LETTER

No evidence for globally coherent warm and cold periods over the preindustrial Common Era

Raphael Neukom[1]*, Nathan Steiger[2], Juan José Gómez-Navarro[3], Jianghao Wang[4] & Johannes P. Werner[5]

Earth's climate history is often understood by breaking it down into constituent climatic epochs[1]. Over the Common Era (the past 2,000 years) these epochs, such as the Little Ice Age[2–4], have been characterized as having occurred at the same time across extensive spatial scales[5]. Although the rapid global warming seen in observations over the past 150 years does show nearly global coherence[6], the spatiotemporal coherence of climate epochs earlier in the Common Era has yet to be robustly tested. Here we use global palaeoclimate reconstructions for the past 2,000 years, and find no evidence for preindustrial globally coherent cold and warm epochs. In particular, we find that the coldest epoch of the last millennium—the putative Little Ice Age—is most likely to have experienced the coldest temperatures during the fifteenth century in the central and eastern Pacific Ocean, during the seventeenth century in northwestern Europe and southeastern North America, and during the mid-nineteenth century over most of the remaining regions. Furthermore, the spatial coherence that does exist over the preindustrial Common Era is consistent with the spatial coherence of stochastic climatic variability. This lack of spatiotemporal coherence indicates that preindustrial forcing was not sufficient to produce globally synchronous extreme temperatures at multidecadal and centennial timescales. By contrast, we find that the warmest period of the past two millennia occurred during the twentieth century for more than 98 per cent of the globe. This provides strong evidence that anthropogenic global warming is not only unparalleled in terms of absolute temperatures[5], but also unprecedented in spatial consistency within the context of the past 2,000 years.

The study of past climate provides an essential baseline from which to understand and contextualize changes in the contemporary climate. Since the formative period of modern Earth sciences in the 1800s, the complex history of Earth's climate has been conceptualized through the construction of distinct climatic periods or epochs[1–7]. Several terms for climatic epochs within the past 2,000 years have come into wide use. Most prominent among these is the 'Little Ice Age' (LIA), a term that was originally created to broadly describe glacier growth in the Sierra Nevada mountains during the late Holocene (the past few thousand years)[2]; later, the LIA was used to describe inferred late Holocene glacial advances in many locations, particularly the European Alps[3,4]. Over the past few decades, this term has been widely used in palaeoclimatology and historical climatology to indicate a nearly global, centuries-long cold climate state that occurred between roughly 1300 AD and 1850 AD (refs [5,8]). This period is often contrasted with the Mediaeval Warm Period, also known as the Mediaeval Climate Anomaly (MCA)[8–10], which is commonly associated with warm temperatures in 800–1200 AD. The first millennium of the Common Era has also been subdivided into the 'Dark Ages Cold Period' (DACP)[11,12], or 'Late Antique Little Ice Age' (LALIA)[13], which occurred within about 400–800 AD, and lastly the 'Roman Warm Period' (RWP)[12,14], which covers the first few centuries of the Common Era. We note that for all of these epochs, no consensus exists about their precise temporal extent.

Each of these climatic epochs has its origin in pieces of palaeoclimatic evidence from the extratropical Northern Hemisphere, particularly Europe and North America[4,9–12]. Climate-epoch narratives were constructed to explain the early palaeoclimatic evidence, and later-developed time series from across the globe were situated within these narrative frameworks. This process probably created the expectation that Common Era climate epochs are global-scale phenomena. Loosely defined epochs based on a few dozen specific proxies were hard to falsify given the inherent noise of natural proxies, with, for example, nearly all annually resolved proxies that cover the Common Era having a signal-to-noise ratio of less than 1, and usually less than 0.5 (ref. [15]). Yet the association of a relatively small number of palaeoclimate proxy records with global-scale phenomena did not come without controversy and the discovery of proxy time series that did not match the standard climate-epoch narratives[4,10,16]. Studies that have attempted to assess the spatial coherence of Common Era climate epochs have used relatively few proxy records (for example, 14 proxy time series[17]), or only continentally averaged temperature reconstructions[18], or only one or two reconstruction methods[8,12]—a choice that has been shown to limit the reliability of the assessment of temperature patterns[19].

Here we test the hypothesis that there were globally coherent climate epochs over the Common Era by using a collection of probabilistic, global temperature reconstructions for the period 1–2000 AD, derived from a set of six different ensemble field reconstruction methodologies (see Methods; we note that we use 'coherence' here in its general, non-signal-processing sense). The reconstructions are based on techniques that vary widely in their assumptions and approaches to the reconstruction problem. They span a broad range of complexity, from basic proxy composites at the one end, to advanced statistical techniques at the other that incorporate physical constraints and forcing information from climate-model simulations. All methods use the same set of input data, namely the annual records from the recent PAGES 2k global temperature-sensitive proxy collection[20] (see Fig. 1 and Methods). This multimethod, probabilistic framework allows us to robustly assess the spatiotemporal homogeneity of climatic variability over the Common Era.

At the original annual resolution, the reconstruction ensemble mean shows no clear indication of a long period of years with globally consistent below-average temperatures relative to the mean for 1–2000 AD (Fig. 2a); the area fraction of warmth and cold shows high interannual variability. Of the years before 1850, 97% had at least 10% of the globe experiencing above-average temperatures, and 10% of the globe experiencing below-average temperatures. It is only if the reconstructed time series are smoothed over multidecadal timescales (see Methods), and if global area is shown in aggregate, that the classical picture of a loosely defined LIA and MCA appears (Fig. 2b and Extended Data Fig. 1). Yet the analysis in Fig. 2 does not include information from individual ensemble members (Extended Data Fig. 1); nor does it indicate spatial patterns of coherence, or provide a precise evaluation of the climate-epochs hypothesis.

[1]Oeschger Centre for Climate Change Research and Institute of Geography, University of Bern, Bern, Switzerland. [2]Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY, USA. [3]Department of Physics, University of Murcia, Murcia, Spain. [4]The MathWorks Inc, Natick, USA. [5]Bjerknes Center for Climate Research, Bergen, Norway. *e-mail: neukom@giub.unibe.ch
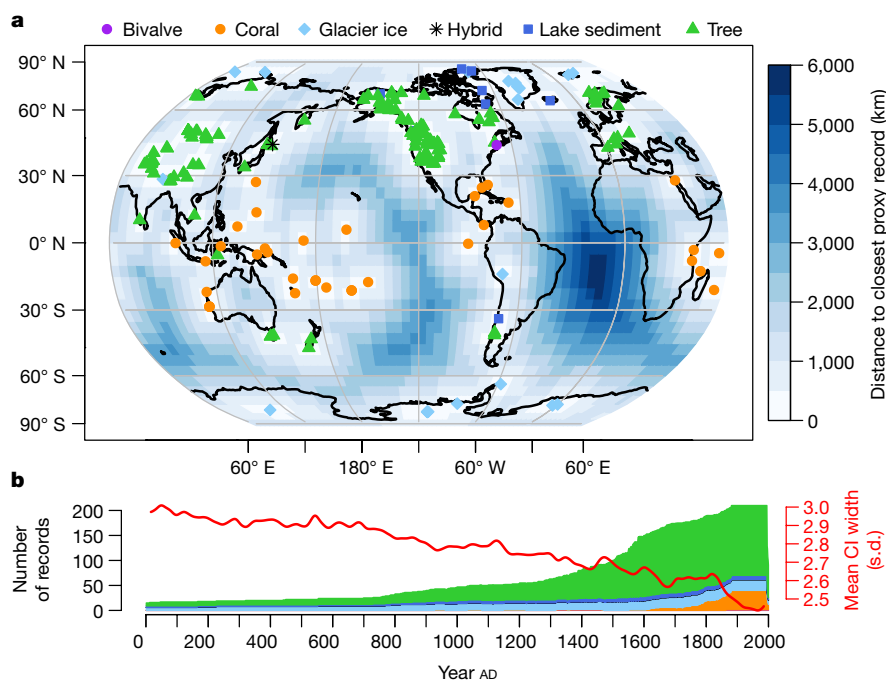
Fig. 1 | Spatiotemporal distribution of proxy data. a, Map showing the locations of the proxy records[20] used in our reconstructions, by archive type (bivalve, coral, glacier ice, hybrid, lake sediment or tree). Shading indicates the distance of a given 5° × 5° grid cell to the closest proxy record(s). **b**, Temporal availability of proxy data for each archive type, colour-coded as in panel **a**. The red line (right-hand y axis) shows the width of the 90% confidence interval (CI) for the unfiltered reconstructions, latitude-weighted and averaged over all methods. Values are relative to the instrumental temperature standard deviation over 1911–1995 AD.

