

Resolving the differences in the simulated and reconstructed climate response to volcanism over the last millennium

Feng Zhu^{1,1}, Julien Emile-Geay^{1,1}, Gregory J. Hakim^{2,2}, Jonathan King^{3,3}, and Kevin John Anchukaitis^{3,3}

¹University of Southern California

²University of Washington

³University of Arizona

November 30, 2022

Abstract

Explosive volcanism imposes impulse-like radiative forcing on the climate system, providing a natural experiment to study the climate response to perturbation. Previous studies have identified disagreements between paleoclimate reconstructions and climate model simulations (GCMs) with respect to the magnitude and recovery from volcanic cooling, questioning the fidelity of GCMs, reconstructions, or both. Using the paleoenvironmental data assimilation framework of the Last Millennium Reanalysis, this study investigates the causes of the disagreements, using both real and simulated data. We demonstrate that discrepancies since 1600 CE can be largely resolved by assimilating tree-ring density records only, targeting growing-season temperature instead of annual temperature, and performing the comparison at proxy locales. Simulations of eruptions earlier in the last millennium may also reflect uncertainties in forcing and modeled aerosol microphysics.

Resolving the differences in the simulated and reconstructed temperature response to volcanism

Feng Zhu¹, Julien Emile-Geay¹, Gregory J. Hakim², Jonathan King^{3,4}, Kevin J. Anchukaitis^{3,4,5}

¹Department of Earth Sciences, University of Southern California, Los Angeles, CA USA

²Department of Atmospheric Sciences, University of Washington, Seattle, WA USA

³Department of Geosciences, University of Arizona, Tucson AZ USA

⁴Laboratory of Tree-Ring Research, University of Arizona, Tucson AZ USA

⁵School of Geography and Development, University of Arizona, Tucson AZ USA

Key Points:

- We explore model–proxy disagreement on the temperature response to volcanic eruptions over the past millennium.
- Using paleoclimate data assimilation with both real and synthetic data, we show that this discrepancy is due to four main factors.
- Over the past 400 years, agreement is found for tree-ring density records at the places and season these proxies record.

Plain Language Summary

The response to volcanic eruptions is a critical benchmark of the performance of climate models. Previous studies of the past millennium have identified discrepancies between model simulations and climate reconstructions regarding the temperature response to volcanic eruptions, raising concerns regarding the source of this mismatch and implications for both models and reconstructions. By evaluating the leading sources of differences between simulations and reconstructions, this study shows that accounting for known factors largely bridges the gap.

Corresponding author: Feng Zhu, fengzhu@usc.edu

Abstract

Explosive volcanism imposes impulse-like radiative forcing on the climate system, providing a natural experiment to study the climate response to perturbation. Previous studies have identified disagreements between paleoclimate reconstructions and climate model simulations (GCMs) with respect to the magnitude and recovery from volcanic cooling, questioning the fidelity of GCMs, reconstructions, or both. Using the paleoenvironmental data assimilation framework of the Last Millennium Reanalysis, this study investigates the causes of the disagreements, using both real and simulated data. We demonstrate that discrepancies since 1600 CE can be largely resolved by assimilating tree-ring density records only, targeting growing-season temperature instead of annual temperature, and performing the comparison at proxy locales. Simulations of eruptions earlier in the last millennium may also reflect uncertainties in forcing and modeled aerosol microphysics.

1 Introduction

Volcanic eruptions influence the climate system through their direct effect on short-wave radiation entering the earth system and the subsequent response of major modes of ocean-atmosphere variability (Handler, 1984; Hirono, 1988; Robock, 2000; Adams et al., 2003; Mann et al., 2005; Emile-Geay et al., 2008; Schneider et al., 2009; Li et al., 2013; Stevenson et al., 2016). Eruptions therefore offer unique natural experiments with which to probe the fidelity of climate model simulations, understand the response of the ocean and atmosphere circulation to changes in radiative forcing, assess climate system feedbacks, and evaluate solar radiation management proposals (Soden et al., 2002; Timmreck, 2012). The sporadic occurrence of large volcanic eruptions means that developing a deeper understanding of their effect on climate necessarily involves analyzing the response to events prior to the instrumental era.

Significant disagreements have been identified between paleoclimate reconstructions of the large-scale temperature response to volcanic eruptions and climate model simulations (D’Arrigo et al., 2013; Schurer et al., 2013; Stoffel et al., 2015). The IPCC AR5 (Masson-Delmotte et al., 2013), which summarized the state of knowledge at the time of publication, highlighted a discrepancy in the intensity and duration of the simulated versus proxy-based reconstructed temperature response to explosive volcanism (Fig. 1b). Coupled Model Intercomparison Project 5 (CMIP5)/Paleoclimate Modeling Intercomparison Project 3 (PMIP3) model simulations for the last millennium experiment (Schmidt et al., 2012a) show more cooling, and for a shorter duration, than paleoclimate reconstructions (Briffa et al., 2001; D’Arrigo et al., 2006; Frank, Esper, & Cook, 2007; Moberg et al., 2005). Compounding this uncertainty, the precise timing and location of some volcanic eruptions over the last millennium remain unknown (Sigl et al., 2015; Stevenson et al., 2017) as does the magnitude of the radiative forcing (Timmreck et al., 2009). A critical question is whether this mismatch is an artifact of uncertainties in (1) the paleoclimate proxy observations, (2) the reconstruction process, (3) the forcing estimates, (4) climate model physics, or (5) a combination thereof (Anchukaitis et al., 2012; Timmreck, 2012; D’Arrigo et al., 2013; LeGrande & Anchukaitis, 2015; Stoffel et al., 2015; Stevenson et al., 2016).

Here we explore four major sources of uncertainty in reconstructions of surface air temperature over the past millennium: spatial coverage, seasonality, biological memory, and proxy noise. We do so in the context of a paleoenvironmental data assimilation (PDA) framework, the Last Millennium Reanalysis (LMR) (Hakim et al., 2016; Tardif et al., 2019), which provides an objective basis for combining information from proxies and models. We show here that the discrepancy in Fig. 1b is present in our reconstruction (Fig. 1c), but that it can be largely reconciled by accounting for the aforementioned sources of uncertainty.

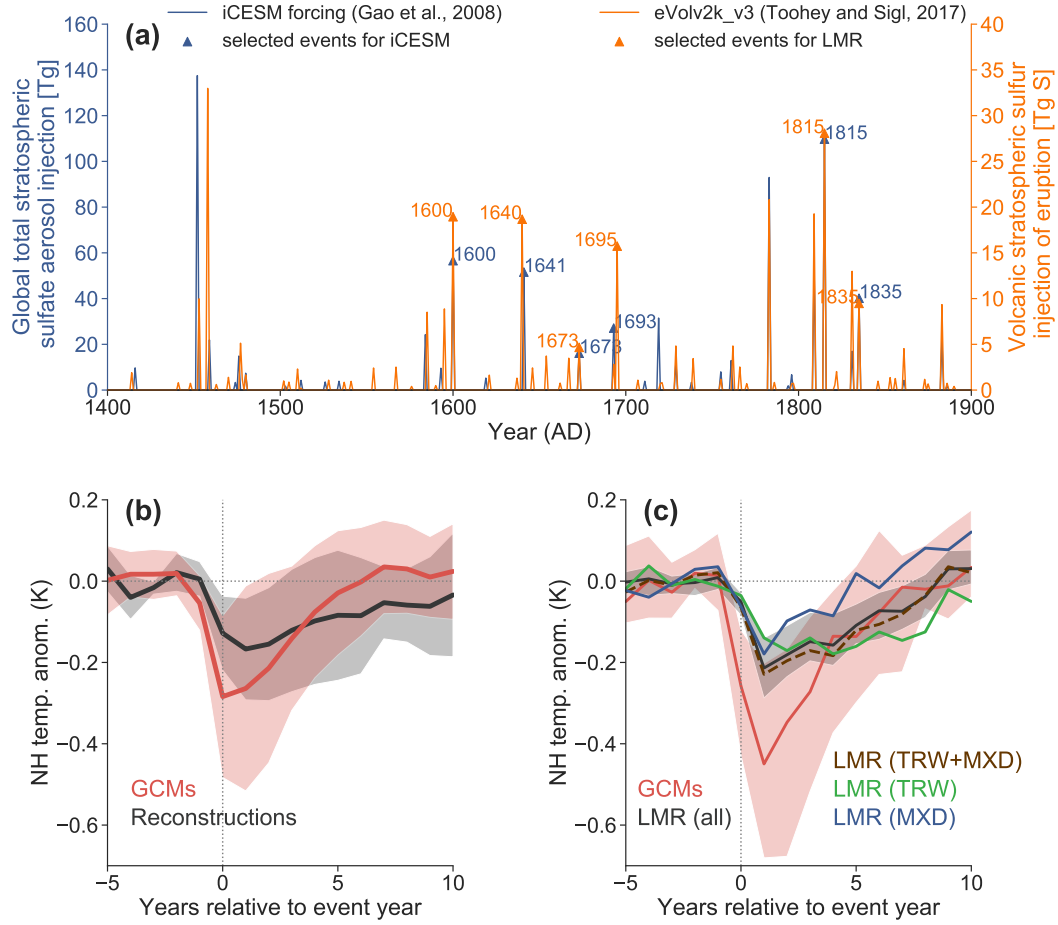


Figure 1. (a) Comparison between the volcanic forcing (Gao et al., 2008) used in the isotope-enabled Community Earth System Model (iCESM) simulation (Stevenson et al., 2019; Brady et al., 2019) and the eVolv2k version 3 Volcanic Stratospheric Sulfur Injection (VSSI) compilation (Toohey & Sigl, 2017). The triangles denote the selected 6 large events between 1400 and 1850 CE. (b) Superposed epoch analysis (SEA) on simulated and reconstructed temperature response to the 12 strongest volcanic eruptions since 1400 AD, reproduced from IPCC AR5 (Masson-Delmotte et al., 2013) Fig. 5.8b. (c) Superposed epoch analysis on annual Northern hemispheric mean temperature (NHMT) simulated by 9 GCMs (Section 2.2, Table S1) and LMR reconstructions assimilating the whole network (solid black curve with shading), the tree-ring network (dashed brown curve), the tree-ring width (TRW) network (solid green curve), and the maximum latewood density (MXD) network (solid blue curve), respectively. The shading encompasses the 5% and 95% quantiles of the ensemble, while the curves indicate the ensemble median (see Text S1 for details about ensemble scheme).

2 Data and methods

2.1 Paleoclimate data assimilation

We apply the paleoenvironmental data assimilation framework of the Last Millennium Reanalysis (LMR) (Hakim et al., 2016; Tardif et al., 2019) to both pseudoproxy and real proxy data networks. LMR uses an offline ensemble data assimilation procedure for multivariate climate field reconstruction (Steiger et al., 2014), where information from a prior expectation of the climate, derived from a climate model, is optimally combined with information from proxy records. The relative weights are determined from the error ratio in these two estimates of the climate, as defined by the update equation in the Kalman filter, which is optimal if the errors are unbiased and normally distributed.

The essential components of the procedure are (1) existing climate model data for the prior expectation, which we take from a last millennium simulation from the isotope-enabled Community Earth System Model (iCESM) (Stevenson et al., 2019; Brady et al., 2019); (2) proxy data networks, which we take from the PAGES 2k phase 2 compilation (PAGES 2k Consortium, 2017, Fig. S1) and the Northern Hemisphere Tree-Ring Network Development (NTREND) compilation (Wilson, Anchukaitis, Briffa, Büntgen, et al., 2016; Anchukaitis et al., 2017, Fig. S7); and (3) a “forward operator” or proxy system model (PSM), which predicts the proxies given the climate state. Here the forward operator is a linear regression procedure, univariate on annual temperature for corals and ice cores, univariate on seasonal temperature for maximum latewood density records, and bivariate on seasonal temperature and seasonal precipitation for tree-ring width records, as in Tardif et al. (2019). Further details of the LMR data assimilation procedure for paleoclimate reconstruction may be found in Hakim et al. (2016).

This study utilizes a fast implementation of the LMR framework, LMRt (Zhu et al., 2019) for computational convenience. As a benchmark, a reconstruction of the spatiotemporal variations of surface temperature over the Common Era is conducted, using iCESM as the model prior and the PAGES 2k network as observations. As expected, the DA procedure yields a substantially better estimate of the temporal variability in the temperature field than the prior, as quantified by the pointwise correlation with an independent instrumental temperature field (see Section 2.2 below) (Fig. S2c, S2d). This reconstruction skill level is comparable to a previous implementation of LMR (Tardif et al., 2019), and supported by the similarity between the reconstructed NHMT using both versions of the code (Fig. S2a). For a more in-depth evaluation of the LMR framework, see Tardif et al. (2019).

To assess the impact of the choice of prior and enable comparison with the LMR version of record (Tardif et al., 2019), we also tested assimilation using the CCSM4 simulation of Landrum et al. (2012) as the model prior. We find virtually identical results, with no significant difference detected in the temperature response to volcanic eruptions after 1400 AD (compare Fig. S2a to Fig. S2b, Fig. 1c to Fig. S14a, and Fig. 4a to Fig. S14b).

2.2 Simulated and instrumental temperature observations

In order to compare paleoclimate reconstructions to climate models, we consider simulations of past millennium climate from the following models: iCESM and CESM1 (Otto-Bliesner et al., 2015), as well as the PMIP3 models (Schmidt et al., 2012b; Brannan et al., 2012), including BCC CSM1.1 (Wu et al., 2014), GISS-E2-R (Schmidt et al., 2006), HadCM3 (Gordon et al., 2000), IPSL-CM5A-LR (Dufresne et al., 2013), MIROC-ESM (Watanabe et al., 2011), MPI-ESM-P (Giorgetta et al., 2013), CSIRO (Rotstayn et al., 2012), and CCSM4, as listed in Table S1.

We also use two sets of instrumental temperature observations, including the Berkeley Earth instrumental temperature analysis (Rohde et al., 2013) and the Goddard Institute for Space Studies (GISS) Surface Temperature Analysis (GISTEMP) (Hansen et al., 2010). GISTEMP and the gridded precipitation dataset (V6) from the Global Precipitation Climatology Centre (GPCC) (Schneider et al., 2014) are also used for PSM calibration in the bivariate framework of Tardif et al. (2019).

2.3 Superposed epoch analysis (SEA)

Superposed epoch analysis (SEA) (Haurwitz & Brier, 1981) is a frequently used technique to assess the temperature response to volcanic eruptions (Adams et al., 2003; Masson-Delmotte et al., 2013; Rao et al., 2019). It consists of aligning temperature anomaly series to the timing of volcanic eruptions within a fixed time window prior to and following the event, and averaging these responses to estimate the typical response to eruptions. The IPCC AR5 (Fig. 1b) considered the reconstructed temperature response to the 12 strongest eruptions since 1400 AD. Based on the temporal coverage of available proxies and model simulations, as well as the scientific knowledge of the eruptions, we selected a smaller set of 6 large and well-dated eruption events over the years 1600-1850 CE that are consistent in timing in both the volcanic forcing used in iCESM (Gao et al., 2008) and the most recent compilation of Volcanic Stratospheric Sulfur Injection (VSSI) (Toohey & Sigl, 2017) (Fig. 1a). For further details about the selection, see Text S3. The LMR response to individual events of the entire millennium is shown in Fig. S10-S12.

3 The discrepancy and its causes

Fig. 1b highlights discrepancies between model simulations and reconstructions in three aspects: (1) the magnitude of the peak cooling (2) the timing of the peak cooling (3) the length of the recovery. Specifically, model simulations show a stronger peak cooling amplitude, a slightly earlier peak cooling, and a shorter recovery interval than the reconstructions. A similar discrepancy pattern can be seen when comparing the LMR reconstruction assimilating the PAGES 2k network to the model simulations (Fig. 1c). Comparing results for solutions assimilating the entire PAGES 2k network [LMR (all), solid dark gray curve] to those assimilating only its tree-ring records [LMR (TRW+MXD), dashed brown curve], we see that most of the reconstructed volcanic cooling originates from the information captured by the tree-ring network. The latter consists of two main observation types: (1) tree-ring width (TRW) and (2) maximum latewood density (MXD). Assimilating these two proxy types separately, however, shows different responses to volcanism: TRW yields a lagged peak cooling year and a more prolonged recovery than MXD. This suggests that the difference between these two proxy types is key to understanding the different volcanic cooling patterns in reconstructions.

