissipation, vegetation, aerosol distributions and river runoff. The palaeogeography and land-sea mask from H14 should be used for the DeepMIP paleo simulations; they are provided digitally in netcdf format in the Supplementary Information of H14, at a resolution of $1° \times 1°$, and are illustrated here in Fig. 1. The palaeogeographic height should be applied as an absolute, rather than as an anomaly to the pre-industrial topography. Most models additionally require some fields related to the subgridscale orography to be provided. Because subgridscale orographies are very sensitive to the resolution of the underlying dataset, the subgridscale orography (if it is required by the model) can be estimated based on fields also provided in Supplementary Information of H14. This can be implemented as the modelling groups see fit, but care should be taken that the pre-industrial and Eocene subgridscale topographies are as consistent as possible. In addition, the code used to calculate the subgridscale orographies in the CESM model is also provided in the Supplementary Information of H14.

Included in the Supplementary Information of this paper are palaeorotations such that the modern location of gridcells in the early Eocene palaeogeography can be identified, as can the early Eocene location of modern gridcells.

We encourage sensitivity studies to the palaeogeography - see Section 4.3.2.

4.2.2 Land surface

(i) vegetation:

The vegetation in the DeepMIP paleo simulations should be prescribed as that in H14, which is included digitally as a netcdf file in the Supplementary Information of H14 (note that the BIOME4 vegetation should be used rather than the Sewall vege-
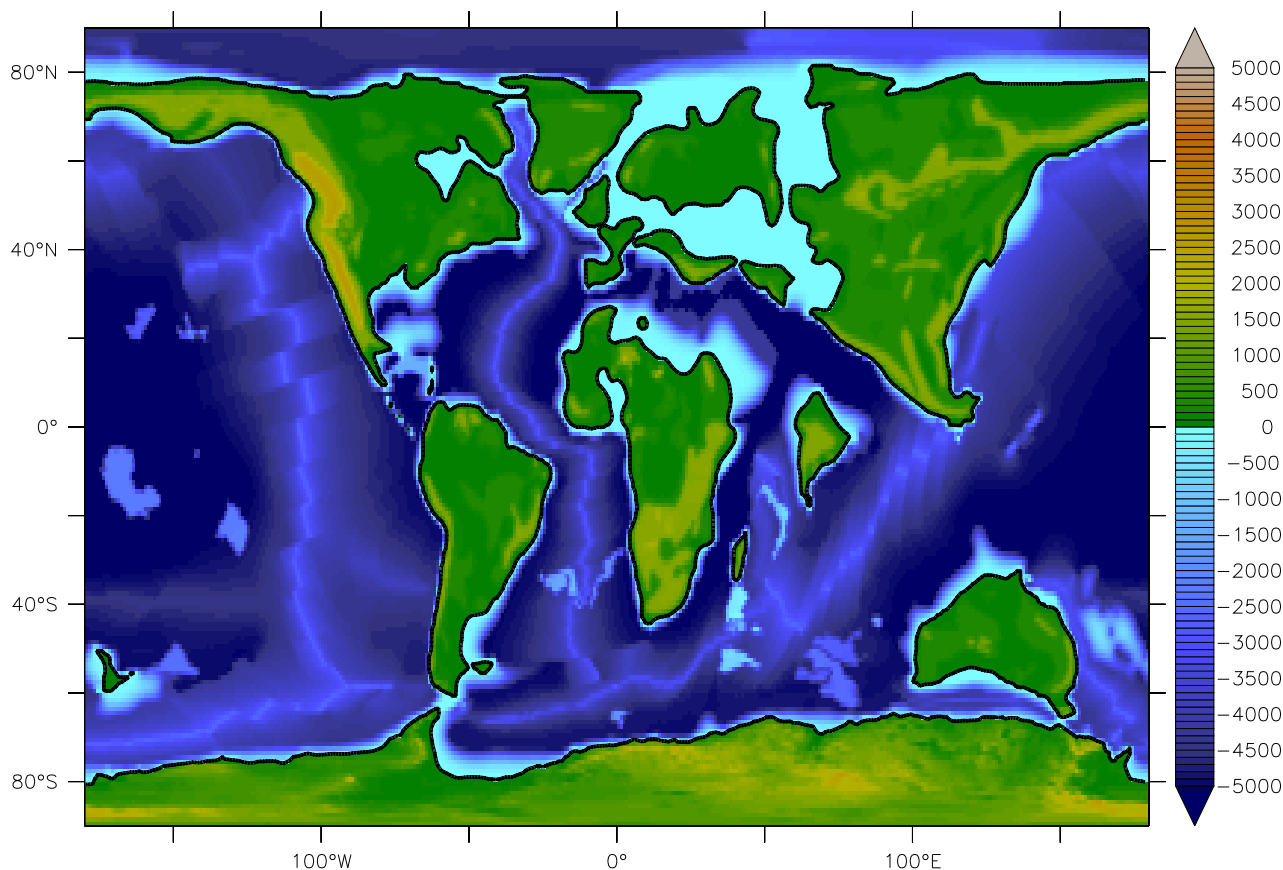
Figure 1. Orography and bathymetry for the paleo simulations in DeepMIP [metres]. A netcdf file of the data at a $1° \times 1°$ resolution is available in the Supplementary data of Herold et al. (2014)

