AGU100 ADVANCING EARTH AND SPACE SCIENCE

Geophysical Research Letters

RESEARCH LETTER
10.1029/2019GL086908

Resolving the Differences in the Simulated and Reconstructed Temperature Response to Volcanism

Feng Zhu[1], **Julien Emile-Geay**[1], **Gregory J. Hakim**[2], **Jonathan King**[3,4], **and Kevin J. Anchukaitis**[3,4,5]

[1]Department of Earth Sciences, University of Southern California, Los Angeles, CA, USA, [2]Department of Atmospheric Sciences, University of Washington, Seattle, WA, USA, [3]Department of Geosciences, University of Arizona, Tucson, AZ, USA, [4]Laboratory of Tree-Ring Research, University of Arizona, Tucson, AZ, USA, [5]School 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

Supporting Information:
- Supporting Information S1

Correspondence to:
F. Zhu,
fengzhu@usc.edu

Citation:
Zhu, F., Emile-Geay, J., Hakim, G. J., King, J., & Anchukaitis, K. J. (2020). Resolving the differences in the simulated and reconstructed temperature response to volcanism. *Geophysical Research Letters*, *47*, e2019GL086908. https://doi.org/10.1029/2019GL086908

Received 31 DEC 2019
Accepted 20 MAR 2020
Accepted article online 28 MAR 2020

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 with respect to the magnitude and recovery from volcanic cooling, questioning the fidelity of climate model simulations, 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.

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.

1. Introduction

Volcanic eruptions influence the climate system through their direct effect on shortwave radiation entering the earth system and the subsequent response of major modes of ocean-atmosphere variability (Adams et al., 2003; Emile-Geay et al., 2008; Handler, 1984; Hirono, 1988; Li et al., 2013; Mann et al., 2005; Robock, 2000; Schneider et al., 2009; 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 (Figure 1b). Coupled Model Intercomparison Project 5 (CMIP5)/Paleoclimate Modeling Intercomparison Project 3 (PMIP3) model simulations for the last millennium experiment (Schmidt et al., 2012) show more cooling, and for a shorter duration, than paleoclimate reconstructions (Briffa et al., 2001; D'Arrigo et al., 2006; Frank, Esper, et al., 2007; Moberg et al., 2005). Compounding this uncertainty, the precise timing and location of some

