Last millennium Northern Hemisphere summer temperatures from tree rings: Part II, spatially resolved reconstructions

Kevin J. Anchukaitis\(^a,b,c,*\), Rob Wilson\(^d,c\), Keith R. Briffa\(^e\), Ulf Büntgen\(^f,g,h\), Edward R. Cook\(^c\), Rosanne D’Arrigo\(^c\), Nicole Davi\(^c,d\), Jan Esper\(^i\), David Frank\(^b,g\), Björn E. Gunnarson\(^b\), Gabi Hegerl\(^i\), Samuli Helama\(^m\), Stefan Klesse\(^b,g\), Paul J. Krusic\(^k,c,n\), Hans W. Linderholm\(^o\), Vladimir Myglan\(^p\), Timothy J. Osborn\(^c\), Peng Zhang\(^o\), Milos Rydval\(^d,o\), Lea Schneider\(^i\), Andrew Schurer\(^i\), Greg Wiles\(^i\), Eduardo Zorita\(^t\)

\(^a\)School of Geography and Development, University of Arizona, Tucson, AZ, USA
\(^b\)Laboratory of Tree-Ring Research, University of Arizona, Tucson, AZ, USA
\(^c\)Lamont-Doherty Earth Observatory of Columbia University, Palisades, NY, USA
\(^d\)School of Geography and Geosciences, University of St Andrews, UK
\(^e\)Climatic Research Unit, School of Environmental Sciences, University of East Anglia, Norwich, UK
\(^f\)Department of Geography, University of Cambridge, Cambridge, UK
\(^g\)Swiss Federal Research Institute WSL, Birmensdorf, Switzerland
\(^h\)Global Change Research Centre and Masaryk University Brno, Czech Republic
\(^i\)Department of Environmental Science, William Paterson University, Wayne, NJ, USA
\(^j\)Department of Geography, Gutenberg University, Mainz, Germany
\(^k\)Department of Physical Geography, Stockholm University, Stockholm, Sweden
\(^l\)School of GeoSciences, University of Edinburgh, Edinburgh, UK
\(^m\)Natural Resources Institute Finland, Rovaniemi, Finland
\(^n\)Navarino Environmental Obs., Messinia, Greece
\(^o\)Department of Earth Sciences, University of Gothenburg, Göteborg, Sweden
\(^p\)Siberian Federal University, Krasnoyarsk, Russia
\(^q\)Faculty of Forestry and Wood Sciences, Czech University of Life Sciences Prague, Czech Republic
\(^r\)Department of Geography, Justus Liebig University, Giessen, Germany
\(^s\)Tree Ring Lab, The College of Wooster, Wooster, OH, USA
\(^t\)Institute of Coastal Research, Helmholtz-Zentrum Geesthacht (HZG), Hamburg, Germany

Abstract

Climate field reconstructions from networks of tree-ring proxy data can be used to characterize regional-scale climate changes, reveal spatial anomaly patterns associated with atmospheric circulation changes, radiative forcing, and large-scale modes of ocean-atmosphere variability, and provide spatiotemporal targets for climate model comparison and evaluation. Here we use a multiproxy

\(^*\)Corresponding author
Email address: kanchukaitis@email.arizona.edu (Kevin J. Anchukaitis)
network of tree-ring chronologies to reconstruct spatially resolved warm season (May-August) mean temperatures across the extratropical Northern Hemisphere (40-90°N) using Point-by-Point Regression (PPR). The resulting annual maps of temperature anomalies (750 to 1988 CE) reveal a consistent imprint of volcanism, with 96% of reconstructed grid points experiencing colder conditions following eruptions. Solar influences are detected at the bicentennial (de Vries) frequency, although at other time scales the influence of insolation variability is weak. Approximately 90% of reconstructed grid points show warmer temperatures during the Medieval Climate Anomaly when compared to the Little Ice Age, although the magnitude varies spatially across the hemisphere. Estimates of field reconstruction skill through time and over space can guide future temporal extension and spatial expansion of the proxy network.

Keywords: Tree-rings, Northern hemisphere, Last millennium, Common Era, Summer temperatures, Reconstruction, spatial

1. Introduction

Global and hemispheric temperature anomalies reflect the influence of both internal variability in the climate system as well as the consequences of changes in radiative forcing, such as insolation, volcanic eruptions, and greenhouse gas concentrations. Surface temperature is determined by the planetary energy balance and serves as a symptom of perturbations to that balance, but also contains variability due to natural climate system dynamics. Rising global mean surface temperature is a key diagnostic for the influence of increasing greenhouse gases on the Earth’s climate system. Yet changes in incoming solar radiation, orbital (Milankovich) changes, albedo and land use alterations, and natural and anthropogenic aerosols also influence surface temperatures. Different radiative forcing mechanisms as well as internal modes of coupled ocean-atmosphere variability may have distinct fingerprints on temperature anomalies across different spatial, temporal, and seasonal scales [Hegerl et al., 1997; Rind et al., 1999; Shindell et al., 2001b; Hegerl et al., 2003; Rind et al., 2004; Shindell et al., 2003, 2004].
Hegerl et al., 2006, 2007; Shindell and Faluvegi, 2009; Shindell et al., 2014; Shindell et al., 2015). Surface temperature anomalies are therefore controlled by the superposition of various external radiative and internal dynamical influences on the climate system. Detection and attribution of the causes of temperature fluctuations, as well as the prediction of future regional-scale changes, thus depend on accurate quantification and understanding of spatial and temporal variations in surface temperature (Hegerl et al., 1997; Stott and Tett, 1998; Meehl et al., 2004; Lean and Rind, 2008; Stott and Jones, 2009; Stott et al., 2010; Solomon et al., 2011; Hegerl and Stott, 2014).

Paleoclimate reconstructions of past temperature extend knowledge of climate system variability beyond that available from the limited instrumental observational record. They offer longer timescales over which to observe a more complete range of variability in solar and volcanic forcing, extended opportunities to characterize internal climate system fluctuations at decadal and longer timescales, and the potential to separate forced and unforced responses to better understand their magnitude and spatiotemporal patterns (Hegerl et al., 2003, 2007). Spatially-explicit reconstructions provide additional opportunities to refine our understanding of fundamental climate system characteristics, diagnose the influence of different forcings on various aspects of the climate system, and provide insight into both regional climate changes and the response of large-scale modes of ocean-atmosphere variability (Seager et al., 2007; Cook et al., 2010a, b; Hegerl and Russon, 2011; Phipps et al., 2013; PAGES 2k-PMIP3 group, 2015; Goosse, 2016). Comparisons between palaeoclimatic data and models also provide out-of-sample tests of the general circulation models (GCMs) used for future climate projections and can indicate where the modeled forced response or internal variability requires further evaluation and continued refinement. Such comparisons may help constrain the probable range of model parameters or identify the forcing configurations most consistent with past climate variability (Edwards et al., 2007; Anchukaitis et al., 2010; Schmidt, 2010; Hegerl and Russon, 2011; Brohan et al., 2012; Schurer et al., 2013; Schmidt et al., 2014; Harrison et al., 2015; Tingley et al., 2015).
Reconstructions of last millennium and Common Era surface temperatures have focused predominantly on single time-series to represent continental- to global-scale variations in mean annual or growing season temperatures aggregated over space (Frank et al., 2010; Masson-Delmotte et al., 2013; PAGES2K, 2013; Stoffel et al., 2015; Smerdon and Pollack, 2016) while fewer have used climate field reconstruction (CFR) methods (Fritts, 1991; Cook et al., 1994; Evans et al., 2001; Tingley et al., 2012) to quantify past temperature anomalies simultaneously through time and across space (c.f. Mann et al., 1998; Tingley and Huybers, 2013; Wang et al., 2015). Spatial field reconstructions offer the benefit of characterizing regional-scale climate changes, can reveal spatial anomaly patterns or fingerprints associated with atmospheric circulation, radiative forcing, and large-scale modes of ocean-atmosphere variability, and provide complete spatiotemporal targets for GCM evaluation (Evans et al., 2001; Anchukaitis and McKay, 2014; Kaufman, 2014; Schmidt et al., 2014).

Here, we develop and evaluate a climate field reconstruction of extratropical Northern Hemisphere summer temperatures using an updated network of temperature-sensitive tree-ring proxy chronologies and existing temperature reconstructions back to 750 CE (Wilson et al., 2016). We are motivated by two fundamental challenges to the development of skillful large-scale last millennium temperature reconstructions revealed over the last two decades (c.f. Frank et al., 2010; Smerdon and Pollack, 2016): First, biases arising from characteristics of the proxies themselves; Second, uncertainties arising from the choice of reconstruction methodologies.

