

Glacial cooling and climate sensitivity revisited

<https://doi.org/10.1038/s41586-020-2617-x>

Received: 20 December 2019

Accepted: 15 June 2020

Published online: 26 August 2020

 Check for updates

Jessica E. Tierney^{1✉}, Jiang Zhu^{2,3}, Jonathan King¹, Steven B. Malevich¹, Gregory J. Hakim⁴ & Christopher J. Poulsen³

The Last Glacial Maximum (LGM), one of the best studied palaeoclimatic intervals, offers an excellent opportunity to investigate how the climate system responds to changes in greenhouse gases and the cryosphere. Previous work has sought to constrain the magnitude and pattern of glacial cooling from palaeothermometers^{1,2}, but the uneven distribution of the proxies, as well as their uncertainties, has challenged the construction of a full-field view of the LGM climate state. Here we combine a large collection of geochemical proxies for sea surface temperature with an isotope-enabled climate model ensemble to produce a field reconstruction of LGM temperatures using data assimilation. The reconstruction is validated with withheld proxies as well as independent ice core and speleothem $\delta^{18}\text{O}$ measurements. Our assimilated product provides a constraint on global mean LGM cooling of -6.1 degrees Celsius (95 per cent confidence interval: -6.5 to -5.7 degrees Celsius). Given assumptions concerning the radiative forcing of greenhouse gases, ice sheets and mineral dust aerosols, this cooling translates to an equilibrium climate sensitivity of 3.4 degrees Celsius (2.4 – 4.5 degrees Celsius), a value that is higher than previous LGM-based estimates but consistent with the traditional consensus range of 2 – 4.5 degrees Celsius^{3,4}.

Palaeoclimatologists have long sought to refine estimates of temperature changes during the LGM, as both a benchmark for climate models and a constraint on Earth's climate sensitivity. In the 1970s, the Climate Long-Range Investigation, Mapping and Prediction (CLIMAP) project collated assemblages of foraminifera, radiolarians and coccolithophores and used transfer functions to create maps of seasonal sea surface temperatures (SSTs) for the LGM¹. Along with geological constraints on sea level and ice sheet extent, these maps were used as boundary conditions for pioneering atmospheric global climate model (GCM) simulations—the first Paleoclimate Modelling Intercomparison Project (PMIP)². Three decades later, the Multiproxy Approach for the Reconstruction of the Glacial Ocean Surface (MARGO) project remapped the LGM oceans using foraminiferal, radiolarian, diatom and dinoflagellate transfer functions and two geochemical proxies—the unsaturation index of alkenones (U_{37}^K) and the Mg/Ca ratio of planktic foraminifera². This product has served as a touchstone for model–data comparison in the second and third phases of PMIP (PMIP2 and PMIP3)⁶ as well as calculations of climate sensitivity⁷.

In spite of this extensive work, estimates of global cooling during the LGM remain poorly constrained due to proxy uncertainties and methodological limitations. Microfossils occasionally present ‘no-analogue’ assemblages; that is, groups of species that are not observed today and therefore are difficult to interpret. In the LGM in particular, no-analogue assemblages appear in north Atlantic dinocysts¹ and tropical Pacific foraminifera⁸, and have cast doubt on the CLIMAP and MARGO inference of relatively mild LGM cooling in the tropics and subtropics^{9–11}. Likewise, geochemical proxies are subject

to seasonal biases and sensitivity to non-thermal controls, all of which affect calculated SSTs^{12,13}. Beyond proxy uncertainties, the data from the LGM present a methodological challenge in that they are not evenly distributed in space; the SST data cluster near coasts, where sediment accumulation rates are high. This heterogeneous sample distribution complicates the calculation of both regional and global average values. Furthermore, the translation of changes in SST to global mean surface air temperature (GMST)—the quantity needed for calculations of climate sensitivity—requires the use of an uncertain scaling factor¹⁴. As a result of these uncertainties, estimates of the change in LGM GMST (ΔGMST) range from -1.7 °C to -8.0 °C (refs. ^{7,14–20}), yielding poorly bounded estimates of climate sensitivity of 1 – 6 °C per doubling of CO_2 (ref. ²¹).

Here, we infer the magnitude and spatial pattern of LGM cooling using geochemical SST proxies, Bayesian proxy forward models, isotope-enabled climate model simulations and offline data assimilation. The SST proxy observations are assimilated with new simulations conducted with the isotope-enabled Community Earth System Model (iCESM)²². The resulting estimates of ΔGMST are combined with published constraints on radiative forcing to produce probabilistic estimates of climate sensitivity based on the LGM climate state.

Paleoclimate data assimilation

Our data collection consists of 956 LGM (23–19 kyr ago, ka) and 879 late Holocene (4–0 ka) SST proxies (Fig. 1). The proxy values are averages over the chosen time intervals and in many cases come from

¹Department of Geosciences, University of Arizona, Tucson, AZ, USA. ²Climate and Global Dynamics Laboratory, National Center for Atmospheric Research, Boulder, CO, USA. ³Department of Earth and Environmental Sciences, University of Michigan, Ann Arbor, MI, USA. ⁴Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA. ✉e-mail: jesst@arizona.edu

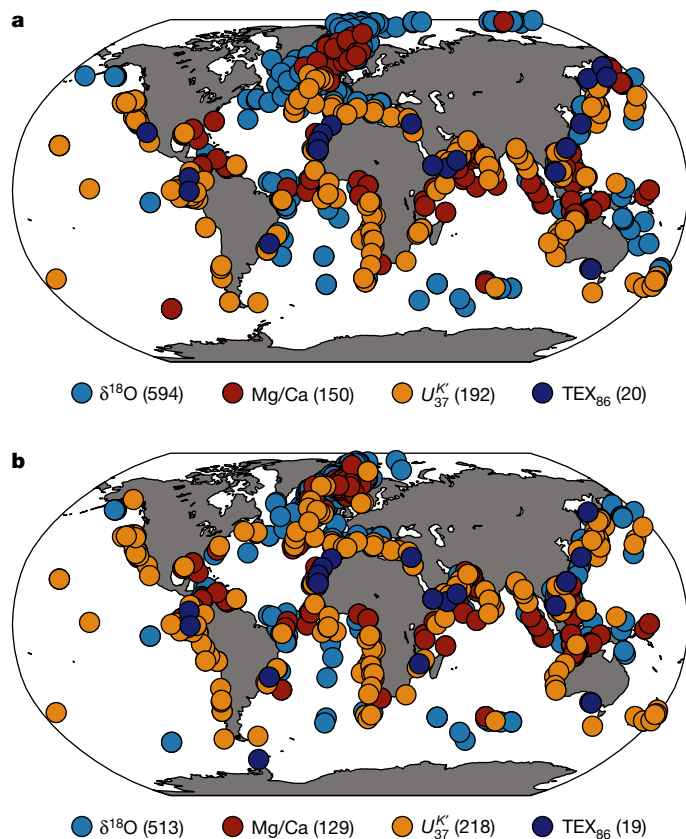


Fig. 1 | Locations of geochemical SST proxies used for the LGM climate reconstruction. a. Proxy sites for the LGM (956). **b.** Proxy sites for the late Holocene (LH; 879). Proxies are colour-coded by type and the number of each is shown in parentheses.

the same location, such that the total numbers of unique locations for each timeslice are 636 and 664, respectively (see Methods for further details). For the purposes of this study, the late Holocene average is interpreted as representative of the pre-industrial (PI) climate state, and is the benchmark from which we compute LGM cooling (see Methods). Distinct from previous work, this study focuses exclusively on geochemical proxies for SST; specifically, U_{37}^K , the TetraEther index of 86 carbons (TEX_{86}), $\delta^{18}\text{O}$ and Mg/Ca. We previously developed Bayesian models for each of these proxy systems^{12,13,23,24}, which enables us to propagate proxy uncertainties and seasonal bias into the data assimilation. Although including marine assemblage data would improve spatial coverage, the outstanding no-analogue problems and lack of comparable Bayesian models prevent us from using these data in the framework presented here. Likewise, we do not assimilate terrestrial temperature estimates, because many of these are also based on pollen or other microfossil assemblages.

To circumvent problems associated with spatial representation and averaging, we use an offline data assimilation technique²⁵ (see Methods) to blend information from proxies with full-field dynamical constraints from iCESM. The assimilation begins with an ensemble 'prior' of possible climate states taken from the model; in our case, these are 50 yr average states from simulations of the glacial state (21 and 18 ka) and the late Holocene (3 ka and PI) (see Methods). The water-isotope-enabled model simulations facilitate the direct assimilation of $\delta^{18}\text{O}$ data (comprising 60% of our collection; Fig. 1), without the need to rely on empirical relationships between seawater $\delta^{18}\text{O}$ and salinity derived from present-day observations. At locations where there are proxy data, values from the ensemble prior are translated into proxy units using

our Bayesian forward models to calculate the 'innovation'—the difference between the observed proxy value and the computed value from the model ensemble. The innovation is weighted by the Kalman gain, which considers both the covariance of the proxy with the rest of the climate fields and the uncertainties in the proxy observation and the model ensemble. The weighted innovation is then added to the prior ensemble to produce an ensemble of assimilated (analysis) climate states (see Methods for a mathematical description). For each time interval (LGM and late Holocene), we conducted 25 assimilation experiments with a 40-member model ensemble in which we withheld 25% of the proxy data at random to calculate verification statistics (see Methods). These collectively yield a total of 1,000 ensemble realizations of LGM and late-Holocene climates.

Spatial patterns in LGM cooling

The assimilated SST field shows distinctive spatial patterns in LGM cooling (Fig. 2a), with changes in SST (ΔSST) in excess of -8°C in the north Atlantic, north Pacific and the Pacific sector of the Southern Ocean; enhanced cooling in eastern boundary upwelling zones; and reduced cooling in the western boundary regions. Many of these patterns are broadly consistent with CLIMAP and MARGO; however, an important difference is that there is no evidence for warming in the subtropical gyres, a feature associated with assemblage data^{1,2} (Fig. 2a). Strong cooling near 40°N in the Pacific reflects a southward shift in the subpolar gyre; this agrees well with independent model simulations of the LGM and a recent regional synthesis of $\delta^{18}\text{O}$ data²⁶. In the Indian Ocean, our reconstructed cooling pattern closely resembles the proxies and reflects the impact of the exposed Sunda and Sahul shelves²⁷. Previous investigations of cooling in the glacial tropical Pacific offer conflicting conclusions: some suggest enhanced cooling in the eastern equatorial Pacific (EEP; as we observe here)⁸, whereas others suggest greater cooling in the warm pool²⁸. Analysis of our proxy collection (separate from the assimilated product) indicates that there is no significant difference in the magnitude of cooling between the warm pool and EEP ($-0.2 \pm 1.0^\circ\text{C}$, 2σ , see Methods). This could reflect a limitation of the proxy network, which is biased towards the coasts (Fig. 1). The stronger cooling in the EEP in our assimilated product thus reflects the CESM prior and possibly the influence of proxies that are teleconnected to the EEP, such as those situated along the California margin.