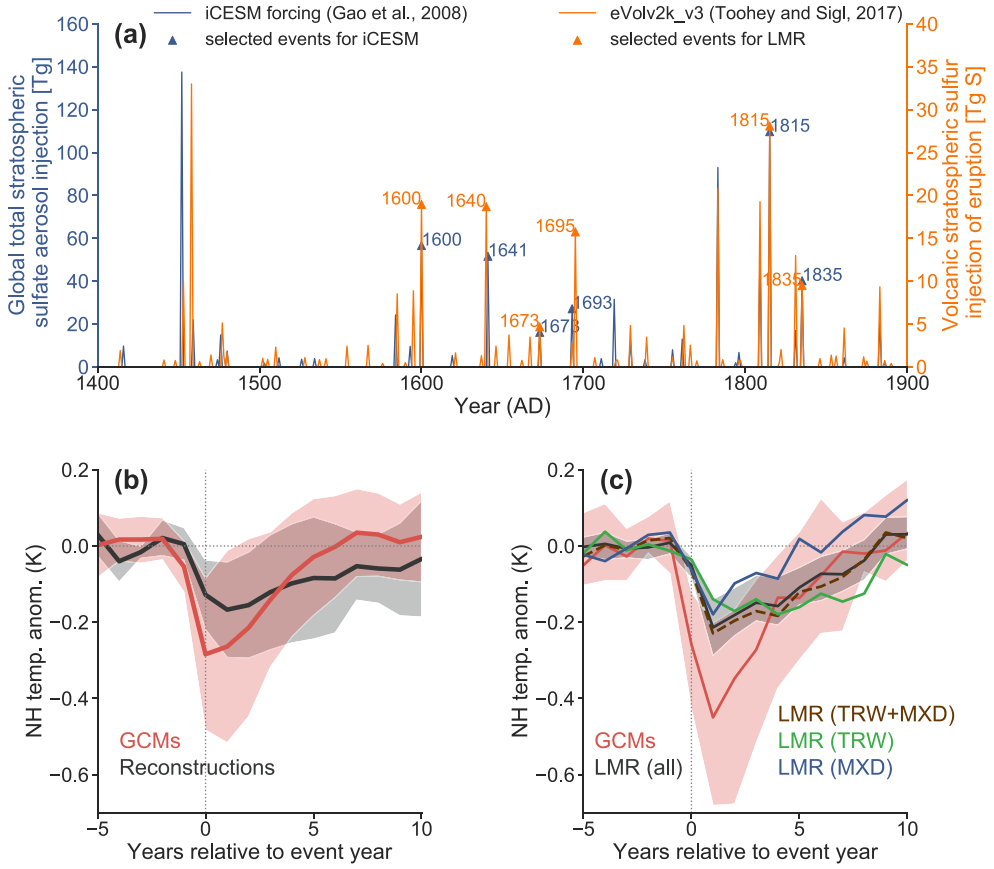

Figure 1. (a) Comparison between the volcanic forcing (Gao et al., 2008) used in the isotope-enabled Community Earth System Model (iCESM) simulation (Brady et al., 2019; Stevenson et al., 2019) and the eVolv2k version 3 Volcanic Stratospheric Sulfur Injection (VSSI) compilation (Toohey & Sigl, 2017). The triangles denote the selected six 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 CE, reproduced from IPCC AR5 (Masson-Delmotte et al., 2013) Figure 5.8b. (c) SEA on annual Northern Hemispheric mean temperature (NHMT) simulated by nine global climate models (GCMs) (section 2.2, supporting information, Table S1) and Last Millennium Reanalysis (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).

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; D'Arrigo et al., 2013; LeGrande & Anchukaitis, 2015; Stevenson et al., 2016; Stoffel et al., 2015; Timmreck, 2012).

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 Figure 1b is present in our reconstruction (Figure 1c), but that it can be largely reconciled by accounting for the aforementioned sources of uncertainty.

2. Data and Methods

2.1. Paleoclimate Data Assimilation

We apply the paleoenvironmental data assimilation framework of 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., 2013), 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) (Brady et al., 2019; Stevenson et al., 2019); (2) proxy data networks, which we take from the PAGES2k phase 2 compilation (PAGES 2k Consortium, 2017, Figure S1) and the Northern Hemisphere Tree-Ring Network Development (NTREND) compilation (Anchukaitis et al., 2017; Wilson et al., 2016a, Figure 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 PAGES2k 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) (Figures S2c and 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 (Figure 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 CE (compare Figure S2a to Figure S2b, Figure 1c to Figure S14a, and Figure 4a to Figure 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 (Braconnot et al., 2012; Schmidt 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

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 (Figure 1b) considered the reconstructed temperature response to the twelve strongest eruptions since 1400 CE. 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 six 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 injections

(Toohey & Sigl, 2017) (Figure 1a). For further details about the selection, see Text S3. The LMR response to individual events of the entire millennium is shown in Figures S10–S12.

3. The Discrepancy and its Causes

Figure 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, and (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 PAGES2k network to the model simulations (Figure 1c). Comparing results for solutions assimilating the entire PAGES2k 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 (LeGrande & Anchukaitis, 2015; LeGrande et al., 2016; Stoffel et al., 2015; Timmreck, 2012; Timmreck et al., 2009) 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 (Anchukaitis et al., 2017; D'Arrigo et al., 2006; Stoffel et al., 2015), (3) biological memory (Esper et al., 2015; Frank, Büntgen, et al., 2007; Fritts, 1966; Krakauer & Randerson, 2003; Lücke et al., 2019; Stoffel et al., 2015; Zhang et al., 2015), and (4) nontemperature "noise" (Neukom et al., 2018; Riedwyl et al., 2009; von Storch et al., 2004).

3.1. Spatial Coverage

The PAGES2k network is composed 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 (Figures 2a and 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 PAGES2k TRW records, (2) the locales of randomly picked 50 PAGES2k TRW records over North America, and (3) the locales of randomly picked 50 PAGES2k TRW records over the NH.

The result of assimilating these three pseudoproxy networks is shown in Figure 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: Figure 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 (Figure 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,