tation, and that groups may choose to base their vegetation either on the 27 biomes or the 10 megabiomes), and shown here in Fig. 2. Groups should make a lookup table for converting the H14 Eocene dataset to a format which is appropriate for their model. To aid in this process, a modern vegetation dataset is also provided in the Supplementary Information of H14, using the same Plant Functional types as in the H14 Eocene reconstruction; in addition, the lookup table for the CLM land model is provided as a guide.

(ii) soils:

Soils should be classified homogeneously over the globe, with properties (e.g. albedo, water-holding capacity etc.) given by the global-mean of the group's pre-industrial simulation.

(iii) lakes:

No lakes should be prescribed in the DeepMIP paleo model simulations, unless these are predicted dynamically by the model.

(iv) river runoff:
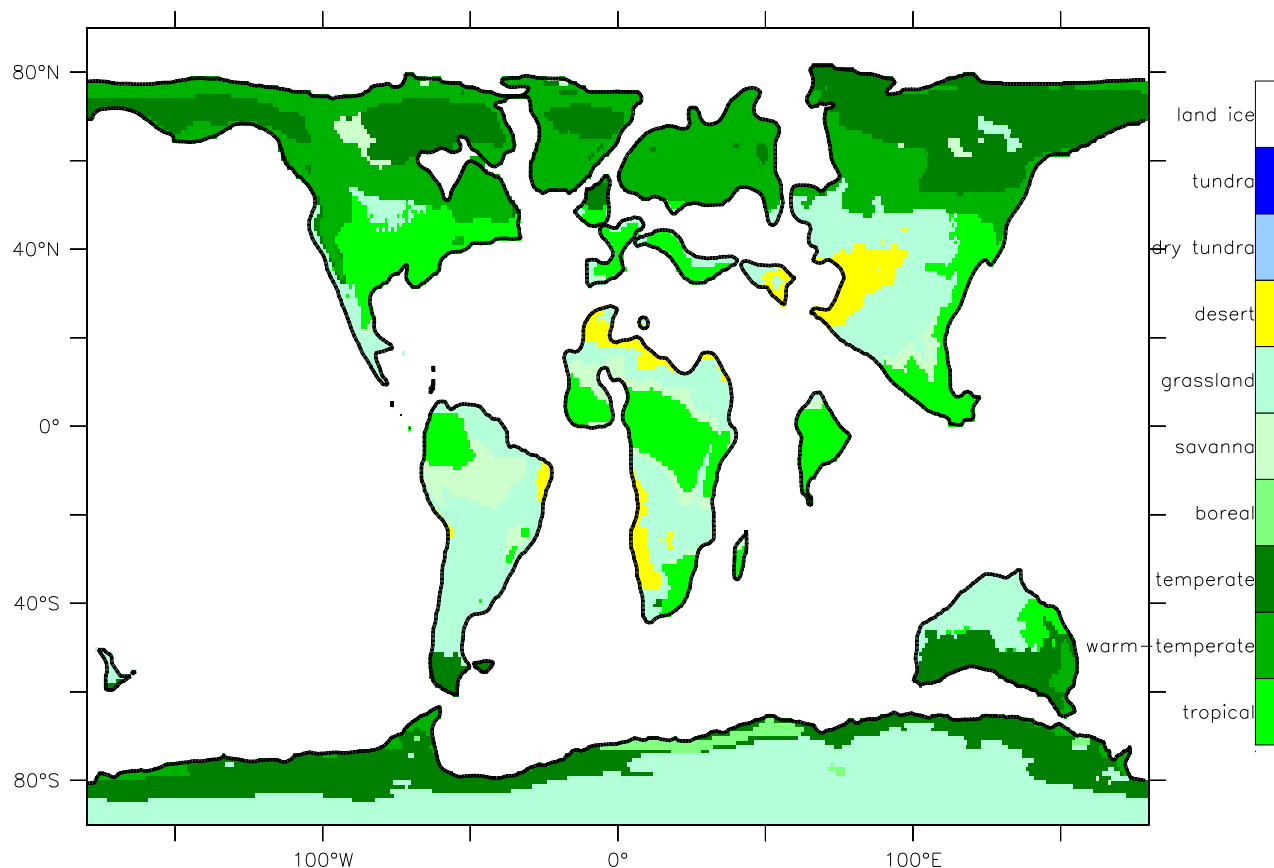
Figure 2. Vegetation, expressed as megabiomes, for the paleo simulations in DeepMIP. A netcdf file of the data at a $1° \times 1°$ resolution is available in the Supplementary data of Herold et al. (2014)

River runoff should be taken from the H14 reconstruction and included digitally as a netcdf file in the Supplementary information of H14.