Tree-ring proxies provide precise annual dating and are broadly distributed across extra-tropical land areas, making them one of the most widely used proxies for climate reconstructions of the Common Era (Hughes, 2002; Jones et al., 2009; Smerdon and Pollack, 2016). Yet despite these advantages, certain challenges or limitations exist: they preferentially reflect growing season temperature conditions, they require some manner of processing to remove non-climatic age or tree geometry related growth trends, and there exist a wide range of climate responses amongst the more than two thousand tree-ring chronologies.
currently archived in public repositories (Briffa, 1995, 2000; Briffa et al., 2002, 2004; St. George, 2014; St. George and Ault, 2014). A decade ago, D’Arrigo et al. (2006) and Wilson et al. (2007) used small high-latitude networks of tree-ring proxy chronologies to reconstruct mean annual Northern Hemisphere temperatures. These and subsequent efforts have illuminated several extant challenges: a relatively limited number of unambiguously temperature-sensitive chronologies, a predominance of ring-width chronologies in comparison to the more temperature-sensitive wood density measurements (D’Arrigo et al., 1992; Schweingruber et al., 1993; Briffa et al., 2004; Frank et al., 2007; D’Arrigo et al., 2009; Esper et al., 2015; Wilson et al., 2016), and the influence of non-stationarity in climate/tree growth associations (‘divergence’; Briffa et al., 1998b; Wilson et al., 2007; D’Arrigo et al., 2008), a particular problem for North American treeline tree-ring width chronologies (Jacoby and D’Arrigo, 1995; Andreu-Hayles et al., 2011; Anchukaitis et al., 2013) and many previously collected wood density chronologies (Briffa et al., 2002). Since the publication of D’Arrigo et al. (2006) and Wilson et al. (2007), dozens of new tree-ring chronologies and local temperature reconstructions have become available, including many new and updated latewood density (MXD) measurement series that do not appear to exhibit any divergence (c.f. D’Arrigo et al., 2009; Esper et al., 2010; Anchukaitis et al., 2013). We draw on these new, published, and updated data here to develop a spatial reconstruction of past summer temperature stretching back to 750 CE. This work extends the non-spatial hemisphere mean reconstruction published by Wilson et al. (2016).

Methodologies for last millennium climate reconstructions have been extensively investigated and tested over the last two decades (Mann and Rutherford, 2002; Rutherford et al., 2003; Zorita et al., 2003; von Storch et al., 2004; Esper et al., 2005; Mann et al., 2005; Smerdon et al., 2008; Lee et al., 2008; Smerdon et al., 2010; Mann et al., 2007; Li et al., 2010; Smerdon et al., 2011; Wang et al., 2015; Smerdon and Pollack, 2016). However, neither reduced space, empirical orthogonal regression methods (c.f. Fritts, 1991; Cook et al., 1994; Mann et al., 1998) nor most variants of regularized expectation maximization (RegEM
Schneider, 2001; Rutherford et al., 2003; Mann et al., 2009) explicitly consider the location of the proxies relative to the reconstruction target, and therefore reconstructions of individual grid points can be significantly influenced by distant proxy sites. While there is potential value in taking into account large-scale teleconnections between the climate in a remote region and the local conditions controlling proxy formation (Evans et al., 2001, 2002), the potential disadvantage to these approaches is that they rely implicitly on the existence, stability, and persistence of those teleconnections (Gershunov and Barnett, 1998; Rimbu et al., 2003; Anchukaitis et al., 2006; Wilson et al., 2010; Lehner et al., 2012; Gallant et al., 2013; Ortega et al., 2015; Wise, 2015; Lewis and LeGrande, 2015). We assume the time-stability of large-scale covariance patterns, and risk admitting distal predictors with spurious relationships to local temperatures. However, several methods exist which do account for the spatial distribution and relationship of the proxy predictor network to the target field (Cook et al., 1999; Tingley and Huybers, 2010; Steiger et al., 2014). Here we use Point-by-Point regression (PPR; Cook et al., 1999) to develop, for the first time, a hemispher-scale spatial temperature reconstruction. For an extensive review of the full range of climate field reconstruction methods, see Tingley et al. (2012).

We seek a spatial reconstruction of past temperatures that utilizes a predictor network with a clear biophysical and statistical relationship with temperature. We therefore apply the expert knowledge of the original developers of the individual chronologies and reconstructions, relying on their experience with respect to climate signal evaluation and statistical treatment (c.f. Esper et al., 2016). PPR is a well-established, transparent, and spatially-explicit methodology for climate field reconstruction. The result of PPR here is a gridded map of summer temperature anomalies for each year of the last \( \sim 1200 \) years with associated validation skill metrics in both space and time. An additional benefit of PPR is that limiting the grid point reconstructions to proximal predictors and avoiding assumptions about the global covariance structure ensures that distant grid points remain independent of one another. The spatial features of the field can therefore be examined and compared without concern that they arise from
We compare our results against radiative forcing over the last millennium (Schmidt et al., 2012), diagnosing the potential role of volcanic eruptions and insolation variability in shaping the Northern Hemisphere extratropical warm season temperature response across space and time. Our field reconstruction can serve as a resource for understanding temperature variability in the past, for comparison with other proxy records of environmental and climate change, and to provide context for coupled human-natural systems response to climate variability (e.g. Buckley et al., 2010; Büntgen et al., 2011; Pederson et al., 2014; Büntgen et al., 2016). Our reconstruction also provides a spatiotemporal target appropriate for formal detection and attribution of the influence of different sources of radiative forcing on the Earth’s climate system.

2. Materials and Methods

2.1. Tree-ring series

Tree-ring based reconstructions of large-scale climate variability, whether a single mean time series or a spatial climate field reconstruction, require selecting potential proxy predictors from the many thousands of chronologies that have been developed over the last century of tree-ring research (St. George, 2014). The majority of these chronologies were not collected or developed with the intended purpose of temperature reconstruction and do not contain a primary temperature signal. It is therefore necessary to apply some selection procedure, lest chronologies lacking temperature information overwhelm the predictor set with ‘noise’ associated with soil moisture, archaeological selection, ecological processes, or spurious non-temperature signals. Two broad categories of approach have been used for initial predictor selection: statistical screening and expert assessment. In statistical screening (c.f. Mann et al., 2008), proxy series are assessed for their correlations against local or regional observational temperature data, with those data passing some significance, sign, or effect size threshold then admitted to the pool of potential predictors. While this
approach clearly has merit as an objective and automated approach, it virtually guarantees the selection of a proportion of proxies without a realistic biophysical or substantial statistical association with the climate variable targeted for reconstruction. A second approach utilizes expert assessment of individual chronologies based on an understanding of whether a proxy has the requisite ecological, biological, geographical, and climatological characteristics to serve as a reasonable temperature proxy. Although this approach is not completely isolated from statistical considerations, the advantage is that it strongly reduces the likelihood that a non-temperature proxy or nonsense predictor will enter into a temperature reconstruction model. Expert assessment can include not only climate signal and ecological criteria, but also the methods used to develop the proxy series (e.g., Esper et al., 2016). A potential disadvantage to this approach is that it is partially subjective and therefore different investigators could make different selections for the predictor pool. Hybrid approaches that combine simple mechanistic or physiological assessment with statistical evaluation have also been applied (c.f. Tierney et al., 2015).

Here, we follow Wilson et al. (2016) and use as our predictor series only published tree-ring chronologies and temperature reconstructions that demonstrate an established and biophysically reasonable association with local temperatures. These include both tree-ring chronologies as well as existing temperature reconstructions from Northern Hemisphere high-latitude and high-altitude locations, where dendrochronological and ecological principles suggest the most limiting factor for growth is temperature (Fritts, 1976). We exclude from consideration chronologies south of 40°N to avoid confounding climate signals associated with moisture-sensitive trees (St. George, 2014) and we also reject chronologies known to demonstrate evidence of the ‘divergence problem’ (Briffa et al., 1998b). D’Arrigo et al. (2008, 2009) and Wilson et al. (2007), a problem previously observed to affect North American Picea glauca tree-ring width chronologies (D’Arrigo et al., 1992, 2009; Andreu-Hayles et al., 2011; Esper et al., 2012) and many MXD chronologies developed in the 1970s and 1980s. We require that our predictors extend back to at least 1750 CE and completely forward to 1988 CE.
The predictors retain the detrending and standardization choices of the original authors. The resulting dataset is designated N-TREND2015 and is composed of a mixture of tree-ring width (TRW), MXD (Schweingruber et al. 1978), and blue intensity (BI; McCarroll et al. 2002) data. N-TREND2015 is archived and publicly available[1] and is the same dataset used in Part 1 of this study (Wilson et al. 2016). N-TREND is intended to be a ‘living dataset’ that will grow or be modified as new proxies become available or are updated. Details of the predictor time series used here and in Wilson et al. (2016) are available in Table 1. As shown in Figure 1, the NTREND-2015 network reflects a mix of proxy types, dominated by MXD or BI (43 series, vs. 11 composed of tree-ring width data only). A total of 54 series are available from 1750 to 1988 CE, the time period of full (denoted ‘BEST’) coverage of the network. The number of sites drops precipitously toward the present, down to 34 by 1990 CE and to 25 series by 2000 CE, with only 3 sites remaining by 2011. There are 23 series at 1000 CE and 4 remain at 750 CE, the limits of our reconstruction. While all the tree-ring chronologies and reconstructions have significant and substantial correlations with local temperatures in one or more months, locations that include MXD and BI data overall have higher correlations compared to sites composed of tree-ring width data alone (Figure 2; see also Wilson et al. (2016)).

2.2. Observational data

Our target field (predictand) for our temperature reconstruction is the interpolated hybrid (surface and satellite information) version of HadCRUT4 from Cowtan and Way (2014). The original HadCRUT4 (Morice et al. 2012) consists of monthly temperature anomalies relative to the mean of the 1961 to 1990 CE period on a regular 5° latitude/longitude grid and combines CRUTEM4 (Jones et al. 2012) over land with HADSST3 (Kennedy et al. 2011a,b) for the oceans. Use of the Cowtan and Way (2014) dataset provides several advantages here: First, this dataset seeks to compensate for observational coverage bias.

and provides gridded estimates of monthly temperature at high latitudes, including portions of our target reconstruction region north of 40°N. Second, it is spatially and temporally complete, allowing us to use the same calibration and validation periods for our reconstruction at every location in the field, which in turn permits straightforward comparisons of reconstruction skill. Following Wilson et al. (2016), we use May through August (MJJA) mean temperature anomalies as our target variable, as this season provides a network-wide balance across the diverse site-local monthly or seasonal climate responses of the individual predictor series (Wilson et al., 2016).

