Data-constrained assessment of ocean circulation changes since the middle Miocene in an Earth system model

Katherine A. Crichton[1]*, Andy Ridgwell[2], Daniel J. Lunt[3], Alex Farnsworth[3], Paul N. Pearson[1]

[1]School of Earth and Ocean Sciences, Cardiff University, UK
5   2Department of Earth and Planetary Sciences, University of California, Riverside, CA 92521, USA
3School of Geographical Sciences, University of Bristol, Bristol, UK
* Now at School of Geography, University of Exeter, EX4 4RJ, UK

Correspondence to: Katherine A. Crichton (k.a.crichton@exeter.ac.uk)

Abstract. Since the middle Miocene, 15 Ma (million years ago), the Earth's climate has undergone a long-term cooling trend,
10   characterised by a reduction in sea surface temperatures by over 6 °C, with 4 to 6 °C cooling occurring in the deep ocean. The causes of this cooling are primarily thought to be linked to changes in ocean circulation due to tectonic plate movements affecting ocean seaways, together with a drop in atmospheric greenhouse gas forcing (and attendant ice-sheet growth and feedback). In this study we assess the potential to constrain, using marine sediment proxy data, the evolving patterns of global ocean circulation and cooling of surface climate over the last 15 million years (Ma) in an Earth system model. We do this by
15   compiling surface and benthic ocean temperature and benthic carbon-13 data in a series of seven time-slices spaced at approximately 2.5 million year intervals. We pair this with a corresponding series of seven tectonic and surface climate boundary condition reconstructions in the cGENIE ('muffin' release) Earth system model. In the cGENIE model, we adjust atmospheric CO_2 together with the magnitude of North Pacific to North Atlantic salinity flux adjustment in a series of 2D parameter ensembles in order to match global temperature and benthic $\delta^{13}C$ patterns in the model to the data. We identify that
20   a relatively high CO_2 equivalent forcing of 1120ppm is required at 15 Ma in cGENIE to reproduce proxy temperature estimates in the model, noting that this CO_2 forcing is dependent on cGENIEs climate sensitivity (which is as the present day) and that it incorporates the effects of all greenhouse gases. The required CO_2 forcing progressively reduces throughout the subsequent six time slices delineating the observed long-term cooling trend. In order to match the evolving patterns of the proxy data, we require fundamental change in the mode of ocean circulation at 12.5 Ma with present-day-like benthic $\delta^{13}C$ trends established
25   by 10 Ma. We also find a general increasing strength of Atlantic overturning despite a reduction in salinity of the surface North Atlantic over the cooling period, attributable to falling intensity of the hydrological cycle and polar cooling caused by CO_2-driven global cooling.

1 Introduction and background

Since the middle Miocene (~ 15 Ma) the Earth has experienced a period of global cooling, with the expansion of ice sheets
30   over Antarctica and later the establishment of the Greenland and North American ice sheet (Zachos et al., 2008; Cramer et al.,

2011) together with the intensification of glacial-interglacial cycles in the Pleistocene. Fundamental changes have also occurred in ocean circulation patterns (Butzin et al., 2011) during this interval of global cooling. Specifically – the Atlantic Meridional Overturning Circulation (AMOC), that today redistributes heat to the Northern hemisphere, became established in its current form sometime after the middle Miocene (Sepulchre et al., 2013; Bell et al., 2015). In turn, this is suspected to have

35 been linked to the closing of the seaway between the Atlantic and the Pacific with the creation of the Isthmus of Panama (Lunt et al., 2008; Montes et al., 2015; O'Dea et al., 2016; Jaramillo et al., 2017) that finally closed the Central American Seaway (CAS).

Other seaways have also been tectonically transformed since the Miocene, particularly: the disappearance of the Tethys Sea due to the northward movement of Africa (Hamon et al., 2013); the restriction of the Indonesian seaway with the

40 northward movement of Australia (Srinivasan et al., 1998); the widening of the Drake Passage driven by the northward movement of South America relative to Antarctica (Lagabrielle et al., 2009). Associated with these plate movements, the Miocene to Holocene was characterised by significant mountain building which may have played a direct role in the draw-down of atmospheric CO_2 via weathering, and hence progressive cooling (Filippelli et al., 1997; Raymo et al., 1998).

Terrestrial proxies for temperature indicate that the Miocene was significantly warmer than the present day (Pound

45 et al., 2012 and references therein). Marine data also indicate a significantly warmer-than-present Miocene climate with surface ocean temperatures over 6 °C warmer than present (Stewart et al., 2004; Herbert et al., 2016), deep ocean temperature 4 to 6 °C warmer (Cramer et al., 2011), and atmospheric CO_2 levels at 470 to 630ppm in the middle Miocene (Sosdian et al., 2018). Modelling efforts of the Miocene find that reconciling the combined constraints of ocean temperature, CO_2 indicators and Antarctic ice sheet dynamics is a non-trivial task (Micheels et al., 2009; Henrot et al., 2010, Bradshaw et al., 2012; Sijp et al.,

50 2014), with vegetation seemingly integral in creating the Miocene climatic conditions (Henrot et al., 2010; Knorr et al., 2011; Krapp and Jungklaus, 2011; Micheels et al., 2011; Bradshaw et al., 2015) along with bathymetry, topography and CO_2.

Sea-floor proxy data have been used extensively to understand patterns of ocean circulation and climate in the Miocene (Butzin et al 2011). Oxygen-18 data has been used as a proxy for temperature and global ice volume, often combined with Mg/Ca (as a further temperature proxy) (Cramer et al., 2011; Hauptvogel et al., 2012; Badger et al., 2013; Lear et al.,

55 2015; Pierce et al., 2017). Carbon-13 data has principally been used as a circulation tracer (Lynch_Stieglitz 2003), with high-resolution data used to understand timing of changes in (local or regional) climate, and linking this to changes in orbital forcing (e.g. Holbourn et al., 2005). With the increasing availability of data, global distributions of $\delta^{13}C$ and differences in ocean basin means, or regional $\delta^{13}C$ has been used as further evidence of changes in circulation since the middle Miocene and of the respective dominance of certain water masses (e.g. Cramer et al., 2011; Butzin et al., 2011). These studies emphasise the

60 importance of changing bathymetry and the closing or opening of ocean seaways, with the Drake Passage (affecting the Antarctic Circumpolar Current, ACC) and the Central American Seaway (affecting the North Atlantic) likely the most important determinants in global circulation patterns since the middle Miocene.

In this paper we aim to explore to what extent proxy data can constrain changing global climate and ocean circulation patterns in an Earth system model. We do not provide an exhaustive sensitivity study of Miocene seaways and the carbon cycle

65 here, but instead aim to combine constraints of temperature and ocean circulation to create plausible and self-consistent paleo realisations of climate and ocean circulation, for each of seven time-slices spanning the middle Miocene to present (late Holocene). Our focus in this paper is hence on the model-data methodology and the self-consistency and plausibility of the outcome. For an overview of Miocene circulation and the effects of seaways on that circulation, e.g. see Butzin et al. (2011) and Sijp et al. (2014).

2 Model-data Methodology

We employ foraminifera proxy data for: surface ocean temperature, benthic ocean temperature, and benthic ocean $\delta^{13}C$, and compile this for seven time-slices, in order to constrain the applied climate forcing and ocean circulation in the cGENIE.muffin Earth system model from the mid Miocene to the present (late Holocene) (Table 1). The model climate forcing we employ is in the form of atmospheric CO_2, that is to be understood as an equivalent CO_2 forcing that encompasses all atmospheric
75 greenhouse gases (esp. methane). Ocean circulation is a function of surface ocean boundary conditions (esp. wind stress) and climate (and the atmospheric CO_2 forcing), ocean bathymetry, and the existence and nature of seaways and gateways – all of which we adjust at each time-slice. We further introduce a variable to the model that alters the salinity transfer between North Pacific and North Atlantic – a classic 'flux adjustment'. This represents the effect of atmospheric moisture transport and the relative precipitation that East vs. West draining watersheds receive on the north American continental land mass, none of
80 which can be reproduced well in the simple 2D energy-moisture balance model based approach employed in the version of the cGENIE model we use (Edwards and Marsh, 2005; Marsh et al., 2011). This 'flux correction' results in an Atlantic Meridional Overturning Circulation (AMOC) being induced (and/or strengthened) in cGENIE and indeed is required in the standard present-day configuration (Ridgwell et al., 2007; Marsh et al., 2011) in order to rectify the aforementioned simplification in modelled atmospheric moisture transport and dynamics.

85 Three different paleo/proxy datasets were compiled: surface temperature, benthic $\delta^{18}O$ and benthic $\delta^{13}C$. The surface and sea-floor temperature data allows us to evaluate the model skill in reproducing ocean heat distribution, and hence the models simulated global-scale pattern of ocean circulation. It also provides a means of determining the atmospheric forcing (in terms of CO_2) that produces surface and deep ocean temperatures that best match the data.

The isotope of carbon, carbon-13, has been used as a tracer for paleo ocean circulation for many years (Lynch-Stieglitz
90 2003). It is a stable isotope, heavier than carbon-12 and accounts for about 1% of all carbon on Earth. The ratio of carbon-13 to carbon-12, designated as "$\delta^{13}C$" – the divergence from a standard in units of parts-per-thousand (‰), can be estimated for paleo ocean waters by measuring the $\delta^{13}C$ in shells of foraminifera formed in those paleo water masses. Different water masses have characteristic $\delta^{13}C$ signatures, which depend on biological activity (in the ocean carbon pumps) and critically for this study, also on ocean circulation patterns. Shells of dead foraminifera gradually accumulate on the ocean floor, thus providing
95 a record of changes in water column chemistry over time at that location in the form of ocean sediments. In general, changes

in $\delta^{13}C$ seen in these records are interpreted as changes of water masses dominant in that location, and therefore as a means of reconstructing paleo ocean circulation.

2.1 Datasets

100 Published surface temperature data selected are those using either alkenones or TEX86 for all seven slices (Figure 1 for surface temperature data locations and Supplementary material [A] for data) noting that these proxies, like any proxy, are subject to uncertainties and limitations (Richey and Tierney 2016). For slices 2.5Ma to 15Ma, published benthic $\delta^{13}C$ and $\delta^{18}O$ data for *Cibicidoides* and *Planulina* foraminifera species were selected, with at least 15 data-points for each time-slice, covering at least the Atlantic and Pacific basins (figure 1 for benthic data locations, figure 2 for $\delta^{13}C$ data plotted by paleo-latitude,

105 Supplementary material [B] for data). These species were selected so that temperature could be calculated from $\delta^{18}O$ using Marchitto et al. (2014). Final benthic temperatures calculated from $\delta^{18}O$ take account of the effect of benthic water salinity on $\delta^{18}O$ which is affected by ocean circulation (the temperatures in Table S2 are uncorrected for salinity). The method for this is described in detail in Appendix A; we use the models benthic salinity field as a correction to the calculated temperature from Marchitto et al (2014). A further dataset for benthic $\delta^{13}C$ for the Holocene from Peterson et al. (2014) was applied for the

110 Holocene time-slice.

Paleo-locations for each data point were found using the reconstructions in www.paleolocation.org (Urban and Hardisty 2013) that provide Zanclean (4.466 Ma), Tortonian (9.427 Ma) and Burdigalian (18.2 Ma) paleo-locations using PLATES reconstructions (University of Texas). Thereafter, the paleo-locations for each data-point at our time-slices (Table 1) were interpolated using these three points together with the present day location in a cubic model regression. To account for

115 age model uncertainties, a window of ±1 million years was inspected for each data-point to ensure the data value was not unrepresentative of the general value at that site around the age point. All the data points and their paleo-locations, and the ±million year time-series plotted for each data location are available in Supplementary Material, Fig. S1.

2.2 Overview of the cGENIE.muffin Earth system model

We run the freely available intermediate complexity Earth system model cGENIE ('muffin' release), which is hosted on

120 GitHub (https://github.com/derpycode/cgenie.muffin). As implemented here, this comprises: (1) a 3D ocean circulation model component, typically configured on a 36×36 equal area grid, with 16 non-equally spaced vertical levels in the ocean, (2) a 2D energy-moisture balance model ('EMBM') component, and (3) a 2D dynamic-thermodynamic sea-ice model component. These three individual components and their coupling are described in Marsh et al. (2011) (and references therein). The basic physics parameter calibration of the climate model component is as per Cao et al. (2009) unless described otherwise (below).

125 For the modern Atlantic basin configuration (and, as per the findings of this paper, also for the last ca. 10 Ma), the lack of atmospheric dynamics and explicit topography on land in the model requires that a freshwater flux adjustment is applied. This

is implemented by transferring salinity to the North Atlantic region from the North Pacific, as described in Edwards and Marsh (2005).

Our implementation of the cGENIE model includes a relatively complete description of the cycling of carbon and
130 oxygen in the ocean plus exchange with the atmosphere, as described in Ridgwell et al. (2007) and Cao et al. (2009). In addition, the carbon isotopic (δ^{13}C) composition of all the carbon pools plus associated fractionations are represented, as described in Ridgwell et al. (2007) with additional description and evaluation in Kirtland Turner and Ridgwell (2016). Given the paucity of constraints on the evolving patterns of aeolian fluxes to the ocean surface, we omit a marine iron cycle and only a single nutrient, phosphate, potentially limiting to biological productivity in the ocean is included in our configuration.

135 **2.3 Model boundary conditions**

For this paper we create a series of new continental and surface boundary condition configurations of the cGENIE model – for seven time-slices spanning the late Holocene through to mid Miocene. In the original modern configuration of the GENIE (and later cGENIE.muffin) model, ocean bathymetry, land-sea mask, and required ocean circulation files (defining 'islands' and circulation paths around islands) were derived from global topographic observations (ETOPO5) and the topography
140 subsequently filtered (Edwards and Marsh, 2005). Because the atmospheric EMBM component lacks clouds and dynamics (e.g. winds), further fields must be provided as fixed annual average boundary conditions. Firstly, a zonally-averaged planetary albedo profile is applied (as a simple cosine function of latitude in Edwards and Marsh (2005)). Secondly, fields for: (a) wind stress on the ocean surface, (b) wind velocity in the atmosphere, and (c) short-term wind speed, are derived from modern observations and applied to (a) drive surface ocean circulation, (b) transport heat and moisture in the atmosphere and for
145 calculation of heat and moisture exchange between ocean surface and atmosphere, and (c) in calculating air-sea gas exchange, respectively.

In this paper and as per in previous (deeper time) paleo applications of the cGENIE.muffin model (e.g. Ridgwell and Schmidt, 2010), rather than observations, we derive the required boundary conditions from a representative fully coupled GCM experiment. The software suite we employ to do this is called 'muffingen' and is hosted on GitHub
150 (https://github.com/derpycode/muffingen). This software takes a specific GCM experiment as input and does the following:

1. Creates input (from GCM) and output (here: 36×36 with 16 levels in the ocean) grids.
2. Using (1), derives a land-sea mask and also ocean bathymetry on the output grid. The land-sea mask is lightly filtered to prevent the occurrence of isolated in-land seas and single-cell width coastal embayment, while the ocean bathymetry is filtered to avoid single cell 'holes' occurring in the ocean floor.
155 3. Generates drainage basins determining where precipitation and hence in which direction runoff is directed towards the ocean. The specific scheme used here is known as a 'roofing scheme' and operates to create a watershed approximately equidistant from the coast.
4. Derives island and ocean paths files required by the ocean circulation model.

