

Supporting Information

Chalk et al. 10.1073/pnas.1702143114

SI Sample Locations and Age Model

We present a highly resolved (one sample per 3.5–4.5 ky) $\delta^{11}\text{B}$ -derived atmospheric CO_2 record from ODP Site 999 (Caribbean Sea; 12°44.64' N and 78°44.36' W) (Fig. S1) that spans the 0–260 (i.e., LP260) and 1,080–1,250 ky (i.e., eMPT) time intervals. Sedimentation rates in these intervals are ~ 3 cm/ky. Across large parts of the oceans (typically the oligotrophic regions of the low latitudes), the CO_2 content of the gas phase of surface seawater approximately equates to the CO_2 content of the overlying atmosphere (i.e., the surface water is in approximate CO_2 equilibrium with the atmosphere). It is from these locations, such as ODP Site 999 (Fig. S1), that atmospheric CO_2 concentrations can be most reliably reconstructed using the $\delta^{11}\text{B}$ method (34). Site 999 is today in near air–sea CO_2 equilibrium (approximately +20 μatm) (35) (Fig. S1) and has likely maintained this near-equilibrium state through the Pleistocene and Pliocene (24, 36). The foraminifera at ODP 999 are well-preserved, meaning that diagenesis is not likely to negatively impact our reconstructions. It is also worth noting that even moderate to severe diagenesis has been shown to have little or no impact on the $\delta^{11}\text{B}$ of planktic foraminifera (37).

For our age model, we generated a benthic $\delta^{18}\text{O}$ record for ODP Site 999 from ~ 600 –1,500 ky at a resolution of 3 ky using *Cibicidoides wuellerstorfi* that includes the same samples used to determine CO_2 during our eMPT interval (Fig. S2). The samples were measured on a Finnigan MAT 253 gas isotope ratio mass spectrometer connected to a Kiel IV automated carbonate preparation device at the Zentrum für Marine Tropenökologie. A detailed age model for Site 999 was generated by graphically aligning our benthic foraminiferal $\delta^{18}\text{O}$ record to the LR04 benthic $\delta^{18}\text{O}$ stack (29) (Fig. S2) using the Analyseries software (38). For the interval 0–500 ky, we generated a detailed age model by aligning the published planktic foraminiferal (*Globigerinoides ruber*) $\delta^{18}\text{O}$ record from Site 999 (at ~ 0.5 - to 2.0-ky resolution) (39) to the LR04 benthic $\delta^{18}\text{O}$ stack (29) (Fig. S2). The resultant age model provides excellent agreement between our lower-resolution benthic $\delta^{18}\text{O}$ data and the LR04 benthic $\delta^{18}\text{O}$ stack in the interval 110–260 ky (Fig. S2).

SI Methodology

Analytical Techniques. Between 140 and 220 individuals of *G. ruber* (white; *sensu stricto*; ~ 10 μg per shell) were picked from the 300- to 355- μm size fraction of the coarse (>63 μm) fraction of washed sediments from ODP 999. Foraminiferal samples were cracked and cleaned in the boron isotope clean laboratory at the University of Southampton. The cracking was done between two glass slides under a microscope, and care was taken to open all major chambers to allow for effective clay removal. The cleaning followed established methods for oxidative cleaning (40–42). After cleaning, *G. ruber* samples were dissolved in weak (~ 0.15 M) Teflon-distilled nitric acid and separated into two fractions: an “isotope” fraction (90%) and a “trace element” fraction (8–10% of sample volume). Boron was separated from the matrix of the isotope fraction with anion exchange resin (Amberlite IRA-743) in purpose-built columns (22, 43).

Boron isotope ratios were measured at the University of Southampton on a Thermo Scientific Neptune multicollector inductively coupled plasma mass spectrometer following established methodology (21, 36, 42, 43). External reproducibility was calculated based on the reproducibility of an in-house standard [Japanese coralline *Porites* – JCP-1 = 24.2‰ (42)] described by

Eq. S1 below, where $[^{11}\text{B}]$ = intensity in volts of ^{11}B ; for this study, the 2σ uncertainty typically corresponds to $\sim 0.2\%$:

$$2\sigma = 1.87 \cdot \exp^{-20.6[^{11}\text{B}]} + 0.22 \cdot \exp^{-0.43[^{11}\text{B}]} \quad \text{[S1]}$$

The trace element fraction was diluted, and Me/Ca ratios of Li, B, Na, Mg, Al, Mn, Fe, Sr, Cd, Ba, Nd, and U were measured on a Thermo Scientific Element 2-XR inductively coupled plasma mass spectrometer at the University of Southampton. Over the period of this study, reproducibility of three internal standards with a range of Me/Ca ratios was 4% for Mg/Ca and 5% for Al/Ca. We use the trace element data to screen for diagenesis and other potential artifacts. Samples with Al/Ca > 100 $\mu\text{mol/mol}$ were strictly removed to preserve data quality, and we find no anomalies in other element ratios (e.g., Mn/Ca, Ba/Ca, and Fe/Ca).

Mg/Ca temperatures were calculated following the approach of Evans and Müller (44) and using an Mg/Ca of seawater ($\text{Mg}/\text{Ca}_{\text{sw}}$) that was calculated from the modeled study of Fantle and DePaolo (45). We use an H value of 0.41, although species-specific to *Trilobatus* (formerly *Globigerinoides*) *sacculifer* (46), as no calibrated H value for *G. ruber* is currently available. Note that the accuracy of this reconstruction is not crucial to our findings, since sea surface temperature has a relatively weak effect on calculated CO_2 (*Determination of CO_2 from $\delta^{11}\text{B}$ -Derived pH*), and also, the potential change in $\text{Mg}/\text{Ca}_{\text{sw}}$ over the last million years is small because of the long residence times of the elements involved.