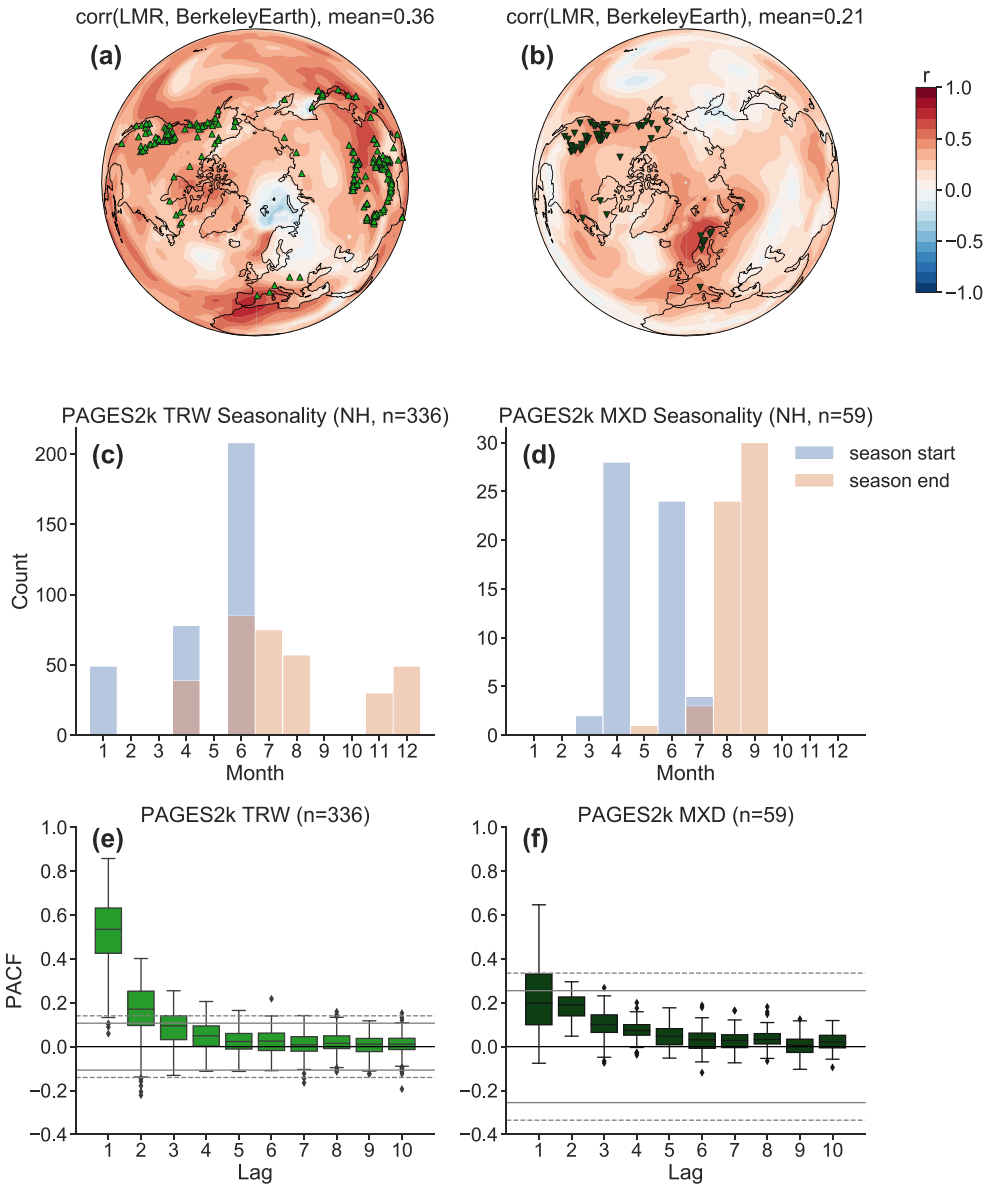

Figure 2. Differences between PAGES2k tree-ring width (TRW) and maximum latewood density (MXD) records regarding (a, b) spatial coverage, (c, d) seasonality detected by the algorithm used in Tardif et al. (2019), and (e, f) biological memory quantified by the partial autocorrelation function (PACF). (a) The spatial coverage of the TRW network. (c) The optimal seasonality of the TRW network. (e) The PACF of the TRW network. (b), (d), and (f) are same as (a), (b), and (e), respectively, but for the MXD network. The color contours in (a, b) indicate the correlation between the Last Millennium Reanalysis (LMR) reconstructions and the Berkeley Earth instrumental temperature analysis (Rohde et al., 2013).

1966; St George, 2014; St George & Ault, 2014; Stoffel et al., 2015; Wilson et al., 2016a). Though the PAGES2k 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 predefined 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 Figures 2c and 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 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

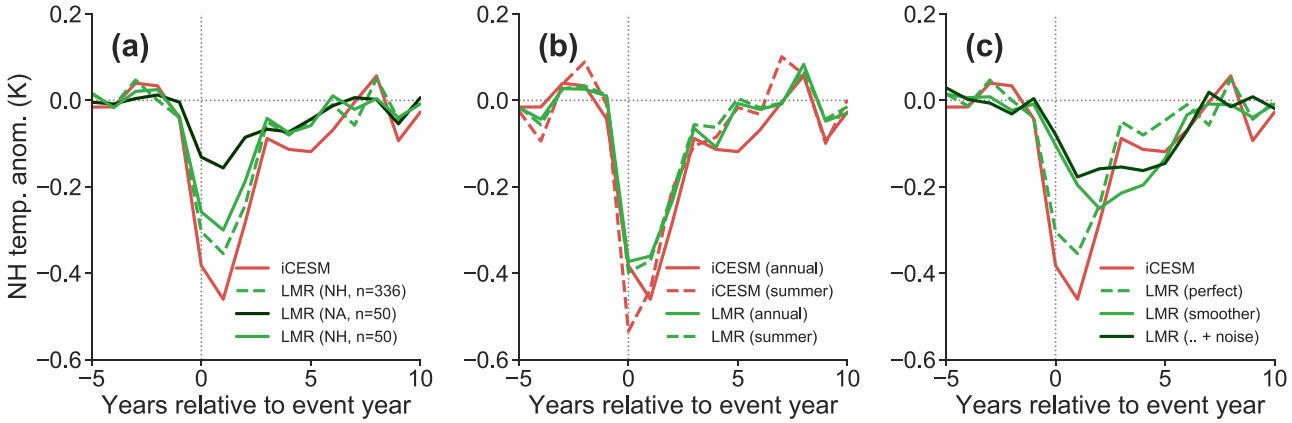

Figure 3. Superposed epoch analysis in pseudoproxy experiments, evaluating the impact of (a) spatial coverage, (b) seasonality, and (c) biological memory and noise. (a) The red curve denotes the target, and the dashed light green curve, the solid dark green curve, and the solid light green curve indicate the Last Millennium Reanalysis (LMR) reconstruction assimilating 336 pseudo-PAGES2k tree-ring width (TRW) records over the Northern Hemisphere (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 (signal-to-noise ratio [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). iCESM = isotope-enabled Community Earth System Model.