4.2.3 Greenhouse gas concentrations

5 Each group should carry out three simulations at three different atmospheric CO_2 concentrations, expressed as multiples of the value in the pre-industrial simulation (typically $280\,\mathrm{ppmv}$, Section 4.1): (i) $3\times$ pre-industrial (typically $840\,\mathrm{ppmv}$), (ii) $6\times$ pre-industrial (typically $1680\,\mathrm{ppmv}$), and (iii) $12\times$ pre-industrial (typically $3360\,\mathrm{ppmv}$). Extant temperature records imply that, within uncertainty of the CO_2 proxies, CO_2 concentrations in the EECO and PETM were similar. As such, whereas the low-CO_2 simulation can be considered as representing the pre-PETM, the two higher CO_2 simulations are intended to

10 represent a range of possible PETM and EECO climate states. The values themselves are based primarily on recent work using boron isotopes (Anagnostou et al., 2016).

It is thought that non-CO_2 greenhouse gases during the early Eocene were elevated relative to pre-industrial, especially CH_4 (e.g., $\sim 3000\,\mathrm{ppbv}$, Beerling et al., 2011). However, there is considerable uncertainty as to exactly how elevated they were. Given these uncertainties, and the fact that we have chosen to use a modern solar constant as opposed to a reduced

5   solar constant (see Section 4.2.5) which would otherwise partially offset the CH_4 increase, all non-CO_2 greenhouse gases and trace gases should be set at the pre-industrial concentrations. In effect we assume that the CO_2 forcing represents the CO_2, CH_4 (and other non-CO_2 greenhouse gases), and solar forcings. Although a solar forcing and a CO_2 forcing have a differing regional expression, the response of the system in terms of surface temperature is similar (Lunt et al., 2008).

Some groups may find the higher CO_2 simulations problematic as some models are known to develop a runaway greenhouse

10   at high CO_2 (Malte Heinemann, pers comm). In this case, in addition to the $3\times$ simulation, groups can carry out simulations at $2\times$ and $4\times$.

If groups only have the computational resource to carry out two simulations, they should carry out the $3\times$ and $6\times$ simulations. For groups that can only carry out a single simulation, the analysis of the runs will be limited due to the focus on anomalies in DeepMIP, but we still encourage such groups to participate; in this case they should just carry out the $3\times$

15   simulation.

For groups with extensive computational resource, we encourage them to carry out additional sensitivity simulations over a range of CO_2 values, and in particular at $1\times$, see Section 4.3.1.

4.2.4 Aerosols

The representation of aerosols (including mineral dust) in Earth system models is undergoing a period of rapid development.

20   Therefore, we leave the implementation of aerosol fields or emissions rather flexible, and give several options. Groups may choose to (i) leave aerosol distributions or emissions identical to pre-industrial (taking account of the changed land-sea mask), or (ii) treat aerosols prognostically, or (iii) use aerosol emissions (including mineral dust) from H14, or (iv) use aerosol distributions from H14, or (v) some combination of the above, depending on the aerosol type. The crucial thing is that groups are asked to document exactly how they have implemented aerosols.

25   ### 4.2.5 Solar constant and orbital parameters

All simulations should be carried out with modern solar constant and orbital parameters. Although the early Eocene solar constant was $\sim 1359\,\mathrm{W\,m^{-2}}$ (Gough, 1981) compared with a modern value of $\sim 1365\,\mathrm{W\,m^{-2}}$, we choose to use a modern value in order to (i) aid comparison of any $1\times$ CO_2 simulations (see Section 4.3.1) with pre-industrial, and (b) to offset the absence of elevated CH_4 in the experimental design (see Section 4.2.3). As with all of Earth history, orbital conditions varied

30   throughout the early Eocene. There is some evidence that the PETM and other Palaeogene hyperthermals may have been paced by orbital forcing (Lourens et al., 2005; Lunt et al., 2011), but the phase of the response relative to the forcing is unknown. The modern orbit has relatively low eccentricity, and so represents a forcing close to the long-term average, and also facilitates comparison with the control pre-industrial simulation. However, we do encourage sensitiity studies to orbital configuration (see Section 4.3.3).

4.2.6 Initial conditions

(i) Atmosphere and land surface:

5 Simulations may be initialised with any state of the atmosphere and land surface, as long as the initial condition would not typically take longer than ∼50 years to spin up in a model with fixed SSTs; for example, initial snow cover should not be hundreds of metres depth.

(ii) Ocean:

Given that even with relatively long simulations, some vestiges of the initial ocean temperature and salinity structure will 10 remain at the end of the simulations, we strongly reccommend that all groups adopt the same initialisation procedure for the ocean, but encourage groups to carry out sensitivity studies to the initialisation (see Section 4.3.6). The ocean should be initialised as stationary, with no initial sea ice, and a zonally symmetric temperature (T, °C) and globally constant salinity (S, psu) distribution given by:

$$T\,[°\mathrm{C}] = \begin{cases} \left(\frac{5000-z}{5000}\,25\cos(\phi)\right) + 10 & \text{if } z \leq 5000\,\mathrm{m} \\ 10 & \text{if } z > 5000\,\mathrm{m} \end{cases}$$

15 $$S\,[\mathrm{psu}] = 34.7 \tag{1}$$

Where ϕ is latitude, and z is depth of the ocean (metres below surface).