Determination of pH from $\delta^{11}\text{B}$ of *G. ruber*. Boron is present in two principal forms in seawater: boric acid $[\text{B}(\text{OH})_3]$ and borate ion $[\text{B}(\text{OH})_4^-]$. The relative proportion of these two species depends on pH and the dissociation constant for boric acid, pK_B^* :

$$\log_{10} \frac{\text{B}(\text{OH})_4^-}{\text{B}(\text{OH})_3} = \text{pH} - \text{pK}_\text{B}^* \quad \text{[S2]}$$

An isotopic fractionation between the two stable isotopes of boron-11 ($\sim 80\%$ abundance) and boron-10 ($\sim 20\%$) is also associated with this equilibrium (above), as the bond strengths between boron and the hydroxyl ion in boric acid and the borate ion differ. The ^{11}B concentrates in the more strongly bonded boric acid, giving it a higher $\delta^{11}\text{B}$ than the borate ion by $\sim 27.2\%$ (47). The delta notation ($\delta^{11}\text{B}$) is used to express differences in boron isotope ratios:

$$\delta^{11}\text{B}(\text{‰}) = \left[\left(\frac{^{11}\text{B}}{^{10}\text{B}} \right)_{\text{smp}} \left(\frac{^{10}\text{B}}{^{11}\text{B}} \right)_{\text{ref}} - 1 \right] \times 1,000, \quad \text{[S3]}$$

where $^{11}\text{B}/^{10}\text{B}_{\text{ref}}$ is the isotopic ratio of NIST SRM 951 boric acid standard ($^{11}\text{B}/^{10}\text{B} = 4.04367$) (48) and $^{11}\text{B}/^{10}\text{B}_{\text{smp}}$ is the isotopic ratio of the sample. The basis of the boron isotope–pH proxy is that, because the borate ion is tetrahedral and charged, it is more readily substituted into CaCO_3 (49, 50). Therefore, we can solve for pH based on the reconstructed isotopic composition of $\delta^{11}\text{B}_{\text{borate}}$:

$$\text{pH} = \text{pK}_\text{B}^* - \log \left(\frac{\delta^{11}\text{B}_{\text{sw}} - \delta^{11}\text{B}_{\text{borate}}}{\delta^{11}\text{B}_{\text{sw}} - (\alpha \cdot \delta^{11}\text{B}_{\text{borate}}) - 1,000 \cdot (\alpha - 1)} \right), \quad \text{[S4]}$$

where pK_B^* is the dissociation constant for boric acid at reconstructed in situ temperature, salinity, and pressure (51); $\delta^{11}B_{sw}$ is the isotopic composition of seawater (39.61‰) (52); $\delta^{11}B_{borate}$ is the isotopic composition of borate ion; and α is the isotopic fractionation factor between the two aqueous species of boron in seawater ($\alpha = 1.0272$, equivalent to an equilibrium isotope fractionation of $\sim 27.2‰$) (47). Boron is well-mixed in the oceans, with a residence time of 10–20 My; to account for likely (small) changes in the boron isotopic composition of seawater ($\delta^{11}B_{sw}$), we use a simple linear extrapolation with a central value at modern $\delta^{11}B_{sw}$ ($39.61 \pm 0.1‰$) (52) with uncertainty increasing by $\pm 0.12‰$ by 1,200 ky ($0.1‰/My$). This central value is consistent with an independent constraint, and any possible changes (53–55) are well-accounted for by the generous uncertainties that we ascribe to all parameters.

As $\delta^{11}B_{ruber}$ (the $\delta^{11}B$ of boron in *G. ruber* calcite) is offset from $\delta^{11}B_{borate}$, it is necessary to first account for this relatively minor deviation to calculate pH from $\delta^{11}B_{ruber}$, and a size-specific core top, culture, and field measurement calibration (Eq. S5) is applied for $\delta^{11}B_{borate}$ reconstruction (33, 39). This relationship has been shown in other studies to provide for accurate atmospheric CO_2 reconstructions (1, 21, 23). Uncertainties in this calibration (*G. ruber* 300–355 μm) are propagated through all relevant calculations (shown below at 2σ):

$$\delta^{11}B_{borate} = \frac{(\delta^{11}B_{ruber} - 8.87 \pm 1.51)}{0.6 \pm 0.08} \quad [S5]$$

Determination of CO_2 from $\delta^{11}B$ -Derived pH. To calculate aqueous CO_2 from pH, a second carbonate system parameter is required (34). Here, we briefly outline two approaches to estimate total alkalinity (ALK) to derive CO_2 from our pH data: (i) ALK remained within a specified range of its modern value, and (ii) ALK change correlates with reconstructed pH change. Below, we describe why both approaches yield very similar results, and we describe a sensitivity test to show that our main conclusions are not affected by the choice of how to estimate ALK. In the main text and our tabulated dataset, we assume that ALK has remained within a generous range around its modern value ($\sim 2,330 \pm 175 \mu mol/kg$) over the investigated period. The outcome of our own model inversion of CO_2 change of the last 1,500 ky (*Model Inversion* has details on the model) suggests a total ALK range of only 130 $\mu mol/kg$, much less than the 350- $\mu mol/kg$ uncertainty range propagated here. This means that we fully explore the likely ALK range based on modeling of whole-ocean ALK variability on G-IG timescales and over the last 1,500 ky (1, 26). This treatment of the second carbonate system parameter essentially assumes that any change in ALK will drive the majority of its impact on CO_2 by changing pH (which we reconstruct directly) rather than by changing the abundance of carbonate and bicarbonate ions. This assertion is based on first principle carbon chemistry and can be verified by comparing reconstructed pH with the ice core CO_2 record (Fig. S3A). When we cross-plot $\delta^{11}B$ -based pH against ice core CO_2 , we find that the data are well-described by the theoretical pH to CO_2 relationships, assuming either constant DIC or constant ALK (Fig. S3A), in support of our assertion. The alternative view is that much of the reconstructed pH change was caused by changes in ALK, such that ALK and pH are correlated (i.e., the ~ 0.2 pH increase during the Last Glacial Maximum corresponds to a +100- $\mu mol/kg$ ALK increase) (26, 56). We carry out a sensitivity test to estimate how much difference in terms of CO_2 that it makes to account for this correlation (Fig. S3B). We find that including the ALK to pH correlation reduces our estimate for additional glacial-stage CO_2 reduction by about 5 μatm , and it reduces our estimated increase of the G-IG CO_2 range by

about 5 μatm (Fig. S3C). As described above, this relatively modest change is because of the fact that ALK change causes most of its impact on CO_2 by changing surface pH, the parameter that we reconstruct (Fig. S3B). We argue that the modest difference between the two ALK assumptions does not change the main conclusions of this study, especially given the agreement between our data, previously published fully independent CO_2 reconstructions, and our carbon cycle modeling results (Fig. 4 and Fig. S4). Since the ALK to pH correlation is itself uncertain, we prefer the calculation using a large and constant ALK uncertainty range that more than covers existing estimates of G-IG ALK change (26, 56). With this approach, the small deviation between the ALK assumptions is effectively covered by our uncertainty propagation, and there is little independent evidence to suggest that this assertion is invalid (57, 58).