(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 (Figures S3d and S3e). This is also evident in reconstructions using real proxies and instrumental temperature (Figure 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 (Figure 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 (Esper et al., 2015; Frank, Büntgen, et al., 2007; Fritts, 1966; Krakauer & Randerson, 2003; Stoffel et al., 2015; Zhang et al., 2015). We quantify the persistence in TRW and MXD in the PAGES2k using the partial autocorrelation function (Figures 2e and 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 (Figures S5a and 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-year moving average of the annual temperature simulated by iCESM, as shown in Figure S5c. Assimilating these smoothed pseudoproxies yields a prolonged temperature recovery and a peak cooling that is both damped and lagged (Figure 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 relationship are observed in temperature-sensitive TRW chronologies irrespective of the potential additional influence of soil moisture (Consortium, 2017; Franke et al., 2013), which we confirm in sensitivity experiments (Figure S15).

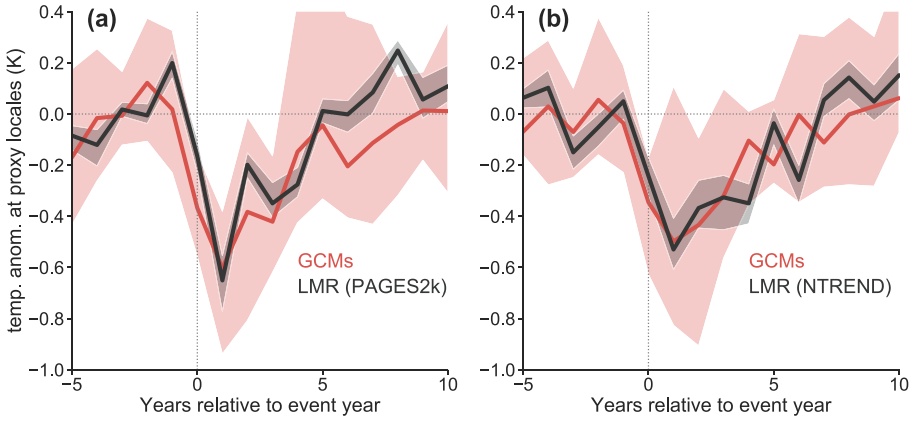

Figure 4. (a) Same as Figure 1c, after resolving differences in the model and proxy domains associated with seasonality, spatial distribution, and biological memory. (b) Same as (a) but using the NTREND maximum latewood density (MXD) network. A version of this figure showing each model simulation is available in Figure S9, and one using more eruption events is available in Figure S13. GCMs = global climate models; LMR = Last Millennium Reanalysis.

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 (Figure 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 (Figure 3c, solid dark green curve) to the real-world case (Figure 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 nontemperature 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 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 Figure 1 with the PAGES2k network (Figures 4a and S11). The same strategy can be used for other proxy networks. For comparison, applying it to the NTREND network (Anchukaitis et al., 2017; Wilson et al., 2016a) (Figure S7) yields similar agreement between simulated and reconstructed temperature (Figures 4b and S12). These results stand in sharp contrast to results where spatial coverage, seasonality, and biological memory are not taken into account (Figure S8).

That the discrepancy in Figure 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 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., 2013) 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 (Figures S11 and S12). While this has little impact in a composite over all events (Figure 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 Figures S2a and 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 (Figures S1b and S16a).
(2) Some CMIP5-era global climate models produce overly strong responses to volcanic forcing due to unrealistic representation of aerosol microphysics (LeGrande et al., 2016; Stoffel et al., 2015; Timmreck, 2012; Timmreck et al., 2009). 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 aeorosol loading was exceedingly large, and has since been revised downward (Jungclaus et al., 2017). There is also lingering uncertainty as to the magnitude, timing, and location of two major events during the 1450s (Hartman et al., 2019; Sigl et al., 2015; Toohey & Sigl, 2017). 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 (Aubry et al., 2019; Toohey & Sigl, 2017), 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 noncontemporaneous 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 (Evans et al., 2013; Tolwinski-Ward et al., 2011), 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 paleotemperature 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, for example, 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 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 associated with this study are available at https://doi.org/10.5281/zenodo.3725030.