Previous studies (Timmreck et al., 2009; Timmreck, 2012; LeGrande & Anchukaitis, 2015; Stoffel et al., 2015; LeGrande et al., 2016) have investigated the components of the PMIP3 models that could potentially result in overestimated cooling in simulations. Here, with a focus on proxies and reconstructions, we investigate four factors that we hypothesize may account for these differences, motivated by prior studies and existing knowledge of the tree-ring proxy network: (1) spatial coverage (Anchukaitis et al., 2012; D’Arrigo et al., 2013) (2) seasonality (D’Arrigo et al., 2006; Stoffel et al., 2015; Anchukaitis et al., 2017) (3) biological memory (Fritts, 1966; Krakauer & Randerson, 2003; Frank, Büntgen, et al., 2007; Esper et al., 2015; Stoffel et al., 2015; Zhang et al., 2015; Lücke et al., 2019), and (4) non-temperature ‘noise’ (von Storch et al., 2004; Riedwyl et al., 2009; Neukom et al., 2018).

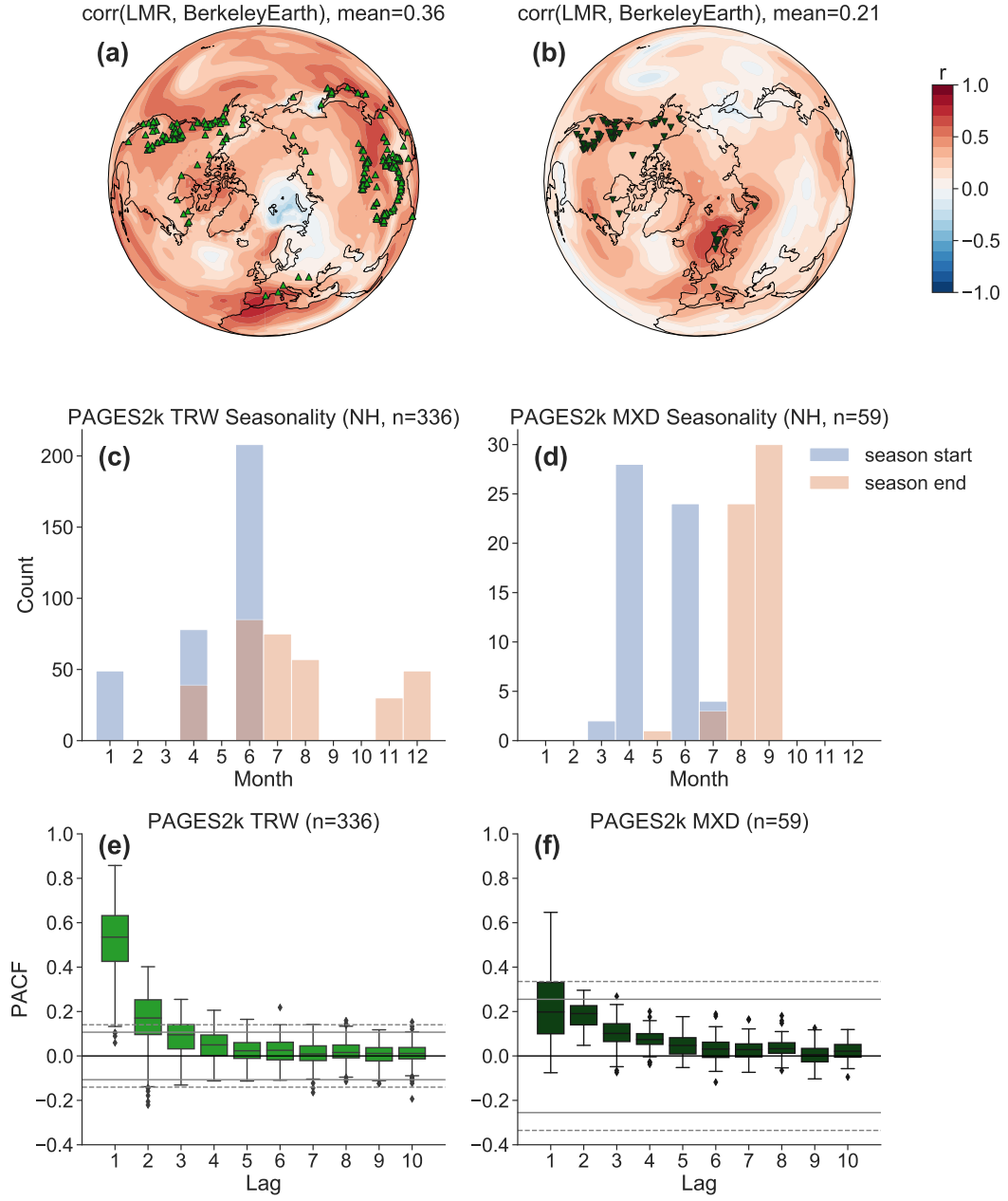


Figure 2. Differences between PAGES 2k TRW and MXD records regarding (a, b) spatial coverage, (c, d) seasonality detected by the algorithm used in Tardif et al. (2019), and (e, f) biological memory quantified by the partial autocorrelation function (PACF). (a) The spatial coverage of TRW network. (c) The optimal seasonality of the TRW network. (e) The PACF of the TRW network. (b), (d), and (f) are as (a), (b), and (e), respectively, but for the MXD network. The color contours in (a, b) indicate the correlation between the LMR reconstructions and the Berkeley Earth instrumental temperature analysis (Rohde et al., 2013).

3.1 Spatial coverage

The PAGES 2k network is comprised of 336 TRW records and 59 MXD records over the Northern Hemisphere (NH). MXD records in PAGES2k are mainly limited to North America and Scandinavia, while the TRW records cover both North America and Asia. Evaluating the correlation between the LMR reconstruction and the Berkeley Earth instrumental temperature analysis (Rohde et al., 2013) over the instrumental era over 1880–2000, we see that assimilating the TRW network yields a greater improvement over the model prior than assimilating the MXD network (Fig 2a, 2b). Is this difference due to the location or the quantity of each type of proxy record? To investigate this question, we use a pseudoproxy experiment (PPE) (Smerdon, 2011). We set the annual iCESM simulated temperature as our truth, and use it as model prior in the DA framework (a “perfect model” scenario). Pseudoproxies are defined as perfect temperature recorders at three sets of locations: (1) the locales of all the 336 NH PAGES 2k TRW records (2) the locales of randomly picked 50 PAGES 2k TRW records over North America and (3) the locales of randomly picked 50 PAGES 2k TRW records over the NH.

The result of assimilating these three pseudoproxy networks is shown in Fig. S3 (a, b, and c), showing that better spatial coverage yields a more accurate reconstruction in the PDA framework, with all other things being equal. This is reflected in SEA as well: Fig. 3a shows that assimilating 50 records spread throughout the NH yields a stronger and more accurate peak cooling amplitude than assimilating 50 records concentrated over North America, suggesting that broad spatial coverage is more important than the sheer number of records for resolving peak cooling amplitude. Location does matter to some degree with regard to the large-scale teleconnection patterns, and optimal placement could be determined with the approach of Comboul et al. (2015), yet this is beyond the scope of this investigation.

3.2 Seasonality

An implicit assumption in reconstructing annual temperature with tree-ring proxies is that growing season temperature is representative of annual temperature (PAGES 2k Consortium, 2017). However, the correlation between summer and annual temperatures in the Northern Hemisphere is high for the oceans but relatively low over continents (Fig. S3f), where the tree-ring records are located. Trees register climate primarily during their growing season, which varies as a function of geography, species, and climate (Fritts, 1966; St. George, 2014; St George & Ault, 2014; Wilson, Anchukaitis, Briffa, Buntgen, et al., 2016; Stoffel et al., 2015). Though the PAGES 2k metadata contain some information about the seasonal sensitivity of all proxies, we follow Tardif et al. (2019) and identify optimal seasonal windows of temperature and precipitation for each proxy record from a pool of pre-defined seasonal windows. The windows are optimal in a least square sense, using calibration over the historical period. The start and end month of the growing season (based on temperature) thus identified are shown in Fig. 2c, 2d. While in the Northern Hemisphere both TRW and MXD proxies record largely boreal summer conditions, the optimal seasonality for TRW is often broader but typically less consistent than that for MXD.

As before, we use a PPE to investigate the impact of growth seasonality on the temperature reconstruction. We generate pseudo-PAGES2k TRW records at their real locations as perfect recorders of local summer (JJA) temperature and perform experiments targeting both JJA temperature and annual temperature. As expected, a much better reconstruction is obtained for the boreal summer temperature field than annual temperature (Fig. S3d, S3e). This is also evident in reconstructions using real proxies and instrumental temperature (Fig. S4). Therefore, summer-sensitive trees can only reconstruct annual temperature to the extent that the summer and annual mean are correlated. While

such seasonal effects result in quite different assessments of reconstruction fidelity, this difference is hardly noticeable in SEA (Fig. 3b).

3.3 Biological memory

Another important difference between TRW and MXD is biological memory, whereby tree growth reflects the influence of climate in previous years (Fritts, 1966; Krakauer & Randerson, 2003; Frank, Büntgen, et al., 2007; Esper et al., 2015; Zhang et al., 2015; Stofel et al., 2015). We quantify the persistence in TRW and MXD in the PAGES2k using the partial autocorrelation function (PACF) (Fig. 2e, 2f). As expected (Breitenmoser et al., 2012; Esper et al., 2015; Lücke et al., 2019), we find that biological memory in TRW across the PAGES2k network is large and significant for lag-1 and lag-2, while for MXD it is limited. Comparing the proxy composites and the corresponding average instrumental temperature at proxy locales, we see that the MXD composite captures contemporaneous temperature variations, including the accurate timing of cooling events, while the TRW composite appears to smooth interannual variability and integrate temperatures over 2 to 5 years (Fig. S5a, S5b), leading to lagged and persistent cooling events (Frank, Büntgen, et al., 2007).

To investigate the impact of such biological memory on the magnitude of reconstructed volcanic cooling, we again turn to PPEs. We simulate a short-term memory effect in TRW by designing pseudoproxies as a 5-yr moving average of the annual temperature simulated by iCESM, as shown in Fig. S5c. Assimilating these smoothed pseudoproxies yields a prolonged temperature recovery and a peak cooling that is both damped and lagged (Fig. 3c, the solid light green curve). We find that this overall result is not sensitive to the precise design of the filter used to construct the smoothed pseudoproxies, so long as it captures this multiple year climate integration in some way. The potential additional influence of soil moisture is not directly modeled here, as these lagged relationships are observed in temperature-sensitive TRW chronologies irrespective of the potential additional influence of soil moisture (Franke et al., 2013; Consortium, 2017), which we confirm in sensitivity experiments (Fig. S15).

3.4 Proxy system noise

So far, our PPEs have employed noiseless temperature recorders for simplicity (a signal-to-noise ratio (SNR) of infinity, wherein SNR is defined as the ratio of the standard deviation of signal and that of noise, following existing practice (Smerdon, 2011)). In reality, of course, proxies are imperfect recorders of climate conditions. To make the PPEs more realistic, we now add uncorrelated Gaussian white noise to the previously described pseudo-PAGES2k TRW records. Using a linear regression procedure, we estimate a SNR around 0.3 (Fig. S6), which is comparable to the estimate of Wang et al. (2014). Since we have already emulated the biological memory utilizing the moving average filter, we consider white noise instead of red noise to avoid adding more memory into the pseudoproxies. The addition of noise to the previous case yields a more similar SEA pattern (Fig. 3c, solid dark green curve) to the real-world case (Fig. 1c, solid green curve): a more damped and prolonged recovery compared to the noiseless case.

Considering the four factors above, we are thus able to simulate the observed discrepancy between modeled and reconstructed NH temperature response to volcanic eruptions. Can this knowledge be used to minimize this discrepancy?

3.5 Reconciling the discrepancy

In the present context, noise reflects any non-temperature influence on the proxy systems, including other climate influences like soil moisture, or biophysical processes that cannot be adequately modeled due to insufficient scientific knowledge, limited

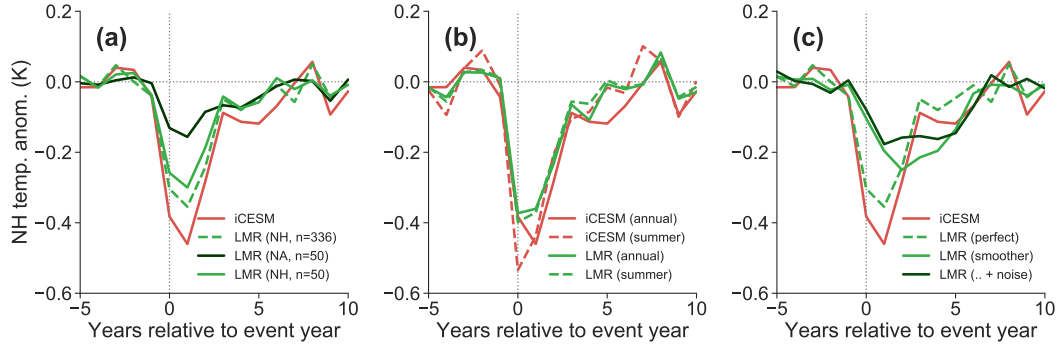


Figure 3. SEA in pseudoproxy experiments, evaluating the impact of (a) spatial coverage, (b) seasonality, and (c) biological memory and noise. (a) the red curve denotes the target, and the dashed light green curve, the solid dark green curve, and the solid light green curve indicate the LMR reconstruction assimilating 336 pseudo-PAGES2k TRW records over the NH, 50 records over North America, and 50 records over the NH, respectively. (b) The solid red curve denotes the annual target, the dashed red curve denotes the boreal summer target, and the green curves indicate the LMR annual and summer reconstructions assimilating the pseudo-PAGES2k TRW network, respectively. (c) The solid red curve denotes the annual target, and the green curves denote the LMR reconstruction assimilating pseudo-PAGES2k TRW as perfect temperature recorders (dashed), and temperature smoothers (solid). The case of smoothed temperature with added Gaussian noise (SNR=0.3) is in dark green. All the reconstruction curves refer to the ensemble median (see Text S1 for details about the ensemble design).

input data, or both. The first three factors can, however, be corrected: to account for the limited spatial coverage, we perform SEA at proxy locations instead of the whole NH; to minimize the seasonal bias, we target boreal summer temperature instead of annual temperature; and to mitigate memory effects, we assimilate MXD records only, leaving out TRW and mixed chronologies. As a result, we are able to almost entirely account for the proxy-model discrepancy in Fig. 1 with the PAGES 2k network (Fig. 4a, Fig S11). The same strategy can be used for other proxy networks. For comparison, applying it to the NTREND network (Wilson, Anchukaitis, Briffa, Büntgen, et al., 2016; Anchukaitis et al., 2017) (Fig. S7) yields similar agreement between simulated and reconstructed temperature (Fig. 4b, Fig S12). These results stand in sharp contrast to results where spatial coverage, seasonality, and biological memory are not taken into account (Fig. S8).

That the discrepancy in Fig. 1b can be largely reconciled by accounting for known characteristics of the proxy data is reassuring, and bodes well for using volcanic eruptions of the past millennium as a test bed for climate models. We now discuss the broader implications of this result.

4 Discussion

Using recent proxy compilations and climate field reconstruction techniques, we have demonstrated that it is possible to largely resolve the discrepancy between the simulated and reconstructed temperature response to explosive volcanism since 1600 CE. We find that this gap was the result of four main factors: spatial coverage, proxy seasonality, biological memory, and proxy noise. While proxy noise is difficult to account for in model-data intercomparisons, the first three factors can be, if care is taken in evaluating comparable quantities. In particular, since our reconstructions are more reliable at locations

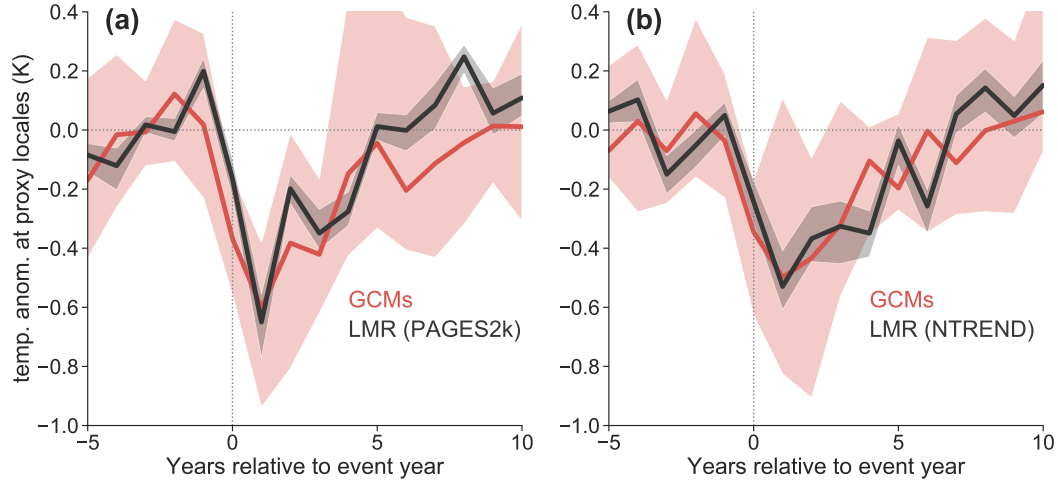


Figure 4. (a) Same as Fig. 1c, after resolving differences in the model and proxy domains associated with seasonality, spatial distribution, and biological memory. (b) Same as (a) but using the NTREND MXD network. A version of this figure showing each model simulation is available in Fig. S9, and one using more eruption events is available in Fig. S13

where proxies are available than at distal locations (Anchukaitis et al., 2017), carrying out the comparison at proxy sites is a simple and effective way to reduce the mismatch. That this is also true in the data assimilation framework (Steiger et al., 2014) suggests that expanding the spatial extent of proxy network is necessary to resolve global-scale patterns. For very large eruptions such as the 1257 Samalas eruption, the 1450s eruptions, and the 1815 Tambora eruption, however, significant mismatches remain between model simulations and reconstructions even when these factors are accounted for (Fig. S11, S12). While this has little impact in a composite over all events (Fig. S13), it warrants discussion.