Using $\delta^{11}B$ -derived pH and the prescribed ALK and modern aqueous CO_2 disequilibrium, uncertainties are propagated via a Monte Carlo simulation ($n = 10,000$) in the statistical analysis program R (59), and 95% confidence intervals are calculated from the variation within these simulations. Two SD uncertainties on the individual input variables are included for $\delta^{11}B$ (with or without analytical uncertainty defined by Eq. S1, normal distribution), Mg/Ca-derived temperature ($\pm 3^\circ C$, normal distribution), salinity (± 3 psu, normal distribution), ALK ($\pm 175 \mu mol/kg$, uniform distribution), and $\delta^{11}B_{sw}$ (39.61 ± 0.10 – $0.22‰$, normal distribution) (52). Atmospheric CO_2 was then calculated from aqueous CO_2 using Henry's Law, subtracting the modern extent of disequilibria with respect to CO_2^{atm} at the site [given a nominal uncertainty at 2σ of $\pm 50\%$ (i.e., $\pm 10 \mu atm$)]. Uncertainties in CO_2 also include propagation of the uncertainties in the $\delta^{11}B$ calibration of *G. ruber* (Eq. S5). All subsequent carbonate system calculations in R were performed using the seacarb package (59, 60).

SI Forcing to SL Relationship

Extraction of the underlying relationship between two empirical datasets requires careful consideration of uncertainty in both x and y coordinates (61). In our case, the x coordinate is CO_2 climate forcing, ΔR_{CO_2} , which is calculated directly from reconstructed CO_2 following ref. 33:

$$\Delta R_{CO_2} = \alpha \times \ln \frac{CO_2}{CO_2^0}; \alpha = 5.35 \frac{W}{m^2}; CO_2^0 = 278 \mu atm. \quad [S6]$$

We calculate mean and SD of ΔR_{CO_2} as the half-point and quarter-width of the $\pm 2\sigma$ interval of our CO_2 reconstruction. As for the y coordinate, SL in our case, we point out that there are three sources of uncertainty: (i) the inherent uncertainty of the SL estimate documented alongside each SL record, (ii) the range of SL within the age window defined by the relative age uncertainty of the SL and the CO_2 age models, and (iii) the systematic discrepancies between different SL records. The latter point we address by carrying out our analyses independently for a number of recent SL records (11, 12, 24, 62), and we find only modest differences, suggesting that our conclusions are robust to these systematic SL discrepancies. To account for both *i* and *ii*, for each CO_2 data point, we estimate SL and its uncertainty by constructing the cumulative distribution function (CDF) of all SL data that fall within the relative age uncertainty window (which we take to be 1.5–2 ky for our LP260 data; an uncertainty ramps from 2 to 4 ky over the course of the 800-ky compilation of ice core CO_2 reconstructions and 6 ky for our eMPT data). For each individual CDF, we estimate the appropriate SL mean and SD as the half-point and half-width of the 16 to 84% interval (corresponding to $\pm 1\sigma$). This procedure yields typical individual data point SL uncertainties about 20 m ($\pm 2\sigma$) and significantly larger uncertainties for a few points that happen to correspond to times of rapid SL change.

As for the regression of the underlying CO₂ forcing to SL relationship, we make use of the approach laid out by York et al. (61), a generalized form of “reduced major axis” regression method that takes full account of normal error in x and y . We carry out this regression technique using four independent SL reconstructions (11, 12, 24, 62) for both our LP260 CO₂ dataset (Fig. S7A) and our eMPT dataset (Fig. S7D). In all cases, we display the regression confidence interval. The confidence intervals of the regressed slopes for LP260 (Fig. S7B1) and eMPT (Fig. S7B3) are of particular interest. We consistently find the CO₂ forcing to SL slope to be significantly steeper during LP260 (about 40–50 m/Wm⁻²) than during eMPT (less than 25 m/Wm⁻²), supporting the main line of argument in our study.

To build confidence in our regression method, we offer three tests. (i) Are the regressions based on our LP260 dataset consistent with regressions based on the compilation of the continuous ice core CO₂ record of the last 800 ky? (ii) Does preindustrial SL fall within the prediction interval of our regressed CO₂ forcing to SL relationships? (iii) Is the regression method by York et al. (61) skillful in extracting the correct CO₂ forcing to SL relationships when put to the test with synthetic data that mimic our eMPT data and their uncertainties?

To address *I*, we carry out the regression using the ice core CO₂ record against all SL records based on a 200-ky sliding window of data. Because the sampling interval of the ice core CO₂ compilation changes dramatically over its 800-ky range (and in particular, over the course of the last deglaciation), we regularize the CO₂ data by binning it into ± 1 -ky increments and estimate bin mean and SD analogous to the SL data. We find that, throughout the last 800 ky, the CO₂ forcing to SL slopes regressed in that way generally fall within a band of 35–55 m/Wm⁻² (gray shading in Fig. S7B2), in good agreement without LP260 data and significantly greater than for our eMPT data. We note that the slope regressed from the SL record that is based on Mg/Ca- $\delta^{18}\text{O}$ deconvolution (11) exhibits a long-term oscillation not seen in any of the other SL regressions, but even at its lowest point, it is still significantly larger than that regressed during the eMPT. Likewise, while the Mediterranean SL reconstruction (12) falls into a constant narrow range for most of the 800 ky, its regressed CO₂ forcing to SL slope yields anomalous variability during a brief interval around 500 ky, where the SL record is known to be overprinted by a series of sapropel layers. The regression based on the Red Sea SL record (25), arguably the best constrained SL reconstruction, is remarkably constant over the entire 500-ky period that it spans, as is the case for the entire last 800 ky for the regression based on simulated SL by De Boer et al. (62). Based on that analysis, we conclude that there is no evidence for substantial changes in the CO₂ forcing to SL slope over the entire 800-ky duration of the ice core CO₂ record and that the significantly lower regressed slopes for our eMPT dataset therefore speak to a change in the CO₂ forcing to SL relationship associated with the MPT. Furthermore, we note that, in Fig. 2, we only present the regression results of the Mediterranean SL record, which this supplementary analysis suggests to be both robust and the most conservative record covering both the LP260 and eMPT intervals. We have further confidence in our conclusion of an MPT reduction in the CO₂ forcing to SL slope, because we find an equivalent change when considering the LR04 benthic foraminifera $\delta^{18}\text{O}$ stack instead of SL (Fig. S8), which is observationally well-constrained and carries a significant component SL signal.