5. Re-grids GCM wind stress and (10 m) wind velocity to the output grid (which for wind stress, means re-gridding to
    both u and v edge grids). Wind speed is calculated from the mean annual wind velocity components.

6. Re-grids the GCM planetary albedo.

The GCM simulations underlying our cGENIE model configurations were carried out using HadCM3LM2.1aE,
which is described in detail in Valdes et al. (2017). The models are constrained with paleogeographies and solar constant
appropriate for each geological stage in the Miocene, and a CO_2 mixing ratio of 400 ppmv. The experimental design is
described in detail in Farnsworth et al. (2019).

For all configurations we generally follow the default re-gridding algorithm of the muffingen software in order not to
impose prior assumptions regarding the importance of specific ocean features, meaning that for most of the reconstructed
continental configurations, a Mediterranean Sea is not present in cGENIE. Only in the 12.5 and 15.0 Ma time-slices is the
remnant Tethys Ocean sufficiently expansive to re-grid as an ocean basin at the selected 36×36 (16 levels) cGENIE model
resolution. We do, however, make the following manual interventions in the generation of the land-sea mask (but not in ocean
bathymetry):

*   For 0 Ma (late Holocene) – the Panama Isthmus as well as the tip of S America is made continuous.
*   For 10.0, 12.5, and 15.0 Ma – the Arctic is opened up (turning land cells to ocean) in order to approximately preserve,
    across all the 7 reconstructions, the global land fraction in the underlying GCM of 0.29.

The resulting cGENIE bathymetry for each time-slice is shown in Fig. 3. The CAS is open for all time slices until the Holocene
in this regridding, with varying bathymetry for each. The Drake Passage is already open at 15 Ma, but widens by 12.5 Ma and
gradually deepens. The Atlantic Ocean widens and deepens throughout, and the Pacific generally deepens. The Australian land
mass moves north and the Indonesian seaway gradually reduces. Africa moves north, reducing the Tethys Sea area, which is
already closed off from the Indian Ocean at 15 Ma. The Tethys disappears by 10 Ma, and the Mediterranean Sea is not included
in the simulations. Greenland is isolated from North America until 7.5 Ma, and the Bering Strait is closed until 7.5 Ma. Only
at the Holocene is both Bering Strait open and Greenland is isolated from N. America once again, allowing mixing between
the Arctic ocean, the North Pacific, and North Atlantic at the same time.

We chose and calculate a zonally average (GCM-derived) planetary albedo profile rather than a 2D re-gridded one,
in order to retain closer back-compatibility with the original GENIE configuration in which an idealised zonal profile is applied
(e.g. Edwards and Marsh, 2005). Different GCMs average and save wind speed differently (or only as velocity vectors), so,
the final re-gridded wind speed product can differ substantially between GCMs and with modern observations. As a result, air-
sea gas exchange is rescaled in the late Holocene configuration in order to give a mean global and annual average modern air-
sea coefficient value for CO_2 of approximately 0.058 mol m^{-2} yr^{-1} µatm^{-1}. This same air-sea gas exchange scaling is then
applied to all older time-slices. As compared to Cao et al. (2009), we also forego the high southern latitude zone of reduced
atmospheric diffusivity, previously used for present-day model configurations (described further in Marsh et al., 2011). Initial
mean salinity was reduced by 1 PSU to 33.9 PSU in all time-slice configurations for simplicity and consistency (although we
recognise that in reality mean ocean salinity should progressively decrease from a modern value of 34.9 back in time with

progressively decreasing global land ice volume). The orbital configuration was kept at its modern settings throughout all time-slices (0.0167 eccentricity, 0.397789 for sine of obliquity, 102.92 for the longitude of perihelion (in degrees)). However, we
195 did vary the following model parameter values and initial conditions as a function of (geological) time (time-slice age):

- The solar constant is assumed to change with time and to follow Gough (1981) (and see: Feulner, 2012), resulting in a small increase between 15.0 and 0 Ma, from 1366.09 Wm^{-2} (a reduction of 0.14% compared to modern) at 15.0 Ma, to 1368.0 Wm^{-2} by the late Holocene.

- The mean Mg/Ca ratio of the ocean is also assumed to change with time, following Tyrrell and Zeebe (2004). The
200 corresponding range is then from 13.15 mmol kg^{-1} Ca^{2+}, 41.21 at mmol kg^{-1} Mg^{2+} at 15.0 Ma, to 10.28, 52.81 mmol kg^{-1} in the present-day ocean. (Changing ocean Mg/Ca with time influences the calculation of carbonate saturation as well as dissociated constants).

All boundary configurations and relevant parameter settings (as well as the cGENIE.muffin model code itself) are available as open-source. Refer to our 'Model code availability' statement regarding obtaining and running any or all of these model
205 configurations.

2.4 Model experimental design

The premise of our experiment design and study is that the observed distribution of ocean temperature proxies and benthic δ^{13}C can be explained by a combination of climate state and the specific pattern of global ocean circulation. For this study, we ignore the strength of the biological pump as an additional and independent control on δ^{13}C (but not temperature). In a 'perfect'
210 paleo model (which does not exist) and under the correct greenhouse gas boundary conditions (which are poorly constrained), climate and ocean circulation would be a correct and emergent property of the model and match the data. In cGENIE, we consider CO_2 as a primary uncertainty in the model – not only in terms of uncertainty in the real past value, but additionally in the radiative forcing required in cGENIE to generate appropriate temperatures. Additionally, because there is no prior expectation that for a 'correct' surface climate the emergent state of ocean circulation is at-all correct in cGENIE, we consider
215 the salinity flux adjustment (hereafter 'FwF') as an 'unknown' and moreover, as a means-to-an-end in adjusting basin (to global) scale patterns and strength of circulation (and specifically Atlantic overturning circulation). The aim in this methodology being to determine a specific circulation pattern and climate (temperature) state that best reproduces the data. Such a general global-scale model state can then provide the boundary conditions for assessing more detailed water-column profiles and modal-data comparison (e.g. in Boscolo-Galazzo and Crichton et al., In Prep.)
220 We hence, for each of the 7 time-slices and model configurations, carry out a 2D parameter space (CO_2 vs. FwF) sweep via an ensemble of model experiments. In each ensemble, we test atmospheric CO_2 values of 280, 400, 560, 800, 1120 and 1600 ppm, and vary the salinity flux correction of the North Atlantic (FwF) between 0.0 Sv and 0.7 Sv, in increments of 0.1 Sv intervals, for a total of 48 members in each ensemble. The CO_2 value are chosen either as simple multiples of 280 ppm (pre-industrial / late Holocene), or commonly assumed GCM values (e.g. the 400 ppm in Farnsworth et al. (2019)) and
225 multiples there-of. We chose the range of FwF values to encompass the equivalent tuned modern cGENIE model value of 0.32

Sv (Edwards and Marsh, 2005). Furthermore, because the land-sea mask progressively changes across the 7 time-slices, we create a common mask for the purpose of salt transfer by identifying all the grid points in the North Pacific and North Atlantic that are ocean ('wet' points) across all configurations. The total FwF value is then divided evenly across all (equal area) grid points in the North Atlantic (positive salinity input) and North Pacific (negative salinity input) such that global ocean salinity

230   is always conserved.

For each ensemble member (48) and each time-slice (7) (a total of 336 model simulations), cGENIE is spun-up for 10,000 years until it reaches steady-state. In the absence of appreciable inter-annual variability and unlike in fully coupled GCM experiments, multi-decadal averaging is not necessary in the cGENIE model and we take the last annual average (year 10,000) of the simulation in order to carry out the model-data comparison.

235   The one caveat and complication to how we conduct the model-data comparison is that in running the parameter ensembles, we identified sustained oscillations in global ocean circulation in many of the simulations that affected both temperature (several degrees) and $\delta^{13}C$ (several tenths of a per mil). These oscillations were of varying period and magnitude and occurred mainly in the mid-range time slices at mid-range CO_2 levels and mid-range flux corrections. Whilst this is extremely interesting from an ocean circulation and dynamics point of view (and will be followed up upon subsequently), it is

240   not the focus of this study and we need to identify a consistent 'representative' climate state in the model in order to make the model-data comparison. To create climate conditions that are representative of each CO_2-flux correction combination we therefore parsed the cGENIE model output, identifying ensemble members characterised by self-sustained oscillations in their output. For these experiments, we further automatically identified the period of the oscillation (typically 2-4 kyr) and averaged the model output over one full period, starting the average from the end of the 10,000 year spin up and working back (towards

245   the start of the model experiment). We thereby created a 'mean' annual average output for the affected ensemble members. Both unmodified run-end and reconstructed mean annual averages were then treated exactly the same in terms of carrying out model-data comparison.

2.5 Model-data comparison

The datasets collated are used as a constraint to determine which combination of CO_2 and North Atlantic salinity flux correction
250   (FwF) produces the best-fit climate to data indicators. This is performed quantitatively by statistically comparing local model output with the data points for all three datasets, and combining these to produce a final "best-fit" model setting for each time-slice. The statistical methods applied are: 1) the difference between the mean of the dataset and the mean of the model output at the data-locations, producing an offset value or mean bias, 2) an overall measure of goodness of fit of the model to the dataset known as "M-score" (Watterson, 1996) that compensates for model-data bias (i.e. bias is not considered in M-score).
255   Where several data-points are located within one model grid square, the mean of the data values is used to compare with the model value. For the benthic data, each data-point is assumed to be on the ocean floor, so the model's bathymetry determines the data depth, and the data value is compared with the model value in the deepest water in that location.

3 Results

3.1 Individual constraints

260   The latitudinal trend for $\delta^{13}C$ data (dotted lines) in Fig. 2 shows that the Atlantic switches trend between 12.5 Ma and 10 Ma. Prior to 12.5 Ma the North Atlantic data indicates a generally more negative $\delta^{13}C$ than in the South Atlantic – a situation that reverses from 10 Ma onwards when more positive $\delta^{13}C$ tend to occur in the North Atlantic. In the Pacific Ocean, the $\delta^{13}C$ gradient from south to north tends to increase towards the present, with the strongest gradient seen at 2.5 Ma. Overall, $\delta^{13}C$ values increase spread from 15 Ma towards the present.

265   Our compiled surface ocean temperatures provide a first order constraint on atmospheric CO_2 (in terms of that required in the cGENIE model to generate appropriate warming). The cGENIE model has a climate sensitivity matching that of the present day for all timeslices, if climate sensitivity was different in the past then this would affect the required CO_2 forcing needed to achieve a certain global temperature (e,g, Bradshaw et al., 2015). The M-score measures how well the model reproduces these surface ocean temperatures. For each time-slice, the model ensemble is plotted as a grid (Fig. 4), with flux

270   correction (FwF) on the x-axis, CO_2 concentration on the y-axis and M-score plotted as contours. The closer the M-score is to 1, the better the model does at reproducing the surface ocean temperature. According to this measure, a CO_2 forcing of 280 ppm at the Holocene, rising to 1600 ppm by 10 Ma, is suggested by tracking the trajectory of maximum M-score values. In older time slices, the maximum M-score reduces significantly, from over 0.8 for the Holocene to less than 0.3 at 15 Ma (Fig. 4), suggesting CO_2 levels higher than 1600 ppm are needed at 12.5 Ma and 15 Ma.

275   The global benthic ocean data estimate from Cramer et al (2011) is plotted over modelled global mean benthic ocean temperature in Fig. 5. In this case, both CO_2 and FwF have a strong effect on benthic temperature (unlike for the surface ocean temperature where FwF has little effect). Bathymetry, albedo and windfields (which differ in our time slices) have a direct effect on benthic ocean temperature, with 15 Ma having a tendency to a warmer deep ocean compared other time slices for the same CO_2 and FwF. A range of combinations of CO_2 and FwF would satisfy the benthic temperature constraint (shown as the

280   white line, the Cramer et al. (2011) benthic global mean temperature estimate). The FwF affects North Atlantic surface salinity and therefore has an impact on thermohaline circulation. Higher FwF values create more saline surface North Atlantic waters of lower buoyancy, transporting more of the surface heat energy to the deep ocean. This results in combinations of high CO_2-low FwF, or low CO_2-high FwF being equally possible, but with a general trend of reducing CO_2 and reducing FwF from 15 Ma towards the Holocene.

285   The benthic temperature model M-score (contours) and bias (showing modelled global mean temperatures within 0.5 degrees of the data, shaded dark grey) are shown in Fig. 6. The mean bias shows a similar pattern to the mean global temperature in Fig. 5, with combinations of CO_2 and flux satisfying the data constraint. The M-score suggests a mid-range CO_2 assumption (560 ppm) at 15 Ma rising to a maximum at 10 Ma of 800 pm, and thereafter falling to 280 ppm by the Holocene. The flux correction that satisfies the benthic temperature constraint is fairly high at all time slices (around 0.4 Sv)

290   until the Holocene, where it drops to 0.2 Sv.

For almost all slices, an elbow-type shape is present in the modelled benthic temperatures. As CO_2 reduces for a given FwF, benthic T is either stable or even increases (Fig 5). At CO_2 levels lower than this elbow, the cooler surface ocean favours the sinking of waters in high latitudes, supporting the thermohaline circulation. From 12.5 Ma this elbow turning point CO_2 value is around 800 ppm and this gradually reduces to about 400 ppm in the Holocene (for settings in which model mean deep

295   ocean temperature agrees with data). At CO_2 higher than this elbow, there is only a very weak (or shallow) AMOC.

The highest M-scores for the older time slices (Fig. 6) do not agree with the surface temperature constraint for CO_2 forcing (Fig. 4), with a CO_2 of 560 ppm at 15 Ma with a high FwF suggested from benthic temperatures compared to a requirement of over 1000 ppm according to surface temperatures. The calculated benthic temperatures are dependent upon the local salinity and on global ice volume. In general, a higher global ice volume (than estimated by Cramer et al., 2011, Eq. 7a therein), means lower benthic temperatures as calculated from $\delta^{18}O$ (Marchitto et al 2014, Eq. 9 therein).

300

3.2 Combining the constraints

All data constraints together are shown in Fig. 7, with the two temperature constraints (plus independent global mean benthic temperature estimate) mapped onto the statistical fit surface (M-score) of model vs. observed $\delta^{13}C$. We mark the best-fit

305   selection with a star and include a range of estimates delineated by error bars on the best fit CO_2 and FwF values. Note that the best-fit is only identified from those specific values that we ran the model ensemble members for (i.e. we do not interpolate between settings). The model fit to benthic data $\delta^{13}C$ are discussed as part of the best-fit selection.

The primary constraint on atmospheric CO_2 in the model in this study is the proxy reconstructed surface ocean temperature. This is much less sensitive to the FwF value than the other, benthic ocean constraints. This is because surface

310   ocean circulation patterns are largely dictated by the wind stress forcing, that we do not vary within any single time-slice. Given a surface ocean circulation pattern that is relatively immune to changes in the applied FwF, the ensemble member CO_2 value and with it the surface climate state control the mean and pole-to-equator gradient of SST and hence model-data surface temperature fit. For all the following analyses of the model-data fit, we hence start with consideration of SSTs and the choice of CO_2 for each time-slice in turn, then bring in additional constraints and consideration of the value of FwF.