Previous work and our own analysis suggests three major causes: (1) proxy attrition, (2) aerosol microphysics in models, and (3) uncertainties in volcanic forcing.

(1) In the absence of reliable proxy data, offline data assimilation reverts to the model prior for its estimate of climate. This results in generally damped variations in periods of reduced and/or noisy coverage, as seen by comparing the first to second millennium CE in Fig S2a, S2b. Here we have mitigated this problem by focusing on the recent period with relatively high proxy coverage, but it is undoubtedly an ingredient in the mismatch observed for earlier eruptions like Samalas, when relative few proxies are available, especially MXD records (Fig. S1b and S16a).

(2) Some CMIP5-era GCMs produce overly strong responses to volcanic forcing due to unrealistic representation of aerosol microphysics (Timmreck et al., 2009; Timmreck, 2012; Stoffel et al., 2015; LeGrande et al., 2016). Both Timmreck et al. (2009) and Stoffel et al. (2015) suggest that the discrepancy is caused by the simplistic assumption used in PMIP3 models that aerosol optical depth is linearly scaled to ice-core sulfate concentration. This assumption uses the 1991 Pinatubo eruption as the reference, and is unlikely to be valid for many significantly larger eruptions. As shown by Stoffel et al. (2015), accounting explicitly for self-limiting aerosol microphysical processes can reconcile this discrepancy, an idea later confirmed by Guillet et al. (2017) with both documentary and tree-ring data.

(3) Many PMIP3 models used the Gao et al. (2008) forcing dataset, where the reconstructed Samalas aerosol loading was exceedingly large, and has since been revised downward (Jungclauss et al., 2017). There is also lingering uncertainty as to the magnitude, timing, and location of two major events during the 1450s (Sigl et al., 2015; Toohey & Sigl, 2017; Hartman et al., 2019). Besides, apparent cooling from a 1761 eruption in some model simulations is actually the result of the misalignment of the 1783 forcing in the uncorrected version of the Gao et al. (2008) forcing (Stevenson et al., 2017). Regardless of changes in physics, the revision in volcanic forcing alone would help to reduce the discrepancy.

Progress in representing volcanic forcing (Toohey & Sigl, 2017; Aubry et al., 2019), as well as improvements in model resolution and processes (e.g. active stratospheric chemistry) in PMIP4 (Kageyama et al., 2018) may lead to closer model-data matches in future work. Regardless of these factors, our analysis suggests that a critical ingredient of minimizing the model-reconstruction mismatch is to evaluate simulated temperature at the times and places where it is recorded by the proxy sensors (Anchukaitis et al., 2012). Naturally, past temperature estimates may be improved as well. While this study has focused on the uncertainties in proxy measurements in the context of paleoenvironmental data assimilation, more work should be done to reduce sources of uncertainty within the data assimilation method itself, such as the forward operator error, the model prior, and the localization scheme, as the coupling of all these uncertainty sources can potentially affect the SEA comparison. In particular, forward operators that allow for non-contemporaneous influences of the state on the proxies [e.g. time-integration, as is believed to be the case for TRW (Fritts et al., 1991; Vaganov et al., 2006)] would enable us to make better use of the information contained in TRW records. While such process-oriented models have been developed (Tolwinski-Ward et al., 2011; Evans et al., 2013), their application to the DA context is contingent upon accurate specification of observation error variance and correcting for biases in the model prior. Both tasks remain active research areas (Dee et al., 2016).

With regard to proxies, we have confirmed that the lagged cooling exhibited in previous reconstructions can be explained as the consequence of their using TRW records. Other proxies that integrate climate information over multiple years (e.g. bioturbated sediments) likely have a similar impact in multiproxy reconstructions. Since MXD records are more faithful paleo-temperature sensors than TRW records (Esper et al., 2015, 2018), we join others in calling for increased collection and development of MXD records (Anchukaitis et al., 2017; St. George & Esper, 2019), particularly at locations where they are presently absent or cover only part of the last millennium, e.g. the North American treeline and at high elevations in Asia (Anchukaitis et al., 2017; Esper et al., 2018).

While our approach reconciles the discrepancy between model and proxy estimates of the surface temperature to moderate eruptions of the last 400 years, important differences remain for large events like Tambora or Samalas. For such eruptions, improved estimates of the forcing, a more realistic model representation of aerosol microphysics, and – for events sparsely sampled by existing proxy network – an expanded proxy coverage may be necessary to resolve extant differences. Future work will help elucidate the relative role of these three factors on this particular comparison.

Acknowledgments

The authors acknowledge support from the Climate Program Office of the National Oceanographic and Atmospheric Administration (grants NA18OAR4310426 to USC, NA18OAR4310422 to UW, and NA18OAR4310420 to UA) GJH also acknowledges support from the NSF through grant AGS-1702423. Code and data are available at: <https://github.com/fzhu2e/lmrvolc> (a placeholder for now, to be filled upon paper acceptance).

References

- Adams, J., Mann, M. E., & Ammann, C. M. (2003, November). Proxy evidence for an El Niño-like response to volcanic forcing. *Nature*, 426(6964), 274–278. doi: 10.1038/nature02101
- Anchukaitis, K. J., Breitenmoser, P., Briffa, K. R., Buchwal, A., Buntgen, U., Cook, E. R., ... Wilson, R. J. S. (2012, 12). Tree rings and volcanic cooling. *Nature Geosci*, 5(12), 836–837. doi: 10.1038/ngeo1645
- Anchukaitis, K. J., Wilson, R., Briffa, K. R., Buntgen, U., Cook, E. R., D'Arrigo, R., ... Zorita, E. (2017, 5 1). Last millennium Northern Hemisphere summer temperatures from tree rings: Part II, spatially resolved reconstructions. *Quaternary Science Reviews*, 163, 1–22. doi: 10.1016/j.quascirev.2017.02.020
- Aubry, T. J., Toohey, M., Marshall, L., Schmidt, A., & Jellinek, A. M. (2019). A new volcanic stratospheric sulfate aerosol forcing emulator (eva_h): Comparison with interactive stratospheric aerosol models. *Journal of Geophysical Research: Atmospheres*. doi: 10.1029/2019JD031303
- Braconnot, P., Harrison, S. P., Kageyama, M., Bartlein, P. J., Masson-Delmotte, V., Abe-Ouchi, A., ... Zhao, Y. (2012, June). Evaluation of climate models using palaeoclimatic data. *Nature Climate Change*, 2(6), 417–424. doi: 10.1038/nclimate1456
- Brady, E., Stevenson, S., Bailey, D., Liu, Z., Noone, D., Nusbaumer, J., ... Zhu, J. (2019). The Connected Isotopic Water Cycle in the Community Earth System Model Version 1. *Journal of Advances in Modeling Earth Systems*, 11(8), 2547–2566. doi: 10.1029/2019MS001663
- Breitenmoser, P., Beer, J., Brönnimann, S., Frank, D., Steinhilber, F., & Wanner, H. (2012, January). Solar and volcanic fingerprints in tree-ring chronologies over the past 2000 years. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 313–314, 127–139. doi: 10.1016/j.palaeo.2011.10.014
- Briffa, K. R., Osborn, T. J., Schweingruber, F. H., Harris, I. C., Jones, P. D., Shiyatov, S. G., & Vaganov, E. A. (2001). Low-frequency temperature variations from a northern tree ring density network. *Journal of Geophysical Research: Atmospheres*, 106(D3), 2929–2941. doi: 10.1029/2000JD900617
- Comboul, M., Emile-Geay, J., Hakim, G. J., & Evans, M. N. (2015). Paleoclimate sampling as a sensor placement problem. *Journal of Climate*, 28, 7717–7740. doi: 10.1175/JCLI-D-14-00802.1
- Consortium, P. H. (2017, December). Comparing proxy and model estimates of hydroclimate variability and change over the Common Era. *Climate of the Past*, 13(12), 1851–1900. doi: https://doi.org/10.5194/cp-13-1851-2017
- D'Arrigo, R., Wilson, R., & Anchukaitis, K. J. (2013). Volcanic cooling signal in tree ring temperature records for the past millennium. *Journal of Geophysical Research-Atmospheres*, 118(16), 9000–9010. doi: 10.1002/jgrd.50692
- D'Arrigo, R., Wilson, R., & Jacoby, G. (2006). On the long-term context for late twentieth century warming. *Journal of Geophysical Research: Atmospheres*, 111(D3). doi: 10.1029/2005JD006352
- Dee, S. G., Steiger, N. J., Emile-Geay, J., & Hakim, G. J. (2016). On the utility of proxy system models for estimating climate states over the common era. *Journal of Advances in Modeling Earth Systems*, 8. doi: 10.1002/2016MS000677
- Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., ... Vuichard, N. (2013, May). Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics*, 40(9–10), 2123–2165. doi: 10.1007/s00382-012-1636-1
- Emile-Geay, J., Seager, R., Cane, M. A., Cook, E. R., & Haug, G. H. (2008, July). Volcanoes and ENSO over the Past Millennium. *Journal of Climate*, 21(13), 3134–3148. doi: 10.1175/2007JCLI1884.1
- Esper, J., George, S. S., Anchukaitis, K., D'Arrigo, R., Ljungqvist, F. C., Luterbacher, J., ... Buntgen, U. (2018). Large-scale, millennial-length temper-

- ature reconstructions from tree-rings. *Dendrochronologia*, 50, 81–90. doi: 10.1016/j.dendro.2018.06.001
- Esper, J., Schneider, L., Smerdon, J. E., Schöne, B. R., & Büntgen, U. (2015, October). Signals and memory in tree-ring width and density data. *Dendrochronologia*, 35, 62–70. doi: 10.1016/j.dendro.2015.07.001
- Evans, M., Tolwinski-Ward, S., Thompson, D., & Anchukaitis, K. (2013). Applications of proxy system modeling in high resolution paleoclimatology. *Quaternary Science Reviews*, 76, 16–28. doi: 10.1016/j.quascirev.2013.05.024
- Frank, D., Büntgen, U., Böhm, R., Maugeri, M., & Esper, J. (2007, December). Warmer early instrumental measurements versus colder reconstructed temperatures: shooting at a moving target. *Quaternary Science Reviews*, 26(25), 3298–3310. doi: 10.1016/j.quascirev.2007.08.002
- Frank, D., Esper, J., & Cook, E. R. (2007). Adjustment for proxy number and coherence in a large-scale temperature reconstruction. *Geophysical Research Letters*, 34(16). doi: 10.1029/2007GL030571
- Franke, J., Frank, D., Raible, C. C., Esper, J., & Brönnimann, S. (2013, April). Spectral biases in tree-ring climate proxies. *Nature Climate Change*, 3(4), 360–364. doi: 10.1038/nclimate1816
- Fritts, H. C. (1966). Growth-rings of trees: their correlation with climate. *Science*, 154(3752), 973–979. doi: 10.1126/science.154.3752.973
- Fritts, H. C., Vaganov, E. A., Sviderskaya, I. V., & Shashkin, A. V. (1991). Climatic variation and tree-ring structure in conifers: Empirical and mechanistic models of tree-ring width, number of cells, cell size, cell-wall thickness and wood density. *Climate Research*, 1, 97–116. doi: 10.3354/cr001097
- Gao, C., Robock, A., & Ammann, C. (2008). Volcanic forcing of climate over the past 1500 years: An improved ice core-based index for climate models. *Journal of Geophysical Research: Atmospheres*, 113(D23). doi: 10.1029/2008JD010239
- Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., ... Stevens, B. (2013). Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems*, 5(3), 572–597. doi: 10.1002/jame.20038
- Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., ... Wood, R. A. (2000, February). The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, 16(2-3), 147–168. doi: 10.1007/s003820050010
- Guillet, S., Corona, C., Stoffel, M., Khodri, M., Lavigne, F., Ortega, P., ... Openheimer, C. (2017, February). Climate response to the Samalas volcanic eruption in 1257 revealed by proxy records. *Nature Geoscience*, 10(2), 123–128. doi: 10.1038/ngeo2875
- Hakim, G. J., Emile-Geay, J., Steig, E. J., Noone, D., Anderson, D. M., Tardif, R., ... Perkins, W. A. (2016). The last millennium climate reanalysis project: Framework and first results. *Journal of Geophysical Research: Atmospheres*, 121, 6745 – 6764. doi: 10.1002/2016JD024751
- Handler, P. (1984, November). Possible association of stratospheric aerosols and El Nino type events. *Geophysical Research Letters*, 11(11), 1121–1124. doi: 10.1029/GL011i011p01121
- Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). Global surface temperature change. *Rev. Geophys.*, 48, RG4004. doi: 10.1029/2010RG000345
- Hartman, L. H., Kurbatov, A. V., Winski, D. A., Cruz-Urbe, A. M., Davies, S. M., Dunbar, N. W., ... Yates, M. G. (2019). Volcanic glass properties from 1459 c.e. volcanic event in south pole ice core dismiss kuwae caldera as a potential source. *Scientific Reports*, 9(1), 14437. doi: 10.1038/s41598-019-50939-x

- Haurwitz, M. W., & Brier, G. W. (1981, October). A Critique of the Superposed Epoch Analysis Method: Its Application to Solar–Weather Relations. *Monthly Weather Review*, 109(10), 2074–2079. doi: 10.1175/1520-0493(1981)109<2074:ACOTSE>2.0.CO;2
- Hirono, M. (1988, May). On the trigger of El Niño Southern Oscillation by the forcing of early El Chichón volcanic aerosols. *Journal of Geophysical Research: Atmospheres*, 93(D5), 5365–5384. doi: 10.1029/JD093iD05p05365
- Jungclauss, J. H., Bard, E., Baroni, M., Braconnot, P., Cao, J., Chini, L. P., ... Zorita, E. (2017, November). The PMIP4 contribution to CMIP6 – Part 3: The last millennium, scientific objective, and experimental design for the PMIP4 *past1000* simulations. *Geoscientific Model Development*, 10(11), 4005–4033. (Publisher: Copernicus GmbH) doi: 10.5194/gmd-10-4005-2017
- Kageyama, M., Braconnot, P., Harrison, S. P., Haywood, A. M., Jungclauss, J. H., Otto-Bliesner, B. L., ... Zhou, T. (2018, March). The PMIP4 contribution to CMIP6 – Part 1: Overview and over-arching analysis plan. *Geoscientific Model Development*, 11(3), 1033–1057. doi: https://doi.org/10.5194/gmd-11-1033-2018
- Krakauer, N. Y., & Randerson, J. T. (2003). Do volcanic eruptions enhance or diminish net primary production? Evidence from tree rings. *Global Biogeochemical Cycles*, 17(4). doi: 10.1029/2003GB002076
- Landrum, L., Otto-Bliesner, B. L., Wahl, E. R., Conley, A., Lawrence, P. J., Rosenbloom, N., & Teng, H. (2012, 2014/05/05). Last millennium climate and its variability in cesm4. *Journal of Climate*, 26(4), 1085–1111. doi: 10.1175/JCLI-D-11-00326.1
- LeGrande, A. N., & Anchukaitis, K. J. (2015). Volcanic eruptions and climate. *PAGES Magazine*, 23(2), 46–47. doi: 10.22498/pages.23.2.46
- LeGrande, A. N., Tsigaridis, K., & Bauer, S. E. (2016, September). Role of atmospheric chemistry in the climate impacts of stratospheric volcanic injections. *Nature Geoscience*, 9(9), 652–655. doi: 10.1038/ngeo2771
- Li, J., Xie, S.-P., Cook, E. R., Morales, M. S., Christie, D. A., Johnson, N. C., ... Fang, K. (2013, September). El Niño modulations over the past seven centuries. *Nature Climate Change*, 3(9), 822–826. doi: 10.1038/nclimate1936
- Lücke, L., Hegerl, G., Schurer, A., & Wilson, R. (2019, September). Effects of memory biases on variability of temperature reconstructions. *Journal of Climate*. doi: 10.1175/JCLI-D-19-0184.1
- Mann, M. E., Cane, M. A., Zebiak, S. E., & Clement, A. (2005, February). Volcanic and Solar Forcing of the Tropical Pacific over the Past 1000 Years. *Journal of Climate*, 18(3), 447–456. doi: 10.1175/JCLI-3276.1
- Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., Rouco, J. G., ... Timmermann, A. (2013). Information from Paleoclimate Archives. In T. F. Stocker et al. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 383–464). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. doi: 10.1017/CBO9781107415324.013
- Moberg, A., Sonechkin, D. M., Holmgren, K., Datsenko, N. M., & Karlén, W. (2005, February). Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data. *Nature*, 433(7026), 613–617. doi: 10.1038/nature03265
- Neukom, R., Schurer, A. P., Steiger, N. J., & Hegerl, G. C. (2018, May). Possible causes of data model discrepancy in the temperature history of the last Millennium. *Scientific Reports*, 8(1), 1–15. doi: 10.1038/s41598-018-25862-2
- Otto-Bliesner, B. L., Brady, E. C., Fasullo, J., Jahn, A., Landrum, L., Stevenson, S., ... Strand, G. (2015). Climate variability and change since 850 CE: An ensemble approach with the community earth system model. *Bull. Amer. Meteor.*