The covariance of SST with surface air temperature (SAT) allows us to recover the latter directly from the assimilation algorithm, rather than having to scale from one to the other¹⁴. Over land, we observe the expected large cooling over the Northern Hemisphere ice sheets, but noticeably little cooling in Alaska and western Beringia (Fig. 2b). These results agree with observations that these locations remained unglaciated during the LGM²⁹ and experienced minimal cooling or even a slight warming³⁰, probably due to dynamical changes induced by the Laurentide Ice Sheet^{31,32}.

Our reconstruction provides updated estimates of tropical cooling. The glacial change in SAT across the tropics (30°S – 30°N) is -3.9°C (-4.2 to -3.7°C ; 95% confidence interval, CI). This estimate is greater than the PMIP2 and PMIP3 multi-model range of -3.2 to -1.6°C (ref. ³³), but is not as large as calculations based on snowlines of tropical glaciers³⁴ and noble gases³⁵ (approximately -5°C). Our reconstruction indicates that the change in tropical SSTs was -3.5°C (-3.8 to -3.3°C , 95% CI). This value is larger than the spatial mean computed from the tropical SST proxies on their own (-2.6°C , -3.0 to -2.3°C , 95% CI), partly due to the enhanced cooling throughout the east-central tropical Pacific in the assimilated field (Fig. 2a). However, the magnitude of both the proxy- and data assimilation-inferred tropical SST cooling is far greater than CLIMAP or MARGO estimates (-0.8°C and -1.5°C , respectively) and closer to the estimates of Ballantyne et al.¹⁰ (-2.7°C) and Lea et al.³⁶ (-2.8°C).

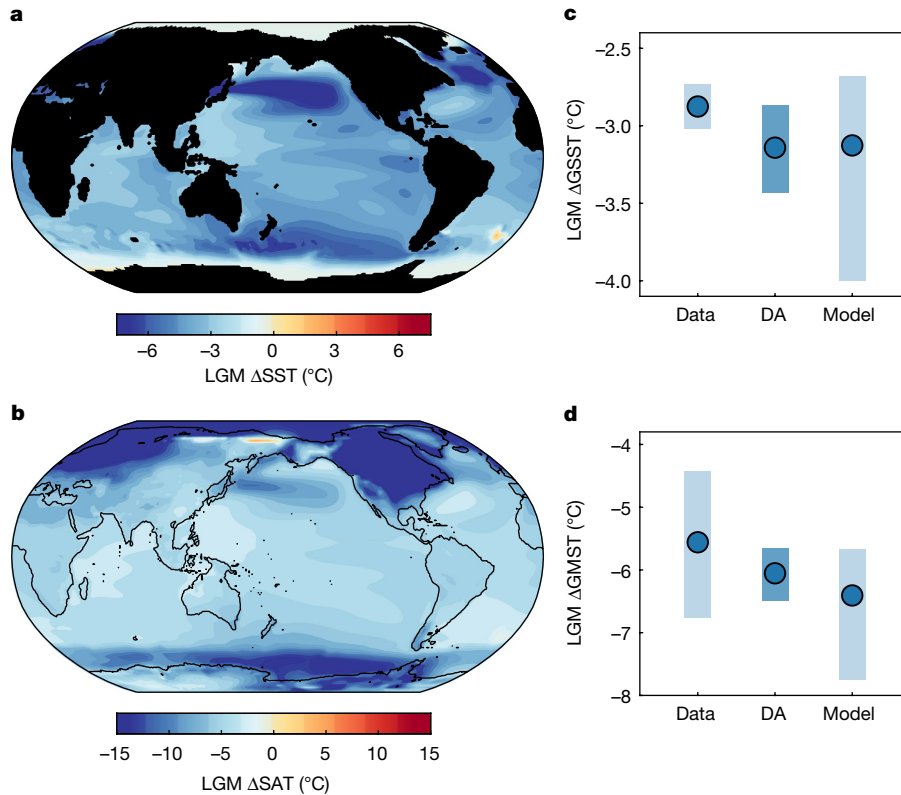


Fig. 2 | Global changes in temperature during the LGM derived from palaeoclimate data assimilation. a, LGM–late Holocene ΔSST. **b**, LGM–late Holocene ΔSAT. **c**, LGM ΔGSST. **d**, LGM ΔGMST. In **c** and **d**, dots represent the

median values and bars the 95% CI for values derived from the data, the data assimilation (DA) and the model prior (Model).

Validation with $\delta^{18}\text{O}$ of precipitation

To assess the reliability of our reconstruction, we included the $\delta^{18}\text{O}$ of precipitation ($\delta^{18}\text{O}_p$) in our model prior to compare the assimilated value to independent ice core and speleothem proxy data (see Methods). Overall, our reconstruction explains 67% of the variance in observed $\Delta\delta^{18}\text{O}_p$ (Fig. 3a). This is a marked improvement over the prior (which explains only 35% of the variance; Extended Data Fig. 1) and suggests that our reconstruction provides a reasonable estimate of global LGM climate. The improvement comes mainly from the ice core sites, which show a better match to observations after the assimilation. A notable feature captured by our reconstruction is the difference in $\Delta\delta^{18}\text{O}_p$ between ice core sites in west and east Antarctica (Fig. 3b); the latter region is warmer (Fig. 2b) and experiences less isotopic depletion. At face value, a warmer east Antarctica contradicts previous work: at Epica Dome C, ice core $\delta^{18}\text{O}$ is interpreted to indicate a change in

SAT (ΔSAT) of approximately $-8\text{ }^\circ\text{C}$ (ref. ³⁷), whereas our assimilated product indicates a more modest $5\text{ }^\circ\text{C}$ of cooling. However, the former estimate assumes that the $\delta^{18}\text{O}$ –SAT relationship remains constant in time³⁷. Isotope-enabled modelling experiments have shown that the $\delta^{18}\text{O}$ –SAT slope in Antarctica may have been different during the LGM, and strongly depends on changes in Southern Ocean SSTs³⁸.

Global temperature change

As the data assimilation technique yields spatially complete fields, we can compute both global SST (GSST) and GMST change during the LGM without needing to consider missing values or use a scaling factor¹⁴. Our calculated change in GSST (ΔGSST) is $-3.1\text{ }^\circ\text{C}$ (-3.4 to $-2.9\text{ }^\circ\text{C}$, 95% CI) (Fig. 2c). This is more tightly constrained than the model prior, which spans -4 to $-2.7\text{ }^\circ\text{C}$ (Fig. 2c), reflecting the influence of the data. The assimilated ΔGSST is slightly larger than the proxy data suggest

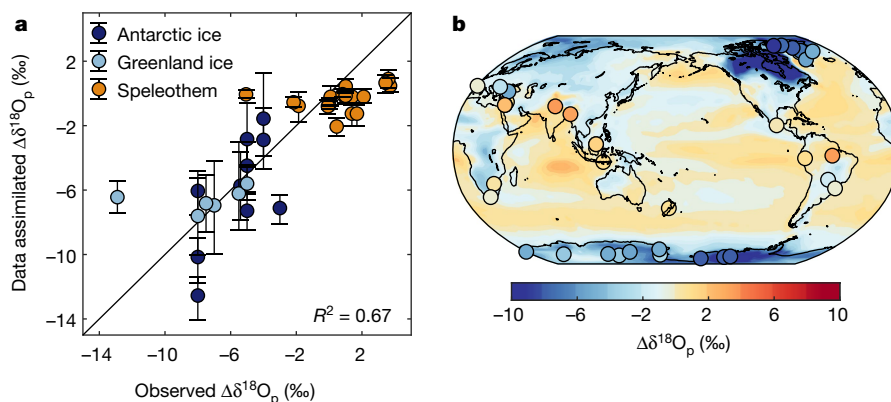


Fig. 3 | Validation of the data assimilation with independent $\delta^{18}\text{O}_p$ data. a, Observed changes in ice core (Antarctica and Greenland) and speleothem LGM–late Holocene $\delta^{18}\text{O}_p$ compared with predicted changes from the data assimilation ensemble. Dots indicate median values, error bars represent the 95% CI. The R^2 value is shown in the lower right corner. **b**, Map of the median changes in LGM–late Holocene $\delta^{18}\text{O}_p$ from the data assimilation, overlain with ice core and speleothem observations (dots). Speleothem $\delta^{18}\text{O}$ values have been converted from $\delta^{18}\text{O}$ of calcite or aragonite to $\delta^{18}\text{O}_p$ (in ‰ Vienna standard mean ocean water, VSMOW) before plotting (see Methods).

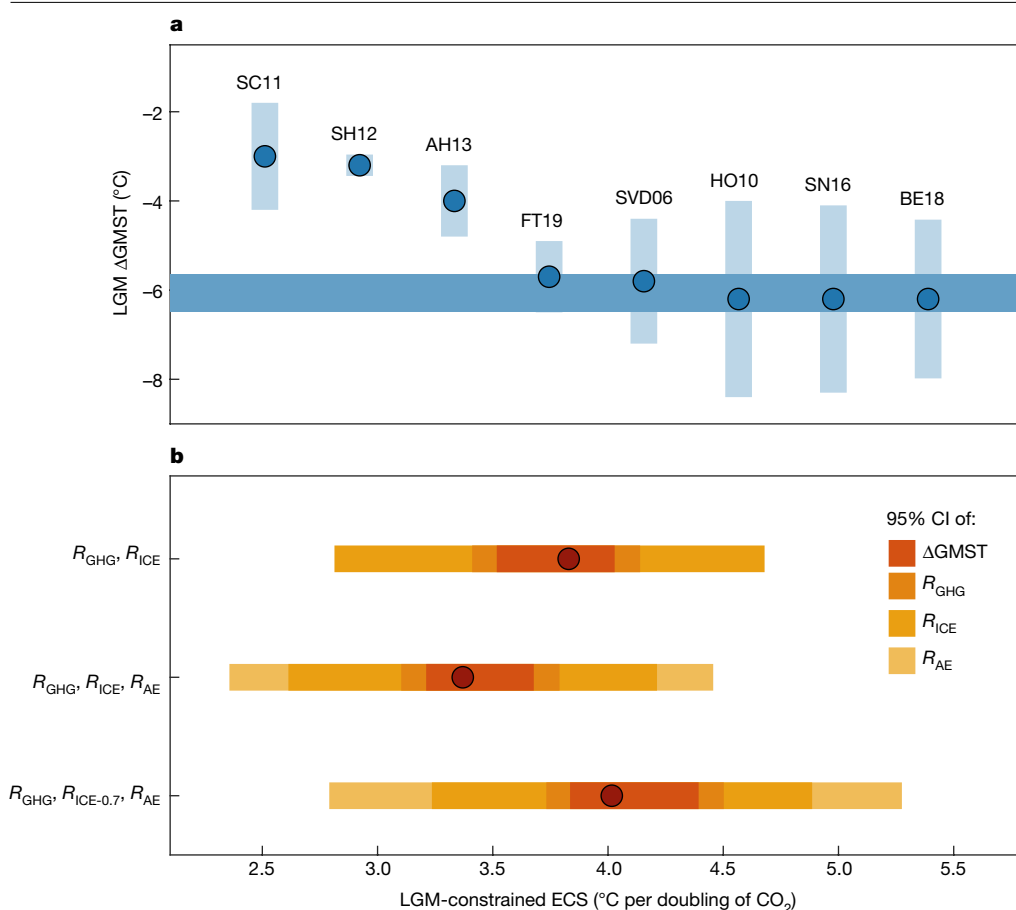


Fig. 4 | LGM global temperature change and climate sensitivity derived from data assimilation. **a**, Estimates of ΔGMST from previous studies (vertical bars represent the 95% CI, dots show the median) compared with the data assimilation result (the height of the horizontal bar represents the 95% CI). SC11, ref. ⁷; SH12, ref. ¹⁷; AH13, ref. ¹⁸; FT19, ref. ²⁰; SVD06, ref. ¹⁵; HO10, ref. ¹⁶; SN16, ref. ¹⁴; BE18, ref. ¹⁹. **b**, LGM-constrained ECS determined using the ΔGMST from data assimilation. Dots indicate median values and coloured bars indicate the 95% CI associated with ΔGMST and R_{GHG} , R_{ICE} and R_{AE} . The lower bar shows the range of ECS with an ice-sheet efficacy of 0.7.