315   1.   For the Holocene slice, we have a priori knowledge that CO_2 at this warm interglacial period was around 280ppm (Indermühle et al., 1999) – a value which in cGENIE agrees with the surface ocean temperature data. This gives us some confidence in the cGENIE Earth system model and the methodology, but caveated by the fact that although we use here a new and different GCM-derived (Farnsworth et al., 2019) modern continental configuration, cGENIE has already previously been calibrated against present day observed ocean temperatures at an atmospheric CO_2

320   value of ca. 280 ppm (e.g. Price et al., 2009; Ridgwell et al., 2007), and so an acceptable fit to SSTs for a CO_2 value of 280 ppm is not necessarily unexpected. The global benthic temperature estimate from Cramer et al. (2011) represents a mean climate for both glacial and interglacials, which is around 2 °C cooler than the warmer Holocene. For a global benthic temperature warmer by 2 °C compared to Cramer et al. (2011), Fig. 5 suggests a relatively low

flux correction of 0.1 to 0.2 Sv is required to fit the data (in comparison, the present-day calibrated value in cGENIE
325   is 0.32 Sv). However, the benthic $\delta^{13}C$ data constraints (Fig. 7) tends towards a higher M-score for a higher flux
correction, so we chose a best fit value of FwF 0.2 Sv.

2. For the 2.5 Ma slice, surface and benthic temperatures both suggests a higher (and larger range for a fit to) CO_2 than
at Holocene, so 400 ppm is selected as the best-fit CO_2 value. Similar as in the Holocene, a higher flux correction
(FwF) at 2.5 Ma tends to show a better model-data benthic $\delta^{13}C$ agreement, although this constraint is less strong
330   than the benthic temperature constraint (maximum M-score for benthic temperature is ~0.4, and for $\delta^{13}C$ at 0.5Sv is
~0.2). We hence select 0.3 Sv as the FwF value, on the higher end of the benthic temperature ($\delta^{18}O$) dataset, and in
agreement with global mean benthic temperature from Cramer et al. (2011).

3. At 4.5 Ma, the surface temperature data supports a higher again CO_2 value in cGENIE, but here the $\delta^{13}C$ constraint
is stronger (as compared to that at 2.5 Ma) and instead supports a lower-CO_2-higher-FwF, as the M-score increases
335   in this direction of parameter space (Fig. 7). We hence place the best-fit in this direction – on the higher end for the
benthic temperature statistical fit range, and with CO_2 at 400 ppm and a FwF value of 0.5 Sv.

4. At 7.5 Ma, matching surface-temperature in cGENIE requires a significantly increased CO_2 compared to 4.5Ma,
with 800 ppm as being clearly the best-fit CO_2 value. A similar tendency (to 4.5 Ma) to lower CO_2 but higher FwF
is also apparent in the $\delta^{13}C$ constraint for this time-slice, although at 800pm this is less strong (with a low M-score
340   for 800 ppm and high FwF). The flux correction is set at 0.4 Sv, near the centre of the benthic temperature
maximum M-score (shown as a dashed blue contour).

5. For 10 Ma and earlier time slices the model-data M-score for surface temperature declines, but with higher scores
for higher CO_2 in all three cases. At 10 Ma, the benthic $\delta^{18}O$ temperature dataset suggests a CO_2 of 800 ppm and
FwF of 0.3 to 0.4 Sv. The $\delta^{13}C$ fit surface tends towards higher M-score values for higher flux, so we set FwF at 0.4
345   Sv.

6. The 12.5 Ma slice seems to show a transition state for the trends in $\delta^{13}C$, with overall low M-score for all
combinations of CO_2 and $\delta^{13}C$. As $\delta^{13}C$ provides a weak constraint, the CO_2 and FwF values are selected as 1120
ppm and 0.2 Sv, respectively, as a compromise between the high CO_2 requirement for surface temperature and the
lower-CO_2 higher-FwF for the benthic temperature.

350   7. At 15 Ma the ensemble surface of M-score for $\delta^{13}C$ is inversed, showing a higher score for (generally) higher CO_2
and combined with lower FwF. Although we have fewer surface temperature data-points, they suggest a high CO_2,
somewhat in disagreement with the benthic temperature dataset that favours a mid-range CO_2 together with higher
FwF. As the $\delta^{13}C$ constraint tends towards a lower FwF (compared to time slices younger than 10 Ma), we select a
CO_2 of 1120ppm and a low flux correction of 0.1 Sv.

355 All these selected settings and the ranges are summarised in Fig. 8. In general, in the older time slices, cGENIE underestimates
the north Atlantic temperature and this results in a generally higher M-score for the highest CO_2 settings for 15, 12.5 and 10
Ma slices. The warm North Atlantic in the data compared to the model may be due to insufficiently simulated increased heat

transport from the lower latitudes. However, benthic temperature patterns do not suggest that a particularly high FwF is appropriate at 15 Ma and 12.5 Ma at the higher CO_2 levels. At 10 Ma, the stronger FwF results in increased sinking of the

360   warmer surface waters to the deep N. Atlantic (Fig. 10), although the model still underestimates surface N. Atlantic temperature according to data. This warm N. Atlantic in older time slices (Figure 9b N. Atlantic locations shown as grey triangles) cannot be explained by a stronger AMOC caused by a more saline surface N. Atlantic in the model. In all older time slices, the model high latitude temperatures tend to be too low compared to data – this is an established characteristic of warm climates, where climate models tend to struggle to reproduce the flatter latitudinal temperature gradients seen in data (Goldner et al., 2014).

365   Taking account of the deep ocean salinity in the temperature calculation from $\delta^{18}O$ data makes a significant difference to deep ocean temperature patterns (Figure 10b). When flux corrections are higher (mainly from 10 Ma), the saltier North Atlantic waters result in higher calculated temperatures, and data-points in shallower waters are most strongly affected. Accounting for water salinity increases some data-points temperature by more than 3°C.

            The picture for $\delta^{13}C$ is more complex than for temperature as it represents the combined effects of ocean circulation

370   and of the action of marine biota and the ocean biological carbon pump (in which respect we do not make any attempt to adjust the model towards the data here). Overall, $\delta^{13}C$ data show a larger range than modelled $\delta^{13}C$ of DIC from 12.5 Ma to 2.5 Ma; at 15 Ma and at the Holocene the ranges are comparable (Fig. 11). At higher flux corrections (10 Ma to 4.5 Ma) the deep ocean is relatively lighter than at lower flux corrections (Holocene, 12.5 Ma and 15 Ma) with respect to the atmospheric value (Fig. 11). For 15 Ma to 12.5 Ma this is due to low ventilation of deep ocean waters, and for the Holocene due to the dominance of

375   AMOC delivering north Atlantic surface waters to the deep. For 15 Ma and 12.5 Ma we would therefore expect a higher surface to deep $\delta^{13}C$ gradient (except for the north Atlantic), a pattern generally reversed in later times slices.

            Modelled surface ocean temperature has fallen more than benthic ocean temperature since 15 Ma, with simulated best-fit declines of ~6 °C and ~3 °C respectively (Fig. 12). From at least 10 Ma onwards, a more saline surface North Atlantic (with higher FwF imposed) in conjunction with the specific configurations of the continents allows the sinking of warm salty

380   waters and the delivery of surface heat to the deep. However, the absolute salinity of the surface north Atlantic reduces consistently from 15 Ma to present (despite global mean salinity in the model being initialised at 33.9 PSU across all time-slices) (Fig. 12). Only the 2.5 Ma slice shows a more saline N. Atlantic than the previous time slice. The benthic N. Atlantic salinity shows a pattern similar to the imposed flux correction, but is also clearly dependent on other boundary conditions; CO_2 and flux correction does not change between 10 Ma and 7.5 Ma but benthic N. Atlantic salinity does. Sea ice cover appears

385   from 10 Ma, but with significantly more sea ice present from 4.5 Ma onwards when best fit CO_2 in the model has dropped to 400ppm, with sea ice both increasing in area and thickening when CO_2 drops further to 280ppm at the Holocene (Fig. 12).

            Every ensemble member, regardless of the specific atmospheric CO_2 assumption, was driven with a $\delta^{13}CO_2$ value of -6.5‰. Comparing the offset of simulated benthic $\delta^{13}C$ values to the data can hence give an estimate of how atmospheric $\delta^{13}CO_2$ has changed through time. This is shown in Fig. 13 and represents the atmospheric $\delta^{13}CO_2$ value that is required to best

390   reproduce the mean benthic $\delta^{13}C$ seen in data measurements. This diagnosed history of atmospheric $\delta^{13}CO_2$ can in turn be used

as a means of identifying changes in the global carbon cycle (e.g. Hilting et al., 2008) or as initial condition values for future model (and model-data) based studies.

4 Discussion

4.1 Climate forcing and CO₂ levels

395   Climate modelling efforts of Miocene time have generally found that warming is insufficient compared to proxy data for an atmospheric CO_2 concentration around 400pm (Goldner et al., 2014; Bradshaw et al 2012; Krapp and Jungclaus, 2011). Greenop et al. (2014) find that CO_2 concentrations of 300ppm to 500ppm are inconsistent with their findings of Antarctic ice sheet stability during the Middle Miocene Climatic Optimum (MMCO, at 17-14 Ma), the warmest part of the middle Miocene. Estimates of Miocene CO_2 in more recent studies have trended higher (MMCO at 470-630 ppm, Sosdian et al., 2018) compared

400   to the earlier estimates (e.g., ~220ppm at 15 Ma, Pagani et al., 1999). Estimates of surface ocean temperature have also tended to trend higher in more recent studies (Pearson, 2012). In both cases (CO_2 and temperature estimates), this trend towards a warmer and higher CO_2 was due to improved scientific methods and understanding (and the identification of new proxies); whilst terrestrial data has fairly consistently suggested this warm Miocene (e.g. Wolfe, 1985). However, considering only the most recent highest CO_2 estimates there is still a discrepancy between the surface ocean temperature constraint (that requires

405   a high CO_2 in this study) and estimated CO_2 levels from Sosdian et al. (2018). As an example, forcing cGENIE with CO_2 levels identified by Sosdian et al. (2018) for 10 Ma (~400ppm) results in a surface ocean that is around 4 °C too cold compared to the surface ocean temperature data. At older time-slices the discrepancy (when assuming 400 ppm throughout) increases further.

At 12.5 Ma and 15 Ma, we find our largest differences in model to data surface temperature in the N. Atlantic (Fig

410   9), with modelled N. Atlantic temperature being lower than sea surface temperature (SST) indicators. Part of this modelled N.Atlantic surface temperature underestimate may be due to a latitudinal temperature gradient being too high in the model (as also identified in previous modelling efforts of past warm climates), with heat transport to the poles being too low. Increased ocean heat transport was found to reduce the latitudinal temperature gradient and the location of Hadley and Ferrel cells in a modelling study by Knietzsch et al (2015), indicating the importance of atmospheric circulation, which in the cGENIE 2D

415   energy-moisture balance atmosphere is simplified and invariant. The interaction between vegetation and atmosphere was found to produce a warmer Miocene independent of CO_2 by Knorr et al. (2011). The interaction between vegetation and paleogeography produced a higher climate sensitivity in the modelled Miocene in Bradshaw et al. (2015) and warming independent of CO_2 by Henrot et al. (2010) due to vegetation and lower-elevation topography. Land surface cover and increased northern transient eddy heat transport were found to result in polar warming in a modelled late Miocene in a fully

420   coupled GCM by Micheels et al. (2010). Chemistry-climate feedbacks linked to vegetation differences compared to the present were found to be as strong as or more important than CO_2 forcing in the Pliocene by Unger and Yue. (2013). The significant changes occurring in global vegetation distribution since the middle Miocene (Pound et al., 2012) then may be critical to fully

reproducing observed SST patterns. The absence of any representation of vegetation and associated feedbacks in the version of cGENIE we employ here may then account for some of the difference in the CO_2 we apply (e.g. 1120ppm at 15 Ma) and

425 that estimated by Sosdian et al. (2018) (e.g. 470-630 ppm at 17-14 Ma).

A portion of the aforementioned CO_2 difference may alternatively (or in tandem) be explained as the contribution of other green gases to climate forcing during the Miocene; in cGENIE only CO_2 is imposed as a greenhouse gas, so the imposed atmospheric CO_2 concentration should be understood as an 'equivalent' climate forcing. If the CO_2 concentration at 15 Ma is as per estimated by Sosdian et al (2018) – 630ppm (their higher estimate), that leaves an equivalent CO_2 forcing of ~500ppm

430 to be explained (as compared to the 1120ppm required in this study for the best-fit model setting at 15 Ma). The 500ppm forcing equates roughly to 1000 GtC as CO_2. If this equivalent forcing was attributable to methane (at 25 times the warming potential of CO_2 at the 100 year timescale), methane concentrations would need to have been 10 times higher (than present) in the mid to late Miocene if they are to explain all the CO_2-forcing difference between this study and Sosdian et al. (2018). Wetlands are a significant source of atmospheric methane in the present day, as well as representing large terrestrial carbon

435 stores. Extensive wetlands existed during the middle Miocene (Eronen and Rossner, 2007; Hoorn et al., 2010; Morley and Morley, 2013), with the generally warmer wetter climate and lower elevation topography being favourable to their creation and persistence. These conditions would be conducive to methane production (Dean et al., 2018).

A large Antarctic ice sheet has been identified for the period soon after the Miocene climatic optimum (Hautvogel et al., 2012; Badger et al., 2013; Pierce et al., 2017). A drop in sea level associated with ice sheet growth would reduce overlying

440 pressure in relatively shallow seas and may have destabilised gas hydrates globally. Gas hydrates were found to have likely destabilised during the sea-level lowering of the Miocene in the present-day Appenines in Italy (Argentino et al., 2019) and in the Black Sea Basin (Kitchka et al., 2016). A release of gas hydrates, if it reached the atmosphere rather than oxidising to CO_2 in overlying waters, would cause a drop in $\delta^{13}CO_2$ that would also be reflected in $\delta^{13}C$ (ocean) data. However, gas hydrate release (as methane) reaching the atmosphere would normally contribute to warming, which seems incompatible with evidence

445 of large ice sheet growth and falling CO_2 soon after the mid-Miocene. Further than this, methane's short lifetime in the atmosphere would suggest that this would need to be a sustained gas-hydrate release (to maintain high atmospheric methane levels) with steady duration of the order of 1 Myr to explain any long-term difference between Sosdian et al (2018) CO_2 estimate and the SST requirement for a warmer climate. It would also require mechanisms promoting ice growth in another way (not greenhouse gas driven cooling), a strong CO_2 drawdown after 15Ma together with the maintenance of a high

450 atmospheric methane levels. Atmospheric chemistry and baseline methane levels does affect the lifetime of atmospheric methane (Schmidt and Shindell, 2003) but it is more likely that destabilisation of gas hydrates would act on shorter timescales (rather than sustained levels) perhaps influencing short term transient climate changes.

Overall, the combination of a flatter latitudinal temperature gradient, the contribution of other (than CO_2) greenhouse gases to climate forcing, and the effect of vegetation and its distribution (which may be seen as part of climate sensitivity),

455 may explain the discrepancy between our modelled CO_2 in this study and the most recent data estimates (Sosdian et al., 2018). However, study of Miocene CO_2 levels is still ongoing and the question remains to be definitively resolved.

