Clim. Past, 21, 973–1000, 2025
https://doi.org/10.5194/cp-21-973-2025
© Author(s) 2025. This work is distributed under
the Creative Commons Attribution 4.0 License.

Climate of the Past
Open Access
EGU

Mean ocean temperature change and decomposition of the benthic δ^{18}O record over the past 4.5 million years

Peter U. Clark[1,2,♀,★], **Jeremy D. Shakun**[3,★], **Yair Rosenthal**[4,5], **Chenyu Zhu**[6], **Patrick J. Bartlein**[7], **Jonathan M. Gregory**[8,9], **Peter Köhler**[10], **Zhengyu Liu**[11], **and Daniel P. Schrag**[12]

[1]College of Earth, Ocean, and Atmospheric Sciences, Oregon State University, Corvallis, OR 97331, USA
[2]School of Geography and Environmental Sciences, University of Ulster, Coleraine, BT52 1SA, Northern Ireland, UK
[3]Department of Earth and Environmental Sciences, Boston College, Chestnut Hill, MA 02467, USA
[4]Department of Marine and Coastal Science, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901, USA
[5]Department of Earth and Planetary Sciences, Rutgers, The State University of New Jersey, New Brunswick, NJ 08901, USA
[6]Institute of Atmospheric Physics, Chinese Academy of Sciences, Beijing, China
[7]Department of Geography, University of Oregon, Eugene, OR 97403-1251, USA
[8]National Center for Atmospheric Science, University of Reading, Reading, UK
[9]Met Office Hadley Centre, Exeter, UK
[10]Alfred-Wegener-Institut Helmholtz-Zentrum für Polar- und Meeresforschung, Bremerhaven, Germany
[11]Department of Geography, The Ohio State University, Columbus, OH 43210, USA
[12]Department of Earth and Planetary Sciences, Harvard University, Cambridge, MA 02138, USA
[♀]*Invited contribution by Peter U. Clark, recipient of the EGU Milutin Milanković Medal 2024.*
[★]These authors contributed equally to this work.

Correspondence: Peter U. Clark (clarkp@onid.orst.edu)

Received: 26 September 2024 – Discussion started: 27 September 2024
Revised: 1 February 2025 – Accepted: 4 March 2025 – Published: 3 June 2025

Abstract. We use a recent reconstruction of global mean sea surface temperature change relative to preindustrial (ΔGMSST) over the last 4.5 Myr together with independent proxy-based reconstructions of bottom water (ΔBWT) or deep-ocean (ΔDOT) temperatures to infer changes in mean ocean temperature (ΔMOT). Three independent lines of evidence show that the ratio of ΔMOT / ΔGMSST, which is a measure of ocean heat storage efficiency (HSE), increased from ∼ 0.5 to ∼ 1 during the Middle Pleistocene Transition (MPT, 1.5–0.9 Ma), indicating an increase in ocean heat uptake (OHU) at this time. The first line of evidence comes from global climate models; the second from proxy-based reconstructions of ΔBWT, ΔMOT, and ΔGMSST; and the third from decomposing a global mean benthic δ^{18}O stack ($\delta^{18}O_b$) into its temperature ($\delta^{18}O_T$) and seawater ($\delta^{18}O_{sw}$) components. Regarding the latter, we also find that further corrections in benthic δ^{18}O, probably due to some combination of a long-term diagenetic overprint and to the carbonate ion effect, are necessary to explain reconstructed Pliocene sea-level highstands inferred from $\delta^{18}O_{sw}$. We develop a simple conceptual model that invokes an increase in OHU and HSE during the MPT in response to changes in deep-ocean circulation driven largely by surface forcing of the Southern Ocean. Our model accounts for heat uptake and temperature in the non-polar upper ocean (0–2000 m) that is mainly due to wind-driven ventilation, while changes in the deeper ocean (> 2000 m) in both polar and non-polar waters occur due to high-latitude deepwater formation. We propose that deepwater formation was substantially reduced prior to the MPT, effectively decreasing HSE. We attribute these changes in deepwater formation across the MPT to long-term cooling which caused a change starting ∼ 1.5 Ma from a highly stratified Southern Ocean due to warm SSTs and reduced sea-ice extent to a Southern Ocean which, due to colder SSTs and increased sea-ice extent, had a greater vertical exchange of water masses.

1 Introduction

The ocean is one of the largest heat reservoirs in the climate system. Changes in ocean heat content play an important role in mitigating the Earth's surface temperature response to a radiative forcing (Gregory et al., 2002) and also influence sea-level change (Church et al., 2013) and ocean stratification, with the latter affecting the rate of ocean heat uptake (OHU) (Newsom et al., 2023), the efficiency of the oceanic carbon sink (Bronselaer and Zanna, 2020), and large-scale circulation (Fox-Kemper et al., 2021). Proxy-based reconstructions of changes in bottom water temperature (ΔBWT), deep-ocean temperature (ΔDOT), and mean ocean temperature (ΔMOT)[1] relative to preindustrial (PI) identify large variations in ocean temperature, ocean heat content, and energy imbalance on 10^3–10^6-year timescales (Baggenstos et al., 2019; Rohling et al., 2022; Shackleton et al., 2023). These include orbital-scale variations of ∼ 2.5 to 3.5 °C over the last ∼ 0.7 Myr (Sosdian and Rosenthal, 2009; Elderfield et al., 2012; Shakun et al., 2015; Haeberli et al., 2021; Shackleton et al., 2023; Martin et al., 2002) and a long-term cooling over much of the past 4.5 Myr (Bates et al., 2014; Hansen et al., 2013; Lear et al., 2003; Rohling et al., 2021; Cramer et al., 2011; de Boer et al., 2014; Rohling et al., 2022; Westerhold et al., 2020; Evans et al., 2024) (Fig. 1).

In detail, however, these long-term reconstructions can differ by as much as 2–3 °C (Fig. 1). Differences among the reconstructions on these longer timescales may reflect some combination of (1) differences in the calibration of Mg / Ca-based reconstructions as well as in the process of accounting for changes in its seawater (sw) ratio Mg / Ca_{sw} and carbonate ion concentration (Cramer et al., 2011; Rosenthal et al., 2022), (2) proxy-based reconstructions that sample local BWT that is not representative of DOT or MOT (Lear et al., 2003; Woodard et al., 2014), (3) an unaccounted-for nonstationarity in the relationship between $\delta^{18}\text{O}_\text{b}$ and sea level that is used to derive BWT or DOT (Bates et al., 2014; Rohling et al., 2021; Rohling et al., 2022; Waelbroeck et al., 2002), and (4) scaling of a $\delta^{18}\text{O}_\text{b}$ range to an inferred DOT range (LGM–Holocene) (Hansen et al., 2013; Westerhold et al., 2020; Hansen et al., 2023) that assumes stationarity in this scaling over the last 4.5 Ma and underestimates the DOT range by ∼ 70 % (Haeberli et al., 2021; Shackleton et al., 2023). It is thus important to narrow these uncertainties in the ocean's temperature history in order to gain a better understanding of the role of the ocean heat reservoir in the Earth's energy budget as well as the processes that contribute to OHU.

Here we find that climate models and proxy data from the last 4.5 Myr show good agreement in the ratio between ΔMOT and changes in global mean sea surface temperature (ΔGMSST), which Zhu et al. (2024) defined as the ocean heat storage efficiency (HSE = ΔMOT / ΔGMSST) when in equilibrium, suggesting that we can use a new ΔGMSST reconstruction (Clark et al., 2024) to derive ΔMOT over this interval.[2] A particularly notable finding is that HSE increased from ∼ 0.5 to ∼ 1 when ΔGMSSTs decreased to below ∼ 0 °C, corresponding to the start of the Middle Pleistocene Transition (MPT, 0.9–1.5 Ma). We use this understanding of ΔMOT to decompose the benthic $\delta^{18}\text{O}$ record ($\delta^{18}\text{O}_\text{b}$) into its temperature ($\delta^{18}\text{O}_\text{T}$) and seawater ($\delta^{18}\text{O}_\text{sw}$) components, with this decomposition providing independent support for the need to increase HSE during the MPT. We then summarize the processes that contribute to changes in ocean heat storage and mean ocean temperature which in turn provide the basis for a simple conceptual model that explains why HSE may have increased during the MPT.

2 Derivation of ocean heat storage efficiency

2.1 Constraints from climate models

We obtained ΔMOTs and ΔGMSSTs from 93 coupled ocean–atmosphere climate model equilibrium simulations that were run with atmospheric CO_2 concentrations ranging from 180 ppm to 9 times PI (Fig. 2) (Alder and Hostetler, 2015; Galbraith and de Lavergne, 2019; Haywood et al., 2020; Clark et al., 2016; Goudsmit-Harzevoort et al., 2023; Rugenstein et al., 2019; He, 2011; Bereiter et al., 2018). Here we use four regression models to describe the relationship

---

[1] The terms "bottom water temperature", "deep-ocean temperature", and "deep-sea temperature" are commonly used interchangeably in paleoceanography to refer to depths > 200 m. In some cases, these terms have been used for the temperature of a specific site (Waelbroeck et al., 2002; Sosdian and Rosenthal, 2009; Elderfield et al., 2012) or for the entire ocean > 200 m (Hansen et al., 2013; Rohling et al., 2022). The latter meaning is not equivalent to "mean ocean temperature" (MOT) since that includes the surface ocean, but the small volume of this surface layer means that the global ocean temperature > 200 m will be comparable to MOT. Finally, these terms do not distinguish between water depths that are commonly used to describe processes and patterns of ocean heat storage: sea surface (skin) with zero heat capacity, upper (0–700 m), intermediate (700–2000 m), deep (2000–4000 m), and abyssal (> 4000 m) (Purkey et al., 2019; Fox-Kemper et al., 2021; Cheng et al., 2022). Here we use the term "bottom water temperature" (BWT) for a site-specific temperature reconstruction and "deep-ocean temperature" (DOT) for temperature of the whole ocean that is > 200 m. We also distinguish among the depth layers (upper, intermediate, deep, and abyssal) when discussing ocean heat storage.

[2] HSE can be interpreted as proportional to the effective heat capacity of the ocean for equilibrium changes, while the change in ocean heat content (in J) can be accounted for by regarding it as the (SST) × (HSE) × (the heat capacity of the ocean). Because HSE refers to equilibrium, it differs from ocean heat uptake efficiency, which specifically describes transient climate states on decadal timescales during which the non-equilibrium of the ocean that causes a substantial rate of ocean heat uptake (in W m^{-2}) is large enough to affect the surface climate through its effect on the energy balance.

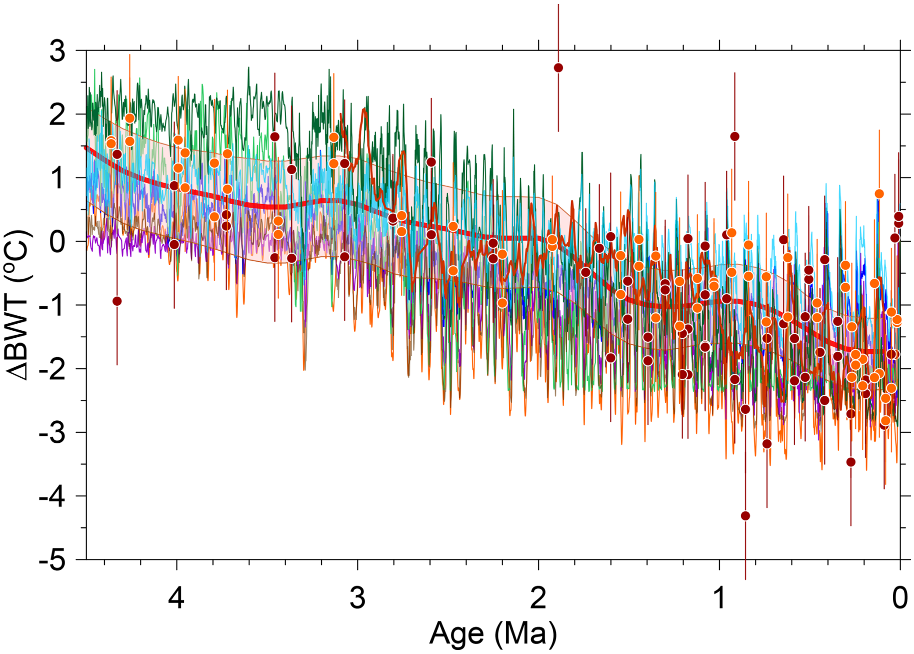

Figure 1. Reconstructions of changes in bottom water temperature (ΔBWT) since 4.5 Ma. Dark-green line is from Evans et al. (2024), referenced to preindustrial (PI). Purple line is from Rohling et al. (2021), referenced to PI. Dark-brown line is from Rohling et al. (2022), referenced to PI. BWTs were mean-shifted by 2.1 °C so that they have the same long-term mean as our ΔMOT reconstruction over the last 400 kyr. Blue line is the ΔBWT reconstruction from Hansen et al. (2013). BWTs were mean-shifted by 1 °C so that they are referenced to the Holocene (i.e., 0 °C for the last 10 kyr). Light-green line is the ΔBWT reconstruction from Bates et al. (2014). BWTs were mean-shifted by 1.46 °C so that they are referenced to the Holocene (i.e., 0 °C for the last 10 kyr). Orange line is the ΔMOT reconstruction from de Boer et al. (2014), referenced to PI. Orange symbols with 1σ uncertainty are the ΔBWT reconstruction from ODP site 806 (Lear et al., 2003). Mg / Ca BWTs from this data set were mean-shifted by −1.8 °C for *O. umbonatus* and by −4 °C for *C. wuellerstorfi* to fall within same range as our ΔMOT reconstruction for the LGM. Red-brown symbols with 1σ uncertainty are the ΔBWT reconstruction from ODP site 926 (Lear et al., 2003). Mg / Ca BWTs from this data set were mean-shifted by −2.2 °C for *O. umbonatus* and by −5 °C for *C. wuellerstorfi* so that they are referenced to the Holocene (i.e., 0 °C for the last 10 kyr). Red bold line with uncertainty in pink (90 % confidence interval) is the smoothed Mg / Ca-based ΔBWT reconstruction from Cramer et al. (2011) removing periods shorter than 2 Myr and their Eq. (7b) for Mg / Ca temperature calibration.

between ΔMOT and ΔGMSST in this synthesis of climate model output. The four models were fit using R *stats* package functions *lm*, *segmented*, *loess*, and *smoothing.spline*, as well as their associated functions for prediction and diagnostic analyses (R, 2024) (see Sect. S1 in Supplement). The relationship between ΔMOT and ΔGMSST is broadly linear, with a slope and intercept close to 1.0 and 0.0, respectively (Figs. 2a, S1A); a high R^2 value (0.924); a highly significant F-test statistic (1104, $p < 0.0001$); and a relatively low residual standard error (SE; 1.184 °C). However, a residual diagnostic analysis of this model reveals a subtle but significant nonlinear lack of fit which can be seen in the residual scatter diagram (Fig. S1B). The nonlinearity in the residual scatter diagram is largely related to a quasi-linear cluster of points that plot below the main cloud of points (Fig. 2a), which include the points with the 10-largest negative residuals. The climate models that produced these data points are generally a mix of higher-than-present CO_2 simulations and sensitivity tests, but not all such experiments produce that pattern of large negative residuals. This eliminates the possibility that there is an underlying family-of-curves explanation for the apparent nonlinearity. The residuals are also not normally distributed (Shapiro–Wilk test $W = 0.932$, $p < 0.001$) and in fact are visibly bimodal, imparted by the cluster of negative residuals.

Inspection of the scatter diagrams suggests that an alternative to an overall straight-line fit might be a segmented linear fit (Muggeo, 2003), with the individual segments approximately defined by values of ΔGMSST less than 1.0 °C, between 1.0 and 5.0 or 6.0 °C, and greater than 6.0 °C. To test that notion, we used the R *segmented* function. The best fit (in terms of goodness-of-fit and residual diagnostics) was a model with two breakpoints, at 1.187 °C (SE = 1.307 °C) and 5.290 °C (SE = 0.855 °C), and slopes (i.e., HSE) for the lower, middle, and upper segments of 1.000 (SE = 0.178), 0.575 (SE = 0.185), and 1.244 (SE = 0.066), respectively (Figs. 2b, S1C). The model fit is slightly better than that for the linear one, with an adjusted R^2 value of

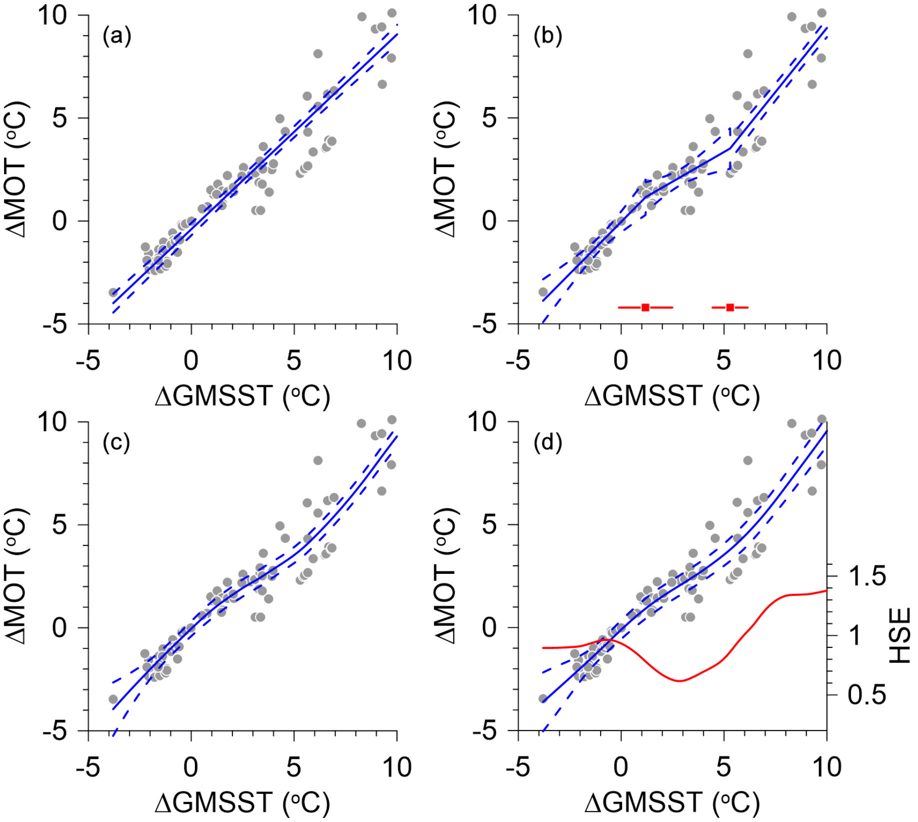

Figure 2. Relationship between equilibrium changes in mean ocean temperature (ΔMOT) and global mean sea surface temperature (ΔGMSST) from preindustrial (PI) established from 93 simulations by climate models. Results shown are from PMIP3 model runs for LGM boundary conditions (Braconnot et al., 2012), model runs for $CO_2 \leq 270$ ppm with no ice sheets and with LGM ice sheets (Galbraith and de Lavergne, 2019), model runs for the last deglaciation (Alder and Hostetler, 2015), model runs for $CO_2 \geq 420$ ppm (Clark et al., 2016), LongRunMIP model runs for $CO_2 \geq 2\times$ PI (Rugenstein et al., 2019), model runs for CO_2 ranging from 1× to 9× PI (Goudsmit-Harzevoort et al., 2023), PlioMIP2 model runs with $CO_2 = 400$ ppm (Haywood et al., 2020), and model runs for $CO_2 \geq 405$ ppm (Galbraith and de Lavergne, 2019). Fits use four different regression models: **(a)** *lm*, **(b)** *segmented*, **(c)** *loess*, and **(d)** *smoothing.spline*. The red dots and lines in panel **(b)** show the estimated breakpoints in the segmented regression (dots) and their uncertainties. The red line in panel **(d)** shows the calculated first derivative of the fitted values, i.e., the heat storage efficiency (HSE = ΔMOT / ΔGMSST).

0.939 (vs. 0.924 for the linear one). The segmented regression thus implies that there are distinct ΔGMSST thresholds (the breakpoints) where HSE changes.

We next considered models that can accommodate smooth variations in slope and can be fit via local- or nonparametric-regression approaches. Of these, we first considered a fit using the R *loess* function (Cleveland et al., 1992). The fitted values (Fig. 2c, S1E) were broadly similar to the individual slopes in the segmented regression (Fig. 2b), but the slope varied smoothly instead of abruptly at ΔGMSST values around 1.0 and 6.0 °C. This model fits better than the two previous ones (AIC = 277.13; adjusted $R^2 = 0.940$) (Fig. S1F). Finally, we considered a fit using a smoothing spline (Hastie, 1992). Among the three alternatives to the linear model, this model has several desirable properties – it fits the data best, with the lowest AIC and highest adjusted R^2 (although the differences are quite small) (Fig. S1G), and is quite smooth while still displaying the changes in slope explicitly represented by the segmented regression model (Fig. 2d). In addition, its associated *predict* function can be used to extract the local numeric first derivative of the fitted curve, which provides a representation of the HSE. This shows that HSE is ∼ 1 when ΔGMSST is < 0 °C and different from 1 when ΔGMSST is between ∼ 0 and 6 °C (red curve in Figs. 2d and S1G), again consistent with the ΔGMSST thresholds identified by the segmented and loess functions. Given these properties, the smoothing spline regression is our preferred choice for representing the relationship between ΔMOT and ΔGMSST in our synthesis of climate model results.

2.2 Constraints from proxy reconstructions

We next derive HSE from proxy reconstructions of ΔMOT, ΔBWT, and ΔGMSST. ΔMOT reconstructions spanning the past 0.7 Myr are based on Antarctic ice-core noble gases. We update Mg / Ca-based ΔBWT reconstructions with newer temperature calibrations (where necessary) and account for variations in the seawater Mg / Ca concentration ratio for long-term temperature reconstructions (> 1 Ma) that extend into the Early Pleistocene and the Pliocene (see Methods in Supplement). Lastly, we use a ΔGMSST reconstruction for the past 4.5 Myr that has negligible loss of variability from stacking of individual SST records with differing resolutions and age-model uncertainties (Clark et al., 2024). This reconstruction suggests that average (401 kyr) ΔGMSSTs decreased from ∼ 3 °C at 4 Ma to 0 °C at 1.5 Ma, falling in the range that climate models suggest should correspond to an HSE smaller than 1 (Figs. 2d, S1). Average cooling then accelerated between 1.5–0.9 Ma, corresponding to the MPT (Clark et al., 2024). ΔGMSSTs then remained largely (> 95 % of the time) below 0 °C over the remainder of the Pleistocene, falling in the range that climate models suggest should correspond to an HSE of ∼ 1. In the following, we use the 0–1 °C ΔGMSST threshold suggested from climate models to derive proxy-based HSE prior to and since the MPT.

2.2.1 Proxy-based HSE prior to the MPT

Figure 3 shows that most ΔBWT and ΔDOT reconstructions covering the interval from 1.5–4.5 Ma fall below ΔGMSSTs, thus identifying an HSE < 1. Differences among the ΔBWT and ΔDOT reconstructions, however, prevent a more accurate derivation of HSE needed to assess climate model predictions and decompose the $\delta^{18}O_b$ record into its $\delta^{18}O_T$ and $\delta^{18}O_{sw}$ components. To first order, these differences reflect the different approaches used for the reconstructions. For example, Waelbroeck et al. (2002) developed an approach that regresses independently known sea-level data (e.g., from corals) on $\delta^{18}O_b$ for the last glacial cycle and then used this regression to reconstruct sea level from $\delta^{18}O_b$ records for the last four glacial cycles. Sea level is converted to $\delta^{18}O_{sw}$ using a constant relation of $\Delta\delta^{18}O_{sw} : \Delta GMSL$. $\delta^{18}O_{sw}$ is then subtracted from the $\delta^{18}O$ record to derive $\delta^{18}O_T$, which is converted to ΔBWT using a constant relation of $\Delta\delta^{18}O_T : \Delta T$ (commonly $0.25‰\,°C^{-1}$). Following this same approach, Siddall et al. (2010) and Bates et al. (2014) extended the regression over the last two glacial cycles to reconstruct ΔDOT from $\delta^{18}O_b$ records over the last 5 Myr (see Fig. 3c for the past 4.5 Myr). Rohling et al. (2021, 2022) further extended the regression over the last 800 kyr using the LR04 $\delta^{18}O_b$ stack and a stack of sea-level records from Spratt and Lisiecki (2016) to reconstruct ΔDOT over the past 40 Myr (see Fig. 3a for the past 4.5 Myr). Rohling et al. (2021, 2022) also accounted for $\delta^{18}O$ variations in land ice ($\delta^{18}O_i$) over the last glacial cycle which were then applied to their sea-level record to derive $\delta^{18}O_{sw}$ from $\delta^{18}O_b$, but this did not include the effect of increasing temperatures on $\delta^{18}O_i$ prior to 0.8 Ma. In any event, applying this regression approach to derive $\delta^{18}O_{sw}$ will, by default, reproduce the variability of the $\delta^{18}O_b$ record, including the increase in the size of glaciations during the MPT, and thus potentially bias the ΔDOT reconstruction.

Another approach scales $\delta^{18}O_b$ to an inferred DOT LGM–Holocene range of −2 °C (Hansen et al., 2013; Westerhold et al., 2020; Hansen et al., 2023) and assumes stationarity in this scaling over the last 4.5 Ma (Fig. 3b). However, the ice-core noble-gas estimates suggest that the prescribed LGM cooling of −2° underestimates the MOT range by ∼ 70 % (Haeberli et al., 2021; Shackleton et al., 2023), and this scaling is likely nonstationary.