(-2.9°C , -3.0 to -2.7°C , 95% CI, Fig. 2c), a difference that arises from biases due to incomplete and uneven proxy data sampling. The ΔGSST from data assimilation agrees well with estimates of glacial cooling based on a subset of SST proxies spanning multiple glacial–interglacial cycles (-3.1°C)²⁰, but is much larger than CLIMAP (-1.2°C) and MARGO (-1.9°C) estimates, as well as previous LGM ocean state estimates that used the MARGO dataset (-2.2 to -2.0°C)^{39,40}.

The change in GMST in the assimilated product is -6.1°C (-6.5 to -5.7°C , 95% CI) (Fig. 2d). This result agrees with a number of previous studies, including those that estimated LGM cooling from a restricted network of proxies spanning multiple glacial–interglacial cycles^{14,20}, noble gas measurements in ice cores¹⁹ and changes in tropical SSTs^{15,16} (Fig. 4a). However, our calculated ΔGMST does not overlap with the estimates that used the MARGO product^{7,18} (-4 to -2°C), or an average of time-continuous marine and terrestrial temperature proxies¹⁷ (Fig. 4a).

One caveat of our results is that they are based on assimilation with an ensemble prior from a single climate model. As discussed above, the structure of the temperature fields in CESM directly shapes our assimilated fields. In addition, CESM1.2 simulates a larger LGM cooling (6.7°C) than the PMIP2 and PMIP3 models (5.9 to 3.1°C)²¹, although CESM ECS in our configuration is well within the multiple model range of the Coupled Model Intercomparison Project Phase 5⁴¹ (see Methods). An assimilation with a model with different covariance structures and a more moderate global cooling could yield a different result. Future work should thus test our results by assimilating data across different models, preferably with water isotope capabilities. This said, several lines of evidence suggest that our inferred median value of approximately -6°C is realistic. First, as already noted, our result agrees well with a number of previous studies (Fig. 4a). Second, our data-only GMST estimate has a median value of -5.6°C and the assimilated estimate falls entirely within its error bounds (Fig. 2d). Finally, an assessment of bias in the model prior, based on rank histograms of withheld proxies,

suggests minimal bias in the mean (Extended Data Fig. 2). However, the uncertainties associated with our ΔGMST estimate could be too narrow due to a lack of structural variability in the model prior (see Methods and Extended Data Fig. 2).

Equilibrium climate sensitivity

Our GMST results can be used to provide updated constraints on ECS. To calculate an ECS that approximates the classical ‘Charney’ definition, we must consider, in addition to GHG forcing (ΔR_{GHG}), the slow feedback processes that affect LGM climate, which (following ref. ⁴²) are treated here as radiative forcings. These include albedo changes associated with the expanded land ice and lowered sea level (ΔR_{ICE}) and increases in mineral dust aerosols (ΔR_{AE}). Although vegetation changes could also impact LGM climate, we do not consider these here because biome reconstructions are poorly defined outside of the northern high latitudes⁴³ and little is known about the associated radiative forcing⁴⁴ and efficacy (effectiveness in changing surface temperature). Likewise, we do not consider non-dust aerosol emissions, as these are not well constrained for the LGM.

We estimate ΔR_{GHG} , ΔR_{ICE} and ΔR_{AE} from published syntheses and propagate uncertainties associated with these into the calculations of ECS (see Methods). As these estimates derive from simplified expressions or multi-model estimates, they are relatively independent from CESM. We show results with and without ΔR_{AE} for comparison with previous work (Fig. 4b). Without dust aerosols, ECS is 3.8°C (2.8 – 4.7°C , 95% CI); with dust aerosols, ECS is 3.4°C (2.4 – 4.5°C , 95% CI). With our ΔGMST from data assimilation, global temperature change is no longer the primary source of uncertainty; it accounts for about 25% of the 95% CI in each estimate (Fig. 4b). Instead, most of the uncertainty comes from the forcings (Fig. 4b). ΔR_{GHG} and the forcing associated with a doubling of CO_2 can be directly estimated from ice core GHG concentrations and

simplified equations (-2.48 W m^{-2} and 3.80 W m^{-2} , respectively)⁴⁵ but still carry uncertainties and account for about 10% of the 95% CI. ΔR_{ICE} is estimated from CESM and available PMIP2 and PMIP3 simulations (see Methods) and its value varies between -2.6 and -5.2 W m^{-2} , thus accounting for ~40% of the 95% CI. Finally, although it is well known that the LGM atmosphere was dustier⁴⁶, the magnitude of ΔR_{AE} also varies widely across models (0 to -2 W m^{-2})⁴⁶ and contributes ~25% of the 95% CI.

It has recently been argued that ice sheets are not as ‘effective’ as GHGs in terms of global radiative forcing, because their impact might be concentrated at high latitudes^{47–49}. We explore this possibility by presenting a distribution of ECS with a reduced ice sheet efficacy (ϵ) of 0.7 (rather than 1.0 as used above; see Methods) (Fig. 4b). Under this scenario, the median ECS rises to $4.0 \text{ }^\circ\text{C}$ with a 95% CI of $2.8\text{--}5.3 \text{ }^\circ\text{C}$. However, these values should be interpreted with caution because ϵ is an under-constrained quantity.

Despite the broad distributions that uncertainties in the radiative forcings yield, we can use these new calculations to make probabilistic statements concerning ECS. Given the uncertainty space explored here, the LGM data suggest that ECS is virtually certain (>99% probability) to be above $2.2 \text{ }^\circ\text{C}$ and below $5.5 \text{ }^\circ\text{C}$, with the latter only plausible under a condition of low ϵ . Assuming $\epsilon = 1$, ECS is very likely (90% probability) to be between $2.5 \text{ }^\circ\text{C}$ and $4.3 \text{ }^\circ\text{C}$. These are tighter constraints on LGM-constrained ECS than those stated in Assessment Report 5 of the Intergovernmental Panel on Climate Change (IPCC) ($1\text{--}6 \text{ }^\circ\text{C}$ per doubling of CO_2)²¹, and are in excellent agreement with the traditional consensus range of $2\text{--}4.5 \text{ }^\circ\text{C}$ (refs.^{3,4}). Unlike previous work, we find little evidence that ECS estimates based on the LGM climate are abnormally low^{7,44}. ECS is unlikely to remain constant across climate states; palaeoclimate and modelling evidence suggests that it scales with background temperature, with lower values during cold climates and higher values during warm states⁵⁰. Taking this into consideration, our LGM results place a strong constraint on the minimum ECS in the climate system, which is almost certainly greater than $2 \text{ }^\circ\text{C}$, and is more likely to lie between 3 and $4 \text{ }^\circ\text{C}$.

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at <https://doi.org/10.1038/s41586-020-2617-x>.