4.2 Benthic Temperatures

The pattern of temperature in the deep ocean arises as a combination of the pattern of ocean surface temperature (and surface climate in general) and the large-scale circulation of the ocean. In this study, we have shown that net salinity transport into the

460   N. Atlantic region can induce and strengthen a meridional overturning circulation that in turn drives a warming of the deep ocean, independent of atmospheric CO_2 forcing. This helps explain a seeming decoupling of surface and deep water total cooling seen in the data since the middle Miocene, where surface temperatures have cooled more than deep ocean temperature according to data. Further, the benthic temperatures that we derive from benthic $\delta^{18}O$ data are dependent upon the global ice volume that we assume (Marchitto et al., 2014), these temperature estimates would be different for differing global ice

465   volumes. For example, at 15 Ma, a lower global ice volume assumption than the one we apply would bring $\delta^{18}O$-based benthic temperatures more in line with the Mg/Ca based temperature estimate from Cramer et al. (2011) (Fig.7). Finally, local salinity has a strong control on $\delta^{18}O$; an increase in measured benthic $\delta^{18}O$ in the N. Atlantic during the late Miocene that may be interpreted as evidence of a strong cooling, may actually be attributable to the increased salinity of the deep sea water, where salinity (rather than temperature) dominates the $\delta^{18}O$ signal recorded in the benthic foraminifera. With the onset of Atlantic

470   overturning circulation during the Miocene, the salinity of deep N. Atlantic waters has a strong control on $\delta^{18}O$, and when included in the temperature calculation results in increases of up to 3°C in some locations (compared to the temperature uncorrected for salinity) (Fig. 10).

4.3 Using carbon-13 to trace ocean circulation patterns

Changes in deep ocean circulation patterns and specifically the aging of water masses and progressive accumulation of

475   isotopically light respired carbon, in theory should be identifiable in global patterns of $\delta^{13}C$ data from benthic foraminifera. However, the processes that control $\delta^{13}C$ are complex, involving both circulation and ocean carbon pumps, which furthermore are not independent – large-scale changes in circulation that affect nutrient return to the surface will hence also modulate the strength of the biological pump in the ocean. Changing ocean interior temperature patterns may further influence where carbon is respired via a temperature control on the rate of carbon respiration (John et al., 2014), further affecting $\delta^{13}C$. For instance,

480   we simulate a flatter (smaller range of) deep water $\delta^{13}C$ than data suggests for time-slices between 10 Ma and the 2.5 Ma, but a comparable range (between model and data) for 15 Ma and the late Holocene (Fig. 11). The similar ranges at these two extremes – one the warmest and the other the coolest – suggest that the issue in between is not primarily with the model representation of the carbon pumps, but with the Central American Seaway (CAS) that may be too deep in the intermediate time-slices. A deep CAS allows the flow of less recently ventilated Pacific waters into the Atlantic basin in the model and

485   reduces the $\delta^{13}C$ gradient between these two ocean basins. The time slice with the closed CAS (Holocene) shows a model range and values of $\delta^{13}C$ in general agreement with the data for the Pacific and the Atlantic basins (Fig. 11), and a closed or restricted CAS in the 10 Ma to 2.5 Ma slices may improve the models fit to data. This tends to support the hypothesis of an early restriction/closure of CAS, as the increasing N. Atlantic to Pacific $\delta^{13}C$ gradient is evident from 10 Ma (Fig. 2). However,

it should be noted that the distribution of $\delta^{13}C$ is also dependent on other factors (such as local water velocity fields and vertical
490   stream functions, as well as effects of the biological pump that affect $\delta^{13}C$).

4.4 North Atlantic salinity and CAS

In our model set-up the CAS allows some mixing of Atlantic and Pacific waters at every time-slice except the Holocene. The
flux correction in the N. Atlantic that we apply may be seen as a combination of compensating for a more restricted (or closed)
CAS, and also for the precipitation gradient across N. America which would be affected by precipitation patterns and
495   topography. With a warmer climate and higher CO_2, the atmosphere is able to hold more water vapour and the result is a
globally wetter climate. However, the North American plains saw the rise of grasslands that favour a drier climate (Janis et al.,
2002) through the Miocene, linked to the uplift of Mountain ranges in the west that created a rain shadow on the central plains.
Over the last 12 million years in the Sierra Nevada a rain shadow similar to present was identified by Mulch et al. (2008),
indicating mountain uplift in this area occurred prior to 12 Ma. Further north in the Cascade Mountains high uplift and erosion
500   rates were dated to late Miocene (12 to 6 Ma), and the Coast Mountains and British Colombia uplift from 10 Ma (Reiners et
al. 2002). We would expect therefore the proportion of a flux correction for rainfall to be higher from around 12.5 Ma to 6 Ma
to take account of this mountain building.

A wash-house (warm and wet) climate was identified in Europe in certain periods of the Miocene, from 10.2 Ma to
9.8 Ma and from 9.0 Ma to 8.5 Ma, which appeared to be linked to deep N. Atlantic Ocean temperatures (Boehme et al., 2008).
505   This was in-turn linked to a possible temporary restriction of the CAS and greater northward heat transport. A middle Miocene,
even if temporary (Jaramillo et al., 2017), closure of the CAS was identified by Montes et al. (2015). We identify a transition
between 12.5 Ma and 10 Ma, and a high flux correction (FwF = 0.4 Sv) for the 10 Ma time-slice which may be representative
of a CAS restriction.

Our modelled global overturning circulation shows a gradual increase in dominance of northern hemisphere deep
510   water, and the development of the AMOC (Fig. 14). Generally in cGENIE, the higher the CO_2 forcing the weaker the AMOC,
with a strong and deep AMOC only present in the Holocene time-slice. The CAS remains open for all the time-slices except
Holocene in this set-up and this may be preventing the establishment of a stronger AMOC in the model at (for example) 2.5
Ma, when the CAS is thought to have already been fully closed. A shallow (but not closed) CAS still allowed the formation of
deep water in the North Atlantic in a GCM model study by Nisancioglu et al. (2003), and our CAS at ~1000m does not fully
515   impede the establishment of a shallow AMOC (if the surface north Atlantic is sufficiently saline). An early Pliocene (4.7 to
4.2 Ma) shoaling of CAS was found to have no profound impact on climate evolution, as North Atlantic deep water formation
was found to be already vigorous by 4.7 Ma by Bell et al. (2015). Vigorous North Atlantic deep water formation appears to
have started by 4.5 Ma in our study (Fig. 14), when CO_2 drops to 400ppm and we increase N. Atlantic salinity (flux correction
FwF at 0.5Sv). A definitive closure of the CAS dates to the late Pliocene (~3Ma, O'Dea et al., 2017). This later final Pliocene
520   closure of CAS had thought to be linked to setting up conditions for northern hemisphere glaciation, although Lunt et al. (2008)

found that an open or closed CAS made little difference to ice sheet size in their model study. Sea ice growth intensifies in our model when CO_2 drops further to 280ppm in the Holocene.

Earlier studies placing Miocene CO_2 levels fairly low, around pre-industrial levels or lower (Pagani et al., 1999; Pearson and Palmer 2000), led research on finding other explanations for the global growth of ice sheets and the onset of
525   glacial cycles since the middle Miocene. With newer estimates for CO_2, a more complete picture emerges, with the combined effects of CO_2 drawdown and ocean seaways influencing global climate. We find that an early CAS restriction probably fits $\delta^{13}C$ data better, and that the draw-down of CO_2 and cooling of polar regions has a more significant role in global circulation patterns and climate (than a late CAS restriction) from the Pliocene onwards.

5 Conclusions and summary

530   In this study, we used proxy data estimates for both surface and benthic temperature ($\delta^{18}O$) as well as benthic $\delta^{13}C$ to constrain the evolution of atmospheric CO_2 and large scale ocean circulation in the cGENIE.muffin Earth system model. We created plausible climatic states in cGENIE for each of 7 time-slices spanning the mid Miocene to Holocene cooling. Constrained by changes in the absolute magnitude and pattern of benthic $\delta^{13}C$, we also diagnosed a plausible history of atmospheric $\delta^{13}C$ over this time interval for use as boundary conditions in future modelling studies, or as a data target for assimilation in geochemical
535   box-models.

In the cGENIE model, we diagnose a progressive reduction in atmospheric greenhouse gas forcing since the mid Miocene, driving global cooling. Simultaneously, we diagnose a gradual strengthening of overturning circulation in the Atlantic that transports heat to the deep Atlantic ocean over the cooling period and leading to a stronger cooling in surface (at ~6°C) than in deep waters (at ~3°C) occurring since the middle Miocene. This onset and strength of the AMOC in cGENIE is
540   controlled by the combined effects of progressive restriction of the Central American Seaway together with our two free parameters – a salinity adjustment (representing mountain building in N. America and an increasing Atlantic-Pacific salinity gradient), and a declining atmospheric CO_2. Declining CO_2 drives cooling which helps to promote the sinking of salty waters in the N. Atlantic, despite surface salinity actually also declining over the cooling period in the model. The net result in the model is a strong and deep AMOC in the Holocene when the CAS is fully closed and atmospheric CO_2 is low.

545   **Appendix A. Correcting benthic Temperature calculated from $\delta^{18}O$ for local salinity**

The Cramer et al. (2011) $\delta^{18}O_{sw}$ estimate for sea water is a global mean value. Applying the Marchitto et al. (2014) paleotemperature calculation using one global mean $\delta^{18}O_{sw}$ therefore results in an uncertainty in temperature depending on location (Fig. A1) due to local differences in $\delta^{18}Osw$ values (Fig. A2).

The seawater $\delta^{18}O_{sw}$ is determined by global ice volume (which we get from Cramer et al., 2011), local temperature
550   and local salinity (Rohling, 2013). In the present day with an active AMOC, the North Atlantic benthos have a more positive

$\delta^{18}O_{sw}$ than, for example, the North Pacific. This is due to both temperature and salinity, with the salty North Atlantic waters transported to the deep by the AMOC. In our model ensembles the benthic salinity is affected by both CO_2 and flux correction as these affect ocean circulation. Therefore, for each simulation we apply a $\delta^{18}O_{sw}$ driven correction to the paleo temperatures due to their local salinity. To do this, we use present day deep-water (2500m) $\delta^{18}O_{sw}$ (LeGrande and Schmidt, 2006) and

555   salinity (WOA 2013, Zweng et al., 2013) to create a general linear model (Eq. A1), where S is salinity.

$$\delta^{18}O_{sw} = 0.8S - 27.7 \tag{A1}$$

The North Atlantic is the region which is most affected by changes in salinity, and therefore the greatest temperature offset will be in this location. We adjusted the salinity model to get best-fit values for the north Atlantic, and a good fit for the difference in $\delta^{18}O_{sw}$ between the north Atlantic and the Pacific. All ocean locations with $\delta^{18}O_{sw}$ between -0.3 and 0.3‰ are

560   shown in a crossplot of data and salinity-model derived $\delta^{18}O_{sw}$ (Fig. A3). The grouping of points clearly offset from the 1:1 line are the high southern latitudes (also clearly visible in Fig. A2). The accuracy of the salinity-derived model in finding $\delta^{18}O_{sw}$ is ±0.03‰ at one standard deviation, and ±0.06‰ at the 95% confidence level excluding latitudes higher than 70°.

The paleotemperature equation we apply is the linear model from Marchitto et al. (2014, Eq. 8 therein), using the global ice volume estimate from Cramer et al. (2011). The $\delta^{18}O_{sw}$ model from salinity is also linear so we apply a simple linear

565   correction to the calculated temperature (that shown in Supplementary Table S2). To obtain the $\delta^{18}O_{sw}$ offset, we first find the modelled global mean salinity (not including latitudes higher than 70°), and subtract this from the modelled benthic salinity, giving a ΔS field (an offset from the mean). We apply Eq. A2 to find ΔT, where 0.8 is the gradient of the linear $\delta^{18}O_{sw}$ model (Eq. A1) and 0.224 is the gradient of the linear paleotemperature model (Marchitto et al., 2014), and use this to correct T for salinity. The uncertainty for the temperature correction is ±0.13°C at one standard deviation and 0.23°C at the 95% confidence

570   level (based on the $\delta^{18}O$ model). The range of temperature correction corresponding to the benthic $\delta^{18}O_{sw}$ spread (of ~0.4‰) in the present day is 1.8°C.

$$\Delta T = \frac{0.8 \, \Delta S}{0.224} \tag{A2}$$

Model code availability

575   The specific version used of the cGENIE.muffin model used in this paper is tagged as release v0.9.8, and has been assigned a DOI: 10.5281/zenodo.3575063

The code is hosted on GitHub and can be obtained by cloning:

https://github.com/derpycode/cgenie.muffin

and then changing the directory to cgenie.muffin and checking out the specific release:

580     $ git checkout v0.9.8

Configuration files for the specific experiments presented in the paper can be found in the subdirectory:

genie-userconfigs\MS\crichtonetal.CP.2019

Details of the experiments, plus the command line needed to run each one, are given in the readme.txt file in that directory. All other configuration files and boundary conditions are provided as part of the release.

585 A manual, detailing code installation, basic model configuration, plus an extensive series of tutorials covering various aspects of muffin capability, experimental design, and results output and processing, is provided on GitHub. The latex source of the manual, along with pre-built PDF file can be obtained, by cloning:

https://github.com/derpycode/muffindoc

The muffin manual contains instructions for obtaining, installing, and testing the code, plus how to run experiments.
590 Specifically:

Section 1.1 – Provides a basic over-view of the software environment required for installing and running muffin. Section 1.2.2 – provides a basic over-view of cloning and testing the code.

Section 15.4 – Provides a detailed guide to cloning the code and configuring an Ubuntu (18.04) software environment including netCDF library installation, plus running a basic test. Section 15.6 – provides a detailed guide to cloning the code and
595 configuring a macOS soft- ware environment including netCDF library installation, plus running a basic test.

Section 1.3 – Provides a basic guide to running experiments (also see 1.6 and 1.7).

Section 1.4 – provides a basic introduction to model output (much more detail is given in Chapter 12).

Author contribution

KAC collated the temperature, δ^{18}O and δ^{13}C data, devised the salinity correction for temperature and carried out the model-
600 data comparison; AR produced the cGENIE time-slice configurations, ran the ensembles and extracted mean-climate data; DJL and AF provided the HadCML configurations; all authors wrote the manuscript.

Competing Interests

All authors declare no competing interests.

Acknowledgments

605 KAC was supported by Natural Environment Research Council (NERC) grant number NE/N001621/1 to PNP. AR acknowledges additional support from the National Science Foundation under Grants 1702913 and 1736771. DJL and AF thank NERC grant NE/K014757/1.

References

Argentino, C., Conti, S., Fioroni, C. and Fontana, D.: Evidences for Paleo-Gas Hydrate Occurrence: What We Can Infer for
the Miocene of the Northern Apennines (Italy). Geosciences, 9(3), 134. doi: 10.3390/geosciences9030134, 2019.

Badger, M.P., Lear, C.H., Pancost, R.D., Foster, G.L., Bailey, T.R., Leng, M.J. and Abels, H.A.: CO_2 drawdown following the
middle Miocene expansion of the Antarctic Ice Sheet. Paleoceanography, 28(1), 42-53, doi: 10.1002/palo.20015, 2013.

Bell et al 2015Bell, D.B., Jung, S.J., Kroon, D., Hodell, D.A., Lourens, L.J. and Raymo, M.E.: Atlantic deep-water response
to the early Pliocene shoaling of the Central American Seaway. Sci. Rep-UK, 5, 12252, doi: 10.1038/srep12252, 2015.

