Supporting Information for "Pliocene warmth and patterns of climate change inferred from paleoclimate data assimilation"

Jessica E. Tierney[1], Jonathan King[1,2], Matthew B. Osman[3], Jordan T. Abell[4], Natalie J. Burls[5], Ehsan Erfani[6], Vincent T. Cooper[7], Ran Feng[8]

[1]Department of Geosciences, The University of Arizona, Tucson, AZ USA

[2]United States Geological Survey, Golden, CO, USA

[3]Department of Geography, The University of Cambridge, Cambridge, UK

[4]Department of Earth & Environmental Sciences, Lehigh University, Bethlehem, PA, USA

[5]Department of Atmospheric, Oceanic and Earth Sciences, George Mason University, Fairfax, VA, USA

[6]Desert Research Institute, Reno, NV USA

[7]Department of Atmospheric and Climate Science, University of Washington, Seattle, WA, USA

[8]Department of Geosciences, University of Connecticut, Storrs, CT, USA

Contents of this file

Additional Supporting Information (Files uploaded separately)

Introduction

The Supporting Information contains additional text on the updates we made to the Mg/Ca proxy forward model as part of this study, as well as the results of the internal and external validation tests applied to plioDA. There are nine supplemental figures including two pertaining to the Mg/Ca proxy (Fig. S1 and S2), two pertaining to the validation tests (Fig. S3 and S4), two showing DA results with alternative prior choices (Fig. S5 and S7), one showing DA results for the mid-Pliocene with a different proxy network (Fig. S6), one showing both mid and early Pliocene results for changes in precipitation minus evaporation ($P - E$) and sea-surface salinity (SSS) (Fig. S8), and one showing the mid-Pliocene posterior mean for $P - E$ compared to the multi-model means for the PlioMIP2 and cloud simulations, respectively (Fig. S9). There are four supplemental tables, three of which are large so are uploaded separately as Excel files. These consist of a description of the proxy data used in plioDA (Table S1), a description of proxy data excluded from plioDA (Table S2), and a description of the model simulations used in plioDA (Table S3). Captions for these three large tables are provided at the end of this document. Table S4, which lists changes in global mean surface temperature (GMST) and the tropical Pacific zonal gradient for different model prior choices, appears after the figures.

Text S1. Updates to the Mg/Ca forward model. Since BAYMAG was published in 2019 (Tierney et al., 2019) the inferred evolution of Mg/Ca_{sw} in the Neogene has changed. Na/Ca ratios in foraminifera provide new constraints on the evolution of Ca concentrations, the impact of which is to increase the rise in Mg/Ca_{sw} over the last 20 Ma (Zhou et al., 2021; Rosenthal et al., 2022). This change affects the forward-modeling of Mg/Ca of foraminifera during our target Pliocene timeslices, so we updated the built-

in Mg/Ca_{sw} function in BAYMAG to include the Na/Ca data. We also excluded the coral Mg/Ca_{sw} inferences (Gothmann et al., 2015) from the last 20 Ma, which have large uncertainties. As expected, the resulting revised Mg/Ca_{sw} reconstruction shows a larger change over the late Neogene (Fig. S1).

We also updated BAYMAG to incorporate a non-linear power component (H) into the Mg/Ca_{sw} - Mg/Ca_{foram} relationship into the BAYMAG model following geochemical expectations (Evans & Müller, 2012); i.e.:

$$Mg/Ca_{foram} = a * Mg/Ca_{sw}^{H} \tag{1}$$

In paleoclimate applications the Mg/Ca_{sw} of past times is ratioed to the modern value, such that the coefficient a cancels out and only H needs to be defined:

$$\frac{a * Mg/Ca_{sw-t}^{H}}{a * Mg/Ca_{sw-t0}^{H}} = \left(\frac{Mg/Ca_{sw-t}}{Mg/Ca_{sw-t0}}\right)^{H} \tag{2}$$

H is user-defined, and in our case since we have multiple foraminiferal species represented in plioDA (*T. sacculifer*, *G. bulloides*, and *N. atlantica*) some of which don't have an experimentally determined H, we sought a general value for H based on multiple studies. We compiled both culture and inorganic precipitation experimental data, and then computed the inter-study mean H (0.74) which is what we use in our forward model application for plioDA (Fig. S2).

Text S2. plioDA validation tests.

To determine the best choices for R and covariance localization, we used both internal and external validation exercises. First, we conducted set of leave-one-out (LOO) proxy internal validation experiments across a range of values for R and the localization radius. In each data assimilation experiment, we systematically left out one proxy at a time and

calculated a posterior, then computed how well the posterior predicted the set of withheld proxies (the skill) using the R^2 metric. We conducted these experiments for each of the three target timeslices (Fig. S3), but we focus on the results for the Pliocene since high skill is expected for the preindustrial slice (the proxy system models are calibrated to preindustrial conditions). In the Pliocene time slices, we found that there was little improvement in skill across different choices of R scalings (the y-axis in Fig. S3), and that skill values were maximized near $R = 1$; hence our decision to use the unscaled values from the Bayesian proxy forward models (as noted in the main text). However, we did find differences in skill when localization was varied (the x-axis in Fig. S3); skill is maximized when a radius of 15,000–27,000 km was applied (Fig. S3).