To quantify the spatial coherence of cold and warm epochs, we consider the time of occurrence of a climate anomaly as the variable to be characterized within a probabilistic framework. We calculate the most probable period of peak warming or cooling during each of the five climatic epochs discussed previously (see Methods). At each grid-point location, we identify the warmest 51-year average within the epochs commonly referred to as warm, namely the RWP, MCA and current warm period (CWP). Analogously, we identify the coldest 51-year average for the DACP and LIA cold epochs. Given the lack of objective definitions for these epochs, we keep a wide window for our search for peak warming or cooling for each period (see Methods and

Fig. 3). To assess the CWP, we search for the warmest peak within the entire Common Era. The century within which we find the highest ensemble-based probability for maximum warming or cooling at each location is shown in Fig. 3.

There is considerable spatial heterogeneity in the timing of temperature maxima and minima (Fig. 3). No preindustrial epoch shows global coherence in the timing of the coldest or warmest periods. There is, however, regional coherence. For example, there are almost continental-scale patterns during many of the periods, and there is a coherent pattern in the tropical Pacific in the RWP, DACP and LIA periods, reminiscent of the El Niño–Southern
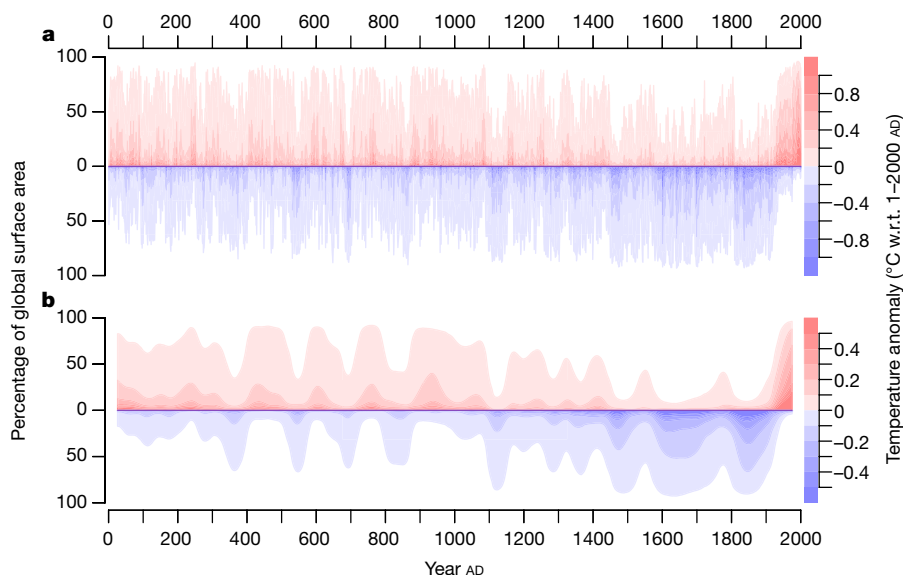

Fig. 2 | Distribution of warm and cold temperatures over the Common Era. The figure shows the percentages of global area with warm (red shading) and cold (blue shading) temperature anomalies with respect to a 1–2000 AD reference period (see Methods). Shading intensity indicates the magnitude of warmth and cold. **a**, Annual unfiltered data. **b**, 51-year lowpass filtered data (see Methods).
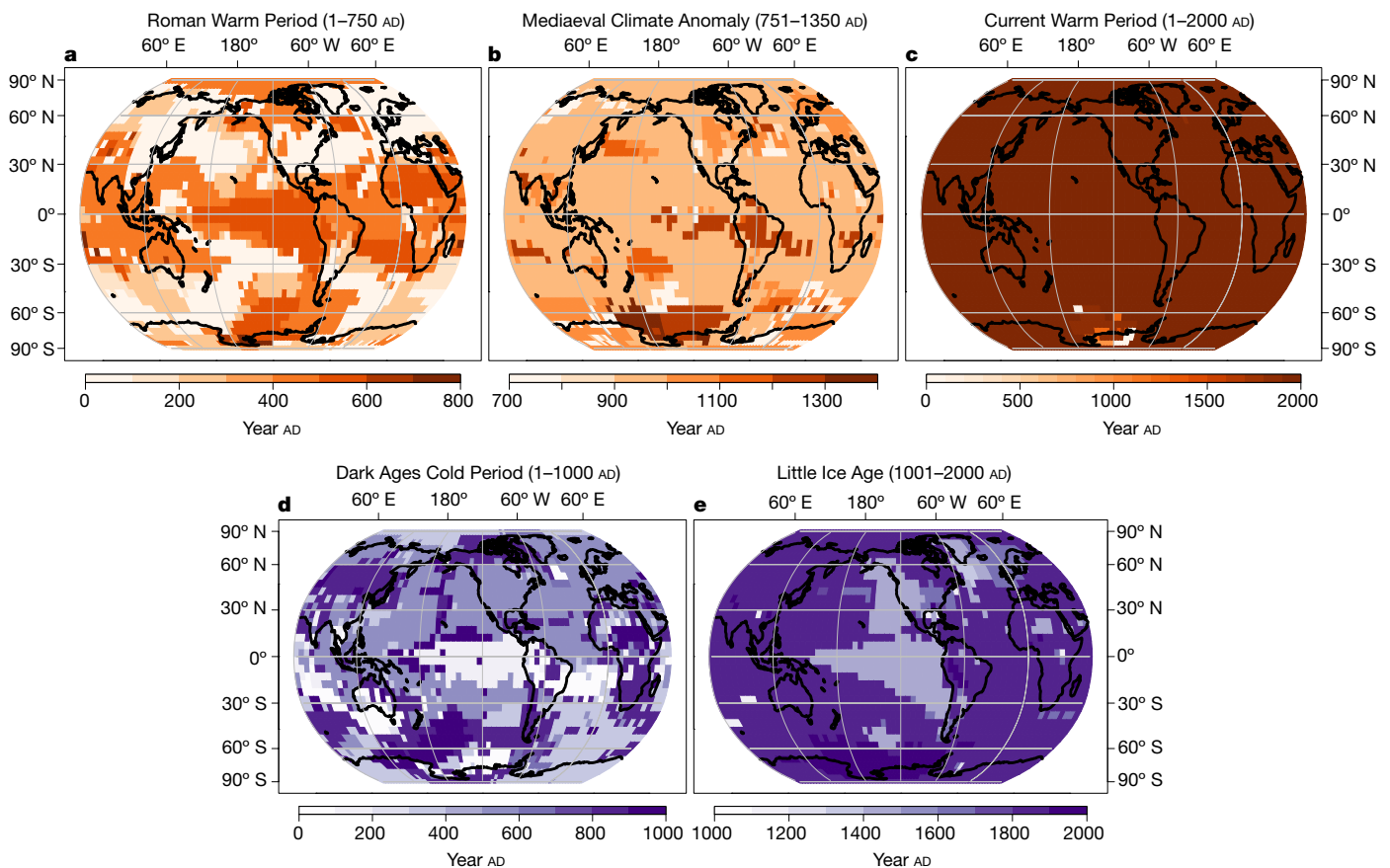
Fig. 3 | Timing of peak warm and cold periods. a–e, Centuries with the highest ensemble probability of containing the warmest (**a–c**) and coldest (**d, e**) 51-year period within each putative climatic epoch (see Methods). The full time ranges over which the search was performed for each epoch are indicated in parentheses. The numbers on the *y* axis and upper *x* axis are degrees latitude and longitude.