- Soc.*, 97(5), 735–754. doi: 10.1175/BAMS-D-14-00233.1
- PAGES 2k Consortium. (2017, 07). A global multiproxy database for temperature reconstructions of the Common Era. *Scientific Data*, 4, 170088 EP. doi: 10.1038/sdata.2017.88
- Rao, M. P., Cook, E. R., Cook, B. I., Anchukaitis, K. J., D’Arrigo, R. D., Krusic, P. J., & LeGrande, A. N. (2019). A double bootstrap approach to Superposed Epoch Analysis to evaluate response uncertainty. *Dendrochronologia*, 55, 119–124. doi: 10.1016/j.dendro.2019.05.001
- Riedwyl, N., Küttel, M., Luterbacher, J., & Wanner, H. (2009). Comparison of climate field reconstruction techniques: Application to Europe. *Climate Dynamics*, 32(2-3), 381–395. doi: 10.1007/s00382-008-0395-5
- Robock, A. (2000). Volcanic eruptions and climate. *Rev. Geophys.*, 38, 191–220. doi: 10.1029/1998RG000054
- Rohde, R., Muller, R., Jacobsen, R., Perlmuter, S., Rosenfeld, A., Wurtele, J., ... Mosher, S. (2013). Berkeley Earth Temperature Averaging Process. *Geoinformatics & Geostatistics: An Overview, 2013*. doi: 10.4172/2327-4581.1000103
- Rotstayn, L. D., Jeffrey, S. J., Collier, M. A., Dravitzki, S. M., Hirst, A. C., Syktus, J. I., & Wong, K. K. (2012, July). Aerosol- and greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: a study using single-forcing climate simulations. *Atmos. Chem. Phys.*, 12(14), 6377–6404. doi: 10.5194/acp-12-6377-2012
- Schmidt, G. A., Jungclauss, J. H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., ... Vieira, L. E. A. (2012a). Climate forcing reconstructions for use in pmip simulations of the last millennium (v1.1). *Geoscientific Model Development*, 5(1), 185–191. doi: 10.5194/gmd-5-185-2012
- Schmidt, G. A., Jungclauss, J. H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., ... Vieira, L. E. A. (2012b, January). Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.1). *Geosci. Model Dev.*, 5(1), 185–191. doi: 10.5194/gmd-5-185-2012
- Schmidt, G. A., Ruedy, R., Hansen, J. E., Aleinov, I., Bell, N., Bauer, M., ... Yao, M.-S. (2006, January). Present-Day Atmospheric Simulations Using GISS ModelE: Comparison to In Situ, Satellite, and Reanalysis Data. *Journal of Climate*, 19(2), 153–192. doi: 10.1175/JCLI3612.1
- Schneider, D. P., Ammann, C. M., Otto-Bliesner, B. L., & Kaufman, D. S. (2009). Climate response to large, high-latitude and low-latitude volcanic eruptions in the Community Climate System Model. *Journal of Geophysical Research: Atmospheres*, 114(D15). doi: 10.1029/2008JD011222
- Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., & Rudolf, B. (2014, January). GPCC’s new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, 115(1), 15–40. doi: 10.1007/s00704-013-0860-x
- Schurer, A. P., Hegerl, G. C., Mann, M. E., Tett, S. F., & Phipps, S. J. (2013). Separating forced from chaotic climate variability over the past millennium. *Journal of Climate*, 26(18), 6954–6973. doi: 10.1175/JCLI-D-12-00826.1
- Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., ... Woodruff, T. E. (2015, 07 30). Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature*, 523(7562), 543–549. doi: 10.1038/nature14565
- Smerdon, J. E. (2011). Climate models as a test bed for climate reconstruction methods: pseudoproxy experiments. *WIREs Clim Change*. doi: 10.1002/wcc.149
- Soden, B. J., Wetherald, R. T., Stenchikov, G. L., & Robock, A. (2002). Global Cooling After the Eruption of Mount Pinatubo: A Test of Climate Feedback by Water Vapor. *Science*, 296(5568), 727–730. doi:

- 10.1126/science.296.5568.727
- 590 St. George, S. (2014). An overview of tree-ring width records across the North-
 591 ern Hemisphere. *Quaternary Science Reviews*, 95, 132–150. doi: 10.1016/j.
 592 .quascirev.2014.04.029
- 593 St George, S., & Ault, T. R. (2014). The imprint of climate within Northern Hemi-
 594 sphere trees. *Quaternary Science Reviews*, 89, 1–4. doi: 10.1016/j.quascirev
 595 .2014.01.007
- 596 Steiger, N. J., Hakim, G. J., Steig, E. J., Battisti, D. S., & Roe, G. H. (2014,
 597 2014/04/08). Assimilation of time-averaged pseudoproxies for climate re-
 598 construction. *Journal of Climate*, 27(1), 426–441. doi: 10.1175/JCLI-D-12
 599 -00693.1
- 600 Stevenson, S., Fasullo, J. T., Otto-Bliesner, B. L., Tomas, R. A., & Gao, C. (2017).
 601 Role of eruption season in reconciling model and proxy responses to tropical
 602 volcanism. *Proceedings of the National Academy of Sciences*, 114(8), 1822–
 603 1826. doi: 10.1073/pnas.1612505114
- 604 Stevenson, S., Otto-Bliesner, B., Fasullo, J., & Brady, E. (2016, February). “El Niño
 605 Like” Hydroclimate Responses to Last Millennium Volcanic Eruptions. *Journal*
 606 *of Climate*, 29(8), 2907–2921. doi: 10.1175/JCLI-D-15-0239.1
- 607 Stevenson, S., Otto-Bliesner, B. L., Brady, E. C., Nusbaumer, J., Tabor, C., Tomas,
 608 R., ... Liu, Z. (2019). Volcanic Eruption Signatures in the Isotope-Enabled
 609 Last Millennium Ensemble. *Paleoceanography and Paleoclimatology*, 0(0). doi:
 610 10.1029/2019PA003625
- 611 St. George, S., & Esper, J. (2019, January). Concord and discord among Northern
 612 Hemisphere paleotemperature reconstructions from tree rings. *Quaternary Sci-*
 613 *ence Reviews*, 203, 278–281. doi: 10.1016/j.quascirev.2018.11.013
- 614 Stoffel, M., Khodri, M., Corona, C., Guillet, S., Poulain, V., Bekki, S., ... Masson-
 615 Delmotte, V. (2015, October). Estimates of volcanic-induced cooling in the
 616 Northern Hemisphere over the past 1,500 years. *Nature Geoscience*, 8(10),
 617 784–788. doi: 10.1038/ngeo2526
- 618 Tardif, R., Hakim, G. J., Perkins, W. A., Horlick, K. A., Erb, M. P., Emile-Geay,
 619 J., ... Noone, D. (2019, July). Last Millennium Reanalysis with an expanded
 620 proxy database and seasonal proxy modeling. *Climate of the Past*, 15(4),
 621 1251–1273. doi: 10.5194/cp-15-1251-2019
- 622 Timmreck, C. (2012). Modeling the climatic effects of large explosive volcanic erup-
 623 tions. *Wiley Interdisciplinary Reviews: Climate Change*, 3(6), 545–564. doi: 10
 624 .1002/wcc.192
- 625 Timmreck, C., Lorenz, S. J., Crowley, T. J., Kinne, S., Raddatz, T. J., Thomas,
 626 M. A., & Jungclaus, J. H. (2009). Limited temperature response to the very
 627 large AD 1258 volcanic eruption. *Geophysical Research Letters*, 36(21). doi:
 628 10.1029/2009GL040083
- 629 Tolwinski-Ward, S. E., Evans, M. N., Hughes, M. K., & Anchukaitis, K. J. (2011).
 630 An efficient forward model of the climate controls on interannual varia-
 631 tion in tree-ring width. *Climate Dynamics*, 36(11-12), 2419–2439. doi:
 632 10.1007/s00382-010-0945-5
- 633 Toohey, M., & Sigl, M. (2017, November). Volcanic stratospheric sulfur injections
 634 and aerosol optical depth from 500 BCE to 1900 CE. *Earth System Science*
 635 *Data*, 9(2), 809–831. doi: 10.5194/essd-9-809-2017
- 636 Vaganov, E. A., Hughes, M. K., & Shashkin, A. V. (2006). *Growth dynamics of*
 637 *conifer tree rings* (Vol. 183). New York, NY: Springer-Verlag. doi: 10.1007/3
 638 -540-31298-6
- 639 von Storch, H., Zorita, E., Jones, J. M., Dimitriev, Y., González-Rouco, F., & Tett,
 640 S. F. B. (2004, October). Reconstructing Past Climate from Noisy Data.
 641 *Science*, 306, 679–682. doi: 10.1126/science.1096109
- 642 Wang, J., Emile-Geay, J., Guillot, D., Smerdon, J. E., & Rajaratnam, B. (2014).
 643 Evaluating climate field reconstruction techniques using improved emu-
 644

- lations of real-world conditions. *Climate of the Past*, 10(1), 1–19. doi: 10.5194/cp-10-1-2014
- Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., ... Kawamiya, M. (2011, January). MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments. *Geoscientific Model Development*, 4(4), 845–872. doi: 10.5194/gmd-4-845-2011
- Wilson, R., Anchukaitis, K., Briffa, K. R., Büntgen, U., Cook, E., D’Arrigo, R., ... Zorita, E. (2016, February). Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context. *Quaternary Science Reviews*, 134, 1–18. doi: 10.1016/j.quascirev.2015.12.005
- Wilson, R., Anchukaitis, K., Briffa, K. R., Büntgen, U., Cook, E., D’Arrigo, R., ... Zorita, E. (2016, 2 15). Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context. *Quaternary Science Reviews*, 134, 1–18. doi: 10.1016/j.quascirev.2015.12.005
- Wu, T., Song, L., Li, W., Wang, Z., Zhang, H., Xin, X., ... Zhou, M. (2014, February). An overview of BCC climate system model development and application for climate change studies. *Journal of Meteorological Research*, 28(1), 34–56. doi: 10.1007/s13351-014-3041-7
- Zhang, H., Yuan, N., Esper, J., Werner, J. P., Xoplaki, E., Büntgen, U., ... Luterbacher, J. (2015, August). Modified climate with long term memory in tree ring proxies. *Environmental Research Letters*, 10(8), 084020. doi: 10.1088/1748-9326/10/8/084020
- Zhu, F., Emile-Geay, J., Hakim, G. J., Tardif, R., & Perkins, A. (2019, December). *LMR Turbo (LMRt): a lightweight implementation of the LMR framework*. Zenodo. doi: 10.5281/zenodo.3590258

Supporting Information for “Resolving the differences in the simulated and reconstructed temperature response to volcanism”

Feng Zhu¹, Julien Emile-Geay¹, Gregory J. Hakim², Jonathan King^{3,5}, Kevin Anchukaitis^{3,4,5}

¹Department of Earth Sciences, University of Southern California, Los Angeles, CA USA

²Department of Atmospheric Sciences, University of Washington, Seattle, WA USA

³Department of Geosciences, University of Arizona, Tucson AZ USA

⁴School of Geography and Development, University of Arizona, Tucson AZ USA

⁵Laboratory of Tree-Ring Research, University of Arizona, Tucson AZ USA

Contents of this file

1. Text S1: Settings of the LMR framework
2. Text S2: Reconstructions using the Northern Hemisphere Tree-Ring Network Development (NTREND) network
3. Text S3: Choice of eruption key dates
4. Text S4: Software

5. Figure S1: Data from the PAGES 2k network (PAGES 2k Consortium, 2017) assimilated in LMR.

6. Figure S2: Comparison between different LMR implementations, and between model prior and reconstruction.

7. Figure S3: The pseudoproxy experiments (PPEs) that indicate the impact of spatial coverage and seasonality on the correlation between reconstruction and the pseudo-truth.

8. Figure S4: Impact of seasonality on the correlation between reconstruction and the Berkeley Earth instrumental temperature analysis.

9. Figure S5: Correlation between instrumental seasonal temperature observations and real proxy and pseudoproxy composites.

10. Figure S6: The detected signal-to-noise ratio (SNR) in PAGES2k TRW and MXD records.

11. Figure S7: The Northern Hemisphere Tree-Ring Network Development (NTREND) network.

12. Figure S8: The comparison between the model simulated temperature response and the LMR reconstruction assimilating the whole NTREND network.

13. Figure S9: Superposed Epoch Analysis (SEA) of model simulations and LMR reconstructions of the selected 6 eruption events, with individual simulation plotted out.

14. Figure S10: The temperature response to individual eruptions in LMR reconstructions assimilating the whole PAGES 2k Network and GCM simulations, targeting NHMT.

15. Figure S11: The temperature response to individual eruptions in LMR reconstructions assimilating the PAGES 2k MXD Network and GCM simulations, targeting mean summer temperature at proxy locales.

16. Figure S12: The temperature response to individual eruptions in LMR reconstructions assimilating the NTREND MXD Network and GCM simulations, targeting mean summer temperature at proxy locales.

17. Figure S13: Superposed Epoch Analysis (SEA) of model simulations and LMR reconstructions of all eruption events listed in Fig. S10.

18. Figure S14: Superposed Epoch Analysis (SEA) of model simulations and LMR reconstructions using CCSM4 as prior.

19. Figure S15: The comparison of the LMR reconstructions using univariate and bivariate forward operator for TRW records.

20. Figure S16: The available PAGES 2k records with start year earlier than 1257 AD, as well as the age of each of the PAGES 2k record.

21. Table S1: Last millennium model simulations considered in this study.

Introduction

This supporting information provides supplementary figures cited in the main text, as well as detailed information about the real- and pseudo-proxy reconstructions mentioned in the main text. Text S1 details the settings of our LMR experiments. Text S2 describes the reconstruction experiment using the Northern Hemisphere Tree-Ring Network Development (NTREND) (Wilson et al., 2016; Anchukaitis et al., 2017). Text S3 justifies the choice of eruption key dates. Software tools used for the analysis in this study are acknowledged in Text S4.

Text S1: Settings of the LMR framework

The reconstruction experiments in this study follow the general settings:

- Model prior: the isotope-enabled Community Earth System Model (iCESM) simulation (Stevenson et al., 2019; Brady et al., 2019) is used as the model prior. We have also tested using the CCSM4 last millennium simulation (Landrum et al., 2012) as model prior (Fig. S2a, S14) and no significant difference is detected in the temperature response to volcanic eruptions after 1400 AD.

- Ensemble design: 50 Monte Carlo iterations, each using a different randomly chosen 100-member ensemble states from the model prior, and 75% of randomly chosen available proxy records for assimilation (25% for independent verification). This scheme was chosen and explained in Hakim et al. (2016) to balance the needs of accuracy and uncertainty quantification.

- Localization scheme: the Gaspari-Cohn localization function (Gaspari & Cohn, 1999) is used with a radius of 25,000 km (Tardif et al., 2019).

- Forward operator: As in Tardif et al. (2019), we use seasonal bivariate (temperature and moisture) linear regression for tree-ring width (TRW) records, seasonal univariate (temperature) linear regression for maximum latewood density (MXD) records, and annual univariate (temperature) linear regression for all other proxy types as the forward operator in real proxy experiments. The forward operator is calibrated against the Goddard Institute for Space Studies (GISS) Surface Temperature Analysis (GISTEMP) (Hansen et al., 2010) instrumental observation and the gridded precipitation dataset from the Global Precipitation Climatology Centre (GPCC) (Schneider et al., 2014) over the timespan 1850-2015. In pseudoproxy experiments (PPE) the forward operator changes according to the experiment (see main text), and is calibrated against the model's true state over the same interval.