The differences in internal validation skill across localization choices are subtle, so we also conducted an external validation test in which we compared our posterior surface air temperature (SAT) spatial fields with the mid-Pliocene terrestrial temperature proxy dataset compiled by Salzmann et al. (2013). These data mostly come from the Northern Hemisphere (Fig. S4a), carry large uncertainties, and many of the estimates are qualitative, but they nevertheless serve as a completely independent check on the range of temperatures produced by plioDA. We find that skill is maximized under moderate localization (18,000–24,000 km) and drops off steeply when localization is set to smaller radii (Fig. S4b), reflecting the fact that smaller radii limit the extent of the DA update over land areas far from the ocean.

We also tested two assumptions regarding the seasonality of the alkenone $U_{37}^{K'}$ signature during the Pliocene. Today, $U_{37}^{K'}$ is seasonally-biased towards the summer months in the North Pacific and North Atlantic, and winter months in the Mediterranean (Tierney &

Tingley, 2018); we follow this guideline when assimilating the $U_{37}^{K'}$ data with the preindustrial simulations. However, $U_{37}^{K'}$ seasonality almost certainly has changed over geological time. If we assume that the modern seasonal patterns hold during the Pliocene, we obtain moderate skill against the Salzmann et al. (2013) dataset at large localization radii ($R^2 = 0.3$; Fig. S4b). However, if we assume that the $U_{37}^{K'}$ represents mean annual temperatures during the Pliocene (in the regions that are presently seasonal), the skill is noticeably improved, rising to $R^2 = 0.53$ at a localization radius of 24,000 km (Fig. S4b). The improved skill reflects the fact that the annual assumption results in more warming over the northern high latitudes, in better agreement with the terrestrial data, which suggest substantial polar amplification. A shift towards annual production is conceivable, especially in the northern high latitudes, since the overall warmer conditions during the Pliocene may have allowed production to continue through the winter months. It is less clear why production would become annual in the Mediterranean, where today it is limited during the summer by stratification (Ternois et al., 1996), but the Pliocene $U_{37}^{K'}$ values from the Mediterranean are very high (translating to a warming of about 7°C) which suggests a shift in seasonality is likely.

Tindall, Haywood, Salzmann, Dolan, and Fletcher (2022) suggested that the terrestrial proxy records in the Salzmann et al. (2013) dataset from the high northern latitudes might not fully capture wintertime temperatures, i.e., that they are warm-season biased. If we conduct our external validation experiment assuming that terrestrial proxies above 55°N represent March–November temperatures (i.e., omitting DJF) the skill of the seasonal $U_{37}^{K'}$ experiment at 24,000 km localization improves markedly (from an R^2 of 0.27 to 0.69). However, the skill of the annual $U_{37}^{K'}$ experiment at 24,000 km localization also improves

(from an R^2 of 0.51 to 0.72). Hence the annual assumption still has higher skill (albeit by a smaller margin). Since there is clear improvement if we assume that the terrestrial proxies are mean-annual indicators, and because the choices of a March–November bias and cutoff at 55°N are somewhat arbitrary, we deem the annual experiment superior and assume that $U_{37}^{K'}$ represents an annual signature in all regions during the two Pliocene timeslices.

Like $U_{37}^{K'}$, the Mg/Ca data can also be seasonally-biased in certain regions depending on the species. Following Tierney et al. (2019) we used the modern thermal tolerances of each of the species in our proxy collection to compute the seasonality of the Mg/Ca proxies (Table S1), and we assumed that this seasonality was the same during the Pliocene. We experimented with changing the Pliocene seasonality by computing the expected seasonal expressions based on the Pliocene model prior mean, but this had no discernible impact on the results (not shown), likely because many of the Mg/Ca data represent mean annual values in both the Pliocene and present day cases. In the modern ocean, TEX_{86} coretop data do not show any strong seasonal expressions (Tierney & Tingley, 2014) so we assume this proxy represents mean annual SST in all of our timeslices.

References

Delaney, M. L., Bé, A. W., & Boyle, E. A. (1985). Li, Sr, Mg, and Na in foraminiferal calcite shells from laboratory culture, sediment traps, and sediment cores. *Geochimica et Cosmochimica Acta*, *49*(6), 1327–1341.

Evans, D., Brierley, C., Raymo, M. E., Erez, J., & Müller, W. (2016). Planktic foraminifera shell chemistry response to seawater chemistry: Pliocene–Pleistocene seawater Mg/Ca, temperature and sea level change. *Earth and Planetary Science*

Letters, *438*, 139–148.

Evans, D., Erez, J., Oron, S., & Müller, W. (2015). Mg/Ca-temperature and seawater-test chemistry relationships in the shallow-dwelling large benthic foraminifera Operculina ammonoides. *Geochimica et Cosmochimica Acta*, *148*, 325–342.

Evans, D., & Müller, W. (2012). Deep time foraminifera Mg/Ca paleothermometry: Nonlinear correction for secular change in seawater Mg/Ca. *Paleoceanography*, *27*(4).