References

Adams, J., Mann, M. E., & Ammann, C. M. (2003). Proxy evidence for an El Niño-like response to volcanic forcing. *Nature*, *426*(6964), 274–278. https://doi.org/10.1038/nature02101

Anchukaitis, K. J., Breitenmoser, P., Briffa, K. R., Buchwal, A., Buntgen, U., Cook, E. R., et al. (2012). Tree rings and volcanic cooling. *Nature Geoscience*, *5*(12), 836–837. https://doi.org/10.1038/ngeo1645

Anchukaitis, K. J., Wilson, R., Briffa, K. R., Büntgen, U., Cook, E. R., D'Arrigo, R., et al. (2017). Last millennium Northern Hemisphere summer temperatures from tree rings: Part II, spatially resolved reconstructions. *Quaternary Science Reviews*, *163*, 1–22. https://doi.org/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*, *125*, e2019JD031303. https://doi.org/10.1029/2019JD031303

Braconnot, P., Harrison, S. P., Kageyama, M., Bartlein, P. J., Masson-Delmotte, V., Abe-Ouchi, A., et al. (2012). Evaluation of climate models using palaeoclimatic data. *Nature Climate Change*, *2*(6), 417–424. https://doi.org/10.1038/nclimate1456

Brady, E., Stevenson, S., Bailey, D., Liu, Z., Noone, D., Nusbaumer, J., et al. (2019). The connected isotopic water cycle in the Community Earth System Model version 1. *Journal of Advances in Modeling Earth Systems*, *11*, 2547–2566. https://doi.org/10.1029/2019MS001663

Breitenmoser, P., Beer, J., Brönnimann, S., Frank, D., Steinhilber, F., & Wanner, H. (2012). Solar and volcanic fingerprints in tree-ring chronologies over the past 2000 years. *Palaeogeography, Palaeoclimatology, Palaeoecology*, *313-314*, 127–139. https://doi.org/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*, *106*(D3), 2929–2941. https://doi.org/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. https://doi.org/10.1175/JCLI-D-14-00802.1

Consortium, P. H. (2017). Comparing proxy and model estimates of hydroclimate variability and change over the Common Era. *Climate of the Past*, *13*(12), 1851–1900. 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*, 9000–9010. https://doi.org/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*, *111*, D03103. https://doi.org/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*, 1164–1179. https://doi.org/10.1002/2016MS000677

Dufresne, J.-L., Foujols, M.-A., Denvil, S., Caubel, A., Marti, O., Aumont, O., et al. (2013). Climate change projections using the IPSL-CM5 Earth System Model: From CMIP3 to CMIP5. *Climate Dynamics*, *40*(9-10), 2123–2165. https://doi.org/10.1007/s00382-012-1636-1

Emile-Geay, J., Seager, R., Cane, M. A., Cook, E. R., & Haug, G. H. (2008). Volcanoes and ENSO over the past millennium. *Journal of Climate*, *21*(13), 3134–3148. https://doi.org/10.1175/2007JCLI1884.1

Esper, J., George, S. S., Anchukaitis, K., D'Arrigo, R., Ljungqvist, F. C., Luterbacher, J., et al. (2018). Large-scale, millennial-length temperature reconstructions from tree-rings. *Dendrochronologia*, *50*, 81–90. https://doi.org/10.1016/j.dendro.2018.06.001

Esper, J., Schneider, L., Smerdon, J. E., Schöne, B. R., & Büntgen, U. (2015). Signals and memory in tree-ring width and density data. *Dendrochronologia*, *35*, 62–70. https://doi.org/10.1016/j.dendro.2015.07.001

Evans, M. N., Tolwinski-Ward, S. E., Thompson, D. M., & Anchukaitis, K. J. (2013). Applications of proxy system modeling in high resolution paleoclimatology. *Quaternary Science Reviews*, *76*, 16–28. https://doi.org/10.1016/j.quascirev.2013.05.024

Frank, D., Büntgen, U., Böhm, R., Maugeri, M., & Esper, J. (2007). Warmer early instrumental measurements versus colder reconstructed temperatures: Shooting at a moving target. *Quaternary Science Reviews*, *26*(25), 3298–3310. https://doi.org/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*, L16709. https://doi.org/10.1029/2007GL030571

Franke, J., Frank, D., Raible, C. C., Esper, J., & Brönnimann, S. (2013). Spectral biases in tree-ring climate proxies. *Nature Climate Change*, *3*(4), 360–364. https://doi.org/10.1038/nclimate1816

Fritts, H. C. (1966). Growth-rings of trees: Their correlation with climate. *Science*, *154*(3752), 973–979. https://doi.org/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. https://doi.org/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*, *113*, D23111. https://doi.org/10.1029/2008JD010239

Giorgetta, M. A., Jungclaus, J., Reick, C. H., Legutke, S., Bader, J., Böttinger, M., et al. (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*, 572–597. https://doi.org/10.1002/jame.20038

Gordon, C., Cooper, C., Senior, C. A., Banks, H., Gregory, J. M., Johns, T. C., et al. (2000). 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. https://doi.org/10.1007/s003820050010