Some groups have previously found that initialising the model with relatively cold ($<10\,°\mathrm{C}$) ocean temperatures at depth results in a relatively long spinup ($> 5000\,\mathrm{years}$), due to the suppression of convection – hence the relatively warm initial temperatures at depth prescribed here. Groups for which the reccommended initial temperature structure still results in a 20 stratififed ocean with little convection, and hence likely long equilibration timescales (for example those with a model with a particularly high climate senstivity), may wish to initialise their model with warmer deep ocean temperatures. If so, this should be clearly documented.

The value of $34.7\,\mathrm{psu}$ is the same as the modern mean ocean value. Although the lack of ice sheets in the Eocene would result in a decrease in mean ocean salinity relative to the modern of about $0.6\,\mathrm{psu}$, on these timescales geological cycling 25 also plays an important role; Hay et al. (2006) estimate mean ocean salinity to be between 35.1 and 36.5 during the Eocene. Given the uncertainties we choose a modern value for simplicity. If groups prefer to initialise salinity with a non-homogeneous distribution, or with a different absolute value, they may do this, but it should be documented

For simulations in which oxygen, carbon or other isotopic systems or passive tracers are included, these can be initialised as each individual group sees fit.

30 ### 4.2.7 Length of simulation

Simulations should be carried out for as long as possible. Ideally, simulations should be (a) at least 1000 years in length, and (b) have an inbalance in the top-of-atmosphere net radiation of less than $0.3\,\mathrm{W\,m^{-2}}$ (or have a similar inbalance to that of the

pre-industrial control), and (c) have SSTs which are not strongly trending (less than $0.1\,°C$ per century in the global mean). Climatologies should be calculated based on the final 100 years of the simulation.

4.2.8 Output format

Ideally, all output should be provided in CMIP6-compliant netcdf format, including the standard PMIP variables, and uploaded to the PMIP database. However, if this is not possible, then netcdf files of the variables in the Appendix should be uploaded to the DeepMIP Modelling Database, which will be set up if and when required. In any case, for the 'highest priority' variables in Appendix 1, all months of the simulations should be retained, such that averages can be calculated from arbitrary years of the simulation, and such that equilibrium states can be estimated using the approach of Gregory et al. (2004).

4.3 Sensitivity Studies

The above gives a summary of the four core simulations (pre-industrial and two early Eocene and one latest Paleocene). Below are some optional sensitivity studies that groups may wish to carry out, although there is no guarantee that other groups will do the same simulations.

4.3.1 Sensitivity to CO_2

Groups may wish to explore more fully the sensitivity of their model to CO_2, and associated non-linearities (Caballero and Huber, 2013), by carrying out additional simulations over a range of CO_2. Normally these would be multiples of the pre-industrial concentration, in addition to the standard $3\times$, $6\times$, and $12\times$ simulations. In particular, we encourage groups to carry out a $1\times$ simulation, for comparison with the pre-industrial control – this simulation enables the contribution of non-CO_2 forcings (palaeogeography and ice sheets) to early Eocene warmth to be evaluated.

4.3.2 Sensitivity to palaeogeography

Getech Plc (www.getech.com) have provided an alternative palaeogeographic reconstruction which may be used for sensitivity studies. It is included digitally in Lunt et al. (2016) as a netcdf file at a resolution of $3.75°$ longitude $\times$ $2.5°$ latitude. Because a high resolution version of this topography is not available, groups will need to use the subgridscale palaeogeography from the H14 reconstruction, and interpolate to the new land-sea mask as appropriate. The vegetation, river routing etc. from H14 will also need to be extrapolated to the new land-sea mask. Ideally groups would carry out these simulations at the same three CO_2 levels as in the standard simulations, but if groups can only carry out a limited number of simulations with this palaeogeography, they should carry them out in the following order of priority (highest priority first): $3\times$, $6\times$, $12\times$.

Both Getech and H14 use the plate rotation model of Müller et al. (2008), which is derived from relative plate motions tied to a mantle reference frame. van Hinsbergen et al. (2015) argue that for paleoclimate studies plate motions should be tied to the spin axis of the Earth using a paleomagnetic reference frame in order to obtain accurate estimates of paleolatitude. For

this reason we will provide an additional palaeogeography based on the methods outlined by van Hinsbergen et al. (2015) and Baatsen et al. (2015).

In addition, groups are encouraged to carry out sensitivity studies around the H14 palaeogeography. This may include
5   widening/constricting and shallowing/deepening key ocean gateways, raising/lowering mountain ranges, and changing the bathymetry of ocean shelves.