Gothmann, A. M., Stolarski, J., Adkins, J. F., Schoene, B., Dennis, K. J., Schrag, D. P., ... Bender, M. L. (2015). Fossil corals as an archive of secular variations in seawater chemistry since the Mesozoic. *Geochimica et Cosmochimica Acta*, *160*, 188–208.

Hauzer, H., Evans, D., Müller, W., Rosenthal, Y., & Erez, J. (2018). Calibration of Na partitioning in the calcitic foraminifer Operculina ammonoides under variable Ca concentration: Toward reconstructing past seawater composition. *Earth and Planetary Science Letters*, *497*, 80–91.

Holland, K., Branson, O., Haynes, L. L., Hönisch, B., Allen, K. A., Russell, A. D., ... Eggins, S. M. (2020). Constraining multiple controls on planktic foraminifera Mg/Ca. *Geochimica et Cosmochimica Acta*, *273*, 116–136.

Mucci, A., & Morse, J. W. (1983). The incorporation of Mg^{2+} and Sr^{2+} into calcite overgrowths: influences of growth rate and solution composition. *Geochimica et Cosmochimica Acta*, *47*(2), 217–233.

Rosenthal, Y., Bova, S., & Zhou, X. (2022). A User Guide for Choosing Planktic Foraminiferal Mg/Ca-Temperature Calibrations. *Paleoceanography and Paleoclimatology*, *37*(6), e2022PA004413.

Salzmann, U., Dolan, A. M., Haywood, A. M., Chan, W.-L., Voss, J., Hill, D. J., ...

others (2013). Challenges in quantifying Pliocene terrestrial warming revealed by data–model discord. *Nature Climate Change*, *3*(11), 969–974.

Segev, E., & Erez, J. (2006). Effect of Mg/Ca ratio in seawater on shell composition in shallow benthic foraminifera. *Geochemistry, Geophysics, Geosystems*, *7*(2).

Ternois, Y., Sicre, M.-A., Boireau, A., Marty, J.-C., & Miquel, J.-C. (1996). Production pattern of alkenones in the Mediterranean Sea. *Geophysical Research Letters*, *23*(22), 3171–3174.

Tierney, J. E., Malevich, S. B., Gray, W., Vetter, L., & Thirumalai, K. (2019). Bayesian calibration of the Mg/Ca paleothermometer in planktic foraminifera. *Paleoceanography and Paleoclimatology*, *34*(12), 2005–2030.

Tierney, J. E., & Tingley, M. P. (2014). A Bayesian, spatially-varying calibration model for the TEX_{86} proxy. *Geochimica et Cosmochimica Acta*, *127*, 83–106.

Tierney, J. E., & Tingley, M. P. (2018). BAYSPLINE: A new calibration for the alkenone paleothermometer. *Paleoceanography and Paleoclimatology*, *33*(3), 281–301.

Tindall, J. C., Haywood, A. M., Salzmann, U., Dolan, A. M., & Fletcher, T. (2022). The warm winter paradox in the Pliocene northern high latitudes. *Climate of the Past*, *18*(6), 1385–1405.

Zhou, X., Rosenthal, Y., Haynes, L., Si, W., Evans, D., Huang, K.-F., ... Erez, J. (2021). Planktic foraminiferal Na/Ca: A potential proxy for seawater calcium concentration. *Geochimica et Cosmochimica Acta*, *305*, 306–322.

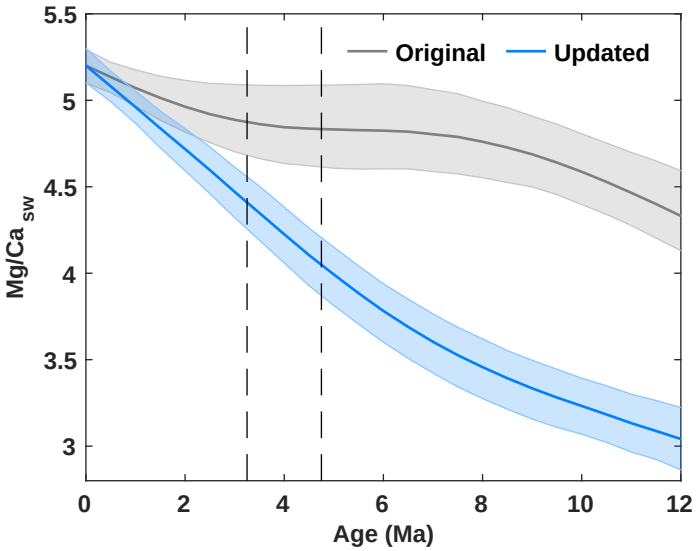

Figure S1. Updated evolution of Mg/Ca_{sw} (in blue) over the late Neogene after inclusion of the Na/Ca data (Zhou et al., 2021; Rosenthal et al., 2022), compared to the original Mg/Ca_{sw} reconstruction released with BAYMAG in 2019 (in gray). Dashed vertical lines denote the two timeslices represented in plioDA, the mid-Pliocene (3.25 Ma) and the early Pliocene (4.75 Ma).