2.3. Reconstruction and statistical methodology

As a prelude to our climate field reconstruction, we assess the spatial characteristics of the temperature signal across our network. We first calculate the Pearson Product Moment correlation between each series and the local gridded MJJA temperatures from Cowtan and Way (2014) as a measure of the correspondence between tree growth and local gridded temperatures. We also calculate, for each site, the field correlation between the individual predictor series and the entire MJJA and annual mean temperature field, using both the original data as well as first-differenced series. For assessing the association between the target field and the predictor network, we follow Schneider et al. (2015) and compute the median correlation coefficient between the temperature record at each grid box and all the predictor times series within 2000 km of the centroid (see below; Briffa and Jones, 1993; Jones et al., 1997; Cook et al., 2013). We perform this procedure for both the best replicated part of our predictor network (1750 to 1988 CE) and at 1000 CE in the midst of the Medieval epoch. Collectively, these statistical assessments provide an evaluation of the climate signal embedded in the predictor network through time and space.

We use Point-by-Point Regression (PPR; Cook et al., 1999) to reconstruct the MJJA surface temperature anomaly field north of 40°N using our predictor network of tree-ring proxy chronologies and temperature reconstructions. We follow the method as developed, tested, and described by Cook et al. (1999).
PPR incorporates the spatial structure of the predictor network and predictand field and confines the potential region of influence of the predictors to a distance estimated from the underlying correlation structure of the temperature field. PPR proceeds by calculating a nested multivariate regression model for each grid point in the target field with the predictors restricted to those within some radius from the grid point centroid. We adopted a dynamic search radius for each grid point in the target field, first identifying predictor series within 1000km. If no chronologies were found within 1000km, the radius was allowed to expand in 500km increments up to a maximum of 2000km to find predictors. These distances are based on the decorrelation decay as a function of distance in the target field data (Cowtan and Way 2014) and are also consistent with the findings of other studies (Briffa and Jones 1993; Jones et al. 1997; Cook et al. 2013). If no predictors were found within 2000km, then no climate reconstruction was produced for that grid point. In evaluating our methods, we found that this dynamic search radius provides an optimal balance between maximizing the number of grid points available for reconstruction while allowing the local predictor series in data-dense regions (for instance, Fennoscandia and western Europe) to provide the paleoclimate information for their neighborhood of grid points.

A multivariate regression model was calibrated for each grid point and its associated predictors over the period 1945 to 1988 CE (the latest date for which all chronologies had data) and then validated on withheld observational data over the period 1901 to 1944 CE (Michaelson 1987). We also checked the sensitivity of our reconstruction to this choice of calibration and validation periods by swapping them and assessing cross-validation. As the number of predictor series declines back through time, the model is newly calibrated and validated at each change in sample depth. In our reconstruction, we use the individual series themselves as predictors as opposed to the leading principal components (PCs). In our sensitivity tests of the PPR method, we discovered that using PCs from a relatively sparse network of chronologies and reconstructions, combined with the expansive target field, created clearly artificial inhomogeneities or disconti-
nunities when predictor numbers declined. Using the individual predictor time
series themselves in a stepwise regression model with an adjusted $R^2$ entry rule
(Meko, 1997) ameliorated such discontinuities. Model skill was assessed using
the calibration $R^2_c$ (adjusted for the number of predictors), the validation $R^2_v$,
the Reduction of Error (RE) and the Coefficient of Efficiency (CE) (c.f. Cook
et al., 1999; Wilson et al., 2006). In addition to the annual maps of reconstructed
temperatures, we calculate an extratropical Northern Hemisphere mean MJJA
time series using a latitude weighted average of all the reconstructed grid cells
where reconstructed values are available back to at least 1000CE and RE is
greater than zero.

Following Masson-Delmotte et al. (2013), we calculated the difference be-
tween reconstructed Medieval Climate Anomaly (MCA) and Little Ice Age
(LIA) temperatures by taking the difference between the mean values over the
field for 950 to 1250 CE (MCA) and 1450 to 1850 CE (LIA). As there is no single
accepted definition of these two periods (Hughes and Diaz, 1994; Bradley et al.,
2001; Matthews and Briffa, 2005; Seager et al., 2008; Mann et al., 2009), we
also tested the sensitivity of the calculated MCA-LIA to differences in the time
definition of these periods. We estimated the temperature response to tropical
explosive volcanism (Robock, 2000; Ammann and Naveau, 2003) by calculating
the composite mean anomaly using superposed epoch analysis (e.g. Haurwitz
and Brier, 1981). Event years were extracted from most recent updated esti-
mates of Common Era volcanic forcing from Sigl et al. (2015), here selecting
those years corresponding to an estimated maximum negative event forcing (in
$Wm^{-2}$) with a magnitude at least as large as Krakatoa in 1883.

3. Results

3.1. Network climate signal

All our predictor series show significant and typically high correlations with
local summer temperatures over one or several months (Figure 2, 3; Wilson
et al., 2016). At sites where the highest local summer temperature signal
in the series is confined to one or two months – for example, at Yakutia in Russia – local correlations with the broader MJJA season are lower ($r = 0.52$, $p < 0.05$ for July vs. $r = 0.09$, $p > 0.05$ for MJJA at Yakutia). The highest individual monthly/seasonal site local temperature correlations \textsuperscript{[Wilson et al. 2016]} range from $r = 0.39$ to $r = 0.84$ (mean $r = 0.63$), while correlations with local MJJA temperatures range from $r = 0.04$ to $r = 0.78$ (mean $r = 0.43$). The highest correlations with local MJJA temperatures are in Fennoscandia and north central Russia, with strong local temperature signals also evident in Scotland, the Alps, the Pyrenees, the Altai, and Japan. In North America, MXD chronologies from western Canada and the northern treeline have the strongest MJJA signals. Tree-ring width only chronologies in North America have a generally weaker association with their local MJJA instrumental temperature than MXD or mixed proxy sites \textsuperscript{[D’Arrigo et al. 1995, D’Arrigo et al. 2009, Andreu-Hayles et al. 2011, Anchukaitis et al. 2013]}. Field correlations between the predictor series and the full MJJA temperature field (Figure 4) suggest that the chronologies and temperature reconstructions reflect climate variability over many hundred or thousands of kilometers, with some exceptions at those sites where the local proxy response to MJJA is already weak (e.g. locations in east Asia, Yakutia, and ring width-only chronologies from the North American treeline). The spatial extent of the large-scale correlation structure is partially related to the common positive trends in predictors and temperatures during the 20th century, as temporarily removing these trends by first differencing both the field and the predictors (Figure 5) reduces the regions of positive correlations to between 500 and 2000 kilometers. In some cases – e.g. the Idaho (USA) chronology – a significant interannual correlation with the MJJA temperature field is entirely absent, indicating the summer association there is driven by common trends and a narrow local monthly temperature response. Annual temperatures are often a target for temperature reconstruction (e.g. \textsuperscript{Esper et al. 2002, D’Arrigo et al. 2006, Mann et al. 2008, 2009}); however, correlations with the annual temperature field herein are uniformly lower and many chronologies show no significant association with an-
nual mean temperatures (Figure 6). In general, higher correlations with annual
mean temperatures are observed for those sites with higher local correlations
with MJJA temperatures ($r_{MJJA, annual} = 0.73, p < 0.01$). The association
between the annual signal and the best (highest) monthly or seasonal local cli-
mate correlations is weaker ($r_{best, annual} = 0.44, p < 0.01$), indicating it is not
a strong climate signal alone that corresponds with a useful annual proxy, but
rather a strong and broad seasonal climate response.

Figure 7 shows the statistical relationships between the predictand target
field and predictor network. Grid points in Asia, Fennoscandia, and Northern
Europe have 15 or more proxy sites that can serve as predictors over the best
replicated epoch (1750 to 1988 CE). In northwestern North America as many as
10 predictors are within 2000km of the grid centroids. In contrast, grid points in
southern Europe and eastern North America have many fewer predictors avail-
able, in some cases only a single series. For the best replicated portion of the
reconstruction, 1750 to 1988 CE, the median correlation between observational
MJJA temperatures at each grid point and the predictors within 2000km of
that grid point range from $r = -0.19$ in northeastern Russia to $r = 0.66$ in
Fennoscandia. In general, the median grid correlations are highest where they
are co-located with one or more chronologies (e.g. interior British Columbia)
and regions with clusters of strong MJJA temperature proxies (e.g. Fennoscandi-
a, the Alps and Pyrenees, northern North American treeline). By 1000 CE,
when the number of predictors is reduced to 23, with only 3 of these in North
America, the possible reconstruction domain is reduced as substantially fewer
predictors are available for each grid point reconstruction. Nevertheless, grid
median correlations between observed temperatures and the available predictors
remain significant and of similar magnitude to the better replicated recent por-
tion of the reconstruction. This is because predictors with strong temperature
signals remain in the Alps, Icefields in British Columbia, the Gulf of Alaska,
Fennoscandia, northern Russia, and the Altai. The reconstruction at 1000 CE
and earlier therefore relies on a reduced number of predictors, but those that
remain contain a significant and substantial temperature signal.
3.2. Field reconstruction, calibration, and validation

The full NTREND spatial reconstruction consists of 1239 yearly fields of MJJA temperature anomalies covering up to a potential 792 grid points each year (40 to 90°N, -180 to 180°E). Figure 8 shows the length of the reconstructed temperature series at each grid point in our domain. In practice the number of grid points with a value in each year of the reconstruction is less than that potential maximum, as the full network can only support reconstruction at 85% (n = 701) of the total grid cells in the domain, declining to 51% (n = 401) by 1000 CE and 24% (n = 190) by 750 CE. The shortest reconstructions are for ocean grid points in the northwestern Atlantic and northeastern Pacific, where in both cases a single and relatively short chronology is the only source of proxy information. Other relatively short portions of the field reconstruction include the central northern treeline in Canada, Japan, northeastern Russia, and the southernmost part of the Eurasian domain in the Mediterranean. By contrast, along the Eurasian treeline in central northern Russia, throughout the Nordic region, and in western Europe, it is possible to reconstruct more than a millennium of past temperatures.