Oscillation—the most dominant mode of interannual variability in the climate system[21].

In contrast to the spatial heterogeneity of the preindustrial era, the highest probability for peak warming over the entire Common Era (Fig. 3c) is found in the late twentieth century almost everywhere (98% of global surface area), except for Antarctica, where contemporary warming has not yet been observed over the entire continent[22]. Thus, even though the recent warming rates are not entirely homogeneous over the globe, with isolated areas showing little warming or even cooling[22,23], the climate system is now in a state of global temperature coherence that is unprecedented over the Common Era.

Through a boostrapping uncertainty analysis (see Methods and Extended Data Fig. 2), we find that the particular spatial patterns shown in Fig. 3 are robust. Furthermore, the heterogeneity in the timing of maxima and minima is an inherent property of the input proxy data, which show a similar lack of global coherence in the timing of each putative climate epoch (Extended Data Fig. 3).

Is the amount of spatial coherence in the preindustrial period consistent with stochastic climate variability? We find that it is (Fig. 4), and that the spatial agreement across all reconstructions is typically low: in 84% of reconstruction ensemble members, less than 50% of the global area fraction agrees on the timing of the warmest or coldest 51-year peaks across all preindustrial epochs (Fig. 4). This supports the results shown in Fig. 3, providing evidence that peak preindustrial warm and cool periods occurred at different times in different locations. By contrast, the CWP shows distinct temporal and spatial agreement, with the warmest multidecadal peak of the Common Era occurring in the late twentieth century. The area fraction agreeing on the timing of the CWP is significantly larger than that expected from stochastic climate variability (Mann–Whitney U-test; *P* < 0.01; see Methods).

In addition to using an unprecedented collection of reconstruction methods to test the climate-epochs hypothesis, we conducted a range of sensitivity experiments, including noise-proxy reconstructions (Methods). These confirm the robustness of our results to the specific proxy network, the choices of reconstruction parameters, potential biases arising from the selection and calibration of proxies over the observational period, and the specific statistical tests of spatiotemporal coherence (see Methods and Extended Data Figs. 4–7). In addition, we confirm the lack of preindustrial spatial coherence in last-millennium climate-model simulations (Extended Data Fig. 8). Moreover, as in the reconstructions, the spatial consistency seen in model simulations over the twentieth century suggests that anthropogenic global warming is the cause of increased spatial temperature coherence relative to prior eras.

An important caveat to our results is that the spatiotemporal distribution of high-resolution proxy data is inherently unequal and often sparse. Future improvements in this regard may lead to better-resolved spatial patterns, especially in the Southern Hemisphere and during the first millennium, where uncertainties in our reconstructions are highest (Fig. 1, Extended Data Figs. 9, 10). However, such improvements are unlikely to lead to greater global coherence when the extant proxy data do not show indications of such (Extended Data Fig. 3).

The results shown here can explain at least two curious facts about climate epochs of the Common Era: the lack of consensus about the timing of climate epochs, and the discovery of records that do not fit the standard narratives. Peak warming and cooling events appear to be regionally constrained. Anomalous globally averaged temperatures during certain periods do not imply the existence of epochs of globally coherent and synchronous climate. This global asynchronicity suggests that multidecadal regional extremes are driven by regionally specific mechanisms, namely either unforced internal climate variability[24,25] or regionally varying responses to external forcing[26–28].
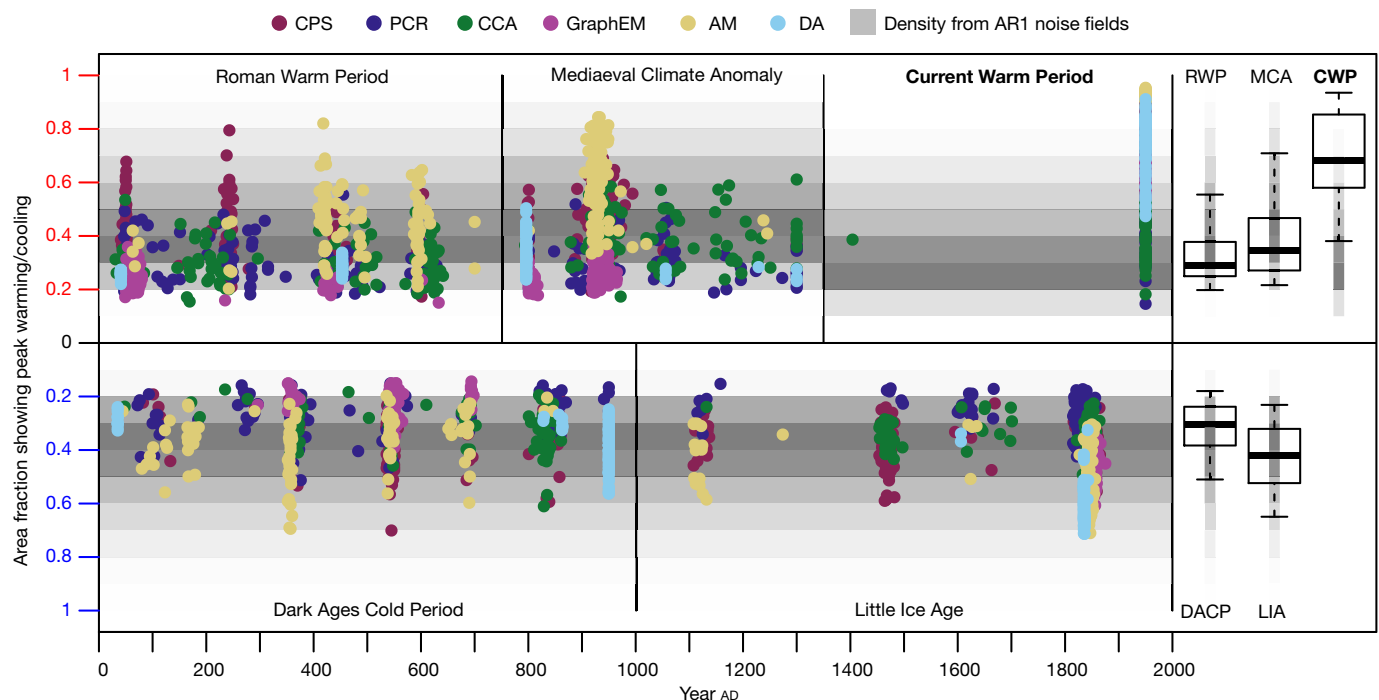
Fig. 4 | Spatial consistency of warm and cold periods. The figure shows the fraction of Earth's surface (*y* axis) that simultaneously experienced the warmest (top panels) or coldest (bottom panels) multidecadal period (51 years) during each of five different epochs (see Methods). Each solid circle represents an ensemble member plotted according to the year (*x* axis) in which the largest area experienced peak warm or cold conditions. Horizontal grey shading represents the distributions from the same analysis but based on multivariate noise fields from a first-order autoregressive (AR1) analysis, with darker colours indicating higher probabilities. Boxplots on the right show area fractions integrated over time (centre line, median; ends of boxes, interquartile range; whiskers,