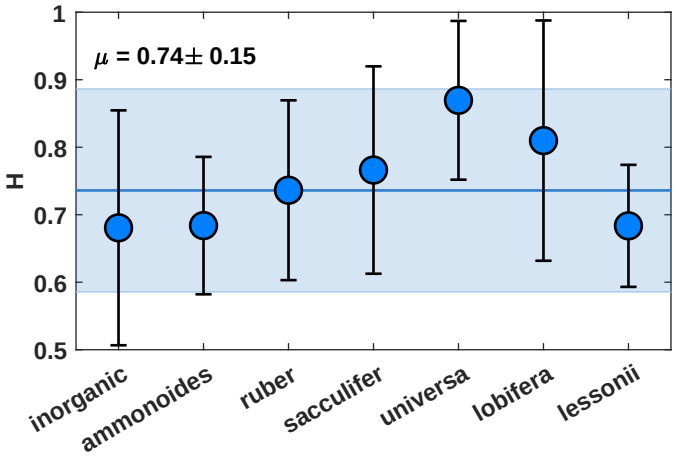

Figure S2. Computed power component H ($Mg/Ca_{foram} = a * Mg/Ca_{sw}^{H}$) from individual studies (blue dots, error bars represent the 95% confidence intervals). The inter-study mean (blue line) and 2σ error (shaded blue area) are given in the upper left corner. inorganic = Mucci and Morse (1983), ammonoides = Evans et al. (2015) and (Hauzer et al., 2018), ruber = Evans et al. (2016), sacculifer = Delaney et al. (1985), universa = Holland et al. (2020), lobifera and lessonii = Segev and Erez (2006).

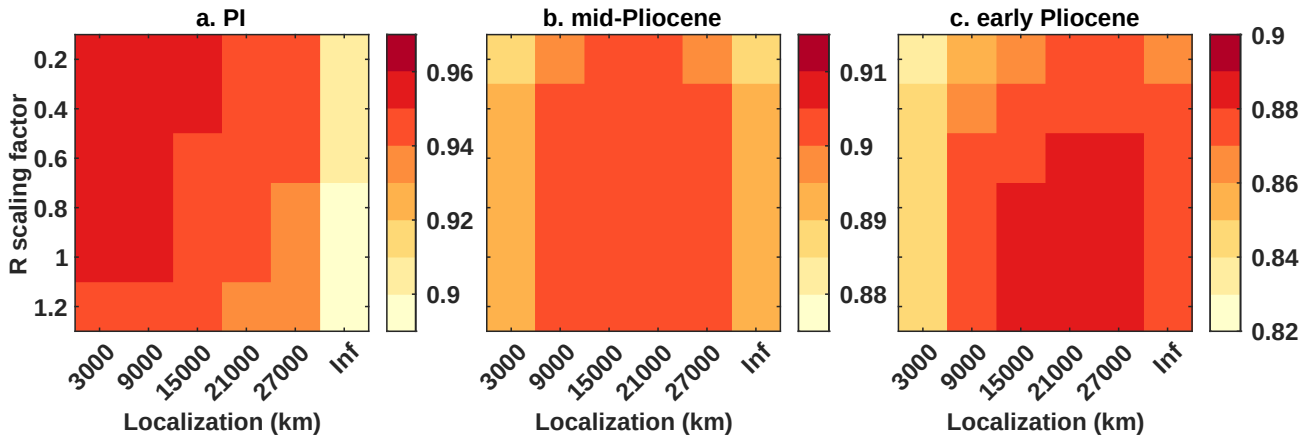

Figure S3. Results of the internal validation leave-one-out experiments for each reconstructed timeslice. Colors represent the R^2 value computed between the actual value and the posterior prediction for the set of excluded proxies. "R scaling factor" represents the number multiplied by R, with 1 indicating no scaling, a value greater than 1 an increase, and a value less than 1 a decrease.