- CLIMAP Project Members The surface of the Ice-Age Earth. *Science* **191**, 1131–1137 (1976).
- MARGO Project Members Constraints on the magnitude and patterns of ocean cooling at the Last Glacial Maximum. *Nat. Geosci.* **2**, 127–132 (2009).
- Charney, J. G. et al. *Carbon Dioxide and Climate: A Scientific Assessment* (National Academy of Sciences, 1979).
- Knutti, R. & Hegerl, G. C. The equilibrium sensitivity of the Earth's temperature to radiation changes. *Nat. Geosci.* **1**, 735–743 (2008).
- Joussaume, S. & Taylor, K. *Status of the Paleoclimate Modeling Intercomparison Project (PMIP)* (WMO, 1995).
- Braconnot, P. et al. Evaluation of climate models using palaeoclimatic data. *Nat. Clim. Change* **2**, 417–424 (2012).
- Schmittner, A. et al. Climate sensitivity estimated from temperature reconstructions of the Last Glacial Maximum. *Science* **334**, 1385–1388 (2011).
- Mix, A. C., Morey, A. E., Pisias, N. G. & Hostetler, S. W. Foraminiferal faunal estimates of paleotemperature: circumventing the no-analog problem yields cool ice age tropics. *Paleoceanography* **14**, 350–359 (1999).
- Crowley, T. CLIMAP SSTs re-visited. *Clim. Dynam.* **16**, 241–255 (2000).
- Ballantyne, A., Lavine, M., Crowley, T., Liu, J. & Baker, P. Meta-analysis of tropical surface temperatures during the Last Glacial Maximum. *Geophys. Res. Lett.* **32**, L05712 (2005).
- Telford, R., Li, C. & Kucera, M. Mismatch between the depth habitat of planktonic foraminifera and the calibration depth of SST transfer functions may bias reconstructions. *Clim. Past* **9**, 859–870 (2013).
- Tierney, J. E. & Tingley, M. P. BAYSPLINE: a new calibration for the alkenone paleothermometer. *Paleoceanogr. Paleoclimatol.* **33**, 281–301 (2018).
- Tierney, J. E., Malevich, S. B., Gray, W., Vetter, L. & Thirumalai, K. Bayesian calibration of the Mg/Ca paleothermometer in planktic foraminifera. *Paleoceanogr. Paleoclimatol.* **34**, 2005–2030 (2019).
- Snyder, C. W. Evolution of global temperature over the past two million years. *Nature* **538**, 226–228 (2016).
- Schneider von Deimling, T., Ganopolski, A., Held, H. & Rahmstorf, S. How cold was the Last Glacial Maximum? *Geophys. Res. Lett.* **33**, L14709 (2006).
- Holden, P. B., Edwards, N., Oliver, K., Lenton, T. & Wilkinson, R. A probabilistic calibration of climate sensitivity and terrestrial carbon change in GENIE-1. *Clim. Dynam.* **35**, 785–806 (2010).
- Shakun, J. D. et al. Global warming preceded by increasing carbon dioxide concentrations during the last deglaciation. *Nature* **484**, 49–54 (2012).
- Annan, J. & Hargreaves, J. C. A new global reconstruction of temperature changes at the Last Glacial Maximum. *Clim. Past* **9**, 367–376 (2013).
- Bereiter, B., Shackleton, S., Baggenstos, D., Kawamura, K. & Severinghaus, J. Mean global ocean temperatures during the last glacial transition. *Nature* **553**, 39–44 (2018).
- Friedrich, T. & Timmermann, A. Using Late Pleistocene sea surface temperature reconstructions to constrain future greenhouse warming. *Earth Planet. Sci. Lett.* **530**, 115911 (2020).
- Masson-Delmotte, V. et al. in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. et al.) 383–464 (IPCC, Cambridge Univ. Press, 2013).
- Brady, E. et al. The connected isotopic water cycle in the Community Earth System Model Version 1. *J. Adv. Model. Earth Syst.* **11**, 2547–2566 (2019).
- Tierney, J. E. & Tingley, M. P. A Bayesian, spatially-varying calibration model for the TEX₈₆ proxy. *Geochim. Cosmochim. Acta* **127**, 83–106 (2014).
- Malevich, S. B., Vetter, L. & Tierney, J. E. Global core top calibration of $\delta^{18}\text{O}$ in planktic foraminifera to sea surface temperature. *Paleoceanogr. Paleoclimatol.* **34**, 1292–1315 (2019).
- Tardif, R. et al. Last millennium reanalysis with an expanded proxy database and seasonal proxy modeling. *Clim. Past* **15**, 1251–1273 (2019).
- Gray, W. R. et al. Wind-driven evolution of the North Pacific subpolar gyre over the last deglaciation. *Geophys. Res. Lett.* **47**, e2019GL086328 (2020).
- DiNezio, P. N. et al. Glacial changes in tropical climate amplified by the Indian Ocean. *Sci. Adv.* **4**, eaat9658 (2018).
- Ford, H. L., Ravelo, A. C. & Polissar, P. J. Reduced El Niño–Southern Oscillation during the Last Glacial Maximum. *Science* **347**, 255–258 (2015).
- Glushkova, O. Y. Geomorphological correlation of Late Pleistocene glacial complexes of Western and Eastern Beringia. *Quat. Sci. Rev.* **20**, 405–417 (2001).
- Bartlein, P. J. et al. Pollen-based continental climate reconstructions at 6 and 21 ka: a global synthesis. *Clim. Dynam.* **37**, 775–802 (2011).
- Roe, G. H. & Lindzen, R. S. The mutual interaction between continental-scale ice sheets and atmospheric stationary waves. *J. Clim.* **14**, 1450–1465 (2001).
- Löfverström, M. & Liakka, J. On the limited ice intrusion in Alaska at the LGM. *Geophys. Res. Lett.* **43**, 11030–11038 (2016).
- Hopcroft, P. O. & Valdes, P. J. How well do simulated Last Glacial Maximum tropical temperatures constrain equilibrium climate sensitivity? *Geophys. Res. Lett.* **42**, 5533–5539 (2015).
- Porter, S. C. Snowline depression in the tropics during the Last Glaciation. *Quat. Sci. Rev.* **20**, 1067–1091 (2000).
- Stute, M. et al. Cooling of tropical Brazil (5°C) during the Last Glacial Maximum. *Science* **269**, 379–383 (1995).
- Lea, D. W., Pak, D. K. & Spero, H. J. Climate impact of late Quaternary equatorial Pacific sea surface temperature variations. *Science* **289**, 1719–1724 (2000).
- Masson-Delmotte, V. et al. EPICA Dome C record of glacial and interglacial intensities. *Quat. Sci. Rev.* **29**, 113–128 (2010).
- Lee, J.-E., Fung, I., DePaolo, D. J. & Otto-Bliesner, B. Water isotopes during the Last Glacial Maximum: new general circulation model calculations. *J. Geophys. Res.* **113**, D19109 (2008).
- Kurahashi-Nakamura, T., Paul, A. & Losch, M. Dynamical reconstruction of the global ocean state during the Last Glacial Maximum. *Paleoceanography* **32**, 326–350 (2017).
- Amrhein, D. E., Wunsch, C., Marchal, O. & Forget, G. A global glacial ocean state estimate constrained by upper-ocean temperature proxies. *J. Clim.* **31**, 8059–8079 (2018).
- Flato, G. et al. in *Climate Change 2013: The Physical Science Basis* (eds Stocker, T. F. et al.) 741–866 (IPCC, Cambridge Univ. Press, 2014).
- PALAEOSSENS Project Members Making sense of palaeoclimate sensitivity. *Nature* **491**, 683–691 (2012); erratum **494**, 130 (2013).
- Prentice, I. C., Jolly, D. & Biome 6000 Participants Mid-Holocene and glacial-maximum vegetation geography of the northern continents and Africa. *J. Biogeogr.* **27**, 507–519 (2000).
- Köhler, P. et al. What caused Earth's temperature variations during the last 800,000 years? Data-based evidence on radiative forcing and constraints on climate sensitivity. *Quat. Sci. Rev.* **29**, 129–145 (2010).
- Etmann, M., Myhre, G., Highwood, E. & Shine, K. Radiative forcing of carbon dioxide, methane, and nitrous oxide: a significant revision of the methane radiative forcing. *Geophys. Res. Lett.* **43**, 12614–12623 (2016).
- Albani, S. et al. Aerosol–climate interactions during the Last Glacial Maximum. *Curr. Clim. Change Rep.* **4**, 99–114 (2018).
- Shakun, J. D. Modest global-scale cooling despite extensive early Pleistocene ice sheets. *Quat. Sci. Rev.* **165**, 25–30 (2017).
- Stap, L., Köhler, P. & Lohmann, G. Including the efficacy of land ice changes in deriving climate sensitivity from paleodata. *Earth Syst. Dynam.* **10**, 333–345 (2019).
- Yoshimori, M., Yokohata, T. & Abe-Ouchi, A. A comparison of climate feedback strength between CO_2 doubling and LGM experiments. *J. Clim.* **22**, 3374–3395 (2009).
- Friedrich, T., Timmermann, A., Tigchelaar, M., Timm, O. E. & Ganopolski, A. Nonlinear climate sensitivity and its implications for future greenhouse warming. *Sci. Adv.* **2**, e1501923 (2016).

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

© The Author(s), under exclusive licence to Springer Nature Limited 2020

SST proxy data collection

We compiled a total of 956 and 879 average proxy values from the LGM and late Holocene (LH) time periods, respectively. In many cases, multiple types of proxy data are available from the same location. Using a cutoff criterion for 'same location' of 0.1° in latitude or longitude, our proxy collection contains 636 and 664 unique locations for the LGM and LH, respectively. Following MARGO, the LGM was defined as 23–19 ka (ref. ²). The LH was defined as 4–0 ka to match the interval of time averaged for the LGM. This choice is consistent with previous LGM–LH comparisons^{12,51}. For the purposes of this study, the LH is considered representative of pre-industrial conditions. Although climate has certainly changed within the past 4 kyr in response to shorter-term forcings such as volcanic eruptions and solar irradiance, we posit that this assumption is reasonable in the context of the large, slow climate changes associated with orbitally driven glacial–interglacial cycles. This differs from CLIMAP and MARGO, in which LGM cooling was calculated relative to twentieth-century observations^{1,2}. However, historical observations carry their own distinctive type of uncertainty⁵² and also include the signature of anthropogenic global warming, so are not ideal in terms of isolating the LGM climatic response. Furthermore, using proxy estimates of the pre-industrial climate, rather than observational estimates, offers an 'apples-to-apples' comparison with the LGM.

The proxy data consist of both continuous time series that pass through the LGM and the LH, and data from 'timeslice' studies that focused on the average LGM climate state. The latter derive in part from the MARGO collection, with additional data from more recent studies. In many cases, LGM slice data were not accompanied by corresponding LH data. To provide matching data for these cases, we searched the core top datasets for each proxy system for a nearby value, with a cutoff radius of 1,000 km. This provided roughly similar spatial coverage between the LGM and LH target slices (Fig. 1). Core tops that were used as LH data were subsequently removed from the calibration datasets, and the parameters for the Bayesian forward models were recalculated without these data, to avoid circularity. However, as the number of core tops used in the data synthesis is small (100 or less) relative to the size of the core top datasets (~1,000) the effect on the model parameters was negligible.

The time series consist of 260 core sites with radiocarbon-based age models. Since we are only reconstructing the mean LGM and LH states, we did not apply any criteria regarding sampling resolution. Before taking time averages from these proxy data, we recalibrated all age models using the Marine13 radiocarbon curve⁵³ and the BACON age modelling software (translated into Python as snakebacon and available at <https://github.com/brews/snakebacon>) to ensure consistent treatment. Not all of the core sites extend through the LGM and/or the LH, so not all of these sites are represented in the analysis. However, many of these sites typically contain more than one proxy for SST, thus the total number of data points from the time series is 335 for the LGM and 273 for the LH.

Model simulations

We employ iCESM²², which is based on CESM version 1.2.54. iCESM has the capability to explicitly simulate the transport and transformation of water isotopes in hydrological processes in the atmosphere, land, ocean and sea ice²². iCESM reproduces instrumental records of both physical climate and water isotopes reasonably well^{22,54}.

We conducted iCESM1.2 simulations of the LH and the LGM, consisting of timeslices of the PI, 3, 18 and 21 ka. These simulations had a horizontal resolution of $1.9^\circ \times 2.5^\circ$ (latitude \times longitude) for the atmosphere and land, and a nominal 1° displaced-pole Greenland grid for the ocean and sea ice. The PI simulation used standard climatic forcings at AD 1850, and the slice simulations of 3, 18 and 21 ka used boundary conditions of GHGs, Earth orbits and ice sheets following the PMIP4 protocol⁵⁵. Specifically, CO_2 , CH_4 and N_2O concentrations were 284.7 ppm,

791.6 ppb and 275.68 ppb for the PI; 275 ppm, 580 ppb and 270 ppb for 3 ka; 190 ppm, 370 ppb and 245 ppb for 18 ka; and 190 ppm, 375 ppb, and 200 ppb for 21 ka. Changes in surface elevation, albedo and land ocean distribution associated with the ice sheets were derived from the ICE-6G reconstruction⁵⁶. For the 18 and 21 ka simulations, global volume mean seawater $\delta^{18}\text{O}$ was set to 1.05‰ in accordance with ice-volume effects. A value of 0.05‰ was used for the PI and 3 ka simulations. The simulations were run with a prescribed satellite phenology in the land model. This choice was based on the overall poorer simulation of vegetation processes with a prognostic phenology, which could potentially be more problematic for palaeoclimate⁵⁷. The LH (3 ka and PI) and the LGM slices (21 and 18 ka) were initialized from equilibrated CESM1.2 simulations with water isotopes^{27,58} and extended for 900 years. All of the slices reached quasi-equilibrium in both the physical climate and the water isotopes with a top-of-atmosphere energy imbalance of less than 0.09 W m^{-2} .