90% range). Bold text indicates epochs with reconstructed area fractions that are significantly higher than those of the noise fields (Mann–Whitney U-test; $\alpha = 0.05$); only for the CWP do the area fractions from the reconstructions exceed those of the noise fields. The figure shows that peak warming and cooling vary temporally across ensemble members (circles distributed over a range of dates on the *x* axis), and affect only a limited fraction of global surface area (low absolute values on the *y* axis) before the industrial era. AM, analogue method; CCA, canonical correlation analysis; CPS, composite plus scale; DA, data assimilation; PCR, principal-component regression.

Given these results, we advocate for a regional framing for understanding the climate variability of the preindustrial Common Era. Likewise, the interpretation of individual palaeoclimate time series should not be forced to fit into global narratives or epochs. Rather, the a priori belief about a given palaeoclimate time series should be that it represents local information, with the extent of its correlation length scale to be justified and not assumed. In this framing, specific individual records can provide regional tests of the mechanisms of climate variability[29,30], while collections of many records can address larger scales. Against this regional framing, perhaps our most striking result is the exceptional spatiotemporal coherence during the warming of the twentieth century. This result provides further evidence of the unprecedented nature of anthropogenic global warming in the context of the past 2,000 years.

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 are available at https://doi.org/10.1038/s41586-019-1401-2.

1. Köppen, W. & Wegener, A. *Die Klimate der Geologischen Vorzeit* (Gebrüder Borntraeger, 1924).
2. Matthes, F. E. Report of Committee on Glaciers, April 1939. *Eos* **20**, 518–523 (1939).
3. Grove, J. M. *The Little Ice Age* (Methuen, 1988).
4. Matthews, J. A. & Briffa, K. R. The 'little ice age': re-evaluation of an evolving concept. *Geogr. Ann. A* **87**, 17–36 (2005).
5. Masson-Delmotte, V. et al. in *Climate Change 2013: The Physical Science Basis Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (eds Stocker, T. F. et al.) 383–464 (Cambridge Univ. Press, 2013).
6. Intergovernmental Panel on Climate Change. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ. Press, 2013).
7. Brückner, E. *Klimaschwankungen seit 1700 nebst Bemerkungen über die Klimaschwankungen der Diluvialzeit* (E. Hölzel, 1890).
8. Mann, M. E. et al. Global signatures and dynamical origins of the Little Ice Age and Medieval Climate Anomaly. *Science* **326**, 1256–1260 (2009).
9. Lamb, H. H. The early medieval warm epoch and its sequel. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **1**, 13–37 (1965).
10. Bradley, R. S., Hughes, M. K. & Diaz, H. F. Climate in medieval time. *Science* **302**, 404–405 (2003).
11. Helama, S., Jones, P. D. & Briffa, K. R. Dark Ages Cold Period: a literature review and directions for future research. *Holocene* **27**, 1600–1606 (2017).
12. Ljungqvist, F. C. A new reconstruction of temperature variability in the extra-tropical Northern Hemisphere during the last two millennia. *Geogr. Ann. A* **92**, 339–351 (2010).
13. Büntgen, U. et al. Cooling and societal change during the Late Antique Little Ice Age from 536 to around 660 AD. *Nat. Geosci.* **9**, 231–236 (2016).
14. Röthlisberger, F. *10,000 Jahre Gletschergeschichte der Erde* (Sauerländer, 1986).
15. Wang, J., Emile-Geay, J., Guillot, D., Smerdon, J. E. & Rajaratnam, B. Evaluating climate field reconstruction techniques using improved emulations of real-world conditions. *Clim. Past* **10**, 1–19 (2014).
16. Bradley, R. 1000 years of climate change. *Science* **288**, 1353–1355 (2000).
17. Osborn, T. J. The spatial extent of 20th-century warmth in the context of the past 1200 years. *Science* **311**, 841–844 (2006).
18. PAGES2k Consortium. Continental-scale temperature variability during the past two millennia. *Nat. Geosci.* **6**, 339–346 (2013); erratum **6**, 503 (2013); corrigendum **8**, 981–982 (2015).
19. Wang, J., Emile-Geay, J., Guillot, D., McKay, N. P. & Rajaratnam, B. Fragility of reconstructed temperature patterns over the Common Era: implications for model evaluation. *Geophys. Res. Lett.* **42**, 7162–7170 (2015).
20. PAGES2k Consortium. A global multiproxy database for temperature reconstructions of the Common Era. *Sci. Data* **4**, 170088 (2017).

21. McPhaden, M. J., Zebiak, S. E. & Glantz, M. H. ENSO as an integrating concept in Earth science. *Science* **314**, 1740–1745 (2006).
22. Stenni, B. et al. Antarctic climate variability on regional and continental scales over the last 2000 years. *Clim. Past* **13**, 1609–1634 (2017).
23. Caesar, L., Rahmstorf, S., Robinson, A., Feulner, G. & Saba, V. Observed fingerprint of a weakening Atlantic Ocean overturning circulation. *Nature* **556**, 191–196 (2018).
24. Wang, J. et al. Internal and external forcing of multidecadal Atlantic climate variability over the past 1,200 years. *Nat. Geosci.* **10**, 512–517 (2017).
25. Delworth, T. L. et al. The North Atlantic Oscillation as a driver of rapid climate change in the Northern Hemisphere. *Nat. Geosci.* **9**, 509–512 (2016).
26. Hegerl, G. C., Brönnimann, S., Schurer, A. & Cowan, T. The early 20th century warming: anomalies, causes, and consequences. *Wiley Interdiscip. Rev. Clim. Change* **9**, e522 (2018).
27. Abram, N. J. et al. Early onset of industrial-era warming across the oceans and continents. *Nature* **536**, 411–418 (2016); corrigendum **545**, 252 (2017).
28. Bindoff, N. L. et al. in *Climate Change 2013: The Physical Science Basis* (eds Intergovernmental Panel on Climate Change) 867–952 (Cambridge Univ. Press, 2013).
29. Collins, M. et al. Challenges and opportunities for improved understanding of regional climate dynamics. *Nat. Clim. Chang.* **8**, 101–108 (2018).
30. Xie, S.-P. et al. Towards predictive understanding of regional climate change. *Nat. Clim. Chang.* **5**, 921–930 (2015).