Given these and other uncertainties and differing assumptions used to derive ΔDOT from $\delta^{18}O_b$ records (Bates et al., 2014; Rohling et al., 2021; Hansen et al., 2013; Westerhold et al., 2020; Rohling et al., 2022; Hansen et al., 2023; Evans et al., 2024), we derive changes in HSE using Mg / Ca-based ΔBWT reconstructions that extend into the Early Pleistocene and the Pliocene (see Sect. S2 in Supplement). These include two low-resolution ΔBWT reconstructions, one from the equatorial Pacific (ODP site 806, 2500 m) (Fig. 3e) that samples Pacific Deep Water and one from the equatorial Atlantic (ODP site 926, 3500 m) that today is in the mixing zone between North Atlantic Deep Water (NADW) and Antarctic Bottom Water (AABW) (Lear et al., 2003) (Fig. 3f). We also use high-resolution Mg / Ca-based ΔBWT reconstructions from DSDP site 607 (3427 m) (Dwyer and Chandler, 2009; Sosdian and Rosenthal, 2009) and nearby IODP U1313 (3426 m) (Jakob et al., 2020) in the North Atlantic (Fig. 3i). Lastly, we use a smoothed BWT reconstruction based on a compilation of Mg / Ca records derived from six species or genera of benthic foraminifera (Cramer et al., 2011) (Fig. 3h). Although the smoothed reconstruction is based on Mg / Ca data from eight different sites, it is largely based on data from sites 806 and 926 for the last 4.5 Myr.

As discussed further in Sect. 4, temperature change of the ocean interior in response to surface forcing is nonuniform, with warming of the upper half of the ocean being stronger than GMSST warming as opposed to weaker in the lower half of the ocean (Bronselaer and Zanna, 2020; Fox-Kemper et al., 2021; Rugenstein et al., 2016; Zhu et al., 2024). To assess the relationship of ΔBWT at the Pacific and Atlantic core sites to ΔMOT, we use results from the Transient Climate Evolution (iTRACE) simulation performed with the isotope-enabled Community Earth System Model version 1.3 that has reproduced the observed evolution of global climate and water masses from 21 ka to the Early Holocene (11 ka) (Gu et al., 2020; He et al., 2021). The simulation is forced by changes in solar insolation from orbital changes (ORB), greenhouse gases (GHG), reconstructed ice sheets (ICE), and meltwater fluxes (MWF), with the latter causing millennial-scale variations in ocean heat storage and MOT through its effect on the Atlantic Meridional Overturning Circulation (AMOC) (Zhu et al., 2024).

Given that the Pacific Ocean constitutes ∼ 50 % of the total ocean volume, Lear et al. (2003) considered site 806 BWT to closely represent MOT. This inference is consistent with

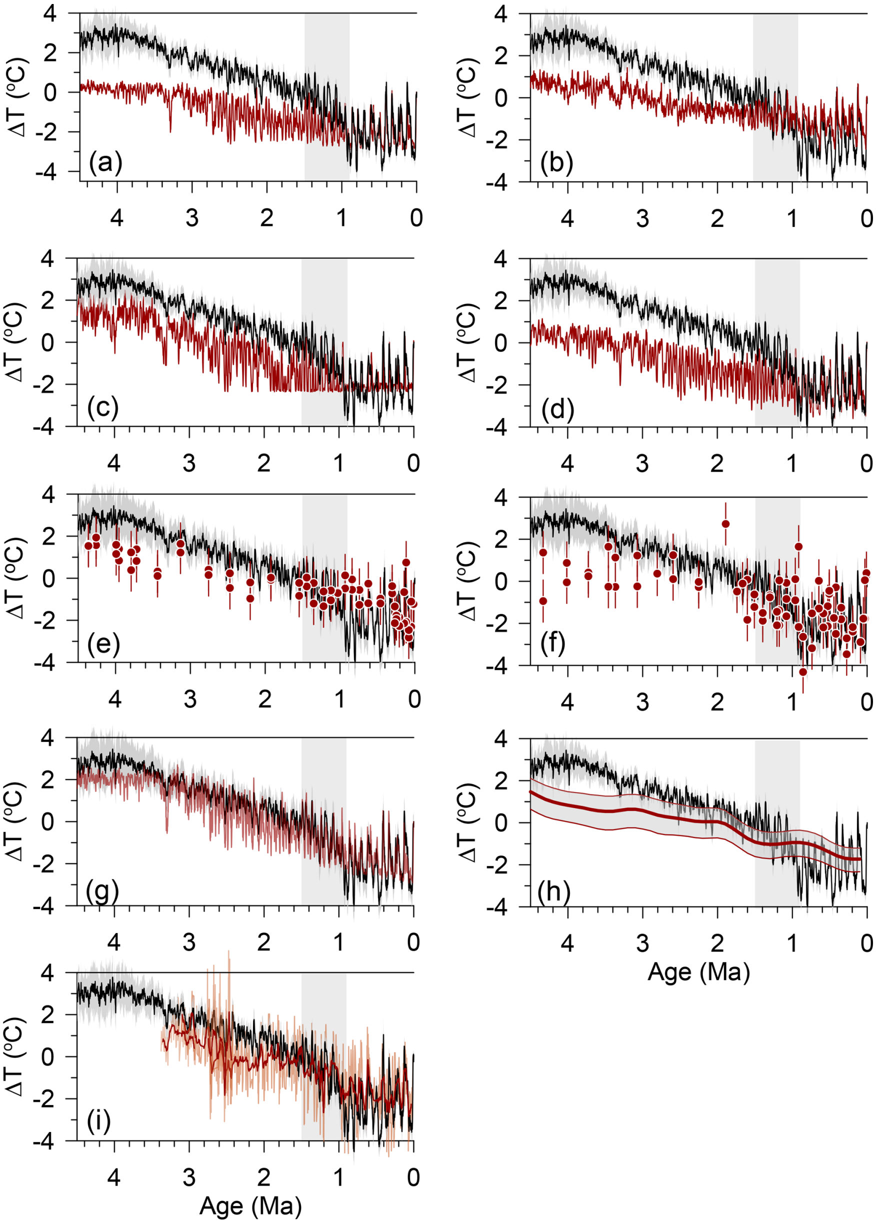

Figure 3. The reconstruction of global mean sea surface temperature change from preindustrial (Clark et al., 2024) (black with 1σ uncertainty in gray) compared to reconstructions of changes in bottom water temperature (ΔBWT, dark red). The vertical gray box represents the Middle Pleistocene Transition (0.9–1.5 Ma). **(a)** ΔBWT reconstruction from Rohling et al. (2022), referenced to preindustrial (PI). **(b)** ΔBWT reconstruction from Hansen et al. (2013). BWTs were mean-shifted by 1 °C so that they are referenced to the Holocene (i.e., 0 °C for the last 10 kyr). **(c)** ΔBWT reconstruction from Bates et al. (2014). BWTs were mean-shifted by 1.46 °C so that they are referenced to the Holocene (i.e., 0 °C for the last 10 kyr). **(d)** ΔMOT reconstruction from de Boer et al. (2014), referenced to PI. **(e)** Mg / Ca ΔBWT reconstruction from ODP site 806 (Lear et al., 2003). Mg / Ca BWTs were mean-shifted by −1.73 °C to fall within same range as our ΔMOT reconstruction for the last 800 kyr. **(f)** Mg / Ca ΔBWT reconstruction from ODP site 926 (Lear et al., 2003). Mg / Ca BWTs were mean-shifted by −1.73 °C so that they are referenced to the Holocene (i.e., 0 °C for the last 10 kyr). **(g)** Deep-ocean temperature reconstruction from Evans et al. (2024), which is the Rohling et al. (2022) reconstruction **(a)** with pH-corrected benthic δ^{18}O, referenced to PI. **(h)** The 2 Myr smoothed ΔBWT reconstruction from Cramer et al. (2011) using their Eq. (7b) with 90 % confidence interval. BWTs were mean-shifted by 2.1 °C so that they have the same long-term mean as our ΔMOT reconstruction over the last 400 kyr. **(i)** Mg / Ca ΔBWT from North Atlantic DSDP site 607 for > 2.9 Ma (Dwyer and Chandler, 2009) and < 2.8 Ma (Sosdian and Rosenthal, 2009) (11-point running average in dark red) and ΔBWT from North Atlantic IODP site U1313 for 2.4–2.75 Ma (Jakob et al., 2020). Site 607 data are mean-shifted by −1.73 °C so that they are referenced to the Early Holocene (i.e., 0 °C at 10 ka). Site U1313 data are referenced to modern BWT derived by Mg / Ca measurements on core-top samples.

iTRACE ΔBWTs at the location of site 806 that closely parallel ΔMOT throughout the simulation (Fig. 4a). In this regard, we note that the Cramer et al. (2011) reconstruction, being largely weighted by site 806 data for the 1–4.5 Ma interval but with a different calibration than used in Lear et al. (2003), also represents a close approximation of ΔMOT. On the other hand, sites 926, 607, and U1313 were largely bathed by NADW prior to ~ 1.5 Ma, after which NADW shoaled and was replaced by AABW at these sites during glaciations (Lisiecki, 2014). In both cases, the associated water masses account for a significantly smaller fraction of total ocean volume, and their BWTs may thus not reflect MOT. Despite being under a greater influence of millennial-scale changes in the AMOC, simulated ΔBWTs at Atlantic sites 926 and 607 warm as much as MOT by 11 ka (Fig. 4a). These similar changes in ΔBWT are further shown when combining the ΔBWT data from sites 806 and 926 (Lear et al., 2003), the smoothed ΔBWT reconstruction from Cramer et al. (2011) (comprised largely of data from sites 806 and 926), and the ΔBWT data from sites 607 (Dwyer and Chandler, 2009; Sosdian and Rosenthal, 2009) and U1313 (Jakob et al., 2020) (Fig. 5b). These results thus suggest that ΔBWTs at the Pacific and Atlantic sites agree and are closely monitoring ΔMOT, suggesting that we can combine their ΔBWT reconstructions to approximate ΔMOT.

We assess which HSE best agrees with the Mg / Ca ΔBWT data before 1.5 Ma by comparing the combined data from sites 806, 926, and 607 (Fig. 5a–c) and the smoothed reconstruction from Cramer et al. (2011) (based on sites 806 and 926 with a different calibration) (Fig. 5d–f) to three different scenarios of ΔMOT derived from our ΔGMSST reconstruction based on HSEs of 1, 0.5, and 0.1. The uncertainty on the ΔMOT-scenario reconstructions is calculated as the square root of the sum of squares of the uncertainty in the ΔGMSST reconstruction and in the HSE derived from models (Fig. 2). We attribute the total spread in the Mg / Ca data around their long-term trend to some combination of climate variability, analytical and calibration errors, and site location.

Based on their similar slopes, we find that the best agreement with the Mg / Ca-based ΔBWT data is for an HSE of ~ 0.5, with long-term rates of cooling between 1.5–4.5 Ma in both the data and our corresponding ΔMOT reconstruction being 0.6–0.8 °C Myr^{-1} (Figs. 5b, e, 6). HSE values of 0.7 and 0.3 closely encompass the uncertainty on our ΔMOT reconstruction based on an HSE $= 0.5$ (Figs. 5b, e, S2A), suggesting an HSE of 0.5 ± 0.2. Any further decrease in HSE would suggest virtually no change in MOT between 1.5–4.5 Ma, in contrast to the long-term cooling trend shown by all ΔBWT and ΔDOT reconstructions (Figs. 1, 3, 5b, e, 6).

2.2.2 Proxy-based HSE since the MPT

The low-resolution Antarctic ice-core noble-gas proxy records of ΔMOT for the past 0.7 Myr (Haeberli et al., 2021; Shackleton et al., 2023, 2020, 2021; Bereiter et al., 2018) provide the most robust measure for deriving HSE. Over this interval, there is good agreement ($R^2 = 0.75$) between the ice-core records of ΔMOT and ΔGMSST with an HSE of ~ 0.9, which, within uncertainties, agrees with an HSE of ~ 1 found in climate models for ΔGMSSTs < 0 °C (Fig. 7a, b, c). The temporal and amplitude changes in ΔBWT and ΔGMSST reconstructions extending back to 0.8 Ma also largely agree (Fig. 7e–k), with dominant orbital-scale variability. We note one exception in the ODP site 1123 reconstruction (Elderfield et al., 2012), where glacial ΔBWTs tend to be warmer than ΔGMSSTs over the past 0.8 Myr (Fig. 7d). We attribute these differences to site 1123 capturing a regional signal of glacial ΔBWT that differs from ΔMOT, although calibration uncertainties for this Mg / Ca record and changes in carbonate saturation, which are greatest at low temperatures, may also play a role. Otherwise, the general agreement between the other ΔBWT reconstructions and the ΔGMSST reconstruction further supports an HSE of ~ 1 over the last 0.8 Myr when ΔGMSSTs were < 0 °C, consistent with model results.

Multiple reconstructions of ΔGMSST and ΔMOT that span the Last Glacial Maximum (LGM, 26–18 ka) provide

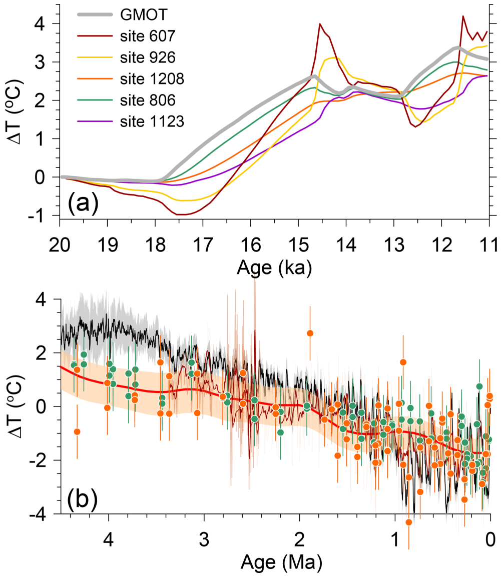

Figure 4. **(a)** Simulated changes in global mean ocean temperature (MOT) relative to 20 ka from the iTRACE experiment as well as for five sites with ΔBWT reconstructions (Zhu et al., 2024). **(b)** Global mean sea surface temperature change (black line, 1σ uncertainty) (Clark et al., 2024) compared to Mg / Ca-based ΔBWT reconstructions from Pacific ODP site 806 (green circles) (Lear et al., 2003), North Atlantic ODP site 926 (orange circles) (Lear et al., 2003), North Atlantic site 607 for > 2.9 Ma (Dwyer and Chandler, 2009) and < 2.8 Ma (Sosdian and Rosenthal, 2009) and site U1313 for 2.4–2.75 Ma (Jakob et al., 2020) (9-, 11-, and 21-point running averages, respectively, in dark red), and smoothed reconstruction from Cramer et al. (2011) using their Eq. (7b) with 90 % confidence interval.

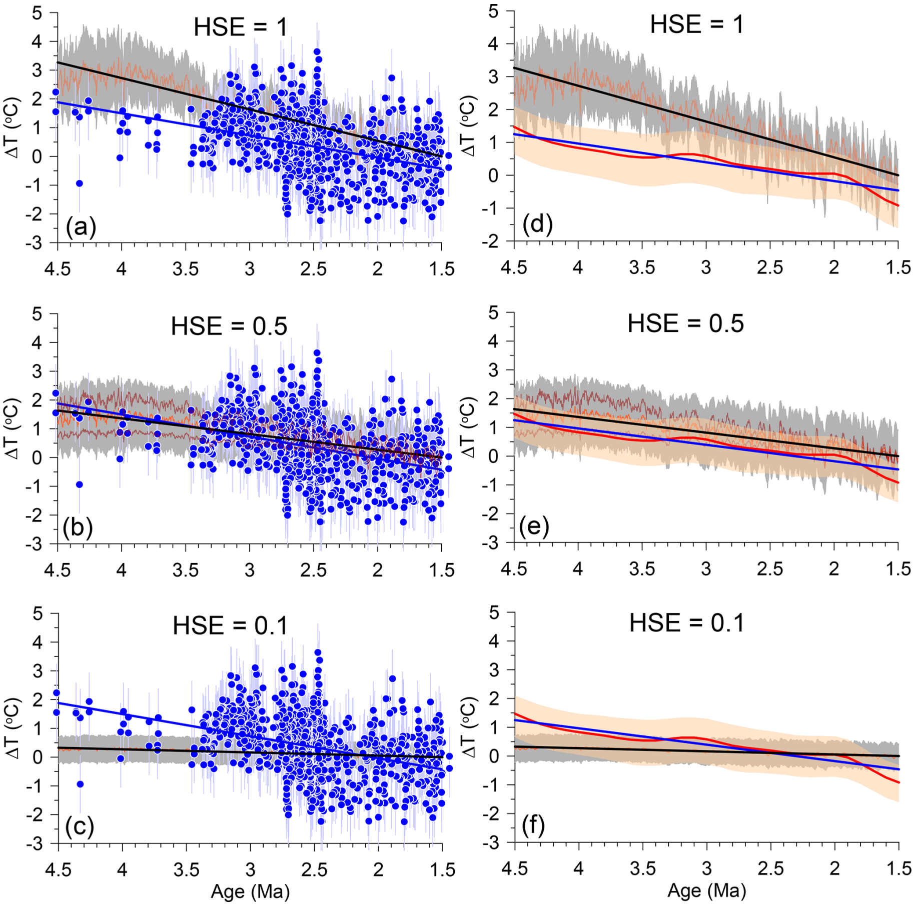

Figure 5. **(a–c)** Our reconstructed global mean ocean temperature anomaly (ΔMOT) for 1.5–4.5 Ma based on three different HSEs (orange lines) compared to proxy-based (Mg / Ca) reconstructions of bottom water temperature anomalies (blue) (1 °C uncertainty) (Sosdian and Rosenthal, 2009; Lear et al., 2003). The 1σ uncertainty on ΔMOT reconstructions (gray fill) derived by combining in quadrature 1σ uncertainties on our ΔSST reconstruction and on the relationship between ΔGMSST and ΔMOT (Fig. 2a). Regression lines plotted through ΔMOT reconstructions (black) and Mg / Ca data (blue). **(a)** Our reconstructed ΔMOT based on HSE = 1. **(b)** Our reconstructed ΔMOT based on HSE = 0.5. Also shown are ΔMOT reconstructions based on HSE = 0.7 and HSE = 0.3 (brown lines), which closely encompass the 1σ uncertainty on our ΔMOT reconstruction based on HSE = 0.5. **(c)** Our reconstructed ΔMOT based on HSE = 0.1 (falls under the corresponding regression line). The decrease in uncertainties on our ΔMOT reconstructions reflects the decrease in uncertainty on our scaled ΔGMSST reconstructions. **(d–f)** Our reconstructed global mean ocean temperature anomaly (ΔMOT) for 1.5–4.5 Ma based on three different HSEs (orange lines) compared to the 2 Myr smoothed ΔBWT reconstruction from Cramer et al. (2011) using their Eq. (7b) (red line) with 90 % confidence interval (orange shading). Their BWTs were mean-shifted by 2.1 °C so that they are the same as the long-term (401 kyr running average) mean as our ΔMOT reconstruction over the last 400 kyr. **(d)** Our reconstructed ΔMOT based on HSE = 1. **(e)** Our reconstructed ΔMOT based on HSE = 0.5. Also shown are ΔMOT reconstructions based on HSE = 0.7 and HSE = 0.3 (brown lines), which closely encompass the 1σ uncertainty on our ΔMOT reconstruction based on HSE = 0.5. **(f)** Our reconstructed ΔMOT based on HSE = 0.1 (falls under the corresponding regression line).

further constraints on HSE at that time (Fig. 7l). There is good agreement among ΔGMSST reconstructions (referenced to PI), with proxy-based estimates of −3.0 ± 0.1 °C (18–22 ka average) (Shakun et al., 2012), −2.9 °C (−3.0 to −2.7 °C, 95 % confidence interval (CI) (23–19 ka average) (Tierney et al., 2020), and −3.3 ± 0.4 °C (26–18 ka average) (Clark et al., 2024), as well as an estimate from a global climate model with data assimilation of −3.1 °C (−3.4 to −2.9 °C, 95 % CI) (23–19 ka average) (Tierney et al., 2020), with these reconstructions suggesting an average LGM ΔGMSST of −3.0 ± 0.2 °C.

There is also good agreement among existing ΔMOT and ΔDOT reconstructions for the LGM (Fig. 7l). The ΔDOT reconstructions by Hansen et al. (2013, 2023) are based on an inferred LGM cooling of −2° and are thus not considered here. Shakun et al. (2015) found an LGM average cooling of −2.63 °C, and Rohling et al. (2022) (updated from Rohling et al., 2021) reconstructed an LGM average cooling of −2.82 ± 0.17 °C. Shackleton et al. (2023) standardized

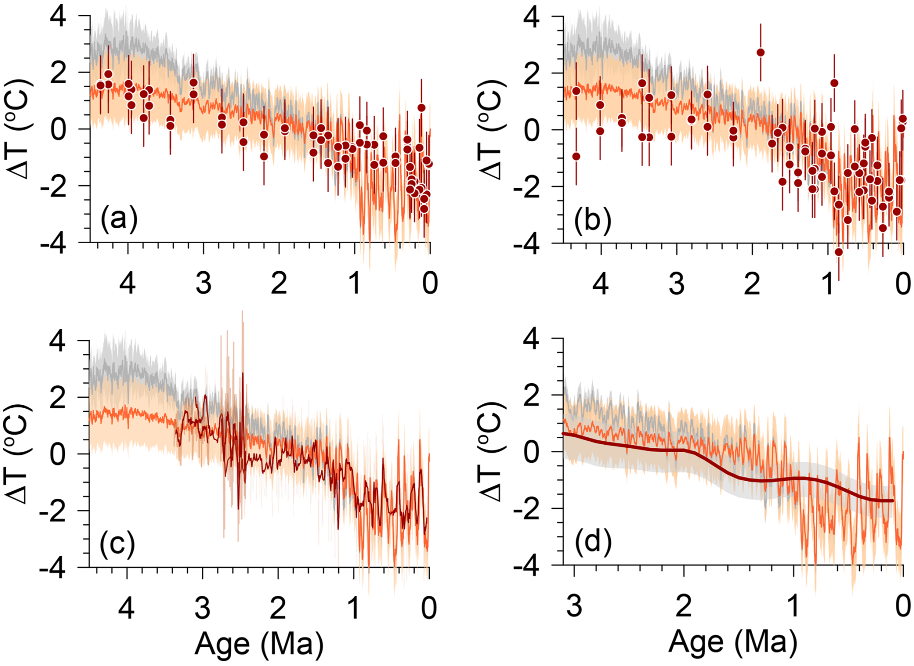

Figure 6. Our reconstructed change in mean ocean temperature (ΔMOT) based on HSE = 0.5 before 1.5 Ma (1σ uncertainty) and HSE = 1 after 0.9 Ma with a linear increase in HSE from 0.5 to 1 between 1.5 and 0.9 Ma (orange) compared to reconstructions of bottom water temperature anomalies (ΔBWT) from Mg / Ca proxies. Also shown is ΔMOT based on HSE = 1 (gray) for the past 4.5 Ma. **(a)** ΔBWT from equatorial Pacific ODP site 806 (Lear et al., 2003). Mg / Ca BWTs were mean-shifted by −1.73 °C to fall within same range as our ΔMOT reconstruction for the last 800 kyr. **(b)** ΔBWT from equatorial Atlantic ODP site 926 (Lear et al., 2003). Mg / Ca BWTs were mean-shifted by −1.73 °C so that they are referenced to the Holocene (i.e., 0 °C for the last 10 kyr). **(c)** ΔBWT from North Atlantic DSDP site 607 for > 2.9 Ma (Dwyer and Chandler, 2009) and < 2.8 Ma (Sosdian and Rosenthal, 2009) (11-point running average in dark red) and ΔBWT from North Atlantic IODP site U1313 for 2.4–2.75 Ma (Jakob et al., 2020). Site 607 data are mean-shifted by 2.71 °C so that they are referenced to the Early Holocene (i.e., 0 °C at 10 ka). Site U1313 data are referenced to modern BWT derived by Mg / Ca measurements on core-top sample. **(d)** Smoothed ΔBWT reconstruction from Cramer et al. (2011) (dark-red line) using their Eq. (7b) with 90 % confidence interval (gray shading). BWTs were mean-shifted by 2.1 °C so that they have the same long-term mean as our ΔMOT reconstruction over the last 400 kyr. The 1σ uncertainty on our ΔMOT reconstructions is derived by combining in quadrature 1σ uncertainties on the ΔGMSST reconstruction and on the relationship between ΔGMSST and ΔMOT (Fig. 2a), which closely corresponds to a range in HSE between 0.7 and 0.3 (Fig. 6b, e).

previously published ice-core noble-gas ΔMOT reconstructions (Bereiter et al., 2018; Baggenstos et al., 2019; Shackleton et al., 2019, 2020) and reported a data-based LGM average ΔMOT of -2.76 ± 0.27 °C and a spline average of -2.83 ± 0.5 °C. We also note that a ΔDOT reconstruction for 0–20 ka derived by subtracting global $\delta^{18}O_{sw}$ from $\delta^{18}O_b$ records (Zhu et al., 2024) is in good agreement with the spline ΔMOT reconstruction (Shackleton et al., 2023), including during the period of LGM overlap (18–20 ka) (Fig. 7l). Taking the average LGM ΔMOT (-2.76 ± 0.31 °C) with the average LGM ΔGMSST (-3.0 ± 0.2 °C) suggests an LGM HSE of 0.9, or the same as when using the ice-core noble-gas ΔMOT reconstructions for the last 0.7 Myr (Fig. 7c) and, within uncertainty, with an HSE of ∼ 1 suggested by climate models (Fig. 2).