4.3.3   Sensitivity to orbit

Evidence of cyclicity during the Paleocene and early Eocene indicates that part of the warmth of the PETM may be orbitally forced on eccentricity timescales (Lourens et al., 2005; Westerhold et al., 2007; Galeotti et al., 2010). This is consistent with
10   the $\sim$50 kyr length of the core of the PETM. As such, we encourage sensitivity studies to orbital configuration. As the standard DeepMIP paleo simulations are with modern orbit, which has relatively low eccentricity, we suggest groups carry out additional simulations with high eccentricity ($e = 0.054$ compared with a modern value of $e = 0.017$), and northern hemisphere winter corresponding with both aphelion and perihelion.

4.3.4   Sensitivity to vegetation

15   For those groups with dynamic vegetation, they may carry out sensitivity studies using a dynamic vegetation component. The initial condition should be broadleaf or needleleaf trees at all locations. Ideally groups would carry out these simulations at the same three CO_2 levels as in the standard simulations, but if groups can only carry out a limited number of simulations with the dynamic vegetation, they should carry them out in the following order of priority (highest priority first): $3\times$, $6\times$, $12\times$. Groups with models which include a dynamic vegetation component can choose to pass to their vegetation model either the ambient
20   atmospheric CO_2, or a lower concentration if required for model stability.

4.3.5   Sensitivity to solar constant

Groups may wish to explore the relative radiative forcing of the solar luminosity compared with other forcings, by carrying out an Eocene simulation with reduced solar luminosity. An appropriate reduction would be from $1365\,\mathrm{W\,m^{-2}}$ in the modern to $1359\,\mathrm{W\,m^{-2}}$ in the Eocene (Gough, 1981). This would typically be carried out at a CO_2 level of $3\times$.

25   ### 4.3.6   Sensitivity to initialisation

We encourage groups to carry out sensitivity studies to the initialisation of the ocean temperature and salinity. It is possible that models will exhibit bistability with respect to initial condition, and as discussed in Section **??** we expect that the speed of equilibrium will a function of the initial conditions and will be different for different models.

4.3.7 'Best in Show'

Participants are invited to carry out simulations in which they attempt to best match existing proxy data. This may be done in a number of ways, for example by modifying the aerosols (Huber and Caballero, 2011), cloud properties (Kiehl and Shields, 2013), physics parameters (Sagoo et al., 2013), using very high CO_2 (Huber and Caballero, 2011), incorporating dynamic vegetation (Loptson et al., 2014), modifying gateways (Roberts et al., 2009), modifying orbital configuration, including non-CO_2 greenhouse gases, or a combination of the above and other modifications.

5   Climate Proxies

A major focus of DeepMIP will be to develop a new synthesis of climate proxy data for the latest Paleocene and early Eocene, focussing on the three targetted time intervals: pre-PETM, PETM and EECO. The main focus of DeepMIP will be on temperature and precipitation proxies. Two working groups have been set up to compile these data from marine and terrestrial records. These groups will also work together to generate new data sets for poorly documented regions, such as the tropics, and will seek multiple lines of evidence for climate reconstructions wherever possible. The marine working group is excited by the possibility of using innovative analytical techniques (e.g. Kozdon et al., 2013) to recover robust estimates for sea surface temperature from planktic foraminiferal assemblages within legacy sediment cores of the International Ocean Discovery Program. Published data sets will be combined into an open-access online database. The EECO and PETM/pre-PETM marine compilations of Lunt et al. (2012), Hollis et al. (2012), and Dunkley Jones et al. (2013), and EECO terrestrial compilations of Huber and Caballero (2011) provide a starting point for this database. One of the great challenges for these working groups will be develop new ways to assess proxy reliability and quantify uncertainties. In some cases, it may be more straightforward to consider relative changes in proxies rather than report absolute values. Proxy system modelling (Evans et al., 2013) coupled with Bayesian analysis (e.g. Khider et al., 2015; Tierney and Tingley, 2014) has great potential for improving estimation of uncertainties and directly linking our proxy compilation with the climate simulations. In addition to these quantitative estimates of uncertainty, all data will be qualitatively assessed based on expert opinion, for example characterising proxies as high, medium, or low confidence (as has been done in PlioMIP, see Dowsett et al. (2012)).

6   Products

In addition to this experimental design paper, and papers describing the new proxy syntheses, once the model simulations are complete we anticipate producing overarching papers describing the 'large-scale features' of the model simulations, and model–data comparisons. Following this, we anticipate a number of spin-off papers looking at various other aspects of the model simulations (e.g., ENSO, ocean circulation, monsoons etc.). Furthermore, we will encourage modelling participants to publish individual papers which describe their own simulations in detail, including how the boundary conditions were implemented. In this respect we are basing our dissemination strategy on that of PlioMIP http://www.geosci-model-dev.net/special_issue5.html (Haywood et al., 2013).

5 # 7 Data availability

The boundary conditions for the DeepMIP paleo simulations are supplied as Suppleemntary Information in H14 (Herold et al., 2014).