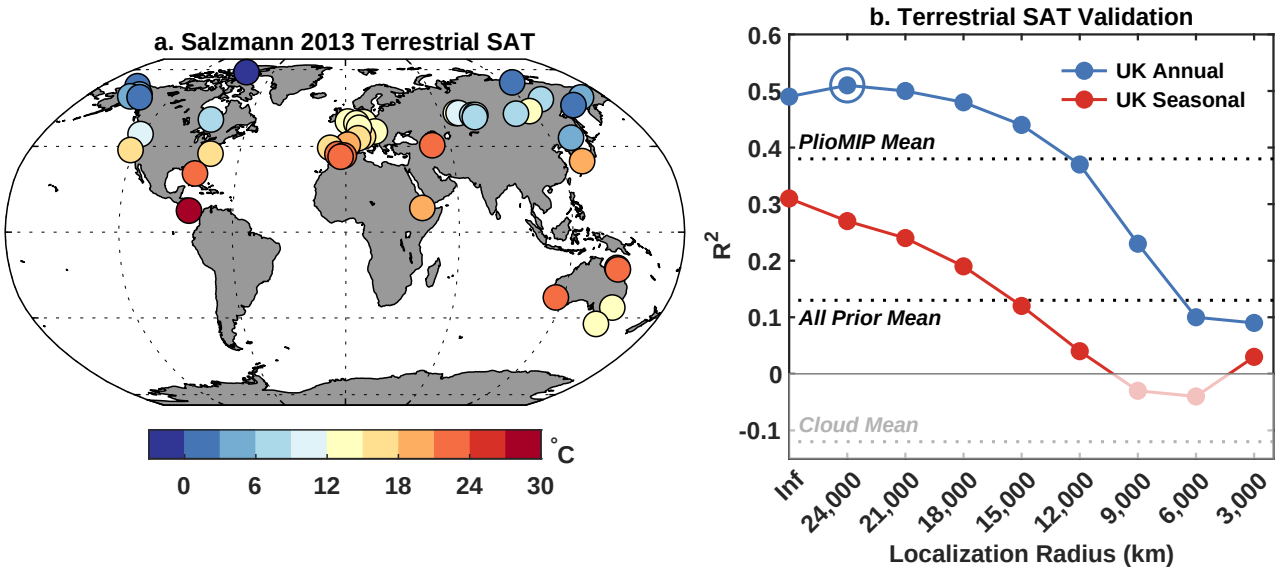

Figure S4. External validation of the PlioDA with terrestrial surface air temperature (SAT) proxies. a) The Salzmann et al. (2013) compilation of mid-Pliocene inferred terrestrial SAT; b) Validation R^2 values for the data in a) versus mid-Pliocene PlioDA posterior SAT across a range of localization radii. Blue values assume $U_{37}^{K'}$ is an annual signal (the configuration used for the final PlioDA is circled); red values assume the same seasonality as today. Dotted lines show the predictive skill of the mean of all priors, the mean of the PlioMIP priors, and the mean of the cloud priors. Values below $R^2 = 0$ indicate solutions that have no effective skill; i.e. their performance is worse than the mean of the data.

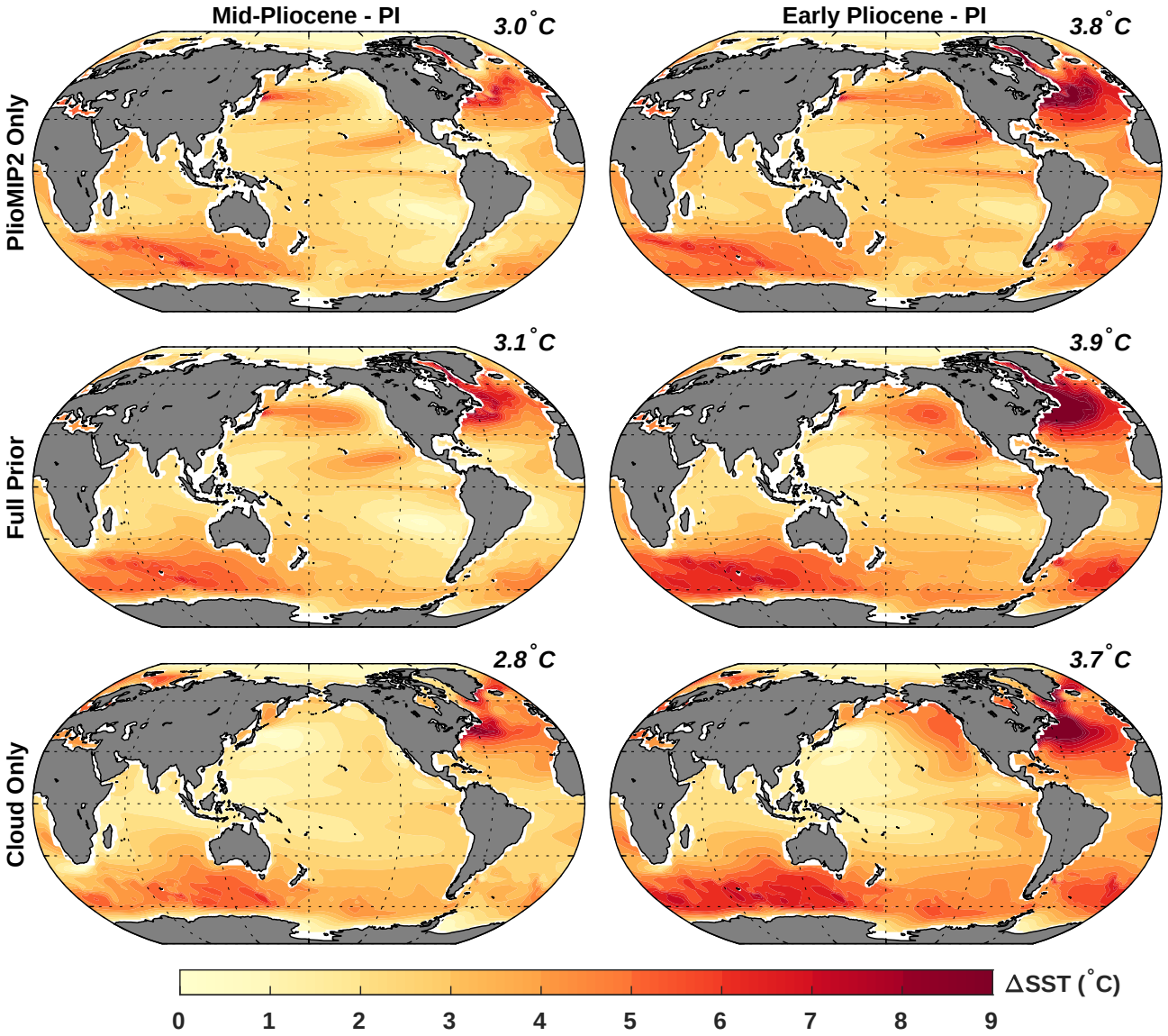

Figure S5. Mean posterior SST changes for the mid-Pliocene and early Pliocene when only the PlioMIP2 simulations are used as the model prior (top row), when the full prior is used (middle row, same as main text), and when only the cloud simulations are used (bottom row). Global mean SST is shown in the upper right hand corner of each plot.