Recent work, however, has argued that changes in noble-gas saturation states of the deep ocean during the LGM may have resulted in a cold bias in the ice-core data equivalent to -0.38 ± 0.37 °C (Seltzer et al., 2024) to -0.50 ± 0.67 °C (Pöppelmeier et al., 2023). Despite these similar results, Seltzer et al. (2024) noted that Pöppelmeier et al. (2023) were unable to reconcile their result from simulated air–sea gas exchange with the ice-core data and thus based their MOT cooling on a climate model. The cold bias identified by Seltzer et al. (2024) is based on an ensemble of five PMIP3 climate models that suggest air–sea disequilibria due to stronger high-latitude winds, from which they argued that LGM ΔMOT was -2.27 ± 0.46 °C, corresponding to an LGM HSE of ∼ 0.8 when using an average LGM ΔGMSST of -3.0 ± 0.2 °C. Seltzer et al. (2024) did not elaborate on how this effect may have differed at other times, but we can assume that high-latitude winds would have weakened from their LGM maxima, resulting in a diminished effect on air–sea disequilibria at those times. As a first-order approximation of how this would be expressed during other glacial maxima, we decreased all glacial maxima ΔMOTs in the 0.7 Myr

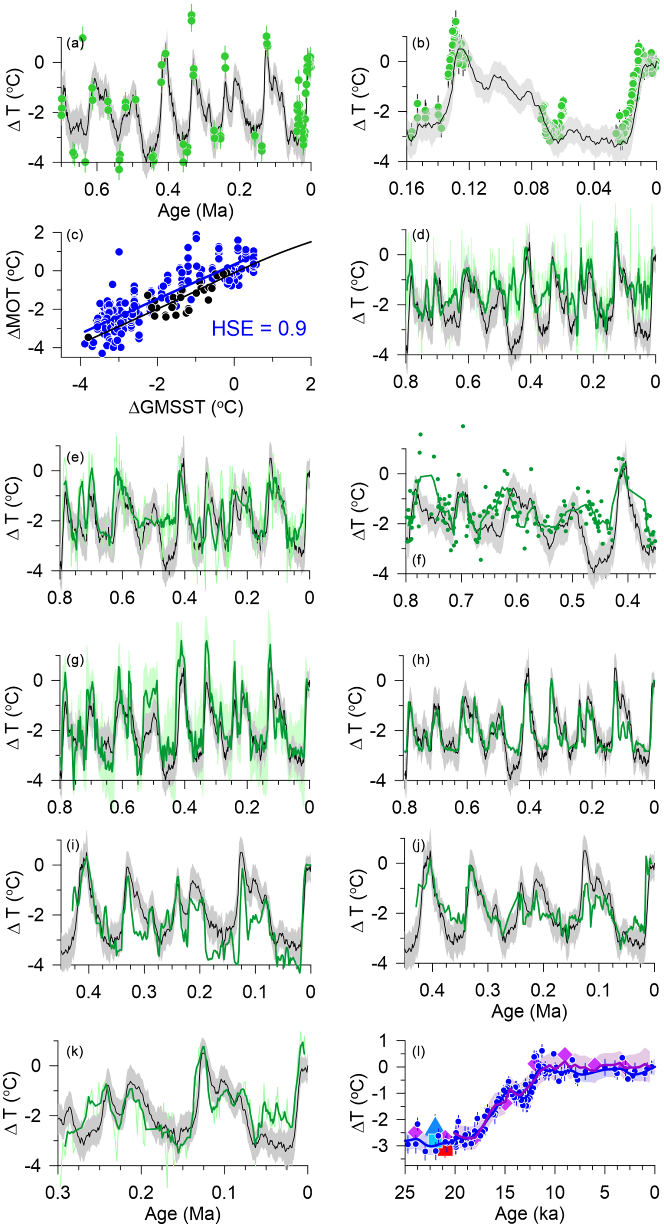

Figure 7. Comparison of the reconstruction of global mean sea surface temperature change (Clark et al., 2024) (black) to proxy reconstructions of bottom water and mean ocean temperature change from preindustrial (green). **(a)** ΔMOT reconstruction from Haeberli et al. (2021). **(b)** ΔMOT reconstructions from Shackleton et al. (2020, 2021, 2023). **(c)** Blue: scatter plot of ΔMOT derived from ice-core noble gases (Bereiter et al., 2018; Haeberli et al., 2021; Shackleton et al., 2019, 2020, 2021) versus the reconstruction of global mean sea surface temperature change from preindustrial for the last 0.7 Myr (Clark et al., 2024). Linear regression through the data identifies an HSE of 0.9. Also shown are climate model data (black circles) for ΔGMSST < 0 °C with smoothing spline fit (black line) (see Fig. 2). **(d)** Mg / Ca-based ΔBWT reconstruction from site 1123 for the last 0.8 Ma (Elderfield et al., 2012) (11-point running average shown by dark-green line). **(e)** Mg / Ca-based ΔBWT reconstruction modified from Sosdian and Rosenthal (2009) (5-point running average shown by dark-green line). **(f)** Mg / Ca-based ΔBWT reconstruction from ODP site 1208 from Ford and Raymo (2020) (5-point running average shown by green line). **(g)** ΔBWT reconstruction from Shakun et al. (2015). **(h)** ΔBWT reconstruction from Rohling et al. (2022). **(i)** ΔBWT reconstruction from ODP site 980 from Waelbroeck et al. (2002). **(j)** ΔBWT reconstruction from site 94-10 from Waelbroeck et al. (2002). **(k)** Mg / Ca-based ΔBWT reconstruction from Martin et al. (2002) (5-point running average shown by dark-green line). **(l)** Reconstructed changes in ΔGMSST, ΔMOT, and ΔDOT since the Last Glacial Maximum. Blue symbols with 1σ uncertainty are ΔMOT derived from ice-core noble gases, with spline fit to these data with 1σ uncertainty shown by the blue curve and 1σ confidence envelope (Shackleton et al., 2023). Purple diamonds are ΔDOT from Shakun et al. (2015). Light-blue square is ΔDOT from Rohling et al. (2022). Thick purple line with 1σ confidence envelope is ΔDOT from Zhu et al. (2024). Light-blue triangle with 1σ uncertainty is ΔMOT from Seltzer et al. (2024). Red square with 1σ uncertainty is ΔGMSST from Tierney et al. (2020).

ice-core noble-gas reconstruction (Haeberli et al., 2021) by 0.38 °C, resulting in a decrease in HSE from 0.9 to 0.85.

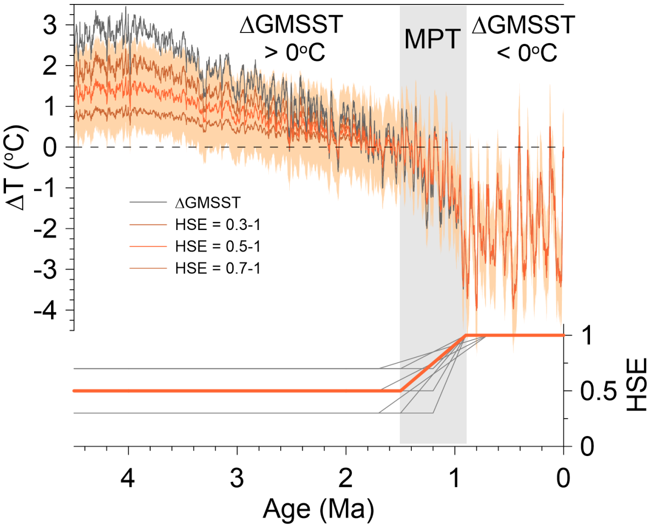

Figure 8. Our reconstructed ΔMOT (orange line with 1 σ uncertainty) that linearly increases from 0.5 to 1 across the MPT. Also shown is ΔGMSST (black line) from (Clark et al., 2024) and ΔMOTs based on increases from 0.3 to 1 and 0.7 to 1 across the MPT (brown lines). Lower panel shows different scenarios (magnitude and duration) for the increase in HSE assessed in Fig. S2, with our preferred scenario shown by the orange line.

We thus conclude that the potential LGM cold bias in ΔMOT does not substantially change the evidence that HSE was ∼ 1 within the uncertainties of the data. At the same time, Seltzer et al. (2024) recognized that current understanding of changes in LGM high-latitude wind speed (and thus air–sea disequilibria) in both observations and models is highly uncertain. For example, more-recent modeling suggests that high-latitude Southern Hemisphere wind speeds during the LGM weakened by 15 % (Zhu et al., 2021) to 25 % (Gray et al., 2023) as opposed to the 11 ± 23 % increase in the PMIP3 model average (largely driven by one model). As Seltzer et al. (2024) note, "Future improvements in our understanding of high-latitude winds in the LGM will help to substantially reduce uncertainties in LGM MOT". In the meantime, the good agreement between independent estimates of ΔDOT from marine proxy records (Shakun et al., 2015; Rohling et al., 2022; Zhu et al., 2024) with the ice-core ΔMOT estimate that neglects air–sea disequilibrium (Shackleton et al., 2023) (Fig. 2c) suggests a negligible effect of air–sea disequilibrium on the ice-core data.

2.3 Summary of changes in MOT and HSE over the past 4.5 million years

In summary, existing constraints from Mg / Ca-based ΔBWT data and ice-core ΔMOT data show that HSE was ∼ 0.5 ± 0.2 for > 1.5 Ma when ΔGMSSTs were > 0 °C and ∼ 1 for the past 0.8 Myr when ΔGMSSTs were < 0 °C (Fig. 8), consistent with climate model results (Fig. 2). These constraints thus suggest that the increase occurred as part of the large-scale changes in ocean circulation during the MPT (0.9–1.5 Ma) (Lisiecki, 2014; Lang et al., 2016; Pena and Goldstein, 2014; Hodell and Venz-Curtis, 2006). Although the data do not identify the exact function by which HSE increased during this time, we make the simplifying assumption that it increased linearly from 0.5 to 1 during the MPT (Fig. 8). As previously noted, HSE increases of 0.3 to 1 and 0.7 to 1 result in ΔMOTs that fall within the 1σ uncertainty of ΔMOTs based on an increase of 0.5 to 1 (Figs. 5, 8, S2). We also find that our ΔMOT reconstruction is insensitive to a shorter (0.9–1.2 Ma) or longer (0.7–1.7 Ma) interval over which HSE increased from 0.5 to 1 (Fig. S2C).

3 Decomposition of the benthic δ^{18}O record

3.1 Derivation of $\delta^{18}O_{sw}$ supports HSE < 1 before MPT

Using our ΔMOT reconstruction to decompose the $\delta^{18}O_b$ record into its temperature ($\delta^{18}O_T$) and seawater ($\delta^{18}O_{sw}$) components provides additional support for the need to decrease HSE to < 1 prior to 0.9 Ma as identified from the proxy ΔBWT data. We use the probabilistic global $\delta^{18}O_b$ stack (Prob-stack) from Ahn et al. (2017), which, compared to the previous LR04 stack (Lisiecki and Raymo, 2005), includes a larger number of records (total = 180) that span a larger depth range of the ocean (500–4500 m) as well as uncertainties in the alignment of these records. Stacking individual $\delta^{18}O_b$ records that span such a geographic and depth range significantly increases the signal-to-noise ratio for global changes in $\delta^{18}O_b$ and its $\delta^{18}O_T$ and $\delta^{18}O_{sw}$ components while minimizing local hydrographic changes that may contribute to a single $\delta^{18}O_b$ record. We convert our ΔMOT reconstruction (Fig. 8) to $\delta^{18}O_T$ using a relation of 0.25 ‰ °C^{-1}, which is appropriate for cold deepwater temperatures (Kim and O'Neil, 1997; Marchitto et al., 2014), and subtract this from the Prob-stack to derive $\delta^{18}O_{sw}$. We estimate the 1σ uncertainty on $\delta^{18}O_{sw}$ from the square root of sum of squares based on the uncertainties in our ΔMOT reconstruction, the HSE from models, and the probabilistic $\delta^{18}O_b$ stack.

Figure 9a shows that when using an HSE of 1 for the last 4.5 Myr, Early Pleistocene and Late Pliocene $\delta^{18}O_{sw}$ values are significantly more positive than the Mg / Ca-based reconstruction from North Atlantic site 607 (Sosdian and Rosenthal, 2009), which is the longest (0–3.2 Ma), orbitally resolved record available. This is to be expected since the Mg / Ca-based ΔBWT reconstruction was one of several used to constrain the reduction in HSE. However, positive $\delta^{18}O_{sw}$ values extend back to 4.5 Ma, with Pliocene interglacial values that are 0.3 ‰ to 0.5 ‰ more positive than average interglacial values over the last 0.8 Myr despite robust evidence for higher-than-present Pliocene sea levels that require values below 0 ‰ (Miller et al., 2012; Raymo et al., 2018; Dumitru et al., 2019; Winnick and Caves, 2015). This suggests that too much of the $\delta^{18}O_b$ signal is being removed by the $\delta^{18}O_T$ component using an HSE of 1.

Applying our ΔMOT reconstruction (Fig. 8) results in $\delta^{18}O_{sw}$ values that are more comparable to Pliocene sea-level reconstructions, with average Pliocene interglacial $\delta^{18}O_{sw}$ values decreasing to 0 ‰ to 0.1 ‰ (Fig. 9b), thus supporting the need for the decrease in HSE prior to 1.5 Ma suggested by proxy data. However, Early Pleistocene and Pliocene $\delta^{18}O_{sw}$ values continue to be more positive than data constraints (Dwyer and Chandler, 2009; Dumitru et al., 2019; Miller et al., 2012; Jakob et al., 2020; Sosdian and Rosenthal, 2009). While further decreases in HSE result in more negative Pliocene $\delta^{18}O_{sw}$ values (Fig. S2B), it would have to decrease to ∼ 0.1 to achieve average Pliocene $\delta^{18}O_{sw}$ values of −0.2 ‰ (not shown). Although this could explain higher sea levels at that time (Raymo et al., 2018; Winnick and Caves, 2015; Dumitru et al., 2019), such a low HSE is ruled out by proxy BWT reconstructions (Figs. 4, 5, 6).

3.2 Other potential factors influencing the long-term $\delta^{18}O_b$ trend

Recent studies have identified offsets between Cenozoic ocean temperatures inferred from the $\delta^{18}O_b$ record as compared to proxy-based temperatures, notably those based on clumped isotope thermometry, and considered possible effects on $\delta^{18}O_b$ such as from diagenesis, carbonate ion changes, or changes in ocean pH; issues with the calibration of the proxy itself may also be involved (Meckler et al., 2022; Rohling et al., 2024). Raymo et al. (2018) suggested that the ∼ 0.3 ‰ decrease in Pliocene $\delta^{18}O_b$ values relative to Late Holocene values is too small to accommodate both the higher sea levels and warmer ocean temperatures inferred for this time. They addressed this discrepancy by proposing that foraminifera tests that recrystallized in pore waters that were colder than those in which they were originally buried would cause precipitation of abiotic calcite with heavier $\delta^{18}O$ values (Schrag, 1999), resulting in 3 Myr benthic foraminifera tests being ∼ 0.25 ‰ heavier than tests with no diagenesis. Decreasing Pliocene $\delta^{18}O_b$ values in the Prob-stack by an additional 0.25 ‰ can then more readily explain the evidence for higher sea levels and warmer ocean temperatures. Applying this diagenetic process using our reconstructed MOT cooling over the past 4.5 Myr suggests that the diagenetic impact is subtle but could account for a shift of 0.2 ‰ to 0.4 ‰ over the 4.5 Myr record, with averaging of the $\delta^{18}O_b$ records in the Prob-stack integrating diagenetic effects to produce a near-constant long-term trend.

We cannot determine exactly how much the effect of diagenesis may have contributed to an increase in $\delta^{18}O_b$ over the least 4.5 Myr, but our assessment suggests that it is sufficient to justify a long-term increase as proposed by Raymo et al. (2018). Given our assessed range, we considered three

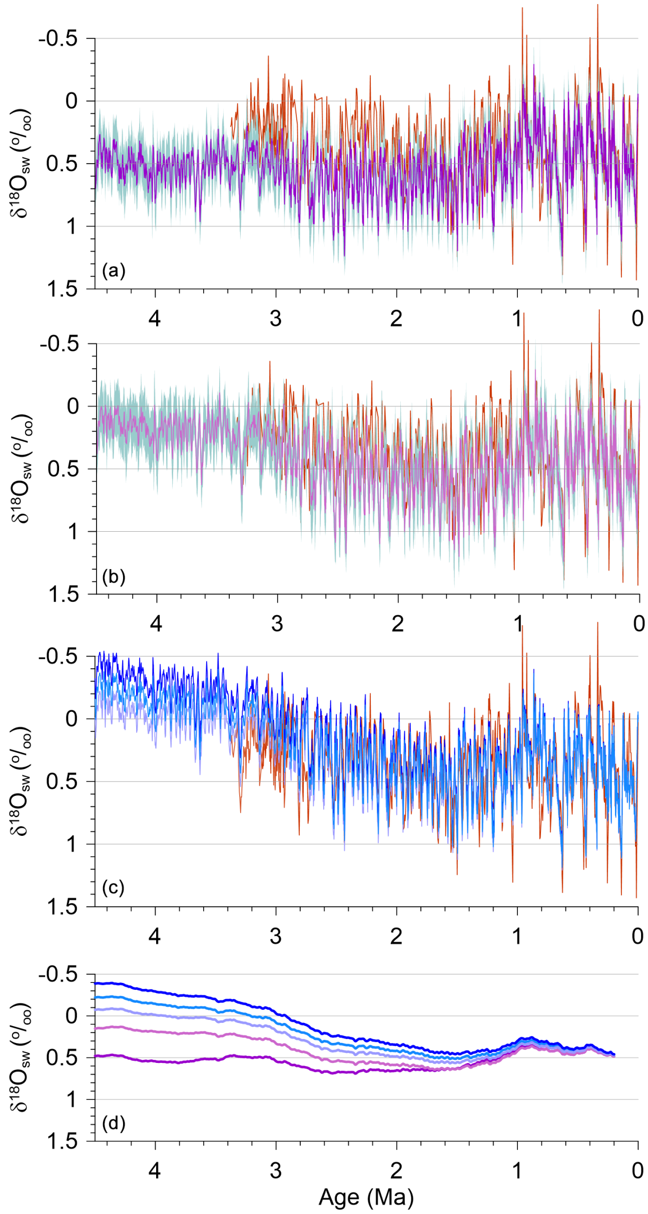

Figure 9. **(a)** Our reconstructed $\delta^{18}O$ of seawater ($\delta^{18}O_{sw}$) (1σ uncertainty) for 0–4.5 Ma based on HSE = 1 (purple) compared to proxy-based (Mg / Ca) reconstruction of $\delta^{18}O_{sw}$ (red) from North Atlantic site 607 (Sosdian and Rosenthal, 2009; Dwyer and Chandler, 2009). **(b)** Our reconstructed $\delta^{18}O$ of seawater ($\delta^{18}O_{sw}$) (1σ uncertainty) for 0–4.5 Ma based on HSE = 0.5 before 1.5 Ma (violet) compared to proxy-based (Mg / Ca) reconstruction of $\delta^{18}O_{sw}$ (red) from North Atlantic site 607 (Sosdian and Rosenthal, 2009; Dwyer and Chandler, 2009). **(c)** Our reconstructed $\delta^{18}O$ of seawater ($\delta^{18}O_{sw}$) (1σ uncertainty) for 0–4.5 Ma based on HSE = 0.5 before 1.5 Ma and removal of long-term trends of 0.05 ‰ Myr^{-1} ($\delta^{18}O_{sw}$-LO, light blue), 0.083 ‰ Myr^{-1} ($\delta^{18}O_{sw}$-INT, medium blue), and 0.12 ‰ Myr^{-1} ($\delta^{18}O_{sw}$-HI, dark blue) compared to proxy-based (Mg / Ca) reconstruction of $\delta^{18}O_{sw}$ (red) from North Atlantic site 607 (Sosdian and Rosenthal, 2009; Dwyer and Chandler, 2009). **(d)** Long-term running averages of the five $\delta^{18}O_{sw}$ scenarios shown in panels **(a)**–**(c)**, color coded in the same way.

scenarios that result in a diagenetic increase in $\delta^{18}O_b$ from today of 0.15 ‰, 0.25 ‰, and 0.35 ‰ at 3 Ma, corresponding to a long-term increase of 0.05 ‰ Myr^{-1} ($\delta^{18}O_{sw}$-LO), 0.083 ‰ Myr^{-1} ($\delta^{18}O_{sw}$-INT), and 0.12 ‰ Myr^{-1} ($\delta^{18}O_{sw}$-HI), respectively. In removing the long-term increase for each of these scenarios from the $\delta^{18}O_b$ Prob-stack, $\delta^{18}O_{sw}$ values based on our ΔMOT reconstruction agree with reconstructed Late Pliocene and Early Pleistocene $\delta^{18}O_{sw}$ values from site 607 (Fig. 9c). Although this agreement hinges on just one $\delta^{18}O_{sw}$ record that may also have experienced diagenesis, we previously showed that this site closely monitors MOT (Fig. 4). Moreover, not accounting for these combined effects (change in HSE, diagenesis) would result in long-term average $\delta^{18}O_{sw}$ values being comparable to Late Pleistocene average values throughout the last 4.5 Myr (Fig. 9d). Finally, our reconstructed long-term Pliocene $\delta^{18}O_{sw}$ values of −0.1 ‰ to −0.4 ‰ (Fig. 9d) are consistent with sea-level highstands of 20–25 m above present (Dumitru et al., 2019) even when accounting for land ice having more positive $\delta^{18}O$ values under warmer surface temperatures (Winnick and Caves, 2015), with the remaining ∼ 0.4 ‰ decrease in the $\delta^{18}O_b$ record relative to the Late Pleistocene consistent with the ∼ 1.5 °C ΔMOT in our reconstruction (Fig. 8). We also evaluated a scenario of having the same magnitude of diagenetic correction (0.375 ‰) over the past 4.5 Myr but with the increase tracking our ΔMOT reconstruction. Figure S3 shows that the faster rate of change across the MPT slightly increases the trend in $\delta^{18}O_{sw}$ over this interval (more depleted before 1 Ma, more enriched after 1 Ma), but the changes are insignificant (≤ 0.1 ‰).

Another potential effect that may be comparable to diagenesis is the impact of changing carbonate ion concentration in seawater on the $\delta^{18}O$ of foraminifera shells (Spero et al., 1997). Due to higher concentrations of atmospheric CO_2 (Köhler, 2023), the 100 kyr mean carbonate ion concentration in Pliocene surface seawater may have been lower than during the Late Pleistocene by on the order of 20 to 50 µmol kg^{-1}. Laboratory experiments (Bijma et al., 1999) and theoretical studies (Zeebe, 1999) on planktic species suggest a species-specific effect which would cause the $\delta^{18}O$ of foraminifera shells in the Early Pliocene to increase by 0.1 ‰ to 0.3 ‰ relative to the Late Pleistocene. Although no laboratory studies on benthic foraminifera species have been conducted, Marchitto et al. (2014) speculated that the $\delta^{18}O$ of some Late Holocene benthic species may have been influenced by pH. On the other hand, a recent study compiling 160 kyr of data from two widely abundant planktic foraminifera species in wider tropical surface waters extracted from 127 sediment cores could not confirm the carbonate ion effect as found in the laboratory (Köhler and Mulitza, 2024). Furthermore, a recent analysis of Late Holocene data from the benthic foraminifera species *Cibicidoides* spp. found that only about a third of the variance in $\delta^{18}O$ can be explained by carbonate chemistry (Nederbragt, 2023). Current understanding of the carbonate ion effect thus remains uncertain but suggests that it may have contributed to the $\delta^{18}O$ corrections in addition to diagenesis. Regardless of the specific combination of potential effects, our results clearly identify the need to account for a long-term increase in $\delta^{18}O_b$ when deriving $\delta^{18}O_{sw}$ and aligning ocean temperature records (Rohling et al., 2024; Meckler et al., 2022).

3.3 Temperature and $\delta^{18}O_{sw}$ controls on the Prob-stack $\delta^{18}O_b$ record

Our reconstruction shows that $\delta^{18}O_{sw}$ increased between 3.0 and 2.5 Ma in its (glacial) maxima to values that, on average, are similar to LGM values. These high glacial values in $\delta^{18}O_{sw}$ persisted throughout much of the Pleistocene (Fig. 10b). The main factor that is modulating the expression of this Late Pliocene/Early Pleistocene $\delta^{18}O_{sw}$ increase in the Prob-stack $\delta^{18}O_b$ is the gradual decrease in long-term average MOT (and thus increase in $\delta^{18}O_T$) relative to the increase in the rate of change in long-term average $\delta^{18}O_b$ values between 3.0–2.5 Ma (Fig. 10b). In other words, since $\delta^{18}O_T$ is only decreasing gradually, the relatively rapid increase in $\delta^{18}O_b$ between 3.0–2.5 Ma must be due to a substantial increase in $\delta^{18}O_{sw}$. After 2.5 Ma, both average $\delta^{18}O_T$ and $\delta^{18}O_{sw}$ values increase at similar rates until the onset of the MPT at 1.5 Ma. The second rapid increase in $\delta^{18}O_b$ during the MPT would then be due to the rapid increase in $\delta^{18}O_T$, in this case resulting in a slight decrease in average $\delta^{18}O_{sw}$ values (Fig. 10b). A subsequent paper will use our $\delta^{18}O_{sw}$-LO, $\delta^{18}O_{sw}$-INT, and $\delta^{18}O_{sw}$-HI reconstructions to derive sea level over the last 4.5 Myr that accounts for land ice having more positive $\delta^{18}O$ values under warmer surface temperatures (Winnick and Caves, 2015; Gasson et al., 2016).