Böhme, M., Ilg, A. and Winklhofer, M.: Late Miocene "washhouse" climate in Europe. Earth Planet Sc. Lett., 275(3-4), 393-
401, doi: 10.1016/j.epsl.2008.09.011, 2008.

Boscolo-Galazzo, F., Crichton, K.A., Ridgwell, A., Mawbey, E.N., Wade, B.S., Pearson, P.N.: Temperature controls carbon
cycling and drives biological evolution in the ocean twilight zone, In Prep.

Bradshaw, C., Lunt, D., Flecker, R., Salzmann, U., Pound, M., Haywood, A. and Eronen, J.: The relative roles of CO_2 and
palaeogeography in determining late Miocene climate: results from a terrestrial model-data comparison. Clim. Past, 8(2),
715-786, doi: 10.5194/cp-8-1257-2012, 2012.

Bradshaw, C.D., Lunt, D.J., Flecker, R. and Davies-Barnard, T.: Disentangling the roles of late Miocene palaeogeography and
vegetation–Implications for climate sensitivity. Palaeogeogr. Palaeocl., 417, 17-34, doi: 10.1016/j.palaeo.2014.10.003,
2015.

Butzin, M., Lohmann, G. and Bickert, T.: Miocene ocean circulation inferred from marine carbon cycle modeling combined
with benthic isotope records. Paleoceanography, 26(1), doi: 10.1029/2009PA001901, 2011.

Cao, L., Eby, M., Ridgwell, A., Caldeira, K., Archer, D., Ishida, A., Joos, F., Matsumoto, K., Mikolajewicz, U., Mouchet, A.,
Orr, J.C., Plattner, G.K., Schlitzer, R., Tokos, K., Totterdell, I., Tschumi, T., Yamanaka, Y., Yool, A.: The role of ocean
transport in the uptake of anthropogenic CO2, Biogeosciences, 6, 375–390, doi: 10.5194/bg-6-375-2009, 2009.

Cerling, T.E., Harris, J.M., MacFadden, B.J., Leakey, M.G., Quade, J., Eisenmann, V. and Ehleringer, J.R.: Global vegetation
change through the Miocene/Pliocene boundary. Nature, 389(6647), 153, doi: 10.1038/38229, 1997.

Cox, P.M., Huntingford, C. and Williamson, M.S.: Emergent constraint on equilibrium climate sensitivity from global
temperature variability. Nature, 553(7688), 319, doi: 10.1038/nature25450, 2018.

Cramer, B.S., Miller, K.G., Barrett, P.J. and Wright, J.D.: Late Cretaceous–Neogene trends in deep ocean temperature and
continental ice volume: Reconciling records of benthic foraminiferal geochemistry (δ18O and Mg/Ca) with sea level
history. J. Geophys. Res-Oceans, 116(C12), doi: 10.1029/2011JC007255, 2011.

Dean, J.F., Middelburg, J.J., Röckmann, T., Aerts, R., Blauw, L.G., Egger, M., Jetten, M.S., de Jong, A.E., Meisel, O.H.,
Rasigraf, O. and Slomp, C.P.: Methane feedbacks to the global climate system in a warmer world. Rev. Geophys., 56(1),
207-250, doi: 10.1002/2017RG000559, 2018.

640   Edwards, N.R. and Marsh, R.: Uncertainties due to transport-parameter sensitivity in an efficient 3-D ocean-climate model, Clim. Dynam., 24(4), 415-433, doi: 10.1007/s00382-004-0508-8, 2005.

Eronen, J.T. and Rössner, G.E.: Wetland paradise lost: Miocene community dynamics in large herbivorous mammals from the German Molasse Basin. Evol. Ecol. Res., 9(3), 471-494, 2007.

Farnsworth, A., Lunt, D.J., Robinson, S.A., Valdes, P.J., Roberts, W.H., Clift, P.D., Markwick, P., Su, T., Wrobel, N., Bragg,
645   F. and Kelland, S.J.: Past East Asian monsoon evolution controlled by paleogeography, not CO2. Sci. Advances, 5(10), p.eaax1697, doi: 10.1126/sciadv.aax1697, 2019.

Feulner, G.: The faint young Sun problem. Rev. Geophys., 50(2), doi: 10.1029/2011RG000375, 2012.

Filippelli, G.M.: Intensification of the Asian monsoon and a chemical weathering event in the late Miocene–early Pliocene: implications for late Neogene climate change. Geology, 25(1), 27-30, doi: 10.1130/0091-
650   7613(1997)025<0027:IOTAMA>2.3.CO;2, 1997.

Goldner, A., Herold, N. and Huber, M.: The challenge of simulating the warmth of the mid-Miocene climatic optimum in CESM1. Climate of the Past, 10(2), 523-536, doi: 10.5194/cp-10-523-2014, 2014.

Gough, D.O.: Solar interior structure and luminosity variations. In Domingo V. (eds) Physics of Solar Variations (21-34). Springer, Dordrecht, doi: 10.1007/978-94-010-9633-1_4, 1981.

655   Greenop, R., Foster, G.L., Wilson, P.A. and Lear, C.H.: Middle Miocene climate instability associated with high-amplitude CO_2 variability. Paleoceanography, 29(9), 845-853, doi : 10.1002/2014PA002653, 2014.

Hamon, N., Sepulchre, P., Lefebvre, V. and Ramstein, G.: The role of eastern Tethys seaway closure in the Middle Miocene Climatic Transition (ca. 14 Ma). Clim. Past, 9(6), 2687-2702, doi : 10.5194/cp-9-2687-2013, 2013.

Hauptvogel, D.W. and Passchier, S.: Early–Middle Miocene (17–14 Ma) Antarctic ice dynamics reconstructed from the heavy
660   mineral provenance in the AND-2A drill core, Ross Sea, Antarctica. Global Planet. Change, 82, 38-50, doi : 10.1016/j.gloplacha.2011.11.003, 2012.

Henrot, A.J., François, L., Favre, E., Butzin, M., Ouberdous, M. and Munhoven, G.: Effects of CO_2, continental distribution, topography and vegetation changes on the climate at the Middle Miocene: a model study. Clim. Past, 6(5), 675-694, doi : 10.5194/cp-5-183-2009, 2010.

665   Herbert, T.D., Lawrence, K.T., Tzanova, A., Peterson, L.C., Caballero-Gill, R. and Kelly, C.S.: Late Miocene global cooling and the rise of modern ecosystems. Nat. Geosci., 9(11), 843, doi : 10.1038/ngeo2813, 2016.

Hilting, A.K., Kump, L.R. and Bralower, T.J.: Variations in the oceanic vertical carbon isotope gradient and their implications for the Paleocene-Eocene biological pump. Paleoceanography, 23(3), doi: 10.1029/2007PA001458, 2008.

Holbourn, A., Kuhnt, W., Schulz, M. and Erlenkeuser, H.: Impacts of orbital forcing and atmospheric carbon dioxide on
670   Miocene ice-sheet expansion, Nature, 438, 483-487, doi: 10.1038/nature04123, 2005.

Hoorn, C., Wesselingh, F.P., Hovikoski, J. and Guerrero, J.: The development of the Amazonian mega-wetland (Miocene; Brazil, Colombia, Peru, Bolivia). Amazonia, landscape and species evolution: a look into the past, 123, 142, doi: 10.1002/9781444306408.ch8, 2010.

Indermühle, A., Stocker, T.F., Joos, F., Fischer, H., Smith, H.J., Wahlen, M., Deck, B., Mastroianni, D., Tschumi, J., Blunier,
675   T. and Meyer, R.: Holocene carbon-cycle dynamics based on CO$_2$ trapped in ice at Taylor Dome, Antarctica. Nature,
        398(6723), 121, doi : 10.1038/18158, 1999.

Janis, C.M., Damuth, J. and Theodor, J.M.: The origins and evolution of the North American grassland biome: the story from
        the hoofed mammals. Palaeogeogr. Palaeocl., 177(1-2), 183-198, doi : 10.1016/S0031-0182(01)00359-5, 2002.

Jaramillo, C., Montes, C., Cardona, A., Silvestro, D., Antonelli, A. and Bacon, C.D.: Comment (1) on "Formation of the
680   Isthmus of Panama" by O'Dea et al. Science Advances, 3(6), p.e1602321, doi : 10.1126/sciadv.1602321, 2017.

John, E.H., Wilson, J.D., Pearson, P.N. and Ridgwell, A.: Temperature-dependent remineralization and carbon cycling in the
        warm Eocene oceans, Palaeogeogr. Palaeocl., 413, 158-166, doi: 10.1016/j.palaeo.2014.05.019, 2014.

Kirtland Turner, S. and Ridgwell, A.: Development of a novel empirical framework for interpreting geological carbon isotope
        excursions, with implications for the rate of carbon injection across the PETM, Earth Planet. Sc. Lett., 435, 1-13., doi:
685   10.1016/j.epsl.2015.11.027, 2016.

Kitchka, A.A., Tyshchenko, A.P. and Lysenko, V.I.: May. Mid-late Miocene sea level falls, gas hydrates decay, submarine
        sliding, and tsunamites in the Black Sea Basin. In 78th EAGE Conference and Exhibition 2016, doi : 10.3997/2214-
        4609.201600609, 2016.

Knietzsch, M.A., Schröder, A., Lucarini, V. and Lunkeit, F.: The impact of oceanic heat transport on the atmospheric
690   circulation. Earth Syst. Dynam., 6(2), 591-615. doi: 10.5194/esd-6-591-2015, 2015.

Knorr, G., Butzin, M., Micheels, A. and Lohmann, G.: A warm Miocene climate at low atmospheric CO$_2$ levels. Geophysical
        Res. Lett., 38(20), doi: 10.1029/2011GL048873, 2011.

Krapp, M. and Jungclaus, J.H.: The Middle Miocene climate as modelled in an atmosphere-ocean-biosphere model. Clim.
        Past, 7(4), 1169-1188, doi: 10.5194/cp-7-1169-2011, 2011.

695 Lagabrielle, Y., Goddéris, Y., Donnadieu, Y., Malavieille, J. and Suarez, M.: The tectonic history of Drake Passage and its
        possible impacts on global climate. Earth Planet. Sc. Lett., 279(3-4), 197-211, doi: 10.1016/j.epsl.2008.12.037, 2009.

Lear, C.H., Coxall, H.K., Foster, G.L., Lunt, D.J., Mawbey, E.M., Rosenthal, Y., Sosdian, S.M., Thomas, E. and Wilson, P.A.:
        Neogene ice volume and ocean temperatures: Insights from infaunal foraminiferal Mg/Ca paleothermometry.
        Paleoceanography, 30(11), 437-1454, doi: 10.1002/2015PA002833, 2015.

700 LeGrande A.N., Schmidt G.A.: Global gridded data set of the oxygen isotopic composition in seawater. Geophysical Res.
        Lett., 33(12), doi: 10.1029/2006GL026011, 2006

Lunt, D.J., Valdes, P.J., Haywood, A. and Rutt, I.C.: Closure of the Panama Seaway during the Pliocene: implications for
        climate and Northern Hemisphere glaciation. Clim. Dynam., 30(1), pp.1-18, doi: 10.1007/s00382-007-0265-6, 2008.

Lynch-Stieglitz, J.: Tracers of past ocean circulation, Treatise on geochemistry, 6, 433-451, doi: 10.1016/B0-08-043751-
705   6/06117-X, 2003.

Marchitto, T.M., Curry, W.B., Lynch-Stieglitz, J., Bryan, S.P., Cobb, K.M. and Lund, D.C.: Improved oxygen isotope temperature calibrations for cosmopolitan benthic foraminifera. Geochim. Cosmochim. Ac., 130, 1-11, doi: 10.1016/j.gca.2013.12.034, 2014.

Marsh, R., Müller, S.A., Yool, A. and Edwards, N.R.: Incorporation of the C-GOLDSTEIN efficient climate model into the
710   GENIE framework:"eb_go_gs" configurations of GENIE. Geosci. Model Dev., 4, 957-992, doi: 10.5194/gmd-4-957-2011, 2011.

Micheels, A., Bruch, A.A. and Mosbrugger, V., 2009. Miocene climate modelling sensitivity experiments for different CO2 concentrations. Palaeontol. Electron., 12(2), 1-19, 2009.

Micheels, A., Bruch, A.A., Eronen, J., Fortelius, M., Harzhauser, M., Utescher, T. and Mosbrugger, V.: Analysis of heat
715   transport mechanisms from a Late Miocene model experiment with a fully-coupled atmosphere–ocean general circulation model. Palaeogeogr. Palaeocl., 304(3-4), 337-350, doi: 10.1016/j.palaeo.2010.09.021, 2011.

Montes, C., Cardona, A., Jaramillo, C., Pardo, A., Silva, J.C., Valencia, V., Ayala, C., Pérez-Angel, L.C., Rodriguez-Parra, L.A., Ramirez, V. and Niño, H.: Middle Miocene closure of the Central American seaway. Science, 348(6231), 226-229, doi: 10.1126/science.aaa2815, 2015.

720 Morley, R.J. and Morley, H.P.: Mid Cenozoic freshwater wetlands of the Sunda region. J. Limnol., e2-e2, doi: 10.4081/jlimnol.2013.s2.e2, 2013.

Mulch, A., Sarna-Wojcicki, A.M., Perkins, M.E. and Chamberlain, C.P.: A Miocene to Pleistocene climate and elevation record of the Sierra Nevada (California). P. Natl. Acad. Sci. USA, 105(19), 6819-6824, doi: 10.1073/pnas.0708811105, 2008.

725 Nisancioglu, K.H., Raymo, M.E. and Stone, P.H.: Reorganization of Miocene deep water circulation in response to the shoaling of the Central American Seaway. Paleoceanography, 18(1), 1006, doi: 10.1029/2002PA000767, 2003.

O'Dea, A., Lessios, H.A., Coates, A.G., Eytan, R.I., Restrepo-Moreno, S.A., Cione, A.L., Collins, L.S., De Queiroz, A., Farris, D.W., Norris, R.D. and Stallard, R.F.: Formation of the Isthmus of Panama. Science Advances, 2(8), e1600883, doi: 10.1126/sciadv.1600883, 2016.

730 Pagani, M., Freeman, K.H. and Arthur, M.A.: Late Miocene atmospheric CO_2 concentrations and the expansion of C4 grasses. Science, 285(5429), 876-879, doi: 10.1126/science.285.5429.876, 1999.

Pearson, P.N. and Palmer, M.R.: Atmospheric carbon dioxide concentrations over the past 60 million years. Nature, 406(6797), 695, doi: 10.1038/35021000, 2000.

Pearson, P.N.: Oxygen isotopes in foraminifera: Overview and historical review. The Paleontological Society Papers, 18, 1-
735   38, doi: 10.1017/S1089332600002539, 2012.

Peterson, C.D., Lisiecki, L.E. and Stern, J.V.: Deglacial whole-ocean $\delta^{13}C$ change estimated from 480 benthic foraminiferal records. Paleoceanography, 29(6), 549-563, doi: 10.1002/2013PA002552, 2014.

Pierce, E.L., van de Flierdt, T., Williams, T., Hemming, S.R., Cook, C.P. and Passchier, S.: Evidence for a dynamic East Antarctic ice sheet during the mid-Miocene climate transition. Earth Planet. Sci. Lett., 478, 1-13, doi: 10.1016/j.epsl.2017.08.011, 2017.

PLATES University of Texas: http://www-udc.ig.utexas.edu/external/plates/, last access 14 December 2019.