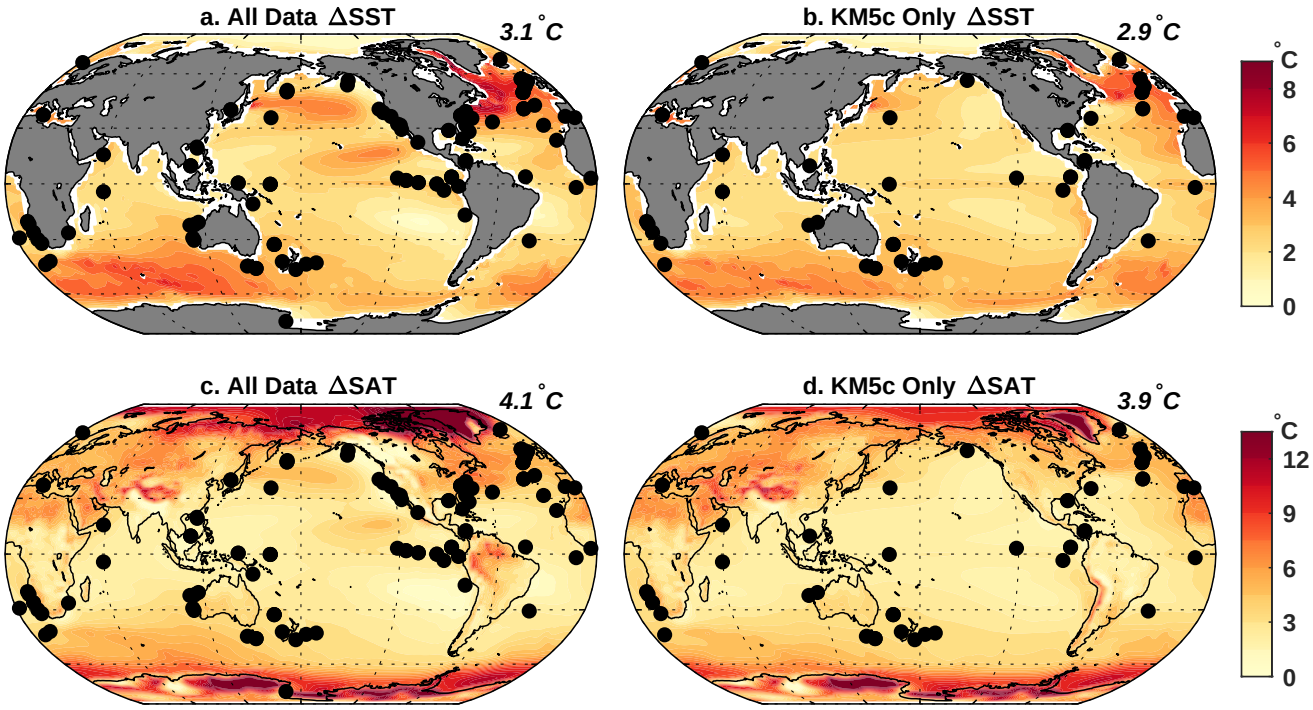

Figure S6. Mean posterior SST changes (a, b) and SAT changes (c, d) for the mid-Pliocene timeslice when all data within our chosen time window of 3–3.5 Ma are used (left) versus when only data securely dated to the KM5c interglacial are used (right). Global mean values for the SST and SAT changes are shown in the upper right corner of each plot. Dots indicate the proxy locations for each DA experiment.