3.4 Assessment of our $\delta^{18}O_{sw}$-INT reconstruction

We assess our $\delta^{18}O_{sw}$-INT reconstruction by comparing them with other $\delta^{18}O_{sw}$ reconstructions that have been derived from two methods. One method is directly comparable to ours in having used an independent reconstruction of BWT which is then subtracted as $\delta^{18}O_T$ from the $\delta^{18}O_b$ record to derive $\delta^{18}O_{sw}$ (Shakun et al., 2015; Sosdian and Rosenthal, 2009; Elderfield et al., 2012; Woodard et al., 2014; Ford and Raymo, 2020; Miller et al., 2020). Other than the Shakun et al. (2015) and Miller et al. (2020) reconstructions, these are based on local records and are thus subject to local temperature and hydrographic effects, and the small number of high-resolution records prevents development of a robust global stack. As summarized in Sect. 2.2.1, the other method takes the opposite approach of ours by first reconstructing sea level and then subtracting it as $\delta^{18}O_{sw}$ from the $\delta^{18}O_b$ record to derive ΔBWT or ΔDOT.

Figure 11 compares our $\delta^{18}O_{sw}$-INT reconstruction with published reconstructions for the last 0.8 Myr, with Fig. 11a–e comparing reconstructions derived from existing BWT re-

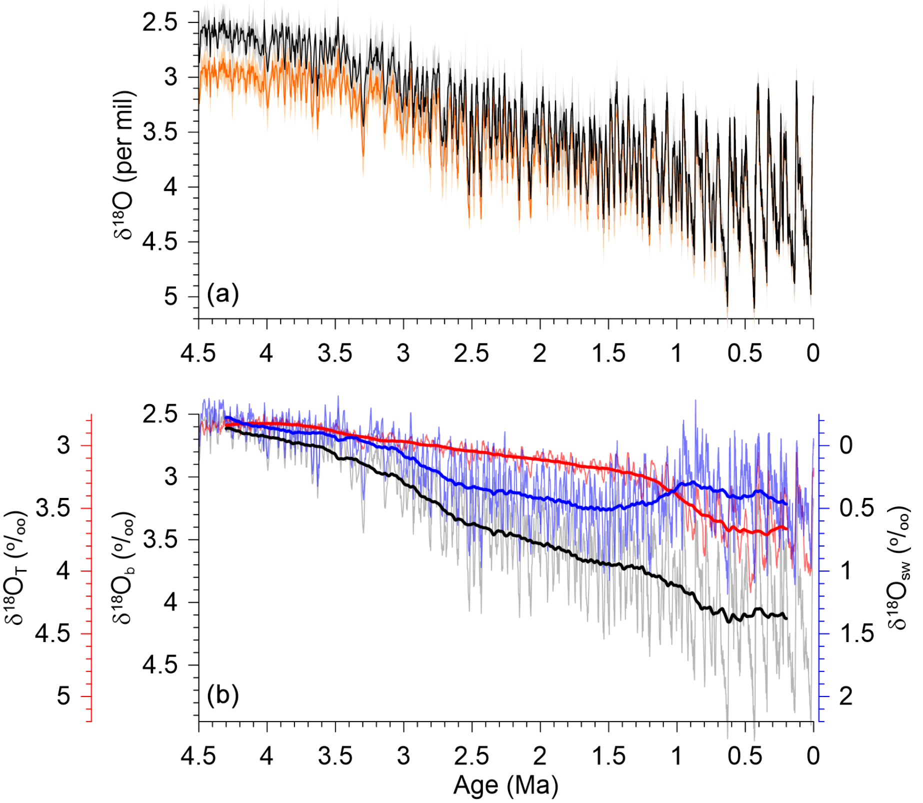

Figure 10. **(a)** Comparison of the Prob-stack $\delta^{18}O_b$ record (orange line with 1σ uncertainty) (Ahn et al., 2017) to the Prob-stack $\delta^{18}O_b$ record with the removal of a long-term increase of 0.083 ‰ Myr^{-1} (black line with 1σ uncertainty). **(b)** Comparison of the Prob-stack $\delta^{18}O_b$ record (Ahn et al., 2017) with the removal of a long-term increase of 0.083 ‰ Myr^{-1} (gray with 251-point running average in black) to our $\delta^{18}O_T$ reconstruction (light red with 251-point running average in red) and our $\delta^{18}O_{sw}$-INT reconstruction (light blue with 251-point running average in blue).

constructions (Ford and Raymo, 2020; Miller et al., 2020; Shakun et al., 2015; Sosdian and Rosenthal, 2009; Elderfield et al., 2012) and Fig. 11f comparing one derived from a regression-based sea-level reconstruction (Rohling et al., 2022). In general, there is good agreement between the BWT-derived reconstructions and our reconstruction, although there are some differences with reconstructions based on individual $\delta^{18}O_b$ records during interglaciations, particularly at site 1208 (Fig. 11e) (Ford and Raymo, 2020), which we attribute to some combination of differences in site-specific $\delta^{18}O_T$ and $\delta^{18}O_{sw}$. If the sea-level-based $\delta^{18}O_{sw}$ reconstruction (Rohling et al., 2021; Rohling et al., 2022) is a global signal not affected by local temperature or hydrography, then the high agreement between the two reconstructions (Fig. 11f) that were derived by completely independent means provides very high confidence in them. We also note that our LGM (19–26 ka) change in $\delta^{18}O_{sw}$ from modern values (0.9 ± 0.1 ‰) is in agreement with a pore-water-based reconstruction (1.0 ± 0.1 ‰) (Schrag et al., 1996, 2002).

Figure 12 compares our $\delta^{18}O_{sw}$-INT reconstruction with published reconstructions that span some part or all of the Late Pliocene and Early Pleistocene. There is a notable difference with the sea-level-based reconstruction for times older than 0.9 Ma (Rohling et al., 2021; Rohling et al., 2022), which is primarily expressed by glacial $\delta^{18}O_{sw}$ values being substantially less positive than in our reconstruction, particularly between 3–0.9 Ma (Fig. 12a). We attribute this difference to the previously noted problem with the regression approach used by Rohling et al. (2022, 2021) to reconstruct sea level that preserves the variability of the $\delta^{18}O_b$ record, including the increase in amplitude during the MPT. The Miller et al. (2020) $\delta^{18}O_{sw}$ reconstruction (Fig. 12b) used the smoothed BWT record from Cramer et al. (2011) to extract the $\delta^{18}O_{sw}$ signal, thus assuming that most of the orbital-scale $\delta^{18}O_b$ variability is comprised of $\delta^{18}O_{sw}$. This results in their $\delta^{18}O_{sw}$ reconstruction being in reasonable agreement with our $\delta^{18}O_{sw}$-INT reconstruction, although their Early Pleistocene $\delta^{18}O_{sw}$ glacial values tend to be 0.25 ‰–0.5 ‰ more negative than in our reconstruction (Fig. 12b). Figure 12c again compares BWT-based $\delta^{18}O_{sw}$ reconstructions from site 607 (Dwyer and Chandler, 2009; Sosdian and Rosenthal, 2009) that we previously used to gauge the sensitivity of our $\delta^{18}O_{sw}$ reconstruction to changes in HSE and long-term trends in $\delta^{18}O_b$ (Fig. 9). The ostracode-based val-

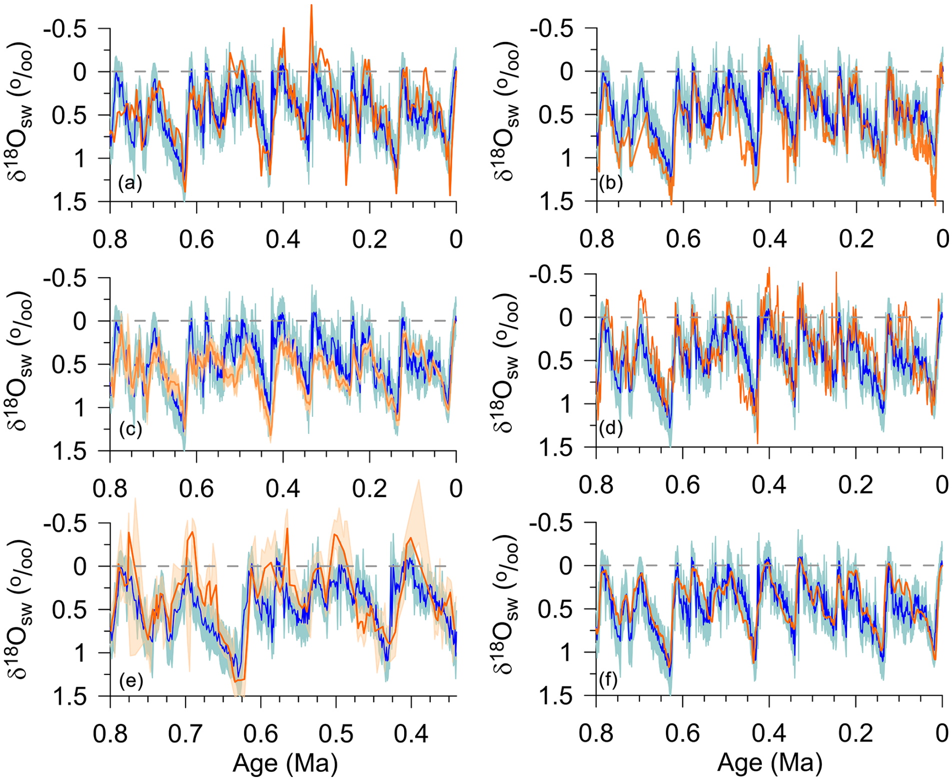

Figure 11. Comparison of our $\delta^{18}O_{sw}$-INT reconstruction (blue) to published $\delta^{18}O_{sw}$ reconstructions (orange). **(a)** Mg / Ca-based $\delta^{18}O_{sw}$ reconstruction from Sosdian and Rosenthal (2009). **(b)** $\delta^{18}O_{sw}$ reconstruction from Miller et al. (2020). **(c)** Proxy-based $\delta^{18}O_{sw}$ reconstruction from Shakun et al. (2015). **(d)** Mg / Ca-based $\delta^{18}O_{sw}$ reconstruction from Elderfield et al. (2012). **(e)** Mg / Ca-based $\delta^{18}O_{sw}$ reconstruction from Ford and Raymo (2020). **(f)** $\delta^{18}O_{sw}$ reconstruction from Rohling et al. (2022). We note that our LGM (19–26 ka) change in $\delta^{18}O_{sw}$ from modern values is 0.9 ± 0.1 ‰, in agreement with a pore-water-based reconstruction of 1.0 ± 0.1 ‰ (Schrag et al., 1996, 2002).

ues for site 607 > 3 Ma are on average slightly more positive than our reconstruction (Fig. 12c), which may reflect uncertainties in the Mg / Ca temperature calibration that requires validation. Otherwise, we emphasize the good agreement between records in their orbital-scale variability. Finally, Fig. 12d compares our reconstruction to two reconstructions from North Atlantic site U1313 (Jakob et al., 2020) that differ based on their Mg / Ca calibrations. As with nearby site 607, this record shows good agreement with our reconstruction in the orbital-scale variability, with differences in amplitude reflecting the different Mg / Ca calibrations.

We next compare our ΔMOT and $\delta^{18}O_{sw}$ reconstructions to those from Pacific sites 1123 and 1208 for two intervals when there are substantial differences between their reconstructions and ours (Fig. 13). Of the nine ΔMOT and ΔBWT reconstructions we had previously compared our ΔMOT reconstruction to for some or all of the last 0.7 Myr (Fig. 7), these two sites showed the largest differences, with site 1123 having good agreement in its temporal variability but having substantially warmer glacial intervals (Fig. 7d), while site 1208 showed some differences in the timing and amplitude of its variability (Fig. 7f). Figure 13a and c also show that temperatures at sites 1123 and 1208 are significantly different than our ΔMOT reconstruction between 0.9–1.4 Ma which spans much of the MPT.

These times of temperature differences between the two Pacific sites and our ΔMOT reconstruction result in significant differences in their site-specific $\delta^{18}O_{sw}$ values and our global $\delta^{18}O_{sw}$ reconstruction. Elderfield et al. (2012) and Ford and Raymo (2020) argued that the more positive $\delta^{18}O_{sw}$ values at sites 1123 and 1208 after 0.9 Ma suggest an increase in ice-sheet volume. Two factors, however, suggest that these changes may instead reflect regional hydrographic changes. The first is that ΔBWTs and $\delta^{18}O_{sw}$ values at the start of the site 1123 record (1.55 Ma) are similar to our reconstructions until ∼ 1.4 Ma, when they depart from our reconstructions until ∼ 0.9 Ma, when they then merge again with our reconstructions (Fig. 13a, b). This implies that site 1123 is recording large ice sheets at ∼ 1.5 Ma, or before the MPT. The second aspect that suggests that the more-negative $\delta^{18}O_{sw}$ values at these two sites between 1.4–0.9 Ma are not representative of global values is based on their implications for sea-level change. As a first-order approximation, we scale $\delta^{18}O_{sw}$ to sea level using 0.008 ‰ m^{-1} as derived from LGM

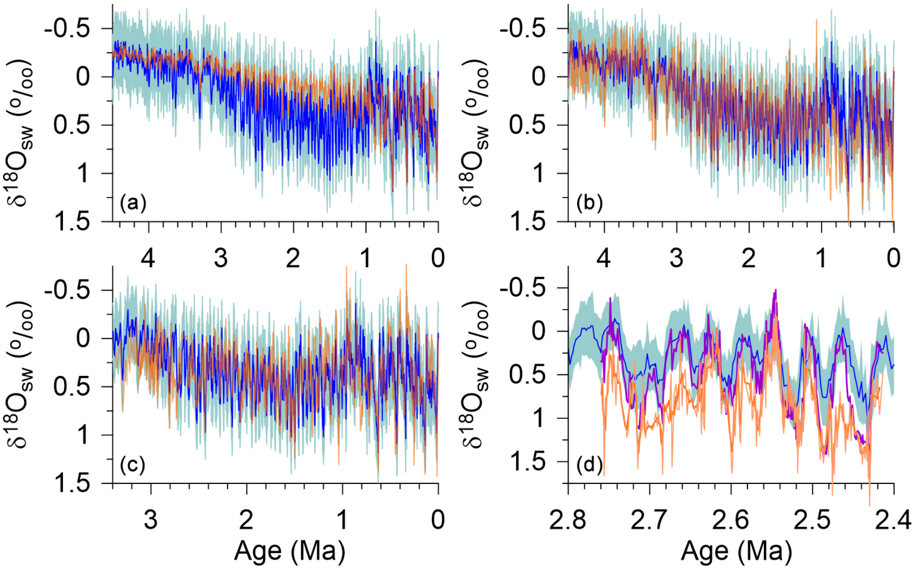

Figure 12. Comparison of our $\delta^{18}O_{sw}$-INT reconstruction (blue) to published $\delta^{18}O_{sw}$ reconstructions (orange). **(a)** $\delta^{18}O_{sw}$ reconstruction from Rohling et al. (2022). **(b)** $\delta^{18}O_{sw}$ reconstruction from Miller et al. (2020). **(c)** Mg / Ca-based $\delta^{18}O_{sw}$ reconstruction from North Atlantic site 607 from Dwyer and Chandler (2009) and Sosdian and Rosenthal (2009). **(d)** Mg / Ca-based $\delta^{18}O_{sw}$ reconstruction from Jakob et al. (2020). The purple line is the published $\delta^{18}O_{sw}$ reconstruction, whereas the orange line is the $\delta^{18}O_{sw}$ reconstruction based on recalibrating the Mg / Ca temperature data using Barrientos et al. (2018).

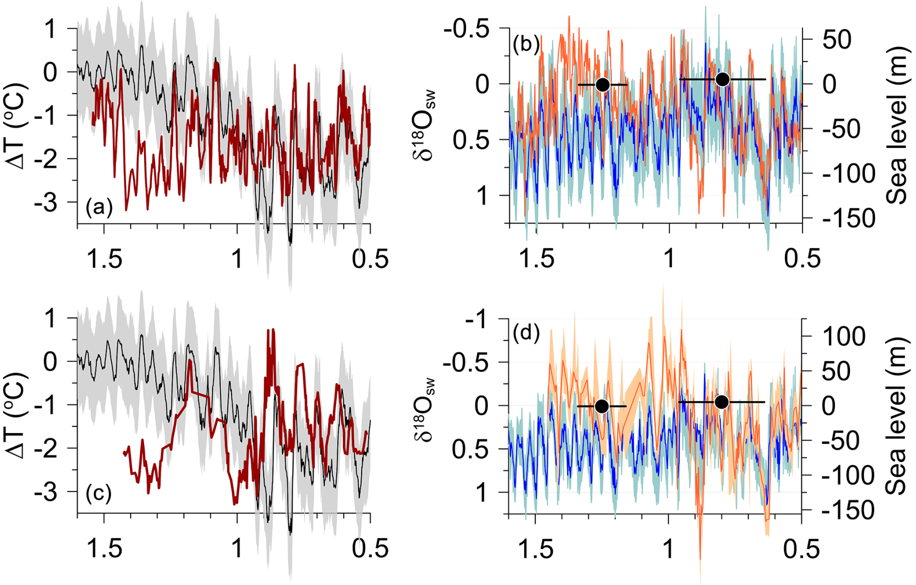

Figure 13. **(a)** Comparison of our ΔMOT reconstruction (black line with 1σ uncertainty) to ΔBWT reconstruction from ODP site 1123 (11-point running average shown by dark-red line) (Elderfield et al., 2012). **(b)** Comparison of our $\delta^{18}O_{sw}$-INT reconstruction (blue line with 1σ uncertainty) to the $\delta^{18}O_{sw}$ reconstruction from ODP site 1123 (orange line) (Elderfield et al., 2012). Black symbols with uncertainties are dated sea-level indicators (Dumitru et al., 2021). **(c)** Comparison of our ΔMOT reconstruction (black line with 1σ uncertainty) to ΔBWT reconstruction from ODP site 1208 (Ford and Raymo, 2020) (5-point running average shown by dark-red line). **(d)** Comparison of our $\delta^{18}O_{sw}$-INT reconstruction (blue line with 1σ uncertainty) to the $\delta^{18}O_{sw}$ reconstruction from ODP site 1208 (orange line with 1σ uncertainty) (Ford and Raymo, 2020). Black symbols with uncertainties are dated sea-level indicators (Dumitru et al., 2021).

pore-water (Schrag et al., 2002) and sea-level (Lambeck et al., 2014) estimates. The scaled pre-MPT $\delta^{18}O_{sw}$ values at the two sites would lead to sea-level highstands that are 50–100 m higher than present throughout much of the 0.9–1.4 Ma interval (Fig. 13b, d), thus implying an essentially ice-free world. We thus conclude that the differences in ΔBWT and $\delta^{18}O_{sw}$ at sites 1123 and 1208 from global values between 1.4-0.9 Ma reflect regional hydrographic changes (i.e., salinity) that were perhaps associated with the large changes in ocean circulation during the MPT (Lisiecki, 2014; Lang et al., 2016).

4 Processes that contribute to changes in ocean heat storage and mean ocean temperature

During time-dependent climate change, the difference between the radiative forcing at the top of the atmosphere and Earth's radiative response leads to an imbalance in the Earth's energy budget, with a positive imbalance causing the climate system to gain energy and warm and a negative imbalance causing a loss of energy that cools the climate system. Under current anthropogenic climate change, the radiative forcing has exceeded the Earth's radiative response, with 90 % of the resulting energy gain being stored in the ocean over the last few decades (von Schuckmann et al., 2023), thus strongly buffering warming of the atmosphere. The greatest ocean storage over the last century has occurred in the upper 700 m, with only weak warming at depths below 2000 m (von Schuckmann et al., 2023; Cheng et al., 2022) because of the slow transfer of energy into the ocean interior by advection, diffusion, and vertical mixing (Rugenstein et al., 2019; Gregory, 2000; Saenko et al., 2021) so that much of the ocean has not yet reached its equilibrium temperature change and HSE is only ∼ 0.1.

Changes in ocean heat storage similarly play an important role in pacing surface temperature change on longer timescales. On orbital timescales (10^4–10^5 years), the contribution to the global energy budget from latent heat fluxes associated with large changes in land ice became comparable to changes in ocean heat storage, with each accounting for ∼ 50 % of the increase in the internal energy of the climate system during the last deglaciation (Baggenstos et al., 2019). Changes in some combination of these two energy reservoirs indicate that the global energy budget has rarely been in balance for any extended period throughout the Plio-Pleistocene glacial–interglacial cycles (Shackleton et al., 2023).

Patterns of ocean heat storage since the late 19th century are largely associated with changes in ocean circulation that redistribute heat but do not change global heat content (Bronselaer and Zanna, 2020; Gregory et al., 2016; Cheng et al., 2022). However, observations over the last few decades and climate models show that large-scale patterns of heat storage are increasingly being determined and sustained by heat from anthropogenic surface warming that is added to the ocean interior predominantly along known water mass pathways, with this added heat dominating ocean heat storage change by 2100 (Bronselaer and Zanna, 2020; Fox-Kemper et al., 2021; Cheng et al., 2022). The patterns of OHU and storage show most warming occurring in the upper 2000 m between 60° S and 60° N, with the majority of heat uptake occurring within wind-driven subduction regions in the Southern Ocean that ventilate the ocean interior (Kuhlbrodt and Gregory, 2012; Gregory et al., 2016), particularly in Subantarctic Mode Water and Antarctic Intermediate Water (Zanna et al., 2019; Cheng et al., 2022), a pattern that persists in equilibrium runs forced by CO_2 quadrupling (Fig. 14a, b) (Li et al., 2013) and in the PlioMIP2 multi-model ($n = 15$) ensemble (Fig. 14g, h) (Haywood et al., 2020). Area-averaged warming in the Pacific Ocean will be smaller than in other basins because of the lack of deepwater formation and limited formation of mode and intermediate water in the North Pacific (Cheng et al., 2022). The increase in Atlantic Ocean heat storage is projected to be nearly equivalent to that of the Pacific but, because of its smaller area, results in a significantly larger area-averaged warming (Cheng et al., 2022).

The temperature of the deep ocean (> 2000 m) is largely associated with deepwater formation at high latitudes. Described in broad terms, this process cools the deep ocean at a rate $v(T_u - T_d)$, where v (in s^{-1}) is the volume rate of deepwater formation divided by the volume of the deep ocean, T_d is the temperature of the water sinking at high latitude following convection, and T_u is the temperature of water upwelling at lower latitude in the basin[3]. The latter temperature also represents the influence of downward mixing of heat from overlying warmer water at low latitudes. By continuously injecting cold water, the overall effect of the overturning circulation associated with NADW and AABW is to keep the deep ocean cooler than overlying intermediate-depth waters, which are ventilated by lower-latitude surface waters. A change in MOT thus occurs from some combination of changes in mid-latitude SSTs (affecting T_u), high-latitude SSTs or temperature of the newly formed and sinking deep water (affecting T_d), and high-latitude deepwater formation rate (ν).

Projected warming in the deep North Atlantic in response to anthropogenic warming is caused by a reduction in surface heat loss in that region and in the formation of NADW in response to anthropogenic warming (Fox-Kemper et al., 2021). However, NADW does not contribute significantly to global

---

[3] If the volume V_d of the ocean (in m^3) occupied by deep water is steady, deepwater formation at a rate r_d (in $m^3\,s^{-1}$) must be balanced by an equal rate of removal of deep water by upwelling and mixing with overlying water masses. The rate (in W) at which heat is removed from the deep ocean by this throughput is $r_d(T_u - T_d)C$, where C is the volumetric heat capacity of seawater (in $J\,m^{-3}\,°C^{-1}$). Since the heat capacity of the deep ocean is V_dC ($J\,°C^{-1}$), its rate of cooling (in $°C\,s^{-1}$) is $r_d(T_u - T_d)C/(V_d C) = \nu(T_u - T_d)$, where $\nu = r_d/V_d$, whose reciprocal $\tau = 1/\nu$ is the time required to renew the entire volume of the deep ocean.