Model ECS was estimated to be 3.6°C and 2.8°C in the PI and LGM simulations, respectively. The model ECS was calculated as the ΔGMST between a baseline simulation (PI or LGM) and a simulation with twice the corresponding atmospheric CO_2 amount in a slab ocean model (SOM) configuration. Ocean dynamics were inactive in SOM simulations and their effects were approximated using prescribed mixed-layer depth and heat transport convergence that were derived from the fully coupled PI or LGM simulation. The PI ECS of 3.6°C is in the middle of the multiple-model range (2.1 – 4.7°C) in the Coupled Model Inter-comparison Project phase 5⁴¹, but slightly smaller than the published values of 4.0 – 4.2°C in CESM1.2^{59,60}. The difference is probably due to the prescribed satellite phenology in our configuration of the land model, which excludes vegetation feedbacks⁶¹.

We also made use of available iCESM1.3 simulations of the LGM and PI⁵⁸. iCESM1.3 differs from iCESM1.2 primarily in the gravity wave scheme, along with a few bug fixes in the cloud microphysics and radiation⁶². The iCESM1.3 PI and LGM simulations have lengths of 400 and 1,000 years, respectively.

Data assimilation

The data assimilation technique used here is the ensemble square-root Kalman filter⁶³; specifically, an offline method that has previously been used to reconstruct climate during the past millennium^{25,64–66}. This data assimilation method solves the following equation to compute an ensemble of climate states (X_{analysis}) from prior states taken from a climate model (X_{prior}) and proxy data observations (y):

$$X_{\text{analysis}} = X_{\text{prior}} + K(y - Y_e) \quad (1)$$

X_{prior} is an ensemble of climate states taken from iCESM, of dimension N state vector elements (SST, SAT and so on collapsed into a vector) by M ensemble members (40, in our case). In a typical data assimilation application, the length of time represented by the ensemble would equal the length of time represented by the data; for example, annual data would be used to update an annual prior. However, in our case the data represent average conditions across 4,000 years. As we cannot run iCESM for 100,000 years or more, we must use a different time average for our model prior. Experimentation with time averaging the model states revealed that once the average exceeded the interannual timescale, the patterns in the assimilated fields, as well as GSST and GMST, were largely insensitive to the length of the average (Extended Data Fig. 3). Thus we chose 50 years—the longest time average that we could use while still retaining enough ensemble members for the assimilation technique (40 members).

$y - Y_e$ is the innovation: the difference between the column vector y (dimensions $P \times 1$) of observed proxy values for each location P and the matrix Y_e (dimensions $P \times M$) of proxy values calculated from the model prior at the same locations. The innovation thus takes place in the native proxy units. Y_e is calculated by feeding model variables

from the prior ensemble (SST, SSS, $\delta^{18}\text{O}$ of seawater) into previously published Bayesian forward models for each proxy type^{12,13,23,24} that convert those variables into an estimate in proxy units. These models take the general form:

$$\begin{aligned} \text{Proxy} &\sim \mathcal{N}(\mu, \sigma^2) \\ \mu &= f(\text{SST}, X_i), \end{aligned} \quad (2)$$

where σ^2 is the residual error variance associated with the function that describes μ , X_i denotes i additional X climatic influences that may contribute to the proxy, and \sim indicates ‘has the distribution of’. For complete details of each model, we refer the reader to the original publications^{12,13,23,24}. Briefly, the $U_{37}^{K'}$ and TEX_{86} forward models are only a function of SST (monthly values for $U_{37}^{K'}$, to account for seasonal responses in the North Atlantic, North Pacific and Mediterranean regions¹²; annual for TEX_{86}). The model for $\delta^{18}\text{O}$ of planktic foraminifera ($\delta^{18}\text{O}_c$) requires monthly SST and the annual $\delta^{18}\text{O}$ composition of seawater²⁴. $\delta^{18}\text{O}_c$ is computed for the optimal growing season using the species-specific hierarchical model described in ref.²⁴. The model for Mg/Ca of planktic foraminifera requires monthly SST and sea surface salinity (SSS), as well as surface water pH, bottom water calcite saturation state (Ω) and the cleaning method used in the laboratory¹³. The latter was recorded as part of the data collection effort, but for pH and Ω we must make assumptions, as iCESM does not simulate the ocean carbonate system. In the absence of good information regarding spatial changes in Ω , we assume that it is the same as today for each given site location, with values drawn from GLODAPv2 (ref.⁶⁷). For pH, we use modern estimates from GLODAPv2 for the LH timeslice, and then for the LGM we add 0.13 units to the modern values to account for the global increase in pH due to lowered CO_2 , following refs.^{13,68}. As with $\delta^{18}\text{O}_c$, Mg/Ca is forward-modelled for the optimal growing season using the species-specific hierarchical model described in ref.¹³. In this manner, the seasonal preferences of foraminifera are explicitly accounted for in the assimilation.

K is the Kalman gain (of dimension $N \times P$), which weights the innovation according to the covariance of the forward-modelled proxy value with the rest of the climate state, as well as the uncertainties of the model ensemble and the proxy observations. Following ref.⁶³, K is defined as:

$$K = W_{\text{loc}} \circ [X'_{\text{prior}} Y'_e{}^T / (n-1)] \times [Y_{\text{loc}} \circ [Y'_e Y'_e{}^T / (n-1)] + R]^{-1}, \quad (3)$$

where X'_{prior} and Y'_e are the matrices of deviations, for example, $X'_{\text{prior}} = X_{\text{prior}} - \bar{X}_{\text{prior}}$, R is a diagonal matrix of the uncertainty (as a variance) of each proxy observation, and W_{loc} and Y_{loc} are weights that apply a covariance localization, a distance-weighted filter that limits the influence of each proxy in space⁶⁹. Covariance localization is applied to minimize spurious relationships from producing artefacts at large distances away from the proxy location. Following ref.²⁵, we use the Gaspari–Cohn fifth-order polynomial⁷⁰ with a specified cutoff radius (outside of which all covariance is eliminated) of 12,000 km (the rationale behind this value is discussed below). W_{loc} contains the localization weights of each N state vector element relative to each proxy estimate in Y_e and is of dimension $N \times P$. Y_{loc} contains the localization weights of each proxy estimate in Y_e relative to the other estimates in Y_e and thus is of dimension $P \times P$. n is the number of ensemble members, and division by $(n-1)$ is applied to obtain an unbiased estimate. \circ denotes element-wise multiplication.

Following ref.⁶³, the update equation (equation (1)) is solved by decomposing the problem into an update of the mean value of the prior state (\bar{X}_{prior}) and the deviations from the mean (X'_{prior}):

$$\bar{X}_{\text{analysis}} = \bar{X}_{\text{prior}} + K(y - \bar{Y}_e), \quad (4)$$

$$X'_{\text{analysis}} = X'_{\text{prior}} - \bar{K} Y'_e, \quad (5)$$

where \bar{K} is defined as:

$$\begin{aligned} \bar{K} &= W_{\text{loc}} \circ [X'_{\text{prior}} Y'_e{}^T / (n-1)] [(\sqrt{Y_{\text{loc}} \circ [Y'_e Y'_e{}^T / (n-1)] + R})^{-1}]^T \\ &\times [\sqrt{Y_{\text{loc}} \circ [Y'_e Y'_e{}^T / (n-1)]} + \sqrt{R}]^{-1}. \end{aligned} \quad (6)$$

The full assimilated ensemble is then recovered through:

$$X_{\text{analysis}} = \bar{X}_{\text{analysis}} + X'_{\text{analysis}} \quad (7)$$

The assimilated ensemble has dimensions $N \times M$, like the prior model ensemble.

The diagonal proxy error matrix R and the cutoff radius for the covariance localization weights are user-defined, so we used validation metrics based on withheld SST proxies and independent proxies for the oxygen isotopic composition of precipitation ($\delta^{18}\text{O}_p$) to guide our choices. The most conservative choice for R is to assign the σ^2 error terms (equation (2)) associated with the Bayesian forward models for each proxy to the diagonal of the matrix. As the forward models are trained on a globally distributed set of data, these terms represent the error in the proxy system on a global level and so we denote these as R_g . In temperature space, these translate to 1.5–4 °C 1 σ errors. However, these are probably too high for an individual proxy location (otherwise, one would expect that the proxies would not be able to detect LGM cooling at all). Rather, at a single site, proxy uncertainty is expected to lie somewhere between analytical precision and the global error. Unfortunately, this can only be directly observed by analysing parallel sediment cores, which is not commonly done in palaeoceanography. Thus, to experimentally determine an optimal value of R , we systematically reduced it from R_g to $R_g/500$ and analysed validation statistics. Values of SST, SSS and $\delta^{18}\text{O}$ of seawater from the assimilated ensemble were then forward-modelled to predict the withheld proxy values. We calculate both the coefficient of efficacy (CE)⁷¹ and the root mean square error (RMSE) between the observed and mean of the predicted proxy values. These were calculated in normalized units to account for the different ranges of absolute proxy values between $U_{37}^{K'}$, TEX_{86} , $\delta^{18}\text{O}$ and Mg/Ca.

In addition, we calculated the R^2 between observed $\Delta\delta^{18}\text{O}_p$ (LGM – PI) derived from ice cores and speleothems and predicted $\Delta\delta^{18}\text{O}_p$ from the assimilation. Model prior $\delta^{18}\text{O}_p$ in the assimilation was calculated as an annual mean weighted by the monthly precipitation amounts. Ice core $\Delta\delta^{18}\text{O}_p$ values were taken from the compilation in ref.⁷² (their table 1). Speleothem $\Delta\delta^{18}\text{O}_p$ values were computed from the SISAL database, version 1b⁷³. We first searched the SISAL database for sites that contained both LH (4–0.2 ka, to exclude anthropogenically influenced values) and LGM (23–19 ka) data and recorded the mean $\delta^{18}\text{O}$ of calcite or aragonite. We then converted these average $\delta^{18}\text{O}$ values to dripwater $\delta^{18}\text{O}$ (considered analogous to $\delta^{18}\text{O}_p$) following the recommendations of ref.⁷⁴ (their equations 1–3). This conversion accounts for the influence of temperature on fractionation as well as kinetic effects, and converts from the VPDB to the VSMOW scale. These calculations require an estimate of cave temperature; for this we use the SAT value from our data assimilation at the grid cells closest to the speleothem locations.