Guillet, S., Corona, C., Stoffel, M., Khodri, M., Lavigne, F., Ortega, P., et al. (2017). Climate response to the Samalas volcanic eruption in 1257 revealed by proxy records. *Nature Geoscience*, *10*(2), 123–128. https://doi.org/10.1038/ngeo2875

Hakim, G. J., Emile-Geay, J., Steig, E. J., Noone, D., Anderson, D. M., Tardif, R., et al. (2016). The last millennium climate reanalysis project: Framework and first results. *Journal of Geophysical Research: Atmospheres*, *121*, 6745–6764. https://doi.org/10.1002/2016JD024751

Handler, P. (1984). Possible association of stratospheric aerosols and El Nino type events. *Geophysical Research Letters*, *11*(11), 1121–1124. https://doi.org/10.1029/GL011i011p01121

Hansen, J., Ruedy, R., Sato, M., & Lo, K. (2010). Global surface temperature change. *Reviews of Geophysics*, *48*, RG4004. https://doi.org/10.1029/2010RG000345

Hartman, L. H., Kurbatov, A. V., Winski, D. A., Cruz-Uribe, A. M., Davies, S. M., Dunbar, N. W., et al. (2019). Volcanic glass properties from 1459 CE volcanic event in South Pole ice core dismiss Kuwae caldera as a potential source. *Scientific Reports*, *9*(1), 14437. https://doi.org/10.1038/s41598-019-50939-x

Haurwitz, M. W., & Brier, G. W. (1981). A critique of the superposed epoch analysis method: Its application to solar–weather relations. *Monthly Weather Review*, *109*(10), 2074–2079. https://doi.org/10.1175/1520-0493(1981)109h2074:ACOTSEi2.0.CO;2

Hirono, M. (1988). On the trigger of El Niño Southern Oscillation by the forcing of early El Chichn volcanic aerosols. *Journal of Geophysical Research*, *93*(D5), 5365–5384. https://doi.org/10.1029/JD093iD05p05365

Hunter, J. D. (2007). Matplotlib: A 2D graphics environment. *Computing in Science Engineering*, *9*(3), 90–95. https://doi.org/10.1109/MCSE.2007.55

Jungclaus, J. H., Bard, E., Baroni, M., Braconnot, P., Cao, J., Chini, L. P., et al. (2017). 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, https://doi.org/10.5194/gmd-10-4005-2017

Kageyama, M., Braconnot, P., Harrison, S. P., Haywood, A. M., Jungclaus, J. H., Otto-Bliesner, B. L., et al. (2018). The PMIP4 contribution to CMIP6 – Part 1: Overview and over-arching analysis plan. *Geoscientific Model Development*, *11*(3), 1033–1057. 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), 1118. https://doi.org/10.1029/2003GB002076

Landrum, L., Otto-Bliesner, B. L., Wahl, E. R., Conley, A., Lawrence, P. J., Rosenbloom, N., & Teng, H. (2012). Last millennium climate and its variability in CCSM4. *Journal of Climate*, *26*(4), 1085–1111. https://doi.org/10.1175/JCLI-D-11-00326.1

LeGrande, A. N., & Anchukaitis, K. J. (2015). Volcanic eruptions and climate. *PAGES Magazine*, *23*(2), 46–47. https://doi.org/10.22498/pages.23.2.46

LeGrande, A. N., Tsigaridis, K., & Bauer, S. E. (2016). Role of atmospheric chemistry in the climate impacts of stratospheric volcanic injections. *Nature Geoscience*, *9*(9), 652–655. https://doi.org/10.1038/ngeo2771

Li, J., Xie, S.-P., Cook, E. R., Morales, M. S., Christie, D. A., Johnson, N. C., et al. (2013). El Niño modulations over the past seven centuries. *Nature Climate Change*, *3*(9), 822–826. https://doi.org/10.1038/nclimate1936

Lücke, L., Hegerl, G., Schurer, A., & Wilson, R. (2019). Effects of memory biases on variability of temperature reconstructions. *Journal of Climate*. *32*, 8713–8731. https://doi.org/10.1175/JCLI-D-19-0184.1

Mann, M. E., Cane, M. A., Zebiak, S. E., & Clement, A. (2005). Volcanic and solar forcing of the tropical Pacific over the past 1000 years. *Journal of Climate*, *18*(3), 447–456. https://doi.org/10.1175/JCLI-3276.1

Masson-Delmotte, V., Schulz, M., Abe-Ouchi, A., Beer, J., Ganopolski, A., Rouco, J. F. G., et al. (2013). Information from Paleoclimate Archives. In Stocker, T. F., 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. https://doi.org/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* (51–56). June 28–July 3. Austin, Texas. https://doi.org/10.25080/Majora-92bf1922-012

Moberg, A., Sonechkin, D. M., Holmgren, K., Datsenko, N. M., & Karln, W. (2005). Highly variable Northern Hemisphere temperatures reconstructed from low- and high-resolution proxy data. *Nature*, *433*(7026), 613–617. https://doi.org/10.1038/nature03265