Appendix A: Output variables

If the PMIP database is not used, the variables below should be submitted to the (yet to exist) DeepMIP Model Database.
10 Climatological averages of the final 100 years of the simulation should be supplied for each month (12 fields for each variable). In addition, for the highest priority variables, all months of the simulation should be supplied.

Furthermore, as many groups are interested in hydrological extremes, groups should aim to produce ten years of hourly precipitation, evaporation and runoff data.

Author contributions. A first draft of this paper was written by Dan Lunt and Matt Huber. It was subsequently edited based on discussions at a DeepMIP meeting in January 2016 at NCAR, Boulder, Colorado, USA., and following further email discussions with the DeepMIP community.

Acknowledgements. We thank NERC grant NE/N006828/1 for providing funds for the first DeepMIP meeting.

Table 1. Atmosphere variables

Variable	Units	Highest priority
Near surface (1.5 m) air temperature	$^\circ$C	X
Surface skin temperature	$^\circ$C	
Precipitation	$\mathrm{kg\,m^2\,s^{-1}}$	X
Total evaporation	$\mathrm{kg\,m^2\,s^{-1}}$	
Total cloud cover	[0,1]	
FLNS	$\mathrm{W\,m^{-2}}$	
FLNT	$\mathrm{W\,m^{-2}}$	X
FSDS	$\mathrm{W\,m^{-2}}$	
FSNS	$\mathrm{W\,m^{-2}}$	
FSNT	$\mathrm{W\,m^{-2}}$	X
FSDT	$\mathrm{W\,m^{-2}}$	
sensible heat flux	$\mathrm{W\,m^{-2}}$	
latent heat flux	$\mathrm{W\,m^{-2}}$	
Near surface (10 m) u wind	$\mathrm{m\,s^{-1}}$	
Near surface (10 m) v wind	$\mathrm{m\,s^{-1}}$	
surface wind stress (x)	$\mathrm{N\,m^{-2}}$	
surface wind stress (y)	$\mathrm{N\,m^{-2}}$	
mean sea-level pressure	Pa	
surface pressure	Pa	
u winds on model atmospheric levels	$\mathrm{m\,s^{-1}}$	
v winds on model atmospheric levels	$\mathrm{m\,s^{-1}}$	
w winds on model atmospheric levels	$\mathrm{m\,s^{-1}}$	
u wind at 200 mbar	$\mathrm{m\,s^{-1}}$	
v wind at 200 mbar	$\mathrm{m\,s^{-1}}$	
u wind at 500 mbar	$\mathrm{m\,s^{-1}}$	
v wind at 500 mbar	$\mathrm{m\,s^{-1}}$	
u wind at 850 mbar	$\mathrm{m\,s^{-1}}$	
v wind at 850 mbar	$\mathrm{m\,s^{-1}}$	
geopotential height at 200 mbar	m	
geopotential height at 500 mbar	m	
geopotential height at 850 mbar	m	
temperature at 200 mbar	$^\circ$C	
temperature at 500 mbar	$^\circ$C	
temperature at 850 mbar	$^\circ$C	
specific humidity at 200 mbar	$\mathrm{kg\,kg^{-1}}$	
specific humidity at 500 mbar	$\mathrm{kg\,kg^{-1}}$	
specific humidity at 850 mbar	$\mathrm{kg\,kg^{-1}}$	

N.B. F*XYZ* notation

F = flux

X = S(hortwave) or L(ongwave)

Y = D(own) or N(et)

Z = S(urface) or T(op of atmosphere)

5 References

Anagnostou, E., John, E., Edgar, K., Foster, G., Ridgwell, A., Inglis, G., Pancost, R., Lunt, D., and Pearson, P.: Changing atmospheric CO2 concentration was the primary driver of early Cenozoic climate, Nature, doi:10.1038/nature17423, 2016.

Table 2. Ocean Variables

Variable	Units	Highest priority
SST	$^{\circ}$C	X
sea-ice fraction	[0,1]	X
u,v,w on model levels	cm s^{-1}	
T on model levels	$^{\circ}$C	
S on model levels	psu	
barotropic streamfunction	cm^3 s^{-1}	
mixed-layer depth	m	
global overturning streamfunction	Sv	

Table 3. Boundary conditions

Variable	Units
land-sea mask	[0,1]
topography	m
bathymetry	m

Baatsen, M., van Hinsbergen, D. J. J., von der Heydt, A. S., Dijkstra, H. A., Sluijs, A., Abels, H. A., and Bijl, P. K.: A generalised approach to reconstructing geographical boundary conditions for palaeoclimate modelling, Climate of the Past Discussions, 11, 4917–4942,

10   doi:10.5194/cpd-11-4917-2015, http://www.clim-past-discuss.net/11/4917/2015/, 2015.

Beerling, D. J., Fox, A., Stevenson, D. S., and Valdes, P. J.: Enhanced chemistry-climate feedbacks in past greenhouse worlds, Proceedings of the national Academy of Sciences, 108, 9770–9775, 2011.