Figure 9 shows reconstruction skill (the adjusted $R^2_c$, $R^2_v$, $RE$, and $CE$) for the best replicated (n = 54) period of the predictor network, 1750 to 1988 CE. Significant skill is observed over the entire domain, but is clearly highest closest to those predictors with the highest correlations to MJJA temperatures, especially where MXD or BI data are available. Grid cells more distal from the predictor network, including cells over the oceans, or where only a single or relatively weak predictor is available, show lower explained variance and in some cases lack positive $RE$ and $CE$ scores. Validation $R^2_v$ scores are lower but largely mirror the calibration $R^2_c$, with the exception of the eastern Mediterranean and Black Sea region, east Asia, and western parts of Central Asia. A similar phenomenon of lower $R^2_v$ was observed by Cook et al. (2013) for Asia, due at least in part to the lack of instrumental climate data from these regions during the reconstruction model validation period (Jones et al. 2012, Harris et al. 2013, Cook et al. 2013, Cowtan and Way 2014). Lack of instrumental data
likely confounds out-of-sample validation in the eastern Mediterranean prior to the 1930s (c.f. Touchan et al., 2014). Skillful regions with $RE$ and $CE$ scores greater than 0 are more spatially confined but likewise show skill with respect to these metrics in regions where chronologies are present or abundant, with the exception once again of the regions mentioned previously. $R^2_c$ values range from \( \sim 0.0 \) to 0.78, $R^2_v$ values range from \( \sim 0.0 \) to 0.76, $RE$ up to 0.72, and $CE$ up to 0.70. A cross-validation (not shown) interchanging the time periods used for calibration and validation reveals that the reconstruction’s skill characteristics are largely insensitive to the choice of these periods.

By 1000 CE, the reduction in the number of predictors and a contraction in their spatial distribution influences both the number of grid points reconstructed and the spatial patterns of skill (Figure 10). The loss of the northern treeline MXD chronologies in North America reduces the reconstructed regions of the continent in the west to coastal Alaska, the Pacific Northwest, Interior British Columbia and parts of Alberta, and in the east to Quebec, Newfoundland, and Labrador. Likewise, the loss of the Scottish and Pyrenees chronologies no longer allow for reconstruction of temperatures over the British Isles and the Iberian peninsula. The lack of Japanese, coastal Russia, and east Asia series at 1000 CE leads to a contraction of the reconstructed spatial domain in the east. However, skillful temperature reconstructions persist over parts of northwestern North America, Fennoscandia, northern Russia, the Alps, and the Altai. $R^2_c$ values range from \( \sim 0.0 \) to 0.73, $R^2_v$ values range from \( \sim 0.0 \) to 0.70, $RE$ up to 0.65, and $CE$ up to 0.61. Interchanging the calibration and validation periods has a minor influence on field skill at 1000 CE, with $n = 229$ grid points with $RE > 0$ for a late calibration (1945 to 1988 CE), and $n = 213$ for an early calibration (1901 to 1944 CE).

Figure 11 shows domain-wide reconstruction and skill metrics over time. For the time span of the full predictor network (1750 to 1988 CE), we can reconstruct a temperature anomaly value for 88.5% of the grid points ($n = 701$) in our extratropical Northern Hemisphere domain. Of those 701 grid points, 63% have $RE > 0$, and 37% have $CE > 0$. These skill percentages remain
remarkably stable ($RE$, 56% to 65%; $CE$, 34% to 37%) even as the number of reconstructed grid points with a reconstructed value declines back through time, a consequence of the shrinking predictor network. Only when the number of reconstructed grid points declines precipitously in the earliest 10th century, falling to 38% of the target domain in 905 CE, do the percent of grid points with $RE$ and $CE$ greater than zero increase substantially. These patterns indicate that the reduction in the number of reconstructed grid cells comes at the cost of locations with already marginal skill scores, while the core reconstruction regions associated with sensitive predictor series persist through much of the length of the reconstruction back to 750 CE.

3.3. Large-scale mean temperature anomalies and climate forcing

We computed a mean summer temperature anomaly series from our domain by calculating a latitude-weighted mean of all gridded values where a reconstruction is available back to 1000CE and $RE$ is greater than zero (Figure 12). This time series is highly and significantly correlated with the comparable observational temperature anomalies averaged over those same grid points (1850–1988, $n = 139$, $r = 0.78$, $p << 0.001$). The mean series indicates a broad warm period from at least 750 CE to the early 1400s, with maximum values centered around the late 900s, the late 1000s and 1100s, and the individual warmest years of the Medieval epoch in 790, 990, 995, 1014, 1016, and 1168 CE. 1168 CE is the warmest year in our reconstruction (750 to 1988 CE), although matched by values in the middle of the 20th century and then exceeded during the comparable filtered instrumental record in the early 21st century. Temperatures decline in the late 13th century, coincident with a series of tropical volcanic eruptions and the Wolf solar irradiance minimum, before warming again during the 14th century. Temperatures then decline sharply in the early 1400s, slightly before the significant volcanic eruptions in the 1450s and the Spörer solar irradiance minimum (1460 to 1550 CE). The mid-1400s through the mid-1800s show cooler conditions during the LIA.
Masson-Delmotte et al. (2013), with local minima in the 1450s, the late 1500s and earliest 1600s, the late 1600s, and the early 1800s. Many of the field-mean coldest years in the reconstruction, including 1259, 1453, 1601, 1643, 1783, 1810, 1817, and 1836 CE are associated with or follow closely after major tropical or Northern Hemisphere volcanic eruptions (Briffa et al., 1998a; Sigl et al., 2015).

In several instances extremely cold years in our reconstruction occur or persist at least one or two years after eruptions (e.g. 1643 and 1644, 1816 and 1817, 1836 and 1837 CE), consistent with the findings of Esper et al. (2013b). The coldest year in our reconstruction is 1601 CE, which also agrees with a prior temperature reconstruction based solely on MXD data by Briffa et al. (1998a), and which follows the eruption of Huaynaputina in Peru in 1600 CE (Verosub and Lippman, 2008). Our ten coldest years include at least 2 (1699 and 1867 CE) that do not appear to be associated with a known volcanic eruption (Briffa et al., 1998a).

Our filtered time series is highly and significantly correlated (750–1988 CE, \( n = 1239, r = 0.71, p << 0.01 \)) with the index reconstruction from Wilson et al. (2016) and perhaps unsurprisingly has many of the same features - a broad warm period during the Medieval until the early 1400s, the LIA from the 15th through early 19th century, warming out of the LIA, a warm mid-20th century, cool 1970s, and recent warming (Figure 13a). There are epochs where our mean time series and that of Wilson et al. (2016) are less strongly correlated (Figure 13b); interestingly, the rapid returns to higher correlation values appear to be associated with the timing of major individual or clusters of volcanic events, suggesting that strong radiative forcing due to volcanism may impose a common spatial forcing, causing the Wilson et al. (2016) mean index reconstruction and the spatial mean from our climate field reconstruction to converge.

Our Northern Hemisphere extratropical time series shows associations with large-scale radiative forcing changes during the last millennium (Figure 13; Schmidt et al., 2012). Colder periods in the late 13th and early 14th, mid-15th, and 19th century occur at the same time as large explosive volcanic eruptions.
and solar minima (Mann et al., 1998; Crowley, 2000; Shindell et al., 2001a, 2003; Wagner and Zorita, 2005; Ammann et al., 2007; Breitenmoser et al., 2012; Anchukaitis et al., 2013; PAGES2k, 2013). Warming after the middle of the 19th century is consistent with a reduced number of volcanic eruptions, increasing insolation, and the rapid rise in greenhouse gases (Andronova and Schlesinger, 2000; Zwiers and Weaver, 2000; Gillett et al., 2012; Jones et al., 2013; Estrada et al., 2013). Although the timing of several epochs of colder temperatures appear to align with solar minima – for instance, the late 13th/early 14th century and the Wolf Minimum or the mid-15th century Spörer Minimum – the correlations between our hemisphere mean reconstruction and estimates of past solar variability (Schmidt et al., 2012) are low (range, $r = 0.10$ to $r = 0.29$, $p < 0.01$).