To address *ii*, the test for SL predicted for preindustrial CO₂ forcing, we replicate our analysis of the CO₂ forcing to SL slope for the intercept regression parameter, shown in an analogous way in Fig. S7C. When using the regressed CO₂ forcing to SL relationship, however, it is the prediction interval (dashed in Fig. S7C) and not the regression confidence interval that is relevant, because the CO₂ forcing to SL regression does not embody re-

sidual SL change caused by factors, such as orbital change. Hence, the prediction interval at the intercept is wider than the intercept confidence interval, and it comfortably includes preindustrial SL at preindustrial CO₂ forcing during LP260.

Finally, to address *iii*, the question as to the skill of York regression to extract the correct CO₂ forcing to SL slope even in the face of substantial x and y uncertainty of individual points and an overall low signal to noise ratio, we construct a synthetic test that mimics the signal to noise ratio of our eMPT data (Fig. S9). That is, we presuppose a known CO₂ forcing to SL relationship and generate randomized data points based on the typical uncertainties of CO₂ forcing and SL. To carry out this analysis, we change the number of synthetic data points used for the regression and find that (i) the true slope of the known synthetic relationship falls within the regressed slope confidence interval; (ii) the width of the regressed slope confidence interval systematically narrows as the number of data points increases; and (iii) at 50 data points, the slope confidence interval is narrow enough to detect the change in slope that we reconstruct between LP260 and eMPT.

SI Carbon Cycle Modeling

Model Details. To simulate the global carbon cycle and atmospheric CO₂ levels across the MPT, we make use of the recently updated CYCLOPS model (27). The model represents the oceanic carbon cycle and physical circulation as the exchange between 18 separate surface, middepth, and deep water reservoirs, as well as carbon fluxes between surface water reservoirs and an atmospheric reservoir. Following previous work (63, 64), the model simulates the open system CaCO₃ cycle through explicit representation of undersaturation-driven seafloor dissolution of biogenic CaCO₃ rain that originates from the surface.

To simulate global carbon cycle variations in the Pleistocene, we apply three fully separate forcings to the model, all of which represent well-established modes of carbon cycle and circulation change previously investigated using CYCLOPS: (i) major nutrient drawdown driven by glacial-stage iron fertilization of the Sub-Antarctic Zone of the Southern Ocean (7, 65); (ii) coupled glacial-stage reduction in vertical exchange, export production, and residual surface nutrient status of the Polar Antarctic Zone of the Southern Ocean (56, 66–68); and (iii) glacial-stage shoaling of the AMOC (69–72). The rationale and evidence base for these changes have been reviewed elsewhere (73), and the model’s atmospheric CO₂ sensitivity to these mechanisms has been evaluated in detail (27). To represent the time evolution in model forcing, we use ODP 1090 iron mass accumulation rate (1), ODP 1094 Ba/Fe (28), and ODP 982/U1313 (29, 30) observational records to adjust model conditions in the Sub-Antarctic Zone, the Polar Antarctic Zone, and the global deep water circulation pattern, respectively. Other drivers of CO₂ change (e.g., silicate weathering feedback, temperature) are not considered here but will be investigated in the future. Similarly, by implementing Atlantic circulation changes based on North Atlantic carbon isotope gradients, our model does not reflect exceptional circulation weakening during MIS 23 inferred from neodymium isotopes (74) or any other modes of AMOC reorganization.

Extension of ODP 1094 Ba/Fe Proxy Forcing. The previously published ODP 1094 Ba/Fe record from 0 to 1 Mya was extended to allow us to start the forced simulation as early as 1.5 My. The composite section for the deeper part of ODP core 1094 (>121 mcd; 1–1.5 My) was mainly reconstructed using magnetic susceptibility (75), Fe counts derived from X-ray fluorescence (XRF) scanning on sediments of Holes A and D, and the $\delta^{18}\text{O}$ of the planktonic foraminiferal species *Neogloboquadrina pachyderma*. The age model of the core is based on tuning the $\delta^{18}\text{O}$ of the benthic foraminifer *C. wuellerstorfi* to the LR04 benthic stack

(29), as will be documented elsewhere. Oxygen isotope analyses were performed with a Thermo GasBench II coupled to a Thermo Delta V Plus mass spectrometer at the Geological Institute, ETH Zurich (1 σ error of 0.07‰). The relative sedimentary concentrations of Ba and Fe were acquired with an AVAATECH profiling XRF core scanner at the MARUM, University of Bremen using the same settings and procedures as in ref. 28 but with a different detector (Canberra X-PIPS Silicon Drift Detector, Model SXD 15C-150-500). To assure that the XRF data are consistent throughout the entire record, four sections between 0 and 1 My have been rescanned. The elemental XRF data of the rescanned sediments are correlated to each other, and their linear fit is used to scale newly acquired data accordingly ($r = 0.97$ for Fe, $r = 0.77$ for Ba).

Model Inversion. The relationship between the observational records and the forcing applied to the model (sub-Antarctic nutrient status, polar Antarctic/deep Southern Ocean exchange, polar Antarctic surface nutrient status, and Atlantic overturning) is represented as four equations with a total of seven free parameters (i.e., intercept and slope, threshold in the case of Atlantic overturning). The Atlantic circulation is simulated as “shallow” only if the reconstructed vertical stable carbon isotope gradient in the deep North Atlantic increases above a threshold value determined as part of the inversion; all other forcing functions are continuous and monotonic. For the Antarctic changes, vertical exchange scales with Ba/Fe, while surface nutrient status scales with the square root of Ba/Fe, thereby in effect, assuming a sublinear scaling between Ba/Fe and export production [i.e., the property arguably recorded by Ba/Fe at ODP 1094 (28)]. The seven forcing function parameters are initially set to arbitrary (but sensible) values.

We invert the model by finding the specific combination of the seven forcing function parameters that minimize the rms error of simulated atmospheric CO₂, as evaluated against a composite reconstruction of atmospheric CO₂ from Antarctic ice cores (19). That is, given the observational forcing records, the forward model, and the CO₂ record, we invert to obtain the model forcing parameters that minimize the rms objective function. To be confident in the outcome of the minimization procedure across 7D parameter space, we incorporate into the model two fully independent methods and verify that they converge onto the same solution: (i) Powell’s conjugate direction method (76–78) and (ii) the Fletcher–Reeves implementation of the Steepest Descent conjugate gradient method (79). Side conditions relating to physically possible parameter solutions are encoded directly into the model, and both algorithms converge to the same solution at an rms of 12.2 μatm .