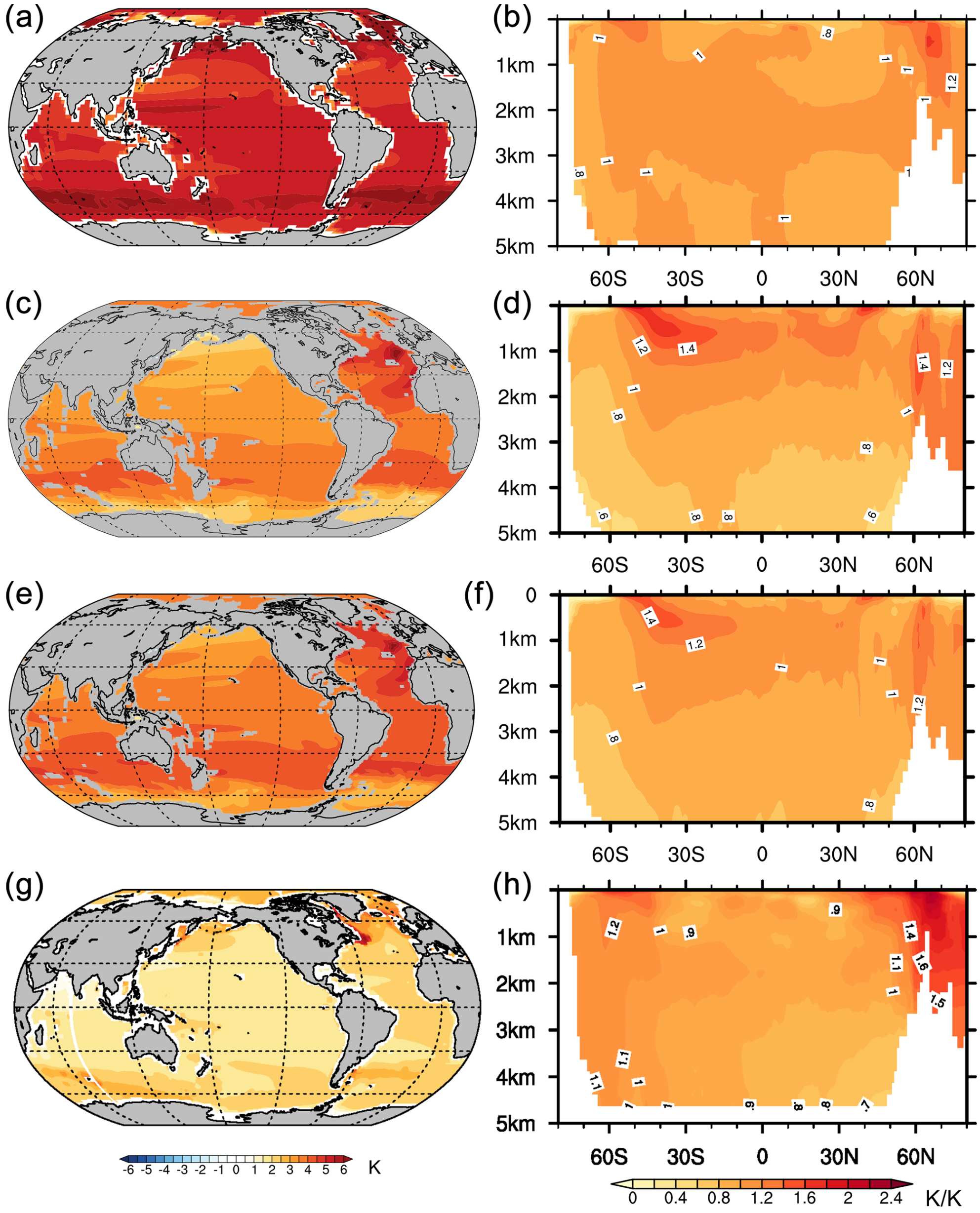

Figure 14. Changes in 0–2000 m mean equilibrium temperature change (left) and zonal average equilibrium ocean warming pattern normalized with ΔMOT (right) for different modeling scenarios. **(a, b)** CESM 1.0.4 for abrupt 4 × CO_2 minus preindustrial after 5900 years, ΔMOT = 5.12 K (Rugenstein et al., 2019). **(c, d)** iTRACE ICE + ORB + GHG simulation for Middle Holocene (6 ka) minus LGM, ΔMOT = 3.02 K (Zhu et al., 2024). **(e, f)** Full iTRACE simulation (ICE + ORB + GHG + AMOC) for Middle Holocene (6 ka) minus LGM, ΔMOT = 3.6 K (Zhu et al., 2024). **(g, h)** Model mean of PlioMIP2 simulations, ΔMOT = 1.86 K (Haywood et al., 2020).

OHU (Saenko et al., 2021), and it makes only a small contribution to changes in MOT because it ventilates only a small volume, which is currently ∼ 20 %–30 % of the global ocean (Johnson, 2008; Khatiwala et al., 2012) that decreases during glaciations (Galbraith and de Lavergne, 2019). The correlation between AMOC and ocean heat uptake efficiency across models seems to be due to a common control, such as vertical stratification of the global ocean, rather than to an effect of AMOC on heat uptake (Gregory et al., 2024; Newsom et al., 2023).

In contrast, AABW has a significantly larger influence on MOT than NADW because it ventilates a larger volume, which is currently ∼ 40 %–50 % of the global ocean (Johnson, 2008; Khatiwala et al., 2012), increasing to as much as 80 % during glaciations (Galbraith and de Lavergne, 2019). AABW is formed by intense heat loss and brine rejection due to sea-ice formation, with the dense waters sinking, spreading northward to fill much of the abyss, and upwelling as they mix with overlying warm waters especially where in contact with areas of rough seafloor topography. At present, because the temperature of its source waters (T_d in the conceptual model above) remains near the freezing point ($\sim -1.8\,^\circ\mathrm{C}$), the influence of AABW on abyssal temperatures has been through a reduction in the volume rate of formation (υ above) in response to freshening associated with increased meltwater from the Antarctic Ice Sheet (Heuzé et al., 2015; Li et al., 2023) or decreased sea-ice formation (Zhou et al., 2023), allowing more heat to diffuse or mix downwards and warm the deep ocean (Purkey et al., 2019; Johnson et al., 2024). Further global warming and associated sea-ice loss will allow AABW source waters to warm, further contributing to warming of abyssal temperatures (Fig. 14b).

This understanding of the major processes involved in OHU and storage in response to GHG emission scenarios over the course of this century (Fox-Kemper et al., 2021; Cheng et al., 2022) or on equilibrium timescales (10^3 years) (Fig. 14a, b, g, h) (Rugenstein et al., 2016; Li et al., 2013) contrasts with the long-standing view in paleoceanography that changes in DOT and MOT result solely from SST changes in high-latitude regions where deep water is formed (Emiliani, 1954; Zachos et al., 2001; Hansen et al., 2023; Evans et al., 2024; Westerhold et al., 2020; Bereiter et al., 2018; Rohling et al., 2022; Hansen et al., 2013). Changes in source water temperature may indeed cause changes in deepwater temperature, for example, during substantially warmer climates without Antarctic sea ice (Goudsmit-Harzevoort et al., 2023; Evans et al., 2024) when AABW could form during Antarctic winter solely by heat loss to the atmosphere without brine rejection, like NADW in the present climate. However, the general notion of a sole control of MOT by SSTs at sites of deepwater formation (affecting T_d) should not be applied regardless of climate state because it neglects the roles of the rate (υ) of deepwater formation and the contribution from OHU in mid-latitudes, both of which can also affect ocean heat storage (Fig. 14a, b) and thus contribute to MOT.

Zhu et al. (2024) used results from the iTRACE simulation of the last deglaciation to assess the role of the primary forcings of Plio-Pleistocene climate change on MOT change. The iTRACE ORB + ICE + GHG simulation includes most of the key forcings during the Plio-Pleistocene ice-age cycles. Orbital forcing has little direct influence on GMSST, leaving ice sheets as the primary forcing that modulates the MOT response to GHG forcing. The effect of lowering of Northern Hemisphere ice sheets as they retreat induces surface warming that is advected downwind to the North Atlantic and North Pacific, further enhancing SST warming (Fig. 15) and OHU through ventilation of intermediate waters at 45° N (Fig. 14d). Ice-sheet forcing in the ORB + ICE + GHG simulation thus significantly enhances intermediate-depth warming at that is otherwise largely occurring through wind-driven ventilation in the Southern Ocean (Fig. 14c).

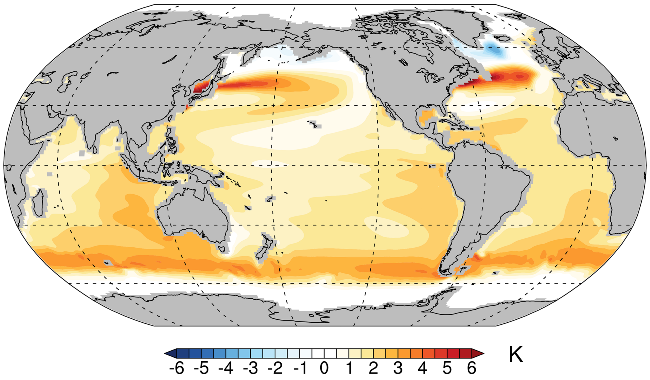

Figure 15. iTRACE ICE-only (ICE run) simulation of ocean surface temperature change (upper 10 m) for the Middle Holocene (6 ka) minus LGM.

Zhu et al. (2024) found that high-latitude source waters where deep water is formed are largely covered by sea ice, resulting in their temperature remaining around the freezing point throughout much of the deglaciation. The presence of sea ice results in peak SST warming during deglaciation occurring in mid-to-subpolar latitudes (e.g., Fig. 15) as opposed to peak surface air temperature warming occurring at high latitudes through polar amplification (Zhu et al., 2024). Strong ventilation regions remained co-located with peak SST warming throughout the deglaciation, particularly in the mid-latitudes of the Southern Ocean associated with ventilation of Antarctic Intermediate Water. Figure 14d shows that this was the primary pathway of warming the global ocean, with a negligible warming contribution from reduced AABW formation (Zhu et al., 2024). During the deglaciation, the iTRACE ORB + ICE + GHG simulation shows that MOT warming lags SST warming by several thousand years, reflecting the timescale of warming the ocean interior by ventilation of intermediate waters and resulting in an average transient HSE ~ 0.5 that disagrees with proxy records (Zhu et al., 2024). However, this simulation shows that the lag is short enough to allow MOT warming to reach equilibrium during the present interglaciation and HSE reaches ~ 1.

The full iTRACE simulation includes meltwater forcing (MWF) associated with Heinrich event 1 and the Younger Dryas (He et al., 2021; Gu et al., 2020) and thus captures MOT changes during times of millennial-scale AMOC variability. The strong AMOC weakening in response to MWF causes northward heat transport to decrease, resulting in the characteristic SST bipolar seesaw pattern of Northern Hemisphere cooling and Southern Hemisphere warming. GMSSTs do not change substantially, however, because the effects in both hemispheres on SST nearly cancel each other. In contrast, a suppression of NADW production and reduction of the AMOC generates a subsurface warming that extends to intermediate depths over much of the global ocean (Fig. 14e) and warmed abyssal waters through circulation and mixing processes (Fig. 14f). The MWF causes the subsurface ocean warming that eventually occurs from orbital forcing and land ice on temperature to occur more rapidly. Thus it eliminates the unrealistic lag of MOT behind GMSST, resulting in HSE ≥ 1 throughout the deglaciation, in agreement with proxy records (Zhu et al., 2024).

5 A general hypothesis for the increase in ocean heat storage efficiency during the Middle Pleistocene Transition

Based on the understanding of controls on OHU and MOT outlined in Sect. 4, we develop a working hypothesis for the reconstructed increase in HSE that occurred during the MPT. The main premise of our simple conceptual model is based on the ocean being comprised of an upper ocean heat reservoir (herein $R_{<2000}$) that extends from roughly 50° S to 50° N and to a depth of about 2000 m and has a volume that in PlioMIP2 is 55 % of the global ocean ($f_{<2000} = 0.55$) and in iTRACE is 43 % of the global ocean ($f_{<2000} = 0.43$), with the heat content and temperature being largely determined by ventilation of mid-latitude surface waters (Fig. 14). The remainder of the ocean (herein $R_{>2000}$), the deeper ocean heat reservoir, is largely below 2000 m, is connected to the surface at latitudes of $> 50°$ S and $> 50°$ N, and represents 45 % (PlioMIP2) and 57 % (iTRACE) of the global ocean volume, with the heat content and temperature being largely determined by high-latitude deepwater formation (some combination of T_d and v). In this simple model, ΔMOT is equal to the mean of the changes in temperatures of the upper reservoir $R_{<2000}(\Delta T_{<2000})$ and deeper reservoir $R_{>2000}(\Delta T_{>2000})$ weighted by their relative ocean volumes ($\Delta\mathrm{MOT} = f_{<2000} \times \Delta T_{<2000} + (1 - f_{<2000}) \times \Delta T_{>2000}$).

We derive parameters from climate models that identify the relation of the temperature of the upper reservoir to ΔGMSST and then calculate the temperature of the two reservoirs prior to the MPT (> 1.5 Ma) and since 1.5 Ma. For > 1.5 Ma, we use the PlioMIP2 results (Haywood et al., 2020) to calculate that the average temperature change of the upper reservoir $\Delta T_{<2000}$ is 80 % that of ΔGMSST, corresponding to a multiplying factor $s = \Delta T_{<2000}/\Delta\mathrm{GMSST}$ of 0.8. For the time since 1.5 Ma, we use the full iTRACE results (Zhu et al., 2024) to calculate that the average temperature change of the upper reservoir $\Delta T_{<2000}$ is 16 % greater than ΔGMSST, corresponding to a multiplying factor $s = \Delta T_{<2000}/\Delta\mathrm{GMSST}$ of 1.16. As a simple scaling analysis, we use these model parameters to derive upper reservoir $\Delta T_{<2000}$ from the ΔGMSST reconstruction (Fig. 16a) by multiplying it with a factor $s = 0.8$ for > 1.5 Ma and $s = 1.16$ for < 1.5 Ma, including $\sigma_{\Delta T<2000} = s \times \sigma_{\Delta\mathrm{GMSST}}$ (Fig. 16c). The temperature change of the deep reservoir $\Delta T_{>2000}$ for these two time intervals is readily derived from the equation for MOT, with its uncertainty being the square root of the sum of squares of the individual uncertainties ($\sigma_{>2000} = (\sigma_{\Delta\mathrm{MOT}}/(1 - f_{<2000}))^2 + (f_{<2000} \times s \times \sigma_{\Delta\mathrm{GMSST}})^2$) (Fig. 16d). For the interval since 1.5 Ma, we assessed the sensitivity of upper reservoir $\Delta T_{>2000}$ to different values of $f_{<2000}$ (in the range 0.4–0.6) and of s (1.11–1.54), with all results falling well within the 1σ uncertainty of $\Delta T_{>2000}$ (not shown).

Albeit highly idealized, this simple analysis suggests that, for the period from 4.5 Ma until the start of the MPT around 1.5 Ma, $\Delta T_{<2000}$ based on the nominal PlioMIP2 parameters ($s = 0.8$, $f_{<2000} = 0.55$) (Fig. 16c) accounts for nearly all of ΔMOT (Fig. 16b) (which in this time window has an HSE of 0.5), leaving only a small decrease in $\Delta T_{>2000}$ (Fig. 16d) and thus providing an explanation for HSE being ~ 0.5. In other words, before the MPT, nearly all of ΔMOT is occurring in the upper reservoir, which is ventilated by the wind-driven circulation, and cools along with ΔGMSST. The global cooling trend throughout this period is assumed to be a response to declining CO_2 (Clark et al., 2024), but the cause of it does not affect our argument. Meanwhile, little change is occurring in the high-latitude deepwater formation rate (v) and sinking water temperature (T_d) and thus in deep reservoir temperature ($\Delta T_{>2000}$).

The lack of substantial deep reservoir $\Delta T_{>2000}$ change on orbital (10^4–10^5 years) and geological (10^6 years) timescales between 4.5–1.5 Ma suggests relatively stable and constant AABW formation, with a subsequent decrease around 1.5 Ma in long-term $\Delta T_{>2000}$ and a rise in its variability, suggesting that significant changes in AABW formation had begun. Two lines of evidence support this scenario. First, multi-model results from PLIOMIP2 found that Southern Ocean ΔSSTs for the KM5c interglaciation at 3.205 Ma were 2.8 ± 1.3 °C (Weiffenbach et al., 2024) (Fig. 16e). These warmer SSTs, combined with a simulated increase in precipitation and decrease in sea-ice cover, resulted in a strongly stratified Southern Ocean with relatively uniform warming of 1.5–2.5 °C throughout much of the water column below the low-density surface layer. In 9 of the 15 PLIOMIP2 models, this increase in stratification led to a decrease in AABW formation, with Weiffenbach et al. (2024) noting that four of the other six models also have greater stratification, but the AABW response may be modulated by interactions with a stronger AMOC in those models.

To examine whether these mid-Pliocene boundary conditions extended into the Pleistocene, we use the SH extratropical ($> 30°$ S) ΔSST stack from Clark et al. (2024) as a proxy for Southern Ocean SSTs. (Note that this stack only extends to 4 Ma because of the limited number of older records available, and so our analysis here only covers the last 4 Myr.) This inference is supported by the good agreement with the ΔSST reconstruction for the Southern Ocean derived from deuterium excess from the Dome Fuji ice core for the last 0.7 Myr (Uemura et al., 2018) and the PLIOMIP2 multi-model mean Southern Ocean ΔSST of 2.8 ± 1.3 °C at 3.205 Ma (Weiffenbach et al., 2024) compared to the stack's 2.4 ± 1.3 °C (Clark et al., 2024) (Fig. 16e). At

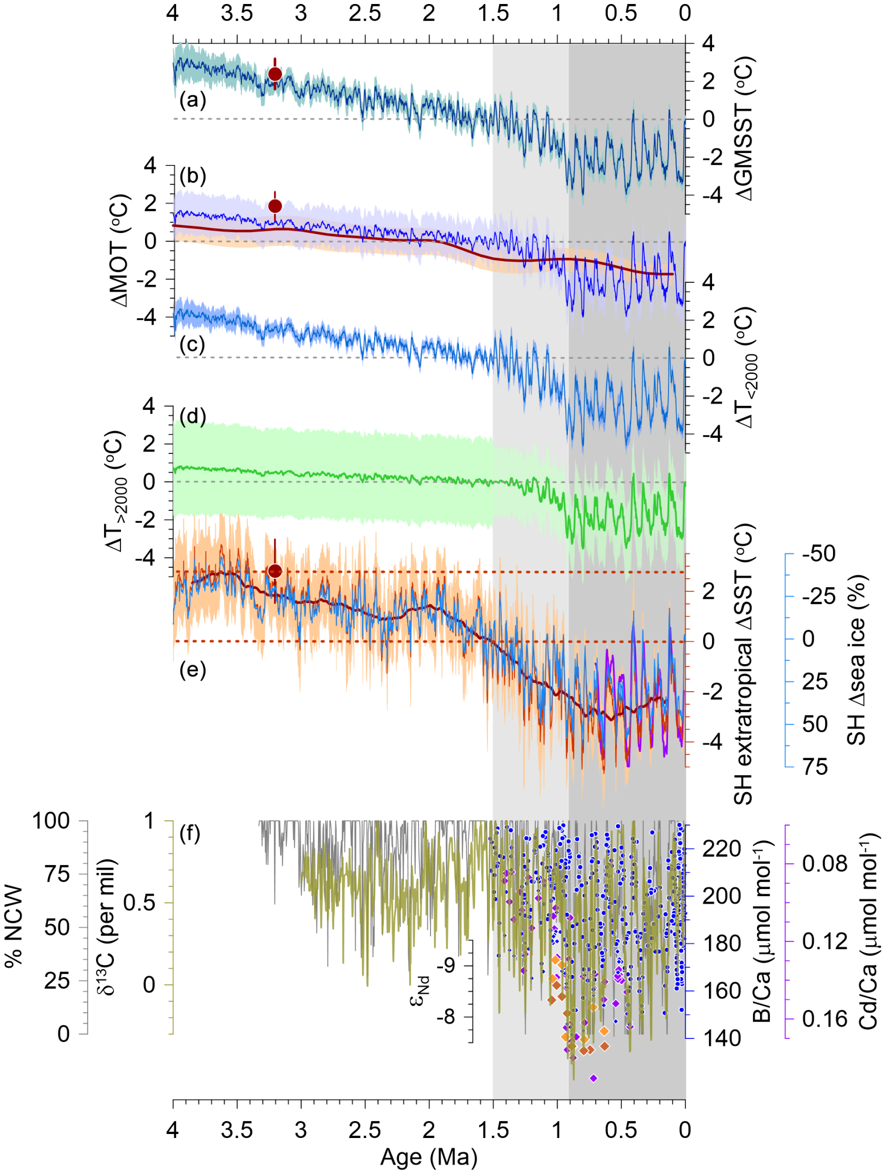

Figure 16. **(a)** Global mean sea surface temperature change from preindustrial (PI) (ΔGMSST) (dark-blue line with 1σ uncertainty) (Clark et al., 2024). Red symbol with 1σ uncertainty is the model mean ΔGMSST during the KM5c time slice at 3.205 Ma from PlioMIP2 (Haywood et al., 2020). **(b)** Mean ocean temperature change from PI (ΔMOT) (blue line with 1σ uncertainty) as derived in this paper. Also shown is the smoothed ΔBWT reconstruction from Cramer et al. (2011) using their Eq. (7b) (brick-red line with 90 % confidence interval). The red symbol with 1σ uncertainty is the model mean MOT during the KM5c time slice at 3.205 Ma from PlioMIP2 (Haywood et al., 2020). **(c)** ΔT for the $R_{<2000}$ area of ocean (43 % of global volume) (lighter blue line with 1σ uncertainty) that is on average 16 % greater than ΔGMSST ($\Delta T\_R_{<2000} = \Delta\mathrm{GMSST} \times 1.16$) (Zhu et al., 2024). **(d)** ΔT for the $R_{>2000}$ area of ocean (57 % of global volume) (green line with 1σ uncertainty) derived by $\Delta T\_R_{>2000} = (\Delta\mathrm{MOT} - 0.43 \times \Delta T\_R_{<2000})/0.57$. **(e)** Southern Hemisphere extratropical ΔSST stack (brick-red line with 1σ uncertainty and 201 kyr running average in dark red) (Clark et al., 2024), Southern Ocean Δsea-ice extent derived from the relation to the Southern Hemisphere extratropical ΔSST stack established by PLIOMIP2 models (Weiffenbach et al., 2024) (light-blue line), ΔSST reconstruction for the Southern Ocean for the last 0.7 Myr derived from deuterium excess from the Dome Fuji Antarctic ice core (purple = 25 kyr running average) (Uemura et al., 2018), and PlioMIP2 simulated multi-model mean ΔSST reconstruction for the Southern Ocean during the KM5c time slice at 3.205 Ma shown by the red symbol and 1σ uncertainty (2.8 ± 1.3 °C) (Weiffenbach et al., 2024). Two dashed red horizontal lines correspond to PLIOMIP2 ΔSST at 3.205 Ma (2.8 °C) and at PI (0 °C). **(f)** δ^{13}C stack of mid-to-deep Atlantic cores (green) (Lisiecki, 2014), percent of northern component waters (NCWs) (gray) (Lang et al., 2016), $\varepsilon_{\mathrm{Nd}}$ data from South Atlantic sites ODP 1088 (light-orange diamonds) and 1090 (red-brown diamonds) (Pena and Goldstein, 2014), and Cd / Ca (purple diamonds) and B / Ca data from North Atlantic cores CHN82-24-23PC and DSDP 607 (blue circles) (Sosdian et al., 2018; Lear et al., 2016) and South Atlantic site ODP 1267 (blue diamonds) (Farmer et al., 2019).

the same time, the PLIOMIP2 models find a linear relation between Southern Ocean ΔSST and Δsea-ice area in percentages relative to preindustrial times where Δsea-ice area = $-12.4\,\%\,^{\circ}\mathrm{C}^{-1} \times \Delta\mathrm{SST}$ (Weiffenbach et al., 2024), which we apply to our ΔSST stack to derive relative changes in sea-ice area over the last 4 Myr (Fig. 16e).

These results suggest that the highly stratified Southern Ocean found in PLIOMIP2 simulations at 3.2 Ma due to warm SSTs and reduced sea-ice extent persisted until ∼ 1.5 Ma. Prior to this, SSTs and sea-ice extent spent 90 % of the time above and below, respectively, their preindustrial values. Apparently, SSTs in the AABW formation regions did not decline sufficiently during this interval to affect $\Delta T_{>2000}$ and MOT substantially by lowering T_{d}. The start of the MPT at 1.5 Ma saw a significant change in the influences on AABW, with SSTs spending increasingly more time below preindustrial levels and sea-ice extent spending more time above preindustrial values (∼ 85 % during the MPT, ∼ 95 % since 0.9 Ma). This might also be related to the proposed decoupling of Southern Ocean vertical mixing and Southern Ocean SST prior to the MPT (Köhler and Bintanja, 2008). We thus conclude that it was the persistence of a highly stratified Southern Ocean that caused a decrease in AABW formation rate and persistently warmer T_{d} than present until ∼ 1.5 Ma, when the gradual decay of stratification and increase in sea-ice extent and variability then enhanced conditions for AABW formation.

Second, deep-ocean water masses show changes that are consistent with changes in AABW formation inferred from our simple model; i.e., v increased at that time. Starting at ∼ 1.5 Ma, there was an increasing frequency of southern component waters (e.g., AABW) (Lang et al., 2016), which are depleted in δ^{13}C (Lisiecki, 2014), recording a growing influence of AABW at the expense of NADW in the Atlantic Ocean (Fig. 16f). This was followed by a further step-change increase in the relative share of δ^{13}C-depleted AABW during glacial climates around 0.9 Ma, as also indicated by a rapid increase in $\varepsilon_{\mathrm{Nd}}$ values (Pena and Goldstein, 2014) and an increase in nutrient content and a decrease in carbonate ion saturation as shown in the Cd / Ca and B / Ca records, respectively (Lear et al., 2016; Sosdian et al., 2018; Farmer et al., 2019) (Fig. 16f).