At centennial timescales, however, there is evidence that solar variability may play a more substantial role. Wavelet coherence (Figure 15; Grinsted et al., 2004) between our hemisphere mean reconstruction and last millennium total solar irradiance estimates assembled by Schmidt et al. (2012) shows high and stable coherence and consistent phasing at bi-centennial time scales (194 to 222 year periods), which bracket and include the $\sim 206$ year ‘de Vries’ (or Suess) solar cycle (Stuiver and Braziunas, 1993; Wagner et al., 2001) and are likely related to the reconstructed temperature response to the major solar minima.

The spectral signal of the bicentennial de Vries cycle has been recognized in numerous tree-ring chronologies and temperature reconstructions (c.f. Raspopov et al., 2008; Breitenmoser et al., 2012; Ogurtsov et al., 2016), and Emile-Geay et al. (2013) also identified bi-centennial periodicity in a reconstruction of eastern tropical Pacific sea surface temperatures. Phase relationships between our hemisphere mean temperature and the total solar irradiance time series suggest a decadal-scale lag of $\sim 11$ years (range, 5 to 20 years), with solar changes leading temperature anomalies, consistent with both climate modeling and prior analysis of tree-ring chronologies and solar variability (Rind et al., 1999; Waple et al., 2002; Breitenmoser et al., 2012). Both the hemisphere mean as well as the spatial grid point sensitivity to solar variability ($C/Wm^{-2}$) are extremely uncertain in this analysis, however, as this quantity is highly sensitive to the
choice of solar reconstruction [Schmidt et al., 2012].

Medieval Climate Anomaly (MCA; 950 to 1250 CE) temperatures compared against those during the Little Ice Age (LIA; 1450 to 1850 CE) show warmer temperature during the MCA at ∼90% of the grid points with minimally skillful (RE > 0) reconstructed values available back to 950 CE (Figure 16). Colder Medieval temperatures are reconstructed over parts of the Altai and Central Asia, and are associated with tree-ring width chronologies from Mongolia that show reduced growth in the 900s and 1100s, despite warmer conditions in the 11th century [Davi et al., 2015] and an MXD chronology from the Altai displaying a cold Medieval epoch and warm LIA [Schneider et al., 2015]. Other grid cells that show a colder Medieval period tend to be distal from the predictor network – for instance in central Greenland – and must be treated with caution. Defining a different MCA or LIA in this case has relatively little effect on the percentage of grids showing warmer vs. colder conditions, as the cooler Central Asia grid cells and marginal cells in Greenland, central Canada, the southern Caspian Sea, and the northeastern Pacific remain irrespective of the specific date range applied. The precise boundaries for both MCA and LIA are, in any case, both arbitrary and uncertain [Hughes and Diaz, 1994; Bradley et al., 2001; Matthews and Briffa, 2005; Seager et al., 2008]. The cause of apparently extremely high (∼3°C) Medieval temperatures in several grid points in northeastern North America is discussed below.

Composite mean MJJA temperature anomaly fields following major volcanic eruptions show coherent, broad-scale cooling associated with large tropical eruptions (Figure 17). 96% of grid points show composite mean colder temperatures compared to the three years prior to the 20 large eruptions considered here. Similar to the MCA-LIA difference discussed above, regions that apparently have an overall composite warming response to volcanic eruptions are largely on the margins of the reconstruction domain, away from the predictor grid, and over the ocean, central Greenland, and the southern Caspian Sea. The coldest grid point composite mean is -1.61°C, and the mean composite response across all grid points and all eruptions is -0.44°C. Closer examination of indi-
individual eruption events (not shown) finds that for some regions, post-volcanic cooling may persist for several years and maximum cold anomalies may be 1 or 2 years after the year of the eruption itself, consistent with observations of other regional temperature reconstructions (D'Arrigo et al., 2013; Cook et al., 2013; Esper et al., 2013b; Davi et al., 2015; Linderholm et al., 2015; Schneider et al., 2015; Wilson et al., 2016). If we consider large Northern Hemisphere high-latitude eruptions only (Figure 18), the large-scale response is likewise toward cold anomalies overall: 89% of grid points have a composite anomaly less than zero. Over the entire field the mean composite response is -0.39°C and the maximum cold composite anomaly is -2.31°C. There is also spatial structure to the temperature anomalies, with the coldest composite conditions over Alaska and the Bering Strait, northeastern North America, parts of western Europe, and central northern Russia, suggestive of a dynamical, in addition to direct radiative, influence of large-magnitude high latitude eruptions (Robock, 2000; Oman et al., 2005; Stenchikov et al., 2006; Schneider et al., 2009; Zanchettin et al., 2012; Pausata et al., 2015). However, the number of radiatively significant high latitude eruptions considered here is smaller (n = 5), and therefore this structure may appear due to the limited sample size.

4. Discussion

4.1. Proxy data and predictor network

Our results here demonstrate that a relatively small (n = 54) network of proxy sites (Table 1, Figure 1) with well-established physically and ecologically reasonable climate signals (Figure 3, 7) can be used to reconstruct the large-scale summer temperature history of the extratropical Northern Hemisphere (Figure 8, 9, 10, 12). While restricting the reconstruction to the higher latitudes of one hemisphere and to only the growing season does not provide a global annual estimate of temperature, it nonetheless accurately reflects the geographic and biological signal that dominates the predictors. Moreover, we have demonstrated here that the reconstruction preserves the signature and influence of external
forcing on the global energy balance. Skill in our reconstruction is, perhaps not
surprisingly, greatest in those locations where high quality temperature-sensitive
proxies are available (Figure 9, 10), and declines at increasing distances from
the predictors themselves. It is clear we could realize substantial benefits in
terms of increased reconstruction skill and spatial extent by developing MXD
and BI chronologies from currently undersampled regions as well as extending
the length of existing MXD and BI chronologies. Although large and useful
multiproxy datasets have resulted from community efforts by the paleoclimate
community (PAGES2K 2013), our analysis here demonstrates that a relatively
small well-distributed network of highly sensitive millennium-length tree ring
chronologies provide skillful reconstructions over a large extratropical region.
Encouragingly, this means that rapid and important gains could be made from
the addition of a relatively small number of new sites and the temporal exten-
sion and recollection of current sites known to contain a strong climate signal.
In particular, a greater number of long MXD and BI chronologies are critically
needed from North America (Figures 1, 10). Fulfilling this need will require a
collaborative and concerted effort to locate subfossil materials and to measure
density proxies, but the potential gain for Northern Hemisphere temperature
reconstructions of the Common Era would be substantial. MXD contains a
stronger temperature signal than TRW alone and is better able to accurately
resolve rapid temperature changes associated with volcanic eruptions (Figure
2; Frank et al. 2007, D’Arrigo et al. 2013, Esper et al. 2015). Continuing
advances in blue intensity (BI) measurements suggest some of the benefits of
wood density analysis can be realized without the expense and difficulty of an-
alyzing MXD itself (Campbell et al. 2007, Wilson et al. 2014, Rydval et al.
2014, Björklund et al. 2014, Björklund et al. 2015) although the low-frequency
characteristics of BI still requires additional exploration.

Detrending and standardization issues for long chronologies remain an ongo-
ing challenge and persistent source of uncertainty (e.g. Cook et al. 1995, Briffa
and Melvin 2011, Melvin and Briffa 2008, Esper et al. 2012, Matskovsky and
Helama 2014, Esper et al. 2016, Matskovsky and Helama 2016). One surpris-
ing feature of the epochal comparison between the MCA and LIA is the large
(> 3°C) difference calculated for northeastern North America. This feature is
due to a single predictor, a black spruce (*Picea mariana*) tree-ring width RCS
chronology developed by Gennaretti et al. (2014). Other, non-tree ring proxies
from the region suggest a lower amplitude of cooling between the MCA and the
LIA, although issues related to time uncertainty, transfer function calibration,
and different seasonal climate signals complicate exact comparisons. A com-
pilation of Holocene paleoenvironmental data for the Arctic (Sundqvist et al.
2014) suggests a range of values for MCA to LIA cooling in northeastern North
America and Greenland of 0 to 1.5°C, which is at least a degree less than the
magnitude inferred from the Quebec MXD record. Alkenone SST reconstruc-
tions near Nova Scotia (Keigwin et al. 2003) suggest a cooling of approximately
0.7°C. The multiproxy PAGES Arctic2k reconstruction (McKay and Kaufman
2014) has a whole-Arctic reconstructed MCA-LIA difference of 0.64°C for the
same time periods used here. Finally, and perhaps most importantly, a tree-
ring oxygen isotope temperature reconstruction from the same site in Quebec
(Naulier et al. 2015) shows a substantially smaller estimated MCA-LIA differ-
ence of 0.4°C. It seems likely that, despite the care taken in applying regional
curve standardization to the Quebec black spruce samples (Autin et al. 2015)
as well as its suitability with respect to other chronology metrics (Esper et al.
2016), artifacts remain in this tree-ring width chronology that unintentionally
but artificially amplify the difference between MCA and LIA temperatures in
this region.