Extended Data Table 1 shows the validation results from scaling R_g . For both the LGM and LH, validation CE and RMSE are relatively insensitive to the choice of R , although there is slight improvement up to $R_g/5$. This is in part because the validation is being calculated across the globe and the spatial variation in proxy values is always large; indeed, even the prior produces decent prediction of the withheld proxies, particularly for the LH (Extended Data Table 1). More useful information can be gleaned from the comparison with independent $\delta^{18}\text{O}_p$ data. Here, we observe a substantial increase in R^2 from the case of R_g to $R_g/5$ from 0.50 to 0.67. This suggests that an increase in proxy precision drives the assimilated product closer to a state that agrees with the $\delta^{18}\text{O}_p$ proxies. However, it is also clear that if the proxy uncertainty is set too

Article

low ($R_g/500$), the R^2 drops back down (0.59). Thus, there appears to be an ideal value of R somewhere between $R_g/5$ and $R_g/100$. We choose the lowest scaling value of $R_g/5$ —the most conservative choice. $R_g/5$ translates to -0.6 – 1.3 °C of 1σ uncertainty, depending on the proxy. This is greater than analytical error (~ 0.3 °C 1σ) and strikes us as a reasonable estimate of site-specific error.

To find an optimal cutoff radius for the covariance localization, we experimented with values from ∞ (for example, no localization) to 6,000 km and analysed the same validation statistics as above. As was the case with varying R , validation CE and RMSE are not very sensitive to localization (Extended Data Table 2). However, comparison with $\delta^{18}\text{O}_p$ shows that the best skill is achieved with a cut-off radius of 12,000 km (Extended Data Table 2), so we use this value for our assimilation.

Each time interval (LGM and LH) is carried out as a separate ensemble assimilation consisting of 25 iterations, in which we withhold 25% of the proxies at random for validation. Across all iterations, the validation CE is 0.94 and 0.92 for the LH and LGM slices respectively, and the validation RMSE is 0.24 and 0.29 in normalized proxy units, respectively. Figure 3 shows the independent validation with the $\delta^{18}\text{O}_p$ proxies. The relatively good fit ($R^2 = 0.67$) is a substantial improvement over the model prior (Extended Data Fig. 1). The improvement mainly comes from the ice core sites, which have an R^2 of 0 in the prior case and an R^2 of 0.18 in the assimilated product. However, these values are skewed by a negative outlier, the Camp Century ice core. Excluding this outlier, the R^2 values are 0 and 0.41 in the prior and assimilated product, respectively, indicating substantial improvement. In contrast, the speleothem sites show less difference between the prior and posterior, with an R^2 of 0.12 and 0.13, respectively. The speleothem group also has an outlier—Sofular Cave. Excluding this site, the R^2 is 0.24 and 0.32 in the prior and assimilated product, respectively, indicating some improvement. The limited change at the speleothem sites may reflect the fact that prior variability in mean annual $\delta^{18}\text{O}_p$ is small at these site locations.

Influence of different proxy types

The advantage of using a multi-proxy approach is that structural biases associated with each proxy type counterbalance each other, resulting in a more robust estimate of LGM cooling. To investigate how each proxy type influences the assimilated result, we conducted experiments in which only $\delta^{18}\text{O}$, U_{37}^K and Mg/Ca data were assimilated, respectively (TEX₈₆ was not investigated as there are only 20 sites available). Extended Data Fig. 4 shows the assimilated ΔGSST and ΔGMST estimates for LGM cooling for each proxy type, alongside maps of proxy locations. Median values for ΔGSST are -3.0 °C, -3.1 °C and -3.3 °C for the U_{37}^K , Mg/Ca and $\delta^{18}\text{O}$ experiments respectively, all of which are very close to the value using all proxy types (-3.1 °C). Error bounds for U_{37}^K and Mg/Ca are larger and most probably reflect the smaller number of total data points assimilated. Median values for GMST are -6.1 °C, -6.6 °C and -6.3 °C for the U_{37}^K , Mg/Ca and $\delta^{18}\text{O}$ experiments respectively. The U_{37}^K and $\delta^{18}\text{O}$ experiments yield results close to the value using all proxy types (-6.1 °C). Assimilation with Mg/Ca yields a slightly cooler result, which might reflect the spatial distribution of the proxies (that is, a concentration of data in the North Atlantic) or an imperfect estimation of non-thermal controls. Although temperature dominates foraminiferal Mg/Ca, salinity, pH and dissolution also exert substantial influence. These secondary factors are accounted for in the forward model (BAYMAG¹³), but we must make simple assumptions about the change in pH and bottom water saturation (see the description in Data assimilation above) for the LGM case that could deviate from actual conditions at the respective core sites.

Assessment of bias in the model prior

As a formal combination of model results and data, our data assimilation product and its uncertainties are inherently influenced by bias in the model's ability to accurately reproduce the true LGM and LH climate states. However, we do not know what these true states were (especially

the LGM), so we can only indirectly assess whether the model prior is 'biased' in this sense. To explore potential bias, we calculated proxy verification rank histograms⁷⁵ for the LGM and LH time slices respectively (Extended Data Fig. 2). Rank histograms are simply counts of the rank of each observation (the proxy data withheld for validation) relative to each of the 40 ensemble members from the assimilated product (forward-modelled into proxy units). If a proxy ranks below all of the ensemble members, it gets a rank of 1; if it ranks above all the ensemble members, it ranks as 41. As the proxy observations themselves are not error-free, following ref. ⁷⁵ we add our inferred proxy uncertainty of $R_g/5$ to the ensembles before calculating ranks.

A perfectly ideal model prior would produce a flat rank histogram, which indicates that the ensemble is drawn from the same distribution as the proxy; however, this is rarely the case as all models carry some structural biases relative to observations. Both the symmetry and the shape of the rank histogram can be used to intuit the nature of the bias⁷⁵. In our case, we observe that our rank histograms are close to symmetrical (particularly for the LGM; Extended Data Fig. 2), which suggests that our model prior does not have too much bias in the mean (otherwise we would expect the proxies to systematically rank higher or lower than the ensemble). However, our histograms have a distinctive U shape (Extended Data Fig. 2), indicating that a number of proxies rank lower or higher than the assimilated ensembles. A U shape can arise if the error on the observations is underestimated; however, our validation against $\Delta\delta^{18}\text{O}_p$ data suggests that our estimate of this value ($R_g/5$) is robust. A U shape can also arise if there are structural biases in the proxies that are not accounted for in our forward models. However, we found that the U shape persists across all of our four proxy types, and that there are no geographical patterns in the locations of high and low ranks (not shown). These observations are encouraging in that they suggest that there is not a systematic error in a single proxy system. Nevertheless, the forward models are fit globally and thus do not account for errors that may arise due to localized non-thermal influences (such as, for example, nutrient availability) so we cannot completely rule out the possibility that structural proxy errors contribute to the shape of the rank histogram.

The most likely explanation for the U shape is that our model prior is underdispersive, or rather, lacks variability relative to what the proxies consider to be the 'true state'. Experimentation with inflation of the model prior variance, and a range of time-averaging (5–50 years) did not substantially reduce the U-shaped rank histograms (not shown), which suggests it is not so much the fact that the 50-yr ensemble members lack variance, but rather that CESM maintains relatively consistent covariance patterns across the pre-industrial and glacial states, respectively. The structure of the SST, SSS and seawater isotope fields, compared to modern observations, are not perfect, with known biases^{22,76,77}. This would cause proxies to consistently rank lower or higher in certain locations.

The implications of an underdispersive prior are mainly that our estimate of uncertainties—both in the assimilated fields and the estimates of global mean temperature—may be too small. As we do not know the true LGM state, we cannot estimate how overconfident our assimilated results are. We need model priors from multiple other isotope-enabled climate models to assess this more thoroughly, an experiment that must be left to future work, when such simulations become available.

Proxy-only estimates of LGM cooling

To provide a point of comparison for the data assimilation results, we computed global and tropical ΔSST and ΔGMST from the proxy data in isolation. The proxy data were calibrated to SST using a suite of Bayesian prediction models^{12,13,23,24} producing 1,000-member ensemble estimates for each data point. To approximate the proxy observational uncertainty used in the data assimilation (see discussion above), data were sorted along the ensemble dimension and normally distributed site-level error ($\mathcal{N}(0, 0.5)$ °C) was added back to the ensemble. GSST

was computed following the method of ref. ⁶⁰—data were first binned and averaged into latitudinal bands, then latitudinal averages were used to calculate an area-weighted global average. As the results are sensitive to the size of the bin⁶⁰, we computed an ensemble of GSST across bin sizes of 2.5° to 15° (at 2.5° intervals). GSST was scaled to GMST using the method of ref. ¹⁴, in which scaling factors (determined from PMIP LGM simulations) were drawn from a uniform distribution spanning values between 1.5 and 2.3. Results for Δ GSST and Δ GMST are shown in Fig. 2c, d.

Analysis of tropical Pacific cooling

As stated in the main text, we analysed LGM cooling across the tropical Pacific in the proxy data alone to compare with the data assimilation result. The proxy data were calibrated to SST as described in the previous subsection, and then the Pacific zonal gradient was computed for both the LGM and LH as the difference between the average SST in the western Pacific (10° S–10° N, 130–170° E) and the eastern Pacific (5° S–5° N, 75–140° W). We then computed the LGM–LH difference in the zonal gradient, yielding a median value of -0.2 °C. These calculations were conducted for all 1,000 ensemble members, yielding an uncertainty of 1.0 °C (2σ).

Climate sensitivity calculations

We calculate ECS as:

$$\text{ECS} = \frac{\Delta\text{GMST}}{\Delta R} \times F_{2\times\text{CO}_2}, \quad (8)$$

where Δ GMST is taken from the data assimilation, $F_{2\times\text{CO}_2}$ is the forcing associated with the doubling of CO_2 from the pre-industrial state and ΔR is the total change in radiative forcing, including the slow feedbacks that affect the LGM climate state. In Fig. 4, we present one solution that includes GHG and ice sheet forcing ($\Delta R = \Delta R_{\text{GHG}} + \Delta R_{\text{ICE}}$) and another that also includes aerosol forcing from mineral dust ($\Delta R = \Delta R_{\text{GHG}} + \Delta R_{\text{ICE}} + \Delta R_{\text{AE}}$). We estimate the mean value of ΔR_{GHG} and $F_{2\times\text{CO}_2}$ to be -2.48 W m^{-2} and 3.80 W m^{-2} , respectively, using published equations⁴⁵. Average GHG concentrations for the LGM and LH are used for the calculation of ΔR_{GHG} . Following the assessment in successive IPCC reports, we assume a 10% (90% CI) uncertainty for $F_{2\times\text{CO}_2}$ (ref. ⁴⁵) ($1\sigma = 0.23$). The 1σ uncertainty in ΔR_{GHG} , which is dependent on $F_{2\times\text{CO}_2}$, is:

$$\sqrt{\left(\frac{\Delta R_{\text{GHG}}}{F_{2\times\text{CO}_2}}\right)^2 \times 0.23^2} = 0.15 \text{ W m}^{-2}. \quad (9)$$

ΔR_{ICE} accounts for the radiative forcing from surface albedo changes associated with the LGM ice sheets and exposed land due to lowered sea level. We calculate ΔR_{ICE} using an approximate partial radiative perturbation method⁷⁸ in CESM and also make use of additional published results in 11 PMIP2 and PMIP3 models^{6,79}. The resulting ΔR_{ICE} has a multi-model ensemble mean of -3.64 W m^{-2} and a range from -2.59 to -5.20 W m^{-2} (Extended Data Table 3). ΔR_{AE} was obtained from a published compilation of the top of atmosphere instantaneous direct radiative forcing of LGM dust in nine modelling studies⁴⁶. ΔR_{AE} has a large spread, ranging from 0 to -2 W m^{-2} .