Text S2: Reconstructions using the Northern Hemisphere Tree-Ring Network Development (NTREND) network

The Northern Hemisphere Tree-Ring Network Development (NTREND) (Wilson et al., 2016; Anchukaitis et al., 2017) consists of 54 tree-ring chronologies spanning parts of North America and Eurasia. Of those 54 chronologies, 18 are pure maximum latewood density (MXD), 13 are pure tree-ring width (TRW), and 23 are mixed composites of MXD and TRW. The spatiotemporal sampling is shown in Fig. S7a, S7b.

As a benchmark, we first assimilate the whole NTREND network using the expert-curated seasonality, and the superposed epoch analysis (SEA) shows a similar discrepancy pattern as in IPCC AR5 Fig. 5.8b (Masson-Delmotte et al., 2013) (Fig. S8).

Applying our strategy for gap-bridging described in the main text, we assimilate only the 18 pure MXD records, and reconstruct the boreal summer temperature field, and then perform SEA at proxy locales. The result is shown in Fig. 4b (main text), which shows a better agreement between model simulations and the LMR reconstruction. Note that since 18 records are very few, we assimilate all the records in each ensemble member of assimilation, yield quite narrow uncertainty bands.

Text S3: Choice of eruption key dates Because superposed epoch analysis is an averaging operation, it involves a tradeoff between, on the one hand, maximizing the number of eruption key dates to reduce uncertainties, and on the other hand considerations particular to each eruption.

We chose to exclude eruptions after 1850 because the PMIP3 **past1000** protocol covers only the period (850-1850), and we wanted to be able to compare the greatest number of simulations to reconstructions.

When a cluster of eruptions are close to each other within 10 yrs, we select only the last one to avoid conflating the response of one eruption within the recovery for a preceding event. Note that not all PMIP3 simulations use the same volcanic forcing dataset (Schmidt et al., 2012), and that all differ from the more recent estimates of (Toohey & Sigl, 2017), which is a source of differences between simulations, and between simulations and reconstructions. Also note that neither 1452 nor 1459 (formerly attributed to

the Kuwae caldera) is selected. According to Toohey and Sigl (2017), the 1452 event in Gao, Robock, and Ammann (2008) was misaligned and is actually the 1459 event, so one should select the 1452 event instead of the 1459 event for GCM simulations. However, considering that the 1452 event is close to the 1459 event, we chose to skip both to avoid introducing an obvious discrepancy source for the comparison between GCM simulations and LMR reconstructions. Additionally, the 1761 and 1783 events are also skipped due to issue of misalignment according to Stevenson, Fasullo, Otto-Bliesner, Tomas, and Gao (2017) and Lücke, Hegerl, Schurer, and Wilson (2019).

Text S4: Software

All the analysis in this study was performed in the open-source Python programming language (Van Rossum & Drake Jr, 1995), version 3.7, with the following packages:

- `numpy` (van der Walt et al., 2011)
- `scipy` (Virtanen et al., 2019)
- `pandas` (McKinney, 2010)
- `statsmodels` (Seabold & Perktold, 2010)
- `matplotlib` (Hunter, 2007)
- `seaborn` (Waskom et al., 2018)

All reconstructions were performed with the Last Millennium Reanalysis fast implementation (`LMRt`), of Zhu, Emile-Geay, Hakim, Tardif, and Perkins (2019). This implementation yields nearly identical results compared to the official reconstruction (Fig. S2a), but with additional features:

- Greater flexibility
 - Easy installation
 - Easy importing and usage in Jupyter notebooks
 - Easy setup for different priors, proxies, and Proxy System Models (PSMs)
- Faster speed
 - Much faster PSM calibration due to optimization of algorithm
 - Easy parallel computing with multiprocessing and other techniques
- More modular code structure

References

- Anchukaitis, K. J., Wilson, R., Briffa, K. R., Büntgen, U., Cook, E. R., D'Arrigo, R., ... Zorita, E. (2017, 5 1). Last millennium Northern Hemisphere summer temperatures from tree rings: Part II, spatially resolved reconstructions. *Quaternary Science Reviews*, 163, 1–22. doi: 10.1016/j.quascirev.2017.02.020
- Brady, E., Stevenson, S., Bailey, D., Liu, Z., Noone, D., Nusbaumer, J., ... Zhu, J. (2019). The Connected Isotopic Water Cycle in the Community Earth System Model Version 1. *Journal of Advances in Modeling Earth Systems*, 11(8), 2547–2566. doi: 10.1029/2019MS001663
- Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., ... Vuichard, N. (2013, May). Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5. *Climate Dynamics*, 40(9-10), 2123–2165. doi: 10.1007/s00382-012-1636-1

- Gao, C., Robock, A., & Ammann, C. (2008). Volcanic forcing of climate over the past 1500 years: An improved ice core-based index for climate models. *Journal of Geophysical Research: Atmospheres*, 113(D23). doi: 10.1029/2008JD010239
- Gaspari, G., & Cohn, S. E. (1999). Construction of correlation functions in two and three dimensions. *Quarterly Journal of the Royal Meteorological Society*, 125(554), 723–757. doi: 10.1002/qj.49712555417
- Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., ... Stevens, B. (2013). Climate and carbon cycle changes from 1850 to 2100 in MPI-ESM simulations for the Coupled Model Intercomparison Project phase 5. *Journal of Advances in Modeling Earth Systems*, 5(3), 572–597. doi: 10.1002/jame.20038
- Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., ... Wood, R. A. (2000, February). The simulation of SST, sea ice extents and ocean heat transports in a version of the Hadley Centre coupled model without flux adjustments. *Climate Dynamics*, 16(2-3), 147–168. doi: 10.1007/s003820050010
- Hakim, G. J., Emile-Geay, J., Steig, E. J., Noone, D., Anderson, D. M., Tardif, R., ... Perkins, W. A. (2016). The last millennium climate reanalysis project: Framework and first results. *Journal of Geophysical Research: Atmospheres*, 121, 6745 – 6764. doi: 10.1002/2016JD024751
- Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). Global surface temperature change. *Rev. Geophys.*, 48, RG4004. doi: 10.1029/2010RG000345
- Hunter, J. D. (2007, May). Matplotlib: A 2d Graphics Environment. *Computing in Science Engineering*, 9(3), 90–95. doi: 10.1109/MCSE.2007.55

- Landrum, L., Otto-Bliesner, B. L., Wahl, E. R., Conley, A., Lawrence, P. J., Rosenbloom, N., & Teng, H. (2012, 2014/05/05). Last millennium climate and its variability in ccsm4. *Journal of Climate*, 26(4), 1085–1111. doi: 10.1175/JCLI-D-11-00326.1
- Lücke, L., Hegerl, G., Schurer, A., & Wilson, R. (2019, September). Effects of memory biases on variability of temperature reconstructions. *Journal of Climate*. doi: 10.1175/JCLI-D-19-0184.1
- Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., Rouco, J. G., ... Timmermann, A. (2013). Information from Paleoclimate Archives. In T. F. Stocker et al. (Eds.), *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 383–464). Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press. doi: 10.1017/CBO9781107415324.013
- McKinney, W. (2010). Data structures for statistical computing in python. In S. van der Walt & J. Millman (Eds.), *Proceedings of the 9th python in science conference* (p. 51 - 56).
- Otto-Bliesner, B. L., Brady, E. C., Fasullo, J., Jahn, A., Landrum, L., Stevenson, S., ... Strand, G. (2015). Climate variability and change since 850 CE: An ensemble approach with the community earth system model. *Bull. Amer. Meteor. Soc.*, 97(5), 735–754. doi: 10.1175/BAMS-D-14-00233.1
- PAGES 2k Consortium. (2017, 07). A global multiproxy database for temperature reconstructions of the Common Era. *Scientific Data*, 4, 170088 EP. doi: 10.1038/sdata.2017.88

- Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S., Rosenfeld, A., Wurtele, J., ... Mosher, S. (2013). Berkeley Earth Temperature Averaging Process. *Geoinformatics & Geostatistics: An Overview, 2013*. doi: 10.4172/2327-4581.1000103
- Rotstayn, L. D., Jeffrey, S. J., Collier, M. A., Dravitzki, S. M., Hirst, A. C., Syktus, J. I., & Wong, K. K. (2012, July). Aerosol- and greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: a study using single-forcing climate simulations. *Atmos. Chem. Phys.*, *12*(14), 6377–6404. doi: 10.5194/acp-12-6377-2012
- Schmidt, G. A., Jungclaus, J. H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., ... Vieira, L. E. A. (2012). Climate forcing reconstructions for use in pmip simulations of the last millennium (v1.1). *Geoscientific Model Development*, *5*(1), 185–191. doi: 10.5194/gmd-5-185-2012
- Schmidt, G. A., Ruedy, R., Hansen, J. E., Aleinov, I., Bell, N., Bauer, M., ... Yao, M.-S. (2006, January). Present-Day Atmospheric Simulations Using GISS ModelE: Comparison to In Situ, Satellite, and Reanalysis Data. *Journal of Climate*, *19*(2), 153–192. doi: 10.1175/JCLI3612.1
- Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., & Rudolf, B. (2014, January). GPCC’s new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, *115*(1), 15–40. doi: 10.1007/s00704-013-0860-x
- Seabold, S., & Perktold, J. (2010). statsmodels: Econometric and statistical modeling with python. In *9th python in science conference*.

- Stevenson, S., Fasullo, J. T., Otto-Bliesner, B. L., Tomas, R. A., & Gao, C. (2017). Role of eruption season in reconciling model and proxy responses to tropical volcanism. *Proceedings of the National Academy of Sciences*, *114*(8), 1822–1826. doi: 10.1073/pnas.1612505114
- Stevenson, S., Otto-Bliesner, B. L., Brady, E. C., Nusbaumer, J., Tabor, C., Tomas, R., ... Liu, Z. (2019). Volcanic Eruption Signatures in the Isotope-Enabled Last Millennium Ensemble. *Paleoceanography and Paleoclimatology*, *0*(0). doi: 10.1029/2019PA003625
- Tardif, R., Hakim, G. J., Perkins, W. A., Horlick, K. A., Erb, M. P., Emile-Geay, J., ... Noone, D. (2019, July). Last Millennium Reanalysis with an expanded proxy database and seasonal proxy modeling. *Climate of the Past*, *15*(4), 1251–1273. doi: 10.5194/cp-15-1251-2019
- Toohey, M., & Sigl, M. (2017, November). Volcanic stratospheric sulfur injections and aerosol optical depth from 500 BCE to 1900 CE. *Earth System Science Data*, *9*(2), 809–831. doi: 10.5194/essd-9-809-2017
- van der Walt, S., Colbert, S. C., & Varoquaux, G. (2011, March). The NumPy Array: A Structure for Efficient Numerical Computation. *Computing in Science Engineering*, *13*(2), 22–30. doi: 10.1109/MCSE.2011.37
- Van Rossum, G., & Drake Jr, F. L. (1995). *Python tutorial*. Centrum voor Wiskunde en Informatica Amsterdam, The Netherlands.
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., ... Contributors, S. . . (2019, Jul). SciPy 1.0–Fundamental Algorithms for Scientific

Computing in Python. *arXiv e-prints*, arXiv:1907.10121.

Waskom, M., Botvinnik, O., O’Kane, D., Hobson, P., Ostblom, J., Lukauskas, S., ...

Qalieh, A. (2018, July). *mwaskom/seaborn: v0.9.0 (july 2018)*. Zenodo. doi: 10.5281/zenodo.1313201

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., ...

Kawamiya, M. (2011, January). MIROC-ESM 2010: Model description and basic results of CMIP5-20c3m experiments. *Geoscientific Model Development*, 4(4), 845–872. doi: 10.5194/gmd-4-845-2011

Wilson, R., Anchukaitis, K., Briffa, K. R., Büntgen, U., Cook, E., D’Arrigo, R., ... Zorita,

E. (2016, 2 15). Last millennium northern hemisphere summer temperatures from tree rings: Part I: The long term context. *Quaternary Science Reviews*, 134, 1–18. doi: 10.1016/j.quascirev.2015.12.005

Wu, T., Song, L., Li, W., Wang, Z., Zhang, H., Xin, X., ... Zhou, M. (2014, February).

An overview of BCC climate system model development and application for climate change studies. *Journal of Meteorological Research*, 28(1), 34–56. doi: 10.1007/s13351-014-3041-7

Zhu, F., Emile-Geay, J., Hakim, G. J., Tardif, R., & Perkins, A. (2019, December). *LMR*

Turbo (LMRt): a lightweight implementation of the LMR framework. Zenodo. doi: 10.5281/zenodo.3590258

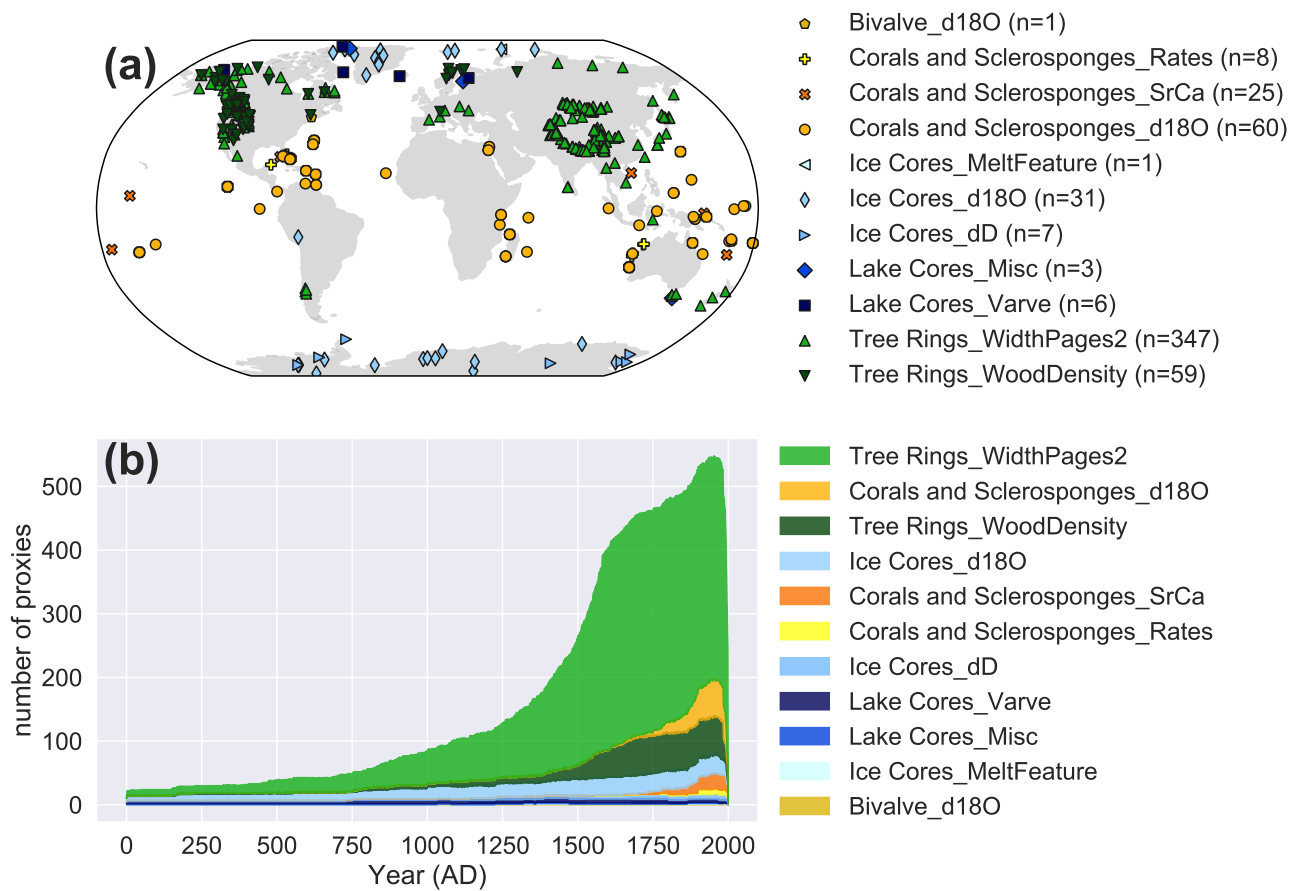


Figure S1. Data from the PAGES 2k network (PAGES 2k Consortium, 2017) assimilated in LMR. (a) Spatial coverage by archive type. (b) Temporal availability by archive type.