More generally, detrending, the removal of non-climatic trends, and there-
fore the retention of low frequency variability, remains an important source
of uncertainty in the amplitude of past temperatures reconstructed from tree
rings (Cook 1987; Briffa et al. 1996), even when conservative detrending tech-
niques have been applied. While regional curve standardization and signal free
methods have been shown to be able to retain the full spectrum of low- and
medium-frequency variability, they are also subject to their own uncertainties
and assumptions (Melvin 2004; Melvin and Briffa 2008; Briffa and Melvin
It is in most cases not possible to know from calibration and validation statistics which detrending method yields the true or most accurate low frequency signal (Cook, 1987; Cook and Kairiukstis, 1990; Wilson et al., 2007). Possible approaches to this problem include both ensemble and simulation-based methods (e.g. Esper et al., 2007; D’Arrigo et al., 2011; Anchukaitis et al., 2013), although these have not yet been applied to large and heterogeneous tree-ring proxy networks.

4.2. Radiative forcing and temperature history

Our reconstruction demonstrates coherent responses to radiative forcing in time and space (Figures 14, 15, 16, 17, 18). Temporal features of the reconstruction are associated with changes in solar irradiance, large and/or clustered tropical volcanic eruptions, and the anthropogenic rise in well-mixed greenhouse gases. Temperatures decline across the field during the Spörer and Maunder Minima, in particular, likely compounded in both cases by a series of volcanic eruptions. Temperatures remained cold during the early 1600s, at least in part due to the eruption of Huaynaputina in Peru in 1600 CE (Verosub and Lippman, 2008). 1601 CE is the coldest year of our entire reconstruction, as it was in the 600 year temperature reconstruction by Briffa et al. (1998a) and the hemisphere mean reconstruction by Wilson et al. (2016). 1601 CE was also one of the coldest years in the Bayesian field reconstruction by Tingley and Huybers (2013), as was 1453 CE, which is the 4th coldest year in our study, associated with the eruption of Kuwae, Vanuatu (but see Plummer et al., 2012; Cole-Dai et al., 2013). Interestingly, Tingley and Huybers (2013) find 1642 CE amongst their coldest years, whereas in our reconstruction it is 1643 CE that is exceptionally cold (5th coldest in our reconstruction). In Wilson et al. (2016), 1641, 1642, and 1643 CE are all amongst the coldest 15 years of their reconstruction. Tingley and Huybers (2013) also find that 1695 CE was anomalously cold, whereas here it is indeed cold but unremarkable (-0.60°C, 284th coldest). These differences highlight extant uncertainties likely related to different reconstruction methods, spatial skill and averaging, and the use of different proxies (D’Arrigo et al., 2011; Anchukaitis et al., 2013).
2013), but also demonstrate that there is no evidence for a one-to-one correspondence between inferred volcanic forcing from ice cores and the magnitude of hemisphere-scale cooling. For instance, the eruption of Huaynaputina in 1600 is believed to have caused a smaller negative radiative forcing anomaly than eruptions in 1458, 1641, 1809, and 1815 CE, let alone the large Medieval eruption of Samalas (1257 CE) (Verosub and Lippman 2008; Lavi-gne et al. 2013; Sigl et al. 2015). Our finding here of a large-scale, coherent cooling in response to explosive volcanism is yet further evidence (Anchukaitis et al. 2012; Brohan et al. 2012; D’Arrigo et al. 2013; Esper et al. 2013b, a; St. George et al. 2013; Büntgen et al. 2014; Jull et al. 2014; Esper et al. 2015; Sigl et al. 2015; Stoffel et al. 2015; Wilson et al. 2016) against the hypothesis that tree-ring proxies are missing the volcanic cooling signal due to undetected absent rings (Mann et al. 2012, 2013).

Low frequency coherence between solar variability and our reconstruction appears to be a stable characteristic through time (Figure 14, 15), likely linked to reconstructed cold anomalies during solar Grand Minima. Because variations in total solar irradiance are relatively small, on the order of a few tenths of a $Wm^{-2}$, the mechanism that could result in a detectable cooling remains uncertain. The most likely connection is via changes in large-scale Northern Hemisphere circulation, which favor colder temperature over continents (e.g. Shindell et al. 2001a, 2003; Swingedouw et al. 2010) and thus would be captured in our reconstruction. Nevertheless, while variability in solar forcing may be important on bicentennial and perhaps at continental scales, fingerprinting suggests that the solar effect in the hemisphere-scale anomalies is otherwise relatively small and that volcanic forcing is more important overall in determining pre-industrial temperature trajectories (Schurer et al. 2014; McGregor et al. 2015). There is no sign in our reconstruction of a discernible temperature response to the shorter 11 and 22 year sunspot cycle (Schwabe/Hale), which is consistent with other investigations of the insolation signal in tree rings (e.g. Briffa 1994). There are a number of possible reasons for the absence of this signal: Internal climate system variability is substantially stronger at interannual
and decadal time scales, which may prevent statistical detection of solar influences with similar frequencies, but still allow it at the centennial scale when the magnitude of internal variability is smaller than the forced signal. Short-term climate anomalies caused by explosive volcanism could also disrupt detection of a decadal solar signal. The temperature response to solar variability at lower frequencies may also reflect slow temperature feedbacks that enhance its direct effect.

Over those grid points available back to 950 CE with minimum level of reconstruction skill ($RE > 0$), ~90% show warmer conditions during the MCA than during the LIA, with a field median difference of 0.32°C. Removing likely individual grid point outliers (Greenland and northeastern North America, see above) results in a slightly smaller epochal field median difference (0.30°C) and a range of grid point values of -0.64 to +1.05°C. Mann et al. (2009) calculated a 0.24°C global summer mean difference between MCA and LIA, but the difference in season, spatial domain and geographic extent, and the ‘fragility’ (Wang et al., 2015) of reconstructing a cold Medieval tropical Pacific make any direct comparison difficult. Calculating the MCA-LIA epochal difference using the spatial mean time series (Figure 12) gives a value of 0.36°C, approximately in the middle of the distribution for other large-scale Northern Hemisphere reconstructions, and within the higher end of the range of values from climate model simulations (Fernández-Donado et al., 2013; Wilson et al., 2016).

5. Conclusions and future work

We have reconstructed the extratropical Northern Hemisphere MJJA temperature anomaly field back to 750 CE using a network of temperature-sensitive predictors. The reconstruction shows significant field skill associated with proximity to the predictors, particularly where proxy density data are available. In other words, we observe the most reconstruction skill and smallest errors where we have the most sensitive tree-ring proxies, whereas higher errors and lower skill are associated with grid points distal from the predictor network or where only
tree-ring width data are available, particularly in North America. These observations will be used to guide future sampling and proxy development priorities, including the development of new sites, efforts to increase the number of MXD and BI series, and the extension in time of existing high quality chronologies.

Our field reconstruction reveals coherent responses to changes in radiative forcing over the last 1200 years, including the influence of solar and volcanic forcing. Future research with our field reconstruction will use fingerprint detection \cite{Hegerl2007, Schurer2013, Schurer2014} to quantitatively assess the role of forcing and internal variability, including identification of spatial patterns linked to large-scale modes of variability and specific forcing agents. Formal, quantitative comparison between our reconstruction and paleoclimate model simulations \cite{Schmidt2012, Kageyama2016} will be used to assess climate model performance and to investigate the dynamical context for reconstructed spatial temperature anomalies. Using proxy system models \cite{Evans2013}, the NTREND network could also be applied within an offline data assimilation framework \cite{Steiger2014, Hakim2016}. Finally, our spatially-explicit reconstructions can be used to explore and understanding the possible role of past temperature variability – especially volcanic eruptions – in contributing to historical societal dynamics, resilience, and change \cite{McCormick2007, Ludlow2013, Sigl2015, Buettgen2016}.

**Acknowledgements**

The N-TREND consortium is not itself funded, but many individuals acknowledge relevant projects, grants, and support; KJA: National Science Foundation Paleoclimate Perspectives on Climate Change NSF AGS-1501856 and NSF AGS-1501834; RW: UK Natural Environment Research Council (NERC - NE/K003097/1) and Leverhulme Trust project (F/00 268/BG); KRB and TO: NERC (Belmont Forum/JPI-Climate: INTEGRATE project NE/P006809/1); KRB also thanks Gaurav Kapur FRCR and Colin Watts FRCS for time; UB: Czech project ‘Building up a multidisciplinary scientific team focused on drought’
No. CZ.1.07/2.3.00/20.0248; EC: National Science Foundation Paleoclimate Perspectives on Climate Change AGS-1502224; RD: National Science Foundation Paleoclimate Perspectives on Climate Change AGS-1159430, AGS-1502150, and AGS-1502224; SH: Academy of Finland; GH and AS: ERC advanced grant TITAN (EC-320691), NCAS GH specifically with a Royal Society Wolfson Research Merit Award (WM130060); GH and AS: PACMEDY, NE/P006752/1; HL: The Swedish Science Council (VR) (2012-5246); MR: The Carnegie Trust for the Universities of Scotland; GW: NSF AGS-1502186. The N-TREND project website, along with the archived TR chronologies and temperature reconstructions can be found at https://www.ncdc.noaa.gov/cdo/f?p=519:1:0::::P1_study_id:19743 and additional information is available at https://ntrenddendro.wordpress.com/. Lamont-Doherty Earth Observatory contribution ####.
Table 1: Tree-ring chronology and temperature reconstruction predictor series used in the reconstruction. Additional information is available in Wilson et al. (2016). Latitude is given in degrees North, and longitude is in degrees East. A range of latitude/longitude indicates the data or reconstruction cover a larger region. Detrending: STD: traditional detrending; RCS: regional curve standardization; SF: signal free extension to STD or RCS standardization. † indicates series used in D’Arrigo et al. (2006) and ‡ those used in Wilson et al. (2007).