METHODS

Instrumental target. We use the HadCRUT4 global temperature grid[31] with a spatial resolution of 5° × 5°, infilled using GraphEM, to obtain complete global coverage over the calibration period[20]. We use annual values aggregated over the April-to-March seasonal window, which reflects the 'tropical year' and, unlike a calendar-year window, does not interrupt the austral growing season.

Proxy data. We use the data from the PAGES 2k temperature database v2.0.0[20] as predictors. For the results shown in the main text, we use a screened network based on regional false-discovery-rate screening (R-FDR)[20], which reduces the number of proxies from the original 688 to 257. The spatial coverages of the full and screened networks are displayed in Extended Data Fig. 3 and Fig. 1, respectively. The screened network yields improved reconstruction skill for most methods over much of the globe. However, our conclusions are robust to the choice of either the full or the screened network (Extended Data Fig. 4). The present implementation of some of the reconstruction methods used herein does not allow us to incorporate proxy data with gaps, or missing values over the calibration period. Therefore we use only records with annual or higher resolution (210 records; Fig. 1), thereby using mainly records with very small to negligible age uncertainties (85% of the records used are from tree or coral archives). Records of subannual resolution are averaged to annual resolution over the April-to-March window. Missing values in the proxy matrix over the calibration/validation window (2.2%) were infilled using DINEOF[32].

Reconstruction parameters. The calibration period used for all methods is 1911–1995 AD. This reflects a compromise between using as many years as possible to cover a large temperature range for calibration, and using a period with sufficient spatial coverage of instrumental data (at the beginning of the calibration period) or proxy records (at the end). We use 100-member ensembles for all methods to generate the analyses and plots presented in the main text, yielding a total ensemble size of 600.

We use the period 1881–1910 AD for validation to compare the relative performance of the reconstruction methods (Extended Data Figs. 9, 10). We use the following metrics to assess the performance of our reconstructions. The continuous ranked probability score[33,34] (CRPS) has been conceptually adapted to mimic the reduction-of-error (RE) and coefficient-of-efficiency (CE) scores[35]. The CRPS of the reconstructions is subtracted from the CRPS generated from surrogates based on instrumental target data according to refs [36,37]. CRPS_RE (CRPS_CE) compares the mean potential CRPS of the reconstruction with the instrumental surrogates over the 1911–1995 AD calibration (1881–1910 AD validation) period. In contrast to the traditional RE and CE measures, these metrics are strictly proper scoring rules[33]. For a detailed description of the metrics, see ref. [37]. The root mean squared error (RMSE) of the ensemble median over the validation period is divided by the instrumental standard deviation at each location to allow a relative comparison at different locations. Finally we calculate the Pearson correlation coefficient of the reconstruction ensemble median with the target over the validation period.

While the spatial distribution of the proxy network is global, the Northern Hemisphere contains more proxies than the Southern Hemisphere, and the number of proxies decreases further back in time (Fig. 1). Consequently, the reconstructions generally have highest confidence and intermethod agreement closer to the present (Fig. 1b, Extended Data Fig. 10) and nearest to the proxy locations (Extended Data Fig. 9). We note that no particular method stands out as being particularly more or less skillful than the others (Extended Data Fig. 9).

Reconstruction methods. *Composite plus scale (CPS).* CPS is a widely used index reconstruction method that has been used to reconstruct local to global mean climate[38,39]. The input proxy data are averaged into a composite time series, which is then scaled to the mean and standard deviation of the reconstruction target over the calibration period. Here we use a point-by-point[40] implementation of CPS, probably the most simple way to reconstruct a climate field. This method does not make use of the spatial covariance structure of the target temperature field, which may lead to unrealistic spatial consistency in the reconstructed fields. On the other hand, the point-by-point approach is, unlike most other CFR techniques, not bound to the problematic assumption that the spatial temperature patterns in the calibration period are stable over time.

Here, we use the CPS implementation of ref. [41], which weights the proxy records by their correlation with the (grid cell) target. We do not limit the number of proxies at each location by using a maximum search radius, as the spatial decorrelation distance is not uniform over the globe; in addition, this allows information from teleconnected areas to be included in the reconstructions. We use an ensemble approach similar to that of ref. [41], combining uncertainties arising from parameter decisions and calibration errors. The following reconstruction parameters are resampled for each reconstruction member: proxy selection (removing 10% of records); calibration period (removing a block of ten years within the 1911–1995 AD calibration period); proxy weight (multiplying the correlation-based weight by a factor between 1/1.5 and 1.5). Ensemble perturbation based on the calibration error is implemented as in ref. [15]: we add to each ensemble member multivariate noise time series with the same standard deviation as the residuals

between the target and reconstructed field over the calibration period. We use first-order autoregressive (AR(1)) noise with the same AR(1) coefficients as the residuals. We use a nested approach, which means that the reconstruction process is repeated for each time period over the Common Era with unique proxy availability. The results of each nest are spliced together to obtain a reconstruction covering the full Common Era.

Principal-component regression (PCR). PCR has been widely used in climate field reconstructions[42–46]. This method reduces the dimensions of both the target field and the proxy matrix using principal-component analysis (PCA). The instrumental principal components are predicted back in time on the basis of regression with the proxy principal components, and then back-transformed to the spatial dimensions of the target grid using the loadings from the PCA. As such, this approach assumes that the covariance structure of the temperature grid remains the same over the reconstruction period as in the time window used for calibration. Here we use the PCR approach introduced in ref. [42] and an ensemble integration similar to that in refs [41,44]. The nested and ensemble approach is identical to that of the CPS (described above), with the following exceptions. The random weight factor is multiplied by the weight of each proxy derived from the PCA. We also resample the principal-component truncation parameters for the proxy (or instrumental) matrices in such a way that the retained principal components explain between 40% and 90% (or between 60% and 99%) of the total variance. We resample this parameter for two reasons: because there are multiple existing approaches to truncating principal components without an objectively discernible best method; and because the truncations are often sensitive to the period over which the PCA is performed.

In contrast to earlier studies using PCR, we do not readjust the variance of the reconstructed field to the target variance over the calibration period. This readjustment is often done to avoid strong variance changes between the reconstructions of the different proxy nests. We do not apply this correction because in our case the differences in variance between the adjusted and non-adjusted reconstructions are relatively small. The largest effect of the readjustment is that it produces a large ratio of low- versus high-frequency variance in the reconstructions, particularly over the high northern latitudes, leading to reconstructed temperatures that are very low compared with those obtained by other methods over the entire reconstruction period.

Canonical correlation analyis (CCA). CCA uses singular-value decomposition (SVD) to carry out dimensional reductions separately on the instrumental temperature matrix, the proxy matrix, and the regression coefficient matrix that describes their relationships[47,48]. The basic assumption, as in most palaeoclimate applications, is that the first few leading modes of empirical orthogonal function (EOF)–principal-component pairs contain most of the variance in the target climate field and the multiproxy network. The algorithm seeks an optimal set of truncation parameters that yields good approximations of the above-mentioned matrices. These truncation parameters are chosen by minimizing the area-weighted RMSE of the reconstruction relative to the target field using a leave-half-out cross-validation procedure. Specifically, a separate set of parameters was obtained for each proxy nest, so that the algorithm accounts for the heterogeneity in data availability and can thus adaptively regularize the regression matrix. This adaptive procedure was developed[15,19] as an improvement to the original algorithm[48]. Ensemble perturbations were done as for PCR above.