Neukom, R., Schurer, A. P., Steiger, N. J., & Hegerl, G. C. (2018). Possible causes of data model discrepancy in the temperature history of the last millennium. *Scientific Reports*, *8*(1), 1–15. https://doi.org/10.1038/s41598-018-25862-2

Otto-Bliesner, B. L., Brady, E. C., Fasullo, J., Jahn, A., Landrum, L., Stevenson, S., et al. (2015). Climate variability and change since 850 CE: An ensemble approach with the community earth system model. *Bulletin of the American Meteorological Society*, *97*(5), 735–754. https://doi.org/10.1175/BAMS-D-14-00233.1

PAGES 2k Consortium (2017). A global multiproxy database for temperature reconstructions of the Common Era. *Scientific Data*, *4*, 170088 EP. https://doi.org/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. https://doi.org/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. https://doi.org/10.1007/s00382-008-0395-5

Robock, A. (2000). Volcanic eruptions and climate. *Reviews of Geophysics*, *38*, 191–220. https://doi.org/10.1029/1998RG000054

Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S., Rosenfeld, A., Wurtele, J., et al. (2013). Berkeley Earth temperature averaging process. Geoinformatics & geostatistics: An overview, https://doi.org/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). Aerosol- and greenhouse gas-induced changes in summer rainfall and circulation in the Australasian region: A study using single-forcing climate simulations. *Atmospheric Chemistry and Physics*, *12*(14), 6377–6404. https://doi.org/10.5194/acp-12-6377-2012

Schmidt, G. A., Jungclaus, J. H., Ammann, C. M., Bard, E., Braconnot, P., Crowley, T. J., et al. (2012). Climate forcing reconstructions for use in PMIP simulations of the Last Millennium (v1.1). *Geoscientific Model Development*, *5*(1), 185–191. https://doi.org/10.5194/gmd-5-185-2012

Schmidt, G. A., Ruedy, R., Hansen, J. E., Aleinov, I., Bell, N., Bauer, M., et al. (2006). Present-day atmospheric simulations using GISS ModelE: Comparison to in situ, satellite, and reanalysis data. *Journal of Climate*, *19*(2), 153–192. https://doi.org/10.1175/JCLI3612.1

Schneider, D. P., Ammann, C. M., OttoBliesner, 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*, *114*, D15101. https://doi.org/10.1029/2008JD011222

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., & Rudolf, B. (2014). 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. https://doi.org/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. https://doi.org/10.1175/JCLI-D-12-00826.1

Sigl, M., Winstrup, M., McConnell, J. R., Welten, K. C., Plunkett, G., Ludlow, F., et al. (2015). Timing and climate forcing of volcanic eruptions for the past 2,500 years. *Nature*, *523*(7562), 543–549. https://doi.org/10.1038/nature14565

Smerdon, J. E. (2011). Climate models as a test bed for climate reconstruction methods: Pseudoproxy experiments. *WIREs Climate Change*, *3*, 63–77. https://doi.org/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. https://doi.org/10.1126/science.296.5568.727

St George, S. (2014). An overview of tree-ring width records across the Northern Hemisphere. *Quaternary Science Reviews*, *95*, 132–150. https://doi.org/10.1016/j.quascirev.2014.04.029

St George, S., & Ault, T. R. (2014). The imprint of climate within Northern Hemisphere trees. *Quaternary Science Reviews*, *89*, 1–4. https://doi.org/10.1016/j.quascirev.2014.01.007

St George, S., & Esper, J. (2019). Concord and discord among Northern Hemisphere paleotemperature reconstructions from tree rings. *Quaternary Science Reviews*, *203*, 278–281. https://doi.org/10.1016/j.quascirev.2018.11.013

Steiger, N. J., Hakim, G. J., Steig, E. J., Battisti, D. S., & Roe, G. H. (2013). Assimilation of time-averaged pseudoproxies for climate reconstruction. *Journal of Climate*, *27*(1), 426–441. https://doi.org/10.1175/JCLI-D-12-00693.1

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. https://doi.org/10.1073/pnas.1612505114

Stevenson, S., Otto-Bliesner, B. L., Brady, E. C., Nusbaumer, J., Tabor, C., Tomas, R., et al. (2019). Volcanic eruption signatures in the isotope-enabled last millennium ensemble. *Paleoceanography and Paleoclimatology*, *34*, 1534–1552. https://doi.org/10.1029/2019PA003625

Stevenson, S., Otto-Bliesner, B., Fasullo, J., & Brady, E. (2016). El Niño like hydroclimate responses to last millennium volcanic eruptions. *Journal of Climate*, *29*(8), 2907–2921. https://doi.org/10.1175/JCLI-D-15-0239.1