SI Quantification of ^GCO₂, ^{IG}CO₂, and ^{IG-G}ΔCO₂

Given our two $\delta^{11}\text{B}$ -based CO₂ reconstructions (eMPT, LP260), we wish to (i) quantify by how much glacial and interglacial CO₂ levels (i.e., ^GCO₂ and ^{IG}CO₂, respectively) are different across the MPT and by how much the magnitude of G-IG CO₂ cycles (i.e., ^{IG-G}ΔCO₂) has changed and (ii) compare our estimates with equivalent estimates derived from existing CO₂ records (1, 2, 19) as well as from our model-derived predictions of CO₂ (*SI Carbon Cycle Modeling*). This section describes the details of the analysis and offers discussion of the rationale behind these calculations. Our analysis of the CO₂ datasets consists of four steps: (i) define which individual data points represent glacial conditions and which data points represent interglacial conditions, (ii) estimate the average CO₂ and its uncertainty from the two sets (i.e., ^GCO₂ and ^{IG}CO₂) of CO₂ data, (iii) estimate the G-IG CO₂ difference (i.e., ^{IG-G}ΔCO₂) before and after the MPT, and (iv) test whether the resulting changes are statistically significant.

To define which individual data points contribute, we subsample each record based on a percentile cutoff criterion using, for example, only the 25% “most interglacial” and the 25% “most glacial” data

points in the quantification of ^GCO₂, ^{IG}CO₂, and ^{IG-G}ΔCO₂. In the case of the $\delta^{11}\text{B}$ -based CO₂ reconstructions, we can rank the individual data points based on $\delta^{18}\text{O}$ measured on the same sample to directly reflect climate state, but in the case of the direct ice core measurements and our model-simulated CO₂, we have to rank the individual data points based on their CO₂ level. The value of 25% for the cutoff criterion is arbitrary, and therefore, we carry out and present the analysis for the full 1–50% range of nonoverlapping subsamples (Fig. S4). A low percentile cutoff implies that only the few most extreme data points are subsampled as glacial and interglacial, which maximizes the glacial/interglacial difference but also leads to a relatively large uncertainty, because only few data contribute to the estimates. Conversely, a high percentile cutoff implies that more individual data contribute to the estimation of CO₂, which reduces the uncertainty and increases robustness but also, progressively averages away the end member changes that we try to quantify. For brevity, we make the tradeoff between robustness and sensitivity and exclusively discuss the 25% percentile cutoff in the main text, and we note that our qualitative results are independent of that choice (Fig. S4).

The second step of our analysis is to estimate ^GCO₂ and ^{IG}CO₂ based on their respective subsamples (see above) and also propagate the uncertainty of the individual data points. To this end, we first sum up and normalize the probability density functions of the individual data points and their normally distributed uncertainty. We assign 1 σ uncertainties of 20 μatm to our CO₂ data, 14 μatm to the Hönisch et al. (11) data (ref. 1 as published; we note that ref. 1 does not include a number of sources of uncertainty that are included in the uncertainty calculations for our data), a nominal 10 μatm to ice core data, and 12.2 μatm to simulated CO₂ based on model inversion residual rms. From these cumulative subsample probability density functions, we estimate the central tendency as the median probability and that normal dispersion as one-half of the central 66% cumulative probability interval. Thus, the dispersion of our estimates represents a combination of inherent measurement uncertainty associated with individual data points and of the true spread in CO₂ levels within the subsample. To estimate ^{IG-G}ΔCO₂, we simply subtract the central estimate of ^GCO₂ from ^{IG}CO₂, whereby the normal dispersion of ^{IG-G}ΔCO₂ is propagated as the root sum of squares of the normal dispersion estimates of ^GCO₂ from ^{IG}CO₂.

The final step in our analysis is to evaluate the significance of the differences that we find in glacial and interglacial CO₂ and glacial/interglacial CO₂ range before and after the MPT (i.e., $\Delta^{\text{G}}\text{CO}_2$, $\Delta^{\text{IG}}\text{CO}_2$, and $\Delta^{\text{IG-G}}\text{CO}_2$ as also visualized in Fig. 4). To this end, we subtract eMPT ^GCO₂, ^{IG}CO₂, and ^{IG-G}ΔCO₂ from the corresponding LP260 estimates, again propagating the estimate dispersion as the root sum of squares of the individual normal dispersion estimates. For all four CO₂ datasets (i.e., two $\delta^{11}\text{B}$ -based CO₂ records, ice core data, and model inversion results), we find that estimated $\Delta^{\text{IG}}\text{CO}_2$ falls onto zero within its 1 σ dispersion (thick black whisker in Fig. 4), and thus, we cannot reject the null hypothesis that interglacial CO₂ levels were identical during our intervals (eMPT and LP260). Conversely, we find that estimated $\Delta^{\text{G}}\text{CO}_2$ is negative for all four CO₂ datasets, with the deviation from zero exceeding the threshold of 1.64 σ for one-sided testing (thin black whisker in Fig. 4). That is, we can reject at 95% confidence level that glacial-stage CO₂ before the MPT was as low or lower than after the MPT. Likewise and largely driven by this significant glacial-stage CO₂ decline across the MPT, the estimated MPT change in the G-IG CO₂ range is significantly greater than zero for all four datasets (at 95% confidence level for one-sided testing) (thin black whisker in Fig. 4). Thus, we can reject at 95% confidence level the null hypothesis that the magnitude of glacial/interglacial CO₂ change before the MPT was as large (or larger) as the glacial/interglacial CO₂ change after the MPT.

Additional results of statistical analysis are tabulated in Table S1.

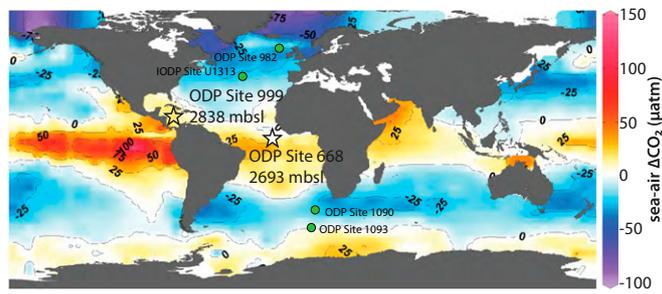


Fig. S1. Map of sediment core sample sites. Color scale represents CO_2 disequilibrium between surface water and air, with surface water supersaturation indicated by red tones (35). The locations of ODP 999 and ODP 668 are indicated. Also indicated are the locations of cores used to inform our modeling results in green: ODP 982, IODP U1313, ODP 1090, and ODP 1093.