GraphEM. GraphEM[49] is based on the theory of Gaussian graphical models (GGMs or Markov random fields). A GGM makes use of the conditional independence structure of the climate field, in order to reduce the dimensionality and obtain a parsimonious estimate of the inverse of the covariance matrix $\hat{\Sigma}$. The conditional independence relations are estimated by solving an l_1 (lasso)-penalized maximum likelihood problem[50]. Σ is then estimated in accordance with these conditional independence relations. The resulting $\hat{\Sigma}$ is sparse and better conditioned, and therefore is applicable within the ordinary-least-squares framework. This procedure is implemented within the standard expectation-maximization algorithm without further need for regularization.

Specifically, the covariance model is chosen via the graphical lasso algorithm[50]. Three sparsity parameters need to be specified to determine the graphical structure, separately for the temperature–temperature (TT), proxy–proxy (PP) and temperature–proxy (TP) parts of the covariance matrix. The values need to be large enough that the true graph is contained within the estimated one, but small enough for the covariance matrix to be well conditioned. We set the parameters to (TT, TP) = (2%, 2%), with a diagonal matrix for PP, reflecting the conditional independence of proxies given the temperature field.

Data assimilation (DA). We use an off-line data-assimilation technique that optimally combines proxy data with climate-model states[51,52]. The climate model provides an initial, or prior, state estimate that is updated on the basis of the proxy observations and an estimate of the errors in both the observations and the prior. For the annual reconstructions here, the reconstruction is made by iteratively computing the state update equations of data assimilation for each year of the existing

proxy data without the need to run an ensemble of online simulations (see ref. [53] for precise mathematical details and links to data-assimilation palaeoclimate reconstruction code). As in previous work[53], we construct the prior using simulation number 10 from the Community Earth System Model Last Millennium Ensemble (CESM LME)[54].

For the prior we specifically used the middle 998 years of the CESM LME simulation, excluding the two simulation endpoints, to create a static 998-member prior ensemble that we used to estimate the climate state in each year of the reconstruction. As in ref. [53], we performed sensitivity tests with different members from the CESM LME and found no discernible differences in the results.

As in ref. [52], the proxies in the data-assimilation framework are modelled as linear univariate responders to temperature; the errors in the model's estimate of the proxies are taken as the residuals of a local linear univariate fit between the proxies and HadCRUT4 over the calibration period. This method naturally provides uncertainty estimates for the target field in the form of an ensemble of equally likely state estimates for each year. For the analyses performed here, we used a random sampling of 100 of these state estimates from the available 998 members.

Analogue method (AM). The analogue method has been successfully used as a CFR technique to reconstruct global temperature fields[55]. The method requires the existence of a pool of plausible climate fields, which can be obtained from computer simulations, observations or combinations of both. As in ref. [55], we use the ensemble of available simulations within the PMIP3 project, which consists of 18,327 annual temperature fields. For direct comparison with the climate fields, each proxy time series is converted into a 'local temperature reconstruction' through a local linear univariate fit with HadCRUT4. We then calculate the distance between each field and the local, proxy-based reconstructions in a given year; the distance is based on minimizing the spatial RMSE between the local reconstructions and the pool sampled at the closest grid point to each proxy location. The full CFR is then the average of the five fields that most closely resemble the local reconstructions at a given time step[55]. Thus the spatial structure of temperature is provided by the different climate models, while the temporal evolution is obtained from the information contained in the network of proxies.

To produce an ensemble that accounts for the probabilistic nature of this reconstruction and their uncertainties, we apply a bootstrap-based approach. We subsample both the pool of analogues and the local reconstructions so that in each instance only half of the information is used. This degradation of the available information naturally leads to spread within the ensemble that is larger around locations in which the pool is loosely constrained by the proxies, and therefore allows us to quantify the uncertainties implicit in the reconstruction.

Spatial anomalies in Fig. 2. To generate Fig. 2, the ensemble median reconstructions of each method and at each grid cell are first centred to the reference period 1–2000 AD. The (area-weighted) fraction of grid cells that exceed the temperature thresholds shown using the right-hand colour bar of Fig. 2 is then calculated. The values from the six methods are then averaged as shown in Fig. 2. In Fig. 2b, the reconstruction time series are smoothed with a 51-year butterworth lowpass filter before analysis[56].

Analysis of cold and warm periods. *Figure 3.* To calculate the timing of peak warming/cooling over the climatic epochs, we first calculate 51-year running temperature averages for each location and ensemble member. For each warm (or cold) epoch and grid cell, we then identify the year with the maximum (or minimum) value, using the centre-year of the warmest (or coldest) 51-year period. In the following we use the term 'peak-year' for this maximum (or minimum). We then identify the century within which the largest number of members (over the combined 600-member reconstruction ensemble) have the peak-year. This century is then indicated in Fig. 3 for each grid point. The maps thus show the century within which the multidecadal peak warming (or largest cooling) is most probable for each epoch. The DACP and LIA minima are searched for within the first and second millennium AD, respectively. RWP maxima are allowed to occur within 1–750 AD, MCA maxima between 751–1350 AD, and the CWP in the full 2,000 years of the Common Era.

Uncertainty analysis for Fig. 3. Uncertainties in the analysis are quantified by bootstrapping. We recalculated the peak warm/cold analysis described above 1,000 times using 600 bootstrap samples drawn from the reconstruction ensemble members (Extended Data Fig. 2). We find that the particular spatial patterns shown in Fig. 3 are robust, with at least 75% of locations having a 1σ range of less than 50 years for all epochs except the DACP. For this epoch, 33% of locations show a 1σ range of over 100 years, mostly concentrated in a few Southern Hemisphere and tropical regions (Extended Data Fig. 2).

Sensitivity tests for Fig. 3. We also conducted a number of sensitivity tests. We tested using an alternative period length of 101 years (instead of 51 years) to calculate running averages (Extended Data Fig. 5) and using the full proxy network instead of the screened proxy network (Extended Data Fig. 4). We computed epoch maps using the raw proxy data instead of the reconstructed fields (Extended Data Fig. 3). We also ran reconstruction experiments using detrended calibration data

to test for potential artefacts arising from the twentieth-century calibration period (Extended Data Fig. 8).

There is also the potential concern that the proxy screening process and the reconstruction methodologies themselves could produce global coherence in the twentieth century (Fig. 3) given noisy proxies (thus no 'real' underlying coherence). In this hypothetical scenario, signal-less proxies that by chance contain a trend in the twentieth century could be selected in the screening process and weighted strongly in the reconstruction process, thus giving rise to a false sense of twentieth-century coherence.

To test this null hypothesis, we generated three kinds of noise proxies that we then used within each of the reconstruction routines. For the first kind, we generated noise proxies that are in the same locations and have the same autoregressive spectrum and temporal coverage as the 210 real proxies used in our reconstructions[57,58]. This kind of noise proxy assesses the role that the reconstruction methodologies themselves have in potentially biasing the result of Fig. 3. The second kind of proxies is the same as the first except that we additionally applied the R-FDR proxy-screening process to noise proxies representing the full PAGES2k database of 515 annually or higher resolved proxies[20]. This screening leaves on average about $n = 66$ noise proxies and may shed light on the influence of the screening approach on the results. As a final, even more conservative noise proxy experiment, we force-screened the noise proxies to have the same number ($n = 210$) as the screened real proxy data used herein. This means that we repeatedly generate noise-proxy time series at each location until the time series passes the R-FDR screening criterion. These second and third kind of proxies test the role that screening plays in the reconstruction process.