Pound, M.J., Haywood, A.M., Salzmann, U. and Riding, J.B.: Global vegetation dynamics and latitudinal temperature gradients during the Mid to Late Miocene (15.97–5.33 Ma). Earth-Sci. Rev., 112(1-2), 1-22, doi: 10.1016/j.earscirev.2012.02.005, 2012.

Price, A. R., Myerscough, R. J., Voutchkov, I. I., Marsh, R., and Cox, S. J.: Multi-objective optimization of GENIE Earth system models, Phil. Trans. R. Soc. A, 367, 2623–2633, 2009.

Raymo, M.E., Ruddiman, W.F. and Froelich, P.N.: Influence of late Cenozoic mountain building on ocean geochemical cycles. Geology, 16(7), 649-653, doi: 10.1130/0091-7613(1988)016<0649:IOLCMB>2.3.CO;2, 1988.

Reiners, P.W., Ehlers, T.A., Garver, J.I., Mitchell, S.G., Montgomery, D.R., Vance, J.A. and Nicolescu, S.: Late Miocene exhumation and uplift of the Washington Cascade Range. Geology, 30(9), 767-770, doi: 10.1130/0091-7613(2002)030<0767:LMEAUO>2.0.CO;2, 2002.

Richey, J.N. and Tierney, J.E.: GDGT and alkenone flux in the northern Gulf of Mexico: Implications for the TEX86 and UK'37 paleothermometers. Paleoceanography, 31(12), 1547-1561, doi: 10.1002/2016PA003032, 2016.

Ridgwell, A. and Schmidt, D.N.: Past constraints on the vulnerability of marine calcifiers to massive carbon dioxide release, Nat. Geosci., 3(3), 196. doi: 10.1038/ngeo755, 2010.

Ridgwell, A., Hargreaves, J.C., Edwards, N.R., Annan, J.D., Lenton, T.M., Marsh, R., Yool, A. and Watson, A.: Marine geochemical data assimilation in an efficient Earth System Model of global biogeochemical cycling. Biogeosciences, 4(1), 87-104, doi: 10.5194/bg-4-87-2007, 2007.

Rohling E.J.: Oxygen isotope composition of seawater, in: The Encyclopedia of Quaternary Science, edited by: Elias S.A., Amsterdam: Elsevier. 915-922, 2013.

Schmidt, G.A. and Shindell, D.T.: Atmospheric composition, radiative forcing, and climate change as a consequence of a massive methane release from gas hydrates. Paleoceanography, 18(1), doi: 10.1029/2002PA000757, 2003.

Sepulchre, P., Arsouze, T., Donnadieu, Y., Dutay, J.C., Jaramillo, C., Le Bras, J., Martin, E., Montes, C. and Waite, A.J.: Consequences of shoaling of the Central American Seaway determined from modeling Nd isotopes. Paleoceanography, 29(3), 176-189, doi: 10.1002/2013PA002501, 2014.

Sijp, W.P., Anna, S., Dijkstra, H.A., Flögel, S., Douglas, P.M. and Bijl, P.K.: The role of ocean gateways on cooling climate on long time scales. Global Planet. Change, 119, 1-22, doi: 10.1016/j.gloplacha.2014.04.004, 2014.

Sosdian, S.M., Greenop, R., Hain, M.P., Foster, G.L., Pearson, P.N. and Lear, C.H.: Constraining the evolution of Neogene ocean carbonate chemistry using the boron isotope pH proxy. Earth Planet. Sci. Lett., 498, 362-376, doi: /10.1016/j.epsl.2018.06.017, 2018.

Srinivasan, M.S. and Sinha, D.K.: Early Pliocene closing of the Indonesian Seaway: evidence from north-east Indian Ocean and Tropical Pacific deep sea cores. J. Asian Earth Sci., 16(1), 29-44, doi: 10.1016/S0743-9547(97)00041-X, 1998.

Stewart, D.R., Pearson, P.N., Ditchfield, P.W. and Singano, J.M.: Miocene tropical Indian Ocean temperatures: evidence from three exceptionally preserved foraminiferal assemblages from Tanzania. Journal of African Earth Sciences, 40(3-4), 173-189, doi: 10.1016/j.jafrearsci.2004.09.001, 2004.

Tyrrell, T. and Zeebe, R.E.: History of carbonate ion concentration over the last 100 million years. Geochim. Cosmochim. Ac., 68(17), 3521-3530, doi: 10.1016/j.gca.2004.02.018, 2004.

Urban, T. and Hardisty, F.: Developing PAMS—A Paleolocation Web Service. In International Cartographic Conference. Dresden, Germany (Vol. 128), 2013.

Valdes, P.J., Armstrong, E., Badger, M.P., Bradshaw, C.D., Bragg, F., Davies-Barnard, T., Day, J.J., Farnsworth, A., Hopcroft, P.O., Kennedy, A.T. and Lord, N.S.: The BRIDGE HadCM3 family of climate models: HadCM3@ Bristol v1. 0. Geosci. Model Dev., 10, 3715-3743, doi: 10.5194/gmd-10-3715-2017, 2017.

Watterson, I.G.: Non-dimensional measures of climate model performance. Int. J. Climatol., 16(4), 379-391, doi: 10.1002/(SICI)1097-0088(199604)16:4<379::AID-JOC18>3.0.CO;2-U, 1996.

Wolfe, J.A.: Distribution of major vegetation types during the Tertiary: 357–375. The carbon cycle and atmospheric CO$_2$: natural variations Archean to present. Am. Geophys. Un., Washington, DC, 1985.

Zachos, J.C., Dickens, G.R. and Zeebe, R.E.: An early Cenozoic perspective on greenhouse warming and carbon-cycle dynamics. Nature, 451(7176), 279, doi: 10.1038/nature06588, 2008.

Zweng, M.M., Reagan, J.R., Antonov, J.I., Locarnini, R.A., Mishonov, A.V., Boyer, T.P., Garcia, H.E., Baranova, O.K., Johnson, D.R., Seidov, D. and Biddle, M.M., 2013. World ocean atlas 2013. Volume 2, Salinity. (eds Levitus, S. & Mishonov, A.) NOAA Atlas NESDIS 74, 2013.

Data slice (Ma)	HadCM3 slice		
	Age name	Start (Ma)	Finish (Ma)
15	Langhian	15.97	13.82
12.5	Serravalian	13.82	11.63
10	Tortonian	11.63	7.246
7.5	Messinian	7.246	5.333
4.5	Zanclean	5.333	3.6
2.5	Piacenzian	3.6	2.58
Core top	Holocene	0.006	

795 **Table 1, Dataslices for the carbon cycling approximate locations (in time) and the HadCM3 setups used to create Genie bathymetry. 7.5Ma and the 2.5Ma timeslice are just outside the ages to which they are assigned, but in order to not have two points in the other ages, they place in the Messinian and Piacenzian respectively**

800

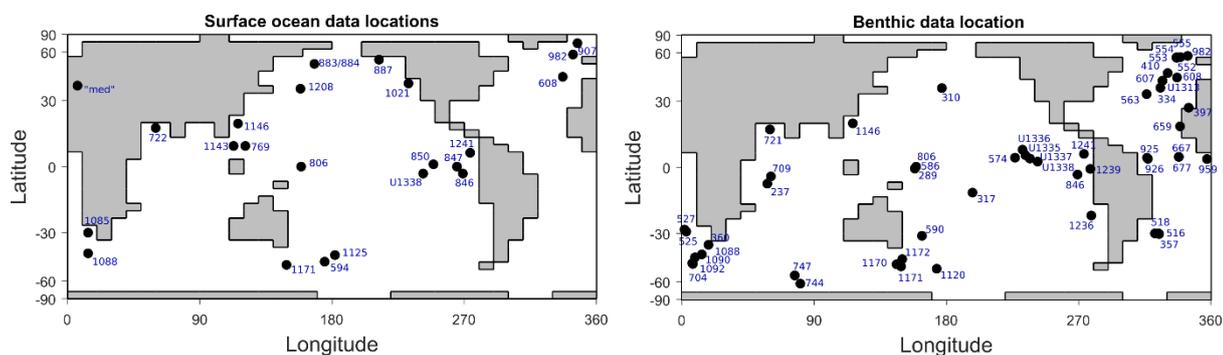



Figure 1 Present day location of surface ocean temperature data points (left) and benthic δ¹⁸O and δ¹³C data points (right).

805

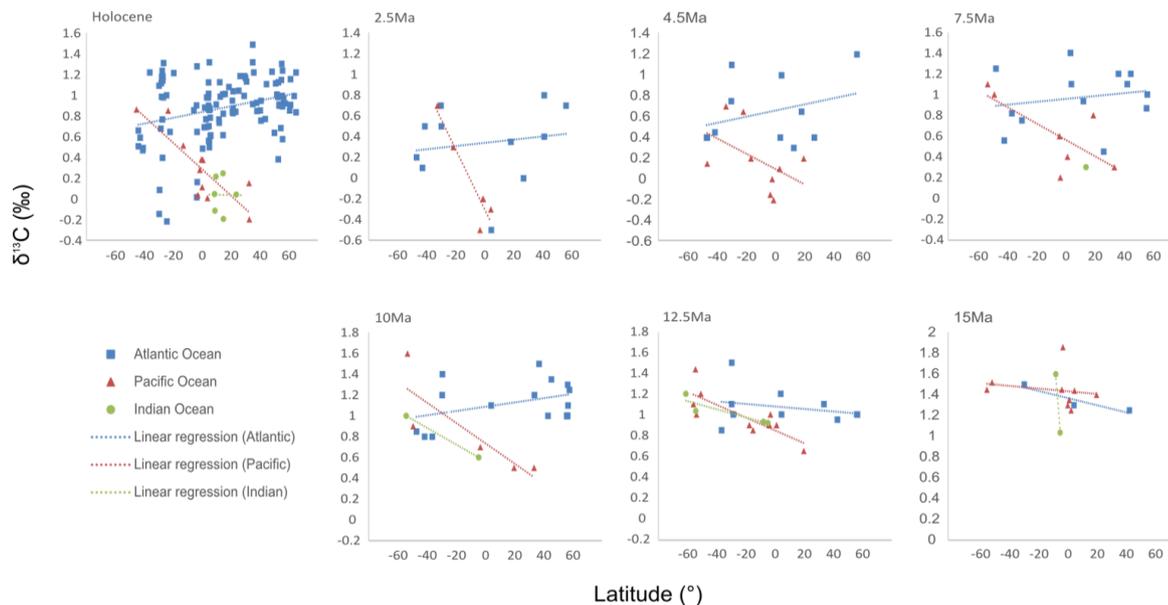



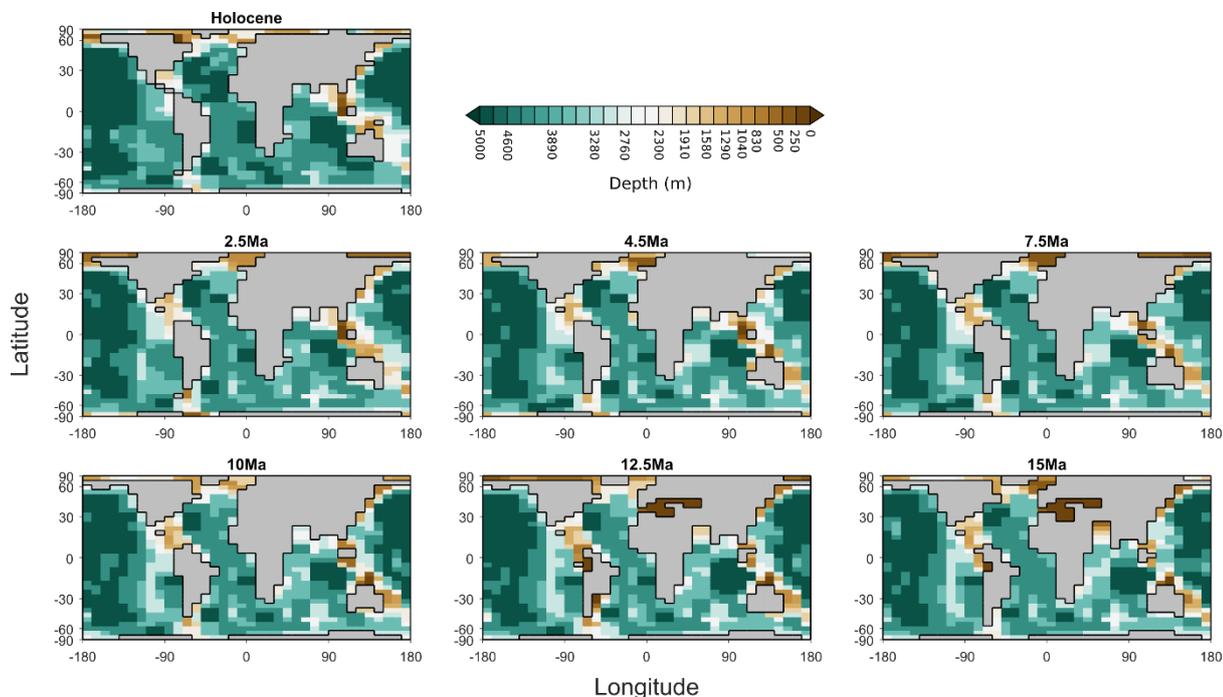
Figure 2, Benthic δ¹³C data plotted against its paleo-latitude, with linear regressions overlaid for each ocean basin.

810

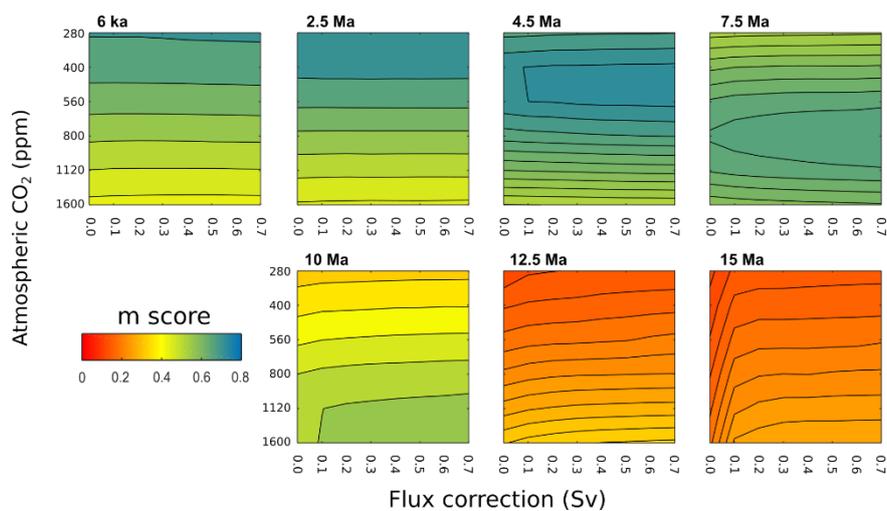
Figure 3, Model bathymetry and coastlines for each time-slice.

815

Figure 4, M-score for fit of model to data for surface temperature, plotted per CO₂ and per flux correction setting as a contour. All model-data pairs are available in supplementary Fig. S2.