After the MPT, ΔMOT variability was greater than before, exceeding the contribution from upper reservoir $\Delta T_{<2000}$. During the glacial cycles of the last 0.8 Myr, proxy records suggest that v has varied along with ΔGMSST (Fig. 16f) (Clark et al., 2024). Without requiring any long-term change in T_{d}, which has remained near freezing, this can explain synchronized variations in $\Delta T_{<2000}$ and $\Delta T_{>2000}$, leading finally to an HSE of ∼ 1 during this period.

6 Summary

When compared to a reconstruction of ΔGMSST over the last 4.5 Myr, high-fidelity proxies of deep (> 200 m) ocean temperature change show good agreement in orbital-scale amplitude and long-term trend over the last 0.7 Ma, but their long-term trends are ∼ 50 % of long-term ΔGMSST beyond 1.5 Ma, suggesting an increase in HSE from ∼ 0.5 to ∼ 1 during the MPT (1.5–0.9 Ma). This increase is further supported when assuming that HSE was 1 throughout the last 4.5 Myr and applying this temperature history (as $\delta^{18}\mathrm{O}_{\mathrm{T}}$) to isolate the seawater component ($\delta^{18}\mathrm{O}_{\mathrm{sw}}$) of a probabilistic global $\delta^{18}\mathrm{O}_{\mathrm{b}}$ stack (Prob-stack). Under this scenario, Pliocene $\delta^{18}\mathrm{O}_{\mathrm{sw}}$ values are 0.3 ‰ to 0.5 ‰ despite robust evidence for higher-than-present Pliocene sea levels that require values in $\delta^{18}\mathrm{O}_{\mathrm{sw}}$ to be smaller than 0 ‰, suggesting that too much of the $\delta^{18}\mathrm{O}_{\mathrm{b}}$ signal is being removed by the $\delta^{18}\mathrm{O}_{\mathrm{T}}$ component using an HSE of 1. Applying our ΔMOT reconstruction where HSE increases from 0.5 to 1 across the MPT results in Early Pleistocene and Pliocene $\delta^{18}\mathrm{O}_{\mathrm{sw}}$ values (0 ‰ to −0.1 ‰) that continue to be more positive than multiple data constraints. While a further decrease in HSE to 0.1 would result in average Pliocene $\delta^{18}\mathrm{O}_{\mathrm{sw}}$ values of −0.2 ‰, which could explain higher sea levels at that time, such a low HSE is ruled out by proxy-based bottom water temperature reconstructions. We therefore adopt the hypothesis that there has been a diagenetic overprint on $\delta^{18}\mathrm{O}_{\mathrm{b}}$ records that average to a long-term increase of between 0.05 ‰ Myr^{-1} and 0.12 ‰ Myr^{-1} which, when removed from the Prob-stack, results in Pliocene $\delta^{18}\mathrm{O}_{\mathrm{sw}}$ values of −0.1 ‰ to −0.4 ‰ that are consistent with sea-level highstands of 20–25 m above present.

To explain the increase in HSE across the MPT, we develop a simple conceptual model that considers the ocean as being comprised of an upper non-polar ocean reservoir with the temperature being largely determined by ventilation of mid-latitude surface waters and a deeper ocean reservoir whose temperature is largely determined by high-latitude deepwater formation. Using results from a transient simulation of the last deglaciation with a global climate model, we develop a simple scaling analysis to derive upper reservoir ΔT from the ΔGMSST reconstruction, which is then subtracted from ΔMOT to derive deep reservoir ΔT. This analysis suggests that before the MPT, nearly all of ΔMOT is occurring in the upper reservoir through changes in wind-driven ventilation and little is occurring in the deep reservoir from changes in deepwater formation, resulting in HSE being ∼ 0.5. Around 1.5 Ma, the amplitude of ΔMOT variability begins to increase and exceeds the contribution from upper reservoir ΔT, thus requiring an increasing contribution of lower reservoir ΔT to ΔMOT through an increase in deepwater formation that leads to an HSE of ∼ 1 over the last 0.8 Myr. We attribute these changes in deepwater formation to long-term cooling which caused a transition starting ∼ 1.5 Ma from a highly stratified Southern Ocean due to warm SSTs and reduced sea-ice extent to colder SSTs with a significant increase in sea-ice extent and more vertical exchange of water masses.

Code and data availability. Our ΔMOT and $\delta 18\mathrm{O}_{\mathrm{sw}}$ data are available in the Supplement. R code that implements the regression analysis between ΔMOT and ΔGMSST can be found at Zenodo (https://doi.org/10.5281/zenodo.14759006, Bartlein, 2025) or in the GitHub repository at (https://github.com/pjbartlein/MOTvsSST, last access: January 2025).

Supplement. The supplement related to this article is available online at https://doi.org/10.5194/cp-21-973-2025-supplement.

Author contributions. Conceptualization: PUC and JDS. Methodology: PUC, JDS, YR, CZ, DPS, and PK. Investigation: PUC, JDS, YR, CZ, JMG, PK, ZL, DPS, and PJB. Writing – original draft: PUC, JDS, and YR. Writing – review and editing: PUC, JDS, YR, JMG, PK, CZ, ZL, PJB, and DPS.

Competing interests. The contact author has declared that none of the authors has any competing interests.

Disclaimer. Publisher's note: Copernicus Publications remains neutral with regard to jurisdictional claims made in the text, published maps, institutional affiliations, or any other geographical representation in this paper. While Copernicus Publications makes every effort to include appropriate place names, the final responsibility lies with the authors.

Acknowledgements. We thank Lorraine Lisiecki and Thomas Stocker for their constructive reviews. We thank the paleoclimate and paleoceanographic communities for making their data sets widely available; the National Centers for Environmental Information of NOAA and the World Data Center PANGAEA for archiving data; and Chris Brierley, Julia Tindall, and Julia E. Weiffenbach for providing climate model data. This publication contributed to Beyond EPICA, a project of the European Union's Horizon 2020 research and innovation program (Oldest Ice Core). This is Beyond EPICA publication number 43. Peter U. Clark is funded by National Science Foundation OPP-2103032, Yair Rosenthal is supported by National Science Foundation OCE-1834208, and Zhengyu Liu is supported by National Science Foundation OCE-1810681.

Financial support. This research has been supported by the National Science Foundation (grant nos. OPP-2103032, OCE-1834208, and OCE-1810681).

Review statement. This paper was edited by Denis-Didier Rousseau and reviewed by Lorraine Lisiecki and Thomas Stocker.

References

Ahn, S., Khider, D., Lisiecki, L. E., and Lawrence, C. E.: A probabilistic Pliocene–Pleistocene stack of benthic δ^{18}O using a profile hidden Markov model, Dynamics and Statistics of the Climate System, 2, 1–16, 2017.

Alder, J. R. and Hostetler, S. W.: Global climate simulations at 3000-year intervals for the last 21 000 years with the GENMOM coupled atmosphere–ocean model, Clim. Past, 11, 449–471, https://doi.org/10.5194/cp-11-449-2015, 2015.

Baggenstos, D., Haberli, M., Schmitt, J., Shackleton, S. A., Birner, B., Severinghaus, J. P., Kellerhals, T., and Fischer, H.: Earth's radiative imbalance from the Last Glacial Maximum to the present, P. Natl. Acad. Sci. USA, 116, 14881–14886, https://doi.org/10.1073/pnas.1905447116, 2019.

Barrientos, N., Lear, C. H., Jakobsson, M., Stranne, C., O'Regan, M., Cronin, T. M., Gukov, A. Y., and Coxall, H. K.: Arctic Ocean benthic foraminifera Mg / Ca ratios and global Mg / Ca-temperature calibrations: New constraints at low temperatures, Geochim. Cosmochim. Ac., 236, 240–259, https://doi.org/10.1016/j.gca.2018.02.036, 2018.

Bartlein, P.: pjbartlein/MOTvsSST: MOT vs SST regressions (v0.1), Zenodo [code], https://doi.org/10.5281/zenodo.14759006, 2025.

Bates, S. L., Siddall, M., and Waelbroeck, C.: Hydrographic variations in deep ocean temperature over the mid-Pleistocene transition, Quaternary Sci. Rev., 88, 147–158, https://doi.org/10.1016/j.quascirev.2014.01.020, 2014.

Bereiter, B., Shackleton, S., Baggenstos, D., Kawamura, K., and Severinghaus, J.: Mean global ocean temperatures during the last glacial transition, Nature, 553, 39–44, https://doi.org/10.1038/nature25152, 2018.

Bijma, J., Spero, H. J., and Lea, D. W.: Reassessing foraminiferal stable isotope geochemistry: Impact of the oceanic carbonate system (experimental results), in: Use of Proxies in Paleoceanography: Examples from the South Atlantic, edited by: Fischer, G. and Wefer, G., Springer-Verlag, Berlin, 489–512, https://doi.org/10.1007/978-3-642-58646-0_20, 1999.

Braconnot, P., Harrison, S. P., Kageyama, M., Bartlein, P. J., Masson-Delmotte, V., Abe-Ouchi, A., Otto-Bliesner, B., and Zhao, Y.: Evaluation of climate models using palaeoclimatic data, Nat. Clim. Change, 2, 417–424, https://doi.org/10.1038/nclimate1456, 2012.

Bronselaer, B. and Zanna, L.: Heat and carbon coupling reveals ocean warming due to circulation changes, Nature, 584, 227–233 , https://doi.org/10.1038/s41586-020-2573-5, 2020.

Cheng, L. J., von Schuckmann, K., Abraham, J. P., Trenberth, K. E., Mann, M. E., Zanna, L., England, M. H., Zika, J. D., Fasullo, J. T., Yu, Y. Q., Pan, Y. Y., Zhu, J., News, E. R., Bronselaer, B., and Lin, X. P.: Past and future ocean warming, Nature Reviews Earth & Environment, 3, 776–794, https://doi.org/10.1038/s43017-022-00345-1, 2022.

Church, J. A., Clark, P. U., Cazenave, A., Gregory, J. M., Jevrejeva, S., Levermann, A., Merrifield, M. A., Milne, G. A., Nerem, R. S., Nunn, P. D., Payne, A. J., Pfeffer, W. T., Stammer, D., and Unnikrishnan, A. S.: Sea Level Change, in: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Stocker, T. F., Qin, D., Plattner, G.-K., Tignor, M., Allen, S. K., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P. M., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1137–1216, https://doi.org/10.1017/9781009157896.011, 2013.

Clark, P. U., Shakun, J. D., Marcott, S. A., Mix, A. C., Eby, M., Kulp, S., Levermann, A., Milne, G. A., Pfister, P. L., Santer, B. D., Schrag, D. P., Solomon, S., Stocker, T. F., Strauss, B. H., Weaver, A. J., Winkelmann, R., Archer, D., Bard, E., Goldner, A., Lambeck, K., Pierrehumbert, R. T., and Plattner, G. K.: Consequences of twenty-first-century policy for multi-millennial climate and sea-level change, Nat. Clim. Change, 6, 360–369, https://doi.org/10.1038/nclimate2923, 2016.

Clark, P. U., Shakun, J. D., Rosenthal, Y., Köhler, P., and Bartlein, P. J.: Global and regional temperature change over the last 4.5 million years, Science, 383, 884–890, 2024.

Cleveland, W. S., Grosse, E., and Shyu, W. M.: Local regression models, in: Statistical Models in S, edited by: Chambers, J. M. and Hastie, T. J., Chapman and Hall, https://doi.org/10.1201/9780203738535, 1992.

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 (delta O-18 and Mg / Ca) with sea level history, J. Geophys. Res.-Oceans, 116, C12023, https://doi.org/10.1029/2011jc007255, 2011.

de Boer, B., Lourens, L. J., and van de Wal, R. S. W.: Persistent 400,000-year variability of Antarctic ice volume and the carbon cycle is revealed throughout the Plio-Pleistocene, Nat. Commun., 5, 2999, https://doi.org/10.1038/ncomms3999, 2014.

Dumitru, O. A., Austermann, J., Polyak, V. J., Fornos, J. J., Asmerom, Y., Gines, J., Gines, A., and Onac, B. P.: Constraints on global mean sea level during Pliocene warmth, Nature, 574, 233–236, https://doi.org/10.1038/s41586-019-1543-2, 2019.

Dumitru, O. A., Austermann, J., Polyak, V. J., Fornos, J. J., Asmerom, Y., Gines, J., Gines, A., and Onac, B. P.: Sea-level stands from the Western Mediterranean over the past 6.5 million years, Sci. Rep.-UK, 11, 6681, https://doi.org/10.1038/s41598-020-80025-6, 2021.

Dwyer, G. S. and Chandler, M. A.: Mid-Pliocene sea level and continental ice volume based on coupled benthic Mg / Ca palaeotemperatures and oxygen isotopes, Philos. T. Roy. Soc. A, 367, 157–168, https://doi.org/10.1098/rsta.2008.0222, 2009.

Elderfield, H., Ferretti, P., Greaves, M., Crowhurst, S., McCave, I. N., Hodell, D., and Piotrowski, A. M.: Evolution of ocean temperature and ice volume through the mid-Pleistocene climate transition, Science, 337, 704–709, https://doi.org/10.1126/science.1221294, 2012.

Emiliani, C.: Temperatures of Pacific bottom waters and polar superficial waters during the Tertiary, Science, 119, 853–855, 1954.

Evans, D., Brugger, J., Inglis, G. N., and Valdes, P.: The temperature of the deep ocean Is a robust proxy for global mean surface temperature during the Cenozoic, Paleoceanography and Paleoclimatology, 39, e2023PA004788, https://doi.org/10.1029/2023PA004788, 2024.

Farmer, J. R., Honisch, B., Haynes, L. L., Kroon, D., Jung, S., Ford, H. L., Raymo, M. E., Jaume-Segui, M., Bell, D. B., Goldstein, S. L., Pena, L. D., Yehudai, M., and Kim, J.: Deep Atlantic Ocean carbon storage and the rise of 100,000-year glacial cycles, Nat. Geosci., 12, 355–360, https://doi.org/10.1038/s41561-019-0334-6, 2019.

Ford, H. L. and Raymo, M. E.: Regional and global signals in seawater δ^{18}O records across the mid-Pleistocene transition, Geology, 48, 113–117, https://doi.org/10.1130/g46546.1, 2020.

Fox-Kemper, B., Hewitt, H., Xiao, C., Aðalgeirsdóttir, G., Drijfhout, S. S., Edwards, T. L., Golledge, N. R., Hemer, M., Kopp, R. E., Krinner, G., Mix, A., Notz, D., Nowicki, S., Nurhati, I. S., Ruiz, L., Sallée, J.-B., Slangen, A. B. A., and Yu, Y.: Ocean, Cryosphere, and Sea Level Change, in: Climate Change 2021: The Physical Science Basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, L., Goldfarb, L., Gomis, M. I., Huang, M., Leitzell, K., Lonnoy, E., Matthews, J. B. R., Maycock, T., Waterfield, T., Yelekçi, O., Yu, R., and Zhou, B., Cambriidge University Press, Cambridge, UK, 1211–1362, https://doi.org/10.1017/9781009157896.011, 2021.

Galbraith, E. and de Lavergne, C.: Response of a comprehensive climate model to a broad range of external forcings: relevance for deep ocean ventilation and the development of late Cenozoic ice ages, Clim. Dynam., 52, 653–679, https://doi.org/10.1007/s00382-018-4157-8, 2019.

Gasson, E., DeConto, R. M., and Pollard, D.: Modeling the oxygen isotope composition of the Antarctic ice sheet and its significance to Pliocene sea level, Geology, 44, 827–830, https://doi.org/10.1130/g38104.1, 2016.

Goudsmit-Harzevoort, B., Lansu, A., Baatsen, M. L. J., von der Heydt, A. S., de Winter, N. J., Zhang, Y. R., Abe-Ouchi, A., de Boer, A., Chan, W. L., Donnadieu, Y., Hutchinson, D. K., Knorr, G., Ladant, J. B., Morozova, P., Niezgodzki, I., Steinig, S., Tripati, A., Zhang, Z. S., Zhu, J., and Ziegler, M.: The Relationship Between the Global Mean Deep-Sea and Surface Temperature During the Early Eocene, Paleoceanography and Paleoclimatology, 38, e2022PA004532, https://doi.org/10.1029/2022pa004532, 2023.

Gray, W. R., de Lavergne, C., Wills, R. J. C., Menviel, L., Spence, P., Holzer, M., Kageyama, M., and Michel, E.: Poleward Shift in the Southern Hemisphere Westerly Winds Synchronous With the Deglacial Rise in CO_2, Paleoceanography and Paleoclimatology, 38, e2023PA004666, https://doi.org/10.1029/2023pa004666, 2023.

Gregory, J. M.: Vertical heat transports in the ocean and their effect on time-dependent climate change, Clim. Dynam., 16, 501–515, 2000.

Gregory, J. M., Stouffer, R. J., Raper, S. C. B., Stott, P. A., and Rayner, N. A.: An observationally based estimate of the climate sensitivity, J. Climate, 15, 3117–3121, https://doi.org/10.1175/1520-0442(2002)015<3117:Aobeot>2.0.Co;2, 2002.

Gregory, J. M., Bouttes, N., Griffies, S. M., Haak, H., Hurlin, W. J., Jungclaus, J., Kelley, M., Lee, W. G., Marshall, J., Romanou, A., Saenko, O. A., Stammer, D., and Winton, M.: The Flux-Anomaly-Forced Model Intercomparison Project (FAFMIP) contribution to CMIP6: investigation of sea-level and ocean climate change in response to CO_2 forcing, Geosci. Model Dev., 9, 3993–4017, https://doi.org/10.5194/gmd-9-3993-2016, 2016.

Gregory, J. M., Bloch-Johnson, J., Couldrey, M. P., Exarchou, E., Griffies, S. M., Kuhlbrodt, T., Newsom, E., Saenko, O. A., Suzuki, T., Wu, Q. R., Urakawa, S., and Zanna, L.: A new conceptual model of global ocean heat uptake, Clim. Dynam., 62, 1669–1713, https://doi.org/10.1007/s00382-023-06989-z, 2024.

Gu, S., Liu, Z., Oppo, D. W., Lynch-Stieglitz, J., Jahn, A., Zhang, J., and Wu, L.: Assessing the potential capability of reconstructing glacial Atlantic water masses and AMOC using multiple proxies in CESM, Earth Planet. Sc. Lett., 541, 116294, https://doi.org/10.1016/j.epsl.2020.116294, 2020.

Haeberli, M., Baggenstos, D., Schmitt, J., Grimmer, M., Michel, A., Kellerhals, T., and Fischer, H.: Snapshots of mean ocean temperature over the last 700 000 years using noble gases in the EPICA Dome C ice core, Clim. Past, 17, 843–867, https://doi.org/10.5194/cp-17-843-2021, 2021.

Hansen, J., Sato, M., Russell, G., and Kharecha, P.: Climate sensitivity, sea level and atmospheric carbon dioxide, Philos. T. Roy. Soc. A, 371, 20120294, https://doi.org/10.1098/rsta.2012.0294, 2013.

Hansen, J. E., Sato, M., Simons, L., Nazarenko, L. S., Sangha, I., Kharecha, P., Zachos, J. C., von Schuckmann, K., Loeb, N. G., Osman, M. B., Jin, Q., Tselioudis, G., Jeong, E., Lacis, A., Ruedy, R., Russell, G., Cao, J., and Li, J.: Global warming in the pipeline, Oxford Open Climate Change, 3, kgad008, https://doi.org/10.1093/oxfclm/kgad008, 2023.

Hastie, T. J.: Generalized additive models, in: Statistical Models in S, edited by: Chambers, J. M., and Hastie, T. J., Chapman and Hall, https://doi.org/10.1201/9780203738535, 1992.

Haywood, A. M., Tindall, J. C., Dowsett, H. J., Dolan, A. M., Foley, K. M., Hunter, S. J., Hill, D. J., Chan, W.-L., Abe-Ouchi, A., Stepanek, C., Lohmann, G., Chandan, D., Peltier, W. R., Tan, N., Contoux, C., Ramstein, G., Li, X., Zhang, Z., Guo, C., Nisancioglu, K. H., Zhang, Q., Li, Q., Kamae, Y., Chandler, M. A., Sohl, L. E., Otto-Bliesner, B. L., Feng, R., Brady, E. C., von der Heydt, A. S., Baatsen, M. L. J., and Lunt, D. J.: The Pliocene Model Intercomparison Project Phase 2: large-scale climate features and climate sensitivity, Clim. Past, 16, 2095–2123, https://doi.org/10.5194/cp-16-2095-2020, 2020.

He, C., Liu, Z., Otto-Bliesner, B. L., Brady, E. C., Zhu, C., Tomas, R., Clark, P. U., Zhu, J., Jahn, A., Gu, S., Zhang, J., Nusbaumer, J., Noone, D., Cheng, H., Wang, Y., Yan, M., and Bao, Y.: Hydroclimate footprint of pan-Asian monsoon water isotope during the last deglaciation, Science Advances, 7, eabe2611, https://doi.org/10.1126/sciadv.abe2611, 2021.

He, F.: Simulating transient climate evolution of the last deglaciation with CCSM3, The University of Wisconsin – Madison, https://www.researchgate.net/profile/Feng-He-27/publication/263618839_Simulating_transient_climate_evolution_of_the_last_deglaciation_with_CCSM3/links/0c96053b5fc619188f000000/Simulating-transient-climate-evolution-of-the-last-deglaciation-with-CCSM3.pdf (last access: 1 January 2025), 2011.

Heuzé, C., Heywood, K. J., Stevens, D. P., and Ridley, J. K.: Changes in global ocean bottom properties and volume transports in CMIP5 models under climate change scenarios, J. Climate, 28, 2917–2944, 2015.

Hodell, D. A. and Venz-Curtis, K. A.: Late Neogene history of deepwater ventilation in the Southern Ocean, Geochem. Geophy. Geosy., 7, Q09001, https://doi.org/10.1029/2005gc001211, 2006.

Jakob, K. A., Wilson, P. A., Pross, J., Ezard, T. H. G., Fiebig, J., Repschlager, J., and Friedrich, O.: A new sea-level record for the Neogene/Quaternary boundary reveals transition to a more stable East Antarctic Ice Sheet, P. Natl. Acad. Sci. USA, 117, 30980–30987, https://doi.org/10.1073/pnas.2004209117, 2020.

Johnson, G. C.: Quantifying Antarctic Bottom Water and North Atlantic Deep Water volumes, J. Geophys. Res., 113, C05027, https://doi.org/10.1029/2007JC004477, 2008.

Johnson, G. C., Sadman Mahmud, A. K. M., Macdonald, A. M., and Twining, B. S.: Antarctic Bottom Water warming, freshening, and contraction in the eastern Bellingshausen Basin, Geophys. Res. Lett., 51, e2024GL109937, https://doi.org/10.1029/2024GL109937, 2024.

Khatiwala, S., Primeau, F., and Holzer, M.: Ventilation of the deep ocean constrained with tracer observations and implications for radiocarbon estimates of ideal mean age, Earth Planet. Sc. Lett., 325–326, 116–125, 2012.

Kim, S. T. and O'Neil, J. R.: Equilibrium and nonequilibrium oxygen isotope effects in synthetic carbonates, Geochim. Cosmochim. Ac., 61, 3461–3475, https://doi.org/10.1016/s0016-7037(97)00169-5, 1997.

Köhler, P.: Atmospheric CO_2 concentration based on boron isotopes versus simulations of the global carbon cycle during the Plio-Pleistocene, Paleoceanography and Paleoclimatology, 38, e2022PA004439, https://doi.org/10.1029/2022PA004439, 2023.

Köhler, P. and Bintanja, R.: The carbon cycle during the Mid Pleistocene Transition: the Southern Ocean Decoupling Hypothesis, Clim. Past, 4, 311–332, https://doi.org/10.5194/cp-4-311-2008, 2008.

Köhler, P. and Mulitza, S.: No detectable influence of the carbonate ion effect on changes in stable carbon isotope ratios (δ^{13}C) of shallow dwelling planktic foraminifera over the past 160 kyr, Clim. Past, 20, 991–1015, https://doi.org/10.5194/cp-20-991-2024, 2024.

Kuhlbrodt, T. and Gregory, J. M.: Ocean heat uptake and its consequences for the magnitude of sea level rise and climate change, Geophys. Res. Lett., 39, L18608, https://doi.org/10.1029/2012gl052952, 2012.

Lambeck, K., Rouby, H., Purcell, A., Sun, Y., and Sambridge, M.: Sea level and global ice volumes from the Last Glacial Maximum to the Holocene, P. Natl. Acad. Sci. USA, 111, 15296–15303, 2014.

Lang, D. C., Bailey, I., Wilson, P. A., Chalk, T. B., Foster, G. L., and Gutjahr, M.: Incursions of southern-sourced water into the deep North Atlantic during late Pliocene glacial intensification, Nat. Geosci., 9, 375–379, https://doi.org/10.1038/ngeo2688, 2016.

Lear, C. H., Rosenthal, Y., and Wright, J. D.: The closing of a seaway: ocean water masses and global climate change, Earth Planet. Sc. Lett., 210, 425–436, https://doi.org/10.1016/s0012-821x(03)00164-x, 2003.

Lear, C. H., Billups, K., Rickaby, R. E. M., Diester-Haass, L., Mawbey, E. M., and Sosdian, S. M.: Breathing more deeply: Deep ocean carbon storage during the mid-Pleistocene climate transition, Geology, 44, 1035–1038, https://doi.org/10.1130/g38636.1, 2016.

Li, C., von Storch, J.-S., and Marotzke, J.: Deep-ocean heat uptake and equilibrium climate response, Clim. Dynam., 40, 1071–1086, 2013.