Caballero, R. and Huber, M.: State-dependent climate sensitivity in past warm climates and its implications for future climate projections, PNAS, 110, 14 162–14 167, 2013.

15   Carmichael, M. J., Lunt, D. J., Huber, M., Heinemann, M., Kiehl, J., LeGrande, A., Loptson, C. A., Roberts, C. D., Sagoo, N., Shields, C., Valdes, P. J., Winguth, A., Winguth, C., and Pancost, R. D.: A model–model and data–model comparison for the early Eocene hydrological cycle, Climate of the Past, 12, 455–481, doi:10.5194/cp-12-455-2016, http://www.clim-past.net/12/455/2016/, 2016.

Dowsett, H. J., Robinson, M. M., Haywood, A. M., Hill, D. J., Dolan, A. M., Stoll, D. K., Chan, W.-L., Abe-Ouchi, A., Chandler, M. A., Rosenbloom, N. A., Otto-Bliesner, B. L., Bragg, F. J., Lunt, D. J., Foley, K. M., and Riesselman, C. R.: Assessing confidence in Pliocene

20   sea surface temperatures to evaluate predictive models, Nature Climate Change, 2, 365–371, 2012.

Dunkley Jones, T., Lunt, D., Schmidt, D., Ridgwell, A., Sluijs, A., Valdes, P., and Maslin, M.: Climate model and proxy data constraints on ocean warming across the Paleocene-Eocene Thermal Maximum, Earth Science Reviews, p. doi.org/10.1016/j.earscirev.2013.07.004, 2013.

Evans, M. N., Tolwinski-Ward, S. E., Thompson, D. M., and Anchukaitis, K. J.: Applications of proxy system modeling in high resolution

5   paleoclimatology, Quaternary Science Reviews, 76, 16–28, 2013.

Galeotti, S., Krishnan, S., Pagani, M., Lanci, L., Gaudio, A., Zachos, J. C., nd G. Morelli, S. M., and Lourens, L.: Orbital chronology of Early Eocene hyperthermals from the Contessa Road section, central Italy, Earth and Planetary Science Letters, 290, 192–200, 2010.

Gasson, E., Lunt, D. J., DeConto, R., Goldner, A., Heinemann, M., Huber, M., LeGrande, A. N., Pollard, D., Sagoo, N., Siddall, M., Winguth, A., and Valdes, P. J.: Uncertainties in the modelled CO_2 threshold for Antarctic glaciation, Climate of the Past, 10, 451–466, doi:10.5194/cp-10-451-2014, http://www.clim-past.net/10/451/2014/, 2014.

Gough, D.: Solar interior structure and luminosity variations, Sol, Phys., 74, 21–34, 1981.

Gregory, J. M., W. J. Ingram, W. J., Palmer, M. A., Jones, G. S., Stott, P. A., Thorpe, R. B., Lowe, J. A., Johns, T. C., and Williams, K. D.: A new method for diagnosing radiative forcing and climate sensitivit, Geophysical Research Letters, 31, L03 205, doi:10.1029/2003gl018747, 2004.

Hay, W. W., Migdisov, A., Balukhovsky, A. N., Wold, C. N., Flögel, S., and Söding, A.: Evaporites and the salinity of the ocean during the Phanerozoic: Implications for climate, ocean circulation and life, Palaeogeography, Palaeoclimatology, Palaeoecology, 240, 3–46, 2006.

Haywood, A. M., Hill, D. J., Dolan, A. M., Otto-Bliesner, B. L., Bragg, F., Chan, W.-L., Chandler, M. A., Contoux, C., Dowsett, H. J., Jost, A., Kamae, Y., Lohmann, G., Lunt, D. J., Abe-Ouchi, A., Pickering, S. J., Ramstein, G., Rosenbloom, N. A., Salzmann, U., Sohl, L., Stepanek, C., Ueda, H., Yan, Q., and Zhang, Z.: Large-scale features of Pliocene climate: results from the Pliocene Model Intercomparison Project, Climate of the Past, 9, 191–209, doi:10.5194/cp-9-191-2013, http://www.clim-past.net/9/191/2013/, 2013.

Herold, N., Buzan, J., Seton, M., Goldner, A., Green, J. A. M., Müller, R. D., Markwick, P., and Huber, M.: A suite of early Eocene (55 Ma) climate model boundary conditions, Geoscientific Model Development, 7, 2077–2090, doi:10.5194/gmd-7-2077-2014, http://www.geosci-model-dev.net/7/2077/2014/, 2014.

Hollis, C. J., Taylor, K. W. T., Handley, L., Pancost, R. D., Huber, M., Creech, J., Hines, B., Crouch, E. M., Morgans, H. E. G., Crampton, J. S., Gibbs, S., Pearson, P., and Zachos, J. C.: Early Paleogene temperature history of the Southwest Pacific Ocean: reconciling proxies and models, Earth and Planetary Science Letters, 349-350, 53–66, 2012.