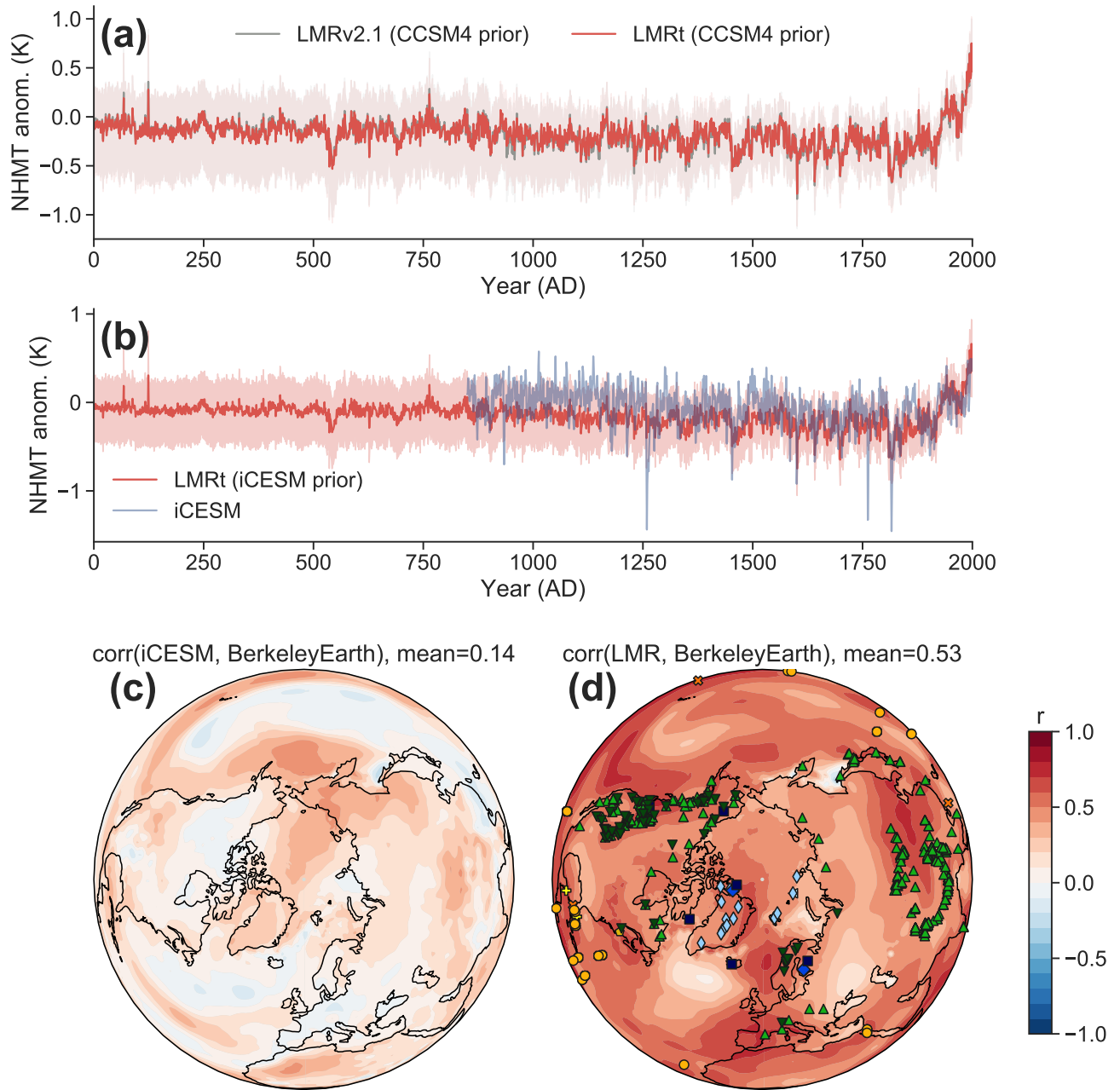


Figure S2. (a) The reconstructed northern hemisphere mean temperature (NHMT) series using the official LMR implementation (Tardif et al., 2019) and the lightweight implementation used in our study LMRt (Zhu et al., 2019), using the same CCSM4 model prior (Landrum et al., 2012) and the PAGES 2k phase 2 dataset (PAGES 2k Consortium, 2017). (b) The LMR reconstructed NHMT series assimilating the PAGES 2k network, along with its model prior, the simulated NHMT series from the isotope-enabled Community Earth System Model (iCESM) (Stevenson et al., 2019; Brady et al., 2019). (c) The correlation between the surface temperature simulated by iCESM and the instrumental observation Berkeley Earth instrumental temperature analysis (Rohde et al., 2013) over 1880 to 2000. (d) Same as (b) but for LMR reconstruction assimilating the PAGES 2k network. The symbols follow that in Fig. S1.

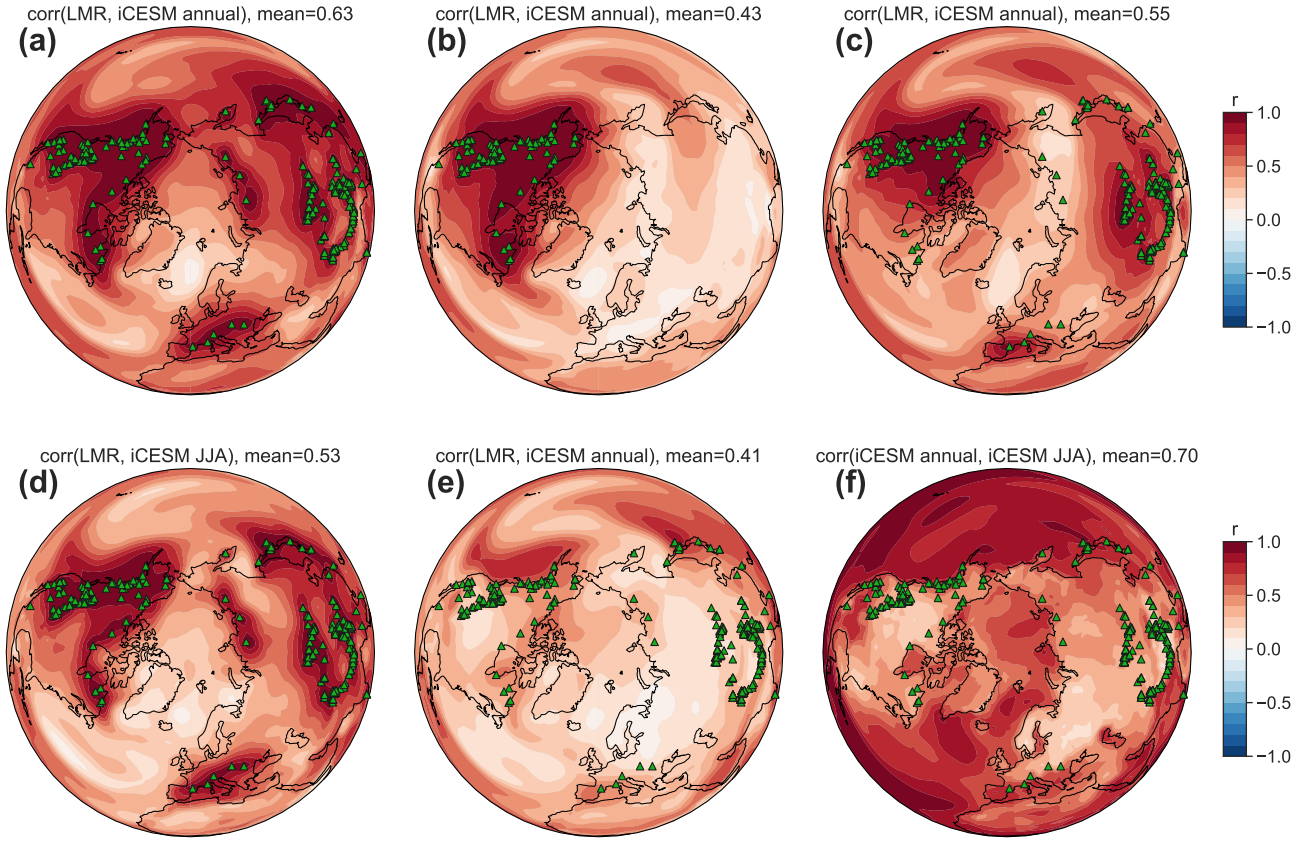


Figure S3. The pseudoproxy experiments (PPEs) that indicate the impact of spatial coverage and seasonality on the correlation between reconstruction and the pseudo-truth. (a) The pseudoproxies are generated as perfect temperature recorders of the annual temperature simulated by iCESM, and the whole network is assimilated. (b) Same as (a), but only 50 records over North America (NA) region are assimilated. (c) Same as (a), but only 50 records over Northern Hemisphere (NH) are assimilated. (d) Same as (a), but the pseudoproxies are generated as perfect temperature recorders of the summer temperature simulated by iCESM, and summer temperature field is reconstructed. (e) Same as (d), but annual temperature field is reconstructed. (f) The correlation between annual temperature and summer temperature simulated by iCESM.

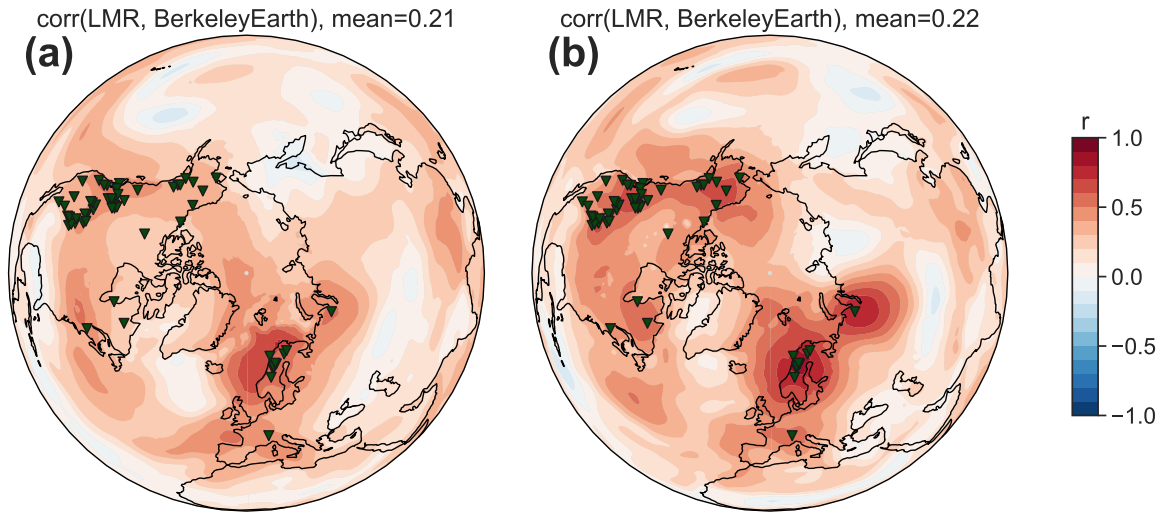


Figure S4. Impact of seasonality on the correlation between the reconstructions assimilating the MXD network and the Berkeley Earth instrumental temperature analysis (Rohde et al., 2013). (a) Reconstructing annual temperature (b) Reconstructing summer temperature. Note that both experiments use real, not pseudo, proxies.

Table S1. Last millennium model simulations considered in this study

Model	Experiment ID
iCESM (Stevenson et al., 2019; Brady et al., 2019)	-
CESM1 (Otto-Bliesner et al., 2015)	b.e11.BLMTRC5CN.f19_g16.001
BCC CSM1.1 (Wu et al., 2014)	past1000_r1ilp1
GISS-E2-R (Schmidt et al., 2006)	past1000_r1ilp1
HadCM3 (Gordon et al., 2000)	past1000_r1ilp1
IPSL-CM5A-LR (Dufresne et al., 2013)	past1000_r1ilp1
MIROC-ESM (Watanabe et al., 2011)	past1000_r1ilp1
MPI-ESM-P (Giorgetta et al., 2013)	past1000_r1ilp1
CSIRO (Rotstayn et al., 2012)	past1000_r1ilp1
CCSM4 (Landrum et al., 2012)	past1000_r1ilp1

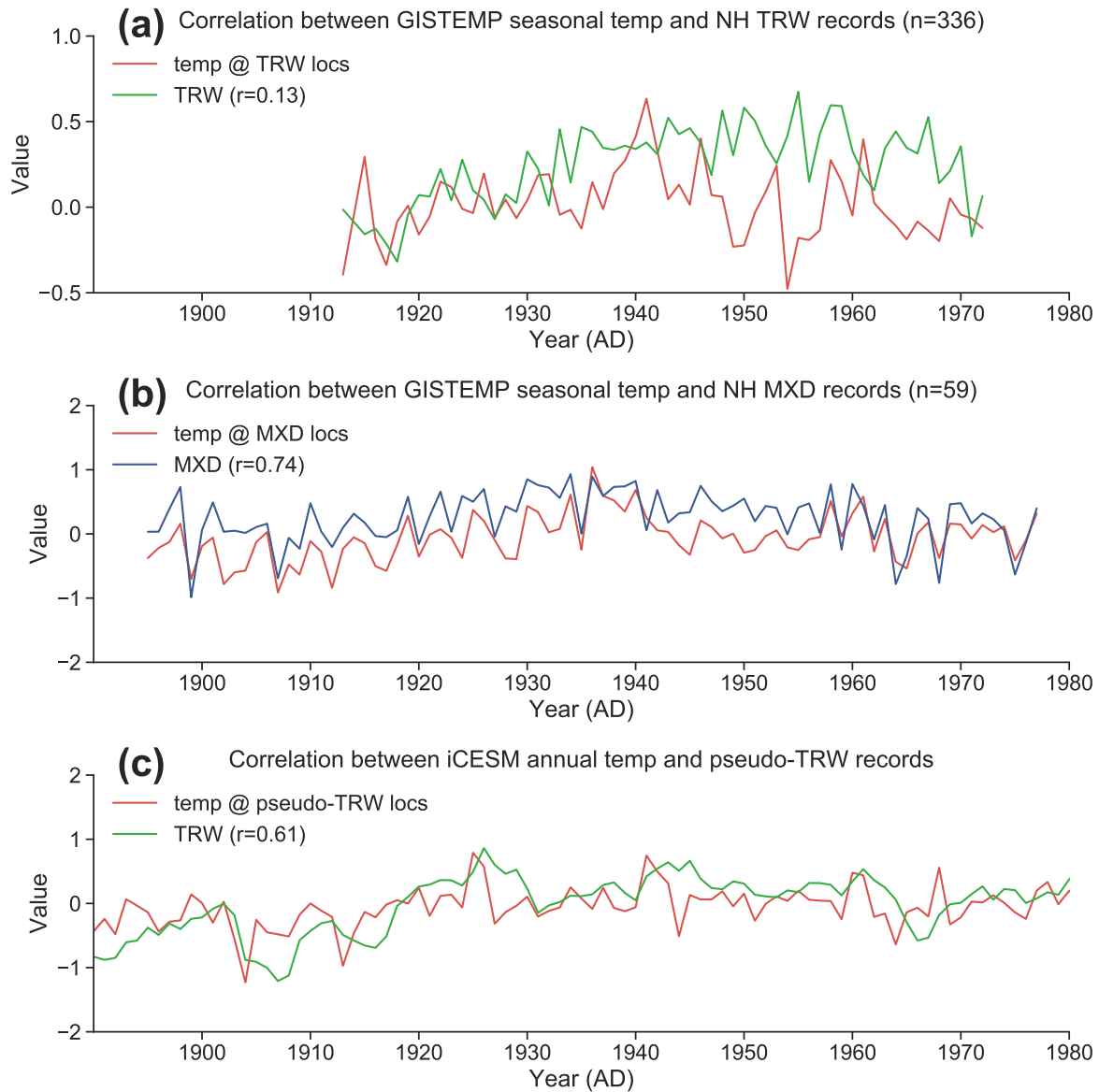


Figure S5. (a) The NH TRW composites compared to the seasonal observational temperature, the Goddard Institute for Space Studies (GISS) Surface Temperature Analysis (GISTEMP) (Hansen et al., 2010), at proxy locales. (b) Same as (a), but for MXD. (c) The composite of the pseudoproxy that is generated as temperature smoother with a 5-yr moving average filter, compared to the iCESM simulated temperature at the proxy locales.