820

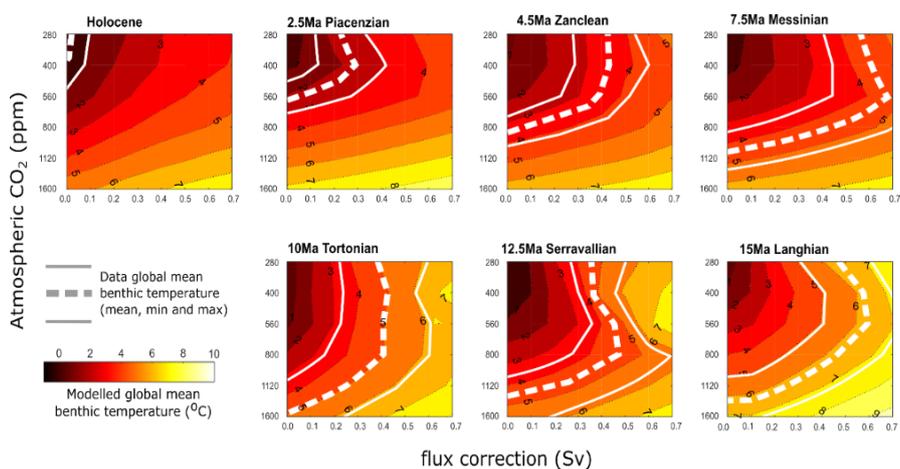



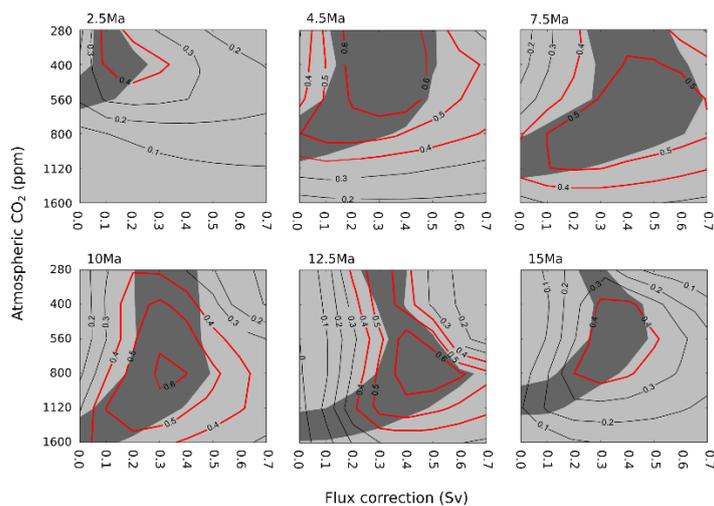
Figure 5, Modelled benthic temperature with global mean temperature estimate (Cramer et al., 2011) overlaid (dashes) and
825 **minimum and maximum estimates (white lines)**

830

Figure 6, M-score (contours) and bias of less than 0.5 °C (dark grey shaded) for model fit to benthic data temperature set derived from δ^{18}O data (see text for method). All model-data pairs are available in supplementary Fig. S3.

835

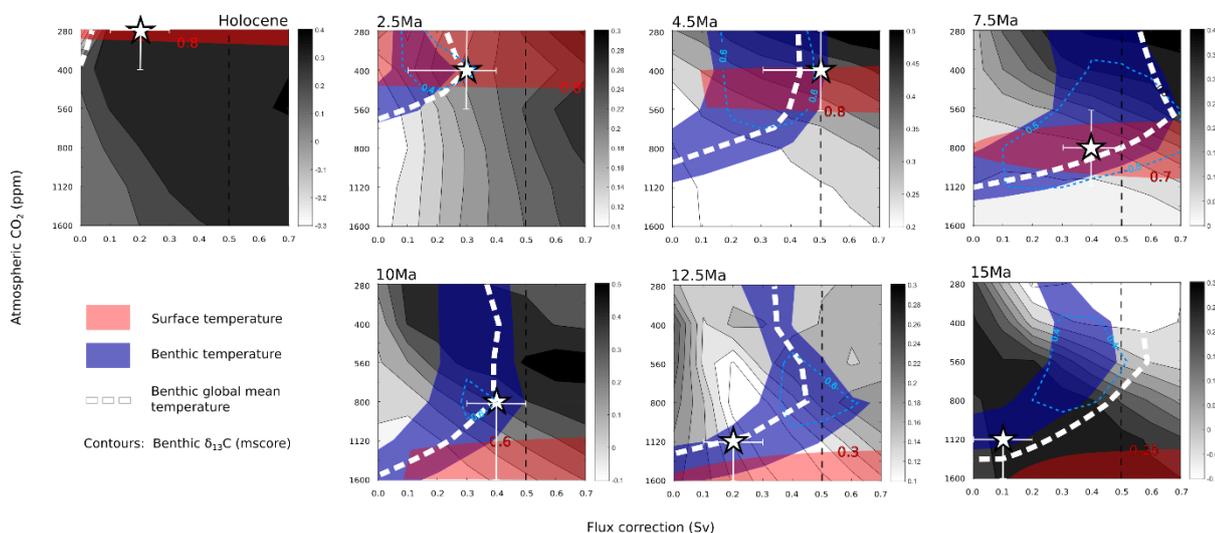



Figure 7, All proxy data constraints. M-score for modelled benthic δ¹³C fit to dataset (filled contours); modelled benthic global mean temperature lowest offset to Cramer et al., 2011 (dashed white line); modelled benthic temperature fit to dataset calculated from δ¹⁸O data (blue filled region where mean bias is ±0.5 degrees, blue dashed line is highest M-score at one decimal place); M-score for modelled surface ocean temperature fit to data (red filled region is highest, to one decimal place, M-score except for 15 Ma where shaded is for M-score > 0.25). Overall best-fit settings (white star, with upper and lower range estimates)

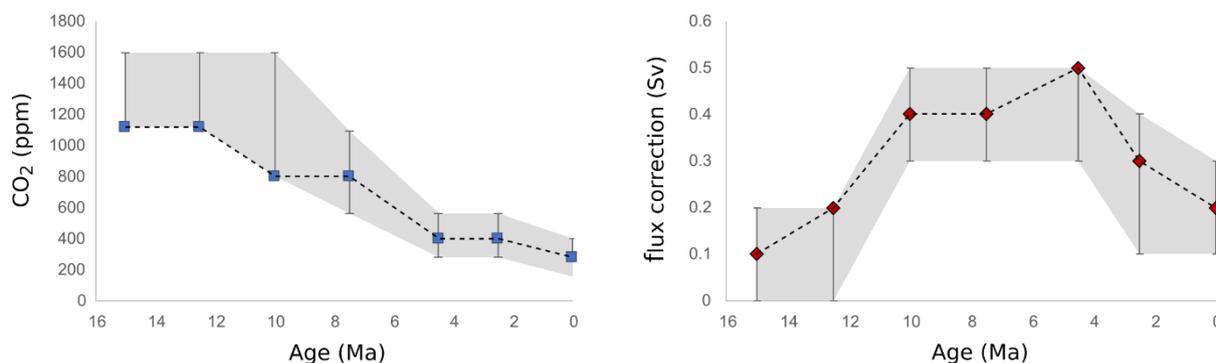



Figure 8, CO₂ and flux correction settings that agree with temperature and benthic δ¹³C data according to this study. Selected settings (shown as stars in Fig. 7) marked with points, grey shaded area is estimated ranges that show some agreement with data.

855

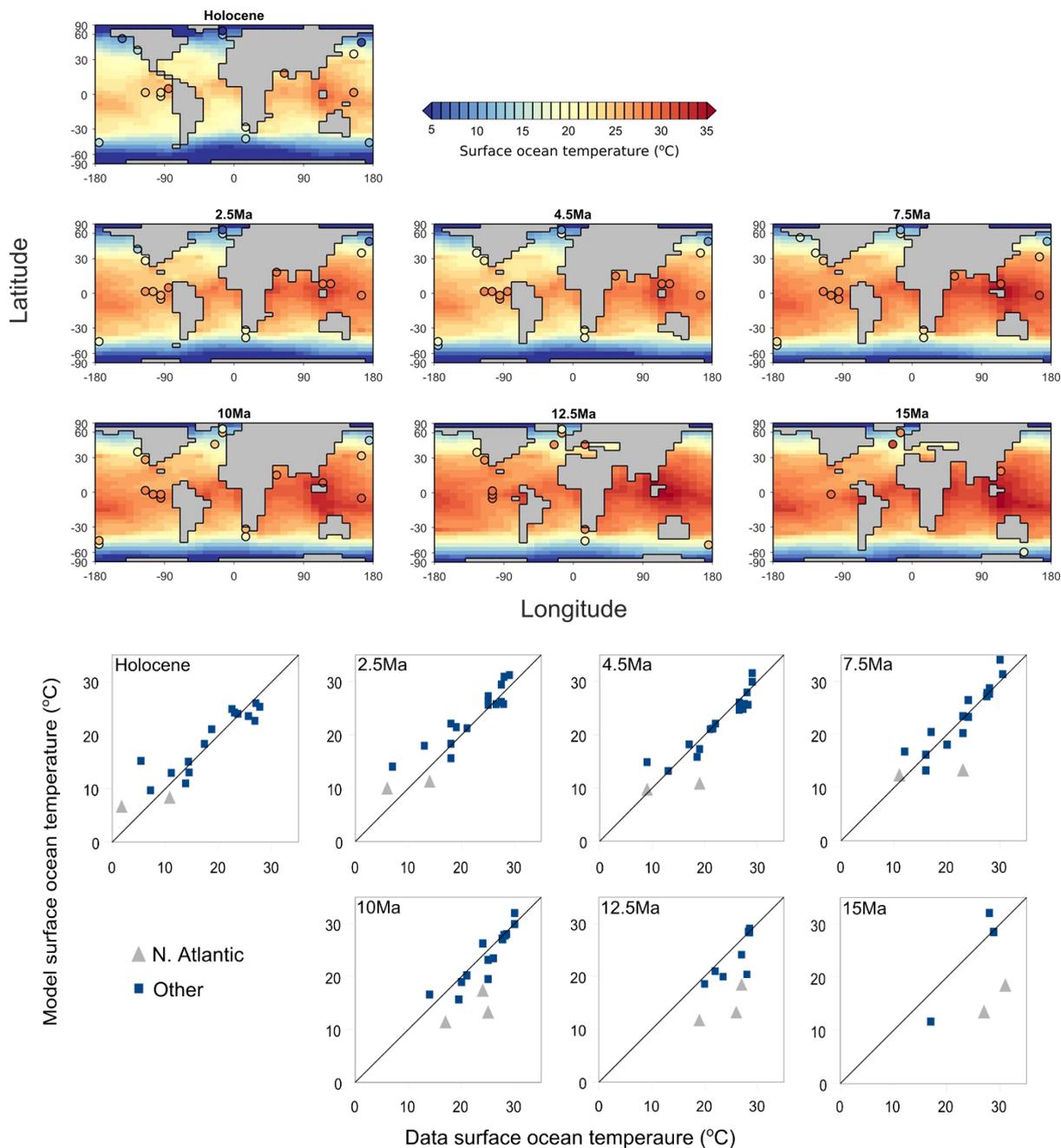



Figure 9, Surface temperature, Top: for best fit model settings with data overlaid (shaded circles), bottom: cross plots of data points
860 **(x axis) and model points (y axis), with North Atlantic data shown as grey triangles, other regions shown as blue squares.**

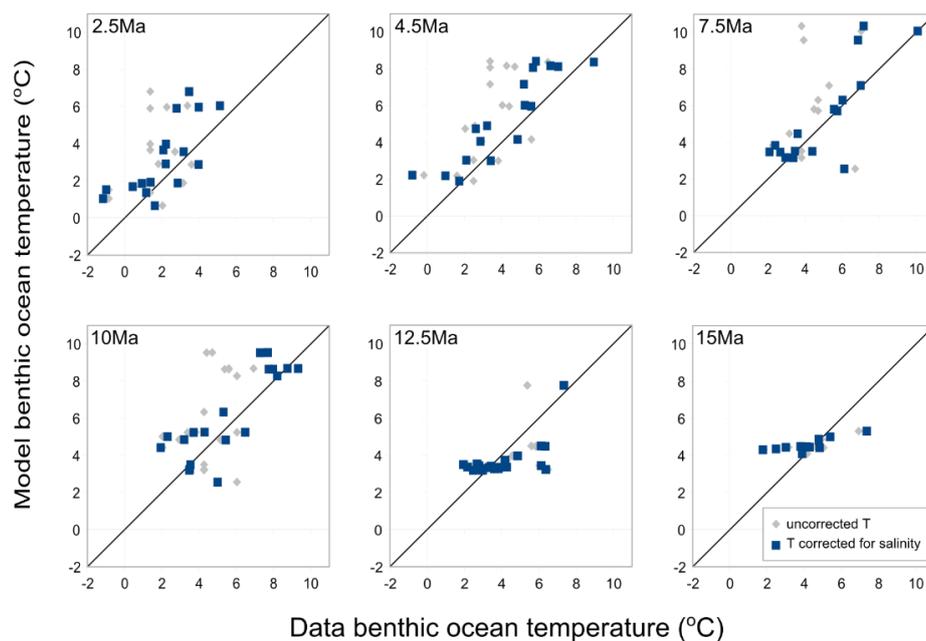
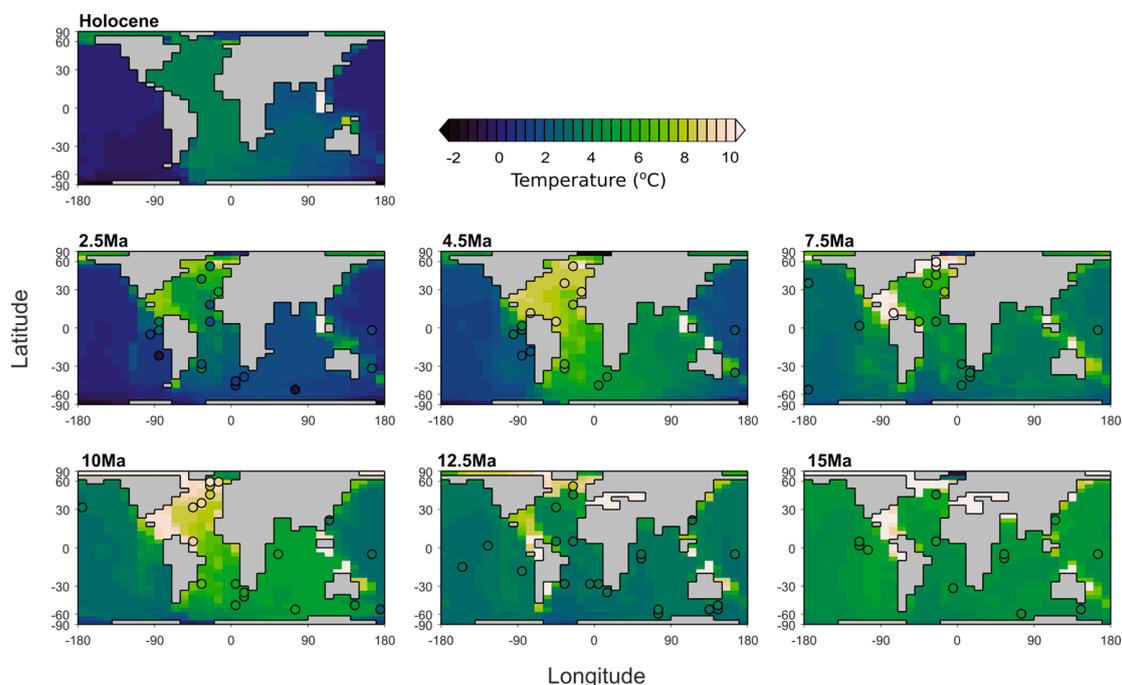



Figure 10, Benthic temperature, Top: for best fit model settings with data overlaid (shaded circles), bottom: cross plots of data points (x axis) and model points (y axis), with temperature calculated with no salinity correction (grey diamonds) and with salinity correction (blue squares).