Li, Q., England, M. H., Hogg, A. M., Rintoul, S. R., and Morrison, A. K.: Abyssal ocean overturning slowdown and warming driven by Antarctic meltwater, Nature, 615, 841–847, 2023.

Lisiecki, L. E.: Atlantic overturning responses to obliquity and precession over the last 3 Myr, Paleoceanography, 29, 71–86, https://doi.org/10.1002/2013pa002505, 2014.

Lisiecki, L. E. and Raymo, M. E.: A Pliocene-Pleistocene stack of 57 globally distributed benthic δ^{18}O records, Paleoceanography, 20, 2004PA001071, https://doi.org/10.1029/2004PA001071, 2005.

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, 2014.

Martin, P. A., Lea, D. W., Rosenthal, Y., Papenfuss, T. P., and Sarnthein, M.: Late Quaternary Deep-Sea Temperatures Inferred From Benthic Foraminiferal Magnesium, Earth Planet. Sc. Lett., 198, 193–209, 2002.

Meckler, A. N., Sexton, P. F., Piasecki, A. M., Leutert, T. J., Marquardt, J., Ziegler, M., Agterhuis, T., Lourens, L. J., Rae, J. W. B., Barnet, J., Tripati, A., and Bernasconi, S. M.: Cenozoic evolution of deep ocean temperature from clumped isotope thermometry, Science, 377, 86–90, https://doi.org/10.1126/science.abk0604, 2022.

Miller, K. G., Wright, J. D., Browning, J. V., Kulpecz, A., Kominz, M., Naish, T. R., Cramer, B. S., Rosenthal, Y., Peltier, W. R., and Sosdian, S.: High tide of the warm Pliocene: Implications of global sea level for Antarctic deglaciation, Geology, 40, 407–410, https://doi.org/10.1130/g32869.1, 2012.

Miller, K. G., Browning, J. V., Schmelz, W. J., Kopp, R. E., Mountain, G. S., and Wright, J. D.: Cenozoic sea-level and cryospheric evolution from deep-sea geochemical and continental margin records, Science Advances, 6, eaaz1346, https://doi.org/10.1126/sciadv.aaz1346, 2020.

Muggeo, V. M. R.: Estimating regression models with unknown break-points, Stat. Med., 22, 3055–3071, 2003.

Nederbragt, A. J.: The effect of seawater carbonate chemistry on the stable isotope composition of Cibicidoides wuellerstorfi and other Cibicidoides species, Paleoceanography and Paleoclimatology, 38, e2023PA004667, https://doi.org/10.1029/2023PA004667, 2023.

Newsom, E., Zanna, L., and Gregory, J.: Background pycnocline depth constrains future ocean heat uptake ffficiency, Geophys. Res. Lett., 50, e2023GL105673, https://doi.org/10.1029/2023gl105673, 2023.

Pena, L. D. and Goldstein, S. L.: Thermohaline circulation crisis and impacts during the mid-Pleistocene transition, Science, 345, 318–322, https://doi.org/10.1126/science.1249770, 2014.

Pöppelmeier, F., Baggenstos, D., Grimmer, M., Liu, Z. J., Schmitt, J., Fischer, H., and Stocker, T. F.: The Effect of Past Saturation Changes on Noble Gas Reconstructions of Mean Ocean Temperature, Geophys. Res. Lett., 50, e2022GL102055, https://doi.org/10.1029/2022gl102055, 2023.

Purkey, S. G., Johnson, G. C., Talley, L. D., Sloyan, B. M., Wijffels, S. E., Smethie, W., Mecking, S., and Katsumata, K.: Unabated bottom water warming and freshening in the South Pacific Ocean, J. Geophys. Res.-Oceans, 124, 1778–1794, 2019.

R: A Language and Environment for Statistical Computing, The R Foundation, https://www.R-project.org/ (last access: 5 January 2025), 2024.

Raymo, M. E., Kozdon, R., Evans, D., Lisiecki, L., and Ford, H. L.: The accuracy of mid-Pliocene δ^{18}O-based ice volume and sea level reconstructions, Earth-Sci. Rev., 177, 291–302, https://doi.org/10.1016/j.earscirev.2017.11.022, 2018.

Rohling, E. J., Yu, J. M., Heslop, D., Foster, G. L., Opdyke, B., and Roberts, A. P.: Sea level and deep-sea temperature reconstructions suggest quasi-stable states and critical transitions over the past 40 million years, Science Advances, 7, eabf5326, https://doi.org/10.1126/sciadv.abf5326, 2021.

Rohling, E. J., Foster, G. L., Gernon, T. M., Grant, K. M., Heslop, D., Hibbert, F. D., Roberts, A. P., and Yu, J.: Comparison and synthesis of sea-level and deep-sea temperature variations over the past 40 million years, Rev. Geophys., 60, e2022RG000775, https://doi.org/10.1029/2022RG000775, 2022.

Rohling, E. J., Gernon, T. M., Heslop, D., Reichart, G. J., Roberts, A. P., and Yu, J.: Reconciling the Apparent Discrepancy Between Cenozoic Deep-Sea Temperatures From Proxies and From Benthic Oxygen Isotope Deconvolution, Paleoceanography and Paleoclimatology, 39, e2024PA004872, https://doi.org/10.1029/2024pa004872, 2024.

Rosenthal, Y., Bova, S., and Zhou, X. L.: A user guide for choosing planktic foraminiferal Mg / Ca-temperature calibrations, Paleoceanography and Paleoclimatology, 37, e2022PA004413, https://doi.org/10.1029/2022pa004413, 2022.

Rugenstein, M., Bloch-Johnson, J., Abe-Ouchi, A., Andrews, T., Beyerle, U., Cao, L., Chadha, T., Danabasoglu, G., Dufresne, J. L., Duan, L., Foujols, M. A., Frolicher, T., Geoffroy, O., Gregory, J., Knutti, R., Li, C., Marzocchi, A., Mauritsen, T., Menary, M., Moyer, E., Nazarenko, L., Paynter, D., Saint-Martin, D., Schmidt, G. A., Yamamoto, A., and Yang, S. T.: LongRunMIP: Motivation and Design for a Large Collection of Millennial-Length AOGCM Simulations, B. Am. Meteorol. Soc., 100, 2551–2570, https://doi.org/10.1175/bams-d-19-0068.1, 2019.

Rugenstein, M. A. A., Sedlácek, J., and Knutti, R.: Nonlinearities in patterns of long-term ocean warming, Geophys. Res. Lett., 43, 3380–3388, https://doi.org/10.1002/2016gl068041, 2016.

Saenko, O., Gregory, J. M., Griffies, S. M., Couldrey, M. P., and Dias, F. B.: Contribution of ocean physics and dynamics at different scales to heat uptaken in low-resolution AOGCMs, J. Climate, 34, 2017–2035, 2021.

Schrag, D. P.: Effects of diagenesis on the isotopic record of late Paleogene tropical sea surface temperatures, Chem. Geol., 161, 215–224, https://doi.org/10.1016/s0009-2541(99)00088-1, 1999.

Schrag, D. P., Hampt, G., and Murray, D. W.: Pore fluid constraints on the temperature and oxygen isotopic composition of the glacial ocean, Science, 272, 1930–1932, 1996.

Schrag, D. P., Adkins, J. F., McIntyre, K., Alexander, J. L., Hodell, D. A., Charles, C. D., and McManus, J. F.: The oxygen isotopic composition of seawater during the Last Glacial Maximum, Quaternary Sci. Rev., 21, 331–342, 2002.

Seltzer, A. M., Davidson, P. W., Shackleton, S. A., Nicholson, D. P., and Khatiwala, S.: Global ocean cooling of 2.3 °C during the Last Glacial Maximum, Geophys. Res. Lett., 51, e2024GL108866, https://doi.org/10.1029/2024GL108866, 2024.

Shackleton, S., Bereiter, B., Baggenstos, D., Bauska, T. K., Brook, E. J., Marcott, S. A., and Severinghaus, J. P.: Is the noble gas-based rate of ocean warming during the Younger Dryas overestimated?, Geophys. Res. Lett., 46, 5928–5936, https://doi.org/10.1029/2019gl082971, 2019.

Shackleton, S., Baggenstos, D., Menking, J. A., Dyonisius, M. N., Bereiter, B., Bauska, T. K., Rhodes, R. H., Brook, E. J., Petrenko, V. V., McConnell, J. R., Kellerhals, T., Haberli, M., Schmitt, J., Fischer, H., and Severinghaus, J. P.: Global ocean heat content in the Last Interglacial, Nat. Geosci., 13, 77–81, https://doi.org/10.1038/s41561-019-0498-0, 2020.

Shackleton, S., Menking, J. A., Brook, E., Buizert, C., Dyonisius, M. N., Petrenko, V. V., Baggenstos, D., and Severinghaus, J. P.: Evolution of mean ocean temperature in Marine Isotope Stage 4, Clim. Past, 17, 2273–2289, https://doi.org/10.5194/cp-17-2273-2021, 2021.

Shackleton, S., Seltzer, A., Baggenstos, D., and Lisiecki, L.: Benthic δ^{18}O records Earth's energy imbalance, Nat. Geosci., 16, 797–802, 2023.

Shakun, J. D., Clark, P. U., He, F., Marcott, S. A., Mix, A. C., Liu, Z. Y., Otto-Bliesner, B., Schmittner, A., and Bard, E.: Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation, Nature, 484, 49–55, 2012.

Shakun, J. D., Lea, D. W., Lisiecki, L. E., and Raymo, M. E.: An 800 kyr record of global surface ocean delta O-18 and implications for ice volume-temperature coupling, Earth Planet. Sc. Lett., 426, 58–68, https://doi.org/10.1016/j.epsl.2015.05.042, 2015.

Siddall, M., Honisch, B., Waelbroeck, C., and Huybers, P.: Changes in deep Pacific temperature during the mid-Pleistocene transition and Quaternary, Quaternary Sci. Rev., 29, 170–181, https://doi.org/10.1016/j.quascirev.2009.05.011, 2010.

Sosdian, S. and Rosenthal, Y.: Deep-sea temperature and ice volume changes across the Pliocene-Pleistocene climate transitions, Science, 325, 306–310, https://doi.org/10.1126/science.1169938, 2009.

Sosdian, S. M., Rosenthal, Y., and Toggweiler, J. R.: Deep Atlantic carbonate ion and $CaCO_3$ compensation during the Ice Ages, Paleoceanography and Paleoclimatology, 33, 546–562, https://doi.org/10.1029/2017pa003312, 2018.

Spero, H. J., Bijma, J., Lea, D. W., and Bemis, B. E.: Effect of seawater carbonate concentration on foraminiferal carbon and oxygen isotopes, Nature, 390, 497–500, 1997.

Spratt, R. M. and Lisiecki, L. E.: A Late Pleistocene sea level stack, Clim. Past, 12, 1079–1092, https://doi.org/10.5194/cp-12-1079-2016, 2016.

Tierney, J. E., Zhu, J., King, J., Malevich, S. B., Hakim, G. J., and Poulsen, C. J.: Glacial cooling and climate sensitivity revisited, Nature, 584, 569–573, https://doi.org/10.1038/s41586-020-2617-x, 2020.

Uemura, R., Motoyama, H., Masson-Delmotte, V., Jouzel, J., Kawamura, K., Goto-Azuma, K., Fujita, S., Kuramoto, T., Hirabayashi, M., Miyake, T., Ohno, H., Fujita, K., Abe-Ouchi, A., Iizuka, Y., Horikawa, S., Igarashi, M., Suzuki, K., Suzuki, T., and Fujii, Y.: Asynchrony between Antarctic temperature and CO_2 associated with obliquity over the past 720,000 years, Nat. Commun., 9, 961, https://doi.org/10.1038/s41467-018-03328-3, 2018.

von Schuckmann, K., Minière, A., Gues, F., Cuesta-Valero, F. J., Kirchengast, G., Adusumilli, S., Straneo, F., Ablain, M., Allan, R. P., Barker, P. M., Beltrami, H., Blazquez, A., Boyer, T., Cheng, L., Church, J., Desbruyeres, D., Dolman, H., Domingues, C. M., García-García, A., Giglio, D., Gilson, J. E., Gorfer, M., Haimberger, L., Hakuba, M. Z., Hendricks, S., Hosoda, S., Johnson, G. C., Killick, R., King, B., Kolodziejczyk, N., Korosov, A., Krinner, G., Kuusela, M., Landerer, F. W., Langer, M., Lavergne, T., Lawrence, I., Li, Y., Lyman, J., Marti, F., Marzeion, B., Mayer, M., MacDougall, A. H., McDougall, T., Monselesan, D. P., Nitzbon, J., Otosaka, I., Peng, J., Purkey, S., Roemmich, D., Sato, K., Sato, K., Savita, A., Schweiger, A., Shepherd, A., Seneviratne, S. I., Simons, L., Slater, D. A., Slater, T., Steiner, A. K., Suga, T., Szekely, T., Thiery, W., Timmermans, M.-L., Vanderkelen, I., Wjiffels, S. E., Wu, T., and Zemp, M.: Heat stored in the Earth system 1960–2020: where does the energy go?, Earth Syst. Sci. Data, 15, 1675–1709, https://doi.org/10.5194/essd-15-1675-2023, 2023.

Waelbroeck, C., Labeyrie, L., Michel, E., Duplessy, J. C., McManus, J. F., Lambeck, K., Balbon, E., and Labracherie, M.: Sea-level and deep water temperature changes derived from benthic foraminifera isotopic records, Quaternary Sci. Rev., 21, 295–305, 2002.

Weiffenbach, J. E., Dijkstra, H. A., von der Heydt, A. S., Abe-Ouchi, A., Chan, W.-L., Chandan, D., Feng, R., Haywood, A. M., Hunter, S. J., Li, X., Otto-Bliesner, B. L., Peltier, W. R., Stepanek, C., Tan, N., Tindall, J. C., and Zhang, Z.: Highly stratified mid-Pliocene Southern Ocean in PlioMIP2, Clim. Past, 20, 1067–1086, https://doi.org/10.5194/cp-20-1067-2024, 2024.

Westerhold, T., Marwan, N., Drury, A. J., Liebrand, D., Agnini, C., Anagnostou, E., Barnet, J. S. K., Bohaty, S. M., De Vleeschouwer, D., Florindo, F., Frederichs, T., Hodell, D. A., Holbourn, A. E., Kroon, D., Lauretano, V., Littler, K., Lourens, L. J., Lyle, M., Palike, H., Rohl, U., Tian, J., Wilkens, R. H., Wilson, P. A., and Zachos, J. C.: An astronomically dated record of Earth's climate and its predictability over the last 66 million years, Science, 369, 1383–1387, https://doi.org/10.1126/science.aba6853, 2020.

Winnick, M. J. and Caves, J. K.: Oxygen isotope mass-balance constraints on Pliocene sea level and East Antarctic Ice Sheet stability, Geology, 43, 879–882, https://doi.org/10.1130/g36999.1, 2015.

Woodard, S. C., Rosenthal, Y., Miller, K. G., Wright, J. D., Chiu, B. K., and Lawrence, K. T.: Antarctic role in Northern Hemisphere glaciation, Science, 346, 847–851, https://doi.org/10.1126/science.1255586, 2014.

Zachos, J., Pagani, M., Sloan, L., Thomas, E., and Billups, K.: Trends, rhythms, and aberrations in global climate 65 Ma to present, Science, 292, 686–693, https://doi.org/10.1126/science.1059412, 2001.

Zanna, L., Khatiwala, S., Gregory, J. M., Ison, J., and Heimbach, P.: Global reconstruction of historical ocean heat storage and transport, P. Natl. Acad. Sci. USA, 116, 1126–1131, https://doi.org/10.1073/pnas.1808838115, 2019.

Zeebe, R. E.: An explanation of the effect of seawater carbonate concentration on foraminiferal oxygen isotopes, Geochim. Cosmochim. Ac., 63, 2001–2007, https://doi.org/10.1016/s0016-7037(99)00091-5, 1999.

Zhou, S., Meijers, A. J. S., Meredith, M. P., Abrahamsen, E. P., Holland, P. R., Silvano, A., Sallée, J. B., and Østerhus, S.: Slowdown of Antarctic bottom water export driven by climatic wind and sea-ice change, Nat. Clim. Change, 13, 701–709, 2023.

Zhu, C., Sanchez, S., Liu, Z., Clark, P. U., He, C., Wan, L., Lu, J., Zhu, C., Li, L., Zhang, S., and Cheng, L.: Enhanced ocean heat storage efficiency during the last deglaciation, Science Advances, 10, eadp55156, https://doi.org/10.1088/1748-9326/ad2b2b, 2024.

Zhu, C. Y., Liu, Z. Y., Zhang, S. Q., and Wu, L. X.: Global oceanic overturning circulation forced by the competition between greenhouse gases and continental ice Sheets during the last deglaciation, J. Climate, 34, 7555–7570, https://doi.org/10.1175/jcli-d-21-0125.1, 2021.

Supplement of Clim. Past, 21, 973–1000, 2025
https://doi.org/10.5194/cp-21-973-2025-supplement
© Author(s) 2025. CC BY 4.0 License.

Supplement of

Mean ocean temperature change and decomposition of the benthic δ^{18}O record over the past 4.5 million years

Peter U. Clark et al.

Correspondence to: Peter U. Clark (clarkp@onid.orst.edu)

- cp-21-973-2025-supplement-title-page.pdf
- Clark_MOT_Draft_SM_v3.docx
- Figure S1.pdf
- Figure S2.pdf
- Figure S3.pdf
- Table S1.xlsx
- Table S2.xlsx

The copyright of individual parts of the supplement might differ from the article licence.

S1. Statistical models for climate model output

We obtain ΔMOTs and ΔGMSSTs from 93 coupled ocean-atmosphere climate model equilibrium simulations that were run with atmospheric CO_2 concentrations ranging from 180 ppm to nine times PI (Fig. 2) (Alder and Hostetler, 2014; Galbraith and De Lavergne, 2019; Haywood et al., 2020; Clark et al., 2016; Goudsmit-Harzevoort et al., 2023; Rugenstein et al., 2019; He, 2011; Bereiter et al., 2018). Here we describe the relationship between ΔMOT and ΔGMSST in this synthesis of climate-model output and four regression models used to describe it. The four models were fit using R *stats* package functions *lm*(), *segmented*(), *loess*(), and *smoothing.spline*(), and their associated functions for prediction and diagnostic analyses (R, 2024). R-squared values are given by 1.0 minus the residual sum of squares divided by the total sum of squares. Adjusted R-squared values are calculated as 1.0 minus the fraction with the residual sum of squares divided by n minus p in the numerator and the total sum of squares divided by n minus 1.0 in the denominator, where p is equal to the number of parameters or equivalent degrees of freedom. The adjusted R-squared value, which is lower than the usual R-squared value, thereby includes a penalty for the number of adjustable parameters.

The relationship between ΔMOT and ΔGMSST is broadly linear, with a slope and intercept close to 1.0 and 0.0, respectively (Fig. 2A), a high R^2 value (0.924), highly significant *F*-test statistic (1104, $p < 0.000$) and a relatively low residual standard error (1.184). However, a residual diagnostic analysis of this model reveals a subtle but significant nonlinear lack of fit which can be seen in the residual scatter diagram (Fig. S1B). The nonlinearity in the residual scatter diagram is largely related to a quasi-linear cluster of points that plot below the main cloud of points (Fig. 2A), which include the points with the 10-largest negative residuals. The models that produced these data points are generally a mix of higher-than-present CO_2 simulations and sensitivity tests, but not all such experiments produce that pattern of large negative residuals. This eliminates the possibility that there is an underlying “family-of-curves” explanation for the apparent nonlinearity. The residuals are also not normally distributed (Shapiro-Wilk test $W = 0.932$, $p < 0.001$), and in fact are visibly bimodal, imparted by the cluster of negative residuals. Large positive residuals also occur for larger values of ΔMOT and ΔGMSST, and those values also have the largest values of Cook’s Distance, a measure of influence on the fitted regression (Cook, 1977).

Inspection of the scatter diagrams suggests that an alternative to an overall straight-line fit might be a segmented linear fit (Muggeo, 2003), with the individual segments approximately defined by values of ΔGMSST less than 1.0 °C, between 1.0 and 5.0 or 6.0 °C, and greater than 6.0 °C. To test that notion, we used the R *segmented*()

function, using 1 to 3 starting breakpoints that ranged between -2.0 to 10.0 °C. The best fit (in terms of goodness-of-fit and residual diagnostics) was a model with two breakpoints, at 1.187 (*s.e.* = 1.307) and 5.290 (*s.e.* = 0.855), and slopes for the lower, middle, and upper segments of 1.000 (*s.e.* = 0.178), 0.575 (*s.e.* = 0.185), and 1.244 (*s.e.* = 0.066), respectively (Fig. S1C). The model fit is slightly better than that for the linear one, with an adjusted R^2 value of 0.939 (vs. 0.924 for the linear one). Although the adjusted R^2 values do not differ much, the Akaike Information Criterion (*AIC*), a measure of the tradeoff between the goodness of fit and the number of adjustable parameters (smaller values are better), is lower here (*AIC* = 281.00) than for the original model (*AIC* = 299.31) despite the larger number of parameters. More importantly, the residual scatter plot shows a smaller indication of non-linearity (Fig. S1D) when compared to the residual scatter plot of the linear fit (Fig. S1B). The distribution of the residuals is still not normal, being a little heavy in the tails, but it is symmetric and shows no signs of bimodality. There are fewer points with extreme values of Cook's distance.

The segmented regression implies that there are distinct thresholds (the breakpoints) where the relationship between ΔMOT and ΔGMSST changes. An alternative model that can accommodate smooth variations in slope can be fit via local- or nonparametric-regression approaches. We also considered polynomial models, but these were inferior overall. The third model we considered was fit using the R *loess*() function (Cleveland et al., 1992). The best overall local regression model had a smoothing parameter (*span*) of 0.8, and polynomial degree of 2. The fitted values (Fig. S1E) were broadly similar to the individual slopes in the segmented regression, but the slope varied smoothly instead of abruptly at ΔGMSST values around 1.0 and 6.0 °C. This model fits better than the two previous ones (*AIC* = 277.13; adjusted R^2 = 0.940), and like the segmented-regression model, the residual scatter diagram (Fig. S1F) shows little evidence of nonlinearity and has a similar distribution to the residuals of the segmented-regression model.

Inspection of the loess curve shows smaller-scale variations in slope in regions of the scatter diagram where data are sparse, and increasing the *span* value did not remove those. This is not unexpected given the nature of local regression. Consequently, we considered a fourth model, a smoothing spline (Hastie, 1992). Smoothing splines are often used as the smoothing function in fitting generalized additive models and in some ways can be seen as a less local, more global, approach than a loess smoother. Smoothing splines require the specification of a smoothing parameter (*spar*). Here we examined a range of *spar* values between 0.80 and 1.00, and selected a value of 0.95 as optimal, using the tradeoff between the goodness of fit and the tendency for low values of *spar* to allow the fitted values to "chase" individual data points. This model fit the data best of all (*AIC* = 272.18; adjusted R^2 = 0.944). The fitted values (Fig. S1G) are quite similar to those of the loess local regression model (Fig. S1E). The residual scatter diagram (Fig. S1H) shows little sign of nonlinearity, and again the distribution of the residuals is approximately normal. The residual scatter diagrams of the segmented-regression, loess, and smoothing-spline regressions are quite similar overall (Figs. S1D, S1F, and S1H).

Among the three alternatives to the linear model, the smoothing-spline model has several desirable properties—it fits the data best, with the lowest *AIC* and highest adjusted R^2 (although the differences are quite small), and is quite smooth while still displaying the changes in slope explicitly represented by the segmented regression model. In addition, its associated *predict*() function can be used to extract the local numeric first derivative of the fitted curve, which provides a representation of the HSE. The values of HSE are plotted as the red curve in Figs. 2D and S1G. Given these properties, the smoothing spline regression is our preferred choice for representing the relationship between ΔMOT and ΔGMSST in our synthesis of climate-model results.

S2. Mg/Ca-based bottom water temperature reconstructions

For ODP sites 806 and 926, we use the original Mg/Ca data from Lear et al. (2003) and the recent calibrations for *C. wuellerstorfi* (Mg/Ca=1.043Exp(0.118BWT)) and *O. umbonatus* (Mg/Ca= 1.317Exp(0.102BWT)) published by Barrientos et al. (2018). We note that temperature estimates derived from *O. umbonatus* using the new exponential calibration are consistent with those calculated using the calibration of Rathmann et al. (2004).

We use the original published ΔBWT reconstruction from ODP site 607 measured on ostracodes by Dwyer and Chandler (2009) which is uncorrected for changes in seawater Mg/Ca. The original Mg/Ca-based ΔBWT record measured on benthic foraminifera from site 607 (Sosdian and Rosenthal, 2009) includes measurements on *C. wuellerstorfi* and *O. umbonatus* using the published calibration Mg/Ca= 0.15*T+1.16, which we update with the exponential calibration of Barrientos et al. (2018). Since the original record is arguably overprinted by the carbonate ion (ΔCO_3^{2-}) effect (Yu and Broecker, 2010), we have added Mg/Ca measurements of *Uvigerina sp.* (Sosdian and Rosenthal, 2010; Ford et al., 2016), an infaunal benthic foraminifera which has been shown to be insensitive to ΔCO_3^{2-} effects (Elderfield et al., 2010; Elderfield et al., 2012). We calculate *Uvigerina sp.* BWTs using the Elderfield et al. (2012) calibration (Mg/Ca=0.1*T+0.94). BWTs from site U1313 (Jakob et al., 2020) were recalculated using the Barrientos et al. (2018) calibration. The new compilation presented here, including new Mg/Ca data from *Uvigerina sp.,* demonstrates that the long-term trends are similar for all three species both in the raw Mg/Ca data and estimated BWTs, lending support to the published BWT records (Ford et al., 2016; Sosdian and Rosenthal, 2009). BWTs from site U1313 (Jakob et al., 2020) are based on *O. umbonatus* Mg/Ca and have been also recalculated using the calibration from Barrientos et al. (2018). The warmer Pliocene BOT at these sites are consistent with independent estimates, based on clumped isotope measurements of benthic foraminifera at the nearby IODP site U1308 (Braaten et al., 2023).