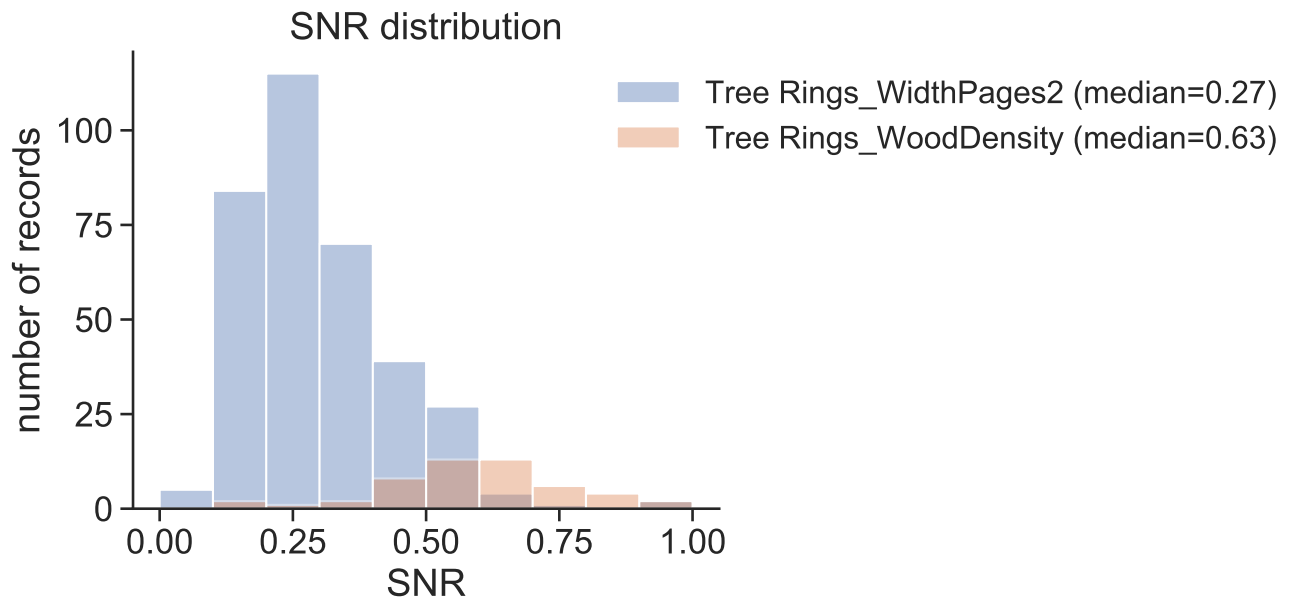


Figure S6. The signal-to-noise ratio (SNR) in TRW (Tree Rings_WidthPages2) and MXD (Tree Rings_WoodDensity) records detected by the forward operator calibration procedure (see Text S1 for details) that follows (Tardif et al., 2019) in LMR, with curated pre-defined seasonal windows. Higher SNR indicates more fraction of signal can be explained by seasonal temperature and moisture via bivariate and univariate linear regression.

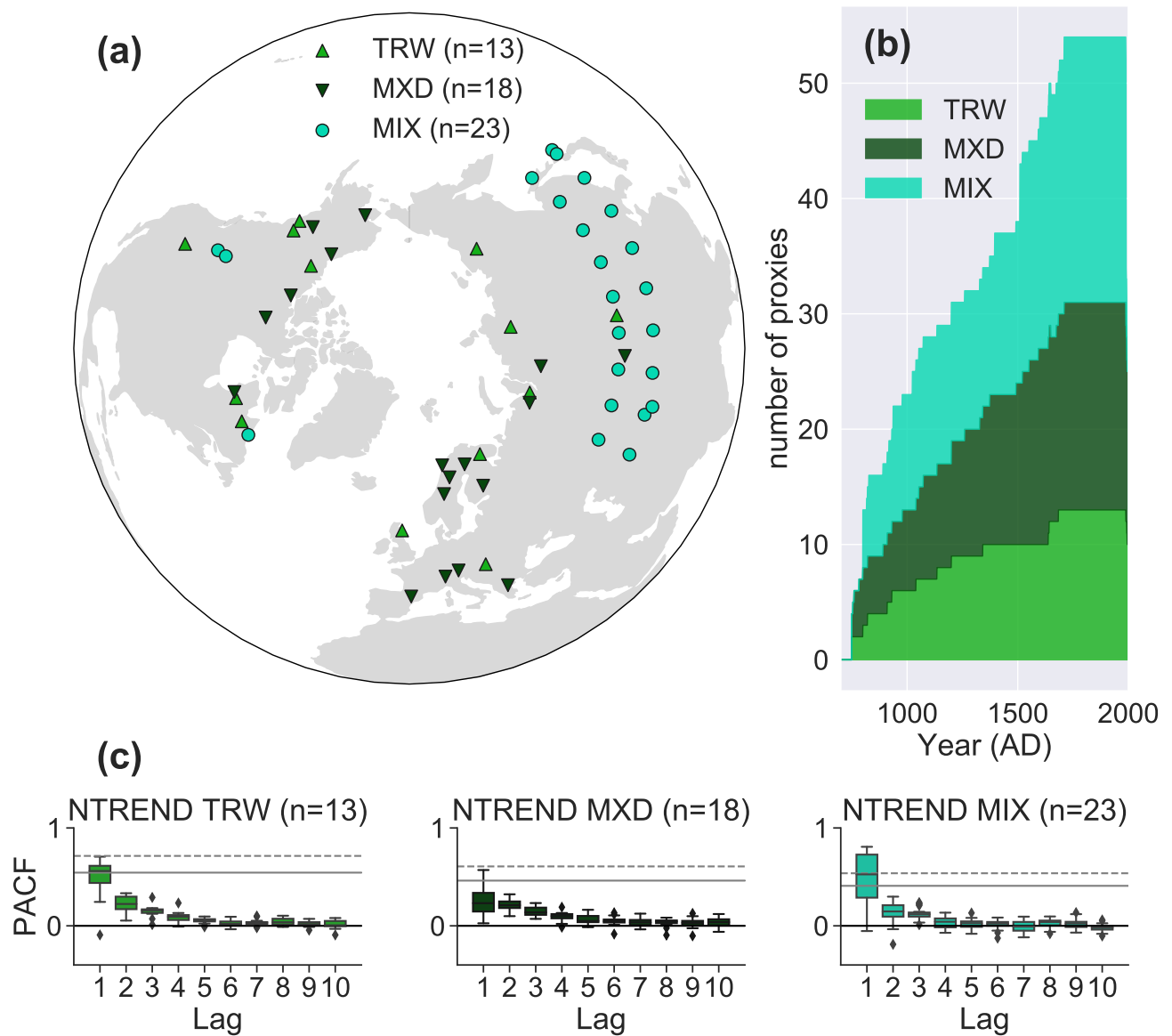


Figure S7. The Northern Hemisphere Tree-Ring Network Development (NTREND) (Wilson et al., 2016; Anchukaitis et al., 2017). (a) The spatial coverage of each proxy type. (b) The temporal availability of each proxy type. (c) The partial autocorrelation function (PACF) up to lag-10 for each proxy type.

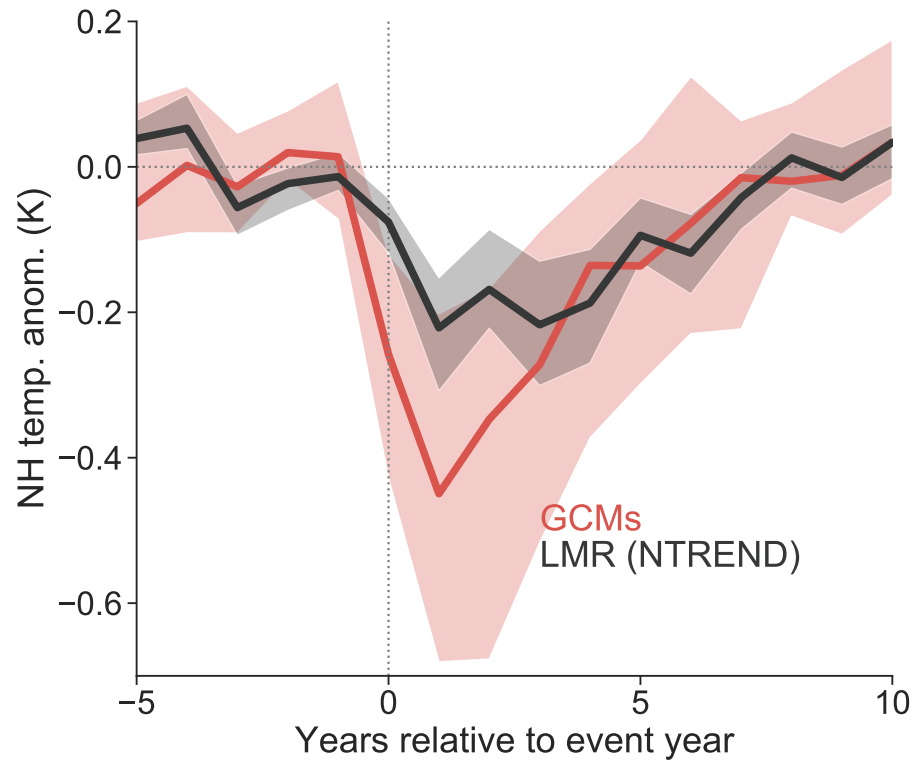


Figure S8. The comparison between the model simulated temperature response and the LMR reconstruction assimilating the whole NTREND network. SEA applied on the annual NHMT over the whole NH.

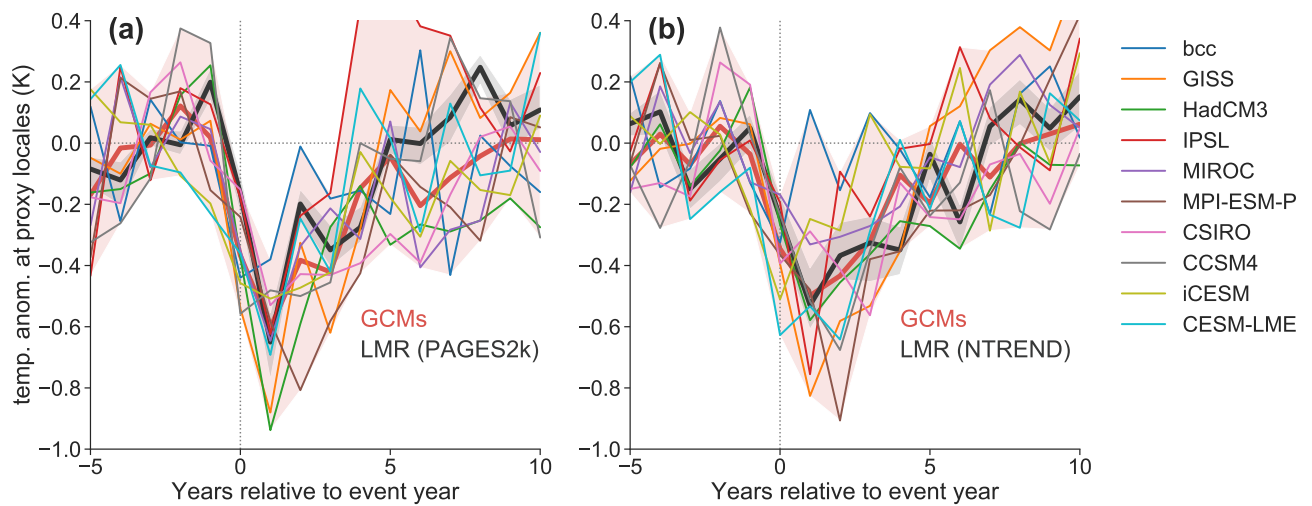


Figure S9. Similar to Fig. 4 (main text), but with the result of each model simulation plotted out.

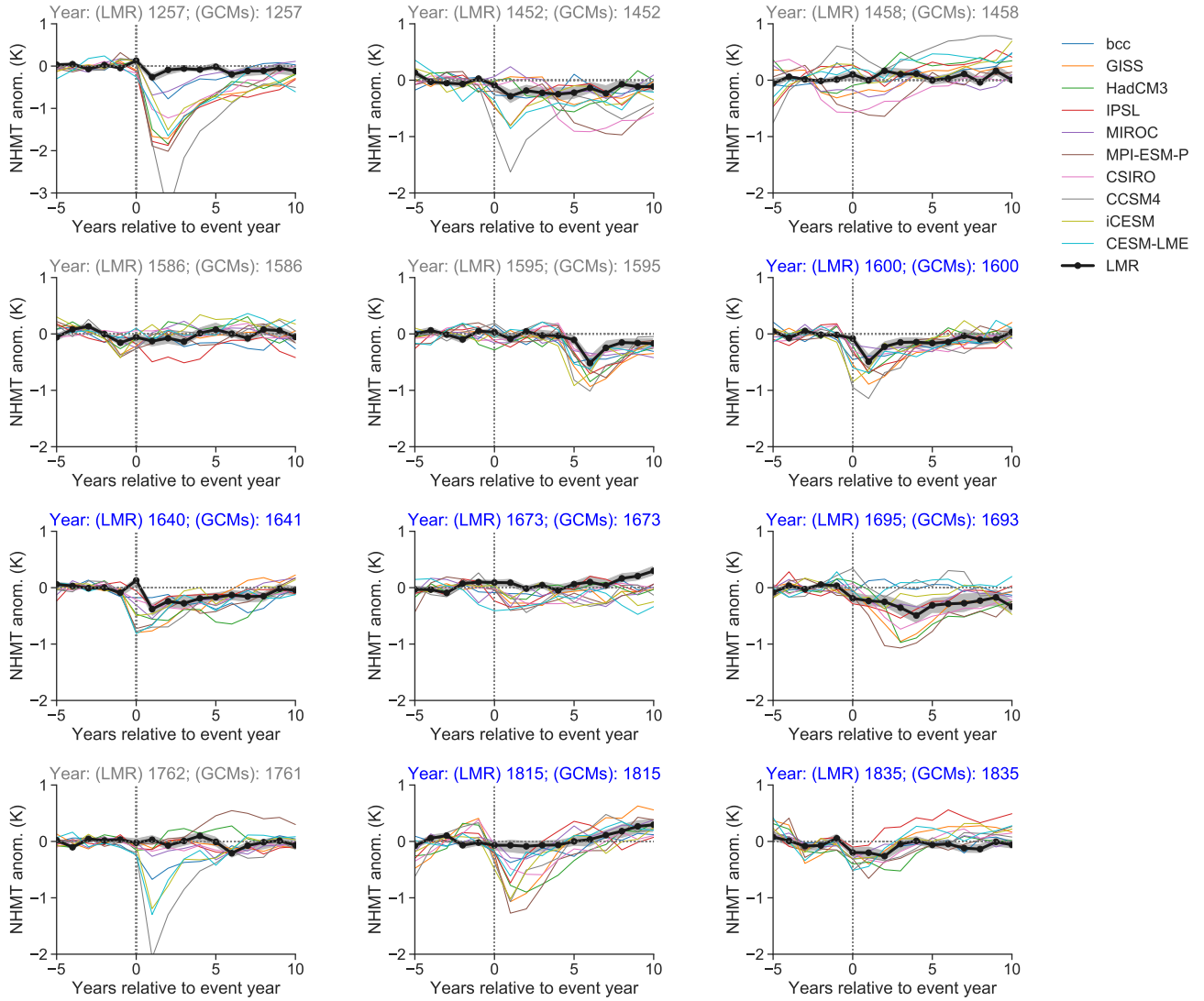


Figure S10. The temperature response to individual eruptions in LMR reconstructions assimilating the whole PAGES 2k Network and GCM simulations, targeting NHMT. The blue title denotes the 6 eruption events that are selected for SEA in our study.

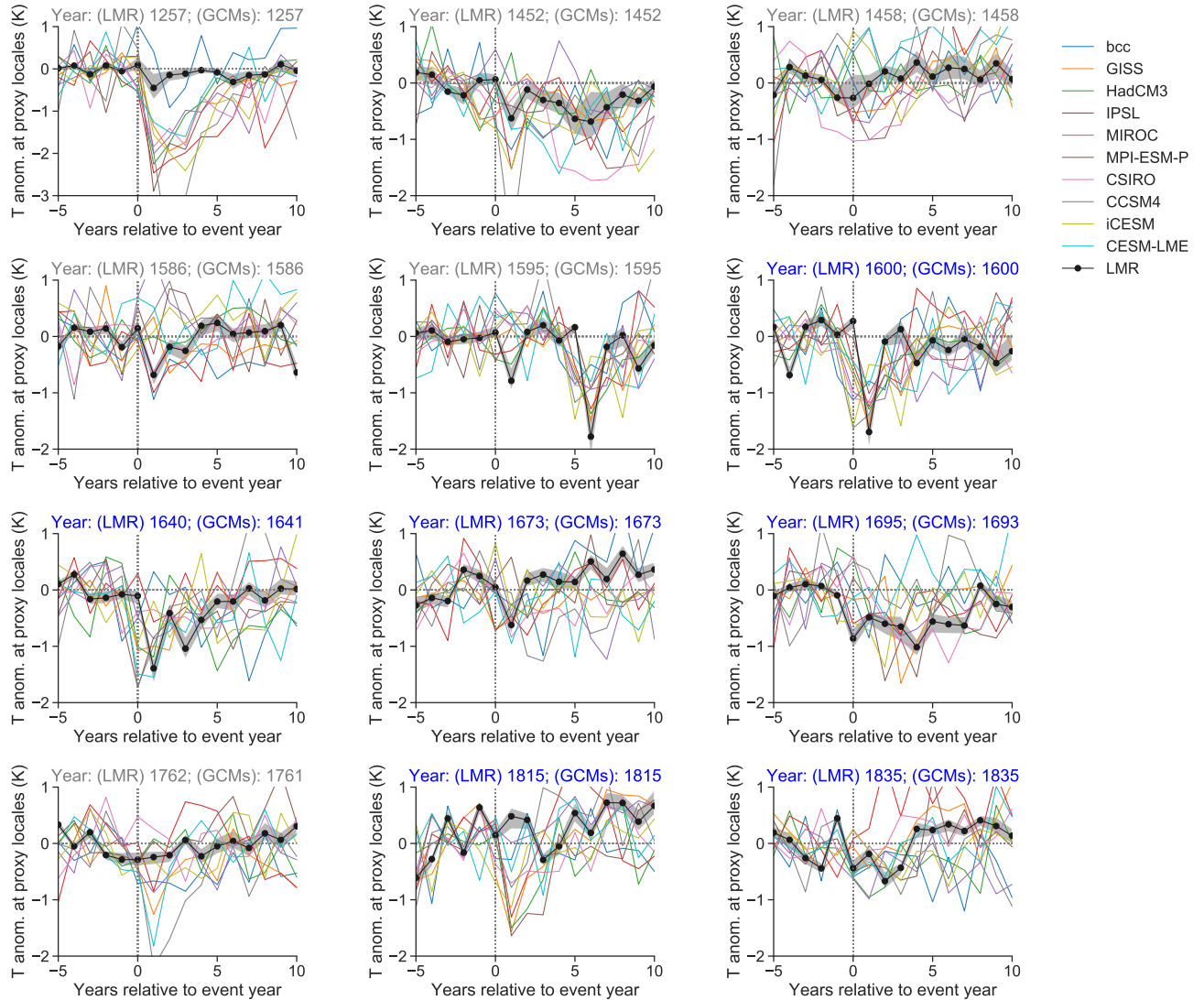


Figure S11. The temperature response to individual eruptions in LMR reconstructions assimilating the PAGES 2k MXD Network and GCM simulations, targeting mean summer temperature at proxy locales. The blue title denotes the 6 eruption events that are selected for SEA in our study.