<table>
<thead>
<tr>
<th>Site Name</th>
<th>Code</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Time Span</th>
<th>Proxy Type</th>
<th>Detrending</th>
<th>Citation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NORTH AMERICA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Seward</td>
<td>NTR</td>
<td>65.11 to 65.22</td>
<td>-162.18 to -162.27</td>
<td>1710-2001</td>
<td>MXD</td>
<td>STD</td>
<td>D’Arrigo et al., 2004</td>
</tr>
<tr>
<td>Coastal Alaska</td>
<td>GOA</td>
<td>60.01 to 60.45</td>
<td>-149.31 to -141.42</td>
<td>800-2010</td>
<td>TRW</td>
<td>RCS</td>
<td>Wilson et al. 2014</td>
</tr>
<tr>
<td>Winsdell</td>
<td>WRA</td>
<td>65-65</td>
<td>-141.00 to -140.00</td>
<td>1393-1982</td>
<td>MXD</td>
<td>STD</td>
<td>Wilson et al. 2013</td>
</tr>
<tr>
<td>Wrangells</td>
<td>WRAx</td>
<td>60-65</td>
<td>-145.00 to -140.00</td>
<td>1593-1992</td>
<td>MXD</td>
<td>RCS-SF</td>
<td>Anchukaitis et al. 2013</td>
</tr>
<tr>
<td>Southern Yukon</td>
<td>YUS</td>
<td>58 to 62</td>
<td>-140.60 to -133.60</td>
<td>1644-2000</td>
<td>TRW</td>
<td>STD</td>
<td>Vogel et al. and Luckman, 2006</td>
</tr>
<tr>
<td>Northern Yukon</td>
<td>YUN</td>
<td>45 to 70</td>
<td>-125.00 to -135.00</td>
<td>1638-1988</td>
<td>TRW</td>
<td>STD</td>
<td>Wilson and Luckman, 1995</td>
</tr>
<tr>
<td>Int. British Columbia</td>
<td>IBC</td>
<td>49.02 to 50.59</td>
<td>-121.43 to -117.03</td>
<td>1600-1995</td>
<td>TRW/MXD/BI</td>
<td>STD/SF</td>
<td>Wilson et al., 2014</td>
</tr>
<tr>
<td>EURASIA</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scotland</td>
<td>SCOT</td>
<td>57.06</td>
<td>-3.44</td>
<td>1290-2010</td>
<td>TRW/BI</td>
<td>STD/RCS</td>
<td>Rydval et al., in review</td>
</tr>
<tr>
<td>Porvoo</td>
<td>PYR</td>
<td>42 to 43</td>
<td>0 to 1</td>
<td>1290-2005</td>
<td>MXD</td>
<td>RCS</td>
<td>Koskolou et al., 2014</td>
</tr>
<tr>
<td>W Alps - Loteinental</td>
<td>ALPS</td>
<td>46.5</td>
<td>9</td>
<td>755-2004</td>
<td>MXD</td>
<td>RCS</td>
<td>Strobl et al., 2004</td>
</tr>
<tr>
<td>E Alps - Tyn</td>
<td>TYR</td>
<td>47.30</td>
<td>12.30</td>
<td>1015-2001</td>
<td>MXD</td>
<td>RCS</td>
<td>Schneider et al., 2015</td>
</tr>
<tr>
<td>Black Forest</td>
<td>JAB</td>
<td>61.30</td>
<td>12.35</td>
<td>783-2011</td>
<td>MXD</td>
<td>RCS</td>
<td>Wiles et al., 2014</td>
</tr>
<tr>
<td>Tarpist, Arizone &amp; Nuevo</td>
<td>TAA</td>
<td>65.14 to 66.16</td>
<td>10.80 to 12.12</td>
<td>1250-2010</td>
<td>MXD</td>
<td>RCS</td>
<td>Anchukaitis et al., 2013</td>
</tr>
<tr>
<td>North Pacifica</td>
<td>FIFAC</td>
<td>65-65</td>
<td>19.50</td>
<td>750-2010</td>
<td>MXD</td>
<td>RCS</td>
<td>Wilson et al., 2014</td>
</tr>
<tr>
<td>Forfjorddalen</td>
<td>FORF</td>
<td>57.08</td>
<td>15.30</td>
<td>978-2005</td>
<td>MXD</td>
<td>RCS</td>
<td>McCarroll et al., 2013</td>
</tr>
<tr>
<td>Tatra</td>
<td>TAT</td>
<td>48 to 49</td>
<td>19 to 20</td>
<td>1090-2010</td>
<td>TRW</td>
<td>RCS</td>
<td>Wiles et al., 2014</td>
</tr>
<tr>
<td>Mt Olympus, Greece</td>
<td>MOG</td>
<td>41.36 to 42.11</td>
<td>22.37</td>
<td>1090-2010</td>
<td>TRW/MXD/BI</td>
<td>STD/RCS</td>
<td>Wilson et al. 2007</td>
</tr>
<tr>
<td>Khibiny (Kola)</td>
<td>KOL</td>
<td>67.38 to 67.50</td>
<td>31.13 to 34.15</td>
<td>821-2005</td>
<td>MXD</td>
<td>STD/RCS</td>
<td>McCarroll et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 1</td>
<td>Grid1</td>
<td>40.15 to 46.15</td>
<td>60.15 to 68.15</td>
<td>817-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 2</td>
<td>Grid2</td>
<td>40.15 to 46.15</td>
<td>60.15 to 68.15</td>
<td>817-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Altai</td>
<td>ALT</td>
<td>50.00</td>
<td>88.00</td>
<td>750-2007</td>
<td>MXD</td>
<td>RCS</td>
<td>Schneider et al., 2015</td>
</tr>
<tr>
<td>Asia Grid 3</td>
<td>Grid3</td>
<td>40.15 to 46.15</td>
<td>80.15 to 88.15</td>
<td>800-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 4</td>
<td>Grid4</td>
<td>40.15 to 46.15</td>
<td>90.15 to 98.15</td>
<td>800-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 5</td>
<td>Grid5</td>
<td>40.15 to 46.15</td>
<td>100.15 to 108.15</td>
<td>800-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 6</td>
<td>Grid6</td>
<td>40.15 to 46.15</td>
<td>110.15 to 118.15</td>
<td>800-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 7</td>
<td>Grid7</td>
<td>40.15 to 46.15</td>
<td>120.15 to 128.15</td>
<td>800-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 8</td>
<td>Grid8</td>
<td>40.15 to 46.15</td>
<td>130.15 to 138.15</td>
<td>800-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 9</td>
<td>Grid9</td>
<td>40.15 to 46.15</td>
<td>140.15 to 148.15</td>
<td>1396-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 10</td>
<td>Grid10</td>
<td>40.15 to 46.15</td>
<td>150.15 to 158.15</td>
<td>1510-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 11</td>
<td>Grid11</td>
<td>40.15 to 46.15</td>
<td>160.15 to 168.15</td>
<td>1510-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 12</td>
<td>Grid12</td>
<td>40.15 to 46.15</td>
<td>170.15 to 178.15</td>
<td>1510-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 13</td>
<td>Grid13</td>
<td>40.15 to 46.15</td>
<td>180.15 to 188.15</td>
<td>1510-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 14</td>
<td>Grid14</td>
<td>40.15 to 46.15</td>
<td>190.15 to 198.15</td>
<td>1510-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 15</td>
<td>Grid15</td>
<td>40.15 to 46.15</td>
<td>200.15 to 208.15</td>
<td>1510-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 16</td>
<td>Grid16</td>
<td>40.15 to 46.15</td>
<td>210.15 to 218.15</td>
<td>1510-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 17</td>
<td>Grid17</td>
<td>40.15 to 46.15</td>
<td>220.15 to 228.15</td>
<td>1510-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>Asia Grid 18</td>
<td>Grid18</td>
<td>40.15 to 46.15</td>
<td>230.15 to 238.15</td>
<td>1510-1989</td>
<td>MXD</td>
<td>RCS/STD</td>
<td>Cook et al., 2013</td>
</tr>
<tr>
<td>North Japan</td>
<td>NJAP</td>
<td>43 to 51</td>
<td>142 to 145</td>
<td>1400-1989</td>
<td>TRW/MXD</td>
<td>STD</td>
<td>D’Arrigo et al., 2006</td>
</tr>
<tr>
<td>Yakutia</td>
<td>YAK</td>
<td>67.27 to 70.33</td>
<td>142.37 to 150.17</td>
<td>1342-1994</td>
<td>TRW</td>
<td>RCS</td>
<td>Hughes et al., 1999</td>
</tr>
</tbody>
</table>