We also calculate a third solution for ECS that assumes a lower ε in which ΔR_{ICE} is multiplied by 0.7, a value consistent with previous work^{48,49}. This value of ε is used merely to show the impact of a smaller-than-unity efficacy on ECS and should not be considered a definitive number. In ref. ⁴⁸ it was demonstrated that there is a wide range of possible values for ε ; however, some of the derived range comes from models of intermediate complexity⁴⁷ that have substantially different physics than the GCMs in PMIP2 and PMIP3 that we use to derive the values of ΔR_{ICE} (Extended Data Table 3). Thus we prefer to show results for a single value of ε as an example rather than sampling from the full range in ref. ⁴⁸.

To propagate uncertainties into the final calculations of ECS, we used a Monte Carlo approach, sampling the full 1,000-member ensemble of Δ GMST from our assimilated product, and combining these with 10,000 samples of each distribution of ΔR as well as $F_{2\times\text{CO}_2}$. ΔR_{GHG} and $F_{2\times\text{CO}_2}$ were assumed to have normal distributions. Since their uncertainties are dependent, $F_{2\times\text{CO}_2}$ was sampled first, then the distribution of ΔR_{GHG} was derived by multiplying the resulting Monte Carlo samples by the ratio of the mean value of ΔR_{GHG} to $F_{2\times\text{CO}_2}$ ($-2.48/3.8 = -0.65$). ΔR_{ICE} and ΔR_{AE} were treated as empirical random distributions because the limited number of samples (derived from modelling experiments) prevents us from knowing the shape of these distributions.

Total ΔR for the case in which only GHG and ice sheet forcing is considered is -6.0 W m^{-2} (-7.9 to -5.0 W m^{-2} , 95%CI). Total ΔR including dust aerosol forcing is -6.8 W m^{-2} (-9.6 to -5.2 W m^{-2} , 95%CI). Total ΔR for the case with reduced ice sheet efficacy is -5.7 W m^{-2} (-8.2 to -4.4 W m^{-2} , 95%CI).

Data availability

The LGM and LH proxy data are available as .csv files (including both raw proxy values and calibrated estimates of SST). We also provide a gridded $5^\circ \times 5^\circ$ map of LGM–LH proxy anomalies in .netcdf format. The fields of the data assimilation product (SST, SAT, SSS, $\delta^{18}\text{O}$ of seawater and $\delta^{18}\text{O}_p$) are also available in .netcdf format. Files are publicly available for download from PANGAEA (<https://doi.org/10.1594/PANGAEA.920596>) and from GitHub (<https://github.com/jesstierney/lgmDA>). Source data are provided with this paper.

Code availability

The data assimilation method used in this paper is publicly available as the Matlab code package DASH on GitHub (<https://github.com/JonKing93/DASH>). The Bayesian forward models, BAYSPAR, BAYSPLINE, BAYFOX and BAYMAG are likewise publicly available on GitHub from <https://github.com/jesstierney>. The iCESM1.2 model code is available at <https://github.com/NCAR/iCESM1.2>.

51. DiNezio, P. N. & Tierney, J. E. The effect of sea level on glacial Indo-Pacific climate. *Nat. Geosci.* **6**, 485–491 (2013).
52. Chan, D., Kent, E. C., Berry, D. I. & Huybers, P. Correcting datasets leads to more homogeneous early-twentieth-century sea surface warming. *Nature* **571**, 393–397 (2019).
53. Reimer, P. J. et al. IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon* **55**, 1869–1887 (2013).
54. Hurrell, J. W. et al. The community earth system model: a framework for collaborative research. *Bull. Am. Meteorol. Soc.* **94**, 1339–1360 (2013).
55. Kageyama, M. et al. The PMIP4 contribution to CMIP6–Part 4: scientific objectives and experimental design of the PMIP4-CMIP6 Last Glacial Maximum experiments and PMIP4 sensitivity experiments. *Geosci. Model Dev.* **10**, 4035–4055 (2017).
56. Peltier, W. R., Argus, D. F. & Drummond, R. Space geodesy constrains ice age terminal deglaciation: the global ICE-6G_C (VM5a) model. *J. Geophys. Res. Solid Earth* **120**, 450–487 (2015).
57. Lawrence, D. M. et al. Parameterization improvements and functional and structural advances in version 4 of the community land model. *J. Adv. Model. Earth Syst.* **3**, M03001 (2011).
58. Zhu, J. et al. Reduced ENSO variability at the LGM revealed by an isotope-enabled Earth system model. *Geophys. Res. Lett.* **44**, 6984–6992 (2017).
59. Gettelman, A., Kay, J. E. & Shell, K. M. The evolution of climate sensitivity and climate feedbacks in the Community Atmosphere Model. *J. Clim.* **25**, 1453–1469 (2012).
60. Zhu, J., Poulsen, C. J. & Tierney, J. E. Simulation of Eocene extreme warmth and high climate sensitivity through cloud feedbacks. *Sci. Adv.* **5**, eaax1874 (2019).
61. Richardson, A. D. et al. Climate change, phenology, and phenological control of vegetation feedbacks to the climate system. *Agric. For. Meteorol.* **169**, 156–173 (2013).
62. Meehl, G. A. et al. Effects of model resolution, physics, and coupling on Southern Hemisphere storm tracks in CESM1.3. *Geophys. Res. Lett.* **46**, 12408–12416 (2019).
63. Whitaker, J. S. & Hamill, T. M. Ensemble data assimilation without perturbed observations. *Mon. Weath. Rev.* **130**, 1913–1924 (2002).
64. Steiger, N. J., Hakim, G. J., Steig, E. J., Battisti, D. S. & Roe, G. H. Assimilation of time-averaged pseudoproxies for climate reconstruction. *J. Clim.* **27**, 426–441 (2014).
65. Hakim, G. J. et al. The last millennium climate reanalysis project: framework and first results. *J. Geophys. Res. Atmos.* **121**, 6745–6764 (2016).
66. Okazaki, A. & Yoshimura, K. Development and evaluation of a system of proxy data assimilation for paleoclimate reconstruction. *Clim. Past* **13**, 379–393 (2017).

67. Lauvset, S. K. et al. A new global interior ocean mapped climatology: the 1 × 1 GLODAP version 2. *Earth Syst. Sci. Data* **8**, 325–340 (2016).
68. Gray, W. R. & Evans, D. Nonthermal influences on Mg/Ca in planktonic foraminifera: a review of culture studies and application to the last glacial maximum. *Paleoceanogr. Paleoeclimatol.* **34**, 306–315 (2019).
69. Houtekamer, P. L. & Mitchell, H. L. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Weath. Rev.* **129**, 123–137 (2001).
70. Gaspari, G. & Cohn, S. E. Construction of correlation functions in two and three dimensions. *Q. J. R. Meteorol. Soc.* **125**, 723–757 (1999).
71. Nash, J. E. & Sutcliffe, J. V. River flow forecasting through conceptual models part I—a discussion of principles. *J. Hydrol.* **10**, 282–290 (1970).
72. Werner, M. et al. Glacial–interglacial changes of H₂¹⁸O, HDO and deuterium excess—results from the fully coupled Earth System Model ECHAM5/MPI-OM. *Geosci. Model Dev.* **9**, 647–670 (2016).
73. Atsawaranunt, K. et al. The SISAL database: a global resource to document oxygen and carbon isotope records from speleothems. *Earth Syst. Sci. Data* **10**, 1687–1713 (2018).
74. Comas-Bru, L. et al. Evaluating model outputs using integrated global speleothem records of climate change since the last glacial. *Clim. Past* **15**, 1557–1579 (2019).
75. Hamill, T. M. Interpretation of rank histograms for verifying ensemble forecasts. *Mon. Weath. Rev.* **129**, 550–560 (2001).
76. Danabasoglu, G. et al. The CCSM4 ocean component. *J. Clim.* **25**, 1361–1389 (2012).
77. Park, T.-W., Deng, Y., Cai, M., Jeong, J.-H. & Zhou, R. A dissection of the surface temperature biases in the Community Earth System Model. *Clim. Dynam.* **43**, 2043–2059 (2014).
78. Taylor, K. et al. Estimating shortwave radiative forcing and response in climate models. *J. Clim.* **20**, 2530–2543 (2007).
79. Braconnot, P. & Kageyama, M. Shortwave forcing and feedbacks in Last Glacial Maximum and Mid-Holocene PMIP3 simulations. *Phil. Trans. R. Soc. A* **373**, 20140424 (2015).

Acknowledgements We thank M. Fox and N. Rapp for assistance with compiling the proxy SST data and P. DiNezio for providing initial and boundary condition files for the CESM simulations. This research was supported by National Science Foundation grant numbers AGS-1602301 and AGS-1602223, and Heising-Simons Foundation grant number 2016-015. CESM computing resources (<https://doi.org/10.5065/D6RX99HX>) were provided by the Climate Simulation Laboratory at NCAR's Computational and Information Systems Laboratory, sponsored by the National Science Foundation and other agencies.

Author contributions J.E.T. designed the study, conducted the data assimilation, analysed the results and led the writing of this paper. J.E.T. and S.B.M. compiled and quality-checked the proxy SST data. S.B.M. designed the proxy database and adapted BACON age modelling software to Python. J.K. wrote the DASH code used for the data assimilation, based on methods developed by G.J.H. J.Z. and C.J.P. planned and conducted the iCESM simulations. All authors contributed to the writing of this manuscript.

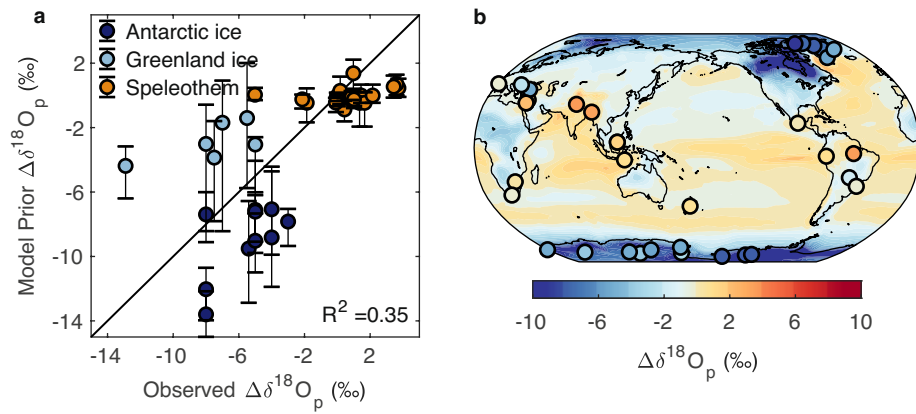
Competing interests The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.E.T.