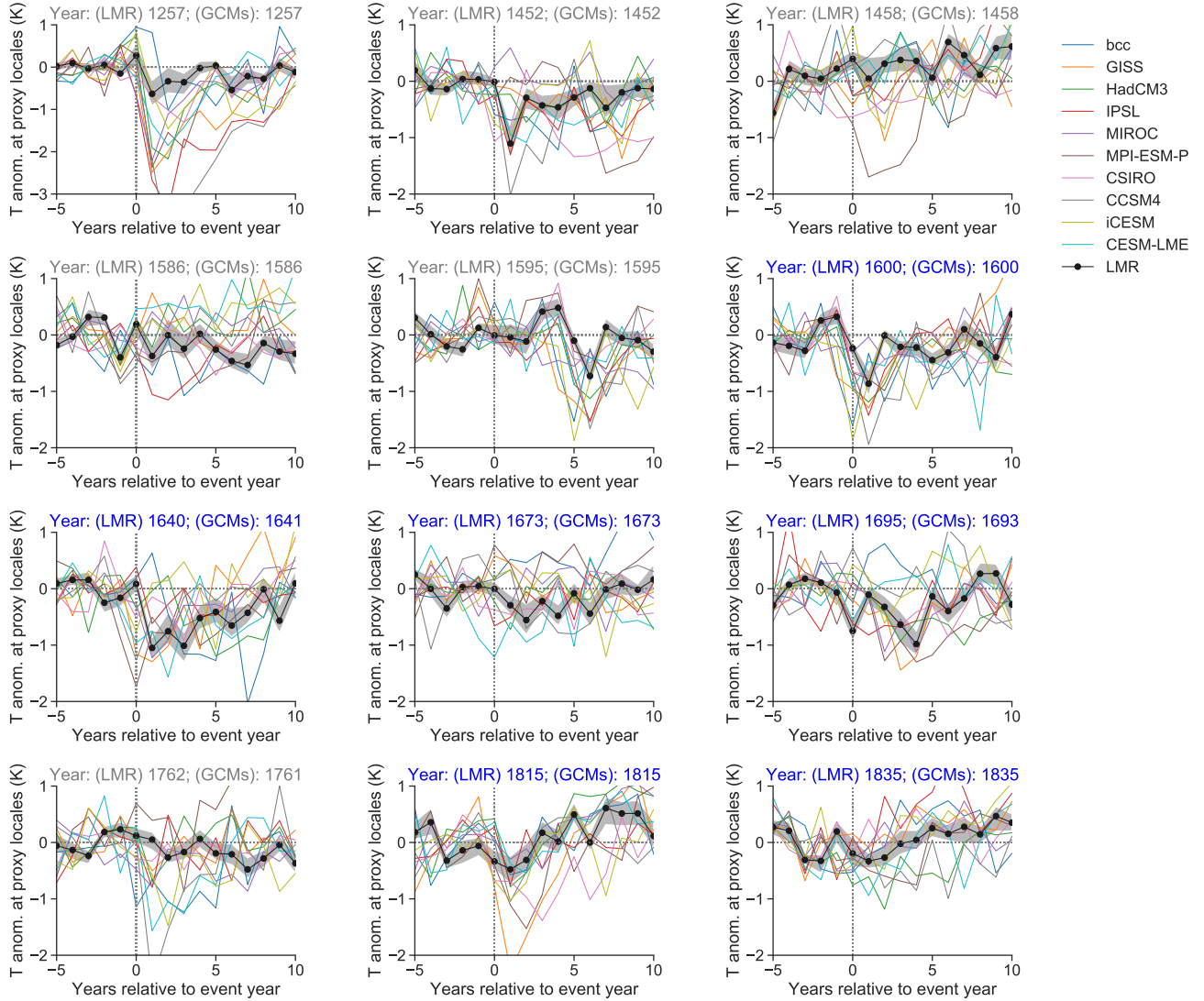


Figure S12. The temperature response to individual eruptions in LMR reconstructions assimilating the NTREND MXD Network and GCM simulations, targeting mean summer temperature at proxy locales. The blue title denotes the 6 eruption events that are selected for SEA in our study.

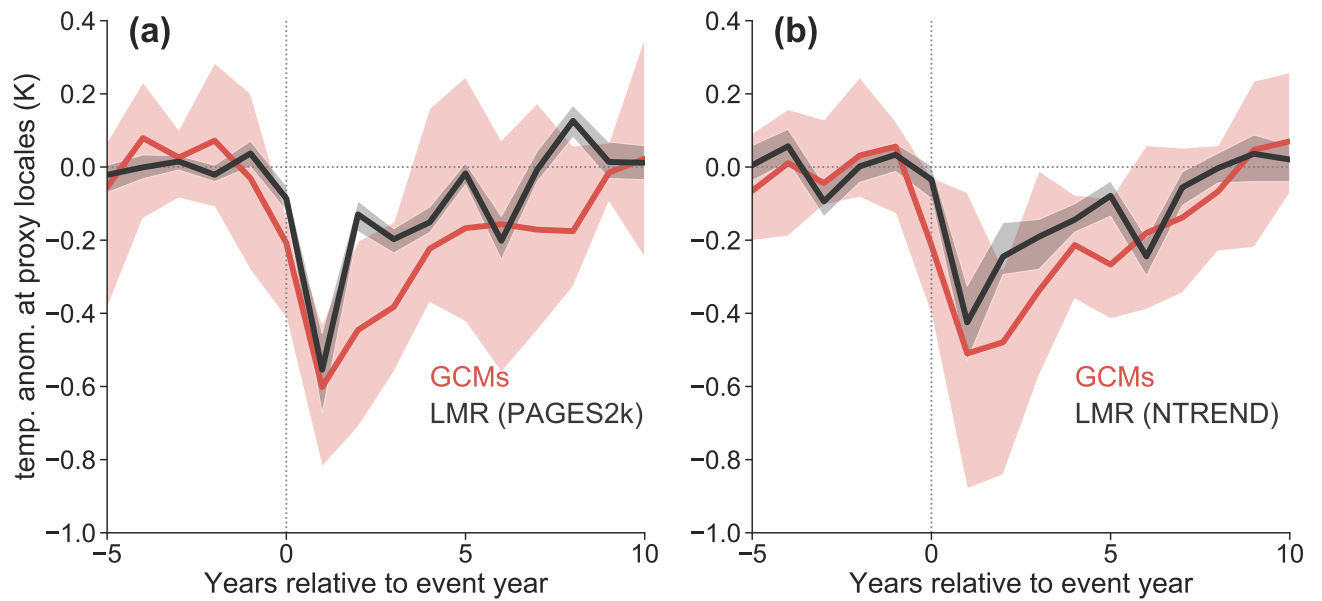


Figure S13. Same as Fig. 4 (main text), but SEA takes all eruption events listed in Fig. S10.

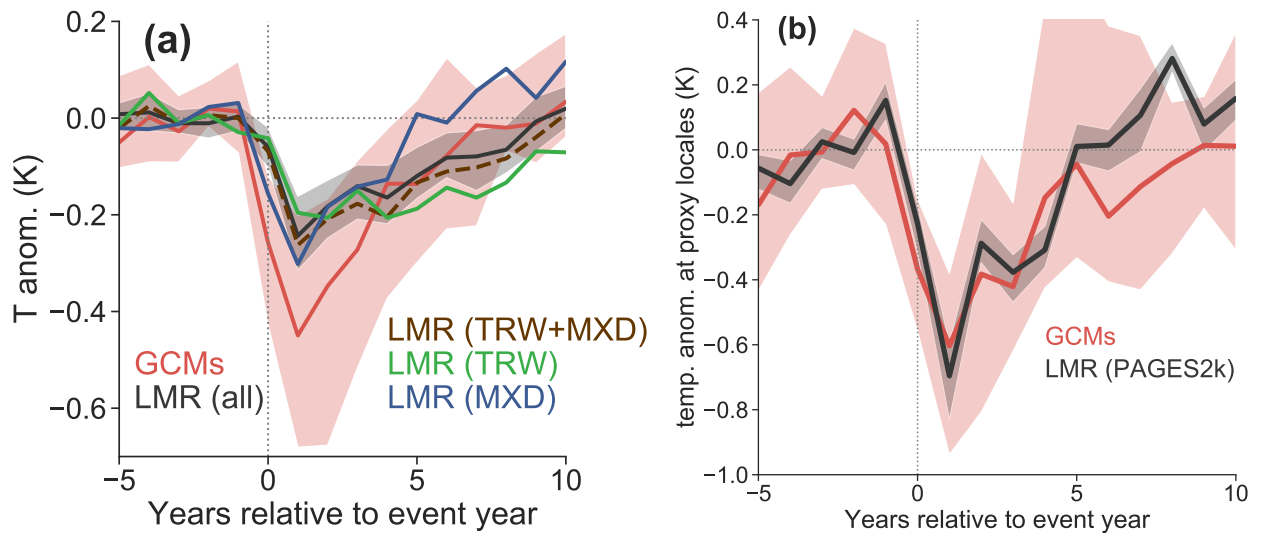


Figure S14. (a) Same as Fig. 1c (main text), but using CCSM4 as prior. (b) Same as Fig. 4a (main text), but using CCSM4 as prior.

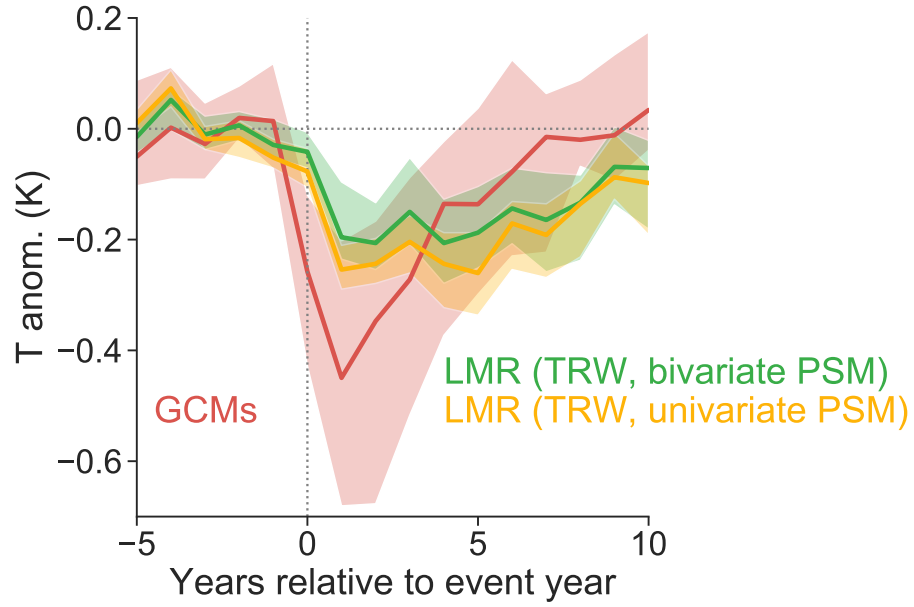


Figure S15. Same as Fig. 1c (main text), but using CCSM4 as prior and only showing the reconstructions assimilating the PAGES2k TRW network, with both bivariate and univariate forward operator calibration. The comparison indicates that moisture information does not alleviate the issue of lagged response to volcanism in TRW records.

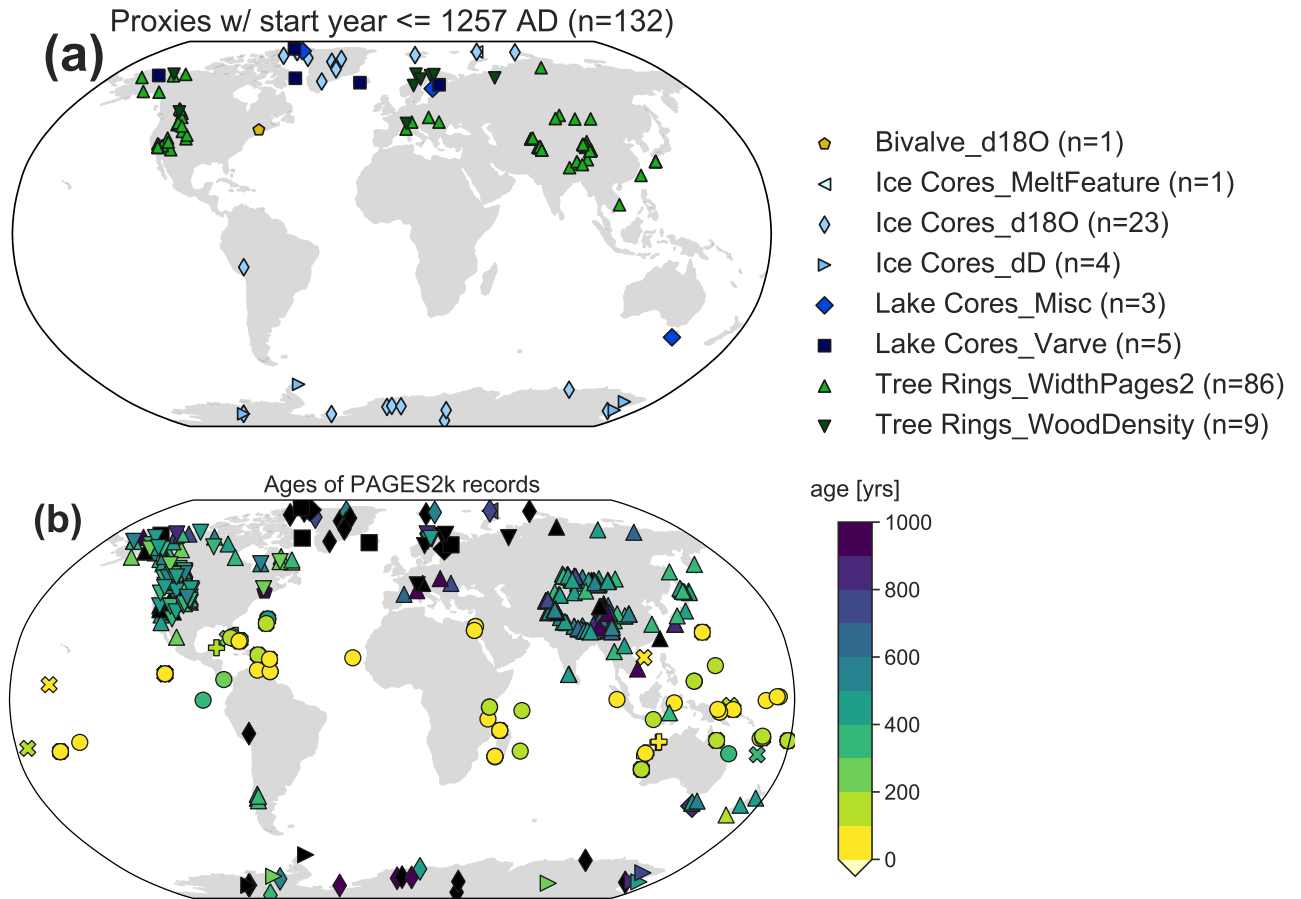


Figure S16. (a) Same as Fig. S1a, but for proxies with start year older than or equal to 1257 AD. The shapes and colors denote each proxy type. (b) Ages of PAGES2k records. The shape is same as in (a), while the colors denote different ranges of age.