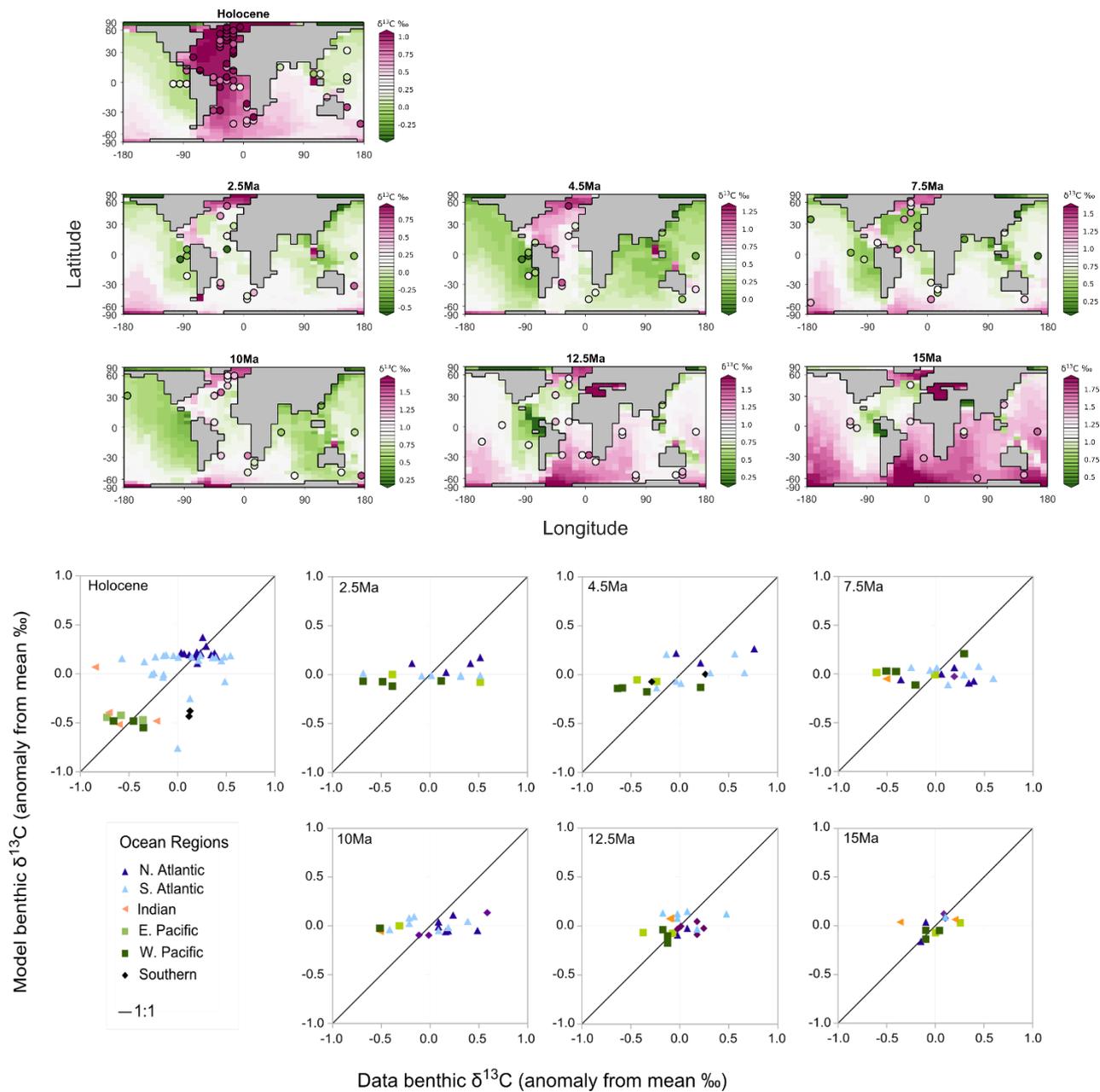



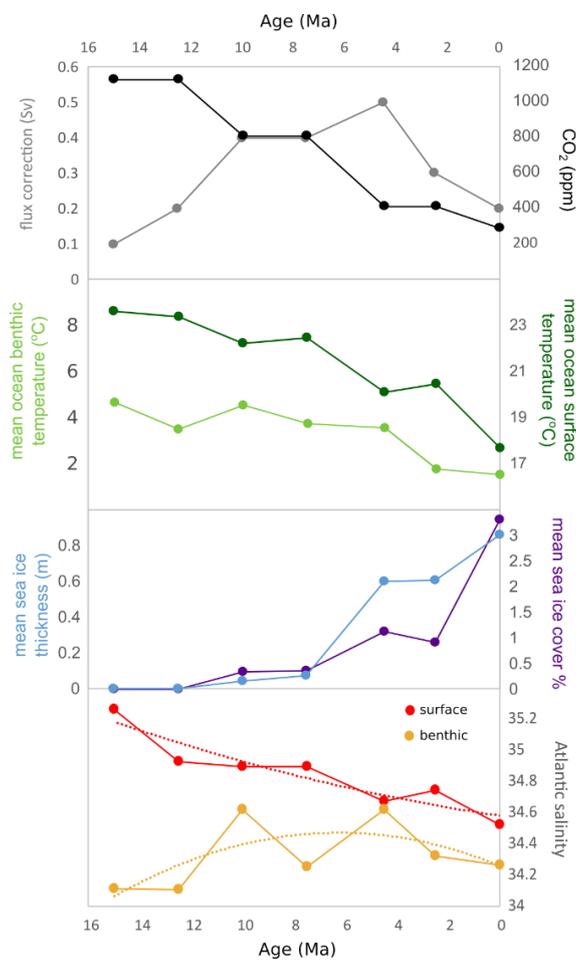
Figure 11, δ¹³C, Top: for best fit model settings with data overlaid (shaded circles). Shading relative to offset from the predicted atmospheric δ¹³CO₂, for example: at 15 Ma the deep ocean is heavier relative to the atmosphere than at 4.5 Ma. Bottom: cross plots of data points (x axis) and model points (y axis), with markers denoting ocean region.

880

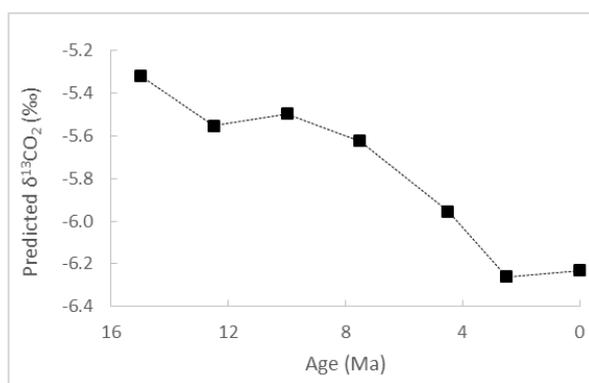
Figure 12, Modelled (from top) CO₂, flux correction (FwF), ocean temperature (surface and benthic global means), sea ice (mean thickness and mean global % cover), and North Atlantic salinity (surface and benthic) for the best-fit settings.

885

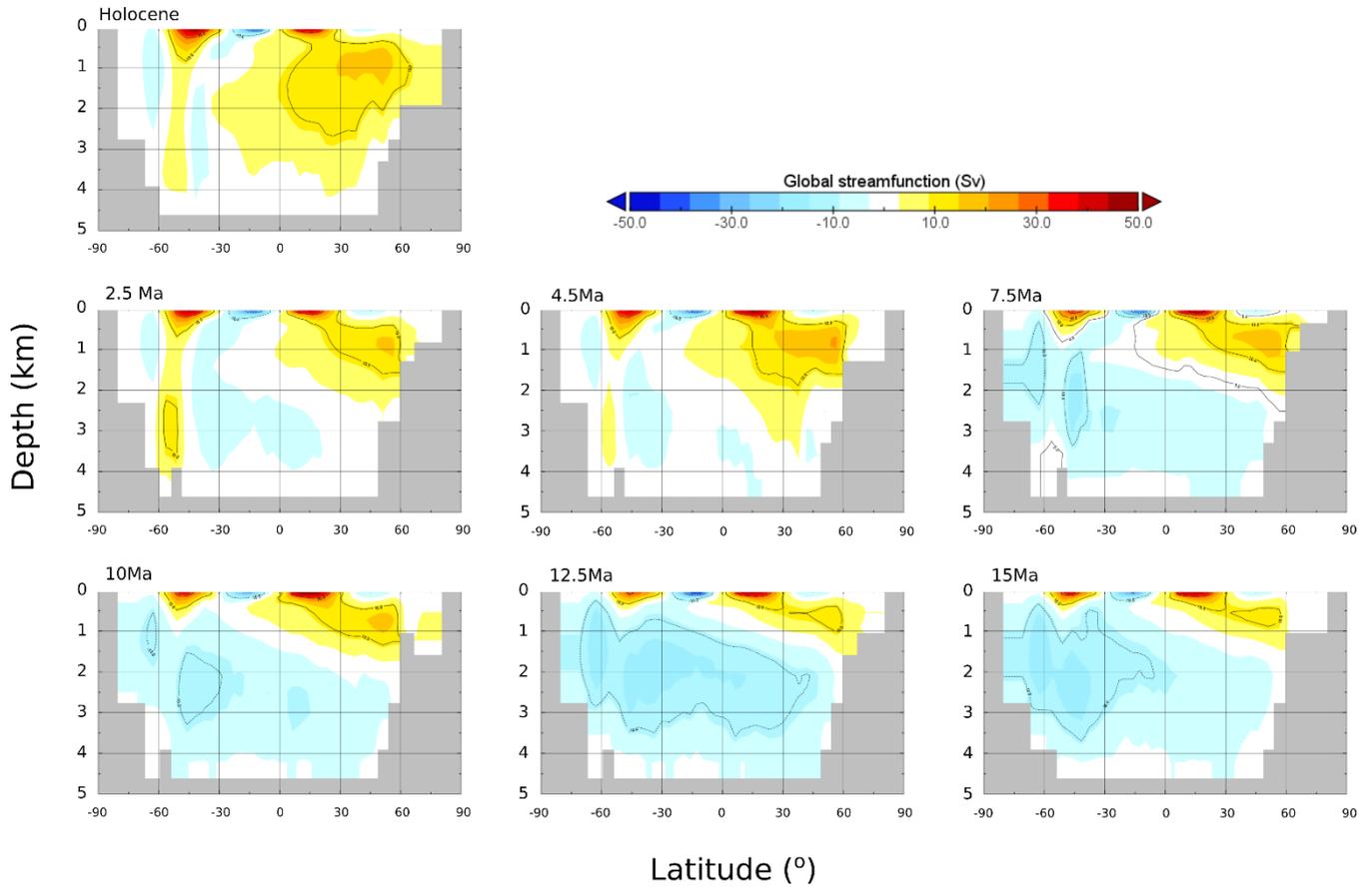
Figure 13, Predicted atmospheric δ¹³CO₂ based on the mean offset of the best fit model to data benthic ocean δ¹³C.

890

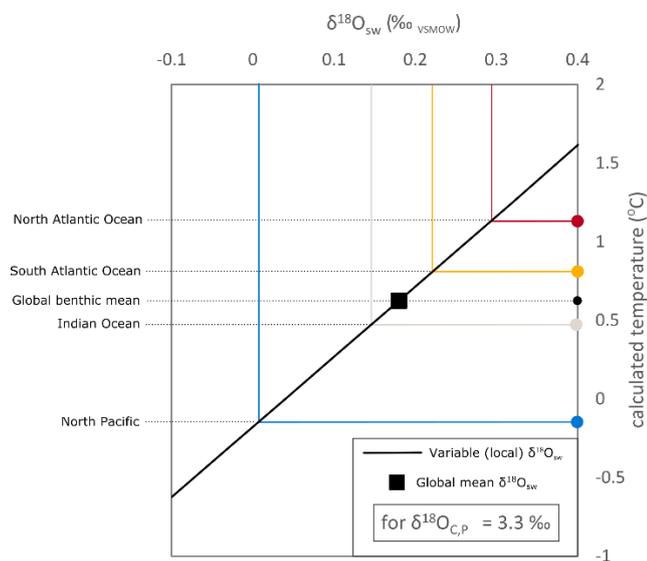
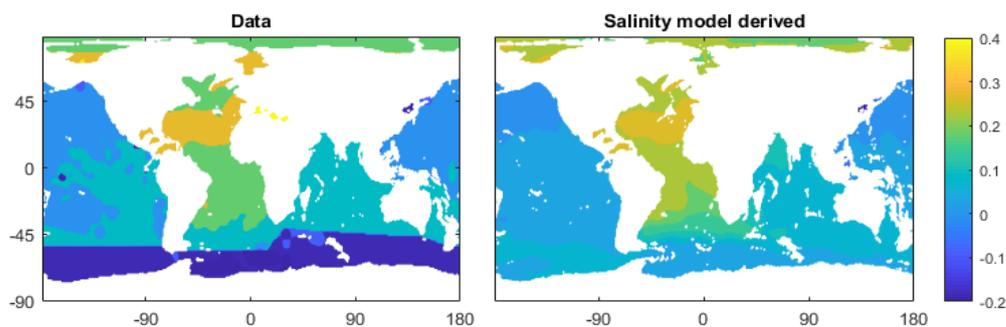
Figure 14, Modelled global stream function (Sv) for each time slice for the best-fit model settings.

895 **Figure A1, Example paleotemperature calculation of Marchitto et al. (2014) and the $\delta^{18}O_{sw}$ from ocean basins (values from Fig. C2 for ocean basins' means), and the global mean $\delta^{18}O_{sw}$ from Cramer et al. (2011) for the Holocene (which may be different to the true present day global benthic mean). The corrected N. Atlantic is ~0.5 °C warmer and the N.Pacific ~0.7 °C cooler than the when using only the global mean $\delta^{18}O_{sw}$.**

900

Figure A2, $\delta^{18}O_{sw}$ data from LeGrande and Schmidt, 2006 (left) and salinity derived $\delta^{18}O_{sw}$ (right, where salinity data from WOA 2013, Zweng et al., 2013, model is Eq. A1) (all in ‰VSMOW).

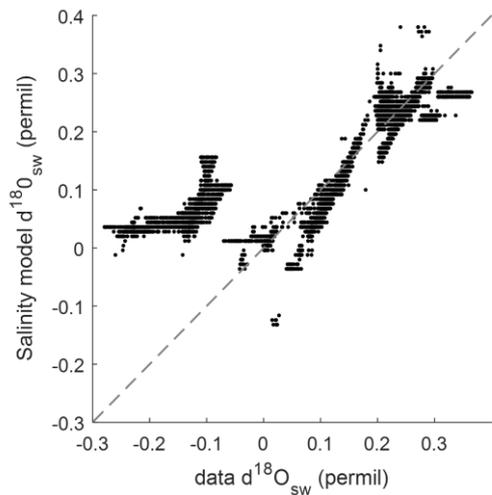



Figure A3, Cross plot of data $\delta^{18}O_{sw}$ and salinity model derived $\delta^{18}O_{sw}$ as shown in Fig. A2.

905