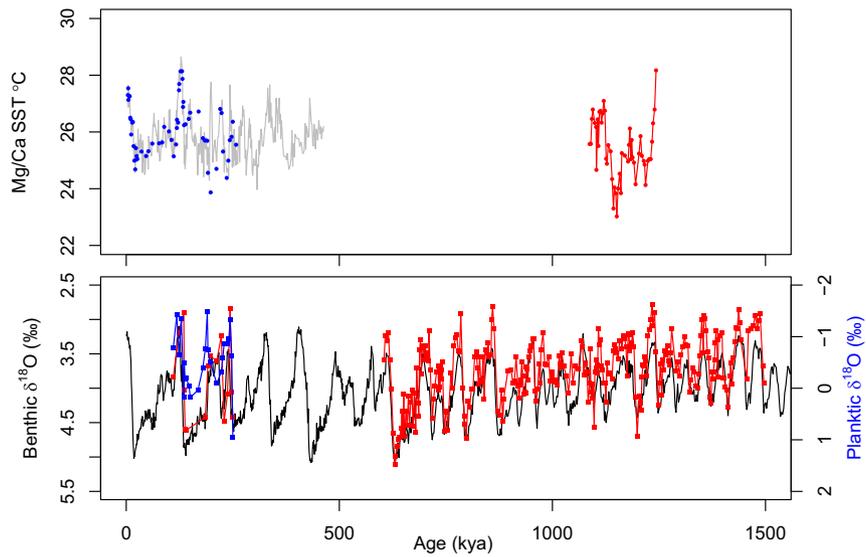


Fig. S2. Foraminiferal Mg/Ca and oxygen isotope measurements from ODP 999. (A) Planktic Mg/Ca temperature estimated from *G. ruber* (blue, LP260; red, eMPT). Temperatures are calculated as stated in *Analytical Techniques*; also shown is the previous record of Mg/Ca sea surface temperature (SST) from this Site (gray) (39). (B) Planktic *G. ruber* (blue) and benthic *C. wuellerstorfi* (red) oxygen isotope results plotted vs. the global benthic foraminiferal oxygen isotope stack (black) (29).

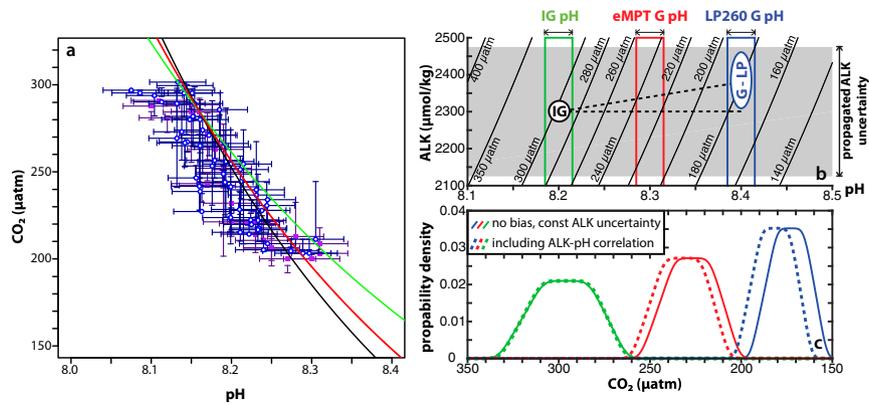


Fig. 53. Assessment of ALK assumption. Our treatment of the second carbonate system parameter asserts that any change in ALK will cause most of its CO_2 impact by changing pH, which we reconstruct directly, rather than by the change in the abundance of carbonate and bicarbonate ion. (A) To test this assertion, we cross-plot reconstructed pH and ice core CO_2 and assess the relationship between these empirical data against our assertion. Blue and purple dots are our LP260 CO_2 data and CO_2 data from ref. 1, respectively, with the corresponding CO_2 interpolated from the ice core CO_2 record (19). The three solid lines represent the theoretical pH– CO_2 relationships under the following assumptions: (red) constant DIC, (black) constant ALK, and (green) DIC and ALK both change with a set ratio of 1:1. Either assumption of constant dissolved inorganic calcite (DIC) or constant ALK gives faithful reproduction of ice core CO_2 from the Site 999 and 668 $\delta^{11}\text{B}$ –pH records, in strong support of our assertion. The CO_2 uncertainty in the empirical data relates to the age model uncertainty (~ 2 ky) of the pH data points and the resulting spread in interpolated ice core CO_2 . (B) To illustrate our ALK assumption further, we show CO_2 as a function of ALK and pH at 25 °C and salinity of 35, as predicted yielding nearly vertical contours that show a strong sensitivity of CO_2 to pH and a weak sensitivity to ALK. As a point of reference, we highlight pH values of 8.2 ± 0.15 , 8.3 ± 0.15 , and 8.4 ± 0.15 to represent interglacial pH values during both eMPT and LP260, eMPT glacial pH, and LP260 glacial pH, respectively. The large ALK uncertainty range that we propagate in our calculations (gray shading) more than covers estimates of ALK change over recent G-IG cycles (i.e., dashed line connecting IG with Last Glacial Maximum). (C) From the sensitivity calculations, we estimate the CO_2 offset between two plausible treatments of ALK uncertainty: (solid lines) assuming ALK remained within constant range (i.e., ± 175 $\mu\text{mol/kg}$) and (dashed lines) assuming an ALK to pH correlation of $+100$ $\mu\text{mol/kg}$ per 0.2 pH increase in addition to the ± 175 $\mu\text{mol/kg}$ ALK uncertainty. We find that the ALK to pH correlation causes modest mean increases of ~ 5 and ~ 8 ppm in the estimates for eMPT and LP260 glacial CO_2 levels, respectively. The implied ~ 3 -ppm reduction in the difference between glacial-stage eMPT and LP260 CO_2 is too small to undermine the main conclusions of our study.