Wilson et al. (2007)
Figure 1: Spatial and temporal distribution of sites by proxy type. Top panel (A) shows all the records available during the best replicated (most recent, 1750 to 1988 CE) nest, while the bottom panel (B) shows the distribution of available proxy sites at 1000 CE. Sites with tree-ring width (TRW) data only are shown in green, while sites that have MXD, BI, or a mix of proxy types are shown in blue. (C) Shows the total number of series through time.
Figure 2: Kernel probability density estimate [Bowman and Azzalini 1997] for the highest local seasonal or monthly coefficient correlations as a function of proxy type. Data are from Wilson et al. (2016). The density estimate is calculated for a support of [-1,1] and the values contributing to the distribution are indicated by symbols along the x-axis.
Figure 3: Local correlations between tree-ring proxy chronologies and local temperature data. (A) Correlations reported by the original authors for local correlations with the optimal seasonal or monthly window (see Wilson et al. 2016, their Table 1). (B) Correlations between each proxy series and the local May through August (MJJA) temperature data from Cowtan and Way 2014 used here.
Figure 4: Field correlations between each tree-ring site (indicated by the black stars) and the MJJA mean temperature field from Cowtan and Way (2014). Labels correspond with the site codes from Table 1. Around each site the black range ring indicates a radii of 2000 km. Only Pearson Product Moment correlation coefficients ($r$) significant at $p < 0.05$ as adjusted for autocorrelation (Trenberth, 1984) are plotted.
Figure 5: Field correlations between each tree-ring site (indicated by the black stars) and the MJJA mean temperature field from Cowtan and Way (2014) after first-differencing each variable to reduce the influence of common trends. Labels correspond with the site codes from Table 1. Around each site the black range ring indicates a radius of 2000 km. Only Pearson Product Moment correlation coefficients \( (r) \) significant at \( p < 0.05 \) as adjusted for autocorrelation (Trenberth, 1984) are plotted.
Figure 6: Field correlations between each tree-ring site (indicated by the black stars) and the annual mean temperature field from Cowtan and Way (2014) after first-differencing each variable to reduce the influence of common trends. Labels correspond with the site codes from Table 1. Around each site the black range ring indicates a radii of 2000 km. Only Pearson Product Moment correlation coefficients (r) significant at $p < 0.05$ as adjusted for autocorrelation (Trenberth, 1984) are plotted.
Figure 7: Proxy and target grid characteristics. Panels (A) and (C) show the number of tree-ring sites within 2000km for each grid point in the target field (Cowtan and Way, 2014), during the best replicated (modern) nest (A) and at 1000 CE (C), respectively. Black circles indicate the location of the available tree-ring sites during in each time period. Panels (B) and (D) show the median value at each target field grid point for the Pearson Product Moment correlation between the MJJA temperatures at that grid point and all the tree-ring chronologies within 2000km of that grid, during the best replicated (modern) nest (B) and at 1000 CE (D), respectively (see Schneider et al. (2015)).
Figure 8: Reconstruction length. The length (in years) of the reconstruction at each target grid point.
Figure 9: Reconstruction skill for the best replicated (modern) nest. Panels show spatial patterns of skill metrics for the best replicated nest (1750 to 1988 CE), as evaluated for the adjusted calibration $R^2$, the validation $R^2$, the reduction of error (RE), and the coefficient of efficiency (CE). Available tree-ring sites during this nest are indicated by black circles.
Figure 10: Reconstruction skill at 1000 CE. Panels show spatial patterns of skill metrics for the reconstructed field at 1000 CE in the midst of the Medieval epoch, as evaluated for the adjusted calibration $R^2$, the validation $R^2$, the reduction of error (RE), and the coefficient of efficiency (CE). Available tree-ring sites at 1000 CE are indicated by black circles.
Figure 11: Aggregate temporal reconstruction skill. The percent of all target grid points that are able to be reconstructed for a given year is shown by the black line. For those grid points with a reconstructed value in a given year, the red and blue lines show the percent of those grid points with RE and CE greater than zero.
Figure 12: Filtered latitude-weighted mean hemisphere MJJA temperature anomaly reconstruction and target MJJA observational time series. Spatial mean values for both the reconstruction (black) and target (red) MJJA fields are calculated from the set of all grid points that have reconstructed values back to at least 1000 CE and which have an RE score greater than zero at 1000 CE ($n = 229$) and are weighted by latitude. Uncertainty in the reconstruction is indicated by the gray shading, and is calculated as the mean latitude-weighted local mean squared error of validation. The reconstruction and target MJJA temperature series are significantly correlated over their common interval (1850–1988, $n = 139$, $r = 0.78$, $p < < 0.001$).
Figure 13: Comparison between time series reconstruction from (blue, Wilson et al., 2016) and the filtered weighted global mean MJJA temperature reconstructed here (black). Time series are shown in (A), and a running correlation (50 year window, 1 year increment) is plotted in (B). The full overall correlation between the two series (750 to 1988 CE, $r = 0.71$) is indicated by the dashed red line in (B).
Figure 14: Radiative forcing and reconstructed Northern Hemisphere warm-season temperatures from this study during the last millennium. All forcing series are those compiled by Schmidt et al. [2012] for PMIP3 simulations of the last millennium (version 1.1). (A) Volcanic forcing following Gao et al. [2008] (black/grey, GRA) and Crowley et al. [2008] (blue/light blue, CEA), with individual years as lighter lines and 30-year Gaussian smoothed values in heavy lines. Note that the magnitude of some individual events exceeds the y-axis limits. (B) Northern Hemisphere mean MJJA temperature anomaly time series as described in the text (black line) and corresponding observed temperatures for the same grid points (red line). Here both reconstructed and observed values have been smoothed with a 30 year Gaussian filter. (C) Solar forcing relative to the period 1976 to 2006 CE, with the pink shaded region showing the range of the forcing reconstructions compiled by Schmidt et al. [2012] including Delaygue and Bard [2011], Muscheler et al. [2007], Steinhilber et al. [2009] and Vieira and Solanki [2010]. Major solar minima are labeled. (D) Forcing due to land use change from Kaplan et al. [2011] (KK10) and Pongratz et al. [2009] (PEA). (E) Well-mixed greenhouse gas forcing.
Figure 15: Wavelet coherence (Torrence and Compo 1998; Grinsted et al. 2004) between our Northern Hemisphere mean MJJA temperature anomaly time series and solar forcing variability from Vieira and Solanki (2010). Arrows indicate the phase of the relationship and for clarity are plotted only where coherence exceeds 0.65. In-phase signals point directly to the right of the plot. Values above the cone of influence (COI; black curve) are potentially influenced by edge effects at that time period and scale.
Figure 16: Medieval Climate Anomaly (MCA; 950-1250 CE) vs Little Ice Age (LIA; 1450-1850 CE) mean temperature anomaly fields (MCA-LIA). Only grid points with values reconstructed at $RE > 0$ at 1000 CE are shown.
Figure 17: Composite mean reconstructed temperatures following major tropical volcanic eruptions (from Sigl et al. (2015)). Eruption years in the composite ($n = 20$) are those with a global forcing magnitude equal to or larger than that associated with Krakatoa (1884), and include 916, 1108, 1171, 1191, 1230, 1258, 1276, 1286, 1345, 1453, 1458, 1595, 1601, 1641, 1695, 1809, 1815, 1832, 1836, and 1884 CE. Event anomalies are calculated by first subtracting the global field mean over the 3 years prior to the eruption. Only grid points with $RE > 0$ in an event year are averaged to form the composite and only those grid points with values for at least 6 eruptions are plotted.
Figure 18: Composite mean reconstructed temperatures following major Northern Hemisphere high latitude volcanic eruptions (from Sigl et al. (2015)). Eruption years in the composite ($n = 5$) are those Northern Hemisphere eruption with a global forcing equal to or larger than the magnitude associated with Katmai (1912), and include 939, 1182 1210, 1783, and 1912 CE. Event anomalies are calculated by first subtracting the global field mean over the 3 years prior to the eruption. Only grid points with $RE > 0$ in an event year are averaged to form the composite and only those grid points with values for at least 2 eruptions are plotted.


tree-ring chronologies for reconstructing past temperature variability. Science
295, 2250–2253.

Northern European summer temperature variations over the Common Era
doi:10.1002/jqs.2726

Esper, J., Frank, D., Büntgen, U., Verstege, A., Hantemirov, R.M., Kirdyanov,

Esper, J., Frank, D., Büntgen, U., Verstege, A., Luterbacher, J., Xoplaki, E.,
34. doi:10.1029/2007GL030844

Esper, J., Frank, D.C., Timonen, M., Zorita, E., Wilson, R.J.S., Luterbacher,
J., Holzkamper, S., Fischer, N., Wagner, S., Niergelt, D., Verstege, A.,
862–866. doi:10.1038/nclimate1589

Esper, J., Krusic, P.J., Ljungqvist, F.C., Luterbacher, J., Carrer, M., Cook, E.,
of tree-ring based temperature reconstructions of the past millennium.
Quaternary Science Reviews 145, 134–151.

Esper, J., Schneider, L., Krusic, P.J., Luterbacher, J., Büntgen, U., Timonen,
M., Sirock, F., Zorita, E., 2013b. European summer temperature response
to annually dated volcanic eruptions over the past nine centuries. Bulletin of
Volcanology 75. doi:10.1007/s00445-013-0736-z

Esper, J., Schneider, L., Smerdon, J.E., Schöne, B.R., Büntgen, U., 2015. Sig-
nals and memory in tree-ring width and density data. Dendrochronologia 35,
62–70.


Kaplan, J.O., Krumhardt, K.M., Ellis, E.C., Ruddiman, W.F., Lemmen, C., Goldewijk, K.K., 2011. Holocene carbon emissions as a result of an-


global ocean cooling trend for the pre-industrial common era. Nature Geoscience 8, 671–677. doi:10.1038/ngeo2510.


Paleoclimate data–model comparison and the role of climate forcings over the past 1500 years. Journal of Climate 26, 6915–6936.


Schneider, L., Smerdon, J.E., Büntgen, U., Wilson, R.J.S., Myglan, V.S., Kirdyanov, A.V., Esper, J., 2015. Revising midlatitude summer tempera-