Of these three noise proxy experiments, we consider the second ($n = 66$) to be the most likely representative null experiment against which to compare the results of Fig. 3. Although the third, $n = 210$ experiment uses the same number of input data, it does not accurately reflect the process of how proxy data are selected for the real data reconstructions. Repeatedly generating noise proxies at each location until a certain number of them pass the screening will increase the potential bias of the screening effect and will further amplify and build the observational signal into the proxy network, thereby eroding the 'noisiness' of the null.

We generated 25 sets of noise proxy networks for the three kinds of noise proxies and performed 100-member ensemble reconstructions for each reconstruction method, thus producing 45,000 global noise-proxy reconstructions. The results (Extended Data Fig. 7) indicate that, although it is possible in some proxy noise realizations to generate artificial twentieth-century warming coherence from the reconstruction algorithms themselves as well as from the proxy screening process, neither of these factors can explain the amount of twentieth-century warming coherence (Extended Data Fig. 7c). We note also that the twentieth-century warming in the noise proxy reconstructions is a product of largely three reconstruction methods (GraphEM, data assimilation and the analogue method, depending on the screening approach). We find that the noise-proxy reconstructions produce coherence that is always less than that seen in the real proxy reconstructions over all noise-proxy experiments and methods. The median global area fraction showing the largest ensemble probability for maximum 51-year temperatures within the twentieth century is between 37% and 67% for the noise proxies, depending on the screening approach, versus a 98% global area fraction for real proxies for both unscreened and screened networks (Extended Data Fig. 7c) and climate-model data (Extended Data Fig. 8).

All of these analyses, together with the independent verification using climate-model data shown in Extended Data Fig. 8, corroborate the results of Fig. 3 and show that they are robust to major technical choices.

Figure 4. While Fig. 3 is based on ensemble probabilities, Fig. 4 addresses the spatial consistency of warm and cool extremes within each ensemble member. To generate Fig. 4, we first calculate 51-year running averages of each reconstruction ensemble member and identify the epoch peak-years (the maximum values for warm epochs or minimum values for cold epochs) at each grid point. Then we find which sliding 51-year period contains the most peak-years in terms of the global area fraction. The centre year of this 51-year period is shown on the x axis of Fig. 4, while the y axis is the area-weighted fraction of peak-years that are contained within the 51-year period. This process thus identifies the period for which there is the largest spatial agreement about multidecadal temperature extremes. For example, 24% of grid cells in PCR ensemble member 1 have the LIA peak-year (that is, the coldest 51-year average during the LIA epoch) within the period 1815–1865 AD. No other 51-year period within the LIA epoch contains a higher area fraction of LIA minima in this ensemble member. The circle for this ensemble member is therefore drawn at the coordinates (1840, −0.24) in Fig. 4. Boxplots on the right-hand side of Fig. 4 integrate the area fractions of all ensemble members independently of the timing.

As a null reference to test the significance of the area fractions, we repeated the Fig. 4 calculations on the basis of multivariate random fields with realistic spatiotemporal covariance properties[15] using the *rmvn* function in the R package

mgcv. First, the covariance matrix V of the ensemble median reconstruction field over the full 2,000 years is calculated for each reconstruction method. Second, a 'square root' of V is generated using pivoted cholesky decomposition[59]. A matrix of normal white noise is then multiplied with the transpose of this square root matrix to obtain a multivariate normal matrix with the same covariance structure as the reconstructed field. Each grid cell in this random field is then modified with an AR(1) model to obtain a time series with the same first-order autoregression coefficient as the corresponding grid cell in the ensemble median reconstruction[60]. This process is repeated 100 times for each reconstruction method to obtain a 600-member ensemble as in the real-world reconstructions. We then perform the search for the peak warming/cooling for each epoch in this noise-field ensemble using the process described above. The area fractions resulting from these noise fields are shown with grey shading in Fig. 4. We use the Mann–Whitney U-test ($\alpha = 0.05$; one-tailed) to test whether the area fractions in the reconstruction ensembles are significantly larger than expected from these noise fields (both $n = 600$). Alternative benchmarks based on noise-proxy reconstructions are shown in Extended Data Fig. 6. As for Fig. 3, all of our sensitivity experiments (Extended Data Figs. 4–6) confirm the robustness of our findings.

Data availability
The PAGES 2k v.2.0.0 dataset is archived at the World Data Service (WDS) for Paleoclimatology (hosted by the National Oceanic and Atmospheric Administration (NOAA)), formatted for both LiPD and WDS ASCII template (https://www.ncdc.noaa.gov/paleo/study/21171). The screened input data matrix and instrumental target grid, as well as the reconstruction outcomes from this study, are available at Figshare (doi:10.6084/m9.figshare.c.4498373.v1) and NOAA WDS Paleoclimatology (www.ncdc.noaa.gov/paleo/study/26850). We strongly recommend using the multimethod ensembles when working with the reconstructions. For analyses of global mean temperatures we recommend using the reconstruction of the PAGES 2k companion project that explicitly targets the global mean[61].

Code availability
The code to generate the figures is available with the output data (see above).