Huber, M. and Caballero, R.: The early Eocene equable climate problem revisited, Climate of the Past, 7, 603–633, 2011.

Khider, D., Huerta, G., Jackson, C., Stott, L. D., and Emile-Geay, J.: A Bayesian, multivariate calibration for Globigerinoides ruber Mg/Ca, Geochem. Geophys. Geosyst., 16, 2916–2932, doi:10.1002/2015GC005844, 2015.

Kiehl, J. T. and Shields, C. A.: Sensitivity of the PalaeoceneEocene Thermal Maximum climate to cloud properties, Phil. Trans. R. Soc. A, 371, doi:10.1098/rsta.2013.0093, 2013.

Kozdon, R., Kelly, D. C., Kitajima, K., Strickland, A., Fournelle, J. H., and Valley, J. W.: In situ d18O and Mg/Ca analyses of diagenetic and planktic foraminiferal calcite preserved in a deep-sea record of the Paleocene-Eocene Thermal Maximum, Paleoceanography, 28, 517–528, doi:10.1002/palo.20048, 2013.

Loptson, C. A., Lunt, D. J., and Francis, J. E.: Investigating vegetation-climate feedbacks during the early Eocene, Clim. Past, 10, 419–436, 2014.

Lourens, L. J., Sluijs, A., Kroon, D., Zachos, J. C., Thomas, E., Rohl, U., Bowles, J., and Raffi, I.: Astronomical pacing of late Palaeocene to early Eocene global warming events, Nature, 435, 1083–1087, 2005.

Lunt, D., Ridgwell, A., Sluijs, A., Zachos, J., Hunter, S., and Haywood, A.: A model for orbital pacing of methane hydrate destabilization during the Palaeogene, Nature Geoscience, 4, 775–778, 2011.

Lunt, D. J., Ridgwell, A. R., Valdes, P. J., and Seale, A.: Sunshade World: A fully coupled GCM evaluation of the climatic impacts of geoengineering, Geophysical Research Letters, 35, doi:10.1029/2008GL033 674, 2008.

Lunt, D. J., Dunkley Jones, T., Heinemann, M., Huber, M., LeGrande, A., Winguth, A., Loptson, C., Marotzke, J., Roberts, C. D., Tindall, J., Valdes, P., and Winguth, C.: A model-data comparison for a multi-model ensemble of early Eocene atmosphere-ocean simulations: EoMIP, Climate of the Past, 8, 1717–1736, 2012.

Lunt, D. J., Elderfield, H., Pancost, R., Ridgwell, A., Foster, G., Haywood, A., Kiehl, J., Sagoo, N., and Stone, E.J. Valdes, P.: Warm climates of the past - a lesson for the future?, Phil. Trans. R. Soc. A, 371, doi:10.1098/rsta.2013.0146, 2013.

10 Lunt, D. J., Farnsworth, A., Loptson, C., Foster, G. L., Markwick, P., O'Brien, C. L., Pancost, R. D., Robinson, S. A., and Wrobel, N.: Palaeogeographic controls on climate and proxy interpretation, Climate of the Past, 12, 1181–1198, doi:10.5194/cp-12-1181-2016, http://www.clim-past.net/12/1181/2016/, 2016.

Müller, R. D., Sdrolias, M., Gaina, C., and Roest, W. R.: Age, spreading rates, and spreading asymmetry of the world's ocean crust, Geochem. Geophys. Geosyst., 9, Q04 006, doi:10.1029/2007GC001743, 2008.

15 Roberts, C. D., LeGrande, A. N., and Tripati, A. K.: Climate sensitivity to Arctic seaway restriction during the early Paleogene, Earth and Planetary Science Letters, 286, 576–585, 2009.

Sagoo, N., Valdes, P., Flecker, R., and Gregoire, L.: The Early Eocene equable climate problem: can perturbations of climate model parameters identify possible solutions?, Phil. Trans. R. Soc. A, 371, doi:10.1098/rsta.2013.0123, 2013.

Tierney, J. E. and Tingley, M. P.: A Bayesian, spatially-varying calibration model for the TEX86 proxy, Geochimica et Cosmochimica Acta, 350   127, 83–106, doi:83-106, 2014.

van Hinsbergen, D. J. J., de Groot, L. V., van Schaik, S. J., Spakman, W., Bijl, P. K., Sluijs, A., Langereis, C. G., and Brinkhuis, H.: A Paleolatitude Calculator for Paleoclimate Studies, PLoS ONE, 10, e0126 946, doi:10.1371/journal.pone.0126946, 2015.

Westerhold, T., Rohl, U., Laskar, J., Raffi, I., Bowles, J., Lourens, L. J., and Zachos, J. C.: On the duration of magnetochrons C24r and C25n and the timing of early Eocene global warming events: Implications from the Ocean Drilling Program Leg 208 Walvis Ridge depth 355   transect, Paleoceanography, 22, PA2201, 2007.