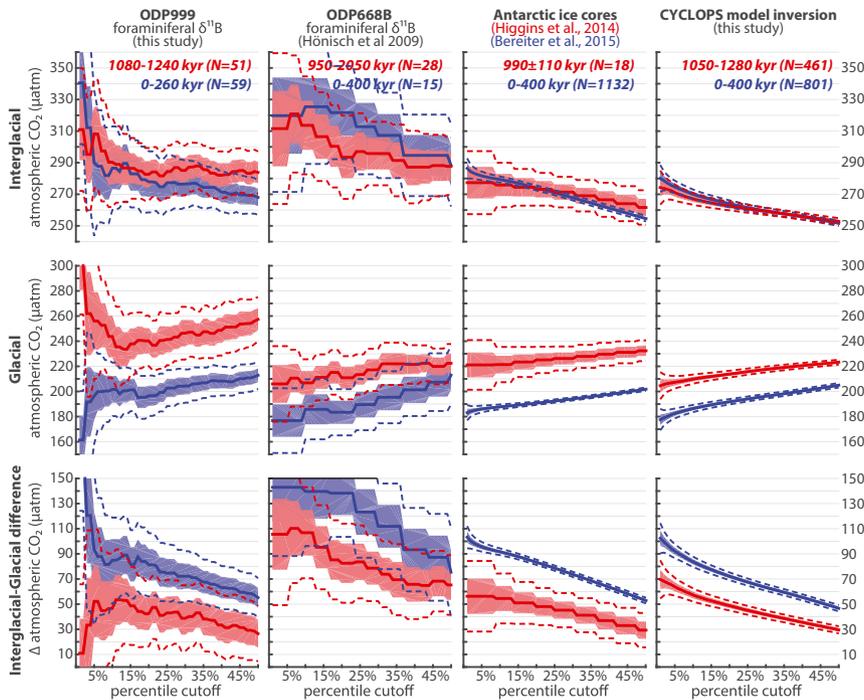


Fig. 54. Quantification of CO_2 change since the MPT. As described in detail in *Materials and Methods*, we quantify (Top) interglacial CO_2 , (Middle) glacial CO_2 , and (Bottom) interglacial – glacial CO_2 for data both from the latest Late Pleistocene (blue) and during or before the MPT (red). From left to right, quantification is carried out on four different datasets: boron isotope data from ODP 999 (this study) and ODP 668B (1), CO_2 directly measured on air trapped in Antarctic ice using the compilation of continuous ice core CO_2 reconstructions (19) and stratigraphically disturbed ~ 1 -My-old “blue ice” from the Allan Hills (2), and CYCLOPS model inversion (this study). To define glacial and interglacial subsets of the datasets, we use on a cutoff criterion, subsampling the data with a given percentile of lowest/highest $\delta^{18}\text{O}$ (marine records) or CO_2 (ice core, model). For this figure, we carry out this analysis while systematically changing the value of the critical percentiles (x axis). The results discussed in the main manuscript and shown in Fig. 4 correspond to a 25% cutoff percentile, but as shown here, our conclusions are robust across a wide range of percentiles. The underlyingly data are compiled for percentile increments of 5% in Table S1.

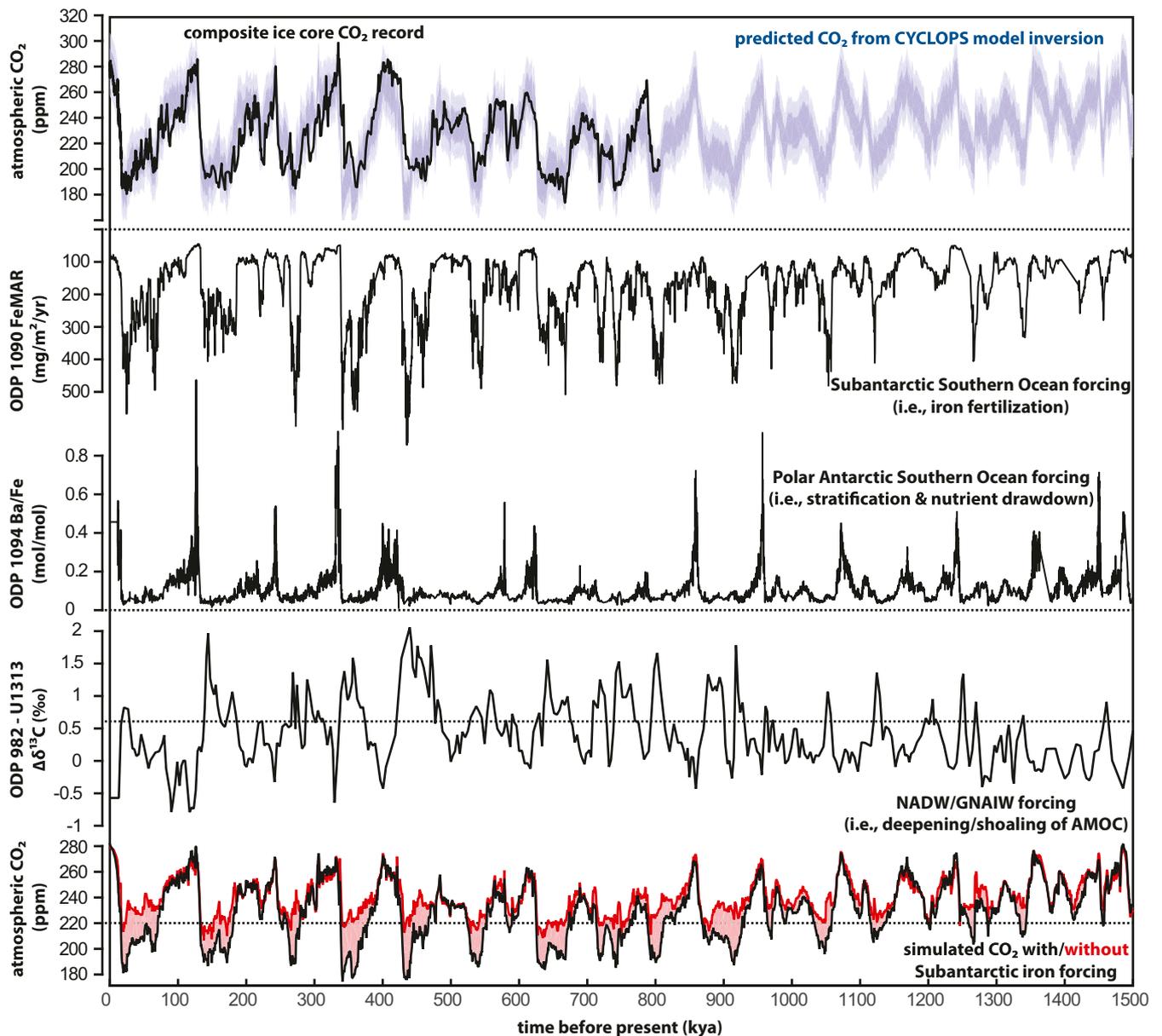


Fig. 55. Overview of CYCLOPS model inversion results. *Top* shows (in black) the composite ice core CO₂ record (19) used as the inversion target and (in blue) the model inversion result, with dark and light shading indicating $\pm 1\times$ and $\pm 2\times$ the rms error of the inversion ($12.3\ \mu\text{atm}$), respectively. *Middle* shows the three paleoceanographic proxy records used to drive the inversion as described in detail in *SI Carbon Cycle Modeling*. *Bottom* illustrates the contribution of dust-borne iron fertilization of the Sub-Antarctic Zone of the Southern Ocean, highlighting the simulated CO₂ change that can be attributed to that forcing. FeMAR, iron mass accumulation rate; NADW/GNAIW, North Atlantic Deep Water/Glacial North Atlantic Intermediate Water.