31. Morice, C. P., Kennedy, J. J., Rayner, N. A. & Jones, P. D. Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: the HadCRUT4 data set. *J. Geophys. Res.* **117**, D08101 (2012).
32. Taylor, M. H., Losch, M., Wenzel, M. & Schröter, J. On the sensitivity of field reconstruction and prediction using empirical orthogonal functions derived from gappy data. *J. Clim.* **26**, 9194–9205 (2013).
33. Gneiting, T. & Raftery, A. E. Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.* **102**, 359–378 (2007).
34. Werner, J. P. & Tingley, M. P. Technical note: Probabilistically constraining proxy age–depth models within a Bayesian hierarchical reconstruction model. *Clim. Past* **11**, 533–545 (2015).
35. Cook, E. R., Briffa, K. R. & Jones, P. D. Spatial regression methods in dendroclimatology: a review and comparison of two techniques. *Int. J. Climatol.* **14**, 379–402 (1994).
36. Tipton, J., Hooten, M., Pederson, N., Tingley, M. & Bishop, D. Reconstruction of late Holocene climate based on tree growth and mechanistic hierarchical models. *Environmetrics* **27**, 42–54 (2016).
37. Werner, J. P., Divine, D. V., Charpentier Ljungqvist, F., Nilsen, T. & Francus, P. Spatio-temporal variability of Arctic summer temperatures over the past 2 millennia. *Clim. Past* **14**, 527–557 (2018).
38. Jones, P. et al. High-resolution palaeoclimatology of the last millennium: a review of current status and future prospects. *Holocene* **19**, 3–49 (2009).
39. Mann, M. E. et al. Proxy-based reconstructions of hemispheric and global surface temperature variations over the past two millennia. *Proc. Natl Acad. Sci. USA* **105**, 13252–13257 (2008).
40. Cook, E. R. et al. Asian monsoon failure and megadrought during the last millennium. *Science* **328**, 486–489 (2010).
41. Neukom, R. et al. Inter-hemispheric temperature variability over the past millennium. *Nat. Clim. Chang.* **4**, 362–367 (2014).
42. Luterbacher, J. et al. Reconstruction of sea level pressure fields over the Eastern North Atlantic and Europe back to 1500. *Clim. Dyn.* **18**, 545–561 (2002).
43. Luterbacher, J., Dietrich, D., Xoplaki, E., Grosjean, M. & Wanner, H. European seasonal and annual temperature variability, trends, and extremes since 1500. *Science* **303**, 1499–1503 (2004).
44. Neukom, R. et al. Multi-centennial summer and winter precipitation variability in southern South America. *Geophys. Res. Lett.* **37**, L14708 (2010).
45. Neukom, R. et al. Multiproxy summer and winter surface air temperature field reconstructions for southern South America covering the past centuries. *Clim. Dyn.* **37**, 35–51 (2011).
46. Smerdon, J. E. & Pollack, H. N. Reconstructing Earth's surface temperature over the past 2000 years: the science behind the headlines. *Wiley Interdiscip. Rev. Clim. Change* **7**, 746–771 (2016).
47. Christiansen, B., Schmith, T. & Thejll, P. A surrogate ensemble study of climate reconstruction methods: stochasticity and robustness. *J. Clim.* **22**, 951–976 (2009).
48. Smerdon, J. E., Kaplan, A., Chang, D. & Evans, M. N. A pseudoproxy evaluation of the CCA and RegEM methods for reconstructing climate fields of the last millennium. *J. Clim.* **23**, 4856–4880 (2010).
49. Guillot, D., Rajaratnam, B. & Emile-Geay, J. Statistical paleoclimate reconstructions via Markov random fields. *Ann. Appl. Stat.* **9**, 324–352 (2015).
50. Friedman, J., Hastie, T. & Tibshirani, R. Sparse inverse covariance estimation with the graphical lasso. *Biostatistics* **9**, 432–441 (2008).
51. 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).
52. Hakim, G. J. et al. The last millennium climate reanalysis project: framework and first results. *J. Geophys. Res. D* **121**, 6745–6764 (2016).
53. Steiger, N. J., Smerdon, J. E., Cook, E. R. & Cook, B. I. A reconstruction of global hydroclimate and dynamical variables over the Common Era. *Sci. Data* **5**, 180086 (2018).
54. Otto-Bliesner, B. L. et al. Climate variability and change since 850 CE: an ensemble approach with the Community Earth System Model. *Bull. Am. Meteorol. Soc.* **97**, 735–754 (2016).
55. Gómez-Navarro, J. J., Zorita, E., Raible, C. C. & Neukom, R. Pseudoproxy tests of the analogue method to reconstruct spatially resolved global temperature during the Common Era. *Clim. Past* **13**, 629–648 (2017).
56. Proakis, J. G. & Manolakis, D. G. Digital Signal Processing Principles, Algorithms, and Applications (Macmillan, 1992).
57. Hosking, J. R. M. Modeling persistence in hydrological time series using fractional differencing. *Wat. Resour. Res.* **20**, 1898–1908 (1984).
58. Wahl, E. R. & Smerdon, J. E. Comparative performance of paleoclimate field and index reconstructions derived from climate proxies and noiseonly predictors. *Geophys. Res. Lett.* **39**, L06703 (2012).
59. Higham, N. J. Cholesky factorization. *Wiley Interdiscip. Rev. Comput. Stat.* **1**, 251–254 (2009).
60. Neukom, R., Schurer, A. P., Steiger, N. J. & Hegerl, G. C. Possible causes of data model discrepancy in the temperature history of the last millennium. *Sci. Rep.* **8**, 7572 (2018).
61. PAGES 2k Consortium Consistent multi-decadal variability in global temperature reconstructions and simulations over the Common Era. *Nature Geosci.* (in the press).
62. Christiansen, B. & Ljungqvist, F. C. Challenges and perspectives for large-scale temperature reconstructions of the past two millennia. *Rev. Geophys.* **55**, 40–96 (2017).
63. Ammann, C. M. & Wahl, E. R. The importance of the geophysical context in statistical evaluations of climate reconstruction procedures. *Clim. Change* **85**, 71–88 (2007).
64. Gergis, J., Neukom, R., Gallant, A. J. E. & Karoly, D. J. Australasian temperature reconstructions spanning the last millennium. *J. Clim.* **29**, 5365–5392 (2016).
65. Wahl, E. R., Ritson, D. M. & Ammann, C. M. Comment on "Reconstructing past climate from noisy data". *Science* **312**, 529 (2006).
66. von Storch, H. Reconstructing past climate from noisy data. *Science* **306**, 679–682 (2004).
67. Bürger, G. & Cubasch, U. Are multiproxy climate reconstructions robust? *Geophys. Res. Lett.* **32**, L23711 (2005).
68. Taylor, K. E., Stouffer, R. J. & Meehl, G. A. An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.* **93**, 485–498 (2012).
69. Xiao-Ge, X., Tong-Wen, W. & Jie, Z. Introduction of CMIP5 experiments carried out with the climate system models of Beijing Climate Center. *Adv. Clim. Chang. Res.* **4**, 41–49 (2013).
70. Landrum, L. et al. Last millennium climate and its variability in CCSM4. *J. Clim.* **26**, 1085–1111 (2013).
71. Phipps, S. J. et al. The CSIRO Mk3L climate system model version 1.0 – Part 2: response to external forcings. *Geosci. Model Dev.* **5**, 649–682 (2012).
72. Schmidt, G. A. et al. Present-day atmospheric simulations using GISS ModelE: comparison to in situ, satellite, and reanalysis data. *J. Clim.* **19**, 153–192 (2006).
73. Schurer, A. P., Hegerl, G. C., Mann, M. E., Tett, S. F. B. & Phipps, S. J. Separating forced from chaotic climate variability over the past millennium. *J. Clim.* **26**, 6954–6973 (2013).
74. Dufresne, J.-L. et al. Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Clim. Dyn.* **40**, 2123–2165 (2013).
75. Jungclaus, J. H. et al. Characteristics of the ocean simulations in the Max Planck Institute Ocean Model (MPIOM) the ocean component of the MPI-Earth system model. *J. Adv. Model. Earth Syst.* **5**, 422–446 (2013).
76. Cowtan, K. & Way, R. G. Coverage bias in the HadCRUT4 temperature series and its impact on recent temperature trends. *Q. J. R. Meteorol. Soc.* **140**, 1935–1944 (2014).
77. Mann, M. E., Rutherford, S., Wahl, E. & Ammann, C. Robustness of proxy-based climate field reconstruction methods. *J. Geophys. Res.* **112**, D12109 (2007).
78. Gallant, A. J. E. & Gergis, J. An experimental streamflow reconstruction for the River Murray, Australia, 1783–1988. *Wat. Resour. Res.* **47**, W00G04 (2011).
79. Gergis, J. et al. On the long-term context of the 1997–2009 'Big Dry' in South-Eastern Australia: insights from a 206-year multi-proxy rainfall reconstruction. *Clim. Change* **111**, 923–944 (2012).
80. Frank, D. C. et al. Ensemble reconstruction constraints on the global carbon cycle sensitivity to climate. *Nature* **463**, 527–530 (2010).