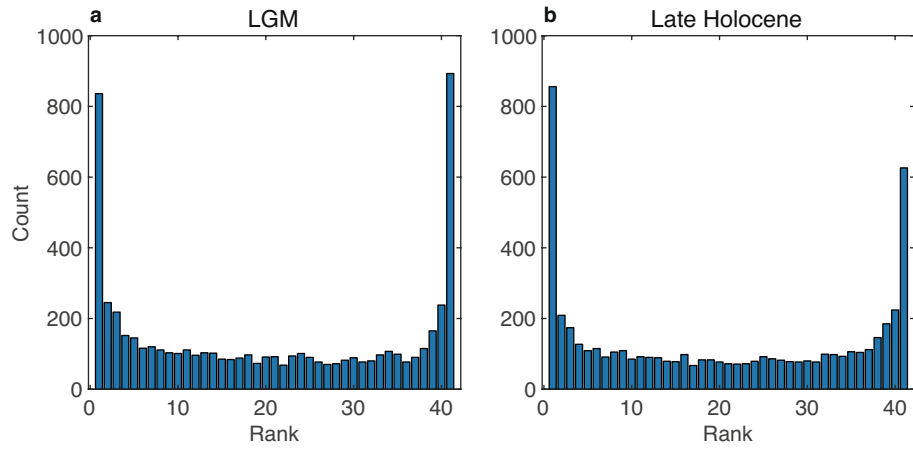
Peer review information *Nature* thanks Bernhard Naafs and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.

Reprints and permissions information is available at <http://www.nature.com/reprints>.

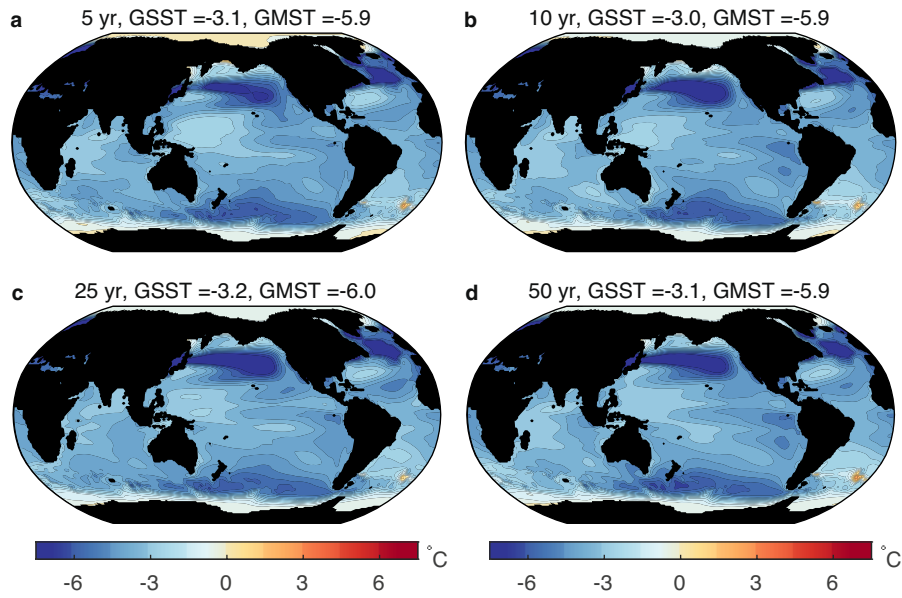


Extended Data Fig. 1 | Comparison of model prior $\delta^{18}\text{O}_p$ with speleothem and ice core proxies. **a**, Observed changes in ice core (Antarctica and Greenland) and speleothem LGM-LH $\delta^{18}\text{O}_p$ compared with the model prior ensemble. The R^2 value is shown in the lower right corner. **b**, Spatial map of

median changes in the $\delta^{18}\text{O}_p$ from the prior ensemble, overlain with LGM-LH ice core and speleothem observations (dots). Speleothem $\delta^{18}\text{O}$ was converted from $\delta^{18}\text{O}$ of calcite or aragonite to $\delta^{18}\text{O}_p$ (in ‰ VSMOW) before plotting (see Methods).

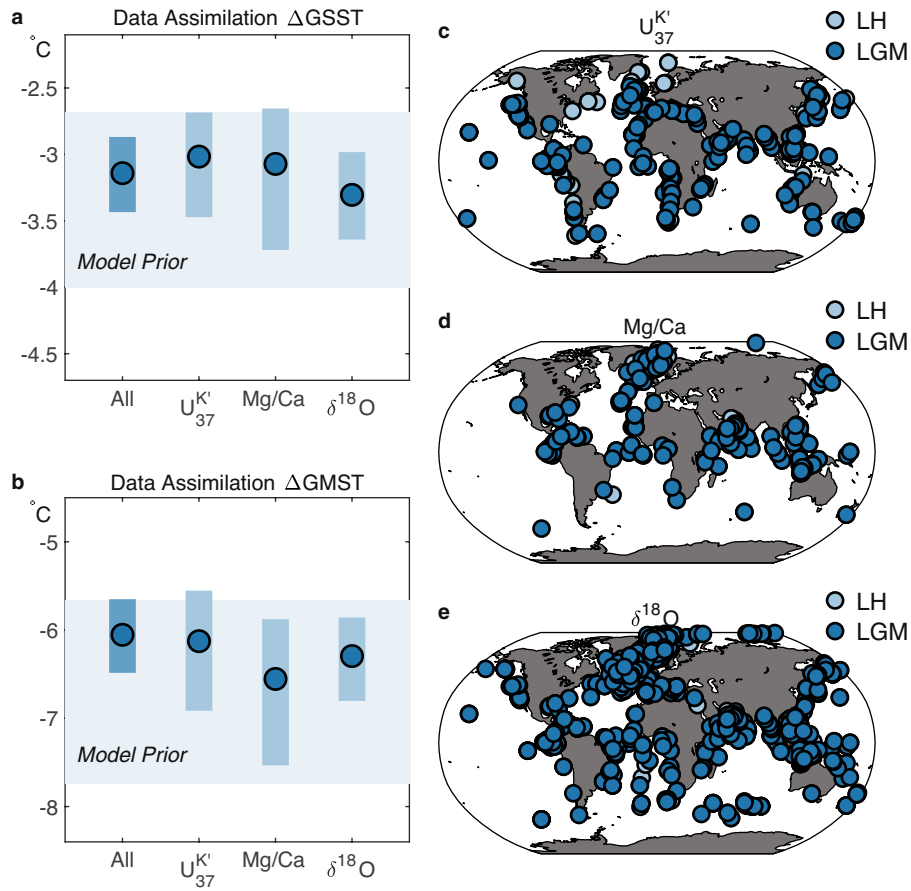


Extended Data Fig. 2 | Assessment of model bias with rank histograms. a, LGM, b, LH. The histograms are generally symmetrical, indicating little bias in the mean, but show a U shape that signals that the model prior may lack variability.



Extended Data Fig. 3 | The impact of time averaging on the assimilation results. a–d, The 40-member ensembles of 5-yr (a), 10-yr (b), 25-yr (c) and 50-yr (d) averages were assimilated with the same set of proxy data. Spatial

structures in the SST fields are largely similar. GSST and GMST values remain identical within uncertainty.



Extended Data Fig. 4 | Data assimilation results for individual proxy types. **a**, Assimilation-derived values for Δ GSST for all proxies combined (All), $U_{37}^{K'}$, Mg/Ca and $\delta^{18}\text{O}$, respectively. Error bars represent the 95% CI. Lightest blue

bounds show the range of the model ensemble prior. **b**, As in **a**, but for Δ GMST. **c–e**, Locations for $U_{37}^{K'}$ (**c**), Mg/Ca (**d**) and $\delta^{18}\text{O}$ (**e**) data. Lighter blue circles are Holocene data points, darker blue circles are LGM data points.

Extended Data Table 1 | Validation statistics associated with scaling the global estimate of the proxy variance (R_g)

		— LGM —						
	Prior	R_g	$R_g/2$	$R_g/5$	$R_g/10$	$R_g/20$	$R_g/100$	$R_g/500$
CE	0.83	0.88	0.90	0.91	0.93	0.93	0.93	0.89
RMSE	0.40	0.34	0.31	0.30	0.26	0.26	0.25	0.32
		— Late Holocene —						
CE	0.92	0.93	0.94	0.95	0.94	0.96	0.95	0.94
RMSE	0.28	0.26	0.25	0.23	0.25	0.21	0.23	0.25
		— LGM – LH vs. $\Delta\delta^{18}\text{O}_p$ —						
R^2	0.37	0.50	0.57	0.67	0.67	0.68	0.68	0.59

CE and RMSE are calculated on the 25% of the proxy data withheld from the assimilation. R^2 is calculated between observed $\Delta\delta^{18}\text{O}_p$ from speleothems and ice cores, and data assimilated $\Delta\delta^{18}\text{O}_p$ at the same locations. Localization was held constant at 12,000 km. Prior denotes comparison with the mean of the prior model ensemble.

Extended Data Table 2 | Validation statistics associated with varying the cutoff radius of the covariance localization

	— LGM —					
	∞	24,000	18,000	12,000	9,000	6,000
CE	0.88	0.92	0.92	0.91	0.92	0.93
RMSE	0.35	0.29	0.27	0.29	0.29	0.26
	— Late Holocene —					
CE	0.93	0.95	0.93	0.95	0.94	0.94
RMSE	0.26	0.23	0.26	0.23	0.24	0.24
	— LGM - LH vs. $\Delta\delta^{18}\text{O}_p$ —					
R^2	0.48	0.61	0.63	0.67	0.46	0.19

CE and RMSE are calculated on the 25% of the proxy data withheld from the assimilation. R^2 is calculated between observed $\Delta\delta^{18}\text{O}_p$ from speleothems and ice cores, and data assimilated $\Delta\delta^{18}\text{O}_p$ at the same locations. Cutoff radii are given in units of km; ∞ denotes no localization. Proxy variance R is held at $R_p/5$.

Extended Data Table 3 | Compilation of estimates of ΔR_{ICE} used for calculations of ECS

Model	ΔR_{ICE}	Reference
CCSM4	-3.79	PMIP3 ⁷⁸
IPSL-CM5A-LR	-4.90	PMIP3 ⁷⁸
MIROC-ESM	-5.20	PMIP3 ⁷⁸
MPI-ESM-P	-4.57	PMIP3 ⁷⁸
MRI-CGCM3	-3.62	PMIP3 ⁷⁸
CCSM3	-2.59	PMIP2 ⁶
CNRM	-2.66	PMIP2 ⁶
HadCM3M2	-3.23	PMIP2 ⁶
HadCM3M2 v	-3.41	PMIP2 ⁶
IPSL-CM4	-3.48	PMIP2 ⁶
MIROC3.2	-2.88	PMIP2 ⁶
CESM1.2	-3.29	This study
Mean	-3.64	
1σ	0.85	