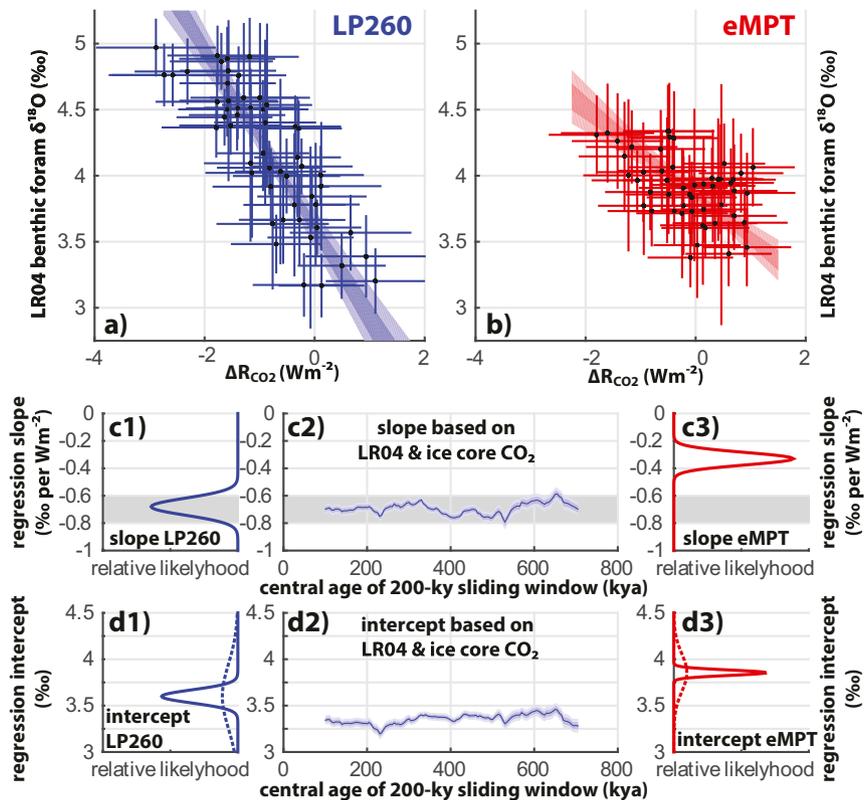


Fig. 58. Regression of $\delta^{18}\text{O}$ and $\delta^{11}\text{B}$ records. *A* and *B* display the regression results of our LP260 and eMPT boron isotope datasets against the LR04 benthic foraminifera oxygen isotope stack (Fig. S7 has details). *C* and *D* compare the results of the regression slope and intercept, respectively, with equivalent analyses using a 200-ky sliding window based on the ice core CO_2 record (*C2* and *D2*). The slopes mimic those seen in the equivalent SL analyses.

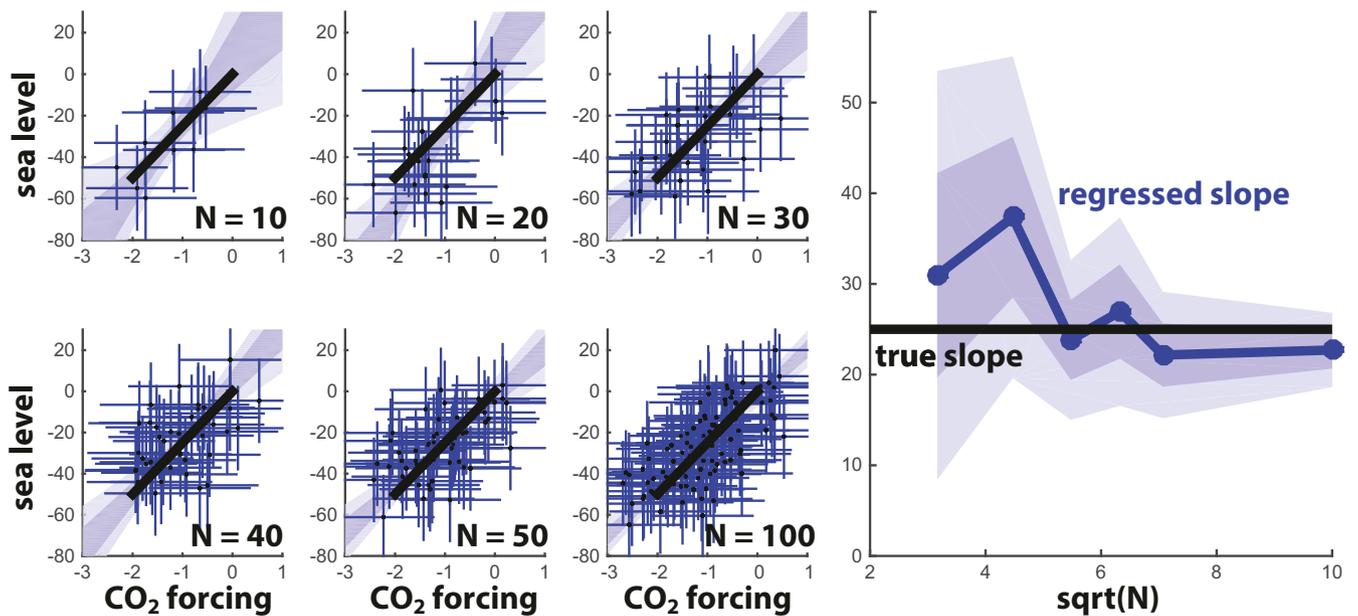


Fig. 59. Synthetic test of the York regression method. Test of regression robustness based on synthetic data that mimic our eMPT dataset. Noise is added to a known relationship of SL and CO_2 radiative forcing ($y = 25x$; shown as the fat black line) before the York regression is applied. We test the performance of the regression as a function of how many data points are contained in the dataset and find that the 51 samples in our eMPT dataset should be sufficient to recover the true CO_2 forcing to SL relationship with the required accuracy.