When reconstructing long-term temperature changes (>1 Ma) from Mg/Ca, it is necessary to account for variations in the seawater Mg/Ca concentration ratio. For the Mg/Ca data from the sites listed above, we follow the correction proposed by Evans and Müller (2012) relating the Mg/Ca of the calcareous shells to the Mg/Ca ratio of seawater through a power equation, where we use the seawater Mg/Ca record from Rosenthal et al. (2022) to correct measured foraminiferal Mg/Ca:

$$(Mg/Ca)_{corr} = (Mg/Ca)_{foram}{}^{*}(5.3/(Mg/Ca)_{sw})^{H}$$

where $(Mg/Ca)_{sw}$ is the seawater ratio at the studied time, 5.3 is the modern seawater Mg/Ca ratio, and H is the species-specific power coefficient. We apply the $(Mg/Ca)_{corr}$ to the calibrations for benthic foraminifera and a power coefficient H=0.1 to correct for seawater changes in Mg/Ca, although choosing values anywhere between 0 and 0.4 would have a negligible effect on Pleistocene BWT. The correction for changes in seawater Mg/Ca for the benthic compilations is only about +1°C. We note that the choice of seawater Mg/Ca or the calibration equation has minimal effect on the temperature reconstructions for the past 5 Myr (Meinicke et al., 2021; Rosenthal et al., 2022).

The smoothed BWT reconstruction from Cramer et al. (2011) is based on a compilation of Mg/Ca records derived from six species or genera of benthic foraminifera. Because *O. umbonatus* is present throughout the study interval (0-60 Ma), Cramer et al. (2011) inferred Mg/Ca offsets between *O. umbonatus* and the other taxa and then applied two published Mg/Ca calibrations derived from core-top *O. umbonatus* to the entire data set (their equations 7a and 7b). It is not clear why the calibrations differ, but they lead to significant differences in the estimated amount and rate of cooling over the last 10 Myr. We use the equation 7b reconstruction in Cramer et al. (2011) based on Rathmann et al. (2004), since it is supported by the newer calibration of Barrientos et al. (2018).

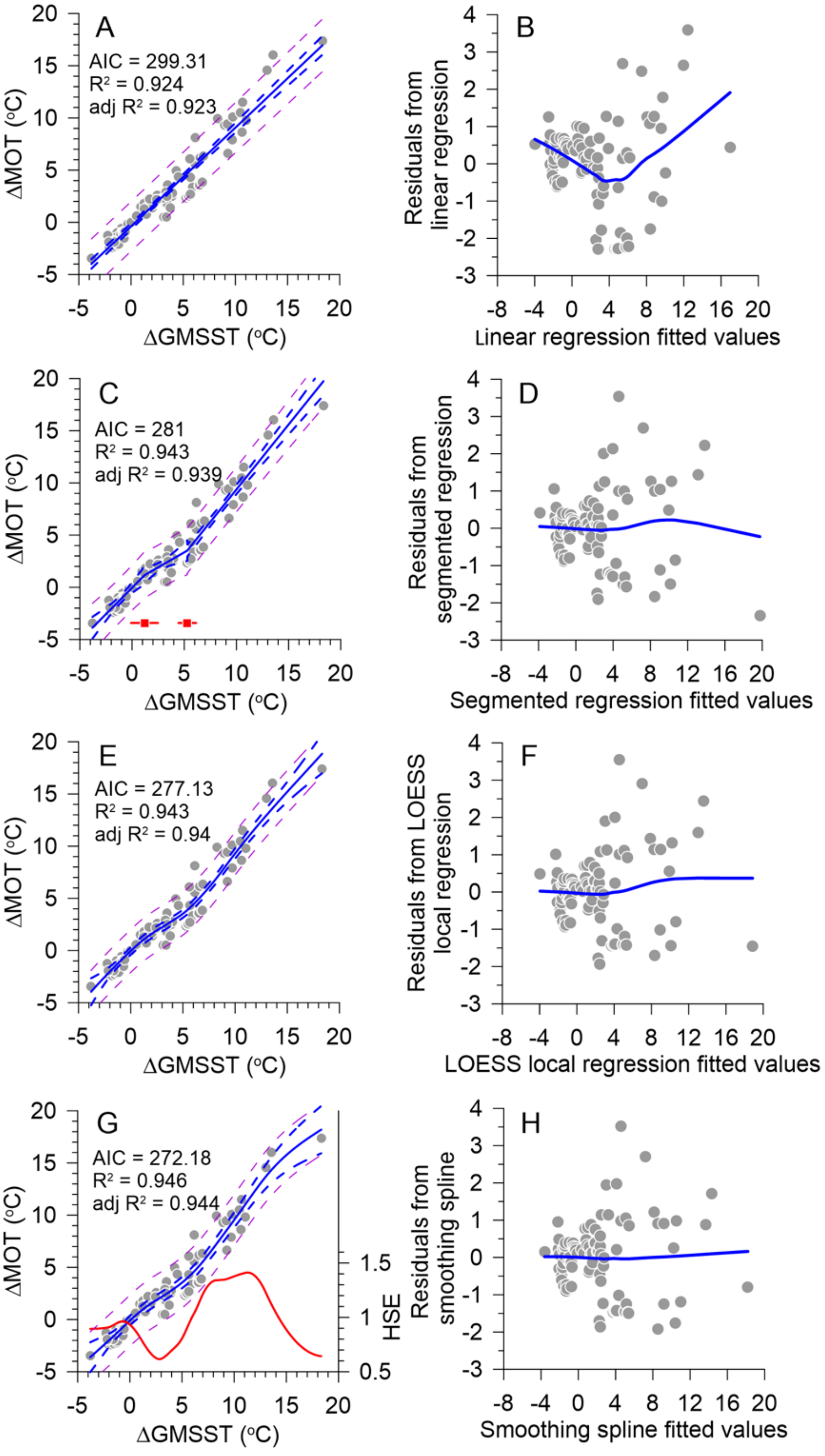
A
AIC = 299.31
R² = 0.924
adj R² = 0.923
ΔMOT (°C)
ΔGMSST (°C)
B
Residuals from linear regression
Linear regression fitted values
C
AIC = 281
R² = 0.943
adj R² = 0.939
ΔMOT (°C)
ΔGMSST (°C)
D
Residuals from segmented regression
Segmented regression fitted values
E
AIC = 277.13
R² = 0.943
adj R² = 0.94
ΔMOT (°C)
ΔGMSST (°C)
F
Residuals from LOESS local regression
LOESS local regression fitted values
G
AIC = 272.18
R² = 0.946
adj R² = 0.944
ΔMOT (°C)
HSE
ΔGMSST (°C)
H
Residuals from smoothing spline
Smoothing spline fitted values
-5
0
5
10
15
20
-8
-4
0
4
8
12
16
20
-3
-2
-1
0
1
2
3
4
0.5
1
1.5

Figure S1. Scatter diagrams and fitted curves showing the relationship between ΔMOT and ΔGMSST from the climate-model output synthesis for four statistical models (left column, panels **A**, **C**, **E**, and **G**) and residual scatter diagrams for each model (right column, panels **B**, **D**, **F**, and **H**). Data shown here differ from Figure 2 in showing the temperature range extending to 20ºC. Results shown are from PMIP3 model runs for LGM boundary conditions (Braconnot et al., 2012), model runs for CO_2 ≤270 ppm with no ice sheets and with LGM ice sheets (Galbraith and De Lavergne, 2019), model runs for the last deglaciation (Alder and Hostetler, 2015), model runs for CO_2 ≥420 ppm (Clark et al., 2016), LongRunMIP model runs for CO_2 ≥2x PI (Rugenstein et al., 2019), model runs for CO_2 ranging from 1x to 9x PI (Goudsmit-Harzevoort et al., 2023), PlioMIP2 model runs with CO_2 = 400 ppm (Haywood et al., 2020), and model runs for CO_2 ≥405 ppm (Galbraith and De Lavergne, 2019). In the left column of panels, the solid blue lines show the fitted values, the dashed blue lines are the 2σ confidence intervals of the fitted values (illustrating the uncertainty in the curves), and the dashed gray lines are the 2σ prediction intervals (illustrating the variability of the data about the fitted curves). In the right column of panels, the values plotted are the ordinary residuals and the fitted values from each model, while the blue lines are loess curves (with span = 0.90 and degree = 1) that summarize the pattern in the residual scatter diagrams. The red dots and lines in panel **C** show the estimated breakpoints in the segmented regression (dots) and their uncertainties. The red line in panel **G** shows the calculated first derivative of the fitted values, i.e. the heat storage efficiency (HSE = ΔMOT/ΔGMSST). R-squared values are given by 1.0 minus the residual sum of squares divided by the total sum of squares. Adjusted R-squared values are calculated as 1.0 minus the fraction with the residual sum of squares divided by n minus p in the numerator and the total sum of squares divided by n minus 1.0 in the denominator, where p is equal to the number of parameters or equivalent degrees of freedom. The adjusted R-squared value, which is lower than the usual R-squared value, thereby includes a penalty for the number of adjustable parameters.

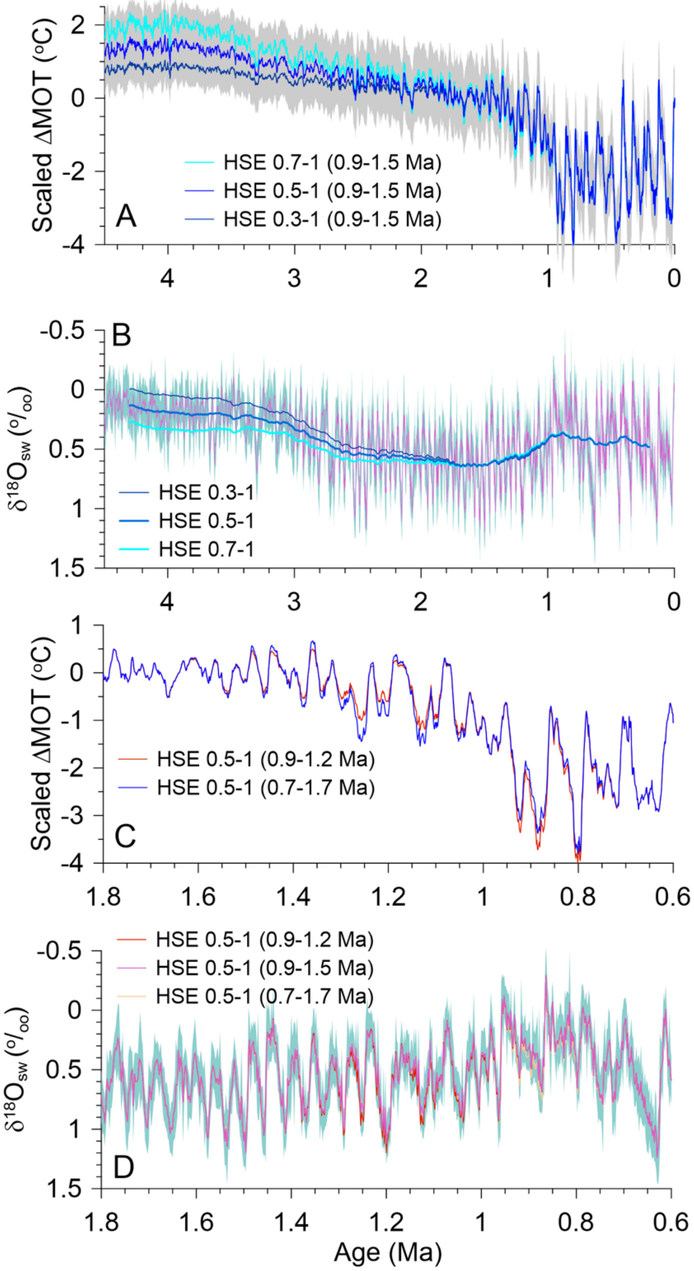

Figure S2. **A**. ΔMOT reconstructions based on HSE = 0.7 and HSE = 0.3 (reproduces Figure 5B) fall within the 1σ uncertainty on our ΔMOT reconstruction based on HSE = 0.5. **B.** Our high-resolution $\delta^{18}O_{sw}$ (violet) reconstruction with 1σ uncertainty compared to long-term (401-kyr running average) $\delta^{18}O_{sw}$ for the three HSE scenarios shown in **A**. The differences during the early Pleistocene from our preferred 0.5-1 HSE scenario are small (<0.1 per mil) and the high-resolution $\delta^{18}O_{sw}$ for the two bracketing HSE scenarios fall within the 1σ uncertainty of high-resolution $\delta^{18}O_{sw}$ based on 0.5-1 HSE scenario. Panels G and H address the timing question by comparing **C.** Sensitivity of ΔMOT in our preferred scenario (HSE increased from 0.5 to 1 between 1.5 Ma and 0.9 Ma) to one scenario where HSE increased between 1.2 Ma and 0.9 Ma and another where it increased from 1.7 Ma to 0.7 Ma. **D.** Sensitivity of our $\delta^{18}O_{sw}$ reconstruction to different durations of the HSE increase from 0.5 to 1.

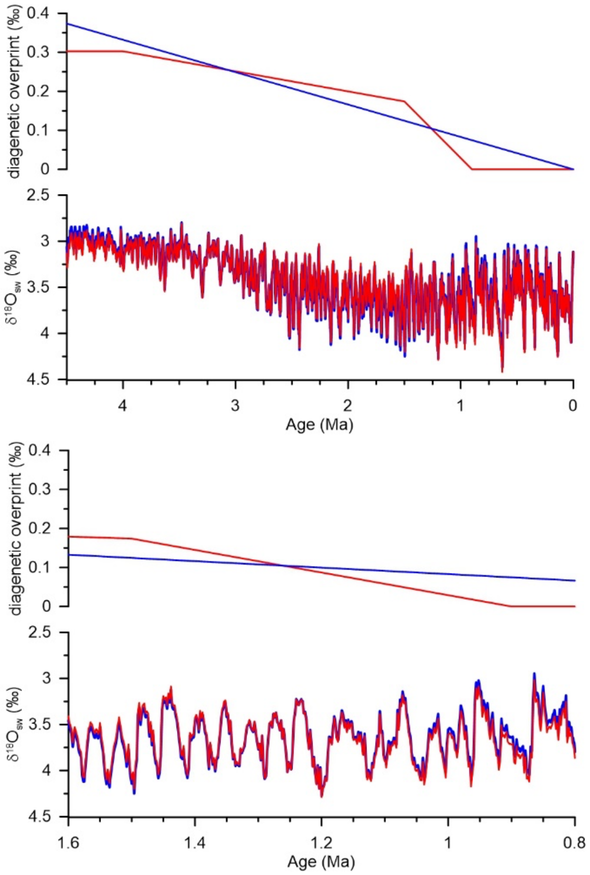

Figure S3. Sensitivity of our $\delta^{18}O_{sw}$ reconstruction (blue lines) to the same magnitude of our diagenetic correction (0.375‰) as used for that reconstruction but that tracks our ΔMOT reconstruction (red lines). The faster rate of change across the MPT slightly increases the trend in $\delta^{18}O_{sw}$ over this interval (more depleted before 1 Ma, more enriched after 1 Ma), but the changes are insignificant (≤ 0.1‰).

References

Alder, J. and Hostetler, S. W.: Global climate simulations at 3000 year intervals for the last 21 000 years with the GENMOM coupled atmosphere-ocean model, Climate of the Past Discussions, 10, 2925-2978, 2014.

Alder, J. R. and Hostetler, S. W.: Global climate simulations at 3000-year intervals for the last 21 000 years with the GENMOM coupled atmosphere-ocean model, Climate of the Past, 11, 449-471, 10.5194/cp-11-449-2015, 2015.

Barrientos, N., Lear, C. H., Jakobsson, M., Stranne, C., O'Regan, M., Cronin, T. M., Gukov, A. Y., and Coxall, H. K.: Arctic Ocean benthic foraminifera Mg/Ca ratios and global Mg/Ca-temperature calibrations: New constraints at low temperatures, Geochimica et Cosmochimica Acta, 236, 240-259, 10.1016/j.gca.2018.02.036, 2018.

Bereiter, B., Shackleton, S., Baggenstos, D., Kawamura, K., and Severinghaus, J.: Mean global ocean temperatures during the last glacial transition, Nature, 553, 39-44, 10.1038/nature25152, 2018.

Braaten, A. H., Jakob, K. A., Ho, S. L., Friedrich, O., Galaasen, E. V., De Schepper, S., Wilson, P. A., and Meckler, A. N.: Limited exchange between the deep Pacific and Atlantic oceans during the warm mid-Pliocene and Marine Isotope Stage M2 “glaciation”, Clim. Past, 19, 2109-2125, 10.5194/cp-19-2109-2023, 2023.

Braconnot, P., Harrison, S. P., Kageyama, M., Bartlein, P. J., Masson-Delmotte, V., Abe-Ouchi, A., Otto-Bliesner, B., and Zhao, Y.: Evaluation of climate models using palaeoclimatic data, Nature Climate Change, 2, 417-424, 10.1038/nclimate1456, 2012.

Clark, P. U., Shakun, J. D., Marcott, S. A., Mix, A. C., Eby, M., Kulp, S., Levermann, A., Milne, G. A., Pfister, P. L., Santer, B. D., Schrag, D. P., Solomon, S., Stocker, T. F., Strauss, B. H., Weaver, A. J., Winkelmann, R., Archer, D., Bard, E., Goldner, A., Lambeck, K., Pierrehumbert, R. T., and Plattner, G. K.: Consequences of twenty-first-century policy for multi-millennial climate and sea-level change, Nature Climate Change, 6, 360-369, 10.1038/nclimate2923, 2016.

Cleveland, W. S., Grosse, E., and Shyu, W. M.: Local regression models, in: Statistical Models in S, edited by: Chambers, J. M., and Hastie, T. J., Chapman and Hall, https://doi.org/10.1201/9780203738535, 1992.

Cook, R. D.: Detection of influential observations in linear regression, Technometrics, 19, 10.2307/1268249, 1977.

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 (delta O-18 and Mg/Ca) with sea level history, Journal of Geophysical Research-Oceans, 116, 10.1029/2011jc007255, 2011.

Dwyer, G. S. and Chandler, M. A.: Mid-Pliocene sea level and continental ice volume based on coupled benthic Mg/Ca palaeotemperatures and oxygen isotopes, Philosophical Transactions of the Royal Society A-Mathematical Physical and Engineering Sciences, 367, 157-168, 10.1098/rsta.2008.0222, 2009.

Elderfield, H., Ferretti, P., Greaves, M., Crowhurst, S., McCave, I. N., Hodell, D., and Piotrowski, A. M.: Evolution of ocean temperature and ice volume through the mid-Pleistocene climate transition, Science, 337, 704-709, 10.1126/science.1221294, 2012.

Elderfield, H., Greaves, M., Barker, S., Hall, I. R., Tripati, A., Ferretti, P., Crowhurst, S., Booth, L., and Daunt, C.: A record of bottom water temperature and seawater delta(18)O for the Southern Ocean over the past 440 kyr based on Mg/Ca of benthic foraminiferal Uvigerina spp, Quaternary Science Reviews, 29, 160-169, 10.1016/j.quascirev.2009.07.013, 2010.

Evans, D. and Müller, W.: Deep time foraminifera Mg/Ca paleothermometry: Nonlinear correction for secular change in seawater Mg/Ca, Paleoceanography, 27, PA4205, 10.1029/2012pa002315, 2012.

Ford, H. L., Sosdian, S. M., Rosenthal, Y., and Raymo, M. E.: Gradual and abrupt changes during the Mid-Pleistocene Transition, Quaternary Science Reviews, 148, 222-233, https://doi.org/10.1016/j.quascirev.2016.07.005, 2016.

Galbraith, E. and de Lavergne, C.: Response of a comprehensive climate model to a broad range of external forcings: relevance for deep ocean ventilation and the development of late Cenozoic ice ages, Climate Dynamics, 52, 653-679, 10.1007/s00382-018-4157-8, 2019.

Goudsmit-Harzevoort, B., Lansu, A., Baatsen, M. L. J., von der Heydt, A. S., de Winter, N. J., Zhang, Y. R., Abe-Ouchi, A., de Boer, A., Chan, W. L., Donnadieu, Y., Hutchinson, D. K., Knorr, G., Ladant, J. B., Morozova, P., Niezgodzki, I., Steinig, S., Tripati, A., Zhang, Z. S., Zhu, J., and Ziegler, M.: The Relationship Between the Global Mean Deep-Sea and Surface Temperature During the Early Eocene, Paleoceanography and Paleoclimatology, 38, 10.1029/2022pa004532, 2023.

Hastie, T. J.: Generalized additive models, in: Statistical Models in S, edited by: Chambers, J. M., and Hastie, T. J., Chapman and Hall, 10.1201/9780203738535, 1992.

Haywood, A. M., Tindall, J. C., Dowsett, H. J., Dolan, A. M., Foley, K. M., Hunter, S. J., Hill, D. J., Chan, W. L., Abe-Ouchi, A., Stepanek, C., Lohmann, G., Chandan, D., Peltier, W. R., Tan, N., Contoux, C., Ramstein, G., Li, X. Y., Zhang, Z. S., Guo, C. C., Nisancioglu, K. H., Zhang, Q., Li, Q., Kamae, Y., Chandler, M. A., Sohl, L. E., Otto-Bliesner, B. L., Feng, R., Brady, E. C., von der Heydt, A. S., Baatsen, M. L. J., and Lunt, D. J.: The Pliocene Model Intercomparison Project Phase 2: large-scale climate features and climate sensitivity, Climate of the Past, 16, 2095-2123, 10.5194/cp-16-2095-2020, 2020.

He, F.: Simulating transient climate evolution of the last deglaciation with CCSM3, The University of Wisconsin - Madison, 2011.

Jakob, K. A., Wilson, P. A., Pross, J., Ezard, T. H. G., Fiebig, J., Repschlager, J., and Friedrich, O.: A new sea-level record for the Neogene/Quaternary boundary reveals transition to a more stable East Antarctic Ice Sheet, Proceedings of the National Academy of Sciences of the United States of America, 117, 30980-30987, 10.1073/pnas.2004209117, 2020.

Lear, C. H., Rosenthal, Y., and Wright, J. D.: The closing of a seaway: ocean water masses and global climate change, Earth and Planetary Science Letters, 210, 425-436, 10.1016/s0012-821x(03)00164-x, 2003.

Meinicke, N., Reimi, M. A., Ravelo, A. C., and Meckler, A. N.: Coupled Mg/Ca and clumped isotope measurements indicate lack of substantial mixed layer cooling in the western Pacific warm pool during the last 5 Million years, Paleoceanography and Paleoclimatology, 36, e2020PA004115, 10.1029/2020pa004115, 2021.

Muggeo, V. M. R.: Estimating regression models with unknown break-points, Statistics in Medicine, 22, 3055-3071, 2003.

R: A Language and Environment for Statistical Computing: https://www.R-project.org/, last

Rathmann, S., Hess, S., Kuhnert, H., and Mulitza, S.: Mg/Ca ratios of the benthic foraminifera <i>Oridorsalis umbonatus</i> obtained by laser ablation from core top sediments:: Relationship to bottom water temperature -: art. no. Q12013, Geochemistry Geophysics Geosystems, 5, 10.1029/2004gc000808, 2004.

Rosenthal, Y., Bova, S., and Zhou, X. L.: A user guide for choosing planktic foraminiferal Mg/Ca-temperature calibrations, Paleoceanography and Paleoclimatology, 37, e2022PA004413, 10.1029/2022pa004413, 2022.

Rugenstein, M., Bloch-Johnson, J., Abe-Ouchi, A., Andrews, T., Beyerle, U., Cao, L., Chadha, T., Danabasoglu, G., Dufresne, J. L., Duan, L., Foujols, M. A., Frolicher, T., Geoffroy, O., Gregory, J., Knutti, R., Li, C., Marzocchi, A., Mauritsen, T., Menary, M., Moyer, E., Nazarenko, L., Paynter, D., Saint-Martin, D., Schmidt, G. A., Yamamoto, A., and Yang, S. T.: LongRunMIP: Motivation and Design for a Large Collection of Millennial-Length AOGCM Simulations, Bulletin of the American Meteorological Society, 100, 2551-2570, 10.1175/bams-d-19-0068.1, 2019.

Sosdian, S. and Rosenthal, Y.: Deep-sea temperature and ice volume changes across the Pliocene-Pleistocene climate transitions, Science, 325, 306-310, 10.1126/science.1169938, 2009.

Sosdian, S. and Rosenthal, Y.: Response to Comment on "Deep-sea temperature and ice volume changes across the Pliocene-Pleistocene climate transitions", Science, 328, 10.1126/science.1186768, 2010.

Yu, J. M. and Broecker, W. S.: Comment on "Deep-sea temperature and ice volume changes across the Pliocene-Pleistocene climate transitions", Science, 328, 10.1126/science.1186544, 2010.