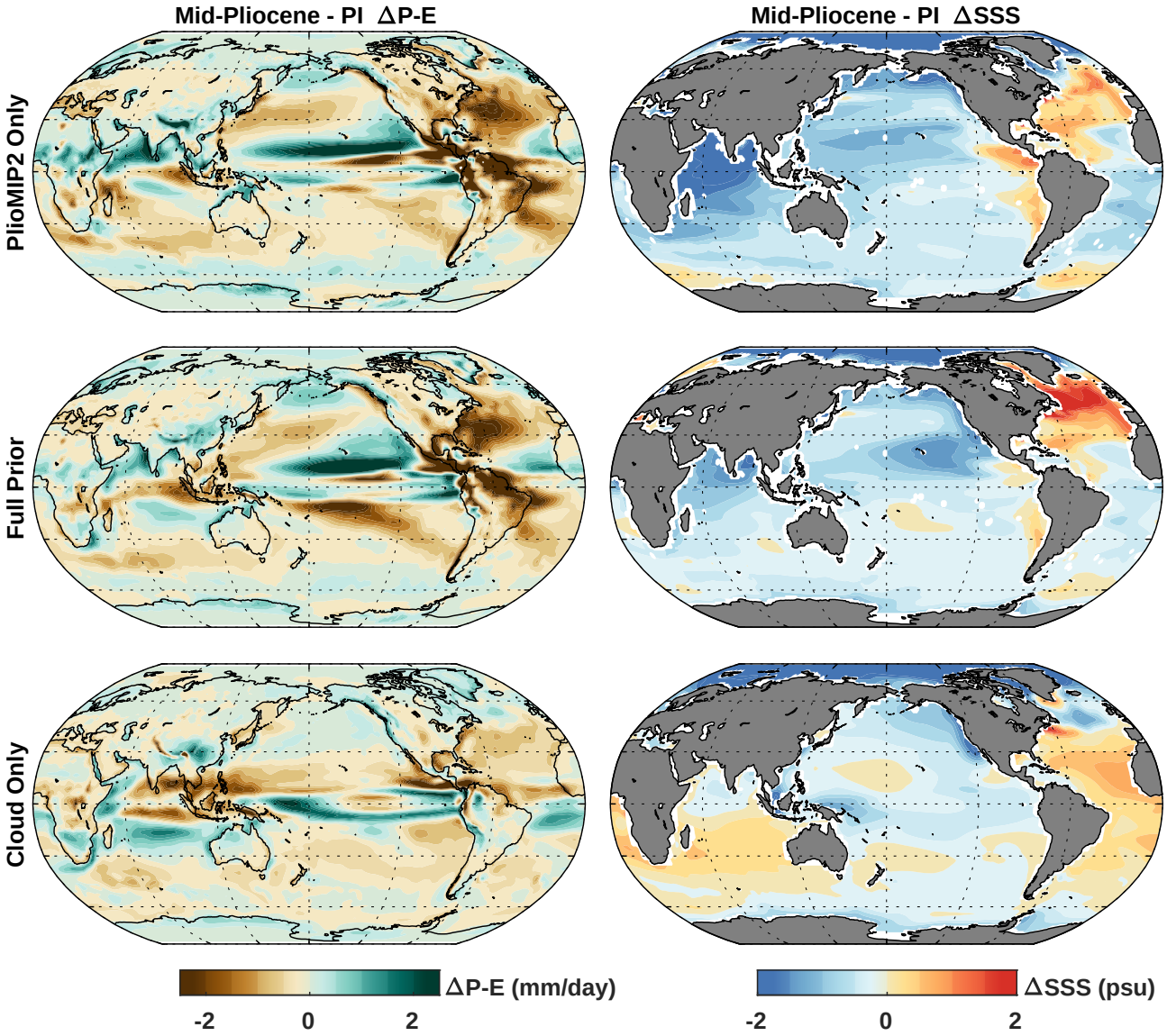

Figure S7. Mean posterior changes in $P-E$ and sea-surface salinity (SSS) for the mid-Pliocene timeslice when only the PlioMIP2 simulations are used as the model prior (top row), when the full prior is used (middle row, same as main text), and when only the cloud simulations are used (bottom row).