Stoffel, M., Khodri, M., Corona, C., Guillet, S., Poulain, V., Bekki, S., et al. (2015). Estimates of volcanic-induced cooling in the Northern Hemisphere over the past 1,500 years. *Nature Geoscience*, *8*(10), 784–788. https://doi.org/10.1038/ngeo2526

Tardif, R., Hakim, G. J., Perkins, W. A., Horlick, K. A., Erb, M. P., Emile-Geay, J., et al. (2019). Last Millennium Reanalysis with an expanded proxy database and seasonal proxy modeling. *Climate of the Past*, *15*(4), 1251–1273. https://doi.org/10.5194/cp-15-1251-2019

Timmreck, C. (2012). Modeling the climatic effects of large explosive volcanic eruptions. *Wiley Interdisciplinary Reviews: Climate Change*, *3*(6), 545–564. https://doi.org/10.1002/wcc.192

Timmreck, C., Lorenz, S. J., Crowley, T. J., Kinne, S., Raddatz, T. J., Thomas, M. A., & Jungclaus, J. H. (2009). Limited temperature response to the very large AD 1258 volcanic eruption. *Geophysical Research Letters*, *36*, L21708. https://doi.org/10.1029/2009GL040083

Tolwinski-Ward, S. E., Evans, M. N., Hughes, M. K., & Anchukaitis, K. J. (2011). An efficient forward model of the climate controls on interannual variation in tree-ring width. *Climate Dynamics*, *36*(11-12), 2419–2439. https://doi.org/10.1007/s00382-010-0945-5

Toohey, M., & Sigl, M. (2017). Volcanic stratospheric sulfur injections and aerosol optical depth from 500 BCE to 1900 CE. *Earth System Science Data*, *9*(2), 809–831. https://doi.org/10.5194/essd-9-809-2017

Vaganov, E. A., Hughes, M. K., & Shashkin, A. V. (2006). *Growth dynamics of conifer tree rings* (Vol. 183). New York, NY: Springer-Verlag. https://doi.org/10.1007/3-540-31298-6

von Storch, H., Zorita, E., Jones, J. M., Dimitriev, Y., González-Rouco, F., & Tett, S. F. B. (2004). Reconstructing past climate from noisy data. *Science*, *306*, 679–682. https://doi.org/10.1126/science.1096109

Wang, J., Emile-Geay, J., Guillot, D., Smerdon, J. E., & Rajaratnam, B. (2014). Evaluating climate field reconstruction techniques using improved emulations of real-world conditions. *Climate of the Pas*, *10*(1), 1–19. https://doi.org/10.5194/cp-10-1-2014

Watanabe, S., Hajima, T., Sudo, K., Nagashima, T., Takemura, T., Okajima, H., et al. (2011). MIROC-ESM 2010: Model description and basic results of CMIP5-20c3 mexperiments. *Geoscientific Model Development*, *4*(4), 845–872. https://doi.org/10.5194/gmd-4-845-2011

Wilson, R., Anchukaitis, K., Briffa, K. R., Büntgen, U., Cook, E., D'Arrigo, R., et al. (2016a). Last millennium Northern Hemisphere summer temperatures from tree rings: Part I: The long term context. *Quaternary Science Reviews*, *134*, 1–18. https://doi.org/10.1016/j.quascirev.2015.12.005

Wu, T., Song, L., Li, W., Wang, Z., Zhang, H., Xin, X., et al. (2014). An overview of BCC climate system model development and application for climate change studies. *Journal of Meteorological Research*, *28*(1), 34–56. https://doi.org/10.1007/s13351-014-3041-7

Zhang, H., Yuan, N., Esper, J., Werner, J. P., Xoplaki, E., Büntgen, U., et al. (2015). Modified climate with long term memory in tree ring proxies. *Environmental Research Letters*, *10*(8), 084,020. https://doi.org/10.1088/1748-9326/10/8/084020

Zhu, F., Emile-Geay, J., Hakim, G. J., Tardif, R., & Perkins, A. (2019). LMR Turbo (LMRt): A lightweight implementation of the LMR framework. https://doi.org/10.5281/zenodo.3590258

References from the Supporting Information

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. https://doi.org/10.1002/qj.49712555417

Seabold, S., & Perktold, J. (2010). Statsmodels: Econometric and statistical modeling with python. In *9th Python in Science Conference*.

van der Walt, S., Colbert, S. C., & Varoquaux, G. (2011). The NumPy array: A structure for efficient numerical computation. *Computing in Science Engineering*, *13*(2), 22–30. https://doi.org/10.1109/MCSE.2011.37

van Rossum, G., & Drake, F. L. Jr. (1995). *Python tutorial*. The Netherlands: Centrum voor Wiskunde en Informatica Amsterdam.
Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods*, *17*(3), 261–272. https://doi.org/10.1038/s41592-019-0686-2
Waskom, M., Botvinnik, O., O'Kane, D., Hobson, P., Ostblom, J., Lukauskas, S., et al. (2018). mwaskom/seaborn: v0.9.0 (july 2018). Zenodo. https://doi.org/10.5281/zenodo.1313201