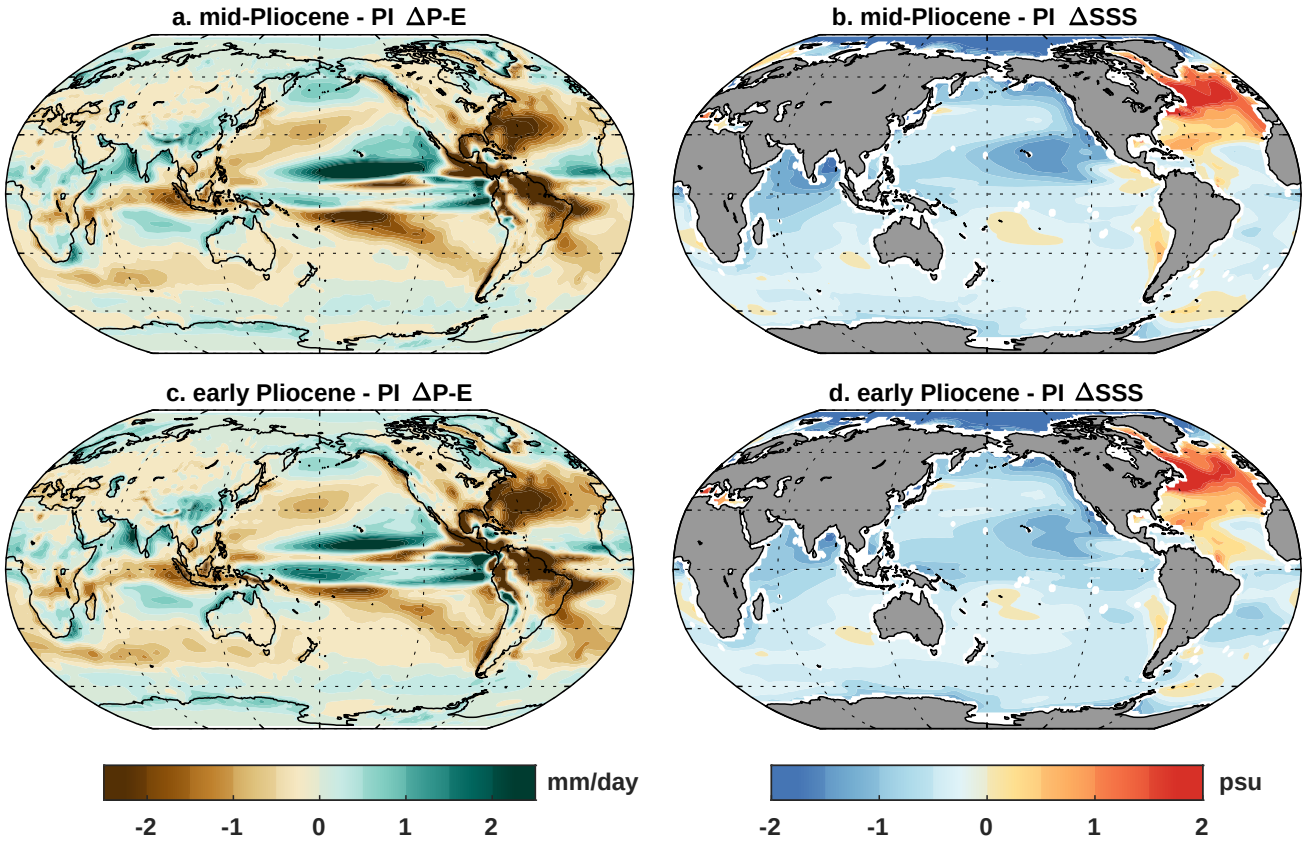

Figure S8. plioDA mid-Pliocene−preindustrial change in precipitation minus evaporation ($\Delta P - E$) and sea-surface salinity (ΔSSS) compared to early Pliocene $\Delta P - E$ and ΔSSS.

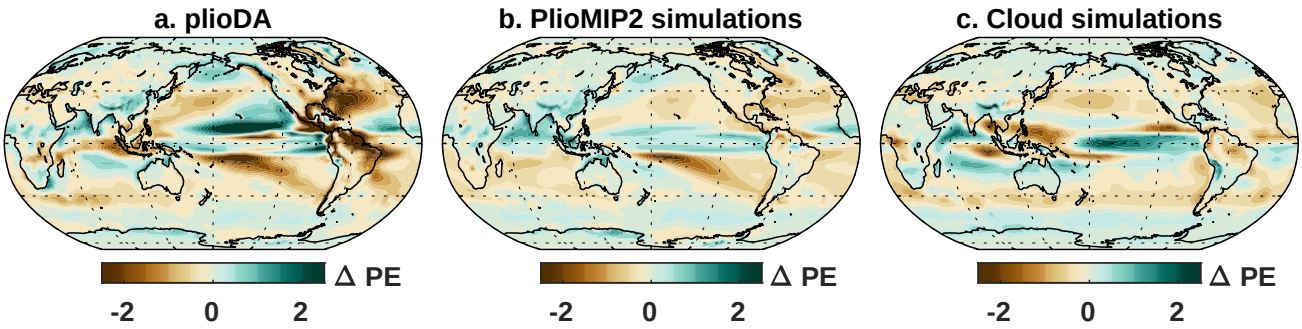

Figure S9. The plioDA mid-Pliocene posterior mean for $\Delta P - E$ (a.) compared to the multi-model means for the PlioMIP2 simulations (b.) and the cloud simulations (c.), respectively.

Table S4. Changes in the global mean temperature (ΔGMST) and tropical Pacific zonal gradient (ΔPZG) for the mid-Pliocene (mP) and early Pliocene (eP) timeslices (relative to preindustrial conditions) using different model priors. Mean values are given with 95% confidence intervals in parentheses.

Model Prior	mP ΔGMST	eP ΔGMST	mP ΔPZG	eP ΔPZG
Full Prior	4.1°C (3.0–5.3°C)	4.8°C (3.6–6.2°C)	−0.8°C (−2.3–0.4°C)	−2.3°C (−3.9–−1.1°C)
PlioMIP2 Only	4.1°C (3.2–5.2°C)	4.8°C (3.8–6.1°C)	−0.4°C (−1.6–0.6°C)	−1.6°C (−2.8–−0.6°C)
Cloud Only	3.8°C (2.9–4.9°C)	4.6°C (3.7–5.9°C)	−1.1°C (−2.1–−0.4°C)	−2.7°C (−3.8–−1.9°C)

Table S1. (Uploaded separately). Description of the paleoclimate sea-surface temperature proxy data used in plioDA, including the site name, location, paleocoordinates, what types of proxies are available at each site for each target time period, the species of foraminifera and assumed seasonality for Mg/Ca proxy data, the method of determination for the preindustrial SST proxy estimate, original study references, and doi links.

Table S2. (Uploaded separately). Description of the paleoclimate sea-surface temperature proxy data that were excluded from plioDA, including the site name, location, what types of proxies are available at each site for each target time period, the species of foraminifera and assumed seasonality for Mg/Ca proxy data, original study references, doi links, the reason for exclusion, and doi of the exclusion reason (if applicable).

Table S3. (Uploaded separately). Description of the model simulations used in plioDA, including the model name, resolution of the atmosphere and ocean components, the boundary conditions used for the Pliocene experiments, run names, the simulated change in global mean surface temperature (GMST) relative to the preindustrial experiment, original study references, and doi links.